UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

Microbial contributions to metabolic oxygen turnover in the central Arctic Ocean

Permalink

https://escholarship.org/uc/item/0k0439q9

Author

Chamberlain, Emelia Janthina

Publication Date

2023

Supplemental Material

https://escholarship.org/uc/item/0k0439q9#supplemental

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA SAN DIEGO

Microbial contributions to metabolic oxygen turnover in the central Arctic Ocean

A Dissertation submitted in partial satisfaction of the requirements
for the degree Doctor of Philosophy

in

Oceanography

by

Emelia J Chamberlain

Committee in charge:

Professor Jeff Bowman, Chair
Professor Andrew Barton
Professor Sara Jackrel
Professor Fiamma Straneo

2023

Copyright

Emelia J Chamberlain, 2023

All rights reserved.

The Dissertation of Emelia J Chamberlain is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2023

DEDICATION

To my parents, who did not live to see me pursue this degree yet were certainly with me every step of the way.

EPIGRAPH

I actually found bacteria — even these regions are not free of them!

-- Fridtjof Nansen, Farthest North

TABLE OF CONTENTS

DISSERTATION APPROVAL PAGE iii

DEDICATION iv

EPIGRAPH v

TABLE OF CONTENTS vi

LIST OF FIGURES viii

LIST OF TABLES x

LIST OF SUPPLEMENTAL FILES xi

ACKNOWLEDGEMENTS xii

VITA xvi

ABSTRACT OF THE DISSERTATION xviii

INTRODUCTION 1

CHAPTER 1 – Impacts of sea ice melting procedure on measurements of microbial community structure 4

Abstract 4

Introduction 5

Methods 8

Results 15

Discussion 26

Conclusions 31

Supplemental Material 33

Acknowledgements 37

CHAPTER 2 – Microbial predictors of net heterotrophic conditions in the pelagic Arctic Ocean 38

Abstract 38

Introduction 39

Materials and Methods 41

Results 48

Discussion 63

Supplemental Material 73

Acknowledgements 87

CHAPTER 3 – Low salinity summer meltwater alters surface Arctic Ocean microbial community structure and metabolism 89

Abstract 89

Introduction 89

Methods 92

Results 105

Discussion 117

Supplemental Material 125

Acknowledgements 127

REFERENCES 129

LIST OF FIGURES

Figure 1-1. Map and images of sampling site. 9

Figure 1-2. Schematic of the experimental set-up. 11

Figure 1-3. Microbial parameters by ice horizon, shaving, and melting method. 18

Figure 1-4. Overview of taxonomic diversity across all samples. 21

Figure 1-5. Significantly differentially abundant bacterial amplicon sequence variants (ASVs) between melting method within each ice horizon. 23

Figure 1-6. Significantly differentially abundant eukaryotic amplicon sequence variants (ASVs) between melting method within each ice horizon. 25

Figure 2-1. Sampling Locations. 43

Figure 2-2. SOM segmentation of the prokaryotic community. 49

Figure 2-3. SOM segmentation of the eukaryotic community. 52

Figure 2-4. Measurements of biological oxygen anomaly ($[O_2]_{bio}$) in μmol kg^{-1} from the upper 11 m of the water column. 54

Figure 2-5. Time series of upper water column biological oxygen anomaly and taxonomic mode during the MOSAiC Drift track. 55

Figure 2-6. Environmental attributes of each mode. 57

Figure 2-7. Final Random Forest (RF) model and validation of predictions for biological oxygen anomaly, $[O_2]_{bio}$ in μmol kg^{-1}. 59

Figure 2-8. Biological oxygen anomaly ($[O_2]_{bio}$) in μmol kg^{-1} from the upper 11 m of the water column, measurements vs. Random Forest (RF) predictions. 60

Figure 2-9. Upper water column discrete biological oxygen anomaly ($[O_2]_{bio}$), measurements vs. Random Forest (RF) predictions. 61

Figure 2-10. Heatmap of key microbial predictors of upper water column biological oxygen anomaly. 62

Figure 3-1. Relevant drift paths and lead sampling sites during the MOSAiC Expedition. ... 93

Figure 3-2. Physical characterization of leads from Legs 4 and 5 of the MOSAiC Expedition. 105

Figure 3-3. Similarities in microbial community structure across meltwater stratification. . 107

Figure 3-4. Dominant prokaryotic community composition in leads. 109

Figure 3-5. Dominant eukaryotic community composition in leads. 111

Figure 3-6. Observed net metabolic oxygen turnover in the meltwater layer. 112

Figure 3-7. Chlorophyll *a* concentrations across salinity treatments and culture strains 113

Figure 3-8. Bacterial and eukaryotic community structure in the algal cultures. 114

Figure 3-9. Oxygen production and aerobic respiration across salinity treatments. 117

Supplemental Figure 1-S1. Impacts of melting method and ice shaving on total microbial community structure. 33

Supplemental Figure 1-S2. Ratio intensity plots of differentially abundant bacterial ASVs between melt method and ice horizon. 34

Supplemental Figure 1-S3. Ratio intensity plots of differentially abundant eukaryotic ASVs between melt method and ice horizon. 35

Supplemental Figure 2-S1. Overview of full dataset taxonomic community structure. 80

Supplemental Figure 2-S2. Overview of prokaryotic SOM taxonomic community structure. 80

Supplemental Figure 2-S3. Overview of eukaryotic SOM taxonomic community structure. 81

Supplemental Figure 2-S4. Comparison between mass spectrometer measurements and CTD sensor measurements of total dissolved oxygen concentration $[O_2]_{tot}$. 81

Supplemental Figure 2-S5. Time series of full water column biological oxygen anomaly and taxonomic mode during the MOSAiC Drift track. 82

Supplemental Figure 2-S6. Training of iterative Random Forest (RF) Regression models using 70 % of available paired data. 83

Supplemental Figure 2-S7. Validation of iterative Random Forest (RF) Regression models using 30 % of available paired data. 85

Supplemental Figure 3-S1. Full overview of microbial community composition in leads. 126

Supplemental Figure 3-S2. Growth curves during the initiation phase of the three algae cultures. 127

LIST OF TABLES

Table 1-1. Procedural overview with averages and standard deviations of key parameters. ..16

Table 1-2. ANOVA summary statistics. ..17

Table 1-3 Kruskal-Wallis summary statistics for pairwise comparisons of melting method. .20

Table 1-4. PERMANOVA summary statistics. ..21

Table 2-1. Summary characteristics for all prokaryotic (PM) and eukaryotic (EM) modes. ..58

Table 3-1. Summary of lead sampling during the MOSAiC. ..94

Table 3-2. Values of chlorophyll *a*, pH, macronutrients, and the percentage of actively respiring cells averaged across triplicate bottles after an 8-day acclimation period to salinity treatments. ..115

Table 3-3. Chlorophyll *a*, pH, and nutrient turnover during the 48 hr oxygen bottle incubations. ..116

Supplemental Table 1-S1. Final dilutions for isohaline melt treatments. ..36

Supplemental Table 1-S2. Phylogenetic placement of significantly differentially abundant bacterial amplicon sequence variants (ASVs) between melt method comparisons. ..36

Supplemental Table 1-S3. Phylogenetic placement of significantly differentially abundant eukaryotic amplicon sequence variants (ASVs) between melt method comparisons. ..37

Supplemental Table 2-S1. Summary statistics and input data for all Random Forest models. ..76

Supplemental Table 2-S2. P-values from the Wilcoxon signed-rank test comparing environmental variables between prokaryotic modes ..78

Supplemental Table 2-S3. P-values from the Wilcoxon signed-rank test comparing environmental variables between eukaryotic modes ..79

Supplemental Table 2-S4. Boruta selected ASVs (final model) ..79

LIST OF SUPPLEMENTAL FILES

Supplemental File 1-S1. Chamberlain_Table_1-S2_16S_deseq_results.xlsx 36
Supplemental File 1-S2. Chamberlain_Table_1-S3_18S_deseq_results.xlsx 37
Supplemental File 2-S1. Chamberlain_Table_2-S4_predictor_ASVs.csv 79
Supplemental File 3-S1. Chamberlain_Video_LeadSurveys.mp4 127

ACKNOWLEDGEMENTS

I am lucky to have known so many wonderful people, without whom this dissertation would not have been possible.

First, I would like to acknowledge my advisor Dr. Jeff Bowman, who took a chance on me despite my earth science background and introduced me to the microbial world. Thank you for challenging me to achieve my best and supporting my many scientific side quests. I am honored to have been among your first students and it has been an absolute joy watching your lab grow and flourish over the years. I would also like to acknowledge my committee members, Drs. Andrew Barton, Fiamma Straneo, and Sara Jackrel. Your guidance and insight helped shape the direction of my PhD and I thank you for your patience and support.

I would additionally like to acknowledge my former scientific mentors at Boston University, Drs. Robinson (Wally) Fulweiler, Alia Al-Haj, and Andrew Christ. Not only did you introduce me to research and academia in the first place, but I have relied on your advice and support so many times in the past years that I truly would not be here without you.

I would be remiss without now acknowledging the mentorship, and great friendship, which I have found among the members of the Bowman Lab. Particularly, Dr. Avishek Dutta, Dr. Jesse Wilson, Dr. Kaycie Lanpher, Elizabeth Connors, Natalia Erazo, Benjamin Klempay, Riley Hale, Matt Heron, and Caitlyn Webster – thank you for your assistance, both in the lab and in life. Additionally, a huge thank you to the master and undergraduate students without who's assistance in the lab Chapter 3 would never have been completed on time – Ruth Varner, Shannon Perry, Grace Wang, Abby Coker, and Sandy Nguyenphuoc. Truly, every member of the Bowman Lab (former, current, and visiting) has been a delight to work with and learn from. You are truly 'lab fam' and I am both proud and grateful to have met you.

There are many others at Scripps to acknowledge for making the past 6 years so impactful (and fun!). Thank you to the members of the 2018 PhD cohort, your support and friendship has meant so much during this time and I hope our paths continue to cross both inside and outside the academic world. Thank you to my peer mentor, Dr. Ashlyn Giddings, for getting me on my feet at Scripps and introducing me to some of the most wonderful people in San Diego. Thank you to past and present coordinators, and all volunteers of the SCOPE program – working with you on science communication has been a highlight of my PhD. Thank you to Cheryl Peach and Donna Shabkie in the Scripps Department for supporting these endeavors in education and outreach. Thank you to the Scripps Polar Center and fellow organizers and regular attendees of the Scripps Polar Hour for your interdisciplinary perspectives and commitment to making the polar sciences a welcoming and inclusive field. Thank you to the UCSD Teaching and Learning Commons for supporting my growth as an educator during my time here. And thank you to other scientific collaborators Tammy Russell (Penguano Project) and Katie Harding (RV *Sikuliaq* cruise 2023). These projects may not be included in this dissertation, but I am glad to have worked with you in the last year(s) and can't wait to see where this work leads.

Most of this dissertation work was completed as part of the MOSAiC Science Consortium, or as a direct result of the MOSAiC International Arctic Drift Expedition (2019-2020) on the RV *Polarstern.* Thank you to all those who made the logistics of this endeavor possible as stated in the extended MOSAiC acknowledgement (Nixdorf et al. 2021). The MOSAiC community has been instrumental in supporting this scientific work, moreover they have been instrumental in supporting my growth as a scientist and a human. Six months is a

long time for your first time at sea and I could not have asked for a more welcoming place to train, better role models to learn from, or more wonderful people to be with. Thank you.

I would also like to acknowledge those who have been with me the longest. I first told my parents I wanted to be an oceanographer when I was 12 years old, and they happily cheered me on in this endeavor ever since. While Darcy and Rick Chamberlain are no longer with us, I am grateful of their legacy and their lessons in self-determination which have led me to this point. I am further lucky to have a remarkable set of close friends and extended family whose wisdom and companionship have shaped me into who I am. Thank you for your never-ending love and support over the years. And thank you CJ Collins, my partner of the last 11, for (along with the cats) providing me an amazing home to return to each day.

Finally, this dissertation work would not have been possible without the many people who contributed to it, as well as funding from a variety of sources. Most prominently, the National Science Foundation (NSF) grant NSF OPP 1821911 awarded to PI Dr. Jeff Bowman. My time at Scripps Institution of Oceanography was additionally funded by an NSF Graduate Research Fellowship (NSF GRFP), the San Diego Cota Robles Fellowship, and the Bob and Betty Beyster Fellowship through my role as a co-coordinator of SCOPE. And thank you to my co-authors, as listed in the following paragraphs. Through the many drafts and presentations which it has taken to get here, I have learned so much about science and scientific writing from you all that I will take with me in all projects from here on out.

Chapter 1, in full, is a reprint of the material as it appears in Elementa: Science of the Anthropocene 2022. Chamberlain, Emelia J; Balmonte, John Paul; Torstensson, Anders; Fong, Allison A; Snoeijs-Leijonmalm, Pauline; Bowman, Jeff., University of California Press, 2022. The dissertation author was the primary investigator and author of this paper.

Chapter 2, in part, has been submitted for publication of the material as it may appear in Limnology & Oceanography, 2024, Chamberlain, Emelia J; Rokitta, Sebastian; Rost, Björn; D'Angelo, Alessandra; Creamean, Jessie; Loose, Brice; Ulfsbo, Adam; Fong, Allison A; Hoppe, Clara JM; Droste, Elise; Nomura, Daiki; Schulz, Kirstin; Bowman, Jeff., Wiley Periodicals LLC, 2024. The dissertation author was the primary researcher and author of this paper.

Chapter 3, in part is currently being prepared for submission for publication of the material. Chamberlain, Emelia J; Webb Alison; Müller, Oliver; Hoppe, Clara JM; Fong, Allison A; Droste, Elise; D'Angelo Alessandra; Varner, Ruth; Smith, Madison M; Nomura Daiki; Kraberg Alexandra; Rost Björn; Schäfer, Hendrik; Loose, Brice; Bowman, Jeff. The dissertation author was the primary researcher and author of this material.

VITA

2018 Bachelor of Arts in Environmental Science and Marine Science, Boston University

2022 Master of Science in Oceanography, University of California San Diego

2023 Doctor of Philosophy in Oceanography, University of California San Diego

PUBLICATIONS

Smith, M.M., Angot, H., **Chamberlain, E.J.**, Droste, E., Karam, S., Muilwijk, M., Webb, A., Archer, S., Beck, I., Blomquist, B., Bowman, J., Boyer, M., Bozzato, D., Chierici, M., Creamean, J., D'Angelo, A., Delille, B., Fer, I., Fong, A.A., Fransson, A., Fuchs, N., Gardner, J., Granskog, M., Hoppe, C., Hoppema, M., Hoppmann, M., Mock, T., Muller, S., Müller, O., Nicolaus, M., Nomura, D., Petäjä, T., Salganik, E., Schmale, J., Schmidt, K., Schulz, K., Shupe, M., Stefels, J., Thielke, L., Tippenhauer, S., Ulfsbo, K.A., van Leeuwe, M., Webster, M., Yoshimura, M., & Xhan, L. (2023). Thin and transient meltwater layers and false bottoms in the Arctic sea ice pack — Recent insights on these historically overlooked features. *Elementa: Science of the Anthropocene, 11*(1), 00025. http://doi.org/10.1525/elementa.2023.00025

Nomura, D., Kawaguchi, Y., Webb, A., Li, Y., Dall'osto, M., Schmidt, K., Droste, E.S., **Chamberlain, E.J.**, Kolabutin, N., Shimanchuk, E., Hoppmann, M., Gallagher, M.R., Meyer, H., Mellat, M., Bauch, D., Gabarró, Smith, MM., Inoue, J., Damm, E., & Delille, B. (2023). Meltwater layer dynamics in a central Arctic lead: Effects of lead width, re-freezing, and mixing during late summer. *Elementa: Science of the Anthropocene*, *11*(1), 00102. http://doi.org/10.1525/elementa.2022.00102

Salganik, E., Katlein, C., Lange, B.A., Matero, I., Lei, R., Fong, A.A., Fons, S.W., Divine, D., Oggier, M., Castellani, G., Bozzato, D., **Chamberlain, E.J.**, Hoppe, C.J.M., Müller, O., Gardner, J., Rinke, A., Simões Pereira, P., Ulfsbo, A., Marsay, C., Webster, M., Maus, S., Høyland, K.V., & Granskog, M.A. (2023). Temporal evolution of under-ice meltwater layers and false bottoms and their impact on summer Arctic sea ice mass balance. *Elementa: Science of the Anthropocene, 11*(1), 00035. http://doi.org/10.1525/elementa.2022.00035

Chamberlain, E.J, Balmonte, J.P., Torstensson ,A., Fong, A.A., Snoeijs-Leijonmalm, P., & Bowman, J.S. (2022). Impacts of sea ice melt procedure on measurements of microbial community structure. *Elementa: Science of the Anthropocene*, *10*(1), 00017. http://doi.org/10.1525/elementa.2022.00017

Mock, T., Boulton, W., Balmonte, J.P., Barry, K., Bertilsson, S., Bowman, J.S., Buck, M., Bratbak, G., **Chamberlain, E.J.**, Cunliffe, M., Creamean, J., Ebenhöh, O., Eggers, S.L., Fong, A.A., Gardner, J., Gradinger, R., Granskog, M.A., Havermans, C., Hill, T., Hoppe, C.J.M., Korte, K., Larsen, A., Müller, O., Nicolaus, A., Oldenburg, E., Popa, O., Rogge, S., Schäfer, H., Shoemaker, K., Snoeijs-Leijonmalm, P., Torstensson, A., Valentin, K., Vader, A., Barry,

K., Chen, I.M.A., Clum, A., Copeland, A., Daum, C., Eloe-Fadrosh, E., Foster, B., Grigoriev, I.V., Huntemann, M., Ivanova, N., Kuo, A., Kyrpides, N.C., Mukherjee, S., Palaniappan, K., Reddy, T.B.K., Salamov, A., Roux, S., Varghese, N., Woyke, T., Wu, D., Leggett, R.M., Moulton, V., & Metfies, K. (2022). Multi-omics in the central Arctic Ocean for benchmarking biodiversity change. *PLOS Biology*, *20*(10), e3001835. http://doi.org/10.1371/JOURNAL.PBIO.3001835

Wilson, J.M., Erazo, N., Connors, E., **Chamberlain, E.J.,** Clements, S.M., Carter, M.L., Smith, J.E., & Bowman, J.S. (2022). Substantial microbial community shifts in response to an exceptional harmful algal bloom in coastal Southern California. *Elementa: Science of the Anthropocene*, *10*(1), 00088. http://doi.org/10.1525/elementa.2021.00088

Creamean, J.M., Barry, K., Hill, T.C.J., Hume, C., DeMott, P.J., Shupe, M.D., Dahlke, S., Willmes, S., Schmale, J., Beck, I., Hoppe, C.J.M., Fong, A., **Chamberlain, E.J.**, Bowman, J., Scharien, R., & Persson, O. (2022). Annual cycle observations of aerosols capable of ice formation in central Arctic clouds. *Nature Communications*, *13*, 3537. http://doi.org/10.1038/s41467-022-31182-x

Nesbit, K.T., Wolfe, W., Mullane, K.K., **Chamberlain, E.J.**, Rasina, B.A., Saberski, E., & Jones-Cervantes, S. (2021). SCOPE: A volunteer-led STEM outreach program connecting communities to research. *Connected Science Learning 3*(6). https://www.nsta.org/connected-science-learning/connected-science-learning-november-december-2021/scope-volunteer-led

Wilson, J.M., **Chamberlain, E.J.**, Erazo, N., Carter, M.L., & Bowman, J.S. (2021). Recurrent microbial community types driven by nearshore and seasonal processes in coastal Southern California. *Environmental Microbiology*, *26*(6), 3225-3239. http://doi.org/10.1111/1462-2920.15548

Chamberlain, E.J., Christ, A.J., & Fulweiler, R.W. (2021). Influence of Late Holocene climate on Lake Eggers hydrology, McMurdo Sound. *Antarctic Science*, *33*(2), 217-229. http://doi.org/10.1017/s0954102021000018

Scavo Lord, K., Burmester, E., Lesneski, K., McPherson, K., Cheung, B., **Chamberlain, E.**, Ewa, R., Fortunato, B., Inge, M., Jacomo, N., Kunis, C., Mather, B., McCarthy, G., McGinnis, L., Morey, J., O'Connor, O., Soukup, J., Sullivan, J., Taylor, J., Velandia, N., Finnerty, F., & Finnerty, J.R. 2018. Using Bioindicator Species to Characterize Distinct Mangrove Habitats on Turneffe Atoll, Belize. REPORT to Belize Fisheries Department: Research Permit #000056-17 https://open.bu.edu/handle/2144/38223

FIELD OF STUDY

Major Field: Biological Oceanography
Studies in Polar Sciences
Professor Jeff Bowman

ABSTRACT OF THE DISSERTATION

Microbial contributions to metabolic oxygen turnover in the central Arctic Ocean

by

Emelia J Chamberlain

Doctor of Philosophy in Oceanography

University of California San Diego, 2023

Professor Jeff Bowman, Chair

The metabolic balance between photosynthesis and respiration imposes a biological control on the ocean's ability to sequester carbon dioxide within the global carbon cycle. In the Arctic, however, the net trophic state of this balance is still uncertain due to sporadic regional data availability, methodological measurement biases, and poor mechanistic constraints on various physical and biological influences. In this dissertation, I used a combination of field observations and experimental approaches to identify mechanistic links between microbial (phytoplankton, bacterial, and archaeal) community structure and biogeochemical oxygen

cycling, resolving uncertainties in the biological influences on net metabolic balance in the surface Arctic Ocean. First, I experimentally investigated the inherent biases of sample melting procedures on measurements of first year sea ice microbial community structure, such as cell abundance, photo-physiology, and community composition. I found that an equivalent-volume 35-ppt buffered melt was best at minimizing cell loss due to osmotic stress, but various buffers were enough to reduce bias in amplicon sequence-based community composition when comparing to direct melting procedures. Next, I explored direct linkages between water column microbial community structure and metabolic balance using nearly a year's worth of amplicon sequencing and biological oxygen utilization observations collected during the MOSAiC International Arctic Drift Expedition (2019-2020). Applying predictive machine learning techniques to this paired timeseries, I identified both general successional patterns and specific taxa within the microbial community that tracked, and could reliably predict, biological oxygen consumption ($\pm$ 5.32 μmol kg^{-1}). Finally, I characterized fine-scale environmental controls on near-surface microbial community structure and metabolic oxygen turnover during the Arctic melt season. Observations from the MOSAiC Expedition indicated that meltwater induced salinity stratification results in ecologically distinct microhabitats with unique microbial assemblages and significant variation in net biological oxygen consumption. Culture-based laboratory experiments further showed that while primary production decreased significantly with exposure to low salinity conditions, community respiration rates persisted indicating increased likelihood for net heterotrophy at the air-sea interface in stratified meltwater layers. Taken together, this dissertation highlights microbial contributions to surface ocean metabolic oxygen turnover and provides information to better study, and model, this key ecological process in a future Arctic.

INTRODUCTION

The central Arctic Ocean is currently undergoing rapid environmental change, with cascading effects on many components of the polar marine ecosystem. Over the last few decades, amplified climate warming (Rantanen et al. 2022) has led to drastic reductions in sea ice thickness (Stroeve and Notz 2018) and extent (Meredith et al. 2019). Sea ice is a key component within the Arctic climate system, modulating the flow of energy, moisture, and gases between the ocean and atmosphere (Lannuzel et al. 2020; Nicolaus et al. 2022) and serving as an important habitat for microbial life (Boetius et al. 2015).

The prokaryotic (*i.e.*, bacteria and archaea) and eukaryotic (*i.e.*, phytoplankton and ice algae) microorganisms growing within and just beneath sea ice and are major contributors to polar biogeochemical cycles (Edwards et al. 2020; Campbell et al. 2022b). An estimated 2–24% of the total primary production in polar seas comes from sea ice associated algae (van Leeuwe et al. 2018). Some amount of this carbon is then remineralized through the microbial loop, consumed by higher trophic levels in the ice, water column, or benthos, or sequestered for long term storage in the deep ocean (Azam et al. 1983). Over long timescales, the total amount of net carbon export is approximately equivalent to the ecosystem's total net community production (NCP), or the metabolic balance between carbon fixation by photosynthetic autotrophs and aerobic respiration by heterotrophs (Volk and Hoffert 1985; Izett et al. 2021).

In situ, net metabolic balance is directly influenced by individual metabolic activity and growth efficiencies of the microbial community, with many taxonomic groups displaying variable primary production rates (Journal and Rao 1988; van Leeuwe et al. 2018), contrasting trophic strategies (Worden et al. 2015; Norris et al. 2021), and respiration rates (Munson-McGee et al., 2022). However, our current understanding of such functional metabolic diversity among the

central Arctic microbial community is still poorly constrained, as well as its ultimate impact on already highly uncertain estimates of net biological carbon drawdown in a changing Arctic (Lannuzel et al. 2020). Sea ice itself is a complex heterogenous matrix which can make physically sampling for microbial and biogeochemical parameters technically difficult (Miller et al. 2015; Campbell et al. 2022b). Additionally, long-term *in situ* biological measurements, particularly for the central Arctic, are scarce and seasonal ice cover prevents remote sensing methods from capturing key variables, i.e., NCP, within and below the ice (Willis et al. 2023). Moreover, concurrent environmental changes to both the physical and biological processes which influence metabolic balance adds complexity to disentangling the relative contributions of process-level mechanisms driving net carbon sequestration in the context of a changing Arctic.

In this thesis, I present work which aims to address these uncertainties and constrain specific microbial contributions to the central Arctic's net metabolic balance, here measured using biological oxygen turnover. In all chapters, prokaryotic and eukaryotic community structure was measured using 16S and 18S rRNA gene amplicon sequencing, respectively.

Chapter 1 quantifies the impacts of methodological biases on measurements of sea ice microbial community structure. Most biological analyses techniques are optimized for liquid samples and certain methodological melting approaches must be applied to conduct meaningful ecological research on sea ice microbial communities. In this chapter, I experimentally compared common buffering methods for sea ice melt procedures and analyzed the resulting biases on microbial measurements. This study offers recommendations for optimal melt approaches in sea ice research to the community and provides an example of the potential biases in such studies.

Chapter 2 examines predictive linkages between microbial community composition and biological oxygen utilization using a year-long paired timeseries from the central Arctic Ocean.

Leveraging machine learning to disentangle seasonal and spatial patterns in oceanographic taxonomic and metabolic community turnover, this work identifies key microbial predictors of metabolic oxygen balance in the surface Arctic Ocean and raises new questions regarding the role of heterotrophic microbes in regulating pelagic NCP and carbon sequestration.

Chapter 3 investigates the impact of seasonal influxes of freshwater from ice and snowmelt during Arctic summer on microbial community structure and biogeochemical cycling at the air-sea interface. Comparing observational data from stratified meltwater habitats to experimental results from lab-based cultures, I found that pooling meltwater on the ocean's surface contained a unique, metabolically active microbial community which was ecologically distinct from, and significantly more net heterotrophic than, the immediately underlying seawater.

Overall, this research emphasizes the importance of microbes, particularly heterotrophic microbes like bacteria and archaea, in regulating net ecosystem function. Bacteria account for most respiration in open ocean environments (Williams 1981; Rivkin and Legendre 2001), and play a critical, yet often underestimated, role in the attenuation of sinking particulate carbon flux under ice cover (Kellogg et al. 2011). How these processes will be impacted by climate change is a pressing research topic (Edwards et al. 2020; Campbell et al. 2022b) with implications far beyond the Arctic system (Cavicchioli et al. 2019). The results of this dissertation highlight the sensitivity of sea ice and Arctic Ocean diversity to changing environmental parameters and offer further evidence of the predictive linkages between community structure and function — providing key baseline information for modeling polar microbial communities under future climate conditions.

CHAPTER 1 – Impacts of sea ice melting procedure on measurements of microbial community structure

Abstract

Microorganisms play critical roles in sea ice biogeochemical processes. However, microbes living within sea ice can be challenging to sample for scientific study. Because most techniques for microbial analysis are optimized for liquid samples, sea ice samples are typically melted first, often applying a buffering method to mitigate osmotic lysis. Here, we tested commonly used melting procedures on three different ice horizons of springtime, first year, land-fast Arctic sea ice to investigate potential methodological impacts on resulting measurements of cell abundance, photophysiology, and microbial community structure as determined by 16S and 18S rRNA gene amplicon sequencing. Specifically, we compared two buffering methods using NaCl solutions ('seawater', melting the ice in an equal volume of 35-ppt solution, and 'isohaline', melting with a small volume of 250-ppt solution calculated to yield meltwater at estimated in situ brine salinity) to direct ice melting (no buffer addition) on both mechanically "shaved" and "non-shaved" samples. Shaving the ice shortened the melting process, with no significant impacts on the resulting measurements. The seawater buffer was best at minimizing cell lysis for this ice type, retaining the highest number of cells and chlorophyll *a* concentration. Comparative measurements of bacterial (16S) community structure highlighted ecologically relevant subsets of the community that were significantly more abundant in the buffered samples. The results for eukaryotic (18S) community structure were less conclusive. Taken together, our results suggest that an equivalent-volume seawater-salinity buffered melt is best at minimizing cell loss due to osmotic stress for springtime Arctic sea ice, but that either buffer will reduce bias in community composition when compared to direct melting. Overall, these findings indicate potential methodological biases that

should be considered before developing a sea ice melting protocol for microbiological studies and afterwards, when interpreting biogeochemical or ecological meaning of the results.

Introduction

Sea ice is a critical component of polar marine ecosystems. Prokaryotic and eukaryotic microorganisms growing in sea ice, particularly at the ice-seawater interface, are significant contributors to polar biogeochemical cycles (Vancoppenolle et al. 2013). Primary production from ice algae is responsible for an estimated 2–24% of the total primary production in sea ice covered regions, where bottom ice can contain up to 10 times more biomass carbon than the underlying seawater (van Leeuwe et al. 2018). The unique ecology of these microbial communities derives in part from the nature of sea ice. When ice forms from seawater, impurities such as salt and particles are excluded from the growing ice crystals, creating a brine that increases in salinity as temperature decreases and the ice continues to grow. This brine accumulates between growing ice lamellae at the ice-seawater interface, forming a habitat for the sea ice microbial community within the ice matrix (Worster and Wettlaufer, 1997; Grossi and Sullivan, 1985; Junge et al., 2004). As air temperature drops and sea ice grows, the ice develops strong vertical gradients in temperature and brine salinity, with pore spaces in the upper ice becoming colder and more saline (Worster and Wettlaufer, 1997; Grossi and Sullivan, 1985). These conditions impose considerable stress on microbial communities. However, sea ice microbial communities appear to be well adapted to such conditions, with a percentage of the community remaining metabolically active even through the polar night (Junge et al. 2004).

The inherently stratified and heterogeneous nature of the sea ice matrix presents a challenge for sampling internal brine communities. Many methods commonly applied to sea ice microbial communities are designed for liquid samples and require melting the ice before analysis (Miller et

al. 2015). Because the bulk salinity of sea ice (melted ice + brine) is lower than the brine salinity of the pore spaces inhabited by the sea ice microbial community, melting can have undesired effects on cell physiology and morphology, including cell lysis due to osmotic stress. Simply melting the ice ("direct" melt) has been shown to result in 13–97% eukaryotic cell loss, with the highest losses in ciliates and flagellates (Garrison and Buck, 1986; Mikkelsen and Witkowski, 2010), and up to 55% bacterial cell loss in winter ice (Ewert et al., 2013). These losses can alter measurements of dissolved inorganic nutrients and other parameters. Recent work identified a significant difference in dissolved ammonium concentration between buffered and direct melting, which the authors hypothesized to be the result of lysis of ammonium-rich flagellates and ciliates (Roukaerts et al. 2019). An alternative explanation to increased ammonium from direct melting is stimulated bacterial respiration of nitrogen-containing compatible solutes, abundantly released from marine microorganisms during sea ice melt (Firth et al. 2016; Torstensson et al. 2019). The relevant explanation may be influenced by ice-melting time, ice community composition, and sampling location (depth) within the sea ice matrix.

To combat these biases when melting sea ice, several techniques have been implemented by the scientific community to control the thermohaline properties of ice during the melting process and minimize osmotic stress. In addition to direct melt methods, common protocols include melting with "seawater" or "isohaline" buffers. A seawater-buffered melt involves melting an ice section in an equivalent (or larger) volume of either filtered seawater or a NaCl solution prepared to near-oceanic salinity (i.e., 34–35 ppt). An isohaline-buffered melt involves melting an ice section in a specific volume of a more concentrated or hypersaline NaCl solution, so that the resulting meltwater has a salinity similar to the estimated internal brine salinity for that ice horizon

(Miller et al. 2015). Internal brine salinities for sea ice are calculated using the measured parameter of ice temperature (Cox and Weeks, 1983).

Mechanical ice crushing, e.g., using a hammer on ice core sections in a plastic bag covered by a towel, is sometimes implemented in sea ice melt protocols to reduce melting time (Junge et al. 2004). A shorter melting time is thought to help minimize shifts in community structure during the melting period; rapid melting also appears to reduce impacts on phytoplankton photophysiology (Campbell et al. 2019). We explored an alternative option for reducing the melting time, using a commercial ice shaver, commonly used in the food and beverage industry for producing ice-cold drinks, to shave the sea ice core sections.

The general recommendation is that ice samples should be melted at low temperature using a buffer compatible with the intended biological analysis. Yet, there is little consensus in the literature on which exact melting procedures to employ, making comparing measurements across studies difficult (Miller et al., 2015) and increasing the complexity of designing field campaigns. Recently, Campbell et al. (2019) found that variations in melt protocols can account for over 60% of the variability within Arctic primary production estimates. Additionally, while biases in phytoplankton community structure due to melt procedures have been well documented (Garrison and Buck, 1986; Mikkelsen and Witkowski, 2010), there has been less effort to quantify the effects on bacterial community structure. By measuring both prokaryotic and eukaryotic communities simultaneously, our data provide a broader perspective to fully understand the impacts and inherent biases of melt procedures on analyses of microbial abundance and community structure in sea ice.

We tested the direct, seawater, and isohaline melt procedures on shaved and non-shaved land-fast spring sea ice collected from the Chukchi Sea to quantify their impacts on microbial (eukaryotic and bacterial) abundance, photophysiology, and community structure. Specifically, we

sought to identify which melting method was best at minimizing cell loss and whether specific bacterial (16S) or eukaryotic (18S) rRNA gene phylotypes were differentially impacted by the melt procedures. The results of this experiment provide a basis for recommending optimal melt approaches in sea ice research and offer an example of potential biases in such studies.

Methods

Site description and ice core processing

A total of 18 first-year sea ice cores were collected from Chukchi Sea land-fast ice approximately 400 m northwest of the Iḷisaġik College (Naval Arctic Research Laboratory campus) in Utqiaġvik, AK (Figure 1-1) on April 10, 2019. The cores were collected approximately 15 cm apart with a 9-cm diameter Kovacs ice corer and sectioned in the field using an ethanol-sterilized hand saw. Snow depth was consistent across the sampling site (5 cm) and total ice thickness was measured at 68 cm. Air temperature at the time of sampling was −11ºC, and ice temperature was measured immediately after collection of the first core at the top and bottom of the core using a thermistor probe. A linear model fit was then used to estimate temperatures throughout the core (ice-seawater interface: −1.8ºC, top: –8.2ºC). For all cores, the bottom 50-cm portion of the ice was sectioned into three distinct depth horizons: 0–10 cm (midpoint temperature: –2.3ºC), 10–30 cm (midpoint temperature: –3.7ºC), and 30–50 cm (midpoint temperature: –5.6ºC), where the ice-seawater interface is 0 cm. Sections were sealed in sterile bags (Whirl-Pak™) before transport in a light-shielded cooler to the Barrow Arctic Research Center/Environmental Observatory, with approximately 2.5 hours between collection and the completion of shaving and buffer addition.

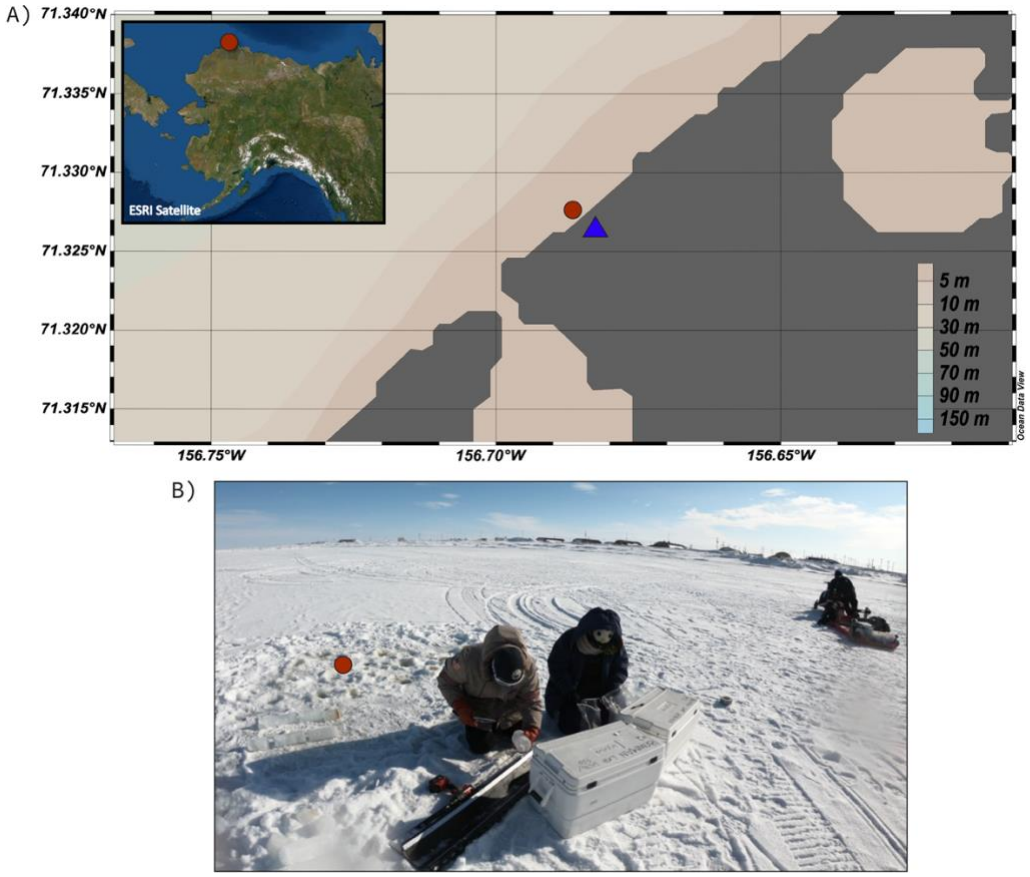


Figure 1-1. Map and images of sampling site.
Ice cores were collected from springtime land-fast sea ice near Utqiaġvik, AK, at a coring location (red circle, all panels) located approximately 400 m northwest of Iḷisaġik College, Naval Arctic Research Laboratory campus (blue triangle, panel A). (A) Map created in Ocean Data View version 5.6.2 (Schlitzer, 2022), where dark grey indicates land and shades of brown indicate ocean bathymetry (IBCAO; Jakobsson et al., 2012); inset created using ESRI satellite imagery in QGIS version 3.18.3 (QGIS.org, 2021). (B) Sampling site determined by field notes and images, one example of which is provided, from the date of sampling. Photo credit: JS Bowman.

Procedural variations in ice-melting method (direct melting and two buffering methods) were assessed for shaved and non-shaved ice sections across all ice horizons (Figure 1-2). First, half of the cores (9 cores, 27 horizons) were shaved using an industrial ice shaver (Swan SI-100E) rinsed with Milli-Q water. Then, shaved and non-shaved cores were split evenly among the three melting methods (6 cores, 18 total horizons per method; Figure 1-2). Samples in the "direct melt" (DM) method received no addition of buffer for an approximate final melt volume of 700 mL in the 0–10 cm ice horizon and 1300 mL in the 10–30 cm and 30–50 cm horizons. In the "isohaline

melt" (IM) method, a volume of autoclaved 250 ppt reagent grade NaCl (Fisher Scientific™) solution (stored at 5ºC) was added to each sample to reach a final melt salinity matching the estimated brine salinity for each ice horizon: 43 ppt, 67 ppt, and 95 ppt, respectively. Brine salinities were estimated as a function of ice temperature. Volumes for brine addition were calculated using the equations outlined in Cox and Weeks (1983) as a function of ice temperature and density (calculated from bulk ice salinity and temperature; Table 1-S1). Bulk ice salinity was measured using a refractometer on a directly melted subsample of the midpoint of each ice section prior to further treatment. Specifically, a 5 cm, 20 cm, and 40 cm subsample were collected into individual 50 mL falcon tubes from a representative ice core at the start of coring and the small volumes allowed for rapid melt at room temperature. Samples in the "seawater melt" (SM) method were weighed on a balance and an equivalent mass of autoclaved 35-ppt NaCl solution (stored at 5ºC) was added, reaching a volumetric dilution factor of 2.06. Approximate final melt volumes were 1300 mL for 0–10 cm ice sections and 2700 mL for 10–30 cm and 30–50 cm sections. Following processing at room temperature, laboratory lighting was turned off and samples were left to melt in their original sample bags for approximately 5 hours before bags were moved into dark refrigerators (1–5ºC) to finish melting overnight. Samples were observed over the course of the next morning (April 11, 2019) and subsampled for the following measurements as quickly as possible following the complete melt of each section.

Cell abundance and photophysiology

Samples for cell abundance (flow cytometry) and photophysiology (active fluorescence) measurements were collected at room temperature in ambient laboratory lighting. After gentle mixing, an aliquot of 4 mL of meltwater was transferred to a polycarbonate cuvette and run immediately on an AquaFlash Handheld Active Fluorometer (Turner Designs™) to measure

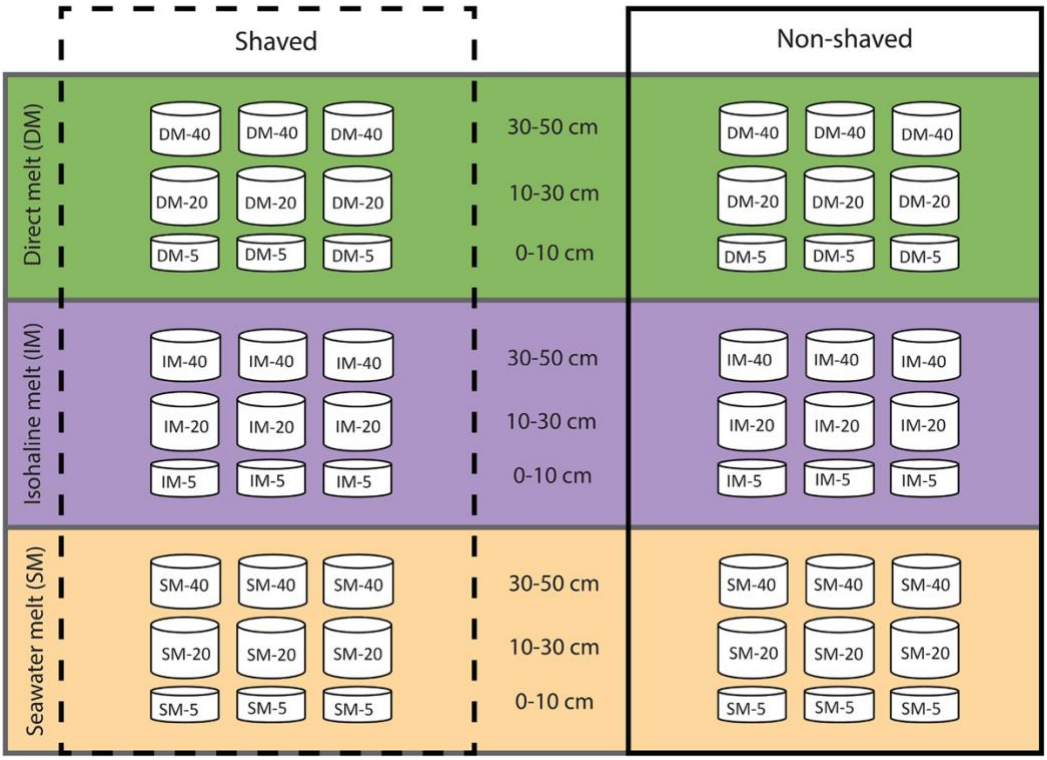


Figure 1-2. Schematic of the experimental set-up.
The bottom 50 cm of 18 ice cores were sectioned into 3 horizons (0–10 cm, 10–30 cm, and 30–50 cm, labeled as 5, 20 and 40, respectively), which were distributed evenly among treatments. Half of the ice sections (27 of 54, indicated by the dashed box) were shaved using an industrial ice shaver. Eighteen sections (9 shaved and 9 non-shaved) were used in each of the 3 melting methods. Direct melt (DM) sections were melted with no osmotic buffer added; isohaline melt (IM) sections were melted in an individualized volume of 250 ppt NaCl solution to reach a final salinity matching estimated internal brine salinities; and seawater melt (SM) sections were melted in an equivalent mass of 35 ppt NaCl solution. All cores were processed further in the same manner.

chlorophyll *a* concentration (µg L^{-1}) and effective quantum yield according to the following equation, where F_m represents the fluorescence maximum and F_0 represents the fluorescence minimum:

$$\frac{\Delta F}{F_m'} = \left(\frac{F_m - F_0}{F_m}\right)$$

We aliquoted 2 mL of meltwater for measurements of cell abundance using a flow cytometer. These samples were stored in plastic cryovials, fixed to a final concentration of 0.25% glutaraldehyde, and shipped frozen on liquid nitrogen to Scripps Institution of Oceanography (SIO) at the University of California San Diego. At SIO, they were stored at −20ºC until processed

(April 23, 2019) on a Cyflow Space flow cytometer (Sysmex America Inc.™). Total cell abundance was measured by staining 1 mL of sample with the nucleic acid stain SYBR-green (Molecular Probes™) which binds to all DNA. Additionally, autofluorescent cell abundance was measured by running the remaining 1 mL of unstained sample to count only the photosynthetic, chlorophyll-containing microbial community. Autofluorescent cell abundance was measured in addition to total cell abundance to assess whether the photosynthetic population was any more or less sensitive to osmotic stress and buffer type than the rest of the microbial community. Forward scatter, side scatter, and fluorescence channels FL1 (ex/em: 488/536 nm), FL2 (ex/em: 488/590 nm), FL3 (ex/em: 488/675nm), FL4 (ex/em: 488/748nm), FL5 (ex/em: 405/455 nm), and FL6 (ex/em: 638/748 nm) were used to characterize and count all (DNA-containing) and autofluorescent (chlorophyll-containing) cells. A standard volume of 1:10 diluted 123-eCount beads (Fisher Scientific™) were used to calculate absolute cell counts. Both absolute cell counts (autofluorescent and total cells) were then normalized to their respective melt-method dilution factors: 1 for the direct-melt method, 2.06 for the seawater-melt method, and between 1.2 and 1.6 for the isohaline-melt method (Table 1-S1).

DNA collection, extraction, and sequencing

Immediately following subsampling for cell abundance and photophysiology, a volume of 500–2000 mL (section- and biomass-dependent) of meltwater was filtered through a sterile 0.2 μm Supor membrane disc filter (Pall Corporation), using a peristaltic pump, for subsequent 16S/18S rRNA gene amplicon sequencing. Tubing and filter housing were flushed with Milli-Q between each melting method/core section combination. Filters were stored at –20ºC until transport to SIO on liquid nitrogen. At SIO, filters were stored at –80ºC until extraction on June 25, 2019. Extractions were performed on a KingFisher™ Flex Purification system using a MagMax

Microbiome Ultra Nucleic Acid Extraction kit (ThermoFisher Scientific). Amplicon library preparation and sequencing were conducted at the Argonne National Laboratory using a 2 x 150 bp library architecture on the Illumina MiSeq platform. Universal primers 515F and 806R (16S, V4; Walters et al., 2016) and 1380F, 1389F, and 1510R (18S, V9; Amaral-Zettler et al., 2009) were used for rRNA gene amplification. Illumina reads were denoised and merged using the R package dada2 (Callahan et al. 2016). The resulting amplicon sequence variants (ASVs) were used as input for the paprica v0.7.0 pipeline, a phylogenetic placement approach for determining community structure and metabolic potential (Bowman and Ducklow, 2015; Erazo et al., 2021; https://github.com/bowmanjeffs/paprica).

For Bacteria and Archaea, paprica places sample reads onto a reference tree created from the concatenated 16S and 23S rRNA genes from all completed genomes in the Genbank RefSeq database (Haft et al. 2018). Each ASV is placed to either an internal or terminal branch on the tree (Nawrocki and Eddy 2013; Barbera et al. 2019; Czech et al. 2020). Taxonomically, ASVs assigned to terminal completed genomes are described by their complete taxonomic name, while ASVs assigned to internal estimated genomes are described by the consensus taxonomy for all downstream nodes. “Placement proportions” represent the percentage of times a unique ASV was assigned to a specific terminal branch. ASVs with lower placement proportions were placed to multiple nodes, indicating lower confidence in the final classification. Regardless, the taxonomic associations derived from terminal placements should not be taken literally; our use here indicates only that the association is the closest relative among genomes in RefSeq. The point of placement in the reference tree is then used to estimate genome size, 16S rRNA gene copy number, and GC content (Bowman and Ducklow 2015; Karp et al. 2021). The 16S rRNA gene copy number was used to normalize the abundance data and combined with total cell counts from flow cytometry

data to calculate normalized read counts for each ASV. For Eukarya, paprica places ASVs onto a reference tree created from the unique 18S rRNA genes present in the PR2 4.13.0 database (Guillou et al. 2013). The 18S rRNA gene ASVs are represented by relative abundance only and were not normalized to 18S rRNA gene copy. ASVs with an abundance of 1, samples with bad library builds (<5000 total ASV counts), and suspected mitochondria/chloroplasts were filtered out prior to all further downstream analyses.

Statistical analyses

All statistical analyses were carried out using the R programming language in R Studio (R Core Team 2023). To assess differences between melting methods or ice shaving on cell abundance (log_{10} transformed), chlorophyll *a* content (log_{10} transformed), effective quantum yield, and microbial diversity (inverse Simpson Index), we first ran a two-way crossed ANOVA with an error factor accounting for ice horizon using all samples collected (vegan package; Oksanen et al., 2015). When statistically significant differences were encountered for either of these procedural variations (with no significant interaction between them), or to further investigate fine scale variability within ice horizons (0–10 cm, 10–30 cm, and 30–50 cm), we then individually tested for significant differences at each horizon using the non-parametric Kruskal-Wallis test with Dunn pair-wise comparisons (vegan package; Oksanen et al., 2015). Only Kruskal-Wallis tests where significant differences were encountered in one or more ice horizon are presented here.

To assess impacts of melt procedures on microbial community structure, a non-metric multidimensional scaling (NMDS) ordination was run based on the Bray-Curtis dissimilarity matrix (Bray and Curtis 1957) of 16S/18S ASV read counts/relative abundances (phyloseq package; McMurdie and Holmes, 2013). A permutational analysis of variance (PERMANOVA) was then used to individually test for significant differences between melting methods, shaved and

non-shaved ice, and ice horizons within each dissimilarity matrix (999 permutations). Confidence in results was confirmed using a beta dispersion test (vegan package; Oksanen et al., 2015). All *p*-values were adjusted to correct for multiple comparisons using the Holm-Bonferroni method with initial significance cut-off of $\alpha = 0.05$ (Holm 1979).

We used the package DeSeq2 to assess if specific 16S or 18S rRNA gene phylotypes were differentially abundant between ice-melting methods (Love et al. 2014). We limited this analysis to procedures that significantly impacted community structure in the PERMANOVA results. DESeq2 calculates dispersion estimates using a parametric, dispersion-mean relation through a robust gamma-family generalized linear model. Size factors were estimated using the median ratio method to control for differences in library size, and the Wald test was used to determine log2fold changes in ASV abundances. *P*-values for significant ASVs were adjusted for false discovery rate using the Benjamini-Hochberg adjustment where values below 0.1 were considered significant (the default setting for DESeq2; Love et al., 2014). Methods for utilizing DeSeq2 on ASVs were adapted from Webb et al. (2019) and Erazo and Bowman (2021).

Results

In this section, comparisons between melting method within individual ice horizons are referred to as method-5, method-20, and method-40, where method indicates direct melt (DM), seawater melt (SM) or isohaline melt (IM) and numbers indicate the midpoint for the 0–10 cm, 10–30 cm, and 30–50 cm ice horizons, respectively. For example, samples from the 0–10 ice horizon that were buffered using the seawater-salinity melt method are designated ‘SM-5’ (Figure 1-2, Table 1-1).

Table 1-1. Procedural overview with averages and standard deviations of key parameters.

Ice horizon (cm)	Melting method (label)[a]	Ice shaving	AF abundance[a] (cells mL^{-1})	TC abundance[a] (cells mL^{-1})	Chlorophyll *a*[a] (μg L^{-1})	Effective quantum yield (%)
0–10	Direct melt (DM-5)	Shaved	$6.7 \pm 4.0 \times 10^4$	$2.2 \pm 0.88 \times 10^5$	131 ± 17	0.09 ± 0.09
		Non-shaved	$6.7 \pm 4.3 \times 10^4$	$2.6 \pm 0.26 \times 10^5$	151 ± 10	0.07 ± 0.01
	Isohaline melt (IM-5)	Shaved	$7.6 \pm 1.4 \times 10^4$	$3.1 \pm 0.70 \times 10^5$	151 ± 42	0.11 ± 0.05
		Non-shaved	$4.4 \pm 0.80 \times 10^4$	$2.3 \pm 0.69 \times 10^5$	180 ± 332	0.04 ± 0.06
	Seawater melt (SM-5)	Shaved	$3.9 \pm 1.3 \times 10^4$	$2.6 \pm 0.10 \times 10^5$	182 ± 22	0.15 ± 0.04
		Non-shaved	$7.7 \pm 3.6 \times 10^4$	$6.5 \pm 1.6 \times 10^5$	213 ± 40	0.04 ± 0.04
10–30	Direct melt (DM-20)	Shaved	$2.4 \pm 0.69 \times 10^3$	$2.4 \pm 0.78 \times 10^4$	12 ± 4	0.12 ± 0.07
		Non-shaved	$3.0 \pm 0.37 \times 10^3$	$2.7 \pm 0.46 \times 10^4$	9 ± 2	0.09 ± 0.07
	Isohaline melt (IM-20)	Shaved	$2.5 \pm 0.35 \times 10^3$	$3.2 \pm 0.62 \times 10^4$	11 ± 2	0.06 ± 0.03
		Non-shaved	$2.6 \pm 0.13 \times 10^3$	$3.4 \pm 1.0 \times 10^4$	18 ± 14	0.08 ± 0.08
	Seawater melt (SM-20)	Shaved	$3.9 \pm 1.5 \times 10^3$	$3.7 \pm 0.85 \times 10^4$	25 ± 9	0.16 ± 0.04
		Non-shaved	$2.7 \pm 0.27 \times 10^3$	$4.9 \pm 0.99 \times 10^4$	19 ± 4	0.10 ± 0.08
30–50	Direct melt (DM-40)	Shaved	$1.2 \pm 0.40 \times 10^3$	$1.8 \pm 1.0 \times 10^4$	7 ± 1	0.12 ± 0.09
		Non-shaved	$2.0 \pm 0.60 \times 10^3$	$2.7 \pm 0.49 \times 10^4$	14 ± 3	0.09 ± 0.05
	Isohaline melt (IM-40)	Shaved	$1.4 \pm 0.68 \times 10^3$	$1.4 \pm 0.49 \times 10^4$	10 ± 4	0.12 ± 0.09
		Non-shaved	$1.1 \pm 0.57 \times 10^3$	$2.2 \pm 1.1 \times 10^4$	11 ± 2	0.10 ± 0.06
	Seawater melt (SM-40)	Shaved	$6.1 \pm 2.4 \times 10^3$	$6.6 \pm 2.6 \times 10^4$	17 ± 4	0.11 ± 0.07
		Non-shaved	$3.0 \pm 0.52 \times 10^3$	$3.6 \pm 0.75 \times 10^4$	28 ± 16	0.08 ± 0.06

[a]Values are the average of meltwater measurements (n = 3) multiplied by the dilution factor for the buffered methods, which for seawater melt (SM) samples was 2.06 and for isohaline melt (IM) samples was 1.2 (IM-5), 1.3 (IM-20), and 1.6 (IM-40). AF indicates autofluorescent cells; TC, total cells

Cell abundance and photophysiology

Average total cell (TC) abundance and autofluorescent cell (AF) abundance did not vary significantly between shaved and non-shaved ice horizons (AF, $p = 0.797$; TC, $p = 0.167$; Table 1-2). However, average TC and AF abundance did vary significantly between melting methods (ANOVA test: TC, $p < 0.0001$; AF, $p = 0.0345$; Table 1-2), and Kruskal-Wallis analyses for each ice horizon revealed that melting method was a statistically significant factor only in the two uppermost sections of the ice column (Figure 1-3A and B, Table 1-3). At 10–30 cm, the SM method resulted in significantly higher TC abundances than the DM method (Figure 1-3A, Table

1-3), with mean TC abundance in the SM-20 samples 75% higher than DM-20 samples (4.3 x 10^4 and 2.6 x 10^4 cells mL^{-1}, respectively; Table 1-1).

At 30–50 cm, the SM method resulted in significantly higher TC and AF abundances than either the DM or IM methods (Figure 1-3A and B, Table 1-3). Mean AF abundance in SM-40 samples was 188% higher than DM-40 and 260% higher than IM-40 samples (Table 1-1). Mean TC abundance in SM-40 samples was 181% higher than IM-40 samples and 128% higher than DM-40 samples (Table 1-1).

Table 1-2. ANOVA summary statistics.

Parameter[a]	Melting method		Ice shaving		Interaction	
	***F* (2, xx)[b]**	***P adj*[c]**	***F* (1, xx)[b]**	***P adj*[c]**	***F* (2, xx)[b]**	***P adj*[c]**
AF abundance (xx = 48)	4.912	0.0345	0.067	0.7965	1.675	0.3960
TC abundance (xx = 48)	13.73	<0.0001	3.122	0.1672	0.550	0.5808
Chlorophyll *a* (xx = 46)	16.6	<0.0001	5.179	0.0552	0.460	0.6345
Effective quantum yield (xx = 52)	0.112	1.000	3.878	0.1629	0.165	1.000
Bacterial diversity (xx = 44)	0.555	1.000	0.000	1.000	0.295	1.000
Eukaryotic diversity (xx = 51)	3.489	0.114	2.345	0.264	0.975	0.384

[a]AF indicates autofluorescent cells; TC, total cells; parenthetic value, degrees of freedom.
[b]Two-way crossed analysis of variance (ANOVA), testing melting method (direct, isohaline, and seawater) and ice shaving (shaved, non-shaved) with an error factor to account for ice horizon.
[c]All *p*-values adjusted to account for multiple comparisons using the Holm-Bonferroni correction, $\alpha = 0.05$ (Holm 1979).

Based on the two-way ANOVA analysis, mean chlorophyll *a* concentration also varied significantly by melting method, but not by ice shaving ($p < 0.0001$ and $p = 0.0552$, respectively; Table 1-2). SM samples retained a significantly higher mean chlorophyll *a* concentration than DM samples across all ice horizons (41% in SM-5, 100% in SM-20, and 109% in SM-40; Figure 1-3C, Tables 1-1 and 1-3). Mean chlorophyll *a* concentration was also significantly higher (109%) in mean SM-40 samples than IM-40 samples (Figure 1-3C, Tables 1-1 and 1-3). Effective quantum yield showed no significant variation between the shave or melt treatments (Figure 1-3D, Table 1-2).

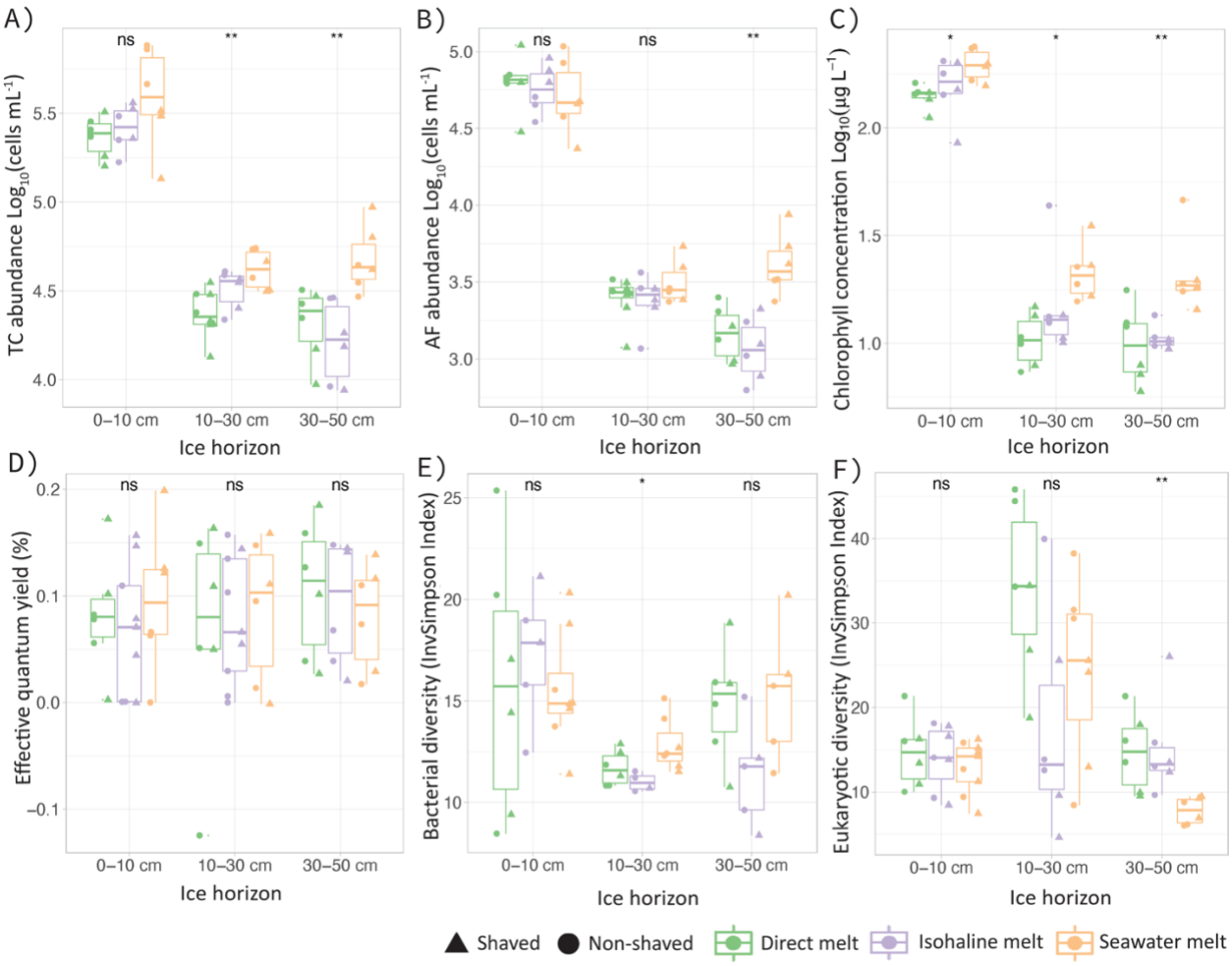


Figure 1-3. Microbial parameters by ice horizon, shaving, and melting method.
(A) Total cell (TC) abundance, (B) autofluorescent (AF) cell abundance, (C) chlorophyll a concentration, (D) effective quantum yield, (E) bacterial alpha diversity, and (F) eukaryotic alpha diversity are plotted by ice horizon and melting procedure, with individual measurements represented by triangles (shaved) or circles (non-shaved) and colored by melting method (direct, isohaline, or seawater melt). The midline of each box represents the median, below which is the first quartile and above, the third quartile. Whiskers represent the minimum and maximum of the spread, and outliers are marked by standalone points. Cell abundances and chlorophyll a concentrations were log-transformed, and alpha diversity was calculated using the inverse Simpson Index. Significance values (ns = $p > 0.05$, * = $p < 0.05$, ** = $p < 0.01$) represent a one-way non-parametric Kruskal-Wallis tests run for melting method on each individual ice horizon; p-values were adjusted to account for multiple comparisons using the Holm-Bonferroni correction, $\alpha = 0.05$ (Holm, 1979).

Microbial community structure

Very few archaeal phylotypes (1.6% of 16S ASVs) were detected following sequencing quality control, so the prokaryotic analysis of this study focuses solely on the bacterial community.

Initially, 42% of the bacterial community was identified as cyanobacteria/chloroplasts and 22% as mitochondrial DNA. After removing these ASVs, bacterial community structure was dominated by Proteobacteria (28% Gamma, 22% Alpha, 0.6% Beta) and Flavobacteriia (33%) taxa across all samples (including all melt procedures and ice horizons, shaved and non-shaved; Figure 1-4A). The parametric two-way crossed ANOVA analysis across all samples yielded no significant variation in bacterial alpha diversity among 16S ASVs within either melting methods or ice shaving (Table 1-2). However, when differences were assessed individually within ice core sections, the non-parametric Kruskal-Wallis and Dunn Pairwise Comparison showed significantly higher alpha diversity in SM-20 than in IM-20 (Figure 1-3E, Table 1-3). These differing results suggest that while there was significant variation between melting method in the 10–30 cm ice horizon, this trend is lost to noise when analyzing all samples in the dataset. While melting methods did not separate visually in the first two dimensions of the NMDS ordination (Figure 1-S1A), the PERMANOVA on the Bray-Curtis dissimilarity matrix of 16S ASV read counts revealed that dispersion among groups was significantly different across both melting method (p = 0.012) and ice horizon (p = 0.004; Table 1-4). This result indicates strong methodological variation in sequence-based bacterial community structure from melting approach. Confidence in PERMANOVA results was confirmed by the beta dispersion test (F = 0.662, Df = 2, p = 0.516).

Eukaryotic community composition across all samples was dominated by diatoms (29% Bacillariophyta), amoeboid Cercozoa (18% Filosa-Imbricatea, 10% Filosa-Thecofilosea), and Preaxostyla (17%; Figure 1-4B). Similar to the bacterial community, there was no significant difference in eukaryotic alpha diversity among 18S relative abundances within the melting methods or ice shaving in the two-way crossed ANOVA (Table 1-2). Yet, the non-parametric Kruskal-Wallis and Dunn Pairwise Comparison showed significantly lower alpha diversity in SM-

Table 1-3 Kruskal-Wallis summary statistics for pairwise comparisons of melting method.

Parameter	Test statistics	Ice horizon		
		0–10 cm (xx = 26, yy = 19, zz = 22)	**10–30 cm (xx = 24, yy = 17, zz = 19)**	**30–50 cm (xx = 18, yy = 18, zz = 18)**
AF abundance	KW[a], *F* (2, xx)	0.573	0.986	11.1
	KW[a], *P*	0.751	0.611	0.00389
	DN[b], *P adj <0.05*	NS[c]	NS	0.0221, SM–DM; 0.00513, SM–IM
TC abundance	KW[a], *F* (2, xx)	3.52	9.75	10.7
	KW[a], *P*	0.172	0.00765	0.00472
	DN[b], *P adj <0.05*	NS	0.00683, SM–DM	0.0347, SM–DM; 0.00513, SM–IM
Chlorophyll *a*	KW[a], *F* (2, xx)	7.03	9.06	10.2
	KW[a], *P*	0.0298	0.0108	0.00624
	DN[b], *P adj <0.05*	0.0242, SM–DM	0.00882, SM–DM	0.0148, SM–DM; 0.0148, SM–IM
Bacterial diversity	KW[a], *F* (2, yy)	1.04	6.94	4.46
	KW[a], *P*	0.594	0.0311	0.107
	DN[b], *P adj <0.05*	NS	0.0265, SM–IM	NS
Eukaryotic diversity	KW[a], *F* (2, zz)	0.708	5.04	11.4
	KW[a], *P*	0.702	0.080	0.00332
	DN[b], *P adj <0.05*	NS	NS	0.00738, SM–DM; 0.00985, SM–IM

[a]One-way, nonparametric Kruskal-Wallis (KW) tests run with melting method factors individually for each ice horizon. DM indicates direct melt; IM, isohaline melt; SM, seawater melt.
[b]*P*-values from Dunn pairwise comparisons (DN) adjusted to account for multiple comparisons using the Holm-Bonferroni correction, $\alpha = 0.05$ (Holm 1979). Only significant values reported.
[c]NS indicates no significant comparisons between melting methods.

40 than either DM-40 or IM-40, again indicating that while some significant variation does occur from melting method within ice horizons (Figure 1-3F, Table 1-3), it is not strong enough for significance when all samples are considered. Clusters by melting method were not identifiable visually in the NMDS ordination (Supplemental Figure 1-S1B), and the PERMANOVA on the Bray-Curtis dissimilarity matrix of transformed 18S ASV relative abundances identified no significant dispersion among groups for melting method ($p = 0.051$), only ice horizon ($p = 0.004$; Table 1-4). This result indicates little methodological variation in sequence-based eukaryotic community structure from melting approach. Confidence in PERMANOVA results was confirmed by the beta dispersion test (F = 0.5035, Df = 2, $p = 0.62$).

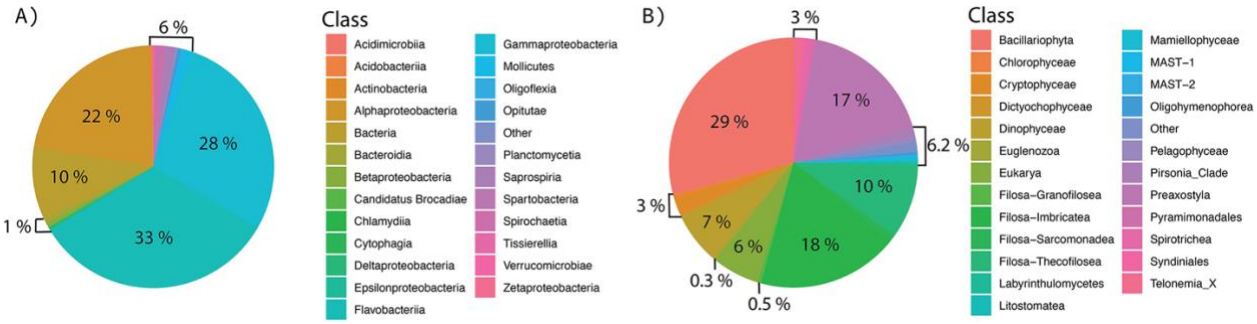


Figure 1-4. Overview of taxonomic diversity across all samples.
Percentage of unique amplicon sequence variants (ASVs) assigned as a completed, or most closely related to, genome within each taxonomic class across all samples (all treatments and ice horizons). (A) Bacterial class percentages calculated using read counts of each 16S ASV normalized to cell count. (B) Eukaryotic class percentages calculated using relative abundance of each 18S ASV.

Ice shaving was not a significant contributor to dispersion within the Bray-Curtis dissimilarity matrix of either bacterial ASV read counts or eukaryotic ASV relative abundances (Table 1-4) and were not considered in further analyses.

Table 1-4. PERMANOVA summary statistics.

Parameter	Melting method		Ice shaving		Ice horizon		Interaction 1:2	
	***F* (2, xx)[a]**	***P adj*[b]**	***F* (1, xx)[a]**	***P adj*[b]**	***F* (2, xx)[a]**	***P adj*[b]**	***F* (2, xx)[a]**	***P adj*[b]**
16S ASV dissimilarity matrix (44)	4.266	0.012	3.016	0.066	46.94	0.004	1.167	0.304
18S ASV dissimilarity matrix (51)	2.438	0.051	2.057	0.364	24.83	0.004	0.940	0.298

[a]Permutational multivariate analysis of variance performed on Bray-Curtis dissimilarity matrices of 16S amplicon sequence variant (ASV) read counts normalized to total cell count and of eukaryotic 18S ASV relative abundances.
[b]*P*-values adjusted using the Holm-Bonferroni correction to account for multiple comparisons, $\alpha = 0.05$ (Holm 1979).

Differential abundances within bacterial taxa

Our DeSeq2 analysis yielded many 16S phylotypes that were differentially abundant between melting methods. Comparisons between each method (SM–DM, IM–DM, and IM–SM) were run independently for each ice horizon, where positive log2-fold changes indicate ASVs that

were significantly ($p < 0.1$) more abundant in the first of the two compared methods and negative log2-fold changes indicate the opposite (Figure 1-S2). A total of 55 ASVs were identified as significantly differentially abundant when comparing the SM and DM methods. Only six of these ASVs were more abundant in DM samples (two belonging to class Spirochaetia, two to Flavobacteriia, one to Alphaproteobacteria, and one to Gammaproteobacteria; Figure 1-S2A and C), while 49 were more abundant in SM samples (26% Gammaproteobacteria, 21% Alphaproteobacteria, 12% Flavobacteriia, and the remaining 41% split among 10 other classes; Figure 1-S2A and C). When comparing IM and SM methods, 71 ASVs were identified as significantly differentially abundant, with 15 being more abundant in SM samples (27% Flavobacteriia, 20% Alphaproteobacteria, 13% Acidimicrobiia, and 40% among other classes; Figure 1-S2D and F) and 56 more abundant in IM samples (25% Gammaproteobacteria, 20% Flavobacteriia, 14% Alphaproteobacteria, 5% unidentified bacteria, and 36% among other classes; Figure 1-S2D–F). When comparing IM and DM methods, 90 ASVs were identified as significantly differentially abundant, with 4 being more abundant in DM samples (one each of Alphaproteobacteria, Flavobacteriia, Spirochaetia, and Mollicutes; Figure 1-S2G and I), and 86 more abundant in IM samples (27% Gammaproteobacteria, 20% Alphaproteobacteria, 14% Flavobacteriia, 5% unidentified bacteria, and 33% among other classes; Figure 1-S2G–I).

When considering all ice horizon and melting method comparisons (removing duplicate ASVs, meaning that the ASV was significant in more than one of the above comparisons), a total of 110 unique bacterial ASVs were identified as significantly differentially abundant (Table 1-S2). Figure 1-5 presents the 50 most abundant of these ASVs, or those with the highest read counts across all samples. These data indicate that the strongest variations in read count occurred in the 0–10 cm

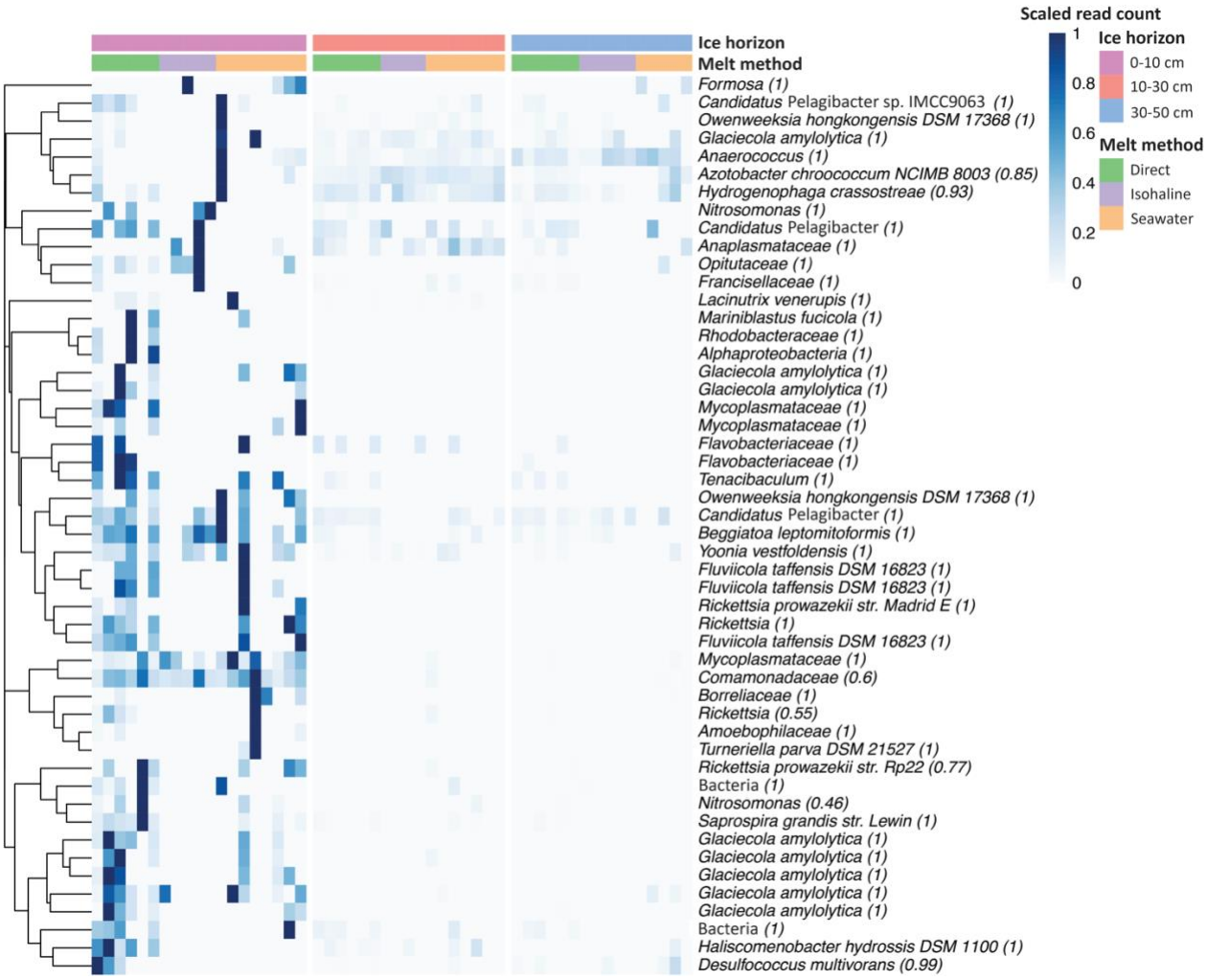


Figure 1-5. Significantly differentially abundant bacterial amplicon sequence variants (ASVs) between melting method within each ice horizon.
Cell count normalized 16S ASV read counts were scaled using a z-score between 0 and 1. Bray-Curtis dissimilarity distances were used to cluster taxa. Columns (samples) are colored by ice horizon and melting method. The 50 most abundant of significantly ($p < 0.1$) differentially abundant ASVs from each comparison are shown with duplicate ASVs (significant in more than one comparison) shown only once. The paprica taxonomic ID (closest relative among RefSeq or PR2 genomes) and placement proportion were used to identify each unique ASV (Bowman and Ducklow, 2015). Heatmaps were generated using R package pheatmap (Kolde, 2019).

ice horizon (Figure 1-5). Many of the significantly differentially abundant ASVs with high read counts across samples appeared more abundant in the buffered SM and/or IM samples than in the DM samples (Figure 1-5). Most of these ASVs had high placement proportions to reference genomes within Proteobacteria such as *Candidatus* Pelagibacter or Bacteroidetes such as

Owenweeksia hongkongensis (DSM 17368), *Fluviicola taffensis* (DSM16823) or other unidentified *Flavobacteriaceae* (Figure 1-5). Some ASVs appeared significantly more abundant in DM samples, placing mostly to a reference genome belonging to *Glaciecola amylolytica* (Figure 1-5). Significant differential abundances between IM-5 and SM-5 (Figure 1-S2F) were among less abundant ASVs or among ASVs placed to the same reference genome (Table 1-S2).

Differential abundances within eukaryotic taxa

Several 18S phylotypes yielded differential abundances between melting methods, with a large proportion of those more abundant in the buffered melts (SM or IM) belonging to classes within the division Cercozoa (i.e., Filosa-Imbricatea, Filosa-Thecofilosea, Filosa-Sarcomonadea, and Filosa-Granofilosea; Figure 1-S3). A total of 24 eukaryotic ASVs were identified as significantly differentially abundant when comparing the SM and DM methods, with 14 being more abundant in DM samples (36% Bacillariophyta, 36% Dinophyceae, and 28% other) and 10 more abundant in SM samples (50% Cercozoa, 30% Bacillariophyta, 20% other; Figure 1-S3A–C). When comparing IM and SM methods, 39 eukaryotic ASVs were identified as significantly differentially abundant, with 16 being more abundant in IM samples (varying classes) and 23 more abundant in SM samples (34% Bacillariophyta, 22% Cercozoa, 17% other; Figure 1-S3D–F). When comparing IM and DM methods, 80 eukaryotic ASVs were identified as significantly differentially abundant, with 38 more abundant in IM samples (13% Cercozoa, and other varying taxa) and 21 more abundant in DM samples (38% Dinophyceae, 33% Bacillariophyta, and 28% other; Figure 1-S3G–I).

When considering all ice horizon and melting method comparisons (again removing duplicates), a total of 81 unique 18S ASVs were significantly differentially abundant (Table 1-S3). Figure 1-6 presents the 50 most abundant of these significant eukaryotic ASVs across all samples.

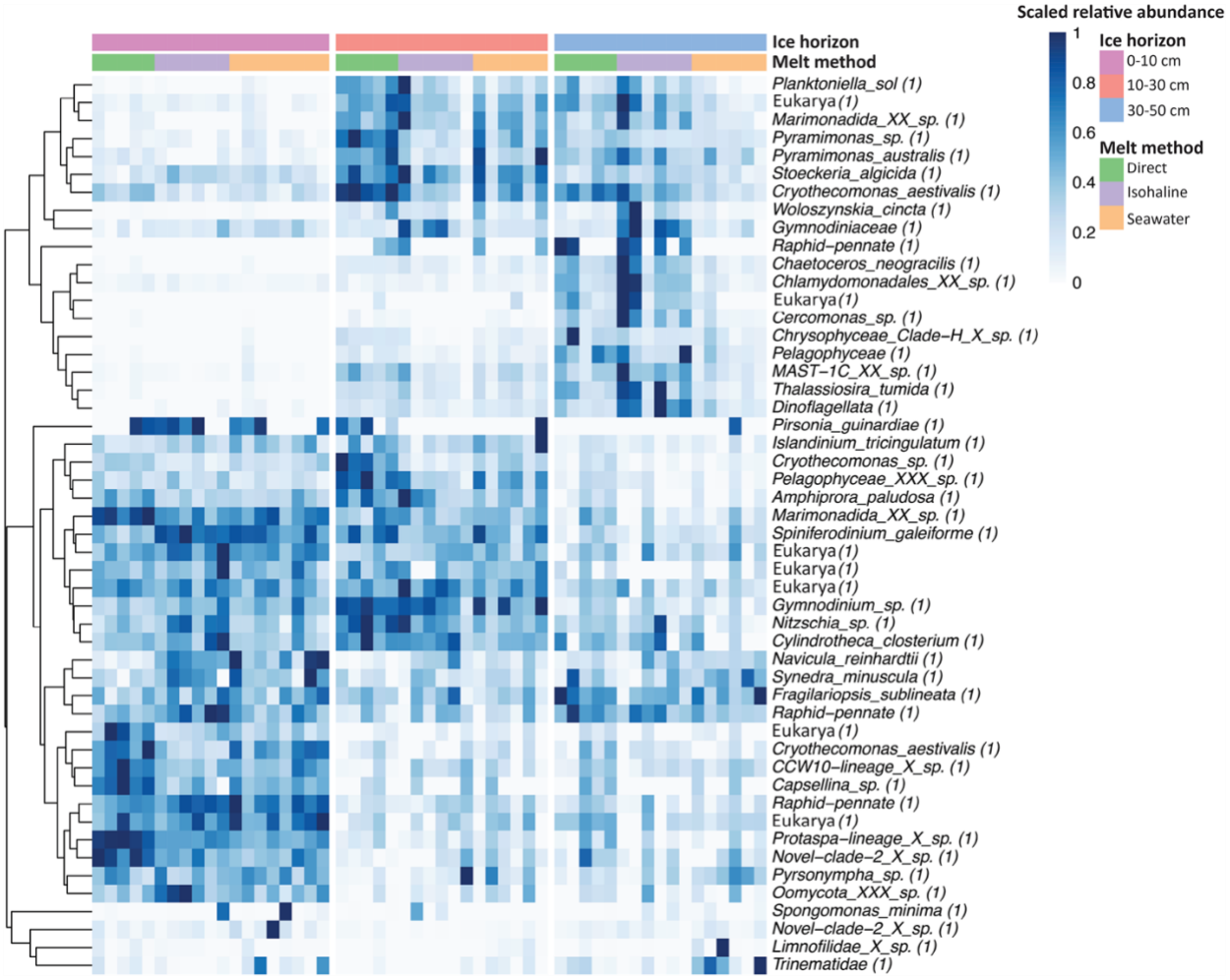


Figure 1-6. Significantly differentially abundant eukaryotic amplicon sequence variants (ASVs) between melting method within each ice horizon.

Relative 18S ASV abundances were scaled using a z-score between 0 and 1. Bray-Curtis dissimilarity distances were used to cluster taxa. Columns (samples) are colored by ice horizon and melting method. The 50 most abundant significantly ($p < 0.1$) differentially abundant ASVs from each comparison are shown with duplicate ASVs (significant in more than one comparison) shown only once. The paprica taxonomic ID (closest relative among RefSeq or PR2 genomes) and placement proportion were used to identify each unique ASV (Bowman and Ducklow, 2015). Heatmaps were generated using R package pheatmap (Kolde, 2019).

A relatively large proportion of these ASVs placed to reference genomes that belonged to the division Cercozoa (i.e., *Protaspa-lineage* sp., Novel clade-2 X sp., *CCW10-lineage* X sp., *Spongomonas minima*, *Marimonadida* XX sp, *Cryothecomonas* sp., *Cercomonas* sp., *Amphiprora paludosa*, *Trinematidae*, and *Capsellina* sp.). However, distinctions were more pronounced among

ice horizons than melting methods, with subsets of ASVs relatively more abundant in the 0–10 cm or 30–50 cm horizon and some shared ASVs of both in the 10–30 cm horizon (Figure 1-6).

Discussion

Our results agree with previous studies that sea ice melting methodology has an impact on resulting microbial parameters (Garrison and Buck 1986; Mikkelsen and Witkowski 2010; Roukaerts et al. 2019; Campbell et al. 2019). They advance understanding of this impact by providing new insights into the effects of melting on bacterial and eukaryotic community structure at the 16S and 18S rRNA gene level.

Impacts of melting method were greatest on measurements of cell abundance and chlorophyll *a* content, where the SM method resulted in significantly higher TC and AF cell abundances (upper ice only) and chlorophyll *a* concentration (all ice horizons) than either the IM or DM methods (Figure 1-3). Significant patterns were consistent between the AF and TC abundance measurements at 30–50 cm, indicating that the SM method buffers equally well for both the photosynthetic and non-photosynthetic communities. However, the photosynthetic community (Figure 1-3B) may be somewhat less sensitive to osmotic stress than the broader microbial community closer to the ice-water interface (SM-5 samples not visually higher and SM-20 samples not significantly higher as observed for TC abundances in Figure 1-3A). When taken together, the greater cell and chlorophyll retention in the SM method of our study agrees well with the results of Campbell et al. (2019), who found that large volume dilutions (8:1 filtered seawater to ice) resulted in highest values of photosynthetic parameters and production in bottom first year sea ice. Both our study and that by Campbell et al. (2019) were based on sea ice collected during Arctic spring and indicate that large volume dilutions with buffers at or close to seawater salinities (29–35 ppt) result in "best" results for the parameters measured. Additionally, because our study

used a sterile NaCl solution as opposed to filtered natural seawater, this conclusion is reached without the potential bias that a natural seawater buffer might bring with its addition of dissolved organic carbon and nutrients to the sample.

However, a wintertime study that included measurements of bacterial abundance in land-fast first year sea ice at a site similar to ours (near Utqiaġvik, AK) yielded somewhat contrasting results. Ewert et al. (2013) found that direct melts resulted in significant (55%) cell loss only in the uppermost 3 cm of this colder ice (approximately –12°C). While their results agree with ours that sea ice communities living closer to the ice-seawater interface appear to be less prone to cell loss, they observed higher bacterial abundances in this upper ice using a near isohaline buffer method, melting the samples with brine salinities of approximately 160 ppt to a meltwater salinity of 100 ppt (Ewert et al. 2013). In our study, the IM buffer method did not result in significantly higher cell counts for any ice horizon (Figure 1-3A), including the coldest upper ice we collected (–5.6°C), and the SM buffer treatment outperformed both IM and DM methods for cell retention in this upper sea ice (Figure 1-3A and B). This performance likely stems from the season and associated differences in the in situ ice conditions.

Our samples were collected during a relatively warm period when the air temperature at the time of sampling (–11ºC) was above the average April high (–13ºC) according to the US Climate database. The underlying seawater temperature was not assessed here, but the ice temperatures for all cores were near or above –5ºC. Constrained by a two-phase equilibrium, the dominant processes responsible for brine channel formation are heat and mass transport (Worster and Wettlaufer, 1997), and some brine loss or ice-ocean exchange may have been occurring, as previously observed during this season (Niedrauer and Martin 1979; Buck et al. 1998; Morawetz et al. 2017). If so, the high eukaryotic diversity seen in the 10–30 cm ice horizon could be from

infiltration or cell movement, preventing the dominance of a single species (Stoecker et al. 1992; van Leeuwe et al. 2018). This potential seawater infiltration, in addition to the relatively large size (10–20 cm length) and warm temperatures (Table 1-S1) of our core sections, could potentially explain the outperformance of the larger volume SM method for our samples. The IM method may have been more osmotically stressful than helpful due to inadequate buffering by the low volumes of hypersaline brine required to match the salinity conditions within our sampled ice. Either way, considering the in situ ice conditions during sampling and the total ice volume of subsamples is important when considering the melting method to apply (Miller et al. 2015).

In addition to total cell abundances and photophysiology, our results suggest that melting method coupled to differential osmotic cell lysis has the potential to impact estimates of bacterial community structure from 16S rRNA gene sequencing. With several taxa significantly more abundant in the buffered SM and IM methods than the unbuffered DM method, where only a few taxa are significantly more abundant in comparison to the buffered methods, our analyses indicate that buffered melts retain a broader spectrum of the bacterial community. However, unlike cell count and chlorophyll *a* concentration, neither buffered method (SM or IM) appears particularly better than the other at ASV retention for our sample set. Differences in ASV read count between the IM and SM comparisons were less striking than the differences between the IM and DM or SM and DM comparisons (Figure 1-5), which accounted for many of the ASVs that were significantly differentially abundant across multiple comparisons (Table 1-S2). While direct measurements of alpha diversity were only significantly impacted in the 10–30 cm layer (Figure 1-3E), dispersion among groups for all samples was significantly different between melting methods (Table 1-4), indicating the potential for community bias throughout the ice column. Indeed, the DeSeq2 analysis resulted in more ASVs that were differentially abundant in the 0–10

cm core section than in the upper ice sections for each melt treatment comparison (Figure 1-S2). This finding contrasts with our cell abundance results, indicating that, while cell loss is reduced at the ice-seawater interface, either the larger community size or stronger seawater influence results in a community subset that is less well adapted to osmotic stress and therefore significantly less abundant using the DM method. This variability in bacterial community structure across melt treatments indicates that cell losses between buffered and unbuffered methods can result in biased abundances of specific subsets of the bacterial community.

Most of the 110 significant ASVs (Table 1-S2) were identified as belonging to classes that comprise the bulk of the community found in our samples (Figure 1-4A) and commonly found in sea ice environments, such as Alpha- and Gammaproteobacteria and members of the Bacteroidetes (Boetius et al. 2015). This finding reduces concern that any differential abundances in phylotype were caused by a potential contamination of the sterile NaCl solution during transit to the field. Although the taxonomic classifications used in this study are not exact and merely represent the closest phylogenetic neighbor based on complete genomes in reference databases, they provide insight into the ecological functions that may be affected by preferential cell lysis due to melting method. Many of the ASVs significantly more abundant in either the IM or SM methods were identified by paprica as members of the ubiquitous Pelagibacterales order (i.e., *Candidatus* Pelagibacter), known for genomic streamlining and their ability to survive across low nutrient regimes (Giovannoni et al., 2005), or members of the diverse phylum Bacteroidetes associated with processing complex organic compounds and byproducts of primary production (e.g., *Flavobacteriaceae* such as *Owenweeksia hongkongensis DSM 17368* or *Cryomorphaceae* such as *Fluviicola taffensis DSM* 1*6823*; Hahnke et al., 2016;Figure 1-5). Why these specific subsets of the microbial community may have been particularly affected by buffering method is unclear, but

in studies based on direct melting, the biased removal of such cells from bacterial community structure may have ramifications for various ecological measurements (carbon or nutrient turnover, etc.) and interpretations when comparing bacterial community composition across different sea ice environments. Other ASVs, identified by paprica as belonging to *Glaciecola amylolytica* of the Gammaproteobacteria, were significantly more abundant in the DM samples (Figure 1-5). Originally isolated from Antarctic sea ice (Bowman et al., 1998), *Glaciecola* is an aerobic, halophilic genus since found in many marine environments (Xiao et al. 2019); perhaps its sea ice members are more suited to surviving the osmotic stress associated with steep salinity shifts.

Contrary to previous studies (Garrison and Buck 1986; Mikkelsen and Witkowski 2010), the melting methods we tested did not appear to bias eukaryotic microbial community structure, as dispersion among groups in our 18S ASVs did not differ significantly among the methods. Nevertheless, the results from our DeSeq2 analysis showed a total of 81 unique ASVs as significantly differentially abundant between melting methods (Table 1-S3). In all but the IM–SM melt comparison, the highest number of differentially abundant ASVs were again found in the 0–10 cm core section at the ice-seawater interface (Figure 1-S3). The IM–DM comparison resulted in the greatest number of significantly differential ASVs, but many of these were shared with the SM–DM comparison (Table 1-S3). While the taxonomic spread of significant ASVs was generally representative of the total community composition (Figure 1-4), one trend was that members of the division Cercozoa, small single-celled protozoan eukaryotes, were more likely to be significantly higher in abundance in the buffered melts than the direct melt. This enrichment could, again, be due to the ice conditions at time of sampling, as well as potentially analysis technique. The relative abundance of protozoa in Arctic sea ice, such as Cercozoa species, is highest in early

Spring (van Leeuwe et al. 2018). As melt begins, the diatom community at the ice-seawater interface is the first to "slough off", while flagellate species endure longer in bottom sea ice communities (Tamelander et al. 2009; Torstensson et al. 2015; van Leeuwe et al. 2018). Previous melt studies have used microscopic analyses to conduct taxonomic assignment and define biases among osmotic buffers and direct melt methods, yet numbers of heterotrophic protozoa are often underestimated in field microscopic analyses because they lack chlorophyll and are small (van Leeuwe et al. 2018). By taking a sequencing approach, the inclusion of more of these protists may have complicated the trends seen in previous studies.

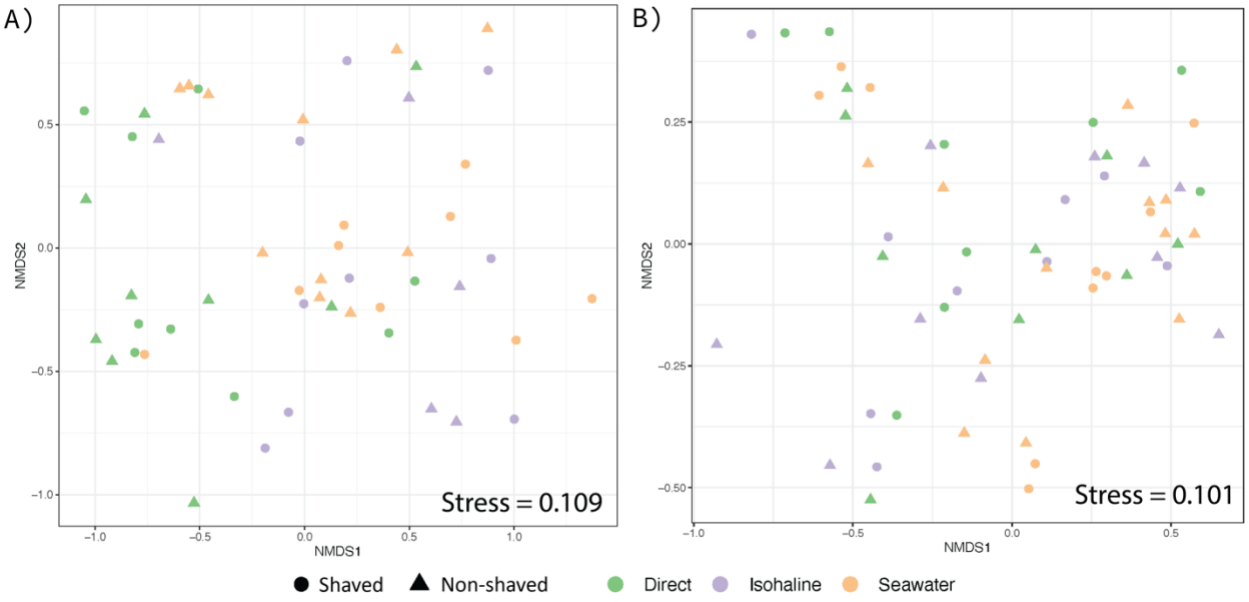
Across all analyses, mechanically shaving the sea ice before melting it had no direct impact on our downstream measurements. However, this procedure added to the overall sample processing time at room temperature, increasing the potential for brine loss prior to melting. While ice-shaving did shorten melting time by a few hours for the buffered melts (shaved DM samples melted more slowly due to snow-like insulation), our core sections were large enough that the degree of difference was smaller than previously determined using thinner ice sections (Campbell et al. 2019).

Conclusions

Based on these results, we conclude that variations in melting methods for sea ice have a significant impact on resulting measurements of bacterial and eukaryotic community structure, while first shaving the ice does not. The degree of bias by melting method depends on the type of measurement pursued and depth horizon of the ice. For spring-time collections of coastal first year sea ice, our results suggest that large (equivalent) volume dilutions with buffers at or near local seawater salinity best minimize osmotic stress and retain the most cell material. However, use of either of the osmotic NaCl buffers we tested — isohaline or seawater salinity — would be

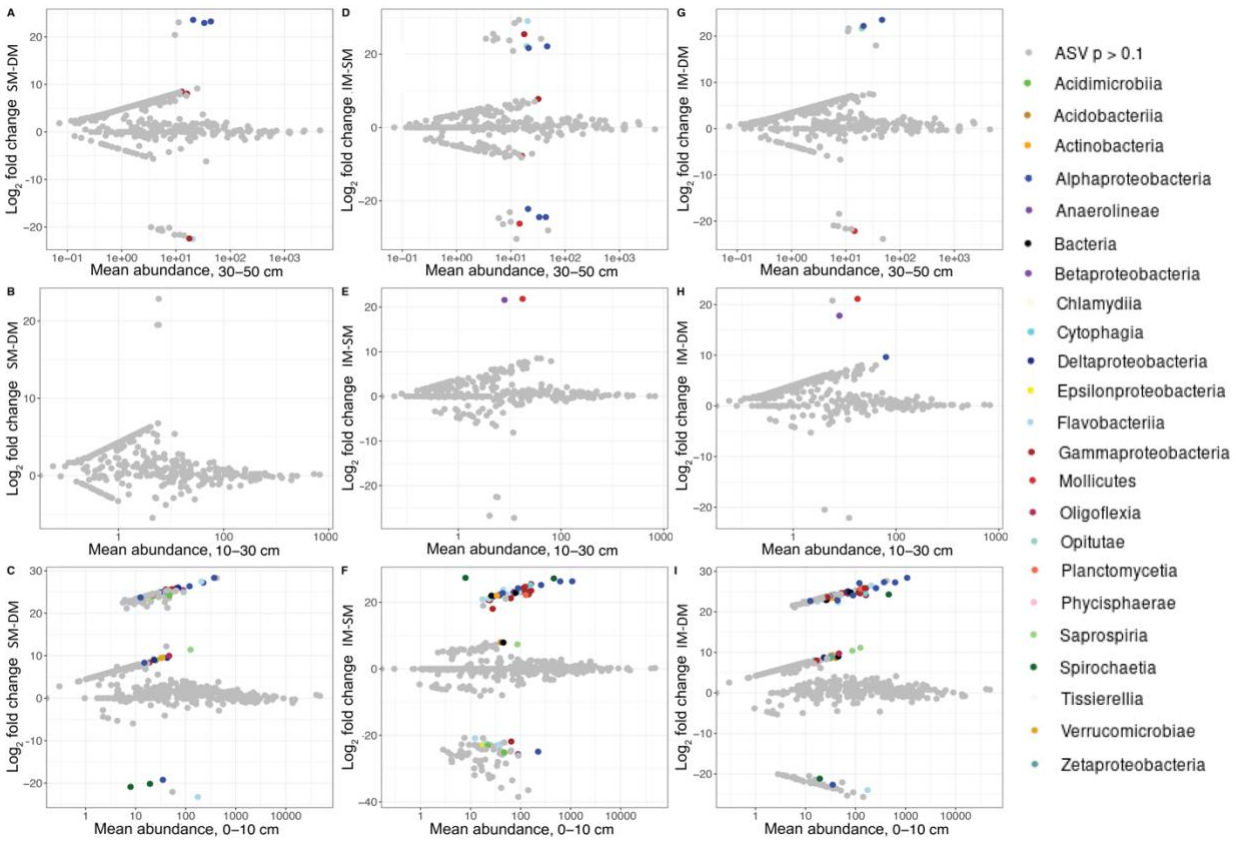
sufficient to reduce biased taxonomic representation in microbial community composition relative to directly melted ice. While cell abundance varied across melting method most strongly in the upper ice (10–30 cm and 30–50 cm core sections), the potential for community level bias at the ice-seawater interface (0–10 cm core section) appears to be higher, particularly for subsets of the bacterial community within the Proteobacteria and Bacteroidetes taxonomic groups. Eukaryotic community level biases among the melting methods were less apparent using our sequencing approach than in previously reported microscope-based work. Ultimately, these results confirm that the "best" melt procedure should be chosen based on research goals, sea ice sampling conditions and final analysis techniques. Repeating similar experiments across varying ice types, seasons, and microbial community compositions would help to identify more conditions in which current methodological practices bias measurements in sea ice and potentially the ensuing biogeochemical and ecological interpretations of this important changing polar habitat.

Supplemental Material



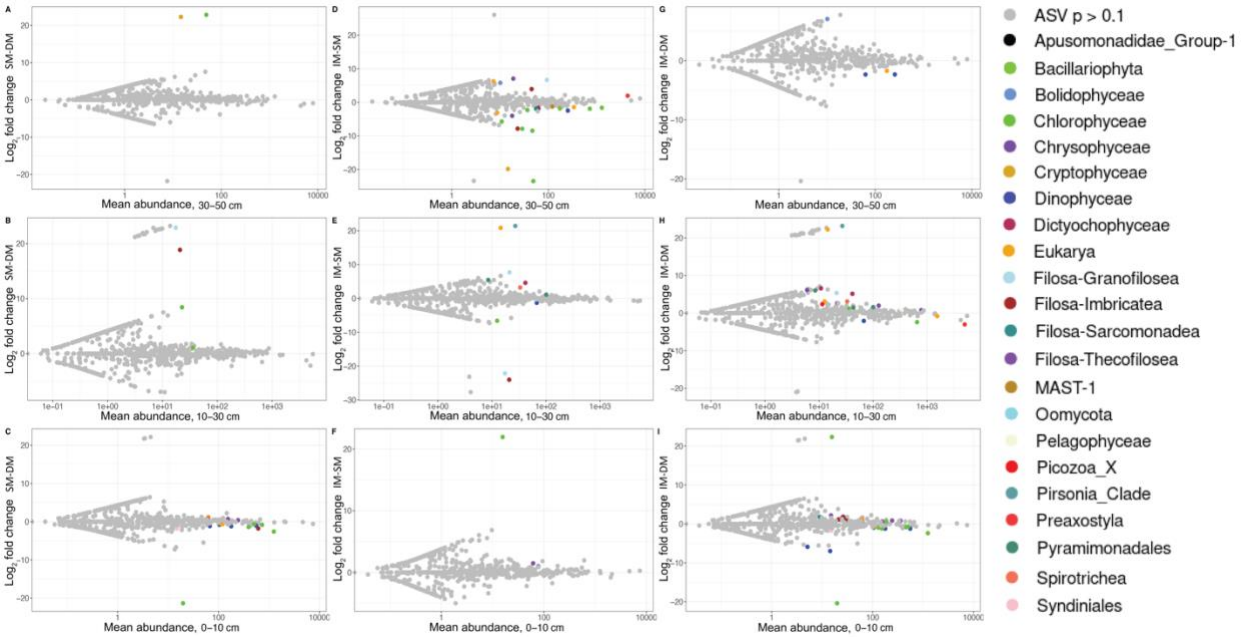
Supplemental Figure 1-S1. Impacts of melting method and ice shaving on total microbial community structure.

Individual samples presented in the first two dimensions of a non-metric multi-dimensional scaling (NMDS) ordination created from (A) bacterial (16S) ASV read counts normalized to cell count and (B) eukaryotic (18S) ASV relative abundances. Samples are distinguished by melt method (color) and whether or not they were shaved prior to melt (shape).



Supplemental Figure 1-S2. Ratio intensity plots of differentially abundant bacterial ASVs between melt method and ice horizon.

The log2-fold change in 16S amplicon sequence variant (ASV) read counts normalized to total cell count between (A) DM-40 and SM-40, (B) DM-20 and SM-20, (C) DM-5 and SM-5, (D) IM-40 and SM-40, (E) IM-20 and SM-20, (F) IM-5 and SM-5, (G) IM-40 and DM-40, (H) IM-20 and DM-20, and (I) IM-5 and DM-5. The labeling system includes melt method (SM for seawater melt, IM for isohaline melt, and DM for direct melt) and ice horizon (5, 20, and 40 for 0–10 cm, 10–30 cm, and 30–50 cm, respectively). When comparing two melt methods, a positive log2-fold change represents higher abundance in the first method listed and a negative log2-fold change indicates higher abundance in the second method. Significantly differential abundances ($p < 0.1$) are displayed in color by taxonomic class.



Supplemental Figure 1-S3. Ratio intensity plots of differentially abundant eukaryotic ASVs between melt method and ice horizon.

The log2-fold change in relative 18S amplicon sequence variant (ASV) read counts between (A) DM-40 and SM-40, (B) DM-20 and SM-20, (C) DM-5 and SM-5, (D) IM-40 and SM-40, (E) IM-20 and SM-20, (F) IM-5 and SM-5, (G) IM-40 and DM-40, (H) IM-20 and DM-20, and (I) IM-5 and DM-5. The labeling system includes melt method (SM for seawater melt, IM for isohaline melt, and DM for direct melt) and ice horizon (5, 20, and 40 for 0–10 cm, 10–30 cm, and 30–50 cm, respectively). When comparing two melt methods, a positive log2-fold change represents a higher abundance in the first method listed and a negative log2-fold change indicates a higher abundance in the second method. Significantly differential abundances ($p < 0.1$) are displayed in color by taxonomic class.

Supplemental Table 1-S1. Final dilutions for isohaline melt treatments.

Core and section	Ice temperature[a] (ºC)	Internal brine salinity (ppt)	Bulk ice salinity (ppt)	Volume sterile NaCl solution added (L)	Approximate total melt volume (L)	Dilution factor
13, 0-10 cm	–2.3	43	8.0	0.088	0.62	1.2
13, 10-30 cm	–3.7	67	6.5	0.34	1.4	1.3
13, 30-50 cm	–5.6	95	5.0	0.56	1.5	1.6
14, 0-10 cm	–2.3	43	8.0	0.077	0.54	1.2
14, 10-30 cm	–3.7	67	6.5	0.34	1.4	1.3
14, 30-50 cm	–5.6	95	5.0	0.59	1.6	1.6
15, 0-10 cm	–2.3	43	8.0	0.084	0.59	1.2
15, 10-30 cm	–3.7	67	6.5	0.33	1.3	1.3
15, 30-50 cm	–5.6	95	5.0	0.58	1.58	1.6
16, 0-10 cm	–2.3	43	8.0	0.086	0.58	1.2
16, 10-30 cm	–3.7	67	6.5	0.33	1.3	1.3
16, 30-50 cm	–5.6	95	5.0	0.58	1.6	1.6
17, 0-10 cm	–2.3	43	8.0	0.08	0.56	1.2
17, 10-30 cm	–3.7	67	6.5	0.32	1.3	1.3
17, 30-50 cm	–5.6	95	5.0	0.57	1.5	1.6
18, 0-10 cm	–2.3	43	8.0	0.086	0.60	1.2
18, 10-30 cm	–3.7	67	6.5	0.33	1.3	1.3
18, 30-50 cm	–5.6	95	5.0	0.59	1.6	1.6

[a]Temperatures calculated for the midpoint of each core section (5 cm, 20 cm, and 40 cm) using a linear model where the ice-water interface (0 cm) is −1.8ºC and the top of the ice core (68 cm) is –8.2ºC (measured value).

Supplemental Table 1-S2. Phylogenetic placement of significantly differentially abundant bacterial amplicon sequence variants (ASVs) between melt method comparisons.

Available online:

Supplemental File 1-S1. Chamberlain_Table_1-S2_16S_deseq_results.xlsx

The first four columns represent taxonomic identification of each significant ASV and the next six columns display the results from the DeSeq2 analysis (Love et al., 2014), including the log2-fold change and significant ($p < 0.1$) adjusted p-values between the seawater melt (SM), direct melt (DM), and isohaline melt (IM) comparisons. When comparing two melt methods, a positive log2-fold change represents a higher abundance in the first method listed and a negative log2-fold change indicates a higher abundance in the second method.

Supplemental Table 1-S3. Phylogenetic placement of significantly differentially abundant eukaryotic amplicon sequence variants (ASVs) between melt method comparisons.

Available online:

Supplemental File 1-S2. Chamberlain_Table_1-S3_18S_deseq_results.xlsx

The first four columns represent taxonomic identification of each significant ASV and the next six columns display the results from the DeSeq2 analysis (Love et al., 2014), including the $\log_2$-fold change and significant ($p < 0.1$) adjusted *p*-values between the seawater melt (SM), direct melt (DM), and isohaline melt (IM) comparisons. When comparing two melt methods, a positive $\log_2$-fold change represents a higher abundance in the first method listed and a negative $\log_2$-fold change indicates a higher abundance in the second method.

Acknowledgements

Chapter 1, in full, is a reprint of the material as it appears in Elementa: Science of the Anthropocene 2022. Chamberlain, Emelia J; Balmonte, John Paul; Torstensson, Anders; Fong, Allison A; Snoeijs-Leijonmalm, Pauline; Bowman, Jeff., University of California Press, 2022. The dissertation author was the primary investigator and author of this paper.

The authors would like to acknowledge Avishek Dutta and Jesse Wilson (University of California San Diego) for their mentorship, along with Natalia Erazo, Elizabeth Connors, and Benjamin Klempay (University of California San Diego) for their input on data analysis.

This work was funded by NSF OPP 1821911 (USA), VR 2018-04685 (Sweden), and Formas 2018-00509 (Sweden). EJ Chamberlain was funded by the NSF Graduate Research Fellowship and the University of California San Diego Cota Robles Fellowship. JP Balmonte was funded by a Carl Tryggers Foundation Fellowship. A Torstensson was funded by the Swedish Research Council (2017-06205).

CHAPTER 2 – Microbial predictors of net heterotrophic conditions in the pelagic Arctic Ocean

Abstract

Microbial diversity and metabolic activity influence rates of net community production (NCP), exerting a direct biological control on marine inorganic and organic carbon fluxes. It is important to understand and quantify the controls on NCP as ecological and oceanographic conditions shift due to changing climate. The 2019–2020 MOSAiC International Arctic Drift Expedition provided a unique opportunity to collect paired observations of microbial diversity and metabolic activity across a seasonally and spatially variable transect of the Arctic Ocean. Community structure was determined using 16S (prokaryotic) and 18S (eukaryotic) rRNA gene amplicon sequencing and metabolic activity was derived using biological oxygen utilization. We then applied predictive machine learning algorithms to disentangle seasonal and spatial patterns in taxonomic and metabolic turnover, extract key microbial predictors, and extend estimates of biological oxygen utilization. We found the water column to be net heterotrophic (minimum: –24.4%, median: –6.0% saturation), except for periodic surface oxygen supersaturation (maximum: 9.8%) within the mixed layer. Using Self Organizing Maps, we identified clear successional patterns in microbial community structure, which were seasonally driven in the upper ocean and uniform at depth. A Random Forest regression model, using DNA sequence data as predictor variables, was used to reliably reconstruct biological oxygen concentrations (RMSE = 5.32 µmol kg^{-1}) and identified heterotrophic taxa as key predictors. Ultimately, this work shows linkages between microbial diversity and upper Arctic Ocean oxygen turnover and highlights diagnostic tools to expand biogeochemical datasets and improve models of net productivity and carbon flux in a changing Arctic.

Introduction

Net community production (NCP), or the metabolic balance between carbon fixation by photosynthetic autotrophs and aerobic respiration by heterotrophs, imposes a biological control on the ocean carbon sink (Volk and Hoffert 1985; Izett et al. 2021). The Arctic Ocean is considered a net carbon sink: abiotically, through low surface seawater temperatures enhancing CO_2 solubility and through the sea ice carbon pump (Rysgaard et al. 2007), and biotically, due to high levels of net primary production (NPP) along coastal shelves (MacGilchrist et al. 2014). However, published estimates of surface Arctic Ocean NCP are relatively low with temporal and spatial variability (Ulfsbo et al. 2014; Eveleth et al. 2014; Schanke et al. 2021; Izett et al. 2021). Understanding how the complex physical and biological controls on NCP interact and shift under changing Arctic conditions, such as ocean warming and sea ice decline (Meredith et al. 2019), is critical for predicting the future of the Arctic Ocean carbon sink.

Microbial community structure exerts a strong control on NCP (Guidi et al. 2016; Wang et al. 2018) and thus a biological control on carbon sequestration. Previous studies have found that observations of microbial community structure can be utilized to estimate metabolic rates, predict key transitions in ecosystem state, and better understand the biological mechanisms underlying variability in observed biogeochemical processes (Bowman et al. 2017; Wilson et al. 2021; Erazo et al. 2021; Dutta et al. 2022). For NCP in polar regions specifically, relationships with taxonomic diversity and shifting community composition have been observed (Campbell et al. 2022a), with some eukaryotic community members acting as viable predictors of measured rates (Lin et al. 2017, 2021). However, prokaryotic predictors of NCP have been less well studied despite their critical role in global respiration (Worden et al. 2015; Kim et al. 2023) and tested predictive capacity for other ecological processes (Bowman et al. 2017; Dutta et al. 2022).

Machine learning tools allow us to extract clear patterns and predictive relationships between biogeochemical processes and diversity data (Qu et al. 2019; Bowman 2021). In recent years an increasing number of studies have leveraged machine learning algorithms when analyzing community structure data for such purposes, with great progress in predicting functional information from the community structure alone (Thompson et al. 2019; Dutta et al. 2022; Ramoneda et al. 2023). To train such models, large, paired datasets of both the taxonomic makeup of the microbial community and the biogeochemical parameter of interest are required. The 2019–2020 Multidisciplinary drifting Observatory for the Study of Arctic Climate Expedition (MOSAiC; Shupe and Rex 2022) provided a novel opportunity to collect paired measurements of both community structure and net productivity estimates across seasonal cycles in an under-sampled region of the central Arctic Ocean. This year-long drift experiment was designed to carry out a comprehensive cross-disciplinary and process-level study of the interconnected central Arctic ocean-ice-atmosphere climate system. In support of this goal, we collected a paired time series of upper ocean microbial community structure, as determined by 16S (bacteria and archaea) and 18S (eukaryotes) rRNA gene amplicon sequencing, and 'biological oxygen anomaly' spanning March–September of 2020. The relative ratio of biologically derived 'anomaly' oxygen (O_2) and biologically inert argon (Ar) in the surface ocean (O_2/Ar) can be used as a signal of biological activity reflecting the net metabolic balance between photosynthesis and respiration, i.e., the net trophic state controlling NCP (Eveleth et al. 2014).

In this study, we investigate the role and predictive capacity of microbial community structure on biological oxygen anomaly (as a proxy of NCP; Seguro et al. 2023) across both seasonal and spatial gradients within the central Arctic. Leveraging the wealth of co-located community structure and biogeochemical measurements collected during 2019–2020 MOSAiC

year, we first examined the seasonal and spatial variability of net metabolic shifts in biological oxygen anomaly and microbial community structure. Using Self Organizing Maps (SOMs; Wehrens and Kruisselbrink 2018), we segmented the prokaryotic (bacteria and archaea) and eukaryotic (phytoplankton and heterotrophic eukaryotes) community structure into coherent community-types, comparing their various genetic characteristics and potential environmental shaping. We then identified key predictors and biological drivers of biological oxygen anomaly during MOSAiC through the construction of Random Forest (RF) regression models (Breiman 2001), attempting both sequencing data and environmental parameters as input.

Materials and Methods

Oceanographic data collection – MOSAiC Expedition

Beginning in September 2019, the German icebreaker RV *Polarstern* (Knust 2017) traveled to the Amundsen Basin, where it was tethered to a multi-year ice floe and an ice camp was established to conduct research while passively drifting with the Transpolar Drift (4 October 2019–31 July 2020; Figure 2-1). A break in this drift occurred when the *Polarstern* was required to temporarily leave, then re-establish, the ice camp during resupply (27 May–15 June 2020). Then, on July 31st, *Polarstern* reached the marginal ice zone and the original ice floe broke apart. *Polarstern* then relocated further north, and a second ice camp was established to continue observations for the remainder of the MOSAiC time series (21 August 2020–20 September 2020). During the MOSAiC Drift, partner vessels supported re-supply and personnel transfers in five coherent "legs" of the expedition (Shupe et al. 2020). There are two gaps in the data presented here, one leg where no project personnel were onboard from December 2019–March 2020 and another during the May–June 2020 resupply.

Ecological and biogeochemical data used in this study were collected from either *Polarstern*'s underway seawater system (inlet at 11 m depth) or from Niskin bottles attached to a

CTD rosette (Figure 2-1). Between March–May 2020, 1000-m deep 12-bottle CTD rosette casts were operated using a mobile winch system at "Ocean City", a dedicated oceanographic sampling ice-hole approximately 300 m away from *Polarstern* within the ice camp's central observatory. Between November–December 2019 and between June–September 2020, full water column 24-bottle CTD rosette casts were operated using the ship's winch from an ice-hole next to *Polarstern* (Rabe et al. 2022). Both the Ocean City and *Polarstern* CTDs were equipped with sensors for measuring pressure, temperature, conductivity, and total dissolved oxygen (Tippenhauer et al. 2023a; b). Temperature and salinity from the duct keel intake (11 m) were collected by the remote temperature sensor SBE38 integrated with the SBE21 Thermosalinograph (Seabird) system installed on *Polarstern* (Haas et al. 2021; Kanzow et al. 2021; Rex et al. 2021a; b; c). Full details of physical oceanographic sampling procedures can be found in Rabe et al. (2022). Calculated oceanographic parameters (Schulz et al. 2023) used in this study, such as mixed layer depth and water mass influence are described further in Schulz et al. (unpubl.). A total of 360 of the water column (CTD rosette or underway system) observations for microbial community structure were sampled at the same time as water column chlorophyll *a* (Chl-a) concentration (Hoppe et al. 2023a; b). Further sampling details of water column ecological properties during MOSAiC can be found in the MOSAiC ECO Team Overview Publication (Fong et al. unpubl.).

Microbial community structure and segmentation

A total of 693 seawater samples (approximately 0.5–1 L) were collected in acid-cleaned and thrice sample-rinsed 1L HDPE plastic bottles once daily from 11 m depth using the ship's underway system and on an approximately weekly basis throughout the water column from Niskin bottles attached to the CTD rosette. Collected samples were filtered immediately or kept dark at

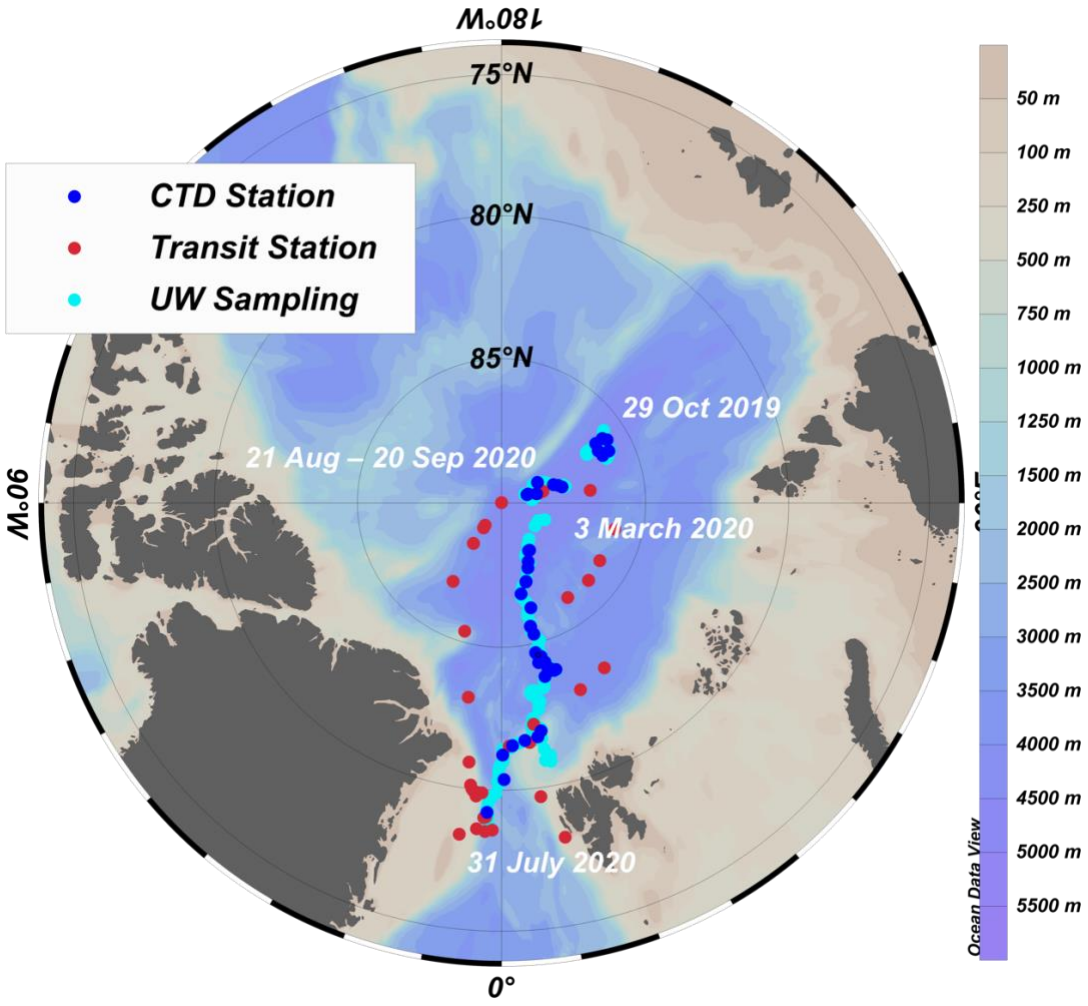


Figure 2-1. Sampling Locations.
Geographic location of daily underway (UW) sample collections (cyan) and CTD stations (navy) during the MOSAiC Drift. Samples collected during active transit marked in red. Map created using Ocean Data View version 5.6.3 (Schlitzer and Reiner 2022) and IBCAO bathymetry data version 3 (Jakobsson et al. 2020).

0ºC until filtered (within 4 hr) through a sterile 0.2 µm Supor membrane disc filter (Pall Corporation) using a vacuum filtration manifold. The manifold was flushed with MilliQ and ethanol sterilized before, in between, and after samples. Filters were stored at −80ºC onboard *Polarstern* until the end of the expedition when they were shipped to Scripps Institution of Oceanography on dry ice and archived at −80ºC.

A KingFisher™ Flex Purification system and MagMax Microbiome Ultra Nucleic Acid Extraction kit (ThermoFisher Scientific) were used for extractions. Extracted material was then sent to Argonne National Laboratory where amplicon library preparation and sequencing were conducted on the Illumina Miseq platform using a 2 x 151 bp library architecture. rRNA gene amplification was conducted using universal primers 515F and 806R (16S, V4; Walters et al., 2016) and 1380F and 1510R (18S, V9; Amaral-Zettler et al., 2009). The R package dada2 (Callahan et al. 2016) was used to denoise and merge Illumina reads prior to taxonomic classification using PAthway PRediction by phylogenetIC placement (paprica v0.7.0; Bowman and Ducklow 2015; Erazo et al. 2021).

Full details on taxonomic assignment and QC can be found in Text 2-S1. Briefly, the paprica pipeline was used to assign individual amplicon sequence variants (ASVs) to its closest relative among completed genomes in either the Genbank RefSeq (16S; Haft et al. 2018) or PR2 (18S; v4.13.0 Guillou et al. 2013) reference databases (which does not necessarily indicate the presence of that exact strain, but a similar organism). For bacteria and archaea, paprica also uses the point of placement in the reference tree to estimate genomic characteristics such as genome size and 16S rRNA gene copy number (Erazo et al. 2021), as well as theoretical growth rate using the gRodon package (Weissman et al. 2021). Growth rate estimates from codon usage do not reflect temperature conditions and therefore represent theoretical maximums only used to compare between samples and not as *in situ* rates. Final read counts for 16S ASVs were normalized to estimated 16S rRNA gene copy number before calculations of relative abundance.

SOMs were run using the 'kohonen' package to reduce dimensionality within the sequence dataset (Wehrens and Kruisselbrink 2018). Following the techniques outlined in Bowman et al. (2017) and Wilson et al. (2021), the Hellinger transformed relative abundance matrices of 16S and

18S ASVs were independently segmented into coherent sample types, here termed "mode." SOM training was conducted on a 6 x 6 toroidal grid with hexagonal map units. The grid size was selected based on the distribution of samples assigned to each map unit in grids up to 10 x 10. Segmentation was achieved through K-means clustering, with a final k selected based on a within-clusters sum of squares scree plot and experimental variation around the perceived optimum, or 'elbow', to account for reasonable distribution within the map grid. Results from SOM segmentation were additionally compared to results of a non-metric multidimensional scaling (NMDS) ordination; run based on the Bray-Curtis dissimilarity matrix (Bray and Curtis 1957) of Hellinger transformed 16S or 18S ASV relative abundances (vegan package; Oksanen et al. 2015). A permutational analysis of variance (PERMANOVA, 999 permutations) was then used to test whether there was a significant ($\alpha = 0.05$) difference between modes (vegan package; Oksanen et al. 2015). Kruskal-Wallis with pair-wise Wilcoxon signed-rank tests were used to test for significant ($\alpha = 0.05$) differences in environmental variables between modes (ggpubr package; Kassambra, 2023). Where relevant, the Holm-Bonferroni method was used to correct p-values for multiple comparisons (Holm 1979). All statistical analyses were conducted using R version 4.3.1 (R Core Team 2023) within the R studio environment (Posit team 2022).

Total oxygen (O_2) and biological oxygen saturation (O_2/Ar)

Biological oxygen saturation, as estimated from the ratio of O_2 and Ar concentrations, was used as a qualitative indicator of recent metabolic activity and functional estimate of trends in NCP (Equation 1). This value was obtained following the methods and assumptions of Craig and Hayward (1987) where the ratio of physical O_2 concentration to O_2 at air-saturation equals the ratio of Ar concentration to Ar at air-saturation.

$$\underline{\text{Equation 1:}}\ \Delta O_2/Ar = \left[\left(\frac{[O_2]_{meas}/[Ar]_{meas}}{[O_2]_{sat}/[Ar]_{sat}}\right)\right] - 1$$

Dissolved O_2 and Ar were analyzed on a Pfeiffer QMG 220 quadrupole membrane inlet mass spectrometer (MIMS) using a flow-through silicon capillary membrane inlet and equipped with a pulse tube cooler for water vapor removal. Underway seawater was measured continuously throughout the entirety of the MOSAiC Expedition, except during daily instrument calibrations (2-point; 0% and 21%/1% O_2/Ar gas) and bottle sampling events. Discrete seawater was collected from the water column in Milli-Q rinsed air-tight 300 mL Biological Oxygen Demand (BOD) bottles at the CTD rosette and were always paired with a measurement for microbial community structure (although not all microbial community structure samples have a paired gas measurement). To align continuous underway data with the discrete underway collections for microbial community structure, we calculated 4–6 hour averages centered on the recorded sampling time. Values were then smoothed using *Loess* local regression (R Core Team 2023) to reduce data noise. Missing values represent periods of the dataset where averaging encompassed periods of missing or poorly calibrated data. Expanded details regarding the MIMS system and bottle sampling at the rosette can be found in the Text 2-S2.

Theoretical concentrations of O_2 and Ar at air saturation ($[O_2]_{sat}$ and $[Ar]_{sat}$) were calculated using equilibration chamber temperature and salinity and the solubility equations from Garcia and Gordon (1992; 1993 – combined fit) and Hamme and Emerson (2004), respectively. Comparing $[O_2]_{sat}$ and $[Ar]_{sat}$ to 21%/1% synthetic air standard calibrations (Text 2-S2), we translated the seawater spectrometer readings into total concentrations of $[O_2]_{meas}$ and $[Ar]_{meas}$ in μmol kg^{-1}. Following the approach used in Eveleth et al. (2014), these values were then used to extract only the biological component of total oxygen concentration, $[O_2]_{bio}$ in μmol kg^{-1} (Equation 2).

Equation 2: $$[O_2]_{bio} = \frac{[Ar]_{meas}}{[Ar]_{sat}}[O_2]_{sat}(\Delta O_2/Ar)$$

Seawater $[O_2]_{sat}$ and $[Ar]_{sat}$ were calculated individually for each sample using *in situ* practical salinity and potential temperature (see section 2.1). Positive (or negative) values of $[O_2]_{bio}$ represent a metabolic tracer for net oxygen production (or consumption), or net productivity (or heterotrophy, respectively).

Random Forest regressions

A RF regression model was used to extract the key microbial predictors of $[O_2]_{bio}$ estimates ('randomForest' package; Liaw and Wiener, 2002). This supervised machine learning algorithm applies an ensemble learning approach where multiple, independent, and randomly constructed decision trees are run simultaneously. The final prediction represents the cumulative average prediction of all trees (Breiman 2001). We applied feature selection using the Boruta algorithm (Kursa and Rudnicki 2010) prior to model construction for select models. This iterative process identifies the predictor variables which will be statistically better for model performance than a randomized version of itself. Predictor variables were further evaluated following model construction by comparing the increase of model mean squared error (% IncMSE) which occurs when the variable is randomly permutated (Liaw and Wiener, 2002). A total of 8 RF models were built and tested using varying combinations of predictor variables (Table 2-S1) from the 217 discrete water column samples collected from the CTD rosette which had paired measurements of community structure and $[O_2]_{bio}$. 70% of these samples (n = 149) were randomly chosen to train the RF models. The number of trees within the model was chosen based upon a plateau of the raw model mean standard error at 500 trees for all models. After validating the accuracy of our models using the remaining 30% (n = 68) of paired CTD community and $[O_2]_{bio}$ samples, all samples were combined and used to train a comprehensive predictive RF model using the feature-selected 16S & 18S ASV relative abundances as predictor variables. Internal model accuracy for the final RF

model was conducted through a bootstrapped cross validation analysis where 10% of the data were randomly withheld for 999 iterations ('rfUtilities' package; https://github.com/jeffreyevans/rfUtilities). This model was then used to predict $[O_2]_{bio}$ values for periods of the drift where non-paired measurements were collected, as well as the daily underway data collections to compare predicted results with the smoothed integrated average of the continuous underway $[O_2]_{bio}$ measurements.

Results

Microbial community composition an SOM segmentation

Denoising yielded a total of 5,266 16S and 17,357 18S ASVs across all samples. A total of 328 of the 16S sequences were identified as suspected mitochondrial DNA and removed from downstream analysis. Additionally, 6.6% of the remaining 16S ASVs were identified as potential chloroplast sequences and 60.4% of 18S ASVs were assigned similar taxonomy to intracellular parasites and excluded from the SOM descriptions of community structure (Figure 2-S1). Overall, the prokaryotic community was dominated by Bacteria with only 5.8 % of the remaining 16S ASVs assigned to the domain Archaea.

After removing suspect organelle and parasite sequences, we identified seven coherent 16S (Figure 2-2a) and six coherent 18S (Figure 2-3a) microbial modes occurring across the time series. This is confirmed by our NMDS ordination analysis, which showed that both the prokaryotic (Figure 2-2B) and eukaryotic (Figure 2-3B) modes are composed of a distinctive community structure, with very little overlap in the 2D ordination space. In both ordinations, distribution

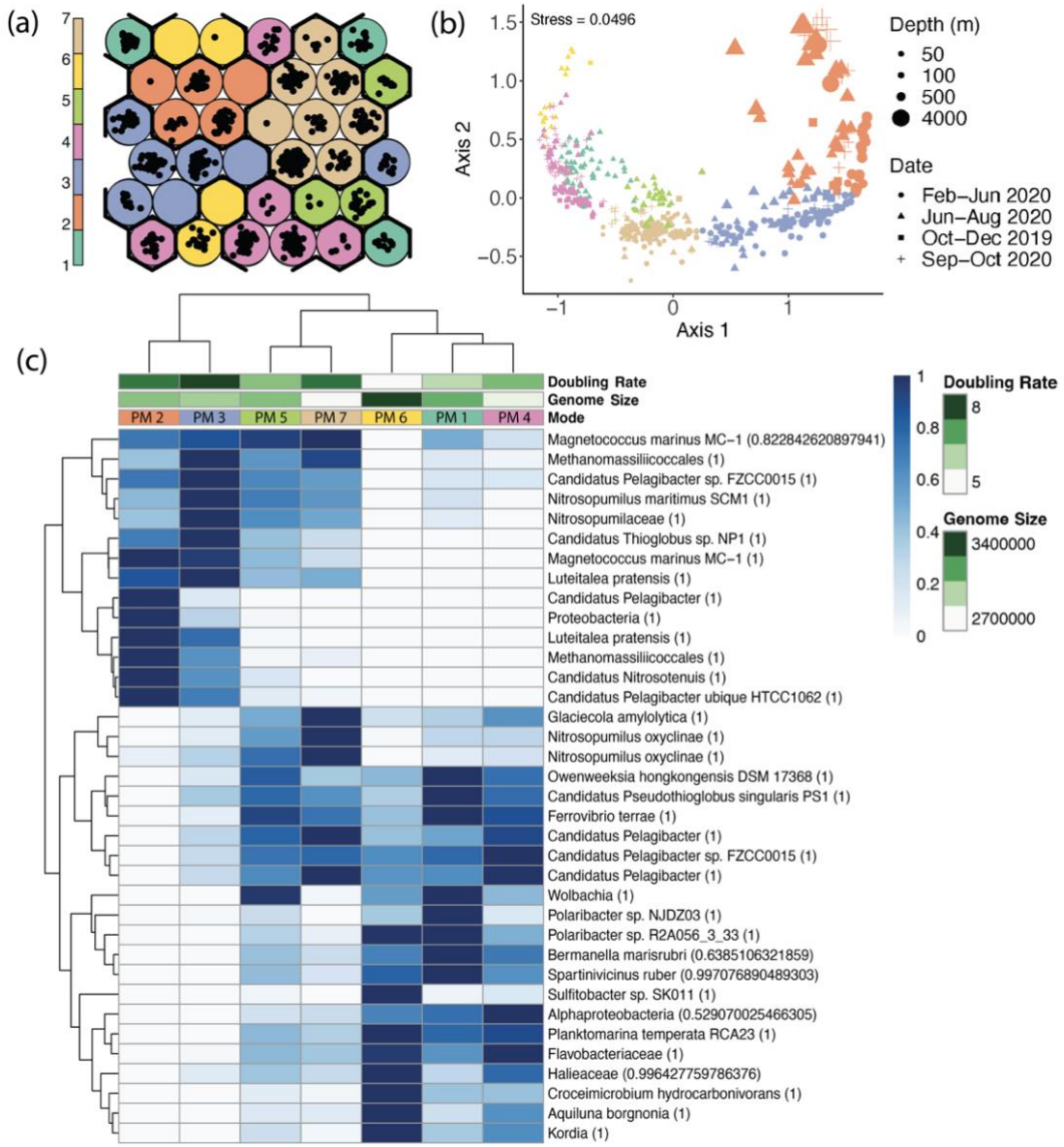


Figure 2-2. SOM segmentation of the prokaryotic community.
(a) The SOM grid (6 x 6) with individual samples (black dots) sorted into their respective map units. Arranged as a toroid, units on opposite sides of the 2-D grid connect. K-means clustering of map units vary by color and thick lines surrounding each prokaryotic mode (PM). (b) Individual samples presented in the first two dimensions of a non-metric multi-dimensional scaling (NMDS) ordination created from the Hellinger transformed relative abundances of 4613 16S ASVs (total community, not including suspected chloroplasts). Samples are colored by their assigned PM from (a), sized by collection depth in the water column, and shaped by collection date. (c) PM community structure as represented by the 10 most abundant ASVs averaged across all samples within each mode. Duplicate ASVs (within the top 10 most abundant of more than one PM) shown only once. Abundances were scaled using a z-score between 0-1. ASVs and columns (PMs) clustered using Bray-Curtis dissimilarity distances. Taxonomic IDs represent the closest relative among Refseq genomes and placement proportion. Columns are additionally colored by average paprica predicted minimum doubling time (in hours) and average genome size (in base pairs).

across the first two axes highlighted significant separation by time and depth within the water column (16S: r^2 = 0.933 and 0.994 respectively, p = 0.001; 18S: r^2 = 0.143 and 0.490 respectively, p = 0.001). The PERMANOVA on the resulting Bray-Curtis dissimilarity matrices confirmed that dispersion among groups was significantly different across modes for both the 16S (p = 0.001; $F(6, 671) = 163.39$) and 18S communities (p = 0.001; $F(5 , 656) = 76.72$).

The 10 most abundant ASVs within each of the 7 prokaryotic modes (PMs) encompassed only 36 unique 16S ASVs, indicating some ASVs were in the top 10 most abundant of more than one mode. Here, we compare the average relative abundance (scaled 0–1) across all samples within each mode. Using Bray-Curtis dissimilarity distances, modes clustered into three distinctive community compositions (Figure 2-2c) across time and space (Figure 2-2b). Deep water samples clustered into either PM-2, sampled 100–4400 m, or PM-3, sampled 50–1000 m. These modes contained high abundances of ASVs identified to reference genomes of Archaea (Methanomassiliicoccales, Nitrosopumilacae sp.), Proteobacteria (Alpha: *Magnetococcus marinus;* Gamma: Candidatus thioglobus, Pelagibacterales sp.), and Acidobacteria (*Luteitalea pratensis*). Sampled above 200 m during spring and early summer, PM-5 and PM-7 also contained high abundances of ASVs identified reference genomes of Archaea (Nitrosopumilus oxyclinae), and Proteobacteria (Alpha: *Magnetococcus marinus, Wolbachia;* Gamma: Pelagibacterales sp.), as well as *Glaciecola amylolytica*. Late summer modes PM-4 and PM-6 and winter mod PM-1 contained high abundances of ASVs identified to reference genomes of Flavobacteria (*Owenweeksia hongkongensis*, *Polaribacter* sp., *Croceimicrobium hydrocarbonivorans*, *Kordia* sp.*),* Proteobacteria (Alpha: *Ferrovibrio terrae, Sulfitobacter* sp., *Planktomarina temperate*; Gamma: Candidatus Pelagibacter and Pseudothioglobus strains, *Bermanella marisrubri*, *Spartinivicinus rube*), and Actinomycetia strain *Aquiluna borgnonia* (Figure 2-2). Full community

composition of each mode at class level can be found in Supporting Information, Figure 2-S2. Estimated theoretical doubling times inferred from the paprica pipeline indicated faster doubling times for PM-6 and PM-1, and slower doubling times for PM-2 and PM-3. PM-6 also had the largest genome size, while PM-7 had the smallest (Figure 2-2c). A full summary of general prokaryotic mode characteristics can be found in Table 2-1.

The 10 most abundant ASVs within each of the 6 eukaryotic modes (EMs) encompassed only 31 unique 18S ASVs (Figure 2-3c). However, those which clustered together (EM-4 and EM-6; EM-3 and EM-5) had less similarity in high abundance taxa than the similarly clustered PMs. Winter mode EM-1 contained high abundances of ASVs assigned to reference genomes of marine stramenopiles (MAST sp.), dinoflagellates (*Prorocentrum texanum, Karenia brevis*), cercozoan (*Spongomonas minma,* Filosa-Imbricatea Novel-clade-2), and diatoms (*Attheya septentrionalis, Chaetoceros sp., Fragilariopsis sublineata).* EM-2 was assigned to samples from almost the entirety of the water column (from 50–4500 m) and contained high abundances of pico-eukaryote ASVs, including dinoflagellates (Syndiniales group, *Herdmania litoralis, Protodinium simplex*), Picozoa, and a *Chaunacanthida* species. Clustered surface water (2–152 m) modes EM-3 and EM-5 had a similar community composition with high abundances of ASVs assigned to reference genomes of dinoflagellates (*Gymnodinium litoralis, Karlodinium veneficum, Palatinus apiculatus, Pentapharsodinium tyrrhenicum*). However, EM-3 also contained high abundance of ASVs assigned to Picozoa and radiolarian (RAD-B Group 1) genomes, while EM-5 was also enriched in diatom ASVs (*Chaetoceros* sp., *Attheya septentrionalis*). EM-4, sampled between 2 and 2643 m, also contained a mixed community with high abundances of ASVs assigned to diatoms (*Thalassiosira tumida, Guinardia delicatula, Fragilariopsis sublineata, Chaetoceros* sp.,), dinoflagellates (*Gymnodiniaceae*), and cercozoan (Filosa-Imbricatea Novel-clade-2) reference

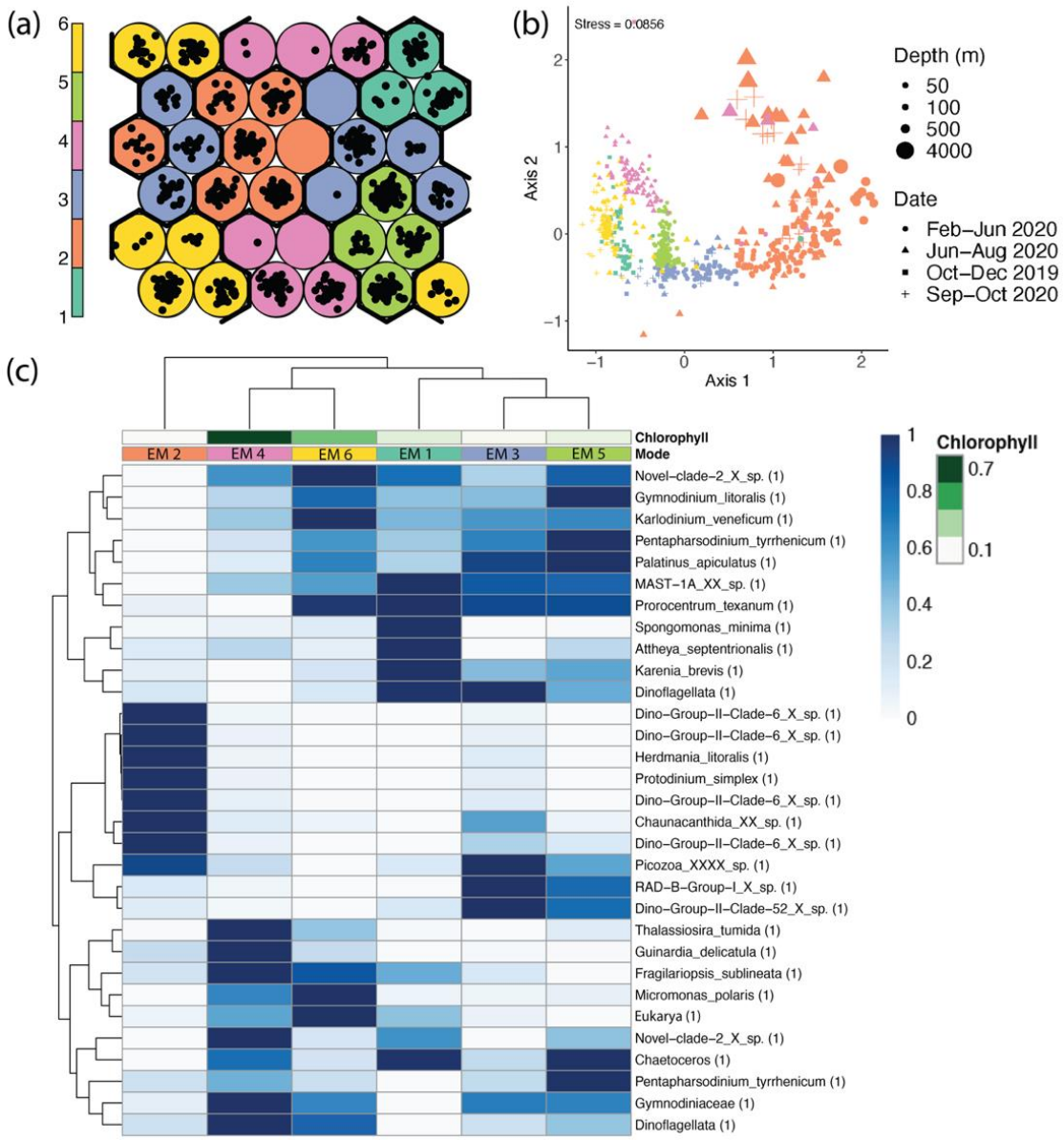


Figure 2-3. SOM segmentation of the eukaryotic community.
(a) The SOM grid (6 x 6) with samples (black dots) sorted into their respective map units. Arranged as a toroid, units on opposite sides of the 2-D grid connect. K-means clustering of map units vary by color and thick lines surrounding each eukaryotic mode (EM). (b) Individual samples presented in the first two dimensions of a non-metric multi-dimensional scaling (NMDS) ordination created from the Hellinger transformed relative abundances of 6870 18S ASVs (total community, not including suspected intracellular parasites). Samples are colored their assigned EM from (a), sized by collection depth in the water column, and shaped by collection date. (c) EM community structure as represented by the 10 most abundant ASVs averaged across all samples within each mode. Duplicate ASVs (within the top 10 most abundant of more than one EM) shown only once. Abundances were scaled using a z-score between 0-1 ASVs and columns (EMs) clustered using Bray-Curtis dissimilarity distances. Taxonomic IDs represent the closest relative among PR2 genomes and placement proportion. Columns are additionally colored by average chlorophyll *a* (Chl-a) concentration for each EM.

genomes. Clustered with EM-4, EM-6 was only sampled between 2 and 100 m but contained high abundances of ASVs assigned to similar dinoflagellate (*Gymnodiniaceae, Karlodinium veneficum, Prorocentrum texanum*) and cercozoan (Filosa-Imbricatea Novel-clade-2) reference genomes as well as reference genomes of *Micromonas polaris, Fragilariopsis sublineata*, and an unidentified unique 18S sequence (Figure 2-3). Full community makeup of each mode can be found in Figure 2-S3. Average Chl-a concentration was highest in EM-4 and lowest in EM-2 (Figure 2-3c). A full summary of general eukaryotic mode characteristics can be found in Table 2-1.

Total and biological oxygen concentration

Water column total dissolved oxygen concentrations $[O_2]_{tot}$ as calculated from the discrete BOD bottle samples ranged between 190.54–470.23 µmol kg^{-1} and generally followed the pattern of $[O_2]_{tot}$ measured by the CTD sensors with an average deviation of –6.68 ± 53 µmol kg^{-1}) (Supporting Information Figure 2-S4). The ratio of $[Ar]_{meas}/[Ar]_{sat}$, representing physical transformations of O_2 as used in Equation 2 (Eveleth et al. 2014), were close to the expected value of 1 (Cassar et al. 2011), with a mean and standard deviation of 1.08 ± 0.16. Water column $[O_2]_{bio}$ (including transit stations) ranged between –79.00 and 33.98 µmol kg^{-1}. $[O_2]_{bio}$ was generally near 0 or negative, indicating oxygen consumption (net respiration) except in early spring and late summer where oxygen accumulation (net photosynthesis) was observed (Figure 2-4).

Below 11 m, negative $[O_2]_{bio}$ values indicated net heterotrophic conditions for most of the upper water column, with stronger undersaturation beneath the mixed layer (approx. 20 µmol kg^{-1} additional consumption; Figure 2-5). Net heterotrophic conditions persisted across the entirety of the deeper water column (below 200 m) with the lowest $[O_2]_{bio}$ values occurring in the deepest water of the Amundsen Basin (Figure 2-S5) or with shoaling near bathymetric features (Figure 2-5). Surface values (11 m and above) from the CTD rosette and averaged values from the underway

system were similar to each other, following similar trends in time and space. Greater variability in magnitude was observed for the averaged underway values (Figure 2-4). Replicate BOD bottles indicated a measurement standard deviation of on average 3.25 µmol kg^{-1}.

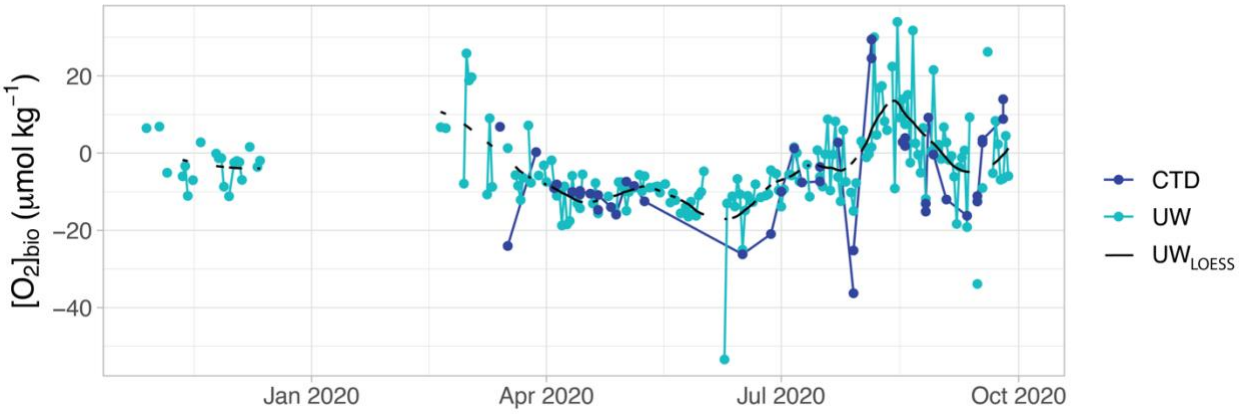


Figure 2-4. Measurements of biological oxygen anomaly ($[O_2]_{bio}$) in µmol kg^{-1} from the upper 11 m of the water column.
Discrete measurements collected from the CTD rosette are shown in red, while continuous underway (UW) measurements integrated across the 4 hours closest to recorded sampling time for discrete community structure are shown in blue. The solid black line shows the *Loess* local regression of underway averages (UW_{LOESS}).

Seasonal and oceanographic context of the MOSAiC Drift

Both magnitude of biological oxygen consumption and the transition between PMs and EMs corresponded with observed oceanographic conditions over the course of MOSAiC, especially vertical stratification within the water column (Figure 2-5). The Kruskal-Wallis test confirmed statistically significant variability in the average environmental and biogeochemical conditions of SOM segmented community modes (Figure 2-6). Many of the pair-wise comparisons (Wilcoxon signed-rank test) between the mean temperature, salinity, $[O_2]_{tot}$, and O_2/Ar of each mode were significantly different as well, for both the prokaryotic (Table 2-S2) and eukaryotic (Table 2-S3) communities. Deep water modes EM-2, PM-2, and PM-3 persisted below the mixed layer over the entirety of MOSAiC and contained, on average, higher temperatures, higher salinities, and lower $[O_2]_{tot}$ than all other modes (Figure 2-6). In the prokaryotic community, PM-

3 had the highest mean salinity (Figure 2-6a) and its vertical stratification in the upper water column closely followed that of Atlantic influenced water (Figure 2-5a), indicating that the variation in deep water prokaryotic community structure was likely related to water mass influence.

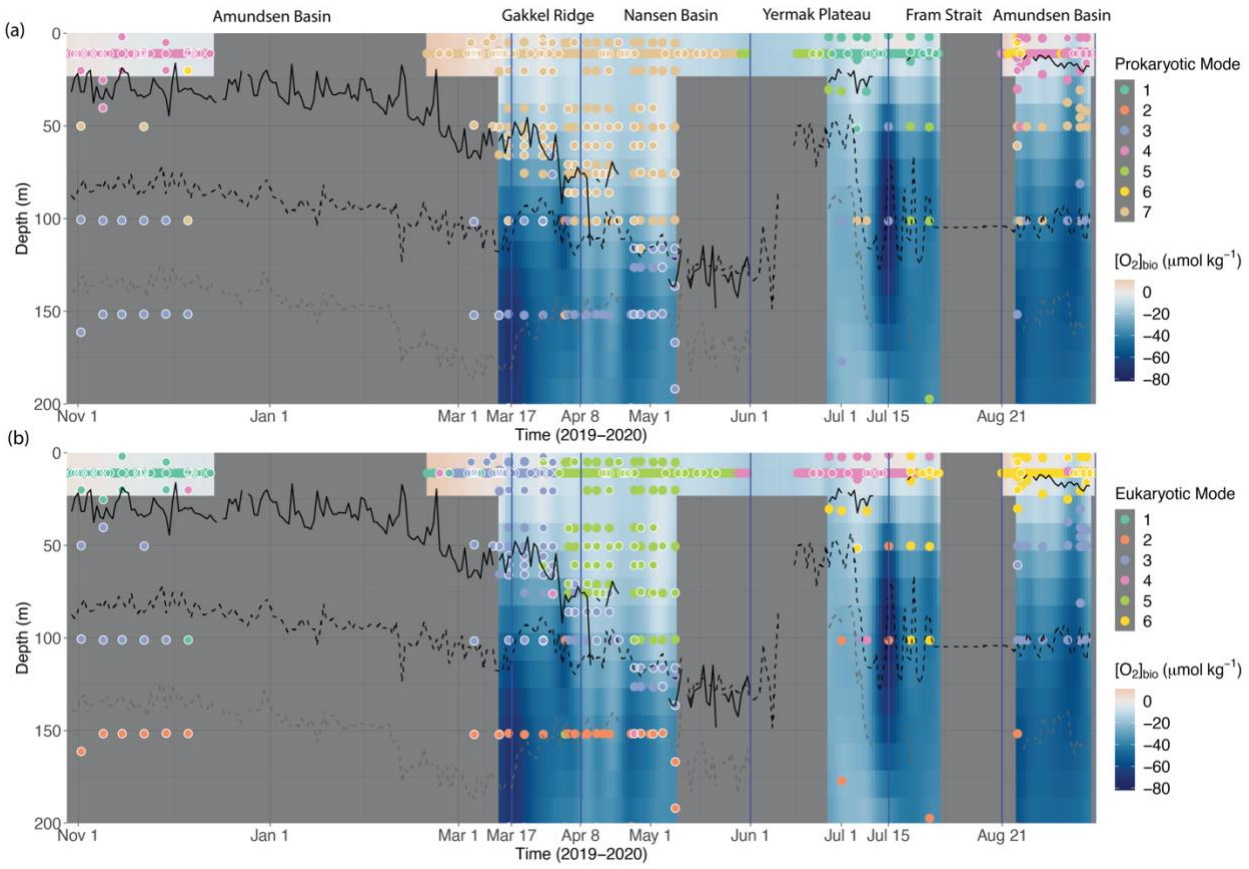


Figure 2-5. Time series of upper water column biological oxygen anomaly and taxonomic mode during the MOSAiC Drift track.
Interpolated biological oxygen anomaly, $[O_2]_{bio}$ (µmol kg^{-1}) with A) Prokaryotic Mode or B) Eukaryotic Mode along the upper 200 m of the MOSAiC Drift. Points representing individual collections for microbial community structure are colored by mode. Points encircled by white indicate no matching measurement of $[O_2]_{bio.}$ to include in the linear interpolation. Large blocks of white indicate periods of no data. Transit stations during periods of ship relocation are not shown. Vertical blue lines represent geographic transitions across the Arctic Ocean basins and are indicated at the top of the chart. The solid black line represents mixed layer depth and the dark grey dashed grey line indicates the top end of Artic Atlantic Water mass influence (zero degree isotherm).

In spring (early March–May), the upper water column (above and below mixed layer) was uniform in community structure (EM-5 and PM-7) until an abrupt transition from EM-3 to EM-5 within the mixed layer in early March — coinciding both with the return of sunlight and when the MOSAiC floe drifted over the Gakkel Ridge, entering the more saline and less stratified Nansen Basin (Figure 2-5b). $[O_2]_{bio}$ was mostly uniform above the mixed layer, hovering near or below 0, but spatially variable between casts (Figure 2-5). After crossing onto the Yermak Plateau, surface waters were undersaturated until late June during which a visible phytoplankton bloom caused local accumulation of $[O_2]_{bio}$ near the surface (Figure 2-4). At this time, surface stratification increased to the point that was no defined mixed layer (Figure 2-5) and strong negative $[O_2]_{bio}$ anomalies reached as shallow as 50 m depth, accompanied by shallow occurrences of deep-water modes EM-2 and PM-3 (Figure 2-4). At the surface (< 11 m), eukaryotic mode remained diatom dominant EM-4 until after reaching the Fram Strait in late July where it switched to the more pico-eukaryote composed EM-6 for the duration of the drift, including relocation to the Amundsen Basin. Following relocation however, EM-3 was once again observed beneath the mixed layer as it was during the previous year. The surface prokaryotic community in the Amundsen Basin was variable during late August, where low average salinity PM-6 was observed (Figure 2-6), but eventually returned to observed fall communities of the previous year (Figure 2-5).

Other primary seasonal and spatial occurrences of all modes are visualized in Figure 2-5 and summarized in Table 2-1. Generally, however, surface communities (< 100 m) contained unique modes above and below the mixed layer (Figure 2-5). The degree of $[O_2]_{bio}$ undersaturation in the water column was not always bound by the mixed layer, but tracked closely with microbial community structure, particularly in the surface (< 11 m; Figure 2-4).

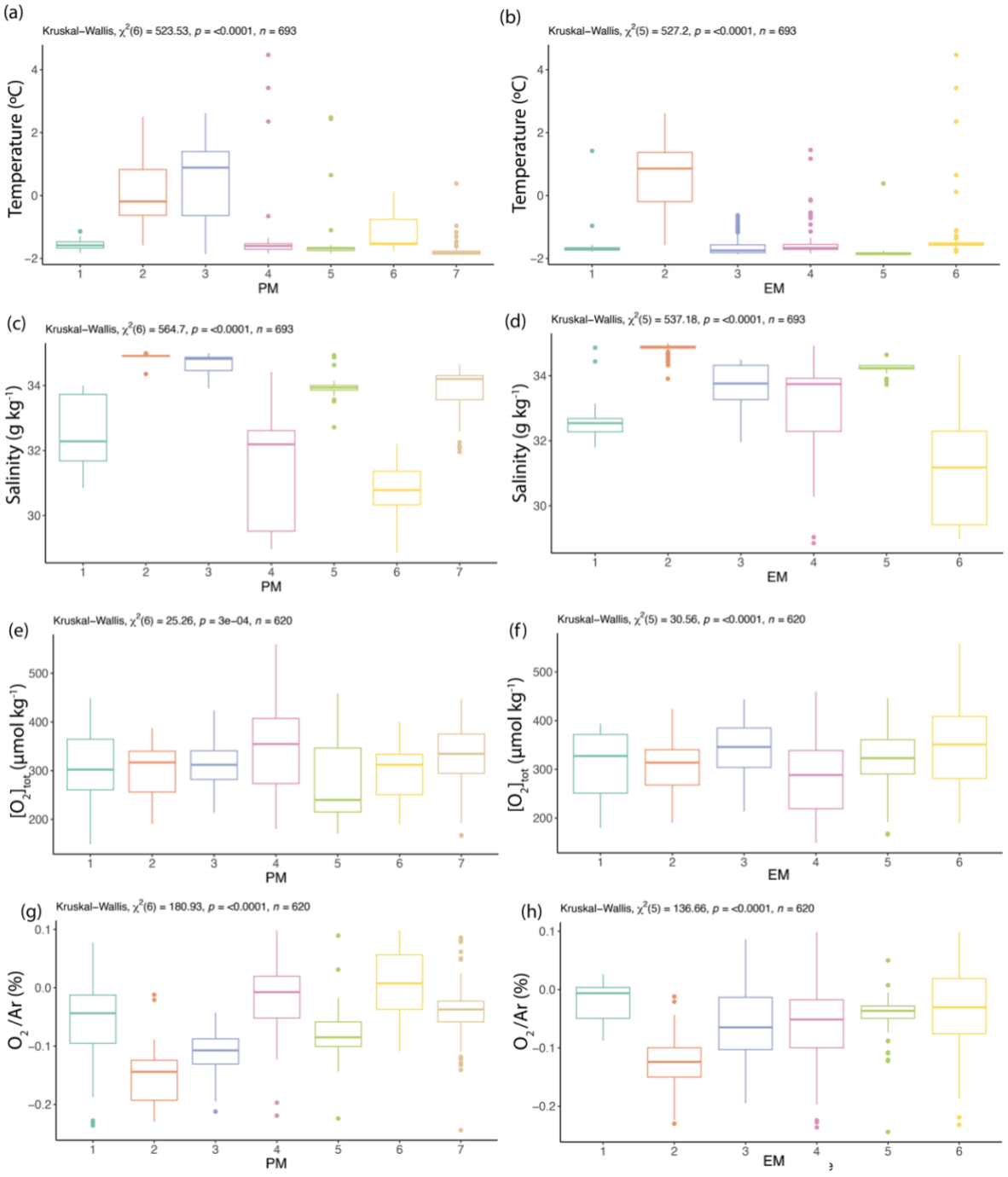


Figure 2-6. Environmental attributes of each mode.
Results of self-organizing map (SOM) segmentation for the prokaryotic and eukaryotic communities with each prokaryotic/eukaryotic mode (PM/EM) related to (a/b) temperature in ºC, (c/d) salinity in g kg^{-1}, (e/f) total dissolved oxygen in μmol kg^{-1}, and (g/h) O_2/Ar ratios, where paired samples were available. In each boxplot, the horizontal line shows the median, the box spans the first quartile to the third quartile, the whiskers show the minimum/maximum, and the dots represent outliers. Significance values for each box plot represent a non-parametric Kruskal-Wallis test.

Table 2-1. Summary characteristics for all prokaryotic (PM) and eukaryotic (EM) modes.

Mode	Seasonal/Spatial Dominance	Estimated Mean (16S)		Mean Chl-a ($\mu g\ L^{-1}$)	Primary Dates Observed	Primary Depths Observed
		Genome size (10^6 bp)	Doubling time (d^{-1})			
PM 1	Summer Bloom	3.13 ± 0.28	5.68 ± 0.53	0.78 ± 0.40	Jul	Surface (<50 m)
PM 2	Deep Water	3.05 ± 0.22	7.44 ± 0.79	0.001 ± 0.00	Year-round	Deep (>500 m)
PM 3	Atlantic Water	2.99 ± 0.12	8.02 ± 0.22	0.01 ± 0.008	Year-round	Mid (100–500 m)
PM 4	Winter Surface	2.78 ± 0.32	6.43 ± 0.66	0.25 ± 0.20	Sep–Dec	Surface (<50 m)
PM 5	Summer Surface	3.07 ± 0.27	6.29 ± 0.68	0.57 ± 0.27	Jun–Jul	Surface–Mid (<100 m)
PM 6	Transpolar Drift[1]	3.43 ± 0.33	4.66 ± 0.71	0.37 ± 0.15	Aug, Dec	Surface (<50 m)
PM 7	Winter-Spring	2.70 ± 0.18	7.51 ± 0.39	0.06 ± 0.08	Feb–May	Surface–Mid (<100 m)
EM 1	Fall-Winter	2.71 ± 0.30	6.78 ± 0.72	0.11 ± 0.15	Nov–Dec	Surface (<50 m)
EM 2	Deep Water	3.01 ± 0.17	7.80 ± 0.63	0.01± 0.01	Year-round	Mid–Deep (>150 m)
EM 3	Spring Dark	2.74 ± 0.19	7.64 ± 0.23	0.02 ± 0.03	Sep–May	Surface–Mid (<150 m)
EM 4	Summer Bloom	3.22 ± 0.29	5.75 ± 0.97	0.73 ± 0.39	Jun–Jul	Surface (<50 m)
EM 5	Spring Light	2.70 ± 0.20	7.47 ± 0.42	0.09 ± 0.09	Apr–May	Surface–Mid (<100 m)
EM 6	Summer-Fall	2.88 ± 0.31	6.16 ± 0.69	0.36 ± 0.18	Jul–Sep	Surface (<50 m)

1. Transpolar Drift here refers to the advective pathway of low salinity, river-rich surface water of Siberian origin

Random Forest regressions

Out of the 8 RF models trained (Figure 2-S6), the one built using only feature-selected 16S and 18S ASV relative abundances performed best on the validation data, explaining 69.31% of variance within the model and containing the smallest mean residuals (12.6; Table 2-S1). The model which statistically performed the least well was built using all 18S ASV relative abundances, but still explained 64.76% of model variance (Table 2-S1). Adding physical and chemical parameters to an ASV model did not increase model fidelity (Table 2-S1) and building models without using ASVs as predictors (Figures 2-S7g and 2-S7i) predicted undersaturation of up to 15 μmol kg^{-1} for net neutral or positive $[O_2]_{bio}$ and had larger average residuals from the 1:1 line (Table 2-S1).

After combining all samples to train the comprehensive ASV-based RF model (Figure 2-7), Boruta feature selection identified 179 ASVs as being statistically significant in making model predictions (Table 2-S4), which were then included as predictor variables in the final model. The cross-validation analysis of this model indicated a median % variance explained of 92.98, a median

average error of 3.42, a mean squared error of 33.94, and a median root mean squared error of 5.32 — indicating a confidence in predictions of approximately ± 5 µmol kg^{-1} (Figure 2-7). When this model was applied to an external dataset, here, the LOESS smoothed underway averages of $[O_2]_{bio}$ data, the model did not perform to this degree of prediction strength, with many values predicted to be 10–20 µmol kg^{-1} more undersaturated than the smoothed, measured values (Figure 2-7b). However, these predictions fell within the spread of values of the raw measured data (Figure 2-8).

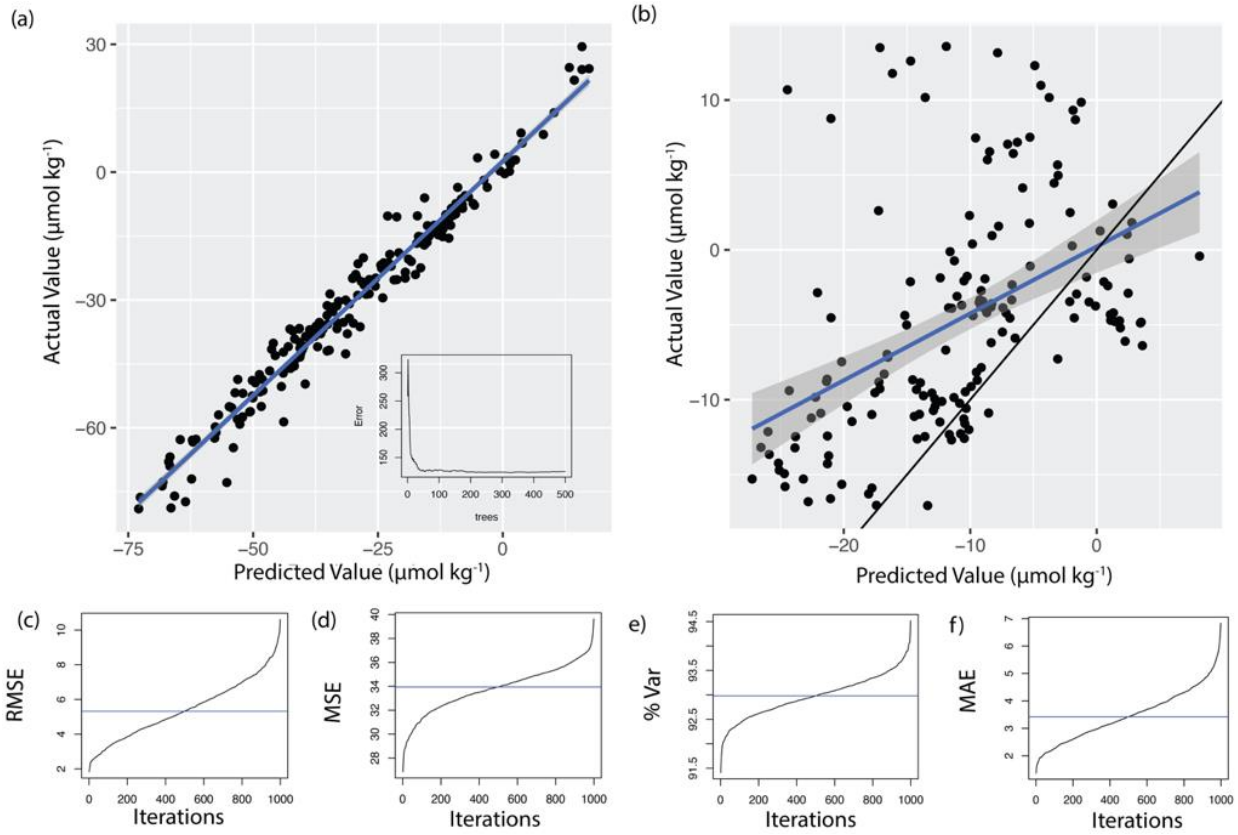


Figure 2-7. Final Random Forest (RF) model and validation of predictions for biological oxygen anomaly, $[O_2]_{bio}$ in µmol kg^{-1}.
(a) Comparison of predicted $[O_2]_{bio}$ and measured $[O_2]_{bio}$ values using a linear regression (p-value and r^2 reported in the top left) from the RF training dataset comprised of all paired discrete water column $[O_2]_{bio}$ measurements collected from the CTD rosette. (b) Comparison of predicted and measured $[O_2]_{bio}$ values using a linear regression (p-value and r^2 reported in the top left) from a validation dataset comprised of LOESS smoothed underway $[O_2]_{bio}$ measurements generated from 4-hour averages around the sampling time of discrete underway community structure measurements. (c) Root Mean-Squared Error (RMSE), (d) Mean Squared Error (MSE), (e) % Variance Explained (% Var), and (f) Mean Average Error (MAE) from a cross-validation analysis of the RF model predicting 10% of randomly withheld training data across 1000 iterations. The median of each statistic is represented by a blue horizontal line.

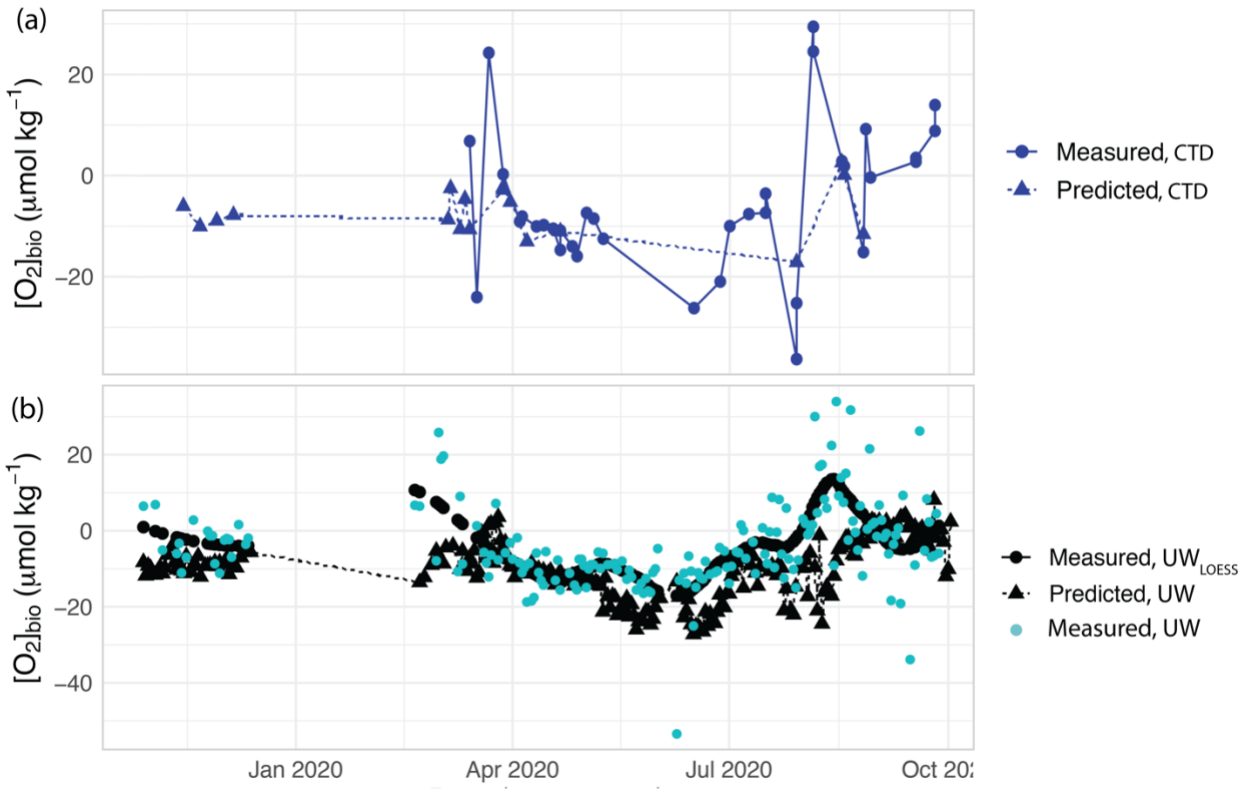


Figure 2-8. Biological oxygen anomaly ($[O_2]_{bio}$) in µmol kg^{-1} from the upper 11 m of the water column, measurements vs. Random Forest (RF) predictions.
(a) Discrete measurements (circles) and predictions (triangles, dashed) collected from the CTD rosette. (b) Continuous underway (UW) measurements integrated across the 4 hours closest to recorded sampling time for discrete community structure plotted in cyan, while the Loess local regression of underway averages (UW_{LOESS} circles) and underway predictions (triangles, dashed) are shown in black.

Values predicted from CTD community samples were closer to their measured counterparts and followed expected oceanographic trends such as Atlantic water influence and mixed layer depth during periods where there is no discrete water column data to compare to (Figure 2-9), effectively expanding this biogeochemical dataset.

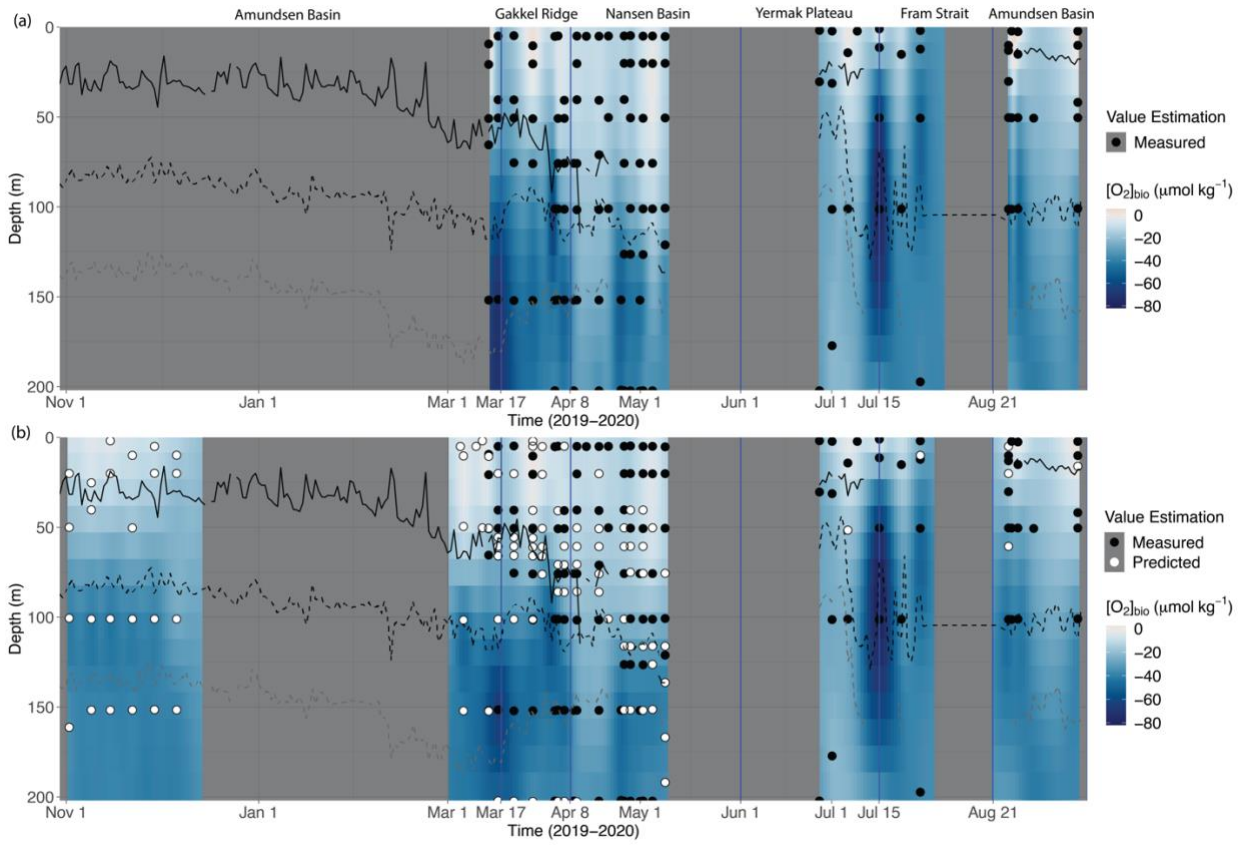


Figure 2-9. Upper water column discrete biological oxygen anomaly ($[O_2]_{bio}$), measurements vs. Random Forest (RF) predictions.
Multilevel B-spline approximation (Finley et al. 2022) interpolation of $[O_2]_{bio}$ measurements in the upper 100 m of the water column. All measured samples (black) were collected from the CTD rosette and predictions (white) were carried out using microbial community structure data run through a RF model trained using 137 paired oxygen anomaly and community structure samples. Both an interpolation of only measured (a) and measured and predicted (b) oxygen anomaly data are shown. Vertical blue lines represent times of geographic transitions across the Arctic Ocean basins, indicated at the top of the chart. The black solid line represents mixed layer depth, the dashed black line indicates the top end of Atlantic influence determined by where temperature begins to increase with depth, and the dashed grey line indicates the upper end of the Artic Atlantic Water mass (zero-degree isotherm).

Out of the 217 ASVs which were feature selected to use as predictor variables, 129 were 16S, and 50 were 18S (Table 2-S4). Similarly, top predictors (%IncMSE of 5 or more, n = 28) were made up of primarily prokaryotic taxa (n = 20; Figure 2-10). The ASV with the highest %IncMSE (8.1) was assigned to a reference genome of Methanomassiliicoccales, followed by the next 4 top predictors of Cytophagales (7.9), Candidatus Pelagibacter (7.8), an unidentified

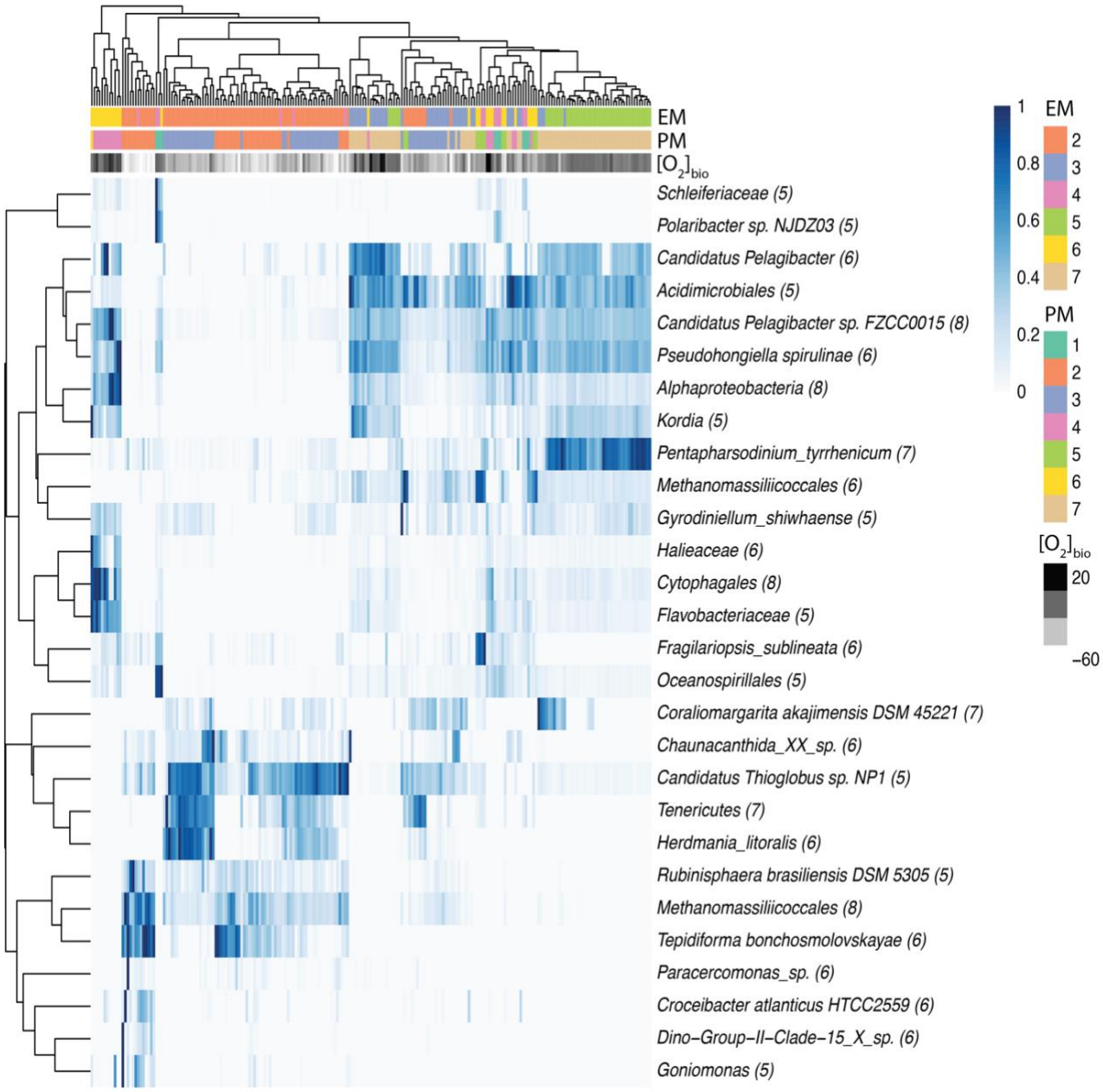


Figure 2-10. Heatmap of key microbial predictors of upper water column biological oxygen anomaly.

Scaled relative abundance of all predictor variables (ASVs) with a percent Increase in Mean Squared Error (%IncMSE) of over 5% when randomized in the final random forest model. Relative abundances were scaled using a z-score between 0–1. Bray-Curtis dissimilarity distances were used to cluster taxa and samples, which are additionally colored by mode (PM or EM) and $[O_2]_{bio}$. The highest resolved paprica taxonomic placement (closest relative among RefSeq or PR2 genomes) is used to identify each unique ASV, with the rounded %IncMSE included in parentheses. Heatmaps generated using R package pheatmap (Kolde 2019).

Alphaproteobacteria (7.7), and the dinoflagellate *Pentapharsodinium tyrrhenicum* (7.2). Overall, Archaea were over-represented in the top predictors, making up 15% of the top predictor ASVs compared to only 5% of the overall prokaryotic community composition. Eukaryotic taxa within the top predictors (n = 8) were 50% dinoflagellates. When looking at the relative abundance of these sequences across all paired samples, two distinctive groups (clustered using Bray-Curtis dissimilarity distances) of taxa emerge: those that were more abundant in samples with strong $[O_2]_{bio}$ undersaturation and those that were more abundant when $[O_2]_{bio}$ was closer to 0 or indicated net production (Figure 2-10). The former was made up of ASVs identified to deep-water taxa with diverse metabolisms (i.e. Thermoplasmata), dinoflagellates, or complex carbon degraders (i.e. Candidatus *Thioglobus*). The latter was made up of commonly particle associated taxa (i.e. Bacterioidetes, Flavobacteria), the one diatom ASV (*Fragilariposis sublineata*), and also dinoflagellates (Figure 2-10).

Discussion

This work reveals some of the key biological and environmental influences on oxygen-based NCP estimates and highlights a unique, year-long time series of paired microbial and biogeochemical data over a spatially and temporally variable drift track across the central Arctic Ocean. By assessing successional patterns in microbial community structure using SOMs, we track where and when key shifts in taxonomy occur and determine which species were dominant in association with specific periods or environmental variables throughout the time series. By assessing predictive relationships between microbial community structure and $[O_2]_{bio}$ using a RF regression, we explore the potential linkages between these observations and highlight key predictor taxa for this biogeochemical variable.

Drivers of observed community structure and potential NCP

In line with previous studies (Carter-Gates et al. 2020; Priest et al. 2023), we found that water mass influence played an important role in the environmental structuring of taxonomic distribution in the pelagic Arctic Ocean. For example, the presence of Atlantic water was an important driver of prokaryotic taxonomic community composition, particularly at depth. PM-3 appeared almost exclusively below the upper boundary of Arctic Atlantic Water mass influence and contained AAW environmental characteristics such as high temperature and salinity (Figure 2-6). PM-3 also had many ASVs identified to the Flavobacteria group (Figure 2-S2), which have been found to compose a large fraction of particle-associated microbes and may support high rates of sinking organic matter remineralization at these depths (Dithugoe et al. 2023). Yet, PM-3's most abundant ASVs were mostly distinct from those most abundant in the deeper PM-2, despite being assigned to many of the same reference genomes (Figure 2-2c) — potentially indicating some species specialization between the two habitats, which is not uncommon for Arctic microbes (Royo-Llonch et al. 2021), and could be driven by substrate availability and utilization or temperature effects (Carter-Gates et al. 2020). The Eukaryotic community was uniform at depth for the entirety of the MOSAiC Drift (Figure 2-S5), with deep-water mode EM-2 (dominated by Telonemia and other flagellates; Figure 2-S3) always present below the true demarcation of the AAW water mass between 100 and 150 m (Schulz et al. 2023; Figure 2-5).

In the water column above Atlantic water influence, disentangling seasonal and spatial influences on community structure was more complicated, with some combination of both likely driving compositional dynamics. PM-6, observed briefly in December 2019 and then again in late August of 2020, was abundant when the MOSAiC floe was in waters influenced by the Transpolar Drift, as indicated by a decrease in $\delta^{18}O$ isotopes (Nomura et al. 2023; Schulz et al. unpubl.). It's low average salinity (Figure 2-6), high abundances of freshwater indicator taxa Actinobacter

(Figure 2-S2), and ASVs assigned to complex carbon degraders (Figure 2-2), suggest a community impacted fresh riverine input. While there was no specific EM associated with the Transpolar Drift, mixed layer eukaryotic communities observed in the Amundsen Basin at the beginning and end of the MOSAiC Drift were segmented into distinct modes (EM 1 and EM 6, respectively), despite similarities in overall composition (Figure 2-3).

Other upper ocean successional patterns appear to be seasonally driven (eukaryotes) and/or substrate driven (prokaryotes), with mode community composition defined by metabolic strategy and often constrained vertically in the water column by the mixed layer depth. The community composition (Figure 2-3) and elevated Chl-a concentrations of EM-1 (Table 2-1) indicate the remnants of a diatom rich under ice bloom with a copiotrophic heterotrophic prokaryotic community (PM-1) and higher capacity for NCP (ΔO_2/Ar, Figure 2-6). Due to the winter data gap, it is unclear whether this community persisted or evolved over the course of the winter. What is clear, however, is that the deepening of the mixed layer in early March brought the primarily heterotrophic and dinoflagellate dominated eukaryotic EM-3 and largely chemotroph dominant PM-7 to the surface, persisting in the mixed layer until crossing over the Gakkel Ridge (EM-3) or Yermak Plateau (PM-7). While the abrupt shift in EMs observed in early April could be due to changes in surface water mass composition (leaving the river-water rich regime), this period also corresponded to the return of sunlight, an exponential increase in Chl-a (Hoppe et al. pers. comm.; Figure 2-2), and a community composition shift from predominantly heterotrophs to predominantly phototrophs (Figure 2-3). This, and the persistence of PM-7 indicating no community separation by basins, supports homogeneous selection pressures across basins (Aalto et al. 2022), and a seasonally driven vernal shift in the eukaryotic community. In addition to Thaumarcheota and other common winter taxonomic groups (Pedrós-Alió et al. 2023), PM-7 also

contained a relatively high proportion of Bacteroidetes taxa (Figure 2-S2) who are generally well adapted to respond quickly to organic matter. However, PM-7 also had the smallest average genome size (Figure 2-2), which is more characteristic of Proteorhodopsin containing Bacteroidetes taxa (Fernández-Gómez et al. 2013), making PM-7 both well suited for oligotrophic conditions during high Arctic winter and the return of light in Arctic spring. Following the onset of the summer melt season in late May (Shupe et al. 2022), we observed a near surface phytoplankton bloom, as described here by diatom rich EM-4 (Figure 2-3) followed by EM-6 which contained greater abundance of pico-plankton, likely due to lower nutrient availabilities in the post-bloom situation together with low salinities and stratification from both melt and increasing river influence within the Fram Strait (Schulz et al. unpubl.), all favoring this functional type (Li et al. 2009). The prokaryotic summer community shifted in response to the bloom conditions of EM-4 (Figure 2-4), although with a slight time lag as has been observed in other polar systems (Bowman et al. 2017). Pre-bloom PM-5 was enriched in Alphaproteobacteria with top ASVs assigned to terrestrial taxa (Figure 2-2), potentially from sediments picked up during ice formation melting out of the ice (Krumpen et al. 2020). PM-1, clustered closely with PM-4 and contained high abundances of similar, particle associated taxa (Figure 2-2), highlighting organic matter availability, here derived from either ice melt (Figure 2-6) or in situ phytoplankton growth (Figure 2-4) as a determining factor for bacterial taxonomic composition (Underwood et al. 2019; Piontek et al. 2021) over the MOSAiC year.

Despite these seasonal signals, the geographic transition between the Amundsen/Nansen Basins and Yermak Plateau may also have contributed to the shift in community composition, likely due to the presence of unmodified Atlantic Water while nearing the marginal ice zone (Schulz et al. unpubl.). Atlantic influenced water masses also reached much shallower depths at

this time (Figure 2-5) resulting in higher concentrations of macronutrients near the surface during a period of high light availability (Fong et al. unpubl.). Given these suitable productivity conditions and the observed phytoplankton bloom (EM-4; Figure 2-3), we would expect to also observe enhanced biological carbon drawdown (Juranek et al. 2019). However, we observed here a shallow and poorly defined mixed layer depth due to rapid ice melt (Salganik et al. 2023). While this stratification can intensify phytoplankton blooms, they can also suppress mixing and delay export by trapping the produced organic material (von Appen et al. 2021), prolonging remineralization at shallow depths. This could explain why bloom mode EM-4, while having the highest Chl-a, did not have the highest $\Delta O_2/Ar$ (Figure 2-6), with respiration from rapidly growing heterotrophs in PM-1 (low estimated doubling time; Table 2-1) outpacing oxygen accumulation in the integrated signal.

Throughout most of the MOSAiC Drift track, the strong density gradient of the halocline acted as a physical barrier for vertical exchange and dispersal of microbial communities (Figure 2-5), indicating that populations above and below this barrier likely developed independently of each other. $[O_2]_{bio}$ values above and below the mixed layer depth were also distinct, with a significant increase in O_2 drawdown below the active mixing layer, where most photosynthetic activity takes place (Figure 2-5). Exceptions to this pattern correspond to periods where active mixing was unusually deep (*i.e.*, eddy events in the Amundsen Basin during Fall 2020; Schulz et al. unpubl.) and is seen in a homogenization of both $[O_2]_{bio}$ and community composition (Figure 2-5). This trend is also followed by the predicted output of our RF model (Figure 2-9), encouraging confidence in these results.

Factors determining distribution and measurements of $[O_2]_{bio}$

In our data, the presence of the mixed layer corresponded with an $[O_2]_{bio}$ that was more uniform and generally higher in concentration within this layer (Figure 2-5). Given water column stratification (Schulz et al. unpubl.) and high ice extent (Nicolaus et al. 2022), it is possible that tracer residence times in this surface water were increased (Eveleth et al. 2014), accounting for the net oxygen accumulation observed at the end of winter, prior to the return of the sun on March 12th (Figure 2-5). If so, this decoupling in tracer $[O_2]_{bio}$ and active biogeochemical cycling could explain some of the poor predictive power in our ASV-based RF model for the higher magnitude and highly variable underway $[O_2]_{bio}$ samples (Figure 2-7).

Generally, $[O_2]_{bio}$ in the water column remained close to 0 or net heterotrophic at all depths for almost the entirety of the MOSAiC Drift (Figure 2-5). Even during summer, $\Delta O_2/Ar$ in the surface ocean (< 11 m) was, on average, undersaturated by 4.7%. These findings contrast with previous studies, which have found spatially variable, but on average supersaturated $\Delta O_2/Ar$ in the surface waters of the Eurasian Arctic (Ulfsbo et al. 2014; Eveleth et al. 2014; Schanke et al. 2021). However, most previous studies only report data from July–October, which is the period in our data that showed the greatest $[O_2]_{bio}$ accumulation from *in situ* production and ice melt (Figure 2-4). It is likely that external carbon sources, such as terrestrial input (Martens et al. 2022), would be required to sustain the levels of remineralization observed over the rest of the drift (Figure 2-5). Additionally, it is important to remember that the ice moved faster than the ocean, particularly upon leaving the Nansen Basin (Krumpen et al. 2021). This combined with greater current velocities and tidal influences (Rabe et al. 2022) removes assumptions of lateral homogeneity beneath the mixed layer (Schulz et al. unpubl.) and could explain the extreme heterogeneity of consumption rates between casts at this time (i.e. < –60 μmol kg^{-1} at 100 m; Figure 2-5). The negative spike on July 15th could be related to enhanced mixing from bottom topography when

passing the western slope of the Yermak Plateau. The positive spikes in $[O_2]_{bio}$ in the Amundsen and Nansen basins during early Spring are also likely linked to physical rather than biogeochemical processes given the decoupling from observed changes in community structure (Figure 2-5), low light levels, and predicted undersaturation of these samples from our ASV-generated RF model (Figure 2-8). Potential explanations include gas entrapment beneath sea ice or frontal dynamics due to crossing bathymetric features, i.e., the Gakkel ridge (Ulfsbo et al. 2014; Eveleth et al. 2014).

$[O_2]_{tot}$ concentrations as measured by the MIMS and presented here are slightly elevated when compared to other O_2 concentration measurements from the CTD rosette (Figure 2-S4). This could be due to various instrumental factors but is likely associated with the temperature offset between *in situ* values and minor warming by the time of sample calibrations and analysis (Text 2-S2). However, such changes would impact O_2 and Ar measurements equally, meaning the net biological oxygen saturation as derived from $\Delta O_2/Ar$ should not be affected. Although this could account for our anonymously high Ar supersaturation values when compared to previous studies (Eveleth et al. 2014), resulting in potentially overestimated magnitudes of net $[O_2]_{bio}$ concentrations. Future work could benefit from accounting for such temperature offsets by comparing MIMS $[O_2]_{tot}$ to other O_2 measurements and correcting values based on a standardized temperature adjustment.

Microbial communities as predictors of water column biogeochemistry

Outside of these features, water column net metabolic balance was tightly coupled with microbial diversity and had high predictability of $[O_2]_{bio}$ from microbial community structure alone (order of $\pm$ 5.32 μmol kg^{-1}, RMSE RF model; Figures 2-9, 2-8a). However, when validating this model with the underway dataset, deviations from measured values were larger (Figure 2-7). We suggest that this disconnect may stem from a few possible issues. First, by integrating the

continuous underway data over a 4–6 hr interval, the original signal may have diluted to the point where it decoupled from the point measurement of microbial community structure—especially as calibrations were often also performed around this time. Alternatively, in situ growth from the underway sampling tube and/or cell damage from pumping may have altered *in situ* community structure enough that the relative abundances of ASVs from the underway data were decoupled from patterns in the CTD data which were used to construct the RF model. And finally, it is possible that because the RF model was built using CTD data alone which, due to the temporal and vertical resolution of CTD casts, had a much smaller fraction of surface (< 15m) water samples and only 5 discrete samples which displayed $[O_2]_{bio}$ supersaturation. So, the undersaturation of predicted values may simply be due to underrepresentation of these sample types in the model training data (Table 2-S1). To better validate and constrain our RF model, we suggest future collections of similarly analyzed depth discrete paired DNA and O_2/Ar data in similar geographic locations to compare to and/or build a better predictive model constrained by multiple years of data. The more training data, the better suited such a RF model will be to predicting $[O_2]_{bio}$ in the highly variable Arctic basin.

Despite these complications in model validation, the overall accuracy of our RF predictions are very promising and support the hypothesis that biogeochemistry (here, $[O_2]_{bio}$) is reflected in microbial community composition and may be accurately predicted from these data, as has been shown in other model systems (Dutta et al. 2022). Since including physical parameters within the model did not significantly improve model accuracy (Table 2-S1), it is likely changes in those variables are naturally accounted for within community structure. However the model built only using physical parameters was not as accurate, indicating the relationship between community composition and $[O_2]_{bio}$ may actually hold greater predictive power than the physical drivers of

these processes. Using this method, we were able to expand our $[O_2]_{bio}$ dataset to periods of the MOSAiC Drift which did not originally have measurements taken (Figure 2-9), expanding this biogeochemical time series and increasing its usability in future studies and for modeling applications. Net biological oxygen production has been shown to be an important predictor in driving the air–sea CO_2 gradient, influencing maximal CO_2 uptake during periods of open water (Juranek et al. 2019) and making it an important variable when it comes to quantifying the Arctic Ocean's potential for net carbon sequestration. ASVs represented the best variables for predicting $[O_2]_{bio}$ across all model iterations that included community composition data (Table 2-S1), additionally highlighting the intricate connections between microbes and this important ecosystem service.

Prokaryotic processes in particular play an important role in the mechanisms driving net carbon export in the pelagic ocean (Worden et al. 2015). Given that 71% of the top predictor taxa in our final RF model belonged to prokaryotes (Figure 2-10), our results corroborate this, and other recent studies that indicate the intensity and signal of biological carbon export (as indicated by surface NCP, or here, $[O_2]_{bio}$) is strongly associated with prokaryotic community composition. Out of the eukaryotic ASVs which were identified as important $[O_2]_{bio}$ predictor taxa, the majority were heterotrophic or mixotrophic. A similar phenomenon was observed in Guidi et al. (2016), which examined planktonic networks driving carbon export in the subtropics and found that driving eukaryotic taxa in subnetworks correlated with carbon export were primarily dinoflagellates, or heterotrophic. Other studies from more nutrient rich environments, such as Lin et al. (2017) from the Southern Ocean and Wang et al. (2018) from the mid-Atlantic Bight, found a mix of both autotrophic and heterotrophic taxa, which were good predictors and/or strongly correlated with NCP estimates. Many of these, however, were associated with bloom conditions and when bloom

samples were excluded from analysis in the Wang et al. study, only cryptophytes and Bacteria remained associated with NCP rates. In our RF model, the key predictor taxa that were most abundant in net productive or net neutral $[O_2]_{bio}$ samples were also primarily made up of either taxa directly associated with CO_2 fixation, i.e. *Fragilariopsis sublineata,* dinoflagellate and Halieaceae ASVs, or, motile heterotrophic taxa associated with high levels of activity when responding to photosynthetically derived organic matter and bloom conditions, i.e. *Polaribacter,* other Flavobacteria, and Cytophagales ASVs (Bernardet and Nakagawa 2006; Simon et al. 2012; Li et al. 2023). An ASV assigned to *Fragilariopsis sublineata* was indeed dominant in the summer bloom-associated EM-4, and several Flavobacteria, including Polaribacter, ASVs were dominant in the bloom response mode PM-1. Flavobacteria and The most abundant predictor taxa in samples with strong oxygen consumption ($[O_2]_{bio}$) were primarily deep water taxa, including several ASVs assigned to the domain Archaea. In these previous studies, Archaea taxa were not clearly associated with NCP rates, however most assessed only surface water within the mixed layer. Our depth discrete water column data therefore provides a new insight into the predictive capacity of a less abundant microbial group. While predictive capacity does not necessarily equal mechanistic importance, several other recent studies have emphasized the role of chemoautotrophic Archaea (in addition to Bacteria) in altering carbon pools (Bayer et al. 2023) and regulating organic carbon export (Dithugoe et al. 2023) in the water column. Therefore, it is highly likely that the relative abundance of these key prokaryotic taxa may closely track changes in oxygen production and consumption rates, suggesting process-level biological mechanisms that directly influence the overall net metabolic balance and carbon sequestration potential in the Arctic Ocean.

Summary and implications

Our findings highlight the influence of microbial community composition on the net trophic state in the central Arctic Ocean. Vertical distributions in community structure were

constrained by water column stratification and mixing, while surface community structure (both prokaryotic and eukaryotic) shifted in response to seasonal and oceanographic changes over the course of the MOSAiC Drift. $[O_2]_{bio}$ values generally tracked changes in community composition and were predicted with high fidelity in our community structure based RF model. Oxygen accumulation (positive $[O_2]_{bio}$ and net photosynthetic activity) was observed only briefly during specific periods and exclusively within the mixed layer indicating a primarily net heterotrophic system. Additionally, most of the key predictor ASVs within our RF model were assigned to heterotrophic or mixotrophic taxa, highlighting the role of O_2 consumption in constraining total NCP within the water column. While there are other O_2 loss mechanisms which can impact this integrated signal but were not measured directly in conjunction with our data (such as zooplankton respiration and grazing), our results show that microbial community composition holds predictive power for this signal. This work can support NCP modelling efforts, critical to understanding the future of a changing Arctic.

Supplemental Material

Text 2-S1. Expanded materials and methods for taxonomic assignment and quality control

The paprica pipeline (Bowman and Ducklow 2015) places bacterial and archaeal reads onto a reference tree created from the concatenated 16S and 23S rRNA genes from all completed genomes in the Genbank RefSeq database (Haft et al. 2018). Eukaryotic reads are placed onto a reference tree created from the unique 18S rRNA genes present in the PR2 4.13.0 database (Guillou et al. 2013). Each amplicon sequence variant (ASV) is placed to either an internal or terminal branch on the tree (Nawrocki and Eddy 2013; Barbera et al. 2019; Czech et al. 2020). ASVs assigned to terminal completed genomes are described taxonomically by the highest resolution taxonomic ID. ASVs assigned to internal estimated genomes are described only by the

consensus taxonomy for all downstream nodes. For bacteria and archaea, paprica also uses the point of placement in the reference tree to estimate genomic characteristics such as genome size and 16S rRNA gene copy number (Erazo et al. 2021), as well as theoretical growth rate using the gRodon package (Weissman et al. 2021). However, growth rate estimates from codon usage do not reflect temperature conditions and therefore represent theoretical maximums only used to compare between samples and not as *in situ* rates. Taxonomic associations as used here indicate only that this is the closest relative among genomes in the database and does not necessarily indicate the presence of that exact strain. Final read counts for 16S ASVs were normalized to estimated 16S rRNA gene copy number before calculations of relative abundance. Rare ASVs with an abundance of <10, samples with bad library builds (<5000 total ASV counts), and suspected mitochondrial sequences were filtered out prior to all downstream analyses. Suspected chloroplasts (16S) and intracellular parasite (18S) sequences were also filtered out prior to SOM segmentation but were included in select downstream RF model designs.

Text 2-S2. Expanded materials and methods for Membrane Inlet Mass Spectrometry (MIMS)

MIMS measurements were logged continuously using Pfeiffer Quadera software. Seawater from the *Polarstern* underway system was run continuously to a multi-cylinder equilibration chamber, which gravity fed prefiltered (100 µm) underway seawater or prefiltered underway seawater bubbled with gas calibration standards at an inflow rate of 11–15 mL min^{-1}. The mesh was changed every 1–2 days to prevent accumulating biological material from affecting gas measurements and the equilibration chamber was disassembled, washed, and thoroughly disinfected in regular intervals of ~3 weeks. Equilibration chamber temperature was monitored

using a temperature probe and was generally 2.5 ± 1°C warmer than the external *in situ* seawater temperatures due to slight warming during the water transport. The instrument was 2-point calibrated daily, and prior to any bottle sampling event. Instrument baselines for O_2 and Ar were obtained by measuring seawater purged with N_2 gas (AirLiquide ALPHAGAZ 1 Nitrogen ≥ 99.999%) for 40 minutes. 21% and 1% abundances for O_2 and Ar, respectively, were obtained by measuring seawater equilibrated with synthetic air (AirLiquide ALPHAGAZ 1; 21.03 ± 0.42% O_2, 1.00 ± 0.02% Ar) for 40 minutes. Baseline values were linearly interpolated between the single determinations and subtracted from all sample measurements and synthetic air calibration measurements to account for instrument drift over time.

During CTD collection of discrete samples, gas samples were always collected first, and clean tubing was used to transfer water from the Niskin. BOD bottles were allowed to overflow 3x to prevent gas exchange during the sampling process. Bottles were stored in ice water in the dark (0 ± 1ºC) until analysis, which either happened concurrently with, or immediately following, sampling of the rosette. During these bottle sampling events, instrument inflow was switched to a peristaltic pumped flow system (13–18 mL min^{-1}), which could draw from both the equilibration chamber for calibration and the discrete bottle samples. Bottle samples were analyzed until a plateau was reached. On 5 occasions, replicate (between 3–5) BOD bottles were collected from the same depth and analyzed to constrain standard deviation between bottle measurements.

Supplemental Table 2-S1. Summary statistics and input data for all Random Forest models.

[1] Relative abundance of all 16S amplicon sequence variants, with potential chloroplast sequences
[2] Relative abundance of all 18S amplicon sequence variants, with intracellular parasite sequences
[3] Relative abundance of QC'd 16S amplicon sequence variants
[4] Relative abundance of QC'd 18S amplicon sequence variants
[5] Sample temperature (ºC), salinity (g/kg), total oxygen concentration (μmol kg^{-1}), depth (m), and latitude (degrees)
[6] Sample mean estimated genome size and growth rate
[7] Sample prokaryotic mode and eukaryotic mode
* With Boruta feature selection
$^{+}$ Final predictions without a true validation value

Predictor Variables (Type)	Predictor Variables (#)	# at Split	# of Samples T–V	Mean squared residual	% Var Explain	Top 5 Predictor Variables (% IncMSE)
[1]	4055	1351	149 –68	175.8	66.36	Flavobacteriaceae (7.45) Cytophagales (6.99) Sedimenticola thiotaurini (6.68) Sulfuriroseicoccus oceanibius (6.55) Coraliomargarita akajimensis (6.40)
[2]	10003	3334	149 –68	187.4	64.13	Eukaryota (10.84) Gyrodiniellum shiwhaense (8.10) Phaeocystis antarctica (7.76) Amphidinium mootonorum (6.82) Cryptomondales (6.08)
[1] + [2]	14057	4685	149 –68	178.6	65.82	Sedimenticola thiotaurini (5.93) Eukaryota (5.83) Cytophagales (5.78) Flavobacteriaceae (5.72) Methanomassiliicoccales (5.58)
[3] + [4]	8534	2844	149 –68	184.2	64.76	Cytophagles (7.16) Candidatus Pelagibacter sp. (6.91) Flavobacteriaceae (6.69) Sedimenticola thiotaurini (6.12) Coraliomargarita akajimensis (5.89)
[3] + [4] *	171	57	149 –68	160.4	69.31	Coraliomargarita akajimensis (8.09) Flavobacteriaceae (6.91) Candidatus Pelagibacter sp. (6.87) Cytophagales (6.76) Alphaproteobacteria (6.2)
[3] + [4] + [5] + [6] + [7]	8545	2850	149 –68	181.8	65.21	Cytophagales (6.94) Alphaproteobacteria (6.36) Flavobactericeae (6.18) Candidatus Pelagibacter sp. (5.84) Eukaryota (5.28)
[5] + [6] + [7]	9	3	149 –68	166.9	68.07	potential temperature (24.32) depth in m (20.75) Prokaryotic Mode (13.13) practical salinity (12.82) Eukaryotic Mode (12.24)
[5]	5	1	149 –68	180.3	65.5	potential temperature (25.29) depth in m (20.26) practical salinity (16.86) latitude (15.67) $[O_2]_{tot}$ in µmol kg^{-1} (11.24)
[3] + [4] *	179	59	217 – NA 157 CTD^{+} 236 UW^{+}	125.2	74.03	Methanomassiliicoccales (8.13) Cytophagales (7.91) Candidatus Pelagibacter sp. (7.77) Alphaproteobacteria (6.3) Pentapharsodinium tyrrhencium (6.2)

Supplemental Table 2-S2. P-values from the Wilcoxon signed-rank test comparing environmental variables between prokaryotic modes.

Temperature						
	PM 2	PM 3	PM 4	PM 5	PM 6	PM 7
PM 1	$p < 0.0001$	$p < 0.0001$	$p = 0.024$	$p = 0.00023$	$p = 0.011$	$p < 0.0001$
PM 2	x	$p = 0.00063$	$p < 0.0001$	$p < 0.0001$	$p < 0.0001$	$p < 0.0001$
PM 3	x	x	$p < 0.0001$	$p < 0.0001$	$p < 0.0001$	$p < 0.0001$
PM 4	x	x	x	$p = 0.0039$	$p = 0.00053$	$p < 0.0001$
PM 5	x	x	x	x	$p < 0.0001$	$p < 0.0001$
PM 6	x	x	x	x	x	$p < 0.0001$
Salinity						
PM 1	$p < 0.0001$	$p < 0.0001$	0.00011	$p < 0.0001$	$p < 0.0001$	$p < 0.0001$
PM 2	x	$p < 0.0001$	$p < 0.0001$	$p < 0.0001$	$p < 0.0001$	$p < 0.0001$
PM 3	x	x	$p < 0.0001$	$p < 0.0001$	$p < 0.0001$	$p < 0.0001$
PM 4	x	x	x	$p < 0.0001$	0.022	$p < 0.0001$
PM 5	x	x	x	x	$p < 0.0001$	0.045
PM 6	x	x	x	x	x	$p < 0.0001$
$[O_2]_{tot}$						
PM 1	0.98	0.52	0.048	0.26	0.91	0.067
PM 2	x	0.41	0.0011	0.19	0.76	0.003
PM 3	x	x	0.01	0.042	0.59	0.0074
PM 4	x	x	x	0.0016	0.093	0.22
PM 5	x	x	x	x	0.56	0.0061
PM 6	x	x	x	x	x	0.13
O_2/Ar						
PM 1	$p < 0.0001$	$p < 0.0001$	0.008	0.12	0.033	0.58
PM 2	x	$p < 0.0001$	$p < 0.0001$	$p < 0.0001$	$p < 0.0001$	$p < 0.0001$
PM 3	x	x	$p < 0.0001$	0.001	$p < 0.0001$	$p < 0.0001$
PM 4	x	x	x	$p < 0.0001$	0.31	0.0012
PM 5	x	x	x	x	0.0041	0.00023
PM 6	x	x	x	x	x	0.023

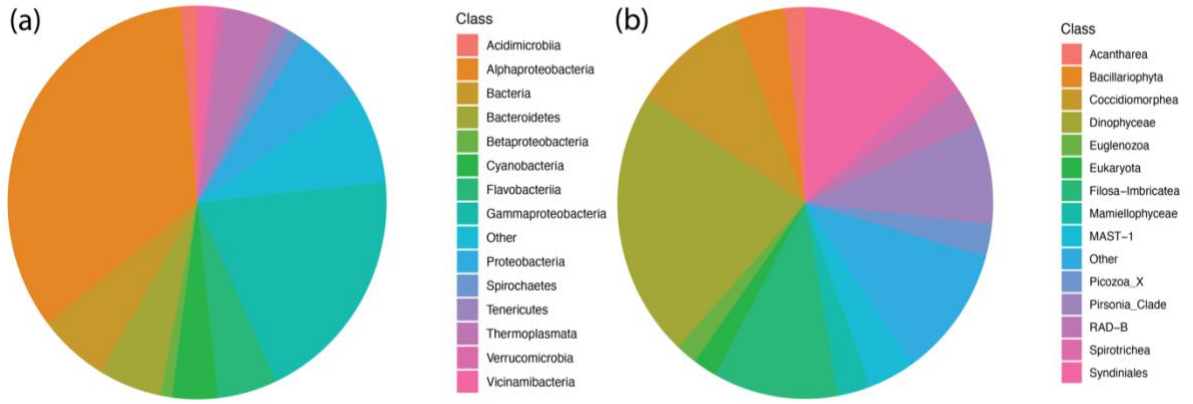
Supplemental Table 2-S3. P-values from the Wilcoxon signed-rank test comparing environmental variables between eukaryotic modes.

Temperature					
	EM 2	EM 3	EM 4	EM 5	EM 6
EM 1	$p < 0.0001$	$p = 0.04$	$p = 0.0018$	$p < 0.0001$	$p < 0.0001$
EM 2	x	$p < 0.0001$	$p < 0.0001$	$p < 0.0001$	$p < 0.0001$
EM 3	x	x	$p < 0.0001$	$p < 0.0001$	$p < 0.0001$
EM 4	x	x	x	$p < 0.0001$	$p < 0.0001$
EM 5	x	x	x	x	$p < 0.0001$
Salinity					
EM 1	$p < 0.0001$	$p < 0.0001$	$p < 0.0001$	$p < 0.0001$	$p < 0.0001$
EM 2	x	$p < 0.0001$	$p < 0.0001$	$p < 0.0001$	$p < 0.0001$
EM 3	x	x	0.02	$p < 0.0001$	$p < 0.0001$
EM 4	x	x	x	$p < 0.0001$	$p < 0.0001$
EM 5	x	x	x	x	$p < 0.0001$
$[O_2]_{tot}$					
EM 1	0.66	0.091	0.39	0.68	0.083
EM 2	x	$p < 0.0001$	0.064	0.063	0.00033
EM 3	x	x	0.00028	0.025	0.86
EM 4	x	x	x	0.012	0.0005
EM 5	x	x	x	x	0.023
O_2/Ar					
EM 1	$p < 0.0001$	0.012	0.017	0.015	0.6
EM 2	x	$p < 0.0001$	$p < 0.0001$	$p < 0.0001$	$p < 0.0001$
EM 3	x	x	0.91	0.07	0.049
EM 4	x	x	x	0.21	0.065
EM 5	x	x	x	x	0.3

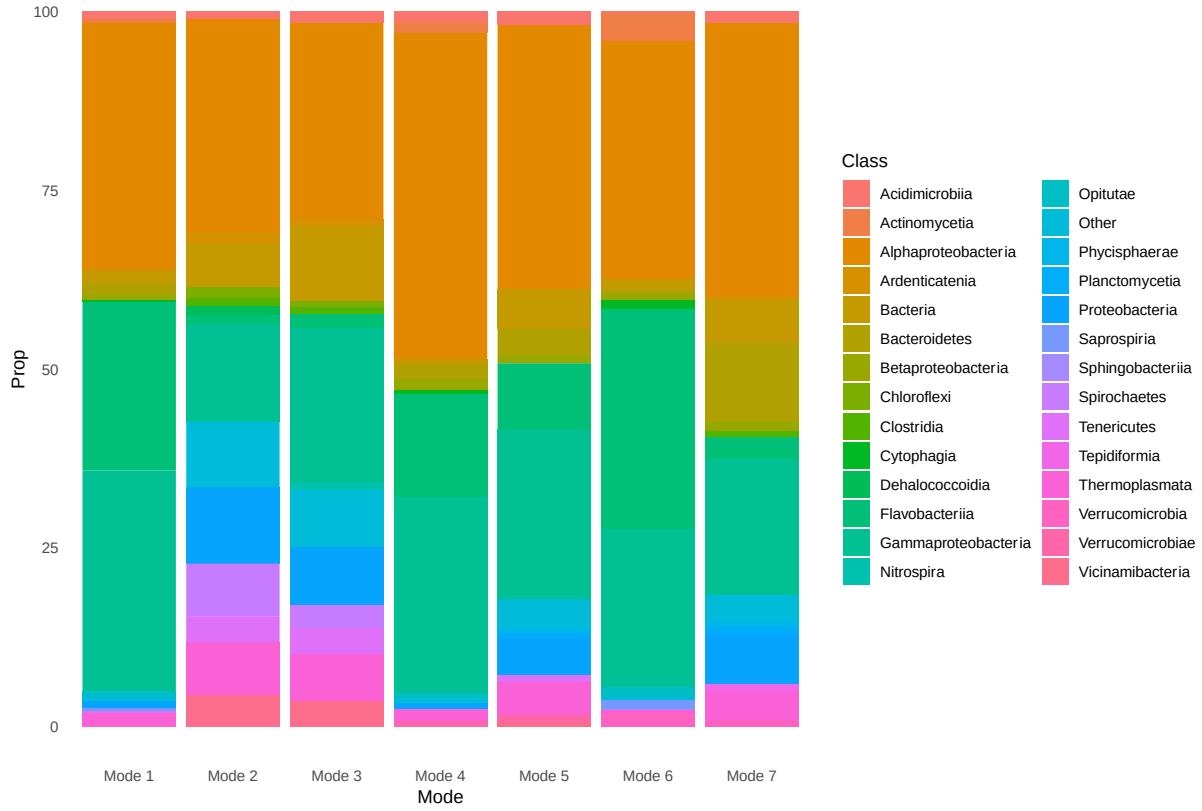
Supplemental Table 2-S4. Boruta selected ASVs (final model).

Available for download:

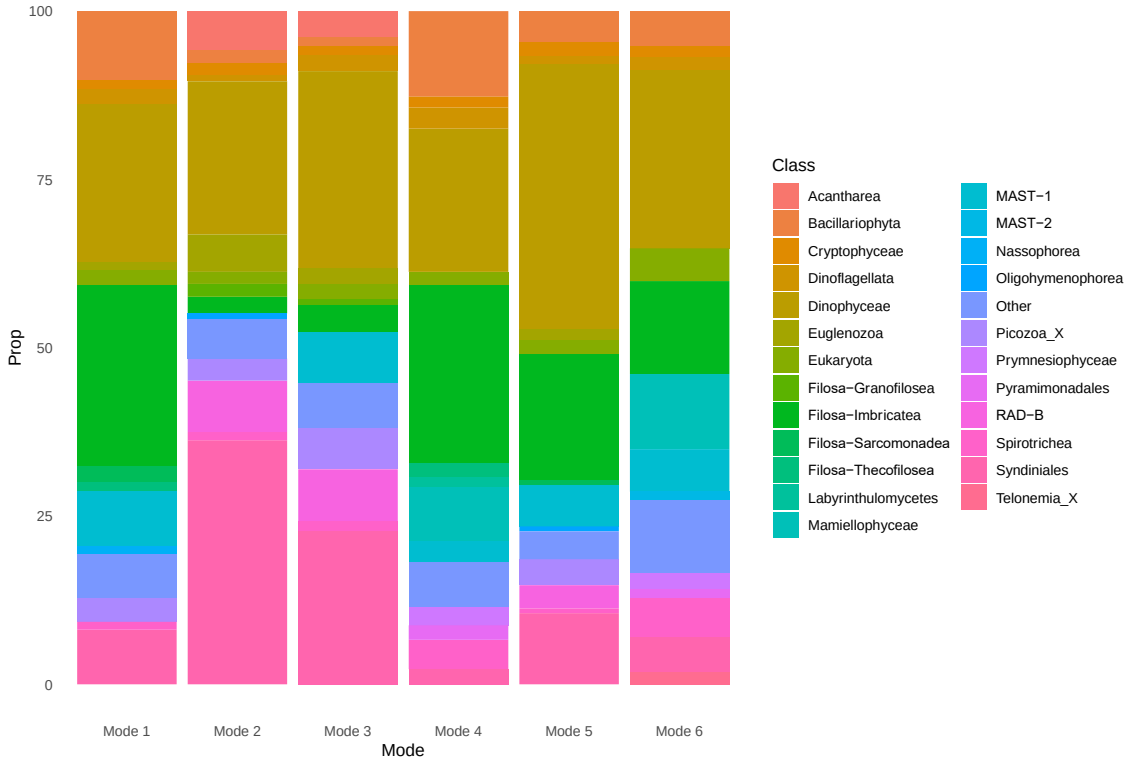
Supplemental File 2-S1. Chamberlain_Table_2-S4_predictor_ASVs.xlsx



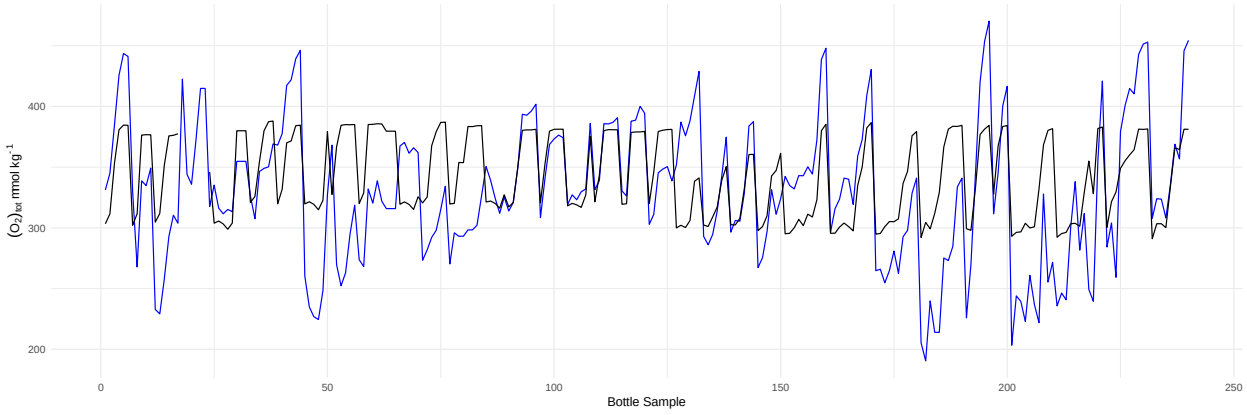
Supplemental Figure 2-S1. Overview of full dataset taxonomic community structure. Relative percentage of most abundant prokaryotic (a) and eukaryotic (b) taxonomic classes (or next lowest order of classification) across all water samples. Calculated using class-rank summed relative abundances from each unique 16S (a) and 18S (b) amplicon sequence variants (ASV), where the taxonomic assignment is representative of the most closely related completed reference genome in either the Refseq (16S) or PR2 (18S) databases.



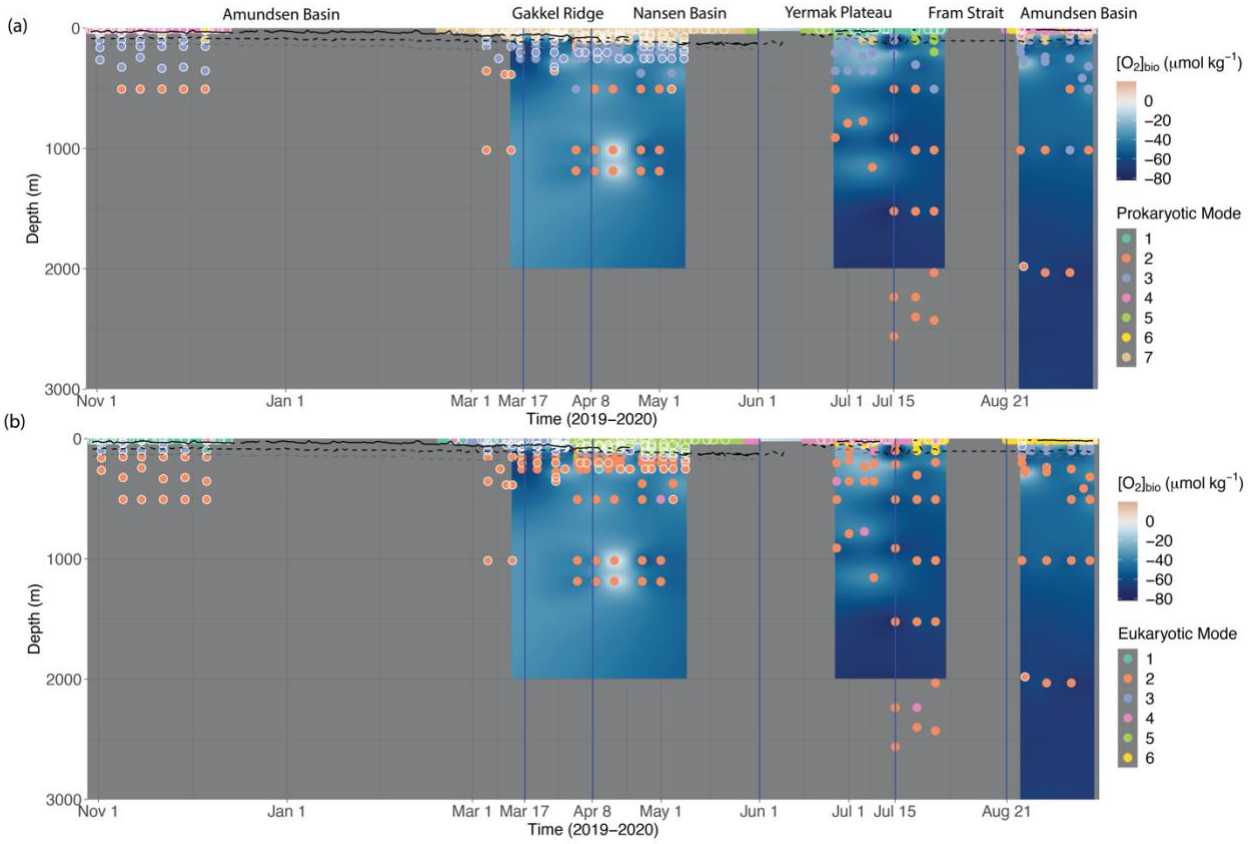
Supplemental Figure 2-S2. Overview of prokaryotic SOM taxonomic community structure. Average percentage of unique 16S amplicon sequence variants within each taxonomic class across each mode. Calculated using average relative abundance summed across each taxonomic identification.



Supplemental Figure 2-S3. Overview of eukaryotic SOM taxonomic community structure. Average percentage of unique 18S amplicon sequence variants within each taxonomic class across each mode. Calculated using average relative abundance summed across each taxonomic identification.



Supplemental Figure 2-S4. Comparison between mass spectrometer measurements and CTD sensor measurements of total dissolved oxygen concentration $[O_2]_{tot}$. Values are plotted sample by sample where black lines indicate $[O_2]_{tot}$ read directly from the CTD sensors and blue lines are calculated from the membrane inlet mass spectrometer from calibration standard derived translations between spectrometry readings and concentration using practical salinity and potential temperature.

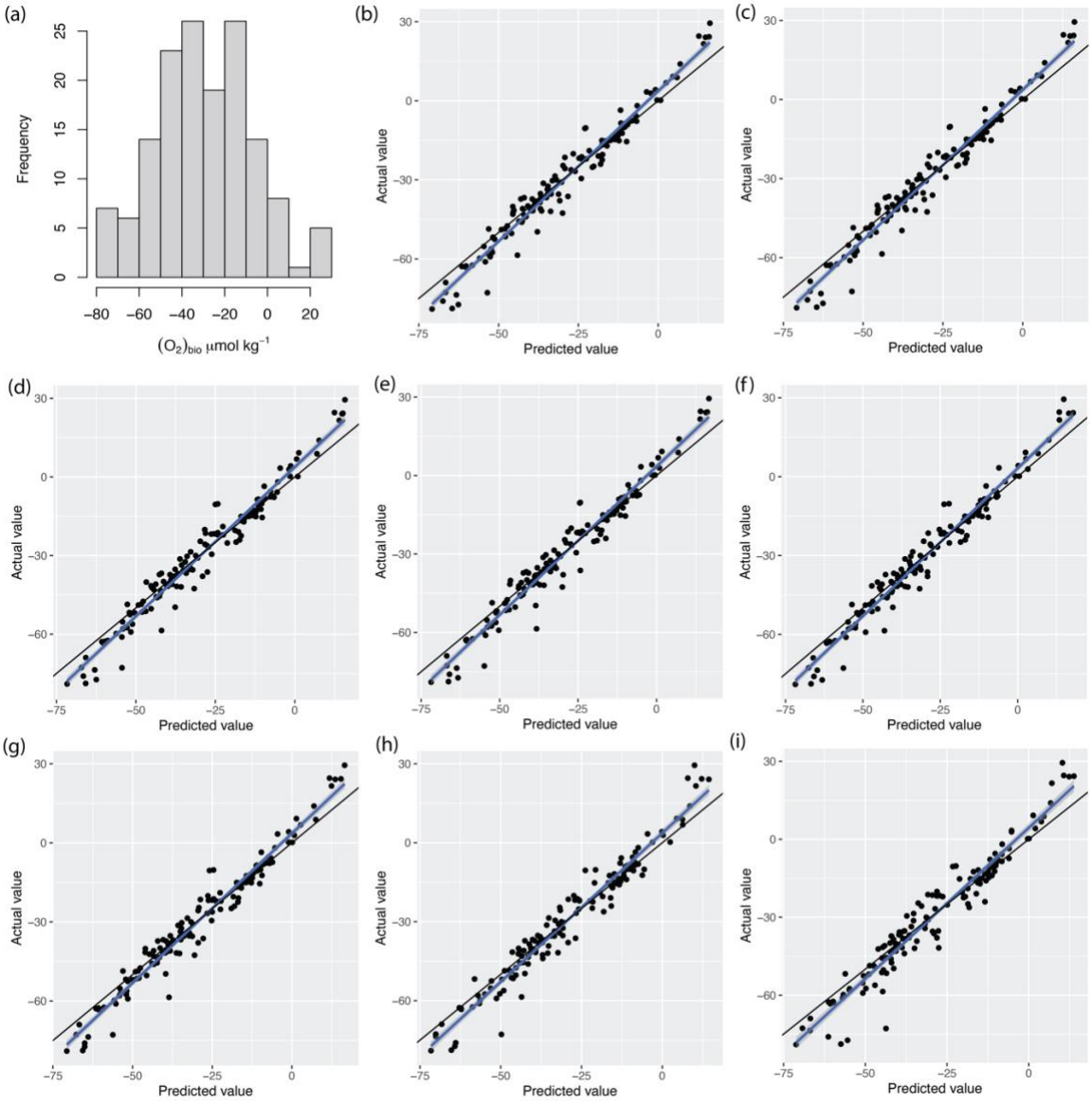


Supplemental Figure 2-S5. Time series of full water column biological oxygen anomaly and taxonomic mode during the MOSAiC Drift track.

Interpolated biological oxygen anomaly, $[O_2]_{bio}$ (μmol kg^{-1}) with Prokaryotic Mode (a) or Eukaryotic Mode (b) along the MOSAiC Drift. Points representing individual collections for microbial community structure are colored by mode. Points encircled by white indicate no matching measurement of $[O_2]_{bio}$. to include in the linear interpolation (package citation). Large blocks of white indicate periods of no data, and transit stations during periods of ship relocation are not shown. Vertical blue lines represent geographic transitions across the Arctic Ocean basins and are indicated at the top of the chart. The solid black line represents mixed layer depth and the dashed grey line indicates the top end of Artic Atlantic Water mass influence (zero isotherm).

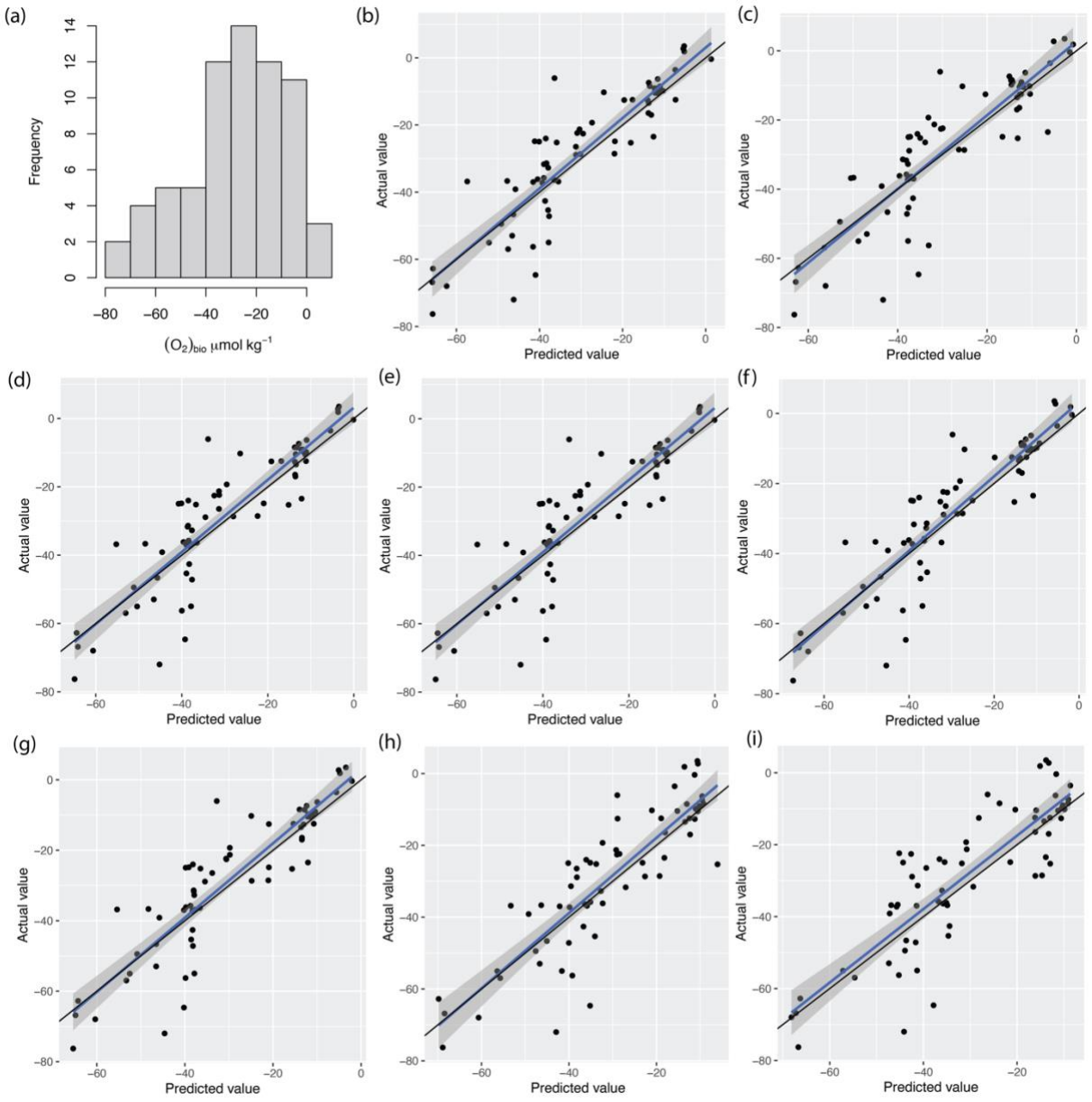
Supplemental Figure 2-S6. Training of iterative Random Forest (RF) Regression models using 70 % of available paired data.

(a) Distribution of biological oxygen anomaly ($[O_2]_{bio}$ in μmol kg^{-1}) values within the training dataset used to construct RF models. Training data comprised of a randomized selection of samples of the discrete surface ocean (≤ 100 m) $[O_2]_{bio}$ measurements with corresponding predictor variables, equating to 70 % of available paired data. (b-i) Comparisons of measured $[O_2]_{bio}$ (μmol kg^{-1}) and model predicted $[O_2]_{bio}$ (μmol kg^{-1}) for the training dataset when run through the trained RF model. Individual samples presented against a 1:1 line (black) and assessed with a linear regression model (blue, p-value and r^2 reported top left of each graph). Each panel shows the results of a model built using the following combinations of predictor variables: (b) the relative abundances (RA) of all 16S ASVs present within the sample sets – (c) the RA of all 18S ASVs present within the sample set – (d) the RA of all 16S and 18S ASVs present with the sample set – (e) the RA of all 16S and 18S ASVs present within the sample set, not including suspected chloroplast and intracellular parasite sequences – (f) the RA of feature selected 16S and 18S ASVs – (g) the RA of all 16S and 18S ASVs present within the sample set, not including suspected chloroplast and intracellular parasite sequences, physiochemical variables of potential temperature, practical salinity, and total oxygen concentration, depth, latitude, prokaryotic mode, eukaryotic mode, and estimated bacterial growth rate and genome size – (h) the physiochemical variables of potential temperature, practical salinity, and total oxygen concentration, depth, latitude, prokaryotic mode, eukaryotic mode, and estimated bacterial growth rate and concentration, depth, and latitude. Insets are representative of model error (y axis) with the experimental addition of more decision trees (x axis) in the model. Cut-offs for all constructed models was 500 decision trees.


(a)
Frequency
0
5
10
15
20
25
−80 −60 −40 −20 0 20
$(O_2)_{bio}$ µmol kg^{-1}
(b)
(c)
(d)
(e)
(f)
(g)
(h)
(i)
Actual value
Predicted value
30
0
−30
−60
−75 −50 −25 0

Supplemental Figure 2-S7. Validation of iterative Random Forest (RF) Regression models using 30 % of available paired data.

(a) Distribution of biological oxygen anomaly ($[O_2]_{bio}$ in) values within the validation dataset used to construct RF models. Validation data comprised of a randomized selection of samples of the discrete surface ocean (≤ 100 m) $[O_2]_{bio}$ measurements with corresponding predictor variables, equating to 30 % of available paired data. (b-i) Comparisons of measured $[O_2]_{bio}$ (μmol kg^{-1}) and model predicted $[O_2]_{bio}$ (μmol kg^{-1}) for the validation dataset when run through the trained RF model. Individual samples presented against a 1:1 line (black) and assessed with a linear regression model (blue, p-value and r^2 reported top left of each graph). Each panel shows the results of a model built using the following combinations of predictor variables: (b) the relative abundances (RA) of all 16S ASVs present within the sample sets – (c) the RA of all 18S ASVs present within the sample set – (d) the RA of all 16S and 18S ASVs present with the sample set – (e) the RA of all 16S and 18S ASVs present within the sample set, not including suspected chloroplast and intracellular parasite sequences – (f) the RA of feature selected 16S and 18S ASVs – (g) the RA of all 16S and 18S ASVs present within the sample set, not including suspected chloroplast and intracellular parasite sequences, physiochemical variables of potential temperature, practical salinity, and total oxygen concentration, depth, latitude, prokaryotic mode, eukaryotic mode, and estimated bacterial growth rate and genome size – (h) the physiochemical variables of potential temperature, practical salinity, and total oxygen concentration, depth, latitude, prokaryotic mode, eukaryotic mode, and estimated bacterial growth rate and genome size – and (i) only the physiochemical parameters of potential temperature, practical salinity, total oxygen concentration, depth, and latitude.


(a)
Frequency
0
2
4
6
8
10
12
14
−80
−60
−40
−20
0
$(O_2)_{bio}$ µmol kg^{-1}
(b)
(c)
(d)
(e)
(f)
(g)
(h)
(i)
Actual value
Predicted value
−80
−60
−40
−20
0

Acknowledgements

Chapter 2, in part, has been submitted for publication of the material as it may appear in Limnology & Oceanography, 2024, Chamberlain, Emelia J; Rokitta, Sebastian; Rost, Björn; D'Angelo, Alessandra; Creamean, Jessie; Loose, Brice; Ulfsbo, Adam; Fong, Allison A; Hoppe, Clara JM; Droste, Elise; Nomura, Daiki; Schulz, Kirstin; Bowman, Jeff., Wiley Periodicals LLC, 2024. The dissertation author was the primary researcher and author of this paper.

The data used in this research were produced as part of the Multidisciplinary Drifting Observatory for the Study of Arctic Climate (MOSAiC) project which are published under the data tag MOSAiC20192020 and project ID AWI_PS122_00. We thank all those who contributed to MOSAiC and made this endeavour possible (Nixdorf et al. 2021). We would specifically like to acknowledge the members of Team Ocean their efforts in processing and making the CTD data and other core hydrographic parameters available. Additionally, Laura Heitmann, Tina Brenneis, and Anja Terbrüggen, Jacqueline Stefels, and Annette Rinke for help with sampling and processing Chl-a and DNA samples. We would also like to thank current and former UCSD Bowman Lab members, particularly Jesse Wilson, Avishek Dutta, Beth Connors, Natalia Erazo, Benjamin Klempay, and Riley Hale, for assistance with extracting and analyzing DNA samples and random forest analysis.

We would additionally like to acknowledge the funding which made this research possible. Primary scientific analyses and EJC, JB, AD, and BL were funded through the US National Science Foundation (NSF) award OPP 1821911 to PI JB. This award also supported travel for JMC. EJC was additionally funded through an NSF Graduate Research Fellowship. JMC was supported by the U.S. Department of Energy's Atmospheric System Research (DOE ASR) program (DE-SC0019745). DN was supported by the Japan Society for the Promotion of Science (grant numbers: JP18H03745; JP18KK0292; JP17KK0083; JP17H04715; JP20H04345) and by a

grant from the Joint Research Program of the Japan Arctic Research Network Center. AU was supported by the Swedish Research Council Formas (grant no. 2018-01398). AAF was supported by MOSAiC_ECO funds provided by the Alfred-Wegener-Institut Helmholtz Zentrun für Polar- und Meeresforschung. ESD was supported by UK Natural Environment Research Council (NERC) through the EnvEast Doctoral Training Partnership (NE/ L002582/1), as the Department for Business, Energy & Industrial Strategy (BEIS) through the UK Arctic Office. This work was also supported by the Swedish Polar Research Secretariat as part of the MOSAiC 2019-2020 expedition.

CHAPTER 3 – Low salinity summer meltwater alters surface Arctic Ocean microbial community structure and metabolism

Abstract

Summer snow and sea-ice melt creates low-salinity meltwater which can locally accumulate between and under ice floes, impacting microbial composition and carbon cycling in the surface Arctic Ocean. During the 2019-2020 MOSAiC Drift Expedition, spatially persistent meltwater layers ($\leq$ 15 g kg^{-1} salinity, 0.1–1 m thick) in leads were observed from July–September. Here, we use both MOSAiC observations and culture-based experiments to investigate how these steep salinity gradients shape summer microbial microhabitats, impacting community structure and ecosystem function. Specifically, we compare *in situ* community structure (16S/18S rRNA gene amplicon sequences) and metabolic oxygen turnover to *in vitro* growth curves and metabolic oxygen turnover from three Arctic algae cultures across observed and artificial salinity gradients. *In situ* observations suggest that before freezing, stratified meltwater layers remain ecologically distinct with a unique microbial assemblage and metabolic potential. The surface meltwater had a community composition indicative of sea ice origin and showed enhanced net oxygen consumption compared to the underlying seawater. Our experiments showed that while algal growth and bacterial biomass were both reduced when exposed to low salinity meltwater, mixed community cultures were more productive following acclimation and heterotrophic activity persisted across all salinity treatments. In a warming Arctic, constraining the ecological impact of summer meltwater stratification will be critical for predicting climate-relevant biogeochemical cycling at the air-sea interface of open-water leads.

Introduction

Summer snow and sea ice melt serves as a local source of fresh water to the Arctic ice-ocean system. This fresh meltwater is either stored in melt ponds or exported to the surface ocean,

stabilizing the water column (Perovich et al. 2021). Under calm meteoric and ocean mixing conditions, this meltwater can become trapped and accumulate into well-defined (i.e., < 2 m thick), low salinity (i.e., 0–15 g kg^{-1}) layers topographically constrained either directly beneath the ice (Mundy et al. 2011; Smith et al. 2022) or in open leads between floes (Richter-Menge et al. 2001; Nomura et al. 2023). The resulting strong salinity gradient at the interface of the meltwater and underlying seawater has the potential to impact many aspects of the surrounding environment. For example, slowing sea ice bottom melt and promoting the formation of platelet ice (Hoppmann et al. 2020) and false bottoms (Salganik et al. 2023), altering the salt and heat balance of the ocean mixed layer (Kadko 2000), and limiting the exchange of gasses, nutrients, and organic material (OM) within the surface ocean and at the air-sea interface (Smith et al. 2023).

The summer meltwater layer thus creates distinctive, and often isolated, environmental niches for microbial (i.e., bacteria, archaea, ice algae, phytoplankton, and other protists) colonization (Smith et al. 2023). The diversity of meltwater assemblages is driven both by the primary sources of the microbes reaching these ephemeral habitats and through physiological acclimation to the corresponding rapid environmental change and its associated stressors such as low salinity and high light (Ralph et al. 2007; Mundy et al. 2011). Ice algae is particularly sensitive to salinity stress caused by freshwater influx, with potential effects on carbon cycling through the biological carbon pump. Pooling melt layers can cause sympagic algae to detach from the ice and sink rapidly (Boetius et al. 2013; Amiraux et al. 2020) or trap smaller algal aggregates near the surface, becoming local OM sources and hot spots for biogeochemical activity (Assmy et al. 2013; Fernández-Méndez et al. 2014).

Even small changes in community composition and individual cellular stress responses can drastically impact ecosystem functionality and contributions to local biogeochemical cycling

(Lannuzel et al. 2020). Smaller autotrophic flagellate species tend to outcompete diatoms as the dominant primary producers in isolated low salinity and/or meltwater environments, (Li et al. 2009; Mundy et al. 2011), which have been observed to contain higher proportions of mixotrophic and heterotrophic eukaryotes overall (Xu et al. 2020), thus altering meltwater metabolic rates of photosynthesis and respiration, and therefore net productivity. Bacterial community composition within isolated meltwater environments (i.e. melt ponds; Brinkmeyer et al. 2004) are similarly distinctive and susceptible to community changes due to salinity stress (Ewert and Deming 2014; Chamberlain et al. 2022), potentially creating an entirely unique microbial loop system within this near-surface feature (Smith et al. 2023). Quantifying the impacts of meltwater stratification on surface ocean microbial community structure and metabolic activity (i.e., photosynthesis and community respiration) will be important to accurately model changes in carbon flux between the ocean and the atmosphere in an increasingly freshening and more stratified central Arctic (Carmack et al. 2016).

While meltwater layers have been documented in the Central Arctic for over a hundred years (Nansen 1902), observations are sporadic and do not effectively capture the potential wide-spread persistence or physical evolution of these ephemeral features. The 2019–2020 Multidisciplinary drifting Observatory for the Study of Arctic Climate (MOSAiC) Expedition provided an invaluable opportunity for extended ice observation during the summer melt period, capturing a high density of discrete interdisciplinary observations of meltwater layers and their impact on a variety of processes across the Arctic ocean-ice-atmosphere continuum (Smith et al. 2023). In this study, we focus on quantifying the impact of meltwater stratification in leads on microbial community structure and function – specifically metabolic oxygen cycling and net productivity within the meltwater layer. First, we characterize the ecological micro-habits

observed during periods of meltwater stratification during the MOSAiC Expedition using measurements of microbial community structure (i.e., amplicon sequences, cell counts, chlorophyll *a*) and oxygen concentration. Next, we interrogate the impact of forced acclimation to low salinity conditions on microbial growth and metabolic activity rates *in vitro* using three distinct and non-axenic algal cultures propagated from samples collected during MOSAiC.

Methods

Field sample collection – MOSAiC Expedition

The year-long MOSAiC Expedition took place aboard the RV *Polarstern* (Knust 2017) from September 2019 to October 2020 and carried out a comprehensive interdisciplinary study of the central Arctic ice-ocean-atmosphere system (Shupe et al. 2020). Overviews of the physical ice, oceanographic, and atmospheric parameters taken over the course of the MOSAiC Drift can be found in Nicolaus et al. (2022), Rabe et al. (2022), and Shupe et al. (2022), respectively. An overview of observed meltwater layers during the MOSAiC Drift and the targeted interdisciplinary sampling strategies of these layers can be found in Smith et al. (2023). Briefly, observations of meltwater layers occurred during the drift tracks of the 4th (June–July) and 5th (August–September) legs of the MOSAiC Expedition (Figure 3-1). The drift tracks for these legs were spatially separated due to the original MOSAiC ice observatory reaching the ice edge in late July. Leads (n=10) sampled generally surrounded the MOSAiC ice observatories but were dynamic and did not always remain open or accessible during the periods of observation. Lead descriptions compiled from field notes and the specific dates they were sampled are described in Table 3-1.

Temperature and salinity profiles for each lead sampled were recorded by hand in the field using a YSI Professional Series with YSI 4-m and 10-m Conductivity/Temperature probe (Smith et al. 2021). Based on salinity profiles, water samples were collected using a peristaltic pump

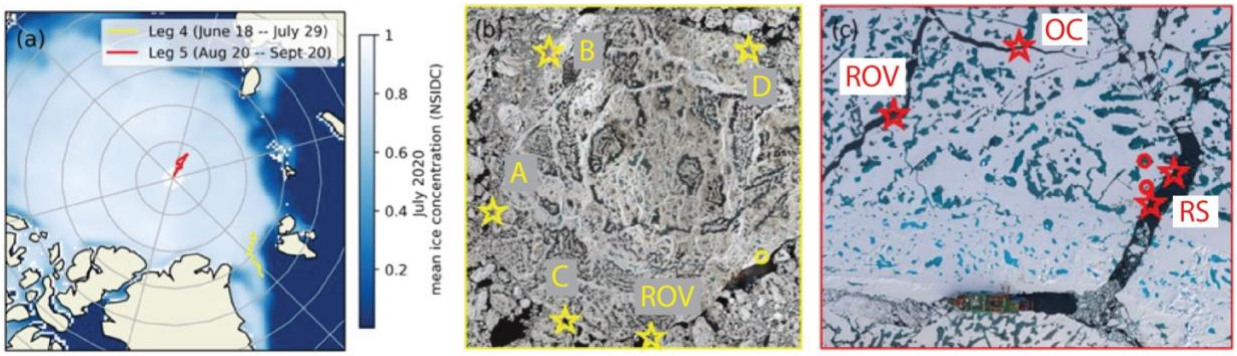


Figure 3-1. Relevant drift paths and lead sampling sites during the MOSAiC Expedition. Figure adapted from Smith et al. (2023), the MOSAiC Leg 4 (b) and Leg 5 (floes) are pictured from above with their respective drift paths shown in (a). Stars represent lead sampling sites, and are labeled by their names as described in Table 3-1. LL (~ 300 m above OC site) and NL (marginal ice zone station) are not pictured.

attached to tubing taped onto a bamboo pole from 3–4 discrete depths corresponding to the measured depth of any surface meltwater layer ('Meltwater'; < 15 g kg^{-1}), interface later ('Interface'; 7–23 g kg^{-1}), or underlying seawater layer(s) ('Seawater'; 27–34 g kg^{-1}). Following the disintegration of the stratified meltwater lens, only surface (0.05 – 0.10 m depth) water was collected ('Surface Seawater'; 21–33 g kg^{-1}). All water samples were kept shaded in coolers during transport back to the *Polarstern* where they were then (within 0–4 hours; held near 0 ºC) processed for ecological measurements.

Water samples for DNA sequencing and cell abundances were collected in acid cleaned and thrice sample-rinsed 1L plastic Nalgene bottles. 1.8 mL of sample was transferred into 4 replicate plastic cryovials and fixed to a 0.5% concentration of 25% glutaraldehyde at 4ºC for a minimum of 2 hours. Cryovials were then flash frozen with liquid nitrogen and stored at –80 ºC or on dry ice until analysis (see section *Laboratory Analyses*). The remaining volume (0.1–1 L; until filter clogged or no remaining volume was left) of sample was then filtered for DNA sequencing using a 0.2 µm supor membrane disc filter (25 mm; Pall Corporation) on an ethanol cleaned vacuum filtration rig. Filters were immediately frozen and stored at –80 ºC until extraction

Table 3-1. Summary of lead sampling during the MOSAiC.
*The MOSAiC Expedition (October 2019 – October 2020) was split into five legs with two discontinuous drift tracks (Legs 1–4 and Leg 5). Measurements from Leg 4 were collected in the Yermak Plateau and Fram Strait region. Measurements from Leg 5 were collected in the Amundsen Basin (Figure 3-1) or from a transit ice station near the marginal ice zone (NL).

Lead Name	Leg*	Location and Description	Dates Sampled
ROV-4	4	Open water lead surrounded by first year ice and located near the remote operated vehicle (ROV) operations running towards the bow of Polarstern. Visible green biomass layer at the interface of the meltwater layer and underlying seawater at 1 m.	2020-07-11
A	4	~15 m wide open water lead containing a mix of second year and/or ridged floating ice blocks. Adjacent to defined transects for albedo measurements (Nicolaus et al. 2022). Visible green biomass layer visible at the interface of the meltwater layer and underlying seawater at 0.40 (July 17) and 0.8 (July 22) m.	2020-07-17 2020-07-22
B	4	~60 m wide open water lead surrounded by level FYI. No visible biomass layer but lots of platelet ice formation around blocks.	2020-07-17
C	4	Open water lead located between the ROV and 'Dronesville' sites (Nicolaus et al. 2022). No visible biomass layer.	2020-07-22
D	4	Open water lead located near previous sampling site Alli's ridge (Nicolaus et al. 2022). Visible green biomass layer at the interface of the meltwater layer and underlying seawater, 0.55 m. Sampling on July 28 aborted due to fog and only surface sample was collected.	2020-07-22 2020-07-28
RS	5	Newly opened lead bisecting the Leg 5 MOSAiC ice observatory near the remote sensing (RS) site (Nicolaus et al. 2022). Physical dynamics of this lead are fully documented in Nomura et al. (2023). Originally opening to ~65 m wide this lead opened and closed several times throughout observation, beginning to freeze over as early as August 29 and became fully frozen over by September 15. Meltwater layer dissipated around September 7.	2020-08-29 2020-09-04 2020-09-06 2020-09-15
OC	5	Newly opened 2.5–3 m wide lead bisecting the Leg 5 MOSAiC ice observatory on the far side of the Ocean City (OC) sampling site (Rabe et al. 2022). This lead then closed and reopened on Setember 7 before freezing over and losing meltwater stratification. By September 12, the lead was 20 m wide at its widest point and frozen with hoar ice features and large integrated floes. Only surface water collected on September 12 and September 18.	2020-08-29 2020-09-12 2020-09-18
ROV-5	5	Refreezing lead near the Leg 5 ROV site (Nicolaus et al. 2022) located near an ASF50 meteorological sled (Shupe et al. 2022).	2020-09-02 2020-09-4
LL	5	Newly opened large refreezing lead originally located ~700 m distance from *Polarstern*, beyond the Met City sampling site (Shupe et al. 2022). Designated 'Luna Lead' (LL). By September 13 there was no meltwater stratification, and the lead was covered with a 9 cm thick, highly deformed grey ice layer and was located ~200 m closer to the ship. Only surface water was collected from this site.	2020-09-06 2020-09-13
NL	5	Re-freezing lead, designated 'Ninja Lead' (NL) sampled during an opportunistic ice station near the Marginal Ice Zone at 82° 07.855N, 070° 42.171 E while transiting to shore after concluding MOSAiC.	2020-09-26

(see section *Laboratory Analyses*). Water samples for chlorophyll *a* (and other parameters; Fong et al. unpubl.) were collected in Milli-Q water cleaned and thrice sample-rinsed 4L shaded plastic bottles. Volumes between 100 and 2300 mL of sample was filtered onto GF/F filters (25 mm; Whatman plc) and frozen at −80 ºC until analysis (see section *Laboratory Analyses*). Water samples for metabolic oxygen turnover were collected with laminar flow and over-fill in gas-tight 300 mL glass Biological Oxygen Demand (BOD) bottles.

Ice-algae collection and culture maintenance

An under-ice algal slush sample was collected on September 26th 2020, from a re-frozen lead in the marginal ice zone during an ice station while returning to land following the conclusion of the Leg 5 drift. Approximately 1 kg of slush from the ice bottom was collected using an ice saw and fine-mesh sieve. The salinity of under-ice water was 32.3 PSU, indicating no meltwater layer at the time of collection. The slush was returned to the ship in a cool box and the salinity of the 'under ice' slush was 29.4 PSU, measured approximately 1 hour after collection. Three subsamples were filtered for community DNA (see previous section). The remainder was decanted into 5x sterilized 50ml tissue culture flasks and stored at approximately 4 ºC in a light cabinet at 15 µmol photons m^{-2} s^{-1} until *Polarstern* returned to shore and they were shipped in a cool box with ice packs to the University of Groningen, NL. Samples were then grown under a light source emitting 150 µmol photons m^{-2} s^{-1} in a natural seawater liquid media enriched with F/2 nutrients; N, P, Si, trace metals, and vitamins (Andersen et al. 2005) for about 4 months.

In February 2021, five 15 ml falcon tubes of mixed culture were shipped on ice to the University of Warwick, UK for further cultivation. Ten ml of the samples were immediately transferred to 50 ml tissue culture flasks in filter sterilized 32 PSU artificial seawater media (Sea salts; Sigma Aldrich) with pre-prepared and 0.2 µm filter sterilized F/2 nutrients, evolving into the

"Brown Mix" culture strains. After another month, several of the original culture communities shifted to become the "Green Mix" culture strains. These were then transferred to a fresh natural seawater liquid media enriched with F/2 nutrients, except with no addition of silica (Si), to continue selection for non-siliceous (*i.e.*, diatom) taxa. Additionally, one uni-algal diatom strain (family Attheya) was isolated from the Brown Mix community culture (see Text 3-S1 for methods). All culture strains were maintained in a 4 °C cabinet under a light source emitting 150 µmol photons m^{-2} s^{-1} and reinoculated into fresh sterile media every 4 weeks. Artificial seawater was used for the Brown Mix and 'Attheya Isolate' culture media due to limited supplies of natural seawater during this time. All media and nutrients were sterilized by filtration through sterile 0.2 µm SUPOR membrane disc filters (47 mm; Pall Corporation) into autoclaved bottles prior to use to avoid altering the carbonate chemistry by autoclaving the liquid media.

In May 2022, representatives of each of these culture strains were shipped on ice to Scripps Institution of Oceanography to conduct the below described experimental work. Once at Scripps, all strains were reinoculated into fresh natural seawater media enriched in pre-prepared 0.2 µm SUPOR filter (25 mm; Pall Corporation) sterilized F/2 nutrients (with or without Si as described above). Liquid media was filter sterilized following the University of Warwick procedure using 0.1 – 0.2 µm Rapid-flow filters with Nalgene attachment (ThermoFisher Scientific) and pre-autoclaved glass bottles. Stocks were maintained in a 1 ºC incubator cabinet with light conditions between 30–50 µmol photons m^{-2} s^{-1} and reinoculated (1:5 dilution) into fresh media approximately once a month.

Experimental design and sample collection

In October 2022, one representative of each strain type — Brown Mix, Green Mix, and Attheya Isolate — was grown into 1250 mL of dense culture (approx. 3.2 x 10^5 cells ml^{-1}) during

an 8-day growth phase to seed three experimental salinity treatments. Triplicate chlorophyll *a* concentration (1 mL) and cell abundance measurements (1 mL) were collected daily. Growth curves for this phase are available in Figure 3-S1. On day 8, 125 mL was aliquoted for Timepoint 0 (T0) measurements for DNA sequencing (40 mL), chlorophyll *a* concentration (1 mL), cell abundances (1 mL), pH (4 mL), and nutrients (20 mL) were collected. The remaining 1125 mL was then dispersed between triplicate 1L glass bottles for 3 experimental salinity treatments of 5 PSU, 15 PSU, and a base media control of approximately 32 PSU. The 5 PSU treatment was created by adding 725 mL of sterile F/2 enriched Milli-Q water (0 PSU) to the 125 mL of culture stock. The 15 PSU treatment was created by adding 475 mL of sterile F/2 enriched Milli-Q water and 250 mL of sterile F/2 enriched natural seawater media to the 125 mL of culture stock. The 32 PSU control treatment was a standard 1:5 stock dilution adding 725 mL of sterile F/2 enriched natural seawater media. The filter sterilized natural seawater used to create the liquid media base was the same across both the initial growth phase (Figure 3-S1) and all experimental salinity treatments. Experiments using the Green Mix community culture did not have Si added during F/2 media enrichment. Initial nutrient concentrations and light availability was otherwise kept constant across both the initial growth phase and all experimental treatments. Bottles were distributed across four shelves with light values ranging from 17–33 μmol photons m^{-2} s^{-1} due to shading on the secondary shelves. Bottles were rotated daily across shelves with random distribution of triplicate treatment bottles. All cultures were maintained at 0–1 ºC throughout the duration of the experiments.

Once diluted into the triplicate experimental salinity treatments, cultures were grown for 8 more days allowing community acclimation to the new salinity conditions. Aliquots of 2 mL were collected from each bottle daily for chlorophyll *a* concentration and cell abundances to track

growth. On day 8, Timepoint "Final" (TF) measurements were collected for DNA sequencing (25–170 mL), chlorophyll *a* (1 mL), cell abundances (1 mL), pH (4 mL), and nutrients (2 mL). Additionally, 1 mL of sample was used to confirm the salinity value for each bottle using a handheld refractometer and 5 mL of sample was saved in a sterile 15 mL Falcon tube and placed back into the incubator for future microscope imaging.

All remaining liquid culture was then used to fill replicate 300 mL glass BOD bottles. Replicate bottles were split between a light and a dark incubation to track net oxygen production and consumption over a 48-hour period. BOD bottles were Milli-Q cleaned prior to use and filled slowly, with overfill, to reduce the incorporation of air bubbles. Once filled, BOD bottles were placed back into the incubator without caps to re-stabilize at incubation temperature and reduce potential temperature induced bubble formation. In each light treatment there was additionally a 0.2 µm filtered 'media blank' which was used to account for any temperature signals in oxygen saturation over the incubation period. After approximately 2 hours, bottles were capped and the incubations started (see section *Laboratory analyses* for full details). At the end of the 48-hr incubation, final measurements for chlorophyll *a* concentration (1 mL), cell abundances (1 mL), pH (4 mL), and nutrients (20 mL) were collected.

Experiments were run staggered for 10 total days for each culture strain. For all experimental samples, water filtration for DNA sequencing was conducted using the same protocol which was used for field samples. chlorophyll *a* concentrations and pH were analyzed immediately on benchtop instruments (see section *Laboratory analyses*). T0 and TF cell abundances were diluted 1:10 with Phosphate Buffered Saline (PBS; Millipore Sigma) and run immediately on a benchtop flow cytometer without fixation (see section *Laboratory* analysis) but daily growth samples were fixed to a final concentration of 0.25 % glutaraldehyde and frozen in plastic cryovials

at –80 ºC until analysis. All were pre-filtered through a 60 µm nylon filter (25 mm; Millipore Sigma) during collection. Nutrients were pre-filtered through a 0.45 µm cellulose acetate membrane filter (25 mm; Advantec) and frozen in sterile plastic falcon tubes at –20 ºC until analysis.

Laboratory analyses

DNA sequencing: DNA from frozen filters (field and experimental collections) was extracted at Scripps Institution of Oceanography using a KingFisher™ Flex Purification system and MagMax Microbiome Ultra Nucleic Acid Extraction kit (ThermoFisher Scientific). Extracted nucleic acid was then sent to Argonne National Laboratory for library preparation and 16S/18S rRNA gene amplicon sequencing. Sequencing was performed on the Illumina Miseq plaform on a 2 x 151 bp library architecture using universal primers 515F and 806R (16S, V4; Walters et al., 2016) and 1380F and 1510R (18S, V9; Amaral-Zettler et al., 2009). The R package dada2 (Callahan et al. 2016) was used to denoise and merge Illumina reads, and the PAthway PRediction by phylogenetIC plAcement pipeline (paprica; v0.7.0) was used for taxonomic identification using reference trees from the PR2 4.13.0 (18S genes; Guillou et al. 2013) and Genbank RefSeq (16S and 23S genes; Haft et al. 2018) databases. Each amplicon sequence variant (ASV) was placed to either an internal (unique sequence; described by consensus taxonomy for all downstream nodes) or terminal branch (edge sequence; described by reference taxonomic ID) on the tree (Nawrocki and Eddy 2013; Barbera et al. 2019; Czech et al. 2020). Field samples are presented at the unique sequence level, while experimental samples are presented at the edge sequence level to provide greater taxonomic resolution and account for lower diversity in culture. For bacteria and archaea, paprica also uses the point of placement in the reference tree to estimate mean genomic characteristics for each sample, such as genome size and 16S rRNA gene copy number (Erazo et

al. 2021). Final read counts for 16S ASVs were normalized to estimated 16S rRNA gene copy number before relative abundance calculations. Samples with < 3000 total ASV counts, singletons, and suspected mitochondrial (16S) and intercellular parasite (18S) sequences were filtered out prior to any statistical analyses following the QC procedures used in Chapter 2 (Text 2-S1).

Cell abundances: Samples for cell abundance collected from the field were analyzed at the University of Bergen for the abundance of viruses, prokaryotes (bacteria and archaea), picophytoplankton, and nanophytoplankton. Autotrophic cell abundance was counted on a Attune® Acoustic Focusing Flow cytometer (Applied Biosystems by ThermoFisher Scientific) immediately following sample thaw. Virus particles and bacteria were measured on a FACS Calibur (Becton Dickinson) flow cytometer after samples were thawed, diluted 1:10 with 0.2 µm filtered TE buffer (Tris 10 mM and EDTA 1 mM, pH 8) and stained with the manufacturers recommended working concentration of SYBR Green (I; Molecular Probes) nucleic acid stain before incubating 10 min at 80 ºC in a water bath (Marie et al. 1997). Phytoplankton groups (autofluorescent cells) were classified using red fluorescence vs. orange fluorescence (Paulsen et al. 2016). Prokaryotic and viral groups were classified using green fluorescence vs. side scatter (Larsen et al. 2008; Paulsen et al. 2015).

Samples for cell abundance collected during the laboratory experiments were analyzed at Scripps Institution of Oceanography for the total abundance of prokaryotes, picophytoplankton, nanophytoplankton, and actively respiring cells on a Guava easyCyte 11HT (Luminex). The Guava 11HT uses both a blue (488 nm) and violet (405 nm) laser to interrogate cells. Autofluorescent cells (< 60 µm) were measured using forward scatter, side scatter, red emission, and yellow emission. To measure prokaryotic cells, samples were stained with SYBR Green (I; Molecular Probes) and measured for forward scatter, side scatter, and green emission. For the 1:10 PBS

diluted and unfixed T0 and TF measurements, samples were additionally run using the BacLight™ RedoxSensor™ Green Vitality Kit (ThermoFisher Scientific) to track bacterial reductase activity and calculate the relative percentage of total cells were actively respiring at the time of analysis. All reagents were added to samples and then incubated for 10 min in a dark 4 ºC fridge. For the experiment using the Attheya Isolate, only carbonyl cyanide m-chlorophenyl hydrazone (CCCP) was added to kill cells and create a 'killed control'. For the two mixed community experiments, two killing agents, sodium azide and CCCP were added to a single killed control sample. For all experiments, the killed control fluorescence signal was subtracted from the RedoxSensor Green stained sample's signal prior to calculating total counts of respiring cells. This value was then divided by the total cell count from a SYBR Green stained sample run concurrently to calculate percentage of active cellular respiration. Quality control for absolute cell counts were confirmed by spiking a standard volume of 1:10 diluted 123-count fluorescent beads (ThermoFisher) to each sample and killed control.

Algal groups and bacterial groups were classified using a self-organizing map (SOM) using forward scatter, side scatter, red emission, and yellow emission (autofluorescent) or forward scatter, side scatter and green emission (SYBR-stained) following the methods outlined in Wilson et al. (2021). Briefly, a training dataset constructed from one growth phase sample from each culture strain was trained using a toroidal map of grid size 41x41 using the 'kohonen' package in R (Wehrens and Kruisselbrink 2018). The number of subgroups for both autofluoresent and SYBR stained cells was identified via k-means clustering with a k=5 based on visual evaluation of a within-cluster sum of squares scree plot and bi-parametric plot of forward scatter vs. red emission or green emission. The two trained cluster models (autofluorescent and SYBR stained) were then used to classify events for each subgroup in both sample types.

Chlorophyll a concentrations: Frozen filters for chlorophyll *a* concentration from the field were extracted and analyzed at Alfred Wegner Institute using a TD-700 fluorometer (Turner Designs™) as described in (Hoppe et al. 2023b). Experimental samples for chlorophyll *a* concentration were transferred to a polycarbonate cuvette and run immediately on an AquaFlash Handheld Active Fluorometer (Turner Designs™).

Metabolic oxygen turnover: Metabolic oxygen turnover in field samples was estimated from the ratio of oxygen (O_2) and Argon (Ar) concentrations (Eq. 1) and is reported as biological oxygen concentration ($[O_2]_{bio}$; Eq. 2) following the approach in Eveleth et al. (2014) and calculated as described in Chapter 2, where positive values represent a metabolic tracer for net oxygen production and negative values represent a metabolic tracer for net oxygen consumption. Physical changes in net oxygen concentration was derived from deviations in theoretical and measured Ar saturation (Eq. 3), which is traditionally assumed to be 1 in seawater (Cassar et al. 2011).

$$\underline{\text{Equation 1.}}\ \Delta O_2/Ar = \left[\left(\frac{[O_2]_{meas}/[Ar]_{meas}}{[O_2]_{sat}/[Ar]_{sat}}\right)\right] - 1$$

$$\underline{\text{Equation 2.}}\ [O_2]_{bio} = \frac{[Ar]_{meas}}{[Ar]_{sat}}[O_2]_{sat}(\Delta O_2/Ar)$$

$$\underline{\text{Equation 3.}}\ [O_2]_{phys} = \left(\frac{[Ar]_{meas}}{[Ar]_{sat}} - 1\right)[O_2]_{sat}$$

O_2 and Ar gasses were measured onboard *Polarstern* using a Pfeiffer QMG 220 quadrupole membrane inlet mass spectrometer. The instrument was 2-point calibrated using N_2 gas (0%) or 21%/1% O_2/Ar synthetic air bubbled underway seawater directly prior to running samples. Details for this instrumentation and onboard measurements during the MOSAiC Expedition can be found in Rokitta et al. (*in prep*) and Chamberlain et al. (*submitted*). Theoretical concentrations of O_2 and Ar at air saturation ($[O_2]_{sat}$ and $[Ar]_{sat}$) were calculated using calibration seawater temperature and

salinity and the solubility equations from Garcia and Gordon (1992; 1993 – combined fit) and Hamme and Emerson (2004) respectively. Comparing $[O_2]_{sat}$ and $[Ar]_{sat}$ to the 21 %/1 % synthetic air standard calibrations, we translated spectrometer readings into total concentrations of $[O_2]_{meas}$ (or $[O_2]_{tot}$) and $[Ar]_{meas}$ in μmol kg^{-1}. Sample $[O_2]_{sat}$ and $[Ar]_{sat}$ were calculated individually using the same solubility equations and recorded *in situ* temperature and salinity values.

Metabolic oxygen turnover in the laboratory experiments was determined by tracking oxygen evolution over a 48 hr incubation period in the 0 ºC incubator. Light bottles were kept in direct light (17 μmol photons m^{-2} s^{-1}). Dark bottles were shaded (0 μmol photons m^{-2} s^{-1}) with a dark metallic bag. Gas tight glass BOD bottles were fitted with an optical Planar oxygen sensor spot (Presens Precision Sensing). Oxygen values were recorded using a Fibox 4 Fiber Optic Oxygen Meter equipped with a PSt3 oxygen sensor and a Pt100 temperature probe (Presens Precision Sensing). Measurements were recorded immediately following bottle capping (T0) and again at 6, 12, 24, 36, and 48 hours (TF) during the incubation period. Bottle measurements were completed in ambient lighting in under 10 minutes. For all bottles, net change in oxygen concentration was calculated by subtracting the initial value (T0) from the final timepoint (TF). For quality control, any signal observed in the 0.2 μm filtered seawater blank was then subtracted from the corresponding light or dark bottles to calculate the final net biological change in oxygen concentration (μmol kg^{-1}) and net productivity rates (μmol kg^{-1} hr^{-1}) over the 48 hour incubation.

Nutrients and pH: Nutrient samples were analyzed for nitrate, nitrite, phosphate, silicate, and ammonia at the University of California San Diego Oceanographic Data Facility using standard analytical methods compatible with the nutrient section of the GO-SHIP repeat hydrography manual (Becker et al. 2020). Samples were thawed in a 50 ºC heating bath and brought to room temperature prior to analysis using a Seal Analytical continuous-flow

AutoAnalyzer 3 (AA3). Due to high nutrient enrichment in the media, a 1:20 dilution was sometimes employed to determine macronutrient concentrations. Experimental pH values were analyzed immediately using a bench-top accumet® AEL50 pH sensor (Fisher Scientific). However, these values are used primarily as relative values between samples, and with some caution due to poor instrument calibration and performance at the time of experimentation.

Statistical analyses

Statistics for this work were calculated in R Studio using the R programming language (R Core Team 2023). A non-metric multidimensional scaling (NMDS) ordination and permutational analysis of variance (PERMANOVA; 999 permutations) were run using Bray-Curtis dissimilarity matrices (Bray and Curtis 1957) of the Hellinger transformed 16S and 18S ASV relative abundances to compare microbial community structure among observed stratified layers and experimental salinity treatments (vegan package; Oksanen et al., 2015). Hierarchical clustering was used in heatmaps generated using the pheatmap package (Kolde 2019) to highlight similarities between sample community composition. Non-parametric Kruskal-Wallis tests with Dunn pair-wise comparisons (vegan package; Oksanen et al., 2015) were used to compare other ecological variables among observed stratified layers and experimental salinity treatments. All p-values were adjusted to correct for multiple comparisons using the Holm-Bonferroni method with initial significance cut-off of $\alpha = 0.05$ (Holm 1979). The inverse Simpson index for alpha-diversity was used to measure diversity within sample types (vegan package; Oksanen et al., 2015).

Results

Meltwater stratification in leads – general observations

Meltwater stratification was observed during late summer (June–Aug) when sampling leads during both Leg 4 (78º – 80º N) and Leg 5 (87.8º – 89º N) of the MOSAiC Expedition (Figure 3-2). On Leg 4, the first visual indication of a lead meltwater layer was observed on July 1st and a

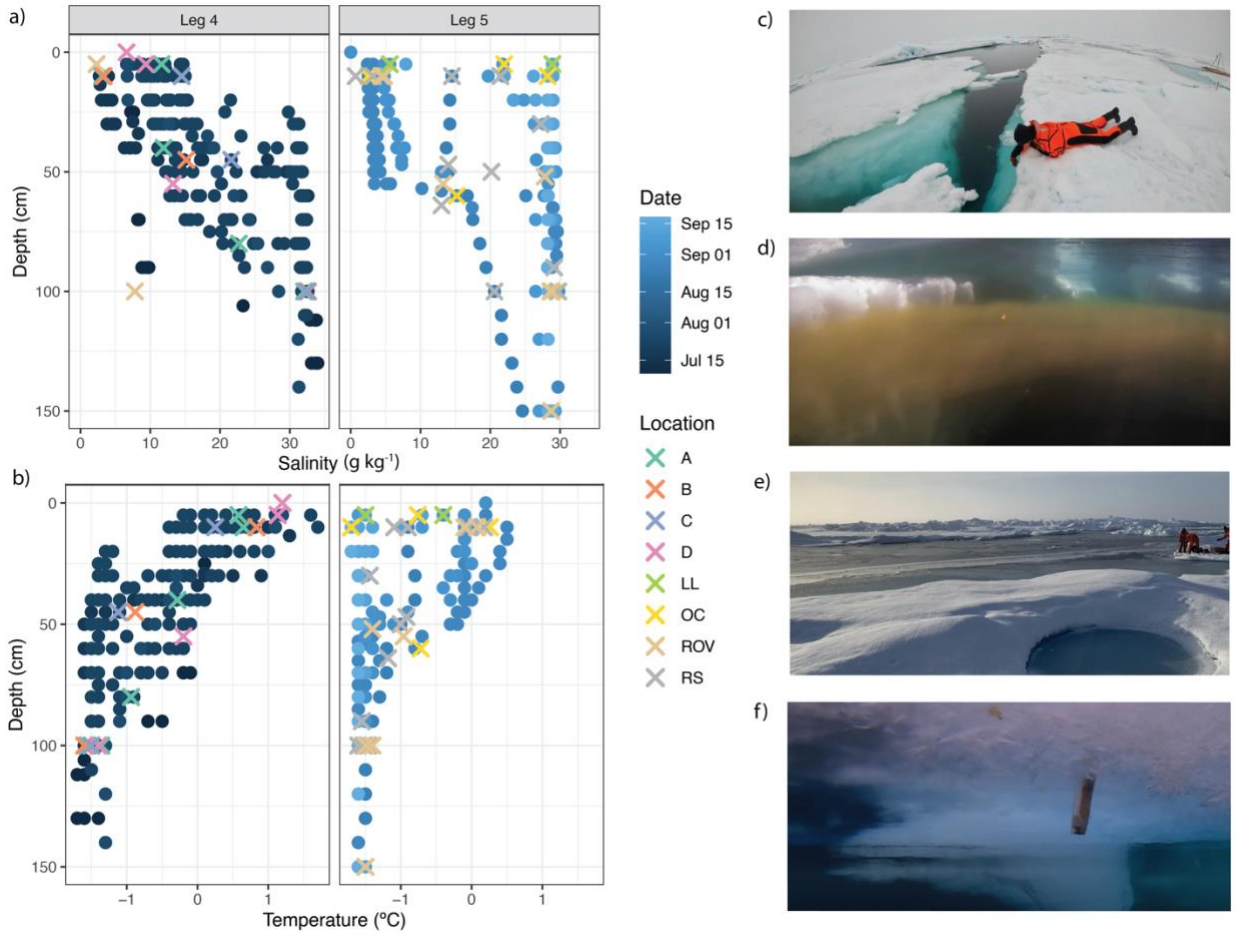


Figure 3-2. Physical characterization of leads from Legs 4 and 5 of the MOSAiC Expedition. All salinity (a) and temperature (b) profiles from sampled leads measured using a YSI probe (Smith et al. 2023). Point measurements are shaded by date sampled (blue), and colored Xs are used to mark where physical water samples were collected. Color here represents the lead site sampled (Table 3-1). Panels c-f show pictures from both above and below the lead surface from two selected sample sites; ROV-4 on July 11 (c) and July 6 (d; Muilwijk et al. 2022) and RS on August 25 (e-f).

1 m thick meltwater lens with salinities < 10 g kg^{-1} was recorded and sampled on July 11th at the ROV lead site (Figure 3-2c). This stratification also contained a dense biomass layer at the meltwater–seawater interface (Figure 3-2d). A visible meltwater layer then persisted in all following lead sampling events albeit with a general trend towards thinning and a decrease in both the salinity and temperature gradient over time (Figure 3-2a). Lead re-freezing was not observed on this leg.

On Leg 5, lead meltwater layers were observed upon arrival at the new ice observatory in early August and first sampled at the RS site on August 25 (Figure 3-2e). Thickness of this meltwater layer over time was found to correlate negatively with lead thickness (Nomura et al. 2023). Surface temperatures were overall lower (Figure 3-2b) with some re-freezing and platelet ice formation at the surface (Figure 3-2f). Visible biomass at these lead sites were not concentrated at the meltwater–seawater interface but instead floating throughout the meltwater layer in aggregates (Supplemental File 3-1). Dissolution of stratification in Leg 5 leads (Figure 3-2) occurred due meteorological changes (Rinke et al. 2021; Shupe et al. 2022) occurring between September 5 and 9 following a storm event and the rapid onset of freeze processes.

Observed microbial community structure and metabolic activity in the meltwater layer

The overall community structure of the meltwater layer was significantly distinct compared to that of both the interface (when present) and underlying seawater. In a 2D NMDS ordination including the community structure of all lead (Leg 4 and 5) samples as well as first year sea ice samples collected June–August (Chamberlain and Bowman 2022), meltwater samples showed no overlap with the underlying seawater samples and were instead clustered more closely to the summer ice communities, particularly for the 18S community (Figure 3-3). Indeed, only 39–42 % of the ASVs present in samples from the meltwater layer were also present in the underlying

seawater (Figure 3-3). Samples collected at the meltwater-seawater interface clustered with either meltwater samples or seawater samples and had a lower number of unique ASVs compared to shared ASVs present in this layer for both legs (Figure 3-3). Surface Seawater samples collected from surface water following disintegration of the meltwater lens clustered with samples collected from the underlying seawater earlier in the observation period (Figure 3-3).

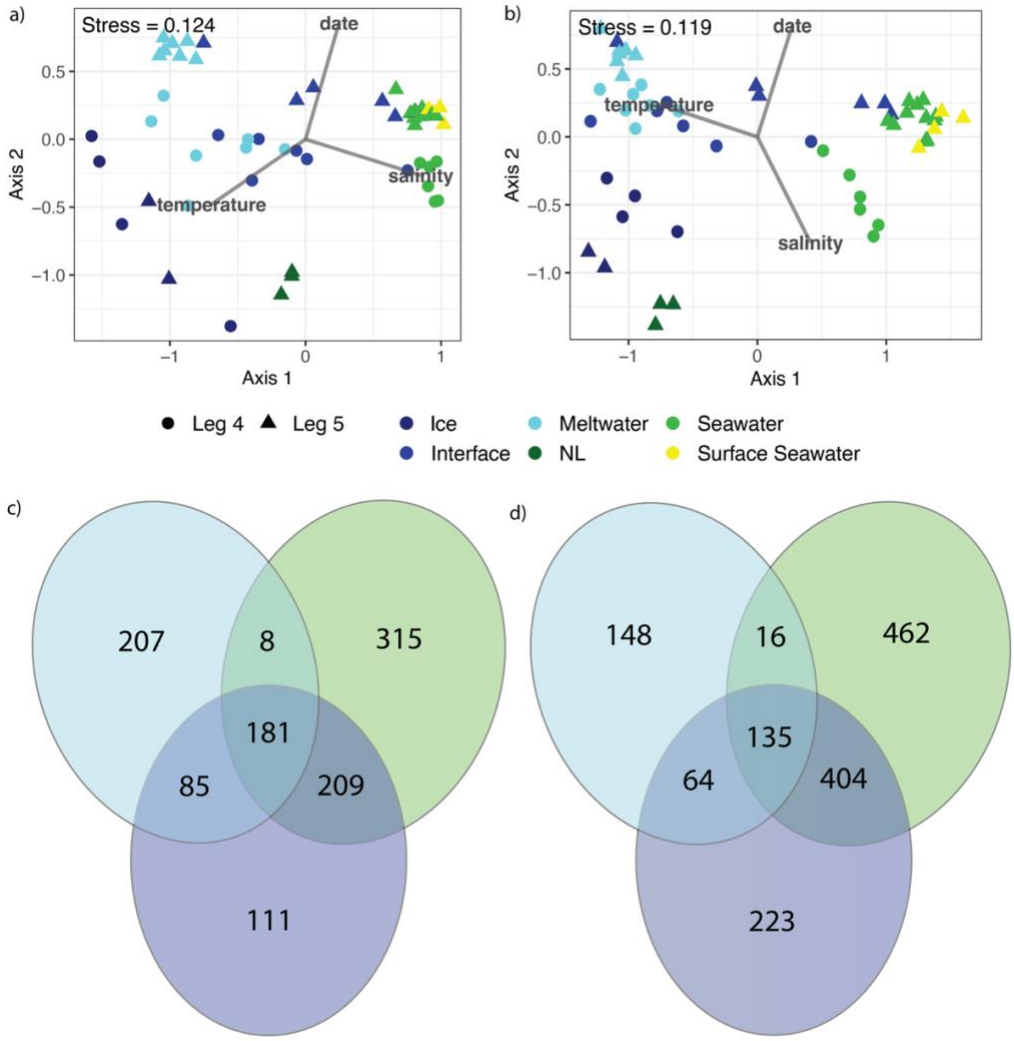


Figure 3-3. Similarities in microbial community structure across meltwater stratification. Individual samples are presented in the first two dimensions of NMDS ordinations to compare similarities in prokaryotic (16S ASVs; a) and eukaryotic (18S ASVs; b) community structure across sample types (color) and Leg collected (shape). Shared and unique 16S and 18S ASVs present in the Meltwater, Interface, and Seawater layers when meltwater stratification was present (Ice and Surface Seawater samples not included) are represented in the two Venn diagrams for Legs 4 (1116 ASVs; c) and 5 samples (1452 ASVs; d) respectively.

A PERMANOVA on the resulting Bray-Curtis dissimilarity matrices confirmed that dispersion among groups was statistically significant among sample types for both the prokaryotic ($p = 0.001$; $F(5 , 56) = 5.69$) and eukaryotic communities ($p = 0.001$; $F(5 , 56) = 5.81$). Additionally, salinity (16S: $r^2 = 0.79$, $p = 0.001$; 18S: $r^2 = 0.78$, $p = 0.001$), temperature (16S: $r^2 = 0.71$, $p = 0.001$; 18S: $r^2 = 0.64$, $p = 0.001$), and date (16S: $r^2 = 0.75$, $p = 0.001$; 18S: $r^2 = 0.79$, $p = 0.001$) were significantly correlated with changes in community structure (Figure 3-3).

The prokaryotic community (Figure 3-4) of the meltwater layer on both MOSAiC legs contained high proportions of ASVs identified to the class Flavobacteria, high nucleic acid containing cells, and large average estimated genome sizes relative to the underlying seawater. Highly abundant ASVs specific to the meltwater layer were assigned to reference genomes of *Aquiluna borgonia, Polaromonas, Nonlabens dokdonensis, Flavobacterium faecale, and Flavobacterium crassostrae* (Figure 3-4). Leg 4 samples (all stratified layers) contained a relatively higher proportion of Flavobacteria when compared to the samples collected on Leg 5, whose meltwater layer also contained high proportions of ASVs assigned to Actinomycetia, and Beta-proteobacteria reference genomes, as well as the greatest total abundance of viral cells (Figure 3-S1). The bottom sea ice slush samples from NL collected to propagate cultures clustered more closely with other sea ice samples (Figure 3-3a) and had a more similar community makeup to the Leg 4 meltwater layer, with high abundances of ASVs assigned to reference genomes of Candidatus *Pelagibacter*, *Rhodoferax ferrireducens*, and an unidentified Proteobacteria ASV with a 100 % BLAST® match to a cultured representative of the genus *Herminiimonas* (Figure 3-4).

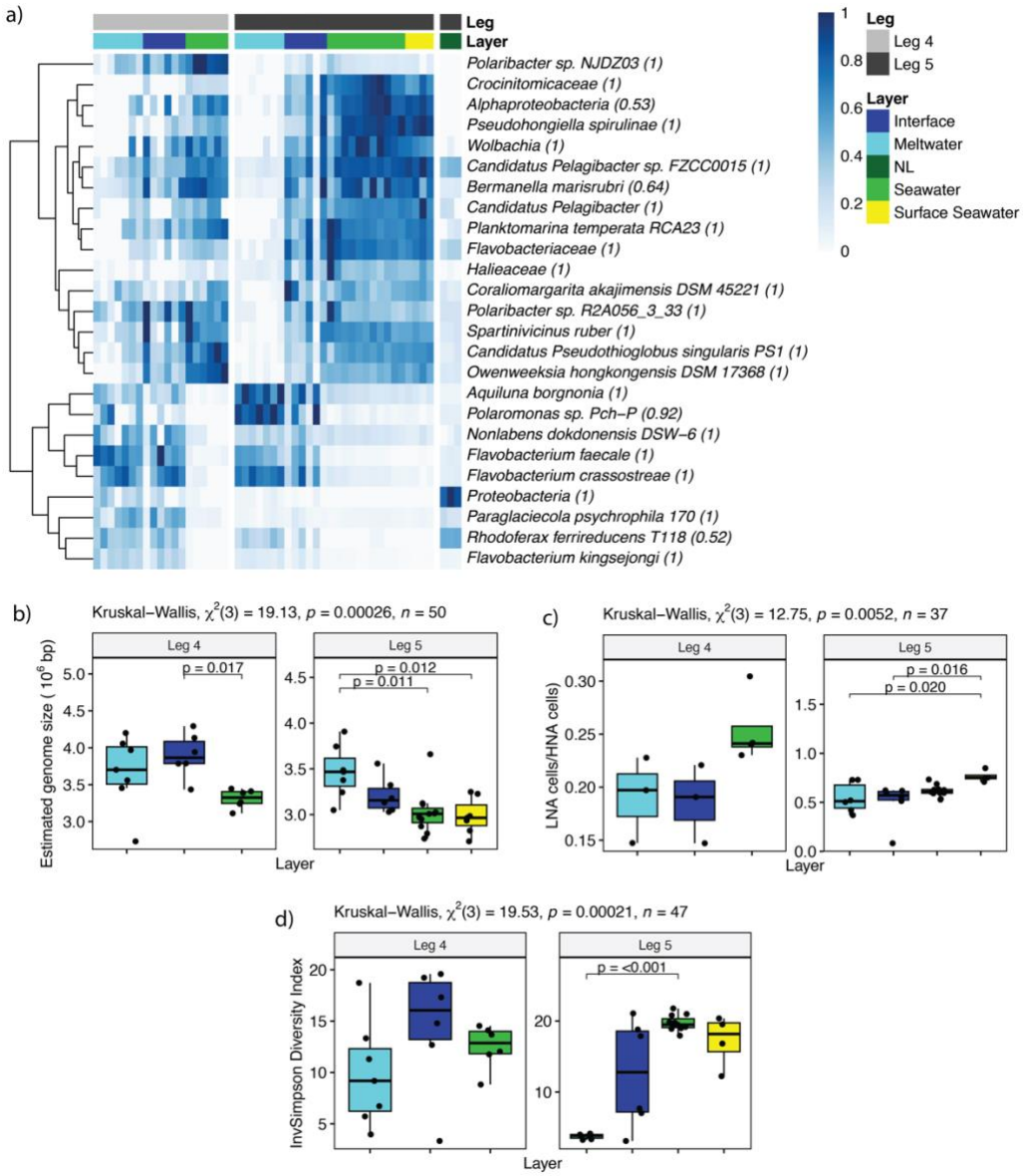


Figure 3-4. Dominant prokaryotic community composition in leads.
The heatmap shows the 25 most abundant 16S ASVs across all lead samples (a). Hellinger transformed read counts were scaled using a z-score between 0 and 1 and Bray-Curtis dissimilarity distances were used to cluster ASVs (rows). The paprica taxonomic ID (closest relative among RefSeq genomes) and placement proportion were used to identify each unique ASV (Bowman and Ducklow, 2015). Each sample (column) is colored by the MOSAiC leg and stratified water layer from which the sample was collected. Meltwater indicates samples collected above 10 cm depth with a salinity of < 15 g kg^{-1}. 'Seawater' indicates samples collected from the underlying seawater. 'Interface' indicates where the two previous layers met, and 'Surface Seawater' indicates surface samples collected above 10 cm depth following disintegration of the meltwater lens (salinity of 21 – 33 g kg^{-1}). 'NL' indicates the triplicate samples collected from the bottom ice slush used to propagate the experimental cultures. The box plots highlight the distribution across layers of each sample's (when measured) mean estimated genome size (b), ratio of Low Nucleic Acid (LNA) to High Nucleic Acid (HNA) bacterial cells (c), and alpha diversity using the InvSimpson Index (d).

The eukaryotic community (Figure 3-5) of the meltwater layer was dominated by low chlorophyll nanophytoplankton, with a high proportion of ASVs assigned to flagellates (specifically classes Filosa-Thecofilosea and Dinophyeae; Figure 3-S1). During Leg 4, the meltwater and, even more so the interface layer, also contained a high abundance of ASVs assigned to pelagophyte reference genomes, (Figure 3-5, 3-S1). The underlying seawater during Leg 4 contained a higher proportion of diatom ASVs (i.e. *Fragilariopsis sublineata*) relative to the seawater community of Leg 5, which had a higher proportion of picophytoplankton ASVs (i.e. *Micromonas polaris*) and dinoflagellates (Figure 3-5). The meltwater community on Leg 5 was similarly distinct with very few pelagophyte ASVs and a greater proportion of protozoa (i.e. Spirotrichea; Figure 3-S1). Like the prokaryotic community, the bottom sea ice slush samples from NL collected to propagate cultures clustered more closely with other sea ice samples (Figure 3-3b). Unlike the prokaryotic community, however, the eukaryotic community was more similar in makeup to the Leg 5 meltwater layer, with high abundances of ASVs assigned to reference genomes of a *Pyramimonas* sp., an unidentified ASV with a 94 % BLAST® similarity to a Chrysophyceae representative, *Cryothecomonas aestivalis*, and *Stoekeria algicida*. The most abundant diatom ASV was classified to *Fragilariopsis sublineata*, which was abundant in both Leg 4 and Leg 5 seawater communities (Figure 3-5).

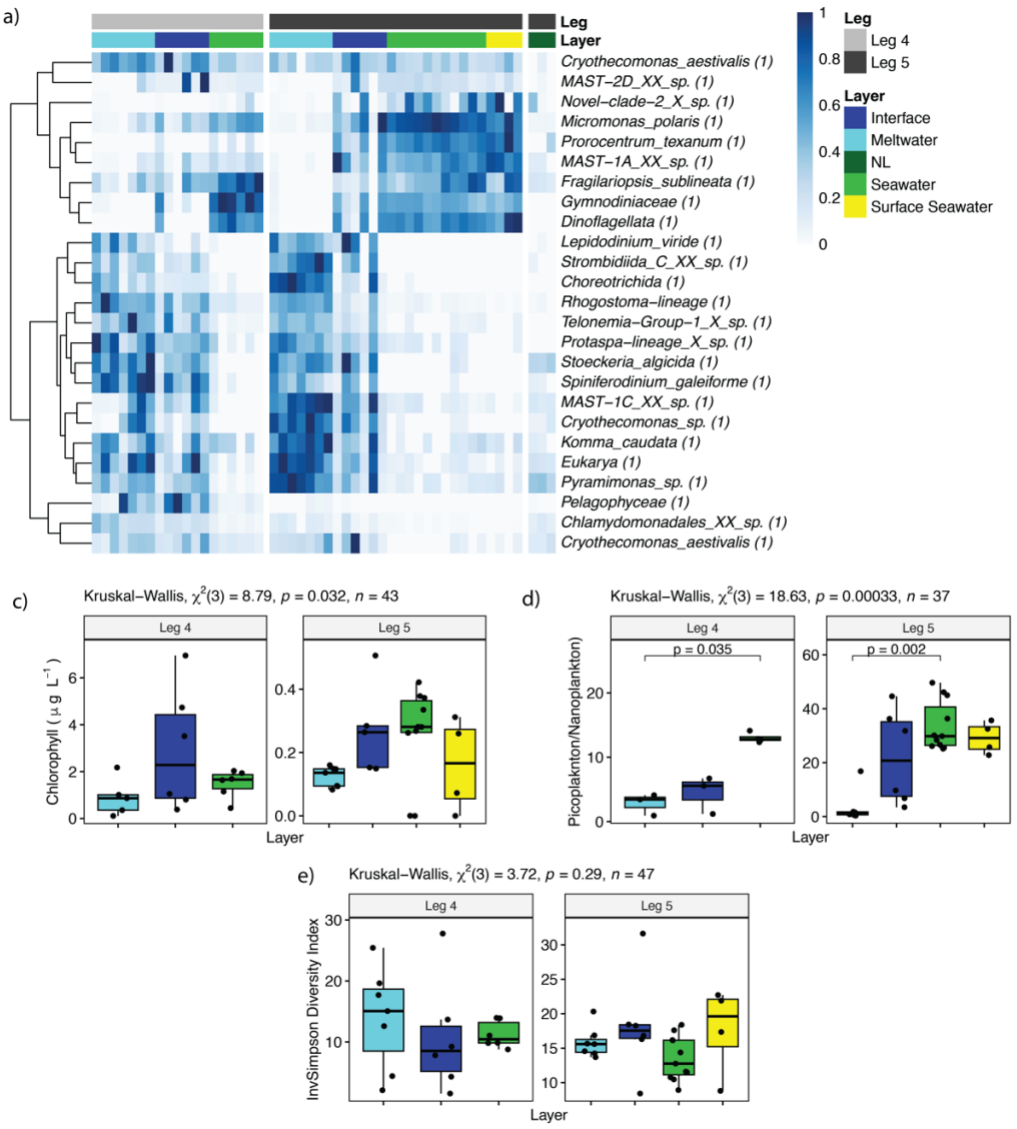


Figure 3-5. Dominant eukaryotic community composition in leads.
The heatmap shows the 25 most abundant 18S ASVs across all lead samples (a). Hellinger transformed read counts were scaled using a z-score between 0 and 1 and Bray-Curtis dissimilarity distances were used to cluster ASVs (rows). The paprica taxonomic ID (closest relative among PR2 genomes) and placement proportion were used to identify each unique ASV (Bowman and Ducklow, 2015). Each sample (column) is colored by the MOSAiC leg and stratified water layer from which the sample was collected. Meltwater' indicates samples collected above 10 cm depth with a salinity of < 15 g kg^{-1}. 'Seawater' indicates samples collected from the underlying seawater. 'Interface' indicates where the two previous layers met, and 'Surface Seawater' indicates surface samples collected above 10 cm depth following disintegration of the meltwater lens (salinity of 21 – 33 g kg^{-1}). 'NL' indicates the triplicate samples collected from the bottom ice slush used to propagate the experimental cultures. The box and whisker plots highlight the distribution across layers of each sample's (when measured) total chlorophyll *a* concentration (b), the ratio of picophytoplankton to nanophytoplankton cells (b), and alpha diversity using the InvSimpson Index (d).

The water samples collected for metabolic oxygen turnover on Leg 5 featured a significant reduction in biological oxygen concentration in samples collected from the meltwater layer relative to both the interface and underlying seawater (Figure 3-6). Net physical changes in total oxygen concentration were most variable in the meltwater layer, but the range was low and the average was not significantly different from either the interface or underlying seawater (p = 0.19; Figure 3-6).

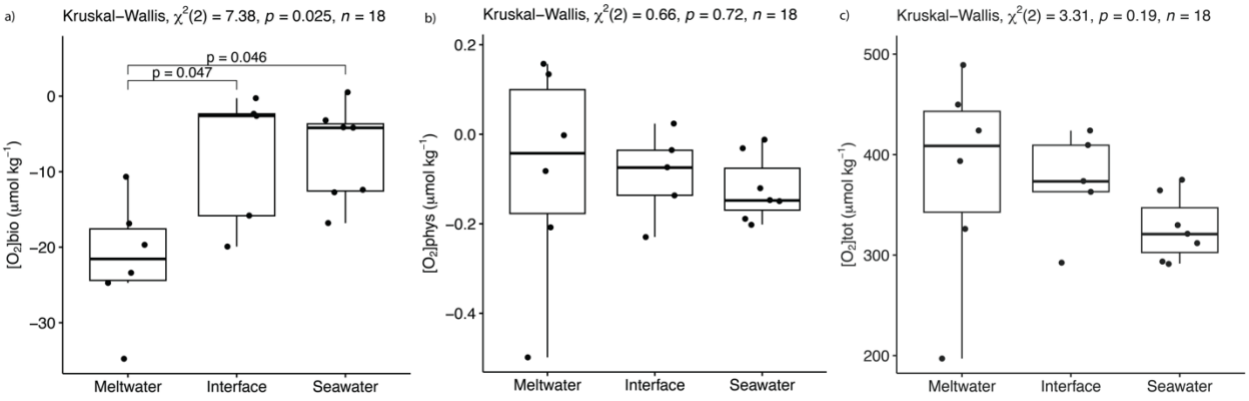


Figure 3-6. Observed net metabolic oxygen turnover in the meltwater layer.
Median changes in biological oxygen concentration, $[O_2]_{bio}$ (a), physical oxygen concentration $[O_2]_{phys}$ (b), and total oxygen concentration, $[O_2]_{tot}$ (c), in µmol kg^{-1} across each water layer collected from leads during Leg 5 of the MOSAiC Expedition. The midline of each box represents the median, with lines below above, representing the first and third quartiles. Whiskers represent the minimum and maximum of the spread with outliers marked by standalone points. The results of a Kruskal-Wallis test comparing layers is printed at the top and significant differences between layers are reported by the p-values from Dunn pair-wise comparisons (vegan package; Oksanen et al., 2015). Only significant comparisons are shown.

Algal culture community composition and acclimation to salinity treatments

The three algal communities each contained a distinct prokaryotic and eukaryotic community composition at the beginning of each experiment. Microscope imaging (Figure 3-7) identified the unialgal culture as an Attheya diatom strain. This was confirmed by sequencing results, with all 18S edge assignments present in more than one replicate assigned to a representative genome of *Attheya septentrionalis* (Figure 3-8). The 'Brown Mix' community

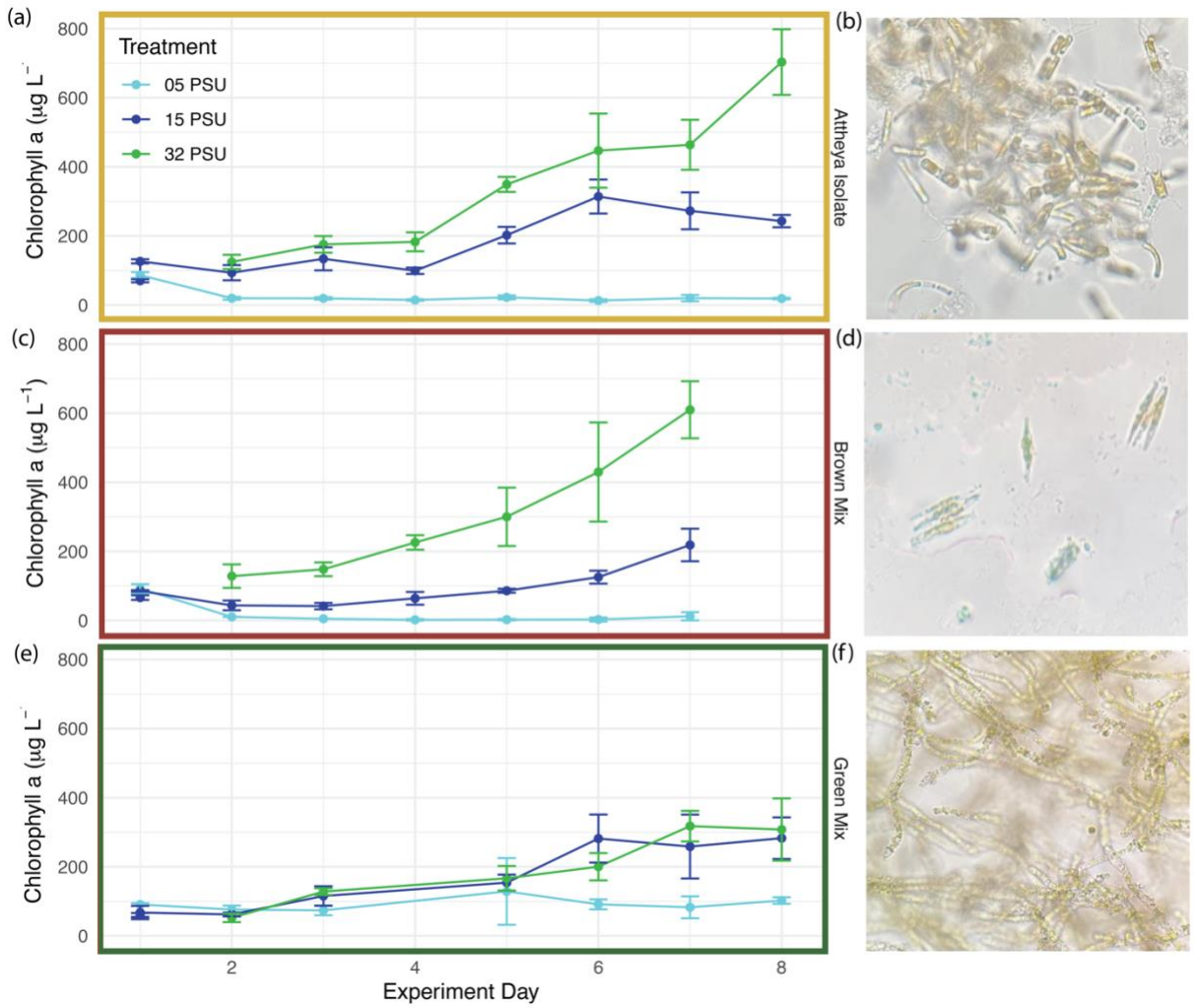


Figure 3-7. Chlorophyll *a* concentrations across salinity treatments and culture strains. Each of the three culture strains, Attheya isolate (a-b), Brown Mix (c-d), and Green Mix (e-f), were grown (staggered starts) in one of three salinity treatments for a total of 8 days during which chlorophyll *a* concentrations were collected daily. Points represent average values and lines the standard deviation between triplicate experimental bottles for each salinity treatment (a,d,e). Data from November 15th are omitted due to an instrument failure on that day. Microscope images (b,d,f) were photographed from the 32 PSU salinity treatment a week following the end of the experiment (R. Varner).

culture was visually marked by its brown pigmentation and a high abundance of singular pennate diatom cells (Figure 3-7d). These were likely *Synedra minescula*, or another unidentified Araphid-pennate diatom, which had 97 % BLAST® similarity to at least three other Fragilariophyceae species. This mixed community culture also had high relative abundances of heterotrophic nano flagellates (i.e. Cercozoa taxa) which became more abundant in the low salinity treatment, creating

a distinctive community cluster for the 5 PSU salinity treatment (Figure 3-8). Unfortunately, all experimental samples from the 'Green Mix' community culture had poor total read counts and had to be excluded from this analysis. Based on the dominant taxa at the end of the growth phase (Figure 3-S3) and microscopic imaging (Figure 3-7F), the dominant algae in this culture were likely species of *Ulothrix* and *Chlamydonomodales*.

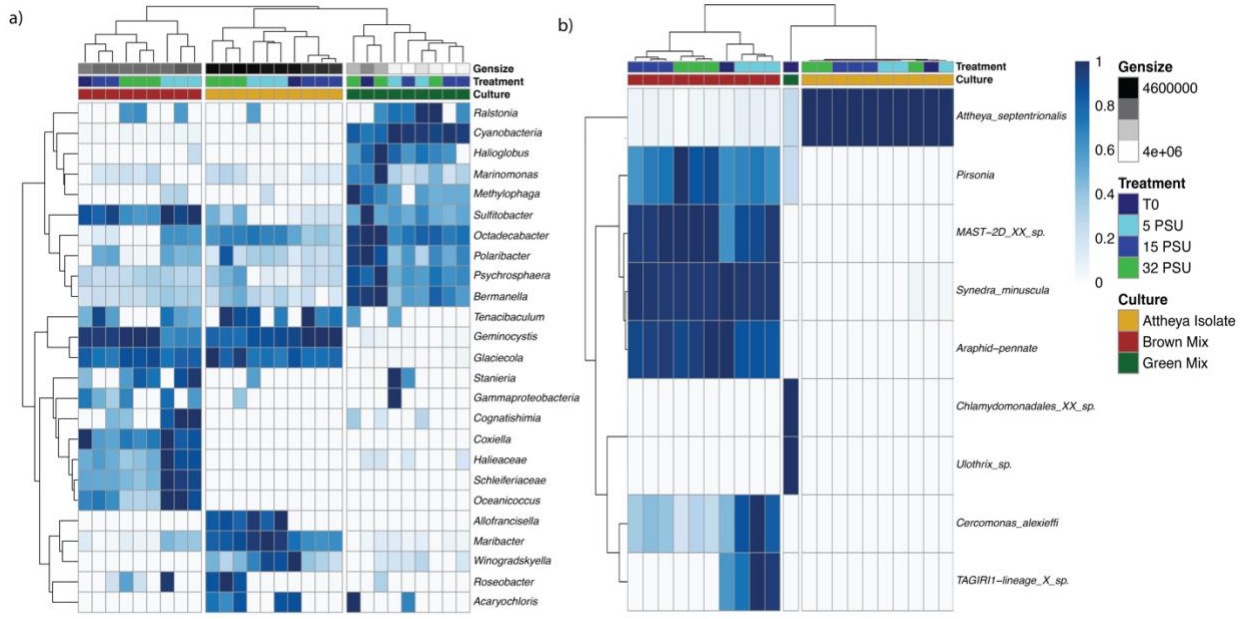


Figure 3-8. Bacterial and eukaryotic community structure in the algal cultures.
Heatmaps shows the most abundant bacterial (a) and eukaryotic (b) taxa present in each community culture strain. Columns (samples) are colored by culture type and salinity treatment (if viable, 3 replicates per treatment), including the initial community sample (labeled T0) collected prior to salinity dilutions. Both rows and columns were clustered using Euclidean distances, highlighting similarities in community composition. For bacterial samples, columns are additionally colored by estimated mean bacterial genome size (bp) for the sample (a).

The associated bacterial communities for each culture were similarly distinctive with the two diatom-based cultures showing the highest amount of overlap in community composition (Figure 3-8). Additionally, salinity treatments from these two experiments clustered together with some taxa showing enhanced relative abundances in the 5 PSU treatment (i.e., *Sulfitobacter, Winograskyella, and Marinobacter* sp.). Treatment clusters in the bacterial community within the

'Green Mix' remained more variable, yet several ASV edges showed decreased relative abundance with the 5 and 15 PSU salinity treatments, i.e. *Polaribacter* and *Psycrosphaera* sp., and one cyanobacterial edge showed increased relative abundance in the 5 PSU treatment (Figure 3-8).

For all cultures, growth curves tracking total chlorophyll *a* concentration (Figure 3-7) and autofluorescent cell counts (Figure 3-8) were lower in the 5 PSU treatment when compared to the 15 or 32 PSU. The Attheya Isolate culture reached the highest overall potential for chlorophyll *a* production in the 32 PSU control treatment, but the Green Mix retained the most chlorophyll in the 5 PSU and 15 PSU treatments. In the 5 PSU treatment, both diatom-based communities reached nearly 0 chlorophyll *a* concentration by the end of the 8-day growth period (Table 3-2). More macronutrients were removed from the media in the 32 PSU salinity treatment when compared to lower salinity treatments for all culture types (except Si in the Green Mix, which did not have silica added to its media base) and pH was lowered in the 5 PSU salinity treatment.

Table 3-2. Values of chlorophyll *a*, pH, macronutrients, and the percentage of actively respiring cells averaged across triplicate bottles after an 8-day acclimation period to salinity treatments.

*Mean treatment salinity values measured in PSU

Culture Treatment*		TF Chl (µg L^{-1})	TF pH	TF NO_3 (µmol L^{-1})	TF PO_4 (µmol L^{-1})	TF Si (µmol L^{-1})	% Respiring Cells
Attheya Isolate	**5.2**	18.6 ± 2.1	9.3 ± 0.1	839.7 ± 4.0	28.8 ± 0.5	85.8 ± 2.5	18.3 ± 17.6
	15	200.7 ± 59.4	10.0 ± 0.02	780.9 ± 48.6	16 ± 3.2	28.8 ± 21.9	9.5 ± 2.8
	32	703.2 ± 95	10.0 ± 0.2	748.3 ± 46.4	12.8 ± 2.6	14.1 ± 17.3	10.4 ± 11.6
Brown Mix	**5.5**	2.1 ± 1.9	9.3 ± 0.2	844.5 ± 3.5	30.6 ± 0	93.1 ± 17.3	56.4 ± 5.6
	14	218.5 ± 47.1	9.9 ± 0.4	796.8 ± 2.3	24.3 ± 0.9	63.4 ± 7.8	62.5 ± 7.1
	32.3	609.7 ± 82.6	9.8 ± 0.1	770.5 ± 6.0	23.2 ± 0.2	58.3 ± 3.5	75.2 ± 7.6
Green Mix	**5**	102.4 ± 9.5	10.4 ± 0.06	773.9 ± 8.3	15.9 ± 0.4	2.4 ± 1.4	63.0 ± 0.85
	15	282.9 ± 60.2	10.6 ± 0.08	695.7 ± 94.1	9.3 ± 1.1	3.5 ± 1.0	63.2 ± 2.7
	32.3	307.9 ± 90.2	10.0 ± 0.2	702.7 ± 9.5	9.1 ± 1.5	6.9 ± 2.6	57.5 ± 2.2

Table 3-3. Chlorophyll *a*, pH, and nutrient turnover during the 48 hr oxygen bottle incubations.

All Δ_{TF-T0} values in concentration were divided by 48 to calculate rates of uptake and production. *Nutrients measured in µmol L^{-1} hr^{-1}

	Culture Treatment		Δ_{TF-T0} Chl (µg $L^{-1}hr^{-1}$)	Δ_{TF-T0} pH (hr^{-1})	Δ_{TF-T0} NO_3 (*)	Δ_{TF-T0} PO_4 (*)	Δ_{TF-T0} Si (*)	Δ_{TF-T0} O_2 (µmol $kg^{-1}hr^{-1}$)
Light	**Attheya Isolate**	**5**	-2.4	0	-0.4	0	0	-0.2
		15	3.3	0	0.4	0	0.2	0.7
		32	11.5	0	-0.9	-0.1	-0.2	3.4
	Brown Mix	**5**	-0.7	0	0.1	0	0	0.2
		15	-5.8	0	-0.3	-0.2	-0.5	3.5
		32	-5.7	0	-0.2	-0.1	-0.3	5.2
	Green Mix	**5**	2.8	0	-0.4	0	0	0.9
		15	3	0	0.9	0	0	1.8
		32	-2.7	0	-0.5	0	0	5.3
Dark	**Attheya Isolate**	**5**	-2.7	0.2	17.2	0.6	1.8	-0.9
		15	-6.5	0.2	16.7	0.3	0.6	-1.3
		32	-6.4	0.2	14.9	0.3	0.1	-2.6
	Brown Mix	**5**	-0.7	0.2	17.6	0.6	1.4	-0.8
		15	-10.6	0.2	16.4	0.5	1.1	-1.6
		32	-15.7	0.1	16	0.3	1.1	-4
	Green Mix	**5**	0.7	0.2	16.6	0.3	0.1	-0.2
		15	-5.5	0.2	16.9	0.2	0.1	-1.3
		32	-12.9	0.1	14.6	0.1	0.1	-2.4

Changes in culture net metabolic activity across salinity treatments

Over the 48-hour incubation period, all salinity treatments had metabolic oxygen turnover with rates ranging from oxygen consumption of –4 µmol kg^{-1} hr^{-1} to 5.3 µmol kg^{-1} hr^{-1} (Table 3-3). The Brown Mix culture showed the highest potential for both total biological oxygen production under light conditions, and oxygen consumption under dark conditions under seawater salinity (Figure 3-9). However, the Green Mix culture had the highest oxygen production in the low salinity 5 PSU salinity treatment, as well as a higher proportion of metabolically active cells when compared to total cell count going into the incubation (Table 3-2). The relative proportion of actively respiring cells in the Attheya Isolate culture was low across all salinity treatments, however direct comparisons to the other two culture communities for this parameter cannot be made due to differences in QC procedures for these samples. Nevertheless, the Attheya Isolate community culture displayed a significant reduction in net oxygen production in the 15 PSU

treatment, and net oxygen consumption even under light conditions in the 5 PSU treatment (Figure 3-9). Patterns of nutrient update and remineralization generally matched trends in oxygen consumption within the bottle incubations (Table 3-3).

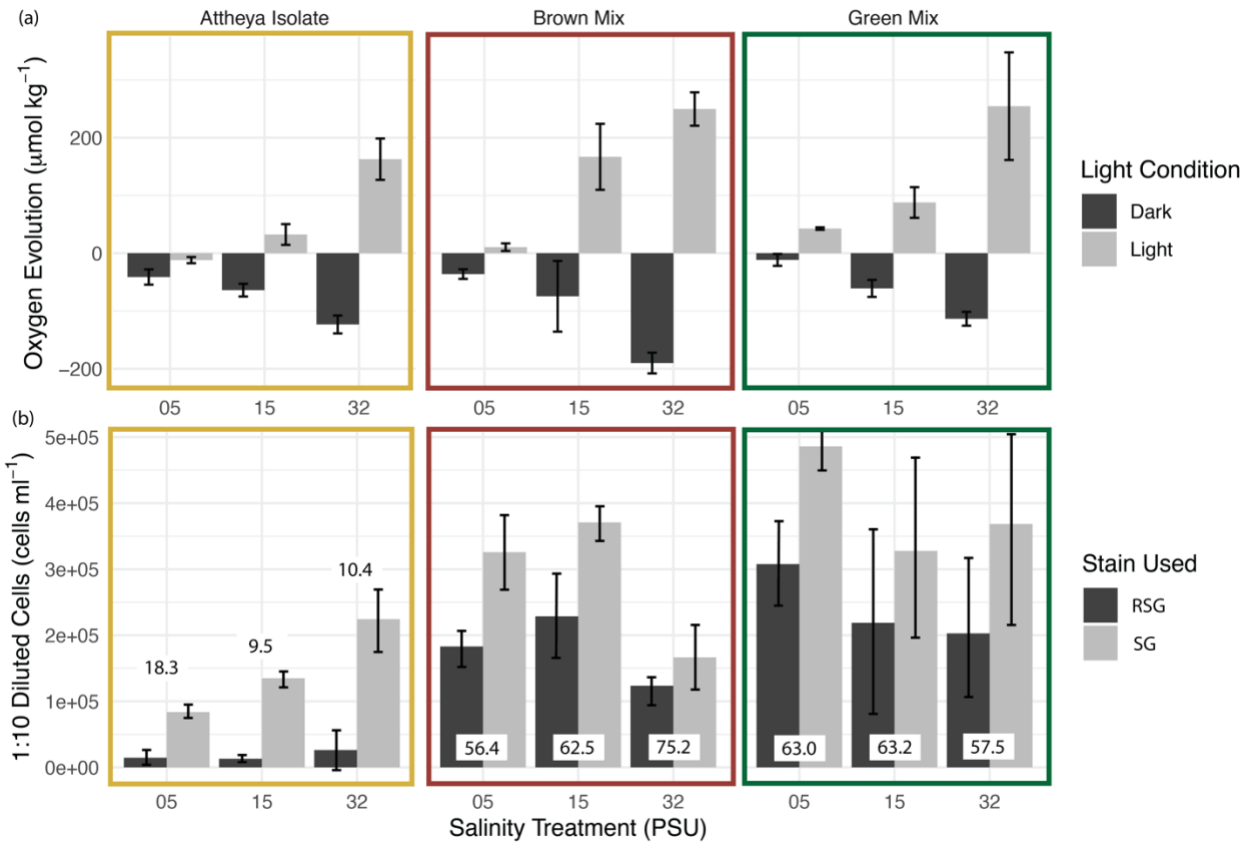


Figure 3-9. Oxygen production and aerobic respiration across salinity treatments. The net change in total oxygen (a) as measured over a 48 hr incubation period for both light and dark bottles across salinity treatment and culture strains. Error bars represent the standard deviation in oxygen turnover between experimental replicates. Measured at the end of the 8-day acclimation period, prior to oxygen incubations, panel (b) indicates cell count (1:10 dilution) for both total (SG stained) cells and cells undergoing active bacterial reductase activity (RSG stained). The number label represents % of actively respiring cells (RSG/SG). Error bars represent the standard deviation in counted cells with each fluorescent stain between experimental replicates.

Discussion

Salinity induced microhabitats: seeding and evolution of the meltwater community

The closer similarity in community structure between samples from the meltwater layer and sea ice compared to the meltwater layer and seawater on both legs of the MOSAiC Expedition

(Figure 3-3) indicate that these communities were likely seeded from the ice, as the primary origin of this meltwater and consistent ice flushing observed over the melt season (Salganik et al. 2023). This supports the hypothesis that sea ice, specifically multi-year ice, serves as a 'seed repository', where algae from the previous winter trapped in the growing ice melts out the following spring to initiate under ice-blooms and algal growth in the newly formed ice of adjacent leads (Olsen et al. 2017). While most of the leads sampled in this study were primarily surrounded by first-year ice (Table 3-1), previous studies have indicated that meltwater accumulated from a large spatial area with contributions from both the sea ice surface and basal melt (Smith et al. 2023; Salganik et al. 2023; Nomura et al. 2023).

When salinity stratification was maintained, both observed meltwater layers then evolved as an isolated system with a subset of the seeded community surviving the osmotic shock of moving from high salinity ice conditions (Torstensson et al. 2015) to the low salinity conditions of the meltwater layer (Figure 3-2). This was explicitly observed in our lab experiments, where the 'Brown Mix' strain showed changes in community composition after only 8 days of acclimation to the low salinity treatment (Figure 3-8). Trapped in the meltwater layer by the density gradient (Assmy et al. 2013), and exposed to additional stressors such as high light and, potentially, nutrient limitation (Mundy et al. 2011) it is likely these communities evolved very rapidly *in situ*, distinguishing themselves from both the original ice communities and the communities present in the underlying seawater (Figure 3-3). On Leg 5, meltwater communities remained distinct even while the meltwater layer thinned with the onset of freezing on September 1 (Nomura et al. 2023). However, after stratification ended following turbulent mixing from storm activity on September 6, surface water communities became almost identical to that of the previously sampled seawater from below 1 m. Either, the meltwater community was mixed down with the remaining meltwater

(Nomura et al. 2023), or, incorporated into the growing lead ice in platelet crystals (Syvertsen and Syoertsen 1991), which is likely considering the amount of observed OM attached to growing platelet ice during the expedition (Supplemental File 3-1).

The "Interface" layer, while containing some unique ASVs (Figure 3-3) did not contain its own 'halocline flora' as has previously been described in other studies (Apollonio 1985). However, the thin nature (~1 cm thick) nature of this layer made it particularly difficult to sample and it is possible that samples were diluted with either water from above or below while pumping. On Leg 4, however, there was a clear functional distinction in the Interface layer, corresponding to the visible biomass layer (Figure 3-2) which was observed at most sites on most dates (Table 3-1). This biomass layer had high diversity for prokaryotes (Figure 3-4d), but low diversity for eukaryotes (3-5d), indicating "bloom" like conditions with a diverse responsive heterotrophic community. The phytoplankton community in the Interface at this time was dominated by Pelagophyceae ASVs (Figure 3-S1) and high standing stocks of chlorophyll as biomass accumulated at the density gradient (Figure 3-5). It is possible that this biomass layer represents dead OM flushed from the ice accumulating at the density gradient. However, Freyria et al. (2022) found Artic Pelagophyte species to be tolerant of salinity fluctuations through differential gene expression resulting in the production of secondary metabolites and lipid secretion and suggesting it is just as likely that they were actively growing within the meltwater and/or the meltwater-seawater interface. Experiments in the Antarctic highlight the potential release of such metabolites into the labile dissolved organic pool during low salinity conditions where they can be readily metabolized by the bacterial community (Dawson et al. 2023), or used as osmoprotectants (Firth et al. 2016). Sample to sample variability in bacterial community composition was high for the Interface layer, but generally the interface bacterial community at this time had a significantly

higher estimated average genome size, a high proportion of HNA cells compared to the underlying seawater, and high relative abundances of copiotrophic Flavobacteria ASVs (Figure 3-4). This, in addition to the high relative abundances of motile heterotrophic protists in the meltwater and Interface layers (Figure 3-5) indicate a quick responding microbial loop within the thin layer of accumulated phytoplankton biomass (Buchan et al. 2014; Norris et al. 2021). Unfortunately, we did not collect oxygen measurements from leads on this leg to confirm net metabolic activity, but whether the Interface microhabitat was net autotrophic or net heterotrophic, the observed microbial community clearly had the potential to impact carbon and nutrient cycling for the nutrient poor surface Arctic Ocean in as much a capacity as the ice algal aggregates observed on Leg 5 and on previous cruises (Fernández-Méndez et al. 2014).

Potential for increased heterotrophy in the meltwater layer

During Leg 5 of MOSAiC, biological oxygen concentration in the meltwater was significantly undersaturated compared to both the interface and underlying seawater (Figure 3-6). A physical explanation of this observation would be gas expulsion as the surface of the leads freeze over, leading to undersaturated meltwater (Eveleth et al. 2014). However, physical oxygen supersaturation calculated from argon measurements were variable across samples with no corresponding pattern, and total oxygen concentration was highest in the meltwater layer (Figure 3–6). Neither of these measurements showed significant ($p > 0.05$) variability in mean values between salinity layers, indicating a true biological signal for net heterotrophy at the surface.

Biomass in the Leg 5 meltwater layer was visually and analytically very different from the meltwater layer observed on Leg 4 – likely due to the seasonal, spatial, and physiochemical differences between ice floes. There was no visual biomass at the interface layer, but instead algal aggregates were observed floating in the meltwater layer during visual surveys (Supplemental File

3-S1) similar to those described in Nansen (1902), Assmy et al. (2013), and Fernández-Méndez et al. (2014). However, these aggregates were all described as primarily composed of diatoms. In the 18S sequencing of bulk water, diatoms represented a very small proportion of sequences in the meltwater layer (Figure 3-S1), suggesting that our sampling method may not have been very effective for capturing in situ measurements of aggregate composition or the aggregates which we observed were pico-eukaryote or dinoflagellate in origin. Regardless, individual meltwater aggregates have previously been observed to be net heterotrophic (Fernández-Méndez et al. 2014) and their accumulation associated with oxygen consumption and elevated ammonium concentrations in the surrounding seawater (Assmy et al. 2013), supporting evidence of rapid remineralization by bacterial colonizers recruited from the sea ice community (Rapp et al. 2018). In addition to high abundances of ASVs related to traditional melt pond and ice taxa (i.e. ß-Proteobacteria and Actinobacteria; Figure 3-S1), the high genome size, high HNA bacterial community in the Leg 5 meltwater layer contained high abundances of Flavobacteria ASVs (Figure 3-4), which can efficiently use high molecular weight ice derived OM (Underwood et al. 2019) and are likely major contributors to the net heterotrophic signal observed in the meltwater layer.

Based on our culture experiments, the net consumption of biological oxygen observed in the Leg 5 meltwater layer was likely due to a combination of reduced net community production and persistent respiration under low salinity stress. All culture strains saw reduced growth and biomass with time in the low salinity treatments (Figure 3-7). Testing across multiple community types, we found that only the Attheya Isolate strain reached total net heterotrophy under light conditions in the 5 PSU treatment (Figure 3-9). Centric diatoms like *Attheya* have a greater tolerance than pennate diatoms to low salinity conditions, however their high abundance has also been linked to abrupt shifts to net heterotrophic conditions in bottom sea ice (Campbell et al. 2017).

Under our 5 PSU treatment, the percentage of actively respiring cells in culture increased by ~8 % compared to the control and 15 PSU treatment. We cannot directly compare total % respiring cells between algal culture strains as it is likely that total counts for the RSG stained cells in the Attheya culture are underestimated due to methodological differences in the killed control treatments for this experiment. However, we can compare to % changes within culture strains. The Brown Mix showed a decrease in the % respiration of cells with decreasing salinity, while the Green Mix showed a ~6 % increase between the control and both lower salinity treatments (Figure 3-9). The Green Mix was also the only community to show an increase in the total number of cells, likely related to the relative increase in Cyanobacteria ASVs (Figure 3-8), which have been observed seasonally in the central Arctic Ocean (Gradinger & Lenz, 1995).

The mixed communities both remained net productive under low salinity stress, however this was likely due to species compositional changes during low salinity acclimation (Figure 3-8). The Green Mix culture showed the highest oxygen production rates under the 5 PSU treatment and had the greatest chlorophyll *a* retention under low salinity. However, it had the lowest chlorophyll *a* concentrations overall (Figure 3-7). This culture was primarily composed of the filamentous green algae Ulothrix, flagellated Chlorophyta, and high relative abundances of Cyanobacteria (Figure 3-8). Autotrophic flagellates tend to dominate meltwater communities (Figure 3-5; Mundy et al. 2011) and if these experimental results were to be scaled up, they would support predictions that under a warming and freshening Arctic, smaller phytoplankton will outcompete diatoms at the ice-water interface with implications for less total production overall and a stronger microbial loop (Li et al. 2009). The Brown Mix Community started with the most similar community composition to the original NL lead ice slush sample (Figure 3-5; Figure 3-8) as well as other observed central Arctic algal aggregates (Assmy et al. 2013) and had the highest potential for both light oxygen

production and dark oxygen consumption (Table 3-3) at rates which were similar to those observed in bottom ice bloom conditions (Campbell et al. 2017). However, production rates dropped rapidly after forced acclimation to the low salinity treatments and community composition within the culture saw a significant increase in the relative abundance of heterotrophic flagellates in the low salinity 5 PSU treatment indicating what would likely become a switch to net heterotrophic conditions over a longer acclimation period.

Summary and implications

Our results regarding microbial community structure in the meltwater layer align well with what has been found in previous studies of Arctic freshwater stratification and isolated meltwater environments, particularly regarding overall community composition. The ice-derived *in situ* eukaryotic community was dominated by smaller phytoplankton (i.e. autotrophic flagellates) and heterotrophic protists (Li et al. 2009; Mundy et al. 2011) while the prokaryotic community was dominated by a combination of traditional melt pond taxa (Boetius et al. 2015) and ice-derived copiotrophs with the ability to respond rapidly to either the phytoplankton bloom observed at the meltwater-seawater interface near the marginal ice zone in July, or congregated ice-algal aggregates bound within the strong density gradient of the more northern site's stratified meltwater layer. This active heterotrophy within the meltwater, as confirmed across multiple bacterial and ice-algal communities by our *in vitro* experiments, leads to a stratified surface community uniquely primed for metabolic oxygen consumption, impacting biogeochemical cycling at the air-sea interface of open and re-freezing leads.

As climate change enhances the dynamic nature of arctic ice cover due to thinning (Kwok et al. 2013) and accelerated motion (Itkin et al. 2017), it is likely that the number and frequency of open leads within the central Arctic pack ice will increase. Leads serve as an important hot spot

for unrestricted biogeochemical exchange with the atmosphere in an otherwise ice-covered ocean basin (Loose et al. 2014). Biological undersaturation at the surface would make stratified leads a significant sink of DO compared to the surrounding sea ice (Smith et al. 2023). Increased respiration of accumulating algal aggregate material within this layer would also increase the likelihood of favorable conditions for the biogenic production of methane and other climate relevant gasses. Previous studies have shown that high respiration rates in surface waters can increase the pCO_2 of growing sea ice (Zhou et al. 2016), impacting sea ice carbon cycling beyond just the microbial loop and net potential for net OM export (Rapp et al. 2018). Trapped OM at the air-sea interface can also lead to increased biogenic ice nucleating particle (INP) production, impacting cloud formation. On MOSAiC, increased fluxes of biogenic INPs were observed during the melt season (Creamean et al. 2022) and future work will look into constraining the microbial communities primarily responsible for contributing to this flux (Mock et al. 2022).

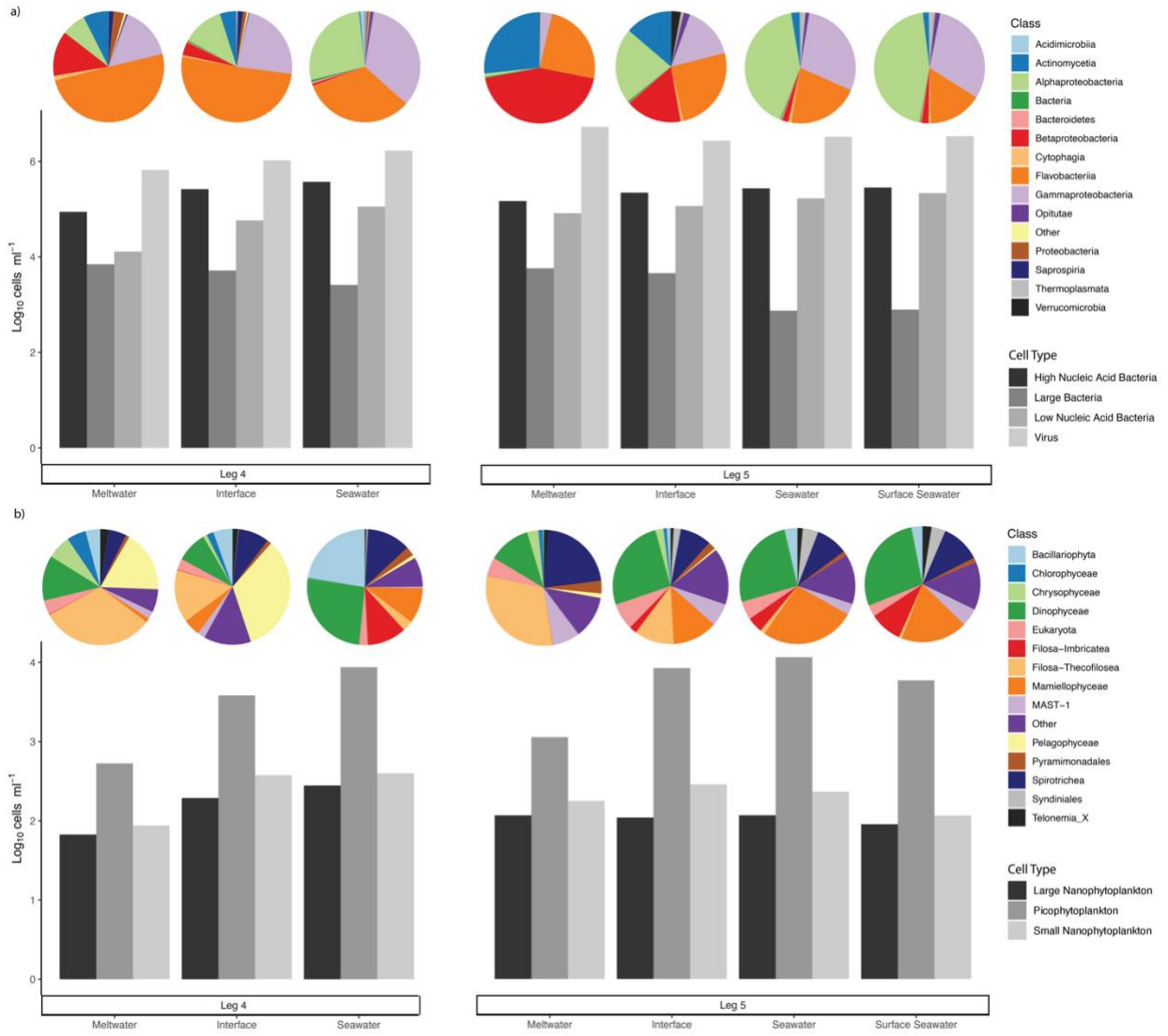
Ultimately, influences of low salinity meltwater on microbial community structure and metabolism have the potential to impact far more than just microbial community dynamics in Arctic leads. Constraining the scale of these impacts as well as the direct mechanisms behind their linkages to other components of the Arctic climate system will be important for determining if, and how, these ephemeral features and their contributions to biogeochemical cycles should be incorporated into ecological models when making predictions for a warmer, meltier, Arctic.

Supplemental Material

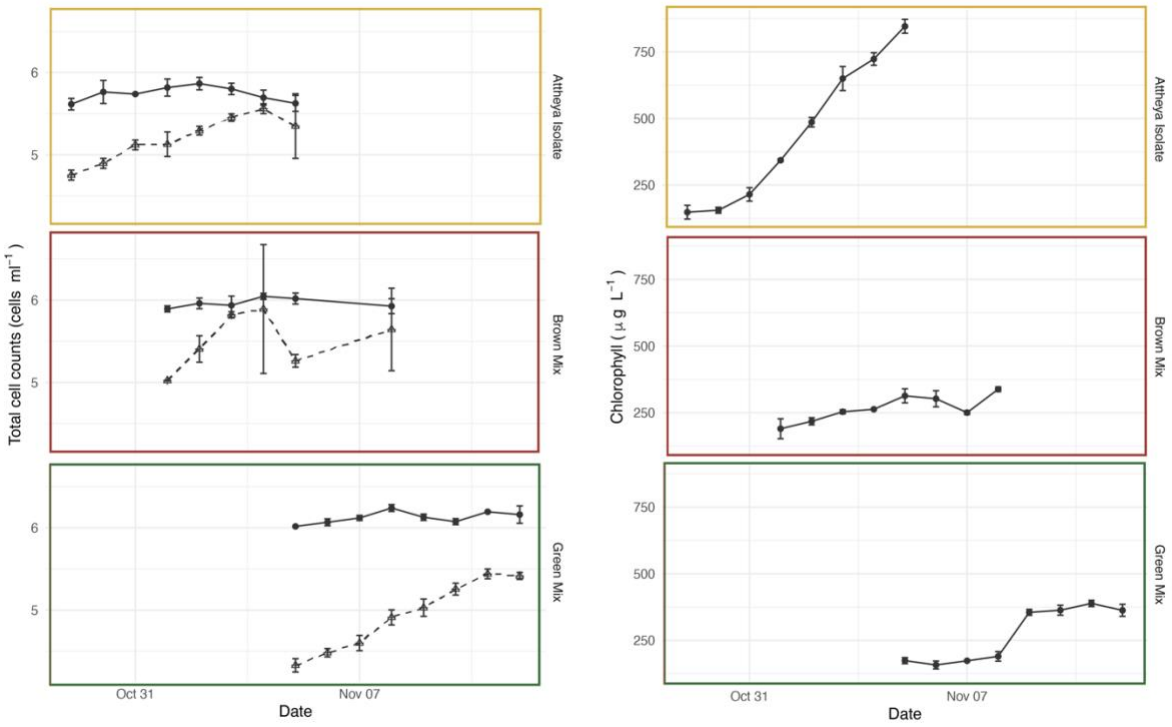
Text 3-S1. Methods: Isolation of Attheya monocultures.

Two- weeks post inoculation into fresh media at University of Warwick, a cell count was taken of each of the five Brown Mix communities, irrespective of cell type or species, using a light microscope (NIKON Eclipse, Japan) and Neubauer haemocytometer with 3x replicate counts. The resultant cell density was diluted with fresh media 1:10 six times, to dilute to extinction and achieve a mean cell density of < 1 cell ml^{-1}. Five 96-well culture plates were filled with 0.25 ml of the dilute cultures, and 12 wells per plate of sterile media as control. Plate lids were sealed with parafilm and stored alongside the parent cultures in the incubator for 3 months without movement. Following this growth period, wells identified with singular 'colonies' of cells were reinoculated into 5 ml of fresh media in a sterile sample tube, and grown for a month before the addition of 10ml of fresh media.

In total, 100 isolates were obtained, and checked to be monoclonal cultures by both examination under light microscopy, and by flow cytometry (CytoFLEX). No cells were identified in any of the control wells, showing that isolates originated from the Arctic culture. In preparation for various experiments, ten of the isolate strains were chosen at random, two from each community culture. These strains were then grown up to larger culture volumes in 250 ml tissue culture flasks. Three sub-samples of each community were prepared, growing at salinities of 20, 32 and 45 PSU, with stepwise salinity changes of 5 PSU weekly. Eventually, following microscope examination and flow cytometry confirmation, two of the strains were selected as suitable for use in further experiments. In these experiments, we use isolates from culture 20 isolated from plate 4, designated "strain 4.20".



Supplemental Figure 3-S1. Full overview of microbial community composition in leads. Pie charts show the relative percentage of assigned prokaryotic (a) or eukaryotic (b) taxonomic class (or next finest taxonomic ID) from combining all 16S (a) or 18S (b) amplicon sequence variants (ASV) read counts from each stratified layer of water collected during lead sampling on Legs 4 and 5 of the MOSAiC Expedition. 'Meltwater' indicates samples collected above 10 cm depth with a salinity of < 15 g kg^{-1}. 'Seawater' indicates samples collected from the underlying seawater. 'Interface' indicates where the two previous layers met, and 'Surface Seawater' indicates surface samples collected above 10 cm depth following disintegration of the meltwater lens (salinity of 21 – 33 g kg^{-1}). Bar charts in (a) show a normalized count of the total cell abundances for generally 'large' bacterial cells, High Nucleic Acid bacterial cells, Low Nucleic Acid bacterial cells, and viruses. Bar charts in (b) show a normalized count of the total cell abundances for large nanophytoplankton, small nanophytoplankton, and picophytoplankton.



Supplemental Figure 3-S2. Growth curves during the initiation phase of the three algae cultures.
Daily total cell counts (SYBR stained; circles with filled lines) and Autofluorescent cell counts (unstained; triangles with dashed lines) are shown on the left for each of the three culture strains. Daily total chlorophyll *a* concentrations for each of the three culture strains are shown on the right. Samples were collected over 8-day growth periods in preparation for salinity experiments.

Supplemental File 3-S1. Chamberlain_Video_LeadSurveys.mp4
This file is a compilation of visual lead surveys collected with a go-pro camera at lead sampling sites during the MOSAiC Expedition. Lead site and date sample notated in the top left corner.

Acknowledgements

Chapter 3, in part is currently being prepared for submission for publication of the material. Chamberlain, Emelia J; Webb Alison; Müller, Oliver; Hoppe, Clara JM; Fong, Allison A; Droste, Elise; D'Angelo Alessandra; Varner, Ruth; Smith, Madison M; Nomura Daiki; Kraberg Alexandra; Rokitta, Sebastian; Rost Björn; Schäfer, Hendrik; Loose, Brice; Bowman, Jeff. The dissertation author was the primary researcher and author of this material.

This work was funded by the National Science Foundation (NSF) Award OPP 1821911 to JB and an NSF Graduate Research Fellowship awarded to EJC.

Field data used in this study were produced as part of the international Multidisciplinary Drifting Observatory for the Study of Arctic Climate (MOSAiC) project with the data tag MOSAiC20192020. I would like to thank all persons involved in the expedition of the Research Vessel *Polarstern* (Knust, 2017) during MOSAiC in 2019-2020 as listed in the MOSAiC extended acknowledgement (Nixdorf et al. 2021). I would specifically like to acknowledge all participants in the MOSAiC Meltwater Working Group, author list Smith et al. (2023), whose many meetings incentivized and encouraged this work. Additionally, I would like to acknowledge Julia Diaz, Maria Vernet, and the Vernet and Diaz Labs (University of California San Diego) for use of their labs and equipment for culture maintenance. Finally, I would like to thank Elizabeth Connors, Kaycie Lanpher, Natalia Erazo, Shannon Perry, Grace Wang, Sandy Nguyenphuoc, and Caitlyn Webster (University of California San Diego), without whom the laboratory component of this work would not have been possible.

REFERENCES

Aalto, N. J., Schweitzer, H. D., Krsmanovic, S., Campbell, K., & Bernstein, H. C. (2022). Diversity and Selection of Surface Marine Microbiomes in the Atlantic-Influenced Arctic. *Frontiers in Microbiology*, *13*, 892634. https://doi.org/10.3389/FMICB.2022.892634/BIBTEX

Amaral-Zettler, L. A., McCliment, E. A., Ducklow, H. W., & Huse, S. M. (2009). A method for studying protistan diversity using massively parallel sequencing of V9 hypervariable regions of small-subunit ribosomal RNA Genes. *PLoS ONE*, *4*(7), e6372. https://doi.org/10.1371/journal.pone.0006372

Amiraux, R., Burot, C., Bonin, P., Massé, G., Guasco, S., Babin, M., Vaultier, F., & Rontani, J.-F. (2020). Stress factors resulting from the Arctic vernal sea-ice melt. *Elementa: Science of the Anthropocene*, *8*(1). https://doi.org/10.1525/elementa.076

Andersen, R. A., Berges, J. A., Harrison, P. J., & Watanabe, M. M. (2005). *Algal culturing techniques* (R. A. Anderson, Ed.). Imprint Academic Press.

Apollonio, S. (1985). Arctic Marine Phototrophic Systems: Functions of Sea Ice Stabilization. *ARCTIC*, *38*(3), 167–173. https://doi.org/10.14430/ARCTIC2129

Assmy, P., Ehn, J. K., Fernández-Méndez, M., Hop, H., Katlein, C., Sundfjord, A., Bluhm, K., Daase, M., Engel, A., Fransson, A., Granskog, M. A., Hudson, S. R., Kristiansen, S., Nicolaus, M., Peeken, I., Renner, A. H. H., Spreen, G., Tatarek, A., & Wiktor, J. (2013). Floating Ice-Algal Aggregates below Melting Arctic Sea Ice. *PLOS ONE*, *8*(10), e76599. https://doi.org/10.1371/JOURNAL.PONE.0076599

Azam, F., Fenchel, T., Field, J., Gray, J., Meyer-Reil, L., & Thingstad, F. (1983). The ecological role of water-column microbes in the sea. *Mar Ecol Prog Ser*, *10*, 257–263. https://doi.org/10.3354/meps010257

Barbera, P., Kozlov, A. M., Czech, L., Morel, B., Darriba, D., Flouri, T., & Stamatakis, A. (2019). EPA-ng: massively parallel evolutionary placement of genetic sequences. *Systematic Biology*, *68*(2), 365–369. https://doi.org/10.1093/sysbio/syy054

Bayer, B., McBeain, K., Carlson, C. A., & Santoro, A. E. (2023). Carbon content, carbon fixation yield and dissolved organic carbon release from diverse marine nitrifiers. *Limnology and Oceanography*, *68*(1), 84–96. https://doi.org/10.1002/LNO.12252

Becker, S., Aoyama, M., Woodward, E. M. S., Bakker, K., Coverly, S., Mahaffey, C., & Tanhua, T. (2020). GO-SHIP repeat hydrography nutrient manual: the precise and accurate determination of dissolved inorganic nutrients in seawater, using continuous flow analysis methods. *Frontiers in Marine Science*, *7*, 581790. https://doi.org/10.3389/FMARS.2020.581790/BIBTEX

Bernardet, J.-F., & Nakagawa, Y. (2006). An introduction to the family Flavobacteriaceae. In *The Prokaryotes* (pp. 455–480). Springer, New York, NY. https://doi.org/10.1007/0-387-30747-8_16

Boetius, A., Albrecht, S., Bakker, K., Bienhold, C., Felden, J., Fernández-Méndez, M., Hendricks, S., Katlein, C., Lalande, C., Krumpen, T., Nicolaus, M., Peeken, I., Rabe, B., Rogacheva, A., Rybakova, E., Somavilla, R., Wenzhöfer, F., & Felden, J. (2013). Export of algal biomass from the melting arctic sea ice. *Science*, *339*(6126), 1430–1432. https://doi.org/10.1126/SCIENCE.1231346/SUPPL_FILE/BOETIUS.SM.PDF

Boetius, A., Anesio, A. M., Deming, J. W., Mikucki, J. A., & Rapp, J. Z. (2015). Microbial ecology of the cryosphere: Sea ice and glacial habitats. In *Nature Reviews Microbiology* (Vol. 13, Issue 11, pp. 677–690). Nature Publishing Group. https://doi.org/10.1038/nrmicro3522

Bowman, J. P., McCammon, S. A., Brown, J. L., & McMeekin, T. A. (1998). Glaciecola punicea gen. nov., sp. nov. and Glaciecola pallidula gen. nov., sp. nov.: psychrophilic bacteria from Antarctic sea-ice habitats. *International Journal of Systematic and Evolutionary Microbiology*, *48*(4), 1213–1222. https://doi.org/10.1099/00207713-48-4-1213

Bowman, J. S. (2021). Making sense of a scent-sensing metaphor for microbes and environmental predictions. *MSystems*, *6*(4). https://doi.org/10.1128/MSYSTEMS.00993-21

Bowman, J. S., Amaral-Zettler, L. A., Rich, J. J., Luria, C. M., & Ducklow, H. W. (2017). Bacterial community segmentation facilitates the prediction of ecosystem function along the coast of the western Antarctic Peninsula. *The ISME Journal 2017 11:6*, *11*(6), 1460–1471. https://doi.org/10.1038/ismej.2016.204

Bowman, J. S., & Ducklow, H. W. (2015). Microbial communities can be described by metabolic structure: A general framework and application to a seasonally variable, depth-stratified microbial community from the coastal West Antarctic Peninsula. *PLOS ONE*, *10*(8), e0135868. https://doi.org/10.1371/journal.pone.0135868

Bray, J. R., & Curtis, J. T. (1957). An ordination of the upland forest communities of southern Wisconsin. *Ecological Monographs*, *27*(4), 325–349. https://doi.org/10.2307/1942268

Breiman, L. (2001). Random Forests. *Machine Learning 2001 45:1*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Brinkmeyer, R., Glöckner, F. O., Helmke, E., & Amann, R. (2004). Predominance of β-proteobacteria in summer melt pools on Arctic pack ice. *Limnology and Oceanography*, *49*(4), 1013–1021. https://doi.org/10.4319/LO.2004.49.4.1013

Buchan, A., LeCleir, G. R., Gulvik, C. A., & González, J. M. (2014). Master recyclers: features and functions of bacteria associated with phytoplankton blooms. *Nature Reviews Microbiology 2014 12:10*, *12*(10), 686–698. https://doi.org/10.1038/nrmicro3326

Buck, K. R., Nielsen, T. G., Hansen, B. W., Gastrup-Hansen, D., & Thomsen, H. A. (1998).

Infiltration phyto- and protozooplankton assemblages in the annual sea ice of Disko Island, West Greenland, spring 1996. *Polar Biology 1998 20:6*, *20*(6), 377–381. https://doi.org/10.1007/S003000050317

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods*, *13*(7), 581–583. https://doi.org/10.1038/nmeth.3869

Campbell, K., Lange, B. A., Landy, J. C., Katlein, C., Nicolaus, M., Anhaus, P., Matero, I., Gradinger, R., Charette, J., Duerksen, S., Tremblay, P., Rysgaard, S., Tranter, M., Haas, C., & Michel, C. (2022). Net heterotrophy in High Arctic first-year and multi-year spring sea ice. *Elementa*, *10*(1). https://doi.org/10.1525/ELEMENTA.2021.00040/119112

Campbell, K., Matero, I., Bellas, C., Turpin-Jelfs, T., Anhaus, P., Graeve, M., Fripiat, F., Tranter, M., Landy, J. C., Sanchez-Baracaldo, P., Leu, E., Katlein, C., Mundy, C. J., Rysgaard, S., Tedesco, L., Haas, C., & Nicolaus, M. (2022). Monitoring a changing Arctic: Recent advancements in the study of sea ice microbial communities. *Ambio*, *51*(2), 318–332. https://doi.org/10.1007/S13280-021-01658-Z/FIGURES/6

Campbell, K., Mundy, C. J., Gosselin, M., Landy, J. C., Delaforge, A., & Rysgaard, S. (2017). Net community production in the bottom of first-year sea ice over the Arctic spring bloom. *Geophysical Research Letters*, *44*(17), 8971–8978. https://doi.org/10.1002/2017GL074602

Campbell, K., Mundy, C. J., Juhl, A. R., Dalman, L. A., Michel, C., Galley, R. J., Else, B. E., Geilfus, N. X., & Rysgaard, S. (2019). Melt Procedure Affects the Photosynthetic Response of Sea Ice Algae. *Frontiers in Earth Science*, *7*, 21. https://doi.org/10.3389/feart.2019.00021

Carmack, E. C., Yamamoto-Kawai, M., Haine, T. W. N., Bacon, S., Bluhm, B. A., Lique, C., Melling, H., Polyakov, I. V., Straneo, F., Timmermans, M. L., & Williams, W. J. (2016). Freshwater and its role in the Arctic Marine System: Sources, disposition, storage, export, and physical and biogeochemical consequences in the Arctic and global oceans. *Journal of Geophysical Research: Biogeosciences*, *121*(3), 675–717. https://doi.org/10.1002/2015JG003140

Carter-Gates, M., Balestreri, C., Thorpe, S. E., Cottier, F., Baylay, A., Bibby, T. S., Moore, C. M., & Schroeder, D. C. (2020). Implications of increasing Atlantic influence for Arctic microbial community structure. *Scientific Reports 2020 10:1*, *10*(1), 1–13. https://doi.org/10.1038/s41598-020-76293-x

Cassar, N., Difiore, P. J., Barnett, B. A., Bender, M. L., Bowie, A. R., Tilbrook, B., Petrou, K., Westwood, K. J., Wright, S. W., & Lefevre, D. (2011). The influence of iron and light on net community production in the Subantarctic and Polar Frontal Zones. *Biogeosciences*, *8*(2), 227–237. https://doi.org/10.5194/bg-8-227-2011

Cavicchioli, R., Ripple, W. J., Timmis, K. N., Azam, F., Bakken, L. R., Baylis, M., Behrenfeld, M. J., Boetius, A., Boyd, P. W., Classen, A. T., Crowther, T. W., Danovaro, R., Foreman, C. M., Huisman, J., Hutchins, D. A., Jansson, J. K., Karl, D. M., Koskella, B., Mark Welch, D. B., … Webster, N. S. (2019). Scientists' warning to humanity: microorganisms and climate change. *Nature Reviews. Microbiology*, *17*(9), 569–586. https://doi.org/10.1038/S41579-019-0222-5

Chamberlain, E., & Bowman, J. (2022). *16S and 18S rRNA amplicon data for the Multidisciplinary drifting Observatory for the Study of Arctic Climate (MOSAiC) expedition in the central Arctic Ocean, 2019-2020.* Arctic Data Center. https://doi.org/10.18739/A2CC0TV5X

Chamberlain, E. J., Balmonte, J. P., Torstensson, A., Fong, A. A., Snoeijs-Leijonmalm, P., & Bowman, J. S. (2022). Impacts of sea ice melting procedures on measurements of microbial community structure. *Elementa: Science of the Anthropocene*, *10*(1). https://doi.org/10.1525/ELEMENTA.2022.00017

Cox, G. F. N., & Weeks, W. F. (1982). Equations for determining the gas and Brine volumes in sea ice samples. *CRREL Report (US Army Cold Regions Research and Engineering Laboratory)*, *29*(102), 306–316. https://doi.org/10.3189/s0022143000008364

Craig, H., & Hayward, T. (1987). Oxygen supersaturation in the ocean: Biological versus physical contributions. *Science*, *235*(4785), 202–205. https://doi.org/10.1126/science.235.4785.199

Creamean, J. M., Barry, K., Hill, T. C. J., Hume, C., DeMott, P. J., Shupe, M. D., Dahlke, S., Willmes, S., Schmale, J., Beck, I., Hoppe, C. J. M., Fong, A., Chamberlain, E., Bowman, J., Scharien, R., & Persson, O. (2022). Annual cycle observations of aerosols capable of ice formation in central Arctic clouds. *Nature Communications 2022 13:1*, *13*(1), 1–12. https://doi.org/10.1038/s41467-022-31182-x

Czech, L., Barbera, P., & Stamatakis, A. (2020). Genesis and Gappa: processing, analyzing and visualizing phylogenetic (placement) data. *Bioinformatics*, *36*(10), 3263–3265. https://doi.org/10.1093/bioinformatics/btaa070

Dawson, H. M., Connors, E., Erazo, N. G., Sacks, J. S., Mierzejewski, V., Rundell, S. M., Carlson, L. T., Deming, J. W., Ingalls, A. E., Bowman, J. S., & Young, J. N. (2023). Microbial metabolomic responses to changes in temperature and salinity along the western Antarctic Peninsula. *The ISME Journal 2023 17:11*, *17*(11), 2035–2046. https://doi.org/10.1038/s41396-023-01475-0

Dithugoe, C. D., Bezuidt, O. K. I., Cavan, E. L., Froneman, W. P., Thomalla, S. J., & Makhalanyane, T. P. (2023). Bacteria and archaea regulate particulate organic matter export in suspended and sinking marine particle fractions. *MSphere*, *8*(3). https://doi.org/10.1128/MSPHERE.00420-22/SUPPL_FILE/MSPHERE.00420-22-S0004.TIF

Dutta, A., Goldman, T., Keating, J., Burke, E., Williamson, N., Dirmeier, R., & Bowman, J. S.

(2022). Machine learning predicts biogeochemistry from microbial community structure in a complex model system. *Microbiology Spectrum*, *10*(1). https://doi.org/10.1128/SPECTRUM.01909-21
Edwards, A., Cameron, K. A., Cook, J. M., Debbonaire, A. R., Furness, E., Hay, M. C., & Rassner, S. M. E. (2020). Microbial genomics amidst the arctic crisis. In *Microbial Genomics* (Vol. 6, Issue 5, pp. 1–20). Microbiology Society. https://doi.org/10.1099/mgen.0.000375

Erazo, N. G., & Bowman, J. S. (2021). Sensitivity of the mangrove-estuarine microbial community to aquaculture effluent. *IScience*, *24*(3), 102204. https://doi.org/10.1016/j.isci.2021.102204

Erazo, N. G., Dutta, A., & Bowman, J. S. (2021). From microbial community structure to metabolic inference using paprica. *STAR Protocols*, *2*(4), 101005. https://doi.org/10.1016/J.XPRO.2021.101005

Eveleth, R., Timmermans, M.-L., & Cassar, N. (2014). Physical and biological controls on oxygen saturation variability in the upper Arctic Ocean. *Journal of Geophysical Research: Oceans*, *119*(11), 7420–7432. https://doi.org/10.1002/2014JC009816

Ewert, M., Carpenter, S. D., Colangelo-Lillis, J., & Deming, J. W. (2013). Bacterial and extracellular polysaccharide content of brine-wetted snow over Arctic winter first-year sea ice. *Journal of Geophysical Research: Oceans*, *118*(2), 726–735. https://doi.org/10.1002/jgrc.20055

Ewert, M., & Deming, J. W. (2014). Bacterial responses to fluctuations and extremes in temperature and brine salinity at the surface of Arctic winter sea ice. *FEMS Microbiology Ecology*, *89*(2), 476–489. https://doi.org/10.1111/1574-6941.12363

Fernández-Gómez, B., Richter, M., Schüler, M., Pinhassi, J., Acinas, S. G., González, J. M., & Pedrós-Alió, C. (2013). Ecology of marine Bacteroidetes: a comparative genomics approach. *The ISME Journal*, *7*(5), 1026. https://doi.org/10.1038/ISMEJ.2012.169

Fernández-Méndez, M., Wenzhöfer, F., Peeken, I., Sørensen, H. L., Glud, R. N., & Boetius, A. (2014). Composition, Buoyancy Regulation and Fate of Ice Algal Aggregates in the Central Arctic Ocean. *PLOS ONE*, *9*(9), e107452. https://doi.org/10.1371/JOURNAL.PONE.0107452

Finley, A., Banerjee, S., & Hjelle, Ø. (2022). MBA: Multilevel B-Spline Approximation. *R Package Version 0.1-0*. https://CRAN.R-project.org/package=MBA

Firth, E., Carpenter, S. D., Sørensen, H. L., Collins, R. E., & Deming, J. W. (2016). Bacterial use of choline to tolerate salinity shifts in sea-ice brines. *Elementa*, *2016*. https://doi.org/10.12952/JOURNAL.ELEMENTA.000120/427625/156-1627-1-CE.PDF

Fong, A. A., Hoppe, C. J., Gradinger, R., & et al. (unpubl.). Overview of the MOSAiC expedition: Ecosystem. *Elementa: Science of the Anthropocene*. Submitted.

Freyria, N. J., Kuo, A., Chovatia, M., Johnson, J., Lipzen, A., Barry, K. W., Grigoriev, I. V., &

Lovejoy, C. (2022). Salinity tolerance mechanisms of an Arctic Pelagophyte using comparative transcriptomic and gene expression analysis. *Communications Biology 2022 5:1*, *5*(1), 1–13. https://doi.org/10.1038/s42003-022-03461-2
Garcia, H. E., & Gordon, L. I. (1992). Oxygen solubility in seawater: Better fitting equations. In *Limnology and Oceanography* (Vol. 37, Issue 6, pp. 1307–1312). John Wiley & Sons, Ltd. https://doi.org/10.4319/lo.1992.37.6.1307

Garrison, D. L., & Buck, K. R. (1986). Organism losses during ice melting: A serious bias in sea ice community studies. *Polar Biology*, *6*(4), 237–239. https://doi.org/10.1007/BF00443401

Giovannoni, S. J., Tripp, H. J., Givan, S., Podar, M., Vergin, K. L., Baptista, D., Bibbs, L., Eads, J., Richardson, T. H., Noordewier, M., Rappé, M. S., Short, J. M., Carrington, J. C., & Mathur, E. J. (2005). Genetics: Genome streamlining in a cosmopolitan oceanic bacterium. *Science*, *309*(5738), 1242–1245. https://doi.org/10.1126/SCIENCE.1114057

Gradinger, R., & Lenz, J. (1995). Seasonal occurrence of picocyanobacteria in the Greenland Sea and central Arctic Ocean. *Polar Biology*, *15*(6), 447–452. https://doi.org/10.1007/BF00239722

Grossi, S. M., & Sullivan, C. W. (2004). Sea ice microbial communities. V. The Vertical zonation of diatoms in an Antarctic fast ice community. *Journal of Phycology*, *21*(3), 401–409. https://doi.org/10.1111/j.0022-3646.1985.00401.x

Guidi, L., Chaffron, S., Bittner, L., Eveillard, D., Larhlimi, A., Roux, S., Darzi, Y., Audic, S., Berline, L., Brum, J. R., Coelho, L. P., Espinoza, J. C. I., Malviya, S., Sunagawa, S., Dimier, C., Kandels-Lewis, S., Picheral, M., Poulain, J., Searson, S., … Gorsky, G. (2016). Plankton networks driving carbon export in the oligotrophic ocean. *Nature 2016 532:7600*, *532*(7600), 465–470. https://doi.org/10.1038/nature16942

Guillou, L., Bachar, D., Audic, S., Bass, D., Berney, C., Bittner, L., Boutte, C., Burgaud, G., De Vargas, C., Decelle, J., Del Campo, J., Dolan, J. R., Dunthorn, M., Edvardsen, B., Holzmann, M., Kooistra, W. H. C. F., Lara, E., Le Bescot, N., Logares, R., … Christen, R. (2013). The Protist Ribosomal Reference database (PR2): A catalog of unicellular eukaryote Small Sub-Unit rRNA sequences with curated taxonomy. *Nucleic Acids Research*, *41*(D1), D597–D604. https://doi.org/10.1093/nar/gks1160

Haas, C., Hoppmann, M., Tippenhauer, S., & Rohardt, G. (2021). Haas, C et al. (2021): *Continuous thermosalinograph oceanography along RV Polarstern cruise track PS122/2.* PANGAEA. https://doi.org/10.1594/PANGAEA.930024

Haft, D. H., DiCuccio, M., Badretdin, A., Brover, V., Chetvernin, V., O'Neill, K., Li, W., Chitsaz, F., Derbyshire, M. K., Gonzales, N. R., Gwadz, M., Lu, F., Marchler, G. H., Song, J. S., Thanki, N., Yamashita, R. A., Zheng, C., Thibaud-Nissen, F., Geer, L. Y., … Pruitt, K. D. (2018). RefSeq: An update on prokaryotic genome annotation and curation. *Nucleic Acids Research*, *46*(D1), D851–D860. https://doi.org/10.1093/nar/gkx1068

Hahnke, R. L., Meier-Kolthoff, J. P., García-López, M., Mukherjee, S., Huntemann, M.,

Ivanova, N. N., Woyke, T., Kyrpides, N. C., Klenk, H. P., & Göker, M. (2016). Genome-based taxonomic classification of Bacteroidetes. *Frontiers in Microbiology*, *7*(12), 2003. https://doi.org/10.3389/fmicb.2016.02003

Hamme, R. C., & Emerson, S. R. (2004). The solubility of neon, nitrogen and argon in distilled water and seawater. *Deep-Sea Research Part I: Oceanographic Research Papers*, *51*(11), 1517–1528. https://doi.org/10.1016/j.dsr.2004.06.009

Holm, S. (1979). A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics*, *6*(2), 65–70. https://www.jstor.org/stable/4615733

Hoppe, C. J. M., Creamean, J., Bowman, J. S., Heitmann, L., Chamberlain, E., Brenneis, T., Terbrüggen, A., & Vortkamp, M. (2023). *Year-round discrete underway water column Chlorophyll a concentrations from the central Arctic.* PANGAEA. https://doi.pangaea.de/10.1594/PANGAEA.962597

Hoppe, C. J. M., Fuchs, N., Notz, D., Anderson, P., Assmy, P., Berge, J., Bratbak, G., Guillou, G., Kraberg, A., Larsen, A., Lebreton, B., Leu, E., Lucassen, M., Müller, O., Oziel, L., Rost, B., Schartmüller, B., Torstensson, A., & Wloka, J. (unpubl.). A new lower limit for photosynthetic light requirements detected in Arctic microalgae. *Nature Ecology & Evolution*. Submitted.

Hoppe, C. J. M., Heitmann, L., Brenneis, T., Terbrüggen, A., Vortkamp, M., Chamberlain, E., Droste, E. S., Eggers, L., Fong, A. A., Gardner, J., Ren, J., Snoeijs-Leijonmalm, P., Svenson, A., & Zhuang, Y. (2023). *Water column Chlorophyll a concentrations during the MOSAiC expedition (PS122) in the Central Arctic Ocean 2019-2020*. PANGAEA. https://doi.org/10.1594/PANGAEA.963277

Hoppmann, M., Richter, M. E., Smith, I. J., Jendersie, S., Langhorne, P. J., Thomas, D. N., & Dieckmann, G. S. (2020). Platelet ice, the Southern Ocean's hidden ice: a review. *Annals of Glaciology*, *61*(83), 341–368. https://doi.org/10.1017/AOG.2020.54

Itkin, P., Spreen, G., Cheng, B., Doble, M., Girard-Ardhuin, F., Haapala, J., Hughes, N., Kaleschke, L., Nicolaus, M., & Wilkinson, J. (2017). Thin ice and storms: Sea ice deformation from buoy arrays deployed during N-ICE2015. *Journal of Geophysical Research: Oceans*, *122*(6), 4661–4674. https://doi.org/10.1002/2016JC012403

Izett, R. W., Hamme, R. C., McNeil, C., Manning, C. C. M., Bourbonnais, A., & Tortell, P. D. (2021). ΔO2/N2′ as a new tracer of marine net community production: Application and evaluation in the subarctic northeast Pacific and Canadian Arctic Ocean. *Frontiers in Marine Science*, *8*, 958. https://doi.org/10.3389/FMARS.2021.718625/BIBTEX

Jakobsson, M., Mayer, L. A., Bringensparr, C., Castro, C. F., Mohammad, R., Johnson, P., Ketter, T., Accettella, D., Amblas, D., An, L., Arndt, J. E., Canals, M., Casamor, J. L., Chauché, N., Coakley, B., Danielson, S., Demarte, M., Dickson, M. L., Dorschel, B., … Zinglersen, K. B. (2020). The International Bathymetric Chart of the Arctic Ocean Version 4.0. *Scientific Data 2020 7:1*, *7*(1), 1–14. https://doi.org/10.1038/s41597-020-0520-9

Journal, B. P., & Rao, D. V. S. (1988). Species specific primary production measurements of Arctic phytoplankton. *British Phycological Journal*, *23*(3), 273–282. https://doi.org/10.1080/00071618800650311

Junge, K., Eicken, H., & Deming, J. W. (2004). Bacterial activity at -2 to -20°C in Arctic wintertime sea ice. *Applied and Environmental Microbiology*, *70*(1), 550–557. https://doi.org/10.1128/AEM.70.1.550-557.2004

Juranek, L., Takahashi, T., Mathis, J., & Pickart, R. (2019). Significant biologically mediated CO_2 uptake in the pacific arctic during the late open water season. *Journal of Geophysical Research: Oceans*, *124*(2), 821–843. https://doi.org/10.1029/2018JC014568

Kadko, D. (2000). Modeling the evolution of the Arctic mixed layer during the fall 1997 Surface Heat Budget of the Arctic Ocean (SHEBA) Project using measurements of 7Be. *Journal of Geophysical Research: Oceans*, *105*(C2), 3369–3378. https://doi.org/10.1029/1999JC900311

Kanzow, T., Hoppmann, M., Tippenhauer, S., & Rohardt, G. (2021*). Continuous thermosalinograph oceanography along RV POLARSTERN cruise track PS122/3.* PANGAEA. https://doi.org/10.1594/PANGAEA.930026

Karp, P. D., Midford, P. E., Billington, R., Kothari, A., Krummenacker, M., Latendresse, M., Ong, W. K., Subhraveti, P., Caspi, R., Fulcher, C., Keseler, I. M., & Paley, S. M. (2021). Pathway Tools version 23.0 update: software for pathway/genome informatics and systems biology. *Briefings in Bioinformatics*, *22*(1), 109–126. https://doi.org/10.1093/bib/bbz104

Kassambra, A. (2023). ggpubr: "ggplot2" Based Publication Ready Plots. *R Package Version 0.6.0*. https://CRAN.R-project.org/package=ggpubr

Kellogg, C. T. E., Carpenter, S. D., Renfro, A. A., Sallon, A., Michel, C., Cochran, J. K., & Deming, J. W. (2011). Evidence for microbial attenuation of particle flux in the Amundsen Gulf and Beaufort Sea: Elevated hydrolytic enzyme activity on sinking aggregates. *Polar Biology*, *34*(12), 2007–2023. https://doi.org/10.1007/S00300-011-1015-0/FIGURES/3

Kim, H. H., Laufkötter, C., Lovato, T., Doney, S. C., & Ducklow, H. W. (2023). Projected 21st-century changes in marine heterotrophic bacteria under climate change. *Frontiers in Microbiology*, *14*, 1049579. https://doi.org/10.3389/FMICB.2023.1049579/BIBTEX

Knust, R. (2017). Polar Research and Supply Vessel POLARSTERN operated by the Alfred-Wegener-Institute. *Journal of Large-Scale Research Facilities JLSRF*, *3*. https://doi.org/10.17815/JLSRF-3-163

Kolde, R. (2019). *pheatmap: Pretty Heatmaps. R package version 1.0.12*. https://CRAN.R-project.org/package=pheatmap

Krumpen, T., Birrien, F., Kauker, F., Rackow, T., Von Albedyll, L., Angelopoulos, M., Jakob Belter, H., Bessonov, V., Damm, E., Dethloff, K., Haapala, J., Haas, C., Harris, C., Hendricks, S., Hoelemann, J., Hoppmann, M., Kaleschke, L., Karcher, M., Kolabutin, N., … Watkins, D. (2020). The MOSAiC ice floe: Sediment-laden survivor from the Siberian shelf. *Cryosphere*, *14*(7), 2173–2187. https://doi.org/10.5194/TC-14-2173-2020

Krumpen, T., Von Albedyll, L., Goessling, H. F., Hendricks, S., Juhls, B., Spreen, G., Willmes, S., Belter, H. J., Dethloff, K., Haas, C., Kaleschke, L., Katlein, C., Tian-Kunze, X., Ricker, R., Rostosky, P., Rückert, J., Singha, S., & Sokolova, J. (2021). MOSAiC drift expedition from October 2019 to July 2020: Sea ice conditions from space and comparison with previous years. *Cryosphere*, *15*(8), 3897–3920. https://doi.org/10.5194/TC-15-3897-2021

Kursa, M. B., & Rudnicki, W. R. (2010). Feature selection with the Boruta Package. *Journal of Statistical Software*, *36*(11), 1–13. https://doi.org/10.18637/JSS.V036.I11

Kwok, R., Spreen, G., & Pang, S. (2013). Arctic sea ice circulation and drift speed: Decadal trends and ocean currents. *Journal of Geophysical Research: Oceans*, *118*(5), 2408–2425. https://doi.org/10.1002/JGRC.20191

Lannuzel, D., Tedesco, L., van Leeuwe, M., Campbell, K., Flores, H., Delille, B., Miller, L., Stefels, J., Assmy, P., Bowman, J., Brown, K., Castellani, G., Chierici, M., Crabeck, O., Damm, E., Else, B., Fransson, A., Fripiat, F., Geilfus, N. X., … Wongpan, P. (2020). The future of Arctic sea-ice biogeochemistry and ice-associated ecosystems. *Nature Climate Change 2020 10:11*, *10*(11), 983–992. https://doi.org/10.1038/s41558-020-00940-4

Larsen, J. B., Larsen, A., Thyrhaug, R., Bratbak, G., & Sandaa, R. A. (2008). Response of marine viral populations to a nutrient induced phytoplankton bloom at different pCO2 levels. *Biogeosciences*, *5*(2), 523–533. https://doi.org/10.5194/BG-5-523-2008

Li, S.-H., Kang, I., & Cho, J.-C. (2023). Metabolic Versatility of the Family Halieaceae Revealed by the Genomics of Novel Cultured Isolates . *Microbiology Spectrum*, *11*(2). https://doi.org/10.1128/SPECTRUM.03879-22

Li, W. K. W., McLaughlin, F. A., Lovejoy, C., & Carmack, E. C. (2009). Smallest algae thrive as the arctic ocean freshens. In *Science* (Vol. 326, Issue 5952, p. 539). Science. https://doi.org/10.1126/science.1179798

Liaw, A., & Wiener, M. (2002). Classification and Regression by randomForest. *R News*, *2*(3), 18–22. https://CRAN.R-project.org/doc/Rnews/

Lin, Y., Cassar, N., Marchetti, A., Moreno, C., Ducklow, H., & Li, Z. (2017). Specific eukaryotic plankton are good predictors of net community production in the Western Antarctic Peninsula. *Scientific Reports 2017 7:1*, *7*(1), 1–11. https://doi.org/10.1038/s41598-017-14109-1

Lin, Y., Moreno, C., Marchetti, A., Ducklow, H., Schofield, O., Delage, E., Meredith, M., Li, Z., Eveillard, D., Chaffron, S., & Cassar, N. (2021). Decline in plankton diversity and carbon flux with reduced sea ice extent along the Western Antarctic Peninsula. *Nature Communications 2021 12:1*, *12*(1), 1–9. https://doi.org/10.1038/s41467-021-25235-w

Loose, B., McGillis, W. R., Perovich, D., Zappa, C. J., & Schlosser, P. (2014). A parameter model of gas exchange for the seasonal sea ice zone. *Ocean Science*, *10*(1), 17–28. https://doi.org/10.5194/OS-10-17-2014

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*(12), 550. https://doi.org/10.1186/s13059-014-0550-8

MacGilchrist, G. A., Naveira Garabato, A. C., Tsubouchi, T., Bacon, S., Torres-Valdés, S., & Azetsu-Scott, K. (2014). The Arctic Ocean carbon sink. *Deep Sea Research Part I: Oceanographic Research Papers*, *86*, 39–55. https://doi.org/10.1016/J.DSR.2014.01.002

Marie, D., Partensky, F., Jacquet, S., & Vaulot, D. (1997). Enumeration and cell cycle analysis of natural populations of marine picoplankton by flow cytometry using the nucleic acid stain SYBR Green I. *Applied and Environmental Microbiology*, *63*(1), 186–193. https://doi.org/10.1128/AEM.63.1.186-193.1997

Martens, J., Wild, B., Semiletov, I., Dudarev, O. V., & Gustafsson, Ö. (2022). Circum-Arctic release of terrestrial carbon varies between regions and sources. *Nature Communications 2022 13:1*, *13*(1), 1–10. https://doi.org/10.1038/s41467-022-33541-0

McMurdie, P. J., & Holmes, S. (2013). phyloseq: An R Package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE*, *8*(4), e61217. https://doi.org/10.1371/journal.pone.0061217

Meredith, M., Sommerkorn, M., Cassotta, S., Derksen, C., Ekaykin, A., Hollowed, A., Kofinas, G., Mackintosh, A., Melbourne-Thomas, J., Muelber, M., Ottersen, G., Pritchard, H., & Schuur, E. (2019). Polar Regions. In H.-O. Pörtner, D. Roberts, V. Masson-Delmotte, P. Zhai, M. Tignor, E. Poloczanska, K. Mintenback, A. Alegría, M. Nicolai, A. Okem, J. Petzold, B. Rama, & N. Weyer (Eds.), *IPCC Special Report on the Ocean and Cryosphere in a Changing Climate* (pp. 203–320). Cambridge University Press. https://doi.org/0.1017/9781009157964.005

Mikkelsen, D. M., & Witkowski, A. (2010). Melting sea ice for taxonomic analysis: a comparison of four melting procedures. *Polar Research*, *29*(3), 451–454. https://doi.org/10.1111/j.1751-8369.2010.00162.x

Miller, L. A., Fripiat, F., Else, B. G. T., Bowman, J. S., Brown, K. A., Collins, R. E., Ewert, M., Fransson, A., Gosselin, M., Lannuzel, D., Meiners, K. M., Michel, C., Nishioka, J., Nomura, D.,

Papadimitriou, S., Russell, L. M., Sørensen, L. L., Thomas, D. N., Tison, J. L., … Zhou, J. (2015). Methods for biogeochemical studies of sea ice: The state of the art, caveats, and recommendations. *Elementa*, *3*, 38. https://doi.org/10.12952/journal.elementa.000038

Mock, T., Boulton, W., Balmonte, J.-P., Barry, K., Bertilsson, S., Bowman, J., Buck, M., Bratbak, G., Chamberlain, E. J., Cunliffe, M., Creamean, J., Ebenhöh, O., Eggers, S. L., Fong, A. A., Gardner, J., Gradinger, R., Granskog, M. A., Havermans, C., Hill, T., … Metfies, K. (2022). Multiomics in the central Arctic Ocean for benchmarking biodiversity change. *PLOS Biology*, *20*(10), e3001835. https://doi.org/10.1371/JOURNAL.PBIO.3001835

Morawetz, K., Thoms, S., & Kutschan, B. (2017). Formation of brine channels in sea ice. *The European Physical Journal E*, *40*(25). https://doi.org/10.1140/epje/i2017-11512-x

Muilwijk, M., Koenig, Z., Fer, I., Schulz, K., Tippenhauer, S., Schuffenhauer, I., Karam, S., Hoppmann, M., & Allerholt, J. (2022). *Videos and extracted frames from Blueye ROV camera during MOSAiC Leg 4 (PS122/4) and Leg 5 (PS122/5). PANGAEA.* http://doi.org/10.1584/PANGAEA.941553

Mundy, C. J., Gosselin, M., Ehn, J. K., Belzile, C., Poulin, M., Alou, E., Roy, S., Hop, H., Lessard, S., Papakyriakou, T. N., Barber, D. G., & Stewart, J. (2011). Characteristics of two distinct high-light acclimated algal communities during advanced stages of sea ice melt. *Polar Biology*, *34*(12), 1869–1886. https://doi.org/10.1007/S00300-011-0998-X/TABLES/1

Nansen, F. (1902). The Norwegian North Polar Epedition, 1893-1898: Scientific Results. *Fridtjof Nansen Fund for the Advancement of Science*, *3*.

Nawrocki, E. P., & Eddy, S. R. (2013). Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics*, *29*(22), 2933–2935. https://doi.org/10.1093/bioinformatics/btt509

Nicolaus, M., Perovich, D. K., Spreen, G., Granskog, M. A., von Albedyll, L., Angelopoulos, M., Anhaus, P., Arndt, S., Jakob Belter, H., Bessonov, V., Birnbaum, G., Brauchle, J., Calmer, R., Cardellach, E., Cheng, B., Clemens-Sewall, D., Dadic, R., Damm, E., de Boer, G., … Wendisch, M. (2022). Overview of the MOSAiC expedition: Snow and sea ice. *Elementa*, *10*(1). https://doi.org/10.1525/ELEMENTA.2021.000046/119791

Niedrauer, T. M., & Martin, S. (1979). An experimental study of brine drainage and convection in Young Sea ice. *Journal of Geophysical Research*, *84*(C3), 1176. https://doi.org/10.1029/JC084iC03p01176

Nixdorf, U., Dethloff, K., Rex, M., Shupe, M., Sommerfeld, A., Perovich, D. K., Nicolaus, M., Heuzé, C., Rabe, B., Loose, B., Damm, E., Gradinger, R., Fong, A., Maslowski, W., Rinke, A., Kwok, R., Spreen, G., Wendisch, M., Herber, A., … Boetius, A. (2021). *MOSAiC Extended Acknowledgement*. https://doi.org/10.5281/ZENODO.5541624

Nomura, D., Kawaguchi, Y., Webb, A. L., Li, Y., Dall'osto, M., Schmidt, K., Droste, E. S., Chamberlain, E. J., Kolabutin, N., Shimanchuk, E., Hoppmann, M., Gallagher, M. R., Meyer, H.,

Mellat, M., Bauch, D., Gabarró, C., Smith, M. M., Inoue, J., Damm, E., & Delille, B. (2023). Meltwater layer dynamics in a central Arctic lead: Effects of lead width, re-freezing, and mixing during late summer. *Elementa*, *11*(1), 23. https://doi.org/10.1525/ELEMENTA.2022.00102/196161

Norris, N., Levine, N. M., Fernandez, V. I., & Stocker, R. (2021). Mechanistic model of nutrient uptake explains dichotomy between marine oligotrophic and copiotrophic bacteria. *PLOS Computational Biology*, *17*(5), e1009023. https://doi.org/10.1371/JOURNAL.PCBI.1009023

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P., O'Hara, R. B., Simpson, G., Solymos, P., Stevens, M. H. H., & Wagner, H. (2015). *Vegan: community ecology package. R package vegan, vers. 2.2-1*.

Olsen, L. M., Laney, S. R., Duarte, P., Kauko, H. M., Fernández-Méndez, M., Mundy, C. J., Rösel, A., Meyer, A., Itkin, P., Cohen, L., Peeken, I., Tatarek, A., Różańska-Pluta, M., Wiktor, J., Taskjelle, T., Pavlov, A. K., Hudson, S. R., Granskog, M. A., Hop, H., & Assmy, P. (2017). The seeding of ice algal blooms in Arctic pack ice: The multiyear ice seed repository hypothesis. *Journal of Geophysical Research: Biogeosciences*, *122*(7), 1529–1548. https://doi.org/10.1002/2016JG003668

Paulsen, M. L., Doré, H., Garczarek, L., Seuthe, L., Müller, O., Sandaa, R. A., Bratbak, G., & Larsen, A. (2016). Synechococcus in the Atlantic gateway to the Arctic Ocean. *Frontiers in Marine Science*, *3*(OCT), 216694. https://doi.org/10.3389/FMARS.2016.00191/BIBTEX

Paulsen, M. L., Riisgaard, K., Thingstad, T. F., John, M., & Nielsen, T. G. (2015). Winter-spring transition in the subarctic Atlantic: microbial response to deep mixing and pre-bloom production. *Aquatic Microbial Ecology*, *76*(1), 49–69. https://doi.org/10.3354/AME01767

Pedrós-Alió, C., Puente-Sánchez, F., Macías, L. A., Redondo, Á., Campbell, K. L., Royo-Llonch, M., Balagué, V., Sánchez, P., Mundy, C.-J., & Tamames, J. (2023). Bacterioplankton succession during the Arctic Ocean spring bloom. *Revista de Derecho Politico*, *117*, 153–196. https://doi.org/10.13039/501100011033

Perovich, D., Smith, M., Light, B., & Webster, M. (2021). Meltwater sources and sinks for multiyear Arctic sea ice in summer. *Cryosphere*, *15*(9), 4517–4525. https://doi.org/10.5194/TC-15-4517-2021

Piontek, J., Galgani, L., Nöthig, E. M., Peeken, I., & Engel, A. (2021). Organic matter composition and heterotrophic bacterial activity at declining summer sea ice in the central Arctic Ocean. *Limnology and Oceanography*, *66*(S1), S343–S362. https://doi.org/10.1002/LNO.11639

Posit team. (2022). *RStudio: Integrated Development Environment for R*. Posit Software, PBC. http://www.posit.co/

Priest, T., von Appen, W. J., Oldenburg, E., Popa, O., Torres-Valdés, S., Bienhold, C., Metfies, K., Boulton, W., Mock, T., Fuchs, B. M., Amann, R., Boetius, A., & Wietz, M. (2023). Atlantic

water influx and sea-ice cover drive taxonomic and functional shifts in Arctic marine bacterial communities. *The ISME Journal 2023 17:10*, *17*(10), 1612–1625. https://doi.org/10.1038/s41396-023-01461-6
Qu, K., Guo, F., Liu, X., Lin, Y., & Zou, Q. (2019). Application of machine learning in microbiology. *Frontiers in Microbiology*, *10*(APR), 451710. https://doi.org/10.3389/FMICB.2019.00827/BIBTEX

R Core Team. (2023). *R: A language and environment for statistical computing* (4.3.1). R Foundation for Statistical Computing. https://www.R-project.org/

Rabe, B., Heuzé, C., Regnery, J., Aksenov, Y., Allerholt, J., Athanase, M., Bai, Y., Basque, C., Bauch, D., Baumann, T. M., Chen, D., Cole, S. T., Craw, L., Davies, A., Damm, E., Dethloff, K., Divine, D. V., Doglioni, F., Ebert, F., … Zhu, J. (2022). Overview of the MOSAiC expedition: Physical oceanography. *Elementa: Science of the Anthropocene*, *10*(1), 1. https://doi.org/10.1525/ELEMENTA.2021.00062

Ralph, P. J., Ryan, K. G., Martin, A., & Fenton, G. (2007). Melting out of sea ice causes greater photosynthetic stress in algae than freezing in1. *Journal of Phycology*, *43*(5), 948–956. https://doi.org/10.1111/J.1529-8817.2007.00382.X

Ramoneda, J., Stallard-Olivera, E., Hoffert, M., Winfrey, C. C., Stadler, M., Niño-García, J. P., & Fierer, N. (2023). Building a genome-based understanding of bacterial pH preferences. *Science Advances*, *9*(17). https://doi.org/10.1126/SCIADV.ADF8998

Rantanen, M., Karpechko, A. Y., Lipponen, A., Nordling, K., Hyvärinen, O., Ruosteenoja, K., Vihma, T., & Laaksonen, A. (2022). The Arctic has warmed nearly four times faster than the globe since 1979. *Communications Earth & Environment 2022 3:1*, *3*(1), 1–10. https://doi.org/10.1038/s43247-022-00498-3

Rapp, J. Z., Fernández-Méndez, M., Bienhold, C., & Boetius, A. (2018). Effects of ice-algal aggregate export on the connectivity of bacterial communities in the central Arctic Ocean. *Frontiers in Microbiology*, *9*(5). https://doi.org/10.3389/FMICB.2018.01035/FULL

Rex, M., Hoppman, M., Tippenhauer, S., & Rohardt, G. (2021). *Continuous thermosalinograph oceanography along RV POLARSTERN cruise track PS122/5*. PANGAEA. https://doi.org/10.1594/PANGAEA.930028

Rex, M., Hoppmann, M., Tippenhauer, S., & Rohardt, G. (2021a). *Continuous thermosalinograph oceanography along RV POLARSTERN cruise track PS122/1*. PANGAEA. https://doi.pangaea.de/10.1594/PANGAEA.930023

Rex, M., Hoppmann, M., Tippenhauer, S., & Rohardt, G. (2021b). *Continuous thermosalinograph oceanography along RV POLARSTERN cruise track PS122/4*. PANGAEA. https://doi.org/10.1594/PANGAEA.930027

Richter-Menge, J. A., Perovich, D. K., & Scott Pegau, W. (2001). Summer ice dynamics during

SHEBA and its effect on the ocean heat content. *Annals of Glaciology*, *33*, 201–206. https://doi.org/10.3189/172756401781818176

Rinke, A., Cassano, J. J., Cassano, E. N., Jaiser, R., & Handorf, D. (2021). Meteorological conditions during the MOSAiC expedition: Normal or anomalous? *Elementa*, *9*(1). https://doi.org/10.1525/ELEMENTA.2021.00023/118092

Rivkin, R. B., & Legendre, L. (2001). Biogenic carbon cycling in the upper ocean: Effects of microbial respiration. *Science*, *291*(5512), 2398–2400. https://doi.org/10.1126/SCIENCE.291.5512.2398

Roukaerts, A., Nomura, D., Deman, F., Hattori, H., Dehairs, F., & Fripiat, F. (2019). The effect of melting treatments on the assessment of biomass and nutrients in sea ice (Saroma-ko lagoon, Hokkaido, Japan). *Polar Biology*, *42*(2), 347–356. https://doi.org/10.1007/s00300-018-2426-y

Royo-Llonch, M., Sánchez, P., Ruiz-González, C., Salazar, G., Pedrós-Alió, C., Sebastián, M., Labadie, K., Paoli, L., M. Ibarbalz, F., Zinger, L., Churcheward, B., Babin, M., Bork, P., Boss, E., Cochrane, G., de Vargas, C., Gorsky, G., Grimsley, N., Guidi, L., … Acinas, S. G. (2021). Compendium of 530 metagenome-assembled bacterial and archaeal genomes from the polar Arctic Ocean. *Nature Microbiology 2021 6:12*, *6*(12), 1561–1574. https://doi.org/10.1038/s41564-021-00979-9

Rysgaard, S., Glud, R. N., Sejr, M. K., Bendtsen, J., & Christensen, P. B. (2007). Inorganic carbon transport during sea ice growth and decay: A carbon pump in polar seas. *Journal of Geophysical Research: Oceans*, *112*(C3), 3016. https://doi.org/10.1029/2006JC003572

Salganik, E., Katlein, C., Lange, B. A., Matero, I., Lei, R., Fong, A. A., Fons, S. W., Divine, D., Oggier, M., Castellani, G., Bozzato, D., Chamberlain, E. J., Hoppe, C. J. M., Müller, O., Gardner, J., Rinke, A., Pereira, P. S., Ulfsbo, A., Marsay, C., Webster, M., Maus, S., Høyland, K.V., & Granskog, M.A. (2023). Temporal evolution of under-ice meltwater layers and false bottoms and their impact on summer Arctic sea ice mass balance. *Elementa*, *11*(1). https://doi.org/10.1525/ELEMENTA.2022.00035/195810

Schanke, N. L., Bolinesi, F., Mangoni, O., Katlein, C., Anhaus, P., Hoppmann, M., Lee, P. A., & DiTullio, G. R. (2021). Biogeochemical and ecological variability during the late summer–early autumn transition at an ice-floe drift station in the Central Arctic Ocean. *Limnology and Oceanography*, *66*(S1), S363–S382. https://doi.org/10.1002/LNO.11676

Schlitzer, & Reiner. (2022). *Ocean Data View*.

Schulz, K., Koenig, Z., & Muilwijk, M. (2023). The Eurasian Arctic Ocean along the MOSAiC Drift (2019-2020): Core hydrographic parameters. *Arctic Data Center*. https://doi.org/10.18739/A21J9790B

Schulz, K., Koenig, Z., Muilwijk, M., Bauch, D., Hoppe, C. J. M., Droste, E., Hoppmann, M., Chamberlain, E. J., Laukert, G., Stanton, T., Zurita, A. Q., Fer, I., Heuzé, C., Karam, S.,

Mieruch-Schnuelle, Baumann, T., Vrendenborg, M., Tippenhauer, S., & Granskog, M. A. (unpubl.). The Eurasian Arctic Ocean along the MOSAiC drift (2019-2020): An interdisciplinary perspective on properties and processes. *Elementa: Science of the Anthropocene*. Submitted. https://doi.org/https://doi.org/10.31223/X5TT2W

Seguro, I., Marca, A. D., Shutler, J., & Kaiser, J. (2023). Different flavours of oxygen help quantify seasonal variations of the biological carbon pump in the Celtic Sea. *Frontiers in Marine Science*, *10*, 698. https://doi.org/10.3389/FMARS.2023.1037470

Shupe, M. D., & Rex, M. (2022). SIDEBAR: A year in the changing arctic sea ice. *Oceanography*, *35*(3–4), 224–225. https://doi.org/10.5670/OCEANOG.2022.126

Shupe, M. D., Rex, M., Blomquist, B., G. Persson, P. O., Schmale, J., Uttal, T., Althausen, D., Angot, H., Archer, S., Bariteau, L., Beck, I., Bilberry, J., Bucci, S., Buck, C., Boyer, M., Brasseur, Z., Brooks, I. M., Calmer, R., Cassano, J., Castro, V., Chu, D., Costa, D., Cox, C.J., Creamean, J., Crewell, S., Dahlke, S., Damm, E., de Boer, G., Deckelmann, H., Dethloff, K., Dütsch, M. ... Yue, F. (2022). Overview of the MOSAiC expedition: Atmosphere. *Elementa: Science of the Anthropocene*, *10*(1). https://doi.org/10.1525/ELEMENTA.2021.00060

Shupe, M. D., Rex, M., Dethloff, K., Damm, E., Fong, A. A., Gradinger, R., Heuzé, C., Loose, B., Makarov, A., Maslowski, W., Nicolaus, M., Perovich, D., Rabe, B., Rinke, A., Sokolov, V., & Sommerfeld, A. (2020). The MOSAiC Expedition: A Year Drifting with the Arctic Sea Ice. *Arctic Report Card 2020.* https://doi.org/10.25923/9G3V-XH92

Simon, M., Billerbeck, S., Kessler, D., Selje, N., & Schlingloff, A. (2012). Bacterioplankton communities in the Southern Ocean: composition and growth response to various substrate regimes. *Aquatic Microbial Ecology*, *68*(1), 13–28. https://doi.org/10.3354/AME01597

Smith, M., Bozzato, D., Chamberlain, E. J., Dietrich, U., Droste, E., Fong, A. A., Light, B., Linhardt, F., Matero, I., Perovich, D., Raphael, I., Regnery, J., Salganik, E., Tavri, A., von Albedyll, L., & Webster, M. (2021). *Near-surface ocean temperature & salinity measurements (using YSI and Castaway) during the summer component of the Multidisciplinary drifting Observatory for the Study of Arctic Climate (MOSAiC) campaign in the Central Arctic Ocean, July – September 2020.* Arctic Data Center. http://doi.org/10.18739/A2TT4FV1G

Smith, M. M., Angot, H., Chamberlain, E. J., Droste, E. S., Karam, S., Muilwijk, M., Webb, A. L., Archer, S. D., Beck, I., Blomquist, B. W., Bowman, J., Boyer, M., Bozzato, D., Chierici, M., Creamean, J., D'Angelo, A., Delille, B., Fer, I., Fong, A. A., Fransson, A., Fuchs, N., Gardner, J., Granskog, M., Hoppe, C., Hoppema, M., Hoppmann, M., Mock, T., Muller, S., Müller, O., Nicolaus, M., Nomura, D., Petäjä, T., Salganik, E., Schmale, J., Schmidt, K., Schulz, K., Shupe, M., Stefels, J., Thielke, L., Tippenhauer, S., Ulfsbo, K.A., van Leeuwe, M., Webster, M., Yoshimura, M., & Xhan, L. (2023). Thin and transient meltwater layers and false bottoms in the Arctic sea ice pack—Recent insights on these historically overlooked features. *Elementa: Science of the Anthropocene*, *11*(1). https://doi.org/10.1525/ELEMENTA.2023.00025

Smith, M. M., von Albedyll, L., Raphael, I. A., Lange, B. A., Matero, I., Salganik, E., Webster, M. A., Granskog, M. A., & Light, B. (2022). Quantifying false bottoms and under-ice meltwater layers beneath Arctic summer sea ice with fine-scale observations. *In Preparation*.

Stoecker, D. K., Buck, K. R., & Putt, M. (1992). Changes in the sea-ice brine community during the spring-summer transition, McMurdo Sound, Antarctica. I. Photosynthetic protists. *Marine Ecology Progress Series*, *84*, 265–278.

Stroeve, J., & Notz, D. (2018). Changing state of Arctic sea ice across all seasons. *Environmental Research Letters*, *13*(10), 103001. https://doi.org/10.1088/1748-9326/AADE56

Syvertsen, E. E., & Syoertsen, E. E. (1991). Ice algae in the Barents Sea: types of assemblages, origin, fate and role in the ice-edge phytoplankton bloom. *Polar Research*, *10*(1), 277–288. https://doi.org/10.1111/J.1751-8369.1991.TB00653.X

Tamelander, T., Reigstad, M., Hop, H., & Ratkova, T. (2009). Ice algal assemblages and vertical export of organic matter from sea ice in the Barents Sea and Nansen Basin (Arctic Ocean). *Polar Biology 2009 32:9*, *32*(9), 1261–1273. https://doi.org/10.1007/S00300-009-0622-5

Thompson, J., Johansen, R., Dunbar, J., & Munsky, B. (2019). Machine learning to predict microbial community functions: An analysis of dissolved organic carbon from litter decomposition. *PLoS ONE*, *14*(7). https://doi.org/10.1371/JOURNAL.PONE.0215502

Tippenhauer, S., Vredenborg, M., Heuze, C., Heuzé, C., Ulfsbo, A., Rabe, B., Granskog, M. A., Allerholt, J., Balmonte, J. P., Campbell, R. G., Castellani, G., Chamberlain, E., Creamean, J., D'Angelo, A., Dietrich, U., Droste, E. S., Eggers, L., Fang, Y.-C., Fong, A. A., Gardner, J., Graupner, R., Grosse, J., He, H., Hildebrandt, N., Hoppe, C.J.M., Hoppmann, M., Kanzow, T., Karam, S., Koenig, Z., Kong, B., Kuhlmey, D., Kuznetsov, I., Lan, M., Liu, H., Mallet, M., Mohrholz, V., Muilwijk, M., Müller, O., Olsen, L.M., Rember, R., Ren, J., Sakinan, S. Schaffer, J., Schmidt, K., Schuffenhauer, I., Schulz, K., Shoemaker, K., Spahic, S., Sukhikh, N., Svenson, A., Torres-Valdés, S., Torstensson, A., Wischnewski, L., & Zhuang, Y. (2023). *Physical oceanography water bottle samples based on Ocean City CTD during POLARSTERN cruise PS122*. PANGAEA. https://doi.org/https://doi.org/10.1594/PANGAEA.959966

Tippenhauer, S., Vredenborg, M., Heuzé, C., Ulfsbo, A., Rabe, B., Granskog, M. A., Allerholt, J., Balmonte, J. P., Campbell, R. G., Castellani, G., Chamberlain, E., Creamean, J., DAngelo, A., Dietrich, U., Droste, E. S., Eggers, L., Fang, Y.-C., Fong, A. A., Gardner, J., Graupner, R., Grosse, J., He, H., Hildebrandt, N., Hoppe, C.J.M., Hoppmann, M., Kanzow, T., Karam, S., Koenig, Zoé; Kong, Bin; Kuhlmey, David; Kuznetsov, Ivan; Lan, Musheng; Liu, Hailong; Mallet, M., Mohrholz, V., Muilwijk, M., Müller, O., Olsen, L.M., Rember, R., Ren, J., Sakinan, S., Schaffer, J., Schmidt, K., Schuffenhauer, I., Schulz, K., Shoemaker, K., Spahic, S., Sukhikh, N., Svenson, A., Torres-Valdés, S., Torstensson, A., Wischnewski, L., & Zhuang, Y. (2023). *Physical oceanography water bottle samples based on ship CTD during POLARSTERN cruise PS122*. PANGAEA. https://doi.org/10.1594/PANGAEA.959965

Torstensson, A., Dinasquet, J., Chierici, M., Fransson, A., Riemann, L., & Wulff, A. (2015).

Physicochemical control of bacterial and protist community composition and diversity in Antarctic sea ice. *Environmental Microbiology*, *17*(10), 3869–3881. https://doi.org/10.1111/1462-2920.12865

Torstensson, A., Young, J. N., Carlson, L. T., Ingalls, A. E., & Deming, J. W. (2019). Use of exogenous glycine betaine and its precursor choline as osmoprotectants in Antarctic sea-ice diatoms 1. *Journal of Phycology*, *55*(3), 663–675. https://doi.org/10.1111/JPY.12839

Ulfsbo, A., Cassar, N., Korhonen, M., van Heuven, S., Hoppema, M., Kattner, G., & Anderson, L. G. (2014). Late summer net community production in the central Arctic Ocean using multiple approaches. *Global Biogeochemical Cycles*, *28*(10), 1129–1148. https://doi.org/10.1002/2014GB004833

Underwood, G. J. C., Michel, C., Meisterhans, G., Niemi, A., Belzile, C., Witt, M., Dumbrell, A. J., & Koch, B. P. (2019). Organic matter from Arctic sea-ice loss alters bacterial community structure and function. *Nature Climate Change 2019 9:2*, *9*(2), 170–176. https://doi.org/10.1038/s41558-018-0391-7

van Leeuwe, M. A., Tedesco, L., Arrigo, K. R., Assmy, P., Campbell, K., Meiners, K. M., Rintala, J.-M., Selz, V., Thomas, D. N., & Stefels, J. (2018). Microalgal community structure and primary production in Arctic and Antarctic sea ice: A synthesis. *Elementa: Science of the Anthropocene*, *6*. https://doi.org/10.1525/ELEMENTA.267

Vancoppenolle, M., Meiners, K. M., Michel, C., Bopp, L., Brabant, F., Carnat, G., Delille, B., Lannuzel, D., Madec, G., Moreau, S., Tison, J. L., & van der Merwe, P. (2013). Role of sea ice in global biogeochemical cycles: emerging views and challenges. *Quaternary Science Reviews*, *79*, 207–230. https://doi.org/10.1016/J.QUASCIREV.2013.04.011

Volk, T., & Hoffert, M. I. (1985). Ocean carbon pumps: Analysis of relative strengths and efficiencies in ocean-driven atmospheric CO_2 changes. *The Carbon Cycle and Atmospheric CO*, 99–110. https://doi.org/10.1029/GM032P0099

von Appen, W. J., Waite, A. M., Bergmann, M., Bienhold, C., Boebel, O., Bracher, A., Cisewski, B., Hagemann, J., Hoppema, M., Iversen, M. H., Konrad, C., Krumpen, T., Lochthofen, N., Metfies, K., Niehoff, B., Nöthig, E. M., Purser, A., Salter, I., Schaber, M., Scholz, D., Soltwedel, T., Torres-Vales, S., Wekerle, C., Wenzhöfer, F., Wietz, M., & Boetius, A. (2021). Sea-ice derived meltwater stratification slows the biological carbon pump: results from continuous observations. *Nature Communications 2021 12:1*, *12*(1), 1–16. https://doi.org/10.1038/s41467-021-26943-z

Walters, W., Hyde, E. R., Berg-Lyons, D., Ackermann, G., Humphrey, G., Parada, A., Gilbert, J. A., Jansson, J. K., Caporaso, J. G., Fuhrman, J. A., Apprill, A., & Knight, R. (2016). Improved Bacterial 16S rRNA Gene (V4 and V4-5) and Fungal Internal Transcribed Spacer Marker Gene Primers for Microbial Community Surveys. *MSystems*, *1*(1). https://doi.org/10.1128/msystems.00009-15

Wang, S., Lin, Y., Gifford, S., Eveleth, R., & Cassar, N. (2018). Linking patterns of net

community production and marine microbial community structure in the western North Atlantic. *The ISME Journal 2018 12:11*, *12*(11), 2582–2595. https://doi.org/10.1038/s41396-018-0163-4

Webb, S. J., Rabsatt, T., Erazo, N., & Bowman, J. S. (2019). Impacts of Zostera eelgrasses on microbial community structure in San Diego coastal waters. *Elementa*, *7*(1). https://doi.org/10.1525/elementa.350

Wehrens, R., & Kruisselbrink, J. (2018). Flexible Self-Organizing Maps in kohonen 3.0. *Journal of Statistical Software*, *87*(7), 1–18. https://doi.org/10.18637/JSS.V087.I07

Weissman, J. L., Hou, S., & Fuhrman, J. A. (2021). Estimating maximal microbial growth rates from cultures, metagenomes, and single cells via codon usage patterns. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(12), e2016810118. https://doi.org/10.1073/PNAS.2016810118/SUPPL_FILE/PNAS.2016810118.SAPP.PDF

Williams, P. J. (1981). Microbial contribution to overall marine plankton metabolism: direct measurements of respiration. *Oceanologica Acta*, *4*(3), 359–364.

Willis, M. D., Lannuzel, D., Else, B., Angot, H., Campbell, K., Crabeck, O., Delille, B., Hayashida, H., Lizotte, M., Loose, B., Meiners, K. M., Miller, L., Moreau, S., Nomura, D., Prytherch, J., Schmale, J., Steiner, N., Tedesco, L., & Thomas, J. (2023). Polar oceans and sea ice in a changing climate. *Elementa: Science of the Anthropocene*, *11*(1), 1. https://doi.org/10.1525/ELEMENTA.2023.00056

Wilson, J. M., Chamberlain, E. J., Erazo, N., Carter, M. L., & Bowman, J. S. (2021). Recurrent microbial community types driven by nearshore and seasonal processes in coastal Southern California. *Environmental Microbiology*. https://doi.org/10.1111/1462-2920.15548

Worden, A. Z., Follows, M. J., Giovannoni, S. J., Wilken, S., Zimmerman, A. E., & Keeling, P. J. (2015). Rethinking the marine carbon cycle: Factoring in the multifarious lifestyles of microbes. *Science*, *347*(6223). https://doi.org/10.1126/SCIENCE.1257594

Worster, M. G., & Wettlaufer, J. S. (1997). Natural convection, solute trapping, and channel formation during solidification of saltwater. *Journal of Physical Chemistry B*, *101*(32), 6132–6136. https://doi.org/10.1021/jp9632448

Xiao, Y.-K., Yan, Z.-F., Kim, Y., Lee, H.-M., Trinh, H., Yang, J.-E., Won, K.-H., Yi, T.-H., & Kook, M. (2019). Glaciecola amylolytica sp. nov., an amylase-producing bacterium isolated from seawater. *International Journal of Systematic and Evolutionary Microbiology*, *69*(4), 957–963. https://doi.org/10.1099/IJSEM.0.003222

Xu, D., Kong, H., Yang, E. J., Li, X., Jiao, N., Warren, A., Wang, Y., Lee, Y., Jung, J., & Kang, S. H. (2020). Contrasting community composition of active microbial eukaryotes in melt ponds and sea water of the Arctic Ocean revealed by high throughput sequencing. *Frontiers in Microbiology*, *11*, 501327. https://doi.org/10.3389/FMICB.2020.01170/BIBTEX

Zhou, J., Kotovitch, M., Kaartokallio, H., Moreau, S., Tison, J. L., Kattner, G., Dieckmann, G., Thomas, D. N., & Delille, B. (2016). The impact of dissolved organic carbon and bacterial respiration on pCO2 in experimental sea ice. *Progress in Oceanography*, *141*, 153–167. https://doi.org/10.1016/J.POCEAN.2015.12.005