DIGITIZING FROZEN EARTH - REVEALING MICROBIAL DIVERSITY AND PHYSIOLOGY IN THE CRYOBIOSPHERE THROUGH ´OMICS´ TOOLS, VOLUME II

EDITED BY: Anne D. Jungblut, Samuel Cirés, Jérôme Comte, Julia Kleinteich, Krishnan Kottekkatu Padinchati, Birgit Sattler and David Velazquez

PUBLISHED IN: Frontiers in Microbiology

frontiers Research Topics

ISSN 1664-8714
ISBN 978-2-83250-506-9
DOI 10.3389/978-2-83250-506-9

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

DIGITIZING FROZEN EARTH - REVEALING MICROBIAL DIVERSITY AND PHYSIOLOGY IN THE CRYOBIOSPHERE THROUGH ´OMICS´ TOOLS, VOLUME II

Topic Editors:
Anne D. Jungblut, Natural History Museum (United Kingdom), United Kingdom
Samuel Cirés, Autonomous University of Madrid, Spain
Jérôme Comte, Université du Québec, Canada
Julia Kleinteich, Bundesanstalt für Gewässerkunde (BFG), Germany
Krishnan Kottekkatu Padinchati, National Centre for Polar and Ocean Research (NCPOR), India
Birgit Sattler, University of Innsbruck, Austria
David Velazquez, Autonomous University of Madrid, Spain

Citation: Jungblut, A. D., Cirés, S., Comte, J., Kleinteich, J., Padinchati, K. K., Sattler, B., Velazquez, D., eds. (2022). Digitizing Frozen Earth - Revealing Microbial Diversity and Physiology in the Cryobiosphere through ´Omics´ Tools, Volume II. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-83250-506-9

Table of Contents

frontiers | Frontiers in Microbiology

TYPE Editorial
PUBLISHED 30 September 2022
DOI 10.3389/fmicb.2022.1013398

OPEN ACCESS

EDITED AND REVIEWED BY
Andreas Teske,
University of North Carolina at Chapel Hill, United States

*CORRESPONDENCE
Jérôme Comte
jerome.comte@inrs.ca
Birgit Sattler
birgit.sattler@uibk.ac.at

SPECIALTY SECTION
This article was submitted to Extreme Microbiology, a section of the journal Frontiers in Microbiology

RECEIVED 07 August 2022
ACCEPTED 26 August 2022
PUBLISHED 30 September 2022

CITATION
Jungblut AD, Velazquez D, Cirés S, Kleinteich J, Kottekkatu Padinchati K, Sattler B and Comte J (2022) Editorial: Digitizing frozen earth—revealing microbial diversity and physiology in the cryobiosphere through "omics" tools, volume II. *Front. Microbiol.* 13:1013398. doi: 10.3389/fmicb.2022.1013398

Editorial: Digitizing frozen earth—revealing microbial diversity and physiology in the cryobiosphere through "omics" tools, volume II

Anne D. Jungblut[1], David Velazquez[2], Samuel Cirés[2], Julia Kleinteich[3], Krishnan Kottekkatu Padinchati[4], Birgit Sattler[5,6]* and Jérôme Comte[7,8]*

[1]Life Sciences Department, Natural History Museum, London, United Kingdom, [2]Biology Department, Universidad Autonoma de Madrid, Madrid, Spain, [3]German Federal Institute of Hydrology, Koblenz, Germany, [4]Arctic Ecology and Biogeochemistry Division, National Institute for Polar and Ocean Research, Vasco da Gama, India, [5]Department of Ecology, University of Innsbruck, Innsbruck, Austria, [6]Austrian Polar Research Institute, Vienna, Austria, [7]Institut National de la Recherche Scientifique, Centre Eau Terre Environnement, Quebec City, QC, Canada, [8]Centre d'Études Nordiques (CEN), Université Laval, Quebec City, QC, Canada

KEYWORDS

Arctic, Antarctica, alpine environments, polar microbiology, high throughput sequencing

Editorial on the Research Topic
Digitizing frozen earth—revealing microbial diversity and physiology in the cryobiosphere through "omics" tools, volume II

Terrestrial, marine and freshwater habitats of the three poles (Arctic, Antarctica, and high-altitude regions) have striking similarities in their environmental properties. The harsh conditions permit the survival of a unique selection of (micro)organisms. Global change is unfolding across polar and alpine regions without precedents in human history (IPCC, 2022). The current human development paradigm has important environmental consequences and all forms of life thriving in the cryosphere face pressures with consequences beyond local scales. Microbes ranging from Bacteria, Archaea, and Eukaryotes (i.e., algae, protists, and fungi) are not only able to thrive in permanently cold environment but also are an important component of foodwebs and players in aquatic and terrestrial ecosystems function, microclimate regimes and any possible response of the polar and alpine regions to anthropogenic warming (Cavicchioli et al., 2019; Tiedje et al., 2022). They are the first responders to environmental changes and act as game-changers of key ecological process kinetics. Microbial ecology research today faces the challenges of still quite complex and unresolved issues to foster its understanding of the natural world.

Effects of global warming on the physical, chemical, ecological structure, and function and biodiversity of polar and alpine ecosystems are not well-understood and there are many opinions on how microbes might respond and adapt to environmental stress and long term changes in environmental conditions in freshwater, ice and soil environments (Guan et al., 2017). At the same time climatic driven environmental change and an increasing number of extreme events are being observed. In order to be able to develop a better understanding of the potential effect of climatic perturbations on ecosystems structure and function at the Three Poles, it requires more details on the ecological drivers of community assemblies and resolution on seasonal and longer time scales that microbial assemblages might respond to environmental change. This might also help to identify sentinel taxa or habitats that could act as indicators of change including ice-based ecosystems as part of the Last Ice Area. The increasing use of "omics" techniques to disentangle microbial diversity and ecosystem function in the cryosphere is now offering new avenues to investigate the response of polar and alpine ecosystems microbial communities to global warming (Abatenh et al., 2018; Edwards et al., 2020).

The aim of this Research Topic "*Digitizing frozen earth—Revealing microbial diversity and physiology in the Cryobiosphere through "Omics" tools, volume II*" is to provide readers with a selection of studies that are using the latest "omics" and multidisciplinary approaches as well as go beyond the description of microbial communities to evaluating mechanistic processes that shape environmental microbial assemblages. Along this Editorial, we present and highlight with concrete examples how this walkthrough—from boots to bytes—is carried out by the scientific community to overcome the hurdles from samples to integration of large ecosystem processes. Contributions include original research studies on understudied taxonomic groups such as protists, fungi, and uncultured candidate bacterial groups as well as evaluation of mechanistic processes driving community assembly across habitats and seasons in polar and alpine environments based on field and laboratory studies. Research featured in this Research Topic also investigates the response of microbial communities to environmental factors and changing conditions along extended temporal scales. The Research Topic brings together studies from the Three Poles; covering Arctic, Antarctic and high altitude environments, to allow a comparison and perspective on the similarities and differences of these permanently cold but geographically separated environments.

The Research Topic includes contributions that demonstrate the presence of key candidate bacteria groups in Antarctic lakes and their potential physiology and contribution to biogeochemical processes based on high-throughput "omics" techniques. Williams, Allen, Berengut et al. outlines a characterization of novel *Cloacimonadota* and *Omnitrophota* from Ace Lake as well as Candidatus *Organicella extenuata*, a Verrucomicrobial endosymbiont with a reduced genome, from Organic Lake, Vestfold Hills in Antarctica (Williams, Allen, Ivanova et al.). While bacteria have been covered by many 16S rRNA gene high throughput sequencing surveys and dominating environmental metagenomic studies, there is still a surprising paucity of data for protists and fungi. Millar et al. showed that 18S rRNA gene based protists and fungi assemblages in supraglacial cryoconite include potential grazers, predators, photoautotrophs that show distinct communities in the Arctic and Antarctic similar to prokaryotic communities. The study by Ilicic et al. focused on microbial eukaryotes and fungi in shallow marine coastal zones in Potters Cove (South Shetland Islands, Western Antarctic Peninsula), and their assessment of fungi suggests that Ascomycota and Chytridiomycota fungi are the most abundant taxa in these benthic microphytobenthic habitats including putative fungal parasites. Furthermore, the review by Gilbertson et al. focuses on polar marine diatoms and evolution and adaptation mechanisms based Omics approaches. All three contributions highlight that there is a need for a more comprehensive assessment of these taxonomic groups.

As the climate continues to warm, the cryospheric environments will experience drastic changes in their structure: ice thickness reduces and aquatic ecosystems become more open, the atmosphere increasingly exchanges with the terrestrial environments, and the active layer of soil is deepening as permafrost thaws. These changes occur at different temporal and spatial scales but collectively represent a strong environmental filter for local microbial communities that have to either respond and adjust to these new conditions or new niches are now available for previously rare taxa or originating from the surrounding environment (Leibold et al., 2004; Hanson et al., 2012; Lindström and Langenheder, 2012). The contributions of this Research Topic explore the mechanisms for selection and habitat filtering of microbial communities and lead to local variation in assemblages in permanently cold environments. The studies also highlight the microbiomes are usually influenced by multiple factors and the difficulty of untangling them. For example, Fillinger et al. describes that bacteria cell number and virus-like particles are potentially influenced by both spatial and temporal drivers in the European Alps. Rapp et al. also found distinct difference in composition and the genomic traits in communities from Arctic seawater-derived subzero hypersaline brines from 1st year ice (near Barrow Sea Ice Balance Site, Alaska) and ancient cryopeg (Barrow Permafrost Tunnel, Alaska) in permafrost.

Marois et al. suggested that local habitat filtering shapes the planktonic microbial community structure in four chemically and physically distinct lakes on northern Ellesmere Island, Canadian Arctic Archipelago that is part of the coastal margin zone of the Last Ice Area. While Blais et al. investigated river bacterial communities along dissolved organic carbon (DOC) and salinity gradients in the Great Whale River Subarctic, Canada, and found that a core microbiome in subarctic rivers

and a coastal plume contained generalists but also taxa with a more limited distribution. However, many studies to date cover mostly the evaluation of the relationship between microbial communities and environmental variables during the short duration of the polar summer due to logistical constraints, and there is therefore a lack of field studies during the remaining and majority of the year between autumn, winter and spring. Extending studies across seasons not only will provide a more complete understanding, but they will also provide valuable data in context of climate change research as it has been predicted that warming in the polar regions will lead to changes in the seasonality. In this context, a comparison between planktonic 18S rRNA gene communities in lakes in Ekaluktutiak (Cambridge Bay, southern Victoria Island, Nunavut, Canada) found differences in the mixotrophic and heterotrophic core communities in open water and under the ice conditions (Potvin et al.).

Environmental gradients and chronosequences provide opportunities for natural laboratories to evaluate the response of microbial communities to a change of conditions across space and time which is explored in several studies on soil ecosystems. The space-for-time substitution was applied to study the response of soil microbial communities to water availability and geochemical soil properties in the McMurdo Dry Valleys as the polar deserts are predicted to experience drastic changes in water availability under current climate change predictions (Monteiro et al.). Liu et al. evaluated the changes of pioneer prokaryotic communities in rhizosphere and bulk soils along the high-elevation glacier retreat chronosequence, the northern Himalayas, Tibetan Plateau, and found that prokaryotic community composition during colonization and succession are shaped by a combination of plants, habitat and duration since glacier retreat. In the permafrost environment, Scheel et al. was able to do first prokaryotic 16S rRNA and fungal ITS2 gene regions sequencing study of an abrupt permafrost erosion microbiome in Northeast Greenland, where a thermal erosion gully collapsed, leading to the thawing of 26,500-year-old permafrost material. Finally, a laboratory-based microcosm experiment was presented by Malard and Pearce to evaluate the colonization of snow derived bacteria deposited onto Arctic soils, demonstrating that potentially successful colonizations of soil by invading bacteria could be influenced by local soil properties rather than by ecosystem disturbances.

In summary, with ongoing climate warming and rapid alteration of the cryosphere at the Three Poles, a mechanistic understanding of the response of microbial communities is greatly needed. The broad-range of articles presented here deepen our current understanding and knowledge of the mechanistic processes that shape environmental microbial assemblages. This Research Topic of studies identified many challenges and questions that yet remain to be addressed. We are hopeful that this Research Topic will stimulate discussions and paved the way for future avenues of research.

Author contributions

AJ, JC, BS, DV, and SC conceived of the manuscript. JK and KK contributed and commented on the article. All authors contributed to the article and approved the submitted version.

Funding

We thank Sentinel North (CFREF) and ArcticNet (NCE) for funding support for the T-MOSAiC workshop. Natural Sciences and Engineering Research Council of Canada (NSERC, #RGPIN-2020-06874). DV and SC were supported by CAM-UAM funds (ref.: SI3-PJI-2021-00461).

Acknowledgments

The Research Topic is part of the programe T-MOSAiC and we thank Warwick F. Vincent, Joao Canario, and Diogo Folhas for support of the Microbiome Action Group. We are grateful to the Editorial Staff at Frontiers in Microbiology for their initial invitation and support throughout. The editors would like to thank all reviewers who evaluated manuscripts for this Research Topic. Their contribution in improving the manuscripts has been greatly appreciated.

Conflict of interest

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

Publisher's note

References

Abatenh, E., Gizaw, B., Tsegaye, Z., and Tefera, G. (2018). Microbial function on climate change - a review. *Environ. Pollut. Climate Change* 2, 147. doi: 10.4172/2573-458X.1000147

Cavicchioli, R., Ripple, W. J., Timmis, K. N., Azam, F., Bakken, L. R., Baylis, M., et al. (2019). Scientists' warning to humanity: microorganisms and climate change. *Nat. Rev. Microbiol.* 17, 569–586. doi: 10.1038/s41579-019-0222-5

Edwards, A., Cameron, K. A., Cook, J. M., Debbonaire, A. R., Furness, E., Hay, M. C., et al. (2020). Microbial genomics amidst the Arctic crisis. *Microb. Genom.* 2020, e000375. doi: 10.1099/mgen.0.000375

Guan, N., Li, J., Shin, H. D., Du, G., Chen, J., and Liu, L. (2017). Microbial response to environmental stresses: from fundamental mechanisms to practical applications. *Appl. Microbiol. Biotechnol.* 101, 3991–4008. doi: 10.1007/s00253-017-8264-y

Hanson, C. A., Fuhrman, J. A., Horner-Devine, M. C., and Martiny, J. B. H. (2012). Beyond biogeographic patterns: processes shaping the microbial landscape. *Nat. Rev. Microbiol.* 10, 497–506. doi: 10.1038/nrmicro2795

IPCC (2022). "Summary for policymakers," in *Climate Change 2022: Mitigation of Climate Change. Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, eds P. R. Shukla, J. Skea, R. Slade, A. Al Khourdajie, R. van Diemen, D. McCollum, et al. (Cambridge; New York, NY: Cambridge University Press), 1–2913. doi: 10.1017/9781009157926.001

Leibold, M. A., Holyoak, M., Mouquet, N., Amarasekare, P., Chase, J. M., Hoopes, M. F., et al. (2004). The metacommunity concept: a framework for multi-scale community ecology. *Ecol. Lett.* 7, 601–613. doi: 10.1111/j.1461-0248.2004.00608.x

Lindström, E. S., and Langenheder, S. (2012). Local and regional factors influencing bacterial community assembly. *Environ. Microbiol. Reports* 4, 1–9. doi: 10.1111/j.1758-2229.2011.00257.x

Tiedje, J. M., Bruns, M. A., Casadevall, A., Criddle, C. S., Eloe-Fadrosh, E., Karl, D. M., et al. (2022). Microbes and climate change: a research prospectus for the future environmental microbiology. *mBio* 13, e0080022. doi: 10.1128/mbio.00800-22

frontiers in Microbiology

ORIGINAL RESEARCH
published: 01 June 2021
doi: 10.3389/fmicb.2021.674758

Genome Analysis of a Verrucomicrobial Endosymbiont With a Tiny Genome Discovered in an Antarctic Lake

*Timothy J. Williams[1], Michelle A. Allen[1], Natalia Ivanova[2], Marcel Huntemann[2], Sabrina Haque[1], Alyce M. Hancock[1†], Sarah Brazendale[1†] and Ricardo Cavicchioli[1]**

[1] School of Biotechnology and Biomolecular Sciences, UNSW Sydney, Sydney, NSW, Australia, [2] U.S. Department of Energy Joint Genome Institute, Berkeley, CA, United States

OPEN ACCESS

Edited by:
Anne D. Jungblut, Natural History Museum, United Kingdom

Reviewed by:
Francisco Rodriguez-Valera, Miguel Hernández University of Elche, Spain
Stefano Campanaro, University of Padua, Italy

***Correspondence:**
Ricardo Cavicchioli
r.cavicchioli@unsw.edu.au

†Present address:
Alyce M. Hancock, Institute for Marine and Antarctic Studies, University of Tasmania, Battery Point, TAS, Australia
Sarah Brazendale, King Island Brewhouse, Pegarah, TAS, Australia

Specialty section:
This article was submitted to Extreme Microbiology, a section of the journal Frontiers in Microbiology

Received: *01 March 2021*
Accepted: *23 April 2021*
Published: *01 June 2021*

Citation:
Williams TJ, Allen MA, Ivanova N, Huntemann M, Haque S, Hancock AM, Brazendale S and Cavicchioli R (2021) Genome Analysis of a Verrucomicrobial Endosymbiont With a Tiny Genome Discovered in an Antarctic Lake. Front. Microbiol. 12:674758. doi: 10.3389/fmicb.2021.674758

Organic Lake in Antarctica is a marine-derived, cold (−13°C), stratified (oxic-anoxic), hypersaline (>200 gl^{-1}) system with unusual chemistry (very high levels of dimethylsulfide) that supports the growth of phylogenetically and metabolically diverse microorganisms. Symbionts are not well characterized in Antarctica. However, unicellular eukaryotes are often present in Antarctic lakes and theoretically could harbor endosymbionts. Here, we describe *Candidatus* Organicella extenuata, a member of the Verrucomicrobia with a highly reduced genome, recovered as a metagenome-assembled genome with genetic code 4 (UGA-to-Trp recoding) from Organic Lake. It is closely related to *Candidatus* Pinguicocccus supinus (163,218 bp, 205 genes), a newly described cytoplasmic endosymbiont of the freshwater ciliate *Euplotes vanleeuwenhoeki* (Serra et al., 2020). At 158,228 bp (encoding 194 genes), the genome of *Ca.* Organicella extenuata is among the smallest known bacterial genomes and similar to the genome of *Ca.* Pinguicoccus supinus (163,218 bp, 205 genes). *Ca.* Organicella extenuata retains a capacity for replication, transcription, translation, and protein-folding while lacking any capacity for the biosynthesis of amino acids or vitamins. Notably, the endosymbiont retains a capacity for fatty acid synthesis (type II) and iron–sulfur (Fe-S) cluster assembly. Metagenomic analysis of 150 new metagenomes from Organic Lake and more than 70 other Antarctic aquatic locations revealed a strong correlation in abundance between *Ca.* Organicella extenuata and a novel ciliate of the genus *Euplotes*. Like *Ca.* Pinguicoccus supinus, we infer that *Ca.* Organicella extenuata is an endosymbiont of *Euplotes* and hypothesize that both *Ca.* Organicella extenuata and *Ca.* Pinguicocccus supinus provide fatty acids and Fe-S clusters to their *Euplotes* host as the foundation of a mutualistic symbiosis. The discovery of *Ca.* Organicella extenuata as possessing genetic code 4 illustrates that in addition to identifying endosymbionts by sequencing known symbiotic communities and searching metagenome data using reference endosymbiont genomes, the potential exists to identify novel endosymbionts by searching for unusual coding parameters.

Keywords: Antarctic microbiology, Bacterial endosymbionts, metagenome, extreme genome reduction, genetic code 4

INTRODUCTION

Bacteria with highly reduced genome sizes are only found as host-restricted symbionts and pathogens (**Supplementary Table 1**; Moran and Bennett, 2014). The smallest bacterial genomes are only known to occur in symbionts that are required by a host (obligate symbionts), with those possessing genomes < 500 kbp being completely dependent on the host while also providing benefit to the host to be retained (mutualistic symbionts) (Moran and Bennett, 2014). Insects that feed on sap (phloem or xylem) rely on endosymbionts to supplement their restrictive or unbalanced diets; these bacteria, either individually or as "patchworks" of metabolically complementary co-symbionts or nested symbionts, provide essential amino acids and/or vitamins for their respective protist hosts (Nakabachi and Ishikawa, 1999; Zientz et al., 2004; Nakabachi et al., 2006; Pérez-Brocal et al., 2006; Bennett and Moran, 2013; Brown et al., 2015; Gil et al., 2018). The cellulolytic protists that reside in the hindguts of termites harbor cytoplasmic endosymbionts that belong to diverse bacterial clades (e.g., Endomicrobia, Deltaproteobacteria, Bacteroidetes, and Actinobacteria) and confer metabolic and nutritional benefits to their respective protist hosts (Stingl et al., 2005; Ohkuma et al., 2007; Hongoh et al., 2008a,b; Sato et al., 2009; Ikeda-Ohtsubo et al., 2016; Strassert et al., 2016; Kuwahara et al., 2017).

Verrucomicrobia is a diverse phylum of bacteria that has been found in a wide array of habitats, with free-living representatives isolated from soils, seawater, marine sediments, lakes, and hot springs (Wagner and Horn, 2006; Dunfield et al., 2007; Yoon et al., 2007a,b). Certain verrucomicrobia live in close association with eukaryotes, including marine sponges (Scheuermayer et al., 2006; Yoon et al., 2008) and tunicates (Lopera et al., 2017), as well as inside the intestinal tracts of humans (Derrien et al., 2004), termites (Wertz et al., 2012), and marine clam worms (Choo et al., 2007). Some verrucomicrobia have entered into very close symbiotic associations with eukaryotic hosts, including anti-predator ectosymbionts (epixenosomes) of the ciliate *Euplotidium* (Petroni et al., 2000) and various endosymbionts, such as inside the cytoplasm of intestinal and ovarial cells of nematode worms (Vandekerckhove et al., 2002), nuclei of cellulolytic protists (Sato et al., 2014), and the cytoplasm of the ciliate *Euplotes vanleeuwenhoeki* (Serra et al., 2020). *Candidatus* Xiphinematobacter, the verrucomicrobial endosymbiont of nematodes, has a 0.916-Mbp metagenome assembled genome (MAG) encoding 817 predicted protein-coding sequences (CDS); compared with free-living relatives, genes are retained for the biosynthesis of amino acids predicted to be required by their nematode hosts (Brown et al., 2015). The unpublished MAG of the intranuclear endosymbiont *Ca.* Nucleococcus (Sato et al., 2014) is ∼1 Mbp and encodes ∼700 CDS (Y. Hongoh, personal communication). The genome of the *Euplotes* endosymbiont *Ca.* Pinguicoccus supinus has an 'extremely reduced genome' at only 0.163 Mbp and encodes 168 CDS (Serra et al., 2020).

Organic Lake is a shallow (∼7 m deep), marine-derived, Antarctic lake formed ∼3,000 years ago (Gibson, 1999). The lake is characterized by a salinity gradient that reaches a maximum of ∼230 g L^{-1} (Gibson, 1999) and has unusual chemistry, with very high levels of dimethylsulfide (Gibson et al., 1991). Temperatures in the upper waters have been recorded as high as 15°C and as low as −14°C (Franzmann et al., 1987), whereas bottom waters (5 – ∼7 m) have typically registered temperatures of −5 to −6°C (Franzmann et al., 1987; Gibson et al., 1991; Roberts et al., 1993; James et al., 1994) but as low as −13°C (Yau et al., 2013). Metaproteogenomic analyses have inferred important roles for virophage-mediated control of algal primary production (Yau et al., 2011) and roles in nutrient cycling by phylogenetically and metabolically diverse bacteria (Yau et al., 2013). The lake is located in the Vestfold Hills, a ∼ 400 km^2 region of East Antarctica that contains hundreds of water bodies, many of which are marine-derived, having been formed ∼3,000–7,000 years ago as a result of the isostatic rebound of the continent (Gibson, 1999; Cavicchioli, 2015; **Supplementary Figure 1**). The water bodies in the Vestfold Hills range in salinity from freshwater to hypersaline, most of which have not been subject to metagenomic analysis of their biota (Cavicchioli, 2015).

During analyses of unusual coding parameters (genetic code 4) in metagenome contigs, we discovered a 158-kbp verrucomicrobial MAG that was assembled from new metagenome data derived from a complete seasonal cycle of Organic Lake. The MAG is comparable in size with the endosymbiont *Ca.* Pinguicoccus supinus (Serra et al., 2020), as well as to obligate mutualistic endosymbionts that belong to the phyla Proteobacteria and Bacteroidetes that also have extremely reduced genomes (Moran and Bennett, 2014). The environmental distribution and inferred *Euplotes* host of the Organic Lake endosymbiont was assessed by analyzing 337 Antarctic metagenomes, including 150 new metagenomes of unstudied Vestfold Hills lakes and neighboring marine locations. The Organic Lake endosymbiont is closely related to *Ca.* Pinguicoccus; herein, we describe the functional traits of this bacterial lineage that seem to underpin the endosymbiosis and discuss the value of searching for unusual coding parameters as a means of identifying endosymbionts.

MATERIALS AND METHODS

Sampling and DNA Extraction

Microbial biomass was obtained and field observations recorded from lakes in the Vestfold Hills, Antarctica (**Supplementary Figure 1**). Sampling at Organic Lake was performed by sequential size fractionation through a 20-μm prefilter onto 3.0-, 0.8-, and 0.1-μm large-format membrane-filters (293 mm diameter polyethersulfone), samples preserved and DNA extracted, as described previously (Yau et al., 2011, 2013; Tschitschko et al., 2018; Panwar et al., 2020).

For other lakes, including Unnamed Lake 18, "Portals" Lake, Unnamed Lake 13, Unnamed Lake 17, "Swamp" Lake, Unnamed Lake 12, and Unnamed Lake 7, biomass was captured using Sterivex cartridges (MilliporeSigma, Burlington, MA, United States) by pumping water from the lake through a 20-μm prefilter using a hand-driven peristaltic pump. After field collection, Sterivex cartridges were kept cold (e.g., in snow)

before transportation to Davis Research Station, where they were cryogenically preserved at −80°C and shipped at −80°C to Australia. To extract DNA, the Sterivex cartridge was removed from −80°C storage and filled with 1.6 ml of freshly prepared "XS" buffer (1% potassium ethyl xanthogenate; 100-mM Tris-hydrochloride, pH 7.4; 20-mM ethylenediamine tetraacetic acid, pH 8; 1% sodium dodecyl sulfate; 800-mM ammonium acetate) (Tillett and Neilan, 2000). Both ends of the cartridge were sealed with parafilm, and the cartridge was placed into an empty 50-ml Falcon tube and incubated in a water bath at 65°C for 2 h. After incubation, 200 μl of 10% sodium dodecyl sulfate and 50 μl of 20 mg ml^{-1} Proteinase K (Thermo Fisher Scientific, Waltham, MA, United States) was added through the Luer-lock end of the cartridge, re-sealed, and returned to the 50-ml Falcon tube for incubation in a water bath at 55°C for 2 h. After incubation, a syringe was attached to the Luer-lock end and air injected to recover the liquid in a 20-ml Falcon tube. The liquid was decanted, placing 500-μl aliquots into 1.5-ml microfuge tubes, 60 μl of phenol added, the tubes inverted several times to mix the solution, 500 μl of chloroform: isoamyl alcohol (24:1) was added, and each tube mixed by inversion. The tubes were centrifuged at 16,800 × *g* for 10 min at room temperature, the aqueous phase of each sample was collected into a fresh 1.5-ml tube, 1.5 μl of GlycoBlue (Thermo Fisher Scientific) was added to each tube, and tubes were left at room temperature for 1 h. Ammonium acetate (3 M, 500 μl) was added to each tube, mixed by inversion, left at room temperature for 30 min, tubes centrifuged at 16,800 × *g* for 15 min, and the supernatant placed into fresh 2-ml tubes. A total of 1 ml of 100% ethanol was added to each tube, and after storage overnight at 4°C, tubes were centrifuged at 14,000 × *g* for 30 min at room temperature and the supernatant carefully discarded. Pellets were washed by adding 500 μl of 70% ethanol and tubes centrifuged at 14,000 × *g* for 5 min. Ethanol was removed, the pellets air-dried on a heating block at 37°C, pellets resuspended in Tris-hydrochloride–ethylenediaminetetraacetic acid buffer (10-mM Tris-hydrochloride, pH 7.4; 1-mM ethylenediamine tetraacetic acid, pH 8) and tubes stored at −80°C. DNA yields were quantified using Qubit dsDNA BR Assay Kit (Thermo Fisher Scientific) and the quality of DNA assessed by agarose gel electrophoresis. DNA was sequenced at the Joint Genome Institute using Hi-Seq2500 (2 × 151 bp run) as described previously (Tschitschko et al., 2018; Panwar et al., 2020) or at the Australian Centre for Ecogenomics using NextSeq500 (on a 2 × 150 bp run) and raw reads filtered using Trimmomatic (Trimmomatic manual: V0.32, no date). Assembly was performed with metaSpades and all contigs > 200 bp uploaded and annotated by the IMG pipeline (Huntemann et al., 2015).

Analyses of DNA Sequence Data

The *Ca.* Organicella MAG was identified using a pipeline to identify stop codon reassignments in metagenomic data (Ivanova et al., 2014). The set of contigs with potential UGA reassignment was identified in Organic Lake metagenomes based on the higher total coding potential as computed by Prodigal upon reannotation with genetic code 4. These contigs had an average GC content of 32%, and they appeared to have characteristics of bacterial genomes, namely, high coding density, typical bacterial gene complement with translation, transcription, and replication machinery, but no multi-subunit NADH dehydrogenase and cytochrome oxidase complexes indicative of mitochondria and no photosynthesis genes indicative of chloroplasts. The longest of these contigs, which were ∼158 kb, turned out to be circular due to an overlap of 100 nt at the ends. No other putative bacterial contigs with UGA reassignment were found in the same metagenomes, suggesting that these circular contigs constituted the entire genome of a bacterium. Because the automated annotations initially performed by IMG used genetic code 11, in which UGA is a stop codon, manual inspection of these contigs identified genes interrupted by stop codons within open reading frames. Re-calling open reading frames and annotating the genome using PROKKA (Seemann, 2014) with codon chart 4, reassigned the opal stop codon (UGA) as tryptophan. This reduced the number of genes from 249 to 193 for the reference *Ca.* Organicella extenuata MAG (contigID Ga0307966_1000010). Total coding density was calculated using all protein-coding genes (CDS), rRNA, tRNA, and tmRNA genes in the genome. Protein identities were determined using ExPASy BLAST for all CDS and, where necessary, InterProScan and HHPred. The isoelectric point (pI) of protein sequences was determined using the Isoelectric Point Calculator (Kozlowski, 2016). The genomic functional potential was assessed by considering cellular and metabolic traits based upon manual examination of genes and pathways performed in a similar way to previous assessments of the veracity of gene functional assignments (Allen et al., 2009; Panwar et al., 2020).

Mapping of reads from 340 Antarctic metagenomes to the *Ca.* Organicella MAG was performed using BWA v0.7.17 (Li and Durbin, 2009). FastANI (Jain et al., 2018) was used to calculate ANI between *Ca.* Organicella MAGs. Multiple alignments were constructed using Clustal (DNA sequences) (Thompson et al., 1994) or MUSCLE (protein sequences) (Edgar, 2004) and used to construct phylogenetic trees (for *Ca.* Organicella and for *Euplotes* sp. AntOrgLke) by the maximum-likelihood method (Tamura and Nei, 1993) in MEGA6 (Tamura et al., 2013) with 1,000 bootstraps. Marker genes predicted from the *Ca.* Organicella MAG were used to place the MAG into a concatenated 43-marker gene tree by CheckM (Parks et al., 2014) using the tree command.

To identify the potential host(s) of *Ca.* Organicella, six metagenomes where *Ca.* Organicella was abundant (Org-646, Org-46, Org-175, Org-784, Portals, and UnnamedLake18) were selected to create a co-assembly using Megahit v.1.2.2b (Li et al., 2016), with contigs binned into 188 MAGs by Metabat v.2.12.1 (Kang et al., 2019) with default settings (min contig length 2,500). During the co-assembly, a single contig representing the *Ca.* Organicella MAG was assembled (k141_311079; 158,131 bp). As this *Ca.* Organicella MAG was not initially binned, due to falling below the default minimum bin size of 200 kb, it was manually assigned to bin189. The ANI between the original *Ca.* Organicella MAG (Ga0307966_1000010) and the co-assembled MAG (k141_311079) was 99.9924%. The *Ca.* Organicella MAG (bin189) and the 188 bins resulting from Metabat binning were screened for contamination, completion, and taxonomic identity using checkM (Parks et al., 2014) and refineM (Parks et al.,

2017). In addition, the abundance of each bin [calculated as the sum of (contig length × contig coverage) for all contigs in the bin] was determined for each of the 29 metagenomes where *Ca.* Organicella was detected by mapping the metagenome reads to the bins with bbmap v38.51 (Bushnell, 2014). These bin abundances were used as input for SparCC (Friedman and Alm, 2012) implemented in python 3[1] to estimate correlation values from the compositional data.

To identify the taxonomy of the bins that were highly correlated to *Ca.* Organicella, MetaEuk v. 20200908 (Levy Karin et al., 2020) was used to identify eukaryotic proteins and assign taxonomy *via* the 2bLCA lowest common ancestor approach. To identify the maximal number of proteins, the larger MERC_MMETSP_Uniclust50_profiles database was used as the reference dataset for MetaEuk, whereas to assign contig taxonomy and putative protein function, the TaxDB_uniclust90_2018_08 database was used. Both databases were obtained from http://wwwuser.gwdg.de/$\sim$compbiol/metaeuk/. The rRNA gene contig missing from the *Euplotes* sp. AntOrgLke MAG was identified as contig k141_859071 by blasting *Euplotes* spp. 18S rRNA genes against the co-assembled contigs used for Metabat. 18S/28S rRNA genes on k141_859071 were predicted using the RNAmmer 1.2 Server at http://www.cbs.dtu.dk/services/RNAmmer/ (Lagesen et al., 2007). The mitochondrial genome of *Euplotes* sp. AntOrgLke was identified by blasting the mitochondrial proteins of *E. vanleeuwenhoeki, Euplotes crassus,* and *Euplotes minuta* against the co-assembled contigs used for the Metabat binning, with the resulting contigs then blasted against the metagenome where *Ca.* Organicella was most abundant.

As few non-mitochondrial *Euplotes* proteins were available in the National Center for Biotechnology Information (NCBI) nr database, additional protein sequences were gathered from five reference *Euplotes* species. The data were obtained from genome-specific databases: *Euplotes octocarinatus*, http://ciliates.ihb.ac.cn/database/home/#eo (Wang et al., 2018); *Euplotes vannus*, http://evan.ciliate.org/ (Chen et al., 2019); from proteins predicted from the Marine Microbial Eukaryote Meta/transcriptome sequencing project (MMETSP): *Euplotes harpa* FSP1.4, IMG ID 3300017294; *Euplotes focardii* TN1, IMG IDs 3300017169 and 3300016941; *E. crassus* CT5, IMG ID 3300017039; all accessed at https://img.jgi.doe.gov/; by manually running MetaEuk for protein prediction on genome sequences held in NCBI Genome: *E. focardii*, GCA_001880345.1 ASM188034v1; *E. crassus* GCA_001880385.1 ASM188038v1; or by manually running MetaEuk on genomic data held in a custom database: *E. vannus*, http://evan.ciliate.org/. Average amino acid identity (AAI) was calculated at http://enve-omics.ce.gatech.edu/aai/index (Rodriguez-R and Konstantinidis, 2016) between the *Euplotes* sp. AntOrgLke and the five reference *Euplotes* species, using the protein sequences downloaded or predicted for their respective genomes (*Euplotes* sp. AntOrgLke – 15328 proteins predicted in this study; *E. octocarinatus* – 29076 proteins obtained from http://ciliates.ihb.ac.cn/database/home/#eo; *E. focardii* TN1, *E. crassus* CT5, and *E. harpa* FSP1.4 – 12634, 12729, and 19386 proteins, respectively, predicted from the Marine Microbial Eukaryote Meta/transcriptome sequencing project (MMETSP) and accessed at http://img.jgi.doe.gov; *E. vannus* – 43338 proteins obtained from http://evan.ciliate.org).

To correctly identify the CDS in the *Ca.* Pinguicoccus genome, Prokka (Seemann, 2014) was used with genetic code 4. To investigate the relationship between *Ca.* Organicella and the newly released *Ca.* Pinguicoccus genome, nucleotide synteny was visualized with progressiveMauve (Darling et al., 2004), a tblastx plot was performed at NCBI[2], and AAI was calculated at http://enve-omics.ce.gatech.edu/aai/index (Rodriguez-R and Konstantinidis, 2016), followed by manual inspection of protein identifications to identify shared and unique metabolic capacities.

RESULTS AND DISCUSSION

Organic Lake MAG Summary and Phylogeny

A MAG (Ga0307966_1000010) representing a complete circular genome with a length of 158,228 bp was identified in new metagenome data from Organic Lake. The MAG encoded 194 bacterial genes, 156 of which were inferred to be CDS (**Supplementary Table 2**) with 145 assigned putative biological functions (**Supplementary Table 3**). Most (76 proteins) were assigned to translation (including tRNA modifications) (**Supplementary Table 3**). Other categories were fatty acid synthesis (including pyruvate oxidation) (18 proteins); cell wall biogenesis including lipopolysaccharides (17), iron–sulfur (Fe-S) cluster assembly (8), protein folding and stability (8), replication and repair (6), and transcription (6). A total of 16 CDS could not be assigned any function, and some or all of these could be pseudogenes. The MAG had one copy each of 23S, 16S, and 5S rRNA genes and 34 identifiable tRNA genes (**Supplementary Table 2**). The highly restricted genomic potential illustrates this bacterium would not be capable of autonomous growth, and we name it *Candidatus* Organicella extenuata gen. et. sp. nov.; the genus name derives from the locality from where the MAG sequence was originally recovered (Organic Lake, Antarctica) with the addition of the diminutive Latin suffix -ella; the species "extenuata" means reduced or diminished in Latin and is in reference to the highly reduced genome.

Additional MAGs for *Ca.* Organicella were generated from a number of Antarctic metagenomes (see section *Ca. Organicella Environmental Distribution and Host* below), enabling the analysis of 23 *Ca.* Organicella 16S rRNA genes (**Supplementary Table 4**). Phylogenetic analysis of these genes found *Ca.* Organicella to be most closely related to *Ca.* Pinguicoccus (Serra et al., 2020), with 85% 16S rRNA gene identity (see section *Comparison of Ca. Organicella and Ca. Pinguicoccus Genomes* below). Both *Ca.* Organicella and *Ca.* Pinguicoccus belong to a cluster of uncultured Verrucomicrobia that also includes *Ca.* Nucleococcus and related endosymbionts of certain amitochondriate protists (*Trichonympha*, *Caduceia*, and *Oxymonas*) present in termite hindguts (Yang et al., 2005;

[1]https://github.com/JCSzamosi/SparCC3

[2]https://blast.ncbi.nlm.nih.gov/

Hongoh et al., 2007; Ikeda-Ohtsubo et al., 2010; Sato et al., 2014; **Figure 1**). This cluster, previously termed the "termite cluster" (Sato et al., 2014), is not closely related to other known verrucomicrobial endosymbionts (Vandekerckhove et al., 2002) or ectosymbionts (Petroni et al., 2000). In view of the cluster now including *Ca.* Organicella and *Ca.* Pinguicoccus, and no longer containing species exclusive to the termite gut, we suggest the cluster be termed the "Nucleococcus cluster." To date, known representatives of this "Nucleococcus cluster" of Verrucomicrobia include both nuclear and cytoplasmic endosymbionts of unicellular eukaryotes.

The *Ca.* Organicella + *Ca.* Pinguicoccus branch within the "Nucleococcus cluster" of the 16S rRNA gene tree was far longer than other branches (**Figure 1**), and similar topology occurred in trees constructed using conserved marker genes (**Supplementary Figure 2**). Such long branches were not evident for any other sequences, including the endosymbionts *Ca.* Nucleococcus and *Ca.* Xiphinematobacter (**Figure 1**), both of which had much larger genomes (∼ 1 Mbp) than *Ca.* Organicella and *Ca.* Pinguicoccus (**Supplementary Table 1**). Long branches likely reflect rapid sequence evolution and are characteristic of degenerate genomes (McCutcheon and Moran, 2012), consistent with *Ca.* Organicella and *Ca.* Pinguicoccus being the only known representatives of Verrucomicrobia with extremely reduced genomes.

Endosymbiont Features

The *Ca.* Organicella MAG exhibits a number of features typical of obligate symbionts that have highly reduced genomes (McCutcheon and Moran, 2010, 2012). The MAG has a high coding density (95% for all genes and 90% for CDS only), with shortened intergenic regions and 23 overlapping genes (**Supplementary Table 1**), which is characteristic of extreme genome reduction (Nakabachi et al., 2006; Moya et al., 2008). The MAG has genetic code 4[3] with UGA stop codons recoded to tryptophan. Of note is that a tRNA-Opal-TCA is also encoded (Ga0307966_1000010189) that has the highest similarity to trnW (UGA) from mitochondria of *Paralemanea* sp. (GenBank accession MG787097.1). UGA-to-Trp recoding is known to occur rarely, having been found in mycoplasmas (Yamao et al., 1985); certain symbiotic bacteria (McCutcheon et al., 2009), including *Ca.* Pinguicoccus (Serra et al., 2020); and several mitochondrial lineages (Knight et al., 2001). The UGA-to-Trp conversion permits the loss of peptide chain release factor 2 (PrfB) (which recognizes UGA codons) through genome erosion (McCutcheon et al., 2009). UGA-to-Trp recoding is typically associated with low GC content (McCutcheon et al., 2009), although some insect endosymbionts with UGA-to-Trp have high GC content (e.g., *Ca.* Hodgkinia cicadicola 58%; *Ca.* Tremblaya princeps PCIT 59%) (McCutcheon and Moran, 2012). The GC content of the *Ca.* Organicella MAG is 32%, compared with 25% for *Ca.* Pinguicoccus (Serra et al., 2020) (also see *Comparison of Ca. Organicella and Ca. Pinguicoccus Genomes*, later). No mobile elements were identified in the *Ca.* Organicella MAG, which is another trait of symbionts with extremely reduced genomes (McCutcheon and Moran, 2012).

The possession of a minimal complement of genes required for transcription and translation (McCutcheon, 2010; McCutcheon and Moran, 2012), and some capacity to perform DNA replication, enables a level of autonomy over cellular processes that distinguishes endosymbiotic bacteria from organelles (McCutcheon and Moran, 2012). *Ca.* Organicella encodes some enzymes involved in DNA replication, including DNA gyrase (GyrAB), DNA primase (DnaG), and replicative DNA helicase (DnaB), but a dedicated DNA polymerase for DNA replication was not identifiable. Although certain insect endosymbionts lack the DNA polymerase III holoenzyme, they at least encode DNA polymerase α-subunit (DnaE), responsible for 5' to 3' polymerization activity of DNA replication (McCutcheon, 2010; McCutcheon and Moran, 2012). In the absence of DnaE, genomic replication is presumably carried out by host proteins (Serra et al., 2020). As in many other reduced endosymbiont genomes, *Ca.* Organicella lacks the DnaA protein for initiation of DNA replication, and this function is presumably carried out by the host (Gil et al., 2003; López-Madrigal et al., 2013), possibly as a mechanism to exercise control over endosymbiont proliferation (e.g., Akman et al., 2002; Gil et al., 2003; Bennett et al., 2014; Bennett and Moran, 2015).

Three subunits of the DNA-directed RNA polymerase (RNAP) for transcription were identified (RpoA, RpoB, and RpoC) as well as a sigma factor (RpoD), components that are typical of endosymbionts (McCutcheon and Moran, 2012). Thus, the components of RNAP retained by *Ca.* Organicella parallel those of unrelated symbionts with genomes of comparable size (McCutcheon, 2010; McCutcheon and Moran, 2012). A total of 34 amino acyl tRNAs for all 20 proteinogenic amino acids were identified, plus aminoacyl tRNA synthetases (aaRS) for 13 of the amino acids (Met, Leu, Ile, Val, Lys, Gly, Ser, Cys, Arg, Tyr, Ala, Phe, and Glu) and a glutamyl/aspartyl-tRNA amidotransferase. The missing aaRS may be provided by the host (Van Leuven et al., 2019), or existing aaRS may catalyze multiple aminoacylation reactions (Moran and Bennett, 2014). *Ca.* Organicella encodes initiation factors IF-1 and IF-2 (but not IF-3); elongation factors EF-G, EF-Ts, and EF-4; translational release factor PrfA (but not PrfB); and ribosome recycling factor. Most, but not all, ribosomal subunits were identified. Known endosymbionts with highly reduced genomes typically do not encode a complete set of ribosomal proteins (McCutcheon, 2010; Moran and Bennett, 2014). Individual ribosomal subunits that could not be identified in the *Ca.* Organicella MAG are also missing from some obligate insect endosymbionts (e.g., RplA, RpmC, RpmD, RpsF, and RpmF) (Moran and Bennett, 2014). Certain tRNA modification enzymes were also evident in the *Ca.* Organicella MAG (e.g., Mnm complex and TsaD) that are usually retained in endosymbionts (McCutcheon and Moran, 2012; Van Leuven et al., 2019) (see ***Supplementary Text*** – *tRNA Modification*).

The only identifiable dedicated DNA repair enzyme in *Ca.* Organicella was a RecA homolog. Depleted DNA repair abilities are typical of bacteria with highly reduced genomes and contribute to the accumulation of deleterious substitutions,

[3]https://www.ncbi.nlm.nih.gov/Taxonomy/Utils/wprintgc.cgi#SG4

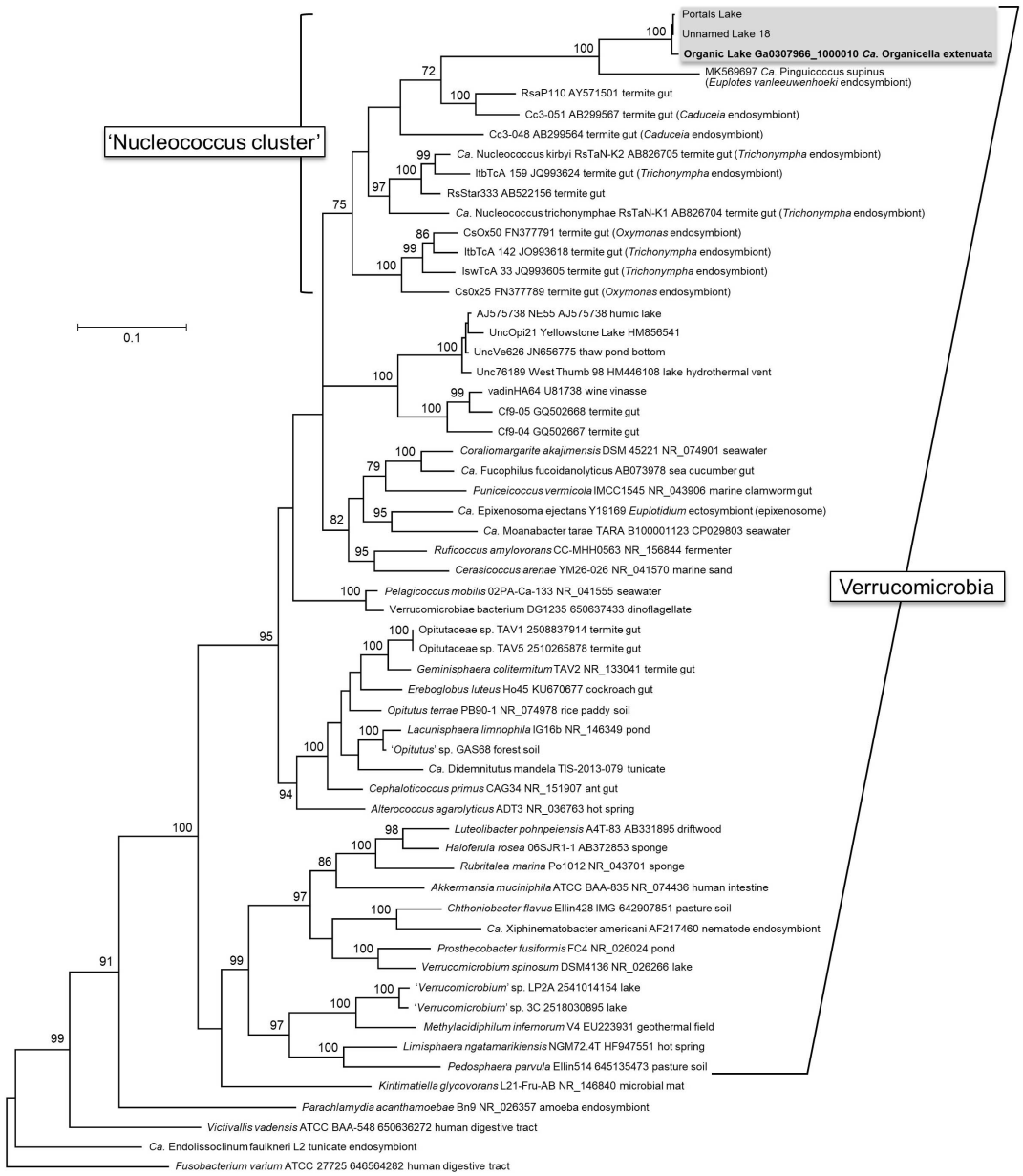

FIGURE 1 | Phylogeny of *Candidatus* Organicella extenuata. Phylogeny of Verrucomicrobia and related bacteria based on 16S rRNA sequences, showing *Ca.* Organicella extenuata nested inside the newly proposed "Nucleococcus cluster"; other than *Ca.* Organicella extenuata, this cluster comprises the cytoplasmic endosymbiont *Ca.* Pinguicoccus supinus from a freshwater ciliate and intranuclear endosymbionts of amitochondriate protists resident in the hindgut of termites. The maximum likelihood tree was constructed with 59 sequences, and positions with less than 80% site coverage were eliminated, resulting in 1,415 positions in the final dataset. Bootstrap values > 70 are shown next to individual nodes. *Fusobacterium varium* is the outgroup. Accessions are given as NCBI Nucleotide accessions or IMG Gene IDs: for *Ca.* Organicella extenuata, sequences were included for the original Organic Lake MAG (contig Ga0307966_1000010, bases 107297..108828), Unnamed Lake 18 (contig Ga0400283_000007, bases 52431..53966), and "Portals" Lake (contig Ga0400669_009478, bases 1..1071 and contig Ga0400669_039189, bases 1314..1821). Sequences identical to the 16S rRNA sequence from the original Organic Lake MAG were represented in metagenome data from 19 other Organic Lake metagenomes and also in Unnamed Lake 13 (**Supplementary Table 4**). Note that nine-digit accessions are IMG Gene IDs, and all others are NCBI Nucleotide accessions.

including in CDS (McCutcheon and Moran, 2012; Bennett and Moran, 2015). The average predicted pI of *Ca.* Organicella proteins was 9.2 (**Supplementary Table 3**). It has been proposed that high (alkaline) pI of the proteome of intracellular parasites and endosymbionts may result from the accumulation of mutations (Kiraga et al., 2007). However, not all *Ca.* Organicella proteins were predicted to have a high pI. Notably, the two most acidic proteins are ferredoxin (pI 4.1) and acyl carrier protein (ACP) (pI 4.2), both of which are naturally acidic proteins (Knaff and Hirasawa, 1991; McAllister et al., 2006). If high pI

does arise from high rates of mutation, the acidic pI of ferredoxin and ACP may be indicative of a strong positive selection to preserve function.

Another trait that is shared between *Ca.* Organicella and known bacterial symbionts with highly reduced genomes is the retention of chaperone proteins (GroES-GroEL; DnaK); these chaperone proteins are thought to ameliorate the adverse effects of accumulated deleterious substitutions on correct protein-folding (Moran, 1996; McCutcheon and Moran, 2012). The bacteria that synthesize these chaperones are therefore heat-sensitive, limiting the thermal tolerance of their hosts (Burke et al., 2010; Fan and Wernegreen, 2013; Moran and Bennett, 2014). Thermal instability would not be expected to be a problem for *Ca.* Organicella in Antarctica (Franzmann et al., 1987; Gibson, 1999; Yau et al., 2013). Proteins that are damaged and cannot be correctly re-folded could be degraded to peptides by the encoded ClpXP (Sabree et al., 2013), although the fate of the peptides is unclear in the absence of identifiable peptidases.

Ca. Organicella Environmental Distribution and Host

To examine the environmental distribution of *Ca.* Organicella, 337 Antarctic lake and marine metagenomes were analyzed, which encompass 77 different Antarctic aquatic locations, including a time (December 2006 to January 2015) and depth series of Organic Lake (**Supplementary Figure 1** and **Supplementary Table 5**). Sequence coverage of *Ca.* Organicella MAGs from Organic Lake was higher at depth in the lake and higher in winter compared with spring or summer (**Supplementary Table 5**). Although the highest abundance of *Ca.* Organicella was from Organic Lake (up to a median read depth of 71), read coverage showed *Ca.* Organicella was also present in seven other lakes in the Vestfold Hills (**Supplementary Figure 1**), including a complete MAG from a small pond ∼15 km away from Organic Lake ("Unnamed Lake 18"), which had a median read depth of 22 and coverage of the original *Ca.* Organicella MAG (Ga0307966_1000010) of 99.97% (**Supplementary Table 5**). The MAGs from Organic Lake (11 close to full length) had an ANI of ≥ 99.5%, with the ANI across all MAGs from Organic Lake, Unnamed Lake 18, Portals Lake, and Unnamed Lake 13, ≥98.1%. Outside of these *Ca.* Organicella MAGs and *Ca.* Pinguicoccus, the best BLAST matches to the *Ca.* Organicella 16S rRNA gene in NCBI-nr and IMG databases were ≤ 82%. This indicates that a single species of *Ca.* Organicella is present in the Vestfold Hills, with *Ca.* Pinguicoccus being the only similar species identifiable elsewhere in the world.

To identify the potential host(s) of *Ca.* Organicella, metagenomes were co-assembled using Metabat, generating a *Ca.* Organicella MAG (k141_311079) plus 188 potential host bins. The abundance of each bin was determined for each of the 29 metagenomes where *Ca.* Organicella was detected by mapping the metagenome reads to the bins, and the correlation of bin abundances was calculated using SparCC. The abundance of *Ca.* Organicella was highly positively correlated with bin81 ($r = 0.89$, $p = 0$), bin149 ($r = 0.95$, $p = 0$), and contig k141_859071 ($r = 0.85$, $p = 0$). The two bins and the contig were also highly positively correlated to each other ($r = 0.94 – 0.99$, $p = 0$). Bin81 (12,580 contigs) and bin149 (18 contigs) were dominated by sequences assigned to the ciliate *Euplotes* (Euplotidae, Spirotrichea, and Ciliophora), and the 8.1-kb contig, k141_859071 contained a 28S rRNA gene (4,455 bp), region of a 5.8S rRNA gene and 18S rRNA gene (1,895 bp) that matched to *Euplotes* (e.g., 28S rRNA, 84.2% identity to *Euplotes aediculatus* across 79% of query length). We infer that bins 81 and 149 plus the rRNA contig represent a MAG that pertains to a single OTU that we refer to as "*Euplotes* sp. AntOrgLke" (**Supplementary Table 6**). The *Euplotes* sp. AntOrgLke MAG (**Supplementary Dataset 1**) comprises 29.98 Mbp across 12,599 contigs (longest contig 19,935 bp, N50 = 2,645, L50 = 3,806, GC = 38.15%), with 6,451 proteins predicted against the TaxDB_uniclust90_2018_08 database (**Supplementary Dataset 2**) and 15,328 proteins predicted against the MERC_MMETSP_Uniclust50_profiles database (**Supplementary Dataset 3**). Of relevance, the abundance of the *Ca.* Organicella MAG was highly positively correlated with the *Euplotes* sp. AntOrgLke MAG ($r = 0.89$, $p = 0$) (**Figure 2**), consistent with this ciliate being the host. Moreover, contigs belonging to the *Euplotes* sp. AntOrgLke mitochondrial genome were also detected (**Supplementary Table 7**; **Supplementary Dataset 4**).

Euplotes sp. AntOrgLke had 97% 18S rRNA identity to *Euplotes* cf. *antarcticus* and *E. vanleeuwenhoeki*. Tree topology was consistent for all three RNA polymerase sequences (**Figure 3**) and 18S rRNA sequence (**Supplementary Figure 3**), and *Euplotes* sp. AntOrgLke seems to be a member of *Euplotes* Clade A (Syberg-Olsen et al., 2016; Boscaro et al., 2018; Serra et al., 2020). The AAI calculated from available *Euplotes* genomic data (six species, including *Euplotes* sp. AntOrgLke) ranged from 49 to 91%, with *Euplotes* sp. AntOrgLke sharing 53–57% with the other five species (**Supplementary Table 8**). Thus, our data indicate *Euplotes* sp. AntOrgLke is likely a novel Antarctic member of the genus *Euplotes*, and *Ca.* Organicella is a verrucomicrobial endosymbiont of a ciliate species known as *Ca.* Pinguicoccus (Serra et al., 2020). *E. vanleeuwenhoeki*, the host of *Ca.* Pinguicoccus, is a freshwater ciliate (Serra et al., 2020), whereas Organic Lake is hypersaline (Franzmann et al., 1987; Yau et al., 2013).

Euplotes is a speciose genus of motile, unicellular ciliate found in many aquatic environments (Boscaro et al., 2019), including Organic Lake, where it was previously detected based on SSU rRNA sequences (Yau et al., 2013). *Euplotes* species have a propensity to harbor one or multiple endosymbiotic bacteria, with at least six genera and 21 species known to date, all of which reside in the cytoplasm (Boscaro et al., 2019; Serra et al., 2020). The majority of reported *Euplotes* endosymbiont species belong to Proteobacteria and are predominantly members of Burkholderiaceae (e.g., *Polynucleobacter*) and the specialized intracellular clades Rickettsiales and Holosporales (Boscaro et al., 2019). The exception is *Ca.* Pinguicoccus, a member of Verrucomicrobia, and the sole known endosymbiont of *E. vanleeuwenhoeki* (Serra et al., 2020). In *E. vanleeuwenhoeki*, *Ca.* Pinguicoccus cells are located free in the cytoplasm and were frequently observed to be in contact with mitochondria and lipid droplets (Serra et al., 2020). The exact benefit of *Ca.* Pinguicoccus

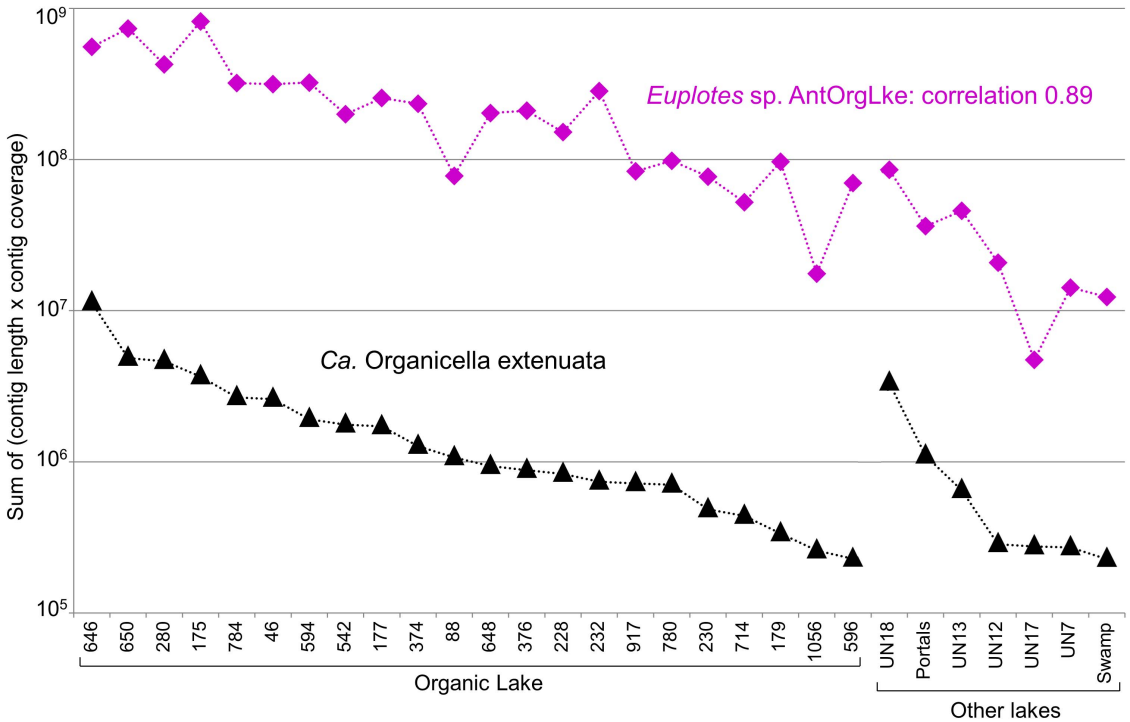

FIGURE 2 | Co-occurrence of *Candidatus* Organicella extenuata and *Euplotes* sp. AntOrgLke in Antarctic metagenomes. The abundance of *Ca.* Organicella extenuata (k141_311079) and *Euplotes* sp. AntOrgLke (bin81 + bin14 + contigk141_859071), calculated as the sum of (contig length × contig coverage) for all contigs, was analyzed using SparCC to determine their co-occurrence (*r,* correlation coefficient). Across 29 metagenomes in which *Ca.* Organicella extenuata was detected, the abundance of *Euplotes* sp. AntOrgLke strongly positively correlated with the abundance of *Ca.* Organicella extenuata ($r = 0.89$, $p = 0$), indicating *Euplotes* sp. AntOrgLke was likely the host of *Ca.* Organicella extenuata. None of the other 187 bins representing other potential hosts exhibited a positive correlation above $r = 0.54$. *X*-axis labels: Organic Lake, metagenome IDs (see **Supplementary Table 5**); Other lakes, lake names (Unnamed abbreviated as UN).

to its ciliate host is unclear, although it is unlikely to be nutritional (see *Ca. Organicella–Euplotes Interactions*, later) (Serra et al., 2020). Similarly, the foundations of the symbiotic relationship between proteobacterial endosymbionts and *Euplotes* are unclear, including those that are essential symbionts (*Polynucleobacter*, *Ca.* Protistobacter, and *Ca.* Devosia) and accessory symbionts, with the latter possibly being parasitic (Boscaro et al., 2013, 2019).

Ca. Organicella–*Euplotes* Interactions

One possibility is that *Ca.* Organicella provides Fe-S clusters and fatty acids to its host as the foundation for a mutualistic symbiosis (**Figure 4**). This is pertinent to *Euplotes*, in which, as in other ciliates, the mitochondrial genome does not encode these functions. We identified 41.8 kb of *Euplotes* sp. AntOrgLke mitochondrial genome sequence—a comparable length to the mitochondrial genome sequences reported for other *Euplotes* species (de Graaf et al., 2009; Serra et al., 2020). Like the mitochondrial genomes of *E. crassus, E. minuta*, and *E. vanleeuwenhoeki*, that of *Euplotes* sp. AntOrgLke has genes that encode electron transport chain proteins, ribosomal proteins, rRNA, tRNA, and a cytochrome *c* assembly protein, along with multiple genes that have no known function, but no identifiable Fe-S cluster or fatty acid synthesis genes (**Supplementary Table 7**; Pritchard et al., 1990; de Graaf et al., 2009; Swart et al., 2011; Johri et al., 2019; Serra et al., 2020). Within the genus *Euplotes,* the mitochondrial genetic code includes a single stop codon (UAA), a single unused codon (UAG), and tryptophan-encoding UGA (Pritchard et al., 1990; Burger et al., 2000; Brunk et al., 2003; de Graaf et al., 2009; Swart et al., 2011). By comparison, in *Ca.* Organicella, UGA is reassigned to Trp, whereas both UAA and UAG are stop codons.

The *Ca.* Organicella MAG encodes ferredoxin and sulfur utilization factor (SUF) proteins involved in Fe-S cluster biogenesis (SufCBD, SufU, and SufT), including cysteine desulfurase (SufS) for the mobilization of sulfur from cysteine (Selbach et al., 2014; **Supplementary Table 3**). In eukaryotes, the iron–sulfur cluster (ISC) and SUF pathways are the dominant Fe-S cluster synthesis pathways, with ISC assembly proteins located in the mitochondria, whereas SUF assembly proteins are localized to plastid organelles (Kispal et al., 1999; Tsaousis, 2019), the latter including chloroplasts and apicoplasts (Takahashi et al., 1986; Lill and Mühlenhoff, 2005; Lim and McFadden, 2010; Gisselberg et al., 2013), although, in certain protists, SUF assembly proteins are located in the cytoplasm (Tsaousis et al., 2012; Karnkowska et al., 2016). Typical of eukaryotes, *Euplotes* sp. AntOrgLke encodes homologs of ISC proteins inferred to be present in the model ciliate *Tetrahymena thermophila*, including cysteine desulfurase (Nfs1), ferredoxin (Yah1), and ferredoxin reductase (Arh1) (**Supplementary Table 9**); ISC assembly would occur in the mitochondrion and depend on nuclear-encoded enzymes

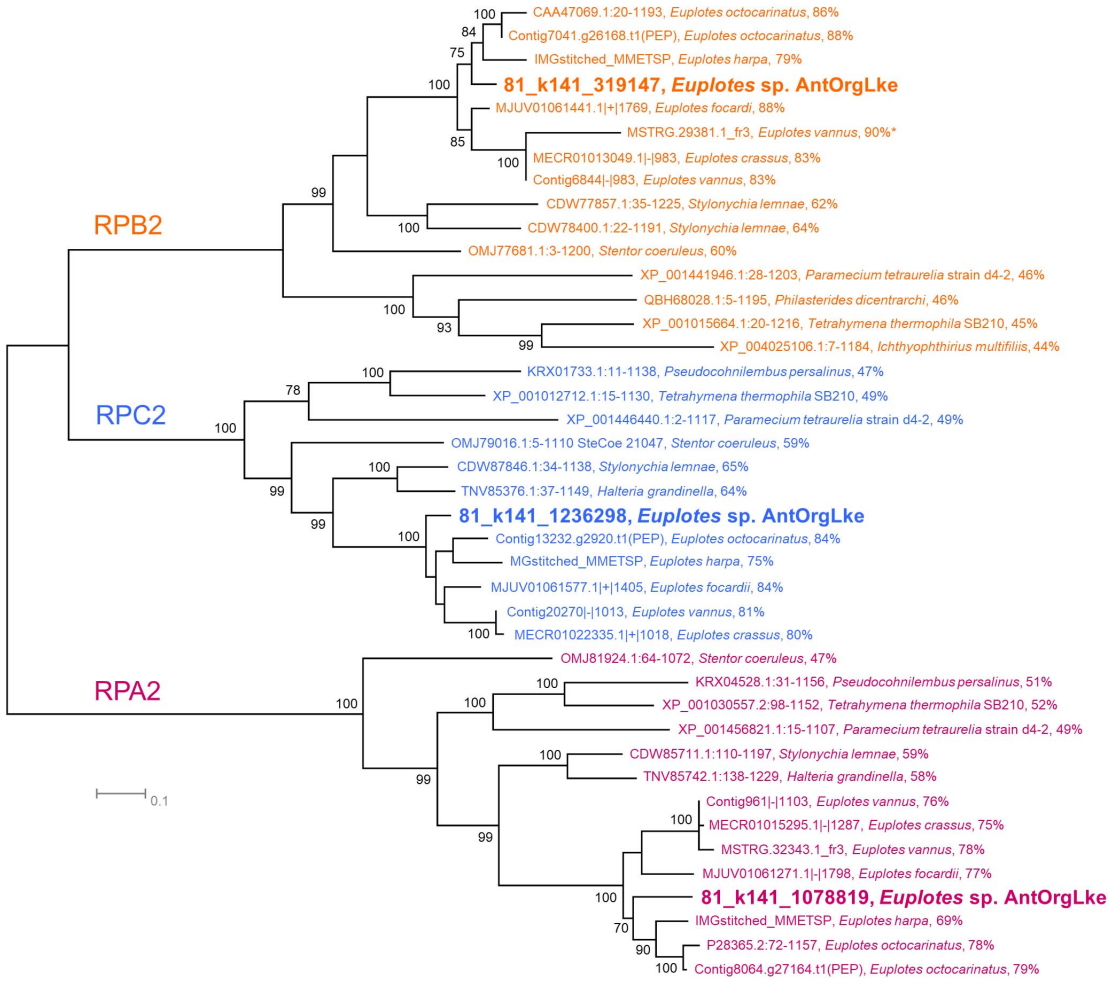

FIGURE 3 | Phylogeny of *Euplotes* sp. AntOrgLke. Unrooted maximum likelihood phylogeny of RNA polymerase subunit II proteins from members of Ciliophora showing *Euplotes* sp. AntOrgLke clustering with members of the *Euplotes* genus. Within the cluster for each RNA polymerase type (RPB, RPC, and RPA), the percent identity between *Euplotes* sp. AntOrgLke protein and an individual protein is shown after the species name. Bootstrap values ≥ 70 are shown next to individual nodes, and protein sequences are available in **Supplementary Dataset 5**. A total of 41 RNA polymerase subunit II amino acid sequences were used in analysis. Positions with less than 80% site coverage were eliminated, and 944 positions remained in the final dataset, with the exception of MSTRG.29381.1_fr3, *Euplotes vannus* which was a partial sequence (283 aa) and is marked with an*.

(Smith et al., 2007). The SUF system of *Ca.* Organicella could therefore function as a complementary Fe-S cluster assembly system to ISC. The SUF system is more resistant to reactive oxygen species than the ISC system (Santos-Garcia et al., 2014); thus, the SUF system encoded by *Ca.* Organicella may be especially important to the host under oxidative stress conditions in response to the degradation of Fe-S clusters of host proteins (Tsaousis, 2019). The SUF system may be especially relevant to *Euplotes* sp. AntOrgLke in Organic Lake and the other Vestfold Hills lakes due to the prevailing environmental conditions (high oxygen concentrations; freezing temperatures; enhanced UV irradiation; **Supplementary Figure 1**) that promote the production of reactive oxygen species (Ricci et al., 2017).

Ca. Organicella also encodes an almost complete suite of genes for bacterial type II fatty acid synthesis (FASII), except for FabH, an enzyme involved in fatty acid elongation (see ***Supplementary Text*** – *Pyruvate Oxidation and Fatty Acid Synthesis*). It is likely that another condensing enzyme involved in fatty acid elongation encoded in *Ca.* Organicella (FabB or FabF) would substitute for FabH, as inferred for *Ca.* Wigglesworthia, which similarly lacks FabH but otherwise encodes a complete FASII pathway (Zientz et al., 2004; Parsons and Rock, 2013). In support of this, *Escherichia coli* and *Lactococcus lactis* mutants that lack *fabH* are still capable of synthesizing fatty acids (Morgan-Kiss and Cronan, 2008; Yao et al., 2012). For *Ca.* Organicella, the acetyl-CoA precursor for straight-chain fatty acid biosynthesis would be generated using a pyruvate dehydrogenase (PDH) complex, presumably using pyruvate acquired from the host (**Figure 4**).

Many protists depend on the fatty acids provided by mitochondrial FASII for processes such as lipoylation of essential enzymes or for incorporation into phospholipids; despite having their own cytoplasmic FAS apparatus (FAS type I), these

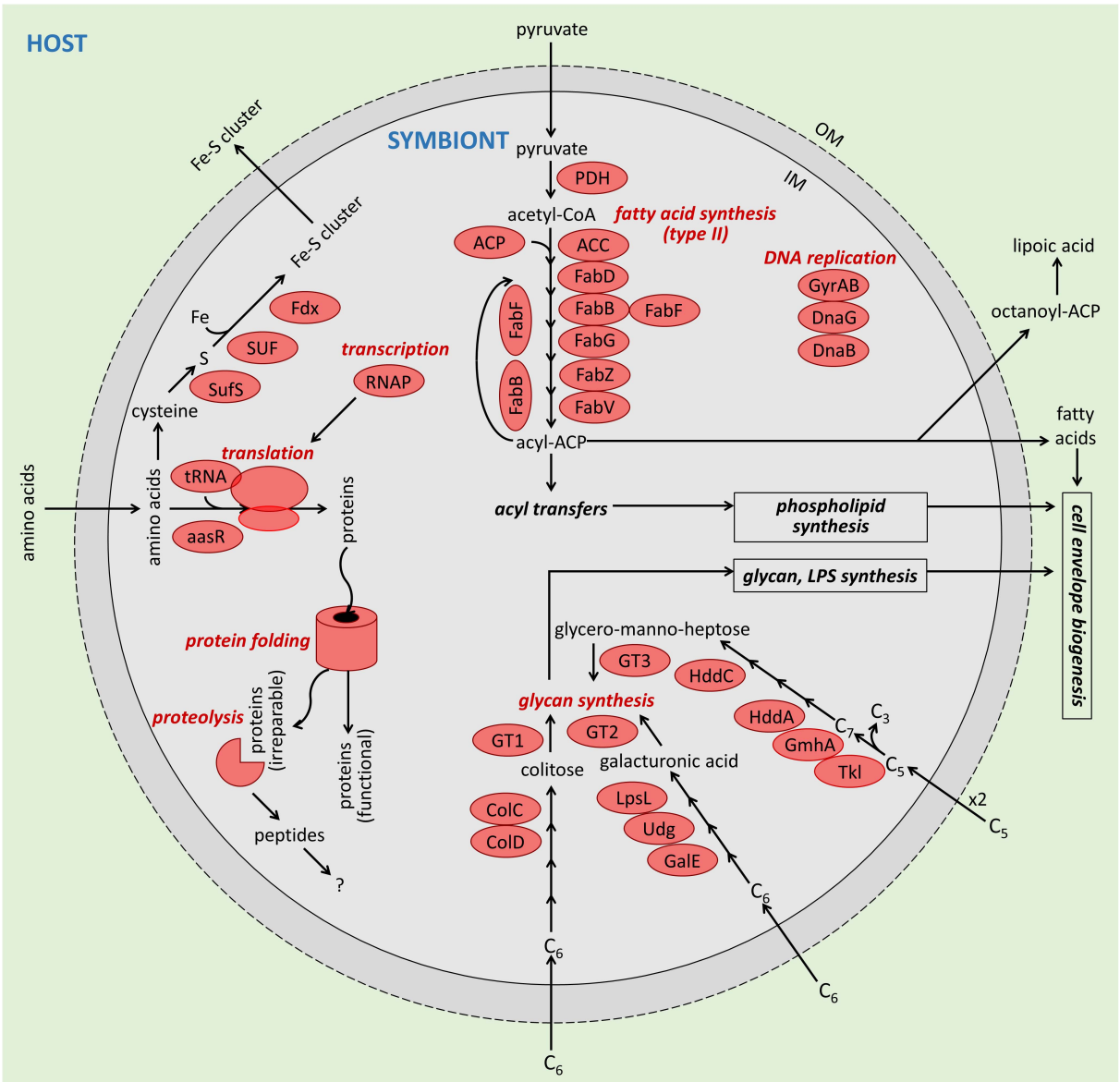

FIGURE 4 | Depiction of function of *Candidatus* Organicella extenuata within *Euplotes* sp. AntOrgLke. The potential metabolic capacities of *Ca.* Organicella extenuata are limited to pyruvate oxidation; type II fatty acid biosynthesis; iron-sulfur (Fe-S) cluster assembly; and conversion, activation, and transfer of hexose and heptose sugars. The *Ca.* Organicella extenuata MAG lacks any identifiable genes for glycolysis, tricarboxylic acid cycle (aside from pyruvate oxidation), pentose phosphate pathway, respiration, fermentation, ATP generation (either by oxidative phosphorylation and ATP synthase, or substrate-level phosphorylation), or synthesis of phospholipids (aside from fatty acids), amino acids, nucleic acids, or vitamins. There were no identifiable transporter genes. Processes, pathways, and enzymes that were inferred to be functional in *Ca.* Organicella extenuata are shaded in red. OM, outer membrane; IM, inner membrane; LPS, lipopolysaccharide. Fatty acid synthesis (type II): PDH, pyruvate dehydrogenase; CoA, coenzyme A; ACP, acyl carrier protein; ACC, acetyl-CoA carboxylase complex; FabD, malonyl-CoA-ACP-transacylase; FabB and FabF, 3-oxoacyl-ACP synthase; FabG, 3-oxoacyl-ACP reductase; FabZ, 3-hydroxyacyl-ACP dehydratase; FabV, enoyl-ACP reductase. Glycan synthesis: GT, glycosyltransferase (three different GT); ColD, GDP-4-keto-6-deoxy-D-mannose 3-dehydratase; ColC, GDP-L-colitose synthase; Udg, UDP-glucose 6-dehydrogenase; LpsL, UDP-glucuronate epimerase; TkI, transketolase; GmhA, phosphoheptose isomerase; HddA, D-glycero-α-D-manno-heptose 7-phosphate kinase; HddC, D-glycero-α-D-manno-heptose 1-phosphate guanylyltransferase. DNA replication: GyrAB, DNA gyrase; DnaG, DNA primase; DnaB, replicative DNA helicase. Transcription: RNAP, RNA polymerase. Translation: tRNA, transfer RNA; aaRS, aminoacyl tRNA synthetases. Fe-S cluster assembly: Fdx, ferredoxin; SufS, cysteine desulfurase; SUF, Fe-S cluster assembly complex (SufCBD, SufU, SufT). In this reconstruction, pyruvate is supplied by the host and the sole purpose of PDH is to provide the acetyl-CoA precursor for fatty acid synthesis. The fatty acid synthesis pathway is functionally complete in *Ca.* Organicella extenuata, with FabF or FabB substituting for missing FabH. There are three pathways involved in synthesis of heptose (glycero-manno-heptose) or hexose (galacturonic acid; colitose) subunits of lipopolysaccharide glycans in *Ca.* Organicella extenuata, but none of these pathways are complete, and all depend on exogenous precursors.

eukaryotes depend on the fatty acids provided by mitochondria (Stephens et al., 2007; Hiltunen et al., 2009). However, as in other ciliates, the mitochondrial genome of *Euplotes* lacks genes associated with FASII (Pritchard et al., 1990; Burger et al., 2000; Brunk et al., 2003; Swart et al., 2011; Johri et al., 2019). Thus, we propose the hypothesis that *Ca.* Organicella provides fatty acids to the host for these essential purposes.

Another possibility is that fatty acids are supplied to the host in a nutritional capacity. For example, there is evidence that fatty acids synthesized by *Ca.* Blochmannia floridanus form part of the symbiont's nutritional support to its host (carpenter ant *Camponotus chromaiodes*) during periods when the insect host is feeding on sugar-rich exudates (Zientz et al., 2004; Fan and Wernegreen, 2013). However, we regard this as unlikely, given that it has been predicted that nutritional symbioses are not likely to be necessary for heterotrophic algal and bacterial feeders such as *Euplotes* that can probably obtain all their required nutrients from their diets (Boscaro et al., 2013, 2019; Serra et al., 2020).

It is also possible that FASII in *Ca.* Organicella contributes to its own cellular requirements, including lipoylation of its own PDH and providing precursors for modification of its own cell envelope (**Figure 4**). In addition to encoding a functionally complete FASII pathway, 17 *Ca.* Organicella genes are predicted to be involved in the biosynthesis of precursors for lipopolysaccharide components: nine proteins are implicated in the biosynthesis of heptose and hexose units (although we could not reconstruct complete pathways), and eight proteins are glycosyltransferases that may be involved in the transfer of nucleotide-activated sugars to construct glycan chains (**Supplementary Table 3**; ***Supplementary Text*** – *Glycan Synthesis*). Obligate endosymbionts with genomes < 500 kbp typically have few if any genes for cell envelope biogenesis, with these pathways being especially prone to loss (McCutcheon and Moran, 2012; Moran and Bennett, 2014; Brown et al., 2015). *Ca.* Organicella lacks acyltransferases necessary for transferring acyl-ACP to glycerol-3-phosphate to produce phosphatidic acid, the phospholipid precursor in bacteria (Yao et al., 2012), and there are no identifiable genes for the biosynthesis of the glycerophosphate backbone or headgroups of phospholipids or for the 3-deoxy-D-manno-octulosonic acid-lipid A (Kdo_2-lipid A) precursor of lipopolysaccharides (Wang et al., 2015).

Thus, *Ca.* Organicella, as in other endosymbionts with highly reduced genomes, is assumed to rely entirely on host-derived membranes (Baumann, 2005; McCutcheon and Moran, 2012; Husnik and McCutcheon, 2016). The presence of lipopolysaccharide- and other cell-wall-related genes is not unusual for symbiotic bacteria with larger genomes (Zientz et al., 2004; Nikoh et al., 2011); for example, the insect endosymbionts *Ca.* Wigglesworthia and *Ca.* Blochmannia (both between 615 and 706 kbp) encode the majority of genes necessary for the synthesis of a normal gram-negative cell wall, including phospholipids and lipopolysaccharides (Akman et al., 2002; Gil et al., 2003; Zientz et al., 2004). Additionally, certain obligately symbiotic bacteria with larger genomes (>600 kb) retain a complete set of FASII genes (Akman et al., 2002; Gil et al., 2003; Nikoh et al., 2011; Lamelas et al., 2011; Chong and Moran, 2018). In these symbionts, the retention of genes necessary for the synthesis of a normal Gram-negative cell wall (including lipopolysaccharides) is possibly for protection against the host and/or reflects a relatively recent symbiotic association (Akman et al., 2002; Gil et al., 2003). The latter does not apply to *Ca.* Organicella, with the extreme reduction in genome size reflecting an ancient symbiosis (Serra et al., 2020).

Nevertheless, *Ca.* Organicella might contribute glycan components to its own cell envelope (including lipopolysaccharides). One possibility is that modifications of the endosymbiont cell wall confer some protection against the host, such as through variation of fatty acid length or altering the glycan moieties of lipopolysaccharides (core and/or O-specific polysaccharides) using modified sugars by the action of glycosyltransferases (Serra et al., 2020). *Ca.* Pinguicoccus has a very similar genome size and gene composition as *Ca.* Organicella, including retaining homologs of the same glycan/lipopolysaccharide-related genes (see *Comparison of Ca. Organicella and Ca. Pinguicoccus Genomes*, later). *Ca.* Pinguicoccus resides free in the cytoplasm of *E. vanleeuwenhoeki*, and it has been proposed that endosymbionts in the host cytoplasm of eukaryote cells face a less stable and possibly hostile environment compared with those endosymbionts that are enclosed within specialized bacteriocytes or host-derived vesicles (Gil et al., 2003; Wu et al., 2004; Serra et al., 2020). For this reason, *Ca.* Pinguicoccus may exercise some control over the composition of its cell envelope because it is in direct contact with the host cytoplasm (Serra et al., 2020). This might also be true of *Ca.* Organicella, which, based on its close phylogenetic affiliation with *Ca.* Pinguicoccus and having *Euplotes* as the putative host, likely lives in the host cytoplasm.

Comparison of *Ca.* Organicella and *Ca.* Pinguicoccus Genomes

The genome sizes of *Ca.* Organicella (158,228 bp, 194 genes, 163 CDS) and *Ca.* Pinguicoccus (163,218 bp, 205 genes, 168 CDS; Serra et al., 2020) are similar; note that the protein sequences for the *Ca.* Pinguicoccus NCBI (Accession CP039370) genome sequence were auto-predicted with genetic code 11, but using genetic code 4, we predict a total of 200 genes [five less than reported in Serra et al. (2020)], consisting of 163 CDS, 34 tRNAs, and the 16S, 5S, and 23S rRNA genes (**Supplementary Table 1**). The two genomes share extensive synteny (**Supplementary Figure 4**). Although the genomic nucleotide sequences were too divergent to calculate ANI, the AAI between the two symbiont genomes was determined to be 46% (two-way AAI based on 134 proteins, all predicted with genetic code 4). Both genomes retain an almost identical small subset of genes represented across Verrucomicrobia (Serra et al., 2020; **Supplementary Table 3**). They also share homologous proteins required for DNA replication, transcription, and translation, in common with other endosymbionts, but both lack the catalytic subunit of DNA polymerase (DnaE), which is exceptional among endosymbionts (Serra et al., 2020).

Ca. Pinguicoccus encodes the same components of the SUF system and a functionally complete FASII pathway as *Ca.* Organicella, suggesting that *Ca.* Pinguicoccus confers the same benefits to its *Euplotes* host that we infer for *Ca.* Organicella. Of interest is that *Ca.* Pinguicoccus cells were often observed associated with lipid droplets in *E. vanleeuwenhoeki* cytoplasm, raising the possibility of a link between the retention of FASII genes and interaction with the host's lipids (Serra et al., 2020). *Ca.* Pinguicoccus also encodes homologs of the same glycosyltransferases and heptose- and hexose-related enzymes encoded in *Ca.* Organicella (**Supplementary Table 3**). Nevertheless, *Ca.* Pinguicoccus retains a putative phospholipid synthesis protein (CDP-diacylglycerol-glycerol-3-phosphate 3-phosphatidyltransferase homolog) not identifiable in *Ca.* Organicella. *Ca.* Pinguicoccus encodes a thioredoxin–thioredoxin reductase system (for maintaining thiol-disulfide redox balance) and NADP-dependent glutamate dehydrogenase (for the reversible oxidative deamination of glutamate), neither of which are identifiable in *Ca.* Organicella. There are also variations between the two genera in the exact complement of ribosomal subunits, aaRS, and initiation factor subunits (**Supplementary Table 3**), with these components being prone to loss among endosymbionts (Moran and Bennett, 2014). However, *Ca.* Organicella and *Ca.* Pinguicoccus possess the same 34 tRNA genes.

Overall, the data suggest that, as their divergence from a common ancestor had a highly reduced genome, further genomic erosion has occurred independently in *Ca.* Organicella and *Ca.* Pinguicoccus, with differential loss of certain genes, especially those involved in translation. By contrast, SUF, FASII, and certain lipopolysaccharide/glycan-related genes are conserved between the two genera. This suggests that these particular genes play important roles in the interactions of these endosymbionts with their ciliate hosts.

CONCLUSION

Many of the smallest bacterial genomes are from insect symbionts that exist as metabolically complementary partnerships within the host (43) (**Supplementary Table 1**). For example, *Ca.* Nasuia deltocephalinicola (∼112 kbp) and *Ca.* Hodgkinia cicadicola (∼144 kbp) are each co-resident with *Ca.* Sulcia (Bennett and Moran, 2013; McCutcheon et al., 2009), whereas *Ca.* Tremblaya princeps (∼139 kbp) contains *Ca.* Moranella endobia to constitute a nested symbiosis (McCutcheon and von Dohlen, 2011). However, *Ca.* Carsonella ruddii (∼160 kbp) is a lone endosymbiont resident in sap-feeding psyllids (Thao et al., 2000; Nakabachi et al., 2006). Unlike known insect symbionts with highly reduced genomes (Nakabachi and Ishikawa, 1999; Zientz et al., 2004; Nakabachi et al., 2006; Pérez-Brocal et al., 2006; Bennett and Moran, 2013; Brown et al., 2015; Gil et al., 2018), *Ca.* Organicella and *Ca.* Pinguicoccus lack any capacity for the biosynthesis of amino acids or vitamins (Serra et al., 2020). Thus, as with the Ca. Pinguicoccus-*E. vanleeuwenhoeki* symbiosis, there is no reason to assume that *Ca.* Organicella exists as part of a co-symbiotic partnership, especially considering that none of the genes encode enzymes involved in amino acid or vitamin biosynthesis, as is typical for such consortia (McCutcheon et al., 2009; McCutcheon and von Dohlen, 2011). The absence of a nutritional basis of a proposed *Ca.* Organicella-*Euplotes* symbiosis likely reflects the algivorous and bacterivorous nature of the ciliate host (Serra et al., 2020), in contrast to insects with specialized and nutritionally unbalanced diets. Instead, we propose that *Ca.* Organicella and *Ca.* Pinguicoccus provide SUF Fe-S clusters and FASII fatty acids as essential molecules to the host, with FASII replacing a lost mitochondrial function in *Euplotes*. Additionally, the ciliate host would possess dual Fe-S cluster biogenesis systems, with the SUF system provided by endosymbiotic Verrucomicrobia.

Ca. Organicella was identified as possibly being an endosymbiont by virtue of having unusual coding parameters (Ivanova et al., 2014) rather than by searching our metagenome data for symbionts. Previous metagenomic screening of multiple *Euplotes* strains and their resident bacteria did not recover symbionts that belong to phylum Verrucomicrobia (Boscaro et al., 2019). In the study, the identification of putative symbionts in *Euplotes*-based metagenome data was based on bacterial taxa that were referrable to known clades of exclusively intracellular bacteria (e.g., Rickettsiales) or related to previously described protist symbionts (e.g., *Polynucleobacter*) (Boscaro et al., 2019); thus, any verrucomicrobial symbionts might have been overlooked, especially if they were present at low coverage. Targeted hosts and/or symbiont reference genomes have been used extensively for identifying both ecto- and endo-symbionts of a broad range of taxa, including magnetotactic bacteria of marine protists (Monteil et al., 2019), gut symbionts of hadal snailfish (Lian et al., 2020) and phytophagous stink bugs (Kashkouli et al., 2020), and symbionts of pea aphids (Guyomar et al., 2018) and scleractinian corals (Shinzato et al., 2014). The discovery of *Ca.* Pinguicoccus arose through the development of a "next-generation taxonomy" approach for assessing symbiont–host associations that combines "bio-taxonomy tools, classical morphology, ultrastructure, molecular phylogeny, genomics, and bioinformatics" (Serra et al., 2020). The study focused on *Euplotes* as a model protist "holobiont," in the process of identifying *Ca.* Pinguicoccus. Being a host-based approach to endosymbiont discovery, the "next-generation taxonomy" approach is applicable to known symbiotic communities. Software (MinYS) has also recently been reported to specifically identify symbionts from genome assemblies of symbiotic communities by using reference genomes (Guyomar et al., 2020). Although genetic code 4 (UGA stop codons recoded to tryptophan) has been reported to only occur rarely (Yamao et al., 1985; Knight et al., 2001; McCutcheon et al., 2009), our findings raise the enticing prospect that searching contigs and MAGs for this recoding may reveal new symbionts, including members of the verrucomicrobial "Nucleococcus cluster" (**Figure 1**).

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The repository and accession numbers are: NCBI

(*Euplotes* sp. AntOrgLke MAG, accession: JAGXKF000000000 and PRJNA720161); IMG (https://img.jgi.doe.gov/) (*Ca.* Organicella extenuata MAG: Scaffold Ga0307966_1000010; *Euplotes* sp. AntOrgLke mitochondrial genome: 2 Scaffolds, Ga0307966_1001133 and Ga0307966_1001206).

ETHICS STATEMENT

Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

TW, MA, NI, MH, and RC conceived the study, analyzed the data, and conducted data interpretation. SH performed Sterivex filter DNA extractions. AH and SB spent 18 months in Antarctica running the 2013–2015 expedition. TW, MA, and RC wrote the manuscript with input from all other co-authors. All authors have read and approved the manuscript submission.

FUNDING

This work was supported by the Australian Research Council (DP150100244) and the Australian Antarctic Science Program (project 4031). The work conducted by the U.S. Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, was supported by the Office of Science of the U.S. Department of Energy under contract no. DE-AC02-05CH11231.

ACKNOWLEDGMENTS

Computational analyses at UNSW Sydney were performed on the computational cluster Katana, supported by Research Technology Services at UNSW Sydney.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.674758/full#supplementary-material

REFERENCES

Akman, L., Yamashita, A., Watanabe, H., Oshima, K., Shiba, T., Hattori, M., et al. (2002). Genome sequence of the endocellular obligate symbiont of tsetse flies. *Wigglesworthia glossinidia*. *Nat. Genet.* 32, 402–407. doi: 10.1038/ng986

Allen, M. A., Lauro, F. M., Williams, T. J., Burg, D., Siddiqui, K. S., DeFrancisci, D., et al. (2009). The genome sequence of the psychrophilic archaeon, *Methanococcoides burtonii*: the role of genome evolution in cold adaptation. *ISME J.* 3, 1012–1035. doi: 10.1038/ismej.2009.45

Baumann, P. (2005). Biology of bacteriocyte-associated endosymbionts of plant sap-sucking insects. *Annu. Rev. Microbiol.* 59, 155–189. doi: 10.1146/annurev.micro.59.030804.121041

Bennett, G. M., McCutcheon, J. P., MacDonald, B. R., Romanovicz, D., and Moran, N. A. (2014). Differential genome evolution between companion symbionts in an insect-bacterial symbiosis. *mBio* 5:e01697-14.

Bennett, G. M., and Moran, N. A. (2013). Small, smaller, smallest: the origins and evolution of ancient dual symbioses in a phloem-feeding insect. *Genome Biol. Evol.* 5, 1675–1688. doi: 10.1093/gbe/evt118

Bennett, G. M., and Moran, N. A. (2015). Heritable symbiosis: the advantages and perils of an evolutionary rabbit hole. *Proc. Natl. Acad. Sci. U.S.A.* 112, 10169–10176. doi: 10.1073/pnas.1421388112

Boscaro, V., Felletti, M., Vannini, C., Ackerman, M. S., Chain, P. S., Malfatti, S., et al. (2013). *Polynucleobacter necessarius*, a model for genome reduction in both free-living and symbiotic bacteria. *Proc Natl Acad Sci U.S.A* 110, 18590–18595. doi: 10.1073/pnas.1316687110

Boscaro, V., Husnik, F., Vannini, C., and Keeling, P. J. (2019). Symbionts of the ciliate *Euplotes*: diversity, patterns and potential as models for bacteria-eukaryote endosymbioses. *Proc. R. Soc. Lond. B* 286:20190693. doi: 10.1098/rspb.2019.0693

Boscaro, V., Syberg-Olsen, M. J., Irwin, N. A. T., del Campo, J., and Keeling, P. J. (2018). What can environmental sequences tell us about the distribution of low-rank taxa? The case of *Euplotes* (Ciliophora, Spirotrichea), including a description of *Euplotes enigma* sp. nov. J. *Eukaryot. Microbiol.* 66, 281–293. doi: 10.1111/jeu.12669

Brown, A. M., Howe, D. K., Wasala, S. K., Peetz, A. B., Zasada, I. A., and Denver, D. R. (2015). Comparative genomics of a plant-parasitic nematode endosymbiont suggest a role in nutritional symbiosis. *Genome Biol. Evol.* 7, 2727–2746. doi: 10.1093/gbe/evv176

Brunk, C. F., Lee, L. C., Tran, A. B., and Li, J. (2003). Complete sequence of Mt genome of *Tetrahymena thermophila* and comparative methods for identifying highly divergent genes. *Nucl. Acid Res.* 31, 1673–1682. doi: 10.1093/nar/gkg270

Burger, G., Zhu, Y., Littlejohn, T. G., Greenwood, S. J., Schnare, M. N., Lang, B. F., et al. (2000). Complete sequence of the mitochondrial genome of *Tetrahymena pyriformis* and comparison with *Paramecium aurelia* mitochondrial DNA. *J. Mol. Biol.* 297, 365–380. doi: 10.1006/jmbi.2000.3529

Burke, G., Fiehn, O., and Moran, N. (2010). Effects of facultative symbionts and heat stress on the metabolome of pea aphids. *ISME J.* 4, 242–252. doi: 10.1038/ismej.2009.114

Bushnell, B. (2014). *BBMap: A Fast, Accurate, Splice-Aware Aligner. Lawrence Berkeley National Laboratory*. LBNL Report #: LBNL-7065E. Berkeley, CA: LBNL.

Cavicchioli, R. (2015). Microbial ecology of Antarctic aquatic systems. *Nat. Rev. Microbiol.* 13, 691–706. doi: 10.1038/nrmicro3549

Chen, X., Jiang, Y., Gao, F., Zheng, W., Krock, T. J., Stover, N. A., et al. (2019). Genome analyses of the new model protist *Euplotes vannus* focusing on genome rearrangement and resistance to environmental stressors. *Mol. Ecol. Resour.* 19, 1292–1308. doi: 10.1111/1755-0998.13023

Chong, R. A., and Moran, N. A. (2018). Evolutionary loss and replacement of *Buchnera*, the obligate endosymbiont of aphids. *ISME J.* 12, 898–908. doi: 10.1038/s41396-017-0024-6

Choo, Y. J., Lee, K., Song, J., and Cho, J. C. (2007). *Puniceicoccus vermicola* gen. nov., sp. nov., a novel marine bacterium, and description of Puniceicoccaceae fam. nov., Puniceicoccales ord. nov., Opitutaceae fam. nov., Opitutales ord. nov. and Opitutae classis nov. in the phylum 'Verrucomicrobia'. *Int. J. Syst. Evol. Microbiol.* 57, 532–537. doi: 10.1099/ijs.0.64616-0

Darling, A. C., Mau, B., Blattner, F. R., and Perna, N. T. (2004). Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res.* 14, 1394–1403. doi: 10.1101/gr.2289704

de Graaf, R. M., van Alen, T. A., Dutilh, B. E., Kuiper, J. W., van Zoggel, H. J., Huynh, M. B., et al. (2009). The mitochondrial genomes of the ciliates *Euplotes minuta* and *Euplotes crassus*. *BMC Genomics* 10:514. doi: 10.1186/1471-2164-10-514

Derrien, M., Vaughan, E. E., Plugge, C. M., and de Vos, W. M. (2004). *Akkermansia muciniphila* gen. nov., sp. nov., a human intestinal mucin-degrading bacterium. *Int. J. Syst. Evol. Microbiol.* 54, 1469–1476. doi: 10.1099/ijs.0.02873-0

Dunfield, P. F., Yuryev, A., Senin, P., Smirnova, A. V., Stott, M. B., Hou, S., et al. (2007). Methane oxidation by an extremely acidophilic bacterium of the phylum Verrucomicrobia. *Nature* 450, 879–882. doi: 10.1038/nature06411

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi: 10.1093/nar/gkh340

Fan, Y., and Wernegreen, J. J. (2013). Can't take the heat: high temperature depletes bacterial endosymbionts of ants. *Microb. Ecol.* 66, 727–733. doi: 10.1007/s00248-013-0264-6

Franzmann, P. D., Deprez, P. P., Burton, H. R., and van den Hoff, J. (1987). Limnology of Organic Lake, Antarctica, a meromictic lake that contains high concentrations of dimethyl sulfide. *Aust. J. Mar. Freshw. Res.* 38, 409–417. doi: 10.1071/mf9870409

Friedman, J., and Alm, E. J. (2012). Inferring correlation networks from genomic survey data. *PLoS Comput. Biol.* 8:e1002687. doi: 10.1371/journal.pcbi.1002687

Gibson, J. A. E. (1999). The meromictic lakes and stratified marine basins of the Vestfold Hills, East Antarctica. *Antarct. Sci.* 11, 175–192. doi: 10.1017/s0954102099000243

Gibson, J. A. E., Garrick, R. C., Franzmann, P. D., Deprez, P. P., and Burton, H. (1991). Reduced sulfur gases in saline lakes of the Vestfold Hills, Antarctica. *Palaeogeo Palaeoclimatol. Palaeoecol.* 84, 131–140. doi: 10.1016/0031-0182(91)90040-x

Gil, R., Silva, F. J., Zientz, E., Delmotte, F., Gonzalez-Candelas, F., Latorre, A., et al. (2003). The genome sequence of *Blochmannia floridanus*: comparative analysis of reduced genomes. *Proc. Natl. Acad. Sci. U.S.A.* 100, 9388–9393. doi: 10.1073/pnas.1533499100

Gil, R., Vargas-Chavez, C., López-Madrigal, S., Santos-García, D., Latorre, A., and Moya, A. (2018). *Tremblaya phenacola* PPER: an evolutionary beta-gammaproteobacterium collage. *ISME J.* 12, 124–135. doi: 10.1038/ismej.2017.144

Gisselberg, J. E., Dellibovi-Ragheb, T. A., Matthews, K. A., Bosch, G., and Prigge, S. T. (2013). The suf iron-sulfur cluster synthesis pathway is required for apicoplast maintenance in malaria parasites. *PLoS Pathog.* 9:e1003655. doi: 10.1371/journal.ppat.1003655

Guyomar, C., Delage, W., Legeai, F., Mougel, C., Simon, J.-C., and Lemaitre, C. (2020). MinYS: mine your symbiont by targeted genome assembly in symbiotic communities. *NAR Genom. Bioinform.* 2:lqaa047.

Guyomar, C., Legeai, F., Jousselin, E., Mougel, C., Lemaitre, C., and Simon, J. C. (2018). Multi-scale characterization of symbiont diversity in the pea aphid complex through metagenomic approaches. *Microbiome* 6:181.

Hiltunen, J. K., Schonauer, M. S., Autio, K. J., Mittelmeier, T. M., Kastaniotis, A. J., and Dieckmann, C. L. (2009). Mitochondrial fatty acid synthesis type II: more than just fatty acids. *J. Biol. Chem.* 284, 9011–9015. doi: 10.1074/jbc.r800068200

Hongoh, Y., Sato, T., Dolan, M. F., Noda, S., Ui, S., Kudo, T., et al. (2007). The motility symbiont of the termite gut flagellate *Caduceia versatilis* is a member of the "*Synergistes*" group. *Environ. Microbiol.* 73, 6270–6276. doi: 10.1128/aem.00750-07

Hongoh, Y., Sharma, V. K., Prakash, T., Noda, S., Taylor, T. D., Kudo, T., et al. (2008a). Complete genome of the uncultured Termite Group 1 bacteria in a single host protist cell. *Proc. Natl. Acad. Sci. U.S.A.* 105, 5555–5560. doi: 10.1073/pnas.0801389105

Hongoh, Y., Sharma, V. K., Prakash, T., Noda, S., Toh, H., Taylor, T. D., et al. (2008b). Genome of an endosymbiont coupling N2 fixation to cellulolysis within protist cells in termite gut. *Science* 322, 1108–1109. doi: 10.1126/science.1165578

Huntemann, M., Ivanova, N. N., Mavromatis, K., Tripp, H. J., Paez-Espino, D., Tennessen, K., et al. (2015). The standard operating procedure of the DOE-JGI Metagenome Annotation Pipeline (MAP v.4). *Stand. Genomic Sci.* 11:17.

Husnik, F., and McCutcheon, J. P. (2016). Repeated replacement of an intrabacterial symbiont in the tripartite nested mealybug symbiosis. *Proc. Natl. Acad. Sci. U.S.A.* 113, E5416–E5424.

Ikeda-Ohtsubo, W., Faivre, N., and Brune, A. (2010). Putatively free-living "*Endomicrobia*" – ancestors of the intracellular symbionts of termite gut flagellates? *Environ. Microbiol. Rep.* 2, 554–559. doi: 10.1111/j.1758-2229.2009.00124.x

Ikeda-Ohtsubo, W., Strassert, J. F., Kohler, T., Mikaelyan, A., Gregor, I., McHardy, A. C., et al. (2016). 'Candidatus Adiutrix intracellularis', an endosymbiont of termite gut flagellates, is the first representative of a deep-branching clade of Deltaproteobacteria and a putative homoacetogen. *Environ. Microbiol.* 18, 2548–2564. doi: 10.1111/1462-2920.13234

Ivanova, N. N., Schwientek, P., Tripp, H. J., Rinke, C., Pati, A., Huntemann, M., et al. (2014). Stop codon reassignments in the wild. *Science* 344:909. doi: 10.1126/science.1250691

Jain, C., Rodriguez-R, L. M., Phillippy, A. M., Konstantinidis, K. T., and Aluru, S. (2018). High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat. Commun.* 9:5114.

James, S. R., Burton, H. R., McMeekin, T. A., and Mancuso, C. A. (1994). Seasonal abundance of *Halomonas meridiana, Halomonas subglaciescola, Flavobacterium gondwanense* and *Flavobacterium salegens* in four Antarctic lakes. *Antarct. Sci.* 6, 325–332. doi: 10.1017/s0954102094000490

Johri, P., Marinov, G. K., Doak, T. G., and Lynch, M. (2019). Population genetics of *Paramecium* mitochondrial genomes: recombination, mutation spectrum, and efficacy of selection. *Genome Biol. Evol.* 11, 1398–1416. doi: 10.1093/gbe/evz081

Kang, D. D., Li, F., Kirton, E., Thomas, A., Egan, R., An, H., et al. (2019). MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ* 7:e7359. doi: 10.7717/peerj.7359

Karnkowska, A., Vacek, V., Zubáèová, Z., Treitli, S. C., Petrželková, R., Eme, L., et al. (2016). A Eukaryote without a mitochondrial organelle. *Curr. Biol.* 26, 1274–1284. doi: 10.1016/j.cub.2016.03.053

Kashkouli, M., Castelli, M., Floriano, A. M., Bandi, C., Epis, S., Fathipour, Y., et al. (2020). Characterization of a novel *Pantoea* symbiont allows inference of a pattern of convergent genome reduction in bacteria associated with Pentatomidae. *Environ. Microbiol.* 23, 36–50. doi: 10.1111/1462-2920.15169

Kiraga, J., Mackiewicz, P., Mackiewicz, D., Kowalczuk, M., Biecek, P., Polak, N., et al. (2007). The relationships between the isoelectric point and: length of proteins, taxonomy and ecology of organisms. *BMC Genomics* 8:163. doi: 10.1186/1471-2164-8-163

Kispal, G., Csere, P., Prohl, C., and Lill, R. (1999). The mitochondrial proteins Atm1p and Nfs1p are required for biogenesis of cytosolic Fe/S proteins. *EMBO J.* 18, 3981–3989. doi: 10.1093/emboj/18.14.3981

Knaff, D. B., and Hirasawa, M. (1991). Ferredoxin-dependent chloroplast enzymes. *Biochim. Biophys. Acta* 1056, 93–125. doi: 10.1016/s0005-2728(05)80277-4

Knight, R. D., Freeland, S. J., and Landweber, L. F. (2001). Rewiring the keyboard: evolvability of the genetic code. *Nat. Rev. Genet.* 2, 49–58. doi: 10.1038/35047500

Kozlowski, L. P. (2016). IPC – Isoelectric point calculator. *Biol. Direct* 11:55.

Kuwahara, H., Yuki, M., Izawa, K., Ohkuma, M., and Hongoh, Y. (2017). Genome of "Ca.*Desulfovibrio trichonymphae*", an H2-oxidizing bacterium in a tripartite symbiotic system within a protist cell in the termite gut. *ISME J.* 11, 766–776. doi: 10.1038/ismej.2016.143

Lagesen, K., Hallin, P., Rødland, E. A., Staerfeldt, H. H., Rognes, T., and Ussery, D. W. (2007). RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.* 35, 3100–3108. doi: 10.1093/nar/gkm160

Lamelas, A., Gosalbes, M. J., Moya, A., and Latorre, A. (2011). New clues about the evolutionary history of metabolic losses in bacterial endosymbionts, provided by the genome of *Buchnera aphidicola* from the aphid *Cinara tujafilina*. *Appl. Environ. Microbiol.* 77, 4446–4454. doi: 10.1128/aem.00141-11

Levy Karin, E., Mirdita, M., and Söding, J. (2020). MetaEuk—sensitive, high-throughput gene discovery, and annotation for large-scale eukaryotic metagenomics. *Microbiome* 8:48.

Li, D., Luo, R., Liu, C. M., Leung, C. M., Ting, H. F., Sadakane, K., et al. (2016). MEGAHIT v1.0: a fast and scalable metagenome assembler driven by advanced methodologies and community practices. *Methods* 102, 3–11. doi: 10.1016/j.ymeth.2016.02.020

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754–1760. doi: 10.1093/bioinformatics/btp324

Lian, C. A., Yan, G. Y., Huang, J. M., Danchin, A., Wang, Y., and He, L. S. (2020). Genomic characterization of a Novel gut symbiont from the Hadal Snailfish. *Front. Microbiol.* 10:2978. doi: 10.3389/fmicb.2019.02978

Lill, R., and Mühlenhoff, U. (2005). Iron-sulfur-protein biogenesis in eukaryotes. *Trends Biochem. Sci.* 30, 133–141. doi: 10.1016/j.tibs.2005.01.006

Lim, L., and McFadden, G. I. (2010). The evolution, metabolism and functions of the apicoplast. *Phil. Trans. R. Soc. Lond. B* 365, 749–763. doi: 10.1098/rstb.2009.0273

Lopera, J., Miller, I. J., McPhail, K. L., and Kwan, J. C. (2017). Increased biosynthetic gene dosage in a genome-reduced defensive bacterial symbiont. *mSystems* 2:e00096-17. doi: 10.1128/mSystems.00096-17

López-Madrigal, S., Latorre, A., Porcar, M., Moya, A., and Gil, R. (2013). Mealybugs nested endosymbiosis: going into the 'matryoshka' system in *Planococcus citri* in depth. *BMC Microbiol.* 13:74. doi: 10.1186/1471-2180-13-74

McAllister, K. A., Peery, R. B., and Zhao, G. (2006). Acyl carrier protein synthases from gram-negative, gram-positive, and atypical bacterial species: biochemical and structural properties and physiological implications. *J. Bacteriol.* 188, 4737–4748. doi: 10.1128/jb.01917-05

McCutcheon, J. P. (2010). The bacterial essence of tiny symbiont genomes. *Curr. Opin. Microbiol.* 13:73. doi: 10.1016/j.mib.2009.12.002

McCutcheon, J. P., McDonald, B. R., and Moran, N. A. (2009). Origin of an alternative genetic code in the extremely small and GC-rich genome of a bacterial symbiont. *PLoS Genet.* 5:e1000565. doi: 10.1371/journal.pgen.1000565

McCutcheon, J. P., and Moran, N. A. (2010). Functional convergence in reduced genomes of bacterial symbionts spanning 200 My of evolution. *Genome Biol. Evol.* 2, 708–718. doi: 10.1093/gbe/evq055

McCutcheon, J. P., and Moran, N. A. (2012). Extreme genome reduction in symbiotic bacteria. *Nat. Rev. Microbiol.* 10, 13–26. doi: 10.1038/nrmicro2670

McCutcheon, J. P., and von Dohlen, C. D. (2011). An interdependent metabolic patchwork in the nested symbiosis of mealybugs. *Curr. Biol.* 21, 1366–1372. doi: 10.1016/j.cub.2011.06.051

Monteil, C. L., Vallenet, D., Menguy, N., Benzerara, K., Barbe, V., Fouteau, S., et al. (2019). Ectosymbiotic bacteria at the origin of magnetoreception in a marine protist. *Nat. Microbiol.* 4, 1088–1095. doi: 10.1038/s41564-019-0432-7

Moran, N. A. (1996). Accelerated evolution and Muller's rachet in endosymbiotic bacteria. *Proc. Natl. Acad. Sci. U.S.A.* 93, 2873–2878. doi: 10.1073/pnas.93.7.2873

Moran, N. A., and Bennett, G. M. (2014). The tiniest tiny genomes. *Annu. Rev. Microbiol.* 68, 195–215. doi: 10.1146/annurev-micro-091213-112901

Morgan-Kiss, R. M., and Cronan, J. E. (2008). The *Lactococcus lactis* FabF fatty acid synthetic enzyme can functionally replace both the FabB and FabF proteins of *Escherichia coli* and the FabH protein of *Lactococcus lactis*. *Arch. Microbiol.* 190, 427–437. doi: 10.1007/s00203-008-0390-6

Moya, A., Pereto, J., Gil, R., and Latorre, A. (2008). Learning how to live together: genomic insights into prokaryote-animal symbioses. *Nat. Rev. Genet.* 9, 218–229. doi: 10.1038/nrg2319

Nakabachi, A., and Ishikawa, H. (1999). Provision of riboflavin to the host aphid, *Acyrothosiphon pisum*, by endosymbiotic bacteria, *Buchnera*. *J. Insect. Physiol.* 45, 1–6. doi: 10.1016/s0022-1910(98)00104-8

Nakabachi, A., Yamashita, A., Toh, H., Ishikawa, H., Dunbar, H. E., Moran, N. A., et al. (2006). The 160-kilobase genome of the bacterial endosymbiont *Carsonella*. *Science* 314:267. doi: 10.1126/science.1134196

Nikoh, N., Hosokawa, T., Oshima, K., Hattori, M., and Fukatsu, T. (2011). Reductive evolution of bacterial genome in insect gut environment. *Genome Biol. Evol.* 3, 702–714. doi: 10.1093/gbe/evr064

Ohkuma, M., Sato, T., Noda, S., Ui, S., Kudo, T., and Hongoh, Y. (2007). The candidate phylum 'Termite Group 1' of bacteria: phylogenetic diversity, distribution, and endosymbiont members of various gut flagellated protists. *FEMS Microbiol. Ecol.* 60, 467–476. doi: 10.1111/j.1574-6941.2007.00311.x

Panwar, P., Allen, M. A., Williams, T. J., Hancock, A. M., Brazendale, S., Bevington, J., et al. (2020). Influence of the polar light cycle on seasonal dynamics of an Antarctic lake microbial community. *Microbiome* 8:116.

Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., and Tyson, G. W. (2014). Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* 25, 1043–1055. doi: 10.1101/gr.186072.114

Parks, D. H., Rinke, C., Chuvochina, M., Chaumeil, P.-A., Woodcroft, B. J., Evans, P. N., et al. (2017). Recovery of nearly 8,000 metagenome-assembled genomes substantially expands the tree of life. *Nat. Microbiol.* 2, 1533–1542. doi: 10.1038/s41564-017-0012-7

Parsons, J. B., and Rock, C. O. (2013). Bacterial lipids: metabolism and membrane homeostasis. *Prog. Lipid Res.* 52, 249–276. doi: 10.1016/j.plipres.2013.02.002

Pérez-Brocal, V., Gil, R., Ramos, S., Lamelas, A., Postigo, M., Michelena, J. M., et al. (2006). A small microbial genome: the end of a long symbiotic relationship? *Science* 314, 312–313. doi: 10.1126/science.1130441

Petroni, G., Spring, S., Schleifer, K. H., Verni, F., and Rosati, G. (2000). Defensive extrusive ectosymbionts of *Euplotidium* (Ciliophora) that contain microtubule-like structures are bacteria related to Verrucomicrobia. *Proc. Natl. Acad. Sci. U.S.A.* 97, 1813–1817. doi: 10.1073/pnas.030438197

Pritchard, A. E., Seilhamer, J. J., Mahalingam, R., Sable, C. L., Venuti, S. E., and Cummings, D. J. (1990). Nucleotide sequence of the mitochondrial genome of *Paramecium*. *Nucleic Acids Res.* 18, 173–180.

Ricci, F., Lauro, F. M., Grzymski, J. J., Read, R., Bakiu, R., Santovito, G., et al. (2017). The anti-oxidant defense system of the marine polar ciliate *Euplotes nobilii*: characterization of the MsrB gene family. *Biology* 6:4. doi: 10.3390/biology6010004

Roberts, N. J., Burton, H. R., and Pitson, G. A. (1993). Volatile organic compounds from Organic Lake, an Antarctic, hypersaline, meromictic lake. *Antarctic Sci.* 5, 361–366. doi: 10.1017/s0954102093000483

Rodriguez-R, L. M., and Konstantinidis, K. T. (2016). The enveomics collection: a toolbox for specialized analyses of microbial genomes and metagenomes. *PeerJ* 4:e1900v1.

Sabree, Z. L., Huang, C. Y., Okusu, A., Moran, N. A., and Normark, B. B. (2013). The nutrient supplying capabilities of *Uzinura*, an endosymbiont of armored scale insects. *Environ. Microbiol.* 15, 1988–1999. doi: 10.1111/1462-2920.12058

Santos-Garcia, D., Latorre, A., Moya, A., Gibbs, G., Hartung, V., Dettner, K., et al. (2014). Small but powerful, the primary endosymbiont of moss bugs, Candidatus Evansia muelleri, holds a reduced genome with large biosynthetic capabilities. *Genome Biol. Evol.* 6, 1875–1893. doi: 10.1093/gbe/evu149

Sato, T., Hongoh, Y., Noda, S., Hattori, S., Ui, S., and Ohkuma, M. (2009). *Candidatus Desulfovibrio trichonymphae*, a novel intracellular symbiont of the flagellate *Trichonympha agilis* in termite gut. *Environ. Microbiol.* 11, 1007–1015. doi: 10.1111/j.1462-2920.2008.01827.x

Sato, T., Kuwahara, H., Fujita, K., Noda, S., Kihara, K., Yamada, A., et al. (2014). Intranuclear verrucomicrobial symbionts and evidence of lateral gene transfer to the host protist in the termite gut. *ISME J.* 8, 1008–1019. doi: 10.1038/ismej.2013.222

Scheuermayer, M., Gulder, T. A. M., Bringmann, G., and Hentschel, U. (2006). *Rubritalea marina* gen. nov., sp. nov., a marine representative of the phylum Verrucomicrobia, isolated from a sponge (Porifera). *Int. J. Syst. Bacteriol.* 56, 2119–2124. doi: 10.1099/ijs.0.64360-0

Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30, 2068–2069. doi: 10.1093/bioinformatics/btu153

Selbach, B. P., Chung, A. H., Scott, A. D., George, S. J., Cramer, S. P., and Dos Santos, P. C. (2014). Fe-S cluster biogenesis in gram-positive bacteria: sufU is a zinc-dependent sulfur transfer protein. *Biochemistry* 53, 152–160. doi: 10.1021/bi4011978

Serra, V., Gammuto, L., Nitla, V., Castelli, M., Lanzoni, O., Sassera, D., et al. (2020). Morphology, ultrastructure, genomics, and phylogeny of *Euplotes vanleeuwenhoeki* sp. nov. and its ultra-reduced endosymbiont "*Candidatus Pinguicoccus supinus*" sp. nov. *Sci. Rep.* 10:20311.

Shinzato, C., Inoue, M., and Kusakabe, M. (2014). A snapshot of a coral "holobiont": a transcriptome assembly of the scleractinian coral, porites, captures a wide variety of genes from both the host and symbiotic zooxanthellae. *PLoS One* 9:e85182. doi: 10.1371/journal.pone.0085182

Smith, D. G., Gawryluk, R. M., Spencer, D. F., Pearlman, R. E., Siu, K. W., and Gray, M. W. (2007). Exploring the mitochondrial proteome of the ciliate protozoon *Tetrahymena thermophila*: direct analysis by tandem mass spectrometry. *J. Mol. Biol.* 374, 837–863. doi: 10.1016/j.jmb.2007.09.051

Stephens, J. L., Lee, S. H., Paul, K. S., and Englund, P. T. (2007). Mitochondrial fatty acid synthesis in *Trypanosoma brucei*. *J. Biol. Chem.* 282, 4427–4436. doi: 10.1074/jbc.m609037200

Stingl, U., Radek, R., Yang, H., and Brune, A. (2005). "Endomicrobia": cytoplasmic symbionts of termite gut protozoa form a separate phylum of prokaryotes. *Appl. Environ. Microbiol.* 71, 1473–1479. doi: 10.1128/aem.71.3.1473-1479.2005

Strassert, J. F., Mikaelyan, A., Woyke, T., and Brune, A. (2016). Genome analysis of '*Candidatus Ancillula trichonymphae*', first representative of a deep-branching clade of Bifidobacteriales, strengthens evidence for convergent evolution in flagellate endosymbionts. *Environ. Microbiol. Rep.* 8, 865–873. doi: 10.1111/1758-2229.12451

Swart, E. C., Nowacki, M., Shum, J., Stiles, H., Higgins, B. P., Doak, T. G., et al. (2011). The *Oxytricha trifallax* mitochondrial genome. *Genome Biol. Evol.* 4, 136–154. doi: 10.1093/gbe/evr136

Syberg-Olsen, M. J., Irwin, N. A. T., Vannini, C., Erra, F., Di Giuseppe, G., Boscaro, V., et al. (2016). Biogeography and character evolution of the ciliate genus *Euplotes* (Spirotrichea, Euplotia), with description of *Euplotes curdsi* sp. nov. *PLoS One* 11:e0165442. doi: 10.1371/journal.pone.0165442

Takahashi, Y., Mitsui, A., Hase, T., and Matsubara, H. (1986). Formation of the iron-sulfur cluster of ferredoxin in isolated chloroplasts. *Proc. Natl. Acad. Sci. U.S.A.* 83, 2434–2437. doi: 10.1073/pnas.83.8.2434

Tamura, K., and Nei, M. (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* 10, 512–526.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30, 2725–2729. doi: 10.1093/molbev/mst197

Thao, M. L., Moran, N. A., Abbot, P., Brennan, E. B., Burckhardt, D. H., and Baumann, P. (2000). Cospeciation of psyllids and their primary prokaryotic endosymbionts. *Appl. Environ. Microbiol.* 66, 2868–2905.

Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22, 4673–4680. doi: 10.1093/nar/22.22.4673

Tillett, D., and Neilan, B. A. (2000). Xanthogenate nucleic acid isolation from cultured and environmental cyanobacteria. *J. Phycol.* 36, 251–258. doi: 10.1046/j.1529-8817.2000.99079.x

Tsaousis, A. D. (2019). On the origin of iron/sulfur cluster biosynthesis in eukaryotes. *Front. Microbiol.* 10:2478. doi: 10.3389/fmicb.2019.02478

Tsaousis, A. D., Ollagnier de Choudens, S., Gentekaki, E., Long, S., Gaston, D., Stechmann, A., et al. (2012). Evolution of Fe/S cluster biogenesis in the anaerobic parasite *Blastocystis*. *Proc. Natl. Acad. Sci. U.S.A.* 109, 10426–10431. doi: 10.1073/pnas.1116067109

Tschitschko, B., Erdmann, S., DeMaere, M. Z., Roux, S., Panwar, P., Allen, M. A., et al. (2018). Genomic variation and biogeography of Antarctic haloarchaea. *Microbiome* 6:113.

Van Leuven, J. T., Mao, M., Xing, D. D., Bennett, G. M., and McCutcheon, J. P. (2019). Cicada endosymbionts have tRNAs that are correctly processed despite having genomes that do not encode all of the tRNA processing machinery. *Mbio* 10:e01950-18.

Vandekerckhove, T. T. M., Coomans, A., Cornelis, K., Baert, P., and Gillis, M. (2002). Use of the verrucomicrobia-specific probe EUB338-III and fluorescent *in situ* hybridization for detection of "*Candidatus* Xiphinematobacter" cells in nematode hosts. *Appl. Environ. Microbiol.* 68, 3121–3125. doi: 10.1128/aem.68.6.3121-3125.2002

Wagner, M., and Horn, M. (2006). The Planctomycetes, Verrucomicrobia, Chlamydiae and sister phyla comprise a superphylum with biotechnological and medical relevance. *Curr. Opin. Biotechnol.* 17, 241–249. doi: 10.1016/j.copbio.2006.05.005

Wang, R. L., Miao, W., Wang, W., Xiong, J., and Liang, A. H. (2018). EOGD: the *Euplotes octocarinatus* genome database. *BMC Genomics* 19:63. doi: 10.1186/s12864-018-4445-z

Wang, X., Quinn, P. J., and Yan, A. (2015). Kdo2-lipid A: structural diversity and impact on immunopharmacology. *Biol. Rev. Camb. Philos. Soc.* 90, 408–427. doi: 10.1111/brv.12114

Wertz, J. T., Kim, E., Breznak, J. A., Schmidt, T. M., and Rodrigues, J. L. (2012). Genomic and physiological characterization of the Verrucomicrobia isolate *Geminisphaera colitermitum* gen. nov., sp. nov., reveals microaerophily and nitrogen fixation genes. *Appl. Environ. Microbiol.* 78, 1544–1555. doi: 10.1128/aem.06466-11

Wu, M., Sun, L. V., Vamathevan, J., Riegler, M., Deboy, R., Brownlie, J. C., et al. (2004). Phylogenomics of the reproductive parasite *Wolbachia pipientis* wMel: a streamlined genome overrun by mobile genetic elements. *PLoS Biol.* 2:E69. doi: 10.1371/journal.pbio.0020069

Yamao, F., Muto, A., Kawauchi, Y., Iwami, M., Iwagami, S., Azumi, Y., et al. (1985). UGA is read as tryptophan in *Mycoplasma capricolum*. *Proc. Natl. Acad. Sci. U.S.A.* 82, 2306–2309. doi: 10.1073/pnas.82.8.2306

Yang, H., Schmitt-Wagner, D., Stingl, U., and Brune, A. (2005). Niche heterogeneity determines bacterial community structure in the termite gut (*Reticulitermes santonensis*). *Environ. Microbiol.* 7, 916–932. doi: 10.1111/j.1462-2920.2005.00760.x

Yao, Z., Davis, R. M., Kishony, R., Kahne, D., and Ruiz, N. (2012). Regulation of cell size in response to nutrient availability by fatty acid biosynthesis in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 109, E2561–E2568.

Yau, S., Lauro, F. M., DeMaere, M. Z., Brown, M. V., Thomas, T., Raftery, M. J., et al. (2011). Virophage control of Antarctic algal host-virus dynamics. *Proc. Natl. Acad. Sci. U.S.A.* 108, 6163–6168. doi: 10.1073/pnas.1018221108

Yau, S., Lauro, F. M., Williams, T. J., Demaere, M. Z., Brown, M. V., Rich, J., et al. (2013). Metagenomic insights into strategies of carbon conservation and unusual sulfur biogeochemistry in a hypersaline Antarctic lake. *ISME J.* 7, 1944–1961. doi: 10.1038/ismej.2013.69

Yoon, J., Matsuo, Y., Katsuta, A., Jang, J. H., Matsuda, S., Adachi, K., et al. (2008). *Haloferula rosea* gen. nov., sp. nov., *Haloferula harenae* sp. nov., *Haloferula phyci* sp. nov., *Haloferula helveola* sp. nov. and *Haloferula sargassicola* sp. nov., five marine representatives of the family Verrucomicrobiaceae within the phylum 'Verrucomicrobia'. *Int. J. Syst. Evol. Microbiol.* 58, 2491–2500. doi: 10.1099/ijs.0.2008/000711-0

Yoon, J., Yasumoto-Hirose, M., Katsuta, A., Sekiguchi, H., Matsuda, S., Kasai, H., et al. (2007a). *Coraliomargarita akajimensis* gen. nov., sp. nov., a novel member of the phylum 'Verrucomicrobia' isolated from seawater in Japan. *Int. J. Syst. Evol. Microbiol.* 57, 959–963. doi: 10.1099/ijs.0.64755-0

Yoon, J., Yasumoto-Hirose, S., Matsuda, S., Nozawa, M., Matsuda, S., Kasai, H., et al. (2007b). *Pelagicoccus mobilis* gen. nov., sp. nov., *Pelagicoccus albus* sp. nov. and *Pelagicoccus litoralis* sp. nov., three novel members of subdivision 4 within the phylum 'Verrucomicrobia', isolated from seawater by *in situ* cultivation. *Int. J. Syst. Evol. Microbiol.* 57, 1377–1385. doi: 10.1099/ijs.0.64970-0

Zientz, E., Dandekar, T., and Gross, R. (2004). Metabolic interdependence of obligate intracellular bacteria and their insect hosts. *Microbiol. Mol. Biol. Rev.* 68, 745–770. doi: 10.1128/mmbr.68.4.745-770.2004

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

frontiers in Microbiology

ORIGINAL RESEARCH
published: 22 July 2021
doi: 10.3389/fmicb.2021.701186

Divergent Genomic Adaptations in the Microbiomes of Arctic Subzero Sea-Ice and Cryopeg Brines

Josephine Z. Rapp[1†], Matthew B. Sullivan[2,3,4,5] and Jody W. Deming[1]*

[1] School of Oceanography, University of Washington, Seattle, WA, United States, [2] Byrd Polar and Climate Research Center, Ohio State University, Columbus, OH, United States, [3] Department of Microbiology, Ohio State University, Columbus, OH, United States, [4] Department of Civil, Environmental and Geodetic Engineering, Ohio State University, Columbus, OH, United States, [5] Center of Microbiome Science, Ohio State University, Columbus, OH, United States

Subzero hypersaline brines are liquid microbial habitats within otherwise frozen environments, where concentrated dissolved salts prevent freezing. Such extreme conditions presumably require unique microbial adaptations, and possibly altered ecologies, but specific strategies remain largely unknown. Here we examined prokaryotic taxonomic and functional diversity in two seawater-derived subzero hypersaline brines: first-year sea ice, subject to seasonally fluctuating conditions; and ancient cryopeg, under relatively stable conditions geophysically isolated in permafrost. Overall, both taxonomic composition and functional potential were starkly different. Taxonomically, sea-ice brine communities ($\sim 10^5$ cells mL^{-1}) had greater richness, more diversity and were dominated by bacterial genera, including *Polaribacter*, *Paraglaciecola*, *Colwellia*, and *Glaciecola*, whereas the more densely inhabited cryopeg brines ($\sim 10^8$ cells mL^{-1}) lacked these genera and instead were dominated by *Marinobacter*. Functionally, however, sea ice encoded fewer accessory traits and lower average genomic copy numbers for shared traits, though DNA replication and repair were elevated; in contrast, microbes in cryopeg brines had greater genetic versatility with elevated abundances of accessory traits involved in sensing, responding to environmental cues, transport, mobile elements (transposases and plasmids), toxin-antitoxin systems, and type VI secretion systems. Together these genomic features suggest adaptations and capabilities of sea-ice communities manifesting at the community level through seasonal ecological succession, whereas the denser cryopeg communities appear adapted to intense bacterial competition, leaving fewer genera to dominate with brine-specific adaptations and social interactions that sacrifice some members for the benefit of others. Such cryopeg genomic traits provide insight into how long-term environmental stability may enable life to survive extreme conditions.

Keywords: cryopeg, sea ice, metagenomics, metatranscriptomics, microbial ecology, hypersalinity, subzero temperature, cryosphere

OPEN ACCESS

Edited by:
Jérôme Comte, Université du Québec, Canada

Reviewed by:
Lyle Whyte, McGill University, Canada
Alon Philosof, California Institute of Technology, United States

****Correspondence:***
Josephine Z. Rapp
josephine.rapp.1@ulaval.ca

†Present address:
Josephine Z. Rapp, Department of Biochemistry, Microbiology and Bioinformatics, Université Laval, Québec, QC, Canada

Specialty section:
This article was submitted to Extreme Microbiology, a section of the journal Frontiers in Microbiology

Received: *27 April 2021*
Accepted: *29 June 2021*
Published: *22 July 2021*

Citation:
Rapp JZ, Sullivan MB and Deming JW (2021) Divergent Genomic Adaptations in the Microbiomes of Arctic Subzero Sea-Ice and Cryopeg Brines. Front. Microbiol. 12:701186. doi: 10.3389/fmicb.2021.701186

INTRODUCTION

The cryosphere, that portion of the planet where most of the water is frozen, includes more than half of land surfaces and 7% of the surface ocean globally (Vaughan et al., 2013). These frozen regions are incurring significant losses due to climate change (Fountain et al., 2012). Cryosphere components like permafrost, relatively stable features that have remained frozen for thousands

or even millions of years (Froese et al., 2008), are warming now (Biskaborn et al., 2019; Nitzbon et al., 2020). Others, like sea ice, are inherently short-lived, with formation and melting occurring annually (Vaughan et al., 2013), yet recent losses in areal extent and volume, particularly for Arctic sea ice, are pronounced (Overland and Wang, 2013; Stroeve and Notz, 2018). Though largely frozen, these at-risk environments contain subzero hypersaline brines that provide interior liquid habitat for diverse microbial life of both ecological relevance and inherent curiosity for their unique adaptations to extreme conditions (Boetius et al., 2015). With continued warming, opportunities to study these microbes and their current roles *in situ* dwindle, yet the molecular microbial ecology of the cryosphere remains underexplored compared to other areas of environmental microbiology (reviewed by Deming and Collins, 2017). Direct comparisons between subzero brines that differ in age or stability are rare (Cooper et al., 2019), and have not yet included an exploration of genomic adaptations to long-term vs. fluctuating extreme conditions. The challenges are many, including safe accessibility, sampling without altering (melting) the habitat, and accounting for the complexities and heterogeneity of the micro-scale living spaces within ice, especially in sea ice (Junge et al., 2001), which is subject to seasonal extremes in temperature and salinity (Ewert and Deming, 2013, 2014) and thus fluctuating porosity (Golden et al., 2007; Petrich and Eicken, 2017).

The habitable space within sea ice or any frozen-water matrix exists in the form of hypersaline liquid inclusions, kept liquid well below 0°C due to freezing point depression by the concentration of dissolved salts (Petrich and Eicken, 2017) and, to a lesser extent, dissolved organics (Deming and Young, 2017). Despite subzero temperatures and high salinities, diverse microbial communities have been reported for even the most extreme cryosphere habitats, including wintertime sea ice at −28°C and 240 ppt salt (Collins et al., 2008, 2010). For a microorganism to inhabit a subzero brine its core cellular processes must be adapted to function under the extreme conditions of temperature and salinity. Classically, research has focused on cellular strategies to counteract these abiotic stressors, particularly through osmolytes (Kempf and Bremer, 1998; Ewert and Deming, 2014; Firth et al., 2016), chemical modification of the cell envelope (Chintalapati et al., 2004), pigment production (Dieser et al., 2010), and synthesis of cold-shock proteins and nucleic acid chaperones (Baraúna et al., 2017). Many organisms in frozen environments also synthesize extracellular polysaccharides (EPS or exopolymers), with chemical characteristics that provide cryoprotection through freezing-point depression or inhibition of ice recrystallization (Carillo et al., 2015) and osmoprotection (Krembs and Deming, 2008; Deming and Young, 2017).

Sea ice may be among the best studied of cryosphere habitats, with winter microbial communities reflecting the parent seawater prior to freezing (Collins et al., 2010) and characteristic Gammaproteobacteria and Bacteroidia reported consistently during spring and later seasons (Brown and Bowman, 2001; Eronen-Rasimus et al., 2016; Rapp et al., 2018). Such studies have been based mostly on 16S rRNA genes, whether from cultured bacteria or melted sea-ice samples; few have used metagenomic sequencing to explore the resident bacteria and archaea (Bowman et al., 2014; Yergeau et al., 2017). We are not aware of metagenomic analyses of microbial communities in subzero sea-ice brines obtained directly, without the compromise of first melting the ice.

In stark contrast to the transient brines of sea ice are the stable, ancient (late Pleistocene) and rarely studied brines of cryopegs, marine-derived sediment layers within permafrost (Gilichinsky et al., 2003, 2005) that contain discrete lenses of brine geophysically isolated at fairly constant temperature year-round, ranging between –9 and –11°C in Siberian cryopegs (Gilichinsky et al., 2005) and –6 and –8°C in Alaskan cryopegs (Colangelo-Lillis et al., 2016; Cooper et al., 2019). A recent first exploration of microbial community composition in cryopeg brines based on 16S rRNA genes, obtained from the same cryopeg system also studied here, revealed a less diverse and distinctive community structure than observed in sea-ice brines (Cooper et al., 2019), with concentrations of total and dividing cells greatly exceeding those in the sampled sea-ice brines (approximately 10^8 vs. 10^5 cells mL^{-1} and up to 29 vs. ∼ 5%, respectively; **Table 1**). Most of the sampled cryopeg brines were dominated by only a few taxa, in particular members of the genus *Marinobacter*, yet the specific adaptations enabling their success under the stably extreme conditions remain unknown.

Because the freezing process that concentrates salts within the brine inclusions also concentrates microorganisms and other particulate and dissolved constituents present in the source water, microbial cell numbers, as well as organic matter and nutrient concentrations in the brines, are elevated compared to their source. As a result, cell-to-cell contact rates within sea ice are much higher than in seawater (Collins and Deming, 2011), and would be even higher in a cryopeg brine containing 10^8 cells mL^{-1} (**Table 1**). Higher cell concentrations and contact rates in a spatially constrained setting present additional challenges in the form of resource and space competition, yet they may also lead to genetic exchange and diversification, as several examples of genes relevant to cold or salt adaptation appear to be the result of horizontal gene transfer (Bowman, 2008; Collins and Deming, 2013; Feng et al., 2014; Zhong et al., 2020). To maintain or increase cell concentrations over time, microbes in subzero brines likely require versatile nutrient acquisition capabilities, given slow rates of diffusion in such viscous fluids (Showalter and Deming, 2021). While microbes in sea ice may be presented with sufficient organic matter from sea-ice algae during the spring and summer months, they must also survive the aphotic winter season. In cryopeg brines, where phototrophic primary production is absent, heterotrophic microbes may survive *via* the products of their extracellular enzymes (Showalter and Deming, 2021), hydrolyzing remnant plant and detrital material from the surrounding permafrost and ice-wedge environments (Iwahana et al., 2021), and *via* dead microbial matter, as proposed for oceanic subsurface marine sediments (Bradley et al., 2018), yet any metagenomic insights into survival mechanisms employed *in situ* are currently lacking.

Here we present a comparative metagenomic analysis of the taxonomic and functional diversity of microbial communities present in brines from first-year sea ice and in ancient

TABLE 1 | Sample characteristics including temperature, salinity, pH, organic matter and nutrient concentrations, prokaryotic cell numbers, and percent dividing cells (data compiled from Cooper et al., 2019; Zhong et al., 2020).

Parameter (unit)	Cryopeg brines					Sea-ice brines[a]		Sea-ice sections[b]	
Sample label	CBIW[c]	CBIW	CBIA	CB1	CB4[d]	SB[c]	SB[d]	SI3U	SI3L
Sampling year	2017	2018	2018	2018	2018	2017	2018	2017	2017
T (°C)	–6	–6	–6	–6	–6	–4	–3	–5	–3
S (ppt)	140	121	112	122	121	78	75	18	18
pH	6.6	–	–	6.6	6.6	–	7.2	–	–
POC (μg C mL^{-1})	23.8	29.0	–	145.1	46.3	0.24	0.22	0.20	2.53
PON (μg N mL^{-1})	3.92	5.32	–	23.99	6.98	0.03	0.04	0.03	0.34
C:N value	6.06	5.46	–	6.08	6.62	8.94	5.66	7.64	7.46
DOC (μM C)	3.00×10^4	8.22×10^4	–	1.02×10^5	8.50×10^4	4.46×10^2	2.00×10^2	2.27×10^2	4.97×10^2
pEPS (μM C)	1.44×10^4	1.14×10^2	–	1.63×10^3	8.91×10^1	1.90×10^0	9.67×10^{-1}	2.98×10^0	2.93×10^1
dEPS (μM C)	8.55×10^3	1.25×10^4	–	1.94×10^4	1.98×10^4	2.61×10^2	bd	–	–
PO_4 (μM)	1.94	0.68	–	1.22	0.6	1.8	1.6	0.21	0.19
NO_3 (μM)	5.56	0.82	–	bd	13.6	0.15	3.63	0.31	0.87
NO_2 (μM)	2.03	0.89	–	16.2	2.96	0.02	0.11	0.03	bd
NH_4 (μM)	1.75×10^3	1.17×10^3	–	3.35×10^3	4.52×10^3	5.50×10^{-1}	3.42×10^0	4.60×10^{-1}	bd
Prokaryotic cells (mL^{-1})	1.39×10^8	1.22×10^8	7.31×10^7	9.57×10^7	1.14×10^7	2.22×10^5	1.11×10^5	7.68×10^4	1.07×10^6
Dividing cells (%)	–	1.3	2.5	1.9	29	–	4.68	–	–

[a]*Collected directly from "sackholes" in the ice.*
[b]*All values determined on melted ice samples and scaled to ice volume; scaling to brine volume (brine fraction of the ice) would increase values by an estimated factor of 2.5 for SI3U and 1.6 for SI3L (Cox and Weeks, 1983).*
[c]*Two size fractions of this sample were included for metagenomic analysis (see* ***Supplementary Table 1*** *for details).*
[d]*Both metagenomic and metatranscriptomic analyses were performed.*
–, not determined; bd, below detection limit.

cryopeg brines found deep within approximately 40,000-year-old permafrost (Meyer et al., 2010; Iwahana et al., 2021) near Utqiaġvik, Alaska. Both of these seawater-sourced habitats share the traits of subzero temperature and hypersalinity (–4 to –3°C and 75–78 ppt for the sea-ice brines at time of sampling, and –6°C and 112–140 ppt for the cryopeg brines; **Table 1**), yet they differ greatly in their environmental age and stability. We used metagenomic and selected metatranscriptomic sequencing to investigate how communities in these two extreme habitats are adapted to life under subzero and hypersaline conditions, and whether they are following similar or different strategies to overcome potential resource and space competition. Our overall aim is to improve the understanding of the impacts of environmental stability vs. transience on the adaptability of life to such extreme conditions.

MATERIALS AND METHODS

Field Sampling

We collected samples near Utqiaġvik, Alaska, during two field seasons in May 2017 and May 2018. Landfast, first-year sea ice was sampled near the Barrow Sea Ice Mass Balance site, operated by the University of Alaska Fairbanks (Druckenmiller et al., 2009), at 71.2223°N, 156.3018°W in 2017 and 71.2238°N, 156.3083°W in 2018. At both times, the sites were covered with a layer of snow, 16–19 cm in 2017 and 6–10 cm in 2018, which we cleared prior to sampling activities. Air temperature at 1 m above the ice varied between –6 and –4°C in both years. Partial ice-core holes (sackholes) were drilled at 1-m distance from each other to a depth of 75 cm in 2017, when ice thickness was 117 cm, and to 55 cm in 2018, when ice thickness was 110 cm. Following Eicken et al. (2009), we covered the sackholes and allowed brine to drain into them for 3–5 h before collection. The first sampling of the sea-ice brine (SB), after sufficient volumes for all other sampling needs had accumulated, was for RNA extraction. We used a sterile syringe to withdraw 50 ml of brine from each of 10 individual sackholes and pooled the brines onto one Sterivex filter (MilliporeSigma), which we treated immediately on site with RNAlater™ (Invitrogen, Thermo Fisher Scientific). For DNA analyses and all ancillary measurements, we then collected 20 L of SB by hand-pumping into an acid-washed cubitainer, first rinsed with sample brine.

In each year, we collected full-length sea-ice cores using a MARK II ice auger (Kovacs Enterprise) with a diameter of 9 cm. For two cores, vertical temperature measurements were made in the field at 5-cm intervals according to Eicken et al. (2009), from which we inferred the sackhole brine temperature of –4 and –3°C in both years (**Table 1**). Sackhole brine salinities were measured by hand-held refractometer, yielding salinities of 75–78 ppt in both years. In 2017, we sectioned one sea-ice core in the field to obtain the upper 25 cm (SI3U), middle 50–75 cm (SI3M) and lowermost 25 cm (SI3L) using a 70% ethanol-rinsed, custom-alloy bow saw. Each ice section was collected in a sterile Whirl-Pak bag. As SI3M failed subsequent sequencing efforts, it does not appear in this study.

Cryopeg brine (CB) was obtained through the Barrow Permafrost Tunnel, located at 71.2944 °N, 156.7153 °W and situated 6 m below the surface. The temperature inside the tunnel held at –6°C during our sampling work (**Table 1**). Brines were extracted from approximately 2 m below the tunnel floor through discrete boreholes, some previously established

(Colangelo-Lillis et al., 2016) and others drilled for this study using a cleaned and ethanol-rinsed ice auger or SIPRE corer (as in Colangelo-Lillis et al., 2016). We used a hand pump and 1.5-m length of sterile Masterflex and Teflon tubing (Cole-Parmer, Vernon Hills, IL, United States) to collect samples in acid-washed 2-L vacuum flasks (initially autoclaved, then acid-washed and ethanol-rinsed between samples). Detailed borehole histories and sampling are provided in Cooper et al. (2019) and Iwahana et al. (2021). Sample names here are based on borehole name and filtration pore size (see below). During the two field campaigns we were able to obtain a total of six cryopeg brine samples for metagenomic analyses: one each from boreholes CB4, CB1 and CBIA and three from borehole CBIW, representing different years and filtration steps. Brine volumes were very limited, but we were able to collect sub-samples for RNA extraction from two of the boreholes (CB4 and CBIW), concentrating 100 ml on a Sterivex™ filter (MilliporeSigma) and immediately stabilizing it with RNAlater™ (Invitrogen, Thermo Fisher Scientific) while still in the tunnel. The metatranscriptome sequencing library for CBIW later failed sequencing efforts and no longer appears in this study.

We stored all samples in insulated coolers on site and during transport to cold rooms set to near *in situ* brine temperatures at the Barrow Arctic Research Center. Cryopeg brines were held at –6°C and sea-ice brines at –1°C until further processing within 2–8 h. Sea-ice sections were melted into a filter-sterilized (0.22-μm pore size) artificial sea salt solution (salinity of 32 g/L, Cat No. S9883, Sigma) at a 1:1 vol:vol ratio. For subsequent DNA extraction, maximum available volumes were filtered onto 47-mm GTTP Isopore filters (**Supplementary Table 1**), after accounting for sufficient minimum volumes to remain for basic sample characterization (**Table 1**). The sea-ice samples and one cryopeg brine sample (CBIW in 2017) were filtered sequentially through filters of 3.0 μm (CBIW_3.0) and 0.2 μm (CBIW_0.2) pore size to obtain operational microeukaryotic and prokaryotic fractions, respectively. Biomass from the other cryopeg brines was collected on a 0.2 μm pore size filter without prefiltration. The artificial sea salt solution was processed in parallel as a blank (**Supplementary Figure 1**). For subsequent RNA extraction, we incubated the Sterivex filters in RNAlater overnight at 4°C, subsequently transferring them to –20°C until extraction according to the manufacturer's instructions.

DNA and RNA Extraction

For DNA extractions, we used the DNeasy PowerSoil kit (QIAGEN, Germantown, MD, United States), following the manufacturer's instructions. To ensure free movement of the samples during bead beating, we cut filters into stripes using sterilized scalpels and forceps, and performed two extractions per sample, each containing half a filter. Isolated DNA was eluted in 60 μl nuclease-free water (Thermo Fisher Scientific, Waltham, MA), and both extracts from the same filter were pooled and stored at –20°C.

Prior to any RNA work, we cleaned all working areas and equipment with 70% Ethanol, followed by a wipe with RNaseZap™ (Thermo Fisher Scientific, Waltham, MA). For RNA extractions, we used the DNeasy PowerWater Sterivex Kit (Qiagen, Germantown, MD), following the manufacturer's instructions, with the following modifications developed by the QIAGEN microbiome team that allowed recovering RNA from the filter membranes without the need to cut open the plastic casing holding the membrane: After removing RNAlater™ from the casing using a sterile 3-ml syringe, we followed protocol, but added 20 μl of β-mercaptoethanol (βME) for every 880 μl of Solution ST1B in step 2. Further, we modified step 7 and incubated the filter unit at 70°C for 10 min, and in step 19 added 1.5 ml of Solution MR, as well as 1.5 ml of 100% ethanol. We eluted the isolated RNA in 50 μl of sterile RNase-free water (Thermo Fisher Scientific, Waltham, MA) and performed a subsequent DNA digest by adding 7 μl DNase buffer (10×) (Sigma-Adrich, St. Louis, MO), 10 μl DNase I recombinant (Sigma-Adrich, St. Louis, MO), and 2 μl RNAsin RNase inhibitor (Promega Corporation, Madison, WI, United States). The mix was then incubated for 20 min at 37°C, and subsequently for 10 min at 56°C before placing it on ice. Finally, RNA was cleaned and concentrated using the RNeasy MinElute Cleanup Kit (Qiagen, Germantown, MD) and the final product was eluted in 40 μl RNase-free water and stored at −80°C.

DNA Sequencing

For sequencing, we shipped nucleic acid extracts on dry ice to the DOE Joint Genome Institute (JGI). Here, DNA was sheared to 300 bp using the Covaris LE220 (Covaris) and size selected using SPRI beads (Beckman Coulter). The fragments were treated with end-repair, A-tailing, and ligation of Illumina compatible adapters (IDT, Inc.) using the KAPA-Illumina library creation kit (KAPA biosystems) and 5–20 cycles of PCR were used to enrich for the final library. Sequencing of metagenomic libraries was performed on an Illumina NovaSeq sequencer to produce 2 × 151 base pair (bp) paired-end reads. Contaminant removal (human, mouse, cat, dog, microbial, synthetic, non-synthetic), trimming of adapters, and right-quality trimming of reads where quality dropped to 0 was done with BBDuk. BBDuk was also used to remove reads that contained four or more "N" bases, had an average quality score across the read of less than 3, or had a minimum length ≤51 bp or 33% of the full read length. Quality control removed between 0.2 and 9% of raw sequences, resulting in a total of 247–661 million quality-controlled reads per sample.

Trimmed, screened, and paired-end Illumina reads were read-corrected using bfc (version r181) with -k 21. Reads with no mate pair were removed. The remaining reads were assembled using SPAdes using the settings "–only-assembler -k 33,55,77,99,127 –meta." The entire filtered read set was mapped to the final assembly and coverage information generated using BBMap with default parameters except ambiguous = random. See **Supplementary Table 2** for details on software versioning and read numbers.

RNA Sequencing

For RNA sequencing, ribosomal RNA was removed from total RNA using a Ribo-Zero™ rRNA Removal Kit (Epicenter). Stranded cDNA libraries were generated using the Illumina Truseq Stranded RNA LT kit. The rRNA depleted RNA

was fragmented and reversed transcribed using random hexamers and SSII (Invitrogen) followed by second strand synthesis. The fragmented cDNA was treated with end-pair, A-tailing, adapter ligation, and 10 cycles of PCR. Low input libraries were sequenced on an Illumina NovaSeq sequencer, generating 2 × 151 bp long paired-end reads. Contaminant removal and quality control was conducted in the same way as for the metagenomic reads (see above). The final filtered fastq contained 48,612,144 and 23,642,934 reads in the cryopeg and the sea-ice metatranscriptome, respectively (**Supplementary Table 2**).

The quality-filtered reads were assembled with MEGAHIT v1.1.2 (kmers: 23,43,63,83,103,123). For coverage information, BBMap was used to map the quality-filtered reads against the metatranscriptome assemblies, as well as against the corresponding metagenomic reference assemblies with default parameters except for ambiguous = random.

Functional Annotation

Functional and structural annotation of the generated metagenomic and metatranscriptomic assemblies was performed through the DOE-JGI Metagenome Annotation Pipeline (MAP) (Huntemann et al., 2016). Information on pipeline, software used, and versioning is provided in **Supplementary Table 2**. Briefly, the pipeline first identifies structural RNAs and regulatory motifs, protein-coding genes, tRNAs and CRISPR arrays, which is followed by functional annotation of protein-coding genes by assignment to 3D fold and functional protein families using various databases. Here we focused on the assignment of KEGG Orthology (KO) Terms. The coverage information obtained from read mapping is used to calculate "estimated gene copies," whereby the number of genes is multiplied by the average coverage of the contigs, on which these genes were predicted. The JGI-IMG Phylogenetic Distribution of Best Hits tool (Chen et al., 2019) assigns taxonomy to genes in an assembly through a homology search using LAST (Frith et al., 2010) against IMG reference isolates (high quality public genomes). Scaffold lineage affiliation is then assigned as the last common ancestor of all best gene hits on the scaffold, provided that at least 30% of the genes have hits.

Taxonomic Composition of Metagenomes and Metatranscriptomes

We used the phyloFlash pipeline (Gruber-Vodicka et al., 2019) to screen the quality-controlled metagenomic reads for small subunit ribosomal RNA (SSU rRNA) reads by mapping against the SILVA SSU Ref NR 99 database 132 (Quast et al., 2013). The top reference hits per read were used to report an approximate taxonomic affiliation, and the mapped read counts were then used to generate an overview of community composition across samples. We used default settings with the parameters "-readlength 150 –almosteverything." For the rRNA depleted metatranscriptomes, we used scaffold taxonomy (see above) to report community composition.

We used SPAdes (as implemented in phyloFlash) for a targeted assembly of full-length SSU rRNA sequences from the identified SSU rRNA reads. From the resulting 460 reconstructed sequences, we selected those of dominant prokaryotic community members (genera of >1% relative abundance) for multiple sequence alignment using MAFFT Q-INS-I (Katoh et al., 2019). We used the resulting alignment to calculate a Neighbor-Joining tree of genus representatives with the Jukes-Cantor model and 1,000 bootstraps. *Candidatus* Nitrosopumilus, the most abundant archaeal community member, was included as tree outgroup despite relative abundance values <1%. The tree and the corresponding community abundance data were visualized with iTOL v4 (Letunic and Bork, 2019).

Comparative Analysis of Prokaryotic Adaptive Genomic Strategies

For comparative analyses of the functional potential encoded in the microbial communities in cryopeg brines and sea ice, we focused on molecular functions represented in the KEGG Orthology (KO) database (Kanehisa et al., 2016). Here, genes are grouped as functional orthologs of experimentally characterized proteins, and further associated with higher-level functions in the context of molecular pathways and modules. We filtered the data for eukaryotic sequences, because eukaryotes, predominantly diatoms, can be a major component of sea-ice communities in lower sea-ice sections and therefore contribute to metagenomic sea-ice data. Although eukaryotes were negligible in cryopeg brines (**Supplementary Figures 2A, 3**), cryopeg data were similarly filtered. We used EukRep, a classifier that utilizes k-mer composition of assembled sequences to differentiate between prokaryotic and eukaryotic genome fragments, with a minimum sequence length cutoff of 1 kb for the identification of eukaryotic contigs in our assemblies (West et al., 2018). We then subtracted the estimated KO abundances for KOs predicted on eukaryotic contigs from the total gene abundances for each sample.

To allow a direct and biologically meaningful comparison between samples, we normalized all metagenomic datasets using Metagenomic Universal Single-Copy Correction (MUSiCC) with the settings "–normalize –correct learn_model" (Manor and Borenstein, 2015). Rather than comparing the relative abundance of KOs between samples, MUSiCC estimates the average genomic copy number of a KO in a given sample using a large set of universal single-copy genes for gene abundance calibration.

As individual KOs can be associated with multiple higher-level functions, we used Evidence-based Metagenomic Pathway Assignment using geNe Abundance Data (EMPANADA) to infer the prevalence of different KO pathways and modules in each sample (Manor and Borenstein, 2017). Here, the average abundance of only non-shared KOs is used to generate support values for each pathway, and to subsequently partition the contribution of shared KOs between pathways. We used the MUSiCC-normalized KO data as input with the settings "–mapping_method by_support -use_only_non_overlapping_genes -v -remove_ko_with_no_abundance_measurement." Abundance

values for pathways and modules therefore represent the average number of genes belonging to a pathway per genome.

Statistical Analyses and Visualization of Results

We performed all statistical analysis and visualization of results in R v3.6.1 (R Core Team, 2017). We calculated alphadiversity indices per sample using the *vegan* package v.2.5.6 (Oksanen et al., 2019) with the *diversity* function, and performed an analysis of variance (ANOVA) followed by a *post-hoc* Tukey pairwise significance test to assess the significance ($p \leq 0.05$) of variance in observed richness and Shannon diversity between both brine environments. We performed a Principal Component Analysis (PCA) for the analysis of beta diversity patterns in our data. We used the *ampvis2* package v.2.5.8 (Andersen et al., 2018), which uses the *vegan* package v.2.5.6 (Oksanen et al., 2019) for ordination and the *ggplot2* package v3.3.2 (Wickham, 2016) for plotting. The taxonomy data has been transformed by applying a Hellinger transformation prior to ordination analysis. By the creation of a biplot and adding species scores to the PCA ordination, we inferred information about taxa and KO functions that contributed most to observed sample spread. Further, we used the Bray-Curtis distance measure on the basis of relative abundances of genus-level taxonomy and normalized KO abundances, and performed an Analysis of Similarity (ANOSIM) with 999 permutations to test whether between-group differences were larger than within-group. For the community composition heatmaps, we used the heatmap function within *ampvis2* with default parameters.

We used the *stats4bioinfo* package v1.1.0 (van Helden, 2016) to perform per-row Welch's *t*-tests to identify genes, modules or pathways where the sample mean differed in abundance between cryopeg brine and sea-ice samples. By applying multiple testing correction, we identified significantly different features as those genes, modules or pathways that passed the false discovery rate (FDR) correction with a corrected $P < 0.05$ (Storey and Tibshirani, 2003). We visualized pathway-level results in a heatmap using the *pheatmap* R package v1.0.12 (Kolde, 2019) and used the euclidean distance measure with ward.D clustering for dendrogram generation. Results were scaled by row to display z-scores depicting the deviation from the mean. For module-level results, we generated box plots that show median and interquartile ranges of average number of genes per module per genome using the *ggplot2* package v3.3.2 (Wickham, 2016).

RESULTS AND DISCUSSION

Distinct Taxonomic and Functional Diversity Patterns

From our combined sampling efforts in 2017 and 2018, we obtained a total of 11 metagenomes derived from cryopeg brine samples ($n = 6$) and sea-ice brines and section melts ($n = 5$). Assemblies of individual datasets resulted in overall larger and more fragmented assemblies from sea ice than from cryopeg, with sea-ice metagenomes having, on average, more but shorter contigs, and being enriched for singleton contigs and unassembled reads (**Supplementary Table 2**). This difference is likely due to input from a taxonomically richer ($p < 0.01$) and more diverse ($p < 0.01$) community in sea ice than in cryopeg brines (**Figure 1A**). Functional alpha diversity in both environments followed similar trends, but was only significantly higher in sea ice for functional richness ($p < 0.01$) though not functional diversity (**Figure 1B**). Presumably the long residence time and environmental stability of cryopeg brines selected for fewer, but more functionally diverse organisms.

Compositionally, brine communities were clearly separated from sea-ice communities both taxonomically (**Figure 1C**, ANOSIM R = 0.93, $p < 0.01$) and functionally (**Figure 1D**, ANOSIM, R = 0.94, $p < 0.01$). In sea ice, the strongest drivers of the taxonomic variance from cryopeg brines were members of *Bacteroidetes*, mainly *Polaribacter* and unclassified *Flavobacteriaceae*, as well as the gammaproteobacterial genera *Paraglaciecola*, *Glaciecola* and *Colwellia*, and the alphaproteobacterial genus *Octadecabacter* and SAR11 clade (**Figure 1C**)—all were dominant in sea ice and near- (<0.05%) or completely absent from the cryopeg brines (**Figure 2**). In cryopeg brines, *Marinobacter* contributed the most strongly to the variance from sea ice (**Figure 1C**), in keeping with its predominance in the cryopeg samples (46–79% of the prokaryotic community) except for CB4 (4%; **Figure 2**). *Marinobacter* representatives have been observed and isolated from a variety of other cold and saline environments, including Blood Falls, the saline subglacial outflow from Taylor Glacier, Antarctica (Mikucki and Priscu, 2007; Chua et al., 2018; Campen et al., 2019), a cold saline spring in the Canadian High Arctic (Niederberger et al., 2010; Trivedi et al., 2020), and Arctic sea ice (Zhang et al., 2008). Experimental work on isolated members (reviewed by Handley and Lloyd, 2013), as well as genomic insights into the functional capabilities of *Marinobacter* members (Singer et al., 2011), attribute to this genus a remarkable versatility in growth and energy acquisition strategies and a high genomic potential to adapt to challenging environmental conditions, which may have enabled *Marinobacter* to compete effectively for resources and reach high relative abundance (**Figure 2**) despite the extreme conditions in cryopeg brine.

Cryopeg brine CB4 differed markedly from the other cryopeg brines, with *Psychrobacter*, *Shewanella*, and *Carnobacterium* as drivers of this compositional distinction (**Figures 1C**, **2**). From previous work, we learned that CB4 was also unique in its high fraction of dividing cells (29% compared to 1.1–2.5% in the other cryopeg brines; **Table 1**), suggesting that the CB4 community was growing more robustly. We were able to obtain a metatranscriptome from CB4 (the only metatranscriptome obtained from cryopeg brines; **Supplementary Table 1**), which revealed that *Shewanella*, *Psychrobacter*, and also *Marinobacter* (despite comparably low relative abundance in the corresponding CB4 metagenome) were the most transcriptionally active members in CB4 (**Supplementary Figure 2B**), and together accounted for >90% of the active community.

In the sea-ice metatranscriptome (SB_2018_MetaT; **Supplementary Table 1**), we found an overall larger number of active taxa, but the most transcriptionally active

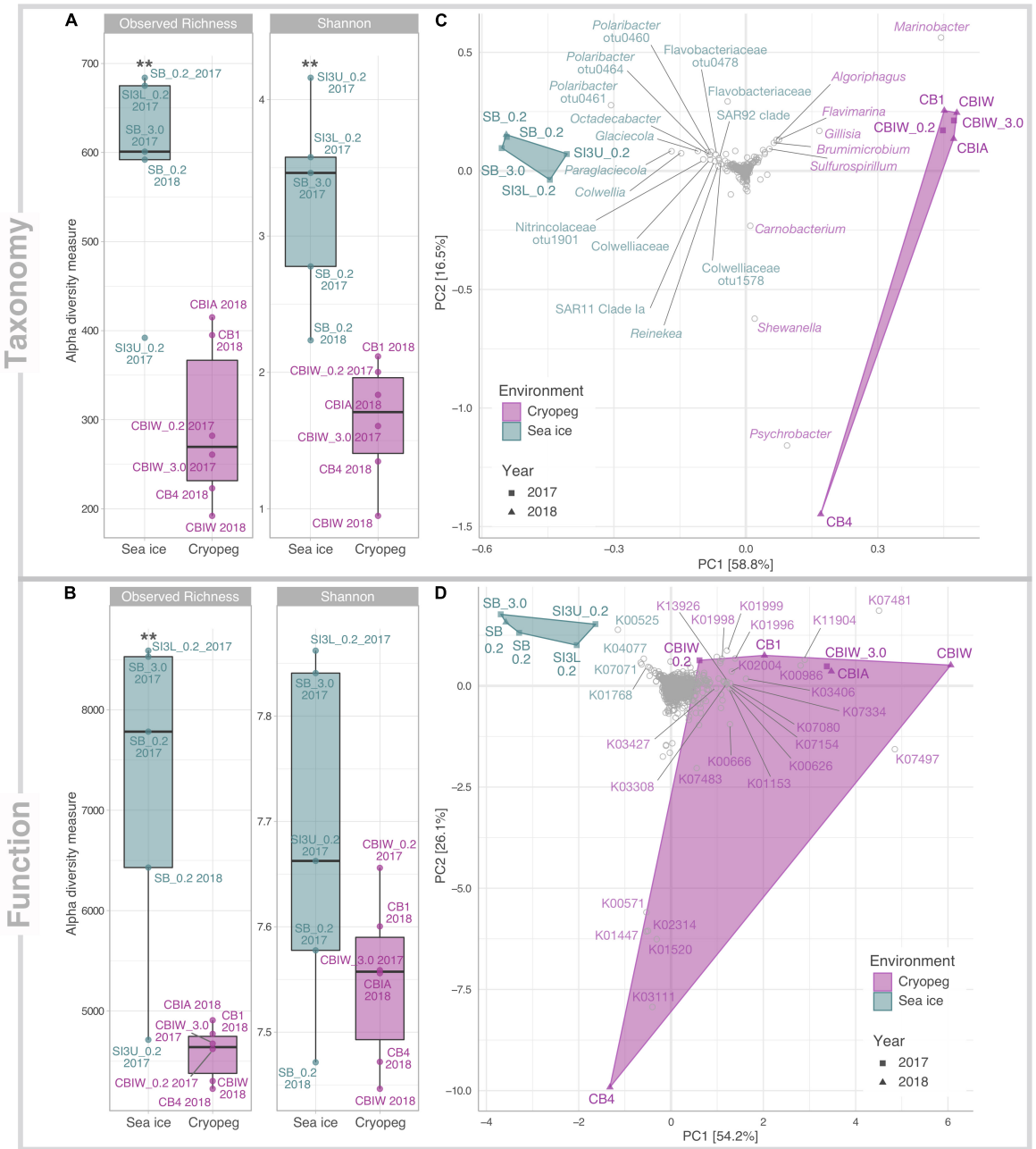

FIGURE 1 | Alpha and beta diversity of prokaryotic community composition and functional potential in sea ice and cryopeg brines. Indices of **(A)** taxonomic observed richness and Shannon diversity on the basis of taxonomic assignments inferred for extracted 16S rRNA gene reads; and **(B)** observed richness and diversity of functional potential on the basis of normalized KO abundances. Asterisks indicate significant variance between means of both environments (** indicates $p < 0.01$). Beta diversity of **(C)** taxonomic composition and **(D)** functional potential was visualized through principal component analysis (PCA). The relative contribution (eigenvalue) of each axis to the total inertia in the data is indicated in percent at the axis titles. For **(C)** data were transformed initially by applying a Hellinger transformation; for **(D)** data were normalized through MUSiCC (see section "Materials and Methods"). An additional layer of information was added in the form of species scores to specify **(C)** prokaryotic taxa and **(D)** molecular functions that contributed most strongly to the variance between samples. K00525: ribonucleoside-diphosphate reductase alpha chain; K00571: site-specific DNA-methyltransferase (adenine-specific); K00626: acetyl-CoA C-acetyltransferase; K00666: fatty-acyl-CoA synthase; K00986: RNA-directed DNA polymerase; K01153: type I restriction enzyme, R subunit; K01447: N-acetylmuramoyl-L-alanine amidase; K01520: dUTP pyrophosphatase; K01768: adenylate cyclase; K01996: branched-chain amino acid transport system ATP-binding protein; K01998: branched-chain amino acid transport system permease protein; K01999:branched-chain amino acid transport system substrate-binding protein; K02004: putative ABC transport system permease protein; K02314: replicative DNA helicase; K03111: single-strand DNA-binding protein; K03308 neurotransmitter:Na + symporter, NSS family; K03406 mcp; methyl-accepting chemotaxis protein; K03427: type I restriction enzyme M protein; K04077: chaperonin GroEL; K07017: uncharacterized protein; K07080: uncharacterized protein; K07154: serine/threonine-protein kinase HipA; K07334: toxin HigB-1; K07481: transposase, IS5 family; K07483: transposase; K07497: putative transposase; K11904: type VI secretion system secreted protein VgrG; K13926: ribosome-dependent ATPase.

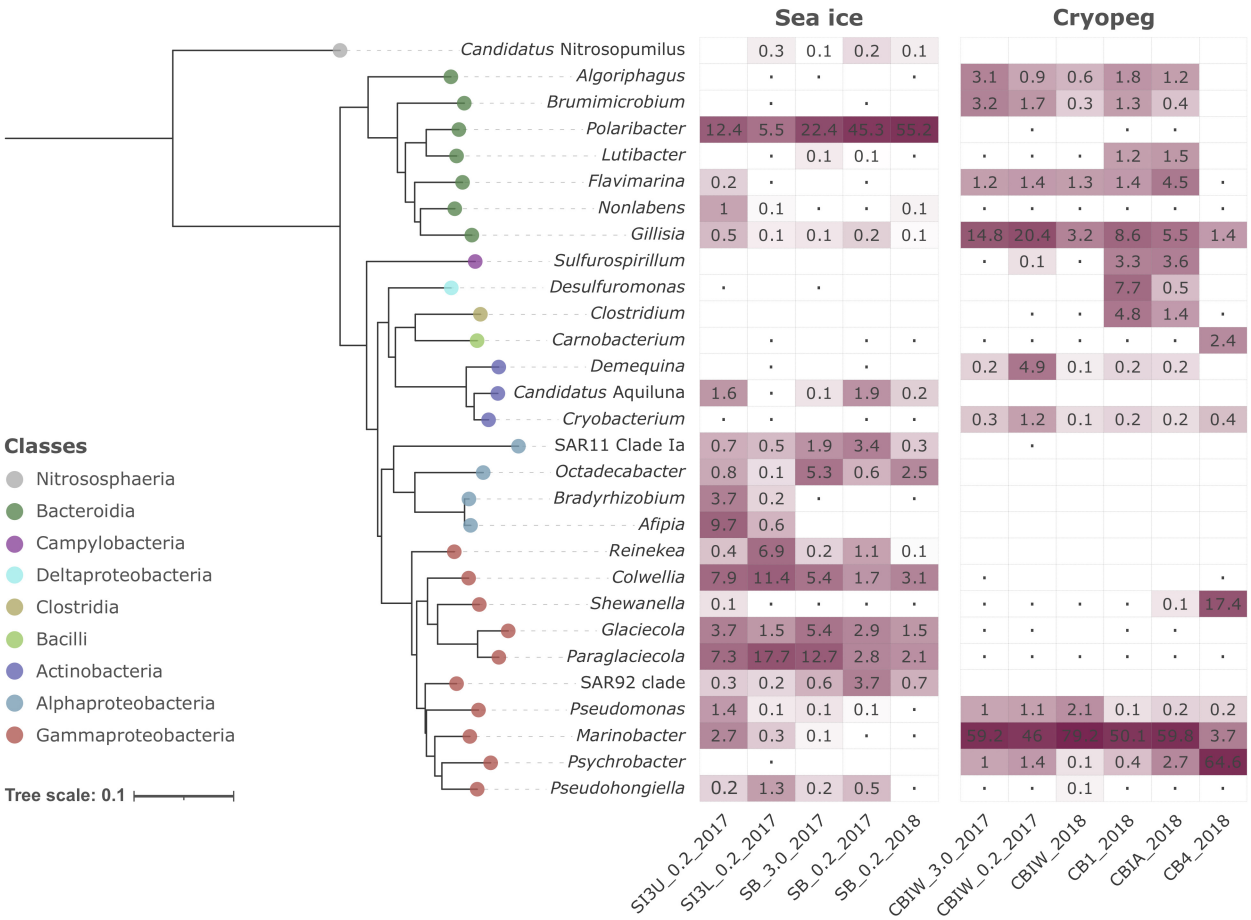

FIGURE 2 | Heatmap of prokaryotic community composition in sea-ice and cryopeg brine metagenomes. Dominant genera (>1% relative abundance in at least one sample) are organized in the form of a Neighbor-Joining phylogenetic tree reconstructed on the basis of full-length 16S rRNA genes. *Candidatus* Nitrosopumilus was chosen as outgroup. The Jukes-Cantor model and 1,000 bootstraps were used for tree calculation. Dots symbolize values <0.05%; empty cells indicate no detection of that genus in a particular sample.

members, *Polaribacter*, *Paraglaciecola*, *Colwellia*, and *Glaciecola* (**Supplementary Figure 2B**), were also among the most abundant in the sea-ice metagenomes (**Figure 2**). The high abundance of the animal-associated *Fusobacterium* genus to the metatranscriptome (SB_2018_MetaT) and not to its corresponding metagenome (SB_0.2_2018; **Figure 2**) may be a result of different treatments, as the sea-ice metatranscriptome was not prefiltered (for rapid treatment in the field), like its paired metagenome, and was thus heavily influenced by eukaryotic signals (**Supplementary Figures 2A, 3**). Archaea were rare in both brine environments (max. 0.2% in SB_0.2_2017, max. 0.02% in CB4_2018; **Supplementary Figure 2A**), which we attribute to the competitive advantage of the heterotrophic bacterial groups, known to be efficient utilizers and recyclers of organic matter, present in these brines of relatively high organic content at the time of sampling (**Table 1**).

Differences in oxygen levels in the two studied brine environments would also contribute to shaping the resident microbial communities and likely add to the observed strong compositional divergence. Sea ice is typically well-oxygenated, and some evidence from previous work suggests that cryopeg habitats are anoxic or microaerophilic (Gilichinsky et al., 2003; Shcherbakova et al., 2009; Colangelo-Lillis et al., 2016). We lack *in situ* oxygen measurements to verify oxygen levels at our sampling sites, but our data on community structure are consistent with these previous findings as most of the dominant cryopeg community members detected here are known to be facultative anaerobes or anaerobes, e.g., *Marinobacter*, *Shewanella*, *Psychrobacter*, *Brumimicrobium*, *Algoriphagus*, *Sulfurospirillum*, *Desulfuromonas*, *Carnobacterium*, and *Demequina*, while the dominant sea-ice members are known to require oxygen, e.g., *Polaribacter*, *Glaciecola*, *Paraglaciecola* and *Colwellia*.

Functionally, there were also clear drivers of structure across these datasets. For sea-ice communities, these included genes involved in DNA synthesis (K00525: ribonucleoside-diphosphate reductase), chaperone-mediated and stress-associated protein folding (K04077: chaperonin GroEL) (Rodrigues et al., 2008), and purine metabolism (K01768: adenylate cyclase) (**Figure 1D**). For cryopeg brines, strong separation resulted from transposases

(e.g., K07481 and K07497) and genes that may function in modulating competition (e.g., K11904: type VI secretion system secreted protein VgrG) and defense (e.g., K03427: type I restriction enzyme M protein). Separation of CB4 from the other cryopeg brines was driven by multiple functions involved in DNA replication, synthesis and repair (K02314: replicative DNA helicase; K03111: single-strand DNA-binding protein; K01520: dUTP pyrophosphatase; K00571: site-specific DNA-methyltransferase, adenine-specific; **Figure 1D**), in further support of a proliferating community in CB4. The following sections explore the broader implications of these functional differences.

Common Adaptations, Different Implementation Strategies

To evaluate genomic adaptations relevant to inhabiting the two subzero environments sampled, we first identified and compared features previously reported for cold-adapted microorganisms. Our analysis approach allowed us to estimate average genomic copy numbers of genes per cell and thus to identify *in silico* features commonly enriched in both sea ice and cryopeg brines. Results confirmed that communities in both brines encode molecular functions relevant for all of the key cold-adaptation features introduced above. We also observed clear differences, however, in average gene abundance and implementation strategies. Compared to sea-ice communities, microbes in cryopeg brines invested significantly more of their genomic content in the transport of phospholipids and biosynthesis of fatty acids, peptidoglycan, lipopolysaccharides and lysine (**Figures 3–5**), all of which represent building blocks of the cell envelope. High numbers of cell envelope genes have been observed in other cold-adapted organisms (Médigue et al., 2005; Ting et al., 2010) and suggested to reflect a mechanism for rapid response to environmental changes and damage to the cell membrane (Tribelli and López, 2018). The permafrost bacterium *Exiguobacterium sibiricum* forms a thicker cell wall under subzero conditions as a result of higher expression of both peptidoglycan and lysine biosynthesis genes (Rodrigues et al., 2008).

A well-described mechanism to maintain membrane fluidity at low temperature involves fatty acid desaturases, which introduce unsaturated, polyunsaturated, and branched-chain fatty acids to membrane lipids (Methé et al., 2005; Ting et al., 2010; Feng et al., 2014). We found that both sea-ice and cryopeg communities encoded such enzymes at similar abundances, though the sea-ice samples displayed greater diversity of desaturases (**Supplementary Table 3**), likely reflective of the greater diversity of taxa (**Figure 1A**). In the directly collected sea-ice brines, we also observed a higher abundance of genes involved in the biosynthesis of carotenoid pigments (**Figure 3**). Such genes may represent another means to regulate membrane fluidity and tolerate freeze-thaw stress (Seel et al., 2020), and coincided with highest relative abundance of *Polaribacter* (**Figure 2**), a genus known for its yellow or orange colored members owing to carotenoid pigment synthesis (Staley and Gosink, 1999; Bowman, 2006). As carotenoid pigments are also involved in adjusting to UV light exposure and protecting against oxidative stress (Dieser et al., 2010), they may be serving multiple functions in the UV radiation-exposed sea-ice environment, but primarily cold-adaptive roles in the permanently dark cryopeg brines (Seel et al., 2020).

In both studied environments, microbes were experiencing subzero temperatures simultaneously with high salt concentrations: 112–140 ppt in cryopeg brines and 75–78 ppt in sea-ice brines (**Table 1**). Although the genetic potential to produce and scavenge compatible solutes for osmoprotection (Kempf and Bremer, 1998; Wood et al., 2001) characterized both systems, we observed different strategies with respect to which osmolytes were favored. While cryopeg communities were well equipped for the synthesis of betaine and ectoine (**Figure 4**), those in sea-ice brine showed a higher potential for synthesizing the polyamines putrescine and spermidine (**Figure 4**). In contrast, cryopeg brine communities appeared to depend on exogenous putrescine and spermidine, being better equipped for transport (**Figure 5**) than synthesis of these compounds (**Figure 4**), as well as for uptake of glycine betaine, proline and various other amino acids from the environment (**Figure 5**). Sea-ice communities encoded on average a significantly higher number of genes for putative transport systems of simple sugars (**Figure 5** and **Supplementary Table 4**), which may serve as compatible solute or substrate (Welsh, 2000). These results from the metagenomes were largely supported by those obtained from the metatranscriptomes, with higher expression of transport systems for glycine betaine, spermidine, putrescine, and various amino acids in the cryopeg brine, and higher expression of glycerol transporters in sea-ice brine (**Supplementary Tables 5, 6**).

We interpret these findings to highlight the contrasting selective pressures in the two environments as follows. The continuous environmental pressure of saltier conditions in the cryopeg brines may favor diverse scavenging capabilities per cell in order to ensure sufficient acquisition of osmolytes. In contrast, microbes in sea ice may be served well enough by a narrower set of solute transport and biosynthesis capabilities, particularly if those solutes are replenished during the seasonal fluctuations that characterize the ice. The exogenous osmolytes (e.g., choline and glycine betaine) readily released by sea-ice algae during seasonal salinity stress (Ikawa et al., 1968; Torstensson et al., 2019; Dawson et al., 2020) would serve to lower the selective pressure on scavenging versatility.

Being Economical and Planning Ahead

Microbes have evolved multiple strategies to overcome unfavorable environmental conditions, including ways to persist through periods of suboptimal nutrition, from compound storage and salvage to the use of metabolic shunts. During evaluation of our samples by epifluorescence microscopy (to obtain the counts reported by Cooper et al., 2019), we observed the presence of bright yellow circular structures evenly spaced within many of the cells in cryopeg brines (**Supplementary Figure 4**) though not in sea ice. We considered that these structures might be polyphosphate granules, known to serve both as versatile modulators of microbial stress responses and for energy and

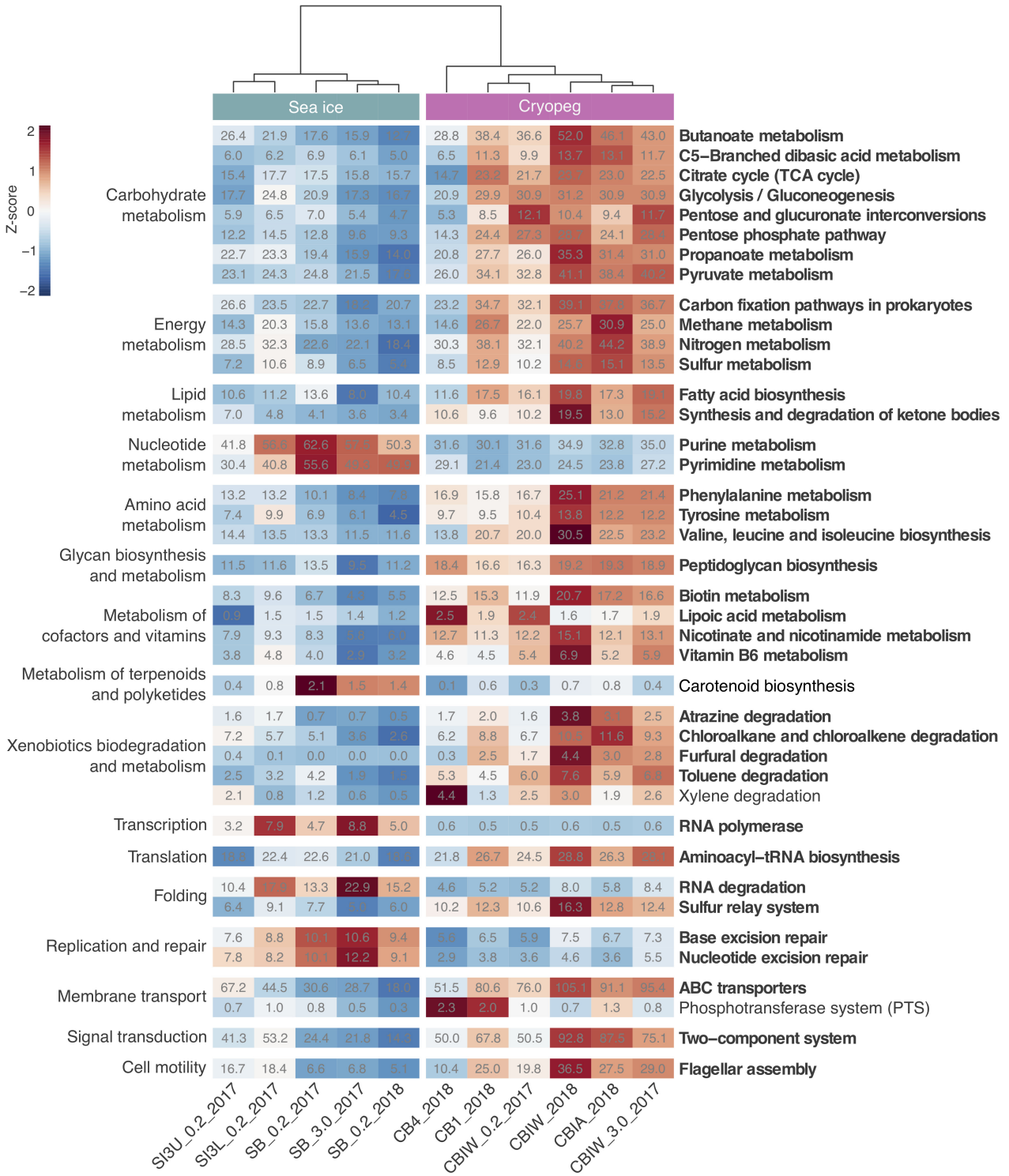

FIGURE 3 | Heatmap of prokaryotic functional potential in sea ice and cryopeg brines summarized at the level of pathways involved in metabolism, genetic and environmental information processing, and cellular processes. Sample order was determined by the calculation of a dendrogram, using the euclidean distance method and ward.D clustering. Cell values display the normalized average number of pathway genes per genome in the sample (see section "Materials and Methods" for details), while cell color represents the value's standard score (z-score), and therefore its deviation from the row mean. Displayed are pathways that were common to sea-ice and/or cryopeg microbes, i.e., detected across all sea-ice and/or all cryopeg samples with a mean abundance of >1. Those that differed significantly between brine types are in bold; pathways that did not differ significantly but showed a log2-fold difference in mean abundance of > 1 between the environments are also shown.

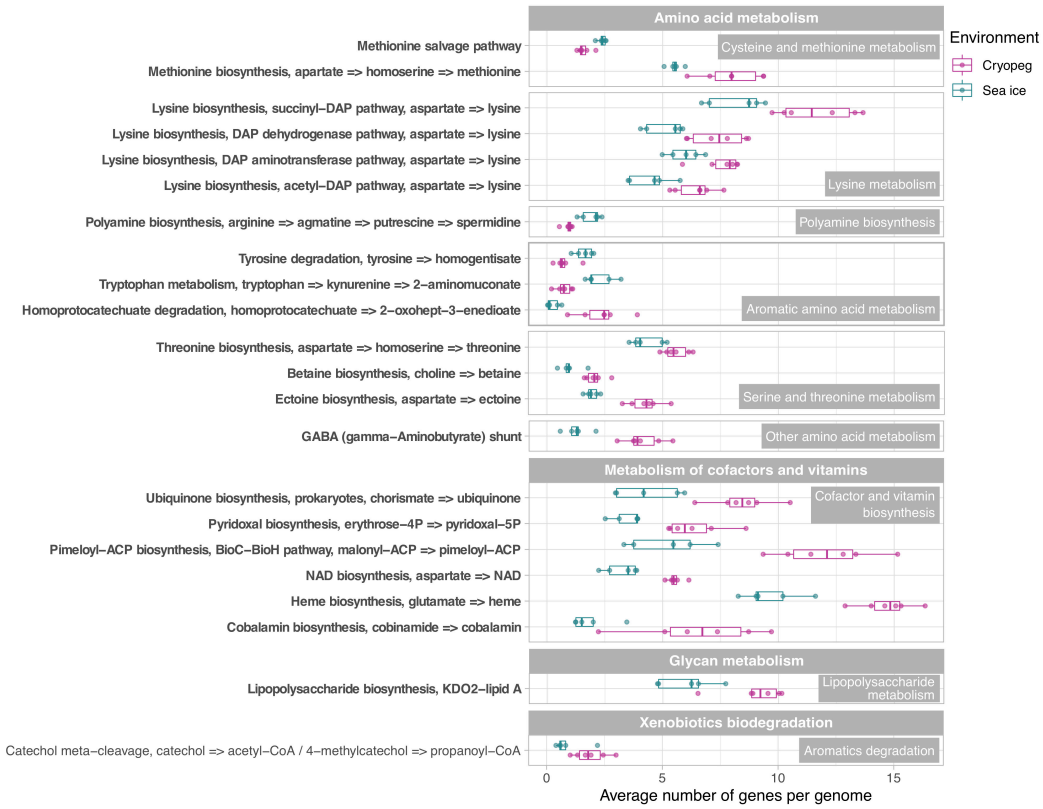

FIGURE 4 | Box and whisker plots summarizing the functional potential of prokaryotic communities in sea ice and cryopeg brines for the metabolism of amino acids, cofactors and vitamins, lipopolysaccharides, as well as the degradation of aromatics. Boxes indicate the median of each group, as well as the first and third quartiles of sample distribution (the 25th and 75th percentiles). Whiskers extend from the hinge to the largest value no further than 1.5 × IQR from the hinge (where IQR is the interquartile range, or distance between the first and third quartiles). Displayed are modules that were common to sea-ice and/or cryopeg microbes, i.e., detected across all sea-ice and/or all cryopeg samples with a mean abundance of >1. Those that differed significantly between brine types are in bold; modules that did not differ significantly but showed a log2-fold difference in mean abundance of >1 between the environments are also shown.

phosphate storage, with links to biofilm development through exopolymer production (reviewed by Seufferheld et al., 2008). We found that the key enzyme for polyphosphate biosynthesis and degradation, polyphosphate kinase 1 (PPK1) (Rao et al., 2009), was significantly more abundant in cryopeg brine than in sea-ice metagenomes (**Supplementary Table 7**), and only expressed in the cryopeg metatranscriptome (**Supplementary Table 5**). Genes for phosphate transport systems, as well as the two-component phosphate starvation response (**Figure 5**), were also significantly more abundant in cryopeg brines and paired with high expression. Although their precise roles require further study, polyphosphates may represent a dual strategy for both intracellular storage and stress tolerance, and the possible production of protective exopolymers (Fraley et al., 2007), which were present at much higher concentrations in the cryopeg brines than in sea ice (millimolar vs. micromolar; **Table 1**).

Other storage polymers of potential relevance to inhabiting subzero brines include polyhydroxyalkanoates (PHAs), nitrogen reserve polymers and cyanophycin. The relevant genes for PHA synthesis and breakdown, PHA synthase, 3−hydroxybutyrate dehydrogenase and PHA depolymerase, were present in both brine types, with no significant differences, but we observed higher gene abundances for the synthesis and degradation of cyanophycins in sea-ice communities. Cyanophycins are amino acid polymers (polyamines) that cyanobacteria and some heterotrophic bacteria, including *Octadecabacter* (Vollmers et al., 2013; **Figure 2**), can accumulate as cytoplasmic granules during light stress and nutrient starvation, serving as nitrogen- and possibly carbon-storage compounds (Krehenbrink et al., 2002). Biosynthesis genes for other polyamines, such as spermidine, were also more abundant in sea ice (**Figure 4**; see above). Genome analysis of cold-adapted *Colwellia psychrerythraea* had underscored the capacity for synthesis and degradation of polyamines (Methé et al., 2005) and its likely importance for microbial persistence in cold environments during unfavorable nutrient conditions. Other compatible solutes that contain nitrogen, including choline, glycine betaine and ectoine, can also serve as valuable organic substrates (Firth et al., 2016). Respiration of such compounds in cryopeg brines might explain the high (millimolar) concentrations of $NH4^+$ measured in the brines (**Table 1**).

Likewise, the use of the γ-aminobutyrate (GABA) shunt may constitute an economical solution to thriving under stressful and potentially unfavorable periods of nutrient limitation in cryopeg

FIGURE 5 | Box and whisker plots summarizing the functional potential of prokaryotic communities in sea ice and cryopeg brines for environmental information processing, including transport of compounds across cell membranes *via* ABC transport systems, use of bacterial secretion systems, and the ability to sense and respond to environmental cues *via* two-component systems. Boxes indicate the median of each group, as well as the first and third quartiles of sample distribution (the 25th and 75th percentiles). Whiskers extend from the hinge to the largest value no further than 1.5 × IQR from the hinge (where IQR is the interquartile range, or distance between the first and third quartiles). Displayed are modules that were common to sea-ice and/or cryopeg microbes, i.e., detected across all sea-ice and/or all cryopeg samples with a mean abundance of >1. Those that differed significantly between brine types are in bold; modules that did not differ significantly but showed a log2-fold difference in mean abundance of >1 between the environments are also shown.

brines, where we observed significantly higher gene abundances for this pathway than in sea ice (**Figure 4**). This shunt allows microbes to bypass two steps of the tricarboxylic acid (TCA) cycle by producing succinate from glutamate (Feehily et al., 2013), while also aiding in osmotic protection, as both glutamate and GABA can serve as compatible solutes (Csonka, 1989).

In sea-ice communities, we found a significantly greater potential for microbes to be economical by recycling sulfur-containing compounds through the methionine salvage pathway (**Figure 4**). Although ionic analyses showed that sulfur was not scarce in either brine type at the time of sampling (**Supplementary Figure 5**), the resident microorganisms may differ in the forms of sulfur utilized, as they differ in their potential use of compatible solutes (see above). For example, members of the SAR 11 clade, a ubiquitous marine taxon abundant in our sea-ice samples (**Figure 2**), lack known genes for assimilatory sulfate reduction and instead rely entirely on exogenous organic sources of reduced sulfur, such as methionine or dimethylsulphoniopropionate (DMSP), to meet their sulfur needs (Tripp et al., 2008). DMSP, often found in high concentrations in springtime sea ice (Kirst et al., 1991), provides a major source of both carbon and reduced sulfur for marine microbes (Reisch et al., 2011; Wilkins et al., 2013). The capacity for methionine salvaging would allow sea-ice microbes to exploit this additional pool of organic sulfur to fulfill their needs when DMSP is scarce. In contrast, we observed significantly higher abundances of genes for sulfate transport (**Figure 5**) and methionine biosynthesis (**Figure 4**) in cryopeg brines, and for sulfur metabolism in general (**Figure 3**), suggesting that economical sulfur recycling may not be required to inhabit cryopeg brines.

Genomic Versatility and Plasticity for Competitive Advantage

While maintenance of essential cellular processes and retention of membrane fluidity for solute exchange ensure survival in subzero brines, competitive advantage may be acquired through accessory traits, e.g., the ability to sense environmental stimuli and changes, and be metabolically versatile to adjust to microniche requirements. We observed strong disparities in the accessory genomic potential encoded by communities in the two brine types, with some of the most marked differences in their environmental information-processing capacities (**Figure 5**). The genomes of cryopeg brine taxa encoded on average 71 genes for two-component regulatory systems, more than twice as many as sea-ice genomes (**Figure 3**). These simple signal transduction pathways allow bacterial cells to sense a variety of environmental stimuli and respond to changes, primarily by regulating transcription and gene expression (Wuichet et al., 2010). Specifically, our analysis showed that gene abundances in cryopeg were significantly higher for the sensing of changes in substrate availability (DctB/DctD, TctE/TctD), cellular redox state (RegB/RegA), quorum signals (QseC/QseB), envelope stress and antibiotics (CpxA/CpxR, RstB/RstA), as well as changes in chemical cues that induce the regulation of alginate, lipopolysaccharide, and hydrogen cyanide production, twitching and swarming motility, biofilm formation (AlgR/AlgZ, ChpA/ChpB), type IV pili expression (PilS/PilR), motility and biofilm formation during nitrate respiration (NarX/NarL), flagellar movement (FlrB/FlrC), potassium homeostasis (KdpD/KdpE), phosphate level regulation and virulence (PhoR/PhoB) by upregulation of bacterial secretion systems type III and VI (**Figure 5**).

Genes encoding two-component systems on average account for 1–2% of a bacterial genome (Salvado et al., 2015), with variation best explained by differences in genome size, lifestyle complexity and exposure to environmental fluctuations (Rodrigue et al., 2000; Alm et al., 2006). For example, the highest number of genes for two-component systems (251 genes corresponding to 3.4% of protein-coding genes) has been reported for free-living bacteria with complex life cycles and large genomes such as *Myxococcus xanthus* (Shi et al., 2008), while few have been detected in the reduced genomes of intracellular pathogens (Alm et al., 2006). By comparing gene abundances determined here (**Figure 3**) to the number of protein-coding genes reported for representative genomes[1] of dominant community members (**Supplementary Figure 6**), we estimated the average fraction of genes for signal transduction to be 0.9% in sea-ice and 2.1% in cryopeg brine genomes. While below average abundances under the fluctuating environmental conditions in sea ice appear counterintuitive, they may be explained by the presence of several oligotrophic taxa with small and streamlined genomes, e.g., *Candidatus* Aquiluna and the SAR11 and SAR92 clades (**Figure 2**), for which a lack or very low numbers of two-component systems have been described (Held et al., 2019). In contrast, above average gene abundances in cryopeg microbes may represent an adaptation to diverse microbial niches, generated in part by the microscale architectures of sediment particles and exopolymers, which may develop and evolve on a shorter time frame than the larger, geophysically stable cryopeg system. A versatile suite of environmental sensing genes would enhance their capacity to exploit localized nutrient patches whenever they become available, e.g., through cell lysis, and facilitate long-term survival in this isolated environment, with little or no external input of organic matter.

Competitive advantage may also be a result of genomic plasticity mediated by mobile genetic elements and transposition processes (Bennett, 2004; Frost et al., 2005; DiCenzo and Finan, 2017). We identified transposase genes as the most abundant genes in cryopeg brines, and among the most abundant in sea ice (**Supplementary Tables 7, 8**), in line with previous research that identified transposases as the most prevalent and ubiquitous genes in nature (Aziz et al., 2010). Overall, the normalized abundance of KOs associated with known and putative transposases was more than twice as high in cryopeg (on average 55 genomic copies or 2.3% of all KOs) than in sea ice (23 genomic copies or 1.5% of all KOs), though differential abundances between individual types of transposases were mostly non-significant due to high within-group variation (**Supplementary Table 8**). Outside these systems, the number of

[1] Obtained from https://ftp.ncbi.nlm.nih.gov/genomes, accessed March 2020.

transposase genes per genome can be highly variable, ranging from zero in the small, streamlined genomes of open ocean picocyanobacteria to over a thousand copies in large genomes of filamentous cyanobacteria (Vigil-Stenman et al., 2015), with 38 copies per genome as an estimated average when present (Aziz et al., 2010). Transposase gene copy numbers are generally higher, however, in microorganisms from extreme, stressful or fluctuating environments, e.g., deep and low oxygen waters (DeLong et al., 2006; Konstantinidis et al., 2009), permafrost and ice wedge settings (Raymond-Bouchard et al., 2018), and the dynamic estuarine waters of the Baltic Sea (Vigil-Stenman et al., 2017), as well as in dense assemblages of deep-sea biofilms (Brazelton and Baross, 2009) and particulate matter (Ganesh et al., 2014; Vigil-Stenman et al., 2017). Transposition events can be stress-induced (Casacuberta and González, 2013), causing the advantageous rearrangement, activation or enrichment of a cell's gene content in response to environmental challenges (Aziz et al., 2010), which helps to explain their higher abundance in extreme environments. We detected active transcription of transposases in both studied brine environments at the time of sampling, but expression values in the cryopeg brine far exceeded those in sea ice, up to 46 times higher (**Supplementary Tables 5, 6**). In highly concentrated communities, as encountered in cryopeg brines (**Table 1**), frequent gene exchange and selection pressure to diversify in order to compete with neighboring cells may favor the retention of high copy numbers of transposase genes, which are at the upper end of abundances reported for metagenomic datasets (Brazelton and Baross, 2009). We further hypothesize that the presumably slow growth rates of cryopeg communities allow for transposases to accumulate, despite the possible detrimental effects of transposition to a genome through gene disruption. Gene duplication resulting from replicative transposition may also explain the observed higher gene copy numbers of several accessory features in cryopeg brine communities (**Figure 5**).

Previous studies have repeatedly reported a significant enrichment of transposons on plasmids, where they are often encoded together with accessory traits, but few core functions (Siguier et al., 2014; DiCenzo and Finan, 2017). We observed several genomic features in our datasets indicating the presence of plasmids in both cryopeg brine and sea ice communities, including genes for at least five of the toxin-antitoxin systems (**Supplementary Table 9**) that have been linked exclusively to plasmids (Jensen and Gerdes, 1995; Hayes, 2003), with several of these being actively transcribed at the time of sampling (**Supplementary Table 9**). Their activities can play important roles in plasmid stabilization through a process called "post-segregational killing," where plasmid loss can result in either growth arrest or cell death and, thus, selects against plasmid-free cells (Jensen and Gerdes, 1995; Harms et al., 2018). Plasmid-carrying representatives have been reported for several of the dominant genera in both brine types[2], including *Marinobacter* in cryopeg brines, though not *Polaribacter* in sea ice (**Supplementary Figure 6**). Plasmids could thus be of greater relevance in cryopeg brine than in sea ice, adding extra plasticity to the cryopeg gene pool and facilitating the close co-occurrence of diverse genotypes (and phenotypes) in this stable but likely heterogeneous environment on the microscale. Both reduced negative selection pressure on plasmids and high occurrence of transposable elements favor the acquisition of novel genes, suggesting that plasmids contribute to an organism's ability to adapt to unique niches (DiCenzo and Finan, 2017). As gained fitness advantages can be transferred horizontally (Frost et al., 2005), the presence of plasmids in cryopeg brine communities may help to explain the higher average genomic copy numbers of individual bacterial secretion (**Figure 5**) and toxin-antitoxin systems (**Supplementary Table 9**), multidrug resistance pumps (**Supplementary Table 10**), and xenobiotic degradation modules (**Figure 3**). Future in-depth analysis of our datasets targeting plasmid sequences may provide a test of this hypothesis.

Elaborate Competitive Strategies at High Cell Densities in Subzero Brines

Within the brine system of a frozen matrix, concentrated microbes may be competing not only for resources but also for space, and therefore could benefit from the ability to sense critical cell density, relocate, and fend off competitors. Indeed, both the sea-ice and cryopeg communities encoded genes to produce, transport and modify diverse antimicrobial compounds, mainly antibacterial (**Supplementary Table 10**). The use of bacteriocins appears particularly important in cryopeg brines where we observed significantly higher numbers of transporter genes for microcin C (**Figure 5**), encoded by the *yejABEF* operon (Piskunova et al., 2017) and fully transcribed in the cryopeg metatranscriptome (**Supplementary Table 5**). Bacteriocins are often produced under stress conditions such as nutrient limitation and overpopulation (Riley and Wertz, 2002), and differ from other antibiotics in their target range, as they are toxic only to bacteria closely related to the producing strain (Riley and Wertz, 2002). We found that the majority of scaffolds encoding the complete operon were affiliated with the dominant *Marinobacter* and *Psychrobacter* genera in the cryopeg brines, and could further link the comparably high numbers in sea-ice sample SI3U to the dominant *Afipia* genus (**Figure 2** and **Supplementary Table 11**). We thus hypothesize that microcin C is a potent means for dominant community members to regulate population densities, and potentially of greater relevance in the relatively low diversity cryopeg brines (**Figure 1**). Besides killing target cells by inhibiting essential enzymes or damaging the inner membrane (Piskunova et al., 2017), an additional relevant role for microcin C may be as a "public good" signal, where exposure induces persistence in sensitive, non-producing cells and thereby increases community resilience to various stressors (Piskunova et al., 2017).

Similarly, the activity of toxin-antitoxin systems serves an autoregulatory purpose in response to stressors (Harms et al., 2018). We detected at least 20 different types of toxin-antitoxin systems across the sea-ice and cryopeg brine metagenomes, with the great majority present in both environments yet of higher average genomic copy number in cryopeg brines (**Supplementary Table 9**). Multiplicity of these systems appears

[2]According to https://ftp.ncbi.nlm.nih.gov/genomes/GENOME_REPORTS/plasmids.txt, accessed March 2020.

to be correlated to a free-living slowly growing lifestyle, and has been previously linked to the presence of mobile genetic elements (Pandey and Gerdes, 2005), in support of our observation of higher genomic copy numbers of transposases in cryopeg brines. Among the most abundant of these systems was the HipA-HipB system, which upon antibiotic exposure exerts reversible growth inhibition that coincides with increased antibiotic tolerance, thus directly implicating this module in persister formation (Harms et al., 2018; Huang et al., 2020). The toxin HipA was significantly more abundant in cryopeg brines, with multiple copies encoded per genome, indicating a greater relevance of persistence. Others, like the MazE-MazF system, which was encoded on average once per genome in cryopeg brines (but on average only once per 12 genomes in sea ice), can ultimately induce programmed cell death in part of the population (Peeters and de Jonge, 2018). In densely populated brines, individual cell death could provide multiple benefits, from limiting the spread of viral infections to releasing nutrition for the remaining population and building blocks for biofilm formation (Peeters and de Jonge, 2018).

Strong selective pressure to evade competition in both systems is also reflected in the diversity of encoded multidrug resistance and efflux pumps (**Supplementary Table 10**), and potential means to modify cell envelope and cell surface. We observed a significantly higher abundance of phosphatidylglycerol lysyltransferase (*mprF*) genes in cryopeg communities, used by bacteria to alter the net charge of their cellular envelope and resist a variety of cationic antimicrobials (Ernst et al., 2009), as well as for osmoprotection (Dare et al., 2014). As phosphatidylglycerol is a major building block of *Marinobacter* membranes (Park et al., 2015; Zhong et al., 2015),

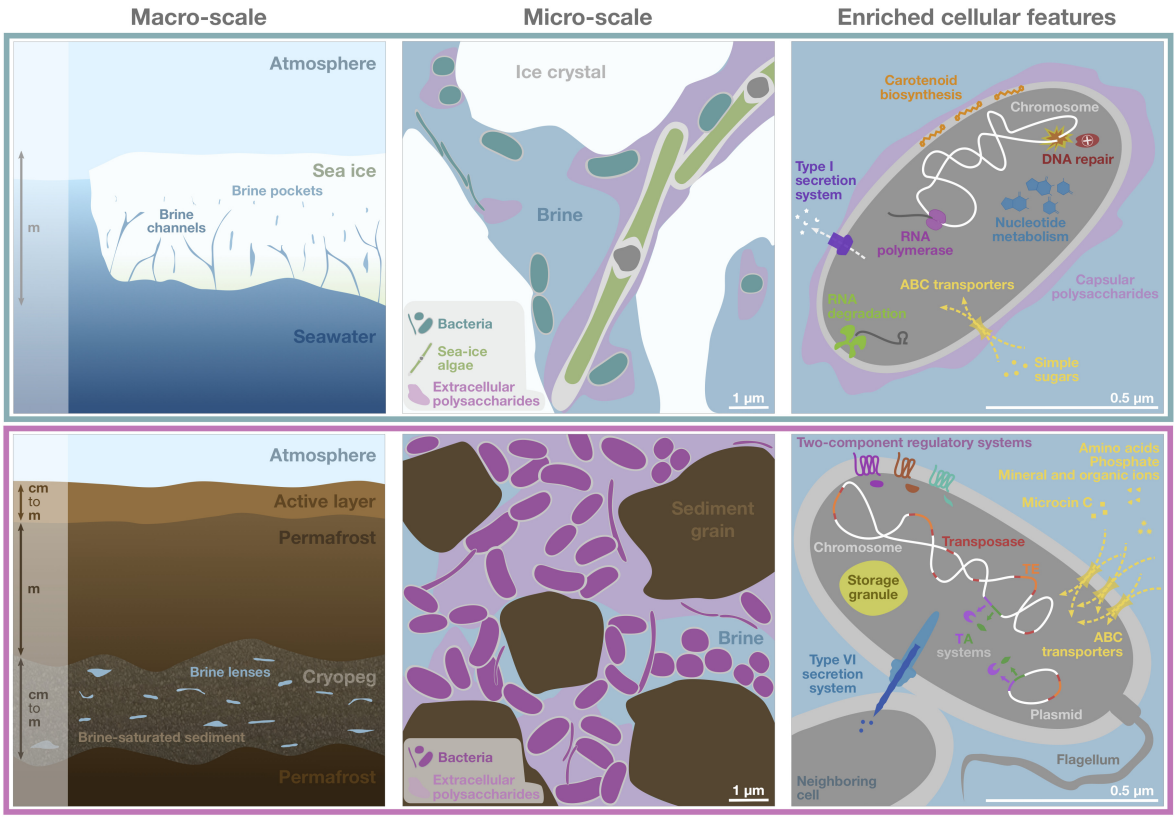

FIGURE 6 | Schematic of genomic adaptations to life in subzero hypersaline sea-ice brines (upper panels) and cryopeg brines (lower panels). Sea ice, exposed to atmospheric conditions and in exchange with underlying seawater (left), is characterized by vertical gradients in temperature, salinity and nutrient availability and subject to temporal disturbances of these conditions. Its internal brine-filled network of pores and channels provides habitat for prokaryotic (primarily bacterial) and eukaryotic life (middle). Bacterial genomes in sea ice, represented by a single generic cell (right), were enriched in functions relevant for DNA replication and repair, transcription, RNA degradation, and nucleotide metabolism, as well as for membrane transport of simple sugars and the export of capsular polysaccharides, secretion of compounds through type I secretion systems and carotenoid biosynthesis. Cryopeg, a subsurface layer of unfrozen sediment within permafrost (left), is isolated from the atmosphere, permanently dark and characterized by relatively stable geophysical conditions across millennia. Brine is present in discrete lenses or as brine-saturated sediment. The cryopeg brines of this study contained high concentrations of bacterial cells, dissolved and particulate organic carbon and extracellular polysaccharides relative to sea-ice brines (middle). Bacterial genomes in cryopeg brines, represented by a generic cell and neighbor (right), were enriched in many accessory features including two-component regulatory and toxin-antitoxin (TA) systems, diverse types of ABC transporters, storage granules, and genes for flagellar assembly. An enrichment of genes for mobile elements, including transposases, transposable elements (TE), and genes exclusively present on plasmids, indicate greater genetic plasticity, while enriched type VI bacterial secretion systems and microcin C transporters attest to complex social interactions, both competitive and beneficial to the community. Canonical genes relevant for cold and salt stress were present in both communities (not shown). Scale bars are estimates for comparative purposes between panels.

including high-salt-adapted members (Zhong et al., 2015), lysylphosphatidylglycerol formation could be an important strategy of the dominant bacteria in cryopeg brines for surviving both strong competition and hypersaline conditions. In sea ice, bacteria may acquire similar protection by producing capsular polysaccharides, as genes involved in transporting and exporting capsular polysaccharide components were more abundant there (**Supplementary Table 4**). These EPS remain intimately associated with the cell surface, where they can act as an antibiotic (and possible viral) shield, while also aiding in surface attachment and biofilm formation (Reckseidler-Zenteno, 2012) or cryoprotection (Carillo et al., 2015).

Cell concentrations in cryopeg brines were orders of magnitude higher than in sea ice (**Table 1**), allowing higher frequency of direct contact between neighboring cells and likely facilitating contact-dependent competitive strategies that circumvent resistance typically provided by cell envelope modifications (Russell et al., 2014). We found significantly higher gene abundances for the type VI secretion system in cryopeg brines than in sea ice (on average six times more; **Figure 5**), which produces a structure often described as a "molecular speargun" that uses a contraction mechanism to fire a toxin-carrying needle into neighboring cells, inducing growth arrest or cell lysis (Russell et al., 2014). Scaffold taxonomy revealed that the vast majority of cryopeg scaffolds encoding a complete type VI system belonged to the *Marinobacter* genus, and that other members of the cryopeg communities, *Shewanella* and *Pseudomonas*, also had the capacity to use this system for attack. In CB4, where no scaffold was assigned to the dominant genus *Psychrobacter* (**Figure 2**), the abundance of type VI genes was much lower than in the rest of the cryopeg samples (**Figure 5**). In sea ice, individual components of the type VI system were found scattered on multiple scaffolds, but we could only detect the complete operon on a few gammaproteobacterial scaffolds, e.g., belonging to *Psychromonas*. Previous research has indicated that sibling cells can transfer and reuse components of the type VI secretion system (Vettiger and Basler, 2016). If also true for *Marinobacter*, this mechanism could have provided members of this genus with increased competitive strength, as both its dominance and high cell densities in cryopeg brine would have increased the number of type VI components among the population. Further, type VI secretion systems have been shown to foster horizontal gene transfer in naturally competent bacteria by releasing DNA from lysed target cells (Borgeaud et al., 2015), including genes encoding new matching sets of toxins and immunity proteins (Thomas et al., 2017). Such a mechanism could have allowed *Marinobacter* and other type VI-encoding cells to successively optimize their strategies for competitive advantage in these extreme subzero brines.

CONCLUSION

Here we have presented the first metagenomic insights into the microbial ecology of cryopeg brines and the first combined exploration of functional potential and active transcription of prokaryotic communities in the subzero brines of both cryopeg and sea-ice environments. Our comparative analyses has provided novel insights into the diverse and divergent molecular strategies used by bacteria in both brine types to withstand the multitude of abiotic and biotic threats posed by their extreme environments (**Figure 6**) and expanded our understanding of the fundamental role that environmental stability plays in microbial adaptation to extreme conditions. A remarkable difference was the higher encoded copy number for many accessory features in cryopeg, particularly traits relevant for sensing and for transport systems. While the two brine types share the selective pressures of hypersalinity and subzero temperature, microbes within the brine inclusions face fundamentally different challenges that we believe to be central to the divergence. Compared to the contemporary sea-ice environment, the ancient cryopeg brines are far less affected by any bulk changes in their physical environment, yet diverse niches can be expected in the form of microscale chemical gradients within a dense EPS matrix and resulting from cell lysis, in many ways resembling the dynamics in bacterial biofilms (Flemming et al., 2016). Rather than a dynamic succession of individual taxa, however, the microorganisms in cryopeg brine communities appear to sustain a versatile genetic makeup that enhances phenotypic plasticity, providing them with the required tools to be highly competitive and dominant, yet also cooperative, i.e., through the sacrifice of individual members for the production of signaling compounds and the greater good nutritionally.

In the densely populated cryopeg brines, the high abundance of transposases together with high genomic copy numbers of numerous accessory traits and indications for the presence of plasmids suggest that these communities may undergo considerable genomic rearrangements, gene duplication events and horizontal gene transfer *via* mobile genetic elements. Horizontal gene transfer, also fostered by the elaborate means of microbial warfare that were detected, could support the inter- and intraspecific co-occurrence of microbial metabolisms despite overall low taxonomic diversity. The constant abiotic pressures in the cryopeg environment, the continuous competition for resources, and the limited pressure for gene deletion (assuming very slow growth) may favor the retention of genomic changes (Guieysse and Wuertz, 2012). We lack *in situ* rate measurements for the cryopeg brine communities, but our metatranscriptomic data paired with previous observations of dividing cells indicate that microorganisms were metabolically active and proliferating *in situ* despite the extreme conditions. Ongoing research aims to resolve microbial metabolisms at the level of individual genomes, and thus provide further insights into taxon-specific phenotypic plasticity, as well as high-resolution strain-level differences. This follow-on work will also address aerobic vs. anaerobic metabolisms as a means to gain insights on the role of oxygen (its presence or absence) in these extreme environments, as obtaining *in situ* oxygen measurements was not possible given the top priority to avoid introducing any contamination, particularly to the geologically isolated cryopeg brines.

In stark contrast are the challenges faced by microorganisms within the brines of sea ice. Frequent disturbance of the physico-chemical conditions imposed by hourly to seasonal fluctuations in atmospheric temperature, paired with steep vertical gradients

in temperature, salinity and nutrient availability, suggests that microbial inhabitants must also be versatile to survive and thrive. However, our results imply that many members of the sea-ice community invest less in the encoding of accessory traits than those in the cryopeg communities, given the lower average copy numbers of individual features. We suggest that a distribution of capabilities across multiple community members, though potentially leading to reduced fitness of some individuals, maintains the versatility of the community as a whole, with the possible benefits of genome streamlining and more rapid replication. The community dynamics in sea ice would then follow a temporal succession, as others have documented (Collins et al., 2010; Eronen-Rasimus et al., 2016), rather than a simultaneous co-occurrence of competing taxa, with only a few members being uniquely adapted to exploit a given set of environmental conditions at a time. The comparatively higher abundance of genes relevant to DNA repair may protect replicating sea-ice taxa against DNA damage induced by exposure to radiation (**Figure 6**), as well as overcome periods of reduced metabolic activity.

Beyond the expected molecular adaptations that allow maintenance of vital cell functions, the survival and prosperity of bacteria in subzero brines appear to be shaped strongly by their abilities to use resources economically, to sense and respond to changes in their microenvironments, and to interact with other community members. The capability to autoregulate and induce growth arrest was a common feature and key mechanism for persistence during unfavorable conditions, including intense competition. An evaluation of the genomic potential of individual community members and efficient techniques to identify actively metabolizing members are needed to better resolve such dynamics in the future. Promising approaches include bioorthogonal non-canonical amino acid tagging (BONCAT), as preliminary results gave positive results in both brine types (Z. Zhong, unpublished data), and subsequent fluorescence-activated cell sorting (FACS) (Hatzenpichler et al., 2020) and sequencing, for specifically interrogating the fraction of brine communities active *in situ*. As environmental conditions change with amplified warming in the Arctic, particularly for historically stable cryopegs in permafrost, knowledge of current microbial dynamics in these unusual communities will be valuable to anticipating their future impacts on the larger Arctic ecosystem. One may speculate that some genes specific to surviving the current extremes and others enabling intensive social interactions, as documented here, may no longer be expressed or maintained. Communities inhabiting fresher brines in thinner sea ice will more closely resemble their source seawater communities, and the inhabitants of cryopeg brines will face new competition as permafrost thaws and the brines begin to merge with the surface environment. We can hypothesize that the strong divergence in functional potential observed between both systems described here may fade as the geophysically isolated and stable cryopeg environment transitions into a more moderate and seasonally fluctuating state. Our metagenomic analyses thus emphasize the ecological importance of habitat stability in the face of climate change.

DATA AVAILABILITY STATEMENT

All generated metagenomes and metatranscriptomes analyzed in this study were made publicly available as raw reads, quality-controlled reads, and assemblies through the IMG system (https://genome.jgi.doe.gov/portal/). For individual accession numbers per sample see **Supplementary Table 2**. Additional environmental metadata are available through our data archives at https://www.ocean.washington.edu/story/Deming_Data_Archives.

AUTHOR CONTRIBUTIONS

JD conceived the study, with input from MS and JR. JR and JD conducted the field work and sampling. JR processed the samples, performed the DNA and RNA extractions, performed all the data analyses, and wrote the first draft of the manuscript. JD and MS supported the interpretation of results and critically revised the manuscript. All authors read and approved the submitted version.

FUNDING

This research was funded by the Gordon and Betty Moore Foundation Grant No. 5488 to JD. Sequencing services and analysis platforms were provided by the U.S. Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, which was supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. Additional High Performance Computing resources were provided by the Ohio Supercomputer Center.

ACKNOWLEDGMENTS

We kindly thank members of the larger team for support in the field and insightful discussion as the results emerged: Shelly Carpenter, Zac Cooper, Max Showalter and Anders Torstensson, Hannah Dawson and Jodi Young, Zhiping Zhong and Dean Vik, Go Iwahana and Hajo Eicken. The Ukpeaġvik Iñupiat Corporation science team provided essential logistical support and helped with access to the tunnel and the sea-ice sampling site. We thank Natalie Solonenko for performing DNA extractions of the samples collected in 2017, Zac Cooper for his help with DNA extractions in 2018, Shelly Carpenter for the micrographs in **Supplementary Figure 4**, and Zhiping Zhong for sharing the unpublished BONCAT results mentioned in our discussion. Ionic analyses depicted in **Supplementary Figure 5** were provided by Dongsen Xue in the University of Washington Analytical Service Center.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.701186/full#supplementary-material

REFERENCES

Alm, E., Huang, K., and Arkin, A. (2006). The evolution of two-component systems in bacteria reveals different strategies for niche adaptation. *PLoS Comput. Biol.* 2:e143. doi: 10.1371/journal.pcbi.0020143

Andersen, K. S., Kirkegaard, R. H., Karst, S. M., and Albertsen, M. (2018). ampvis2: An R package to analyse and visualise 16S rRNA amplicon data. *bioRxiv* 2018:299537. doi: 10.1101/299537

Aziz, R. K., Breitbart, M., and Edwards, R. A. (2010). Transposases are the most abundant, most ubiquitous genes in nature. *Nucleic Acids Res.* 38, 4207–4217. doi: 10.1093/nar/gkq140

Baraúna, R., Freitas, D., Pinheiro, J., Folador, A., and Silva, A. (2017). A proteomic perspective on the bacterial adaptation to cold: Integrating OMICs data of the psychrotrophic bacterium *Exiguobacterium antarcticum* B7. *Proteomes* 5, 9–20. doi: 10.3390/proteomes5010009

Bennett, P. M. (2004). Genome plasticity: insertion sequence elements, transposons and integrons, and DNA rearrangement. *Methods Mol. Biol.* 266, 71–113. doi: 10.1385/1-59259-763-7:071

Biskaborn, B. K., Smith, S. L., Noetzli, J., Matthes, H., Vieira, G., Streletskiy, D. A., et al. (2019). Permafrost is warming at a global scale. *Nat. Commun.* 10, 1–11. doi: 10.1038/s41467-018-08240-4

Boetius, A., Anesio, A. M., Deming, J. W., Mikucki, J., and Rapp, J. Z. (2015). Microbial ecology of the cryosphere: sea ice and glacial habitats. *Nat. Rev. Microbiol.* 13, 677–690. doi: 10.1038/nrmicro3522

Borgeaud, S., Metzger, L. C., Scrignari, T., and Blokesch, M. (2015). The type VI secretion system of *Vibrio cholerae* fosters horizontal gene transfer. *Science* 347, 63–67. doi: 10.1126/science.1260064

Bowman, J. P. (2006). "The marine clade of the family *Flavobacteriaceae*: The genera *Aequorivita, Arenibacter, Cellulophaga, Croceibacter, Formosa, Gelidibacter, Gillisia, Maribacter, Mesonia, Muricauda, Polaribacter, Psychroflexus, Psychroserpens, Robiginitalea, Salegentibacter*," in *The Prokaryotes*, eds M. Dworkin, S. Falkow, E. Rosenberg, K. Schleifer, and E. Stackebrandt (New York, NY: Springer), 677–694. doi: 10.1007/0-387-30747-8_26

Bowman, J. P. (2008). "Genomic analysis of psychrophilic prokaryotes," in *Psychrophiles: From Biodiversity to Biotechnology*, eds R. Margesin, F. Schinner, J.-C. Marx, and C. Gerday (Berlin: Springer Verlag), 265–284. doi: 10.1007/978-3-540-74335-4_16

Bowman, J. S., Berthiaume, C. T., Armbrust, E. V., and Deming, J. W. (2014). The genetic potential for key biogeochemical processes in Arctic frost flowers and young sea ice revealed by metagenomic analysis. *FEMS Microbiol. Ecol.* 89, 376–387. doi: 10.1111/1574-6941.12331

Bradley, J. A., Amend, J. P., and LaRowe, D. E. (2018). Necromass as a limited source of energy for microorganisms in marine sediments. *J. Geophys. Res. Biogeosci.* 123, 577–590. doi: 10.1002/2017JG004186

Brazelton, W. J., and Baross, J. A. (2009). Abundant transposases encoded by the metagenome of a hydrothermal chimney biofilm. *ISME J.* 3, 1420–1424. doi: 10.1038/ismej.2009.79

Brown, M. V., and Bowman, J. P. (2001). A molecular phylogenetic survey of sea-ice microbial communities (SIMCO). *FEMS Microbiol. Ecol.* 35, 267–275. doi: 10.1016/S0168-6496(01)00100-3

Campen, R., Kowalski, J., Lyons, W. B., Tulaczyk, S., Dachwald, B., Pettit, E., et al. (2019). Microbial diversity of an Antarctic subglacial community and high—resolution replicate sampling inform hydrological connectivity in a polar desert. *Environ. Microbiol.* 21, 2290–2306. doi: 10.1111/1462-2920.14607

Carillo, S., Casillo, A., Pieretti, G., Parrilli, E., Sannino, F., Bayer-Giraldi, M., et al. (2015). A unique capsular polysaccharide structure from the psychrophilic marine bacterium *Colwellia psychrerythraea* 34H that mimics antifreeze (glyco)proteins. *J. Am. Chem. Soc.* 137, 179–189. doi: 10.1021/ja5075954

Casacuberta, E., and González, J. (2013). The impact of transposable elements in environmental adaptation. *Mol. Ecol.* 22, 1503–1517. doi: 10.1111/mec.12170

Chen, I. M. A., Chu, K., Palaniappan, K., Pillay, M., Ratner, A., Huang, J., et al. (2019). IMG/M v.5.0: An integrated data management and comparative analysis system for microbial genomes and microbiomes. *Nucleic Acids Res.* 47, D666–D677. doi: 10.1093/nar/gky901

Chintalapati, S., Kiran, M. D., and Shivaji, S. (2004). Role of membrane lipid fatty acids in cold adaptation. *Cell. Mol. Biol.* 50, 631–642.

Chua, M. J., Campen, R. L., Wahl, L., Grzymski, J. J., and Mikucki, J. A. (2018). Genomic and physiological characterization and description of *Marinobacter gelidimuriae* sp. nov., a psychrophilic, moderate halophile from Blood Falls, an antarctic subglacial brine. *FEMS Microbiol. Ecol.* 94:fiy021. doi: 10.1093/femsec/fiy021

Colangelo-Lillis, J., Eicken, H., Carpenter, S. D., and Deming, J. W. (2016). Evidence for marine origin and microbial-viral habitability of sub-zero hypersaline aqueous inclusions within permafrost near Barrow, Alaska. *FEMS Microbiol. Ecol.* 92:fiw053. doi: 10.1093/femsec/fiw053

Collins, R. E., and Deming, J. W. (2011). Abundant dissolved genetic material in Arctic sea ice Part II: Viral dynamics during autumn freeze-up. *Polar Biol.* 34, 1831–1841. doi: 10.1007/s00300-011-1008-z

Collins, R. E., and Deming, J. W. (2013). An inter-order horizontal gene transfer event enables the catabolism of compatible solutes by *Colwellia psychrerythraea* 34H. *Extremophiles* 17, 601–610. doi: 10.1007/s00792-013-0543-7

Collins, R. E., Carpenter, S. D., and Deming, J. W. (2008). Spatial heterogeneity and temporal dynamics of particles, bacteria, and pEPS in Arctic winter sea ice. *J. Mar. Syst.* 74, 902–917. doi: 10.1016/j.jmarsys.2007.09.005

Collins, R. E., Rocap, G., and Deming, J. W. (2010). Persistence of bacterial and archaeal communities in sea ice through an Arctic winter. *Environ. Microbiol.* 12, 1828–1841. doi: 10.1111/j.1462-2920.2010.02179.x

Cooper, Z. S., Rapp, J. Z., Carpenter, S. D., Iwahana, G., Eicken, H., and Deming, J. W. (2019). Distinctive microbial communities in subzero hypersaline brines from Arctic coastal sea ice and rarely sampled cryopegs. *FEMS Microbiol. Ecol.* 95:fiz166. doi: 10.1093/femsec/fiz166

Cox, G. F. N., and Weeks, W. F. (1983). Equations for determining the gas and brine volumes in sea-ice samples. *J. Glaciol.* 29, 306–316. doi: 10.3189/s0022143000008364

Csonka, L. N. (1989). Physiological and genetic responses of bacteria to osmotic stress. *Microbiol. Rev.* 53, 121–147. doi: 10.1128/mmbr.53.1.121-147.1989

Dare, K., Shepherd, J., Roy, H., Seveau, S., and Ibba, M. (2014). LysPGS formation in *Listeria monocytogenes* has broad roles in maintaining membrane integrity beyond antimicrobial peptide resistance. *Virulence* 5, 534–546. doi: 10.4161/viru.28359

Dawson, H. M., Heal, K. R., Boysen, A. K., Carlson, L. T., Ingalls, A. E., and Young, J. N. (2020). Potential of temperature- and salinity-driven shifts in diatom compatible solute concentrations to impact biogeochemical cycling within sea ice. *Elem. Sci. Anth.* 8:25. doi: 10.1525/elementa.421

DeLong, E. F., Preston, C. M., Mincer, T., Rich, V., Hallam, S. J., Frigaard, N. U., et al. (2006). Community genomics among stratified microbial assemblages in the ocean's interior. *Science* 311, 496–503. doi: 10.1126/science.1120250

Deming, J. W., and Collins, R. E. (2017). "Sea ice as a habitat for Bacteria, Archaea and viruses," in *Sea Ice*, ed. D. N. Thomas (Chichester: John Wiley & Sons, Ltd), 326–351. doi: 10.1002/9781118778371.ch13

Deming, J. W., and Young, J. N. (2017). "The role of exopolysaccharides in microbial adaptation to cold habitats," in *Psychrophiles: From Biodiversity to Biotechnology*, ed. R. Margesin (Cham: Springer), 259–284. doi: 10.1007/978-3-319-57057-0_12

DiCenzo, G. C., and Finan, T. M. (2017). The divided bacterial genome: Structure, function, and evolution. *Microbiol. Mol. Biol. Rev.* 81:17. doi: 10.1128/mmbr.00019-17

Dieser, M., Greenwood, M., and Foreman, C. M. (2010). Carotenoid pigmentation in Antarctic heterotrophic bacteria as a strategy to withstand environmental stresses. *Arctic Antarct. Alp. Res.* 42, 396–405. doi: 10.1657/1938-4246-42.4.396

Druckenmiller, M. L., Eicken, H., Johnson, M. A., Pringle, D. J., and Williams, C. C. (2009). Toward an integrated coastal sea-ice observatory: System components and a case study at Barrow, Alaska. *Cold Reg. Sci. Technol.* 56, 61–72. doi: 10.1016/j.coldregions.2008.12.003

Eicken, H., Gradinger, R., Salganek, M., Shirasawa, K., Perovich, D., and Leppäranta, M. (2009). *Field techniques for sea-ice research*, 2nd Edn. Fairbanks: University of Alaska Press.

Ernst, C. M., Staubitz, P., Mishra, N. N., Yang, S.-J., Hornig, G., Kalbacher, H., et al. (2009). The bacterial defensin resistance protein MprF consists of separable domains for lipid lysinylation and antimicrobial peptide repulsion. *PLoS Pathog.* 5:e1000660. doi: 10.1371/journal.ppat.1000660

Eronen-Rasimus, E., Piiparinen, J., Karkman, A., Lyra, C., Gerland, S., and Kaartokallio, H. (2016). Bacterial communities in Arctic first-year drift ice

during the winter/spring transition. *Environ. Microbiol. Rep.* 8, 527–535. doi: 10.1111/1758-2229.12428

Ewert, M., and Deming, J. W. (2013). Sea ice microorganisms: Environmental constraints and extracellular responses. *Biology* 2, 603–628. doi: 10.3390/biology2020603

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

Feehily, C., O'Byrne, C. P., and Karatzas, K. A. G. (2013). Functional γ-aminobutyrate shunt in *Listeria monocytogenes*: Role in acid tolerance and succinate biosynthesis. *Appl. Environ. Microbiol.* 79, 74–80. doi: 10.1128/AEM.02184-12

Feng, S., Powell, S. M., Wilson, R., and Bowman, J. P. (2014). Extensive gene acquisition in the extremely psychrophilic bacterial species *Psychroflexus torquis* and the link to sea-ice ecosystem specialism. *Genome Biol. Evol.* 6, 133–148. doi: 10.1093/gbe/evt209

Firth, E., Carpenter, S. D., Sørensen, H. L., Collins, R. E., and Deming, J. W. (2016). Bacterial use of choline to tolerate salinity shifts in sea-ice brines. *Elem. Sci. Anth.* 2016:000120. doi: 10.12952/journal.elementa.000120

Flemming, H. C., Wingender, J., Szewzyk, U., Steinberg, P., Rice, S. A., and Kjelleberg, S. (2016). Biofilms: An emergent form of bacterial life. *Nat. Rev. Microbiol.* 14, 563–575. doi: 10.1038/nrmicro.2016.94

Fountain, A. G., Campbell, J. L., Schuur, E. A. G., Stammerjohn, S. E., Williams, M. W., and Ducklow, H. W. (2012). The disappearing cryosphere: Impacts and ecosystem responses to rapid cryosphere loss. *Bioscience* 62, 405–415. doi: 10.1525/bio.2012.62.4.11

Fraley, C. D., Rashid, P. H., Lee, S. S. K., Gottschalk, R., Harrison, J., Wood, P. J., et al. (2007). A polyphosphate kinase 1 (ppk1) mutant of *Pseudomonas aeruginosa* exhibits multiple ultrastructural and functional defects. *Proc. Natl. Acad. Sci. U S A.* 104, 3526–3531. doi: 10.1073/pnas.0609733104

Frith, M. C., Hamada, M., and Horton, P. (2010). Parameters for accurate genome alignment. *BMC Bioinformatics* 11:80. doi: 10.1186/1471-2105-11-80

Froese, D. G., Westgate, J. A., Reyes, A. V., Enkin, R. J., and Preece, S. J. (2008). Ancient permafrost and a future, warmer Arctic. *Science* 321:1648. doi: 10.1126/science.1157525

Frost, L. S., Leplae, R., Summers, A. O., and Toussaint, A. (2005). Mobile genetic elements: The agents of open source evolution. *Nat. Rev. Microbiol.* 3, 722–732. doi: 10.1038/nrmicro1235

Ganesh, S., Parris, D. J., Delong, E. F., and Stewart, F. J. (2014). Metagenomic analysis of size-fractionated picoplankton in a marine oxygen minimum zone. *ISME J.* 8, 187–211. doi: 10.1038/ismej.2013.144

Gilichinsky, D., Rivkina, E., Bakermans, C., Shcherbakova, V., Petrovskaya, L., Ozerskaya, S., et al. (2005). Biodiversity of cryopegs in permafrost. *FEMS Microbiol. Ecol.* 53, 117–128. doi: 10.1016/j.femsec.2005.02.003

Gilichinsky, D., Rivkina, E., Shcherbakova, V., Laurinavichuis, K., and Tiedje, J. (2003). Supercooled water brines within permafrost - An unknown ecological niche for microorganisms: A model for astrobiology. *Astrobiology* 3, 331–341. doi: 10.1089/153110703769016424

Golden, K. M., Eicken, H., Heaton, A. L., Miner, J., Pringle, D. J., and Zhu, J. (2007). Thermal evolution of permeability and microstructure in sea ice. *Geophys. Res. Lett.* 34:2007GL030447. doi: 10.1029/2007GL030447

Gruber-Vodicka, H. R., Seah, B. K., and Pruesse, E. (2019). phyloFlash — Rapid SSU rRNA profiling and targeted assembly from metagenomes. *bioRxiv* 2019:521922. doi: 10.1101/521922

Guieysse, B., and Wuertz, S. (2012). Metabolically versatile large-genome prokaryotes. *Curr. Opin. Biotechnol.* 23, 467–473. doi: 10.1016/j.copbio.2011.12.022

Handley, K. M., and Lloyd, J. R. (2013). Biogeochemical implications of the ubiquitous colonization of marine habitats and redox gradients by *Marinobacter* species. *Front. Microbiol.* 4:136. doi: 10.3389/fmicb.2013.00136

Harms, A., Brodersen, D. E., Mitarai, N., and Gerdes, K. (2018). Toxins, targets, and triggers: An overview of toxin-antitoxin biology. *Mol. Cell* 70, 768–784. doi: 10.1016/j.molcel.2018.01.003

Hatzenpichler, R., Krukenberg, V., Spietz, R. L., and Jay, Z. J. (2020). Next-generation physiology approaches to study microbiome function at single cell level. *Nat. Rev. Microbiol.* 18, 241–256. doi: 10.1038/s41579-020-0323-1

Hayes, F. (2003). Toxins-antitoxins: Plasmid maintenance, programmed cell death, and cell cycle arrest. *Science* 301, 1496–1499. doi: 10.1126/science.1088157

Held, N. A., McIlvin, M. R., Moran, D. M., Laub, M. T., and Saito, M. A. (2019). Unique patterns and biogeochemical relevance of two-component sensing in marine bacteria. *mSystems* 4, 317–318. doi: 10.1128/msystems.00317-18

Huang, C. Y., Gonzalez-Lopez, C., Henry, C., Mijakovic, I., and Ryan, K. R. (2020). hipBA toxin-antitoxin systems mediate persistence in *Caulobacter crescentus*. *Sci. Rep.* 10, 1–15. doi: 10.1038/s41598-020-59283-x

Huntemann, M., Ivanova, N. N., Mavromatis, K., Tripp, H. J., Paez-Espino, D., Tennessen, K., et al. (2016). The standard operating procedure of the DOE-JGI Metagenome Annotation Pipeline (MAP v.4). *Stand. Genomic Sci.* 11:17. doi: 10.1186/s40793-016-0138-x

Ikawa, M., Borowski, P. T., and Chakravarti, A. (1968). Choline and inositol distribution in algae and fungi. *Appl. Environ. Microbiol.* 16, 620–623. doi: 10.1128/am.16.4.620-623.1968

Iwahana, G., Cooper, Z. S., Carpenter, S. D., Deming, J. W., and Eicken, H. (2021). Intra-ice and intra-sediment cryopeg brine occurrence in permafrost near Utqiaġvik (Barrow). *Permafr. Periglac. Process.* 2021, 1–20. doi: 10.1002/ppp.2101

Jensen, R. B., and Gerdes, K. (1995). Programmed cell death in bacteria: Proteic plasmid stabilization systems. *Mol. Microbiol.* 17, 205–210. doi: 10.1111/j.1365-2958.1995.mmi_17020205.x

Junge, K., Krembs, C., Deming, J., Stierle, A., and Eicken, H. (2001). A microscopic approach to investigate bacteria under in situ conditions in sea-ice samples. *Ann. Glaciol.* 33, 304–310. doi: 10.3189/172756401781818275

Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M., and Tanabe, M. (2016). KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* 44:gkv1070. doi: 10.1093/nar/gkv1070

Katoh, K., Rozewicki, J., and Yamada, K. D. (2019). MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief. Bioinform.* 20, 1160–1166. doi: 10.1093/bib/bbx108

Kempf, B., and Bremer, E. (1998). Uptake and synthesis of compatible solutes as microbial stress responses to high-osmolality environments. *Arch. Microbiol.* 170, 319–330. doi: 10.1007/s002030050649

Kirst, G. O., Thiel, C., Wolff, H., Nothnagel, J., Wanzek, M., and Ulmke, R. (1991). Dimethylsulfoniopropionate (DMSP) in ice-algae and its possible biological role. *Mar. Chem.* 35, 381–388. doi: 10.1016/S0304-4203(09)90030-5

Kolde, R. (2019). *pheatmap: Pretty Heatmaps*. Vienna: R Core Team.

Konstantinidis, K. T., Braff, J., Karl, D. M., and DeLong, E. F. (2009). Comparative metagenomic analysis of a microbial community residing at a depth of 4,000 meters at station ALOHA in the North Pacific Subtropical Gyre. *Appl. Environ. Microbiol.* 75, 5345–5355. doi: 10.1128/AEM.00473-09

Krehenbrink, M., Oppermann-Sanio, F. B., and Steinbüchel, A. (2002). Evaluation of non-cyanobacterial genome sequences for occurrence of genes encoding proteins homologous to cyanophycin synthetase and cloning of an active cyanophycin synthetase from *Acinetobacter* sp. strain DSM 587. *Arch. Microbiol.* 177, 371–380. doi: 10.1007/s00203-001-0396-9

Krembs, C., and Deming, J. W. (2008). "The role of exopolymers in microbial adaptation to sea ice," in *Psychrophiles: From Biodiversity to Biotechnology*, eds R. Margesin, F. Schinner, J.-C. Marx, and C. Gerday (Heidelberg: Springer Verlag), 247–264. doi: 10.1007/978-3-540-74335-4_15

Letunic, I., and Bork, P. (2019). Interactive Tree of Life (iTOL) v4: Recent updates and new developments. *Nucleic Acids Res.* 47, W256–W259. doi: 10.1093/nar/gkz239

Manor, O., and Borenstein, E. (2015). MUSiCC: a marker genes based framework for metagenomic normalization and accurate profiling of gene abundances in the microbiome. *Genome Biol.* 16:53. doi: 10.1186/s13059-015-0610-8

Manor, O., and Borenstein, E. (2017). Revised computational metagenomic processing uncovers hidden and biologically meaningful functional variation in the human microbiome. *Microbiome* 5:19. doi: 10.1186/s40168-017-0231-4

Médigue, C., Krin, E., Pascal, G., Barbe, V., Bernsel, A., Bertin, P. N., et al. (2005). Coping with cold: The genome of the versatile marine Antarctica bacterium *Pseudoalteromonas haloplanktis* TAC125. *Genome Res.* 15, 1325–1335. doi: 10.1101/gr.4126905

Methé, B. A., Nelson, K. E., Deming, J. W., Momen, B., Melamud, E., Zhang, X., et al. (2005). The psychrophilic lifestyle as revealed by the genome sequence of *Colwellia psychrerythraea* 34H through genomic and proteomic analyses. *Proc. Natl. Acad. Sci. U S A.* 102, 10913–10918. doi: 10.1073/pnas.0504766102

Meyer, H., Schirrmeister, L., Andreev, A., Wagner, D., Hubberten, H. W., Yoshikawa, K., et al. (2010). Lateglacial and Holocene isotopic and

environmental history of northern coastal Alaska - Results from a buried ice-wedge system at Barrow. *Quat. Sci. Rev.* 29, 3720–3735. doi: 10.1016/j.quascirev.2010.08.005

Mikucki, J. A., and Priscu, J. C. (2007). Bacterial diversity associated with Blood Falls, a subglacial outflow from the Taylor Glacier, Antarctica. *Appl. Environ. Microbiol.* 73, 4029–4039. doi: 10.1128/AEM.01396-06

Niederberger, T. D., Perreault, N. N., Tille, S., Lollar, B. S., Lacrampe-Couloume, G., Andersen, D., et al. (2010). Microbial characterization of a subzero, hypersaline methane seep in the Canadian High Arctic. *ISME J.* 4, 1326–1339. doi: 10.1038/ismej.2010.57

Nitzbon, J., Westermann, S., Langer, M., Martin, L. C. P., Strauss, J., Laboor, S., et al. (2020). Fast response of cold ice-rich permafrost in northeast Siberia to a warming climate. *Nat. Commun.* 11, 1–11. doi: 10.1038/s41467-020-15725-8

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., et al. (2019). *vegan: Community Ecology Package*. Vienna: R Core Team.

Overland, J. E., and Wang, M. (2013). When will the summer Arctic be nearly sea ice free? *Geophys. Res. Lett.* 40, 2097–2101. doi: 10.1002/grl.50316

Pandey, D. P., and Gerdes, K. (2005). Toxin-antitoxin loci are highly abundant in free-living but lost from host-associated prokaryotes. *Nucleic Acids Res.* 33, 966–976. doi: 10.1093/nar/gki201

Park, S., Kim, S., Kang, C. H., Jung, Y. T., and Yoon, J. H. (2015). *Marinobacter confluentis* sp. nov., a lipolytic bacterium isolated from a junction between the ocean and a freshwater lake. *Int. J. Syst. Evol. Microbiol.* 65, 4873–4879. doi: 10.1099/ijsem.0.000659

Peeters, S. H., and de Jonge, M. I. (2018). For the greater good: Programmed cell death in bacterial communities. *Microbiol. Res.* 207, 161–169. doi: 10.1016/j.micres.2017.11.016

Petrich, C., and Eicken, H. (2017). "Overview of sea ice growth and properties," in *Sea Ice*, ed. D. N. Thomas (Chichester: John Wiley & Sons, Ltd), 1–41. doi: 10.1002/9781118778371.ch1

Piskunova, J., Maisonneuve, E., Germain, E., Gerdes, K., and Severinov, K. (2017). Peptide-nucleotide antibiotic Microcin C is a potent inducer of stringent response and persistence in both sensitive and producing cells. *Mol. Microbiol.* 104, 463–471. doi: 10.1111/mmi.13640

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2013). The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res.* 41, D590–D596. doi: 10.1093/nar/gks1219

R Core Team (2017). *R: A language and environment for statistical computing*. Vienna: R Core Team.

Rao, N. N., Gómez-García, M. R., and Kornberg, A. (2009). Inorganic polyphosphate: Essential for growth and survival. *Annu. Rev. Biochem.* 78, 605–647. doi: 10.1146/annurev.biochem.77.083007.093039

Rapp, J. Z., Fernández-Méndez, M., Bienhold, C., and Boetius, A. (2018). Effects of ice-algal aggregate export on the connectivity of bacterial communities in the central Arctic Ocean. *Front. Microbiol.* 9:1035. doi: 10.3389/fmicb.2018.01035

Raymond-Bouchard, I., Goordial, J., Zolotarov, Y., Ronholm, J., Stromvik, M., Bakermans, C., et al. (2018). Conserved genomic and amino acid traits of cold adaptation in subzero-growing Arctic permafrost bacteria. *FEMS Microbiol. Ecol.* 94:fiy023. doi: 10.1093/femsec/fiy023

Reckseidler-Zenteno, S. L. (2012). "Capsular polysaccharides produced by the bacterial pathogen *Burkholderia pseudomallei*," in *The Complex World of Polysaccharides*, ed. D. N. Karunaratne (Rijeka: InTech), 127–152. doi: 10.5772/50116

Reisch, C. R., Moran, M. A., and Whitman, W. B. (2011). Bacterial catabolism of dimethylsulfoniopropionate (DMSP). *Front. Microbiol.* 2:172. doi: 10.3389/fmicb.2011.00172

Riley, M. A., and Wertz, J. E. (2002). Bacteriocins: Evolution, ecology, and application. *Annu. Rev. Microbiol.* 56, 117–137. doi: 10.1146/annurev.micro.56.012302.161024

Rodrigue, A., Quentin, Y., Lazdunski, A., Méjean, V., and Foglino, M. (2000). Two-component systems in *Pseudomonas aeruginosa*: Why so many? *Trends Microbiol.* 8, 498–504. doi: 10.1016/S0966-842X(00)01833-3

Rodrigues, D. F., Ivanova, N., He, Z., Huebner, M., Zhou, J., and Tiedje, J. M. (2008). Architecture of thermal adaptation in an *Exiguobacterium sibiricum* strain isolated from 3 million year old permafrost: A genome and transcriptome approach. *BMC Genomics* 9:547. doi: 10.1186/1471-2164-9-547

Russell, A. B., Peterson, S. B., and Mougous, J. D. (2014). Type VI secretion system effectors: Poisons with a purpose. *Nat. Rev. Microbiol.* 12, 137–148. doi: 10.1038/nrmicro3185

Salvado, B., Vilaprinyo, E., Sorribas, A., and Alves, R. (2015). A survey of HK, HPt, and RR domains and their organization in two-component systems and phosphorelay proteins of organisms with fully sequenced genomes. *PeerJ* 3:e1183. doi: 10.7717/peerj.1183

Seel, W., Baust, D., Sons, D., Albers, M., Etzbach, L., Fuss, J., et al. (2020). Carotenoids are used as regulators for membrane fluidity by *Staphylococcus xylosus*. *Sci. Rep.* 10, 57006–57005. doi: 10.1038/s41598-019-57006-5

Seufferheld, M. J., Alvarez, H. M., and Farias, M. E. (2008). Role of polyphosphates in microbial adaptation to extreme environments. *Appl. Environ. Microbiol.* 74, 5867–5874. doi: 10.1128/AEM.00501-08

Shcherbakova, V. A., Chuvil'Skaya, N. A., Rivkina, E. M., Pecheritsyna, S. A., Suetin, S. V., Laurinavichius, K. S., et al. (2009). Novel halotolerant bacterium from cryopeg in permafrost: Description of *Psychrobacter muriicola* sp. nov. *Microbiology* 78, 84–91. doi: 10.1134/S0026261709010111

Shi, X., Wegener-Feldbrügge, S., Huntley, S., Hamann, N., Hedderich, R., and Søgaard-Andersen, L. (2008). Bioinformatics and experimental analysis of proteins of two-component systems in *Myxococcus xanthus*. *J. Bacteriol.* 190, 613–624. doi: 10.1128/JB.01502-07

Showalter, G. M., and Deming, J. W. (2021). Extracellular enzyme activity under subzero hypersaline conditions by model cold-adapted bacteria and Arctic sea-ice microbial communities. *Aquat. Microb. Ecol.* (in press). doi: 10.3354/AME01974

Siguier, P., Gourbeyre, E., and Chandler, M. (2014). Bacterial insertion sequences: Their genomic impact and diversity. *FEMS Microbiol. Rev.* 38, 865–891. doi: 10.1111/1574-6976.12067

Singer, E., Webb, E. A., Nelson, W. C., Heidelberg, J. F., Ivanova, N., Pati, A., et al. (2011). Genomic potential of *Marinobacter aquaeolei*, a biogeochemical "Opportunitroph.". *Appl. Environ. Microbiol.* 77, 2763–2771. doi: 10.1128/AEM.01866-10

Staley, J. T., and Gosink, J. J. (1999). Poles apart: Biodiversity and biogeography of sea ice bacteria. *Annu. Rev. Microbiol.* 53, 189–215. doi: 10.1146/annurev.micro.53.1.189

Storey, J. D., and Tibshirani, R. (2003). Statistical significance for genomewide studies. *Proc. Natl. Acad. Sci. U S A.* 100, 9440–9445. doi: 10.1073/pnas.1530509100

Stroeve, J., and Notz, D. (2018). Changing state of Arctic sea ice across all seasons. *Environ. Res. Lett.* 13:103001. doi: 10.1088/1748-9326/aade56

Thomas, J., Watve, S. S., Ratcliff, W. C., and Hammer, B. K. (2017). Horizontal gene transfer of functional type VI killing genes by natural transformation. *MBio* 8, 654–617. doi: 10.1128/mBio.00654-17

Ting, L., Williams, T. J., Cowley, M. J., Lauro, F. M., Guilhaus, M., Raftery, M. J., et al. (2010). Cold adaptation in the marine bacterium, *Sphingopyxis alaskensis*, assessed using quantitative proteomics. *Environ. Microbiol.* 12, 2658–2676. doi: 10.1111/j.1462-2920.2010.02235.x

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

Tribelli, P. M., and López, N. I. (2018). Reporting key features in cold-adapted bacteria. *Life* 8:life8010008. doi: 10.3390/life8010008

Tripp, H. J., Kitner, J. B., Schwalbach, M. S., Dacey, J. W. H., Wilhelm, L. J., and Giovannoni, S. J. (2008). SAR11 marine bacteria require exogenous reduced sulphur for growth. *Nature* 452, 741–744. doi: 10.1038/nature06776

Trivedi, C. B., Stamps, B. W., Lau, G. E., Grasby, S. E., Templeton, A. S., and Spear, J. R. (2020). Microbial metabolic redundancy is a key mechanism in a sulfur-rich glacial ecosystem. *mSystems* 5, e504–e520. doi: 10.1128/msystems.00504-20

van Helden, J. (2016). *stats4bioinfo: Utilities for the book "Statistics for bioinformatics."*. San Francisco: Github.

Vaughan, D. G., Comiso, J. C., Allison, I., Carrasco, J., Kaser, G., Kwok, R., et al. (2013). "Observations: Cryosphere," in *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, eds T. F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, et al. (Cambridge: Cambridge University Press), 317–382. doi: 10.1017/CBO9781107415324.012

Vettiger, A., and Basler, M. (2016). Type VI secretion system substrates are transferred and reused among sister cells. *Cell* 167, 99.e–110.e. doi: 10.1016/j.cell.2016.08.023

Vigil-Stenman, T., Ininbergs, K., Bergman, B., and Ekman, M. (2017). High abundance and expression of transposases in bacteria from the Baltic Sea. *ISME J.* 11, 2611–2623. doi: 10.1038/ismej.2017.114

Vigil-Stenman, T., Larsson, J., Nylander, J. A. A., and Bergman, B. (2015). Local hopping mobile DNA implicated in pseudogene formation and reductive evolution in an obligate cyanobacteria-plant symbiosis. *BMC Genomics* 16:193. doi: 10.1186/s12864-015-1386-7

Vollmers, J., Voget, S., Dietrich, S., Gollnow, K., Smits, M., Meyer, K., et al. (2013). Poles apart: Arctic and Antarctic *Octadecabacter* strains share high genome plasticity and a new type of xanthorhodopsin. *PLoS One* 8:e63422. doi: 10.1371/journal.pone.0063422

Welsh, D. T. (2000). Ecological significance of compatible solute accumulation by micro-organisms: from single cells to global climate. *FEMS Microbiol. Rev.* 24, 263–290. doi: 10.1111/j.1574-6976.2000.tb00542.x

West, P. T., Probst, A. J., Grigoriev, I. V., Thomas, B. C., and Banfield, J. F. (2018). Genome-reconstruction for eukaryotes from complex natural microbial communities. *Genome Res.* 28, 569–580. doi: 10.1101/gr.228429.117

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Texas: ggplot2.

Wilkins, D., Lauro, F. M., Williams, T. J., Demaere, M. Z., Brown, M. V., Hoffman, J. M., et al. (2013). Biogeographic partitioning of Southern Ocean microorganisms revealed by metagenomics. *Environ. Microbiol.* 15, 1318–1333. doi: 10.1111/1462-2920.12035

Wood, J. M., Bremer, E., Csonka, L. N., Kraemer, R., Poolman, B., Van der Heide, T., et al. (2001). Osmosensing and osmoregulatory compatible solute accumulation by bacteria. *Comparat. Biochem. Physiol. Mol. Integrat. Physiol.* 2001, 437–460. doi: 10.1016/S1095-6433(01)00442-1

Wuichet, K., Cantwell, B. J., and Zhulin, I. B. (2010). Evolution and phyletic distribution of two-component signal transduction systems. *Curr. Opin. Microbiol.* 13, 219–225. doi: 10.1016/j.mib.2009.12.011

Yergeau, E., Michel, C., Tremblay, J., Niemi, A., King, T. L., Wyglinski, J., et al. (2017). Metagenomic survey of the taxonomic and functional microbial communities of seawater and sea ice from the Canadian Arctic. *Sci. Rep.* 7:42242. doi: 10.1038/srep42242

Zhang, D. C., Li, H. R., Xin, Y. H., Chi, Z. M., Zhou, P. J., and Yu, Y. (2008). *Marinobacter psychrophilus* sp. nov., a psychrophilic bacterium isolated from the Arctic. *Int. J. Syst. Evol. Microbiol.* 58, 1463–1466. doi: 10.1099/ijs.0.65690-0

Zhong, Z. P., Liu, Y., Liu, H. C., Wang, F., Zhou, Y. G., and Liu, Z. P. (2015). *Marinobacter halophilus* sp. nov., a halophilic bacterium isolated from a salt lake. *Int. J. Syst. Evol. Microbiol.* 65, 2838–2845. doi: 10.1099/ijs.0.000338

Zhong, Z.-P., Rapp, J. Z., Wainaina, J. M., Solonenko, N. E., Maughan, H., Carpenter, S. D., et al. (2020). Viral ecogenomics of Arctic cryopeg brine and sea ice. *mSystems* 5, e246–e220. doi: 10.1128/msystems.00246-20

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

frontiers in Microbiology

ORIGINAL RESEARCH
published: 08 October 2021
doi: 10.3389/fmicb.2021.736407

Prokaryotic Community Succession in Bulk and Rhizosphere Soils Along a High-Elevation Glacier Retreat Chronosequence on the Tibetan Plateau

Jinbo Liu[1,2,3], Weidong Kong[3], Pinhua Xia[4], Chunmao Zhu[5] and Xiangzhen Li[6]*

[1] Department of Hepatobiliary Surgery, The Affiliated Hospital of Southwest Medical University, Luzhou, China, [2] Academician (Expert) Workstation of Sichuan Province, The Affiliated Hospital of Southwest Medical University, Luzhou, China, [3] Key Laboratory of Alpine Ecology, Institute of Tibetan Plateau Research, Chinese Academy of Sciences, Beijing, China, [4] Guizhou Key Laboratory for Mountainous Environmental Information and Ecological Protection, Guizhou Normal University, Guiyang, China, [5] Research Institute for Global Change, Japan Agency for Marine–Earth Science and Technology (JAMSTEC), Yokohama, Japan, [6] Key Laboratory of Environmental and Applied Microbiology, Chengdu Institute of Biology, Chinese Academy of Sciences, Chengdu, China

OPEN ACCESS

Edited by:
Anne D. Jungblut, Natural History Museum, United Kingdom

Reviewed by:
Igor S. Pessi, University of Helsinki, Finland
Shuo Jiao, Northwest A and F University, China

***Correspondence:**
Weidong Kong
wdkong@itpcas.ac.cn

Specialty section:
This article was submitted to Extreme Microbiology, a section of the journal Frontiers in Microbiology

Received: *05 July 2021*
Accepted: *30 August 2021*
Published: *08 October 2021*

Citation:
Liu J, Kong W, Xia P, Zhu C and Li X (2021) Prokaryotic Community Succession in Bulk and Rhizosphere Soils Along a High-Elevation Glacier Retreat Chronosequence on the Tibetan Plateau. Front. Microbiol. 12:736407. doi: 10.3389/fmicb.2021.736407

Early colonization and succession of soil microbial communities are essential for soil development and nutrient accumulation. Herein we focused on the changes in pioneer prokaryotic communities in rhizosphere and bulk soils along the high-elevation glacier retreat chronosequence, the northern Himalayas, Tibetan Plateau. Rhizosphere soils showed substantially higher levels of total organic carbon, total nitrogen, ammonium, and nitrate than bulk soils. The dominant prokaryotes were Proteobacteria, Actinobacteria, Acidobacteria, Chloroflexi, Crenarchaeota, Bacteroidetes, and Planctomycetes, which totally accounted for more than 75% in relative abundance. The dominant genus *Candidatus Nitrososphaera* occurred at each stage of the microbial succession. The richness and evenness of soil prokaryotes displayed mild succession along chronosequene. Linear discriminant analysis effect size (LEfSe) analysis demonstrated that Proteobacteria (especially Alphaproteobacteria) and Actinobacteria were significantly enriched in rhizosphere soils compared with bulk soils. Actinobacteria, SHA_109, and Thermoleophilia; Betaproteobacteria and OP1.MSBL6; and Planctomycetia and Verrucomicrobia were separately enriched at each of the three sample sites. The compositions of prokaryotic communities were substantially changed with bulk and rhizosphere soils and sampling sites, indicating that the communities were dominantly driven by plants and habitat-specific effects in the deglaciated soils. Additionally, the distance to the glacier terminus also played a significant role in driving the change of prokaryotic communities in both bulk and rhizosphere soils. Soil C/N ratio exhibited a greater effect on prokaryotic communities in bulk soils than rhizosphere soils. These results indicate that plants, habitat, and glacier retreat chronosequence collectively control prokaryotic community composition and succession.

Keywords: prokaryote, deglaciated soil, alpine ecology, microbial community composition, Tibetan Plateau

INTRODUCTION

Glaciers cover ∼10% of the land surface of the Earth and are rapidly shrinking in most parts of the world, leading to significant impacts on terrestrial ecosystems (Jain, 2014; Milner et al., 2017). The biological, physical, and chemical characteristics of the deglaciated soil are closely linked to the deglaciation chronosequence (Bernasconi et al., 2011). These newly exposed substrates represent natural laboratories to study primary succession of the microbial community and the concomitant development of new soil (Schuette et al., 2010). Deglaciated soil typically has low nutrient status and an absence of organic carbon (C) (Strauss et al., 2009). Microbial succession is highly correlated with soil C and nitrogen (N) contents along the deglaciation chronosequence (Zumsteg et al., 2012), and the microbial communities are the main drivers that build the soil organic matter pool, expediting pedogenesis for ecosystem succession (Sun et al., 2016). Early soil formation processes should be related to the composition of the microbial communities, the primary substrate structure, and available water (Górniak et al., 2017). The establishment of pioneering microbial communities is the key determinant of deglaciated soil development and its ecosystem function and stability (Schmidt et al., 2008; Kabala and Zapart, 2012) and facilitates the colonization of pioneering plants (Bradley et al., 2014). In a high Arctic glacier forefield, genomic data analysis showed that bacteria derived from the glacial environment was the dominant initial microbial community; a mixed community of autotrophic and heterotrophic bacteria was hosted in older soils (Bradley et al., 2016). In a high-Andean chronosequence, photosynthetic bacteria, and N-fixing bacteria, such as cyanobacteria, play important roles in the acquisition of nutrients and ecological succession in recently deglaciated soils (Schmidt et al., 2008). Cyanobacterial diversity and evenness increase in young deglaciated soils (<6 years old) before becoming stable along a chronosequence of the Tibetan Plateau (Liu et al., 2016).

Plants, during primary succession, and microbial activity in rhizosphere soil can be limited by N (Castle et al., 2017); in the case of the latter, it was found this limitation could last for 123 years in the Hailuogou Glacier forefield (Li et al., 2020). However, the initial microbial community of newly exposed soils, including those of the High Arctic, can change rapidly, suggesting that some key soil processes, such as C cycling, can also shift within a relatively short period after rapid glacial retreat (Yoshitake et al., 2018). Specific rhizobacterial communities can be selected by pioneer plants of different species in high mountain ecosystems during early primary succession (Ciccazzo et al., 2014). The stage of soil development modulates rhizosphere effect along a high Arctic desert chronosequence (Mapelli et al., 2018). Additionally, it has been found that plant–microbe interactions are important as a driver of community assembly and ecosystem succession (Bueno de Mesquita et al., 2017; Knelman et al., 2018).

The climate-induced glacier melting revealed a primary ecological succession (Cazzolla Gatti et al., 2018). Studying the succession of deglaciated soil microbial communities is important to fully understand the impact of climate change on soil system stability in alpine areas. The community structures of glacier foreland soils are strongly correlated with climatic, vegetation, and soil properties and thus closely mirror the complexity and small-scale heterogeneity of alpine soils (Donhauser and Frey, 2018). The Tibetan Plateau, with an average elevation of over 4,000 m above sea level (a.s.l.) and an area of 2.5×10^6 km^2, is the highest and most extensive highland in the world and has been named "the Third Pole" (Kang et al., 2010). Tibetan Plateau glaciers have exhibited rapid retreat due to climate warming after the little ice age (extended from AD 1400 to AD 1700) (Mann et al., 2009) and particularly since the 1980s (Yao et al., 2012). For the last years, several studies have been focused on the melt glacier foreland bulk soil community, such as an autotrophic community in Zhadang Glacier (30°28.540′N, 90°38.362′E, 5,200 m at glacier termini) (Liu et al., 2016), a bacterial community in the foreland of Baishui Glacier No. 1 (27°06′16″ N, 100°11′44″E, 4,395 m at glacier termini) (Sajjad et al., 2021), a bacterial and fungal community in Hailuogou Glacier (29°34′N, 102°00′E, 2,951 m at glacier termini) (Bai et al., 2019, 2020; Jiang et al., 2019; Li et al., 2020), and a bacterial community in Muztag Ata Glacier (38°16′N, 75°0′E, 4,350 m at glacier termini) (Khan et al., 2020). These studies were not enough as yet to understand prokaryotic community succession and the effect of plants on soil microorganism in this region. These studies were mainly focused on the bacterial succession in bulk soils from glacier forelands, few studies took into account the role of plants. These glaciers were in different altitudes and spread out in different areas of the Tibetan Plateau with different retreat times. Altitude is a sensitive environment selector of plant growth (Ma et al., 2010), not to mention the extreme oligotrophic environment of the new terrestrial habitats. Research on the high-elevation glacier retreat area is rare, and the plant effect is not clear. We hypothesized that (1) the high-elevation glacier foreland may have a special prokaryotic community structure, and (2) the pioneer plants appearing in a high-elevation glacier retreat area may be favorable for special microbes and have an effect on the microbial community structure and soil nutrient accumulation along the chronosequence.

Qiangyong Glacier is a high-elevation glacier (the terminus is at an altitude of 5,000–5,100 m) located on the northern Himalayas, Tibetan Plateau. New terrestrial habitats have emerged, and a primary succession has developed in the retreat area after the glacier retreated. Herbs and shrubs appeared along the successional chronosequence. Recently, two reports showed the bacterial community succession in Qiangyong Glacier terminus to the downstream water (Kong et al., 2019; Gu et al., 2021); the prokaryotic community in this glacier foreland and the plant effects were still unknown. To test the hypotheses addressed above, both bulk and rhizosphere soils during different succession stages were collected in Qiangyong Glacier. Primer sets specifically designed to target part of the 16S rRNA gene

were used to explore the structure, diversity, and succession of prokaryotic communities in deglaciated soils along a ∼259-year chronosequence. Soil physicochemical properties (C, N, and water content) were measured, and attempts were made to identify which factor may be the key one in community succession.

MATERIALS AND METHODS

Study Site and Soil Sampling

Qiangyong Glacier is located on the north side of the Himalayas, with a length of 4.6 km, a maximum width of 2.8 km, and an area of 7.7 km^2 (Luo et al., 2003). It is a continental glacier; the snow line is at an altitude of 5,600 m, and the terminus is at an altitude of 5,000–5,100 m (Li et al., 2012). From 1976 to 2006, the glacier has retreated at an average rate of 4 m $year^{-1}$ (Yao et al., 2012). In July 2013, soil samples were collected from three sample sites (IS, SM, and BD) along two transects [line middle (LM) and line Top (LT)], on the east-facing side of the glacier movement residual moraine ridges (**Figure 1**). The three sample sites were as follows: the middle of the glacier termini and small Qiangyong lake upstream edge (IS), midway along the edge of the smaller lake (SM), and approximately two-thirds of the way along the edge of the larger lake (BD). At the time of sampling, sites IS_M(96a), SM_M(160a), and BD_M(259a) corresponded to a time of exposure after ice melting of 96, 160, and 259 years, respectively; sites IS_T(92a), SM_T(159a), and BD_T(258a) corresponded to a time of exposure after ice melting of 92, 159, and 258 years, respectively. The main plant located at IS and SM was *Kobresia homilies*, which then transitioned *Potentilla fruticose* at the BD sites. Bulk soil not in contact with the root system and located at 50–100 cm from each sampled abundant plant was collected from 3 to 5 cm after removing the top 1–2 cm of large sand grains. Soil samples near the plant root (less than 1 cm) were collected as rhizosphere soil from a similar depth as the bulk soils. Soil samples were stored in sterile sampling bags (Labplas, Canada) and transferred to the laboratory on ice. One part of each sample was stored at −80°C until DNA extraction was performed; the other part was used for physicochemical property analyses.

Soil Physicochemical Property Determination

Soil factors were measured using the methods described in Guo et al. (2015) and will only be summarized here. Soil total C and total organic C (TOC) were measured in the solid state using a TOC analyzer (TOC-L, Shimadzu, Japan); soil total nitrogen (TN) was determined by elemental analyzer (vario MAX, Elementar, Germany). Soil nitrate (NO_3^--N) and ammonium (NH_4^+-N) were extracted with 1 M KCL and determined using an Automated Discrete Analyzer (AQ2, SEAL Analytical Inc., England). Soil water content (WC) was gravimetrically determined after drying at 105°C for 12 h.

Sample DNA Extractions, PCR Amplification, and High-Throughput Sequencing

Soil genomic DNA was extracted from 0.5 g of frozen soil using the FastDNA® spin kit for soil (MP Biomedicals, Solon, OH, United States) following the manufacturer's protocol. The quantity of the DNA was determined using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies Inc., Wilmington, DE, United States). The primers 515F (5′-GTG YCA GCM GCC GCG GTA-3′) and 806R (5′-GGA CTA CHV GGG TWT CTA AT-3′) were used for Illumina MiSeq sequencing at the Chengdu Institute of Biology, Chinese Academy of Sciences. PCR amplification and high-throughput sequencing method was according to Li et al. (2015).

Sequence Analysis

The raw sequences were sorted based on the unique sample barcodes, and then quality control was performed for the sequences using QIIME pipeline (Caporaso et al., 2010). All sequence reads were trimmed and assigned to each sample based on their barcodes. The sequences with high quality (length > 150 bp, without ambiguous base "N," and average base quality score > 30) were used for downstream analysis. Chimeric sequences were identified and removed using the Uchime algorithm (Edgar et al., 2011). After filtering and chimera removal, *de novo* operational taxonomic unit (OTU) picking was performed using uclust at 97% sequence identity, and subsequently, taxonomy was assigned to OTU based on the Greengenes database at a confidence level of 80% (version 13.8). All low-abundance Archaeal and Bacterial sequences that were only detected in one of the samples were culled. To avoid the influence of sequencing depth, rarefaction was performed with 9,000 sequences per sample for the diversity of microorganisms.

Statistical Analyses

Community α-diversity was estimated using Shannon, Evenness, and Richness diversity indices. The β-diversity was calculated based on Bray–Curtis distances, and the differences in community structure between samples were visualized using principal coordinates analysis (PCoA) plot. Paired-samples *t*-tests were used for comparing data of bulk and rhizosphere soil with the SPSS 18.0 statistical software package. Adonis analysis was used to compare the difference between groups, and redundancy analysis (RDA) based on Bray–Curtis distances was used to determine the most significant environmental variables that might influence the prokaryotic community structure using the vegan and picante packages of the statistical platform R. Linear discriminant analysis (LDA) effect size (LEfSe) analyses between different sample groups utilized an online platform (Segata et al., 2011) and was applied to the OTU table to identify discriminant prokaryotic taxa.

Data Submission

The raw sequence reads generated in the present study are available in the National Center for Biotechnology

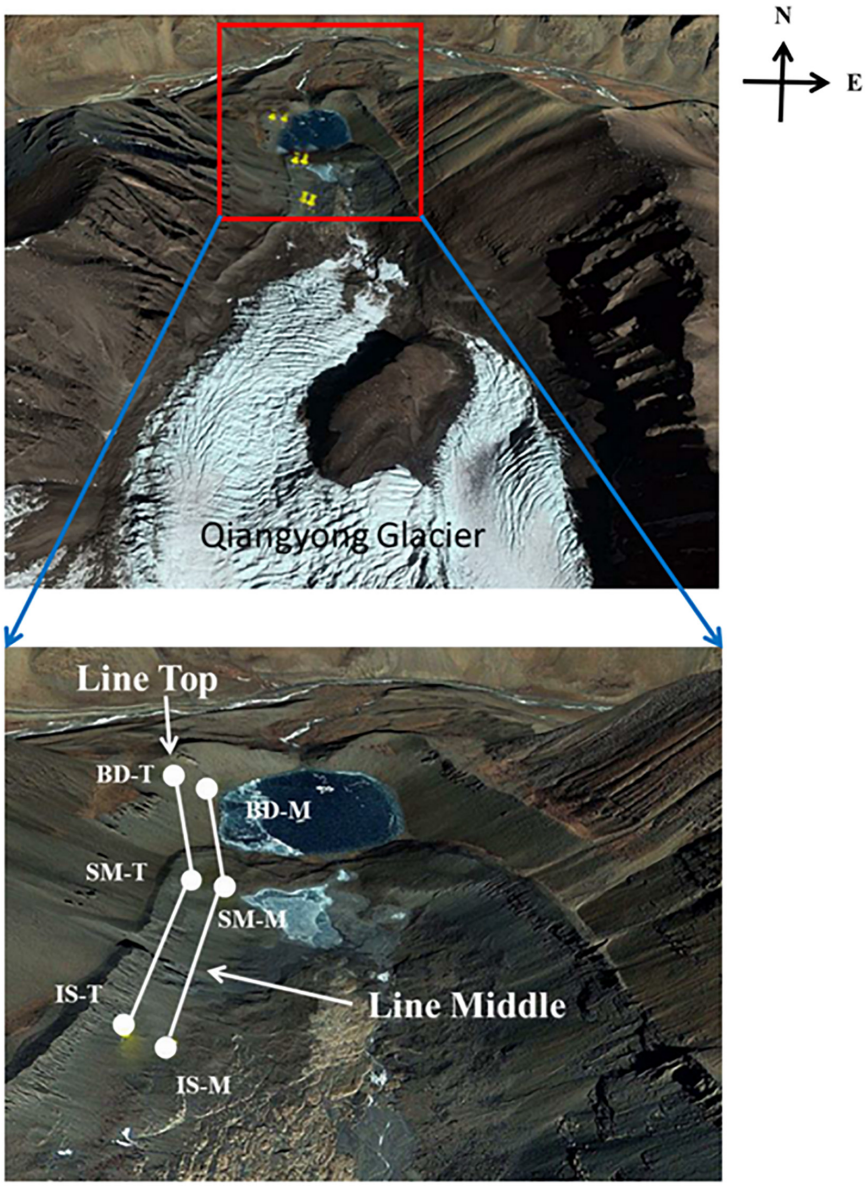

FIGURE 1 | Map of sampling locations along the Qiangyong Glacier retreat chronosequence. Sites: IS, middle of glacier terminal and small Qiangyong lake; SM, midway along the edge of the smaller lake; BD, approximately one-third of the way along the edge of the larger lake; M and T, the sampling lines, Middle and Top, respectively.

Information (NCBI) Sequence Read Archive under the project ID PRJNA595717.

RESULTS

Alpine Plants Altered the Accumulation of Soil Carbon and Nitrogen Content Along the Chronosequence

In total, the contents of TOC, TN, NH_4^+-N, and NO_3^--N in the rhizosphere were all significantly higher than in bulk soil ($n = 18$, $p < 0.01$). The TOC and TN contents showed an increasing trend along the glacier chronosequence, while NH_4^+-N and NO_3^--N contents showed a decreasing trend, at line middle and line top, respectively (**Supplementary Figure 1**).

Alpine Plant Effects on Prokaryotic Community Composition and Succession

In the phyla and genus level community composition, there were some clades detected in different abundance and succession trends between bulk and rhizosphere soil. In the phylum level,

the dominant phyla (relative abundance > 5%) across all samples were Proteobacteria (21.95%), Actinobacteria (18.11), Acidobacteria (9.50%), Chloroflexi (7.9%), Crenarchaeota (6.81%), Bacteroidetes (6.27%), and Planctomycetes (5.69%), accounting for more than 75% of the total prokaryotic sequences. The relative abundance of the phyla Gemmatimonadetes, Verrucomicrobia, Euryarchaeota, and Firmicutes were > 1% (**Figure 2A**). The relative abundances of Proteobacteria and Bacteroidetes in rhizosphere soil was significantly higher than that in bulk soil; Crenarchaeota and Gemmatimonadetes were on the opposite (n = 36, $p < 0.01$). In bulk soil, the relative abundance of Acidobacteria showed an increasing trend along the chronosequence, while Actinobacteria displayed the opposite trend. A similar trend was also found in rhizosphere soil of the top line samples.

The dominant genera (average relative abundance > 0.5%) across all samples were *Candidatus Nitrososphaera* (5.7%), *Arthrobacter* (0.83%), *Rubrobacter* (0.78%), *Modestobacter* (0.55%), *Flavobacterium* (0.53%), *Pseudomonas* (0.53%), and *Gemmata* (0.52%). Seventeen other genera were found in the majority of samples with an average relative abundance of more than 0.18% (**Figure 2B**). Among them, the relative abundance of the predominant genus *Candidatus Nitrososphaera* in rhizosphere soil (3.69%) was significantly lower than that in bulk soil (7.89%) (n = 36, $p < 0.001$). The relative abundance of *Modestobacter* and *Flavobacterium* in rhizosphere soil was significantly higher than in bulk soil (n = 36, $p < 0.001$, $p < 0.05$). The relative abundances of *Rubrobacter* was on the opposite (n = 36, $p < 0.05$). Flavobacterium has an increased trend along the chronosequence. Modestobacter was decreased from site IS to SM and then increased from SM to BD in bulk soil, while in the rhizosphere they were on the opposite trend. *Rubrobacter* was relatively stable from site IS to SM, and then decreased from site SM to BD in bulk soil but increased in rhizosphere soil.

Alpine Plants Altered Prokaryotic Community α-Diversity Succession Along the Chronosequence

Prokaryotic community α-diversity indices showed mild succession in the glacier retreat chronosequence (**Figure 3**). In total, the averages of the three indices in the rhizosphere soil were slightly higher than those in bulk soils. In bulk soils, the indices were decreased from site IS to SM and then increased from SM to BD; the final value in BD was higher than that in IS in the line Top sample (**Figures 3A,C,E**). The opposite trend of succession was found in the rhizosphere soil in the line Top sample (**Figures 3B,D,F**). In line middle, the richness in bulk soil was relatively stable in bulk soil, but that has a drop in the rhizosphere soil, especially at site BD (**Figures 3E,F**).

Alpine Plant Effects on Prokaryotic Community β-Diversity

Alpine plants affect the prokaryotic community β-diversity (**Figure 4**). Bulk and rhizosphere soils separated well from each other along the first axis (**Figure 4A**). This was in agreement with the Adonis analysis, which showed that the community composition of bulk and rhizosphere soils were significantly different (n = 18, $p < 0.001$). In bulk soil, each of the sampling sites (IS, SM, and BD) grouped separately (**Figure 4B**). In rhizosphere soil, the IS samples were close to SM along the positive direction of the second PCoA axis, but two subgroups were found in the BD samples, one of which (three samples all from the middle line) was much closer to the other two sampling sites (**Figures 4C,E,F**). According to Adonis analysis, the community compositions of IS, SM, and BD were significantly different in total, bulk, and rhizospheres soils, respectively (n = 12, 6, 6, $p < 0.001$). For each of the three sample locations (IS, SM, and BD), Adonis analysis showed that the bulk and rhizosphere soils within the same site were significantly different from each other (n = 6, $p < 0.01$) (**Figures 4D–F**). From IS to BD, the distance between the group centroids of the middle and top lines showed an increasing trend and no significant difference (n = 6, p = 0.576, 0.314, and 0.053, respectively) (**Figures 4D–F**). Together, these results suggested that the prokaryotic community composition was different at all three locations.

Special Selection of Microorganisms by Alpine Plants Along the Chronosequence

Across the glacier successional stages, different clades were detected in the bulk and rhizosphere soil fractions, which explained the statistically significant differences between their respective microbial communities (**Figure 5**). The numbers of discriminant clades were 24 and 11 from bulk and rhizosphere soils, respectively (**Figure 5A**). The number of discriminant clades increased in bulk soils from site IS to BD (6, 9, and 12), while the opposite trend was observed in rhizosphere soils (5, 4, and 3) (**Figures 5E,F**). In general, the phyla Proteobacteria, especially class Alphaproteobacteria, and Actinobacteria were significantly enriched in rhizosphere soil; meanwhile, Acidimicrobiia and Gemmatimonadetes were significantly enriched in bulk soil (**Figure 5A**). The number of discriminant clades in line top bulk and rhizosphere soils were both higher than in the respective line middle soils. In line middle, Alphaproteobacteria were significantly enriched in the rhizosphere soil, while Acidimicrobiia were enriched in bulk soil (**Figure 5B**). In line top, the phylum Proteobacteria, and specifically the class Alphaproteobacteria, was significantly enriched in rhizosphere soil. Meanwhile, Tenericute and Mollicutes were more enriched in bulk soil (**Figure 5C**). Among three sample sites in the total sample, site IS was enriched with Actinobacteria, SHA_109, and Thermoleophilia. Site SM was enriched with Betaproteobacteria and OP1.MSBL6. Site BD was enriched with Planctomycetia and Verrucomicrobia (**Figure 5D**).

Effect of Environmental Factors of Microbial Community Composition

RDA analysis was used to assess which factors affected the microbial community structure of the samples (**Figure 6**). In total, eight factors were detected: TOC, TN, NH_4^+-N, NO_3^--N, C/N ratio, WC, plants, and distance. For total soil samples, all factors except TOC had significant impact on the

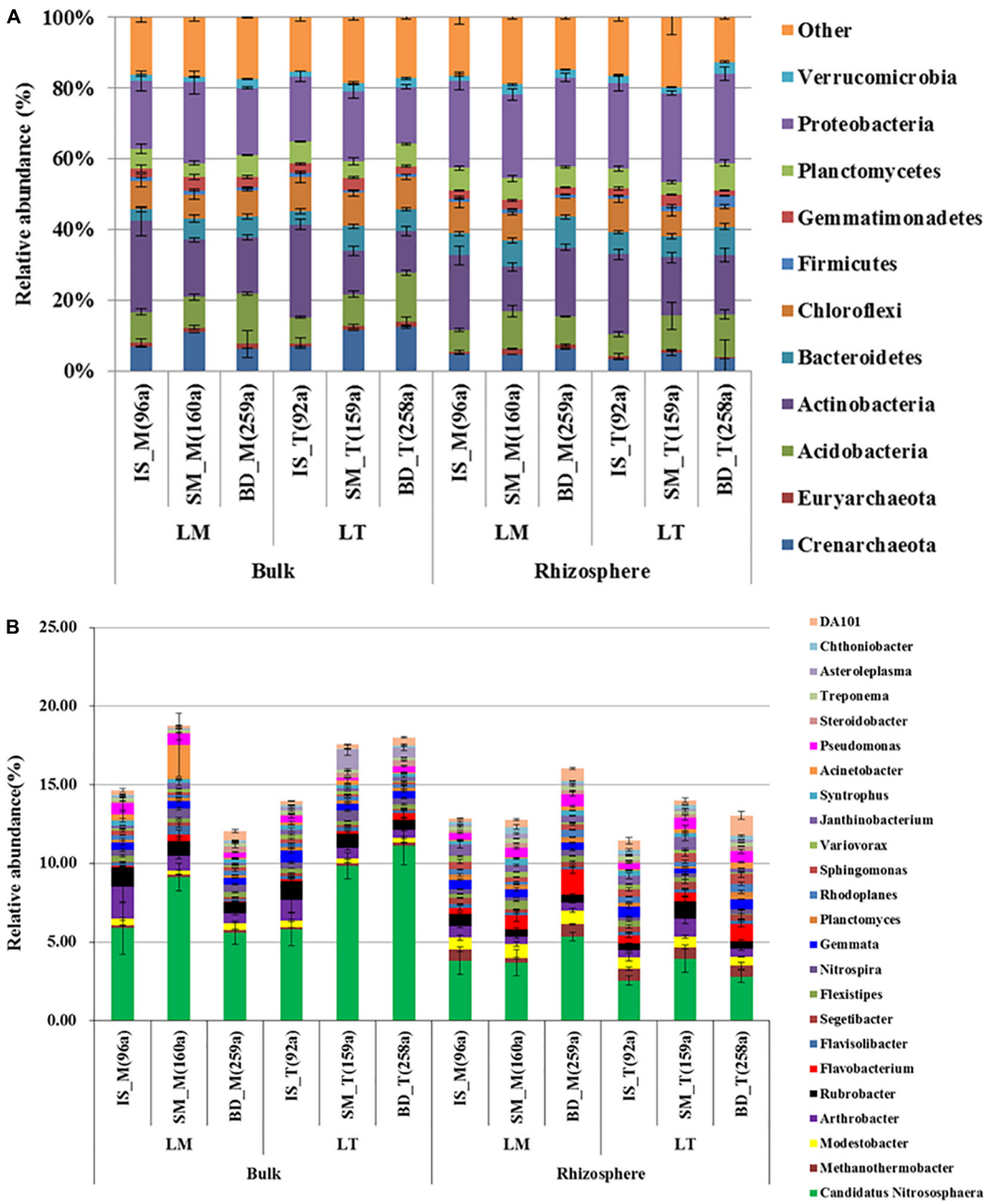

FIGURE 2 | The relative abundance of prokaryotic phyla **(A)** and genera (**B**). Sites: IS, middle of glacier terminal and small Qiangyong lake; SM, midway along the edge of the smaller lake; BD, approximately one-third of the way along the edge of the larger lake; M and T, the sampling lines, Middle and Top, respectively; S and R, bulk and rhizosphere soils, respectively. The relative abundances are given as a percentage of each phylum or genus against the total 16S rRNA gene sequences. Only phyla with average relative abundance > 1% are shown. Only genera with average relative abundance > 0.18% are shown. Data are shown as average ± SE (n = 3).

prokaryotic community structure ($n = 36, p < 0.001$) (**Figure 6A**). Among them, all the rhizosphere soils were significantly affected by plants. Distance affects both bulk and rhizosphere soil prokaryotic community structure; its effect on different samples sites was in the order of BD > SM > IS. The C/N affected the bulk soil microbial community structure more than that of the rhizosphere soil, while NH_4^+-N, NO_3^--N, and WC had the opposite trend (**Figures 6A,D–F**). In bulk soil, all factors except plants had significant effect on the prokaryotic community structure (n = 18, $p < 0.001$); the effect of C/N from different sample sites was in the order of SM > IS > BD (**Figure 6B**). In rhizosphere samples, NH_4^+-N, NO_3^--N, C/N ratio, WC,

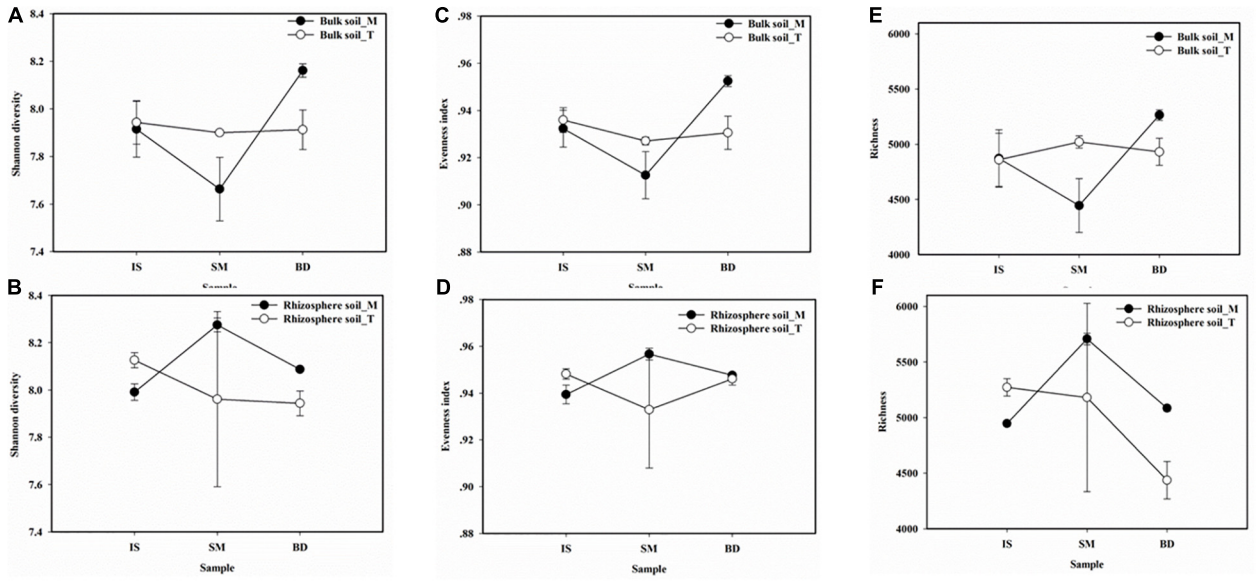

FIGURE 3 | Prokaryotic community α-diversity succession along the chronosequence. **(A)** and **(B)** were Shannon diversity in bulk and rhizoshpher soils, respectively. **(C)** and **(D)** were the Evenness index in bulk and rhizoshpher soils, respectively. **(E)** and **(F)** were the Richness in bulk and rhizoshpher soils, respectively. Data were shown as average ± se (n = 3). Sites: IS, middle of glacier terminal and small Qiangyong lake; SM, midway along the edge of the smaller lake; BD approximately one-third of the way along the edge of the larger lake. M and T indicate the sampling lines, Middle and Top, respectively.

and distance were the main factors affecting the prokaryotic community structure (n = 18, $p < 0.05$); the effect of NH_4^+-N, NO_3^--N, and WC from different sample sites was in the similar order of IS > SM > BD (**Figure 6C**).

DISCUSSION

A Special Prokaryote Community Structure Has Been Found in the High-Elevation Glacier Chronosequences

As expected, the special prokaryote community structures in the high-elevation Qiangyong Glacier foreland soil were revealed in this study. The relative abundance of seven phyla was above 5%; in a total of 11 phyla the relative abundance was above 1% (**Figure 2A**). The number of phyla was higher than that found in the frozen soil [five most abundant phyla (>2%)] from the glacier of the northwestern Himalayas of Jammu and Kashmir, India (Gupta et al., 2020); the two studies both were conducted in Himalayas. Furthermore, the three top abundant phyla detected in this research were also found in other glacier forelands of the Tibetan Plateau but with different order and abundance. For example, Proteobacteria (43%) and Actinobacteria (16%) were the predominant bacteria phyla in soils from Hailuogou Glacier retreat region. Similarly, they were also reported on the Muztag Ata Glacier chronosequence (Khan et al., 2020) and the frontier of Baishui Glacier No.1 (Sajjad et al., 2021). In this research, the Proteobacteria average relative abundance was 21.95% in total and comparable to the data from Gupta's report (Gupta et al., 2020). The reason may be the effect of elevation, as in this research the glacier termini were at the elevation of 5,000–5,100 m; the height of the glacier is 4,700 m a.s.l. in Gupta's report. The glaciers in other reports cited above were all below 4,400 m (Sun et al., 2016; Khan et al., 2020; Sajjad et al., 2021). These results suggested that as the elevation increased, the abundance of Proteobacteria decreased; at the same time, other phyla were enriched. As the microbial community reports on high-elevation glacier foreland were few, further research in this area is being looked forward to.

Plant Rhizosphere Selected Specific Prokaryotic Communities and Shifted Nutrient Accumulation Along the Qiangyong Chronosequence

The relative abundance of Proteobacteria, Bacteroidetes, *Modestobacter*, and *Flavobacterium* in rhizosphere soil was significantly higher than that in bulk soil; some genera were in the opposite succession trend compared with bulk and rhizosphere soil (**Figure 2**). The LEfse detected differential clades infractions, which consistently explained the statistically significant difference between the bulk soil and rhizosphere soil prokaryotic communities (**Figure 5**). In general, the phyla Proteobacteria, especially class Alphaproteobacteria, and Actinobacteria were significantly enriched in rhizosphere soil; meanwhile, Acidimicrobiia, and Gemmatimonadetes were significantly enriched in bulk soil. At three sites, bulk and rhizosphere soil also had enriched clades. This result was different from reports found for low elevation, such as the finding that rhizobacterial communities were mainly composed

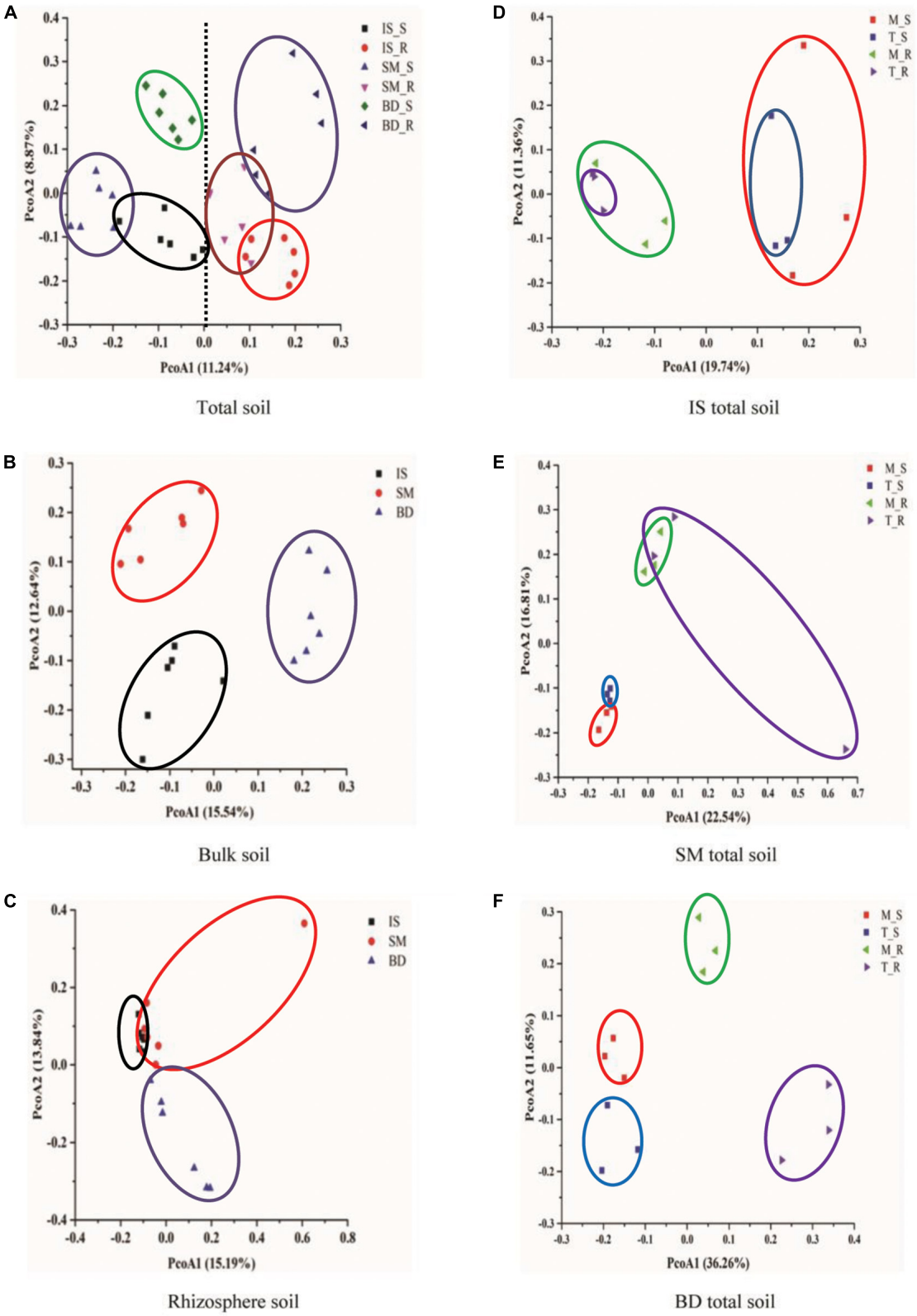

FIGURE 4 | PCoA plot of soil bacterial community composition. **(A)** For total soil, **(B)** for bulk soil, **(C)** for rhizosphere soil, **(D)** for IS total soil, **(E)** for SM total soil, and **(F)** for BD total soil. Sites: IS, middle of glacier terminal and small Qiangyong lake; SM, midway along the edge of the smaller lake; BD approximately one-third of the way along the edge of the larger lake. M and T indicate the sampling lines, Middle and Top, respectively; S and R represent bulk and rhizosphere soils, respectively.

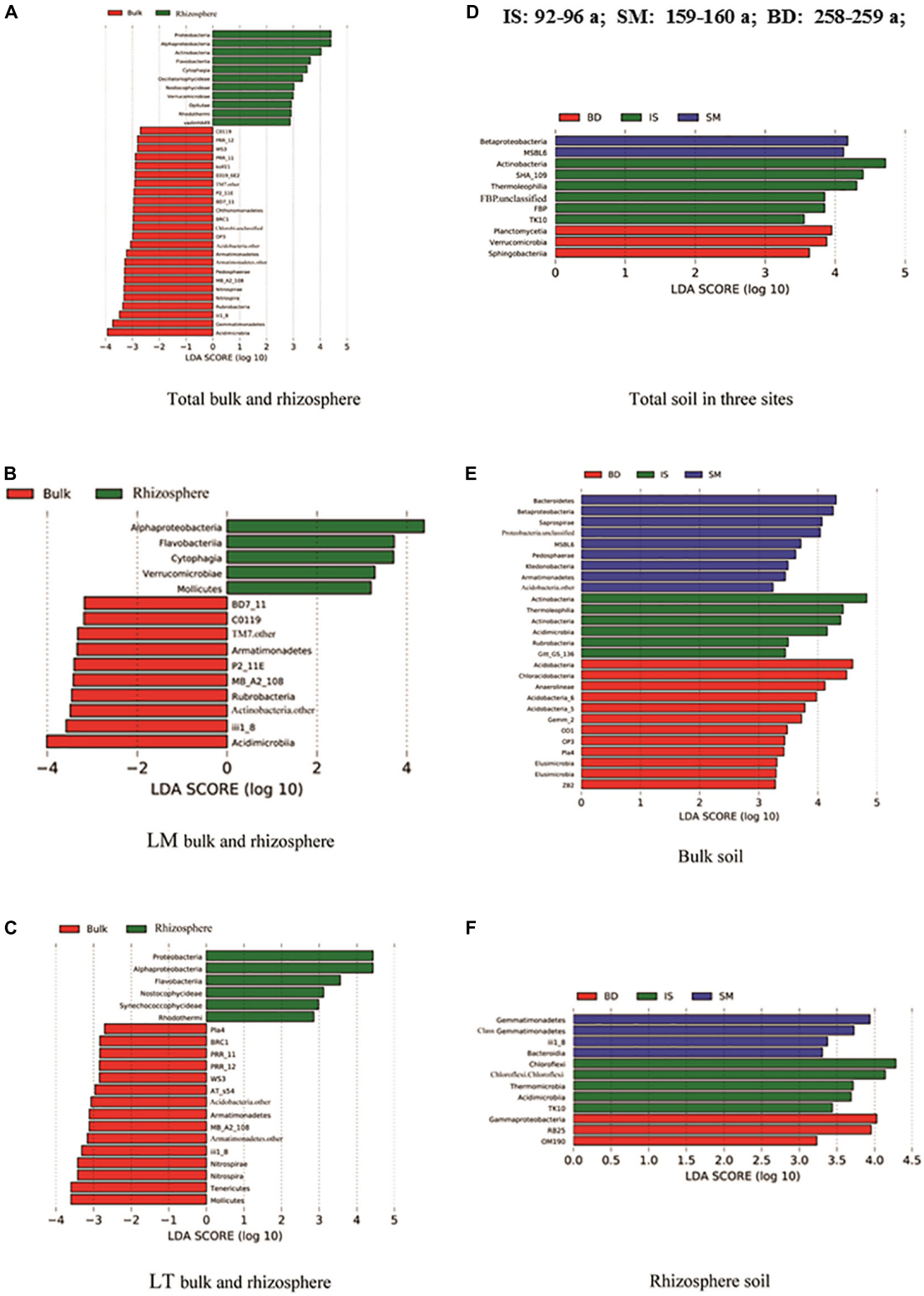

FIGURE 5 | LEfSe analyses of different sample groups. **(A)** For tobal bulk and rhizosphere soil, **(B)** for LM bulk and rhizosphere soil, **(C)** for LT bulk and rhizosphere soil, **(D)** for total soil in three sites, **(E)** for bulk soil, and **(F)** for rhizosphere soil. Sites: IS, middle of glacier terminal and small Qiangyong lake; SM, midway along the edge of the smaller lake; BD approximately one-third of the way along the edge of the larger lake. LM and LT indicate the sampling lines, Middle and Top, respectively; the different colors in the LE analysis represent different groups. The closer the sample distance is, the more similar the microbial compositions of the samples are, and the smaller the difference is ($P \leq 0.01$).

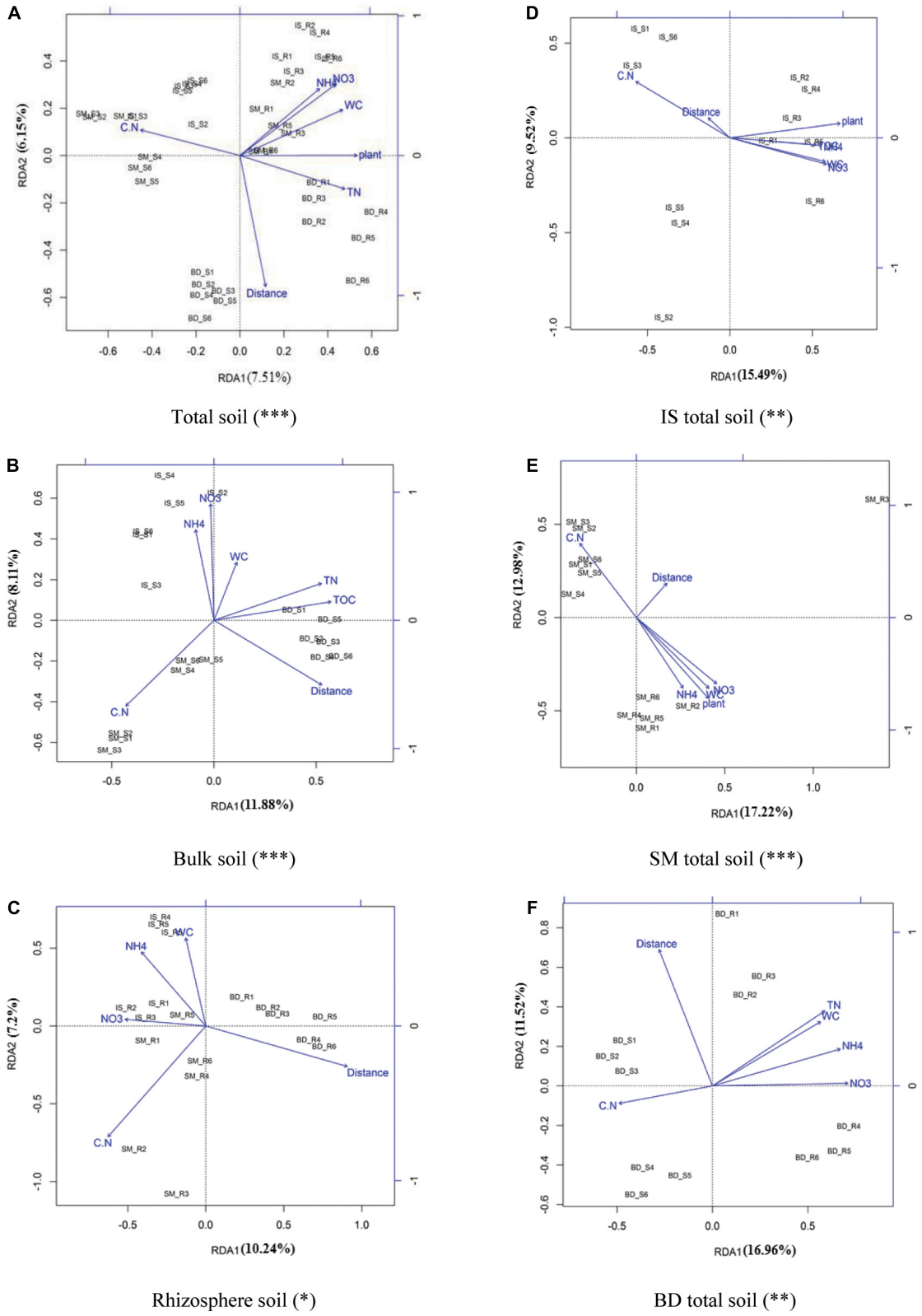

FIGURE 6 | Bacterial community and environment factor RDA analysis. **(A)** For total soil, **(B)** for bulk soil, **(C)** for rhizosphere soil, **(D)** for IS total soil, **(E)** for SM total soil, and **(F)** for BD total soil. Sites: IS, middle of glacier terminal and small Qiangyong lake; SM, midway along the edge of the smaller lake; BD approximately one-third of the way along the edge of the larger lake. Significance: ***0.001; **0.01; and *0.05. C:N, total Carbon to total Nitrogen ratio; WC, water content.

of Acidobacteria and Proteobacteria, whereas bare soil was colonized by Acidobacteria and Clostridia in the foreland of Weisskugel Glacier (2,400 m a.s.l.) (Ciccazzo et al., 2014).

Meanwhile, in this research, the detected nutrient contents in rhizosphere soil were significantly higher than in bulk soil (**Supplementary Figure 1**). The α-diversity in rhizosphere soil was higher than in bulk soil (**Figure 3**); PCoA and Adonis analysis showed the bulk and rhizosphere soil microbial community was significantly different along the chronosequence (**Figure 4**). Based on the RDA analysis, plants also showed a significant effect on microbial community structures in Qiangyong Glacier foreland (**Figure 6**). These results suggest that the plants altered the soil C and N accumulation by shifting the soil prokaryotic community structure. This consisted of the most enriched phyla in rhizosphere soil (phyla Proteobacteria, especially class Alphaproteobacteria, and Actinobacteria) (**Figure 5**). Proteobacteria were detected as autotrophic microbial communities in Qiangyong Glacier-originated water (Kong et al., 2019). They were also supposed to have the ability for N fixation at a glacial foreland on Anvers Island (Strauss et al., 2012). Actinobacteria and Betaproteobacteria have the potential to photosynthesize; Betaproteobacteria and Alphaproteobacteria have the potential to perform nitrification and denitrification of microbes from supraglacial cryoconite of polar regions (Cameron et al., 2012). Furthermore, the dominance of Actinobacteria in Antarctic soils has been linked to specific trace elements (magnesium, calcium, and potassium) and salts in the associated glacier forelands (Bajerski and Wagner, 2013). Actinobacteria are adapted to oligotrophic environments where their hyphae allow them to restore nutrients and moisture through pores in the soil (Arocha-Garza et al., 2017; Zhang et al., 2019). The rhizosphere soil microbial development significantly affected soil organic C and total N accumulation in the Hailuogou Glacier forefield (Li et al., 2020). The results above suggest that the emergence of plants had a selective effect on microorganisms and promoted the accumulation of C and N in this high-elevation glacier retreat in rhizosphere soils.

Environmental Parameters Related to Pedogenesis Shape the Prokaryotic Communities in Bulk and Rhizosphere Soils

The concentration of key nutrients related to soil fertility changed from site IS to BD along the Qiangyong chronosequence (**Supplementary Figure 1**). All detected environment parameters had effects on the high-elevation prokaryotic community composition succession (**Figure 6**). This result means that the distance, plant, and soil physicochemical properties together determine microbial community structures in the Qiangyong Glacier foreland. This result was partly similar to the reports of others. Microbial community structure is strongly conditioned by the successional stage, deglaciation time, water content, plants, spruce leachate, and the C and N contents of the foreland soil (Tscherko et al., 2004; Noll and Wellinger, 2008; Göransson et al., 2011; Knelman et al., 2012, 2018; Górniak et al., 2017; Kim et al., 2017; Bai et al., 2020). Most of the above reports did not include rhizosphere soil, and these studies were on elevations of < 1,000 m (Knelman et al., 2012, 2018; Kim et al., 2017), 1,780 m (Górniak et al., 2017), 1,920–2,054 m (Göransson et al., 2011), 1,950–2,050 m (Noll and Wellinger, 2008), 2,280–2,450 m (Tscherko et al., 2004), and 2,951 m at glacier termini (Bai et al., 2020). The effects of the high-elevation environmental parameters on the glacier foreland prokaryotic community were seldom found in field research. In this research, the effect of C/N on the bulk soil microbial community was greater than that in rhizosphere soil, while NH_4^+-N, NO_3^--N, and WC had the opposite trend. The above detected different physicochemical properties in each research, and most of them were in a relatively low elevation glacier retreat area. These results suggest that the distinct habitat was not ignorable. All the above results showed that plants, habitat, and glacier retreat chronosequence collectively control prokaryotic community composition and succession.

CONCLUSION

The soil TOC and TN contents showed an increasing trend along the deglaciation chronosequence, while ammonium and nitrate content showed the opposite trend. TOC, TN, ammonium, and nitrate contents in rhizosphere soil were significantly higher than in bulk soil ($p < 0.05$). In total, 11 phyla (relative abundance > 1%) and 7 dominant genera (relative abundance > 0.5%) were observed in all samples. According to LEfSe analyses, the Proteobacteria, Alphaproteobacteria, and Actinobacteria were significantly enriched in rhizosphere soil. Acidimicrobiia and Gemmatimonadetes were significantly enriched in bulk soil. No statistically significant difference was observed in the values from multiple α-diversity indices. Bacterial β-diversity showed that bulk and rhizosphere soils were obviously separate. Distance, soil C/N ratio, plants, and physicochemical properties all affected the soil bacterial community composition along the chronosequence. These results indicated that plants, habitat, and glacier retreat chronosequence collectively control prokaryotic community composition and succession.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

JL and WK contributed to conception and design of the study. PX and CZ collected the samples and determined soil physicochemical properties. JL conducted soil DNA extractions, performed the statistical analysis, and wrote the first draft of the manuscript. XL conducted the high-throughput sequencing. All

authors contributed to manuscript revision and approved the submitted version.

FUNDING

This project was financially supported by the Chinese Academy of Sciences (KZZD-EW-TZ-14 and XDB15010203 to WK), the National Natural Science Foundation of China (41471054 and 42177101 to WK), and the China Postdoctoral Science Foundation (2014M550849 to JL).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.736407/full#supplementary-material

REFERENCES

Arocha-Garza, H. F., Canales-Del Castillo, R., Eguiarte, L. E., Souza, V., and De la Torre-Zavala, S. (2017). High diversity and suggested endemicity of culturable Actinobacteria in an extremely oligotrophic desert oasis. *PeerJ* 5:e3241. doi: 10.7717/peerj.3247

Bai, Y., Huang, X. Y., Zhou, X. R., Xiang, Q. J., Zhao, K., Yu, X. M., et al. (2019). Variation in denitrifying bacterial communities along a primary succession in the Hailuogou Glacier retreat area, China. *PeerJ* 7:7356. doi: 10.7717/peerj.7356

Bai, Y., Xiang, Q., Zhao, K., Yu, X., Chen, Q., Ma, M., et al. (2020). Plant and soil development cooperatively shaped the composition of the phoD-harboring bacterial community along the primary succession in the Hailuogou glacier chronosequence. *mSystems* 5:e00475-20. doi: 10.1128/mSystems.00475-20

Bajerski, F., and Wagner, D. (2013). Bacterial succession in Antarctic soils of two glacier forefields on Larsemann Hills, East Antarctica. *FEMS Microbiol. Ecol.* 85, 128–142. doi: 10.1111/1574-6941.12105

Bernasconi, S. M., Bauder, A., Bourdon, B., Brunner, I., Bunemann, E., Christl, I., et al. (2011). Chemical and biological gradients along the Damma Glacier Soil chronosequence, Switzerland. *Vadose Zone J.* 10, 867–883. doi: 10.2136/vzj2010.0129

Bradley, J. A., Arndt, S., Sabacká, M., Benning, L. G., Barker, G. L., Blacker, J. J., et al. (2016). Microbial dynamics in a high Arctic glacier forefield: a combined field, laboratory, and modelling approach. *Biogeosciences* 13, 5677–5696. doi: 10.5194/bg-13-5677-2016

Bradley, J. A., Singarayer, J. S., and Anesio, A. M. (2014). Microbial community dynamics in the forefield of glaciers. *Proc. R. Soc. B Biol. Sci.* 281:20140882. doi: 10.1098/rspb.2014.0882

Bueno de Mesquita, C. P., Knelman, J. E., King, A. J., Farrer, E. C., Porazinska, D. L., Schmidt, S. K., et al. (2017). Plant colonization of moss-dominated soils in the alpine: microbial and biogeochemical implications. *Soil Biol. Biochem.* 111, 135–142. doi: 10.1016/j.soilbio.2017.04.008

Cameron, K. A., Hodson, A. J., and Osborn, A. M. (2012). Carbon and nitrogen biogeochemical cycling potentials of supraglacial cryoconite communities. *Polar Biol.* 35, 1375–1393. doi: 10.1007/s00300-012-1178-3

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., et al. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nat. methods* 7, 335–336. doi: 10.1038/nmeth.f.303

Castle, S. C., Sullivan, B. W., Knelman, J., Hood, E., Nemergut, D. R., Schmidt, S. K., et al. (2017). Nutrient limitation of soil microbial activity during the earliest stages of ecosystem development. *Oecologia* 185, 513–524. doi: 10.1007/s00442-017-3965-6

Cazzolla Gatti, R., Dudko, A., Lim, A., Velichevskaya, A. I., Lushchaeva, I. V., Pivovarova, A. V., et al. (2018). The last 50 years of climate-induced melting of the Maliy Aktru glacier (Altai Mountains, Russia) revealed in a primary ecological succession. *Ecol. Evol.* 8, 7401–7420. doi: 10.1002/ece3.4258

Ciccazzo, S., Esposito, A., Rolli, E., Zerbe, S., Daffonchio, D., and Brusetti, L. (2014). Different pioneer plant species select specific rhizosphere bacterial communities in a high mountain environment. *SpringerPlus* 3:391. doi: 10.1186/2193-1801-3-391

Donhauser, J., and Frey, B. (2018). Alpine soil microbial ecology in a changing world. *FEMS microbiol. Ecol.* 94:fiy099. doi: 10.1093/femsec/fiy099

Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C., and Knight, a.R (2011). UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* 27, 2194–2200. doi: 10.1093/bioinformatics/btr381

Göransson, H., Venterink, H. O., and Bååth, E. (2011). Soil bacterial growth and nutrient limitation along a chronosequence from a glacier forefield. *Soil Biol. Biochem.* 43, 1333–1340. doi: 10.1016/j.soilbio.2011.03.006

Górniak, D., Marszałek, H., Kwaśniak-Kominek, M., Rzepa, G., and Manecki, M. (2017). Soil formation and initial microbiological activity on a foreland of an Arctic glacier (SW Svalbard). *Appl. Soil Ecol.* 114, 34–44. doi: 10.1016/j.apsoil.2017.02.017

Gu, Z. Q., Liu, K. S., Pedersen, M. W., Wang, F., Chen, Y. Y., Zeng, C., et al. (2021). Community assembly processes underlying the temporal dynamics of glacial stream and lake bacterial communities. *Sci. Total Environ.* 761:143178. doi: 10.1016/j.scitotenv.2020.143178

Guo, G. X., Kong, W. D., Liu, J. B., Zhao, J. X., Du, H. D., Zhang, X. Z., et al. (2015). Diversity and distribution of autotrophic microbial community along environmental gradients in grassland soils on the Tibetan Plateau. *Appl. Microbiol. Biotechnol.* 99, 8765–8776. doi: 10.1007/s00253-015-6723-x

Gupta, V., Singh, I., Rasool, S., and Verma, V. (2020). Next generation sequencing and microbiome's taxonomical characterization of frozen soil of north western Himalayas of Jammu and Kashmir, India. *Electron. J. Biotechnol.* 45, 30–37. doi: 10.1016/j.ejbt.2020.03.003

Jain, S. (2014). "Glaciers," in *Fundamentals of Physical Geology*, ed. S. Jain (New Delhi: Springer India), 241–262. doi: 10.1007/978-81-322-1539-4_11

Jiang, Y., Lei, Y., Qin, W., Korpelainen, H., and Li, C. (2019). Revealing microbial processes and nutrient limitation in soil through ecoenzymatic stoichiometry and glomalin-related soil proteins in a retreating glacier forefield. *Geoderma* 338, 313–324. doi: 10.1016/j.geoderma.2018.12.023

Kabala, C., and Zapart, J. (2012). Initial soil development and carbon accumulation on moraines of the rapidly retreating Werenskiold Glacier, SW Spitsbergen, Svalbard archipelago. *Geoderma* 175, 9–20. doi: 10.1016/j.geoderma.2012.01.025

Kang, S., Xu, Y., You, Q., Flügel, W.-A., Pepin, N., and Yao, T. (2010). Review of climate and cryospheric change in the Tibetan Plateau. *Environ. Res. Lett.* 5:015101. doi: 10.1088/1748-9326/5/1/015101

Khan, A., Kong, W., Ji, M., Yue, L., Xie, Y., Liu, J., et al. (2020). Disparity in soil bacterial community succession along a short time-scale deglaciation chronosequence on the Tibetan Plateau. *Soil Ecol. Lett.* 2, 83–92. doi: 10.1007/s42832-020-0027-5

Kim, M., Jung, J. Y., Laffly, D., Kwon, H. Y., and Lee, Y. K. (2017). Shifts in bacterial community structure during succession in a glacier foreland of the High Arctic. *FEMS Microbiol. Ecol.* 93:fiw213. doi: 10.1093/femsec/fiw213

Knelman, J. E., Graham, E. B., Prevéy, J. S., Robeson, M. S., Kelly, P., Hood, E., et al. (2018). Interspecific plant interactions reflected in soil bacterial community structure and nitrogen cycling in primary succession. *Front. Microbiol.* 9:128. doi: 10.3389/fmicb.2018.00128

Knelman, J. E., Legg, T. M., O'Neill, S. P., Washenberger, C. L., González, A., Cleveland, C. C., et al. (2012). Bacterial community structure and function change in association with colonizer plants during early primary succession in a glacier forefield. *Soil Biol. Biochem.* 46, 172–180. doi: 10.1016/j.soilbio.2011.12.001

Kong, W. D., Liu, J. B., Ji, M. K., Yue, L. Y., Kang, S. C., and Morgan-Kiss, R. M. (2019). Autotrophic microbial community succession from glacier terminus to downstream waters on the Tibetan Plateau. *Fems Microbiol. Ecol.* 95:fiz074. doi: 10.1093/femsec/fiz074

Li, J., Rui, J., Yao, M., Zhang, S., Yan, X., Wang, Y., et al. (2015). Substrate type and free ammonia determine bacterial community structure in full-scale mesophilic anaerobic digesters treating cattle or swine manure. *Front. Microbiol.* 6:1337. doi: 10.3389/fmicb.2015.01337

Li, Q., Liu, Y., Gu, Y., Guo, L., Huang, Y., Zhang, J., et al. (2020). Ecoenzymatic stoichiometry and microbial nutrient limitations in rhizosphere soil along the Hailuogou Glacier forefield chronosequence. *Sci. Total Environ.* 704:135413. doi: 10.1016/j.scitotenv.2019.135413

Li, Y. F., Tian, L. D., Shi, X. L., and Yao, T. D. (2012). The composition characteristics and environmental significance of trace elements in the firn core of Qiangyong glacier, Southern Qinghai-Tibetan Plateau. *Geochimica* 41, 181–187.

Liu, J., Kong, W., Zhang, G., Khan, A., Guo, G., Zhu, C., et al. (2016). Diversity and succession of autotrophic microbial community in high-elevation soils along deglaciation chronosequence. *FEMS Microbiol. Ecol.* 92:fiw160. doi: 10.1093/femsec/fiw160

Luo, R., Cao, J., Liu, G., and Cui, Z. (2003). Characteristics of the subglacially-formed debris-rich chemical deposits and related subglacial processes of Qiangyong Glacier, Tibet. *J. Geogr. Sci.* 13, 455–462. doi: 10.1007/BF02837884

Ma, W., Shi, P., Li, W., He, Y., Zhang, X., Shen, Z., et al. (2010). Changes in individual plant traits and biomass allocation in alpine meadow with elevation variation on the Qinghai-Tibetan Plateau. *Sci. China Life Sci.* 53, 1142–1151. doi: 10.1007/s11427-010-4054-9

Mann, M. E., Zhang, Z., Rutherford, S., Bradley, R. S., Hughes, M. K., Shindell, D., et al. (2009). Global signatures and dynamical origins of the little ice age and medieval climate anomaly. *Science* 326, 1256–1260. doi: 10.1126/science.1177303

Mapelli, F., Marasco, R., Fusi, M., Scaglia, B., Tsiamis, G., Rolli, E., et al. (2018). The stage of soil development modulates rhizosphere effect along a High Arctic desert chronosequence. *ISME J.* 12, 1188–1198. doi: 10.1038/s41396-017-0026-4

Milner, A. M., Khamis, K., Battin, T. J., Brittain, J. E., Barrand, N. E., Füreder, L., et al. (2017). Glacier shrinkage driving global changes in downstream systems. *Proc. Natl. Acad. Sci.U.S.A.* 114, 9770–9778. doi: 10.1073/pnas.1619807114

Noll, M., and Wellinger, M. (2008). Changes of the soil ecosystem along a receding glacier: testing the correlation between environmental factors and bacterial community structure. *Soil Biol. Biochem.* 40, 2611–2619. doi: 10.1016/j.soilbio.2008.07.012

Sajjad, W., Ali, B., Bahadur, A., Ghimire, P. S., and Kang, S. C. (2021). Bacterial diversity and communities structural dynamics in soil and meltwater runoff at the frontier of Baishui Glacier No.1, China. *Microb. Ecol.* 81, 370–384. doi: 10.1007/s00248-020-01600-y

Schmidt, S., Reed, S. C., Nemergut, D. R., Grandy, A. S., Cleveland, C. C., Weintraub, M. N., et al. (2008). The earliest stages of ecosystem succession in high-elevation (5000 metres above sea level), recently deglaciated soils. *Proc. R. Soc. Lond. B Biol. Sci.* 275, 2793–2802. doi: 10.1098/rspb.2008.0808

Schuette, U. M., Abdo, Z., Foster, J., Ravel, J., Bunge, J., Solheim, B., et al. (2010). Bacterial diversity in a glacier foreland of the high Arctic. *Mol. Ecol.* 19(Suppl. 1), 54–66. doi: 10.1111/j.1365-294X.2009.04479.x

Segata, N., Izard, J., Waldron, L., Gevers, D., Miropolsky, L., Garrett, W. S., et al. (2011). Metagenomic biomarker discovery and explanation. *Genome Biol.* 12:R60. doi: 10.1186/gb-2011-12-6-r60

Strauss, S. L., Garcia-Pichel, F., and Day, T. A. (2012). Soil microbial carbon and nitrogen transformations at a glacial foreland on Anvers Island, Antarctic Peninsula. *Polar Biol.* 35, 1459–1471. doi: 10.1007/s00300-012-1184-5

Strauss, S. L., Ruhland, C. T., and Day, T. A. (2009). Trends in soil characteristics along a recently deglaciated foreland on Anvers Island, Antarctic Peninsula. *Polar Biol.* 32, 1779–1788. doi: 10.1007/s00300-009-0677-3

Sun, H.-y, Wu, Y.-h., Zhou, J., and Bing, H.-j (2016). Variations of bacterial and fungal communities along a primary successional chronosequence in the Hailuogou glacier retreat area (Gongga Mountain, SW China). *J. Mt. Sci.* 13, 1621–1631. doi: 10.1007/s11629-015-3570-2

Tscherko, D., Hammesfahr, U., Marx, M.-C., and Kandeler, E. (2004). Shifts in rhizosphere microbial communities and enzyme activity of Poa alpina across an alpine chronosequence. *Soil Biol. Biochem.* 36, 1685–1698. doi: 10.1016/j.soilbio.2004.07.004

Yao, T., Thompson, L., Yang, W., Yu, W., Gao, Y., Guo, X., et al. (2012). Different glacier status with atmospheric circulations in Tibetan Plateau and surroundings. *Nat. Clim. Chang.* 2, 663–667. doi: 10.1038/nclimate1580

Yoshitake, S., Uchida, M., Iimura, Y., Ohtsuka, T., and Nakatsubo, T. (2018). Soil microbial succession along a chronosequence on a High Arctic glacier foreland, Ny-Ålesund, Svalbard: 10 years' change. *Polar Sci.* 16, 59–67. doi: 10.1016/j.polar.2018.03.003

Zhang, B., Wu, X., Tai, X., Sun, L., Wu, M., Zhang, W., et al. (2019). Variation in actinobacterial community composition and potential function in different soil ecosystems belonging to the arid Heihe river basin of Northwest China. *Front. Microbiol.* 10:2209. doi: 10.3389/fmicb.2019.02209

Zumsteg, A., Luster, J., Göransson, H., Smittenberg, R. H., Brunner, I., Bernasconi, S. M., et al. (2012). Bacterial, archaeal and fungal succession in the forefield of a receding glacier. *Microbial. Ecol.* 63, 552–564. doi: 10.1007/s00248-011-9991-8

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

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 11 October 2021
doi: 10.3389/fmicb.2021.741077

Shedding Light on Microbial "Dark Matter": Insights Into Novel Cloacimonadota and Omnitrophota From an Antarctic Lake

*Timothy J. Williams[1], Michelle A. Allen[1], Jonathan F. Berengut[2] and Ricardo Cavicchioli[1]**

[1] School of Biotechnology and Biomolecular Sciences, UNSW Sydney, Sydney, NSW, Australia, [2] EMBL Australia Node for Single Molecule Science, School of Medical Sciences, UNSW Sydney, Kensington, NSW, Australia

The potential metabolism and ecological roles of many microbial taxa remain unknown because insufficient genomic data are available to assess their functional potential. Two such microbial "dark matter" taxa are the *Candidatus* bacterial phyla Cloacimonadota and Omnitrophota, both of which have been identified in global anoxic environments, including (but not limited to) organic-carbon-rich lakes. Using 24 metagenome-assembled genomes (MAGs) obtained from an Antarctic lake (Ace Lake, Vestfold Hills), novel lineages and novel metabolic traits were identified for both phyla. The Cloacimonadota MAGs exhibited a capacity for carbon fixation using the reverse tricarboxylic acid cycle driven by oxidation of hydrogen and sulfur. Certain Cloacimonadota MAGs encoded proteins that possess dockerin and cohesin domains, which is consistent with the assembly of extracellular cellulosome-like structures that are used for degradation of polypeptides and polysaccharides. The Omnitrophota MAGs represented phylogenetically diverse taxa that were predicted to possess a strong biosynthetic capacity for amino acids, nucleosides, fatty acids, and essential cofactors. All of the Omnitrophota were inferred to be obligate fermentative heterotrophs that utilize a relatively narrow range of organic compounds, have an incomplete tricarboxylic acid cycle, and possess a single hydrogenase gene important for achieving redox balance in the cell. We reason that both Cloacimonadota and Omnitrophota form metabolic interactions with hydrogen-consuming partners (methanogens and Desulfobacterota, respectively) and, therefore, occupy specific niches in Ace Lake.

Keywords: Cloacimonadota, Omnitrophota, cellulosome, autotrophy, metagenome, Antarctic bacteria

OPEN ACCESS

Edited by:
Anne D. Jungblut, Natural History Museum, United Kingdom

Reviewed by:
Tobias Goris, German Institute of Human Nutrition Potsdam-Rehbruecke (DIfE), Germany
Magdalena R. Osburn, Northwestern University, United States

****Correspondence:***
Ricardo Cavicchioli
r.cavicchioli@unsw.edu.au

Specialty section:
This article was submitted to Extreme Microbiology, a section of the journal Frontiers in Microbiology

Received: *14 July 2021*
Accepted: *13 September 2021*
Published: *11 October 2021*

Citation:
Williams TJ, Allen MA, Berengut JF and Cavicchioli R (2021) Shedding Light on Microbial "Dark Matter": Insights Into Novel Cloacimonadota and Omnitrophota From an Antarctic Lake. Front. Microbiol. 12:741077. doi: 10.3389/fmicb.2021.741077

INTRODUCTION

Microorganisms make up the majority of the biomass of the planet, yet the genomic potential of many microbial species remains elusive. The existence of many hitherto unknown taxa has only been revealed through cultivation-independent approaches, particularly from 16S rRNA gene libraries, and metagenome data and the analysis of metagenome-assembled genomes (MAGs) (Hugenholtz et al., 1998; Rinke et al., 2013; Momper et al., 2017; Parks et al., 2017, 2020;

Nayfach et al., 2020; Zamkovaya et al., 2021). Uncultivated clades, referred to as "microbial dark matter," include lineages that are inferred to play key roles in ecosystem formation and nutrient cycling (Rinke et al., 2013; Parks et al., 2017, 2020; Nayfach et al., 2020; Zamkovaya et al., 2021), including in Antarctica (Cavicchioli, 2015; Panwar et al., 2020). The phyla *Candidatus* Cloacimonadota and *Candidatus* Omnitrophota (hereafter Cloacimonadota and Omnitrophota, respectively) are inferred to contribute to anaerobic recycling of organic matter, although their ecophysiological traits remain largely undetermined (Baricz et al., 2020; Suominen et al., 2021).

Phylum Cloacimonadota [originally WWE1 ("Waste Water of Evry 1"); Chouari et al., 2005a,b] belongs to the "Fibrobacteres-Chlorobia-Bacteroidetes" (FCB) superphylum of bacteria (Rinke et al., 2013). Cloacimonadota can be a major component of anaerobic digestors and especially important in lipid-rich waste (Toth and Gieg, 2018; Saha et al., 2019; Shakeri Yekta et al., 2019). The first named member of this phylum, *Ca.* Cloacimonas acidaminovorans, is based on a MAG from an anaerobic digester of a municipal wastewater treatment plant (Chouari et al., 2005a,b; Pelletier et al., 2008). Based on single-cell amplified genome (SAG) and metatranscriptomic analyses, syntrophic propionate oxidation was inferred for a novel *Ca.* Cloacimonas species from a terephthalate-degrading bioreactor (Nobu et al., 2015) and for *Ca.* Syntrophosphaera thermopropionivorans based on a MAG from a thermophilic biogas reactor (Dyksma and Gallert, 2019). As well as being recovered from anaerobic digesters, abundant Cloacimonadota have also been detected in natural environments, including in anoxic and sulfidic water layers of the Black Sea (Suominen et al., 2021; Villanueva et al., 2021), Ursu Lake, Romania (Baricz et al., 2020), and from the Thuwal cold seep brine pool of the Red Sea (Zhang et al., 2016). Based on incubations with complex carbon substrates and analysis of MAGs, the Black Sea Cloacimonadota were inferred to be fermentative heterotrophic generalists capable of assimilating diverse carbon sources, including proteins (Suominen et al., 2021).

Phylum Omnitrophota [originally candidate division OP3 (Obsidian Pool 3)] was first discovered in 16S rRNA gene libraries generated from a hot spring sediment at the Yellowstone National Park (Hugenholtz et al., 1998). Further Omnitrophota sequences were detected in anoxic environments such as terrestrial subsurface fluids, flooded paddy soils, marine sediments, lagoon sediments, hypersaline deep sea waters, freshwater lakes, aquifers, methanogenic bioreactors, and acidic peatland soils (Derakshani et al., 2001; Glöckner et al., 2010; Rinke et al., 2013; Dombrowski et al., 2017; Momper et al., 2017; Lin et al., 2020; Santos et al., 2020). Phylogenetically, Omnitrophota has been assigned to the "Planctomycetes-Verrucomicrobia-Chlamydiae" superphylum based on 16S rRNA analysis (Wagner and Horn, 2006; Pilhofer et al., 2008; Glöckner et al., 2010), which was confirmed using metagenome-based analysis (Rinke et al., 2013). The nominative species *Candidatus* Omnitrophus fodinae SCGC AAA011-A17 is based on a SAG from groundwater (Homestake Mine, South Dakota); genome analysis of this SAG, and other MAGs from the same deep subsurface locality, indicated capacities for carbon fixation by the Wood–Ljungdahl (WL) pathway (reductive acetyl-CoA pathway) (Rinke et al., 2013; Momper et al., 2017). The latter MAGs also possessed genes for hydrogen (H_2) oxidation, methane oxidation, and dissimilatory nitrate reduction (Momper et al., 2017). Single-cell analysis of *Ca.* Omnitrophus magneticus SKK-01 isolated from the suboxic layer of lake sediments (Lake Chiemsee, Bavaria) revealed ovoid, flagellated cells that harbored intracellular sulfur inclusions and chains of magnetite (Kolinko et al., 2012); analysis of genome sequences identified genes associated with magnetosome biosynthesis, sulfur oxidation, and carbon fixation (Kolinko et al., 2016). Genomic analysis of an Omnitrophota MAG ("bin146") from the Black Sea inferred a fermentative heterotroph that scavenged low-molecular-weight organic substrates and was capable of glycolysis to acetate as well as H_2 production (Suominen et al., 2021).

Both Cloacimonadota and Omnitrophota were detected in Ace Lake (Panwar et al., 2020), a marine-derived, meromictic (stratified) system in the Vestfold Hills of Antarctica (Rankin et al., 1999; Lauro et al., 2011). The interface of the lake (12–15 m) is defined by a strong halocline and oxycline, dominated in the austral summer months by a species of the green sulfur bacterium *Chlorobium* (Ng et al., 2010; Lauro et al., 2011; Panwar et al., 2020). The oxic–anoxic interface and lower anoxic zone (16–24 m) support anaerobes, including members of Cloacimonadota and Omnitrophota, which were among the most abundant taxa, with peak relative abundances of 16 and 5%, respectively (Panwar et al., 2020). Preliminary analysis of the Ace Lake Cloacimonadota MAGs inferred a chemolithoautotrophic carbon fixation capacity driven by H_2 oxidation, while the functional potential of the Ace Lake Omnitrophota MAGs was not examined (Panwar et al., 2020). The Ace Lake data represents 120 metagenomes generated from size-fractionated samples representing a depth profile and a 10-year sampling period (Panwar et al., 2020). The large metagenome dataset provided a unique opportunity to reconstruct the metabolisms of these two "dark matter" candidate phyla, infer their ecophysiology, and consider the potential ecological niches they occupy in Ace Lake.

MATERIALS AND METHODS

Microbial biomass was sampled from Ace Lake in austral summers of 2006/2007 and 2008/2009, and a full Antarctic seasonal cycle of summer 2013/2014 to summer 2014/2015. Biomass was collected by sequential size fractionation through a 20 μm prefilter onto 3.0, 0.8-, and 0.1-μm pore-sized, large format (293-mm polyethersulfone membrane) filters, and DNA was extracted from the biomass as described previously (Ng et al., 2010). Six depths were sampled (surface, 5, 11.5–13, 12.7–14.5, 14–16, 18–19, and 23–24 m) with the precise depths varying depending on the water level in the lake (Panwar et al., 2020). In winter 2014, samples were not taken below the oxic–anoxic interface (Panwar et al., 2020). DNA was sequenced and the sequences uploaded to Integrated Microbial Genomes (IMG) (Huntemann et al., 2015) generating 120 individual metagenomes, as described previously (Panwar et al., 2020). High- and medium-quality MAGs were auto-generated from

individual metagenomes during the IMG pipeline process. QC-filtered raw reads from the individual Ace Lake metagenomes were co-assembled using Megahit v1.1.1 (Li et al., 2016) with a setting of meta-large, and MAGs were generated from the co-assembly using MetaBAT v2.12.1 with minContig length 2,500 bp (Kang et al., 2019). MAGs from the co-assembly (available in IMG as Metagenome ID 3300035698) were assessed for completeness and contamination using CheckM v1.0.7 (Parks et al., 2015), for taxonomic identity using RefineM v 0.0.23 (Parks et al., 2017), and for phylogenetic placement using Genome Taxonomy Database Toolkit (GTDB-Tk) v.1.4.0 with GTDB release R95 (Chaumeil et al., 2019; Parks et al., 2020). The GTDB-Tk dependencies were pplacer (Matsen et al., 2010), FastANI (Jain et al., 2019), Prodigal (Hyatt et al., 2010), FastTree 2 (Price et al., 2010), HMM (Eddy, 2011), and Mash (Ondov et al., 2016).

Metagenome-assembled genomes from the individual Ace Lake metagenomes and from the co-assembly that belonged to the phyla Cloacimonadota (one high- and 21 medium-quality MAGs) and Omnitrophota (9 high- and 72 medium-quality MAGs) were grouped based on average nucleotide identity (FastANI v 1.32; Jain et al., 2019), and average amino acid identity (CompareM v 0.1.1[1]) to identify representative MAGs for further examination (**Supplementary Table S1**).

Phylogenetic trees showing all Cloacimonadota and Omitrophota MAGs were generated by GTDB-Tk, based on a ∼5,000 amino acid-long concatenated multiple sequence alignment of 120 bacterial reference genes, and viewed with Dendroscope 3.5.7 (Huson and Scornavacca, 2012). Maximum likelihood phylogenies of selected novel and reference Cloacimonadota and Omnitrophota taxa were generated from the same GTDB-Tk concatenated multiple sequence alignments using W-IQ-Tree (Nguyen et al., 2015; Trifinopoulos et al., 2016) with autoselection of the best-fit model and 1,000 ultrafast bootstraps (Minh et al., 2013).

Of the total of 22 Cloacimonadota and 81 Omnitrophota MAGs identified in Ace Lake, certain MAGs were chosen for in-depth genomic examination, based on the aim of sampling the total known phylogenetic diversity of the respective phyla, as well as completeness of the MAGs (**Supplementary Table S1**). On this basis, 10 Cloacimonadota MAGs and 14 Omnitrophota MAGs were chosen. The genomic functional potential of the MAGs was assessed by considering cellular and metabolic traits based upon manual examination of proteins and pathways that was performed in a similar way to previous assessments of the veracity of gene functional assignments (Allen et al., 2009; Panwar et al., 2020; Williams et al., 2021). This method included the vetting via manual curation of the IMG protein annotations used in this study. All protein sequences were submitted to ExPASy BLAST (using the "UniProtKB/Swiss-Prot only" option) (Gasteiger et al., 2003); proteins needed to show ≥ 35% sequence identity to an experimentally verified protein in the ExPASy BLAST database for the functional annotation to be considered valid. If this threshold was not reached, protein sequences were submitted to InterProScan (Blum et al., 2020) to identify functional domains (e.g., catalytic domains; dockerin and cohesin domains) and potential subcellular locations (e.g., using signal peptides for an extracytoplasmic location; transmembrane helices for a membrane location). IMG annotations that could not be verified using this process were discarded. All of our protein identifications are considered putative. GH families were identified according to the CAZy (Carbohydrate-Active enZymes) classification (Lombard et al., 2014). Protein sequences that were identified as hydrogenases based on catalytic domains were classified further using the hydrogenase classifier HydDB (Søndergaard et al., 2016). Only those MAGs that were subjected to in-depth examination are named here; these were named according to recommendations for describing novel *Candidatus* species (Konstantinidis et al., 2017; Chuvochina et al., 2019; Murray et al., 2020).

[1]https://github.com/dparks1134/CompareM

RESULTS AND DISCUSSION

Genomic Assemblies and Phylogenetic Analysis

The genomes of 10 Cloacimonadota MAGs were interrogated (58–97% completeness; 0–4.4% contamination), which represent four novel genus-level and eight novel species-level taxa (**Table 1** and **Supplementary Table S2**). Based on phylogenetic analysis and GTDB taxonomy, the four novel genera are deeply nested within the phylum Cloacimonadota, within the class *Candidatus* Cloacimonadia. None of the four genera were found to be closely related to *Ca.* Cloacimonas or *Ca.* Syntrophosphaera, both of which belong to the family *Candidatus* Cloacimonadaceae (Dyksma and Gallert, 2019; **Figure 1** and **Supplementary Figure S1**).

The genomes of 14 Omnitrophota MAGs were interrogated (61–93% completeness; 0–9.1% contamination) that represent 11 novel genus-level and 13 novel species-level taxa (**Table 2** and **Supplementary Table S2**). Phylogenetic analysis and GTDB taxonomy revealed that the 11 genera represent two class-level and eight order-level clades. None of the Ace Lake MAGs were closely related to *Ca.* Omnitrophus (class *Candidatus* Omnitrophia, order *Candidatus* Omnitrophales), which the phylogenetic analysis recovered in a relatively basal position within the phylum (**Figure 2** and **Supplementary Figure S2**).

For both Cloacimonadota and Omnitrophota, individual genera and species are provided along with etymologies of all proposed names (**Tables 1**, **2** and **Supplementary Table S2**). Proteins and pathways discussed for all MAGs assigned to each genus are provided for Cloacimonadota (**Supplementary Tables S3, S4**) and Omnitrophota (**Supplementary Tables S5, S6**).

Cloacimonadota: Biopolymer Degradation and the Cell Envelope

Cloacimonadota are here inferred to be capable of degrading recalcitrant organic matter under anoxic conditions in Ace Lake. The MAGs of the four Ace Lake genera encode multiple glycoside hydrolase (GH) enzymes with signal peptides (indicating an extracytoplasmic location) and include enzymes that degrade polysaccharides and certain glycoconjugates,

TABLE 1 | Metabolic capacity of the Ace Lake Cloacimonadota inferred from metagenome-assembled genomes (MAGs).

Cloacimonadota		**Metabolic capacity**						
All genera and species		Anaerobic heterotroph secreted glycoside hydrolases and peptidases F-type ATP synthase						
***Candidatus* Genus**	***Candidatus* Species**	**MAGs (% completeness)**	**Extracellular features + secreted GHs**	**Fermentation substrates**	**Carbon fixation**	**Sulfur metabolism**	**Other bioenergetic complexes/enzymes**	**Hydrogenases**
Tenebribacter	burtonii[T] davisii mawsonii	3300035698_1346 (97%) 3300035698_1468 (97%) 3300035698_1174 (91%)	Poly-γ-glutamate synthesis; β-glucanase, β-glucosidase, glucosylceramidase, chitinase	Sugars, amino acids, 2-oxoacids, aldehydes, alcohols, glycerol, formate	Reverse tricarboxylic acid cycle	Oxidation of sulfur compounds (including thiosulfate), linked to Hdr reduction	Rnf, Nqr, Sud, HppA	Membrane-bound, H_2-evolving NiFe hydrogenase (Group 4g) linked to Mrp; FeFe hydrogenase (Group C1) for redox balance
Stygibacter	australis[T] frigidus	3300025642_13 (88%), 3300035698_2003 (69%) 3300035698_198 (58%)	Poly-γ-glutamate synthesis; halomucin-like protein; cellulosome-like; β-glucanase, β-glucosidase, α-amylase, glucosylceramidase, chitinase	Sugars, amino acids, 2-oxoacids, aldehydes, alcohols, glycerol, formate	Reverse tricarboxylic acid cycle	Oxidation of sulfur compounds (including thiosulfate), linked to Hdr reduction	Rnf, Nqr, Sud, HppA	Membrane-bound, H_2-evolving NiFe hydrogenase (Group 4g) linked to Mrp; FeFe hydrogenases (Groups A3 and C1) for redox balance
Zophobacter	franzmannii [T]	3300035698_360 (80%)	Poly-γ-glutamate synthesis; β-glucanase, β-glucosidase, chitinase	Sugars, amino acids, 2-oxoacids, aldehydes, glycerol, formate	–	–	Rnf, Nqr, Sud, HppA	–
Celaenobacter	antarcticus [T] polaris	3300025698_8 (92%), 3300035698_1683 (65%) 3300035698_1703 (91%)	β-glucosidase, β-galactosidase	Sugars, amino acids, 2-oxoacids, aldehydes, formate	Reverse tricarboxylic acid cycle	Assimilatory sulfate reduction	Rnf, Nqr, Sud, HppA	Membrane-bound, NiFe H_2-evolving hydrogenase (Group 4g) linked to Mrp; H_2-oxidizing, Hdr-linked cytoplasmic NiFe hydrogenase (Group 3c)

GH, glycoside hydrolase; Hdr, heterodisulfide reductase; HppA, pyrophosphate-energized sodium pump; Mrp, multicomponent Na^+:H^+ antiporter; Nqr, sodium-translocating NADH:quinone oxidoreductase; Rnf, ferredoxin:NAD^+-oxidoreductase complex; Sud, bifunctional sulfide dehydrogenase/ferredoxin:NADP oxidoreductase. [T]*Indicates type species.*

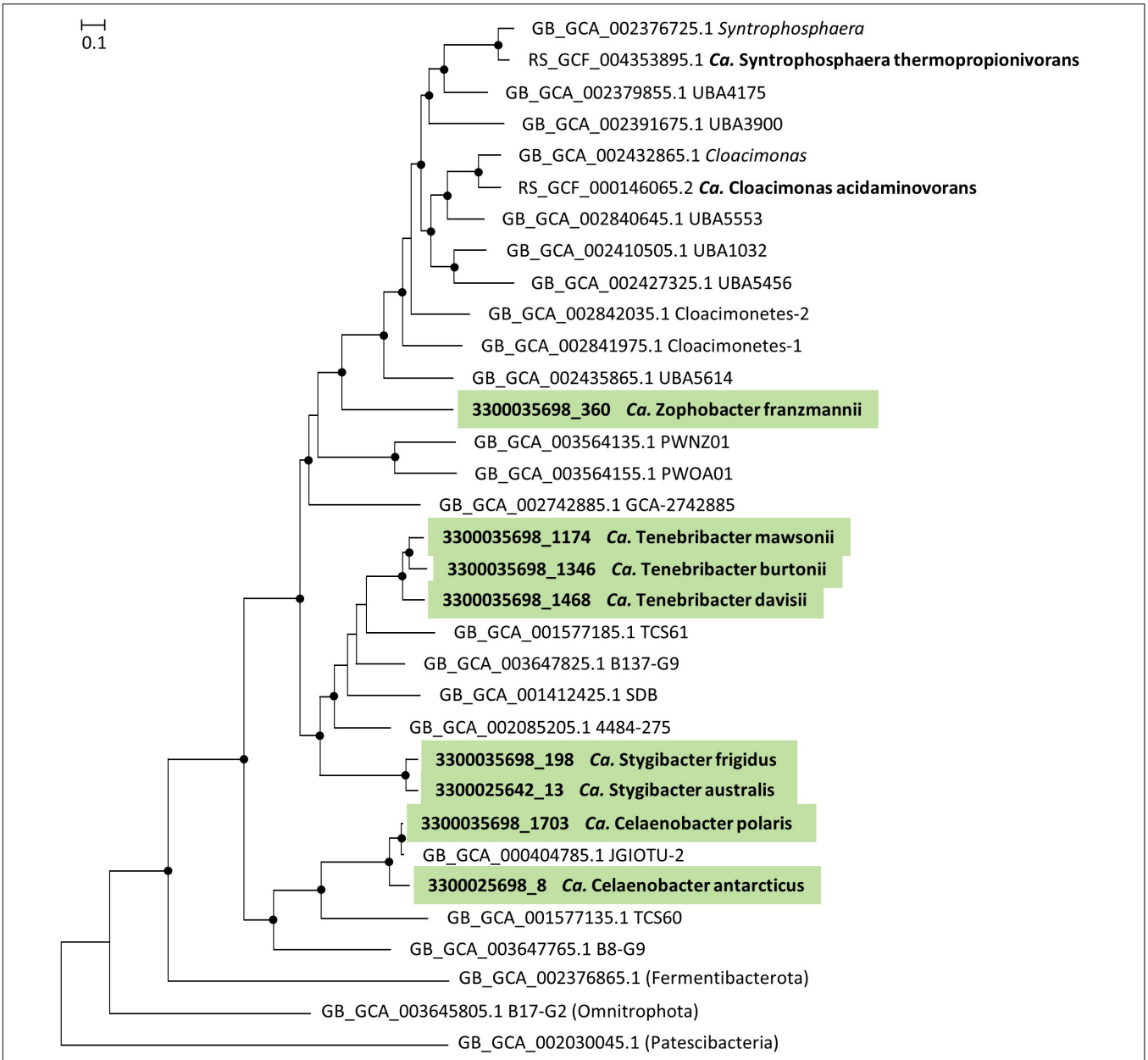

FIGURE 1 | Phylogeny of phylum *Candidatus* Cloacimonadota. Maximum likelihood tree constructed in IQ-Tree with autoselection of the best-fit model (LG + F + I + G4) and 1,000 ultra-fast bootstraps. UFBootstraps ≥ 95% (black dot); Metagenome-assembled genomes (MAGs) featured in this study (green) with their IMG MAG ID and proposed *Candidatus* genus and species name. Reference Cloacimonadota MAGs are shown with their Genome Taxonomy Database (GTDB) accession and GTDB taxonomy, except for *Candidatus* Cloacimonas acidaminovorans and *Candidatus* Syntrophosphaera thermopropionivorans, which already have names. The tree is rooted using a representative of the Patescibacteria.

which would release oligosaccharides and simple sugars such as glucose (**Table 1**). These hydrolytic enzymes belong to various GH families, indicating a range of potential substrates, such as starch, β-glucans, β-glucosides, chitin, and glucosylceramides (**Table 1** and **Supplementary Tables S3, S4**). The Ace Lake Cloacimonadota also encode diverse proteases and peptidases, including both secreted and cytoplasmic, indicating that polypeptides could be used as amino acid sources (**Supplementary Tables S3, S4**). Histidine degradation pathways are encoded in MAGs of all four genera, and *Ca.* Celaenobacter *gen. nov.* encodes proteins for tryptophan degradation (**Supplementary Table S3**). The abilities of Cloacimonadota to use complex sugars and proteins as organic substrates have been previously reported for this clade in both anaerobic digestors and lakes (Pelletier et al., 2008; Limam et al., 2014; Suominen et al., 2021).

However, in MAGs of the Ace Lake genus *Ca.* Stygibacter *gen. nov.* we identified genes for components of a putative

TABLE 2 | Metabolic capacity of the Ace Lake Omnitrophota inferred from metagenome-assembled genomes (MAGs).

Omnitrophota		**Metabolic capacity**				
All genera and species		Anaerobic heterotroph incomplete TCA cycle, terminating at fumarate (reductive branch) and 2-oxoglutarate (oxidative branch) secreted glycoside hydrolases and peptidases V-type ATP synthase				
***Candidatus* Genus**	***Candidatus* Species**	**MAGs (% completeness)**	**Fermentation substrates**	**Other carbon metabolism**	**Other bioenergetic complexes/enzymes**	**Hydrogenase**
Aceula	lacicola^{T}	3300035698_1100 (92%)	Sugars, glycerol	Glycogen synthesis	Rnf, Sud, HppA	NiFe hydrogenase (Group 3d)
	meridiana	3300035698_985 (92%)				
Zapsychrus	exili^{T}	3300035698_1555 (89%)	Sugars, glycerol	Glycogen synthesis	Rnf, Nqr, Sud, HppA	NiFe hydrogenase (Group 3d)
Gygaella	obscura^{T}	3300035698_1934 (85%)	Sugars, 2-oxoacids	Glycogen synthesis	Rnf, Sud, HppA	FeFe hydrogenase (Group A3)
Susulua	stagnicola^{T}	3300035698_1005 (82%)	Sugars, 2-oxoacids	Glycogen synthesis	Rnf, Sud, HppA	FeFe hydrogenase (Group A3)
Saelkia	tenebricola^{T}	3300035698_749 (91%)	Sugars, alcohols	Glycogen synthesis	Rnf, Mrp, Sud, HppA	NiFe hydrogenase (Group 4g), Mrp-linked
Kaelpia	aquatica^{T}	3300035698_2000 (93%)	Sugars, 2-oxoacids, alcohols	Glycogen synthesis	Rnf, Sud, HppA	FeFe hydrogenase (Group A3)
	imicola	3300035698_1655 (92%)				
Kappaea	frigidicola^{T}	3300035698_1500 (76%)	Sugars	Trehalose synthesis	Rnf, HppA	FeFe hydrogenase (Group A3)
Tantalella	remota^{T}	3300035698_1097 (93%)	Sugars, glycerol, 2-oxoacids, alcohols	Glycogen synthesis, trehalose synthesis	Rnf, Sud, HppA	NiFe hydrogenase (Group 3b)/sulfhydrogenase
Aadella	gelida^{T}	3300035698_848 (91%)	Sugars, glycerol, 2-oxoacids	Trehalose synthesis	Rnf, Sud	NiFe hydrogenase (Group 3b)/sulfhydrogenase
Gorgyraea	atricola^{T}	3300035698_32 (93%)	Sugars, 2-oxoacids	Wood–Ljungdahl pathway, glycogen synthesis, trehalose synthesis	Rnf, Mrp, Sud, HppA	NiFe hydrogenase (Group 4g), Mrp-linked
Orphnella	occulta^{T}	3300025586_21 (85%), 3300035698_104 (61%)	Sugars, 2-oxoacids	Trehalose synthesis	Rnf, Sud, HppA	NiFe hydrogenase (Group 3b)/sulfhydrogenase

HppA, pyrophosphate-energized sodium pump; Mrp, multicomponent Na^{+}:H^{+} antiporter; Nqr, sodium-translocating NADH:quinone oxidoreductase; Rnf, ferredoxin:NAD^{+}-oxidoreductase complex; Sud, bifunctional sulfide dehydrogenase/ferredoxin:NADP oxidoreductase. ^{T}*Indicates type species.*

extracellular, multienzyme complex for the binding and degradation of biopolymers (**Table 1** and **Figure 3A**), not previously reported for Cloacimonadota. In *Ca.* Stygibacter, certain signal-peptide-bearing enzymes contain C-terminal dockerin domains: chitinase homolog (GH18), α-amylase/α-mannosidase homolog (GH57), serine peptidase (Peptidase S8/S53 domain), and gingipain-like peptidase (Peptidase C25). *Ca.* Stygibacter also encodes a large (3,755 amino acids) non-catalytic scaffoldin-like protein (Artzi et al., 2017) that contains tandemly repeated cohesin and carbohydrate-binding (CBM2/CBM3) domains. We infer that these dockerin- and cohesin-domain proteins combine to produce a cellulosome-like structure, with biopolymer-degrading enzymes integrated into this scaffoldin-like protein via complementary cohesin–dockerin interactions (Artzi et al., 2017).

Canonical cellulosomes are extracellular complexes equipped with cellulose-, hemicellulose-, and polypeptide-targeting enzymes that are used by cellulolytic Clostridia (Firmicutes) to bind and degrade plant cell walls (Schwarz and Zverlov, 2006; Peer et al., 2009; Fontes and Gilbert, 2010). As in clostridial cellulosomes, the *Ca.* Stygibacter complex has a scaffoldin-like protein with tandem cohesin domains for integration of multiple dockerin-containing enzymes, and a conserved C-terminal domain (CTD) for direct attachment to its own cell surface (Lasica et al., 2017). In common with the clostridial cellulosome, we infer that the cellulosome-like structure in *Ca.* Stygibacter can mediate attachment of cells to insoluble substrates and promote degradation to soluble products that are taken up by the cell (Lamed et al., 1983). Unlike cellulosomes, we propose that, based on the annotated GHs and peptidases, this complex in *Ca.* Stygibacter is utilized for the attachment to and degradation of starch (via endohydrolysis) and chitinous and proteinaceous material rather than cellulose-rich plant-derived material. A non-canonical cellulosome-like structure ("planctosome") has also been reported in certain freshwater Planctomycetes (Nemodlikiaceae) for polypeptide degradation (Andrei et al., 2019), also distinct from the *Ca.* Stygibacter structure described here; our finding for the Ace Lake *Ca.* Stygibacter therefore adds to the repertoire of cellulosome-like complexes represented across the domain Bacteria.

The Ace Lake Cloacimonadota MAGs also encode other putative extracellular structures not previously reported for this phylum, all of which indicate an elaborate cell envelope (**Figure 3A**). Three genera (*Ca.* Stygibacter, *Ca.* Tenebribacter *gen. nov.*, *Ca.* Zophobacter *gen. nov.*) encode poly-γ-glutamate synthetase (CapBC) and other proteins required for synthesis and transport of poly-γ-glutamate (PGA), a biopolymer involved in capsule formation or released extracellularly as a water-binding component of a biofilm matrix (Rehm, 2010). The water-binding properties of PGA allow it to locally decrease high salt concentrations, allowing survival of the cell in high-salt environments (Kandler et al., 1983; Rehm, 2010). Additionally, the *Ca.* Stygibacter MAGs encode a glycine-rich protein (944 amino acids), with the N-terminal half containing a ∼440 amino acid region that has 27–30% identity to a ∼400–500 amino acid repeat sequence in halomucin (a very large protein in the halophilic archaeon *Haloquadratum walsbyi*) (Bolhuis et al., 2006), and the C-terminal half includes a dockerin domain. As proposed for *H. walsbyi*, it is possible that both PGA and the halomucin-like protein form a water-enriched capsule around the cell that facilitates growth in high

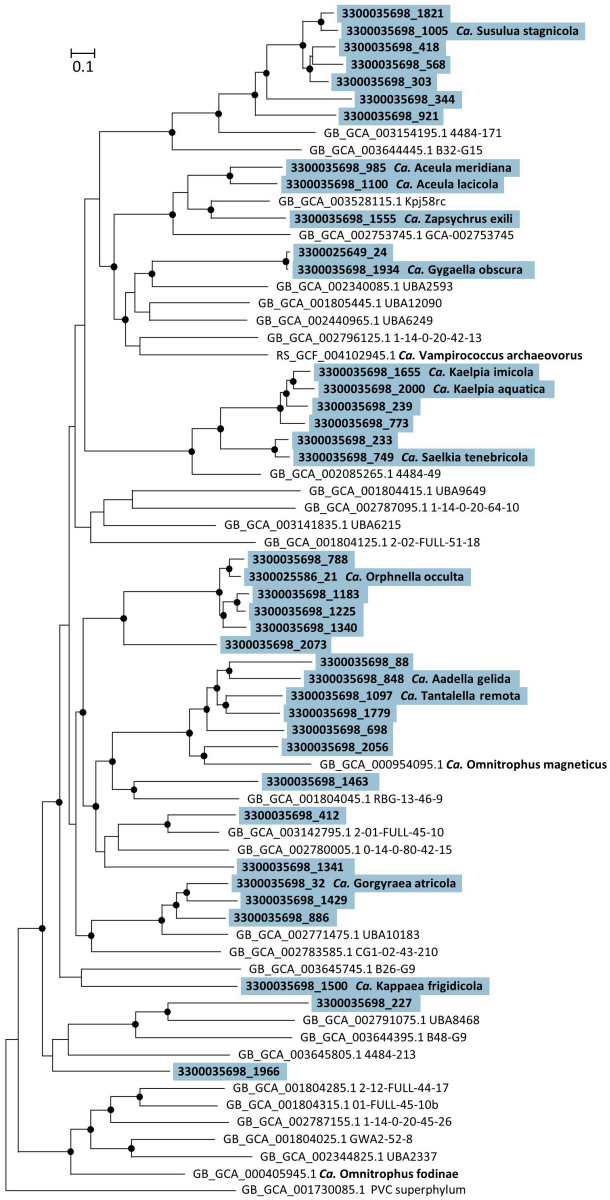

FIGURE 2 | Phylogeny of phylum *Candidatus* Omnitrophota. Maximum likelihood tree constructed in IQ-Tree with autoselection of the best-fit model (LG + F + I + G4) and 1,000 ultra-fast bootstraps. UFBootstraps ≥ 95% (black dot); Metagenome-assembled genomes (MAGs) featured in this study (blue) with their IMG MAG ID and proposed *Candidatus* genus and species names. Reference Omnitrophota MAGs are shown with their Genome Taxonomy Database (GTDB) accession and GTDB taxonomy, except for *Candidatus* Omnitrophus fodinae and *Candidatus* Omnitrophus magneticus, which already have names. The tree was rooted with a basal representative of the "Planctomycetes-Verrucomicrobia-Chlamydiae" (PVC) superphylum.

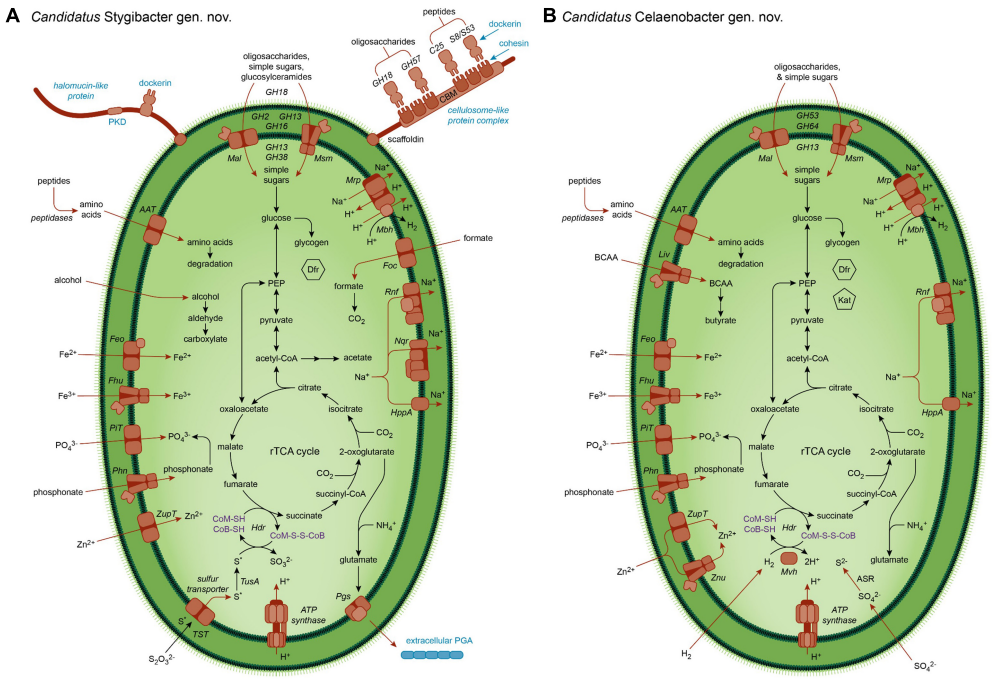

FIGURE 3 | Metabolic capacity of Cloacimonadota. Metabolic capacities inferred from MAGs for **(A)** *Candidatus* Stygibacter and **(B)** *Candidatus* Celaenobacter. AAT, amino acid transporter; ABC, ATP-binding cassette transport system; BCAA, branched-chain amino acids; CoB-SH, coenzyme B; CoM-SH, coenzyme M; CoM-S-S-CoB, heterodisulfide; Dfr, desulfoferrodoxin; Feo, ferrous iron transporter; Fhu iron ABC transporter; Foc, formate transporter; GH, glycoside hydrolase; Hdr, heterodisulfide reductase; HppA, pyrophosphate-energized sodium pump; Mal, maltooligosaccharide transport system (permease + solute-binding protein); Kat, catalase; Liv, branched-chain amino acid ABC transporter; Mbh, membrane-bound [NiFe] hydrogenase; Mrp, multicomponent Na^+:H^+ antiporter; Msm, multiple sugar ABC transporter; Mvh, cytoplasmic [NiFe] hydrogenase; Nqr, sodium-translocating NADH:quinone oxidoreductase; PEP, phosphoenolpyruvate; PGA, poly-γ-glutamate; Pgs, poly-γ-glutamate synthase; Phn, phosphonate ABC transporter; PiT, inorganic phosphate transporter; PKD, Polycystic Kidney Disease domain (implicated in adhesion); Rnf, ferredoxin:NAD^+-oxidoreductase complex; rTCA cycle, reverse tricarboxylic acid cycle; S*, unknown sulfur species; TST, thiosulfate:sulfurtransferase (rhodanese-like) (periplasmic); Znu, zinc ABC transporter; ZupT, zinc transporter. Note that the precise sulfur oxidation pathway for **(B)** is speculative; thiosulfate is shown here as the source of reductant, but the identity of the sulfur species is unclear.

concentrations of salt (Bolhuis et al., 2006). These cell envelope features combined with Na^+-translocating homeostasis abilities (see section "Cloacimonadota: Fermentation") may assist growth of Cloacimonadota at the bottom of the Ace Lake water column where salinity is at its highest (43 g L^{-1}) (Rankin et al., 1999; Lauro et al., 2011).

Cloacimonadota: Fermentation

A fermentative metabolism is inferred for the Ace Lake Cloacimonadota, by which energy is derived from the oxidation of organic substrates (sugars, amino acids, 2-oxoacids, aldehydes, alcohols, and formate), and protons are used as the electron acceptor (Calusinska et al., 2010; **Table 1** and **Figures 3A,B**). Simple sugars and amino acids, including those generated by extracellular degradation of biopolymers, appear to be principally imported by secondary transport, although a complete primary transporter for branched-chain amino acids (BCAAs) was identified in *Ca.* Celaenobacter MAGs. For simple sugars, the Embden–Meyerhof–Parnas (EMP) pathway for glycolysis generates NADH and reduced ferredoxin, as well as anabolic precursors (such as for the pentose phosphate pathway). There is no evidence in any of the 10 Ace Lake Cloacimonadota MAGs of a capacity for propionate oxidation, unlike members of *Ca.* Cloacimonadaceae (Pelletier et al., 2008; Dyksma and Gallert, 2019).

Ca. Tenebribacter, *Ca.* Stygibacter, and *Ca.* Zophobacter MAGs encode a membrane-bound [NiFe] hydrogenase (Mbh) (Group 4g) (Søndergaard et al., 2016). It has been proposed that Mbh transfers electrons from reduced ferredoxin to protons, thereby producing H_2 gas; this would generate a Na^+ gradient across the cell membrane via a Mrp-type Na^+/H^+ antiporter module (Mayer and Müller, 2014; Søndergaard et al., 2016; Yu et al., 2018). All four genera encode the Rnf complex, which couples electron transfer from reduced ferredoxin to NAD^+ to generate NADH, with concomitant translocation of Na^+ ions across the membrane (Biegel et al., 2011). The Na^+ gradient drives ATP synthesis via a Na^+-dependent F-type ATP synthase (Meier et al., 2009). This Na^+ gradient can also be used for other purposes, such as phosphate uptake via a Na^+/phosphate cotransporter (NptA-like). NADH can also be used for anabolic purposes. Additionally, *Ca.* Tenebribacter, *Ca.* Stygibacter, and *Ca.* Zophobacter encode a Na^+-translocating NADH:quinone oxidoreductase complex (NQR), which couples NADH re-oxidation to Na^+-extrusion, as well as maintaining ionic balance inside the cell (Verkhovsky and Bogachev, 2010).

All 10 Ace Lake Cloacimonadota MAGs lack genes for two essential enzymes of the oxidative tricarboxylic acid (TCA) cycle: citrate synthase and succinate dehydrogenase. Thus, we infer that they cannot operate a complete oxidative TCA cycle. All four genera encode phosphoenolpyruvate (PEP) carboxykinase, which converts oxaloacetate to PEP. The Ace Lake Cloacimonadota MAGs encode a pyrophosphate-dependent phosphofructokinase (PP_i-PFK) as well as the more widely distributed ATP-dependent 6-phosphofructokinase (ATP-PFK), the former of which can reversibly function in both glycolysis and gluconeogenesis (Mertens, 1991; Kemp and Tripathi, 1993). Because PP_i is a byproduct of biosynthetic reactions, the use of PPi-PFK rather than ATP-PFK increases the energetic efficiency of glycolysis, especially during fermentation (Mertens, 1991; Reshetnikov et al., 2008). The presence of the reversible enzyme PP_i-PFK is consistent with the absence of the gluconeogenesis-specific enzyme fructose-1,6-bisphosphatase from the Ace Lake Cloacimonadota MAGs. Having dual enzymes for the conversion of fructose-6-phosphate to fructose 1,6-bisphosphate might allow the Ace Lake Cloacimonadota to respond to the flux of high-energy phosphoryl donors in the cell (ATP vs. PP_i). Furthermore, PP_i may also be diverted directly to energy conservation using a PP_i-dependent Na^+ pump (HppA) that utilizes the energy of PP_i hydrolysis as the driving force for Na^+ translocation.

In addition to the abilities to ferment sugars, the Ace Lake Cloacimonadota MAGs encode multiple ferredoxin oxidoreductases that oxidize 2-oxoacids (including products of amino acid degradation), as inferred for *Ca.* Cloacimonas (Pelletier et al., 2008). These include pyruvate:ferredoxin oxidoreductase (POR); 2-oxoglutarate:ferredoxin oxidoreductase (OGOR); branched-chain 2-oxoacid (2-oxoisovalerate):ferredoxin oxidoreductase (VOR); indolepyruvate:ferredoxin oxidoreductase; and aldehyde:ferredoxin oxidoreductase (**Supplementary Tables S3, S4**). In addition to VOR, the Ace Lake Cloacimonadota encode phosphate butyryltransferase and butyrate kinase, suggesting the potential for further catabolism of branched-chain 2-oxoacids derived from degradation of BCAAs.

Ca. Stygibacter and *Ca.* Zophobacter MAGs encode phosphate acetyltransferase and acetate kinase for the conversion of acetyl-CoA to acetate *via* acetyl phosphate with concomitant production of ATP *via* substrate-level phosphorylation (Sapra et al., 2003), also inferred for Cloacimonadota MAG TCS47 (Zhang et al., 2016). In addition to Mbh, two reversible [FeFe] cytoplasmic hydrogenases were identified in certain Ace Lake Cloacimonadota: a tetrameric Group A3 hydrogenase (*Ca.* Stygibacter) and a monomeric Group C1 hydrogenase (*Ca.* Tenebribacter and *Ca.* Stygibacter) (Søndergaard et al., 2016). As a bidirectional hydrogenase, the Group A3 hydrogenase could use H_2 as an energy source through the bifurcation of electrons from H_2 to ferredoxin and NAD^+ (Poudel et al., 2016; Søndergaard et al., 2016; Kpebe et al., 2018), or it could serve as a confurcating hydrogenase to dissipate surplus reductant (from both NADH and reduced ferredoxin) that is generated during fermentation (Schut and Adams, 2009; Poudel et al., 2016). For the latter, substrate-level phosphorylation in the conversion of glucose to acetate would be facilitated by the dissipation of both reducing equivalents (NADH and reduced ferredoxin) as H_2 (Herrmann et al., 2008). These findings indicate that the Ace Lake Cloacimonadota would generate H_2 and acetate as byproducts of carbohydrate fermentation.

In both the *Ca.* Tenebribacter and *Ca.* Stygibacter MAGs, the Group C1 [FeFe] hydrogenase gene is immediately downstream of a gene for a histidine kinase domain protein, providing support for a putative sensory function (Greening et al., 2016). However, in one *Ca.* Tenebribacter MAG (3300035698_1346) the same gene cluster also encodes homologs of hydrogenase subunits associated with electron bifurcation (Poudel et al., 2016), which raises the possibility of a metabolic role for the Group C1 [FeFe] hydrogenase.

Cloacimonadota: Carbon Fixation Using a Reverse Tricarboxylic Acid Cycle

The gene inventories of certain Ace Lake Cloacimonadota suggest that they are capable of operating the reverse tricarboxylic acid (rTCA) cycle for carbon fixation, driven by sulfur oxidation (*Ca.* Tenebribacter and *Ca.* Stygibacter) (**Figure 3A**) or H_2 oxidation (*Ca.* Celaenobacter) (**Figure 3B**). The MAGs of these three genera encode ATP citrate lyase (ACL), thiol:fumarate reductase (TFR), and OGOR; these three enzymes allow the TCA cycle to proceed in the reductive direction (Rubin-Blum et al., 2019). PEP carboxykinase would connect the rTCA cycle to gluconeogenesis (Marietou et al., 2020) (see section "Cloacimonadota: Fermentation"). In the *Ca.* Celaenobacter MAGs, the genes for TFR and OGOR are part of a gene cluster that also includes fumarate hydratase, succinyl-CoA synthetase, [NiFe] hydrogenase (Mvh) (Group 3c), and heterodisulfide reductase (Hdr), which is consistent with a functional link between all these proteins. ACL is encoded elsewhere in the *Ca.* Celaenobacter genome, in a gene cluster that also includes the TCA cycle enzymes aconitase and isocitrate dehydrogenase. In general, the cytoplasmic Mvh hydrogenase forms a complex with Hdr, and bifurcates electrons from H_2 to heterodisulfide (CoM-S-S-CoB) and ferredoxin; the Mvh–Hdr complex couples the exergonic reduction of heterodisulfide with the endergonic reduction of ferredoxin with H_2 (Heim et al., 1998; Kaster et al., 2011; Greening et al., 2016). Thus, in *Ca.* Celaenobacter, carbon fixation using the rTCA cycle would be driven by H_2 oxidation (**Figure 3B**), as in *Aquifex aeolicus* (Brugna-Guiral et al., 2003; Guiral et al., 2005), although the latter have been inferred to use a Group 2d cytoplasmic hydrogenase for carbon fixation (Greening et al., 2016). ACL, OGOR, TFR, and Hdr genes were also identified in MAGs assigned to *Ca.* Tenebribacter and *Ca.* Stygibacter. Hdr genes in one *Ca.* Stygibacter MAG (3300025642_13) are within a gene cluster that also contains genes implicated in sulfur metabolism, including thiosulfate:sulfurtransferase (TST) (with a predicted signal peptide), a cytoplasmic sulfur relay protein TusA, and a sulfur compound transporter (Gristwood et al., 2011; Tanaka et al., 2020). Homologs of these four proteins are also encoded in MAGs assigned to *Ca.* Tenebribacter. Thus, energy required for carbon fixation in *Ca.* Tenebribacter and *Ca.* Stygibacter may be derived from sulfur oxidation catalyzed by TST and

Hdr, as in certain other autotrophic bacteria (Boughanemi et al., 2016; Koch and Dahl, 2018; Wang et al., 2019). Based on the presence of a TST homolog in these Ace Lake Cloacimonadota MAGs, the electron donor may be thiosulfate, with the initial reaction occurring in the periplasm (**Figure 3A**); however, elemental sulfur might also be utilized, as in *A. aeolicus* (Boughanemi et al., 2016).

Omnitrophota Ecophysiology

Based on interrogation of 14 Ace Lake Omnitrophota MAGs, this candidate phylum possesses a heterotrophic and fermentative metabolism. None of the MAGs possess any genes necessary for motility or magnetotaxis. We infer the Ace Lake Omnitrophota to be heterotrophs that are capable of fermenting a narrow range of substrates for energy conservation. All MAGs encode proteases and peptidases (some with signal peptides) to degrade proteins to amino acids. However, there are very few enzymes encoded for the catabolism of amino acids, and there is no evidence in any of the 14 MAGs of genes required for amino acid fermentation, unlike *Clostridium* spp. (Herrmann et al., 2008; Perret et al., 2011). Thus, we posit that these Ace Lake Omnitrophota use amino acids derived from peptide hydrolysis as nitrogen sources (especially by deamination or transamination) or for protein synthesis.

The Ace Lake Omnitrophota MAGs encode ABC transporter systems for sugars (disaccharides and/or oligosaccharides) and a number of GHs (including β-glucosidases and sugar phosphorolytic enzymes) to break down di- and oligosaccharides into simpler sugars such as glucose and/or glucose-1-phosphate (**Table 2** and **Supplementary Tables S5, S6**). Enzymes necessary for the initial depolymerization of polysaccharides are absent from all MAGs, which suggests that the Ace Lake Omnitrophota are dependent on other microorganisms for initial degradation of biopolymers, as inferred for the Black Sea Omnitrophota (Suominen et al., 2021). Simple sugars imported into the cell could also be utilized by the Ace Lake Omnitrophota for the synthesis of the compatible solute trehalose, or for the synthesis of glycogen for carbon and energy storage; enzymes for both processes were encoded across the Omnitrophota MAGs (**Table 2** and **Supplementary Table S5**).

The Ace Lake Omnitrophota MAGs also encode the capacity for fermentation of glucose to acetyl-CoA *via* the EMP pathway. The Rnf complex is predicted to couple the reduction of NAD^+ to the oxidation of reduced ferredoxin and translocation of protons across the cell membrane, allowing ATP generation via a V-type ATP synthase. The majority of MAGs also encode the enzymes phosphate acetyltransferase and acetate kinase for the subsequent conversion of acetyl-CoA to acetate with concomitant production of ATP via substrate-level phosphorylation (Sapra et al., 2003).

Excess reductant generated during fermentation could be dissipated as H_2 using a cytoplasmic hydrogenase (Dombrowski et al., 2017; Suominen et al., 2021). The Ace Lake Omnitrophota MAGs encode various hydrogenases, although it is noteworthy that each MAG has only one identifiable hydrogenase: Group 3d [NiFe] hydrogenase (*Ca.* Aceula *gen. nov*. and *Ca.* Zapsychrus *gen. nov.*); Group 3b [NiFe] hydrogenase (*Ca*. Tantalella *gen. nov*., *Ca.* Aadella *gen. nov*., and *Ca.* Orphnella *gen. nov.*); Group 4g [NiFe] hydrogenase (*Ca.* Saelkia *gen. nov*., and *Ca.* Gorgyraea *gen. nov.*), and Group A3 [FeFe] hydrogenase (*Ca.* Gygaella. *gen. nov*, *Ca.* Susulua *gen. nov*., *Ca.* Kaelpia *gen. nov*., and *Ca.* Kappaea *gen. nov.*). We infer that these hydrogenases are used for redox balance, associated with the need to dispose of surplus reductant. MAGs of both genera that encode Group 4g Mbh also encode Mrp (*Ca.* Saelkia and *Ca.* Gorgyraea) (**Figure 4A**); this hydrogenase may therefore function in a complex with the Mrp antiporter to generate an ionic gradient across the cell membrane, as in the other Mbh (Mayer and Müller, 2014; Søndergaard et al., 2016; Yu et al., 2018). The Group 3b hydrogenases of *Ca*. Tantalella, *Ca.* Aadella, and *Ca.* Orphnella were annotated as a bifunctional sulfhydrogenase (Shy) with dual hydrogenase and sulfur reductase activity (**Figure 4B**), meaning that excess reductant generated during fermentation can be disposed of as H_2 and sulfide, respectively (Ma et al., 1993; Silva et al., 1999; Ma et al., 2000).

None of the 14 Ace Lake Omnitrophota MAGs encode a complete TCA cycle (**Figures 4A,B**), either in the oxidative or reverse directions, with OGOR, succinyl-CoA synthetase, succinate dehydrogenase, ACL, and fumarate reductase absent from all MAGs (**Table 2**). We infer that the Ace Lake Omnitrophota possess an incomplete, "horse-shoe"-type TCA cycle as found in certain other anaerobic bacteria (e.g., Herlemann et al., 2009; Marco-Urrea et al., 2011). The type of citrate synthase varies, with either (but never both) citrate (*Si*)-synthase or citrate (*Re*)-synthase (Li et al., 2007; Marco-Urrea et al., 2011) encoded in individual MAGs, with the distribution of the functional analogs mostly conforming to separate Omnitrophota clades (**Supplementary Table S6**). The right branch of the incomplete TCA pathway of Omnitrophota is inferred to occur in the oxidative direction and commence at citrate synthase and terminate at 2-oxoglutarate. The left branch allows the interconversion of oxaloacetate, malate, and fumarate (Herlemann et al., 2009). This could proceed in the oxidative direction, with fumarate (such as generated as a byproduct of arginine synthesis) converted to oxaloacetate and used for gluconeogenesis (van Vugt-Lussenburg et al., 2009). Alternatively, this left branch may operate in the reductive direction, and be initiated by PEP carboxykinase (Herlemann et al., 2009); the subsequent reduction of oxaloacetate to fumarate would provide redox balance to the oxidative branch (Meléndez-Hevia et al., 1996). The "horseshoe-type" TCA cycle has no energy conservation function but serves solely for biosynthesis (Wood et al., 2004). The carbon skeleton 2-oxoglutarate is required for ammonia assimilation, and all 14 MAGs encode enzymes for this process (**Supplementary Tables S5, S6**). However, the fate of fumarate in Omnitrophota is unclear; there is no identifiable fumarate reductase (for anaerobic respiration), fumarate-adding enzymes (for hydrocarbon degradation), or aspartase (for synthesis of aspartate directly from fumarate) in any of the 14 Omnitrophota MAGs.

Although autotrophic pathways have been inferred in other Omnitrophota (Rinke et al., 2013; Kolinko et al., 2016) (see section "Cloacimonadota and Omnitrophota in the Ace Lake Ecosystem"), the Ace Lake Omnitrophota appear to be obligate heterotrophs. *Ca.* Gorgyraea encodes WL pathway genes

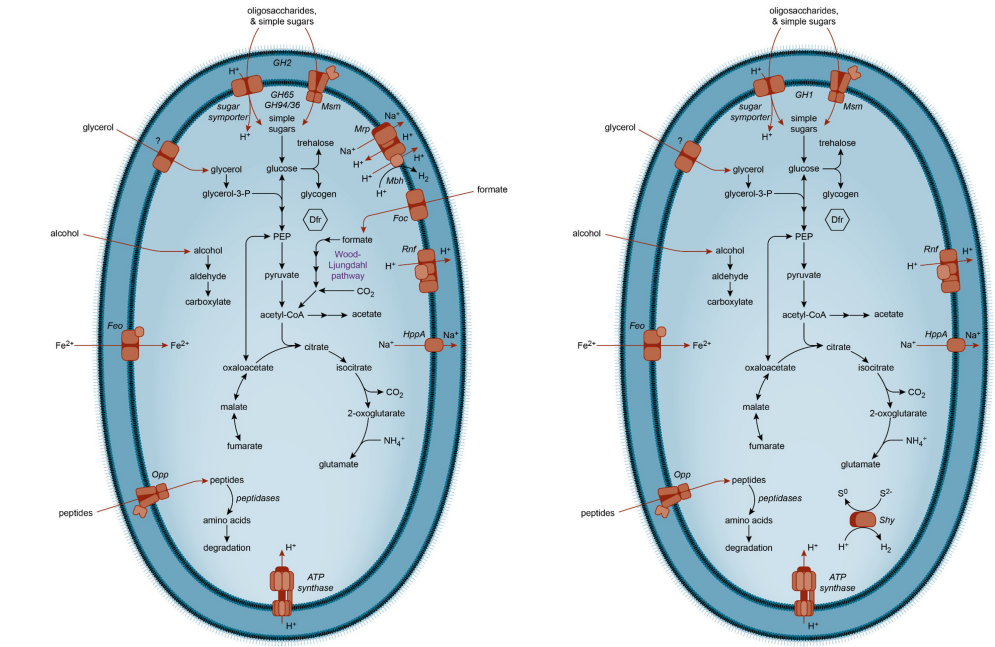

FIGURE 4 | Metabolic capacity of Omnitrophota. Metabolic capacities inferred from MAGs for **(A)** *Candidatus* Gorgyraea and **(B)** *Candidatus* Tantalella. ABC, ATP-binding cassette transport system; Dfr, desulfoferrodoxin; Feo, ferrous iron transporter; Foc, formate transporter; GH, glycoside hydrolase; HppA, pyrophosphate-energized sodium pump; Mbh, membrane-bound [NiFe] hydrogenase; Mrp, multicomponent Na^+:H^+ antiporter; Msm, multiple sugar ABC transporter; Opp, oligopeptide ABC transporter; PEP, phosphoenolpyruvate; Rnf, ferredoxin:NAD^+-oxidoreductase complex; Shy, sulfhydrogenase.

(**Table 2**, **Figure 4A**, and **Supplementary Table S5**), but in the absence of rTCA cycle genes that link acetyl-CoA to central biosynthetic pathways (Youssef et al., 2019), we infer that the WL pathway does not function in autotrophic CO_2 fixation. Instead, we propose that the WL pathway, in combination with the Rnf complex, functions in the reductive direction as an electron sink during homoacetogenic glucose fermentation, to maintain redox balance (Schuchmann and Müller, 2016; Youssef et al., 2019).

Overall, we infer only minor differences in the metabolic strategies among the 14 Ace Lake Omnitrophota MAGs (e.g., ability to catabolize glycerol, alcohol, or certain sugars; mechanisms for redox balance) (**Table 2**). Thus, all 11 Omnitrophota genera described here appear to conform to the same metabolic template: fermentative heterotrophs capable of degradation of a narrow range of organic compounds (especially simple sugars), with a hydrogenase for redox balance. The biosynthetic potential of the Ace Lake Omnitrophota MAGs are impressive, with the genomic capacity to synthesize nucleosides, fatty acids, the majority of proteinogenic amino acids, and essential cofactors (**Supplementary Table S5**).

Cloacimonadota and Omnitrophota in the Ace Lake Ecosystem

Although the ecophysiology of Ace Lake Cloacimonadota broadly agree with the fermentative, heterotrophic generalists inferred for members of this phylum from the Black Sea (Suominen et al., 2021), we infer a number of traits in certain Ace Lake Cloacimonadota that have not been previously reported for this candidate phylum. These include the presence of an extracellular cellulosome-like structure for the binding and degradation of biopolymers, PGA synthesis, a halomucin-like protein, and a chemolithoautotrophic pathway for carbon fixation *via* the rTCA cycle, fueled by oxidation of H_2 or sulfur compounds. These abilities attest to the physical and metabolic diversity of the Cloacimonadota, and emphasize the potential importance of this group in cycling of carbon, hydrogen, and sulfur in Ace Lake.

Chlorobium, which is the dominant organism in Ace Lake, also employs the rTCA cycle; this anaerobic photoautotroph grows at the limits of the penetration of photosynthetically active radiation at the oxic–anoxic interface (Panwar et al., 2020; **Figure 5**). As a consequence, the abundance of *Chlorobium* in this lake system is dictated by the polar light cycle, with a peak relative abundance of 83% at the interface in summer, and a marked decline in winter (6%) to spring (1%) (Panwar et al., 2020). *Chlorobium* is not metabolically active in the perennially dark anoxic zone of Ace Lake, and sinks to the bottom as particulate matter (Rankin et al., 1999; Lauro et al., 2011; Panwar et al., 2020). By contrast, light-independent, facultative chemolithoautotrophs, including the Ace Lake Cloacimonadota, would not be directly impacted by the marked seasonal variation in light availability. Although Cloacimonadota were most abundant in the deepest part of the anoxic zone of Ace Lake, they were detected throughout the

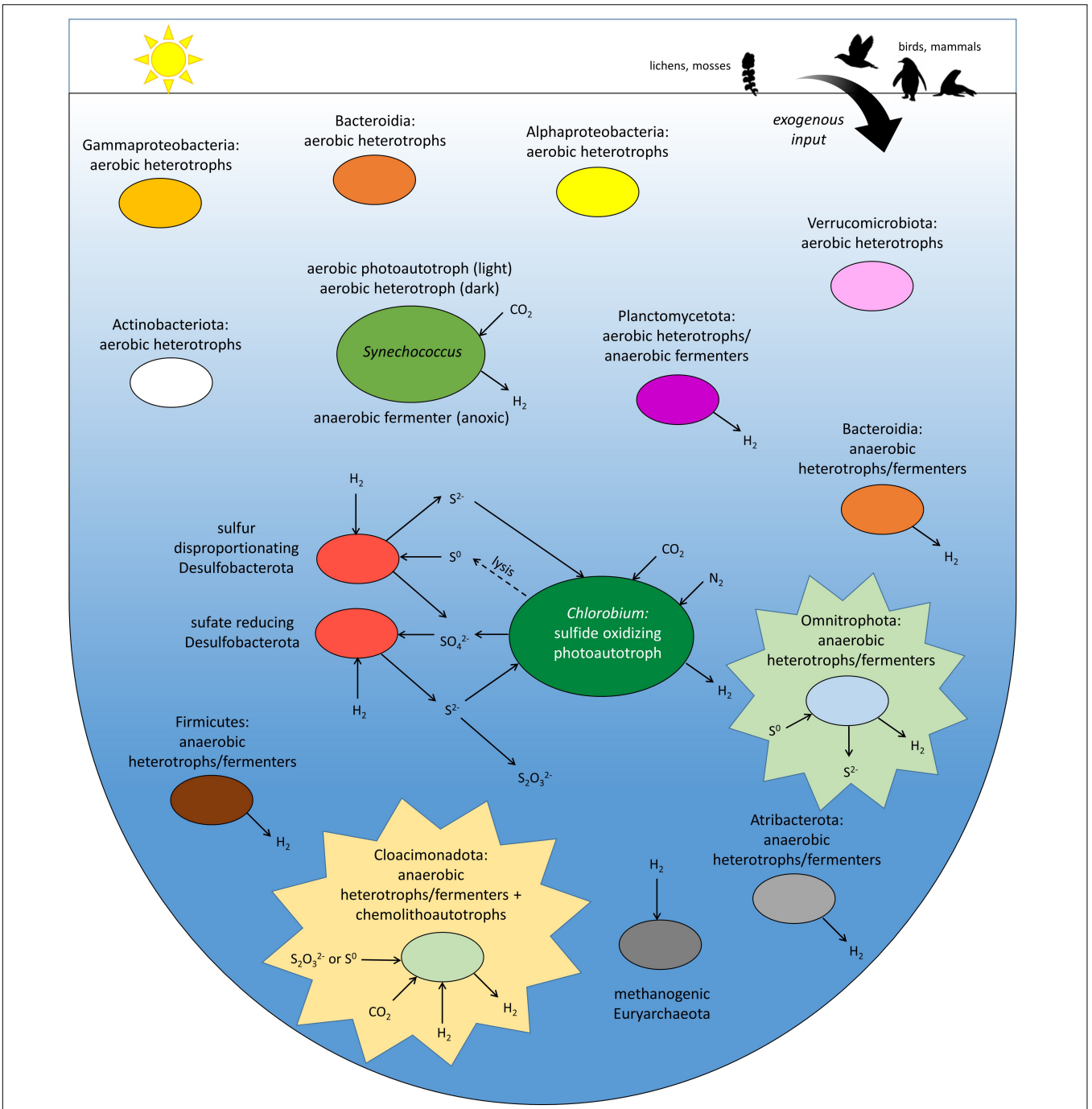

FIGURE 5 | Depiction of the most abundant bacteria and archaea in Ace Lake in the austral summer. Sulfur and hydrogen cycles are emphasized. Sizes of the ellipses that represent individual clades are approximately proportional to their peak relative abundance in the lake (Panwar et al., 2020), with *Chlorobium* (Chlorobia) the most abundant and *Synechococcus* (Cyanobacteria) the second most abundant. *Chlorobium* marks the oxic–anoxic interface of the lake. The candidate phyla Cloacimonadota and Omnitrophota are highlighted. Different taxa within candidate phylum Cloacimonadota are inferred to be capable of hydrogen and sulfur oxidation for chemolithoautotrophic growth; hydrogen is produced during heterotrophic growth. All taxa within the Omnitrophota candidate phylum are inferred to generate hydrogen, but only some are inferred to generate sulfide. Animal and plant silhouettes are courtesy of PhyloPic (http://phylopic.org/).

anoxic zone, as well as at the interface (Panwar et al., 2020). For those Cloacimonadota that we infer to use H_2 oxidation for carbon fixation (*Ca.* Celaenobacter), *Chlorobium* is potentially a major source of H_2, as a byproduct of nitrogen fixation by a membrane-bound nitrogenase (Lauro et al., 2011; Panwar et al., 2020). There are other bacteria present throughout the anoxic zone of Ace Lake that have been inferred to be H_2-evolving obligate anaerobes; aside from Omnitrophota, these

include members of Bacteroidia, Firmicutes, and Atribacterota (Panwar et al., 2020). The cyanobacterium *Synechococcus*, the most abundant phototrophic bacterium in the oxic zone of Ace Lake, was also abundant in the interface and anoxic zone, and previously inferred to be capable of fermentation coupled to H_2 production (Panwar et al., 2020); thus, *Synechococcus* is also a potential source of H_2 for H_2-oxidizing Cloacimonadota. For those Ace Lake Cloacimonadota that we infer to rely on sulfur oxidation for carbon fixation (*Ca.* Tenebribacter and *Ca.* Stygibacter), *Chlorobium* cells could be a source of elemental sulfur and polysulfide, as intermediates in sulfide oxidation (Marnocha et al., 2016), whereas thiosulfate is likely generated from the chemical reaction of sulfide (such as those generated by sulfate-reducing Desulfobacterota) with dissolved oxygen (Kondo et al., 2000). However, unlike sulfate-reducing and sulfur-disproportionating Desulfobacterota in Ace Lake, which were previously inferred to be metabolically linked to *Chlorobium* via sulfur cycling (Lauro et al., 2011; Panwar et al., 2020; **Figure 5**), no seasonal variation was observed for Cloacimonadota in Ace Lake. Although Cloacimonadota and Omnitrophota in Ace Lake are inferred to be obligate anaerobes, both are found in the oxic–anoxic interface; the MAGs encode mechanisms to help protect against oxidative stress (e.g., desulfoferrodoxin in both phyla, catalase in Cloacimonadota), allowing survival at minimal oxygen concentrations (Pelletier et al., 2008). Thus, we infer that these anaerobes are aerotolerant, as previously inferred for *Chlorobium* in Ace Lake (Ng et al., 2010).

During heterotrophic growth, the anaerobic degradation of hexoses to acetate, CO_2, and H_2 does not yield sufficient energy to support growth unless H_2 levels are sufficiently low (Thauer et al., 1977; Schink, 1997; Morris et al., 2013). Acetate and H_2 would be beneficial to hydrogenotrophic and acetoclastic methanogens (Euryarchaeota), which, like Cloacimonadota, are most abundant in the deepest waters of Ace Lake (Panwar et al., 2020; **Figure 5**); as such, we posit that Cloacimonadota and methanogens may be metabolically linked. The extracellular GHs and peptidases, including the extracellular cellulosome-like structure inferred for the Cloacimonadota genus *Ca.* Stygibacter, suggest that Cloacimonadota are among the "first responders" in deconstructing and assimilating recalcitrant particulate organic matter, including microbial aggregates that sink to the bottom from higher in the water column (Rankin et al., 1999; Lauro et al., 2011).

All the Ace Lake Omnitrophota MAGs appear to be obligate heterotrophs, with no evidence of autotrophic capacity. Furthermore, these anaerobes appear to be dependent on hydrogenotrophic microbes to consume H_2 released *via* anaerobic glucose fermentation. In Ace Lake, Omnitrophota were most abundant at the oxic–anoxic interface and the water column immediately below the interface, the depths at which sulfate-reducing and sulfur-disproportionating Desulfobacterota were also most abundant (Panwar et al., 2020; **Figure 5**). These Desulfobacterota encode H_2-uptake hydrogenases (Panwar et al., 2020), and would therefore be capable of consuming H_2.

The ability to infer the ecophysiology of the members of Cloacimonadota and Omnitrophota emphasizes the importance of having metagenome data and accompanying analyses detailing the metabolisms of numerous other members of the microbial community (Ng et al., 2010; Lauro et al., 2011; Panwar et al., 2020). Here, we added to the understanding of the system by analyzing a total of 24 MAGs for these "microbial dark matter" lineages. Both of these candidate phyla are predicted to engage in metabolic associations with other Ace Lake microorganisms. The specific involvement of hydrogen is noteworthy in view of hydrogen cycling being previously identified as pivotal to multiple nutrient cycles in Ace Lake (Panwar et al., 2020), as well as being increasingly recognized as important to ecosystem function in global anoxic and oxic environments (Greening et al., 2016).

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online public repositories. The repository and accession numbers are: IMG (https://img.jgi.doe.gov/) (Metagenomes 3300035698, 3300025586, 3300025698, and 3300025642; MAGs 3300035698_1174, 3300035698_1346, 3300035698_1468, 3300035698_198, 3300035698_2003, 3300025642_13, 3300035698_360, 3300035698_1703, 3300035698_1683, 3300025698_8, 3300035698_1100, 3300035698_985, 3300035698_1555, 3300035698_1934, 3300035698_1005, 3300035698_749, 3300035698_2000, 3300035698_1655, 3300035698_1500, 3300035698_1097, 3300035698_848, 3300035698_32, 3300025586_21, and 3300035698_104).

AUTHOR CONTRIBUTIONS

TW, MA, and RC conceived the study, analyzed the data, and/or conducted the data interpretation. JB provided illustrations. TW, MA, and RC wrote the manuscript with input from all authors. All authors have read and approved the manuscript submission.

FUNDING

This work was supported by the Australian Research Council (DP150100244) and the Australian Antarctic Science program (project 4031).

ACKNOWLEDGMENTS

Computational analyses at UNSW Sydney were performed on the computational cluster Katana, supported by Research Technology Services at UNSW Sydney.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.741077/full#supplementary-material

REFERENCES

Allen, M. A., Lauro, F. M., Williams, T. J., Burg, D., Siddiqui, K. S., DeFrancisci, D., et al. (2009). The genome sequence of the psychrophilic archaeon, *Methanococcoides burtonii*: the role of genome evolution in cold adaptation. ISME J. 3, 1012–1035. doi: 10.1038/ismej.2009.45

Andrei, A.-S., Salcher, M. M., Mehrshad, M., Rychtecký, P., Znachor, P., and Ghai, R. (2019). Niche-directed evolution modulates genome architecture in freshwater Planctomycetes. *ISME J.* 13, 1056–1071. doi: 10.1038/s41396-018-0332-5

Artzi, L., Bayer, E. A., and Moraïs, S. (2017). Cellulosomes: bacterial nanomachines for dismantling plant polysaccharides. *Nat. Rev. Microbiol.* 15, 83–95. doi: 10.1038/nrmicro.2016.164

Baricz, A., Chiriac, C. M., Andrei, A.-S., Bulzu, P.-A., Levei, E. A., Cadar, O., et al. (2020). Spatio-temporal insights into microbiology of the freshwater-to-hypersaline, oxic-hypoxic-euxinic waters of Ursu Lake. *Environ. Microbiol.* 23, 3523–3540. doi: 10.1111/1462-2920.14909

Biegel, E., Schmidt, S., González, J. M., and Müller, V. (2011). Biochemistry, evolution and physiological function of the Rnf complex, a novel ion-motive electron transport complex in prokaryotes. *Cell. Mol. Life Sci.* 68, 613–634. doi: 10.1007/s00018-010-0555-8

Blum, M., Chang, H. Y., Chuguransky, S., Grego, T., Kandasaamy, S., Mitchell, A., et al. (2020). The InterPro protein families and domains database: 20 years on. *Nucleic Acids Res.* 49, D344–D354. doi: 10.1093/nar/gkaa977

Bolhuis, H., Palm, P., Wende, A., Falb, M., Rampp, M., Rodriguez-Valera, F., et al. (2006). The genome of the square archaeon *Haloquadratum walsbyi* : life at the limits of water activity. *BMC Genomics* 7:169. doi: 10.1186/1471-2164-7-169

Boughanemi, S., Lyonnet, J., Infossi, P., Bauzan, M., Kosta, A., Lignon, S., et al. (2016). Microbial oxidative sulfur metabolism: biochemical evidence of the membrane-bound heterodisulfide reductase-like complex of the bacterium *Aquifex aeolicus*. *FEMS Microbiol. Lett.* 363:fnw156. doi: 10.1093/femsle/fnw156

Brugna-Guiral, M., Tron, P., Nitschke, W., Stetter, K.-O., Burlat, B., Guigliarelli, B., et al. (2003). [NiFe] hydrogenases from the hyperthermophilic bacterium *Aquifex aeolicus*: properties, function, and phylogenetics. *Extremophiles* 7, 147–157. doi: 10.1007/s00792-002-0306-3

Calusinska, M., Happe, T., Joris, B., and Wilmotte, A. (2010). The surprising diversity of clostridial hydrogenases: a comparative genomic perspective. *Microbiology* 156, 1575–1588. doi: 10.1099/mic.0.032771-0

Cavicchioli, R. (2015). Microbial ecology of Antarctic aquatic systems. *Nat. Rev. Microbiol.* 13, 691–706. doi: 10.1038/nrmicro3549

Chaumeil, P.-A., Mussig, A. J., Hugenholtz, P., and Parks, D. H. (2019). GTDB-Tk: a toolkit to classify genomes with the genome taxonomy database. *Bioinformatics* 36, 1925–1927. doi: 10.1093/bioinformatics/btz848

Chouari, R., Le Paslier, D., Daegelen, P., Ginestet, P., Weissenbach, J., and Sghir, A. (2005a). Novel predominant archaeal and bacterial groups revealed by molecular analysis of an anaerobic sludge digester. *Environ. Microbiol.* 7, 1104–1115. doi: 10.1111/j.1462-2920.2005.00795.x

Chouari, R., Le Paslier, D., Dauga, C., Daegelen, P., Weissenbach, J., and Sghir, A. (2005b). Novel major bacterial candidate division within a municipal anaerobic sludge digester. *Appl. Environ. Microbiol.* 71, 2145–2153. doi: 10.1128/AEM.71.4.2145-2153.2005

Chuvochina, M., Rinke, C., Parks, D. H., Rappé, M. S., Tyson, G. W., Yilmaz, P., et al. (2019). The importance of designating type material for uncultured taxa. *Syst. Appl. Microbiol.* 42, 15–21. doi: 10.1016/j.syapm.2018.07.003

Derakshani, M., Lukow, T., and Liesack, W. (2001). Novel bacterial lineages at the (sub)division level as detected by signature nucleotide-targeted recovery of 16S rRNA genes from bulk soil and rice roots of flooded rice microcosms. *Appl. Environ. Microbiol.* 67, 623–631. doi: 10.1128/AEM.67.2.623-631.2001

Dombrowski, N., Seitz, K. W., Teske, A. P., and Baker, B. J. (2017). Genomic insights into potential interdependencies in microbial hydrocarbon and nutrient cycling in hydrothermal sediments. *Microbiome* 5:106. doi: 10.1186/s40168-017-0322-2

Dyksma, S., and Gallert, C. (2019). *Candidatus* Syntrophosphaera thermopropionivorans: a novel player in syntrophic propionate oxidation during anaerobic digestion. *Environ. Microbiol. Rep.* 11, 558–570.

Eddy, S. R. (2011). Accelerated profile HMM searches. *PLoS Comp. Biol.* 7:e1002195. doi: 10.1371/journal.pcbi.1002195

Fontes, C. M. G. A., and Gilbert, H. J. (2010). Cellulosomes: highly efficient nanomachines designed to deconstruct plant cell wall complex carbohydrates. *Annu. Rev. Biochem.* 79, 655–681. doi: 10.1146/annurev-biochem-091208-085603

Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D., and Bairoch, A. (2003). ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* 31, 3784–3788. doi: 10.1093/nar/gkg563

Glöckner, J., Kube, M., Shrestha, P. M., Weber, M., Glöckner, F. O., Reinhardt, R., et al. (2010). Phylogenetic diversity and metagenomics of candidate division OP3. *Environ. Microbiol.* 12, 1218–1229.

Greening, C., Biswas, A., Carere, C. R., Jackson, C. J., Taylor, M. C., Stott, M. B., et al. (2016). Genomic and metagenomic surveys of hydrogenase distribution indicate H2 is a widely utilised energy source for microbial growth and survival. *ISME J.* 10, 761–777. doi: 10.1038/ismej.2015.153

Gristwood, T., McNeil, M. B., Clulow, J. S., Salmond, G. P., and Fineran, P. C. (2011). PigS and PigP regulate prodigiosin biosynthesis in *Serratia* via differential control of divergent operons, which include predicted transporters of sulfur-containing molecules. *J. Bacteriol.* 193, 1076–1085. doi: 10.1128/JB.00352-10

Guiral, M., Aubert, C., and Giudici-Orticoni, M. T. (2005). Hydrogen metabolism in the hyperthermophilic bacterium *Aquifex aeolicus*. *Biochem. Soc. Trans.* 33, 22–24. doi: 10.1042/BST0330022

Heim, S., Kunkel, A., Thauer, R. K., and Hedderich, R. (1998). Thiol:fumarate reductase (Tfr) from *Methanobacterium thermoautotrophicum* - identification of the catalytic sites for fumarate reduction and thiol oxidation. *Eur. J. Biochem.* 253, 292–299. doi: 10.1046/j.1432-1327.1998.2530292.x

Herlemann, D. P. R., Geissinger, O., Ikeda-Ohtsubo, W., Kunin, V., Sun, H., Lapidus, A., et al. (2009). Genomic analysis of "*Elusimicrobium minutum*," the first cultivated representative of the phylum "Elusimicrobia" (formerly Termite Group 1). *Appl. Environ. Microbiol.* 75, 2841–2849. doi: 10.1128/AEM.02698-08

Herrmann, G., Jayamani, E., Mai, G., and Buckel, W. (2008). Energy conservation via electron-transferring flavoprotein in anaerobic bacteria. *J. Bacteriol.* 190, 784–791. doi: 10.1128/JB.01422-07

Hugenholtz, P., Pitulle, C., Hershberger, K. L., and Pace, N. R. (1998). Novel division level bacterial diversity in a Yellowstone hot spring. *J. Bacteriol.* 180, 366–376. doi: 10.1128/JB.180.2.366-376.1998

Huntemann, M., Ivanova, N. N., Mavromatis, K., Tripp, H. J., Paez-Espino, D., Tennessen, K., et al. (2015). The standard operating procedure of the DOE-JGI metagenome annotation pipeline (MAP v.4). *Stand. Genomic Sci.* 11:17. doi: 10.1186/s40793-016-0138-x

Huson, D., and Scornavacca, C. (2012). Dendroscope 3: an interactive tool for rooted phylogenetic trees and networks. *Syst. Biol.* 61, 1061–1067. doi: 10.1093/sysbio/sys062

Hyatt, D., Chen, G.-L., LoCascio, P. F., Land, M. L., Larimer, F. W., and Hauser, L. J. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11:119. doi: 10.1186/1471-2105-11-119

Jain, C., Rodriguez-R, L. M., Phillippy, A. M., Konstantinidis, K. T., and Aluru, S. (2019). High-throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat. Commun.* 9:5114. doi: 10.1038/s41467-018-07641-9

Kandler, O., König, H., Wiegel, J., and Claus, D. (1983). Occurrence of poly-γ-D-glutamic acid and poly-α-L-glutamine in the genera *Xanthobacter*, *Flexithrix*, *Sporosarcina* and *Planococcus*. *Syst. Appl. Microbiol.* 4, 34–41. doi: 10.1016/S0723-2020(83)80032-0

Kang, D. D., Li, F., Kirton, E., Thomas, A., Egan, R., An, H., et al. (2019). MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ* 7:e7359. doi: 10.7717/peerj.7359

Kaster, A. K., Moll, J., Parey, K., and Thauer, R. K. (2011). Coupling of ferredoxin and heterodisulfide reduction via electron bifurcation in hydrogenotrophic methanogenic archaea. *Proc. Natl. Acad. Sci. U.S.A.* 108, 2981–2986.

Kemp, R. G., and Tripathi, R. L. (1993). Pyrophosphate-dependent phosphofructo-1-kinase complements fructose 1,6-bisphosphatase but not phosphofructokinase deficiency in *Escherichia coli*. *J. Bacteriol.* 175, 5723–5724. doi: 10.1128/jb.175.17.5723-5724.1993

Koch, T., and Dahl, C. (2018). A novel bacterial sulfur oxidation pathway provides a new link between the cycles of organic and inorganic sulfur compounds. *ISME J.* 12, 2479–2491. doi: 10.1038/s41396-018-0209-7

Kolinko, S., Jogler, C., Katzmann, E., Wanner, G., Peplies, J., and Schüler, D. (2012). Single-cell analysis reveals a novel uncultivated magnetotactic bacterium within the candidate division OP3. *Environ. Microbiol.* 14, 1709–1721. doi: 10.1111/j.1462-2920.2011.02609.x

Kolinko, S., Richter, M., Glöckner, F. O., Brachmann, A., and Schüler, D. (2016). Single-cell genomics of uncultivated deep-branching magnetotactic bacteria reveals a conserved set of magnetosome genes. *Environ. Microbiol.* 18, 21–37. doi: 10.1111/1462-2920.12907

Kondo, R., Kasashima, N., Matsuda, H., and Hata, Y. (2000). Determination of thiosulfate in a meromictic lake. *Fish. Sci.* 66, 1076–1081. doi: 10.1046/j.1444-2906.2000.00171.x

Konstantinidis, K. T., Rosselló-Móra, R., and Amann, R. (2017). Uncultivated microbes in need of their own taxonomy. *ISME J.* 11, 2399–2406. doi: 10.1038/ismej.2017.113

Kpebe, A., Benvenuti, M., Guendon, C., Rebai, A., Fernandez, V., Le Laz, S., et al. (2018). A new mechanistic model for an O2-protected electron-bifurcating hydrogenase, Hnd from *Desulfovibrio fructosovorans*. *Biochim. Biophys. Acta Bioenerget.* 1859, 1302–1312. doi: 10.1016/j.bbabio.2018.09.364

Lamed, R., Setter, E., and Bayer, E. A. (1983). Characterization of a cellulose-binding, cellulase-containing complex in *Clostridium thermocellum*. *J. Bacteriol.* 156, 828–836. doi: 10.1128/jb.156.2.828-836.1983

Lasica, A. M., Ksiazek, M., Madej, M., and Potempa, J. (2017). The type IX secretion system (T9SS): highlights and recent insights into its structure and function. *Front. Cell. Infect. Microbiol.* 7:215. doi: 10.3389/fcimb.2017.00215

Lauro, F. M., Demaere, M. Z., Yau, S., Brown, M. V., Wilkins, D., Ng, S. C., et al. (2011). An integrative study of a meromictic lake ecosystem in Antarctica. *ISME J.* 5, 879–895. doi: 10.1038/ismej.2010.185

Li, D., Luo, R., Liu, C. M., Leung, C. M., Ting, H. F., Sadakane, K., et al. (2016). MEGAHIT v1.0: a fast and scalable metagenome assembler driven by advanced methodologies and community practices. *Methods* 102, 3–11. doi: 10.1016/j.ymeth.2016.02.020

Li, F., Hagemeier, C. H., Seedorf, H., Gottschalk, G., and Thauer, R. K. (2007). Re-citrate synthase from *Clostridium kluyveri* is phylogenetically related to homocitrate synthase and isopropylmalate synthase rather than to Si-citrate synthase. *J. Bacteriol.* 189, 4299–4304. doi: 10.1128/JB.00198-07

Limam, R. D., Chouari, R., Mazéas, L., Wu, T.-D., Li, T., Grossin-Debattista, J., et al. (2014). Members of the uncultured bacterial candidate division WWE1 are implicated in anaerobic digestion of cellulose. *Microbiol. Open* 3, 157–167. doi: 10.1002/mbo3.144

Lin, W., Zhang, W. S., Paterson, G. A., Zhu, Q. Y., Zhao, X., Knight, R., et al. (2020). Expanding magnetic organelle biogenesis in the domain Bacteria. *Microbiome* 8:152. doi: 10.1186/s40168-020-00931-9

Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M., and Henrissat, B. (2014). The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* 42, D490–D495. doi: 10.1093/nar/gkt1178

Ma, K., Schicho, R. N., Kelly, R. M., and Adams, M. W. W. (1993). Hydrogenase of the hyperthermophilic *Pyrococcus furiosus* is an elemental sulfur reductase or sulfhydrogenase: evidence for a sulfur-reducing hydrogenase ancestor. *Proc. Natl. Acad. Sci. U.S.A.* 90, 5341–5344. doi: 10.1073/pnas.90.11.5341

Ma, K., Weiss, R., and Adams, M. W. W. (2000). Characterization of hydrogenase II from the hyperthermophilic archaeon *Pyrococcus furiosus* and assessment of its role in sulfur reduction. *J. Bacteriol.* 182, 1864–1871. doi: 10.1128/JB.182.7.1864-1871.2000

Marco-Urrea, E., Paul, S., Khodaverdi, V., Seifert, J., von Bergen, M., Kretzschmar, U., et al. (2011). Identification and characterization of a Re-citrate synthase in *Dehalococcoides* strain CBDB1. *J. Bacteriol.* 193, 5171–5178. doi: 10.1128/JB.05120-11

Marietou, A., Lund, M. B., Marshall, I. P. G., Schreiber, L., and Jørgensen, B. B. (2020). Complete genome sequence of *Desulfobacter hydrogenophilus* AcRS1. *Mar. Genomics* 50:100691. doi: 10.1016/j.margen.2019.05.006

Marnocha, C. L., Levy, A. T., Powell, D. H., Hanson, T. E., and Chan, C. S. (2016). Mechanisms of extracellular S0 globule production and degradation in *Chlorobaculum tepidum* via dynamic cell–globule interactions. *Microbiology* 162, 1125–1134. doi: 10.1099/mic.0.000294

Matsen, F. A., Kodner, R. B., and Armbrust, E. V. (2010). pplacer: linear time maximum-likelihood and Bayesian phylogenetic placement of sequences onto a fixed reference tree. *BMC Bioinformatics* 11:538. doi: 10.1186/1471-2105-11-538

Mayer, F., and Müller, V. (2014). Adaptations of anaerobic archaea to life under extreme energy limitation. *FEMS Microbiol. Rev.* 38, 449–472. doi: 10.1111/1574-6976.12043

Meier, T., Krah, A., Bond, P., Pogoryelov, D., Diederichs, K., and Faraldo-Gómez, J. (2009). Complete ion-coordination structure in the rotor ring of Na+-dependent F-ATP synthases. *J. Mol. Biol.* 391, 498–507. doi: 10.1016/j.jmb.2009.05.082

Meléndez-Hevia, E., Waddell, T. G., and Cascante, M. (1996). The puzzle of the Krebs citric acid cycle: assembling the pieces of chemically feasible reactions, and opportunism in the design of metabolic pathways during evolution. *J. Mol. Evol.* 43, 293–303. doi: 10.1007/BF02338838

Mertens, E. (1991). Pyrophosphate-dependent phosphofructokinase, an anaerobic glycolytic enzyme? *FEBS Lett.* 285, 1–5. doi: 10.1016/0014-5793(91)80711-B

Minh, B. Q., Nguyen, M. A. T., and von Haeseler, A. (2013). Ultrafast approximation for phylogenetic bootstrap. *Mol. Biol. Evol.* 30, 1188–1195. doi: 10.1093/molbev/mst024

Momper, L., Jungbluth, S. P., Lee, M. D., and Amend, J. P. (2017). Energy and carbon metabolisms in a deep terrestrial subsurface fluid microbial community. *ISME J.* 11, 2319–2333. doi: 10.1038/ismej.2017.94

Morris, B. E. L., Henneberger, R., Huber, H., and Moissl-Eichinger, C. (2013). Microbial syntrophy: interaction for the common good. *FEMS Microbiol. Rev.* 37, 384–406. doi: 10.1111/1574-6976.12019

Murray, A. E., Freudenstein, J., Gribaldo, S., Hatzenpichler, R., Philip Hugenholtz, P., Kämpfer, P., et al. (2020). Roadmap for naming uncultivated Archaea and Bacteria. *Nat. Microbiol.* 5, 987–994. doi: 10.1038/s41564-020-0733-x

Nayfach, S., Roux, S., Seshadri, R., Udwary, D., Varghese, N., Schulz, F., et al. (2020). A genomic catalog of Earth's microbiomes. *Nat. Biotechnol.* 39, 499–509. doi: 10.1038/s41587-020-0718-6

Ng, S. C., De Maere, M., Williams, T. J., Lauro, F. M., Raftery, M. J., Gibson, J., et al. (2010). Metaproteogenomic analysis of a dominant green sulfur bacterium from Ace Lake, Antarctica. *ISME J.* 4, 1002–1019. doi: 10.1038/ismej.2010.28

Nguyen, L.-T., Schmidt, H. A., von Haeseler, A., and Minh, B. Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum likelihood phylogenies. *Mol. Biol. Evol.* 32, 268–274. doi: 10.1093/molbev/msu300

Nobu, M. K., Narihiro, T., Rinke, C., Kamagata, Y., Tringe, S. G., Woyke, T., et al. (2015). Microbial dark matter ecogenomics reveals complex synergistic networks in a methanogenic bioreactor. *ISME J.* 9, 1710–1722. doi: 10.1038/ismej.2014.256

Ondov, B. D., Treangen, T. J., Melsted, P., Mallonee, A. B., Bergman, N. H., Koren, S., et al. (2016). Mash: fast genome and metagenome distance estimation using MinHash. *Genome Biol.* 17:132. doi: 10.1186/s13059-016-0997-x

Panwar, P., Allen, M. A., Williams, T. J., Hancock, A. M., Brazendale, S., Bevington, J., et al. (2020). Influence of the polar light cycle on seasonal dynamics of an Antarctic lake microbial community. *Microbiome* 8, 1–24. doi: 10.1186/s40168-020-00889-8

Parks, D. H., Chuvochina, M., Chaumeil, P.-A., Rinke, C., Mussig, A. J., and Hugenholtz, P. (2020). A complete domain-to-species taxonomy for Bacteria and Archaea. *Nat. Biotechnol.* 38, 1079–1086. doi: 10.1038/s41587-020-0501-8

Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., and Tyson, G. W. (2015). CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* 25, 1043–1055. doi: 10.1101/gr.186072.114

Parks, D. H., Rinke, C., Chuvochina, M., Chaumeil, P.-A., Woodcroft, B. J., Evans, P. N., et al. (2017). Recovery of nearly 8,000 metagenome-assembled genomes substantially expands the tree of life. *Nat. Microbiol.* 2, 1533–1542. doi: 10.1038/s41564-017-0012-7

Peer, A., Smith, S. P., Bayer, E. A., Lamed, R., and Borovok, I. (2009). Noncellulosomal cohesin- and dockerin-like modules in the three domains of life. *FEMS Microbiol. Lett.* 291, 1–16. doi: 10.1111/j.1574-6968.2008.01420.x

Pelletier, E., Kreimeyer, A., Bocs, S., Rouy, Z., Gyapay, G., Chouari, R., et al. (2008). "*Candidatus* Cloacamonas acidaminovorans": genome sequence reconstruction provides a first glimpse of a new bacterial division. *J. Bacteriol.* 190, 2572–2579. doi: 10.1128/JB.01248-07

Perret, A., Lechaplais, C., Tricot, S., Perchat, N., Vergne, C., Pelle, C., et al. (2011). A novel acyl-CoA beta-transaminase characterized from a metagenome. *PLoS One* 6:e22918. doi: 10.1371/journal.pone.0022918

Pilhofer, M., Rappl, K., Eckl, C., Bauer, A. P., Ludwig, W., Schleifer, K. H., et al. (2008). Characterization and evolution of cell division and cell wall synthesis

genes in the bacterial phyla *Verrucomicrobia*, *Lentisphaerae*, *Chlamydiae*, and *Planctomycetes* and phylogenetic comparison with rRNA genes. *J. Bacteriol.* 190, 3192–3202. doi: 10.1128/JB.01797-07

Poudel, S., Tokmina-Lukaszewska, M., Colman, D. R., Refai, R., Schut, H. J., King, P. W., et al. (2016). Unification of [FeFe]-hydrogenases into three structural and functional groups. *Biochim. Biophys. Acta* 1860, 1910–1921. doi: 10.1016/j.bbagen.2016.05.034

Price, M. N., Dehal, P. S., and Arkin, A. P. (2010). FastTree 2 - approximately maximum-likelihood trees for large alignments. *PLoS One* 5:e9490. doi: 10.1371/journal.pone.0009490

Rankin, L. M., Gibson, J. A. E., Franzmann, P. D., and Burton, H. R. (1999). The chemical stratification and microbial communities of Ace Lake, Antarctica: a review of the characteristics of a marine-derived meromictic lake. *Polarforschung* 66, 33–52.

Rehm, B. H. A. (2010). Bacterial polymers: biosynthesis, modifications and applications. *Nat. Rev. Microbiol.* 8, 578–592. doi: 10.1038/nrmicro2354

Reshetnikov, A. S., Rozova, O. N., Khmelenina, V. N., Mustakhimov, I. I., Beschastny, A. P., Murrell, J. C., et al. (2008). Characterization of the pyrophosphate-dependent 6-phosphofructokinase from *Methylococcus capsulatus* Bath. *FEMS Microbiol. Lett.* 288, 202–210. doi: 10.1111/j.1574-6968.2008.01366.x

Rinke, C., Schwientek, P., Sczyrba, A., Ivanova, N. N., Anderson, I. J., Cheng, J. F., et al. (2013). Insights into the phylogeny and coding potential of microbial dark matter. *Nature* 499, 431–437. doi: 10.1038/nature12352

Rubin-Blum, M., Dubilier, N., and Kleiner, M. (2019). Genetic evidence for two carbon fixation pathways (the Calvin-Benson-Bassham cycle and the reverse tricarboxylic acid cycle) in symbiotic and free-living bacteria. *mSphere* 4:e00394-18. doi: 10.1128/mSphere.00394-18

Saha, S., Jeon, B. H., Kurade, M. B., Chatterjee, P. K., Chang, S. W., Markkandan, K., et al. (2019). Microbial acclimatization to lipidic-waste facilitates the efficacy of acidogenic fermentation. *Chem. Eng. J.* 358, 188–196. doi: 10.1016/j.cej.2018.09.220

Santos, A., Rachid, C., Pacheco, A. B., and Magalhães, V. (2020). Biotic and abiotic factors affect microcystin-LR concentrations in water/sediment interface. *Microbiol. Res.* 236:126452. doi: 10.1016/j.micres.2020.126452

Sapra, R., Bagramyan, K., and Adams, M. W. W. (2003). A simple energy-conserving system: proton reduction coupled to proton translocation. *Proc. Natl. Acad. Sci. U.S.A.* 100, 7545–7550. doi: 10.1073/pnas.1331436100

Schink, B. (1997). Energetics of syntrophic cooperation in methanogenic degradation. *Microbiol. Mol. Biol. Rev.* 61, 262–280. doi: 10.1128/.61.2.262-280.1997

Schuchmann, K., and Müller, V. (2016). Energetics and application of heterotrophy in acetogenic bacteria. *Appl. Environ. Microbiol.* 82, 4056–4069. doi: 10.1128/AEM.00882-16

Schut, G. J., and Adams, M. W. (2009). The iron-hydrogenase of *Thermotoga maritima* utilizes ferredoxin and NADH synergistically: a new perspective on anaerobic hydrogen production. *J. Bacteriol.* 191, 4451–4457. doi: 10.1128/JB.01582-08

Schwarz, W. H., and Zverlov, V. V. (2006). Protease inhibitors in bacteria: an emerging concept for the regulation of bacterial protein complexes? *Mol. Microbiol.* 60, 1323–1326. doi: 10.1111/j.1365-2958.2006.05181.x

Shakeri Yekta, S., Liu, T., Bjerg, M. A., Šafarič, L., Karlsson, A., Björn, A., et al. (2019). Sulfide level in municipal sludge digesters affects microbial community response to long-chain fatty acid loads. *Biotechnol. Biofuels* 12, 1–15. doi: 10.1186/s13068-019-1598-1

Silva, P. J., de Castro, B., and Hagen, W. R. (1999). On the prosthetic groups of the NiFe sulfhydrogenase from *Pyrococcus furiosus*: topology, structure, and temperature-dependent redox chemistry. *J. Biol. Inorg. Chem.* 4, 284–291. doi: 10.1007/s007750050314

Søndergaard, D., Pedersen, C. N. S., and Greening, C. (2016). HydDB: a web tool for hydrogenase classification and analysis. *Sci. Rep.* 6:34212. doi: 10.1038/srep34212

Suominen, S., Dombrowski, N., Sinninghe Damsté, J. S., and Villanueva, L. (2021). A diverse uncultivated microbial community is responsible for organic matter degradation in the Black Sea sulphidic zone. *Environ. Microbiol.* 23, 2709–2728. doi: 10.1111/1462-2920.14902

Tanaka, Y., Yoshikaie, K., Takeuchi, A., Ichikawa, M., Mori, T., Uchino, S., et al. (2020). Crystal structure of a YeeE/YedE family protein engaged in thiosulfate uptake. *Sci. Adv.* 6:eaba7637. doi: 10.1126/sciadv.aba7637

Thauer, R. K., Jungermann, K., and Decker, K. (1977). Energy conservation in chemotrophic anaerobic bacteria. *Bacteriol. Rev.* 41, 100–180. doi: 10.1128/br.41.1.100-180.1977

Toth, C. R. A., and Gieg, L. M. (2018). Time course-dependent methanogenic crude oil biodegradation: dynamics of fumarate addition metabolites, biodegradative genes, and microbial community composition. *Front. Microbiol.* 8:2610. doi: 10.3389/fmicb.2017.02610

Trifinopoulos, J., Nguyen, L.-T., von Haeseler, A., and Minh, B. Q. (2016). W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Res.* 44, W232–W235. doi: 10.1093/nar/gkw256

van Vugt-Lussenburg, B. M. A., van der Weel, L., Hagen, W. R., and Hagedoorn, P.-L. (2009). Identification of two [4Fe–4S]-cluster-containing hydro-lyases from *Pyrococcus furiosus*. *Microbiology* 155, 3015–3020. doi: 10.1099/mic.0.030320-0

Verkhovsky, M., and Bogachev, A. V. (2010). Sodium-translocating NADH:quinone oxidoreductase as a redox-driven ion pump. *Biochim. Biophys. Acta* 1797, 738–746. doi: 10.1016/j.bbabio.2009.12.020

Villanueva, L., Bastiaan von Meijenfeldt, F. A., Westbye, A. B., Yadav, S., Hopmans, E. C., Dutilh, B. E., et al. (2021). Bridging the membrane lipid divide: bacteria of the FCB group superphylum have the potential to synthesize archaeal ether lipids. *ISME J.* 15, 168–182. doi: 10.1038/s41396-020-00772-2

Wagner, M., and Horn, M. (2006). The *Planctomycetes*, *Verrucomicrobia*, *Chlamydiae* and sister phyla comprise a superphylum with biotechnological and medical relevance. *Curr. Opin. Biotechnol.* 17, 241–249. doi: 10.1016/j.copbio.2006.05.005

Wang, R., Lin, J.-Q., Liu, X.-M., Pang, X., Zhang, C.-J., Yang, C.-L., et al. (2019). Sulfur oxidation in the acidophilic autotrophic *Acidithiobacillus* spp. *Front. Microbiol.* 9:3290. doi: 10.3389/fmicb.2018.03290

Williams, T. J., Allen, M. A., Ivanova, N., Huntemann, M., Haque, S., Hancock, A. M., et al. (2021). Genome analysis of a verrucomicrobial endosymbiont with a tiny genome discovered in an Antarctic lake. *Front. Microbiol.* 12:674758. doi: 10.3389/fmicb.2021.674758

Wood, A. P., Aurikko, J. P., and Kelly, D. P. (2004). A challenge for 21st century molecular biology and biochemistry: what are the causes of obligate autotrophy and methanotrophy? *FEMS Microbiol. Rev.* 28, 335–352. doi: 10.1016/j.femsre.2003.12.001

Youssef, N. H., Farag, I. F., Rudy, S., Mulliner, A., Walker, K., Caldwell, F., et al. (2019). The Wood–Ljungdahl pathway as a key component of metabolic versatility in candidate phylum Bipolaricaulota (Acetothermia, OP1). *Environ. Microbiol. Rep.* 11, 538–547. doi: 10.1111/1758-2229.12753

Yu, H., Wu, C. H., Schut, G. J., Haja, D. K., Zhao, G., Peters, J. W., et al. (2018). Structure of an ancient respiratory system. *Cell* 173, 1636–1649. doi: 10.1016/j.cell.2018.03.071

Zamkovaya, T., Foster, J. S., de Crécy-Lagard, V., and Conesa, A. (2021). A network approach to elucidate and prioritize microbial dark matter in microbial communities. *ISME J.* 15, 228–244. doi: 10.1038/s41396-020-00777-x

Zhang, W., Ding, W., Yang, B., Tian, R., Gu, S., Luo, H., et al. (2016). Genomic and transcriptomic evidence for carbohydrate consumption among microorganisms in a cold seep brine pool. *Front. Microbiol.* 7:1825. doi: 10.3389/fmicb.2016.01825

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

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 25 November 2021
doi: 10.3389/fmicb.2021.738451

Polar Cryoconite Associated Microbiota Is Dominated by Hemispheric Specialist Genera

Jasmin L. Millar[1,2], Elizabeth A. Bagshaw[1], Arwyn Edwards[3], Ewa A. Poniecka[1†] and Anne D. Jungblut[2]*

[1]School of Earth and Environmental Sciences, Cardiff University, Cardiff, United Kingdom, [2]Department of Life Sciences, The Natural History Museum, London, United Kingdom, [3]Institute of Biological, Environmental and Rural Sciences, Aberystwyth University, Ceredigion, United Kingdom

OPEN ACCESS

Edited by:
Virginia P. Edgcomb, Woods Hole Oceanographic Institution, United States

Reviewed by:
Nozomu Takeuchi, Chiba University, Japan
Antonio Quesada, Autonomous University of Madrid, Spain

***Correspondence:**
Jasmin L Millar
millarjl@cardiff.ac.uk

†Present address:
Ewa A. Poniecka, Department of Environmental Microbiology and Biotechnology, Institute of Microbiology, Faculty of Biology, University of Warsaw, Warsaw, Poland

Specialty section:
This article was submitted to Extreme Microbiology, a section of the journal Frontiers in Microbiology

Received: *08 July 2021*
Accepted: *11 October 2021*
Published: *25 November 2021*

Citation:
Millar JL, Bagshaw EA, Edwards A, Poniecka EA and Jungblut AD (2021) Polar Cryoconite Associated Microbiota Is Dominated by Hemispheric Specialist Genera. Front. Microbiol. 12:738451. doi: 10.3389/fmicb.2021.738451

Cryoconite holes, supraglacial depressions containing water and microbe-mineral aggregates, are known to be hotspots of microbial diversity on glacial surfaces. Cryoconite holes form in a variety of locations and conditions, which impacts both their structure and the community that inhabits them. Using high-throughput 16S and 18S rRNA gene sequencing, we have investigated the communities of a wide range of cryoconite holes from 15 locations across the Arctic and Antarctic. Around 24 bacterial and 11 eukaryotic first-rank phyla were observed in total. The various biotic niches (grazer, predator, photoautotroph, and chemotroph), are filled in every location. Significantly, there is a clear divide between the bacterial and microalgal communities of the Arctic and that of the Antarctic. We were able to determine the groups contributing to this difference and the family and genus level. Both polar regions contain a "core group" of bacteria that are present in the majority of cryoconite holes and each contribute >1% of total amplicon sequence variant (ASV) abundance. Whilst both groups contain Microbacteriaceae, the remaining members are specific to the core group of each polar region. Additionally, the microalgal communities of Arctic cryoconite holes are dominated by *Chlamydomonas* whereas the Antarctic cryoconite holes are dominated by *Pleurastrum*. Therefore cryoconite holes may be a global feature of glacier landscapes, but they are inhabited by regionally distinct microbial communities. Our results are consistent with the notion that cryoconite microbiomes are adapted to differing conditions within the cryosphere.

Keywords: cryoconite, illumina sequencing, Antarctic microbiology, Arctic microbiology, pole-to-pole, 16S rRNA gene, 18S rRNA gene

INTRODUCTION

Cold climate habitats are present across the globe, sustaining a surprising abundance of life above, within, and below the ice (Boetius et al., 2015). Approximately 10% of the Earth's surface is covered by glacial ice (Hornberger and Winter, 2009), and glaciers and ice sheets are now regarded as a distinct biome (Hodson et al., 2008; Anesio and Laybourn-Parry, 2012). On the surface of glaciers and ice sheets, particular hotspots of diversity occur when these living cells and the associated organic matter they produce accumulate. The term "cryoconite"

refers to the microbially-aggregated wind-blown dust, organic and mineral matter that forms on glacial surfaces, particularly in the ablation zone (Cook et al., 2015). The dark colour of the cryoconite depresses ice surface albedo, resulting in the localised melting of surface ice, often forming near-cylindrical holes containing meltwater in ice surfaces known as cryoconite holes (Nordenskjold, 1875). Cryoconite provides a habitat in ice for a plethora of organisms (Vincent and Laybourn-Parry, 2009; Cameron et al., 2012; Lutz et al., 2015), and is common to glaciers worldwide, including those of the Arctic and Antarctic (Cook et al., 2015). Many of the microbes present produce extracellular polymeric substances that allow them to form biofilms and increase the habitability of their surrounding environment. In cryoconite holes, these polymeric substances cause other materials such as sediment to adhere to the cells (Langford et al., 2010), forming a stabilised habitat (Langford et al., 2010; Takeuchi et al., 2010; Webster-Brown et al., 2015; Cook et al., 2016).

The structure of cryoconite holes varies according to the prevailing physical conditions. In cold, dry climates such as continental Antarctica, cryoconite holes are usually covered by an ice lid through most or all of the year (Tranter et al., 2004; Webster-Brown et al., 2015). In the McMurdo Dry Valleys, ice lids of up to 30 cm have been observed on the majority of cryoconite holes year-round (Fountain et al., 2004; Tranter et al., 2004). Some of these thick lids may melt during particularly warm periods, years apart (Foreman et al., 2007). Around 50% of lidded holes are hydrologically connected under the ice surface; the other 50% are completely isolated (Fountain et al., 2004). The holes may melt under the ice lid, connecting to one another on a seasonal or sub-seasonal timescale (Bagshaw et al., 2007). These closed holes are a contrasting environment to the seasonally "open" cryoconite holes found elsewhere, particularly in many Arctic and mountainous regions (Cook et al., 2015). Open holes do not have a permanent ice lid and can hence exchange gas with the atmosphere. These holes primarily occur on glacial ablation zones that exhibit seasonal melting (Bagshaw et al., 2012), experiencing regular flushing by stream flow which distributes cryoconite across the glacier surface (Irvine-Fynn et al., 2010).

Varying cryoconite environments lead to a range of cryoconite microbial ecosystems (Edwards et al., 2011). Sequences extracted from cryoconite communities tend to be dominated by Proteobacteria, Bacteroidetes, Cyanobacteria, and microalgae (Cameron et al., 2012; Edwards et al., 2013; Sommers et al., 2018). They also harbour fungi, protists, and micro-animals (meiofauna; Zawierucha et al., 2015). The relative abundance of these organisms varies with biogeography (Liu et al., 2017; Darcy et al., 2018). Research to date has revealed that communities are more similar within glaciers than between glaciers, but lack sufficient detail to determine whether trends are local, regional, or global. In a study using community fingerprinting and clone library analysis, it was discovered that there is a divide between the bacterial communities of the Arctic and Antarctic (Cameron et al., 2012). These results raised important questions about the potentially distinct biomes of the poles and the variability of cryoconite. However, the methods used could only yield limited detail in comparison to insights arising from the rapid evolution of high throughput DNA sequencing technologies in the ensuing decade. It has not yet been established which taxa are contributing to this difference in community composition.

In this study, we used 16S and 18S rRNA gene high throughput sequencing to cover Archaea, Bacteria, and Eukarya. Through these methods, we show the presence of distinct communities of the Arctic and Antarctic in significantly more detail than previous studies (Mueller and Pollard, 2004; Cameron et al., 2012; Hell et al., 2013; Sommers et al., 2018). We also reveal a genus-level breakdown of the composition of cryoconite communities. The breadth of geographic coverage in our sample set allowed comparison not only between the polar regions, but also between different environments within each polar region. We compare Greenland cryoconite from cryoconite holes formed the year of collection, and "ice core" cryoconite frozen the previous year. We also compare samples from ice sheet interiors with marginal locations. Together, these results provide an in-depth analysis of cryoconite ecology at a global scale, and improve understanding of the distinct communities of the polar regions.

MATERIALS AND METHODS

Cryoconite Sampling Sites and Sample Collection

Samples were collected from 85 individual cryoconite holes across 15 sites (**Table 1**). Ten locations were in the Antarctic and five were in the Arctic, and all were collected during the regional summer melt season (**Figure 1**). The Antarctic samples were collected in the McMurdo Dry Valleys, with the exception of the Utsteinen samples, which were collected in Queen Maud Land near the Utsteinen Nunatak (Lutz et al., 2019). Collection technique was determined by the presence and thickness of ice lid or ice layers above. Three sample types were collected from the southwest of Greenland. Those labelled "Greenland Margin" were collected within 2 km of the ice edge in 2014. Samples labelled "Greenland Core" were frozen cryoconite layers collected from shallow subsurface layers of 1 m ice cores on inland ice in the "Dark Zone," at Camp Black and Bloom (Poniecka et al., 2019; Williamson et al., 2020). Typically the layers are located around 10–30 cm depth of the core and are believed to be retained from the previous year's cryoconite holes following burial under frozen snowpack (Nicholes et al., 2019). Greenland core samples were drilled in 2015 using a hand auger and the sediments separated from the meltwater (Poniecka, 2020). Greenland surface samples were also collected at Camp Black and Bloom 2016. Two sample sets were collected from elsewhere in the Arctic at Storglaciaren in Sweden (2017), and Midtre Lovénbreen in Svalbard (2016). Exposed cryoconite was sampled using a syringe or spoon and stored in clean, sterile Whirlpak bags or tubes, and frozen until analysis (Edwards et al., 2011; Poniecka et al., 2019; Poniecka, 2020). Antarctic Dry Valley samples from Canada, Taylor and Commonwealth glaciers were collected from the base of frozen cores (20–50 cm

TABLE 1 | GPS locations and site names for cryoconite samples from the Arctic and Antarctic.

Sample site	Location	No. of samples	Width (cm)	Ice Lid
Canada Glacier, Antarctica	77.6°S, 163.0°E	9	23–57	Present
Commonwealth Glacier, Antarctica	77.6°S, 163.3°E	9	28–64	Present
Taylor Glacier, Antarctica	77.7°S, 162.0°E	9	41–52	Present
Upper Wright Glacier, Antarctica	77.5°S, 162.9°E	7	37–57	Present
Lower Wright Glacier, Antarctica	77.5°S, 160.7°E	6	27–43	Present
Diamond Glacier, Antarctica	79.8°S, 159.0°E	4	41–52	Present
Miers Glacier, Antarctica	78.8°S, 163.7°E	5	14–39	Present
Upper Koettlitz Glacier, Antarctica	78.3°S, 163.63°E	9	28–97	Present
Lower Koettlitz Glacier, Antarctica	78.1°S, 164.2°E	5	35–55	Present
Utsteinen Scoop, Antarctica	72.0°S, 23.3°E	2	5–10	Present
Kangerlussuaq, Greenland	67.1°N, 50.7°E	3	15–40	Absent
Ice Margin, Greenland	67.2°N, 50.0°E	4	5–99	Absent
Ice Core, Greenland	67.1°N, 50.7°E	3	1–15	Present (layers of ice above)
Storglaciären, Sweden	67.1°N, 18.6°E	3	Unknown	Absent
Midtre Lovénbreen, Svalbard	78.8°N, 12.1°E	5	1–15	Absent

deep) drilled from cryoconite holes between 2005 and 2009 (Bagshaw et al., 2012). The cores were melted at room temperature, the meltwater removed, and the sediment transferred to Nalgene bottles previously rinsed six times with deionised water, then refrozen at −20°C. The cryoconite from Koettlitz, Wright and Darwin Glaciers were collected using a sterile spatula as described in Webster-Brown et al. (2015). Utsteinen Nunatak samples were accessed in 2017 using a Kovacs drill, and the sediment removed using a sterilised scoop (Lutz et al., 2019). Greenland margin and 'Kangerlussuaq' samples were collected from melted cryoconite holes in summer 2012 and 2014, where clean nitrile gloves were used to scoop sediment into Ziploc bags previously rinsed with deionised water. Samples were frozen within a few hours of collection. All samples were transported frozen to home laboratories and remained frozen at −20°C until analysis.

DNA Extraction, PCR, Purification, Quantification, and Illumina Miseq Sequencing

Subsamples of each cryoconite sample were melted at 4°C. Genomic DNA was extracted using the DNeasy PowerSoil Kit (QIAGEN) according to manufacturer's instructions. The prokaryotic 16S (V4 region) and eukaryotic 18S (V9 region) rRNA genes were amplified by PCR. DNA volumes of 0.5, 1.0, and 1.5 µl DNA were added to 19 µl of GoTaq Polymerase reaction mix. In cases where the DNA concentration was too low to yield detectable quantities of PCR product, volumes of 1.5, 2.0, and 2.5 µl were used. The final master mix contained 4 µl 5x GoTaq Flexi buffer, 2 µl MgCl2 (25 µM, Promega, Madison, United States), 0.8 µl Bovine Serum Albumin (BSA, 20 mg/ml BSA, NEB, United Kingdom), 0.16 µl of 200 µM dnTPs (Bioline, United Kingdom), 9.84 µl H2O, 0.2 µl Taq polymerase (5 U/µl, Promega, Madison, United States), and 1 µl of each forward and reverse primer (10 µM). The forward primer 515 F (GTGCCAGCMGCCGCGGTAA) and reverse primer 806R (GGACTACHVGGGTWTCTAAT) containing the MiSeq sequencing adapters and 12-nucleotide Golay barcodes were used to amplify the V4 hyper-variable region of the bacteria and archaea 16S rRNA gene (260 bp, Caporaso et al., 2011). The primers 1391F and EukBr were used to amplify the V9 hypervariable region of the eukaryote 18S rRNA genes containing MiSeq sequencing adapters, a 12-nucleotide Golay barcode on the reverse primer (130 bp, Amaral-Zettler et al., 2009; Caporaso et al., 2011, 2012). The 16S rRNA gene was amplified in an thermocycler using the following program: initial denaturation at 94°C for 2 min followed by 30 cycles of denaturation at 94°C for 45 s, annealing at 50°C for 60 s, and elongation at 72°C for 90 s; and then, a final extension of 72°C for 10 min. The 18S rRNA gene was amplified using the following program: initial denaturation at 94°C for 3 min followed by 35 cycles of denaturation at 94°C for 45 s, annealing at 57°C for 60 s, and elongation at 72°C for 90 s; then, a final extension of 72°C for 10 min. The PCR products along with a negative control were verified by gel electrophoresis using 1.5% agarose for 18S rRNA gene and 1% agarose gel for 16S rRNA gene PCR-products. Following purification according to AxyPrep Mag PCR clean-up protocol (Axygen), the triplicate PCR-products per sample were combined and concentrations determined using a Qubit 2.0 Fluorometer (ThermoFisher Scientific, Waltham, MA, United States) and the manufacturer's protocol. The 16S and 18S rRNA gene amplicons of each sample were separately in preparation for sequencing. The PCR products were sequenced at the Natural History Museum sequencing facility using an Illumina MiSeq platform (Illumina, San Diego, CA, United States).

16S and 18S rRNA Gene Sequence Analysis

The raw sequence data were processed using QIIME2 v2018.8 (Caporaso et al., 2012; Bolyen et al., 2019). Sequences were demultiplexed based on Golay barcodes as a pre-processing step on the Illumina Miseq platform. Reads were quality-filtered, joined, chimeras were removed, and amplicon sequence variants (ASVs) were generated using DADA2 (Callahan et al., 2016). Alignment was performed with MAFFT (Katoh and Standley, 2013), and low complexity and repeating sequences were removed using the mask function in QIIME2. Phylogenetic trees were

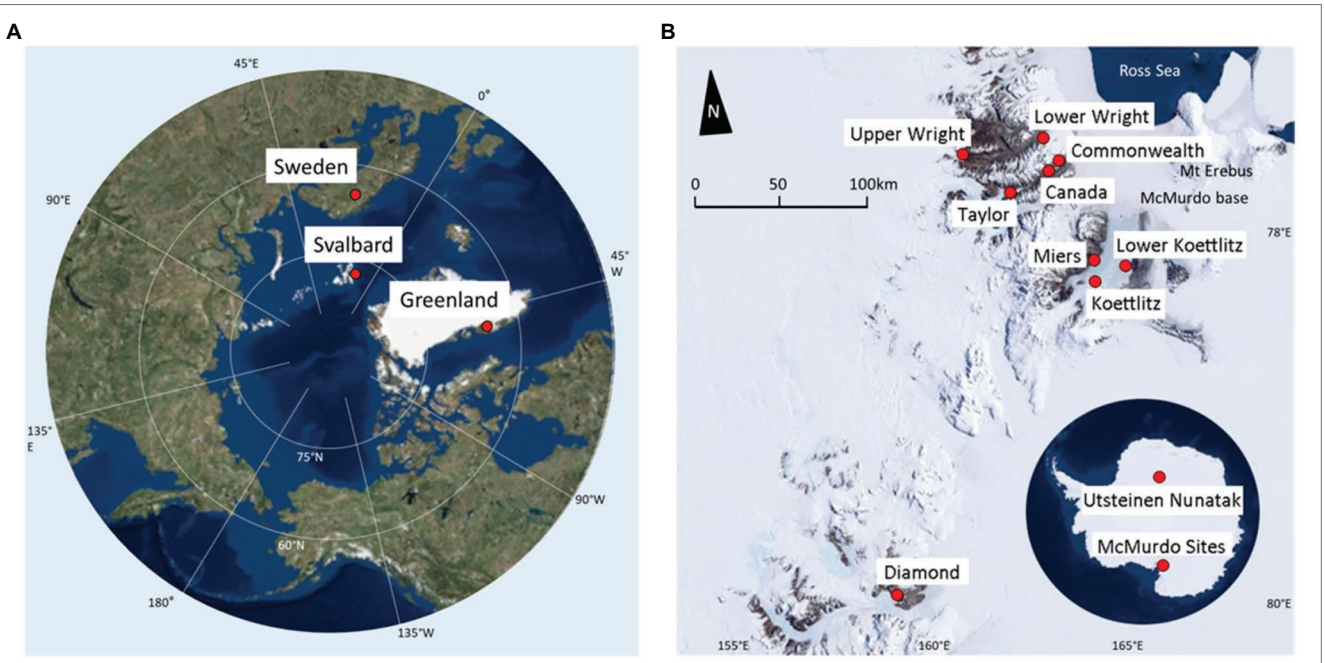

FIGURE 1 | Sampling locations. **(A)** Arctic map, highlighting regions sampled. Points marked indicate sample site locations and names. **(B)** Map of Antarctic sampling locations. Antarctic continental map highlighting the Utsteinen Scoop and the McMurdo Dry Valleys. Underlying is a McMurdo Valleys map, which shows a detailed view of those sites. Base maps were created using ArcGIS® software by Esri. ArcGIS® and ArcMap™ are the intellectual property of Esri and are used herein under license. Source: ArcWorld Supplement.

constructed with Fasttree (Price et al., 2009). Taxonomy was assigned with sklearn-based taxonomy classifier using the SILVA 138 database (Quast et al., 2013). Representative sequences for each ASV were assigned to the highest confidence and identity match on the SILVA 138 database. ASVs assigned to the same taxon were grouped for relative abundance analyses. Chloroplast and mitochondrial DNA were excluded from the prokaryote dataset. ASVs with a frequency < 3 were removed and the dataset was rarefied to 13,044 16S rRNA gene sequences and 6,261 18S rRNA gene sequences. Relative taxa abundance and ASV counts were then generated using QIIME2. About 0.04% of the 16S rRNA gene assignment output, and 12.62% of 18S rRNA gene assignment output were unassigned to a domain or any lower classification. These were removed as they are unlikely to be relevant and correct sequences. The highest Genbank BLASTn match was also obtained for top most abundant 20 16S rRNA gene features and 20 18S rRNA gene features for verification. Alpha diversity was calculated using both Shannon's diversity (Shannon, 1948), and the Simpson index (Simpson, 1949). Beta diversity was tested using R packages vegan and phyloseq, and ggplot2 was used to visualise the results (McMurdie and Holmes, 2013; Wickham, 2016; Oksanen et al., 2019). Non-metric dimensional scaling of ASV relative abundance was performed using Bray-Curtis distances. The species richness correlation was calculated using Pearson's test for correlation. Occupancy of ASVs in each sample was plotted based on mean relative abundance of each ASV per sample against presence of ASV in samples using ggplot2 (Wickham, 2016).

RESULTS

Bacteria and Archaea Composition and Community Assembly in Arctic and Antarctic Cryoconite

Following filtering and rarefaction, 13,044 16S rRNA gene sequences per sample were obtained and 24 bacterial and one archaeal phyla were identified. About 4,497 distinct 16S rRNA gene ASVs were identified, 313 of which were shared between the Arctic and Antarctic. Only nine archaeal sequences were found, appearing in seven samples in the McMurdo Dry Valleys. Cyanobacteria, Proteobacteria, Bacteroidetes, and Actinobacteria were the most abundant bacterial phyla, accounting for 65% of the total 16S rRNA gene ASVs (**Figure 2**). These phyla were present across all samples. All other phyla contributed to <6% of the total 16S rRNA gene ASVs.

Although the same phyla had the highest relative abundance across all locations, there are noticeable differences between the Arctic and Antarctic 16S rRNA gene cryoconite composition. Arctic cryoconite communities had a higher relative abundance of Chloroflexi and Armatimonadetes, and a lower relative abundance of Cyanobacteria. Cyanobacteria accounted for 9% of ASVs across Arctic locations compared to 24% in cryoconite from Antarctica. The variation between Arctic and Antarctic taxonomic diversity within the datasets was explored by analysis of beta diversity. We found a compositional divide between Arctic and Antarctic cryoconite ecosystems based on the relative abundance of taxa. An ANOSIM test of the dissimilarity between poles using the Bray-Curtis method produced an *R* value of

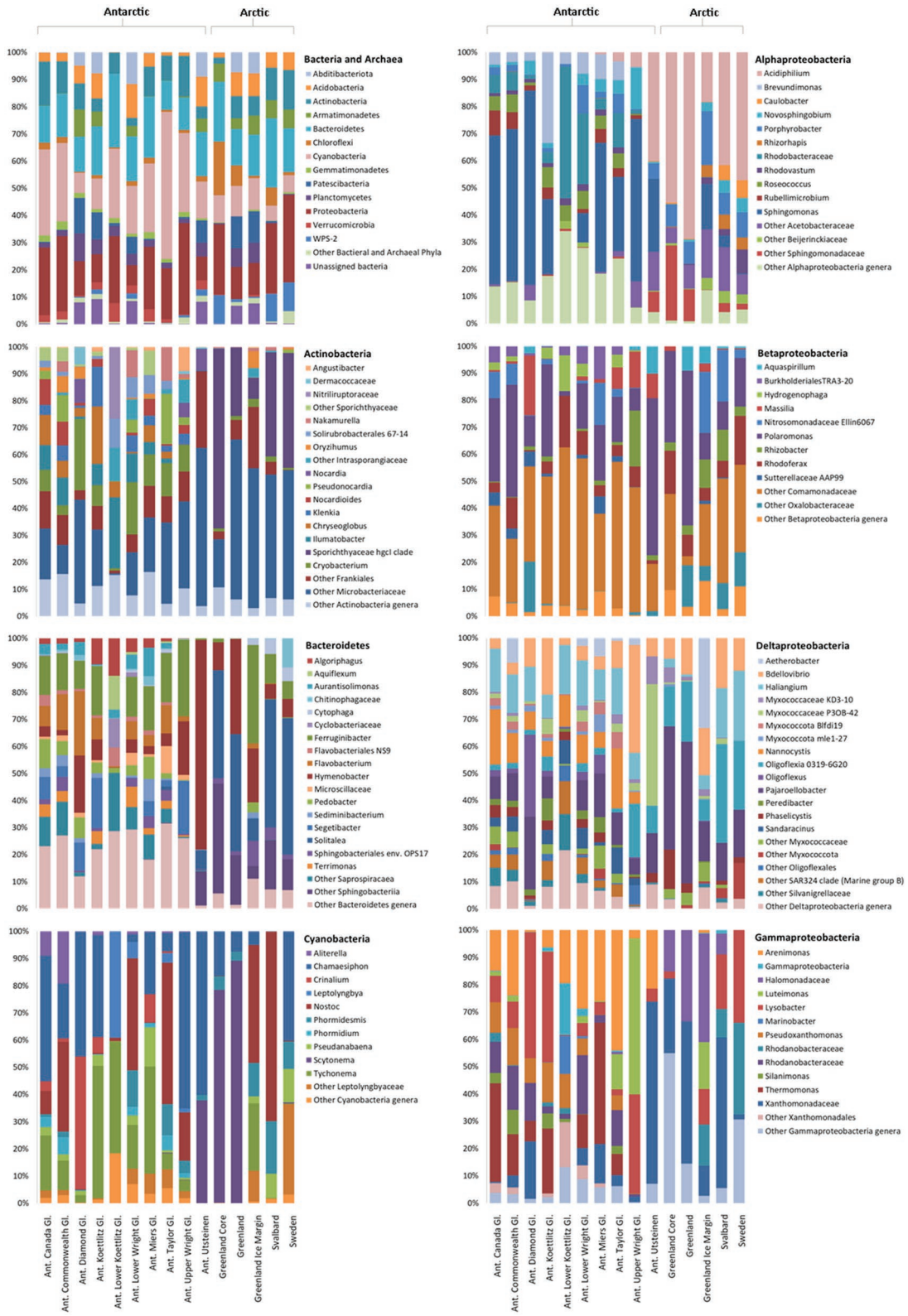

FIGURE 2 | Relative abundance of bacteria in Arctic and Antarctic cryoconite, averaged by glacier. Top left: Bacterial phyla. Phyla contributing to <1% of the total abundance are grouped as "Other." Remaining panels: Relative abundance of genera within the top four most abundant phyla. Proteobacteria have been divided into alpha-, beta-, delta-, and gammaproteobacteria. Genera contributing to <1% of the total abundance are grouped as "Other." Where a genus was unknown, lowest rank known is shown. Taxa were assigned according to the SILVA database. "Ant. Gl." denotes Antarctic Glaciers.

0.723 at $p = 0.0001$ (pairwise community dissimilarity), suggesting a significant level of dissimilarity (**Supplementary Table 2**). Non-metric multi-dimensional scaling of Bray-Curtis distances ordinates this separation between Arctic and Antarctic cryoconite communities (**Figure 3**). The Greenland margin and ice core samples cluster closest with the other Greenland samples, and then other Arctic sites. While there was more variation between the types of Greenland samples (margin, ice core, and ice sheet surface) than within groups, none were outliers within the grouping of Arctic samples. Therefore, Greenland margin and Greenland ice core samples have been grouped with the other Arctic samples during polar region comparisons. The range and mean number of distinct ASVs was considerably lower in the Arctic than Antarctic (**Supplementary Table 1**). Alpha diversity and evenness of bacteria and archaea in the samples was investigated further using the Shannon diversity indices (**Figure 4A**). The lowest richness was identified in cryoconite samples from the Antarctic Upper Wright (4.76)

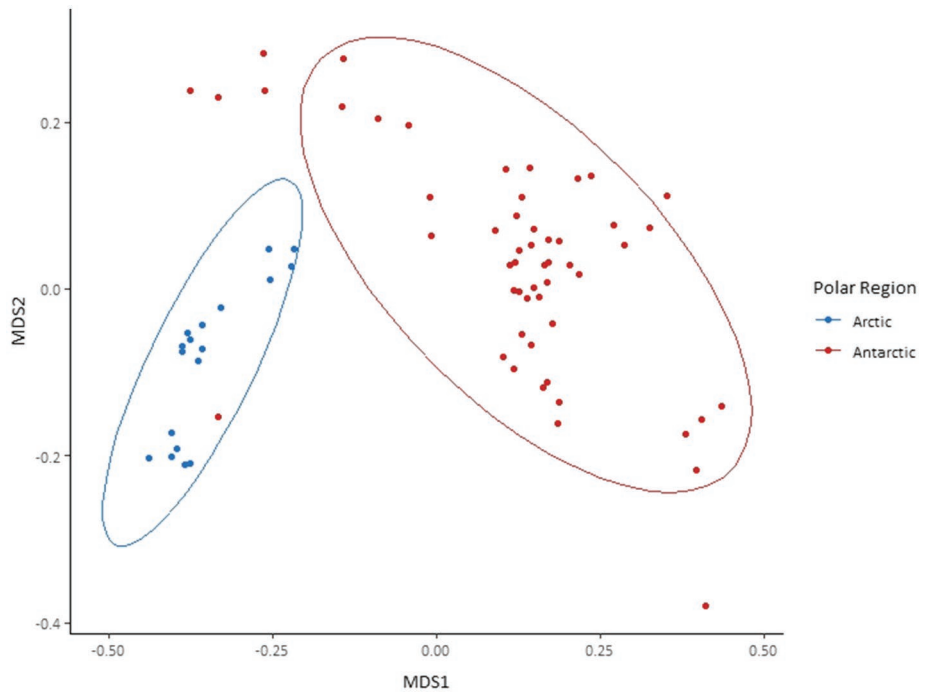

FIGURE 3 | Bray-Curtis dissimilarity of 16S rRNA gene amplicon sequence variants (ASVs) found in each cryoconite and visualised by non-metric multidimensional scaling (NMDS) ordination. Cryoconite holes are grouped by polar region. Ellipses represent 95% CIs.

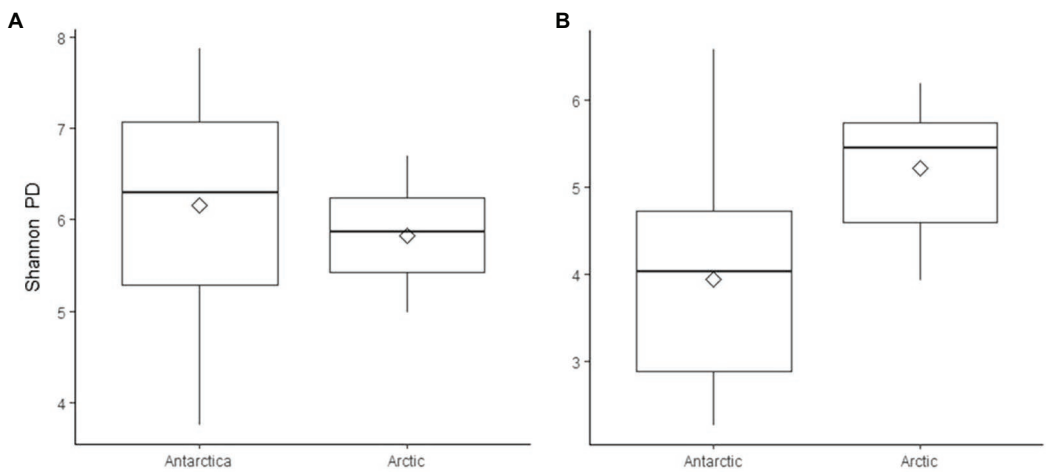

FIGURE 4 | Shannon's phylogenetic diversity of **(A)** 16S rRNA gene ASVs and **(B)** 18S rRNA gene ASVs in cryoconite samples, grouped by polar region. Diamonds indicate mean values.

and Lower Wright (4.85) glaciers. The highest values were also from the Antarctic such the Commonwealth (7.15) and Canada (6.66) glaciers. The mean values were 6.22 for the Antarctic and 5.62 for the Arctic.

The four most abundant bacterial phyla were investigated further to elucidate the composition of their genera and contribution to the dissimilarity between the poles (**Figure 2**). In the Actinobacteria, there is a striking difference between Arctic and Antarctic samples. While the Antarctic samples, with the exception of the Upper Wright glacier, contain a diverse array of Actinobacteria, the Arctic samples are dominated by Sporichthyaceae hgcl clade and Microbacteriaceae which make up 77% of Arctic Actinobacteria sequences. In the Antarctic samples, the Sporichthyaceae hgcl clade contributes 0.04% and the Microbacteriaceae contributes 27%. The psychrophile *Cryobacterium* contributes to 9% of total Actinobacteria sequences and 12% of Antarctic Actinobacteria sequences. In the Bacteroidetes, the Arctic cryoconite holes vary from the Antarctic cryoconite holes in their high relative abundance of *Solitalea* (39% of Bacteroidetes sequences in the Arctic compared to 1% in the Antarctic). The Bacteroidetes ASVs cover 155 genera. Around 139 of these were present in Antarctic and only 51 in Arctic cryoconite communities. Of these 37 present in Arctic samples, 11were only present in the Greenland margin site. These included the *Segetibacter* and *Spirosoma*. The most abundant Cyanobacterial genera were *Tychonema*, *Chamaesiphon*, *Nostoc*, *Scytonema*, and *Tychonema*. Together these contributed to 71% of the total cyanobacterial ASVs. All other groups each contributed <6%. Around 28 genera were found in total, all of which were present in the Antarctic cryoconite communities, but only 11 of these were found in the Arctic samples. The largest difference in percentage proportion between the Arctic and Antarctic was *Scytonema* (42% of Arctic and 3% of Antarctic samples). This genus is absent from the Greenland margin sediment but dominated the other Greenland samples. It was also absent in cryoconite from Svalbard.

The Alphaproteobacteria made up 38% of Proteobacteria 16rRNA gene sequences in the Antarctic and 66% of Proteobacteria sequences in Arctic cryoconite, the Betaproteobacteria comprised 33% of Antarctic and 18% of Arctic Proteobacteria sequences, and the Gammaproteobacteria comprised 22% of Antarctic and 6% of Arctic Proteobacteria sequences. Other 16S rRNA gene sequences belonged to Deltaproteobacteria (7 and 10% of Proteobacteria 16S rRNA gene sequences in Antarctic and Arctic cryoconite, respectively) and unassigned Proteobacteria (0.02 and 0.05% of Proteobacteria 16S rRNA gene sequences in the Antarctic and Arctic cryoconite, respectively). The composition of Betaproteobacteria genera was similar between the Arctic and Antarctic, but there were more differences between the polar regions in the Alphaproteobacteria, Deltaproteobacteria, and Gammaproteobacteria. In the Alphaproteobacteria, *Sphingomonas* had a high relative abundance in the Antarctic (41% of Alphaproteobacteria compared to 3% in the Arctic) whereas *Acidiphilium* had higher relative abundant in the Arctic (50% of Alphaproteobacteria compared to 7% in the Antarctic). The genus *Nannocystis*, which contributed 10% to the Antarctic Deltaproteobacteria, was absent from the Arctic samples. Oligoflexia 0319-6G20 contributed to 4% of Antarctic Deltaproteobacteria sequences and 4% of Arctic Deltaproteobacteria. Of the 31 genera that contributed to >1% of Deltaproteobacterial 16S rRNA gene sequences, 28 were present in the Antarctic samples and nine in the Arctic samples. A large proportion of the 16S rRNA gene sequences assigned as Gammaproteobacteria in the Arctic was assigned to the Halomonadaceae (21%), a group of halophiles, and Xanthomonadaceae (24%). Interestingly, the Gammaproteobacteria 16S rRNA gene composition in samples from the Greenland margin samples bore a greater similarity to the Svalbard samples than the other Greenland samples, largely consisting of, *Lysobacter* and unassigned Rhodanobacteraceae in addition to Halomonadaceae and Xanthomonadaceae. About 49 of the 51 Gammaproteobacteria genera were found in the Antarctic, 17 were present in the Arctic, and 15 were present in both the polar regions.

Separate non-metric multidimensional scaling (NMDS) analysis of the four most abundant phyla Cyanobacteria, Bacteroides, Proteobacteria, and Actinobacteria 16S rRNA gene sequences suggested that the presence of distinct communities in the Arctic and Antarctica (**Figure 5**). ANOSIM analyses produce R values of 0.5 or above for Actinobacteria and Proteobacteria at a significance level of $p = 0.0001$. However, the R values are greater when testing for differences in phyla composition between locations (0.24–0.66 between poles, 0.24–0.75 between locations; **Supplementary Table 2**). The contribution of each prokaryotic phylum to the dissimilarity between poles was investigated using SIMPER analysis (**Supplementary Table 3**). The phyla with a significant contribution to the dissimilarity between the poles were less abundant groups including Chloroflexi and the WPS-2 group. To further investigate the distribution of ASVs and its differences between the poles, occupancy of each ASV across the cryoconite holes was plotted against mean relative abundance of that ASV for each polar region. Both the polar regions contain a "core group" of ASVs that are present in the majority of cryoconite holes and each contribute >1% of total abundance. However, the genera in these groups differ between the Arctic and Antarctic. Whilst both groups contain Microbacteriaceae, the remaining members are specific to the core group of each polar region. All core taxa contained in the Arctic and Antarctic groups are present in some proportion on both poles with the exception of Blastocatellaceae.

Eukaryote Composition and Community Assembly

Following filtering, 6,261 18S rRNA gene sequences were analysed from each Arctic and Antarctic cryoconite sample, which were assigned to a total of 11 eukaryotic phyla. Around 6,865 distinct 18S rRNA gene ASVs were found, 82 of which were shared between the Arctic and Antarctic. Metazoa accounted for 31% of 18S rRNA gene ASVs. Parachela tardigrades were the most abundant metazoans in the Arctic (64% of Arctic

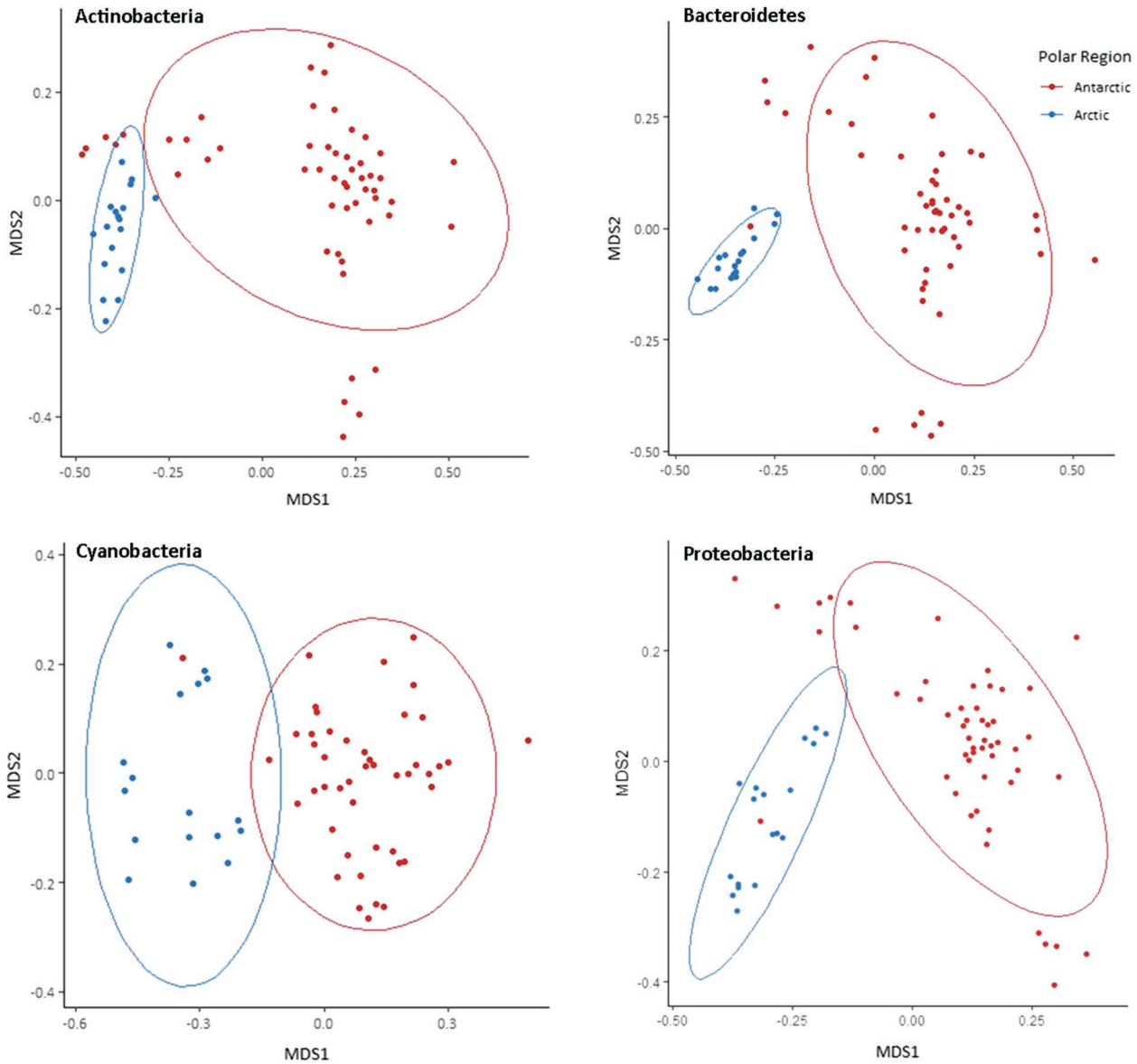

FIGURE 5 | Bray-Curtis dissimilarity of Arctic and Antarctic 16S rRNA gene ASVs in the most abundant four bacterial phyla. Visualised by NMDS ordination. Holes are grouped by polar region. Ellipses represent 95% CIs.

metazoan ASVs and 39% of Antarctic metazoan ASVs) and Adinetida rotifers were the most abundant metazoans in the Antarctic (36% of Arctic metazoan ASVs and 59% of Antarctic metazoan ASVs). The remaining metazoa recovered were Monogononta Ploimida and the platyhelminth Rhabdocoela Neodalyellida, both of which were only present in Antarctic cryoconite holes (contributing to 0.23 and 1.34% of Antarctic metazoan ASVs, respectively). **Figure 6** shows microbial eukaryotes without metazoan 18S rRNA gene sequences. About 94% ASVs assigned to microbial eukaryotic phyla were assigned to the SAR group, Opisthokonta and Archaeplastida.

The 18S rRNA gene eukaryote taxonomic community structure showed less a clear divide between the Arctic and Antarctic cryoconite samples. However, when examining weighted UniFrac distances and abundance of taxa we observed variation of microbial eukaryote composition with polar region (**Supplementary Figure 1**). A much larger proportion of microbial Opisthokonts was found in the Arctic, 32% of assigned Arctic sequences belonged to the Opisthokonta compared to only 6% in the Antarctic. SIMPER analysis showed microbial Opisthokonts contributed significantly to the dissimilarity between poles (**Supplementary Table 3**). Several eukaryotic phyla present in the Antarctic were absent in Arctic cryoconites, but all of these phyla contribute less than 0.1% to the total number 18S rRNA gene sequences across both poles with exception for 18S rRNA gene sequences assigned to

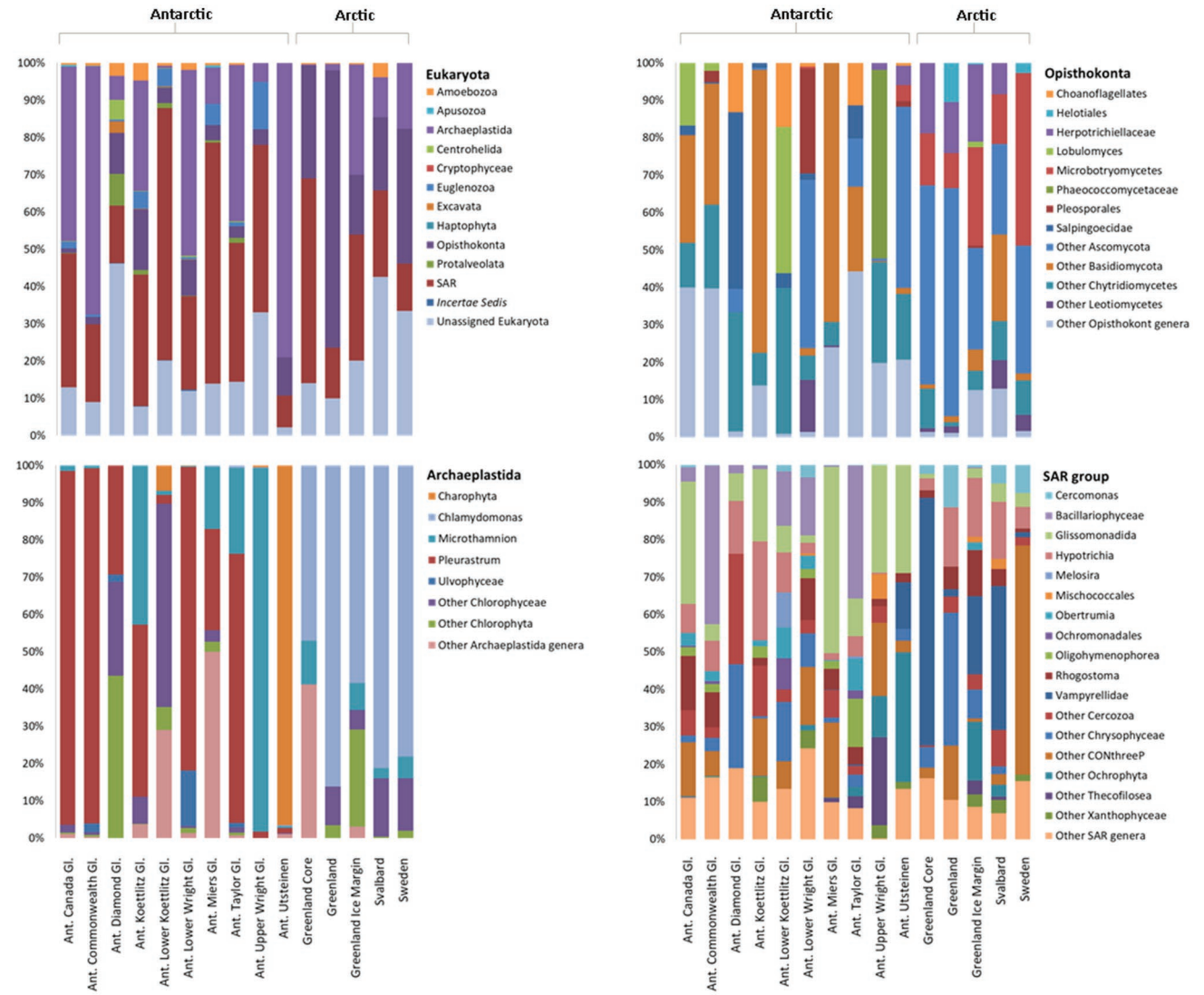

FIGURE 6 | Relative abundance of microbial eukaryotes in Arctic and Antarctic cryoconite, averaged by glacier. Metazoa were excluded from these data. Top left: eukaryotic phyla. Phyla contributing to <1% of the total abundance are not included. Remaining panels: Relative abundance of genera within the top three most abundant phyla. Genera contributing to <1% of the total abundance are grouped as "Other." Where a genus was unknown, lowest rank known is shown. Taxa were assigned according to the SILVA database. "Ant. Gl." denotes Antarctic Glaciers.

Euglenozoawhich were present in all Antarctic locations but the Upper Wright glacier. The Euglenozoa made up 2.5% of the total ASVs.

The most abundant eukaryotic phyla were Opisthokonta, Archaeplastida, and SAR group, which were investigated in further detail. The most striking difference between the Arctic and Antarctic compositions of the SAR group is that the Vampyrellidae, a predatory family of cercozoans, made up 32% of the Arctic SAR 18S rRNA gene sequences but less than 1% of Antarctic 18S rRNA gene sequences in Greenland and Svalbard. The sites in Sweden had 61% of its SAR 18S rRNA genes sequences belonging to CONthreeP ciliates. Bacillariophyceae (diatoms) represented 12% of Antarctic SAR sequences but 0.1% of Arctic sequences (an average of 4.75 ASVs assigned to Bacillariophyceae present, found in the Greenland Margin samples).

Around 89% of Archaeplastida were assigned to one of only four genera. Around 3% of the remaining sequences were unassigned Chlorophyta, 3% to unassigned Chlorophyceae, 3% to unassigned Uvlophyceae, and the remaining 18S rRNA gene sequences each contributed to >1% of the community composition. In the Archaeplastida, 18S rRNA gene sequence from the Arctic, 67% belonged to the snow algae genus *Chlamydomonas*, and 16% of 18S rRNA gene sequences were attributed to unassigned Chlorophyta. In contrast, the Antarctic cryoconite samples contained 0.1% *Chlamydomonas* 18S rRNA gene sequences. Around 52% belonged to *Pleurastrum*, 13% to *Microthamnion*, and 26% to an unassigned group of

Charophyta. No Arctic 18S rRNA gene ASVs were assigned to *Pleurastrum* or Charophyta.

Of the 52 genera assigned to the Opisthokonta, only 11 appear in both the Arctic and Antarctic samples. Overall, 66% of Opisthokonta sequences belonged to the metazoa. The other metazoa contributed less than 1% to the total number 18S rRNA gene sequences. To better distinguish the microbial community, metazoa were excluded from the Opisthokonta in **Figure 6**. Of the microbial Opisthokonts, all samples were dominated by a group of unassigned Ascomycota fungi (39% of total microbial Opisthokont sequences). Microbotryomycetes (19% of Arctic and 1% of Antarctic Opisthokonta) and Herpotrichiellaceae (12% of Arctic and 2% of Antarctic microbial Opisthokonta) were also significant contributors to the Arctic microbial Opisthokonta, whereas Chytridiomycetes were more abundant in the Antarctic (19% of Antarctic sequences and

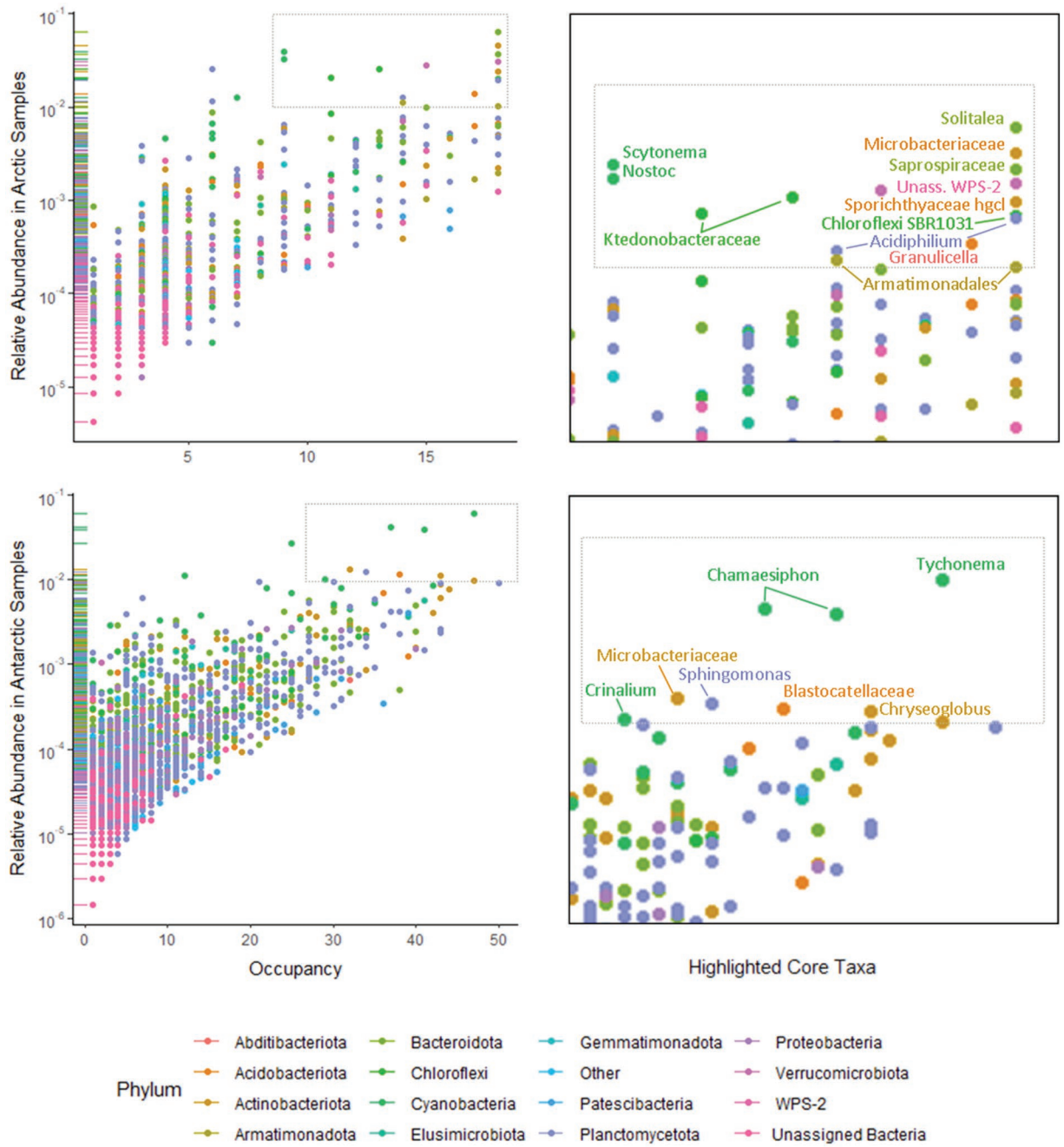

FIGURE 7 | Occupancy of 16S rRNA ASVs in Arctic (upper left) and Antarctic (lower left) cryoconite holes by phyla. Phyla contributing to <1% of the total 16S rRNA gene sequences were grouped as "Other." Taxa present in ≥75% of samples from that polar region that contributed to ≥1% of relative abundance are highlighted as core taxa (upper and lower right).

5% of Arctic sequences). All other assigned groups contributed to >4% of total microbial Opisthokont ASVs.

The mean number of different ASVS was 102 in the Arctic samples and 125 in the Antarctic samples (**Supplementary Table 1**). Alpha diversity and evenness of eukaryotes in the samples was further investigated using the Shannon diversity indices (**Figure 4B**). Similar to 16S rRNA gene communities, the lowest richness was in cryoconite from the Antarctic Lower Wright (2.30) and Upper Wright (3.32) glaciers. However, the highest values were from the Arctic: Svalbard cryoconite (5.94) and the Greenland margin sediment samples (5.87). The average values were 4.00 for the Antarctic and 5.02 for the Arctic sites. Only five of the total assigned 18S rRNA gene ASVs was present in more than one sample. These belonged to (by lowest rank assigned) *Monomastix minuta*, Herpotrichiellaceae, Microbotryomycetes, and Ascomycota.

DISCUSSION

Differences in Community Composition Between Arctic and Antarctic Cryoconite Holes

The use of Illumina sequencing enabled the examination of community composition across more locations and to a greater depth than previous comparisons of Arctic-Antarctic cryoconite microbial communities (e.g., Mueller et al., 2001; Cameron et al., 2012). The comparison of 16S and 18S rRNA gene communities shows that variation between poles was greater than between glaciers or individual cryoconite holes. The bacterial and archaea assemblages in particular clustered according to polar region (**Figures 3, 5**). Additionally, of the "core group" of ASVs that were both more abundant (contributing to >1% of total abundance) and present in the majority of cryoconite holes in each pole, only two genera were shared between the Arctic and Antarctic (**Figure 7**). Greater than 99% of eukaryotic ASVs were only present in one sample, illustrating that unlike the biogeography of bacteria in cryoconite, highly localised variation between cryoconite communities are predominant over regional and hemispheric differences. However, within each of the most abundant eukaryotic phyla (Archaeplastida, Opisthokonta, and the SAR group), there were genera contributing to a compositional divide between the cryoconite of the two polar regions. The results agree with findings by Cameron et al. (2012) that Arctic and Antarctic cryoconite holes harbour distinct bacterial and eukaryotic communities.

One possible explanation for this is that the glacier surface environments of the two poles have different physical characteristics governed by differences in temperature, radiation, and surrounding landscapes, and so may exhibit different selection pressures. For example, lower air and ice temperatures mean that many of the Antarctic cryoconite holes were closed (ice lidded), as is typical for the McMurdo Dry Valleys (Fountain et al., 2004). Most of these holes had a thick layer of sediment at the bottom (between 1 cm and 10 cm). By contrast, the Arctic cryoconite holes tend to be open and contain thinner, more aggregated sediment (<1 cm thickness), which is often clustered into granules (Cook et al., 2016). The ice lid also results in lower levels of light reaching the cryoconite. These differences in habitat may create preferential conditions for some species, altering the community composition and selecting for particular species. Indeed, abiotic variability between cryoconite holes (e.g., pCO2, mineral availability, and temperature) has been shown to impact community structure (Edwards et al., 2011) and Antarctic cryoconite communities are adapted to low light availability (Bagshaw et al., 2016).

Limitations of transport and dispersal have also been shown to contribute to community differences (Telford et al., 2006). The Antarctic soil ecosystem, which is one important source of cryoconite matter, has limited connectivity to the airborne non-polar microbial pool (Pearce et al., 2009; Bottos et al., 2014; Archer et al., 2019). Similar dispersal limitation has been found in the Arctic soils through studies on Actinobacteria community assemblage (Eisenlord et al., 2012). Therefore, differences in community composition may not be solely caused by differing environments, but selection due to transport. There are also strong local winds in Antarctica and transport of biological material have been documented for the McMurdo Dry Valleys (Šabacká et al., 2012; Michaud et al., 2012), which may overshadow lower levels of biological material from long-range transport. If limitations of aeolian transport leads to selection, it follows that there should be segregation between the Arctic and Antarctic microbial communities. Through biogeographical analyses it may be possible for future studies to determine the contribution of these abiotic factors to community dissimilarity, although this is beyond the scope of our analysis. Co-correlation mapping between taxa and environmental variables, such as pCO2, mineral availability, temperature, and cryoconite hole physical parameters would be beneficial, as has been carried out previously in soil microorganisms (King et al., 2010). It would also be valuable to obtain transcriptomic data in addition to gene metabarcoding or genomic data to ascertain differences in active communities (Shakya et al., 2019). A comparison of active and legacy genes in cryoconite holes may also lend insight into which organisms successfully disperse within regions.

It is most likely that both geographical separation and the environment within the poles contribute the community differences we have found. Our 16S and 18S rRNA gene results show that a number of ASVs were present across all locations on one pole and absent on the other. The proportion of highly abundant groups such as Proteobacteria and Cyanobacteria vary between poles, but the significant differences include low abundance phyla: Chloroflexi and the WPS-2 supergroup (**Supplementary Table 3**). Surprisingly, the proportion of photosynthetic organisms was lower in the Arctic than the Antarctic samples (14% of Arctic 16S rRNA gene sequences, 32% of Antarctic 16S rRNA gene sequences, 20% of Arctic 18S rRNA gene sequences, and 49% of Antarctic 18S rRNA gene sequences). This may have implications for the overall autotrophy of the system. It should be noted that the present study uses ribosomal RNA gene sequences, which naturally cannot differentiate between living, dormant, or legacy sources

of ribosomal RNA gene fragments (Blazewicz et al., 2013; Edwards et al., 2020). Such legacy genes may be stored inside cryoconite granules, which show stratification between a photoautotroph-rich exterior and the storage of degraded organic matter within their interiors (Takeuchi et al., 2001; Langford et al., 2010) and is potentially consistent with the enhanced accumulation of phylotypes within larger cryoconite granules (Uetake et al., 2016). By contrast, ribosomal RNA (cDNA) based analyses of communities in western Greenland return highly distinctive (potentially) active communities notable for the dominance of photoautotrophic bacterial lineages (Stibal et al., 2015; Gokul et al., 2019). It is therefore possible that differences in the retention and flushing of legacy DNA between cryoconite habitats may contribute to some difference in gene community composition; however, this is a factor that varies within the polar regions as well as between them.

Environmental differences and the limits of transport contribute to biogeographical clustering within glaciers (Edwards et al., 2011), and this is demonstrated in our data. Cryoconite communities within individual glaciers also tended to cluster, and so the differences between the glaciers could be viewed as the same contributing factors as the differences between poles on a smaller scale. In the Arctic and in Queen Maud Land in Antarctica, geographical clustering is stronger in the bacteria than eukaryotes (Cameron et al., 2012; Lutz et al., 2019), but curiously this result was not previously found in the McMurdo Dry Valleys (Cameron et al., 2012). Our results show clustering to be stronger in the bacteria across all locations on both poles, including the McMurdo Dry Valley sites.

The major phyla present (Actinobacteria, Bacteroidetes, Cyanobacteria, Proteobacteria, Archaeplastida, Opisthokonta, and the SAR group) are consistent with prior studies (Cameron et al., 2012; Kaczmarek et al., 2016) but there were some notable differences. *Mrakia*, a psychrophilic genus were absent in the Antarctic cryoconite and one of the more abundant opisthokont groups in the Arctic samples, despite having been discovered in the Antarctic (Xin and Zhou, 2007). In the eukaryotic microalgae, a higher proportion of *Chlamydomonas* and lower proportion of unknown Chloroccocales algae have been reported in Antarctic cryoconites (Christner et al., 2003; Sommers et al., 2018). Similarly, Bacillariophyceae (diatoms) were found in the Antarctic cryoconite as has been reported previously (Stanish et al., 2013) but, with the exception of one ASV recovered from the Greenland Margin, were absent in the Arctic cryoconite in contrast with previous findings from Greenland and Svalbard (Yallop et al., 2012; Vinšová et al., 2015). Although we found a clear split between the microalgae of the Arctic and Antarctic, it does not follow that those algae are only present on one pole. In light of previous studies, it is more likely that cryoconite holes became dominated by the residents of algal blooms flushed into those cryoconite holes in the season were collected (Yallop et al., 2012; Williamson et al., 2020). Sampling at other times of year and in other locations may yield a different algal community. In addition, 6% of the Archaeplastida could not be assigned to a genus or family (**Figure 6**), so could potentially belong to groups mentioned above or others that have been found in polar cryoconite but not detected here such as Zygnematophyceae (Vonnahme et al., 2016).

16S rRNA and 18S rRNA Gene High Throughput Sequencing Reveals High Taxonomic and Functional Diversity Across Both Poles

Through 16S and 18S rRNA gene sequencing we recovered a total of 35 phyla and superphyla; a considerably higher taxonomic diversity compared to previous studies that did not use next generation sequencing (Mueller et al., 2001; Christner et al., 2003; Mueller and Pollard, 2004; Porazinska et al., 2004; Hodson et al., 2010; Cameron et al., 2012). The most abundant phyla (16S rRNA gene: Bacteroidetes, Actinobacteria, Cyanobacteria, and Proteobacteria; 18S rRNA gene: SAR group, Archaeplastida, and Opisthokonta) were present in all samples, although the relative abundance and genera present varied. The average species richness of the Arctic was lower than that of the Antarctic, however, species richness varied widely between glaciers within each polar region. The Antarctic cryoconite assemblages showed more variation, both when measured by ASVs and by the Shannon diversity index (**Supplementary Table 1**). In contrast to the findings reported by Sommers et al., 2018, only a weak correlation was found between bacterial and eukaryotic diversity in cryoconite holes ($r = 4.1$). The environmental conditions on individual glaciers, such as position within the valley, may contribute to variations in species richness (Stanish et al., 2013).

Bacterial and eukaryotic photoautotrophs were found across all locations and several mixotrophs were recovered, including the *Chlamydomonas* which dominate the Arctic eukaryotic algal population. We also detected chemotrophic bacteria such as *Thiobaccillus*, which plays an important part in sulphur cycling in subglacial environments (Harrold et al., 2015), and may have a role in sulphate reduction in the anoxic zone of cryoconite holes (Bagshaw et al., 2007; Poniecka et al., 2018). There is also a range of organisms predating and grazing on the community. The Vampyrellidae are likely microbial predators in the Arctic cryoconite habitats. Other heterotrophic and predatory groups such as the Ciliophora were found across both poles. Tardigrades and rotifers were found on glaciers with the exception of Sweden. Sweden may be outside the range of Arctic metazoa, though the microbial community is otherwise remarkably similar to that of the Svalbard cryoconite. The tardigrades and rotifers are likely to have been alive and active in the community, as these metazoans are commonly present in cryoconite communities (Porazinska et al., 2004; Zawierucha et al., 2021), and are key constituents of adjacent soil communities (Treonis et al., 1999; Virginia and Wall, 1999). Together these data suggest a complex trophic web within cryoconite holes across both poles, with metazoans as the top level grazer, microbial predation and heterotrophy, chemotrophy, diverse bacterial photoautotrophy, and microalgal assemblages but dominated by a small number of families.

Archaea were detected in very low relative abundances (<0.001% of 16S rRNA gene sequences) in the Arctic and

Antarctic cryoconite holes. Other 16S rRNA gene sequencing surveys on cryoconite have found comparably low relative abundances of archaea (Lutz et al., 2015, 2017; Sommers et al., 2018) when using universal 16S rRNA gene primers (Caporaso et al., 2012). This may be due to these primers being less well suited for the amplification of the archaeal 16S rRNA gene (Parada et al., 2016). Investigations that have used specific primers designed for archaea detected more archaeal ASVs richness, though they are still minor component in comparison to the overall bacteria diversity recovered from cryoconite holes (Cameron et al., 2012; Weisleitner et al., 2020).

Microbial Community Difference Between Glacial Environments

While cryoconite aggregate material can remain in a site for some years, the cryoconite holes may be flushed, buried, and otherwise deformed (Hodson et al., 2008; Bagshaw et al., 2012). In the Arctic sites this may happen more regularly, on a seasonal or sub seasonal timescale, whereas McMurdo Dry Valley cryoconite holes may not be completely flushed for several years (Foreman et al., 2007). Cryoconite is also washed towards the glacial margins over time, in the direction of melt. Previous studies have found retention of a local foundational community on glacier sites over time, but also selection based on highly localised environments (Edwards et al., 2011; Gokul et al., 2016; Segawa et al., 2017). As well as comparing supraglacial habitats between poles, we were also able to compare the communities of ice margin open cryoconite holes, mid-ice sheet open cryoconite holes, and ice covered cryoconite holes formed the previous year in Greenland to broaden our representation of Arctic cryoconite and examine the impact of local habitat within glacier sites. In Svalbard, it has been established that the microbial communities of cryoconite holes were distinct from those of the ice margins, with only a minority of phylotypes appearing in both habitats (Edwards et al., 2013). We confirmed that there is also a distinct difference between cryoconite communities found 60 km onto the ice sheet and those within a few hundred metres of the ice margin in Greenland. There was significantly more variation between the types of Greenland samples (margin, ice core, and ice sheet surface) than within groups. The communities from cryoconite core samples were similar to the communities obtained from the surface Greenland cryoconite. The samples from Greenland core cryoconite and surface cryoconite formed two distinct clusters, but were considerably more similar to each other than to the other Arctic and Antarctic cryoconite communities (**Supplementary Figure 2**). This suggests consistency over the subsequent year as well as location in the Greenland cryoconite, and reinforces its distinction from the community present at the Greenland margin, which is likely influenced by uprafted subglacial debris (Knight et al., 2002). Mixing with a distinctive subglacial microbial community results in a higher abundance of methanogenic and sulphate reducing groups in margin samples when comparing to interior samples (Poniecka, 2020).

The Utsteinen region is situated in Queen Maud Land, an understudied region considerably far removed from the McMurdo Dry Valleys sites (Lutz et al., 2019). Despite the distance between the McMurdo Dry Valley and Utsteinen glaciers, the microbial community composition showed a high similarity. This suggests that there might be similar environmental drivers and sources for microbial communities in cryoconite holes across the Antarctic continent. Both regions are arid inland and glaciers will likely support closed-lidded cryoconite holes. They house recognisable hemisphere-specific cryoconite communities despite the limitations of aeolian transport across the Antarctic (Pearce et al., 2009). However, the Upper Wright Glacier samples bore a closer resemblance to the Arctic samples in several metrics and in some aspects, such the presence of Charophyta in several cryoconite holes, they were unique. There is no clear explanation for this. The Upper Wright sampling site was somewhat distinctive, as it is far from the sea and high altitude (950 m) compared to the other Dry Valley samples. However, there is no certainty in the contributing factors to the Upper Wright's distinct ecology.

CONCLUSION

The use of 16S and 18S rRNA gene high throughput sequencing enabled a more comprehensive examination of taxonomic diversity of bacteria, archaea, and eukaryotes across Arctic and Antarctic cryoconite ecosystems than previous studies. We were able to resolve community composition to the family and often genus level, revealing a diverse community of microbes that contribute to a complex tropic web. Most significantly, it allowed for the direct comparison of microbial assemblages in cryoconite holes from both the Arctic and Antarctic. Our findings suggest that the Arctic and Antarctic cryoconite holes harbour distinct microbial communities, but the various biotic niches (grazer, predator, photoautotroph, and chemotroph) are filled in every location. The "core taxa," which are numerous in both abundance and occupancy, share little similarity between the poles. The characteristics of the local environment and neighbouring habitats play a distinct role in the community composition. Therefore while cryoconite holes may be a global feature of glacier landscapes, they are inhabited by regionally distinct microbial communities.

DATA AVAILABILITY STATEMENT

The dataset presented in this study can be found at NCBI, accession project: PRJNA744712. https://www.ncbi.nlm.nih.gov/bioproject/744712.

AUTHOR CONTRIBUTIONS

JM, AJ, EB, and AE devised the project outline. JM conducted DNA extraction, sequencing, and analysis according to aims set by AJ. AJ, EB, and EP collected and provided cryoconite samples. JM wrote the resulting manuscript with contribution and approval from all authors. All authors contributed to the article and approved the submitted version.

FUNDING

JM was funded by a UK Natural Environmental Research Council Great Western 4+ Doctoral Training Partnership, which also funded the DNA sequence analysis. McMurdo Dry Valley (MCM) samples were collected with the support of the US National Science Foundation MCM Long Term Ecological Research site. Greenland interior samples were obtained during the Black and Bloom fieldwork campaign,[1] funded by UK NERC NE/M021025/1. Collection of the Utsteinen scoop samples was funded by the Baillet Latour Antarctica Fellowship (2016–2018) awarded to Lori Ziolkowski and by the Helmholtz Recruiting Initiative (award no: I-044-16-01) granted to Liane G. Benning.

[1]https://blackandbloom.org/

ACKNOWLEDGMENTS

Illumina sequencing was carried out at the Natural History Museum London Molecular Labs sequencing facility. Samples were kindly provided as follows: Utsteinen Nunatak samples were provided courtesy of Liane G. Benning, Jenine McCutcheon, and Lori Ziolkowski, Sweden samples were provided to EP by Alexandre Anesio, and some Greenland margin samples were collected by James Bradley. We thank the reviewers for their insightful comments and suggestions.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.738451/full#supplementary-material

REFERENCES

Amaral-Zettler, L. A., McCliment, E. A., Ducklow, H. W., and 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:e6372. doi: 10.1371/annotation/50c43133-0df5-4b8b-8975-8cc37d4f2f26

Anesio, A. M., and Laybourn-Parry, J. (2012). Glaciers and ice sheets as a biome. *Trends Ecol. Evol.* 27, 219–225. doi: 10.1016/j.tree.2011.09.012

Archer, S. D. J., Lee, K. C., Caruso, T., Maki, T., Lee, C. K., Cary, S. C., et al. (2019). Airborne microbial transport limitation to isolated Antarctic soil habitats. *Nat. Microbiol.* 4, 925–932. doi: 10.1038/s41564-019-0370-4

Bagshaw, E. A., Stibal, M., Anesio, A. M., Bellas, C., Tranter, M., Telling, J., et al. (2012). "Glacier surface habitats," in *Life at Extremes: Environments, Organisms, and Strategies for Survival*. ed. E. Bell (Wallington, United Kingdom: CABI), 155–175.

Bagshaw, E. A., Tranter, M., Fountain, A. G., Welch, K. A., Basagic, H., and Lyons, W. B. (2007). Biogeochemical evolution of cryoconite holes on Canada glacier, Taylor Valley, Antarctica. *J. Geophys. Res.* 112, 4–35. doi: 10.1029/2007JG000442

Bagshaw, E. A., Wadham, J. L., Tranter, M., Perkins, R., Morgan, A., Williamson, C. J., et al. (2016). Response of Antarctic cryoconite microbial communities to light. *FEMS Microbiol. Ecol.* 92:fiw076. doi: 10.1093/femsec/fiw076

Blazewicz, S. J., Barnard, R. L., Daly, R. A., and Firestone, M. K. (2013). Evaluating rRNA as an indicator of microbial activity in environmental communities: limitations and uses. *ISME J.* 7, 2061–2068. doi: 10.1038/ismej.2013.102

Boetius, A., Anesio, A. M., Deming, J. W., Mikucki, J. A., and Rapp, J. Z. (2015). Microbial ecology of the cryosphere: sea ice and glacial habitats. *Nat. Rev. Microbiol.* 13, 677–690. doi: 10.1038/nrmicro3522

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., et al. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* 37, 852–857. doi: 10.1038/s41587-019-0209-9

Bottos, E. M., Woo, A. C., Zawar-Reza, P., Pointing, S. B., and Cary, S. C. (2014). Airborne bacterial populations above desert soils of the McMurdo dry valleys, Antarctica. *Microb. Ecol.* 67, 120–128. doi: 10.1007/s00248-013-0296-y

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

Cameron, K. A., Hodson, A. J., and Osborn, A. M. (2012). Structure and diversity of bacterial, eukaryotic and archaeal communities in glacial cryoconite holes from the Arctic and the Antarctic. *FEMS Microbiol. Ecol.* 82, 254–267. doi: 10.1111/j.1574-6941.2011.01277.x

Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Huntley, J., Fierer, N., et al. (2012). Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J.* 6, 1621–1624. doi: 10.1038/ismej.2012.8

Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Lozupone, C. A., Turnbaugh, P. J., et al. (2011). Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc. Natl. Acad. Sci. U. S. A.* 108, 4516–4522. doi: 10.1073/pnas.1000080107

Christner, B. C., Kvitko, B. H. 2nd, and Reeve, J. N. (2003). Molecular identification of bacteria and Eukarya inhabiting an Antarctic cryoconite hole. *Extremophiles* 7, 177–183. doi: 10.1007/s00792-002-0309-0

Cook, J. M., Edwards, A., Bulling, M., Mur, L. A. J., Cook, S., Gokul, J. K., et al. (2016). Metabolome-mediated biocryomorphic evolution promotes carbon fixation in Greenlandic cryoconite holes. *Environ. Microbiol.* 18, 4674–4686. doi: 10.1111/1462-2920.13349

Cook, J., Edwards, A., Takeuchi, N., and Irvine-Fynn, T. (2015). Cryoconite: The dark biological secret of the cryosphere. *Prog. Phys. Geogr.* 40, 66–111. doi: 10.1177/0309133315616574

Darcy, J. L., Gendron, E. M. S., Sommers, P., Porazinska, D. L., and Schmidt, S. K. (2018). Island biogeography of Cryoconite hole bacteria in Antarctica's Taylor Valley and Around the world. *Front. Ecol. Evol.* 6:180. doi: 10.3389/fevo.2018.00180

Edwards, A., Anesio, A. M., Rassner, S. M., Sattler, B., Hubbard, B., Perkins, W. T., et al. (2011). Possible interactions between bacterial diversity, microbial activity and supraglacial hydrology of cryoconite holes in Svalbard. *ISME J.* 5, 150–160. doi: 10.1038/ismej.2010.100

Edwards, A., Cameron, K. A., Cook, J. M., Debbonaire, A. R., Furness, E., Hay, M. C., et al. (2020). Microbial genomics amidst the Arctic crisis. *Microb. Genom.* 6:e000375. doi: 10.1099/mgen.0.000375

Edwards, A., Pachebat, J. A., Swain, M., Hegarty, M., Hodson, A. J., Irvine-Fynn, T. D. L., et al. (2013). A metagenomic snapshot of taxonomic and functional diversity in an alpine glacier cryoconite ecosystem. *Environ. Res. Lett.* 8:035003. doi: 10.1088/1748-9326/8/3/035003

Eisenlord, S. D., Zak, D. R., and Upchurch, R. A. (2012). Dispersal limitation and the assembly of soil Actinobacteria communities in a long-term chronosequence. *Ecol. Evol.* 2, 538–549. doi: 10.1002/ece3.210

Foreman, C. M., Sattler, B., Mikucki, J. A., Porazinska, D. L., and Priscu, J. C. (2007). Metabolic activity and diversity of cryoconites in the Taylor Valley. *Antarctica. J. Geophys. Res.* 112, 4–32. doi: 10.1029/2006jg000358

Fountain, A. G., Tranter, M., Nylen, T. H., Lewis, K. J., and Mueller, D. R. (2004). Evolution of cryoconite holes and their contribution to meltwater runoff from glaciers in the McMurdo dry valleys, Antarctica. *J. Glaciol.* 50, 35–45. doi: 10.3189/172756504781830312

Gokul, J. K., Cameron, K. A., Irvine-Fynn, T. D. L., Cook, J. M., Hubbard, A., Stibal, M., et al. (2019). Illuminating the dynamic rare biosphere of the

Greenland ice Sheet's dark zone. *FEMS Microbiol. Ecol.* 95:fiz177. doi: 10.1093/femsec/fiz177

Gokul, J. K., Hodson, A. J., Saetnan, E. R., Irvine-Fynn, T. D. L., Westall, P. J., Detheridge, A. P., et al. (2016). Taxon interactions control the distributions of cryoconite bacteria colonizing a high Arctic ice cap. *Mol. Ecol.* 25, 3752–3767. doi: 10.1111/mec.13715

Harrold, Z. R., Skidmore, M. L., Hamilton, T. L., Desch, L., Amada, K., van Gelder, W., et al. (2015). Aerobic and anaerobic thiosulfate oxidation by a cold-adapted, subglacial chemoautotroph. *Appl. Environ. Microbiol.* 82, 1486–1495. doi: 10.1128/AEM.03398-15

Hell, K., Edwards, A., Zarsky, J., Podmirseg, S. M., Girdwood, S., Pachebat, J. A., et al. (2013). The dynamic bacterial communities of a melting high Arctic glacier snowpack. *ISME J.* 7, 1814–1826. doi: 10.1038/ismej.2013.51

Hodson, A., Anesio, A. M., Tranter, M., Fountain, A., Osborn, M., Priscu, J., et al. (2008). Glacial Ecosystems. *Ecol. Monogr.* 78, 41–67. doi: 10.1890/07-0187.1

Hodson, A., Bøggild, C., and Hanna, E. (2010). "The cryoconite ecosystem on the Greenland ice sheet," in *Annals of Glaciology*. Cambridge University Press.

Hornberger, G., and Winter, T. C. (2009). "Snow and ice," in *Encyclopedia of Inland Waters*. ed. G. E. Likens (Oxford: Academic Press), 773.

Irvine-Fynn, T. D. L., Bridge, J. W., and Hodson, A. J. (2010). Rapid quantification of cryoconite: granule geometry and in situ supraglacial extents, using examples from Svalbard and Greenland. *J. Glaciol.* 56, 297–308. doi: 10.3189/002214310791968421

Kaczmarek, Ł., Jakubowska, N., Celewicz-Gołdyn, S., and Zawierucha, K. (2016). The microorganisms of cryoconite holes (algae, archaea, bacteria, cyanobacteria, fungi, and Protista): a review. *Polar Rec.* 52, 176–203. doi: 10.1017/S0032247415000637

Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. doi: 10.1093/molbev/mst010

King, A. J., Freeman, K. R., McCormick, K. F., Lynch, R. C., Lozupone, C., Knight, R., et al. (2010). Biogeography and habitat modelling of high-alpine bacteria. *Nat. Commun.* 1:53. doi: 10.1038/ncomms1055

Knight, P. G., Waller, R. I., Patterson, C. J., Jones, A. P., and Robinson, Z. P. (2002). Discharge of debris from ice at the margin of the Greenland ice sheet. *J. Glaciol.* 48, 192–198. doi: 10.3189/172756502781831359

Langford, H., Hodson, A., and Banwart, S. (2010). "The microstructure and biogeochemistry of Arctic cryoconite granules," in *Annals of Glaciology*. Cambridge University Press.

Liu, Y., Vick-Majors, T. J., Priscu, J. C., Yao, T., Kang, S., Liu, K., et al. (2017). Biogeography of cryoconite bacterial communities on glaciers of the Tibetan plateau. *FEMS Microbiol. Ecol.* 93. doi: 10.1093/femsec/fix072

Lutz, S., Anesio, A. M., Edwards, A., and Benning, L. G. (2015). Microbial diversity on Icelandic glaciers and ice caps. *Front. Microbiol.* 6:307. doi: 10.3389/fmicb.2015.00307

Lutz, S., Anesio, A. M., Edwards, A., and Benning, L. G. (2017). Linking microbial diversity and functionality of arctic glacial surface habitats. *Environ. Microbiol.* 19, 551–565. doi: 10.1111/1462-2920.13494

Lutz, S., Ziolkowski, L. A., and Benning, L. G. (2019). The biodiversity and geochemistry of Cryoconite holes in Queen Maud Land, East Antarctica. *Microorganisms* 7:160. doi: 10.3390/microorganisms7060160

McMurdie, P. J., and Holmes, S. (2013). Phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* 8:e61217. doi: 10.1371/journal.pone.0061217

Michaud, A. B., Šabacká, M., and Priscu, J. C. (2012). Cyanobacterial diversity across landscape units in a polar desert: Taylor Valley, Antarctica. *FEMS Microbiol. Ecol.* 82, 268–278. doi: 10.1111/j.1574-6941.2012.01297.x

Mueller, D. R., and Pollard, W. H. (2004). Gradient analysis of cryoconite ecosystems from two polar glaciers. *Polar Biol.* 27, 66–74. doi: 10.1007/s00300-003-0580-2

Mueller, D. R., Vincent, W. F., Pollard, W. H., and Fritsen, C. H. (2001). Glacial cryoconite ecosystems: A bipolar comparison of algal communities and habitats. *Nova Hedwigia* 123, 173–197.

Nicholes, M. J., Williamson, C. J., Tranter, M., Holland, A., Poniecka, E., Yallop, M. L., et al. (2019). Bacterial dynamics in supraglacial habitats of the Greenland ice sheet. *Front. Microbiol.* 10:1366. doi: 10.3389/fmicb.2019.01366

Nordenskjold, E. A. (1875). Cryoconite found 1870, July 19th-25th, on the inland ice, east of Auleitsivik Fjord, Disco Bay, Greenland. *Geol. Mag. Decade* 2, 157–162.

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., et al. (2019). vegan: Community Ecology Package. Available at: https://CRAN.R-project.org/package=vegan (Accessed October 25, 2021).

Parada, A. E., Needham, D. M., and Fuhrman, J. A. (2016). Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environ. Microbiol.* 18, 1403–1414. doi: 10.1111/1462-2920.13023

Pearce, D. A., Bridge, P. D., Hughes, K. A., Sattler, B., Psenner, R., and Russell, N. J. (2009). Microorganisms in the atmosphere over Antarctica. *FEMS Microbiol. Ecol.* 69, 143–157. doi: 10.1111/j.1574-6941.2009.00706.x

Poniecka, E. (2020). *The Role of Heterotrophs in Glacier Surface Ecosystem Productivity*. Cardiff University.

Poniecka, E., Bagshaw, E., Sass, H., Williamson, C., Anesio, A., and Tranter, M. (2019). The Secrets of Black Holes on Ice: Eco-physiology of Microorganisms in Cryoconite Holes. in Geophysical Research Abstracts.

Poniecka, E. A., Bagshaw, E. A., Tranter, M., Sass, H., Williamson, C. J., Anesio, A. M., et al. (2018). Rapid development of anoxic niches in supraglacial ecosystems. *Arct. Antarct. Alp. Res.* 50:S100015. doi: 10.1080/15230430.2017.1420859

Porazinska, D. L., Fountain, A. G., Nylen, T. H., Tranter, M., Virginia, R. A., and Wall, D. H. (2004). The biodiversity and biogeochemistry of Cryoconite holes from McMurdo Dry Valley glaciers, Antarctica. *Arct. Antarct. Alp. Res.* 36, 84–91. doi: 10.1657/1523-0430(2004)036[0084:TBABOC]2.0.CO;2

Price, M. N., Dehal, P. S., and Arkin, A. P. (2009). FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Mol. Biol. Evol.* 26, 1641–1650. doi: 10.1093/molbev/msp077

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* 41, D590–D596. doi: 10.1093/nar/gks1219

Šabacká, M., Priscu, J. C., Basagic, H. J., Fountain, A. G., Wall, D. H., Virginia, R. A., et al. (2012). Aeolian flux of biotic and abiotic material in Taylor Valley, Antarctica. *Geomorphology* 155–156, 102–111. doi: 10.1016/j.geomorph.2011.12.009

Segawa, T., Yonezawa, T., Edwards, A., Akiyoshi, A., Tanaka, S., Uetake, J., et al. (2017). Biogeography of cryoconite forming cyanobacteria on polar and Asian glaciers. *J. Biogeogr.* 44, 2849–2861. doi: 10.1111/jbi.13089

Shakya, M., Lo, C.-C., and Chain, P. S. G. (2019). Advances and challenges in Metatranscriptomic analysis. *Front. Genet.* 10:904. doi: 10.3389/fgene.2019.00904

Shannon, C. E. (1948). A mathematical theory of communication. *Bell Syst. Tech. J.* 27, 379–423. doi: 10.1002/j.1538-7305.1948.tb01338.x

Simpson, E. H. (1949). Measurement of diversity. *Nature* 163:688. doi: 10.1038/163688a0

Sommers, P., Darcy, J. L., Gendron, E. M. S., Stanish, L. F., Bagshaw, E. A., Porazinska, D. L., et al. (2018). Diversity patterns of microbial eukaryotes mirror those of bacteria in Antarctic cryoconite holes. *FEMS Microbiol. Ecol.* 94. doi: 10.1093/femsec/fix167

Stanish, L. F., Bagshaw, E. A., McKnight, D. M., Fountain, A. G., and Tranter, M. (2013). Environmental factors influencing diatom communities in Antarctic cryoconite holes. *Environ. Res. Lett.* 8:045006. doi: 10.1088/1748-9326/8/4/045006

Stibal, M., Schostag, M., Cameron, K. A., Hansen, L. H., Chandler, D. M., Wadham, J. L., et al. (2015). Different bulk and active bacterial communities in cryoconite from the margin and interior of the Greenland ice sheet. *Environ. Microbiol. Rep.* 7, 293–300. doi: 10.1111/1758-2229.12246

Takeuchi, N., Kohshima, S., and Seko, K. (2001). Structure, formation, and darkening process of albedo-reducing material (Cryoconite) on a Himalayan glacier: A granular algal mat growing on the glacier. *Arct. Antarct. Alp. Res.* 33, 115–122. doi: 10.1080/15230430.2001.12003413

Takeuchi, N., Nishiyama, H., and Li, Z. (2010). Structure and formation process of cryoconite granules on Ürümqi glacier no. 1, Tien Shan, China. *Ann. Glaciol.* 51, 9–14. doi: 10.3189/172756411795932010

Telford, R. J., Vandvik, V., and Birks, H. J. B. (2006). Dispersal limitations matter for microbial morphospecies. *Science* 312:1015. doi: 10.1126/science.1125669

Tranter, M., Fountain, A. G., Fritsen, C. H., Berry Lyons, W., Priscu, J. C., Statham, P. J., et al. (2004). Extreme hydrochemical conditions in natural microcosms entombed within Antarctic ice. *Hydrol. Process.* 18, 379–387. doi: 10.1002/hyp.5217

Treonis, A. M., Wall, D. H., and Virginia, R. A. (1999). Invertebrate biodiversity in Antarctic Dry Valley soils and sediments. *Ecosystems* 2, 482–492. doi: 10.1007/s100219900096

Uetake, J., Tanaka, S., Segawa, T., Takeuchi, N., Nagatsuka, N., Motoyama, H., et al. (2016). Microbial community variation in cryoconite granules on Qaanaaq glacier, NW Greenland. *FEMS Microbiol. Ecol.* 92:fiw127. doi: 10.1093/femsec/fiw127

Vincent, W. F., and Laybourn-Parry, J. (2009). *Polar Lakes and Rivers—Limnology of Arctic and Antarctic Aquatic Ecosystems*. United Kingdom: Cambridge University Press (CUP).

Vinšová, P., Pinseel, E., Kohler, T. J., Van de Vijver, B., Žárskı`y, J. D., Kavan, J., et al. (2015). Diatoms in cryoconite holes and adjacent proglacial freshwater sediments, Nordenskiöld glacier (Spitsbergen, high Arctic). *Czech Polar Rep.* 5, 112–133. doi: 10.5817/CPR2015-2-11

Virginia, R. A., and Wall, D. H. (1999). How soils structure communities in the Antarctic dry valleys. *Bioscience* 49, 973–983. doi: 10.2307/1313731

Vonnahme, T. R., Devetter, M., Žárský, J. D., Šabacká, M., and Elster, J. (2016). Controls on microalgal community structures in cryoconite holes upon high-Arctic glaciers, Svalbard. *Biogeosciences* 13, 659–674. doi: 10.5194/bg-13-659-2016

Webster-Brown, J. G., Hawes, I., Jungblut, A. D., Wood, S. A., and Christenson, H. K. (2015). The effects of entombment on water chemistry and bacterial assemblages in closed cryoconite holes on Antarctic glaciers. *FEMS Microbiol. Ecol.* 91:fiv144. doi: 10.1093/femsec/fiv144

Weisleitner, K., Perras, A. K., Unterberger, S. H., Moissl-Eichinger, C., Andersen, D. T., and Sattler, B. (2020). Cryoconite hole location in East-Antarctic Untersee oasis shapes physical and biological diversity. *Front. Microbiol.* 11:1165. doi: 10.3389/fmicb.2020.01165

Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis. Available at: https://ggplot2.tidyverse.org (Accessed October 25, 2021).

Williamson, C. J., Cook, J., Tedstone, A., Yallop, M., McCutcheon, J., Poniecka, E., et al. (2020). Algal photophysiology drives darkening and melt of the Greenland ice sheet. *Proc. Natl. Acad. Sci. U. S. A.* 117, 5694–5705. doi: 10.1073/pnas.1918412117

Xin, M.-X., and Zhou, P.-J. (2007). Mrakia psychrophila sp. nov., a new species isolated from Antarctic soil. *J Zhejiang Univ Sci B* 8, 260–265. doi: 10.1631/jzus.2007.B0260

Yallop, M. L., Anesio, A. M., Perkins, R. G., Cook, J., Telling, J., Fagan, D., et al. (2012). Photophysiology and albedo-changing potential of the ice algal community on the surface of the Greenland ice sheet. *ISME J.* 6, 2302–2313. doi: 10.1038/ismej.2012.107

Zawierucha, K., Kolicka, M., Takeuchi, N., and Kaczmarek, Ł. (2015). What animals can live in cryoconite holes? A faunal review. *J. Zool.* 295, 159–169. doi: 10.1111/jzo.12195

Zawierucha, K., Porazinska, D. L., Ficetola, G. F., Ambrosini, R., Baccolo, G., Buda, J., et al. (2021). A hole in the nematosphere: tardigrades and rotifers dominate the cryoconite hole environment, whereas nematodes are missing. *J. Zool.* 313, 18–36. doi: 10.1111/jzo.12832

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

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 29 November 2021
doi: 10.3389/fmicb.2021.781904

Spatial and Annual Variation in Microbial Abundance, Community Composition, and Diversity Associated With Alpine Surface Snow

Lucas Fillinger[1], Kerstin Hürkamp[2], Christine Stumpp[3†], Nina Weber[3], Dominik Forster[3†], Bela Hausmann[4,5], Lotta Schultz[1] and Christian Griebler[3*†]*

[1] Department of Functional and Evolutionary Ecology, University of Vienna, Vienna, Austria, [2] Institute of Radiation Medicine, Helmholtz Zentrum München, Neuherberg, Germany, [3] Institute of Groundwater Ecology, Helmholtz Zentrum München, Neuherberg, Germany, [4] Joint Microbiome Facility of the Medical University of Vienna and the University of Vienna, Vienna, Austria, [5] Department of Laboratory Medicine, Medical University of Vienna, Vienna, Austria

Understanding microbial community dynamics in the alpine cryosphere is an important step toward assessing climate change impacts on these fragile ecosystems and meltwater-fed environments downstream. In this study, we analyzed microbial community composition, variation in community alpha and beta diversity, and the number of prokaryotic cells and virus-like particles (VLP) in seasonal snowpack from two consecutive years at three high altitude mountain summits along a longitudinal transect across the European Alps. Numbers of prokaryotic cells and VLP both ranged around 10^4 and 10^5 per mL of snow meltwater on average, with variation generally within one order of magnitude between sites and years. VLP-to-prokaryotic cell ratios spanned two orders of magnitude, with median values close to 1, and little variation between sites and years in the majority of cases. Estimates of microbial community alpha diversity inferred from Hill numbers revealed low contributions of common and abundant microbial taxa to the total taxon richness, and thus low community evenness. Similar to prokaryotic cell and VLP numbers, differences in alpha diversity between years and sites were generally relatively modest. In contrast, community composition displayed strong variation between sites and especially between years. Analyses of taxonomic and phylogenetic community composition showed that differences between sites within years were mainly characterized by changes in abundances of microbial taxa from similar phylogenetic clades, whereas shifts between years were due to significant phylogenetic turnover. Our findings on the spatiotemporal dynamics and magnitude of variation of microbial abundances, community diversity, and composition in surface snow may help define baseline levels to assess future impacts of climate change on the alpine cryosphere.

Keywords: alpine cryobiosphere, microbial biogeography, virus-like particles, VLP, European Alps, Jungfraujoch, Sonnblick, Zugspitze

OPEN ACCESS

Edited by:
Jérôme Comte, Institut National de la Recherche Scientifique, Université du Québec, Canada

Reviewed by:
Trinity L. Hamilton, University of Minnesota Twin Cities, United States
Mincheol Kim, Korea Polar Research Institute, South Korea

****Correspondence:***
Lucas Fillinger
lucas.fillinger@univie.ac.at
Christian Griebler
christian.griebler@univie.ac.at

†Present address:
Christine Stumpp, Institute for Soil Physics and Rural Water Management, University of Natural Resources and Life Sciences, Vienna, Austria
Dominik Forster, Institute of Immunology and Genetics, Kaiserslautern, Germany
Christian Griebler, Department of Functional and Evolutionary Ecology, University of Vienna, Vienna, Austria

Specialty section:
This article was submitted to Extreme Microbiology, a section of the journal Frontiers in Microbiology

Received: *23 September 2021*
Accepted: *08 November 2021*
Published: *29 November 2021*

Citation:
Fillinger L, Hürkamp K, Stumpp C, Weber N, Forster D, Hausmann B, Schultz L and Griebler C (2021) Spatial and Annual Variation in Microbial Abundance, Community Composition, and Diversity Associated With Alpine Surface Snow. Front. Microbiol. 12:781904. doi: 10.3389/fmicb.2021.781904

INTRODUCTION

The alpine cryosphere is among the ecosystems that are the most affected by climate change. In regions like the European Alps, surface air temperatures have increased at a faster rate over the past decades compared to the global average (Hock et al., 2019). The total glacier volume in the European Alps is predicted to decrease by 63–94% compared to present-day levels by the end of this century (Zekollari et al., 2019). Snow cover duration and snow mass are projected to decline by 30–80% and 10–40%, respectively, over the same time period, while winter precipitation extremes will occur more frequently. Overall, these changes are expected to alter the amount, quality, and seasonality of runoff (Hock et al., 2019). As mountain snow and ice are important sources of freshwater for downstream environments like rivers and lakes, these changes can have far reaching consequences for biogeochemical processes and biodiversity on a larger scale (Boetius et al., 2015; Hotaling et al., 2017; Hock et al., 2019; Ren et al., 2019; Zhou et al., 2019; Elser et al., 2020).

Microbial communities play a key role in the nutrient composition of mountain surface ice and snowpack. Despite extreme conditions characterized by low temperatures and intense UV radiation, microbial growth and metabolic activity have repeatedly been documented in such environments (Carpenter et al., 2000; Anesio et al., 2010; Stibal et al., 2012b; Lopatina et al., 2013), supporting primary production via fixation of atmospheric carbon and nitrogen, and heterotrophic processing of nutrients produced *in situ* and deposited from the atmosphere as dust (Stibal et al., 2012a; Maccario et al., 2015, 2019; Antony et al., 2016; Franzetti et al., 2016; Ren et al., 2019; Zhou et al., 2019). Additionally, viruses have been proposed to be an important factor in the regulation of nutrient cycles by modulating host metabolism and releasing labile organic matter from microbial cells through lysis (Anesio and Bellas, 2011; Zhong et al., 2021). Most of these observations stressing the importance of viruses in snow and ice come from cryoconite holes of arctic glaciers (Anesio et al., 2007; Anesio and Bellas, 2011; Maccario et al., 2015), where viral lysis has been estimated to release about one third of the carbon from microbial production (Rassner et al., 2016). Organic matter produced and processed in the alpine cryosphere is transported downstream with meltwater, together with microbial cells and viruses, which can be infectious also for cells in meltwater-fed aquatic environments (Anesio et al., 2007). Hence, the functioning and activity of microorganisms as well as microbial and viral abundances in the alpine cryosphere can have an impact on nutrient levels and stoichiometries in other environments downstream (Ren et al., 2019; Zhou et al., 2019).

Characterizing microbial and viral abundances as well as microbial community composition and diversity associated with surface snow and ice therefore is an important step toward a better understanding of current and future changes in the alpine cryosphere, and concomitant effects on meltwater-fed environments. Previous studies have identified atmospheric deposition as the main mechanism that seeds mountain snow surfaces with microbial cells (Boetius et al., 2015). The frequency and intensity of this deposition is spatiotemporally variable, for instance depending on the occurrence of dust events, which in turn can result in variation of microbial cell abundances, and cause significant differences in community composition across space and time (Segawa et al., 2005; Chuvochina et al., 2011; Meola et al., 2015; Els et al., 2019; Courville et al., 2020). Following deposition, microbial taxa are subject to selection imposed by environmental factors such as high solar radiation as well as low temperatures and nutrient levels, which only allow a certain fraction of the initially deposited taxa to persist and remain active (Segawa et al., 2005; Chuvochina et al., 2011; Lopatina et al., 2013; Els et al., 2019, 2020; Maccario et al., 2019).

Knowledge on dynamics and the magnitude of variation in microbial abundance, community composition, and diversity is necessary for establishing a baseline of natural variation against which future changes can be compared in order to assess climate change impacts. Recording spatiotemporally data is an important step toward achieving this goal. However, so far, studies on microbial communities associated with mountain surface snow and ice have either focused on temporal changes at single locations or small spatial scales (Segawa et al., 2005; Chuvochina et al., 2011; Pittino et al., 2018; Els et al., 2019, 2020), or differences between locations at single time points (Wunderlin et al., 2016; Azzoni et al., 2018; Brown and Jumpponen, 2019), whereas data covering both spatial and temporal variation are scarce. As a first initiative to fill this gap, we collected samples across depth profiles of seasonal snowpack at the end of snow accumulation periods in two consecutive years (2015 and 2016) from three mountain summits along a longitudinal transect across the European Alps (i.e., Jungfraujoch, Switzerland; Zugspitze, Germany; Sonnblick, Austria) (**Figure 1**). In this study, we report spatiotemporal variation in abundances of prokaryotic cells and virus-like particles (VLP) measured by flow cytometry in addition to differences in microbial community composition and diversity inferred from 16S rRNA gene amplicon sequence variants (ASVs).

MATERIALS AND METHODS

Site Descriptions and Sample Collection

The Jungfraujoch site was located in the Swiss Alps (46.5475°N, 7.9851°E, 3572 m.a.s.l.) directly south of the High Altitude Research Station Jungfraujoch in the firn of the Aletsch glacier. Mean air temperatures are −7.2°C with solid precipitation and strong winds throughout the year. Sampling campaigns were conducted on June 25, 2015, and June 1, 2016. Mean snow water equivalents (SWE) in 2015 and 2016 were 1994 mm and 1408.6 mm, respectively.

The Sonnblick is one of the highest mountains in the Austrian Alps (47.0536°N, 12.9578°E, 3106 m.a.s.l.) belonging to the Goldberg mountain range. Mean air temperatures are −5.1°C with an average precipitation of 2263 mm. The area is strongly wind exposed and snowfall occurs throughout the whole year. The sampling location was situated at the northern margin of the Kleinfleißkees glacier on a small rock outcrop, directly south of the mountain summit close to the Sonnblick Observatory research station. Sampling

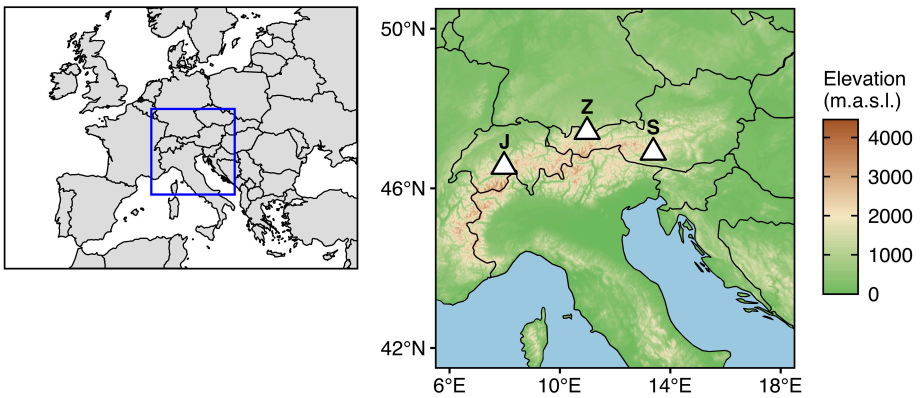

FIGURE 1 | Geographic locations of the study sites (*J*: Jungfraujoch, 3572 m.a.s.l.; *Z*: Zugspitze, 2420 m.a.s.l.; *S*: Sonnblick, 3106 m.a.s.l.). Shapefiles for country borders were retrieved from Natural Earth (https://www.naturalearthdata.com/). Topographic data were retrieved from ETOPO1 Global Relief Model (Amante and Eakins, 2009) using metR (v. 0.9.2; Campitelli, 2021).

campaigns were conducted on July 2, 2015, and June 8, 2016. Mean SWE in 2015 and 2016 were 1587 mm and 2125 mm, respectively.

The Zugspitze situated in the southern Bavarian Alps is Germany's highest mountain (47.4211°N, 10.9854°E, 2420 m.a.s.l.). Underneath the summit and surrounded by a horseshoe-shaped ridge of further summits, a flat karstic plateau serves as a popular skiing area and drains the whole Zugspitze region to the Partnach river in the east. The area is strongly wind exposed and receives intense solar radiation due to its south-eastern exposition. The average amount of precipitation is 2085 mm with liquid precipitation only during <10 weeks in summer. Mean air temperatures are −4.8°C. The sampling location was selected in the central western part of the plateau in a fenced area used by the Bavarian Avalanche Warning Service and situated close to the Environmental Research Station Schneefernerhaus (further details on the site are described in Hürkamp et al., 2019). Sampling campaigns were conducted on May 13, 2015, and May 3, 2016. Mean SWE in 2015 and 2016 were 1222 mm and 1521 mm, respectively.

The sampling locations at all three sites were selected to provide a flat rock surface to avoid potential interflow on sloping snow layer boundaries. The sampling location at each study site was the same in both years. To ensure that the sampled snow profiles covered the whole seasonal snowpack, vertical snow pits at Zugspitze and Sonnblick were dug all the way to the bottom on rock outcrops that had been free of snow cover during the previous summer season. At Jungfraujoch, the profiles were confined by a massive ice layer at the bottom, preventing further digging. Descriptions of the snow structures at the different layers according to Fierz et al. (2009) are listed for all sites in **Supplementary Tables 1–6**. Snow density and SWE were measured directly in the field by determination of snow mass of a defined volume using a snow cylinder.

Samples for the quantification of prokaryotic cells and VLP were collected separately into sterile 50 mL Falcon tubes at ∼10 cm intervals, taking into account differences in snow texture. Samples for cell and VLP counting were fixed on-site with 2.5% v/v and 0.5% v/v glutardialdehyde (final concentration), respectively. Vertical slot samples were taken over the total depth of a snow layer or divided into more than one sample for layers >20 cm. For layers with a thickness <5 cm, for example thin ice layers as well as the entire profile at Zugspitze in 2015, horizontal slots of about 10 cm in length were sampled at the center of each layer parallel to the snow surface. Samples for stable isotope analyses were collected into 50 mL polyethylene bottles at the same positions as samples for prokaryotic cell and VLP counting. The bottles were kept closed during snow melting to prevent isotope fractionation for example due to evaporation. Slot samples for DNA extraction were collected from selected layers of up to 100 cm in sterile 2 L wide neck polyethylene bottles. In general, the sample volume was reduced to ∼10% of the initially sampled snow volume after melting. Samples were kept cooled in polystyrene boxes for transport to the lab. Samples for prokaryotic cell counting and isotope analyses were kept at 4°C until analysis; samples for VLP counting and DNA extraction were stored at −20°C.

Determination of Prokaryotic Cell and Virus-Like Particle Numbers by Flow Cytometry

Numbers of prokaryotic cells and VLP were measured separately in 500 μL snow meltwater aliquots. Samples were mixed with 500 μL suspension of reference beads as internal standard (Trucount Tubes; Becton-Dickinson, Franklin Lakes, NJ, United States). Fluorescent nucleic acid staining was used to distinguish cells and VLP from inorganic particles. Prokaryotic cells were stained with 1x (final concentration) SYBR Green I (Invitrogen, Darmstadt, Germany) during incubation at room temperature for 10 min in the dark. VLP were stained with 1x (final concentration) SYBR Gold (Invitrogen) during incubation at 80°C for 5 min followed by 10 min at room temperature in the dark. Cells and VLP were counted in duplicate on a FC500 CYTOMICS flow cytometer equipped with a 488 nm argon ion laser (Beckman Coulter, Brea, CA, United States).

Determination of Stable Isotope Ratios

Comparing stable isotopes of oxygen ($\delta^{18}O$) and hydrogen ($\delta^{2}H$) in snow with the isotopic composition in precipitation provides information on potential transport processes within the accumulated snowpack and the temporal origin of the snow (Moser and Stichler, 1974). For example, lower isotope ratios in snow are associated with precipitation origination from colder temperatures or months. Additionally, freezing, thawing, water vapor exchange, and sublimation processes can influence the isotopic composition of the snow (Stichler, 1987; Gat, 1996).

$\delta^{18}O$ and $\delta^{2}H$ were analyzed in melted snow samples using laser spectroscopy (Picarro L2120-i and Picarro L2140-i, Picarro, Santa Clara, CA, United States). Stable isotope compositions in precipitation and snow are presented in delta notation as δ-value (‰). Precision of the instrument (1σ) was better than 0.6‰ and 0.15‰ for $\delta^{2}H$ and $\delta^{18}O$, respectively. For comparisons of the isotopic composition of the snow profiles with precipitation, monthly or biweekly data were taken from nearby national monitoring sites at Zugspitze (own data), Jungfraujoch (personal communication: M. Leuenberger, High Altitude Research Station Jungfraujoch, February 9, 2016), and Feuerkogel as the closest station to Sonnblick (Environment Agency Austria [EAA], 2016; **Supplementary Table 7**). Elevation gradients were used to account for differences in the isotope ratios due to different altitudes of isotope monitoring sites and snow profiles. Winter precipitation data were normalized to minimize differences in the amounts of precipitation and snow, and SWE (Hürkamp et al., 2019). Snow accumulation periods for different depth zones were approximated from the comparison of the normalized data. Larger, unexpected differences between isotopic composition in precipitation and snow profiles can be caused by redistribution of the snowpack, for example caused by wind drift. Only $\delta^{18}O$ values are reported here due to the linear relationship between $\delta^{18}O$ and $\delta^{2}H$ and because fractionation processes due to sublimation are of minor importance as shown for Zugspitze by Hürkamp et al. (2019).

DNA Extraction

Snow meltwater samples were pooled from several adjacent layers to an equal volume of 2 L per sample. Cells for DNA extraction were collected from snow meltwater on 0.2 μm polycarbonate membrane filters (Merck-Millipore, Carrigtwohill, Ireland). DNA was extracted from the filters using bead beating and phenol–chloroform extraction as described by Pilloni et al. (2012). DNA was precipitated with 30% w/v polyethylene glycol containing 3 M NaCl for 2 h at 4°C followed by centrifugation at 20,000 × *g* for 30 min. The DNA pellet was washed twice with 150 μL 80% v/v ice-cold ethanol and resuspended in 15 μL EB buffer (Qiagen, Hilden, Germany). DNA was stored at −80°C.

16S rRNA Gene Amplicon Sequencing and Data Processing

The V4 region of the 16S rRNA gene was amplified using a unique dual barcoding two-step PCR approach (UDB-H12) as described by Pjevac et al. (2021) with primers 515F (Parada et al., 2016) and 806R (Apprill et al., 2015). Amplicons were sequenced in paired-end mode (2 × 300 bp) on a MiSeq platform (Illumina, San Diego, CA, United States) at the Joint Microbiome Facility of the Medical University of Vienna and the University of Vienna under project ID JMF-2006-1. Sequence data were processed in R (version 3.6.1; R Core Team, 2021) using DADA2 (v. 1.14.1; Callahan et al., 2016a) following the workflow by Callahan et al. (2016b). ASVs were inferred across all samples in pooled mode (further details on sequence trimming and settings for quality filtering are described in Pjevac et al., 2021). Taxonomic assignment was done by mapping ASV sequences against the SILVA SSU reference database (release 138; McLaren, 2020) using the RDP Naive Bayesian Classifier implemented in the DADA2 "assignTaxonomy" function with default confidence settings (i.e., minimum bootstrap confidence set to 50). ASVs classified as unclassified domain, eukaryotes, mitochondria, or chloroplasts were discarded from the dataset, as well as samples with less than 100 sequence reads. The final ASV table used for downstream analyses contained 427 ASVs with numbers of reads per sample ranging between a minimum and maximum of 214 and 9172, respectively (median: 1549; mean: 2773.2). For phylogenetic community analyses, sequences were aligned using the "AlignSeqs" function of the DECIPHER package (v. 2.16.1; Wright, 2016). A phylogenetic tree was constructed from the alignment using FastTree with a GTR + CAT model (v. 2.1.11; Price et al., 2010), and subsequently midpoint-rooted in R using the phytools package (v. 0.7.80; Revell, 2012). The 16S rRNA gene amplicon sequence data are publicly available at the NCBI Sequence Read Archive (accession number PRJNA756880).

Data Analysis

All analyses were done in R (v. 4.0.2). Differences in numbers of prokaryotic cells, VLP, and VLP-to-prokaryotic cell ratios between sites within years were tested using Kruskal–Wallis tests and Dunn's *post hoc* test for pairwise comparisons with Bonferroni *p*-value adjustment for multiple testing using the PMCMRplus package (v. 1.9.0; Pohlert, 2021). Differences between years within sites were tested using Wilcoxon rank sum tests.

Taxonomic and phylogenetic microbial community alpha diversity were estimated based on Hill numbers of orders $q = 0$ (taxon richness/Faith's phylogenetic diversity), $q = 1$ (exponential of Shannon entropy/phylogenetic entropy), and $q = 2$ (Gini-Simpson index/Rao's Q) inferred from interpolation and extrapolation of rarefaction curves (Chao et al., 2014a,b). To account for differences in the number of sequence reads between samples, diversity was calculated for all samples at an equal coverage level of 92.8%, which was the highest possible level for this dataset based on twice the smallest reference sample size (Chao et al., 2014b). Taxonomic and phylogenetic Hill numbers were calculated using the iNEXT (v. 2.0.20; Chao et al., 2014b) and iNextPD (v. 0.3.1; Hsieh and Chao, 2017) packages, respectively.

For calculating microbial community beta diversity, ASV abundances were converted to relative abundances within samples. Taxonomic and phylogenetic beta diversity were calculated as Bray–Curtis dissimilarity and relative abundance-weighted β-mean pairwise distance (β-MPD) using the vegan

(v. 2.5-7; Oksanen et al., 2020) and picante (v. 1.8.2; Kembel et al., 2010) packages, respectively. PCoA ordinations for visualizing differences in community composition were generated using the ape package with Lingoes correction for negative eigenvalues (v. 5.4.1; Paradis and Schliep, 2019). Differences in community composition between sites and years were assessed using PERMANOVA ("adonis2" function, vegan) with 999 permutations for significance testing.

We took precautions to ensure that results of the beta diversity analyses obtained from relative abundance data were not biased by the variation in the number of sequence reads between samples. Bray–Curtis dissimilarity and β-MPD were additionally calculated on 100 rarefied ASV abundance tables that were generated in independent iterations of random subsampling without replacement ("rrarefy" function, vegan) to 214 reads per samples (i.e., the smallest number of reads per sample across all samples in the dataset). For both beta diversity measures, we found strong agreement between the values obtained from non-rarefied relative abundances and those obtained from rarefied abundance data based on Mantel correlation analyses and Gaussian generalized linear regression (**Supplementary Figure 1**). Furthermore, all PERMANOVA tests were additionally performed on each of the 100 Bray–Curtis dissimilarity and β-MPD matrices, respectively, calculated with the rarefied abundances. In all cases, results regarding effect sizes and significance of explanatory variables were highly similar and led to identical conclusions as those obtained from tests based on non-rarefied relative abundance data (**Supplementary Figures 2–5**). Finally, we directly tested for the effect of differences in the number sequence reads on the observed beta diversity patterns obtained from relative abundance data by repeating the PERMANOVA tests including the number of reads per sample as additional explanatory variable, following the recommendation by Weiss et al. (2017). The number of sequence reads did not have a significant effect in any of the analyses after controlling for site or year (all partial $R^2 < 0.08$; all $p > 0.18$; results not shown), and were therefore not considered further for the final analyses presented below.

RESULTS

Snow Profile Characteristics and Dynamics of Prokaryotic Cell and Virus-Like Particle Numbers

The snowpack at Jungfraujoch in 2015 was 377 cm deep and was confined by a massive ice layer at the bottom. Near isothermal conditions with temperatures around 0°C across the entire profile indicated the end of the snowpack ripening and beginning of the melting phase. Several thin ice layers occurred throughout the depth profile, resulting in snow densities of 353–588 kg m^{-3} (**Figure 2**). A sandy layer was observed at ∼260 cm above ground (**Supplementary Table 3**). In 2016, snowpack height was 440 cm, with snow densities ranging between 341 and 544 kg m^{-3}. Sandy layers were observed at heights of ∼20 cm, ∼130 cm, and ∼190 cm above ground (**Supplementary Table 6**).

At Sonnblick, snowpack height in 2015 was 404 cm with several thin ice layers in between. Snow densities were relatively high, especially in the top layers, ranging between 507 and 682 kg m^{-3}. A mineral-rich, brown layer was visible at a height of ∼70 cm, which likely was an old surface layer possibly derived from the previous summer season (**Supplementary Table 2**). In 2016, the snowpack height was 357 cm and the snow was less compacted. The top 30 cm consisted of fresh snow with a lower density compared to deeper layers. The overall snow density ranged between 251 and 554 kg m^{-3}. A visible sand layer was observed at ∼220 cm above ground, which was proven to originate from a Sahara dust event that occurred on April 4, 2016 (**Supplementary Table 5**).

At Zugspitze, snowpack height in 2015 was 233 cm with snow densities ranging between 440 and 567 kg m^{-3}. In 2016, snowpack height was 270 cm. The profile consisted of a less dense, fresh snow layer of ∼50 cm at the top; overall snow densities across the whole profile ranged between 214 and 589 kg m^{-3}. A sandy layer was observed at ∼180 cm above ground that derived from a Sahara dust event occurring on April 6, 2016 (**Supplementary Tables 1**, **4**).

Numbers of prokaryotic cells and VLP showed significant variation between sites and years, albeit over a relatively modest range of one order of magnitude in general (**Figure 2**). Across all samples, the median number of prokaryotic cells was 7.41×10^3 per mL of snow meltwater with an interquartile range (IQR) of 9.98×10^3. For both years, we found significant differences in cell numbers between sites (Kruskal–Wallis rank sum tests; 2015: $H = 20.2$, $p < 0.0001$; 2016: $H = 16.5$, $p = 0.0003$). Closer inspection of these differences by pairwise comparisons using Dunn's *post hoc* tests showed that in 2015 prokaryotic cell numbers at Zugspitze (median = 7.35×10^3 cells mL^{-1}, IQR = 2.82×10^3) were significantly lower compared to Jungfraujoch (median = 1.19×10^4 cells mL^{-1}, IQR = 1.48×10^4, $p = 0.011$) and Sonnblick (median = 1.77×10^4 cells mL^{-1}, IQR = 3.91×10^3, $p < 0.0001$). In 2016, cell numbers were significantly lower at Jungfraujoch (median = 2.98×10^3 cells mL^{-1}, IQR = 1.64×10^3) compared to Sonnblick (median = 5.96×10^3 cells mL^{-1}, IQR = 3.13×10^3, $p = 0.0037$) and Zugspitze (median = 6.81×10^3 cells mL^{-1}, IQR = 6.05×10^3, $p = 0.0005$). Within sites, cell numbers were significantly higher in 2015 compared to 2016 at Jungfraujoch (Wilcoxon rank sum test; $W = 724$, $p << 0.0001$) and Sonnblick ($W = 646.5$, $p << 0.0001$).

Numbers of VLP were in a similar range as prokaryotic cell numbers (overall median = 9.90×10^3 VLP mL^{-1}, IQR = 1.00×10^4) and showed significant variation between sites within years (Kruskal-Wallis rank sum tests; 2015: $H = 36.2$, $p < 0.0001$; 2016: $H = 10.5$, $p = 0.0053$). Multiple pairwise comparisons showed that these differences were due to significantly higher VLP numbers at Sonnblick in 2015 (median = 4.66×10^4 VLP mL^{-1}, IQR = 4.49×10^4) compared to Jungfraujoch (median = 1.17×10^4 VLP mL^{-1}, IQR = 1.22×10^4, $p = 0.0002$) and Zugspitze (median = 6.96×10^3 VLP mL^{-1}, IQR = 4.28×10^3, $p << 0.0001$), while in 2016 VLP numbers were significantly lower at Zugspitze (median = 6.85×10^3 VLP mL^{-1}, IQR = 2.79×10^2) compared to Jungfraujoch

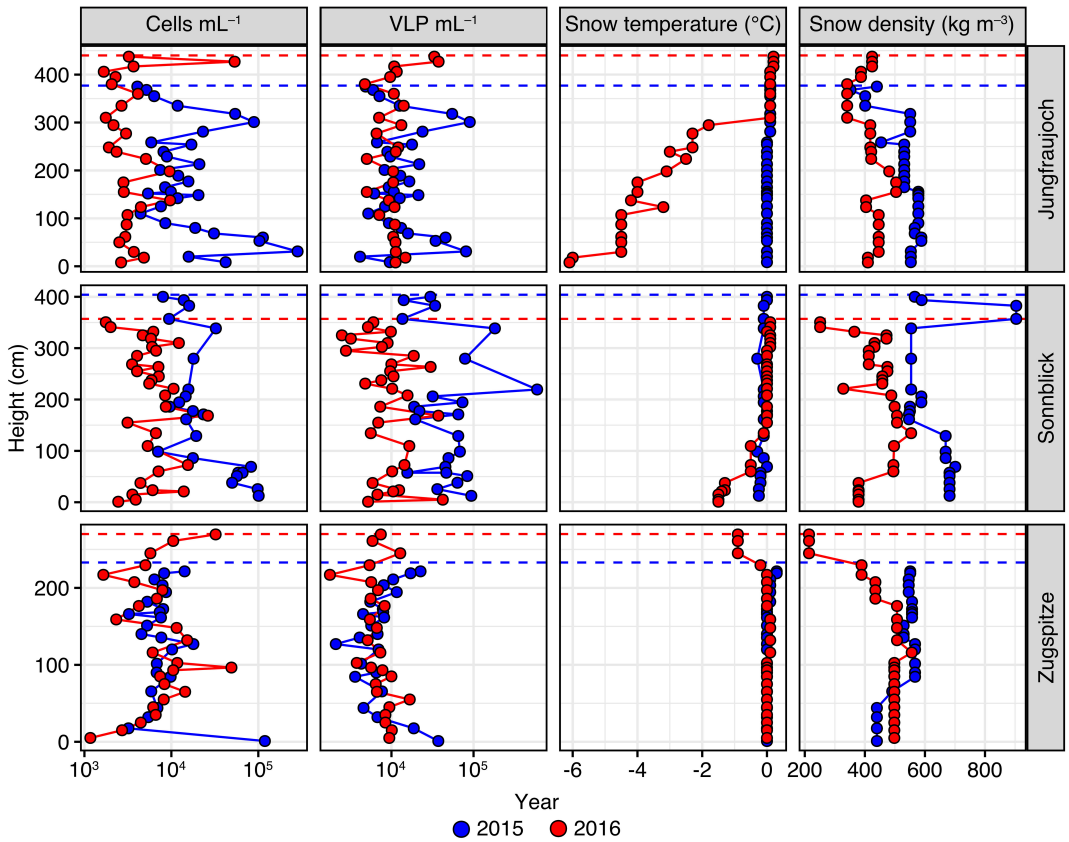

FIGURE 2 | Numbers of prokaryotic cells and VLP per mL snow meltwater, in addition to snow temperature and snow density measured across the snowpack profiles per site and year. Height of the snow layer is given in cm above the bottom of the snowpack. Dashed lines represent the maximum snowpack height for each year.

(median = 1.09×10^4 VLP mL^{-1}, IQR = 2.14×10^3, $p = 0.0036$) but not Sonnblick (median = 9.31×10^3 VLP mL^{-1}, IQR = 6.03×10^3, $p = 0.24$). Significant differences within sites between years were only found for Sonnblick (Wilcoxon rank sum test; $W = 652$, $p << 0.0001$).

Although VLP and prokaryotic cell numbers also varied across the depth of the snow profiles, there was no significant monotonic relationship between either parameter and height of the snow layer above bottom across all samples (cell numbers: Spearman's $rho = -0.140$, $p = 0.0810$; VLP: $rho = -0.150$, $p = 0.061$). When tested within sites and years, significantly decreasing cell numbers with snow layer height were found only at Jungfraujoch ($rho = -0.402$, $p = 0.0274$) and Sonnblick ($rho = -0.712$, $p = 0.0001$) in 2015 (all others: $|rho| < 0.14$, $p > 0.37$). VLP numbers showed a significant decrease with snow layer height only at Zugspitze in 2016 ($rho = -0.482$, $p = 0.0147$; all others: $|rho| < 0.31$, $p > 0.096$).

Numbers of prokaryotic cells and VLP were moderately positively correlated only across all samples without distinguishing between sites and years ($rho = 0.472$, $p << 0.0001$). However, this relationship was not generally observed when considering the data within sites and years separately, where significant positive correlations were observed only at Jungfraujoch in 2015 ($rho = 0.880$, $p << 0.0001$) and Sonnblick in 2016 ($rho = 0.389$, $p = 0.034$) (**Supplementary Figure 6**).

Virus-like particle-to-prokaryotic cell ratios varied over two orders of magnitude (minimum: 0.12, maximum: 37.90) with an overall median of 1.24 (IQR = 1.37) (**Figure 3**). Differences within sites between years were observed only at Jungfraujoch ($W = 77$, $p << 0.0001$; all others: $W > 270$, $p > 0.35$). Within years, VLP-to-cell ratios were significantly higher at Sonnblick in 2015 and Jungfraujoch in 2016 compared to the other two respective sites ($p < 0.003$). We further analyzed correlations between VLP-to-cell ratios and numbers of prokaryotic cells and VLP, respectively, to test whether changes in VLP-to-cell ratios were caused by changes in either only cell numbers or VLP alone, or concomitant changes of both. In the majority of cases, VLP-to-cell ratios were significantly positively correlated with VLP numbers, and significantly negatively correlated with prokaryotic cell numbers. The only exception was observed at Jungfraujoch in 2015, where both numbers of prokaryotic cells and VLP were significantly negatively correlated with VLP-to-cell ratios (**Supplementary Figure 7**).

In summary, we found significant albeit relatively small variation (within one order of magnitude) in the numbers of

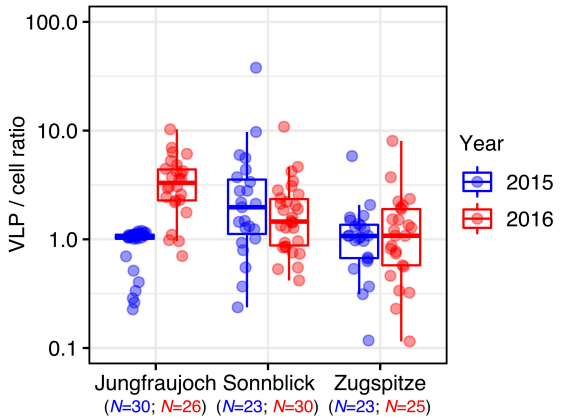

FIGURE 3 | VLP-to-prokaryotic cell ratios per site and year.

prokaryotic cells, VLP, and VLP-to-cell ratios between sites and years. However, we could not detect consistent patterns indicating generally increased numbers within a particular year or at any of the given sites. Moreover, occasional peaks in prokaryotic cell and VLP numbers were not always consistent with observed sandy layers in the snow profiles.

Stable Isotope Ratios in Snow Profiles

To estimate approximate time points of snow accumulation (**Figure 4**), stable isotopes in snow were compared to isotope ratios measured in precipitation during snow accumulation periods (**Supplementary Table 7**). This was only possible by also considering the amount of snow in the snowpack calculated from individual depths and snow density (**Figure 2**), and the amount of precipitation (**Supplementary Table 7**). Overall, the annual snowpack at the three sites covered snow accumulation periods of about 6–10 months. Time differences between the oldest bottom layers and the youngest top layers from the previous season were approximately 4–8 months. It is worth mentioning that those are rough estimates, due to the coarse monthly or biweekly sampling of isotopes in precipitation and differences in snow and precipitation amounts. The dating of two snow layers in the snow profiles at Sonnblick and Zugspitze in 2016 could be verified by the presence of Saharan dust. The dust events occurred from April 4 to 6, 2016, (**Supplementary Tables 4, S5**) as proven by aerosol measurements of the German Federal Environment Agency (personal communication by L. Ries). Nevertheless, most of the isotope patterns from precipitation resembled the distribution in the snow profiles except for the middle and upper parts of profiles in 2016 from Jungfraujoch and Sonnblick, respectively. The isotope ratios between 0 and 100 cm and between 100 and 300 cm in the snowpack from Jungfraujoch in 2016 were about 2.5‰ lower and 5‰ greater, respectively, than expected from isotope ratios in precipitation, indicating either snow redistribution or melting (100–300 cm), and refreezing processes (0–100 cm). At Sonnblick in 2016, there were major differences in the water balance (450 mm more water in the snow profile than in precipitation). When comparing isotope ratios in the snow and precipitation, the difference in the water balance is associated with time points after March 2016 and snow accumulated >165 cm, still having almost similar isotope averages (−14.4‰ in precipitation and −14.1‰ in snow). This discrepancy can be explained by either erroneous precipitation amount analysis, or redistribution of snow due to wind drift.

Dynamics of Microbial Community Composition and Diversity

Microbial communities displayed similar compositions within years across sites (**Figure 5**). In 2015, communities at Jungfraujoch and Sonnblick were dominated by Gammaproteobacteria (mainly *Massilia*, *Polaromonas*) and Bacteroidia (*Ferruginibacter*, *Hymenobacter*). These classes were replaced to a large extent in 2016 by Alphaproteobacteria (*Sphingomonas*, *Methylobacterium-Methylorubrum*, *Acidiphilum*), diverse Actinobacteria, Bacilli (*Staphylococcus*, *Streptococcus*), and diverse Cyanobacteria (abundances of the most dominant genera within the most abundant classes are shown in **Supplementary Figures 8–13**). In line with this qualitative observation, PCoA ordinations showed stronger clustering of samples by year than by site based on Bray–Curtis dissimilarity and especially β-MPD (**Figure 6**).

We assessed marginal effect sizes and significance of differences between sites and years using PERMANOVA considering only samples from Jungfraujoch and Sonnblick, since samples from both years were not available for the Zugspitze. The analysis showed that differences in taxonomic community composition measured as Bray–Curtis dissimilarity were significant between sites and years, although variation between years was larger [partial R^2 = 0.1781, $F_{(1,17)}$ = 3.67, p = 0.001] than between sites [partial R^2 = 0.094, $F_{(1,17)}$ = 1.94, p = 0.011]. However, differences in phylogenetic community composition based on β-MPD were significant only between years [partial R^2 = 0.2255, $F_{(1,17)}$ = 4.70, p = 0.001] but not between sites [partial R^2 = 0.0543, $F_{(1,17)}$ = 1.13, p = 0.26]. Comparable results were obtained from analyzing differences between all three sites considering only samples from 2016 [Bray–Curtis dissimilarity: R^2 = 0.2522, $F_{(2,13)}$ = 1.86, p = 0.001; β-MPD: R^2 = 0.1822, $F_{(2,13)}$ = 1.23, p = 0.131]. Taken together, the analyses indicate that microbial communities were composed of phylogenetically similar ASVs across sites within years, and that differences between sites were due to shifts in abundances of ASVs within phylogenetically similar clades.

Microbial community alpha diversity inferred from Hill numbers generally decreased with higher orders of q, indicating relatively low contributions of common (q = 1) and abundant ASVs (q = 2) to the total ASVs richness (q = 0) in the communities (**Figure 7**). Across all samples, average diversity was 59.5, 22.0, and 11.1 (minima: 2.6, 2.1, 1.4; maxima: 186.0, 81.9, 39.0) for orders q = 0, q = 1, and q = 2, respectively. Significant differences between sites were only observed in 2016 with 5–8 times higher diversity at Sonnblick compared to Jungfraujoch for all orders of q (Wilcoxon rank sum tests; all W = 0, p < 0.008), while in 2015 no significant differences between sites were observed for

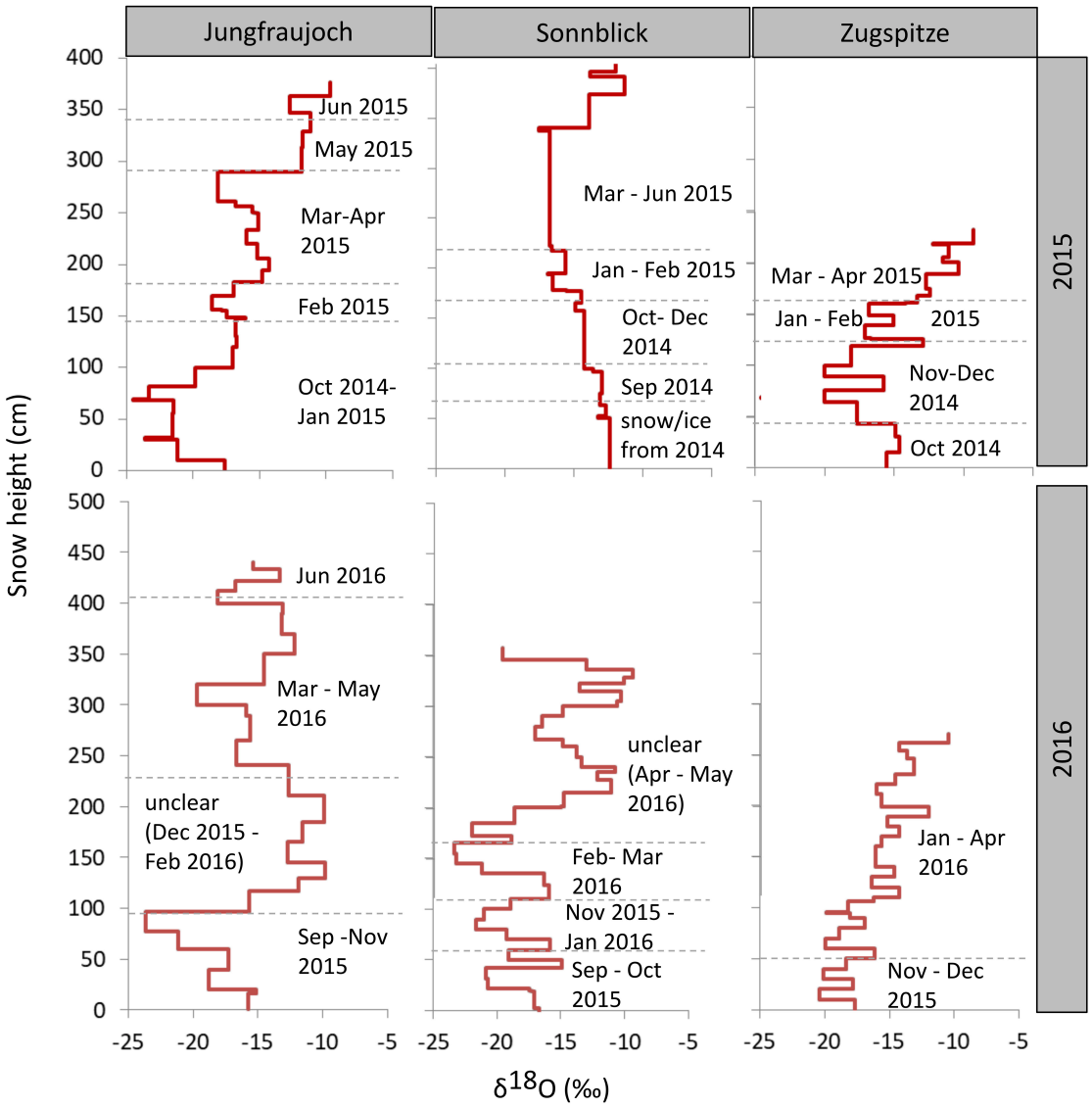

FIGURE 4 | $\delta^{18}O$ values across the snowpack profiles per site and year. Times of accumulation periods were approximated by considering isotope ratios in precipitation during the accumulation period as well as precipitation amount and snow water equivalents. Height of the snow layer is given in cm above the bottom of the snowpack.

any order of q (all $W \geq 1$, $p > 0.05$) (samples from the Zugspitze were not included in these analyses due to the lack of samples from 2015, and the generally large variation and small number of samples for 2016). Although community alpha diversity showed some variation between years within sites, significant year-to-year variation was only observed at Sonnblick for $q = 0$ ($W = 0$, $p = 0.016$; all others: $W \geq 2$, $p > 0.28$) but not for higher orders of q. Similar results were obtained based on phylogenetic community alpha diversity (**Supplementary Figure 14**). Thus, significant variation in community alpha diversity between sites appeared to have affected rare as well as common and abundant ASVs, whereas variation within sites between years was less pronounced, especially considering the more abundant fraction of the microbial communities.

Unfortunately, the relatively coarse pooling of samples that was necessary to ensure sufficient sample volumes and DNA yield did not allow for direct analyses of differences in microbial community composition or diversity in relation to snow layer height, time point of snow deposition inferred from stable isotopes, or snow temperature and density.

DISCUSSION

In this study, we set out to characterize spatiotemporal variation in microbial abundance, community composition, and diversity associated with seasonal surface snow in high altitudes of the European Alps, as a first step toward establishing a

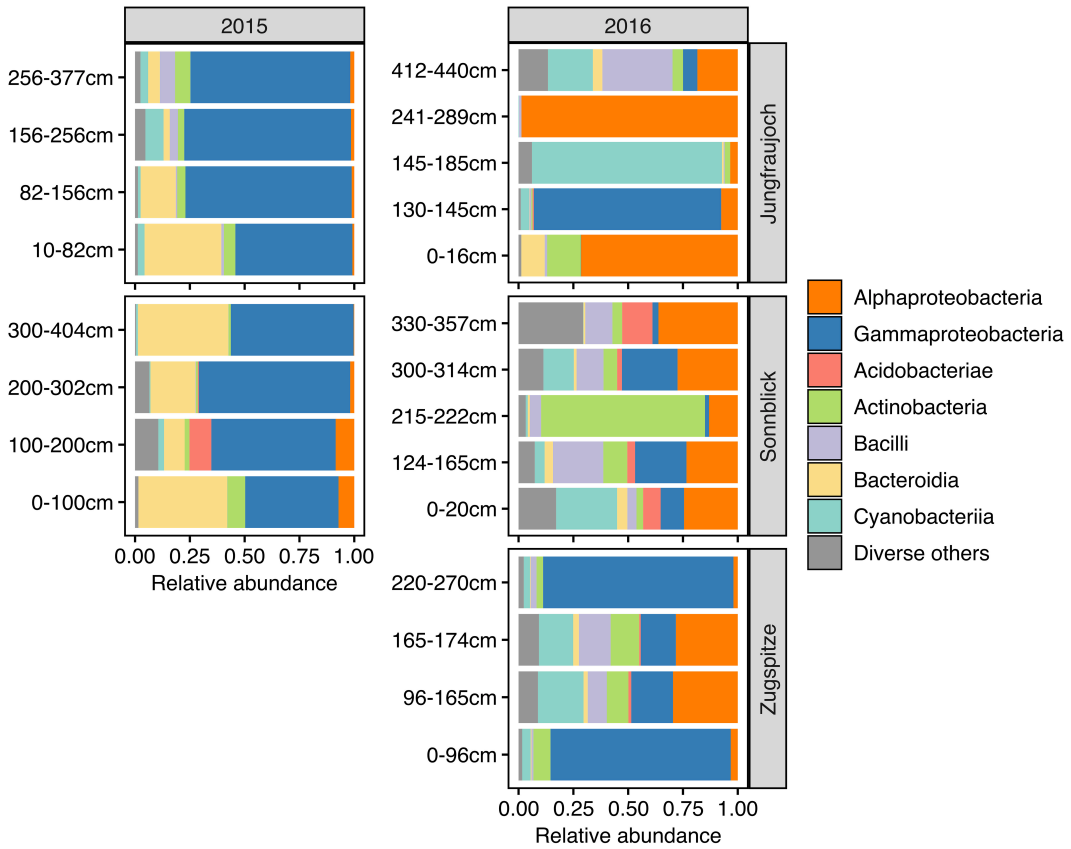

FIGURE 5 | Microbial community composition across snowpack profiles per site and year based on relative abundances summarized at class level. The range of pooled layers used for DNA extraction is indicated in cm above the bottom of the snowpack. For clarity of display, classes with a combined relative abundance <5% across all samples are shown collectively as "Diverse others."

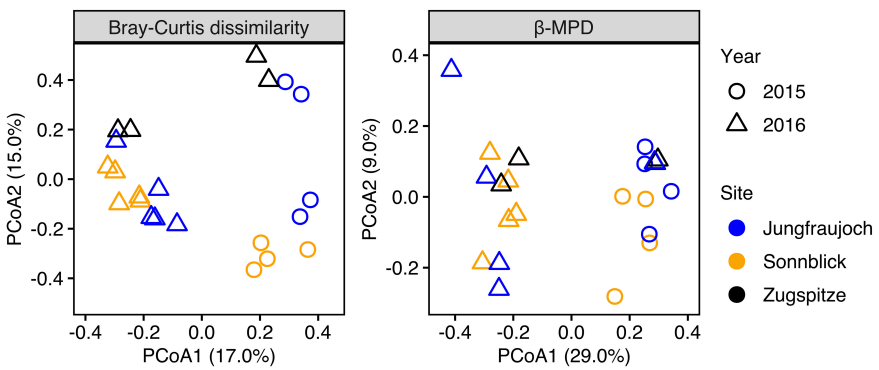

FIGURE 6 | PCoA ordination plots of differences in taxonomic community composition based on Bray–Curtis dissimilarity calculated from relative ASV abundances, and differences in phylogenetic community composition calculated as relative abundance-weighted β-MPD.

baseline for natural variation of microbial communities in these environments. Prokaryotic cells and VLP occurred in similar numbers ranging between 10^4 and 10^5 per mL of snow meltwater on average, with relatively little—albeit occasionally significant—variation between sites and years within one order of magnitude. These numbers are in a similar range as previously reported for alpine surface snow and ice (Segawa et al., 2005; Boetius et al., 2015; Lazzaro et al., 2015), including an earlier study at Jungfraujoch conducted in 2015 (Wunderlin et al., 2016), as well as numbers commonly found in cold, oligotrophic freshwater environments (Wommack and Colwell, 2000). In comparison to glacial cryoconite, these numbers are up to four orders of

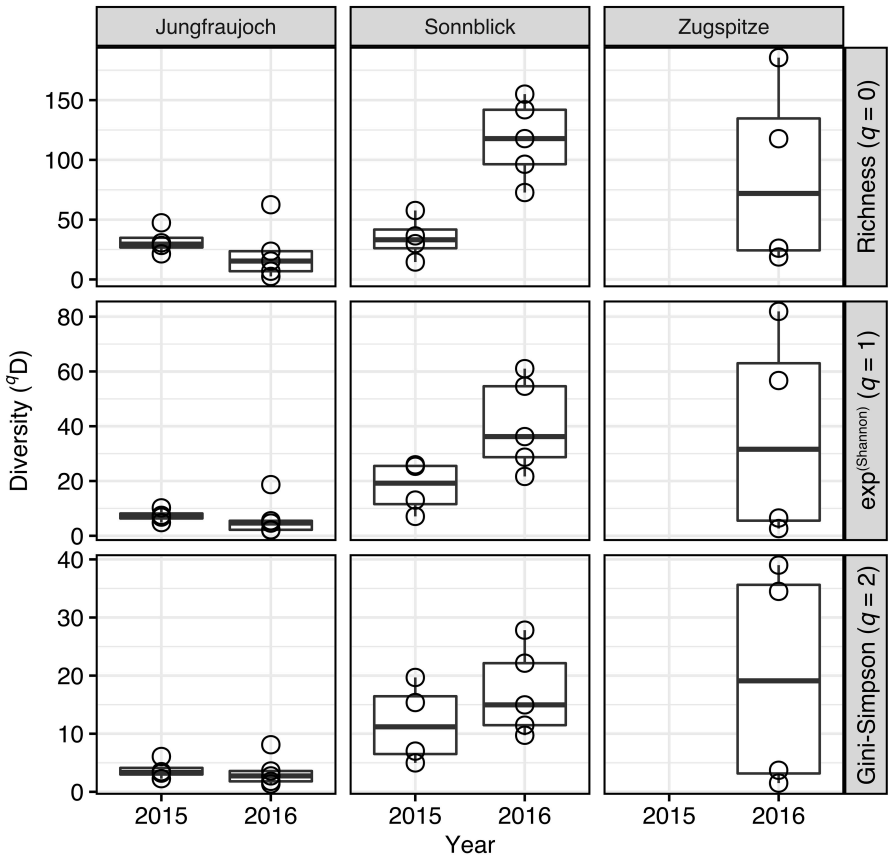

FIGURE 7 | Microbial community alpha diversity estimates inferred from Hill numbers per site and year.

magnitude lower (Anesio et al., 2007; Bellas et al., 2013; Maccario et al., 2015), which is not surprising given the higher nutrient levels and microbial productivity in cryoconite compared to clean surface snow as investigated in our study.

Meola et al. (2015) have reported distinctly elevated abundances of prokaryotic cells in Sahara dust layers contained in surface snow collected at Jungfraujoch in 2014. While snow profiles in our study also contained visible sand layers likely derived from Sahara dust events, peaks with elevated numbers of prokaryotic cells or VLP compared to clean snow above or below were not always consistent with these layers. Similarly, we did not find a consistent trend regarding changes in prokaryotic cell and VLP numbers with snow layer height. The lack of a noticeable, consistent relationship between cell and VLP numbers, respectively, and observed dust layers or snow layer height in our study might be explained by findings by Lazzaro et al. (2015). Here, the authors showed that microbial cells are mobilized more readily than dust particles during intermediate snow melting, which can lead to homogenization of cells across the depth profile of snowpack along meltwater pores over time. Single instances where significantly increasing cell numbers toward the bottom of the snowpack were observed in our study (Jungfraujoch and Sonnblick, 2015) suggest accumulation of cells at greater depths due to snow melting and downward transport.

Since prokaryotic cells and VLP were present in similar numbers, VLP-to-cell ratios ranged slightly above 1 on average and rarely exceeded 10, even though the total variation spanned two orders of magnitude. Except for the considerable difference at Jungfraujoch between the 2 years, the variation between sites and years was comparatively low. Overall, variation of environmental VLP-to-prokaryotic cell ratios over several orders of magnitude are not uncommon (Parikka et al., 2017). However, compared to other oligotrophic environments or glacial cryoconite (Anesio et al., 2007; Rassner et al., 2016; Parikka et al., 2017), VLP-to-cell ratios observed in our snow samples were about one order of magnitude lower. VLP-to-cell ratios are generally expected to decrease with decreasing microbial productivity common for oligotrophic environments, and high viral decay for example caused by intense solar radiation (Wommack and Colwell, 2000; Parikka et al., 2017). Both factors likely play a larger role in clean surface snow compared to cryoconite, which might explain the lower ratios observed in our study. Comparatively high VLP-to-cell ratios and strong positive correlations between the numbers of prokaryotic cells and VLP in cryoconite holes and glacial ice have been suggested to indicate strong control of viruses over microbial communities in these environments (Anesio et al., 2007; Anesio and Bellas, 2011; Bellas et al., 2013). The relatively low VLP-to-cell ratios and generally weak or lack of correlations

between cell and VLP numbers observed in our study seem to suggest that such a strong viral control might not apply to alpine surface snow. Nevertheless, the single instance of a strong positive correlation between cell and VLP numbers observed for Jungfraujoch in 2015, and the large variation of VLP-to-cell ratios within the individual snow profiles might indicate that the strength of viral control is dynamic even in these environments, and might change over relatively small time scales, for instance with changes in nutrient inputs from the atmosphere (Rassner et al., 2016).

Microbial communities were composed of similar groups of organisms as have been found in previous studies on mountain surface snow and ice in different parts of the globe (Segawa et al., 2005; Boetius et al., 2015; Maccario et al., 2015; Antony et al., 2016; Carey et al., 2016; Wunderlin et al., 2016; Azzoni et al., 2018; Zhou et al., 2019), including the European Alps (Chuvochina et al., 2011; Meola et al., 2015; Wunderlin et al., 2016; Azzoni et al., 2018; Courville et al., 2020; Els et al., 2020), and >10000-year-old Himalayan glacial ice (Zhong et al., 2021). Dominant genera included known psychrophilic or psychrotolerant, UV-resistant organisms typical for cryosphere environments such as heterotrophic *Polaromonas* (Gammaproteobacteria; Darcy et al., 2011) and *Hymenobacter* (Bacteroidia; Dai et al., 2009; Klassen and Foght, 2011; Sedláček et al., 2019), or photoautotrophs within the *Chroococcidiopsidaceae* (Cyanobacteria; Baqué et al., 2013). Additionally, communities contained more generalist organisms such as *Massilia* (Gammaproteobacteria), *Sphingomonas* (Alphaproteobacteria), *Streptococcus* and *Staphylococcus* (Bacilli), or *Ferruginibacter* (Bacteroidia), which are frequently encountered in temperate freshwater and soil, but have also been found in mountain snow and ice by others (Segawa et al., 2005; Meola et al., 2015; Zhou et al., 2019; Zhong et al., 2021). The composition of surface snow-associated microbial communities in alpine environments can differ from those found at lower altitudes in the Arctic or Antarctic (Boetius et al., 2015), in part likely due to contributions of additional sources of microorganisms absent from alpine regions such as seawater and brine (Harding et al., 2011; Maccario et al., 2019). However, certain organisms like *Polaromonas*, but also other Proteobacteria, Cyanobacteria, Actinobacteria, Firmicutes, and Bacteroidetes have frequently been observed in snowpack in alpine regions like in our study and the ones cited above, as well as in the Arctic and Antarctic (e.g., see Harding et al., 2011; Hell et al., 2013; Antony et al., 2016). This suggests a wide distribution of these taxa across broad geographical ranges and altitudes.

Similar to previous studies, we observed significant variation in microbial community composition between years (Chuvochina et al., 2011; Pittino et al., 2018; Els et al., 2020) and sites (Azzoni et al., 2018; Courville et al., 2020). However, our analyses showed that differences in community composition between years were stronger compared to differences between sites within years, especially considering phylogenetic community composition. Hence, despite differences in the presence and abundance of individual ASVs, the microbial communities at different sites were composed of phylogenetically similar taxa. This observation is in line with previous studies that found similar phylogenetic groups in mountain surface snow across different locations despite overall significant differences in community composition (Wunderlin et al., 2016; Brown and Jumpponen, 2019). Els et al. (2019, 2020) showed that microbial community composition in tropospheric air and precipitation can vary significantly over time, which in turn contributes to temporal variation in community composition in mountain surface snow as microbes are deposited from the atmosphere (see also Chuvochina et al., 2011).

Interestingly, the stable isotope analyses suggested that the time spans of snow accumulation periods covered by the snow profiles in each year (6–10 months) were usually larger than time differences between the youngest top layers of the old snowpack and the oldest bottom layers of new snow accumulation periods in the following year (4–8 months). We may speculate that the seasonal snow surfaces at the different sites in our study had been initially colonized by microorganisms from similar sources, such as atmospheric dust and aerosols of similar origins, and that the composition of these seed communities varied strongly between years. The subsequent community succession following the initial colonization could have been determined to some degree by the composition of the seed communities through priority effects (Fukami, 2015), as well as similar environmental conditions across sites selecting for phylogenetically similar organisms (Maccario et al., 2019; Els et al., 2020). In addition to selection, we can expect that communities were additionally subject to ecological drift, causing random extinction and arrival of individual ASVs, thus leading to the significant albeit comparatively small differences in taxonomic community composition between sites within years.

Our estimates of community alpha diversity were in agreement with those observed in previous studies at Jungfraujoch from 2014 (Meola et al., 2015) and 2015 (Wunderlin et al., 2016), and estimates reported for Sonnblick for three consecutive seasons between 2016 and 2017 (Els et al., 2020). In line with our results, communities in these studies contained approximately 50–250 distinct microbial taxa on average, with a similarly low contribution of abundant taxa to the total richness, resulting in low community evenness. Comparable ranges have further been reported for surface snow from different alpine regions spanning from Anatolia to the Himalayas (Azzoni et al., 2018), or the Australian Alps (Wunderlin et al., 2016) (however, about 5–10 times higher estimates have been reported for mountain surface snow in the Sierra Nevada, United States; Carey et al., 2016). The modest variation in community alpha diversity between years in our study, and in comparison to estimates from earlier studies at the same sites, indicates that community alpha diversity—in contrast to community composition—varies within a relatively narrow range from year to year. Taking into account a certain bias due to methodological differences in amplicon sequencing and sequence data processing between studies, collectively these results suggest that alpine surface snow environments may have similar microbial community diversity equilibria (also referred to as carrying capacity for species richness; Storch and Okie, 2019) across wide geographical ranges that are relatively stable over consecutive seasons.

The predicted decrease in snow cover duration and snow mass combined with higher frequencies of precipitation extremes

in the alpine cryosphere due to climate change (Hock et al., 2019) can be expected to affect the spatiotemporal dynamics and magnitude of variation in the microbiological parameters observed over the relatively short timescale of our study. We may assume that shorter snow cover duration might leave less time for selection processes to operate on microbial communities in the snowpack, which in turn could amplify the impact of stochastic drift and thereby cause larger variation in microbial community composition and diversity between seasons as well as locations. Similarly, abundances of and dynamics between microorganisms and viruses may be affected by changes in the quantity and quality of precipitation. As variation increases, uncertainties in predictions of the amounts and types of microorganisms and viruses that are transported with meltwater to downstream environments would also increase. Hence, we argue that knowledge on the spatiotemporal dynamics of microbial communities associated with mountain snow and changes in the magnitude of variation over time is an important factor for our understanding of climate change impacts on the alpine cryosphere and beyond. Our study provides a first glimpse at the temporal variation in microbial abundance, community composition, and diversity between two consecutive seasons at the three investigated sites. However, longer time series over several seasons are certainly necessary to assess whether this variation remains stable around a certain baseline like the values reported here, or rather shows an increasing trend over longer time scales in the future.

CONCLUSION

Our study suggests that numbers of prokaryotic cells and VLP as well as microbial community alpha diversity in alpine surface snow show relatively little variation across space and snow accumulation periods on average. This is further supported by the agreement between values observed in our study and those previously reported for the same sites from different time points, as well as observations from studies at different high altitude locations. Relatively low VLP-to-prokaryotic cell ratios, and mostly weak or lacking correlations between numbers of prokaryotic cells and VLP observed in our study may indicate a weaker viral control over microbial communities in clean alpine surface snow compared to glacial ice or cryoconite reported by others. However, more direct research on this issue would be required. While microbial community composition was comparable between sites within years, we showed that in contrast to prokaryotic cell numbers and microbial community alpha diversity, community composition was highly dynamic and changed significantly between years. Based on previous findings, it seems likely that the strong year-to-year variation in community composition could have been determined by differences in the composition of seed communities derived from the atmosphere and precipitation, although our data do not allow any direct inferences in that regard. All in all, our findings add to the knowledge on spatiotemporal dynamics and magnitude of variation of microbial abundances, community composition, and diversity in alpine surface snow. This knowledge may aid defining baselines to assess future impacts of climate change on the alpine cryosphere.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/ **Supplementary Material**.

AUTHOR CONTRIBUTIONS

LF analyzed the data and wrote the manuscript. KH organized and conducted the sampling campaigns and sample collection. CS analyzed the stable isotope data. NW performed the flow cytometry measurements and participated in sampling campaigns. DF did the DNA extraction. BH performed the bioinformatic processing of raw amplicon sequence data. LS contributed to the data curation and literature research. CG organized and supervised sample processing for microbial community analyses. All authors have contributed to and provided feedback on the manuscript.

FUNDING

This study was performed within the framework of the research project "Virtual Alpine Observatory," funded by the Bavarian State Ministry of the Environment and Consumer Protection, Germany (71_1d-U8729-2013/193-24). Financial support for publication was provided by the Open Access Publishing Fund of the University of Vienna.

ACKNOWLEDGMENTS

We gratefully acknowledge the support by Eva-Maria Schiestl, Rainer Lindner, Jan Ruhland, and Aleksandra Kiecak (Helmholtz Zentrum München), who joined the snow sampling campaigns, and Nadine Zentner and Sigrid Kaschuba for helping with the isotope analyses and DNA extractions, respectively. We also kindly thank the staff of the Environmental Research Station Schneefernerhaus (M. Neumann and colleagues), the High Altitude Research Station Jungfraujoch (M. Leuenberger and colleagues), and the Sonnblick Observatory (E. Ludewig and colleagues) for their kind accommodation and technical support for the field work.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.781904/full#supplementary-material

REFERENCES

Amante, C., and Eakins, B. W. (2009). *ETOPO1 1 Arc-Minute Global Relief Model: Procedures, Data Sources and Analysis. NOAA Technical Memorandum NESDIS NGDC-24. National Geophysical Data Center, NOAA*. Washington, DC: NOAA.

Anesio, A. M., and Bellas, C. M. (2011). Are low temperature habitats hot spots of microbial evolution driven by viruses? *Trends Microbiol.* 19, 52–57. doi: 10.1016/j.tim.2010.11.002

Anesio, A. M., Mindl, B., Laybourn-Parry, J., Hodson, A. J., and Sattler, B. (2007). Viral dynamics in cryoconite holes on a high Arctic glacier (Svalbard). *J. Geophys. Res.* 112:G04S31. doi: 10.1029/2006jg000350

Anesio, A. M., Sattler, B., Foreman, C., Telling, J., Hodson, A., Tranter, M., et al. (2010). Carbon fluxes through bacterial communities on glacier surfaces. *Ann. Glaciol.* 51, 32–40. doi: 10.3189/172756411795932092

Antony, R., Sanyal, A., Kapse, N., Dhakephalkar, P. K., Thamban, M., and Nair, S. (2016). Microbial communities associated with Antarctic snow pack and their biogeochemical implications. *Microbiol. Res.* 192, 192–202. doi: 10.1016/j.micres.2016.07.004

Apprill, A., McNally, S., Parsons, R., and Weber, L. (2015). Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. *Aquat. Microb. Ecol.* 75, 129–137. doi: 10.3354/ame01753

Azzoni, R. S., Tagliaferri, I., Franzetti, A., Mayer, C., Lambrecht, A., Compostella, C., et al. (2018). Bacterial diversity in snow from mid-latitude mountain areas: alps, Eastern Anatolia, Karakoram and Himalaya. *Ann. Glaciol.* 59, 10–20. doi: 10.1017/aog.2018.18

Baqué, M., de Vera, J.-P., Rettberg, P., and Billi, D. (2013). The BOSS and BIOMEX space experiments on the EXPOSE-R2 mission: endurance of the desert cyanobacterium Chroococcidiopsis under simulated space vacuum, Martian atmosphere, UVC radiation and temperature extremes. *Acta Astronaut.* 91, 180–186. doi: 10.1016/j.actaastro.2013.05.015

Bellas, C. M., Anesio, A. M., Telling, J., Stibal, M., Tranter, M., and Davis, S. (2013). Viral impacts on bacterial communities in Arctic cryoconite. *Environ. Res. Lett.* 8:045021. doi: 10.1088/1748-9326/8/4/045021

Boetius, A., Anesio, A. M., Deming, J. W., Mikucki, J. A., and Rapp, J. Z. (2015). Microbial ecology of the cryosphere: sea ice and glacial habitats. *Nat. Rev. Microbiol.* 13, 677–690. doi: 10.1038/nrmicro3522

Brown, S. P., and Jumpponen, A. (2019). Microbial ecology of snow reveals taxa-specific biogeographical structure. *Microb. Ecol.* 77, 946–958. doi: 10.1007/s00248-019-01357-z

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., and Holmes, S. P. (2016a). DADA2: high-resolution sample inference from Illumina amplicon data. *Nat. Methods* 13, 581–583. doi: 10.1038/nmeth.3869

Callahan, B. J., Sankaran, K., Fukuyama, J. A., McMurdie, P. J., and Holmes, S. P. (2016b). Bioconductor workflow for microbiome data analysis: from raw reads to community analyses. *F1000Research* 5:1492. doi: 10.12688/f1000research.8986.2

Campitelli, E. (2021). *metR: Tools for Easier Analysis of Meteorological Fields. R Package*.

Carey, C. J., Hart, S. C., Aciego, S. M., Riebe, C. S., Blakowski, M. A., and Aronson, E. L. (2016). Microbial community structure of subalpine snow in the Sierra Nevada, California. *Arct. Antarct. Alp. Res.* 48, 685–701. doi: 10.1657/AAAR0015-062

Carpenter, E. J., Lin, S., and Capone, D. G. (2000). Bacterial activity in South Pole snow. *Appl. Environ. Microbiol.* 66, 4514–4517. doi: 10.1128/AEM.66.10.4514-4517.2000

Chao, A., Chiu, C.-H., and Jost, L. (2014a). Unifying species diversity, phylogenetic diversity, functional diversity, and related similarity and differentiation measures through hill numbers. *Annu. Rev. Ecol. Evol. Syst.* 45, 297–324. doi: 10.1146/annurev-ecolsys-120213-091540

Chao, A., Gotelli, N. J., Hsieh, T. C., Sander, E. L., Ma, K. H., Colwell, R. K., et al. (2014b). Rarefaction and extrapolation with Hill numbers: a framework for sampling and estimation in species diversity studies. *Ecol. Monogr.* 84, 45–67. doi: 10.1890/13-0133.1

Chuvochina, M. S., Alekhina, I. A., Normand, P., Petit, J.-R., and Bulat, S. A. (2011). Three events of Saharan dust deposition on the Mont Blanc glacier associated with different snow-colonizing bacterial phylotypes. *Microbiology* 80, 125–131. doi: 10.1134/S0026261711010061

Courville, Z. R., Lieblappen, R. M., Thurston, A. K., Barbato, R. A., Fegyveresi, J. M., Farnsworth, L. B., et al. (2020). Microorganisms associated with dust on alpine snow. *Front. Earth Sci.* 8:122. doi: 10.3389/feart.2020.00122

Dai, J., Wang, Y., Zhang, L., Tang, Y., Luo, X., An, H., et al. (2009). Hymenobacter tibetensis sp. nov., a UV-resistant bacterium isolated from Qinghai-Tibet plateau. *Syst. Appl. Microbiol.* 32, 543–548. doi: 10.1016/j.syapm.2009.09.001

Darcy, J. L., Lynch, R. C., King, A. J., Robeson, M. S., and Schmidt, S. K. (2011). Global distribution of Polaromonas phylotypes–evidence for a highly successful dispersal capacity. *PLoS One* 6:e23742. doi: 10.1371/journal.pone.0023742

Els, N., Greilinger, M., Reisecker, M., Tignat-Perrier, R., Baumann-Stanzer, K., Kasper-Giebl, A., et al. (2020). Comparison of bacterial and fungal composition and their chemical interaction in free tropospheric air and snow over an entire winter season at Mount Sonnblick, Austria. *Front. Microbiol.* 11:980. doi: 10.3389/fmicb.2020.00980

Els, N., Larose, C., Baumann-Stanzer, K., Tignat-Perrier, R., Keuschnig, C., Vogel, T. M., et al. (2019). Microbial composition in seasonal time series of free tropospheric air and precipitation reveals community separation. *Aerobiologia* 35, 671–701. doi: 10.1007/s10453-019-09606-x

Elser, J. J., Wu, C., González, A. L., Shain, D. H., Smith, H. J., Sommaruga, R., et al. (2020). Key rules of life and the fading cryosphere: impacts in alpine lakes and streams. *Glob. Chang. Biol.* 26, 6644–6656. doi: 10.1111/gcb.15362

Environment Agency Austria [EAA] (2016). *Österreichisches Messnetz für Isotope im Niederschlag und in Oberflächengewässern (ANIP)*. Available online at: https://www.umweltbundesamt.at/wasser/informationen/isotope/isotopenmessnetz-anip (accessed May 12, 2016).

Fierz, C., Armstrong, R. L., Durand, Y., Etchevers, P., Greene, E., McClung, D. M., et al. (2009). "The international classification for seasonal snow on the ground," in *Technical Documents in Hydrology No 83, IACS Contribution No 1*, ed. UNESCO (Paris: UNESCO-IHP).

Franzetti, A., Tagliaferri, I., Gandolfi, I., Bestetti, G., Minora, U., Mayer, C., et al. (2016). Light-dependent microbial metabolisms drive carbon fluxes on glacier surfaces. *ISME J.* 10, 2984–2988. doi: 10.1038/ismej.2016.72

Fukami, T. (2015). Historical contingency in community assembly: integrating niches, species pools, and priority effects. *Annu. Rev. Ecol. Evol. Syst.* 46, 1–23. doi: 10.1146/annurev-ecolsys-110411-160340

Gat, J. R. (1996). Oxygen and hydrogen isotopes in the hydrologic cycle. *Annu. Rev. Earth Planet. Sci.* 24, 225–262. doi: 10.1146/annurev.earth.24.1.225

Harding, T., Jungblut, A. D., Lovejoy, C., and Vincent, W. F. (2011). Microbes in high arctic snow and implications for the cold biosphere. *Appl. Environ. Microbiol.* 77, 3234–3243. doi: 10.1128/AEM.02611-10

Hell, K., Edwards, A., Zarsky, J., Podmirseg, S. M., Girdwood, S., Pachebat, J. A., et al. (2013). The dynamic bacterial communities of a melting High Arctic glacier snowpack. *ISME J.* 7, 1814–1826. doi: 10.1038/ismej.2013.51

Hock, R., Rasul, G., Adler, C., Cáceres, B., Gruber, S., Hirabayashi, Y., et al. (2019). "Chapter 2: high mountain areas," in *IPCC Special Report on the Ocean and Cryosphere in a Changing Climate*, et al Edn, eds H. O. Pörtner, D. C. Roberts, V. Masson-Delmotte, P. Zhai, M. Tignor, and E. Poloczanska (Geneva: Intergovernmental Panel on Climate Change), 131–202.

Hotaling, S., Hood, E., and Hamilton, T. L. (2017). Microbial ecology of mountain glacier ecosystems: biodiversity, ecological connections and implications of a warming climate. *Environ. Microbiol.* 19, 2935–2948. doi: 10.1111/1462-2920.13766

Hsieh, T. C., and Chao, A. (2017). Rarefaction and extrapolation: making fair comparison of abundance-sensitive phylogenetic diversity among multiple assemblages. *Syst. Biol.* 66, 100–111. doi: 10.1093/sysbio/syw073

Hürkamp, K., Zentner, N., Reckerth, A., Weishaupt, S., Wetzel, K.-F., Tschiersch, J., et al. (2019). Spatial and temporal variability of snow isotopic composition on Mt. Zugspitze, Bavarian alps, Germany. *J. Hydrol. Hydromech.* 67, 49–58. doi: 10.2478/johh-2018-0019

Kembel, S. W., Cowan, P. D., Helmus, M. R., Cornwell, W. K., Morlon, H., Ackerly, D. D., et al. (2010). Picante: r tools for integrating phylogenies and ecology. *Bioinformatics* 26, 1463–1464. doi: 10.1093/bioinformatics/btq166

Klassen, J. L., and Foght, J. M. (2011). Characterization of Hymenobacter isolates from Victoria Upper Glacier, Antarctica reveals five new species and substantial non-vertical evolution within this genus. *Extremophiles* 15, 45–57. doi: 10.1007/s00792-010-0336-1

Lazzaro, A., Wismer, A., Schneebeli, M., Erny, I., and Zeyer, J. (2015). Microbial abundance and community structure in a melting alpine snowpack. *Extremophiles* 19, 631–642. doi: 10.1007/s00792-015-0744-3

Lopatina, A., Krylenkov, V., and Severinov, K. (2013). Activity and bacterial diversity of snow around Russian Antarctic stations. *Res. Microbiol.* 164, 949–958. doi: 10.1016/j.resmic.2013.08.005

Maccario, L., Carpenter, S. D., Deming, J. W., Vogel, T. M., and Larose, C. (2019). Sources and selection of snow-specific microbial communities in a Greenlandic sea ice snow cover. *Sci. Rep.* 9:2290. doi: 10.1038/s41598-019-38744-y

Maccario, L., Sanguino, L., Vogel, T. M., and Larose, C. (2015). Snow and ice ecosystems: not so extreme. *Res. Microbiol.* 166, 782–795. doi: 10.1016/j.resmic.2015.09.002

McLaren, M. R. (2020). *Silva SSU Taxonomic Training Data Formatted for DADA2 (Silva version 138). Zenodo*.

Meola, M., Lazzaro, A., and Zeyer, J. (2015). Bacterial composition and survival on sahara dust particles transported to the European Alps. *Front. Microbiol.* 6:1454. doi: 10.3389/fmicb.2015.01454

Moser, H., and Stichler, W. (1974). "Deuterium and oxygen-18 contents as an index of the properties of snow covers," in *Proceedings of the Grindelwald Symposium: Snow Mechanics*, Vol. 144, (Wallingford: International Association of Hydrological Sciences), 122–135.

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., et al. (2020). *vegan: Community Ecology Package. R package*. Available online at: https://CRAN.R-project.org/package=vegan

Parada, A. E., Needham, D. M., and Fuhrman, J. A. (2016). Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environ. Microbiol.* 18, 1403–1414. doi: 10.1111/1462-2920.13023

Paradis, E., and Schliep, K. (2019). ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* 35, 526–528. doi: 10.1093/bioinformatics/bty633

Parikka, K. J., Le Romancer, M., Wauters, N., and Jacquet, S. (2017). Deciphering the virus-to-prokaryote ratio (VPR): insights into virus-host relationships in a variety of ecosystems. *Biol. Rev.* 92, 1081–1100. doi: 10.1111/brv.12271

Pilloni, G., Granitsiotis, M. S., Engel, M., and Lueders, T. (2012). Testing the limits of 454 pyrotag sequencing: reproducibility, quantitative assessment and comparison to T-RFLP fingerprinting of aquifer microbes. *PLoS One* 7:e40467. doi: 10.1371/journal.pone.0040467

Pittino, F., Maglio, M., Gandolfi, I., Azzoni, R. S., Diolaiuti, G., Ambrosini, R., et al. (2018). Bacterial communities of cryoconite holes of a temperate alpine glacier show both seasonal trends and year-to-year variability. *Ann. Glaciol.* 59, 1–9. doi: 10.1017/aog.2018.16

Pjevac, P., Hausmann, B., Schwarz, J., Kohl, G., Herbold, C. W., Loy, A., et al. (2021). An economical and flexible dual barcoding, Two-Step PCR approach for highly multiplexed amplicon sequencing. *Front. Microbiol.* 12:669776. doi: 10.3389/fmicb.2021.669776

Pohlert, T. (2021). *PMCMRplus: Calculate Pairwise Multiple Comparisons of Mean Rank Sums Extended. R Package*. Available online at: https://CRAN.R-project.org/package=PMCMRplus

Price, M. N., Dehal, P. S., and Arkin, A. P. (2010). FastTree 2 – approximately maximum-likelihood trees for large alignments. *PLoS One* 5:e9490. doi: 10.1371/journal.pone.0009490

R Core Team, (2021). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Rassner, S. M. E., Anesio, A. M., Girdwood, S. E., Hell, K., Gokul, J. K., Whitworth, D. E., et al. (2016). Can the bacterial community of a high arctic glacier surface escape viral control? *Front. Microbiol.* 7:956. doi: 10.3389/fmicb.2016.00956

Ren, Z., Martyniuk, N., Oleksy, I. A., Swain, A., and Hotaling, S. (2019). Ecological stoichiometry of the mountain cryosphere. *Front. Ecol. Evol.* 7:360. doi: 10.3389/fevo.2019.00360

Revell, L. J. (2012). phytools: an R package for phylogenetic comparative biology (and other things). *Methods Ecol. Evol.* 3, 217–223. doi: 10.1111/j.2041-210x.2011.00169.x

Sedláček, I., Pantůček, R., Králová, S., Mašlaňová, I., Holochová, P., Staňková, E., et al. (2019). Hymenobacter amundsenii sp. nov. resistant to ultraviolet radiation, isolated from regoliths in Antarctica. *Syst. Appl. Microbiol.* 42, 284–290. doi: 10.1016/j.syapm.2018.12.004

Segawa, T., Miyamoto, K., Ushida, K., Agata, K., Okada, N., and Kohshima, S. (2005). Seasonal change in bacterial flora and biomass in mountain snow from the Tateyama Mountains, Japan, analyzed by 16S rRNA gene sequencing and real-time PCR. *Appl. Environ. Microbiol.* 71, 123–130. doi: 10.1128/AEM.71.1.123-130.2005

Stibal, M., Telling, J., Cook, J., Mak, K. M., Hodson, A., and Anesio, A. M. (2012b). Environmental controls on microbial abundance and activity on the greenland ice sheet: a multivariate analysis approach. *Microb. Ecol.* 63, 74–84. doi: 10.1007/s00248-011-9935-3

Stibal, M., Šabacká, M., and Žárský, J. (2012a). Biological processes on glacier and ice sheet surfaces. *Nat. Geosci.* 5, 771–774. doi: 10.1038/ngeo1611

Stichler, W. (1987). "Snowcover and snowmelt processes studied by means of environmental isotopes," in *Seasonal Snowcovers: Physics, Chemistry, Hydrology*, eds H. G. Jones and W. J. Orville-Thomas (Dordrecht: Springer Netherlands), 673–726.

Storch, D., and Okie, J. G. (2019). The carrying capacity for species richness. *Glob. Ecol. Biogeogr.* 28, 1519–1532. doi: 10.1111/geb.12987

Weiss, S., Xu, Z. Z., Peddada, S., Amir, A., Bittinger, K., Gonzalez, A., et al. (2017). Normalization and microbial differential abundance strategies depend upon data characteristics. *Microbiome* 5:27. doi: 10.1186/s40168-017-0237-y

Wommack, K. E., and Colwell, R. R. (2000). Virioplankton: viruses in aquatic ecosystems. *Microbiol. Mol. Biol. Rev.* 64, 69–114. doi: 10.1128/MMBR.64.1.69-114.2000

Wright, E. S. (2016). Using DECIPHER v2.0 to analyze big biological sequence data in R. *R J.* 8, 352–359. doi: 10.32614/rj-2016-025

Wunderlin, T., Ferrari, B., and Power, M. (2016). Global and local-scale variation in bacterial community structure of snow from the Swiss and Australian Alps. *FEMS Microbiol. Ecol.* 92:fiw132. doi: 10.1093/femsec/fiw132

Zekollari, H., Huss, M., and Farinotti, D. (2019). Modelling the future evolution of glaciers in the European Alps under the EURO-CORDEX RCM ensemble. *Cryosphere* 13, 1125–1146. doi: 10.5194/tc-13-1125-2019

Zhong, Z.-P., Tian, F., Roux, S., Gazitúa, M. C., Solonenko, N. E., Li, Y.-F., et al. (2021). Glacier ice archives nearly 15,000-year-old microbes and phages. *Microbiome* 9:160. doi: 10.1186/s40168-021-01106-w

Zhou, L., Zhou, Y., Hu, Y., Cai, J., Liu, X., Bai, C., et al. (2019). Microbial production and consumption of dissolved organic matter in glacial ecosystems on the Tibetan Plateau. *Water Res.* 160, 18–28. doi: 10.1016/j.watres.2019.05.048

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

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 03 January 2022
doi: 10.3389/fmicb.2021.760282

Size-Fractionated Microbiome Structure in Subarctic Rivers and a Coastal Plume Across DOC and Salinity Gradients

Marie-Amélie Blais[1,2], Alex Matveev[1,2], Connie Lovejoy[1,3] and Warwick F. Vincent[1,2]**

[1]*Département de Biologie, Institut de Biologie Intégrative et des Systèmes (IBIS) and Takuvik Joint International Laboratory, Université Laval, Quebec City, QC, Canada,* [2]*Centre for Northern Studies (CEN), Université Laval, Quebec City, QC, Canada,* [3]*Québec-Océan, Université Laval, Quebec City, QC, Canada*

Little is known about the microbial diversity of rivers that flow across the changing subarctic landscape. Using amplicon sequencing (rRNA and rRNA genes) combined with HPLC pigment analysis and physicochemical measurements, we investigated the diversity of two size fractions of planktonic Bacteria, Archaea and microbial eukaryotes along environmental gradients in the Great Whale River (GWR), Canada. This large subarctic river drains an extensive watershed that includes areas of thawing permafrost, and discharges into southeastern Hudson Bay as an extensive plume that gradually mixes with the coastal marine waters. The microbial communities differed by size-fraction (separated with a 3-μm filter), and clustered into three distinct environmental groups: (1) the GWR sites throughout a 150-km sampling transect; (2) the GWR plume in Hudson Bay; and (3) small rivers that flow through degraded permafrost landscapes. There was a downstream increase in taxonomic richness along the GWR, suggesting that sub-catchment inputs influence microbial community structure in the absence of sharp environmental gradients. Microbial community structure shifted across the salinity gradient within the plume, with changes in taxonomic composition and diversity. Rivers flowing through degraded permafrost had distinct physicochemical and microbiome characteristics, with allochthonous dissolved organic carbon explaining part of the variation in community structure. Finally, our analyses of the core microbiome indicated that while a substantial part of all communities consisted of generalists, most taxa had a more limited environmental range and may therefore be sensitive to ongoing change.

Keywords: bacteria, microbial eukaryotes, permafrost, river microbiomes, climate change, salinity, dissolved organic carbon, northern ecosystems

OPEN ACCESS

Edited by:
David Velazquez, Autonomous University of Madrid, Spain

Reviewed by:
Veljo Kisand, University of Tartu, Estonia
Christian Wurzbacher, Technical University of Munich, Germany

****Correspondence:***
Marie-Amélie Blais
marie-amelie.blais.1@ulaval.ca
Warwick F. Vincent
warwick.vincent@bio.ulaval.ca

Specialty section:
This article was submitted to Extreme Microbiology, a section of the journal Frontiers in Microbiology

Received: *17 August 2021*
Accepted: *01 December 2021*
Published: *03 January 2022*

Citation:
Blais MA, Matveev A, Lovejoy C and Vincent WF (2022) Size-Fractionated Microbiome Structure in Subarctic Rivers and a Coastal Plume Across DOC and Salinity Gradients. Front. Microbiol. 12:760282. doi: 10.3389/fmicb.2021.760282

INTRODUCTION

River ecosystems connect biogeochemical cycles across terrestrial, lacustrine and marine biomes (Aufdenkampe et al., 2011) and are major features of the northern landscape (Vincent and Laybourn-Parry, 2008; Pekel et al., 2016). These northern landscapes are experiencing climate warming and associated transformation through vegetation changes (Jia et al., 2019), hydrological

shifts (Wrona et al., 2016) and permafrost degradation (Biskaborn et al., 2019). Rivers act as vectors of these terrestrial changes to the marine environment, and northern seas are highly influenced by their freshwater inflows. The Arctic Ocean receives close to 11% of global river discharge while representing only 1% of the global ocean volume (McClelland et al., 2012), and many biological processes in Arctic marine ecosystems are closely linked to carbon and nutrient inputs from rivers and coastal erosion (Vallières et al., 2008; Terhaar et al., 2021). Additionally, northern rivers are biogeochemical conduits from land to the atmosphere *via* the decomposition of terrestrial carbon and greenhouse gas fluxes across the water–air interface (Kling et al., 1991; Karlsson et al., 2021).

Despite the wide-ranging importance of northern rivers and their biogeochemistry, little is known about their microbiomes, defined as the assemblage of microbial eukaryotes, Bacteria, Archaea and viruses (Grossart et al., 2020). These communities underpin biogeochemical processes and food webs in the aquatic environment, but most analyses of high latitude rivers to date have focused on specific subcomponents. For example, early studies of Archaea in the Mackenzie River, Canada, revealed an unexpected diversity, which was attributed to heterogeneous substrates and to different archaeal populations transported into the river from different parts of the flooded permafrost catchment (Galand et al., 2006). A major transect analysis along 1800 km of the Yenisei River (Russia) showed that bacterial assemblages differed among three sections of the river, and partitioned according to the terrestrial ecozones of mountain taiga, plain taiga and the downstream permafrost region of forest-tundra and tundra (Kolmakova et al., 2014). Bacterial community structure in tundra rivers on Svalbard (Spitzbergen) provided evidence of seasonal changes, attributed in part to increased organic carbon supply in late summer (Kosek et al., 2019). A multi-component analysis of microbiomes along a hydrological continuum (soil water, stream waters, terminal lake) in the Alaskan Arctic tundra showed the effects of dispersal and species sorting, with the downstream diversity of microbial eukaryotes less dependent on dispersal from terrestrial sources than for Bacteria and Archaea (Crump et al., 2012).

The influx of freshwater from rivers into the marine environment creates a transition zone that is characterized by pronounced gradients in chemical properties, including terrestrially derived organic matter, nutrients and salinity (Eyre and Balls, 1999; Dagg et al., 2004; Gebhardt et al., 2004). Salinity is known to act as a major filter in aquatic microbial dispersion, with a controlling effect on the shift in riverine microbiome structure towards brackish and marine taxa (Crump et al., 2004; Lozupone and Knight, 2007; Fortunato and Crump, 2015). Increasing attention has been given to the microbial ecology of these freshwater–saltwater transition zones of northern rivers. In the transition zone of the Mackenzie River, for example, Garneau et al. (2006) found strong gradients in bacterial community structure that correlated with salinity, with a large fraction of the prokaryotic community production associated with particles >3 μm. In a set of transects in Hudson Bay, high-throughput amplicon sequencing showed that colonization of marine coastal waters by freshwater protists was restricted by salinity effects, and that the estuarine communities were a mixture of estuarine specialists, freshwater taxa and marine species (Jacquemot et al., 2021). A study of the prokaryotic and eukaryotic microbiome of coastal lagoons along the eastern Alaskan Beaufort Sea revealed a strong seasonality in community structure, with shifts in energy acquisition pathways associated with seasonal changes in sea ice cover and terrestrial carbon inputs (Kellogg et al., 2019).

Thawing permafrost is likely to affect the microbial community structure of northern rivers and their receiving coastal marine waters in a variety of ways. By increasing hydrological connectivity, permafrost thaw could favor immigration of soil microbes into the river plankton (Crump et al., 2012; Vonk et al., 2015). Degrading permafrost releases soil particles to the aquatic environment, and additional particles may form by the flocculation of dissolved organic matter (DOM); these particles (total suspended sediments, TSS) provide potential substrates for microbial colonization and growth (Deshpande et al., 2016; Cai, 2020). The mobilization of nutrients and organic matter, including dissolved organic carbon (DOC), previously locked within the frozen permafrost changes water geochemistry, hence the resources available to microbes (Frey and McClelland, 2009; Kendrick et al., 2018). Since bacterial taxa differ in their DOM processing capacities, any changes in chemical composition and concentration of DOM are likely to affect bacterial community composition (Amaral et al., 2016; Roiha et al., 2016). Other changes such as nutrient input could increase primary productivity, stimulating other components of the biota, whereas increased DOM and TSS (turbidity) result in decreased light availability in the water column, potentially limiting photosynthesis and favoring heterotrophy.

The various size components of aquatic microbiomes often differ in their species composition and ecology, and therefore size fractionation allows for a more detailed, ecologically relevant analysis. This approach can be used to distinguish picoeukaryotes (cell size 0.22–3 μm) from larger microbial eukaryotes, as well as free-living bacteria from those associated with particles, often assumed to be attached bacteria. These different size-fractions are likely to represent divergent microbial lifestyles and assemblages that play different roles in biogeochemical processes and food webs. Picocyanobacteria and picoeukaryotes may be more efficient in nutrient acquisition and light capture than larger cells in the same phyla, and are more likely to enter the aquatic food web *via* protist grazing (Stockner, 1988). There are pronounced ecological differences among picophytoplankton taxa, however, including in their potential responses to climate change (Flombaum and Martiny, 2021). Particles contain heterogeneous microhabitats, with localized nutrient and organic matter concentrations that may be higher than in the surrounding water (Simon et al., 2002), and they can be important hubs for biogeochemical processes. For example, they contribute to nitrogen loss in the Yangtze and Yellow rivers (China), with low redox conditions in the center of the particles allowing denitrification to occur in oxygenated waters (Xia et al., 2017). Particle-associated communities are often more diverse (Mohit et al., 2014; Bižić-Ionescu et al., 2015; Savio et al., 2015), more metabolically versatile (Lyons

and Dobbs, 2012), more productive (Crump et al., 1998; Garneau et al., 2006; Ortega-Retuerta et al., 2013) and taxonomically different (Rieck et al., 2015; Payne et al., 2017; Liu et al., 2020) than their free-living counterparts. However, this is not always the case (Hollibaugh et al., 2000; Ghiglione et al., 2007) and bacteria may also alternate between the two lifestyles (Grossart, 2010).

Our aims in the present study were to characterize the environmental gradients in a set of high latitude rivers, tributaries and coastal receiving waters that are influenced by permafrost degradation; to determine the taxonomic composition and diversity of size-fractionated riverine and coastal microbiomes across these gradients; and to identify relationships between the microbiome variables and potential environmental drivers. Specifically, we focused on the Great Whale River (GWR) region in subarctic Québec, Canada, at the southern limit of permafrost soils where thawing and erosion are proceeding rapidly (Bhiry et al., 2011). The southern part of the catchment lies near James Bay, where the permafrost boundary receded northwards by around 130 km over a recent 50 year period (Thibault and Payette, 2009). The downstream region of the catchment contains eroding permafrost, including a valley sub-catchment containing lithalsa thaw lakes (Bouchard et al., 2014) and an adjacent catchment near the coast of palsa thaw lakes (Arlen-Pouliot and Bhiry, 2005). The GWR discharges into Hudson Bay, and we also examined how the bacterioplankton and microbial eukaryotes differed among size-fractions and shifted along salinity gradients in the river-influenced coastal waters. We hypothesized that variables related to permafrost thaw (TSS, DOM/DOC concentrations, DOM quality, total nutrients) would differ among sub-catchments and cause environmental gradients that act in concert with salinity in the plume to control variations in microbial community structure. We additionally posed the questions: are there systematic changes in alpha and beta diversity across these gradients; is there a core microbiome of prokaryotes and eukaryotes across all environmental conditions; and does the taxonomic makeup and diversity of the microbiome differ between size fractions, including between particle-associated and free-living bacteria.

To address these hypotheses and questions, we sampled a 150-km downstream section of the GWR, along with several of its tributaries, an adjacent river system passing through eroding organic-rich permafrost, and the river plume in Hudson Bay, along its salinity gradient at the mouth of the GWR. Fractionated microbial community composition (0.22–3 μm and 3–30 μm) was determined by high throughput sequencing of the V4 region of 16S and 18S rRNA and rRNA genes. Although our main focus here is on rRNA genes, we used both rRNA and rRNA genes, since the latter can be detected in dead cells, potentially biasing the results, while the former have a higher turnover and can be used as an indicator of living cells (Blazewicz et al., 2013). The concentrations of Bacteria and chlorophyll-containing cells were determined by flow cytometry, and photosynthetic pigments were quantified by high pressure liquid chromatography (HPLC). Colored dissolved organic matter (CDOM) absorbance was used to evaluate DOM quality, and to define the environmental gradients we measured nutrients (total phosphorus, total nitrogen, total dissolved nitrogen), dissolved organic and inorganic carbon, and total suspended sediments.

MATERIALS AND METHODS

Study Sites and Sample Collection

The Great Whale River is 726 km long and discharges into southeastern Hudson Bay near the communities of Whapmagoostui (Cree First Nation) and Kuujjuarapik (Inuit; Commission de toponymie du Gouvernement du Québec, 2021). It is one of the largest rivers of Nunavik (subarctic Québec) and a major freshwater source for the bay, with an annual discharge of 19.77 km^3 (Déry et al., 2005). The river water creates a large freshwater plume on the surface of the Hudson Bay (Ingram, 1981) and discharges annually 21,000 t of particulate organic matter and 90×10^3 t of dissolved organic carbon to the bay (Hudon et al., 1996). Its watershed of 44,735 km^2 is predominantly located in a sporadic permafrost area (<2% permafrost coverage) within the lichen woodland zone, except for the coastal area, which is under the climate influence of Hudson Bay and is characterized by discontinuous and scattered permafrost (<50%) and forest-tundra vegetation (Allard and Seguin, 1987; Payette and Rochefort, 2001; Bhiry et al., 2011).

The GWR landscape is heterogeneous and responding to climate change (Allard and Lemay, 2012; Vincent et al., 2017; Kuzyk and Candlish, 2019). The Kwakwatanikapistikw River (hereafter KWK) is a small tributary that enters the GWR approximately 15 km upstream from Hudson Bay. Its valley contains numerous lithalsa thaw lakes (Bouchard et al., 2011), in an area that has recently changed substantially. Between 1959 and 2006, there was a greater than threefold increase in tree vegetation and a near disappearance of permafrost mounds within this lake study area (Bouchard et al., 2014). The nearby Sasapimakwananistikw River (hereafter SAS) flows through another watershed experiencing rapid change and permafrost degradation. This small river passes through a valley of degrading palsa mounds and associated thaw lakes (Arlen-Pouliot and Bhiry, 2005; Figure 7 in Vincent et al., 2017), and discharges into Hudson Bay approximately 1 km southwest of the mouth of the GWR. Additional background information about the GWR and its associated landscapes and marine coastal habitats is given in Nozais et al. (2021).

The 150-km lower reach of GWR, its tributaries (including KWK, Coats River, and Denys River), SAS River and the plume from GWR flowing into Hudson Bay were sampled from 2 to 10 August 2018 (**Figure 1**). Sampling in Hudson Bay and at the mouth of SAS and GWR was by boat, and all other sites were from the riverbank, accessed by helicopter. Physicochemical properties were measured *in situ* with an RBR Concerto CTD logger and a Hydrolab DS5X multiparameter probe. Within the plume, the sampling sites were based on the salinity values measured on-site with the Hydrolab. Given the rapid changes of salinity in the plume, conductivity was re-measured in the individual 10 L Cubitainers™ that were used to collect the water samples. For all sites, triplicate surface water samples were collected

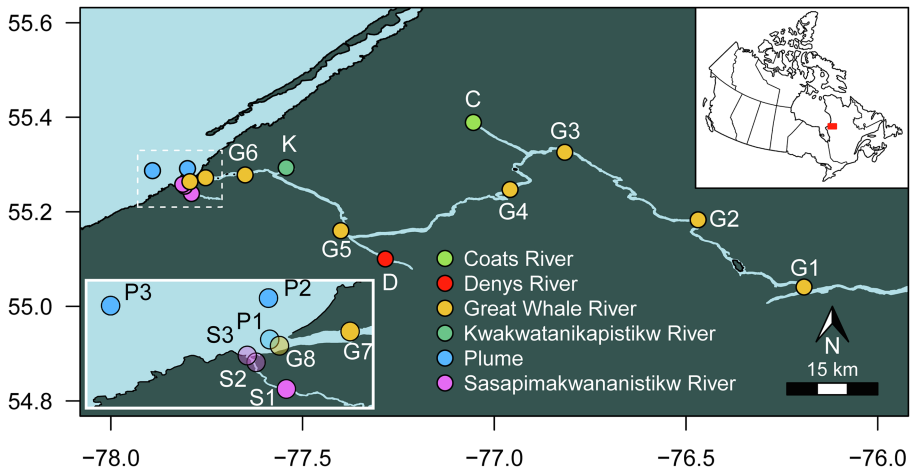

FIGURE 1 | Location of the sampling sites.

5–20 m apart into cleaned Cubitainers™, that were rinsed three times with surface waters and opened submerged 5–10 cm below the surface and then capped. The samples were kept in the dark during transport to the nearby research station in this subarctic region (Center for Northern Studies (CEN), Whapmagoostui-Kuujjuarapik), where they were filtered and subsampled within hours of collection for the molecular, HPLC pigment and water chemistry analyses. The relatively large volume (10 L) would have acted to minimize bottle effects during transport, although the brief delay between collection, filtration or preservation may have allowed some alteration of the microbial communities and chemical parameters. The filtered and preserved samples for chemical measurements were kept in the dark at 4°C until analysis at Laval University and INRS (Quebec City, Canada).

At the subarctic research station, water samples for nucleic acid analysis were prefiltered through a 30-µm mesh to remove zooplankton and then filtered sequentially through a 3-µm pore size, 47-mm diameter polycarbonate filters (large fraction) and 0.22-µm Sterivex™ filter units (Millipore; small fraction) using a peristaltic pump. The volume of water filtered varied (**Supplementary Table S1**). This serial filtration separated putative free-living from particle-associated bacteria, and for microbial eukaryotes, picophytoplankton were enriched in the smaller size fraction and could be analyzed separately, without dilution by the large fraction. The filters were preserved in RNAlater™ solution (Invitrogen™) before being frozen below −50°C until nucleic acid extraction. Unfiltered water samples (1.8 ml) for flow cytometry (FCM) analysis of bacterial and phytoplankton abundance were transferred to Cryovials and preserved by adding 180 µl of glycerol-TE cryoprotectant solution before being stored frozen below −50°C until analysis.

Laboratory and Analysis

Water samples for dissolved organic carbon (DOC), total dissolved nitrogen (TDN) and colored dissolved organic matter (CDOM) were filtered through pre-rinsed 0.2-µm cellulose acetate filters (Advantech MFS) and kept in the dark at 4°C in acid-washed glass bottles (60 ml for DOC and TDN, and 120 ml for CDOM). To correct for any DOC released by the cellulose acetate filters, two blanks were run along with samples, after rinsing the filters with ultrapure water, and these blank values (0.12 and 0.16 mg $C L^{-1}$) were subtracted from the DOC results. TDN (detection limit of 0.02 mg $N L^{-1}$) and DOC (detection limit of 0.05 mg $C L^{-1}$) samples were acidified and analyzed by high temperature catalytic oxidation with non-dispersive infrared detection for the DOC (Standard Methods 5310 B) and with chemiluminescence detection for TDN, using a Shimadzu VCPH analyzer. CDOM absorbance was measured between 200 and 800 nm using a Cary 300 Bio UV–Visible Spectrophotometer (Agilent Technologies). All spectra were blank-corrected using ultrapure water and null-point adjustments were made using the mean value from 750 to 800 nm. Absorbance units were converted to absorption coefficients using:

$$a_{\mathrm{CDOM}}(\lambda) = \frac{2.303 \times A(\lambda)}{L}$$

where $a_{\mathrm{CDOM}}(\lambda)$ is the absorption coefficient (m^{-1}) at wavelength λ, $A(\lambda)$ is the absorbance at that wavelength and L is the path length of the optical cell (m). Specific ultraviolet absorbance at 254 nm normalized to DOC ($SUVA_{254}$, L mg C^{-1} m^{-1}) was calculated by dividing the UV absorbance (m^{-1}) by the DOC concentration (mg L^{-1}) and used as proxy for organic matter aromaticity (Weishaar et al., 2003). The absorption coefficient at 320 nm (a_{320}, m^{-1}) was used as indicator of CDOM concentration (Retamal et al., 2007). Spectral slopes (S, Loiselle et al., 2009) for the intervals 279–299 (S_{289}), 275–295 (S_{285}), 350–400 nm (S_{375}) and the spectral slope ratio (S_R, S_{285}/S_{375}) were calculated using the abs_parms function of the staRdom package (Pucher et al., 2019) for R. S_R was shown to be inversely related to CDOM molecular weight (Helms et al., 2008) and S_{289} is an index of autochthonous carbon (Loiselle et al., 2009; Roiha et al., 2016).

Unfiltered water samples for total phosphorus (TP) and for total nitrogen (TN) were acidified with H_2SO_4 (0.1% final

concentration) and kept at 4°C in 50 ml Falcon tubes. TP and TN samples were digested with alkaline persulphate and analyzed on a Lachat Autoanalyzer using the ascorbic acid colorimetric method (Standard Methods 4500-P E) for TP (detection limit 5 µg P L^{-1}) and the hydrazine reduction followed by sulfanilamide colorimetric method (Standard Methods 12-107-04-1-E) for TN (detection limit 15 µg N L^{-1}). Unfiltered water samples for dissolved inorganic carbon (DIC) were acidified (HCl, 0.05 M) and kept in the dark at 4°C in gastight borosilicate glass vials. DIC (detection limit 0.02 mM) was determined with the headspace gas chromatography method using a GC Trace 1310 (Thermo Scientific™). Total suspended sediments (TSS) samples were collected by filtration until clogging, onto a precombusted, preweighed 47-mm GF/F filters (0.7-µm), kept at −60°C at the CEN research station and transferred to −80°C at Laval University (as for pigments, cytometry and amplicon samples), and determined by weighing of dried filters (70°C for 19 h).

Cells for pigment analysis were collected by filtration until clogging onto 0.7-µm GF/F filters, which were stored below −50°C until analysis. The pigments were extracted from the filters with 95% MeOH, analyzed by HPLC as described in Fournier et al. (2021), and attributed to specific taxonomic groups according to Roy et al. (2011). Bacteria and phytoplankton FCM samples were analyzed with a BD Accuri™ C6 flow cytometer (BD Biosciences). For bacteria, 0.5 µl of SYBR Green (1,000X) was added to 200 µl of the sample followed by 15 min incubation at room temperature before being analyzed at a slow flow rate (16 µl min^{-1}) for 5 min with the threshold set at 800 for FL1 (green fluorescence). Phytoplankton counts were obtained using their natural fluorescence, with 1 ml of sample analyzed at a fast flow rate (66 µl min^{-1}) for 10 min with the threshold set at 800 for FL3 (red fluorescence). Trucount™ tubes (BD Biosciences) were used for the validation and calculation of the flow rate. Fluorescent beads (1 and 3 µm) were used as an internal size standard by adding the beads to GF/F-filtered water samples from each sampling site of different salinities and from the SAS river, with filtered water samples used as blanks.

Nucleic Acid Extraction

Nucleic acids (RNA and DNA) were extracted from each filter (one large and one small fraction from each site, except for the plume samples and the most upstream site in the GWR (sample G1) for which replicates were extracted and for the plume site at salinity 14.86 for which no filters were extracted). For the extractions, we used the AllPrep DNA/RNA Mini Kit (Qiagen) following a modified version of the manufacturer's protocol. Briefly, the RNAlater™ solution (Invitrogen™) was removed prior to the extraction and for an optimal cell lysis, lysozyme was added to the mix of Buffer RLT Plus/β-ME before incubation of the filters at 37°C for 45 min. Proteinase K and SDS 10% were then added to the mix and filters were incubated at 65°C for 15 min. A final modification to the protocol was made to ensure the purity of the RNA extract; after the washing of the column with Buffer RW1, DNase (Qiagen RNase-Free DNase set) was added to the RNeasy spin columns (Qiagen) and incubated at room temperature for 15 min. After extraction, the RNA was tested for DNA contamination by polymerase chain reaction (PCR), using one of the DNA extracted samples as positive control. RNA was then converted to cDNA with a High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems).

Illumina MiSeq Amplicon Library Preparation, Sequencing and Analysis

The microbial community composition of the small and large filtration fractions was determined by amplification and sequencing of the V4 region of the 18S rRNA (cDNA) and 18S rRNA genes (DNA) using primers E572F/E1009R (Comeau et al., 2011) for microbial eukaryotes, and the V4 region of the 16S rRNA (cDNA) and 16S rRNA genes (DNA) using primers 515F (Parada)/806R (Apprill) for Bacteria and Archaea (Apprill et al., 2015; Parada et al., 2016). The amplicon library was prepared by using two single PCR amplification steps, the first to amplify the gene fragment and the second to add Illumina MiSeq adapters and sample indexes. The reaction mix for the first PCR contained 5 µl of Q5® reaction buffer (New England BioLabs), 0.5 µl of Deoxynucleotide Solution Mix (dNTP, New England BioLabs), 0.5 or 1.25 µl of forward and reverse primers, 0.25 µl of Q5® High-Fidelity DNA Polymerase (New England BioLabs), 1, 2 or 3 µl, depending on the concentration, of DNA or cDNA templates and the mix was completed to 25 µl with UltraPure DNase/RNase-Free Distilled Water (Invitrogen™). Successful amplification of some of the extracts required dilution 10 to 50 times to overcome PCR inhibition. PCR conditions for the 18S were 98°C for 30 s; 30 cycles including 98°C for 10 s, 55°C for 30 s and 72°C for 30 s; and finally 4.5 min at 72°C. The PCR conditions for the 16S were 98°C for 30 s; 30 cycles including 98°C for 10 s, 50°C for 30 s and 72°C for 30 s; and finally 5 min at 72°C. The second PCR reactions contained 10 µl of Q5® reaction buffer (New England BioLabs), 1 µl of dNTP (New England BioLabs), 1 µl of each Illumina Index, 0.5 µl of Q5® High-Fidelity DNA Polymerase (New England BioLabs), 1 or 2 µl of PCR 1 product and was completed to 50 µl with UltraPure DNase/RNase-Free Distilled Water (Invitrogen™). The PCR conditions were 98°C for 30 s, 13 cycles including 98°C for 10 s, 55°C for 30 s and 72°C for 30 s; and finally 4.5 min at 72°C. PCR products were verified on a 1% agarose gel and purified using magnetic beads (AMPure XP, Beckman Coulter) after each step. Each PCR included a negative control using the reaction mix without DNA template. Quantification and quality check of the purified second PCR product were done using a Spark® multimode microplate reader (Tecan). PCR products were then pooled equimolarly separately for 18S or 16S, purified and then sequenced on an Illumina MiSeq system, using V3 sequencing chemistry, at the Plateforme d'Analyses Génomiques (IBIS, Laval University, Québec).

Sequences were demultiplexed by the Plateforme d'Analyses Génomiques. Forward and reverse read pairs were merged using BBMerge (v.38.44, Bushnell et al., 2017). Primers were trimmed from the merged reads using fastx_truncate on

USEARCH (v.11.0.667, Edgar, 2010), followed by a quality check using FastQC (v.0.11.8, Andrews, 2018) and quality filtering (maxee: 0.5) using fastq_filter on vsearch (v.2.5.0, Rognes et al., 2016). The 16S sequences were quality filtered with Trimmomatic (Bolger et al., 2014) after the removal of primers and before the quality filtering on vsearch. Sequences were dereplicated and sorted using vsearch (v.2.5.0, Rognes et al., 2016), and chimera and singletons were removed from the sequences before being clustered into Operational Taxonomic Units (OTUs) with a similarity threshold of 99% for the eukaryotes and 97% for Bacteria and Archaea, using USEARCH (v.11.0.667, Edgar, 2010). Taxonomy was assigned with the Wang method using mothur (v.1.41.3, Schloss et al., 2009) on the Silva database (v.138, Pruesse et al., 2007; Quast et al., 2013) for Archaea and Bacteria, and the PR2 database (v.4.12.0, Guillou et al., 2013) for microbial eukaryotes. OTUs identified as eukaryotes (204 OTUs), chloroplasts (347 OTUs), domain unknown (379 OTUs) and mitochondria (57 OTUs) were removed from the final 16S OTU table, and those identified as metazoa (55 OTUs) and Embryophyceae (7 OTUs) were removed from the final 18S OTU table. OTUs originating from DNA sequencing are referred to here as rDNA, those from cDNA sequencing as rRNA and those coming from the 0.22-µm and 3-µm filters as small and large, respectively. The sequences reported in this paper have been deposited in NCBI Sequence Read Archive under the BioProject accession number PRJNA744875. OTU tables and fasta files are provided in the **Supplementary Material** (**Supplemental Data Files 1–3**).

Statistical Analysis

All statistical analyses were performed using the software R (v.3.6.3, R Core Team, 2020) in RStudio environment (v.1.2.5001, RStudio Team, 2019). All randomized calculations were made using a set.seed value of 007. Chemical values below the level of detection were considered to be 0 for all statistical analyses. Principal component analysis (PCA) was computed using centered and scaled environmental parameters and was represented using scaling type 1, which preserved the distance between sampling points. Missing values (dissolved oxygen for P2, TSS for G3 and $SUVA_{254}$, a_{320} normalized to DOC for P2) were estimated prior to computed to PCA (function imputePCA, missMDA).

The datasets were rarified for the alpha diversity estimates and we note that this alpha diversity underestimates the full diversity in the samples given sequencing bias and the rarefaction itself. All other statistical analysis were performed on the non-rarefied dataset, after confirming that use of this dataset produced the same beta diversity results, as visualized on the NMDS, as using the rarefied dataset. The microbial eukaryote dataset was rarefied to 3,182 reads per sample and bacterioplankton dataset was rarefied to 34,665 reads per sample (with an average of 3,353 and 8,513 total OTUs respectively). The rarefaction and alpha diversity were calculated 100 times using the rarefy_even_depth and estimate_richness function of the phyloseq package (McMurdie and Holmes, 2013) and we report the average diversity estimate. To evaluate relationships between physicochemical parameters or alpha diversity to salinity (in the plume) or physical distance between sampling points (in the GWR), we performed Spearman correlations using the cor.test function. To compare alpha diversity between size fractions, we used Kruskal–Wallis and *post hoc* Dunn's test with adjusted *p* values using the Benjamini–Hochberg correction (FSA package, Ogle et al., 2020).

Beta diversity indices for the microbial eukaryote and bacterioplankton communities were determined for the rDNA reads using a Bray–Curtis dissimilarity matrix and were visualized on a non-metric multidimensional scaling plot (NMDS). The Bray–Curtis dissimilarity was calculated using proportional abundance of each OTU within each sample. Analysis of similarity (ANOSIM) and a permutational multivariate analysis of variance (ADONIS, 999 permutations) were used to test differences between fractions. Prior to performed ADONIS, the homogeneity of dispersion within a group was evaluated using the vegan betadisper function. Bray–Curtis dissimilarity, ANOSIM and ADONIS were calculated using the vegan package (Oksanen et al., 2019). Bray–Curtis and ANOSIM gave value ranging from 0 to 1, with 0 meaning the same community composition and 1 a completely different community.

The distance–decay relationship in the GWR was determined with a linear regression between the log-transformed (log10) Bray–Curtis values and the distances between pairs of samples. We excluded dissimilarities between G1 triplicates, which were collected within a few meters of each other, to avoid giving weight to small distances in the regression. The regression was performed using the lm function and the assumptions tested using the package gvlma (Peña and Slate, 2006). We calculated the Bray–Curtis dissimilarity on the GWR samples separately for each fraction and nucleic acid. A distance-based redundancy analyses (db-RDA) using the Bray–Curtis dissimilarity matrix was used to quantify the variation in the small and large fraction of the bacterioplankton and the microbial eukaryotes communities (rDNA and rRNA) that could be explained by the physicochemical parameters (capscale function, vegan), followed by a variation partitioning of the significant factors (varpart function, vegan). There was some disparity between the salinity values from *in-situ* probe measurements and those from the Cubitainers™, determined back in the laboratory. We interpreted this as the variability in the plume position with the added complication of boat drift between environmental profiling and sample collection. For this reason, the profiling variables were excluded from the model (temperature, pH and dissolved oxygen). Environmental data were standardized by scaling to zero mean and unit variance (z-score standardization) and a Spearman correlation matrix was examined to reduced multicollinearity (cor function). Selection of explanatory environmental variables was made using a forward selection on adjusted R^2 values (ordiR2step function, vegan) and the variance inflation factor (VIF) was calculated to verify the collinearity between environmental data in the model. An ANOVA like permutation test (999 permutations) was used to assess the significance of the model. To identify OTUs differing in abundance ($p_a < 0.01$, adjusted using Benjamini–Hochberg correction) between small and large fractions, we performed a differential

abundance analysis using the DESeq2 package (Love et al., 2014) and used a script from Monteux et al. (2020) to present the results.

To identify the links between the microbial OTUs and environmental variables, we used a weighted correlation network analysis (WGCNA, Langfelder and Horvath, 2008) as previously described for amplicon datasets (Guidi et al., 2016; Henson et al., 2018; Horton et al., 2019). Prior to the analysis, OTU abundances were Hellinger-transformed and WGCNA was performed separately for the rDNA and rRNA, and for each fraction. A signed scale-free topology adjacency matrix, representing the strength of connection between two OTUs, was created based on OTU pairwise Pearson correlations across all samples raised to a soft power threshold. This adjacency matrix was used to construct a topological overlap matrix for hierarchical clustering analysis to cluster OTUs in modules (subnetworks) of highly connected OTUs. As recommended in the WGCNA tutorial, we merged modules with similar profiles (eigenvalue correlation higher than 0.7) since their OTUs have high co-occurrence. To identify modules most related to environmental variables, each module eigenvalue was pairwise correlated against environmental variables. We selected modules that were significantly and positively correlated with environmental variables and that explained most of the variation in community structure, as determined by the db-RDA, and identified OTUs with a significant ($p_a < 0.05$) positive correlation. All *p*-values were adjusted using the Benjamini–Hochberg correction for multiple correlations (p.adjust function, stats).

RESULTS

Environmental Characteristics

The measured physicochemical variables (**Supplementary Figure S1**) separated the sampling sites into three distinct groups (**Figure 2A**). The SAS and KWK rivers showed a greater influence of terrestrial organic matter input, as indicated by their high $SUVA_{254}$ values. Their low S_R and S_{289} values suggest a higher DOM molecular weight and a smaller contribution of autochthonous DOM. KWK and SAS were also characterized by higher DOC, CDOM (a_{320}) and nutrients (TP, TN, TDN) than the GWR and its coastal plume, and had higher bacterial and phytoplankton cell concentrations. In SAS, a decrease in concentrations of these variables was observed along the river. Several of these environmental variables were correlated, including DOC with CDOM, nutrients, chlorophyll-*a* and cell concentrations (**Supplementary Figure S2**).

The plume showed the effects of mixing of freshwater from GWR with seawater in the bay (**Supplementary Figure S3**). These coastal waters had the highest DIC concentrations of all samples analyzed, with salinity values that ranged from 1.16 to 14.86. The plume samples had higher TN, but lower TP and bacterial cell concentrations than the GWR. CDOM concentrations correlated negatively with the salinity, while its molecular weight (as indicated by S_R values) correlated positively ($\rho = -0.93$, $p < 0.001$, and $\rho = 0.80$, $p = 0.01$, respectively; **Supplementary Figure S4**). There were higher concentrations of suspended sediments in the plume, and pH increased with increasing salinity.

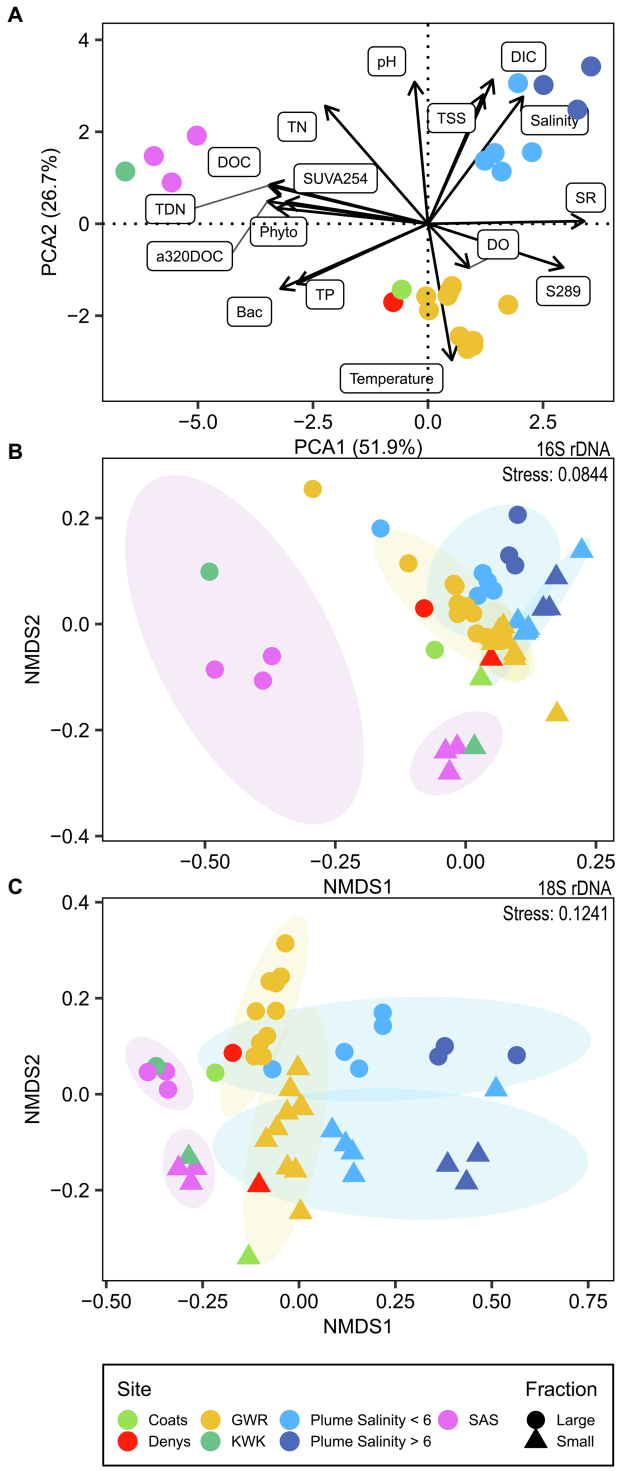

FIGURE 2 | Ordination of environmental variables and microbial communities. **(A)** Principal component analysis, scaling type 1, of standardized physico-chemical parameters; **(B)** NMDS based on Bray–Curtis distance for the bacterioplankton community (16S rDNA); **(C)** NMDS based on Bray–Curtis distance for the microbial eukaryotes (18S rDNA).

An increase in TSS concentrations was observed with distance downstream in the GWR, while there was little change in DIC, DOC, CDOM, nutrients and bacterial cell concentrations throughout the 150 km freshwater reach of the river. Phytoplankton cell concentrations in the GWR increased after the confluence of the Coats and Denys rivers and increased over the sampled reach of the river by a factor of two. The environmental characteristics of Coats and Denys rivers were similar to GWR except for higher DOC, chlorophyll-*a* and phytoplankton cell concentrations, and lower bacterial cell concentrations. Values of pH increased along the SAS river and the GWR until near Kuujjuarapik-Whapmagoostui (G7 and G8), where a decrease was observed. Oxygen concentrations were generally near or slightly above saturation at all sampled stations, except for the SAS mouth and site P3 in the plume (**Supplementary Table S2**).

Phototrophic Pigment Composition

The HPLC analyses showed that a broad range of pigments were present at all stations (**Supplementary Figures S5**, **S6**), specifically (in addition to chlorophyll-*a*): chlorophyll-*b* and -*c* (c_1 and c_2), Mg 2,4-divinylpheoporphyrin (MgDVP), β,Ɛ-carotene, β,β-carotene, peridinin, fucoxanthin, 9-cis-neoxanthin, violaxanthin, diadinoxanthin, antheraxanthin, alloxanthin, zeaxanthin, and lutein with trace levels also of echinenone in many samples. The dinoflagellate pigment peridinin, and the cryptophyte pigments crocoxanthin and alloxanthin were conspicuously higher (per unit chlorophyll-*a*; concentration ratio) in the plume at intermediate salinities (salinities 7.59, 8.39, 10.07; site P2). Lutein (green algae) and zeaxanthin (green algae and cyanobacteria) occurred in peak concentrations at the KWK river station, while the SAS samples had higher concentrations of violaxanthin (green algae but also other groups) compared to other sites. Fucoxanthin dropped to low values with an increase in salinity up to 10.07, rose again at salinity 14.86, and was also at high concentrations in the Coats River. Violaxanthin and lutein decreased with increasing salinity. The freshwater cyanobacterial pigment canthaxanthin showed a general trend of increasing concentration down the GWR to a maximum at G3, but was absent from the plume at salinities higher than 5.09. Similarly, the related pigment astaxanthin occurred throughout the GWR, but was absent from the plume and the SAS River. In general, highest concentrations of photoprotective carotenoids and photosynthetic pigments (including chlorophyll-*a*) occurred in the SAS and KWK rivers. The full HPLC pigment data set is archived in Blais et al. (2021).

Microbial Community Composition and Diversity

After the quality filtering, we retained 11,696,390 reads (121,837 ± 48,965, mean ± SD, reads per sample; $n = 96$) for bacterioplankton (here including Archaea and Bacteria) and 2,312,415 reads (22,213 ± 12,774; $n = 96$) for microscopic eukaryotes and fungi (collectively referred to as microbial eukaryotes), which clustered into 8,673 Bacteria and Archaea and 3,383 eukaryote OTUs, respectively. The observed richness values (number of OTUs) were typically at least two times greater for the bacterioplankton than the microbial eukaryotes (**Supplementary Figure S7**; **Supplementary Tables S3** and **S4**). The rDNA bacterioplankton richness of the large fraction was systematically greater than in the small fraction ($p < 0.01$, Kruskal–Wallis), and the small fraction diversity increased along the GWR (rDNA: $\rho = 0.64$, $p = 0.04$, not significant for rRNA; **Figure 3A**). For the microbial eukaryotes, the small fraction richness was greater for SAS and KWK rivers ($p = 0.02$), but no consistent trend was observed in the plume nor in the GWR. The large fraction richness increased along the GWR ($\rho = 0.85$, $p < 0.01$; **Figure 3A**) and decreased with the salinity in the plume ($\rho = -0.83$, $p = 0.02$; **Figure 3B**); these trends were similar for the rRNA but not significant.

Archaea accounted for less than 0.08% of the reads for each sample, with a total of 94 OTUs. These were mainly dominated by the Order Woesearchaeales, with a markedly higher proportion in the SAS river (**Supplementary Figure S8**). The archaeal reads included the nitrifier *Nitrosopumilus* and methanogens such as *Methanoregula*, *Methanosaeta* and *Methanobacterium*.

The bacterial community was composed of 51 phyla, of which 22 each accounted for more than 0.1% of the reads (**Supplementary Figure S9**). Bacterial communities (both rRNA and rDNA) were predominantly composed of Proteobacteria (mostly Burkholderiaceae and Comamonadaceae), Actinobacteriota (mostly Sporichthyaceae), Bacteroidota (previously known as Bacteroidetes; mostly Chitinophagaceae) and Verrucomicrobiota (mostly Pedosphaeraceae). The relative abundance of Actinobacteriota was consistently lower in rRNA, potentially due to their low growth rates. Cyanobacteria contributed less than 0.5% of the bacterial reads at most sites, but contained diverse taxa including the picocyanobacteria *Cyanobium* and *Synechococcus*, filamentous oscillatorians such as *Pseudanabaena*, *Tychonema* and *Leptolyngbya*, colonial taxa such as *Snowella*, and nitrogen fixing genera such as *Aphanizomenon*, *Anabaena* (*Dolichospermum*), *Nostoc*, *Rivularia* and *Calothrix* (**Supplementary Figure S10**).

The microbial eukaryote communities (both rRNA and rDNA) were composed of nine major taxonomic groups dominated by Alveolata (mostly Dinoflagellata, Ciliphora and *Perkinsea*) and Stramenopiles (mostly Ochrophyta) and to a lesser extent Hacrobia (mostly Cryptophyta, Centroheliozoa and Katablepharidophyta). In addition, there were representatives of Opisthokonta (mostly Choanoflagellida), Rhizaria (mostly Cercozoa) and Archaeplastida (mostly Chlorophyta; **Supplementary Figure S11**). OTUs unique to the plume included the Mamiellales picochlorophytes *Bathycoccus prasinos* and *Micromonas* (clade B3), and the diatom *Skeletonema*.

Core Microbiome

The bacterial component of the core microbiome was defined as OTUs present in both the rRNA and the rDNA reads at all sites (**Figure 4**, Core OTU tables are available in the **Supplementary Data File 3** as **Supplementary Tables S5–S8**). This consisted of 198 OTUs for the small fraction (representing 15–77% of the reads and 2.42% of the total OTUs of this fraction) and 322 OTUs for the large fraction (36–79% of the

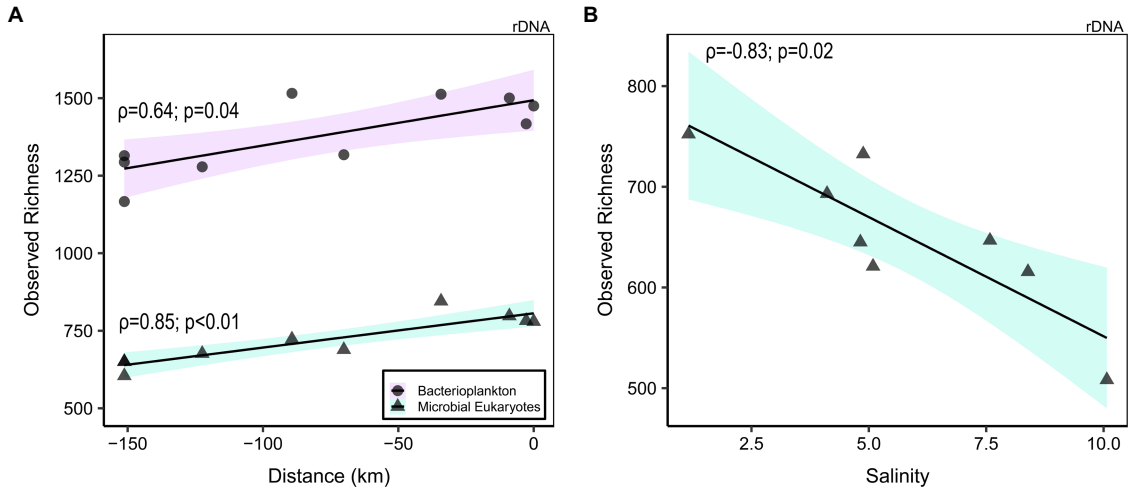

FIGURE 3 | Richness of bacterioplankton (rDNA small fraction; circles) and microbial eukaryotes (rDNA large fraction; triangles) in the Great Whale River as a function of distance downstream (**(A)**; 0 km = river mouth) and salinity in the plume **(B)**. Linear regression was plotted to highlight the correlation.

reads and 3.77% of the total OTUs of this fraction). This core microbiome was represented by 11 phyla: Proteobacteria, Bacteroidota, Actinobacteriota, Bdellovibrionota, Myxococcota, Planctomycetota, SAR324 clade (Marine group B), Verrucomicrobiota for both fractions and Cyanobacteria (*Pseudanabaena PCC-7429* and *Snowella*), Desulfobacterota, Chloroflexi, for the large fraction. The core was predominantly represented by *Polynucleobacter*, *Sediminibacterium* (Chitinophagaceae), unclassified Sporichthyaceae, unclassified Comamonadaceae and the NS11-12 marine group.

The microbial eukaryote communities showed much greater variability in OTU composition, and our bacterial definition of a core microbiome could not be applied. This variability included the over-representation of Ciliophora for some rRNA samples, and high read proportions for unclassified eukaryotes in the Coats River (rDNA, small) and in some GWR samples (G6 rRNA, small, G8 rRNA small and large). To adjust for these outliers, we therefore defined the microbial eukaryote core microbiome as OTUs present in the both the rRNA and the rDNA reads in 90% of the samples. The microbial eukaryote core microbiome was classified into six broad groups (Alveolata, Archeaplastida, Opisthokonta, Rhizaria, Stramenopiles, and the polyphyletic Hacrobia) and consisted of 57 OTUs for the small fraction (0.01–43% of the reads and 1.74% of the total OTUs of this fraction) and 48 OTUs for the large fraction (0.47–57% of the reads and 1.45% of the total OTUs of this fraction; **Figure 5**). In the plume, the proportion of core reads generally decreased with increasing salinity. The most abundant OTUs were members of the class Perkinsida and Chrysophyceae in the small fraction, and Chrysophyceae and the genus *Choanocystis* in the large.

Beta Diversity and Distance–Decay Relationship

Beta diversity of the rDNA OTUs revealed a weak separation between small and large fractions of the bacterioplankton and microbial eukaryote communities (ADONIS $p < 0.001$ for both and ANOSIM $R = 0.2754$, $p = 0.001$ for bacterioplankton and $R = 0.3021$, $p = 0.001$ for microbial eukaryotes). There was clustering of beta diversity among sites (**Figures 2B, 2C**). The GWR communities tended to cluster together, as did the plume communities where they were more similar within the same salinity range, except for the lowest salinity (1.16) communities where the large fraction clustered with the GWR and the small fraction was apart from the others. SAS and KWK communities clustered, together except for the large fraction of the bacterioplankton community, while communities from Coats and Denys rivers clustered between GWR and SAS/KWK.

Distance–decay relationships based on the transformed Bray–Curtis dissimilarity (log10) and the distance between the sampling points of the GWR were investigated with linear models. Both of the rDNA fractions for microbial eukaryote communities (rDNA large fraction $R^2 = 0.63$, $p < 0.001$; small fraction $R^2 = 0.34$, $p < 0.001$; **Supplementary Figure S12**) showed a significant distance–decay relationship, meaning that community dissimilarity increased with increasing distance between sampling points. The microbial eukaryote rRNA showed a similar relationship, but the high proportion of Ciliophora reads (especially in the large fraction) had a strong influence. As for bacterioplankton, the distance–decay relationship was significant for the rDNA large fraction ($R^2 = 0.16$, $p = 0.008$), however, the linear regression assumptions were not satisfied for the rDNA small fraction, mainly because of the sampling site G1.1 that had a higher dissimilarity irrespective of distance to other sites. For the rRNA sequences, the distance–decay relationship was significant for both fractions (rRNA small fraction $R^2 = 0.28$, $p < 0.001$ and large fraction $R^2 = 0.32$, $p < 0.001$; **Supplementary Figure S13**).

Differences Between Fractions

To evaluate differences between size fractions, we performed a differential abundance analysis on the rDNA sequences. There were 1,309 OTUs (16% of rDNA OTUs) that were significantly different between fractions for the bacterioplankton and 648

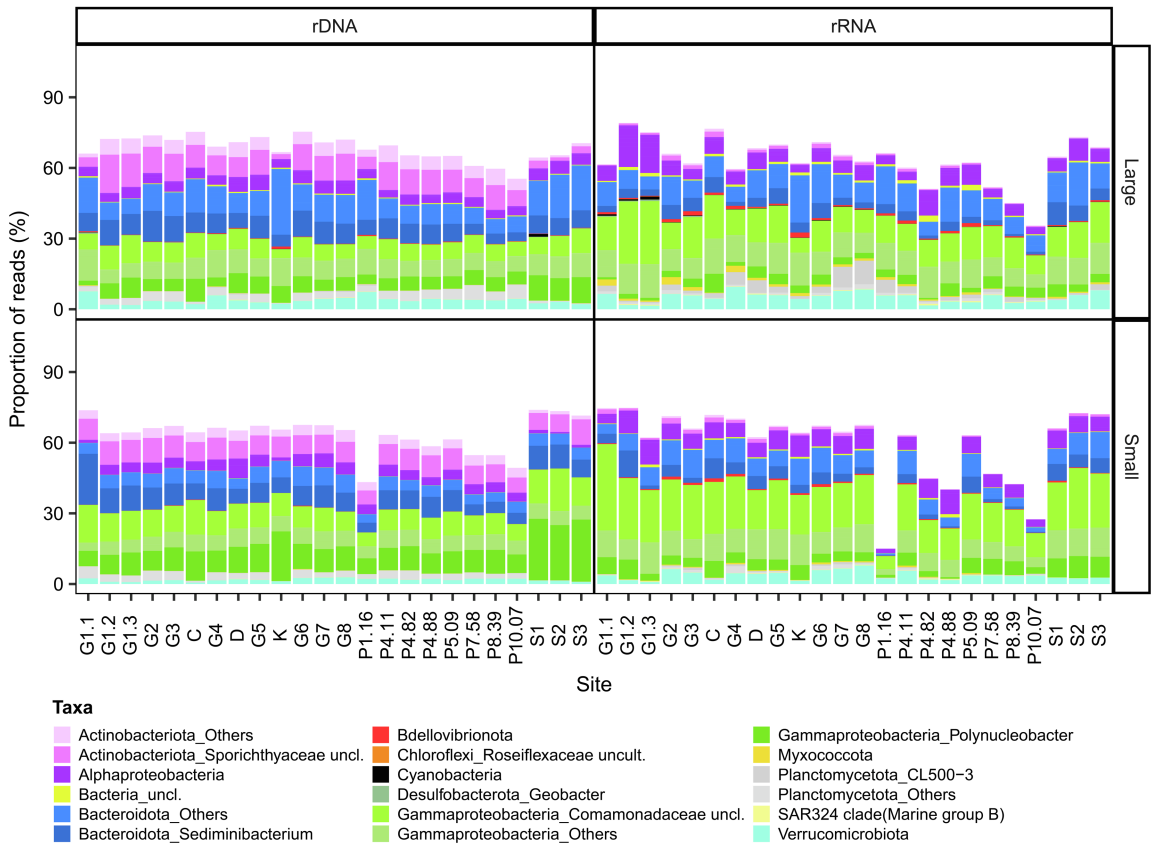

FIGURE 4 | Bacterial composition of the core microbiome in GWR and associated waters. The values are % reads for 16S rDNA or rRNA in the large **(top)** and small **(bottom)** fractions.

OTUs (20% of rDNA OTUs) for the microbial eukaryotes (**Figure 6**). The bacterioplankton OTUs that were differentially abundant were mainly enriched in the large fraction, often to a substantial extent. Phyla enriched in the large fraction included Acidobacteriota, Armatimonadota, Bdellovibrionota, Desulfobacterota and Planctomycetota, while unclassified Burkholderiales (Gammaproteobacteria) had a higher number of OTUs enriched in the small fraction. For the microbial eukaryotes, OTUs with the greatest differences between fractions were mainly from the Centroheliozoa, Ciliophora, Dinoflagellata and Ochrophyta taxa, which were more represented in the large fraction, and *Perkinsea*, Chlorophyta, Dinoflagellata (*Prorocentrum*) and Cercozoa (Filosa-Thecofilosea) in the small fraction.

Environmental Drivers of Microbial Community Structure

Distance-based redundancy analysis (db-RDA) showed that salinity and DOC were environmental factors that contributed significantly to the variation in microbial eukaryote and bacterioplankton community composition across all sampling sites ($p < 0.001$). The variation explained (adjusted R^2) was higher for the rDNA and for the large fraction in both the bacterioplankton (rDNA large 51%, small 48%; rRNA large 37%, small 27%) and the microbial eukaryotes (rDNA large 52%, small 40%; rRNA large and small 28%). DOC explained a greater proportion of the variation in bacterioplankton rDNA and the small fraction of microbial eukaryote rRNA, whereas salinity explained more variation in the bacterioplankton rRNA and all other microbial eukaryote samples (**Figure 7**).

To identify groups of OTUs (subnetworks) correlating with these potential environmental drivers, we used a weighted correlation network analysis (WGCNA) that was computed separately for bacterioplankton and microbial eukaryotes, and for each fraction and nucleic acid. We then considered OTUs only if they had a significant correlation with DOC or salinity ($p_a < 0.05$) and if they were part of a subnetwork that positively and significantly correlated with DOC or salinity (**Figure 8, Supplementary Figures S14–S21**). The number of OTUs correlated with salinity for the bacterioplankton large fraction (rDNA, 273; rRNA, 321 OTUs) was more than double that for the small fraction (rDNA, 106; rRNA, 117 OTUs), with a greater number of OTUs principally in the phyla Bacteroidota and Proteobacteria (Gammaproteobacteria and Alphaproteobacteria). The phylum diversity of OTUs correlated to DOC was greater than for salinity (mean 18.5 vs. 10), and was mainly taxa in the Bacteroidota, Proteobacteria (Gammaproteobacteria) and Verrucomicrobiota. More eukaryotic OTUs were correlated with DOC in both the small fraction (rDNA, 208; rRNA, 232) and

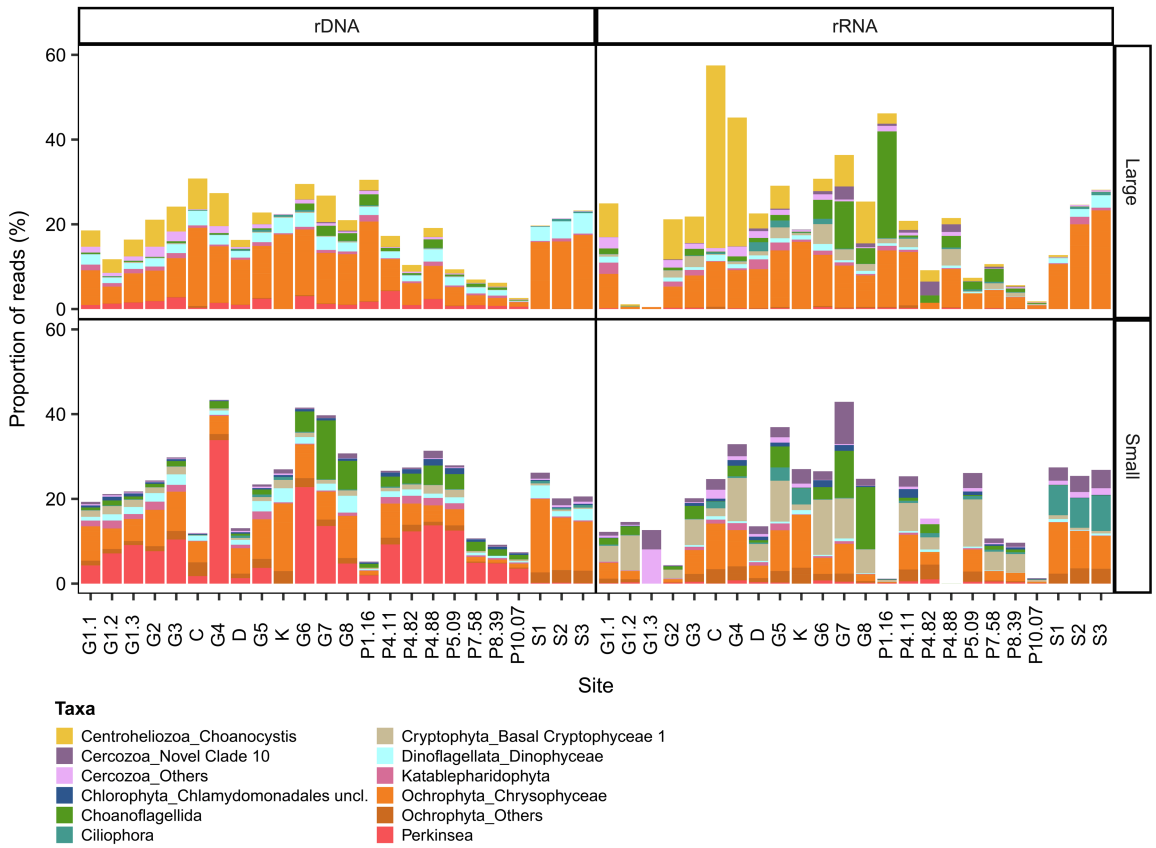

FIGURE 5 | Eukaryotic composition of the core microbiome in GWR and associated waters. The values are % reads for 18S rDNA or rRNA in the large **(top)** and small **(bottom)** fractions.

large fraction (rDNA, 135; rRNA, 130 OTUs), with greater representation of the divisions Ochrophyta and Cercozoa. Dinoflagellata, Ochrophyta and Chlorophyta represented more than half of the eukaryotic OTUs that correlated with salinity.

DISCUSSION

Environmental Characteristics

Our sampled environments encompassed a broad range of salinities (0.01 to 14.86), temperatures (11.65 to 19.48°C), TSS (0.96 to 10.89 mg L^{-1}), DOC (2.86 to 8.16 mg C L^{-1}), TP (<5 to 26.14 µg P L^{-1}) and TN (170.30 to 404.48 µg N L^{-1}). As hypothesized, permafrost thaw related variables differed among subcatchments, with higher values (except for the TSS) in the SAS and KWK rivers that pass through degraded permafrost landscapes. There was little change over the 150-km reach of GWR in TN, TP and DOC (including CDOM optical variables), suggesting relatively homogenous landscape conditions over much of this section of the catchment, despite some shifts in vegetation and local fire effects (Payette et al., 1989; Bhiry et al., 2011). This relative stability of the GWR environment indicates substantial dilution of the three tributary rivers (KWK, Coats and Denys), which all had higher DOC concentrations than the main stem of the river. There was a downstream increase in TSS along the GWR transect (especially at site G6) that may indicate decreased landscape stability and increased erosion in the lower reaches of the river. Geomorphological analysis of the downstream river channel shows evidence of increased recent landslide activity, possibly linked to increasing extreme rainfall events (Owczarek et al., 2020), and a major landslide occurred 8 km upstream from the river mouth in April 2021 (Nozais et al., 2021). There was a further upshift in TSS between the most downstream freshwater sites (G7, G8) and the plume in coastal Hudson Bay, which may be the result of tidal resuspension of shallow sediments (Ingram, 1981), as well as flocculation of particles induced by the mixing with saline waters (Mosley and Liss, 2020).

The two small rivers (SAS and KWK) differed from the other waters in their physicochemical characteristics, which may in part be due to their lesser size (Wologo et al., 2021), but most likely reflected the influence of severely degrading permafrost in their catchments (Arlen-Pouliot and Bhiry, 2005; Bouchard et al., 2014). These rivers had the highest nutrient, DOC and CDOM (a_{320}) values, and optical analysis of the CDOM (S_{289}, $SUVA_{254}$, S_R) showed that this carbon was dominated by aromatic compounds of higher molecular weight, indicative of terrestrial rather than autochthonous sources. This allochthonous dominance of DOC is consistent with the allochthonous CDOM previously

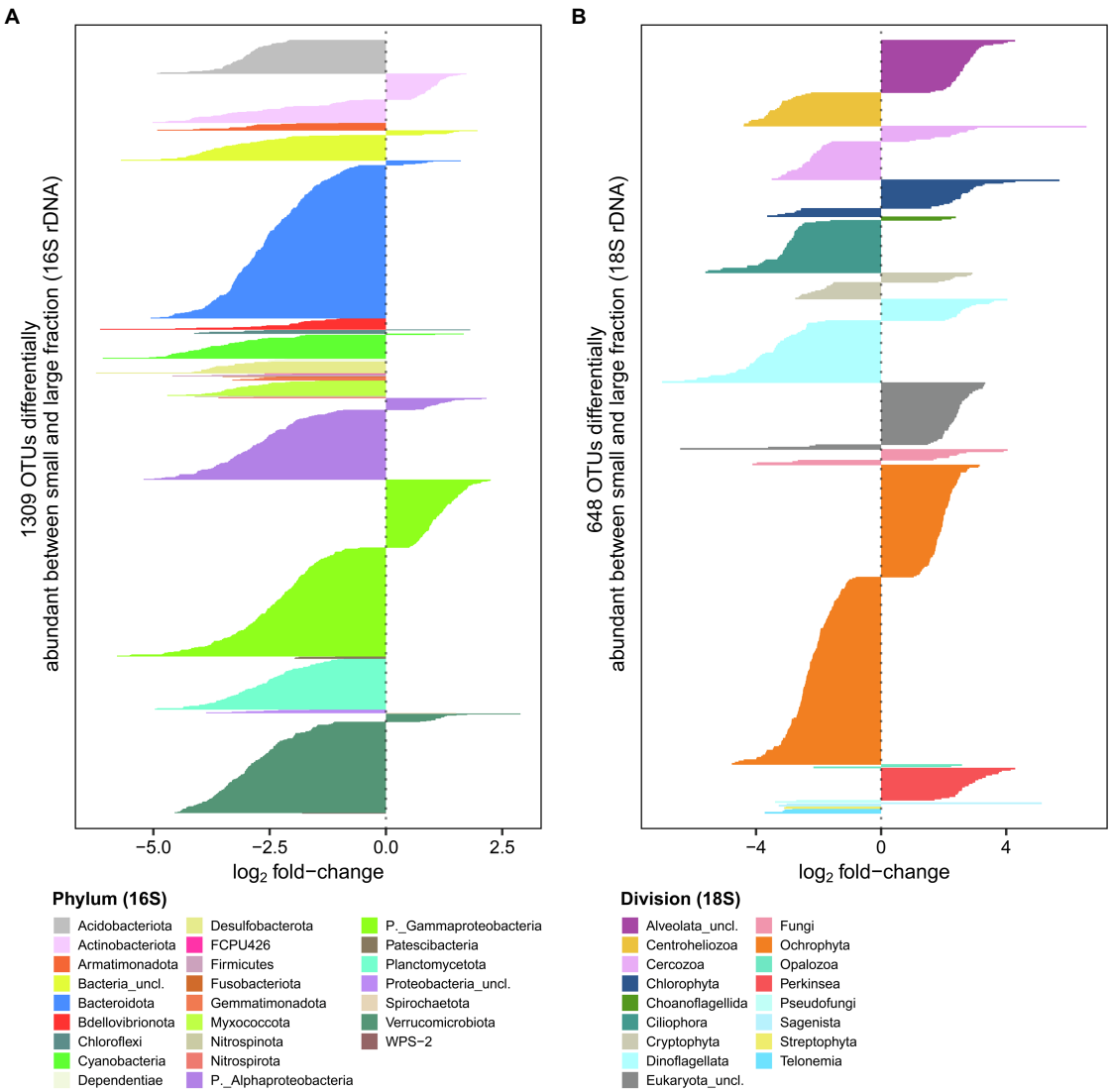

FIGURE 6 | Comparison of large and small fractions for bacterioplankton **(A)** and microbial eukaryotes **(B)**. Negative values represent rDNA OTUs more abundant in the large fraction, while positive values represent OTUs more abundant the small fraction.

observed in thermokarst lakes of the SAS and KWK river valleys and attributed to permafrost degradation (Wauthy et al., 2018). The SAS valley contains permafrost peatlands, and peatlands in general have higher DOC export and concentrations than mineral soils (Dillon and Molot, 1997). Radiocarbon analysis of DOC in the SAS River during winter has indicated much older carbon than in nearby non-peatland rivers, and thermokarst lakes in the KWK River watershed (Gonzalez Moguel et al., 2021), suggesting the mobilization of old frozen peat deposits. In a study of streams flowing through discontinuous permafrost, higher concentrations of DON, as well as higher DOC and lower DOC quality, were associated with a greater extent of permafrost in the watershed (Balcarczyk et al., 2009). The underlying permafrost at KWK and SAS is composed of nutrient-enriched marine clays (Table 2 in Deshpande et al., 2016 for a SAS palsa), and thawing of these soils may have contributed to the higher TN and TP values observed in the KWK and SAS rivers.

Pigment Composition

A wide range of accessory pigments occurred in all samples, indicating that the phototrophs at all sites had major contributions of taxa from diverse phyla, including cyanobacteria (canthaxanthin), dinoflagellates (peridinin), chlorophytes (lutein), prasinophytes (MgDVP), ochrophytes (fucoxanthin) and cryptophytes (crocoxanthin and alloxanthin). This diversity is consistent with HPLC pigment studies in the lakes of the region, including thaw lakes in the KWK and SAS valleys (Przytulska et al., 2016). The plume itself had a distinctive set of pigment signatures that indicated an abundance of cryptophytes and peridinin-containing dinoflagellates. This is in accord with our molecular analyses

in which we detected a greater abundance of dinoflagellates in the plume, particularly the peridinin-rich *Heterocapsa pygmaea*, and although the relative sequence abundance of cryptophytes was not greater in the plume, the molecular data showed the presence of the kleptoplastidic species *Mesodinium rubrum*, which contains cryptophyte chloroplasts and their associated pigments (Gustafson et al., 2000). The conspicuous absence of canthaxanthin in the plume at salinities higher than 5.09 suggests the loss of a subset of freshwater cyanobacteria in response to mixing with seawater. This pigment is especially common in filamentous freshwater taxa of cyanobacteria including members of the Nostocales (Przytulska et al., 2016), and such filaments could be differentially lost by grazing or flocculation in the plume. However, cyanobacteria of the family Nostocaceae were detected in the plume in the molecular data, albeit with a decrease in relative abundance, suggesting a limitation of the HPLC analysis. Chlorophyll-*a* concentrations in all samples were within the oligotrophic to mesotrophic range (0.36–3.08 µg chl *a* L^{-1}, trophic state index 21.14–40.96; Carlson, 1977). Maximum chlorophyll-*a* concentrations were in the SAS and KWK rivers, consistent with their higher phytoplankton cell concentrations as measured by flow cytometry.

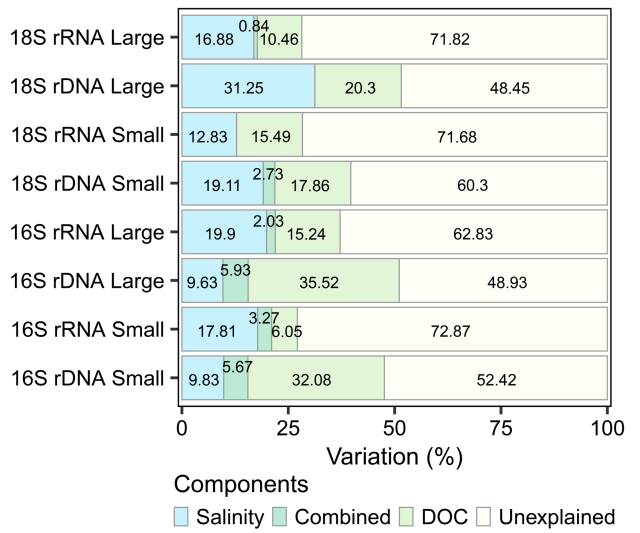

FIGURE 7 | Variation partitioning of the significant environmental drivers (salinity and dissolved organic carbon) for each dataset.

Core Microbiome

Despite the broad environmental gradients sampled here, a large fraction of the bacterioplankton reads consisted of bacterial

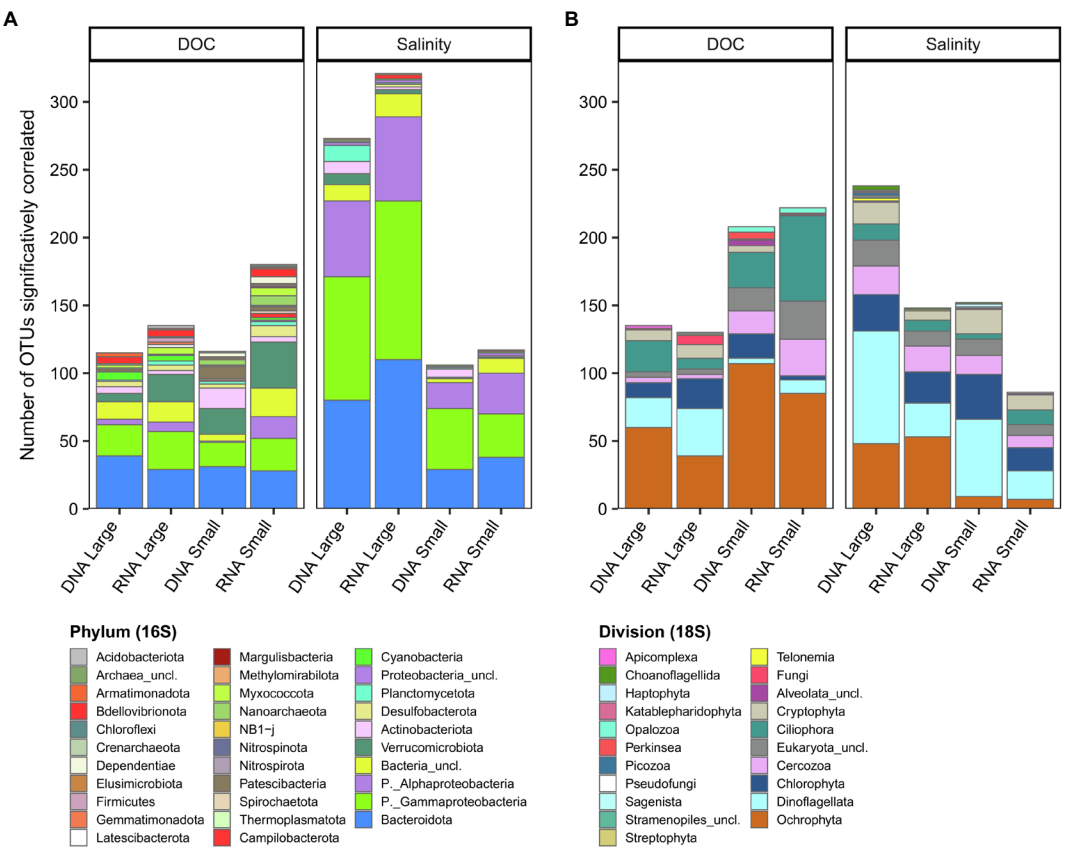

FIGURE 8 | Number of OTUs significantly correlated to DOC or salinity for each rDNA or rRNA fraction for bacterioplankton **(A)** and microbial eukaryotes **(B)**. The OTUs are grouped at the phylum level for bacterioplankton **(A)** and division level for microbial eukaryotes **(B)**.

OTUs that were ubiquitously distributed across all sites. This core microbiome represented a small percentage of the total richness (198 core OTUs for the small fraction and 322 core OTUs for the large fraction, versus 8,673 total OTUs), implying that a subset of generalist taxa can persist in the chemically diverse environments. These core taxa that were retrieved from the rRNA as well as rDNA indicate the potential for protein synthesis (Blazewicz et al., 2013) across this wide range of conditions. However, there was a decrease in the relative abundance of rRNA OTUs in this core assemblage with increasing salinity, which may indicate a decline in the physiological state of many taxa when they encounter seawater. The combined use of rRNA and rDNA provided a more reliable identification of this core assemblage, as rRNA is considered a more accurate guide to the presence of living cells (Li et al., 2017). Among the core microbiome OTUs, there were some well-known cosmopolitan taxa that are found in a wide range of freshwater environments, including *Polynucleobacter* (Hahn and Hoetzinger, 2019), *Limnohabitans* (Kasalický et al., 2013) and CL500-3 (Andrei et al., 2019), also identified in the core community of the Mississippi River (Henson et al., 2018). In addition to freshwater taxa, terrestrial taxa such as *Geobacter*, *Dinghuibacter*, and *Chthoniobacter*, and sediment taxa such as *Sediminibacterium* were also represented in the core, indicating a microbial contribution from soils and sediment to the river.

The eukaryotic core microbiome consisted of fewer OTUs compared to bacterioplankton and accounted for a small fraction of the total eukaryote reads (57 core OTUs for the small fraction and 48 core OTUs for the large fraction versus 3,383 total OTUs). This is consistent with the view that microbial eukaryotes are less metabolically plastic than bacteria and have a narrower niche breadth (Wu et al., 2018). Species sorting in the GWR plume with rising salinity may have especially restricted the number of taxa in this core subset. Eukaryotic OTUs in this core microbiome included taxa found commonly in freshwater such as the putative parasite *Perkinsea* (Lefèvre et al., 2008; Jobard et al., 2020), many Chrysophyceae such as the mixotrophic genus *Dinobryon*, which produces stomatocysts (Piątek et al., 2020) ensuring their survival in unfavorable conditions, and unclassified members of the freshwater Novel Clade 10 (Cercozoa), which contains biflagellate species that can feed on other flagellates (Bass et al., 2018).

Comparison of Size Fractions

The bacterial and microbial eukaryote communities showed differences between fractions, similar to aquatic ecosystems elsewhere (e.g., Crump et al., 1999; Savio et al., 2015; Henson et al., 2018; Wang et al., 2020), and many OTUs were significantly more abundant in one of the two fractions. Cyanobacteria in the genera *Synechococcus* and *Cyanobium* often account for a large proportion of the freshwater picophytoplankton (Callieri, 2008), and in the present study two cyanobacterial OTUs, belonging to the Synechococcales order (*Cyanobium PCC-6307* and unclassified Cyanobiaceae) were enriched in the small fraction. Previous work on the GWR has identified phycocyanin-rich cells of this type in concentrations of around 10^4 cells ml^{-1} and although comparable to other phytoplankton, these cell counts were low relative to the bacterial cell counts in the river of around 10^6 cells ml^{-1} (Rae and Vincent, 1998), as expected. In the present study, most of the cyanobacteria were associated with the large fraction, in keeping with their tendency to form colonies (e.g., *Snowella*) or to be in filamentous form (e.g., *Aphanizomenon*). Eight Synechococcales OTUs (*Cyanobium PCC-6307*) were also more abundant in the large fraction, indicating their tendency to from aggregates (Jezberová and Komárková, 2007; Huber et al., 2017). Certain well-known soil bacteria such as *Citrifermentans*, *Geothrix* and *Geobacter* were enriched in the larger fraction, suggesting an attached lifestyle. These taxa may also have entered the river along with soil particles washed in from the surrounding catchment.

Eukaryotic taxa that sorted according to differences in cell size included picoeukaryotes such as the chlorophyte *Micromonas*, which occurred in the small fraction. In the study by Jacquemot et al. (2021), this common Arctic Ocean genus (Lovejoy et al., 2007) was abundant throughout coastal Hudson Bay. However, larger taxa were also found in the small fraction, which could relate to cell breakup and transfer of cellular debris through the 3 µm filters, or for some species, may be the result of life cycle stages in the smaller fraction. For example, *Perkinsea* produces zoospores in the size range 2–5 µm (Jobard et al., 2020), which would pass through the larger filter.

There was also substantial overlap of OTUs between size fractions for bacteria (rDNA, 77%; rRNA, 86%). It has been suggested that bacteria can alternate between free-living and particle-associated lifestyles (Grossart, 2010), causing similarity between those two communities (Hollibaugh et al., 2000). However, given that we observed the same high overlap for microbial eukaryotes (rDNA 90%, rRNA 87%), the co-occurrence of OTUs in both fractions could also possibly be explained by filtration artifacts such as filter failure, blockage of the 3-µm filters or cellular break-up (Padilla et al., 2015; Cruaud et al., 2019). It should also be noted that for our comparisons of size fractions, we pooled the entire datasets for each fraction to improve the statistical power. This approach may have masked stronger site-specific differences between fractions, as reported elsewhere. For example, there was a significant difference between size fractions among Amazon River plume samples, but no significant difference among river samples (Doherty et al., 2017). Similarly, in a transect of the Mackenzie River to the Beaufort Sea, a significant difference between fractions was only observed in the open sea samples (Ortega-Retuerta et al., 2013). Additionally, our 30 µm prefiltration may have biased comparison between fractions, as large particles and their associated microbial communities were excluded as well as large microbial aggregates and phytoplankton.

Diversity Patterns

In the GWR, microbial community structure became less similar as the geographic distance between sampling points increased. This distance–decay relationship is often caused by drift and selection, counteracted by dispersal (Hanson et al., 2012). Since most of the environmental characteristics of the GWR were homogeneous over the 150-km reach sampled, selection probably played a limited role. However, neutral processes such as immigration

of taxa from sub-catchment inputs (tributaries, groundwater and overland runoff), combined with high dispersal may have contributed to the modest distance–decay relationship observed. These inputs also likely drove the increase in richness in the free-living bacteria and large microbial eukaryote fraction with distance downstream, rather than growth, competition and replacement processes, given that the total transit time over this reach is relatively short (1 to 3.5 days at flows in the range 0.5–1.5 $m\,s^{-1}$). This wholesale mixing of different sub-communities from different source locations would be consistent with the concept of longitudinal coalescence in fluvial microbiomes (Mansour et al., 2018). No downstream trends in richness were observed in the particle-associated bacterial fraction, nor in the small eukaryotes, suggesting less heterogeneity in subcatchment inputs for these fractions. Decreases in bacterial diversity have been observed downstream of the headwaters of some systems (Crump et al., 2012; Savio et al., 2015), and such changes could potentially occur at the outflow of the lakes in the headwaters of the GWR, several hundred km upstream of our transect (7 days upstream at 1 $m\,s^{-1}$).

There was a decrease in richness in the large microbial eukaryote fraction with increasing salinity in the GWR plume. A similar decrease has been reported in the Chesapeake Bay and the Baltic Sea where diversity was higher in both fresh and marine waters and decreased at salinity around 7 to 9, following the Remane curve (Olli et al., 2019). Salinity is a strong filter for dispersal (Lozupone and Knight, 2007) and this observed decrease is likely the result of species sorting of freshwater taxa. Microbial eukaryotes appear to be more subject to species sorting relative to dispersal limitation than bacteria (Wu et al., 2018), which could potentially explain our observed absence of a diversity pattern for bacteria across the plume salinity gradient. Differences in our filtered water volumes could have affected community diversity (Padilla et al., 2015), however we found no correlation between richness and filtered volume in the GWR or plume.

Bacterial richness was higher for the particle-associated fraction than for the free-living fraction at all sites as reported elsewhere and has been attributed to environmental heterogeneity such as redox gradients within the suspended particles (Savio et al., 2015; Payne et al., 2017; Wang et al., 2020). Microbial eukaryote richness was higher for the small fraction in the SAS and KWK rivers suggesting that smaller eukaryotes could potentially be favored in these CDOM-rich waters. All of these analyses are subject to the usual limitations of primer mismatches and amplicon biases (e.g., Burki et al., 2021; Vaulot et al., 2021).

Salinity Effects

The plume community consisted of OTUs originating from the river, but also specific to the plume, with changes in community composition at salinities as low as 1.16, and a marked change between low (1.16 to 5.09) and moderate (7.58 to 10.07) salinity samples for some eukaryotic taxa. Since we did not sample offshore marine waters in Hudson Bay, we cannot differentiate brackish from marine taxa. However, the earlier study by Jacquemot et al. (2021) in the GWR estuary reported that the eukaryotic microbial community consisted of marine, freshwater, and specialized estuarine taxa. This mixing of taxa from different water sources and the establishment of a brackish population has also been observed for bacteria in transition zones elsewhere, such as in Parker River estuary and Plum Island Sound (Crump et al., 2004) and in the Baltic Sea (Herlemann et al., 2011).

Microbiome structure was significantly related to salinity, as hypothesized, but with less responsiveness of bacteria relative to eukaryotes. Many bacterial taxa occurred across the full range of salinities, however there was a greater number of bacterial OTUs correlated to salinity in the large fraction, suggesting some differentiation of particle-associated bacterial assemblages with increasing saltwater influence in the plume. Bacterial OTUs correlated with salinity were primarily members of the taxonomic groups Bacteroidota, Gammaproteobacteria and Alphaproteobacteria, which often reported as more abundant in the sea compared to freshwater (Bouvier and del Giorgio, 2002; Herlemann et al., 2011). In the GWR system, members of the Rhodobacteraceae family (Alphaproteobacteria), unclassified taxa and the genus *Plantomarina*, were highly correlated with salinity and were almost exclusively in our plume samples. Rhodobacteraceae sometimes dominate coastal marine particles (Bižić-Ionescu et al., 2015) and are abundant over a wide range of salinities including mesohaline waters (Zhu et al., 2018), a hypersaline coastal lagoon (Ghai et al., 2012), and in biofilms over marine macroalgae (Bengtsson et al., 2011; Dogs et al., 2017).

Bacteroidota in the families Cryomorphaceae and Flavobacteriaceae also correlated with salinity in the GWR system. Comparative genomic studies have revealed that adhesion to particles is a major property of marine Flavobacteria (Fernández-Gomez et al., 2013), and they are often abundant in marine particle-associated communities (DeLong et al., 1993; Ortega-Retuerta et al., 2013). These groups have the ability to degrade high molecular weight materials (Cottrell and Kirchman, 2000; Kirchman, 2002; Buchan et al., 2014) and are present during marine phytoplankton blooms and their decay, where they colonize aggregates and senescent phytoplankton cells (Abell and Bowman, 2005; Teeling et al., 2012). Moreover, Flavobacteria have been identified as indicator species for the plume environment in a Colombia River-plume-estuary study (Fortunato et al., 2013).

In the third bacterial group correlated with salinity, Gammaproteobacteria, we identified the OM43 clade, a methylotroph in the Methylophilaceae family (Giovannoni et al., 2008), and the SAR92 clade (family Porticoccaceae), which possess genes for the degradation of complex polysaccharides (Klindworth et al., 2014). These two clades are common in coastal waters (Rappé et al., 1997; Morris et al., 2006; Stingl et al., 2007), and a culture study has shown that their growth is stimulated by high molecular weight organic matter (Sosa et al., 2015).

Only one archaeal OTU, identified as an unclassified member of the Marine Group II (Thermoplasmatota), correlated with salinity. This group is the most abundant planktonic archaeal group in surface marine waters and likely plays a role in degradation of high molecular weight proteins and fatty acids (Pereira et al., 2019; Rinke et al., 2019; Tully, 2019).

There were strong relationships between eukaryotic OTUs and salinity, especially for Dinoflagellata, Ochrophyta and Chlorophyta. Diatoms (within the Ochrophyta) were among

the most important plume taxa in terms of relative abundance in the large fraction, notably the species *Skeletonema marinoi*. The genus *Skeletonema* is present in coastal and brackish waters (Kooistra et al., 2008; Paulino et al., 2018) and the ability of *Skeletonema marinoi* to growth at low salinity is variable among strains and populations (Balzano et al., 2011). There was a major shift in Dinoflagellata in the plume, with high relative abundance of unclassified Gymnodiniaceae at low salinity, and dominance of *Heterocapsa pygmaea* at moderate salinity. *H. pygmaea* is thought to be mixotrophic (Bretherton et al., 2020), and grazing by this species along with other phagotrophic protists may contribute towards the decline of bacterial cells in the plume. Unlike many marine dinoflagellates, *H. pygmaea* is well known to contain peridinin as an important light-harvesting pigment (Carbonera et al., 1999), consistent with our HPLC observation of this pigment in the plume.

At moderate salinity, OTUs affiliated with *Pyramimonas australis* dominated the Chlorophyta in the large fraction, and were also detected in the small fraction. *P. australis* is a marine species first identified in Terra Nova Bay, Antarctica (Moro et al., 2002) and subsequently identified in Baffin Bay in the Arctic (Gérikas Ribeiro et al., 2020). In addition, two marine picochlorophytes - the cosmopolitan *Bathycoccus prasinos* (Vannier et al., 2016) and *Micromonas* clade B3 - had highest relative abundances in the small fraction of the plume samples.

The mixotrophic ciliate *Mesodinium rubrum* also showed a relationship with salinity. This taxon was in highest abundance at salinity 10.07 and has been previously identified as the most abundant OTU in the GWR estuary at salinity 12.5 (Jacquemot et al., 2021). This species captures and retains plastids from the cryptophytes *Teleaulax gracilis* and *Plagioselmis prolonga* (Rial et al., 2015); these two cryptophytes also occurred in our sequences, mainly at moderate salinity in the GWR plume. As noted above, this presence of *Mesodinium* and the two cryptophytes is consistent with the HPLC data, with the peak in cryptophyte pigments crocoxanthin and alloxanthin at moderate salinity. *P. prolonga* has been recently identified as the haploid life stage of *T. amphioxeia* (Altenburger et al., 2020), also an important plastid donor to *M. rubrum* (Hansen et al., 2012). This haploid life stage appears to be more abundant in summer when irradiance, temperature and grazing pressure are high, but dissolved inorganic nitrogen concentrations are low (Altenburger et al., 2020).

Permafrost-Related Effects

We hypothesized that in addition to salinity control, there would be a significant effect of permafrost related variables on microbiome structure in the GWR system and its associated waters. Although TSS and TN did not explain any of the community variation, there were significant relationships with DOC. In our constrained ordination, we considered DOC as a proxy for the cluster of permafrost-related variables since it correlated with CDOM concentration (a_{320}) and CDOM quality (S_R, S_{289}, $SUVA_{254}$), as well as with TP and TDN (the latter two variables explaining less variation than DOC). Although we focus on DOC in this discussion, it should be kept in mind that its correlates, notably changes in DOC composition thus bioavailability, would also have some impact on community composition.

Unlike the salinity-dependent OTUs that were largely restricted to the plume, most of the OTUs that correlated with DOC occurred in all waters, but were especially abundant in the KWK and SAS rivers that pass through rapidly degrading permafrost catchments. Increased CDOM would result in decreased light availability, although light may not be a limiting factor in this environment, particularly given the regular exposure of cells to high irradiances by mixing to the surface in these relatively shallow turbulent rivers. The fact that SAS and KWK rivers had the highest phytoplankton cell concentrations, along with the absence of a positive relationship between chlorophyll-*a* per unit cell volume and DOC, also argue against CDOM-induced light limitation. In part, the positive relationship between OTUs and DOC-correlated variables may reflect the increased supply of land-derived nitrogen and phosphorus from permafrost soils, but the primary effect is likely to be *via* the availability of organic carbon as an energy source for microbial heterotrophs. Although this DOC is largely allochthonous and of high molecular weight, breakdown by photochemical processes at the surface of the water column (Cory et al., 2014) may ensure a continuous supply of more labile substrates. Also, some bacterial taxa such as certain Bacteroidetes and Verrucomicrobia can break down larger organic polymers (Kirchman, 2002; Cabello-Yeves et al., 2017).

The eukaryotic taxa correlated with DOC were mostly Ochrophyta, Chlorophyta, Dinoflagellata, Ciliophora and Cercozoa. Phagotrophy, in the form of mixotrophy or obligate heterotrophy, is a common feature among these groups, which would give them an advantage in higher DOC environments, where bacteria use DOC as a substrate. The ochrophytes that correlated with DOC were mainly identified as unclassified Chrysophyceae; this family includes many mixotrophs (Sanders and Porter, 1988) that can be stimulated by DOM enrichment (Daggett et al., 2015). The dinoflagellate OTUs correlated with DOC were mostly unclassified taxa, and also a species affiliated to *Asulcocephalium miricentonis* that was first described from a Japanese pond (Takahashi et al., 2015). *A. miricentonis* is a photosynthetic dinoflagellate containing peridinin, and is within the family Suessiaceae that contains mixotrophic taxa (Stoecker, 1999).

Bacterial OTUs correlated with DOC were more diverse than those correlated with salinity, including many OTUs that occurred in low relative abundance. The latter included methylotrophs and methanotrophic genera such as *Methylomonas*, *Methylotenera*, *Crenothrix*, *Candidatus methylopumilus* and *pLW-20*. Among the bacterial phyla, Gammaproteobacteria was among those with the highest number of OTUs correlated with DOC along with Bacteroidota and Verrucomicrobiota. Gammaproteobacteria are favored by high DOC concentrations and show preferences for terrestrial organic matter of high molecular weight (Amaral et al., 2016). The genus *Polynucleobacter* was part of the core microbiome, however 3 of the 19 OTUs significantly correlated to DOC with a greater relative abundance of these sequences in the small fraction. It has been hypothesized that some *Polynucleobacter* species are favored by humic matter degradation products (Hutalle-Schmelzer et al., 2010). A species affiliated to *Armatimonas* was among the most common OTUs (in relative abundance) in the rDNA large fraction of the SAS and KWK rivers, although it was in low relative abundance

in the rRNA, suggesting a slow growth rate. This group (Armatimonadota) has been detected in a groundwater culture enriched with a mixture of sediment organic matter and bacterial cell lysate, and may possess the metabolic potential to utilize recalcitrant organic matter (Wu et al., 2020).

Archaea were in greater proportion in the SAS River. Fourteen OTUs, mostly belonging to the order Woesearchaeales and identified as unclassified or as the family SCGC AAA011-D5, were correlated to DOC. Woesearchaeota are present in a variety of environments including freshwater lakes (Ortiz-Alvarez and Casamayor, 2016), permafrost (Shcherbakova et al., 2016), wetlands (Narrowe et al., 2017), and sediments (Liu et al., 2021). They have been reported as the dominant archaeal group in Swedish boreal lakes with higher DOM aromaticity (Juottonen et al., 2020), and in Greenland ponds and lakes characterized by higher organic and inorganic carbon and total nitrogen concentrations (Bomberg et al., 2019). Their small genome and limited metabolic capacities suggest a symbiotic or parasitic lifestyle (Castelle et al., 2015, 2021). In anoxic environments, they have a potential role in nitrogen and sulfur cycling, and a syntrophic relationship with methanogenic archaea has been suggested based on their high co-occurrence (Liu et al., 2018, 2021).

CONCLUSION

Numerous rivers drain the vast subarctic landscape, and our study of the Great Whale River and associated flowing waters draws attention to their importance as biological habitats for species-rich microbiomes. Our results show the presence of a core microbiome with less commonality among sites in the eukaryotes than in the prokaryotes. The high diversity of taxa that were more site-specific and of much lower relative abundance relative to this core assemblage suggests that these rarer, more specialized taxa may be potentially more vulnerable to ongoing changes in the catchment and river.

Our observations revealed differences between size-fractions in microbial community structure, and trends in richness along the GWR and its plume that differed among microbial groups, implying taxon-specific responses to the environment. The results highlight the distinctive nature of the coastal microbiome, with specific photosynthetic pigment characteristics and a strong influence of freshwater taxa as well as brackish water and fully marine species. Additionally, these results indicate the effects of degrading permafrost, with DOC explaining part of the variation in microbiome structure, and imply that mixotrophic and methanotrophic species may be favored by ongoing landscape thawing and erosion.

We analyzed the GWR microbiome in late summer, the time of maximum temperatures and high biological activity, and before the onset of fall cooling and freeze-up. Subarctic thaw lakes show marked changes in microbiome composition between summer and winter seasons (Vigneron et al., 2019), and the microbiome structure of Arctic rivers also changes markedly through the year (Crump et al., 2009). There is therefore a need to extend these summer observations of a subarctic river and its coastal plume to other seasons, including the winter period of prolonged ice-cover and the spring conditions of ice break-up and peak discharge.

DATA AVAILABILITY STATEMENT

The molecular datasets generated in this study are available in the NCBI online repository (https://www.ncbi.nlm.nih.gov/; accession number PRJNA744875) and the environmental and HPLC data are deposited in the northern environmental data repository Nordicana D (http://www.cen.ulaval.ca/nordicanad/en_index.aspx; doi: 10.5885/45741CE-38138EC6C8E849AD and doi: 10.5885/45660CE-8B92339884C146D0).

AUTHOR CONTRIBUTIONS

MB formulated the research and sampling design with input from WV. MB, WV, and AM conducted the field sampling. WV obtained the funding and logistic support. CL provided infrastructure and advised on molecular approaches. MB analyzed the data. Writing of the manuscript was led by MB with contributions from WV and input from all authors.

FUNDING

This research was supported by the Sentinel North program of Université Laval, funded by the Canada First Research Excellence Fund (CFREF). Additional funding and support were provided by the Natural Sciences and Engineering Research Council of Canada (NSERC), the Canada Research Chair program, the Canada Network of Excellence ArcticNet, and the Centre for Northern Studies (CEN).

ACKNOWLEDGMENTS

We thank the communities of Kuujjuarapik and Whapmagoostui, Sydney Arruda for the help at the field station, Marc-Antoine Bansept for his assistance in the field, Marianne Potvin for her invaluable guidance, Marie-Josée Martineau for technical assistance with the HPLC analysis, Aurélie Rivard for HPLC chromatogram analysis, and our helicopter pilot Yancy Yergeau. We also thank Lise Rancourt and the Institut national de la recherche scientifique, Centre Eau-Terre-Environnement (INRS-ETE) for chemical analyses and the Plateforme d'Analyses Génomiques (IBIS, Laval University, Québec) for the amplicon sequencing. This is a contribution to the project Terrestrial Multidisciplinary distributed Observatories for the Study of Arctic Connections (T-MOSAiC), under the auspices of the International Arctic Science Committee (IASC).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.760282/full#supplementary-material

REFERENCES

Abell, G. C. J., and Bowman, J. P. (2005). Colonization and community dynamics of class Flavobacteria on diatom detritus in experimental mesocosms based on Southern Ocean seawater. *FEMS Microbiol. Ecol.* 53, 379–391. doi: 10.1016/j.femsec.2005.01.008

Allard, M., and K.-Seguin, M., (1987). Le pergélisol au Québec nordique: bilan et perspectives. *Géo. Phys. Quatern.* 41, 141–152. doi: 10.7202/032671ar

Allard, M., and Lemay, M. (2012). *Nunavik and Nunatsiavut: From Science to Policy. An Integrated Regional Impact Study (IRIS) of Climate Change and Modernization*. ArcticNet Inc., Quebec City, Canada, 303 p.

Altenburger, A., Blossom, H. E., Garcia-Cuetos, L., Jakobsen, H. H., Carstensen, J., Lundholm, N., et al. (2020). Dimorphism in cryptophytes—The case of *Teleaulax amphioxeia*/*Plagioselmis prolonga* and its ecological implications. *Sci. Adv.* 6:eabb1611. doi: 10.1126/sciadv.abb1611

Amaral, V., Graeber, D., Calliari, D., and Alonso, C. (2016). Strong linkages between DOM optical properties and main clades of aquatic bacteria. *Limnol. Oceanogr.* 61, 906–918. doi: 10.1002/lno.10258

Andrei, A. Ş., Salcher, M. M., Mehrshad, M., Rychtecký, P., Znachor, P., and Ghai, R. (2019). Niche-directed evolution modulates genome architecture in freshwater Planctomycetes. *ISME J.* 13, 1056–1071. doi: 10.1038/s41396-018-0332-5

Andrews, S. (2018). FastQC – a quality control tool for high throughput sequence data. Babraham Bioinformatics, The Babraham Institute, Cambridge. Available at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Apprill, A., McNally, S., Parsons, R., and Weber, L. (2015). Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. *Aquat. Microb. Ecol.* 75, 129–137. doi: 10.3354/ame01753

Arlen-Pouliot, Y., and Bhiry, N. (2005). Palaeoecology of a palsa and a filled thermokarst pond in a permafrost peatland, subarctic Québec, Canada. *The Holocene* 15, 408–419. doi: 10.1191/0959683605hl818rp

Aufdenkampe, A. K., Mayorga, E., Raymond, P. A., Melack, J. M., Doney, S. C., Alin, S. R., et al. (2011). Riverine coupling of biogeochemical cycles between land, oceans, and atmosphere. *Front. Ecol. Environ.* 9, 53–60. doi: 10.1890/100014

Balcarczyk, K. L., Jones, J. B., Jaffé, R., and Maie, N. (2009). Stream dissolved organic matter bioavailability and composition in watersheds underlain with discontinuous permafrost. *Biogeochemistry* 94, 255–270. doi: 10.1007/s10533-009-9324-x

Balzano, S., Sarno, D., and Kooistra, W. H. (2011). Effects of salinity on the growth rate and morphology of ten *Skeletonema* strains. *J. Plankton Res.* 33, 937–945. doi: 10.1093/plankt/fbq150

Bass, D., Tikhonenkov, D. V., Foster, R., Dyal, P., Janouškovec, J., Keeling, P. J., et al. (2018). Rhizarian 'novel clade 10' revealed as abundant and diverse planktonic and terrestrial flagellates, including *Aquavolon* n. gen. *J. Eukaryot. Microbiol.* 65, 828–842. doi: 10.1111/jeu.12524

Bengtsson, M. M., Sjøtun, K., Storesund, J. E., and Øvreås, L. (2011). Utilization of kelp-derived carbon sources by kelp surface-associated bacteria. *Aquat. Microb. Ecol.* 62, 191–199. doi: 10.3354/ame01477

Bhiry, N., Delwaide, A., Allard, M., Bégin, Y., Filion, L., Lavoie, M., et al. (2011). Environmental change in the Great Whale River region, Hudson Bay: five decades of multidisciplinary research by Centre d'études nordiques (CEN). *Ecoscience* 18, 182–203. doi: 10.2980/18-3-3469

Biskaborn, B. K., Smith, S. L., Noetzli, J., Matthes, H., Vieira, G., Streletskiy, D. A., et al. (2019). Permafrost is warming at a global scale. *Nat. Commun.* 10:264. doi: 10.1038/s41467-018-08240-4

Bižić-Ionescu, M., Zeder, M., Ionescu, D., Orlić, S., Fuchs, B. M., Grossart, H. P., et al. (2015). Comparison of bacterial communities on limnic versus coastal marine particles reveals profound differences in colonization. *Environ. Microbiol.* 17, 3500–3514. doi: 10.1111/1462-2920.12466

Blais, M.-A., Matveev, A., and Vincent, W. F. (2021). Limnological and pigment data from the Great Whale River and surrounding surface waters, v. 1 (2018). Nordicana D92.

Blazewicz, S. J., Barnard, R. L., Daly, R. A., and Firestone, M. K. (2013). Evaluating rRNA as an indicator of microbial activity in environmental communities: limitations and uses. *ISME J.* 7, 2061–2068. doi: 10.1038/ismej.2013.102

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170

Bomberg, M., Claesson Liljedahl, L., Lamminmäki, T., and Kontula, A. (2019). Highly diverse aquatic microbial communities separated by permafrost in Greenland show distinct features according to environmental niches. *Front. Microbiol.* 10:1583. doi: 10.3389/fmicb.2019.01583

Bouchard, F., Francus, P., Pienitz, R., and Laurion, I. (2011). Sedimentology and geochemistry of thermokarst ponds in discontinuous permafrost, subarctic Quebec, Canada. *J. Geophys. Res. Biogeosci.* 116:G00M04. doi: 10.1029/2011JG001675

Bouchard, F., Francus, P., Pienitz, R., Laurion, I., and Feyte, S. (2014). Subarctic thermokarst ponds: investigating recent landscape evolution and sediment dynamics in thawed permafrost of northern Québec (Canada). *Arct. Antarct. Alp. Res.* 46, 251–271. doi: 10.1657/1938-4246-46.1.251

Bouvier, T. C., and del Giorgio, P. A. (2002). Compositional changes in free-living bacterial communities along a salinity gradient in two temperate estuaries. *Limnol. Oceanogr.* 47, 453–470. doi: 10.4319/lo.2002.47.2.0453

Bretherton, L., Hillhouse, J., Kamalanathan, M., Finkel, Z. V., Irwin, A. J., and Quigg, A. (2020). Trait-dependent variability of the response of marine phytoplankton to oil and dispersant exposure. *Mar. Pollut. Bull.* 153:110906. doi: 10.1016/j.marpolbul.2020.110906

Buchan, A., LeCleir, G. R., Gulvik, C. A., and González, J. M. (2014). Master recyclers: features and functions of bacteria associated with phytoplankton blooms. *Nat. Rev. Microbiol.* 12, 686–698. doi: 10.1038/nrmicro3326

Burki, F., Sandin, M. M., and Jamy, M. (2021). Diversity and ecology of protists revealed by metabarcoding. *Curr. Biol.* 31, R1267–R1280. doi: 10.1016/j.cub.2021.07.066

Bushnell, B., Rood, J., and Singer, E. (2017). BBMerge–accurate paired shotgun read merging via overlap. *PLoS One* 12:e0185056. doi: 10.1371/journal.pone.0185056

Cabello-Yeves, P. J., Ghai, R., Mehrshad, M., Picazo, A., Camacho, A., and Rodriguez-Valera, F. (2017). Reconstruction of diverse Verrucomicrobial genomes from metagenome datasets of freshwater reservoirs. *Front. Microbiol.* 8:2131. doi: 10.3389/fmicb.2017.02131

Cai, Y. M. (2020). Non-surface attached bacterial aggregates: A ubiquitous third lifestyle. *Front. Microbiol.* 11:557035. doi: 10.3389/fmicb.2020.557035

Callieri, C. (2008). Picophytoplankton in freshwater ecosystems: the importance of small-sized phototrophs. *Fr. Rev.* 1, 1–28. doi: 10.1608/FRJ-1.1.1

Carbonera, D., Giacometti, G., Segre, U., Hofmann, E., and Hiller, R. G. (1999). Structure-based calculations of the optical spectra of the light-harvesting peridinin− chlorophyll− protein complexes from *Amphidinium carterae* and *Heterocapsa pygmaea*. *J. Phys. Chem. B* 103, 6349–6356. doi: 10.1021/jp9846429

Carlson, R. E. (1977). A trophic state index for lakes. *Limnol. Oceanogr.* 22, 361–369. doi: 10.4319/lo.1977.22.2.0361

Castelle, C. J., Méheust, R., Jaffe, A. L., Seitz, K., Gong, X., Baker, B. J., et al. (2021). Protein family content uncovers lineage relationships and bacterial pathway maintenance mechanisms in DPANN archaea. *Front. Microbiol.* 12:1233. doi: 10.3389/fmicb.2021.660052

Castelle, C. J., Wrighton, K. C., Thomas, B. C., Hug, L. A., Brown, C. T., Wilkins, M. J., et al. (2015). Genomic expansion of domain archaea highlights roles for organisms from new phyla in anaerobic carbon cycling. *Curr. Biol.* 25, 690–701. doi: 10.1016/j.cub.2015.01.014

Comeau, A. M., Li, W. K., Tremblay, J. É., Carmack, E. C., and Lovejoy, C. (2011). Arctic Ocean microbial community structure before and after the 2007 record sea ice minimum. *PLoS One* 6:e27492. doi: 10.1371/journal.pone.0027492

Commission de toponymie du Gouvernement du Québec (2021). Grande Rivière de la Baleine. Available at: https://toponymie.gouv.qc.ca/ct/ToposWeb/Fiche.aspx?no_seq=3224 (Accessed February 10, 2021).

Cory, R. M., Ward, C. P., Crump, B. C., and Kling, G. W. (2014). Sunlight controls water column processing of carbon in arctic fresh waters. *Science* 345, 925–928. doi: 10.1126/science.1253119

Cottrell, M. T., and Kirchman, D. L. (2000). Natural assemblages of marine proteobacteria and members of the Cytophaga-Flavobacter cluster consuming low-and high-molecular-weight dissolved organic matter. *Appl. Environ. Microbiol.* 66, 1692–1697. doi: 10.1128/AEM.66.4.1692-1697.2000

Cruaud, P., Vigneron, A., Fradette, M. S., Dorea, C. C., Culley, A. I., Rodriguez, M. J., et al. (2019). Annual protist community dynamics in a freshwater ecosystem undergoing contrasted climatic conditions: The Saint-Charles River (Canada). *Front. Microbiol.* 10:2359. doi: 10.3389/fmicb.2019.02359

Crump, B. C., Amaral-Zettler, L. A., and Kling, G. W. (2012). Microbial diversity in arctic freshwaters is structured by inoculation of microbes from soils. *ISME J.* 6, 1629–1639. doi: 10.1038/ismej.2012.9

Crump, B. C., Armbrust, E. V., and Baross, J. A. (1999). Phylogenetic analysis of particle-attached and free-living bacterial communities in the Columbia River, its estuary, and the adjacent coastal ocean. *Appl. Environ. Microbiol.* 65, 3192–3204. doi: 10.1128/AEM.65.7.3192-3204.1999

Crump, B. C., Baross, J. A., and Simenstad, C. A. (1998). Dominance of particle-attached bacteria in the Columbia River estuary, USA. *Aquat. Microb. Ecol.* 14, 7–18. doi: 10.3354/ame014007

Crump, B. C., Hopkinson, C. S., Sogin, M. L., and Hobbie, J. E. (2004). Microbial biogeography along an estuarine salinity gradient: combined influences of bacterial growth and residence time. *Appl. Environ. Microbiol.* 70, 1494–1505. doi: 10.1128/AEM.70.3.1494-1505.2004

Crump, B. C., Peterson, B. J., Raymond, P. A., Amon, R. M., Rinehart, A., McClelland, J. W., et al. (2009). Circumpolar synchrony in big river bacterioplankton. *Proc. Natl. Acad. Sci. U. S. A.* 106, 21208–21212. doi: 10.1073/pnas.0906149106

Dagg, M., Benner, R., Lohrenz, S., and Lawrence, D. (2004). Transformation of dissolved and particulate materials on continental shelves influenced by large rivers: plume processes. *Cont. Shelf Res.* 24, 833–858. doi: 10.1016/j.csr.2004.02.003

Daggett, C. T., Saros, J. E., Lafrancois, B. M., Simon, K. S., and Amirbahman, A. (2015). Effects of increased concentrations of inorganic nitrogen and dissolved organic matter on phytoplankton in boreal lakes with differing nutrient limitation patterns. *Aquat. Sci.* 77, 511–521. doi: 10.1007/s00027-015-0396-5

DeLong, E. F., Franks, D. G., and Alldredge, A. L. (1993). Phylogenetic diversity of aggregate-attached vs. free-living marine bacterial assemblages. *Limnol. Oceanogr.* 38, 924–934. doi: 10.4319/lo.1993.38.5.0924

Déry, S. J., Stieglitz, M., McKenna, E. C., and Wood, E. F. (2005). Characteristics and trends of river discharge into Hudson, James, and Ungava Bays, 1964–2000. *J. Clim.* 18, 2540–2557. doi: 10.1175/JCLI3440.1

Deshpande, B. N., Crevecoeur, S., Matveev, A., and Vincent, W. F. (2016). Bacterial production in subarctic peatland lakes enriched by thawing permafrost. *Biogeosciences* 13, 4411–4427. doi: 10.5194/bg-13-4411-2016

Dillon, P. J., and Molot, L. A. (1997). Effect of landscape form on export of dissolved organic carbon, iron, and phosphorus from forested stream catchments. *Water Resour. Res.* 33, 2591–2600. doi: 10.1029/97WR01921

Dogs, M., Wemheuer, B., Wolter, L., Bergen, N., Daniel, R., Simon, M., et al. (2017). Rhodobacteraceae on the marine brown alga *Fucus spiralis* are abundant and show physiological adaptation to an epiphytic lifestyle. *Syst. Appl. Microbiol.* 40, 370–382. doi: 10.1016/j.syapm.2017.05.006

Doherty, M., Yager, P. L., Moran, M. A., Coles, V. J., Fortunato, C. S., Krusche, A. V., et al. (2017). Bacterial biogeography across the Amazon River-ocean continuum. *Front. Microbiol.* 8:882. doi: 10.3389/fmicb.2017.00882

Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26, 2460–2461. doi: 10.1093/bioinformatics/btq461

Eyre, B., and Balls, P. (1999). A comparative study of nutrient behavior along the salinity gradient of tropical and temperate estuaries. *Estuaries* 22, 313–326. doi: 10.2307/1352987

Fernández-Gomez, B., Richter, M., Schüler, M., Pinhassi, J., Acinas, S. G., González, J. M., et al. (2013). Ecology of marine Bacteroidetes: a comparative genomics approach. *ISME J.* 7, 1026–1037. doi: 10.1038/ismej.2012.169

Flombaum, P., and Martiny, A. C. (2021). Diverse but uncertain responses of picophytoplankton lineages to future climate change. *Limnol. Oceanogr*. 22, 4171–4181. doi: 10.1002/lno.11951

Fortunato, C. S., and Crump, B. C. (2015). Microbial gene abundance and expression patterns across a river to ocean salinity gradient. *PLoS One* 10:e0140578. doi: 10.1371/journal.pone.0140578

Fortunato, C. S., Eiler, A., Herfort, L., Needoba, J. A., Peterson, T. D., and Crump, B. C. (2013). Determining indicator taxa across spatial and seasonal gradients in the Columbia River coastal margin. *ISME J.* 7, 1899–1911. doi: 10.1038/ismej.2013.79

Fournier, I. B., Lovejoy, C., and Vincent, W. F. (2021). Changes in the community structure of under-ice and open-water microbiomes in urban lakes exposed to road salts. *Front. Microbiol.* 12:660719. doi: 10.3389/fmicb.2021.660719

Frey, K. E., and McClelland, J. W. (2009). Impacts of permafrost degradation on arctic river biogeochemistry. *Hydrol. Process.* 23, 169–182. doi: 10.1002/hyp.7196

Galand, P. E., Lovejoy, C., and Vincent, W. F. (2006). Remarkably diverse and contrasting archaeal communities in a large arctic river and the coastal Arctic Ocean. *Aquat. Microb. Ecol.* 44, 115–126. doi: 10.3354/ame044115

Garneau, M. È., Vincent, W. F., Alonso-Sáez, L., Gratton, Y., and Lovejoy, C. (2006). Prokaryotic community structure and heterotrophic production in a river-influenced coastal arctic ecosystem. *Aquat. Microb. Ecol.* 42, 27–40. doi: 10.3354/ame042027

Gebhardt, A. C., Gaye-Haake, B., Unger, D., Lahajnar, N., and Ittekkot, V. (2004). Recent particulate organic carbon and total suspended matter fluxes from the Ob and Yenisei Rivers into the Kara Sea (Siberia). *Mar. Geol.* 207, 225–245. doi: 10.1016/j.margeo.2004.03.010

Gérikas Ribeiro, C., dos Santos, A. L., Gourvil, P., Le Gall, F., Marie, D., Tragin, M., et al. (2020). Culturable diversity of Arctic phytoplankton during pack ice melting. *Elem. Sci. Anthropocene* 8:6. doi: 10.1525/elementa.401

Ghai, R., Hernandez, C. M., Picazo, A., Mizuno, C. M., Ininbergs, K., Díez, B., et al. (2012). Metagenomes of Mediterranean coastal lagoons. *Sci. Rep.* 2:490. doi: 10.1038/srep00490

Ghiglione, J. F., Mevel, G., Pujo-Pay, M., Mousseau, L., Lebaron, P., and Goutx, M. (2007). Diel and seasonal variations in abundance, activity, and community structure of particle-attached and free-living bacteria in NW Mediterranean Sea. *Microb. Ecol.* 54, 217–231. doi: 10.1007/s00248-006-9189-7

Giovannoni, S. J., Hayakawa, D. H., Tripp, H. J., Stingl, U., Givan, S. A., Cho, J. C., et al. (2008). The small genome of an abundant coastal ocean methylotroph. *Environ. Microbiol.* 10, 1771–1782. doi: 10.1111/j.1462-2920.2008.01598.x

Gonzalez Moguel, R., Bass, A. M., Garnett, M. H., Pilote, M., Keenan, B., Matveev, A., et al. (2021). Radiocarbon data reveal contrasting sources for carbon fractions in thermokarst lakes and rivers of eastern Canada (Nunavik, Quebec). *J. Geophys. Res. Biogeosci.* 126:e2020JG005938. doi: 10.1029/2020JG005938

Grossart, H. P. (2010). Ecological consequences of bacterioplankton lifestyles: changes in concepts are needed. *Environ. Microbiol. Rep.* 2, 706–714. doi: 10.1111/j.1758-2229.2010.00179.x

Grossart, H. P., Massana, R., McMahon, K. D., and Walsh, D. A. (2020). Linking metagenomics to aquatic microbial ecology and biogeochemical cycles. *Limnol. Oceanogr.* 65, S2–S20. doi: 10.1002/lno.11382

Guidi, L., Chaffron, S., Bittner, L., Eveillard, D., Larhlimi, A., Roux, S., et al. (2016). Plankton networks driving carbon export in the oligotrophic ocean. *Nature* 532, 465–470. doi: 10.1038/nature16942

Guillou, L., Bachar, D., Audic, S., Bass, D., Berney, C., Bittner, L., et al. (2013). The Protist ribosomal reference database (PR2): a catalog of unicellular eukaryote small sub-unit rRNA sequences with curated taxonomy. *Nucleic Acids Res.* 41, D597–D604. doi: 10.1093/nar/gks1160

Gustafson, D. E., Stoecker, D. K., Johnson, M. D., Van Heukelem, W. F., and Sneider, K. (2000). Cryptophyte algae are robbed of their organelles by the marine ciliate *Mesodinium rubrum*. *Nature* 405, 1049–1052. doi: 10.1038/35016570

Hahn, M. W., and Hoetzinger, M. (2019). "*Polynucleobacter*" in *Bergey's Manual of Systematics of Archaea and Bacteria*. eds. W. B. Whitman, F. Rainey, P. Kämpfer, M. Trujillo, J. Chun and P. DeVoset al. (New York, NY: John Wiley & Sons, Inc.; in Association with Bergey's Manual Trust), 1–24.

Hansen, P. J., Moldrup, M., Tarangkoon, W., Garcia-Cuetos, L., and Moestrup, Ø. (2012). Direct evidence for symbiont sequestration in the marine red tide ciliate *Mesodinium rubrum*. *Aquat. Microb. Ecol.* 66, 63–75. doi: 10.3354/ame01559

Hanson, C. A., Fuhrman, J. A., Horner-Devine, M. C., and Martiny, J. B. (2012). Beyond biogeographic patterns: processes shaping the microbial landscape. *Nat. Rev. Microbiol.* 10, 497–506. doi: 10.1038/nrmicro2795

Helms, J. R., Stubbins, A., Ritchie, J. D., Minor, E. C., Kieber, D. J., and Mopper, K. (2008). Absorption spectral slopes and slope ratios as indicators of molecular weight, source, and photobleaching of chromophoric dissolved organic matter. *Limnol. Oceanogr.* 53, 955–969. doi: 10.4319/lo.2008.53.3.0955

Henson, M. W., Hanssen, J., Spooner, G., Fleming, P., Pukonen, M., Stahr, F., et al. (2018). Nutrient dynamics and stream order influence microbial community patterns along a 2914 kilometer transect of the Mississippi River. *Limnol. Oceanogr.* 63, 1837–1855. doi: 10.1002/lno.10811

Herlemann, D. P., Labrenz, M., Jürgens, K., Bertilsson, S., Waniek, J. J., and Andersson, A. F. (2011). Transitions in bacterial communities along the 2000 km salinity gradient of the Baltic Sea. *ISME J.* 5, 1571–1579. doi: 10.1038/ismej.2011.41

Hollibaugh, J. T., Wong, P. S., and Murrell, M. C. (2000). Similarity of particle-associated and free-living bacterial communities in northern San Francisco Bay, California. *Aquat. Microb. Ecol.* 21, 103–114. doi: 10.3354/ame021103

Horton, D. J., Theis, K. R., Uzarski, D. G., and Learman, D. R. (2019). Microbial community structure and microbial networks correspond to nutrient gradients

within coastal wetlands of the Laurentian Great Lakes. *FEMS Microbiol. Ecol.* 95:fiz033. doi: 10.1093/femsec/fiz033

Huber, P., Diovisalvi, N., Ferraro, M., Metz, S., Lagomarsino, L., Llames, M. E., et al. (2017). Phenotypic plasticity in freshwater picocyanobacteria. *Environ. Microbiol.* 19, 1120–1133. doi: 10.1111/1462-2920.13638

Hudon, C., Morin, R., Bunch, J., and Harland, R. (1996). Carbon and nutrient output from the Great Whale River (Hudson Bay) and a comparison with other rivers around Quebec. *Can. J. Fish. Aquat. Sci.* 53, 1513–1525. doi: 10.1139/f96-080

Hutalle-Schmelzer, K. M. L., Zwirnmann, E., Krüger, A., and Grossart, H. P. (2010). Enrichment and cultivation of pelagic bacteria from a humic lake using phenol and humic matter additions. *FEMS Microbiol. Ecol.* 72, 58–73. doi: 10.1111/j.1574-6941.2009.00831.x

Ingram, R. G. (1981). Characteristics of the Great Whale River plume. *J. Geophys. Res.* 86, 2017–2023. doi: 10.1029/JC086iC03p02017

Jacquemot, L., Kalenitchenko, D., Matthes, L. C., Vigneron, A., Mundy, C. J., Tremblay, J.-É., et al. (2021). Protist communities along freshwater–marine transition zones in Hudson Bay (Canada). *Elem. Sci. Anth.* 9:00111. doi: 10.1525/elementa.2021.00111

Jezberová, J., and Komárková, J. (2007). Morphological transformation in a freshwater *Cyanobium* sp. induced by grazers. *Environ. Microbiol.* 9, 1858–1862. doi: 10.1111/j.1462-2920.2007.01311.x

Jia, G., Shevliakova, E., Artaxo, P., De Noblet-Ducoudré, N., Houghton, R., House, J., et al. (2019). "Land–climate interactions," in *Climate Change and Land: An IPCC Special Report on Climate Change, Desertification, Land Degradation, Sustainable Land Management, Food Security, and Greenhouse Gas Fluxes in Terrestrial Ecosystems*. eds. P. R. Shukla, J. Skea, E. C. Buendia, V. Masson-Delmotte, H.-O. Pörtner and D. C. Robertset al. (Geneva: IPCC), 131–247.

Jobard, M., Wawrzyniak, I., Bronner, G., Marie, D., Vellet, A., Sime-Ngando, T., et al. (2020). Freshwater *Perkinsea*: diversity, ecology and genomic information. *J. Plankton Res.* 42, 3–17. doi: 10.1093/plankt/fbz068

Juottonen, H., Fontaine, L., Wurzbacher, C., Drakare, S., Peura, S., and Eiler, A. (2020). Archaea in boreal Swedish lakes are diverse, dominated by Woesearchaeota and follow deterministic community assembly. *Environ. Microbiol.* 22, 3158–3171. doi: 10.1111/1462-2920.15058

Karlsson, J., Serikova, S., Vorobyev, S. N., Rocher-Ros, G., Denfeld, B., and Pokrovsky, O. S. (2021). Carbon emission from Western Siberian inland waters. *Nat. Commun.* 12:825. doi: 10.1038/s41467-021-21054-1

Kasalický, V., Jezbera, J., Hahn, M. W., and Šimek, K. (2013). The diversity of the *Limnohabitans* genus, an important group of freshwater bacterioplankton, by characterization of 35 isolated strains. *PLoS One* 8:e58209. doi: 10.1371/journal.pone.0058209

Kellogg, C. T., McClelland, J. W., Dunton, K. H., and Crump, B. C. (2019). Strong seasonality in arctic estuarine microbial food webs. *Front. Microbiol.* 10:2628. doi: 10.3389/fmicb.2019.02628

Kendrick, M. R., Huryn, A. D., Bowden, W. B., Deegan, L. A., Findlay, R. H., Hershey, A. E., et al. (2018). Linking permafrost thaw to shifting biogeochemistry and food web resources in an arctic river. *Glob. Chang. Biol.* 24, 5738–5750. doi: 10.1111/gcb.14448

Kirchman, D. L. (2002). The ecology of Cytophaga–Flavobacteria in aquatic environments. *FEMS Microbiol. Ecol.* 39, 91–100. doi: 10.1111/j.1574-6941.2002.tb00910.x

Klindworth, A., Mann, A. J., Huang, S., Wichels, A., Quast, C., Waldmann, J., et al. (2014). Diversity and activity of marine bacterioplankton during a diatom bloom in the North Sea assessed by total RNA and pyrotag sequencing. *Mar. Genomics* 18, 185–192. doi: 10.1016/j.margen.2014.08.007

Kling, G. W., Kipphut, G. W., and Miller, M. C. (1991). Arctic lakes and streams as gas conduits to the atmosphere: implications for tundra carbon budgets. *Science* 251, 298–301. doi: 10.1126/science.251.4991.298

Kolmakova, O. V., Gladyshev, M. I., Rozanov, A. S., Peltek, S. E., and Trusova, M. Y. (2014). Spatial biodiversity of bacteria along the largest Arctic river determined by next-generation sequencing. *FEMS Microbiol. Ecol.* 89, 442–450. doi: 10.1111/1574-6941.12355

Kooistra, W. H., Sarno, D., Balzano, S., Gu, H., Andersen, R. A., and Zingone, A. (2008). Global diversity and biogeography of *Skeletonema* species (Bacillariophyta). *Protist* 159, 177–193. doi: 10.1016/j.protis.2007.09.004

Kosek, K., Luczkiewicz, A., Kozioł, K., Jankowska, K., Ruman, M., and Polkowska, Ż. (2019). Environmental characteristics of a tundra river system in Svalbard. Part 1: bacterial abundance, community structure and nutrient levels. *Sci. Total Environ.* 653, 1571–1584. doi: 10.1016/j.scitotenv.2018.11.378

Kuzyk, Z. A., and Candlish, L. M. (2019). *From Science to Policy in the Greater Hudson Bay Marine Region: An Integrated Regional Impact Study (IRIS) of Climate Change and Modernization*. ArcticNet, Québec City, 424 pp.

Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinform.* 9:559. doi: 10.1186/1471-2105-9-559

Lefèvre, E., Roussel, B., Amblard, C., and Sime-Ngando, T. (2008). The molecular diversity of freshwater picoeukaryotes reveals high occurrence of putative parasitoids in the plankton. *PLoS One* 3:e2324. doi: 10.1371/journal.pone.0002324

Li, R., Tun, H. M., Jahan, M., Zhang, Z., Kumar, A., Fernando, W. D., et al. (2017). Comparison of DNA-, PMA-, and RNA-based 16S rRNA Illumina sequencing for detection of live bacteria in water. *Sci. Rep.* 7:5752. doi: 10.1038/s41598-017-02516-3

Liu, X., Li, M., Castelle, C. J., Probst, A. J., Zhou, Z., Pan, J., et al. (2018). Insights into the ecology, evolution, and metabolism of the widespread Woesearchaeotal lineages. *Microbiome* 6:102. doi: 10.1186/s40168-018-0488-2

Liu, Y., Lin, Q., Feng, J., Yang, F., Du, H., Hu, Z., et al. (2020). Differences in metabolic potential between particle-associated and free-living bacteria along Pearl River estuary. *Sci. Total Environ.* 728:138856. doi: 10.1016/j.scitotenv.2020.138856

Liu, X., Wang, Y., and Gu, J. D. (2021). Ecological distribution and potential roles of Woesearchaeota in anaerobic biogeochemical cycling unveiled by genomic analysis. *Comput. Struct. Biotechnol. J.* 19, 794–800. doi: 10.1016/j.csbj.2021.01.013

Loiselle, S. A., Bracchini, L., Dattilo, A. M., Ricci, M., Tognazzi, A., Cózar, A., et al. (2009). The optical characterization of chromophoric dissolved organic matter using wavelength distribution of absorption spectral slopes. *Limnol. Oceanogr.* 54, 590–597. doi: 10.4319/lo.2009.54.2.0590

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 1–21. doi: 10.1186/s13059-014-0550-8

Lovejoy, C., Vincent, W. F., Bonilla, S., Roy, S., Martineau, M. J., Terrado, R., et al. (2007). Distribution, phylogeny and growth of cold-adapted picoprasinophytes in Arctic seas. *J. Phycol.* 43, 78–89. doi: 10.1111/j.1529-8817.2006.00310.x

Lozupone, C. A., and Knight, R. (2007). Global patterns in bacterial diversity. *Proc. Natl. Acad. Sci. U. S. A.* 104, 11436–11440. doi: 10.1073/pnas.0611525104

Lyons, M. M., and Dobbs, F. C. (2012). Differential utilization of carbon substrates by aggregate-associated and water-associated heterotrophic bacterial communities. *Hydrobiologia* 686, 181–193. doi: 10.1007/s10750-012-1010-7

Mansour, I., Heppell, C. M., Ryo, M., and Rillig, M. C. (2018). Application of the microbial community coalescence concept to riverine networks. *Biol. Rev.* 93, 1832–1845. doi: 10.1111/brv.12422

McClelland, J. W., Holmes, R. M., Dunton, K. H., and Macdonald, R. W. (2012). The Arctic Ocean estuary. *Estuar. Coasts* 35, 353–368. doi: 10.1007/s12237-010-9357-3

McMurdie, P. J., and Holmes, S. (2013). phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* 8:e61217. doi: 10.1371/journal.pone.0061217

Mohit, V., Archambault, P., Toupoint, N., and Lovejoy, C. (2014). Phylogenetic differences in attached and free-living bacterial communities in a temperate coastal lagoon during summer, revealed via high-throughput 16S rRNA gene sequencing. *Appl. Environ. Microbiol.* 80, 2071–2083. doi: 10.1128/AEM.02916-13

Monteux, S., Keuper, F., Fontaine, S., Gavazov, K., Hallin, S., Juhanson, J., et al. (2020). Carbon and nitrogen cycling in Yedoma permafrost controlled by microbial functional limitations. *Nat. Geosci.* 13, 794–798. doi: 10.1038/s41561-020-00662-4

Moro, I., La Rocca, N., Dalla Valle, L., Moschin, E., Negrisolo, E., and Andreoli, C. (2002). *Pyramimonas australis* sp. nov. (Prasinophyceae, Chlorophyta) from Antarctica: fine structure and molecular phylogeny. *Eur. J. Phycol.* 37, 103–114. doi: 10.1017/S0967026201003493

Morris, R. M., Longnecker, K., and Giovannoni, S. J. (2006). *Pirellula* and OM43 are among the dominant lineages identified in an Oregon coast diatom bloom. *Environ. Microbiol.* 8, 1361–1370. doi: 10.1111/j.1462-2920.2006.01029.x

Mosley, L. M., and Liss, P. S. (2020). Particle aggregation, pH changes and metal behaviour during estuarine mixing: review and integration. *Mar. Freshw. Res.* 71, 300–310. doi: 10.1071/MF19195

Narrowe, A. B., Angle, J. C., Daly, R. A., Stefanik, K. C., Wrighton, K. C., and Miller, C. S. (2017). High-resolution sequencing reveals unexplored archaeal diversity in freshwater wetland soils. *Environ. Microbiol.* 19, 2192–2209. doi: 10.1111/1462-2920.13703

Nozais, C., Vincent, W. F., Belzile, C., Gosselin, M., Blais, M. A., Canário, J., et al. (2021). The Great Whale River ecosystem: ecology of a subarctic river and its receiving waters in coastal Hudson Bay, Canada. *Écoscience* 28, 1–20. doi: 10.1080/11956860.2021.1926137

Ogle, D. H., Wheeler, P., and Dinno, A. (2020). FSA: Fisheries Stock Analysis. R package version 0.8.30. Available at: https://github.com/droglenc/FSA

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., et al. (2019). vegan: Community Ecology Package. R package version 2.5–6. Available at: https://CRAN.R-project.org/package=vegan

Olli, K., Ptacnik, R., Klais, R., and Tamminen, T. (2019). Phytoplankton species richness along coastal and estuarine salinity continua. *Am. Nat.* 194, E41–E51. doi: 10.1086/703657

Ortega-Retuerta, E., Joux, F., Jeffrey, W. H., and Ghiglione, J. F. (2013). Spatial variability of particle-attached and free-living bacterial diversity in surface waters from the Mackenzie River to the Beaufort Sea (Canadian Arctic). *Biogeosciences* 10, 2747–2759. doi: 10.5194/bg-10-2747-2013

Ortiz-Alvarez, R., and Casamayor, E. O. (2016). High occurrence of Pacearchaeota and Woesearchaeota (Archaea superphylum DPANN) in the surface waters of oligotrophic high-altitude lakes. *Environ. Microbiol. Rep.* 8, 210–217. doi: 10.1111/1758-2229.12370

Owczarek, P., Opała-Owczarek, M., Boudreau, S., Lajeunesse, P., and Stachnik, Ł. (2020). Re-activation of landslide in sub-Arctic areas due to extreme rainfall and discharge events (the mouth of the Great Whale River, Nunavik, Canada). *Sci. Total Environ.* 744:140991. doi: 10.1016/j.scitotenv.2020.140991

Padilla, C. C., Ganesh, S., Gantt, S., Huhman, A., Parris, D. J., Sarode, N., et al. (2015). Standard filtration practices may significantly distort planktonic microbial diversity estimates. *Front. Microbiol.* 6:547. doi: 10.3389/fmicb.2015.00547

Parada, A. E., Needham, D. M., and Fuhrman, J. A. (2016). Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environ. Microbiol.* 18, 1403–1414. doi: 10.1111/1462-2920.13023

Paulino, A. I., Larsen, A., Bratbak, G., Evens, D., Erga, S. R., Bye-Ingebrigtsen, E., et al. (2018). Seasonal and annual variability in the phytoplankton community of the Raunefjord, west coast of Norway from 2001–2006. *Mar. Biol. Res.* 14, 421–435. doi: 10.1080/17451000.2018.1426863

Payette, S., Morneau, C., Sirois, L., and Desponts, M. (1989). Recent fire history of the northern Quebec biomes. *Ecology* 70, 656–673. doi: 10.2307/1940217

Payette, S., and Rochefort, L. (2001). *Écologie des tourbières du Québec-Labrador*. Québec: Presses Université Laval.

Payne, J. T., Millar, J. J., Jackson, C. R., and Ochs, C. A. (2017). Patterns of variation in diversity of the Mississippi River microbiome over 1,300 kilometers. *PLoS One* 12:e0174890. doi: 10.1371/journal.pone.0174890

Pekel, J. F., Cottam, A., Gorelick, N., and Belward, A. S. (2016). High-resolution mapping of global surface water and its long-term changes. *Nature* 540, 418–422. doi: 10.1038/nature20584

Peña, E. A., and Slate, E. H. (2006). Global validation of linear model assumptions. *J. Am. Stat. Assoc.* 101, 341–354. doi: 10.1198/016214505000000637

Pereira, O., Hochart, C., Auguet, J. C., Debroas, D., and Galand, P. E. (2019). Genomic ecology of marine group II, the most common marine planktonic Archaea across the surface ocean. *MicrobiologyOpen* 8:e00852. doi: 10.1002/mbo3.852

Piątek, J., Lenarczyk, J., and Piątek, M. (2020). Assessing morphological congruence in *Dinobryon* species and their stomatocysts, including a newly established *Dinobryon pediforme*–stomatocyst connection. *Sci. Rep.* 10:9779. doi: 10.1038/s41598-020-65997-9

Pruesse, E., Quast, C., Knittel, K., Fuchs, B. M., Ludwig, W., Peplies, J., et al. (2007). SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res.* 35, 7188–7196. doi: 10.1093/nar/gkm864

Przytulska, A., Comte, J., Crevecoeur, S., Lovejoy, C., Laurion, I., and Vincent, W. F. (2016). Phototrophic pigment diversity and picophytoplankton abundance in subarctic permafrost thaw lakes. *Biogeosciences* 13, 13–26. doi: 10.5194/bg-13-13-2016

Pucher, M., Wünsch, U., Weigelhofer, G., Murphy, K., Hein, T., and Graeber, D. (2019). staRdom: versatile software for analyzing spectroscopic data of dissolved organic matter in R. *Water* 11, 2366. doi: 10.3390/w11112366

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* 41, D590–D596. doi: 10.1093/nar/gks1219

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available at: https://www.R-project.org/

Rae, R., and Vincent, W. F. (1998). Effects of temperature and ultraviolet radiation on microbial foodweb structure: potential responses to global change. *Freshw. Biol.* 40, 747–758. doi: 10.1046/j.1365-2427.1998.00361.x

Rappé, M. S., Kemp, P. F., and Giovannoni, S. J. (1997). Phylogenetic diversity of marine coastal picoplankton 16S rRNA genes cloned from the continental shelf off Cape Hatteras, North Carolina. *Limnol. Oceanogr.* 42, 811–826. doi: 10.4319/lo.1997.42.5.0811

Retamal, L., Vincent, W. F., Martineau, C., and Osburn, C. L. (2007). Comparison of the optical properties of dissolved organic matter in two river-influenced coastal regions of the Canadian Arctic. *Estuar. Coast. Shelf Sci.* 72, 261–272. doi: 10.1016/j.ecss.2006.10.022

Rial, P., Laza-Martínez, A., Reguera, B., Raho, N., and Rodríguez, F. (2015). Origin of cryptophyte plastids in *Dinophysis* from Galician waters: results from field and culture experiments. *Aquat. Microb. Ecol.* 76, 163–174. doi: 10.3354/ame01774

Rieck, A., Herlemann, D. P., Jürgens, K., and Grossart, H. P. (2015). Particle-associated differ from free-living bacteria in surface waters of the Baltic Sea. *Front. Microbiol.* 6:1297. doi: 10.3389/fmicb.2015.01297

Rinke, C., Rubino, F., Messer, L. F., Youssef, N., Parks, D. H., Chuvochina, M., et al. (2019). A phylogenomic and ecological analysis of the globally abundant marine group II archaea (Ca. Poseidoniales Ord. Nov.). *ISME J.* 13, 663–675. doi: 10.1038/s41396-018-0282-y

Rognes, T., Flouri, T., Nichols, B., Quince, C., and Mahé, F. (2016). VSEARCH: a versatile open source tool for metagenomics. *PeerJ* 4:e2584. doi: 10.7717/peerj.2584

Roiha, T., Peura, S., Cusson, M., and Rautio, M. (2016). Allochthonous carbon is a major regulator to bacterial growth and community composition in subarctic freshwaters. *Sci. Rep.* 6:34456. doi: 10.1038/srep34456

Roy, S., Llewellyn, C. A., Skarstad, E., and Johnsen, G. (2011). *Phytoplankton Pigments - Characterization Chemotaxonomy and Applications in Oceanography*. Cambridge: Cambridge University Press.

RStudio Team (2019). RStudio: Integrated Development for R. RStudio, PBC, Boston, MA. Available at: http://www.rstudio.com/

Sanders, R. W., and Porter, K. G. (1988). "Phagotrophic phytoflagellates," in *Advances in Microbial Ecology*. ed. K. C. Marshall (Boston, MA: Springer), 167–192.

Savio, D., Sinclair, L., Ijaz, U. Z., Parajka, J., Reischer, G. H., Stadler, P., et al. (2015). Bacterial diversity along a 2600 km river continuum. *Environ. Microbiol.* 17, 4994–5007. doi: 10.1111/1462-2920.12886

Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., et al. (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol.* 75, 7537–7541. doi: 10.1128/AEM.01541-09

Shcherbakova, V., Yoshimura, Y., Ryzhmanova, Y., Taguchi, Y., Segawa, T., Oshurkova, V., et al. (2016). Archaeal communities of Arctic methane-containing permafrost. *FEMS Microbiol. Ecol.* 92, 1–11. doi: 10.1093/femsec/fiw135

Simon, M., Grossart, H. P., Schweitzer, B., and Ploug, H. (2002). Microbial ecology of organic aggregates in aquatic ecosystems. *Aquat. Microb. Ecol.* 28, 175–211. doi: 10.3354/ame028175

Sosa, O. A., Gifford, S. M., Repeta, D. J., and DeLong, E. F. (2015). High molecular weight dissolved organic matter enrichment selects for methylotrophs in dilution to extinction cultures. *ISME J.* 9, 2725–2739. doi: 10.1038/ismej.2015.68

Stingl, U., Desiderio, R. A., Cho, J. C., Vergin, K. L., and Giovannoni, S. J. (2007). The SAR92 clade: an abundant coastal clade of culturable marine bacteria possessing proteorhodopsin. *Appl. Environ. Microbiol.* 73, 2290–2296. doi: 10.1128/AEM.02559-06

Stockner, J. G. (1988). Phototrophic picoplankton: an overview from marine and freshwater ecosystems. *Limnol. Oceanogr.* 33, 765–775. doi: 10.4319/lo.1988.33.4part2.0765

Stoecker, D. K. (1999). Mixotrophy among dinoflagellates. *J. Eukaryot. Microbiol.* 46, 397–401. doi: 10.1111/j.1550-7408.1999.tb04619.x

Takahashi, K., Moestrup, Ø., Jordan, R. W., and Iwataki, M. (2015). Two new freshwater woloszynskioids *Asulcocephalium miricentonis* gen. Et sp. nov. and *Leiocephalium pseudosanguineum* gen. Et sp. nov. (Suessiaceae, Dinophyceae) lacking an apical furrow apparatus. *Protist* 166, 638–658. doi: 10.1016/j.protis.2015.10.003

Teeling, H., Fuchs, B. M., Becher, D., Klockow, C., Gardebrecht, A., Bennke, C. M., et al. (2012). Substrate-controlled succession of marine bacterioplankton populations induced by a phytoplankton bloom. *Science* 336, 608–611. doi: 10.1126/science.1218344

Terhaar, J., Lauerwald, R., Regnier, P., Gruber, N., and Bopp, L. (2021). Around one third of current Arctic Ocean primary production sustained by rivers and coastal erosion. *Nat. Commun.* 12:169. doi: 10.1038/s41467-020-20470-z

Thibault, S., and Payette, S. (2009). Recent permafrost degradation in bogs of the James Bay area, Northern Quebec, Canada. *Perm. Periglac. Proc.* 20, 383–389. doi: 10.1002/ppp.660

Tully, B. J. (2019). Metabolic diversity within the globally abundant marine group II Euryarchaea offers insight into ecological patterns. *Nat. Commun.* 10:271. doi: 10.1038/s41467-018-07840-4

Vallières, C., Retamal, L., Ramlal, P., Osburn, C. L., and Vincent, W. F. (2008). Bacterial production and microbial food web structure in a large arctic river and the coastal Arctic Ocean. *J. Mar. Syst.* 74, 756–773. doi: 10.1016/j.jmarsys.2007.12.002

Vannier, T., Leconte, J., Seeleuthner, Y., Mondy, S., Pelletier, E., Aury, J. M., et al. (2016). Survey of the green picoalga *Bathycoccus* genomes in the global ocean. *Sci. Rep.* 6:37900. doi: 10.1038/srep37900

Vaulot, D., Geisen, S., Mahe, F., and Bass, D. (2021). pr2-primers: an 18S rRNA primer database for protists. *Mol. Ecol. Resour.* 21, 1–12. doi: 10.1111/1755-0998.13465

Vigneron, A., Lovejoy, C., Cruaud, P., Kalenitchenko, D., Culley, A., and Vincent, W. F. (2019). Contrasting winter versus summer microbial communities and metabolic functions in a permafrost thaw lake. *Front. Microbiol.* 10:1656. doi: 10.3389/fmicb.2019.01656

Vincent, W. F., and Laybourn-Parry, J. (2008). *Polar Lakes and Rivers - Limnology of Arctic and Antarctic Aquatic Ecosystems*. New York: Oxford University Press.

Vincent, W. F., Lemay, M., and Allard, M. (2017). Arctic permafrost landscapes in transition: towards an integrated earth system approach. *Arctic Sci.* 3, 39–64. doi: 10.1139/as-2016-0027

Vonk, J. E., Tank, S. E., Bowden, W. B., Laurion, I., Vincent, W. F., Alekseychik, P., et al. (2015). Reviews and syntheses: effects of permafrost thaw on Arctic aquatic ecosystems. *Biogeosciences* 12, 7129–7167. doi: 10.5194/bg-12-7129-2015

Wang, Y., Pan, J., Yang, J., Zhou, Z., Pan, Y., and Li, M. (2020). Patterns and processes of free-living and particle-associated bacterioplankton and archaeaplankton communities in a subtropical river-bay system in South China. *Limnol. Oceanogr.* 65, S161–S179. doi: 10.1002/lno.11314

Wauthy, M., Rautio, M., Christoffersen, K. S., Forsström, L., Laurion, I., Mariash, H. L., et al. (2018). Increasing dominance of terrigenous organic matter in circumpolar freshwaters due to permafrost thaw. *Limnol. Oceanogr. Lett.* 3, 186–198. doi: 10.1002/lol2.10063

Weishaar, J. L., Aiken, G. R., Bergamaschi, B. A., Fram, M. S., Fujii, R., and Mopper, K. (2003). Evaluation of specific ultraviolet absorbance as an indicator of the chemical composition and reactivity of dissolved organic carbon. *Environ. Sci. Technol.* 37, 4702–4708. doi: 10.1021/es030360x

Wologo, E., Shakil, S., Zolkos, S., Textor, S., Ewing, S., Klassen, J., et al. (2021). Stream dissolved organic matter in permafrost regions shows surprising compositional similarities but negative priming and nutrient effects. *Glob. Biogeochem. Cycles* 35:e2020GB006719. doi: 10.1029/2020GB006719

Wrona, F. J., Johansson, M., Culp, J. M., Jenkins, A., Mård, J., Myers-Smith, I. H., et al. (2016). Transitions in Arctic ecosystems: ecological implications of a changing hydrological regime. *J. Geophys. Res. Biogeosci.* 121, 650–674. doi: 10.1002/2015JG003133

Wu, W., Lu, H. P., Sastri, A., Yeh, Y. C., Gong, G. C., Chou, W. C., et al. (2018). Contrasting the relative importance of species sorting and dispersal limitation in shaping marine bacterial versus protist communities. *ISME J.* 12, 485–494. doi: 10.1038/ismej.2017.183

Wu, X., Spencer, S., Gushgari-Doyle, S., Yee, M. O., Voriskova, J., Li, Y., et al. (2020). Culturing of "Unculturable" subsurface microbes: natural organic carbon source fuels the growth of diverse and distinct bacteria from groundwater. *Front. Microbiol.* 11:610001. doi: 10.3389/fmicb.2020.610001

Xia, X., Liu, T., Yang, Z., Michalski, G., Liu, S., Jia, Z., et al. (2017). Enhanced nitrogen loss from rivers through coupled nitrification-denitrification caused by suspended sediment. *Sci. Total Environ.* 579, 47–59. doi: 10.1016/j.scitotenv.2016.10.181

Zhu, J., Hong, Y., Zada, S., Hu, Z., and Wang, H. (2018). Spatial variability and co-acclimation of phytoplankton and bacterioplankton communities in the Pearl River estuary, China. *Front. Microbiol.* 9:2503. doi: 10.3389/fmicb.2018.02503

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

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 31 January 2022
doi: 10.3389/fmicb.2021.783767

Geochemically Defined Space-for-Time Transects Successfully Capture Microbial Dynamics Along Lacustrine Chronosequences in a Polar Desert

*Maria R. Monteiro[1,2], Alexis J. Marshall[1,2], Ian Hawes[2], Charles K. Lee[1,2], Ian R. McDonald[1,2] and Stephen Craig Cary[1,2]**

[1] International Centre for Terrestrial Antarctic Research, University of Waikato, Hamilton, New Zealand, [2] Te Aka Matuatua—School of Science, University of Waikato, Hamilton, New Zealand

OPEN ACCESS

Edited by:
Jérôme Comte, Université du Québec, Canada

Reviewed by:
Catherine Girard, Université du Québec à Chicoutimi, Canada
Mincheol Kim, Korea Polar Research Institute, South Korea

***Correspondence:**
Stephen Craig Cary
caryc@waikato.ac.nz

Specialty section:
This article was submitted to Extreme Microbiology, a section of the journal Frontiers in Microbiology

Received: *26 September 2021*
Accepted: *15 December 2021*
Published: *31 January 2022*

Citation:
Monteiro MR, Marshall AJ, Hawes I, Lee CK, McDonald IR and Cary SC (2022) Geochemically Defined Space-for-Time Transects Successfully Capture Microbial Dynamics Along Lacustrine Chronosequences in a Polar Desert. Front. Microbiol. 12:783767. doi: 10.3389/fmicb.2021.783767

The space-for-time substitution approach provides a valuable empirical assessment to infer temporal effects of disturbance from spatial gradients. Applied to predict the response of different ecosystems under current climate change scenarios, it remains poorly tested in microbial ecology studies, partly due to the trophic complexity of the ecosystems typically studied. The McMurdo Dry Valleys (MDV) of Antarctica represent a trophically simple polar desert projected to experience drastic changes in water availability under current climate change scenarios. We used this ideal model system to develop and validate a microbial space-for-time sampling approach, using the variation of geochemical profiles that follow alterations in water availability and reflect past changes in the system. Our framework measured soil electrical conductivity, pH, and water activity *in situ* to geochemically define 17 space-for-time transects from the shores of four dynamic and two static Dry Valley lakes. We identified microbial taxa that are consistently responsive to changes in wetness in the soils and reliably associated with long-term dry or wet edaphic conditions. Comparisons between transects defined at static (open-basin) and dynamic (closed-basin) lakes highlighted the capacity for geochemically defined space-for-time gradients to identify lasting deterministic impacts of historical changes in water presence on the structure and diversity of extant microbial communities. We highlight the potential for geochemically defined space-for-time transects to resolve legacy impacts of environmental change when used in conjunction with static and dynamic scenarios, and to inform future environmental scenarios through changes in the microbial community structure, composition, and diversity.

Keywords: space-for-time (SFT) substitution, climate change, polar desert environments, microbial communities, wetness gradients, Antarctica

INTRODUCTION

Long-term ecological observations provide valuable information for studying the impacts of climate change. They form excellent resources to detect climate trends and patterns over time, study slow or highly variable ecological processes, validate modeled predictions of change, and support environmental policies (Kratz et al., 2003; Rustad, 2008). However, the maintenance of these continuous observations is generally dependent on long-term financial and logistic security from local institutions and governments. When time and funding are a constraint, or when long-term studies are not feasible, space-for-time substitution models are an attractive alternative to forecast long-term climate impacts on ecosystems (Blois et al., 2013).

Space-for-time substitution approaches, such as ecological chronosequences, rely on the assumption that factors responsible for spatial turnover in species abundance are similar to those responsible for temporal turnover (Pickett, 1989; Wogan and Wang, 2018). First used to study plant succession and soil development (Pickett, 1989; Johnson and Miyanishi, 2008; Walker et al., 2010), it has recently been adapted to predict impacts of climate change on microbial communities (Wilhelm et al., 2013; Yang et al., 2014; Yan et al., 2017; Colby et al., 2020). However, despite its common use, the reliability of this approach has been questioned, particularly when its primary assumption is not met or tested (Blois et al., 2013; Damgaard, 2019). For instance, deterministic and stochastic processes can both affect how different microbial groups assemble over different space and time scales (Caruso et al., 2011). Therefore, space-for-time sampling designs should consider the scale and history of the sampling site and discuss the legacy impacts left by historical processes that have known lasting effects on microbial communities (Chase and Myers, 2011; Dini-Andreote et al., 2015; Zhou and Ning, 2017). Without *a priori* knowledge of the ecological niche characteristics or empirical measurements that constrain the deterministic drivers of species variation, space-for-time sampling approaches may lead to a naive interpretation of the community, missing spatial/temporal context, and impairing any comparisons between replicated spatial transects. Ideally, to validate the use of this approach to assess the temporal effects of climate change on an ecosystem's microbiome, a baseline study is required in an ecosystem that lacks trophic complexity and includes well-characterized deterministic gradients of species distribution.

The McMurdo Dry Valleys (MDV) are the largest ice-free area in Antarctica and represent one of Earth's coldest and driest regions (Cary et al., 2010). These polar deserts exhibit high spatial variability in geochemistry, climate, and landscape characteristics, resulting in a patchy distribution of a simple and unique biota functionally dominated by microbial communities (Cary et al., 2010; Lee et al., 2012; Kwon et al., 2017; Feeser et al., 2018; Bottos et al., 2020). Since the environmental conditions select against the establishment of vascular plants and limit complex trophic interactions, the MDV represents an ideal natural laboratory to validate space-for-time as a tool to study climate-related disturbances on microbial communities. Recent changes in the local climate have triggered hydrologic responses across the MDV (Castendyk et al., 2016; Fountain et al., 2016; Levy et al., 2018). This is especially relevant in closed-basin lakes, which are described as lakes that do not have an outlet channel for water to flow out. Examples of these type of lakes include Lake Vanda, Lake Bonney, and Lake Fryxell in the Taylor and Wright Valleys, where the water level has risen (Castendyk et al., 2016; Levy et al., 2018). Water level rise triggers the expansion of the adjacent wetted margins, which imposes selective pressures on the established microbial communities adapted to long-term dry conditions (Van Horn et al., 2014; Niederberger et al., 2015, 2019; Buelow et al., 2016; Lee et al., 2018; Coyne et al., 2020; Ramoneda Massague et al., 2021). Our work and others have previously used simplistic sampling approaches based on physical distance to better understand how microbial community diversity responds to geochemical changes along an environmental gradient (Yang et al., 2014; Niederberger et al., 2015; Yan et al., 2017; Feeser et al., 2018; Lee et al., 2018). These studies assume that geochemical variables along a gradient change gradually and linearly with distance. However, interacting environmental factors may not continuously change along a distance-based gradient (Kappes et al., 2010), nor are microbial communities randomly distributed along natural gradients, being continuously under the influence of deterministic, stochastic, or a combination of both processes.

In this study, our goal was to develop and validate a space-for-time sampling approach to assess the impacts of climate-related hydrological changes on the terrestrial microbiome in a polar desert, using a wetness gradient. We achieved this by determining whether extant microbial community attributes (e.g., changes in structure and diversity) across replicated geochemically defined space-for-time transects could reconstruct past wetting events within the MDV. To ensure the robustness of the transects, we first methodically characterized the spatial variability of local geochemical parameters (water activity, electrical conductivity, and pH) at the chosen sites. Transects were established across static (open-basin with outflow) and dynamic (closed-basin with no outflow) lakes to identify if the structure, diversity, and composition of the microbial community could be used to assess historic vs. more recent impacts of water availability. Open-basin lakes are expected to be less impacted by increased glacial meltwater and groundwater flow due to the presence of an outflow channel which drains the overflow. Therefore, wetted areas surrounding these lakes are expected to expand at lower rates and be more stable, comparatively to those surrounding closed-basin lakes which reflect more recent and predicted hydrological disturbances in the system (Castendyk et al., 2016). We demonstrate the strength of extant microbial community attributes for reconstructing past impacts of hydrological changes and support the capacity for space-for-time to be developed as a robust tool to

understand future impacts of change under current climate warming scenarios.

MATERIALS AND METHODS

Site Description, Space-for-Time Transects, and Sampling

We geochemically defined seventeen space-for-time soil transects, representing a wetness gradient from the shores of six lakes during the 2016 and 2017 Antarctic field seasons. The lakes span the length of the Wright (Lake Brownworth and Lake Vanda) and Taylor (Spaulding Pond, Lake Fryxell, Lake Hoare, and Lake Bonney) valleys (**Figure 1** and **Table 1**). Lake Brownworth and Spaulding Pond are open-basin lakes, meaning that they have an outlet channel through which water flows out. The presence of the outlet regulates the water level in these lakes (Levy et al., 2018). Transects defined from these lakes are referred to in this study as static (as having more static water levels). Lakes Vanda, Fryxell, Bonney, and Hoare are closed-basin lakes, which do not have an outlet channel for water to flow out of the lake. The water levels in these lakes are dynamic on multiple time scales but, over the last 100 years, are thought to have been intermittently rising (Castendyk et al., 2016; Levy et al., 2018; Ramoneda Massague et al., 2021). Transects defined from these lakes are referred to in this study as dynamic (as having more dynamic water levels). For each lake, except for Lake Hoare, three space-for-time transects were defined along the wetness gradient. At Lake Hoare, only two transects were sampled due to snowfall while sampling.

In these soils, pH, electrical conductivity (EC) and water activity (WA) have been described in several studies as being the key drivers of community assembly in the MDV (Barrett et al., 2009; Van Horn et al., 2014; Niederberger et al., 2015, 2019; Feeser et al., 2018; Bottos et al., 2020; George et al., 2021). These metrics were measured in the field to identify *in situ* geochemical gradients across transects from all water bodies within which three geochemical zones were identified (**Table 1** and **Supplementary Figures 1**, **2**). This process involved performing measurements with an average spacing of 50 cm from the existing water level within a transition zone to observe the natural variation of these geochemical variables along the wetness gradient (**Supplementary Figures 1**, **2**). The first zone represented a water-saturated wet zone characterized by comparatively higher water activity (average of 1) and the lowest electrical conductivity (138 ± 81 μS) (sampling point 1). With increasing distance from the lake shore, water content remained high, but conductivity increased several times (938 ± 1,591 μS)

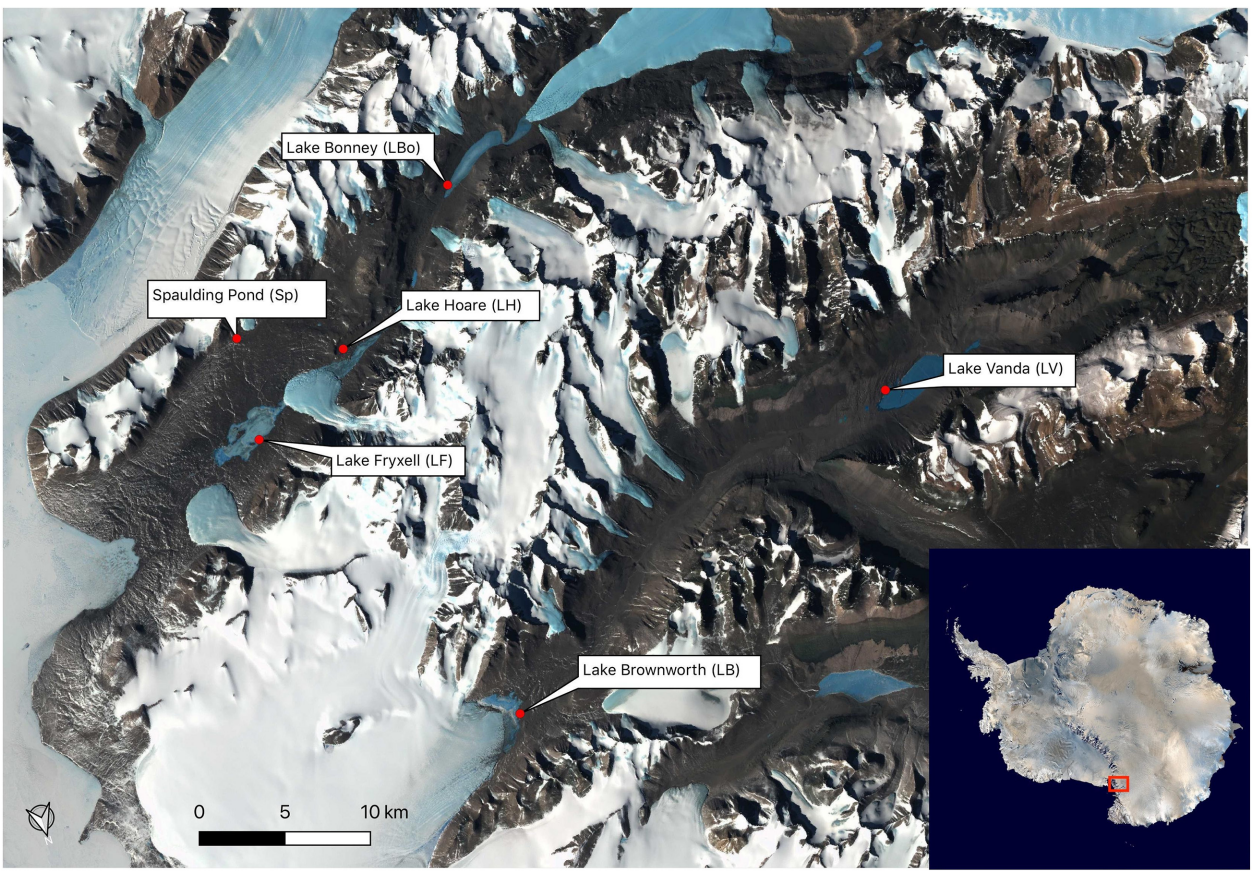

FIGURE 1 | Sampling locations in Wright Valley and Taylor Valley.

TABLE 1 | Soil samples collected across all space-for-time transects and associated geochemical data.

Lake	Lat and long	Valley	Transect section	Gradient	Sampling point	Elevation (cm)	Distance from shoreline (m)	EC (μS)	WA	pH	Moisture (%)
Lake brownworth (LB)	S77.42414 E162.73761	Wright	Wet	Static	1	0	0	25.45 (10.9)	1	7.3 (0.42)	16.77 (3.28)
			Transition		2	14.33 (1.5)	5 (3.6)	71.7 (36.6)	1	7.8 (0.11)	11.30 (1.58)
			Transition		3	36.13 (3.53)	7.35 (3.44)	321 (194.7)	0.77 (0.25)	7.89 (0.10)	6.40 (2.90)
			Dry		4	45.73 (1.41)	8.25 (3.27)	205.7 (87.2)	0.43 (0.11)	7.87 (0.16)	0.74 (0.46)
			Dry		5	92.17 (6.17)	13.7 (3.05)	61.2 (54.5)	0.33 (0.05)	7.87 (0.32)	0.35 (0.6)
Lake vanda (LV)	S77.53408 E161.62372	Wright	Wet	Dynamic	1	0	0	160 (42.23)	1	8.06 (0.11)	18.65 (1.58)
			Transition		2	15.8 (3.34)	2.3 (0.58)	796.8 (154)	0.93 (0.11)	7.9 (0.10)	8.74 (2.9)
			Transition		3	35 (3.37)	3.7 (0.64)	2,340 (2,303)	0.46 (0.38)	7.63 (0.11)	1.06 (0.55)
			Dry		4	45.63 (9.27)	4.83 (1.04)	156.2 (155)	0.26 (0.06)	7.57 (0.15)	0.14 (0.12)
			Dry		5	138.85 (6.15)	11 (1.41)	72.1 (40.44)	0.25 (0.07)	7.45 (0.07)	0.11 (0.14)
Lake bonney (LBo)	S77.69933 E162.53149	Taylor	Wet	Dynamic	1	0	0	83.1 (19.1)	1	8 (0.24)	17.07 (0.28)
			Transition		2	9.83 (0.29)	1	705.7 (302.2)	1	8.44 (0.41)	11.30 (17.63)
			Transition		3	84.67 (2.57)	5.97 (0.05)	5,300 (1,503)	0.31 (0.02)	8.10 (0.15)	2.06 (0.6)
			Dry		4	105 (8.41)	7.5 (0.87)	1269.77 (344.24)	0.40 (0.26)	8.19 (0.07)	1.75 (0.94)
			Dry		5	116.25 (2.48)	8	1217.9 (712.90)	0.28 (0.05)	8.25 (0.13)	0.63 (0.40)
Lake hoare (LH)	S77.63271 E162.93942	Taylor	Wet	Dynamic	1	0	0	160 (61.05)	1	8.47 (0.03)	11.93 (0.51)
			Transition		2	13.25 (1.06)	1	788.25 (761.90)	1	8.76 (0.27)	8.60 (2.58)
			Transition		3	43.75 (1.06)	2.5 (0.7)	475.85 (136.11)	0.46 (0.64)	9.05 (0.16)	2.94 (1.16)
			Dry		4	57.5 (0.70)	3 (0.70)	267.4 (154)	0.6	8.75 (0.30)	1.19 (0.12)
			Dry		5	136 (14.85)	7	133.6 (31.96)	0.8	8.75 (0.20)	0.42 (0.35)
Lake fryxell (LF)	S77.60336 E163.13921	Taylor	Wet	Dynamic	1	0	0	159 (82.53)	1	8.22 (0.12)	16.90 (1.59)
			Transition		2	6 (4.58)	1	633.20 (275.07)	0.98 (0.04)	8.25 (0.30)	12.82 (3.58)
			Transition		3	52.17 (2.57)	7.67 (1.15)	1582.33 (302)	0.98 (0.01)	8.20 (0.02)	6.45 (2.38)
			Dry		4	64.33 (2.52)	8.83 (1.04)	106.26 (34.93)	1 (0.05)	8.47 (0.03)	1.86 (0.53)
			Dry		5	161 (0.12)	19.83 (0.76)	161.12 (99.83)	1 (0.02)	8.69 (0.20)	1.42 (0.37)
Spaulding pond (Sp)	S77.65936 E163.12224	Taylor	Wet	Static	1	0	0	210.63 (107.97)	1	7.69 (0.24)	20.63 (4.44)
			Transition		2	14.83 (2.47)	1.83 (1.04)	765.63 (403.19)	1	8.71 (0.25)	11.87 (2.52)
			Transition		3	27.17 (8.46)	3.67 (1.15)	469.17 (304.77)	1	8.69 (0.19)	4.85 (3.07)
			Dry		4	68.83 (3.25)	7 (2.65)	198.60 (88.51)	0.78 (0.14)	8.53 (0.11)	0.40 (0.16)
			Dry		5	122.33 (20.14)	14.67 (1.15)	355.47 (231.73)	0.88 (0.16)	8.44 (0.27)	0.56 (0.10)

Values represent the average and standard deviations of the measurements across replicated transects (n = 3), except for Lake Hoare (n = 2). Elevation from the shoreline (cm), distance from the shoreline (m), electrical conductivity (EC) (μS), water activity (WA), pH, and soil moisture (%).

(sampling points 2 and 3). This was used to define the second zone, representative of a transition zone. The third zone represented a dry zone, characterized by the lowest measured water activity and lower electrical conductivity (341 ± 443 μS) (sampling points 4 and 5).

Briefly, electrical conductivity and pH were determined from a 1:5 slurry of soil to Milli-Q water (Thermo Scientific Orion meter, United States) (Lee et al., 2012). Water activity was measured using a PawKit meter (AquaLab, NZ), following the manufacturer's instructions. Infield measurements of electrical conductivity and pH were confirmed under laboratory conditions using Thermo Scientific Orion meter and the same slurry technique, but with the addition of 0.5 mL of 0.01M $CaCl_2$ per sample (Minasny et al., 2011; Lee et al., 2012). Soil moisture content was determined according to Lee et al. (2012). The elevation of each sampling point along the transects was surveyed in relation to the lakeshore (sampling point 1) using a laser leveling system (Topcon Laser Systems, United States).

For microbial community analysis, we aseptically collected surface soil samples (top 2 cm) with a sterile spatula from the five geochemically predefined sampling points. Spatulas were washed and then sterilized with ethanol wipes between sampling points and sampling sites. In total, across the six lakes, seventeen space-for-time transects were sampled, resulting in 85 soil samples. Individual samples were thoroughly mixed in a sterile Whirl-Pak® bag and then distributed into sterile 50 mL Falcon tubes. Samples were initially stored in the field on ice for 24 h before being transferred to dry ice and transported to Scott Base (the New Zealand Antarctic Research Station), where they were maintained a −60°C until DNA extraction. The remaining samples in each Whirl-Pak® bag were used to confirm moisture content, electrical conductivity, and pH analysis under laboratory conditions. These samples were kept cold at 4°C until processed.

Microbial Community Analysis

Total DNA was extracted from approximately 1 g of soil using a modified version of the CTAB (cetyl-trimethyl-ammonium-bromide) bead-beating method (Coyne et al., 2001). For each batch of DNA extractions, a negative control was included to ensure the identification of possible contamination. DNA concentration was determined using the Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific, United States). The bacterial and archaeal microbial community was targeted through the V4 region of the 16S rRNA gene was amplified in triplicate using the fusion-primer set 515F/806R (Parada et al., 2016) and sequenced using Ion PGM chemistry. Briefly, 20 μL PCR reactions each contained: 1 ng of total DNA, dNTPs (240 μM), $MgCl_2$ (6 mM), bovine serum albumin (0.24 μM), forward and reverse fusion primers (0.2 μM), 1 U of Platinum Taq polymerase (Invitrogen Inc., Carlsbad, California) and PCR buffer (1.2×). The following PCR conditions were used: 94°C for 3 min, followed by 30 cycles at 94°C for 45 s, 50°C for 1 min, 72°C for 1.5 min, and a final extension step at 72°C for 10 min. For each PCR run, negative amplification was confirmed for each DNA extraction batch control. Triplicate PCR amplicons were pooled, and the expected amplicon size was confirmed via electrophoresis with a 1% agarose TAE gel. PCR products were cleaned and each sample concentration was normalized with SequalPrep^{TM} (Thermo Fisher Scientific, United States). Normalized samples were pooled at an equimolar concentration into a single library for sequencing. Amplicon sequencing was performed using the Ion PGM^{TM} System for Next Generation Sequencing (Thermo Fisher Scientific, United States) at the Waikato DNA Sequencing Facility (University of Waikato, New Zealand).

Raw sequences were filtered with Ion PGM^{TM} software to remove low-quality and polyclonal reads. The remaining sequences were processed using a combination of Mothur (v.1.40.5) and USEARCH 10 (v10.0.24) software (Schloss et al., 2009; Edgar, 2010, 2013). Forward and reverse primers were identified within the sequences and trimmed using the python script fastq_strip_barcode_relabel.py supplied by UPARSE (v10.0.240). Sequences without forward and reverse primers were discarded. The remaining sequences were trimmed based on the length and the number of homopolymers sourced by Mothur script (Schloss et al., 2009). All reads with expected error rates higher than 2.5 were discarded using USEARCH (Edgar, 2010), and all reads were truncated to 350 bp. Through dereplication, unique sequences were identified and their abundances quantified. Sequences were sorted, singletons removed, and the remaining sequences were clustered into representative OTUs using the UPARSE-OTU algorithm combined with the GOLD database to detect and remove chimeras (Bernal et al., 2001; Edgar, 2013). Reads were finally clustered into operational taxonomic units (OTUs) using UCLUST with a similarity threshold of 97%. Sequences that did not map to any OTU were discarded. Taxonomy was inferred using SINA (v1.2.11) and the SILVA SSU database (v 138) (Pruesse et al., 2012). A 0.005% cutoff was applied to the raw OTU counts across the dataset to remove poorly represented OTUs (Bokulich et al., 2013). Filtered OTUs were subsequently removed from the sequence Fasta file. Sequences were aligned using a Multiple Alignment using Fast Fourier Transform with default settings (MAFFT v7.429-gimkl-2020a) and a phylogenetic tree was generated using FastTree (v2.1.11).

Statistical Analysis

All statistical and visualization analyses were computed in R (v3.5.2, R Core Team, 2000) using the following packages: phyloseq (v1.26.1) (McMurdie and Holmes, 2013), vegan (v2.5) (Oksanen et al., 2012), picante (v1.8) (Kembel et al., 2010), ggplot2 (v3.2.1) (Wickham, 2016), ggpubr (v0.4.0), compositions (van den Boogaart and Tolosana-Delgado, 2008), and randomForest (v4.6) (Liaw and Wiener, 2002). Alpha diversity was calculated on a non-rarefied dataset using richness and phylogenetic diversity (PD) indexes. Differences in library sizes were tested on both rarefied and non-rarefied data with no significant impact on the data interpretation. Diversity differences between transect zones and lakes were tested using one-way ANOVA. Normality assumptions were tested using a Shapiro-Wilk test, and the assumption of homogeneity of variances was tested using a Levene's test. For the relationship between the microbial diversity and the log transformed moisture content and electrical conductivity, correlation

analyses (Pearson) were applied. For beta diversity analysis we transformed the data using a total sum normalization. Beta diversity was calculated using weighted UniFrac phylogenetic pairwise distances (Lozupone and Knight, 2015) and visualized in a principal coordinates analysis (PCoA). A PERMANOVA analysis, using Adonis function in vegan package (Oksanen et al., 2012), tested the dissimilarities among communities from different groups (lakes, and zones within the wetness gradients). The variation between samples from the three wetness zones described along the transects was tested using a permutational multivariate analyses of dispersion (betadisp) from vegan package (Anderson et al., 2006). A Random Forest model (Liaw and Wiener, 2002) was used to detect the most reliable and relevant top 20 OTUs to predict the different zones along the wetness gradients. Raw counts were transformed using a centered log-ratio (clr) before classification and regression models. OTUs were selected based on the MeanDecreaseAccuracy values, which measures the extent to which a variable (OTU) improves the accuracy of the forest in predicting the classification. Higher values indicate that the OTU improves the prediction. To evaluate the phylogenetic community assembly among species within each sample we calculated the mean nearest taxon distance (MNTD) and the nearest taxon index (NTI) using "taxa.labels" null model, with 999 iterations, using the function "ses.mntd" from the R package picante. The NTI was quantified as the number of standard deviations that the observed MNTD was from the mean of the MNTD null distribution, multiplying by –1. For a single community, observed NTI values $> +2$ or < -2 indicate phylogenetic clustering of species or phylogenetic overdispersion, respectively. NTI values between -2 and 2 usually indicates the influence of a stochastic assembly (Stegen et al., 2012). To compare community assembly processes along the moisture gradients, we calculated the Beta Nearest Taxon Index (βNTI). Following Stegen et al. (2013), the βNTI is the number of standard deviations that the observed beta mean nearest taxon distance (βMNTD) is from the mean of the null distribution. It indicates how much the observed difference between a pair of communities differs from a null distribution. Similarly to NTI, values $> +2$ or < -2 indicate phylogenetic clustering of species or phylogenetic overdispersion, respectively, while values between -2 and 2 usually indicates the influence of a stochastic assembly (Stegen et al., 2013). β-NTI pairwise comparisons were plotted against increasing differences in moisture content. A Euclidean distance matrix was calculated using pH, electrical conductivity, water activity, moisture content, and elevation log-transformed and normalized data. ANOSIM was performed on the resemblance matrix to test the significance of the dissimilarities between the predefined zones of the wetness gradients. Sampling locations were plotted using Quantarctica (v3.1) (Matsuoka et al., 2021).

RESULTS

Characterization of Space-for-Time Wetness Transects

Seventeen geochemically defined space-for-time transects were parameterized using measurements of elevation, soil moisture content, electrical conductivity, pH, and water activity from both dynamic (n = 11 transects; Lake Vanda, Lake Bonney, Lake Hoare, and Lake Fryxell) and static (n = 6 transects; Lake Brownworth and Spaulding Pond) lakes across the Wright and Taylor Valleys (**Figure 1**, **Table 1**, and **Supplementary Figure 2**). Across all transects, elevation increased from wet (sampling point 1) to dry (sampling points 4 and 5) soils (from 0 to 135 ± 19 cm), and soil moisture content decreased in the same direction (from 17.25 ± 3.07 to 0.54 ± 0.54%) (**Figure 2** and **Table 1**). Electrical conductivity was lowest in wet and dry soil samples (138 ± 81 μS and 341 ± 443 μS, respectively), achieving the highest values and variation within the transition zone (938 ± 1,591 μS) (**Figure 2** and **Table 1**). Water activity decreased from wet to dry soils in Lake Vanda, Lake Bonney, Lake Brownworth, and Lake Hoare and remained stable across the wetness gradient in Spaulding Pond and Lake Fryxell (**Figure 2** and **Table 1**). The variability in soil pH across the wetness gradients was lake-specific (**Figure 2** and **Table 1**). The significance of the distinction between the three wetness zones (wet, transition, and dry) was confirmed with non-parametric testing (ANOSIM, R = 0.69, p-value < 0.01, 999 permutations) (**Supplementary Figure 3**).

Sequencing Results and Quality Control

After filtering out low-quality and short sequence reads, we obtained a total of 2,084,605 sequence reads and 4,098 OTUs at 97% sequence similarity (**Supplementary Table 1**). We then filtered the low abundant OTUs using a 0.005% relative abundance cutoff across the entire OTU dataset and removed 72% of the initial OTUs. The remaining 28% (1,133 OTUs) comprised 97% of the initial reads. A significant correlation exists between the ordinations of the initial and filtered datasets (m12 squared: < 0.001; Procrustes Correlation: 0.99; p-value < 0.01, 999 permutations). We removed thirty-two OTUs classified as Chloroplast using the SILVA database, leaving a dataset containing 1,101 OTUs for downstream analysis. Sixteen percent of the remaining OTUs were unclassified at the phylum level. The Procrustes Correlation between the ordination matrices with and without the unclassified phyla was significant (m12 squared: < 0.001; Procrustes Correlation: 0.99; p-value < 0.01, 999 permutations). As a result, we retained the unclassified OTUs in the analysis.

Microbial Community Diversity Along Space-for-Time Wetness Transects

We observed significant differences in species phylogenetic diversity (PD) within the dry, transition, and wet transect zones from each lake (**Supplementary Figure 4**; p-value < 0.01). Spaulding Pond consistently presented the most diverse community along the wetness gradients and Lake Vanda the least diverse community (**Supplementary Figure 4**). Except for Lake Vanda and Lake Bonney, PD correlated positively with soil moisture content (p-value < 0.05; **Figure 3A**). Soil electrical conductivity correlated negatively with PD (p-value < 0.05; **Figure 3B**) only in Lake Bonney transects, likely driven by the comparatively high electrical conductivity (5,300 μS) measurements.

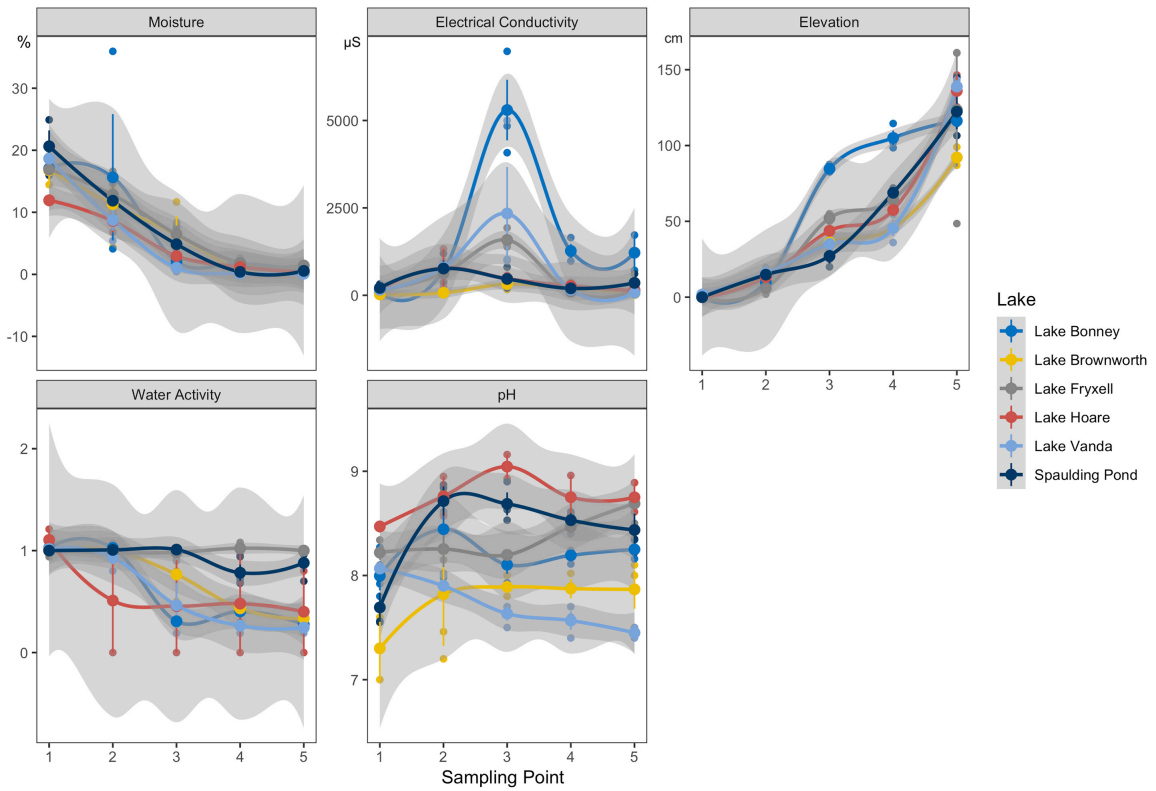

FIGURE 2 | Soil moisture content (%), electrical conductivity (µS), elevation (cm), water activity (WA), and pH profiles along the space-for-time transects defined from all major lakes (n = 3, except for Lake Hoare n = 2) in Taylor and Wright valleys.

Microbial Community Composition Along Space-for-Time Wetness Transects

The soil microbial community was represented by 24 phyla, with 98% of reads classified within the top 10 phyla (**Supplementary Figure 5**). These dominant 10 phyla maintained their representation across all lakes sampled, with a variation in relative abundance reflective of the moisture zone. *Bacteroidota* (0.31 ± 0.08), *Proteobacteria* (0.24 ± 0.08), and *Cyanobacteria* (0.19 ± 0.11) were consistently abundant in wet soils. *Bacteroidota* (0.28 ± 0.09), *Proteobacteria* (0.18 ± 0.10), and *Actinobacteriota* (0.15 ± 0.13) were highly abundant in the transition zones. Dry zones were dominated by members of *Actinobacteriota* (0.30 ± 0.07), *Bacteroidota* (0.20 ± 0.05), *Proteobacteria* (0.11 ± 0.03) and *Acidobacteriota* (0.10 ± 0.06) phyla. In Lake Bonney, taxa affiliated with the phyla *Cyanobacteria* and *Firmicutes* presented a relatively high abundance in transition zone (0.16 ± 0.08 and 0.14 ± 0.16, respectively) and dry samples (0.36 ± 0.09 and 0.09 ± 0.09, respectively).

Microbial Community Structure Along Space-for-Time Wetness Transects

Community beta-diversity from communities from the dry zones of space-for-time transects varied significantly between lakes (PERMANOVA, R^2 = 0.69, F = 11.55, p-value < 0.01; **Supplementary Figure 6A**), sharing only 32% of the OTUs (**Supplementary Figure 7A**). Microbial communities from the wet zones of space-for-time transects were significantly different between lakes (PERMANOVA, R^2 = 0.73 F = 4.47, p-value < 0.01; **Supplementary Figure 6B**), sharing only 16% of the OTUs (**Supplementary Figure 7B**).

Along the space-for-time transects, communities collected in the dry zone were structurally distinct from those collected in the wet zone (PERMANOVA, R^2 = 0.27, F = 14.11, p-value < 0.01) (**Figure 4**). For each lake, significant clustering of the communities by zone was identified (**Figure 5**). However, only the Lake Brownworth Spaulding Pond and Lake Hoare transects met the premise for homogenous dispersion (**Figures 5A,B,F**; beta-dispersion > 0.05). For all dynamic lake transects (apart from Lake Hoare), the dispersion of communities within the transition zone was significantly different (beta-dispersion < 0.05) (**Supplementary Table 2**). Except for Lake Hoare, the transition zone community across all dynamic lake transects was segregated into two clusters, with one of the clusters intercepting with the dry soil community cluster on sampling point 3 (ellipses at 95% confidence level) (**Figures 5C–E**). The latter was not observed in the stable lake transects collected from Lake Brownworth (**Figure 5A**) or Spaulding Pond (**Figure 5B**).

Phylogenetic community composition within each sample revealed a trend leaning toward a deterministic assembly of the community governed by environmental selection (NTI > 2)

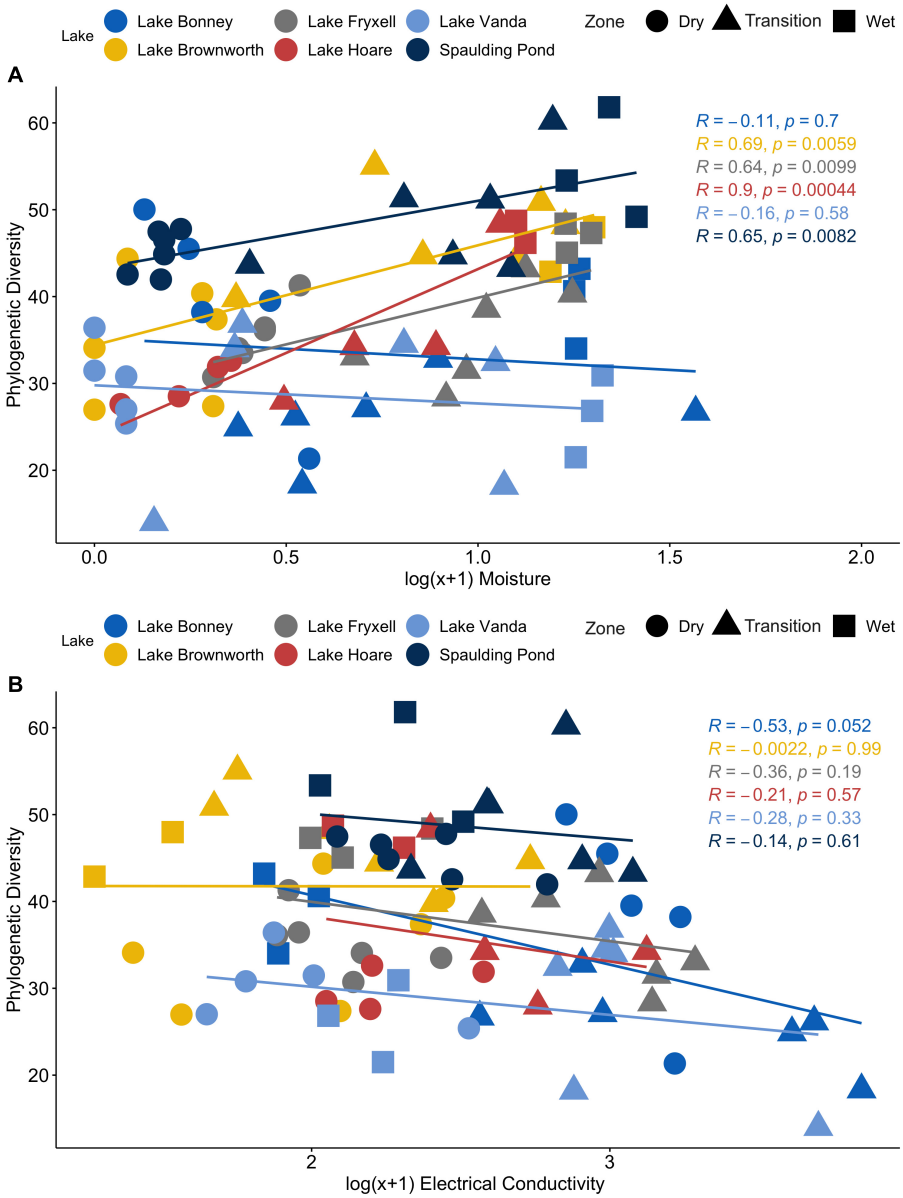

FIGURE 3 | (A) Relationship between Phylogenetic Diversity (PD) and log transformed moisture content measurements across the SFT transects from Lake Brownworth, Lake Bonney, Lake Fryxell, Lake Hoare, Spaulding Pond, and Lake Vanda. **(B)** Relationship between Phylogenetic Diversity (PD) and log transformed electrical conductivity measurements across the SFT transects from Lake Brownworth, Lake Bonney, Lake Fryxell, Lake Hoare, Spaulding Pond, and Lake Vanda.

(**Supplementary Figure 8A**). That effect was higher, and consistent in the dry soils (sampling point 5) and tended to decrease toward the wet zone (sampling point 1) of the gradient, except for Lake Bonney and Lake Vanda, where no trend was identified (**Figure 6**). In contrast to NTI values, βNTI values distribution showed a median close to 0 but with skewed distribution stretching beyond the + 2 significant threshold (**Supplementary Figure 8A**). This reflects the influence of deterministic processes for a given pairwise comparison. However, pairwise comparisons regressed against differences in moisture content indicates a large influence of stochasticity with a weak relation to moisture content (**Supplementary Figure 8B**).

Microbial Taxa as Sentinels for Change

Soil moisture content was the primary non-categorical variable driving microbial community structure patterns (PERMANOVA, R^2 = 0.25, F = 26.48, p-value < 0.01; **Table 2**). A random forest approach identified that 64.8% of the variance in the community could be explained by moisture content (RF, no of trees: 10001, mean of squared residuals 19.88). Taxa affiliated to *Actinobacteriota,*

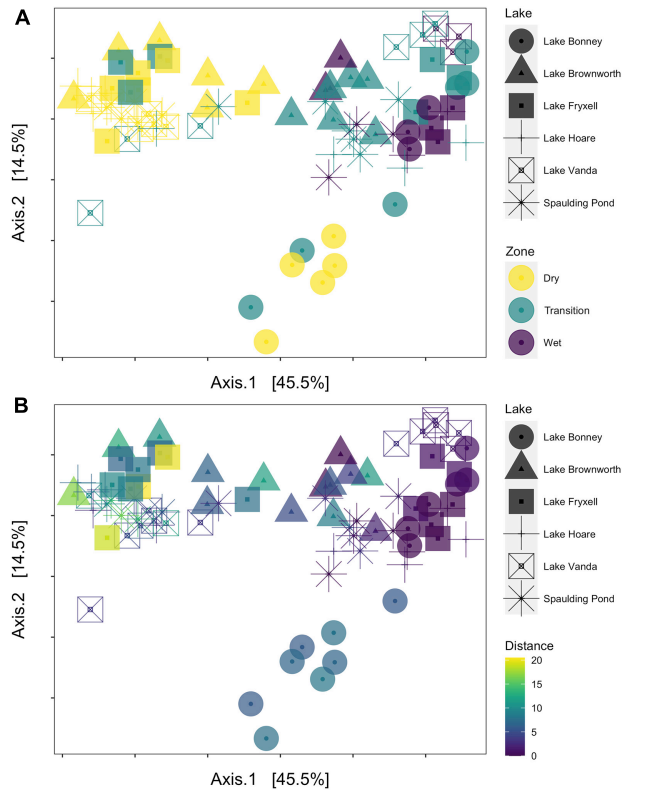

FIGURE 4 | Principal coordinate analysis (PCoA) of microbial community compositional data based on a weighted UniFrac distance matrix. **(A)** The samples are colored by transect zone. **(B)** The samples are colored according to their distance from the lake shore.

Chloroflexi, Deinococcota, Verrucomicrobiota, Proteobacteria, Acidobacteriota, Abditibacteriota, and *Bacteroidota* were the major taxa driving differences between wet and dry soils (out-of-bag estimate of error rate for the transect section classification: 24.4%) (**Figure 7A**). OTUs affiliated to *Actinobacteria,* order *Solirubrobacterales*; *Chloroflexi*, order *Kallotenuales*; *Deinococcota*, order *Deinococcales;* *Acidobacteria*, order Blastocaellales, and *Abditibacteriota,* order *Abditibacteriales,* were exclusively associated to dry and transition zone soils (**Figure 7B** and **Supplementary Table 3**). Our model also identified the OTU assigned to *Verrucomicrobiota*, order *Verrucomicrobiales* (genus *Luteolibacter)* as an important OTU to classify wet soils (**Figure 7B** and **Supplementary Table 3**). Members of *Bacteroidota* and *Proteobacteria* phyla were associated with both dry or wet soil conditions. *Bacteroidota* taxa related to wet soil conditions belonged to the *Flavobacterium* and *Ferruginobacter* genus (order *Flavobacteriales* and *Chitinophagales),* whereas the genus *Segitobacter* (order *Chitinophagales*) was more abundant in dry soils. Within *Proteobacteria*, 4 OTUs affiliated to the genus *Sphingorhabdus* (order *Sphingomonadales*), *Thermomonas and Pseudoxanthomonas* (order *Xanthomonadales*), and genus *Brevundimonas* (order *Caulobacterales*) were associated with wet soil conditions, whereas the genus *Sphingomonas* (order *Sphingomonadales*) were associated with the dry zone of the gradient (**Figure 7B** and **Supplementary Table 3**).

TABLE 2 | PERMANOVA test using soil moisture content, pH, and electrical conductivity on microbial communities along the wetness transects.

	F model	Statistic R2	Significance (*p*-value)
Moisture (%)	26.48	0.25	<0.01
pH	1.36	0.02	0.22
Electrical conductivity (μS)	5.28	0.06	<0.01

DISCUSSION

Here we sought to validate the application of space-for-time as a reliable approach to forecast the impacts of climate-related hydrological changes on the terrestrial microbiome in a polar desert. We chose the soil ecosystem of the MDV as a natural, trophically simple, and well-characterized system where single environmental drivers (such as water availability) have a profound and quantifiable impact on the resident microbial community (Van Horn et al., 2014; Niederberger et al., 2015, 2019; Buelow et al., 2016; Lee et al., 2018; Bottos et al., 2020; Coyne et al., 2020; Ramoneda Massague et al., 2021). Our sampling approach starts by highlighting the importance of selection of sites along spatial gradients based on field determinations of critical environmental drivers of species distributions and their patterns across space. This will help to select sites at consistent points along different ecotones, which can be missed when arbitrary measures such as distance are chosen (**Figure 2**). Secondly, we point out the importance of repeating space-for-time transects across multiple units within the ecosystem to fully harness the characteristics of environmental variability. Lastly, we considered the detailed historical background on ecosystem response to climate change and defined transects in relatively stable and dynamic environmental scenarios. The MDV offers access to multiple lake systems with variable rates and timing of change. Closed-basin lakes, in particular, are quite sensitive to climate change, which is reflected by historic and current changes in water level and salinity (Castendyk et al., 2016; Levy et al., 2018). Across all sampled lakes Lake Vanda and Lake Bonney have experienced the biggest water level rise, followed by Lake Fryxell and Lake Hoare (Levy et al., 2018). Open-basin lakes, in contrast, may change at slower rates due to an established outlet that rapidly drains glacial meltwater and groundwater flow (Levy et al., 2018). Transects defined in open-basin lakes set the background state for longer periods of stability, whereas transects defined in closed-basin lakes mirror more recent and predicted hydrological disturbances in the system.

Historical Legacy of Water Availability Reflected by the Structure and Composition of Microbial Communities

By comparing geochemically defined space-for-time transects from lake systems with relatively stable or dynamic water levels, we were able to resolve the legacy impacts of water presence on microbial community structure (**Figure 4**). Only within transects defined from Lake Brownworth, and to a lesser extent Spaulding

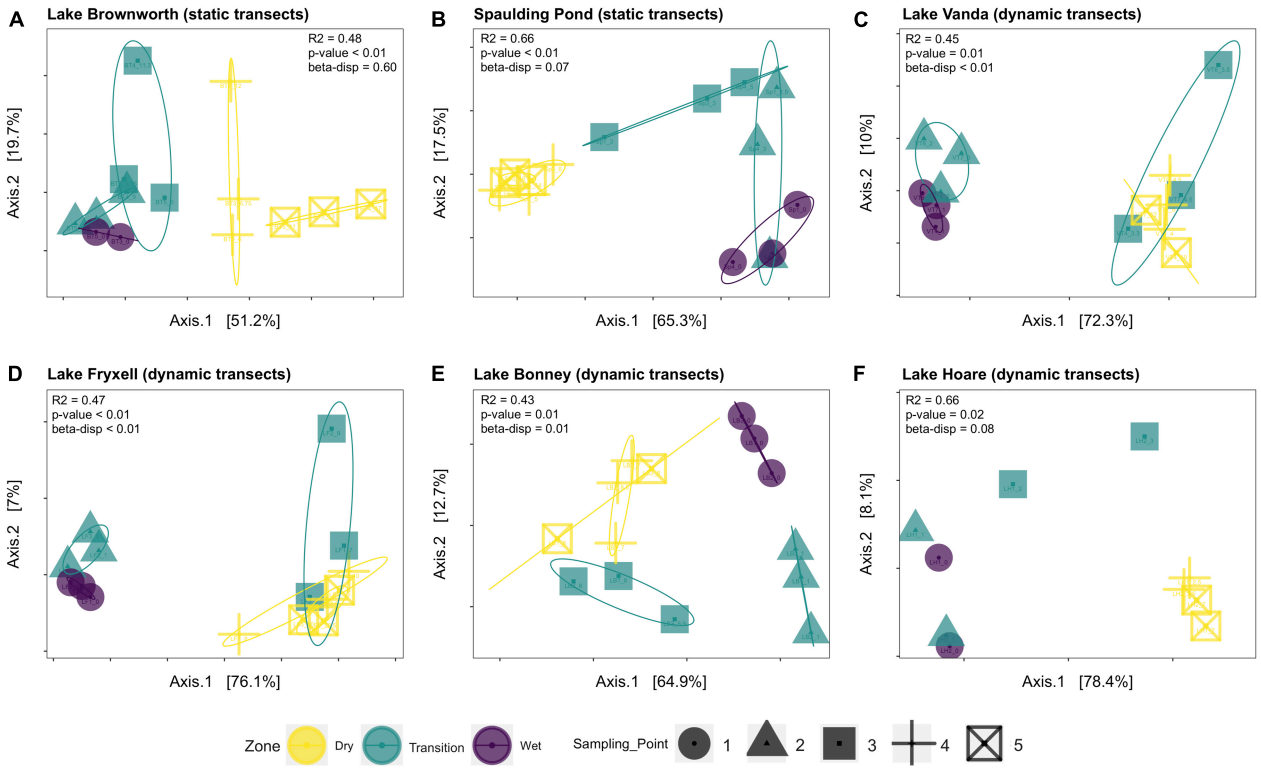

FIGURE 5 | Principal coordinate analysis (PCoA) of microbial community data based on a weighted UniFrac distance matrix for each lake along each transect: **(A)** Lake Borwnworth; **(B)** Spaulding Pond; **(C)** Lake Vanda; **(D)** Lake Fryxell; **(E)** Lake Bonney; **(F)** Lake Hoare. The significance between weighted UniFrac distance was calculated using PERMANOVA. The assumption for homogeneity of group dispersion was tested with betadisper function. Ellipses on the PCoA plot highlight the deviations from the mean for each clustered group of samples representative of different zones of the transects (95% confidence level).

Pond (static transects), did the microbial community display specific structural patterns reflecting the three geochemical zones identified (**Figures 5A,B**). The high degree of zone specificity (wet, transition, and dry) in community structuring from these lakes aligns with an extended period of more stability due to the presence of an outflow (Levy et al., 2018). In contrast, from the dynamic closed-basin lake transects, the clear spatial segregation of the transition zone communities into one of "dry" or "wet"-like community could reflect a response related to legacy of water level changes in these lakes, not detected by the defined geochemical metrics, which respond on a different time scale. Particularly in Lake Vanda, Lake Bonney, and Lake Fryxell (**Figure 5C–E**), which have risen between 1.6 and 4.1 m from 2001 to 2017 (Levy et al., 2018), no distinct transitional zone community was recognizable (**Figure 7A**). Our findings support recent monitoring of Lake Vanda water levels, which indicate that this lake has been continually rising at a rate of 22 cm/year since 1947 (Castendyk et al., 2016). The lack of zone specificity in the microbial community structure at this lake likely reflects that at the time of sampling, microbial communities collected at the geochemically defined transition between transition and dry zones were likely to have been recently wet (potentially hours to days) and remain similar to a dry soil community. In contrast, communities exposed to wetness for more extended periods resemble a more long-term wet soil community. Similar temporal shifts were observed by Ramoneda Massague et al. (2021) in moat samples collected along a depth gradient in Lake Vanda. Microbial assemblages collected from the higher depths diverged from those collected at the lake shore, being likely influenced by pre-existing terrestrial microbes in recently inundated soils (Ramoneda Massague et al., 2021). The consistency of our results across closed-basin lake transects suggests that a geochemically defined space-for-time sampling approach can be applied across polar systems to monitor change through the lasting legacy impacts of water availability on assembly mechanisms that shape extant microbial community structure.

Upon observing different spatial structuring of microbial communities, we evaluated whether microbial community assembly patterns within and between communities along the moisture gradient were driven by deterministic or stochastic processes, or by a combination of both. Contrary to a previous observation (Lee et al., 2018), the NTI values observed for each sample across all 17 geochemically defined space-for-time transects indicate that the degree of phylogenetic relatedness among the individuals in a community is likely driven by environmental selection (NTI > 2) (**Figure 6**). In the MDV, the dry soils are typically oligotrophic, with minimal water content (<2%) and high salt concentrations (Cary et al., 2010), which collectively exclude a significant range

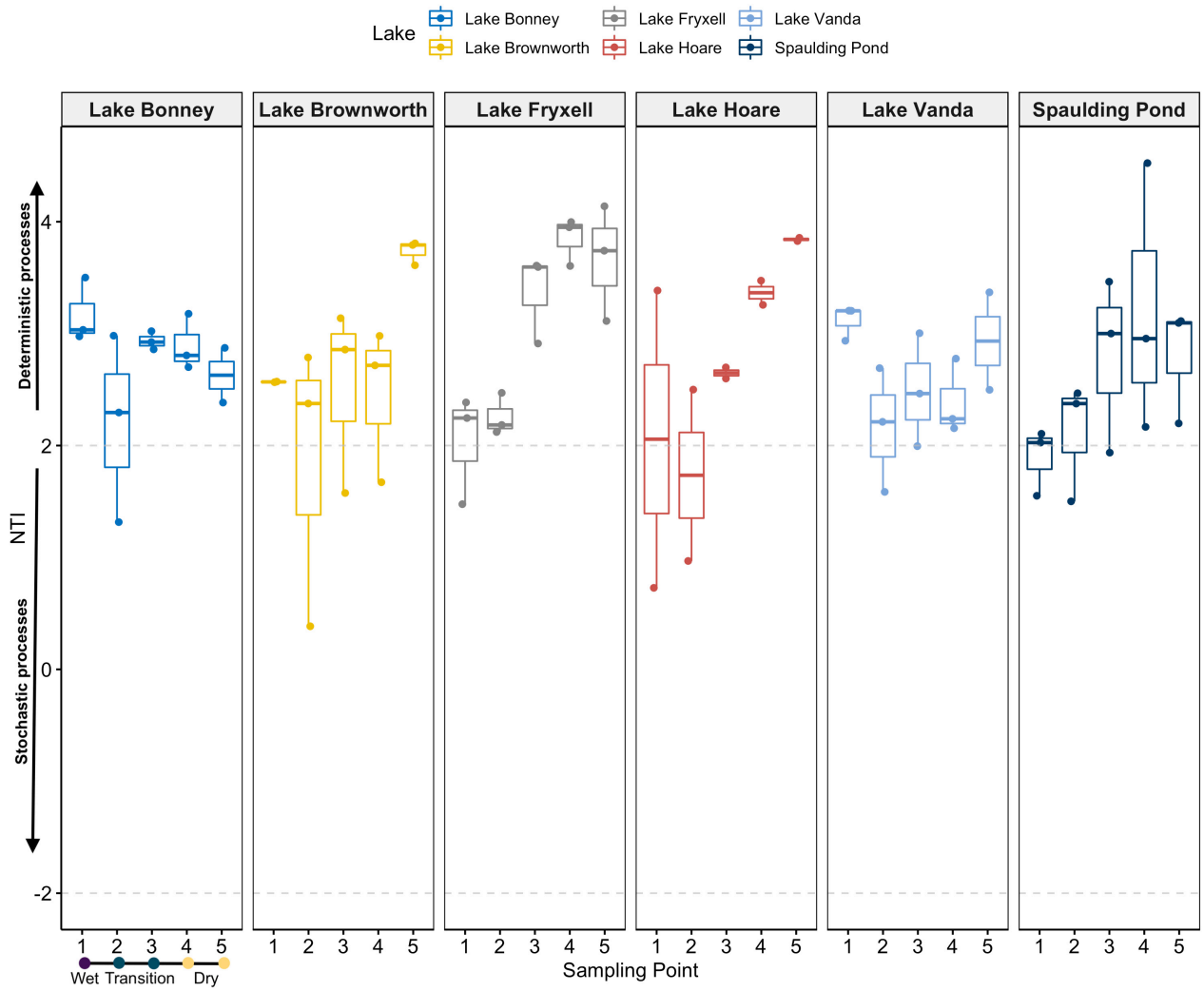

FIGURE 6 | NTI values for each single community sampled along space-for-time transects at each lake. Dashed gray lines at the −2 and + 2 values delimitate the significance thresholds from the null expectation. NTI values < −2 or > + 2 suggests that phylogenetic turnover is less or greater than the null expectation, and it is related to deterministic processes. NTI values ranging between –2 and 2 is usually considered signifying the influence of stochastic assembly.

of taxa that cannot tolerate such selective pressures (Chase, 2007; Stegen et al., 2012). Nonetheless, overall patterns of community assembly along increasing differences in moisture content (βNTI values) suggest a considerable but variable influence of stochastic and, to a lesser extent, deterministic processes, with a weak relation to differences in moisture content (**Supplementary Figure 8**). While this was surprising considering the NTI values, it suggests that, at a regional scale, community assembly is driven by stochastic processes, but phylogenetic differences within communities are driven by deterministic processes that favor the occurrence of closely related taxa. Such a pattern could indicate that microbial traits that confer adaptability are phylogenetically not well-conserved across the region. Alternatively, it could also result from the distribution of spore-forming, or dormant bacteria in dry or highly conductive soils (Lee et al., 2018), or from differences in diversity between lakes or along the gradients. Overall, we suggest that deterministic and stochastic processes play a role in community assembly, however, further extensive analyses are required to understand how deterministic and stochastic components change communities as moisture content changes in the soils.

The impact of moisture as a key driver of microbial diversity was evident at all lakes studied except in Lake Bonney and Lake Vanda (**Figure 3A**). The presence of water in an available state for life has quantifiable impacts on soil microbial diversity within polar deserts (Niederberger et al., 2015, 2019; Buelow et al., 2016; Bottos et al., 2020). Nonetheless, its positive impact can be suppressed by the effect of soil conductivity (Van Horn et al., 2014; Feeser et al., 2018; Bottos et al., 2020; George et al., 2021). This was the case of Lake Bonney, where the high levels of soil conductivity

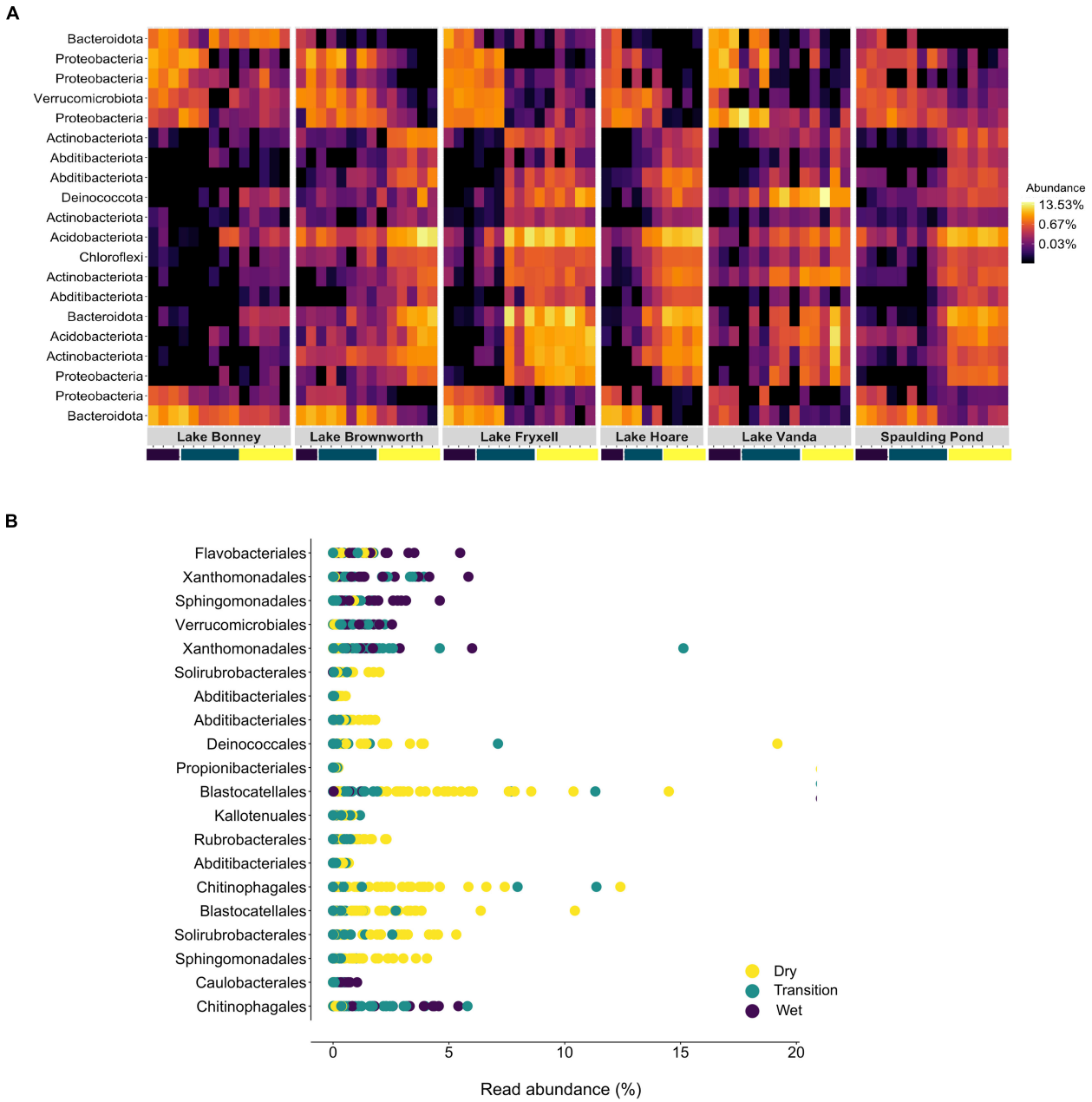

FIGURE 7 | (A) Heatmap of the top 20 bacterial OTUs classified by Random Forest classification analysis as the most important to discriminate between the different zones of lake transects. Each row represents each OTU at the phylum level, and the colour of the box indicates the relative abundance of each OTU, with yellow depicting high relative abundance and black low relative abundance. **(B)** Rank abundance plot of the correspondent top 20 OTUs identified across the different zones of the SFT transects at the order level.

affected the microbial diversity along the wetness gradient (**Figure 3B**), creating a niche that only allows for taxa capable of tolerating high osmotic stress (e.g., *Firmicutes*) to thrive. Lake level rise caused by increased glacial runoff in Lake Bonney facilitates the leaching and movement of accumulated meteoric salts, altering soil geochemistry, particularly pH and conductivity (Barrett et al., 2009). High conductivity measurements in these soils support our evidence for the greater prevalence of deterministic processes driving *within* community assembly in Lake Bonney and, to a lesser extent, in Lake Vanda, where the moisture content and relatively higher salt concentration in the soils may have a constant selective effect on the community. Both lakes are located further inland, where soils are characteristically older, more developed, and with higher salt concentrations and with different chemistry, compared to coastal soils (Bockheim, 2002). These could also

be factors for such disparities in community diversity and composition along transects sampled from these two inland lakes. Although pH has been suggested to be a critical factor driving spatial patterns of microbial communities (Tripathi et al., 2018), in this study, pH changes within and between space-for-time transects did not have a significant effect on community diversity.

Specific phyla and orders of microbial taxa associated with long-term dry and wet conditions can provide a metric to assess early signs of change (**Figure 7**). Despite the variability in the structure of microbial communities among wet soil samples, OTUs affiliated with *Verrucomicrobiota* were consistent indicators of long-term wet conditions, whereas OTUs affiliated with *Proteobacteria* and *Bacteroidetes* were associated with both wet and dry conditions. On the other hand, *Actinobacteria*, *Acidobacteria*, *Chloroflexi,* and *Deinococcota* taxa significantly and consistently enriched communities from dry soils, corroborating our previous observations conducted in the MDVs (Niederberger et al., 2015, 2019; Zhang et al., 2020). A recent study by Lee et al. (2018) found similar taxonomic trends, with *Acidobacteria* and *Actinobacteria* being less abundant with higher moisture content. We then provide strong evidence that the presence of discrete taxa within the phyla *Actinobacteria* (family *Solirubrobacteraceae*), *Acidobacteria* (subdivision 4 genera *Blastocatella),* and *Deinococcota* (genus *Truepera*) are ubiquitously associated with long-term dry edaphic conditions. This could be a result of metabolic specialization in the shape of very limiting long-term dry, oligotrophic, and ionizing (UV radiation) conditions to life. Taxa belonging to the family *Solirubrobacteraceae* are currently known for their capability to metabolize atmospheric trace gases contributing to primary production in the dry oligotrophic soils (Ji et al., 2017), and members of the genus *Truepera* are known for being extremely resistant to ionizing radiation, which is considered a strong stress factor in arid desert soils (Cary et al., 2010). The alleviation of specific stress conditions will impose selective pressures on these highly adapted taxa. As such, any quantifiable taxonomic changes among these taxa can be developed as sentinels for increased temporal exposure to moisture within the MDVs soils, which from a structural standpoint shows a heightened sensitivity of the terrestrial Antarctic microbiome to hydrological changes.

Validation of the Space-for-Time Approach

The space-for-time approach employs ecological structures within contemporary environmental gradients to project ecological responses to changes in the environment over time. Primarily applied to study slow ecological processes, this approach is informative when factors that drive community change are equivalent across space and time (Blois et al., 2013). From a microbial ecology perspective, the geochemical profile of a system is often highlighted as an important, if not the primary deterministic force driving community assembly, particularly in ecosystems with low trophic complexity or where diversity is environmentally constrained (Chase, 2007). Yet, sampling designs along spatial gradients are often conducted following distance points (Yang et al., 2014; Niederberger et al., 2015; Yan et al., 2017; Feeser et al., 2018; Lee et al., 2018) rather than being based on the spatial variability of geochemical factors, to which biological communities respond to, along a constrained gradient. The patterns of geochemical drivers across space may not be continuous along a gradient, therefore without characterizing them, geochemical boundaries that may drive community assembly within the gradient are likely to be missed, contributing to extant microbial community heterogeneity.

In an environmentally constrained system, such as the Antarctic polar deserts, we demonstrate that contemporary and historical environmental changes (e.g., water availability) can be resolved by subtle structural and compositional re-arrangements of the microbial communities along geochemically defined environmental gradients. These observations could only be depicted when sampling approaches are determined by the spatial variability of geochemical drivers along environmental gradients and replicated across comparable environments that differ in historical exposure to the driving factor. The validation of our space-for-time transects relies on the observed interdependence between structural and compositional community shifts across space, with the temporal exposure to wet conditions.

Given the accelerated rate of climate change across the globe (IPCC, 2021), the ongoing increasing lake levels (Castendyk et al., 2016; Levy et al., 2018), and the expected increase in snow and ground ice melt (Linhardt et al., 2019) will lead to drastic changes in the MDVs landscape, hydrology and ecology (Fountain et al., 2014). For instance, a recent study conducted in a high Arctic lake demonstrated that the taxonomic and functional diversity of dominant microbes can be impacted by an increase in glacial runoff as a consequence of rapid climate change (Colby et al., 2020). Our study, demonstrates that predictions of a wetter system will directly affect the stability of microbial communities currently adapted to dry and ultra-oligotrophic conditions, resulting in significant compositional and diversity changes across the system. Such environmental and taxonomical changes have the potential to lead to alterations in primary productivity, carbon and nitrogen cycling, and energy fluxes across the system (Coyne et al., 2020; Monteiro et al., 2020; Karen et al., 2021). Whilst microorganisms underpin the stability and functionality of polar deserts ecosystems, these are still rarely considered in policy development and lack any protective status within the Antarctic Treaty (Hughes et al., 2015). This study highlights that microbial communities are highly sensitive to change and closely reflect changes in the environment. Therefore, as in more temperate environments (Cavicchioli et al., 2019), microbial communities should be considered by policy-makers as primary sentinels for current and historical changes in the Antarctic systems.

DATA AVAILABILITY STATEMENT

The raw DNA sequences presented in this study can be found in the online repository: https://www.ncbi.nlm.nih.gov/bioproject/ under the BioProject accession number PRJNA764757.

AUTHOR CONTRIBUTIONS

SC, CL, and IH conceived and designed the project. SC and IH support for the project obtained. MM and IM performed by the all fieldwork analyses and sample collection. MM and AM conducted all subsequent laboratory work, DNA sequencing, and analyses. The manuscript was written by MM, AM, and SC with contribution and editing by CL, IM, and IH. All authors contributed to the article and approved the submitted version.

FUNDING

This research was conducted as part of the Resilience in Antarctic Biota and Ecosystems Project funded by the New Zealand Antarctic Research Institute to SC, IH, and CL (NZARI 2016-2). SC, CL, and IH were partially supported by the Dry Valley Ecosystem Resilience (DryVER) Programme funded by the New Zealand Ministry of Business, Innovation and Employment (UOWX1401). AM was supported by a Smart Ideas award (UOWX1602) from the New Zealand Ministry of Business, Innovation and Employment and the Rutherford Foundation Royal Society Te Aparangi Postdoctoral Fellowship (20-UOW006). MM was supported by a University of Waikato Doctoral Scholarship.

ACKNOWLEDGMENTS

We thank Antarctica New Zealand for the critical logistical support needed to carry out the field operations.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.783767/full#supplementary-material

REFERENCES

Anderson, M. J., Ellingsen, K. E., and McArdle, B. H. (2006). Multivariate dispersion as a measure of beta diversity. *Ecol. Lett.* 9, 683–693. doi: 10.1111/j.1461-0248.2006.00926.x

Barrett, J. E., Gooseff, M. N., and Takacs-Vesbach, C. (2009). Spatial variation in soil active-layer geochemistry across hydrologic margins in Polar desert ecosystems. *Hydrol. Earth Syst. Sci.* 13, 2349–2358. doi: 10.5194/hess-13-2349-2009

Bernal, A., Ear, U., and Kyrpides, N. (2001). Genomes OnLine Database (GOLD): a monitor of genome projects world-wide. *Nucleic Acids Res.* 29, 126–127. doi: 10.1093/nar/29.1.126

Blois, J. L., Williams, J. W., Fitzpatrick, M. C., Jackson, S. T., and Ferrier, S. (2013). Space can substitute for time in predicting climate-change effects on biodiversity. *Proc. Natl. Acad. Sci. U.S.A.* 110, 9374–9379. doi: 10.1073/pnas.1220228110

Bockheim, J. G. (2002). Landform and soil development in the McMurdo Dry valleys, Antarctica: a regional synthesis. *Arct. Antarc. Alp. Res.* 34, 308–317. doi: 10.1080/15230430.2002.12003499

Bokulich, N. A., Subramanian, S., Faith, J. J., Gevers, D., Gordon, J. I., Knight, R., et al. (2013). Quality-filtering vastly improves diversity estimates from Illumina amplicon sequencing. *Nat. Methods* 10, 57–59. doi: 10.1038/nmeth.2276

Bottos, E. M., Laughlin, D. C., Herbold, C. W., Lee, C. K., McDonald, I. R., and Cary, S. C. (2020). Abiotic factors influence patterns of bacterial diversity and community composition in the Dry valleys of Antarctica. *FEMS Microbiol. Ecol.* 96:fiaa042. doi: 10.1093/femsec/fiaa042

Buelow, H. N., Winter, A. S., Van Horn, D. J., Barrett, J. E., Gooseff, M. N., Schwartz, E., et al. (2016). Microbial community responses to increased water and organic matter in the arid soils of the McMurdo Dry valleys, Antarctica. *Front. Microbiol.* 7:1040. doi: 10.3389/fmicb.2016.01040

Caruso, T., Chan, Y., Lacap, D. C., Lau, M. C. Y., McKay, C. P., and Pointing, S. B. (2011). Stochastic and deterministic processes interact in the assembly of desert microbial communities on a global scale. *ISME J.* 5, 1406–1413. doi: 10.1038/ismej.2011.21

Cary, S., Mcdonald, I., Barrett, J., and Cowan, D. (2010). On the rocks: the microbiology of Antarctic dry valley soils. *Nat. Rev. Microbiol.* 8, 129–138. doi: 10.1038/nrmicro2281

Castendyk, D. N., Obryk, M. K., Leidman, S. Z., Gooseff, M., and Hawes, I. (2016). Lake Vanda: a sentinel for climate change in the McMurdo sound region of Antarctica. *Glob. Planet. Change* 144, 213–227. doi: 10.1016/j.gloplacha.2016.06.007

Cavicchioli, R., Ripple, W. J., Timmis, K. N., Azam, F., Bakken, L. R., Baylis, M., et al. (2019). Scientists' warning to humanity: microorganisms and climate change. *Nat. Rev. Microbiol.* 17, 569–586. doi: 10.1038/s41579-019-0222-5

Chase, J. M. (2007). Drought mediates the importance of stochastic community assembly. *Proc. Natl. Acad. Sci. U.S.A.* 104, 17430–17434. doi: 10.1073/pnas.0704350104

Chase, J. M., and Myers, J. A. (2011). Disentangling the importance of ecological niches from stochastic processes across scales. *Philos. Trans. R. Soc. B Biol. Sci.* 366, 2351–2363. doi: 10.1098/rstb.2011.0063

Colby, G. A., Ruuskanen, M. O., St.Pierre, K. A., St.Louis, V. L., Poulain, A. J., and Aris-Brosou, S. (2020). Warming climate is reducing the diversity of dominant microbes in the largest high Arctic lake. *Front. Microbiol.* 11:2316. doi: 10.3389/fmicb.2020.561194

Coyne, K. J., Hutchins, D. A., and Hare, C. E. (2001). Assessing temporal and spatial variability in *Pfiesteria piscicida* distributions using molecular probing techniques. *Aquat. Microb. Ecol.* 24, 275–285. doi: 10.3354/ame024275

Coyne, K. J., Parker, A. E., Lee, C. K., Sohm, J. A., Kalmbach, A., Gunderson, T., et al. (2020). The distribution and relative ecological roles of autotrophic and heterotrophic diazotrophs in the McMurdo Dry valleys, Antarctica. *FEMS Microbiol. Ecol.* 96:fiaa010. doi: 10.1093/femsec/fiaa010

Damgaard, C. (2019). A critique of the space-for-time substitution practice in community ecology. *Trends Ecol. Evol.* 34, 416–421. doi: 10.1016/j.tree.2019.01.013

Dini-Andreote, F., Stegen, J. C., van Elsas, J. D., and Salles, J. F. (2015). Disentangling mechanisms that mediate the balance between stochastic and deterministic processes in microbial succession. *Proc. Natl. Acad. Sci. U.S.A.* 112, E1326–E1332. doi: 10.1073/pnas.1414261112

Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26, 2460–2461. doi: 10.1093/bioinformatics/btq461

Edgar, R. C. (2013). UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat. Methods* 10, 996–998. doi: 10.1038/nmeth.2604

Feeser, K. L., Van Horn, D. J., Buelow, H. N., Colman, D. R., McHugh, T. A., Okie, J. G., et al. (2018). Local and regional scale heterogeneity drive bacterial community diversity and composition in a Polar desert. *Front. Microbiol.* 9:1928. doi: 10.3389/fmicb.2018.01928

Fountain, A. G., Levy, J. S., Gooseff, M. N., and Van Horn, D. (2014). The McMurdo Dry Valleys: a landscape on the threshold of change. *Geomorphology* 225, 25–35. doi: 10.1016/j.geomorph.2014.03.044

Fountain, A. G., Saba, G., Adams, B., Doran, P., Fraser, W., Gooseff, M., et al. (2016). The impact of a large-scale climate event on Antarctic ecosystem processes. *Bioscience* 66, 848–863. doi: 10.1093/biosci/biw110

George, S. F., Fierer, N., Levy, J. S., and Adams, B. (2021). Antarctic water tracks: microbial community responses to variation in soil moisture, pH, and salinity. *Front. Microbiol.* 12:616730. doi: 10.3389/fmicb.2021.616730

Hughes, K. A., Cowan, D. A., and Wilmotte, A. (2015). Protection of Antarctic microbial communities - "out of sight, out of mind". *Front. Microbiol.* 6:151. doi: 10.3389/fmicb.2015.00151

IPCC (2021). *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, eds V. Masson-Delmotte, P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, et al. (Cambridge: Cambridge University Press).

Ji, M., Greening, C., Vanwonterghem, I., Carere, C. R., Bay, S. K., Steen, J. A., et al. (2017). Atmospheric trace gases support primary production in Antarctic desert surface soil. *Nature* 552, 400–403. doi: 10.1038/nature25014

Johnson, E. A., and Miyanishi, K. (2008). Testing the assumptions of chronosequences in succession. *Ecol. Lett.* 11, 419–431. doi: 10.1111/j.1461-0248.2008.01173.x

Kappes, H., Sundermann, A., and Haase, P. (2010). High spatial variability biases the space-for-time approach in environmental monitoring. *Ecol. Indic.* 10, 1202–1205. doi: 10.1016/j.ecolind.2010.03.012

Karen, J., Rachael, L., Xiyang, D., Ian, J. A., Sean, K. B., Eleonora, C., et al. (2021). Hydrogen-oxidizing bacteria are abundant in desert soils and strongly stimulated by hydration. *mSystems* 5, e01131-20. doi: 10.1128/mSystems.01131-20

Kembel, S. W., Cowan, P. D., Helmus, M. R., Cornwell, W. K., Morlon, H., Ackerly, D. D., et al. (2010). Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* 26, 1463–1464. doi: 10.1093/bioinformatics/btq166

Kratz, T. K., Deegan, L. A., Harmon, M. E., and Lauenroth, W. K. (2003). Ecological variability in space and time: insights gained from the US LTER program. *Bioscience* 53, 57–67. doi: 10.1641/0006-3568(2003)053[0057:evisat]2.0.co;2

Kwon, M., Kim, M., Takacs-Vesbach, C., Lee, J., Hong, S. G., Kim, S. J., et al. (2017). Niche specialization of bacteria in permanently ice-covered lakes of the McMurdo Dry valleys, Antarctica. *Environ. Microbiol.* 19, 2258–2271. doi: 10.1111/1462-2920.13721

Lee, C. K., Barbier, B. A., Bottos, E. M., McDonald, I. R., and Cary, S. C. (2012). The inter-valley soil comparative survey: the ecology of Dry valley edaphic microbial communities. *ISME J.* 6, 1046–1057. doi: 10.1038/ismej.2011.170

Lee, K. C., Caruso, T., Archer, S. D. J., Gillman, L. N., Lau, M. C. Y., Cary, S. C., et al. (2018). Stochastic and deterministic effects of a moisture gradient on soil microbial communities in the McMurdo Dry valleys of Antarctica. *Front. Microbiol.* 9:2619. doi: 10.3389/fmicb.2018.02619

Levy, J. S., Fountain, A. G., Obryk, M., Telling, J., Glennie, C., Pettersson, R., et al. (2018). Decadal topographic change in the McMurdo Dry valleys of Antarctica: Thermokarst subsidence, glacier thinning, and transfer of water storage from the cryosphere to the hydrosphere. *Geomorphology* 323, 80–97. doi: 10.1016/j.geomorph.2018.09.012

Liaw, A., and Wiener, M. (2002). Classification and regression by randomForest. *R News* 2, 18–22.

Linhardt, T., Levy, J. S., and Thomas, C. K. (2019). Water tracks intensify surface energy and mass exchange in the Antarctic McMurdo Dry valleys. *Cryosphere* 13, 2203–2219. doi: 10.5194/tc-13-2203-2019

Lozupone, C. A., and Knight, R. (2015). The unifrac significance test is sensitive to tree topology. *BMC Bioinformatics* 16:211. doi: 10.1186/s12859-015-0640-y

Matsuoka, K., Skoglund, A., Roth, G., de Pomereu, J., Griffiths, H., Headland, R., et al. (2021). Quantarctica, an integrated mapping environment for Antarctica, the Southern Ocean, and sub-Antarctic islands. *Environ. Model. Softw.* 140:105015. doi: 10.1016/j.envsoft.2021.105015

McMurdie, P. J., and Holmes, S. (2013). phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* 8:e61217. doi: 10.1371/journal.pone.0061217

Minasny, B., McBratney, A. B., Brough, D. M., and Jacquier, D. (2011). Models relating soil pH measurements in water and calcium chloride that incorporate electrolyte concentration. *Eur. J. Soil Sci.* 62, 728–732. doi: 10.1111/j.1365-2389.2011.01386.x

Monteiro, M., Baptista, M. S., Séneca, J., Torgo, L., Lee, C. K., Cary, S. C., et al. (2020). Understanding the response of nitrifying communities to disturbance in the McMurdo Dry valleys, Antarctica. *Microorganisms* 8:404. doi: 10.3390/microorganisms8030404

Niederberger, T. D., Bottos, E. M., Sohm, J. A., Gunderson, T., Parker, A., Coyne, K. J., et al. (2019). Rapid microbial dynamics in response to an induced wetting event in Antarctic Dry valley soils. *Front. Microbiol.* 10:621. doi: 10.3389/fmicb.2019.00621

Niederberger, T. D., Sohm, J. A., Gunderson, T. E., Parker, A. E., Tirindelli, J., Capone, D. G., et al. (2015). Microbial community composition of transiently wetted Antarctic Dry valley soils. *Front. Microbiol.* 6:9. doi: 10.3389/fmicb.2015.00009

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P., O'Hara, R. B., et al. (2012). *Vegan: Community Ecology Package. R Package Version 2.0-2*.

Parada, A. E., Needham, D. M., and Fuhrman, J. A. (2016). Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environ. Microbiol.* 18, 1403–1414. doi: 10.1111/1462-2920.13023

Pickett, S. T. A. (1989). "Space-for-time substitution as an alternative to long-term studies," in *Long-Term Studies in Ecology: Approaches and Alternatives*, ed. G. E. Likens (New York, NY: Springer New York), 110–135. doi: 10.1007/978-1-4615-7358-6_5

Pruesse, E., Peplies, J., and Glöckner, F. O. (2012). SINA: accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics* 28, 1823–1829. doi: 10.1093/bioinformatics/bts252

Ramoneda Massague, J., Hawes, I., Pascual-García, A., Mackey, T. J., Sumner, D. Y., and Jungblut, A. (2021). Importance of environmental factors over habitat connectivity in shaping bacterial communities in microbial mats and bacterioplankton in an Antarctic freshwater system. *FEMS Microbiol. Ecol.* 97:fiab044. doi: 10.1093/femsec/fiab044

Rustad, L. E. (2008). The response of terrestrial ecosystems to global climate change: towards an integrated approach. *Sci. Total Environ.* 404, 222–235. doi: 10.1016/j.scitotenv.2008.04.050

R Core Team (2000). *R Language Definition*. Vienna: R Foundation for Statistical Computing. Available online at: https://cran.r-project.org/doc/manuals/r-release/R-lang.pdf

Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., et al. (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol.* 75, 7537–7541. doi: 10.1128/AEM.01541-09

Stegen, J. C., Lin, X., Fredrickson, J. K., Chen, X., Kennedy, D. W., Murray, C. J., et al. (2013). Quantifying community assembly processes and identifying features that impose them. *ISME J.* 7, 2069–2079. doi: 10.1038/ismej.2013.93

Stegen, J. C., Lin, X., Konopka, A. E., and Fredrickson, J. K. (2012). Stochastic and deterministic assembly processes in subsurface microbial communities. *ISME J.* 6, 1653–1664. doi: 10.1038/ismej.2012.22

Tripathi, B. M., Stegen, J. C., Kim, M., Dong, K., Adams, J. M., and Lee, Y. K. (2018). Soil pH mediates the balance between stochastic and deterministic assembly of bacteria. *ISME J.* 12, 1072–1083. doi: 10.1038/s41396-018-0082-4

van den Boogaart, K. G., and Tolosana-Delgado, R. (2008). "compositions": A unified R package to analyze compositional data. *Comput. Geosci.* 34, 320–338. doi: 10.1016/j.cageo.2006.11.017

Van Horn, D., Okie, J. G., Buelow, H. N., Gooseff, M. N., Barrett, J. E., Takacs-Vesbach, C. D., et al. (2014). Soil microbial responses to increased moisture and organic resources along a salinity gradient in a Polar desert. *Appl. Environ. Microbiol.* 80, 3034–3043. doi: 10.1128/AEM.03414-13

Walker, L. R., Wardle, D. A., Bardgett, R. D., and Clarkson, B. D. (2010). The use of chronosequences in studies of ecological succession and soil development. *J. Ecol.* 98, 725–736. doi: 10.1111/j.1365-2745.2010.01664.x

Wickham, H. (2016). *ggplot: Elegant Graphics for Data Analysis*. New York, NY: Springer.

Wilhelm, L., Singer, G. A., Fasching, C., Battin, T. J., and Besemer, K. (2013). Microbial biodiversity in glacier-fed streams. *ISME J.* 7, 1651–1660. doi: 10.1038/ismej.2013.44

Wogan, G. O. U., and Wang, I. J. (2018). The value of space-for-time substitution for studying fine-scale microevolutionary processes. *Ecography* 41, 1456–1468. doi: 10.1111/ecog.03235

Yan, W., Ma, H., Shi, G., Li, Y., Sun, B., Xiao, X., et al. (2017). Independent shifts of abundant and rare bacterial populations across East Antarctica glacial foreland. *Front. Microbiol.* 8:1534. doi: 10.3389/fmicb.2017.01534

Yang, Y., Gao, Y., Wang, S., Xu, D., Yu, H., Wu, L., et al. (2014). The microbial gene diversity along an elevation gradient of the Tibetan grassland. *ISME J.* 8, 430–440. doi: 10.1038/ismej.2013.146

Zhang, E., Thibaut, L. M., Terauds, A., Raven, M., Tanaka, M. M., van Dorst, J., et al. (2020). Lifting the veil on arid-to-hyperarid Antarctic soil microbiomes: a tale of two oases. *Microbiome* 8:37. doi: 10.1186/s40168-020-00809-w

Zhou, J., and Ning, D. (2017). Stochastic community assembly: does it matter in microbial ecology? *Microbiol. Mol. Biol. Rev.* 81:e00002-17. doi: 10.1128/MMBR.00002-17

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

frontiers in Microbiology

ORIGINAL RESEARCH
published: 11 February 2022
doi: 10.3389/fmicb.2022.779505

Local Habitat Filtering Shapes Microbial Community Structure in Four Closely Spaced Lakes in the High Arctic

*Catherine Marois[1,2,3], Catherine Girard[2,4], Yohanna Klanten[2,5], Warwick F. Vincent[2,3,6], Alexander I. Culley[1,2,3] and Dermot Antoniades[2,5]**

[1] Département de Biochimie, Microbiologie et Bio-Informatique, Université Laval, Québec, QC, Canada, [2] Centre d'Études Nordiques (CEN), Université Laval, Québec, QC, Canada, [3] Institut de Biologie Intégrative des Systèmes (IBIS), Université Laval, Québec, QC, Canada, [4] Département des Sciences Fondamentales, Université du Québec à Chicoutimi, Chicoutimi, QC, Canada, [5] Département de Géographie, Université Laval, Québec, QC, Canada, [6] Département de Biologie, Université Laval, Québec, QC, Canada

OPEN ACCESS

Edited by:
David Velazquez,
Autonomous University of Madrid, Spain

Reviewed by:
Carlos Rochera Cordellat,
University of Valencia, Spain
Byron C. Crump,
Oregon State University, United States

***Correspondence:**
Dermot Antoniades
dermot.antoniades@cen.ulaval.ca

Specialty section:
This article was submitted to Extreme Microbiology, a section of the journal Frontiers in Microbiology

Received: 18 September 2021
Accepted: 20 January 2022
Published: 11 February 2022

Citation:
Marois C, Girard C, Klanten Y, Vincent WF, Culley AI and Antoniades D (2022) Local Habitat Filtering Shapes Microbial Community Structure in Four Closely Spaced Lakes in the High Arctic. Front. Microbiol. 13:779505. doi: 10.3389/fmicb.2022.779505

Arctic lakes are experiencing increasingly shorter periods of ice cover due to accelerated warming at northern high latitudes. Given the control of ice cover thickness and duration over many limnological processes, these changes will have pervasive effects. However, due to their remote and extreme locations even first-order data on lake ecology is lacking for many ecosystems. The aim of this study was to characterize and compare the microbial communities of four closely spaced lakes in Stuckberry Valley (northern Ellesmere Island, Canadian Arctic Archipelago), in the coastal margin zone of the Last Ice Area, that differed in their physicochemical, morphological and catchment characteristics. We performed high-throughput amplicon sequencing of the V4 16S rRNA gene to provide inter- and intra-lake comparisons. Two deep (>25 m) and mostly oxygenated lakes showed highly similar community assemblages that were distinct from those of two shallower lakes (<10 m) with anoxic bottom waters. *Proteobacteria*, *Verrucomicrobia,* and *Planctomycetes* were the major phyla present in the four water bodies. One deep lake contained elevated proportions of *Cyanobacteria* and *Thaumarchaeota* that distinguished it from the others, while the shallow lakes had abundant communities of predatory bacteria, as well as microbes in their bottom waters that contribute to sulfur and methane cycles. Despite their proximity, our data suggest that local habitat filtering is the primary determinant of microbial diversity in these systems. This study provides the first detailed examination of the microbial assemblages of the Stuckberry lakes system, resulting in new insights into the microbial ecology of the High Arctic.

Keywords: diversity, connectivity, predatory bacteria, Ellesmere Island, Stuckberry Valley, amplicon sequence variant (ASV), Arctic lake

INTRODUCTION

The Arctic is among the regions on Earth most affected by the rapid acceleration of global warming. It is estimated that air temperature is increasing more than twice as fast as the global average due to Arctic amplification (Meredith et al., 2019). The region is facing changes in its oceanic and atmospheric circulation, ice and snow extent and duration, permafrost, hydrology, vegetation, and carbon cycling (Wrona et al., 2016; Vincent, 2020; Cai et al., 2021). These changes have repercussions for ecosystem function at different spatial and temporal scales that are still not fully understood.

Lakes are sentinels of environmental change because they are integrators of watershed and airshed processes (e.g., plant and soil dynamics, nutrient loading, climate shifts, anthropogenic impacts; Williamson et al., 2008; Vincent, 2018). They are important and highly diverse features of the Arctic landscape, ranging from deep proglacial lakes to shallow thermokarst ponds (Pienitz et al., 2008). Due to warming climates, the abundance and area of glacial lakes are increasing (Shugar et al., 2020), while other types of lakes are disappearing (Smith et al., 2005). With their simplified food webs and unique physical and chemical characteristics (low water temperature, prolonged ice cover and highly seasonal photoperiods; Vincent, 2018), these extreme waterbodies represent key habitats that can be used to better understand the polar biome in transition.

Microorganisms dominate the food webs of Arctic lakes (Vincent, 2010), and their physiological diversity underpins the four major biogeochemical cycles that occur in aquatic ecosystems: carbon (including methane), sulfur, phosphorus, and nitrogen (Falkowski et al., 2008). Microbial distribution is determined by competition for heterogeneous niches created throughout the water column by chemical and physical gradients and other controlling forces such as predation and viral lysis (e.g., Shade et al., 2008; Comeau et al., 2012; Somers et al., 2020). Despite their extreme conditions, High Arctic Lakes host a relatively high diversity of microbial phyla, many of which are also found in temperate environments. For example, *Actinobacteria*, *Bacteroidetes*, *Cyanobacteria*, *Proteobacteria,* and *Verrucomicrobia* are phyla that are commonly found in temperate freshwater microbial communities (Zwart et al., 2002; Newton et al., 2011), but have also been identified in Arctic lakes such as Fuglebekken and Revvatnet (Norway) (Ntougias et al., 2016), Lake Hazen (Canada) (Cavaco et al., 2019), and Ward Hunt Lake (Canada) (Comte et al., 2018), as well as in Antarctic lakes (Byers Peninsula) (Picazo et al., 2019). Mat-producing and free-living (picoplanktonic) *Cyanobacteria* are the major bacterial primary producers in high-latitude lakes, and they conduct oxygenic photosynthesis (Vincent, 2018). Green and purple sulfur and non-sulfur bacteria carry out photosynthesis in low oxygen hypolimnia (Lizotte, 2008). Methanogens and methanotrophs, as well as sulfur reducers, sulfur oxidizers, and nitrifiers, are distributed above, across, or below oxyclines according to their specific redox requirements, for example as demonstrated by Schutte et al. (2016) in Potentilla Lake, a seasonally ice-covered dimictic lake in Greenland, and Vigneron et al. (2021) in meromictic Lake A in the Canadian High Arctic.

Microorganisms are responding rapidly to changing conditions in the warming Arctic (Yadav et al., 2019). For example, psychrophilic microorganisms may be replaced by psychrotrophic ones (Vincent, 2010), and the northward expansion of the boreal and tundra vegetation boundaries may contribute greater inputs of organic matter and increase lake productivity (Wrona et al., 2016). Climate change is also expected to modify microbially driven biogeochemical cycles. Indeed, mainly due to increased ice-free periods, perturbations have been observed, including the intensification of methane emissions (Tan and Zhuang, 2015) and the accumulation of reduced sulfur in sediments (Drevnick et al., 2010). In short, in addition to being key components of lake ecosystems, Arctic lacustrine microbial communities are greatly affected by climate change. However, the way they are affected, and the consequences of these changes, remain poorly understood. Although we were unable to determine the response of these microbial communities to climate change in this study, it is urgent to characterize the diversity of current communities before they are transformed.

To better understand microbial ecology in High Arctic lakes, we examined four lakes in Stuckberry Valley that are closely spaced and physicochemically and morphologically diverse (Top, Y, 2FB, and Bottom lakes). Located on the northern coast of Ellesmere Island (Nunavut, Canada), approximately 780 km from the North Pole, these lakes are under great pressure due to climate change (Mueller et al., 2009; White and Copland, 2018; Moore et al., 2019). They lie in the coastal margin zone of the Last Ice Area, which contains the thickest sea ice in the Arctic Ocean, but is now subject to rapid warming and ice loss. The vulnerability of this area to environmental change is evidenced by the recent loss of 43% of the Milne Ice Shelf, the last intact Canadian ice shelf (Vincent and Mueller, 2020), and the over 90% reduction of the coastal ice shelves of Ellesmere Island since the beginning of the twentieth century (Copland and Mueller, 2017). The four Stuckberry Valley lakes were successively isolated from the Arctic Ocean by glacioisostatic rebound, a process in which glacial retreat leads to rising of the Earth's crust, trapping water, and thus creating lakes.

Previous comparisons of morphologically distinct lakes in the Byers Peninsula (Antarctica) have shown that nutrients, pigments, and bacterial communities follow a gradient from inland to coastal sites (Rochera et al., 2013) and that this shift was reflected in predicted microbial metabolism (Picazo et al., 2021). The objective of this study was to characterize and compare the microbial communities within four lacustrine ecosystems in the Canadian High Arctic. We hypothesized that local habitat filtering based on the distinct chemical and physical gradients, as well as the unique catchment and morphological characteristics of each waterbody, would be the primary determinant of microbial diversity despite their close proximity. To test this hypothesis, we performed deep sequencing of the V4 region of the 16S rRNA gene on samples collected at several depths in each water column. We then determined richness and evenness within communities, the identified dominant phyla, and similarities in microbial community composition in the context of physicochemical water column variability. These data also enabled us to measure the impacts of local and regional forcing,

and thus lay the foundation for a better understanding of the ecological impacts of a warming Arctic.

RESULTS AND DISCUSSION

Limnological Characteristics

The Stuckberry Valley lakes fell into two categories based on their limnological and physical characteristics: deep and shallow (**Figure 1**). The deep lakes (Top and Y) had maximum depths of 49 and 28 m, respectively, and larger watershed areas (14.4 and 17.5 km^2). At the time of sampling, dissolved oxygen (DO,% saturation) levels were between 85 and 100% at the surface and declined steadily with depth, reaching hypoxic conditions in Top Lake (5.6%) at 45 m (**Supplementary Figure 1**). Hypoxic conditions were not reached in Y Lake. Both lakes had low specific conductivities (mean water column values of 143 and 139 μScm^{-1} for Top and Y lakes, respectively) and DOC concentrations (respective means of 0.4 and 0.5 mgL^{-1}). A stream directly connected Top Lake with Y Lake, and is likely a major reason for the similar characteristics of the two lakes. The shallow lakes (2FB and Bottom) were less than 10 m deep and had smaller watershed areas (each ∼0.4 km^2). They had steep oxyclines, with DO concentrations of ∼50% just below the surface that decreased sharply at a depth of ∼2 m before becoming anoxic near the lake bottoms. The shallow lakes had notably higher specific conductivities (mean water column values of 387 and 401 μScm^{-1} in 2FB and Bottom lakes, respectively) and DOC concentrations (means = 1.3 and 2.2 mgL^{-1}, respectively) than the deeper lakes. The deeper, more dilute lakes were also more transparent, with 1% of solar radiation that arrived at the surface of the water (i.e., below the ice and snow) reaching 30.8 and 19.3 m in Top and Y lakes, but only 5.9 and 6.4 m in 2FB and Bottom lakes, respectively. Over 0.1% of PAR reached the sediment surface in all lakes except Top Lake.

Although the lakes were separated into two groups (shallow and deep) based on morphological properties, one of the shallow lakes (Bottom Lake) had unique characteristics, including a chlorophyll *a* peak at 4 m and higher total phosphorus (TP) and nitrogen (TN) concentrations (**Table 1**). In contrast, nutrient concentrations in Top, Y, and 2FB were typical of Arctic oligotrophic lakes (Lizotte, 2008). Viral particle and heterotrophic bacterial counts from flow cytometry were approximately five and four times higher, respectively, in Bottom Lake than in the others. We also observed a peak in photosynthetic eukaryotes at 4 m in Bottom Lake (**Table 1**), which was much deeper in the water column relative to the other lakes.

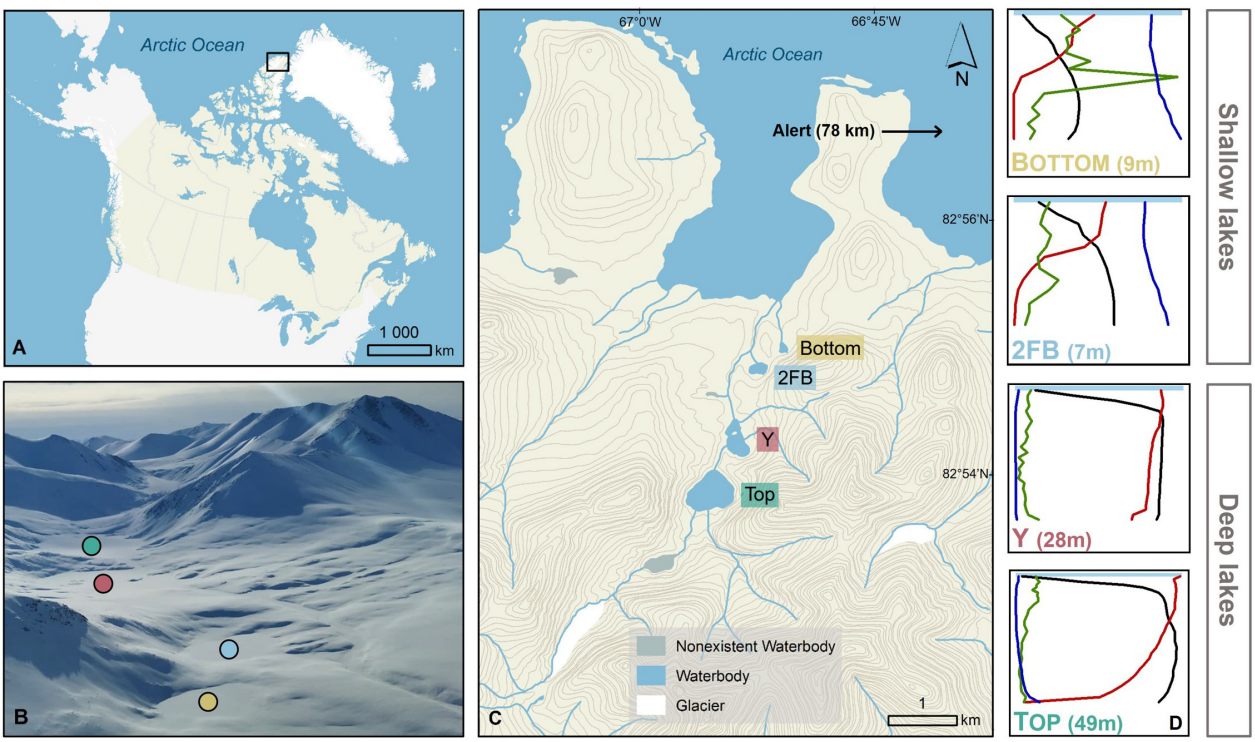

FIGURE 1 | Stuckberry Valley lakes. **(A)** Stuckberry Valley is located on the northern coast of Ellesmere Island on the edge of the Arctic Ocean **(B)** Aerial photograph of the four lakes surrounded by mountains. In late May, lakes were still covered with ice and snow (photo credit: Addison Gilpin-Payne). **(C)** The valley includes four lakes. **(D)** Chemical and physical gradients in the water columns of the four Stuckberry Valley lakes. For all four lakes, the maximum depth is in parentheses. Each profile parameter is associated with a color: green, chlorophyll *a*; blue, specific conductivity; black, temperature; red, dissolved oxygen. Complete profiles with values are available in **Supplementary Figure 1**. Maps were created using ESRI ArcMap (v.10.7). The digital elevation model comes from Natural Resources Canada (2001) and the boundary files are from Statistics Canada (2016).

TABLE 1 | Limnological characteristics of sampled depths for each studied lake.

		Limnological properties						Microbial abundance				
Lake	**Depth**	**DOC**	**TP**	**TN**	**Chlorophyll *a***	**SO_4^{2-}**	**PAR**	**Photosynthetic eukaryotes**	**Photosynthetic bacteria**	**Heterotrophic bacteria**	**Viruses**	**VPR**
	(m)	**(mgL^{-1})**	**(μgL^{-1})**	**(μgL^{-1})**	**(μgL^{-1})**	**(mgL^{-1})**	**%**	**(10^3mL^{-1})**	**(10^3mL^{-1})**	**(10^3mL^{-1})**	**(10^3mL^{-1})**	
Top	0	0.3	12.5	236	1.08	22.5	100	6.1 (7)	15.3 (16)	338.9 (15)	1137.9 (65)	3.2
	10	0.3	7.8	218	1.02	21.9	NA	4.1 (16)	6.8 (3)	172.3 (15)	1125.6 (19)	6.3
	20	0.3	3.2	231	0.60	22.1	NA	1.6 (18)	4.9 (10)	241.1 (5)	879.6 (17)	3.6
	45	0.7	5.6	285	0.42	22.1	NA	1.2 (3)	2.3 (1)	333.9 (8)	807.3 (7)	2.4
Y	0	0.9	4.0	237	0.84	23.0	100	2.4 (0)	1.0 (3)	359.5 (1)	944.0 (21)	2.6
	10	0.3	7.1	236	0.53	21.6	NA	1.2 (15)	1.5 (26)	349.2 (1)	1720.3 (47)	4.9
	25	0.6	4.4	280	1.27	21.9	NA	1.2 (12)	9.2 (8)	275.3 (19)	1147.8 (59)	4.0
2FB	0	1.4	8.7	154	1.91	42.8	100	1.8 (20)	0.6 (22)	358.5 (22)	2153.3 (10)	6,0
	2	1.4	8.5	176	1.74	42.6	22.31	2.9 (38)	0.4 (28)	327.6 (8)	2791.4 (46)	8.5
	3	1.2	4.4	183	1.39	41.2	14.23	2.4 (61)	0.7 (7)	314.2 (39)	2384.3 (29)	7.6
	5	1.3	7.1	235	0.82	38.6	3.46	2.7 (13)	11.7 (35)	419.1 (15)	1570.5 (47)	3.6
Bottom	0	2.1	23.8	310	2.54	32.9	100	6.8 (23)	3.4 (4)	1027.7 (2)	8084.5 (13)	7.8
	3	2.1	19.3	329	4.10	33.0	6.93	7.4 (29)	2.3 (16)	1056.0 (19)	7640.5 (9)	7.2
	4	1.9	22.8	362	8.78	30.9	3.80	10.4 (0)	11.7 (11)	1581.2 (10)	7215.8 (3)	4.5
	7.5	2.4	37.5	1020	1.13	20.1	0.15	1.8 (43)	103.6 (14)	936.4 (4)	7215.1 (6)	6.9

The limits of detection were 0.1 mgL^{-1} for dissolved organic matter (DOC), 0.5 μgL^{-1} for total phosphorus (TP), 15 μgL^{-1} for total nitrogen (TN), 0.01 μgL^{-1} for chlorophyll a and 0.1 mgL^{-1} for the sulfate concentration (SO_4^{2-}). Photosynthetically active radiation (PAR) could not be measured (NA) below a depth of 7.5 m. The virus-to-prokaryote ratio (VPR) was calculated by dividing the virus counts by the sum of photosynthetic and heterotrophic bacteria counts. Values represent the mean of two replicates with percent coefficient of variation (CV) in parentheses.

In all lakes, viral particles were approximately two to nine times more abundant than cellular microbes (**Table 1**), which corresponds with the averages observed in Arctic and Antarctic lakes (Sawstrom et al., 2008; Rochera et al., 2017). The virus-to-prokaryote ratio (VPR) can be influenced by multiple biological factors that affect viral and prokaryotic abundances (e.g., depth, season) (Wommack and Colwell, 2000), as well as methodological biases such as a high abundance of viruses (virus-like particles) that infect eukaryotes.

Intra-Lake Comparisons

All four lakes had chemical and physical gradients in their water column (**Figure 1** and **Table 1**). This heterogeneity suggests that there were distinct ecological niches for organisms (e.g., Shade et al., 2008; Comeau et al., 2012; Somers et al., 2020). Others have observed distinct bacterial communities at different depths in both non-Arctic and Arctic lakes with gradients (e.g., Shade et al., 2008; Schutte et al., 2016; Vigneron et al., 2021) and our results are consistent with these studies.

We investigated whether the microbial community composition, which contained a total of 4,155 unique amplicon sequence variants (ASVs) (**Supplementary Table 1**), reflected the observed partitioning along depth and chemical gradients. We did this by estimating the microbial diversity and characterizing the major phyla at each depth for the four lakes. We also identified features (or microbial biomarkers), which are differentially abundant features that most likely explain the differences noted between lakes, using linear discriminant analysis effect size (LEfSe) analysis (Segata et al., 2011; see below). A total of 20 features were identified among the four lakes (5, 3, 8, and 4 in Top, Y, 2FB and Bottom lakes, respectively).

ASV richness increased significantly with depth in Y, 2FB and Bottom lakes (Observed species index, **Figure 2A**). 2FB Lake had the richest microbial communities with 1,706 different ASVs, vs. 610, 1,093, and 909 for Top, Y and Bottom lakes, respectively. When evenness was considered (Shannon diversity index), only microbial communities in the bottom waters of the shallow lakes were significantly more diverse compared to the surface (**Figure 2B**). In Top Lake, we observed a decrease in evenness at depths of 20 m and below, while richness increased. Overall, the microbial community evenness was lower in Bottom than in 2FB and Y lakes, while richness remained similar; however, none of these differences were significant. One explanation for increases in microbial diversity with depth may be that the ice cover prevents wind-driven mixing, thereby allowing chemical gradients to persist in the deeper waters, favoring a wider range of species and fewer dominants. A similar trend was observed by Schutte et al. (2016), who reported richer and more even communities at depth in Potentilla Lake in Greenland, which had strongly stratified oxygen conditions similar to 2FB and Bottom lakes. Driven by the internal waves in ice-covered lakes, the resuspension of sediments, including microbes, could also explain the increase in microbial diversity with depth (Bouffard et al., 2016).

Hallmark phyla of freshwater environments were all observed in the Stuckberry Lakes system, including *Actinobacteria*, *Bacteroidetes*, *Cyanobacteria*, *Proteobacteria,* and *Verrucomicrobia* (**Figure 3**). *Actinobacteria* and *Bacteroidetes* maintained relatively constant relative abundances throughout the water columns of all the lakes, averaging 16 and 9% of reads, respectively. These phyla are widespread in freshwater lakes (Zwart et al., 2002; Newton et al., 2011), including High Arctic (Ntougias et al., 2016; Comte et al., 2018; Cavaco et al., 2019) and Antarctic lakes (Picazo et al., 2019).

Top Lake

Cyanobacteria, often major primary producers in Arctic and Antarctic inland waters (Vincent, 2018), were relatively more abundant at the surface of Top Lake than in the other waterbodies (**Figure 3A**). Picocyanobacteria are often the dominant phytoplankton in terms of cell concentrations in cold, oligotrophic lakes, where their abundance can be increased by nutrient supply, (e.g., by mixing and nutrient entrainment in High Arctic Lake A; Veillette et al., 2011). The relatively higher TP yet overall oligotrophic conditions (**Table 1**) combined with high water transparency (Klanten et al., 2021) may have favored their relative abundance in Top Lake. *Chloroflexi* were mostly found at 20 m in Top Lake. This phylum is known to include aerobic phototrophs that do not produce oxygen but use bacteriochlorophyll *a* to perform anoxygenic photosynthesis and photoheterotrophy (Ward et al., 2018). Representatives are commonly found in northern lakes with high light penetration and oligotrophic conditions (Fauteux et al., 2015), and in the microbial mats that often coat the sediments of high latitude waterbodies (Vigneron et al., 2018). Similarly, *Rhodobacterales* were identified as important features in Top Lake (**Figure 3A**). This order also contains aerobic anoxygenic phototrophs (Imhoff et al., 2018) and chemoheterotrophs (Kersters et al., 2006). Both *Rhodobacterales* and *Chloroflexi* may contribute to total primary production and carbon fluxes in this deep oligotrophic lake.

We observed a high relative abundance of *Thaumarchaeota* in the hypolimnion of Top Lake. The fact that this is the only archaeal phylum identified in our dataset is likely because our marker gene approach targeted a bacteria-specific gene (V4 region of 16S). These ammonia-oxidizing archaea prefer aquatic habitats with low dissolved oxygen, pH, nutrients, and light (Erguder et al., 2009; Hatzenpichler, 2012; Auguet and Casamayor, 2013; Hollibaugh, 2017; Juottonen et al., 2020). These conditions occurred only at the bottom of Top Lake (**Figure 1** and **Table 1**), in which pH was 7.59 at 45 m (in 2019) and photosynthetically active radiation (PAR) was ∼0% at 45 m (Klanten et al., 2021). The greater pH (7.94 and 8.15 for 2FB and Bottom lakes, respectively) and lower PAR values under the ice (3.46% at 5 m and 0.15% at 7 m for 2FB and Bottom lakes, respectively) may explain why *Thaumarchaeota* relative abundance was not as high at the bottom of 2FB and Bottom lakes. The importance of this group as ammonia oxidizers has been shown in multiple studies (Erguder et al., 2009; Hatzenpichler, 2012; Stahl and de la Torre, 2012), and their ammonium oxidation potential has been previously recorded across the oxycline of High Arctic meromictic lakes (Pouliot et al., 2009). The results from Top Lake suggest that *Thaumarchaeota* drove nitrogen cycling in this part of the lake. *Diplosphaera*, an abundant feature in Top Lake (**Figure 3A**), has been found to be

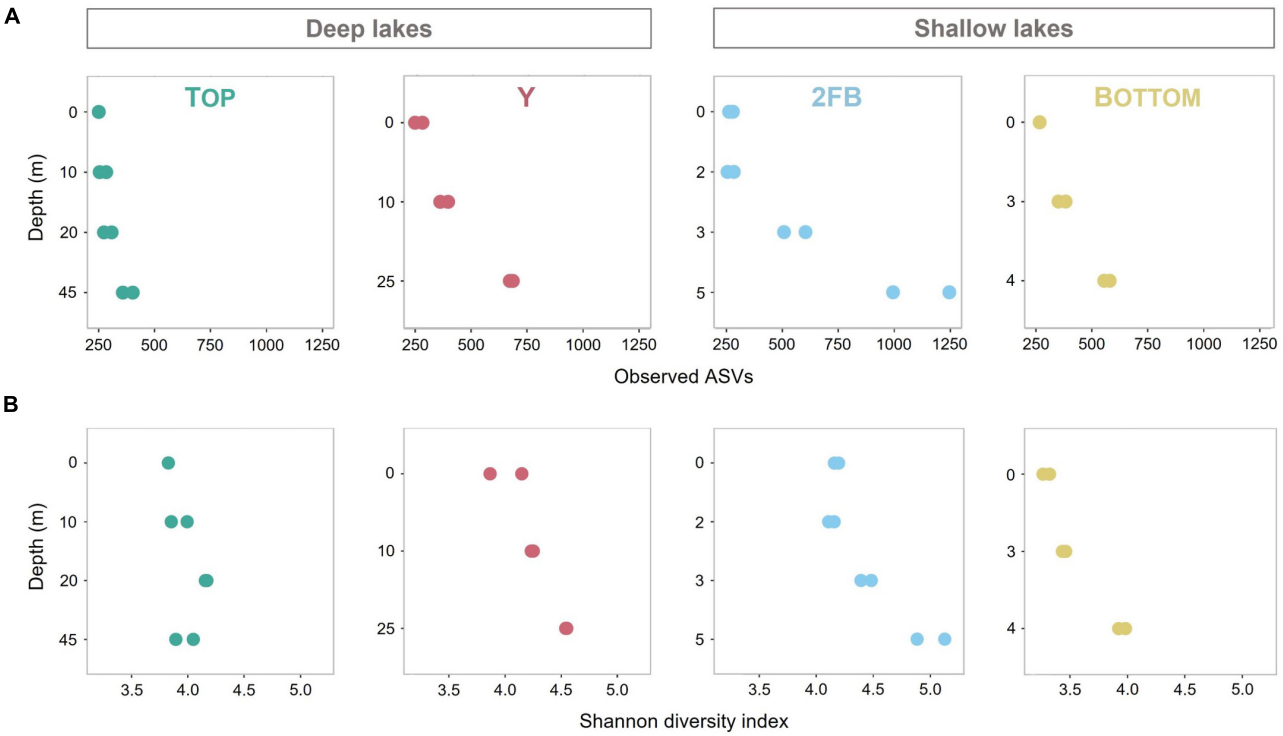

FIGURE 2 | Observed diversity and Shannon diversity in all four Stuckberry Valley lakes. **(A)** Microbial richness calculated with Observed species index. **(B)** Microbial richness and evenness calculated with Shannon diversity index. Colors are associated with each lake: turquoise, Top; pink, Y; blue, 2FB; yellow, Bottom. Note that each sample had two replicates (except 0 m in Top Lake) and the changing graduations of the y-axis established according to the maximum depth of the four lakes.

capable of nitrogen fixation in microaerophilic conditions (Wertz et al., 2012). While our taxonomic approach cannot confirm this, the overrepresentation of this taxa suggests it may contribute to nitrogen cycling in the lake.

Verrucomicrobia and its representatives (e.g., *Opitutales*) were identified as features only in Top Lake (**Figure 3A**). *Verrucomicrobia* is a diverse phylum that includes members that play a role in nutrient and carbon cycling in eutrophic and DOC-rich lakes (He et al., 2017). Their relative abundance here in perennially cold, oligotrophic, DOC-poor Top Lake extends the range of known habitats for these bacteria.

Y Lake

In Y Lake, all the major phyla had relatively constant relative abundances throughout the water column, reflecting the minimal variations in the limnological profile (**Figure 3B**). Here, three features from the *Alphaproteobacteria* were identified: *Caulobacterales, Sphingorhabdus planktonica,* and Sphingomonadales).

2FB Lake

Planctomycetes were present in all four Stuckberry Lakes (**Figure 3**), as expected given the wide freshwater distribution of this group (Zwart et al., 2002; Newton et al., 2011). These heterotrophs have various important ecosystem functions, including contributing to the degradation of complex organic compounds (Schlesner et al., 2004; Fuchsman et al., 2012), and they are mostly found in the hypolimnia of deep oxygenated lakes (Okazaki and Nakano, 2016; Karlov et al., 2017; Okazaki et al., 2017; Storesund et al., 2020). By contrast, 2FB Lake, where *Planctomycetes* reached their greatest relative abundances in our dataset (40%) (**Figure 3C**), was shallow and hypoxic. However, there appeared to be a relationship between *Planctomycetes* communities and DO concentrations in Stuckberry Valley lakes, as proportions were relatively constant throughout the water columns of the deep, oxygenated lakes, and declined markedly below the oxyclines of the two shallow lakes.

Proteobacteria constitute the largest and most phenotypically diverse phylum of bacteria. Members of this taxon are essential contributors to key lacustrine biogeochemical cycles (Kersters et al., 2006). This explains their dominance among the features that are responsible for differences between lakes (**Figure 3**). Some of these bacteria had high relative abundances in the Stuckberry Valley lakes, but mainly in the deep anoxic waters of 2FB Lake (46%) (**Figure 3C**). Key *Proteobacteria* in this shallow lake are discussed below.

In 2FB Lake, as in Bottom Lake, sulfate concentrations decreased with depth (**Table 1**) suggesting the presence of sulfate-reducing bacteria. This is consistent with the pronounced odor of hydrogen sulfide detected in the anoxic deep waters during sampling. *Deltaproteobacteria*, which include sulfate-reducing bacteria (e.g., *Desulfuromonadales)* (Kersters et al., 2006), were present in the hypolimnia of 2FB and Bottom lakes

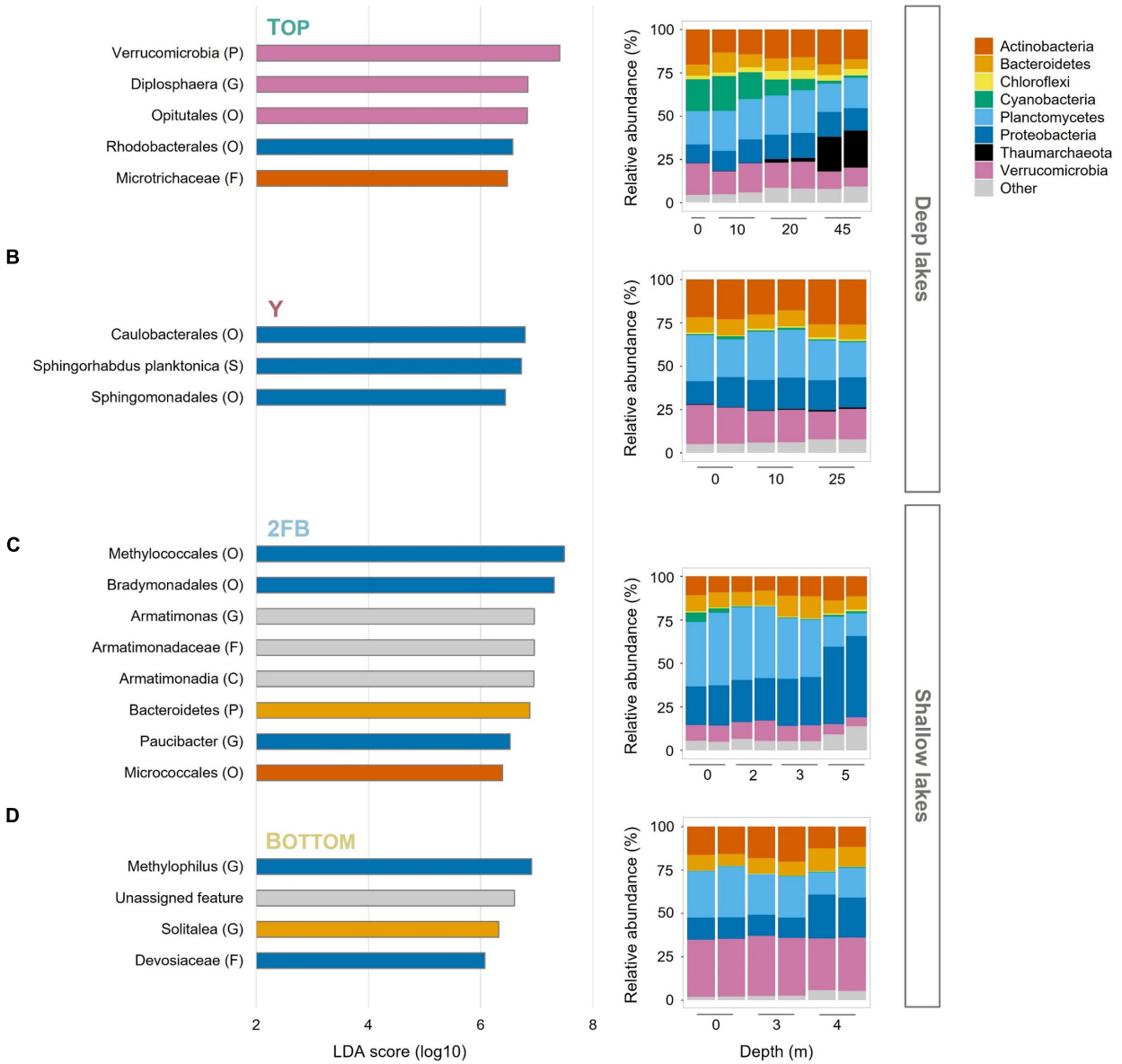

FIGURE 3 | Features and relative abundances of major phyla in the four Stuckberry Valley lakes. **(A)** Top Lake, **(B)** Y Lake, **(C)** 2FB Lake, and **(D)** Bottom Lake. The linear discriminant analysis (LDA) score (log 10) was used to estimate the size effect of each biomarker. The letters in parentheses correspond to the taxonomic rank (P, phyla; C, class; O, order; F, family; G, genus; S, species) and colors correspond to phyla. The color of the lake name is consistent with the map in **Figure 1**. In Bottom Lake, the unassigned feature represents an ASV without any taxonomic assignment. In the relative abundance graphs, phyla that represent less than 5% of reads in at least one sample were grouped together as "Other." Note that each sample had two replicates (except 0 m in Top Lake).

(**Supplementary Figure 2A**). These bacteria have also been shown to dominate the lower water columns in polar lakes with anoxic bottoms, such as Arctic Potentilla Lake (Schutte et al., 2016), and Antarctic Lake Fryxell (Karr et al., 2005).

Gammaproteobacteria (35%) were among the *Proteobacteria* that dominated the bottom of 2FB Lake. These included *Methylococcales,* which use methane and methanol as carbon and energy sources (Kersters et al., 2006; **Figure 3C** and **Supplementary Figure 2B**), and *Chromatiales*, which are photosynthetic purple sulfur bacteria that typically grow under anoxic conditions (Kersters et al., 2006). The high relative abundance of sulfate-reducing (**Supplementary Figure 2A**) and purple sulfur bacteria (*Chromatiales*) (**Supplementary Figure 2B**) suggest that these groups were important contributors to the cycling of sulfur in the anoxic bottom waters of 2FB Lake. The major role of the sulfur cycle in this lake is also

indicated by the association of its microbial communities with the sulfate vector in the ordination analysis (**Figure 4**). Schutte et al. (2016) found a similar distribution in Potentilla Lake, where microbes involved in the sulfur and methane cycles were most abundant below the oxycline, where they likely compete for resources. Despite the presence of methane oxidizers, we did not detect methanogenic archaea. This may reflect limitations of the primer set used, or competitive exclusion of methanogens by sulfate reducers, as observed elsewhere (e.g., Pester et al., 2012). It is also possible that methanogenesis is restricted to the bottom sediments. In High Arctic Ward Hunt Lake, for example, methanogenic Archaea were detected in microbial mats that coated the bottom sediments and experienced anoxic conditions in winter (Mohit et al., 2017).

Viruses and eukaryotic grazers are well-known agents of bacterial mortality in aquatic ecosystems (Fuhrman and Noble, 1995; Ram et al., 2014). However, additional sources of top-down control, such as predatory bacteria, have remained relatively poorly characterized in freshwater lakes. In Stuckberry Valley, we observed high relative abundances of the *Deltaproteobacteria* orders *Bdellovibrionales,* and *Bradymonadales* in 2FB Lake (**Supplementary Figure 2A**), which are known predators of other Gram-negative bacteria (Mu et al., 2020). Interestingly, higher relative abundances of these taxa appeared to correspond with lower counts of heterotrophic bacteria which include their prey (Kersters et al., 2006), and vice versa (**Table 1**). The identification of *Bradymonadales* as a features also supports the biological significance of these taxa as a top-down control in this shallow lake (**Figure 3C**). The opportunistic predators *Myxococcales* (Perez et al., 2016) were also detected in the bottom waters of the shallow lakes (**Supplementary Figure 2A**), and their relative abundances proportionally increased as *Bdellovibrionales* and *Bradymonadales* decreased. Abundant communities of predatory bacteria in lakes have been observed elsewhere, including in Lake Geneva (Paix et al., 2019; Ezzedine et al., 2020). The data from Lake Geneva and from our study suggest that the community composition of prey bacteria (Chen et al., 2011) and the competition between predators (Chase et al., 2002; Johnke et al., 2017) may shape the community structure of these predators and their prey. High Arctic lakes are also known to contain a wide range of phagotrophic protists as well as rotifers (e.g., ten species of rotifers and numerous mixotrophic and heterotrophic protists in ice-covered Ward Hunt Lake; Bégin et al., 2021a,b), which likely exert a top-down grazing control on the bacterial communities. However, eukaryotic grazers are likely to have a minimal grazing impact in anoxic waters (e.g., Oikonomou et al., 2014), where predatory bacteria and viruses could play the greater role.

Bottom Lake

Verrucomicrobia were present in all Stuckberry Valley lakes at all depths, despite the pronounced limnological differences. These bacteria are common in freshwater ecosystems (Zwart et al., 2002; Newton et al., 2011), including in the Arctic (Ntougias et al., 2016; Comte et al., 2018; Cavaco et al., 2019), and are found throughout water columns regardless of nutrient concentrations (Kolmonen et al., 2011). This group includes aerobic, facultative anaerobic, and obligate anaerobic heterotrophs. With their high potential for degrading polysaccharides, they can use multiple carbon sources (He et al., 2017). This metabolic plasticity likely explains why they were so widely distributed in Stuckberry Valley lakes and why they were identified as features in Top Lake (**Figure 3**)—this is likely due to *Diplosphaera* and its order *Opitutales*, which have also been identified as features of this lake. However, we observed the highest relative abundances of *Verrucomicrobia* in Bottom Lake (33%) (**Figure 3D**), where phosphorus and nitrogen concentrations reached their maxima in our dataset (**Table 1**) and where the microbial assemblages were associated with the TP vector (**Figure 4**). This is consistent with the prevalence of this phylum in high-latitude lakes such as Lake Siggeforasjön (Sweden), Lake Vesijärvi (Finland), and Lake Joutikas (Finland) (Kolmonen et al., 2004; Lindstrom et al., 2004; Haukka et al., 2006). With longer ice-free periods due to warming (Magnuson et al., 2000; Vincent, 2018) that will increase organic carbon and nutrient inputs into lakes (Luoto et al., 2019), we hypothesize that generalist taxa such as *Verrucomicrobia* will make up a larger portion of the bacterial flora in Arctic aquatic ecosystems.

As in 2FB Lake, the high abundance of sulfate-reducing bacteria and decrease in the sulfate concentration in Bottom Lake (**Table 1** and **Supplementary Figure 2A**) suggests that sulfur was being actively metabolized at depth. The differentially abundant taxa *Methylophilus* and *Devosiaceae* (**Figure 3D**) may contribute to the carbon (methane) and nitrogen cycles, respectively (Kersters et al., 2006). Although microbial activities can often be inferred from taxonomy, we did not measure transcription or microbial processes directly and note that our inferences are limited to the genetic potential for biogeochemical functions.

Inter-Lake Comparisons

Constrained Analysis of Principal Coordinates (CAP) showed that differences between lakes explained a high proportion of community composition variance (R^2 = 0.90), with non-significant within-lake dispersal (**Figure 4**). A high proportion of total variance (63.9%) was explained by the first axis, and samples were closely grouped on this axis according to lake type (deep vs. shallow). However, while samples from Top and Y lakes overlapped on the second axis (18.8% of variance explained), microbial communities in Bottom and 2FB lakes showed clear separation. This separation appears driven by dissolved oxygen (**Supplementary Figure 3**). In Stuckberry Valley, the outflow stream associated with Top Lake flows into Y Lake, which likely contributes to their similar microbial communities. Studies of microbial connectivity across subarctic and Arctic watersheds have shown that some microbial communities in upstream habitats, such as snow and rivers, were also present in downstream microbiomes, suggesting a landscape-level microbial seeding effect (Ruiz-González et al., 2015; Cavaco et al., 2019). In some locations, up to 30% of phylotypes were shared along the hydrologic continuum (Comte et al., 2018). In contrast to Top and Y lakes, 2FB and Bottom lakes have no hydrological connection (**Figure 1**), and this likely contributes to their marked separation. Lake metabolic processes are depth-dependent, with smaller ecosystems showing greater

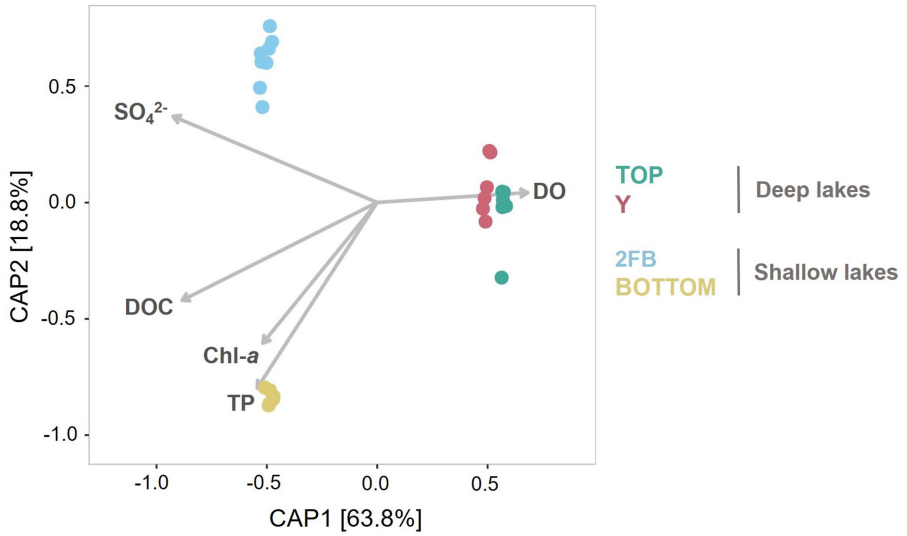

FIGURE 4 | Constrained analysis of principal coordinates of the lake microbial community composition constrained to five environmental variables. Vectors for dissolved oxygen (DO), dissolved organic carbon (DOC), total phosphorus (TP), chlorophyll a (Chl-*a*) and the sulfate concentrations (SO_4^{2-}) illustrate correlations between samples and each of these limnological properties. Colors are associated with each lake: turquoise, Top; pink, Y; blue, 2FB; yellow, Bottom. Note that each sample had two replicates (except 0 m in Top Lake).

variability than larger ecosystems (Staehr et al., 2012), further explaining why the shallow lakes sampled in the present study had fewer similarities between them. In small productive lakes, these depth-dependent changes can be accompanied by sharp vertical gradients in chemical conditions (chemoclines), which increase the diversity of potential bacterial niches, and thus variability, within and between lakes.

Although the Stuckberry Valley lakes could be easily divided into two groups according to their limnological and physical properties (as suggested in **Supplementary Figure 3**), no such clear divisions were evident in their microbial community compositions due to differing relative abundances of major phyla and microbial features (**Figure 3**). Previous studies of lakes with distinct limnological characteristics in larger regions (Lindstrom, 2000; Van der Gucht et al., 2001, 2005; Somers et al., 2020) have reached similar conclusions about the primary role of physicochemical and morphological parameters in determining microbial communities. Those studies identified patterns like those observed in **Figure 4**, where samples clustered by lake type (deep vs. shallow). However, within these broader trends, our study found that each Stuckberry Valley lake also had an individual microbial community structure that was influenced by contrasting environmental conditions within the lakes. This is similar to observations in a study of 67 Finnish lakes, which suggests that individual responses toward environmental factors may occur among the bacterioplankton (Kolmonen et al., 2011).

CONCLUSION

The northern coast of Ellesmere Island is undergoing fundamental changes due to accelerated regional warming. The lakes of Stuckberry Valley in this region are natural laboratories to better understand the impacts of climate change on aquatic ecosystems in the High Arctic. The limnological and microbial stratification we observed in all lakes is likely linked to ice cover which prevents mixing of the water column, and as ice cover decreases in thickness and duration with climate change, vertical stratification is likely to weaken, altering microbial communities at each depth. For the ensemble of Stuckberry Valley lakes, a key driver of microbial community structure appears to be local habitat filtering, i.e., the selection of taxa by unique characteristics of each catchment, and the physical and chemical environment of each lake. Examples include the high relative abundance of *Cyanobacteria* and *Thaumarchaeota* at the surface and at depth, respectively, in Top Lake, the high relative abundance of predatory bacteria in 2FB Lake, and the dominance of *Verrucomicrobia* throughout the Bottom Lake water column. Although more direct measurements of microbial function and more detailed molecular studies are needed to fully elucidate the microbial ecology of this system, these data establish an important baseline characterization for lakes in this remote, extreme, and vulnerable area.

MATERIALS AND METHODS

Study Site

Stuckberry Valley (82° 54′N, 66° 56′W) is located on the northern coast of Ellesmere Island in the Canadian High Arctic (**Figure 1**). The four study sites (unofficially named Top, Y, 2FB, and Bottom lakes) were formed gradually by glacioisostatic rebound following glacial retreat (Bednarski, 1986; Dyke, 2004). Top Lake is the highest above sea level (56 m asl) and Bottom Lake is the

closest to the ocean (31 m asl). About 4 km separates the far end of Top Lake from the opposite end of Bottom Lake. Top Lake, whose catchment contains a small glacier remnant, has an outlet stream that flows into Y Lake, while 2FB and Bottom are isolated from the other lakes and from each other. Y, 2FB, and Bottom lakes have outlets that flow directly into the Arctic Ocean. Based on regional emergence curves, the four lakes were formed approximately 5.6, 5.4, 4.0, and 4.0 ka calibrated years before present (cal BP), respectively (Klanten et al., 2021). The two oldest lakes (Top and Y) are deeper (49 and 28 m, respectively) and mostly oxygenated throughout their water columns. The two younger lakes (Bottom and 2FB) are shallow (< 10 m) and have anoxic bottom waters (Klanten et al., 2021). The deeper lakes are located at the head of the Stuckberry Valley, and presumably have different hydrological and nutrient inputs than the shallower lakes located at the mouth of the valley, closer to the ocean (**Figure 1**).

The nearest weather station is located at Alert (78 km to the east, Ellesmere Island). From 1981 to 2010, the annual average air temperature was −17.7°C (mean minimum in February of −37°C and mean maximum in July of 6.1°C), and annual precipitation averaged 158 mm. From November to May, the average snow cover reached 33 cm (Environment Canada).[1] The polar night of continuous total darkness occurs from October to late February, while the polar summer of continuous daylight lasts from April to August.

Sampling

Samples were collected from May 28th to June 3rd, 2018. Parameters for physicochemical profiles of the water column were measured using an EXO2 Multiparameter Sonde (YSI). Temperature, specific conductivity, chlorophyll *a* fluorescence, and dissolved oxygen profiles are summarily presented in **Figure 1** (complete profiles available in **Supplementary Figure 1**). Photosynthetically active radiation (PAR) was measured using a cosine-corrected underwater quantum sensor (Li-189, LI-CO), as described in Klanten et al. (2021). The full data set for physical and chemical variables is archived in Antoniades et al. (2021).

Three 20-cm-diameter holes (1 m apart) were bored through the snow and ice cover with a manual ice auger at the deepest known point of each lake. The ice on Top Lake was 111 cm thick, while it was 90 cm for the other three lakes. Snow was more variable, with snow covers of 51, 73, 60, and 69 cm depth on Top, Y, 2FB and Bottom lakes, respectively. Water samples were collected in triplicate with a 6.2L-Kemmerer bottle at multiple depths within the water column that were selected according to the physicochemical profiles (**Figure 1** and **Supplementary Figure 1**). Samples were stored in cubitainers that had been previously washed with 2% (vol/vol) Contrad™ 70 liquid detergent (DeconLabs), 10% (vol/vol) ACS-grade HCl (Sigma-Aldrich), and distilled water, and then rinsed three times with water from the lake. Three to four depths were sampled in each lake (0, 10, 20, and 45 m for Top Lake; 0, 10, and 25 m for Y Lake; 0, 2, 3, and 5 m for 2FB Lake; 0, 3, 4, and 7.5 m for Bottom Lake). Cubitainers were kept in the dark during sampling and transportation to the field laboratory.

Within 20 h of sampling, whole water samples were filtered at the field camp on Sterivex™ 0.22 μm capsule filters (Millipore) using a Masterflex® Peristaltic Tubing Pump (Cole-Parmer) and Contrad-washed tubing. Samples were preserved by adding 2 mL RNA*later*™ Solution (Thermo Fisher Scientific) to the filters and then frozen. They were stored subsequently at −80°C once back from the field until extraction.

Water subsamples were mixed with EM grade glutaraldehyde (Canemco) to obtain a final concentration of 1% (vol/vol) for flow cytometry counts, which were analyzed as described in Belzile et al. (2008) and Brussaard et al. (2010). For dissolved organic carbon (DOC), dissolved inorganic carbon (DIC), particulate organic carbon (POC), particulate organic nitrogen (PON), total phosphorus (TP), total nitrogen (TN), major ions and metals analyses, subsamples of water were collected, treated, and stored as described in Klanten et al. (2021). Partial data are listed in **Table 1**, with complete data available in Klanten et al. (2021).

Nucleic Acid Extraction and Processing

After removing RNA*later*™ Solution from Sterivex™ filters, nucleic acids were extracted using the AllPrep DNA/RNA Mini kit (QIAGEN). The manufacturer's protocol was modified as described in Cruaud et al. (2017). DNA concentrations were quantified using the Qubit 3.0 Fluorometer (Thermo Fisher Scientific) as per the manufacturer's protocol. Out of three replicates, the two with the highest DNA concentrations were used for the next steps. Library preparation and sequencing were performed by the Plateforme d'analyse génomique at the Institut de biologie intégrative et des systèmes (IBIS, Université Laval, Québec, Canada). The V4 region of the 16S rRNA gene was amplified by a 2-step PCR using the 515F forward primer (5′-GTGYCAGCMGCCGCGGTAA-3′) (Parada et al., 2016), the 806R reverse primer (5′-GGACTACNVGGGTWTCTAAT-3′) (Apprill et al., 2015), and the Q5 High-fidelity polymerase (NEB). Amplicons were purified on sparQ PureMag beads (QuantaBio) and sequenced on an Illumina MiSeq by paired-end sequencing (2 × 300 pb). The total sequencing yield was 8,447,001 reads (**Supplementary Table 1**).

Sequencing Processing

Sequences were processed using the DADA2 package (v.1.14.0) (Callahan et al., 2016) in R (v.3.5.0) (R Core Team, 2020). Within the package, reads were subjected to quality filtering, trimming, error-rate learning, dereplication, amplicon sequence variant (ASV) inference (Callahan et al., 2017), paired-read merging, chimera removal and taxonomy assignment. The default parameters were used except for the *filterAndTrim()* function with a *truncLen* = (220,195) and *trimLeft* = (19,20). Taxonomy was assigned using the SILVA reference database (v.132) (Quast et al., 2013; Yilmaz et al., 2014). 30 samples were sequenced in total, but due to low counts of initial reads (< 2,000 reads compared to > 240,000 reads for others) (**Supplementary Table 1**), three were discarded (Top_0B, Bot_7.5A and Bot_7.5B). The DNA extracts from the two replicates at 7.5 m from Bottom Lake could not be re-sequenced and

[1] http://climate.weather.gc.ca

compared to the other samples due to the different methods that should have been used to remove PCR inhibitors.

Species richness (α diversity) estimates were calculated and plotted using the *plot_richness()* function with the Observed and Shannon diversity metrics in the package phyloseq in R (v.1.32.0) (McMurdie and Holmes, 2013). The Observed metric considers the numbers of different ASVs (richness) and the Shannon diversity index estimates the ASV richness and evenness (Hill et al., 2003). For both metrics, a *t*-test was used to compare surface and bottom water diversity estimates using the *T.TEST* formula (two-sample assuming equal variances) in Microsoft® Excel® (Microsoft 365; v. 2103) for all lakes separately except for Top Lake. The test could not be performed for Top Lake because there was only a single surface sample. Grouped samples within each lake were compared to others using a Kruskal-Wallis test [stats{}; v.4.0.2] (R Core Team, 2020) followed by a Dunn's test using the FSA package (v.0.8.30) (Ogle et al., 2020), since parametric tests could not be carried out. The assumptions of normality and homoscedasticity were tested with Shapiro-Wilk and Bartlett's tests (stats{}). The data met the assumption of normality but not homoscedasticity.

A bar plot was produced using ggplot2{} (v.3.2.2) (Wickham, 2016). The ASV table was converted to relative abundances and manipulated using tidyr{} (v.1.1.2) (Wickham, 2020), dplyr{} (v.1.0.2) (Wickham et al., 2020) and tibble{} (v.3.0.3) (Müller and Wickham, 2020) in R.

To identify microbiome features that characterized the differences within each lake, we performed linear discriminant analysis (LDA) using LEfSe (LDA Effect Size) (Segata et al., 2011) within the Galaxy web application and workflow framework of the Huttenhower laboratory.[2] The LEfSe analysis comprises three steps. First, it uses the non-parametric factorial Kruskal-Wallis sum-rank test to identify discriminant features (ASVs) with significant differential abundances between the four Stuckberry Valley lakes. Second, to investigate the biological consistency of these features, pairwise tests are performed using the unpaired Wilcoxon rank-sum test. Finally, a LDA estimates the effect size of each differentially abundant feature. The Kruskal-Wallis, Wilcoxon, and LDA test cut-offs were set at 0.05, 0.05, and 2.0, respectively (the default values).

Community-wide diversity (β diversity) was calculated with the Jenson-Shannon divergence (JSD) method (phyloseq{}) using a filtered relative abundance ASV table (ASVs with a mean relative abundance across all samples $< 0.00001\%$ were rejected). Vectors showing limnological characteristics within the water columns were calculated with vegan{} (v.2.5.6) (Oksanen et al., 2019). Metadata collinearity was verified by the use of backward selection and the *vif.cca()* function. Adjusted R^2 values were calculated with the *RsquareAdj()* function. They were visualized along with β diversity in Constrained Analysis of Principal Coordinates (CAP) using Jenson-Shannon divergence with ggplot2{}, as well as through a redundancy analysis with water chemistry parameters using Bray-Curtis dissimilarity. Groups were compared by permutational analysis of variance with *adonis(),* and within-group dispersion homogeneity was verified with *betadisper()*, both functions in the vegan package.

Statistical significance was determined using $\alpha = 0.05$, and only statistically significant results are reported, unless stated otherwise.

[2]https://huttenhower.sph.harvard.edu/galaxy/

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.ncbi.nlm.nih.gov/, PRJNA726255.

AUTHOR CONTRIBUTIONS

CM conceived the study, with input from DA, CG, and AC. CM, YK, and DA conducted the fieldwork and sampling. CM processed the samples and performed the lab work and the data analyses, with assistance from YK and CG. CM wrote the first draft of the manuscript which was edited and revised by CM, DA, CG, AC, and WV. All the authors approved the submitted version.

FUNDING

We thank the Polar Continental Shelf Program (PCSP, Natural Resources Canada) and Parks Canada for logistical and in-kind support. This research was funded by the Natural Sciences and Engineering Research Council of Canada (NSERC), the Fonds de Recherche du Québec—Nature et technologies (FRQNT), the Networks of Centres of Excellence program ArcticNet, and the Sentinel North program at Université Laval (Canada First Research Excellence Fund). CM and YK were supported by the Northern Scientific Training Program. CM was supported by MSc scholarships from NSERC and FRQNT.

ACKNOWLEDGMENTS

We thank Jérôme St-Cyr from the Plateforme d'analyse génomique at IBIS (Université Laval) for library preparation and sequencing, Claude Belzile for flow cytometry, Marianne Potvin for laboratory support, Derek Muir and Xiaowa Wang from the CCIW for chemical and nutrient analyses and Katherine Triglav for help in the field.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2022.779505/full#supplementary-material

REFERENCES

Antoniades, D., Klanten, Y., Lapointe, A.-M., Marois, C., Triglav, K., Muir, D. C. G., et al. (2021). Limnological data from Stuckberry Valley, northern Ellesmere Island, Nunavut, v. 1.2 (2017-2019). *Nordicana* D83. doi: 10.5885/45690CE-5714F25354274D62

Apprill, A., McNally, S., Parsons, R., and Weber, L. (2015). Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. *Aquat. Microb. Ecol.* 75, 129–137. doi: 10.3354/ame01753

Auguet, J. C., and Casamayor, E. O. (2013). Partitioning of Thaumarchaeota populations along environmental gradients in high mountain lakes. *FEMS Microbiol. Ecol.* 84, 154–164. doi: 10.1111/1574-6941.12047

Bednarski, J. (1986). Late Quaternary glacial and sea-level events, Clements Markham Inlet, northern Ellesmere Island, Arctic Canada. *Can. J. Earth Sci.* 23, 1343–1355. doi: 10.1139/e86-129

Bégin, P. N., Rautio, M., Tanabe, Y., Uchida, M., Culley, A. I., and Vincent, W. F. (2021a). The littoral zone of polar lakes: Inshore-offshore contrasts in an ice-covered High Arctic lake. *Arct. Sci.* 6, 158–181. doi: 10.1139/as-2020-0026

Bégin, P. N., Tanabe, Y., Rautio, M., Wauthy, M., Laurion, I., Uchida, M., et al. (2021b). Water column gradients beneath the summer ice of a High Arctic freshwater lake as indicators of sensitivity to climate change. *Sci. Rep.* 11:2868. doi: 10.1038/s41598-021-82234-z

Belzile, C., Brugel, S., Nozais, C., Gratton, Y., and Demers, S. (2008). Variations of the abundance and nucleic acid content of heterotrophic bacteria in Beaufort Shelf waters during winter and spring. *J. Mar. Syst.* 74, 946–956. doi: 10.1016/j.jmarsys.2007.12.010

Bouffard, D., Zdorovennov, R. E., Zdorovennova, G. E., Pasche, N., Wüest, A., and Terzhevik, A. Y. (2016). Ice-covered Lake Onega: Effects of radiation on convection and internal waves. *Hydrobiologia* 780, 21–36. doi: 10.1007/s10750-016-2915-3

Brussaard, C., Payet, J. P., Winter, C., and Weinbauer, M. (2010). "Quantification of aquatic viruses by flow cytometry," in *Manual of aquatic viral ecology*, eds S. W. Wilhelm, M. G. Weinbauer, and C. A. Suttle (Waco, USA: ASLO), 102–109. doi: 10.4319/mave.2010.978-0-9845591-0-7.102

Cai, Q., Wang, J., Beletsky, D., Overland, J., Ikeda, M., and Wan, L. (2021). Accelerated decline of summer Arctic sea ice during 1850–2017 and the amplified Arctic warming during the recent decades. *Environ. Res. Lett.* 16:034015. doi: 10.1088/1748-9326/abdb5f

Callahan, B. J., McMurdie, P. J., and Holmes, S. P. (2017). Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME J.* 11, 2639–2643. doi: 10.1038/ismej.2017.119

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

Cavaco, M. A., St Louis, V. L., Engel, K., St Pierre, K. A., Schiff, S. L., Stibal, M., et al. (2019). Freshwater microbial community diversity in a rapidly changing High Arctic watershed. *FEMS Microbiol. Ecol.* 95:fiz161. doi: 10.1093/femsec/fiz161

Chase, J. M., Abrams, P. A., Grover, J. P., Diehl, S., Chesson, P., Holt, R. D., et al. (2002). The interaction between predation and competition: A review and synthesis. *Ecol. Lett.* 5, 302–315. doi: 10.1046/j.1461-0248.2002.00315.x

Chen, H., Athar, R., Zheng, G. L., and Williams, H. N. (2011). Prey bacteria shape the community structure of their predators. *ISME J.* 5, 1314–1322. doi: 10.1038/ismej.2011.4

Comeau, A. M., Harding, T., Galand, P. E., Vincent, W. F., and Lovejoy, C. (2012). Vertical distribution of microbial communities in a perennially stratified Arctic lake with saline, anoxic bottom waters. *Sci. Rep.* 2:604. doi: 10.1038/srep00604

Comte, J., Culley, A. I., Lovejoy, C., and Vincent, W. F. (2018). Microbial connectivity and sorting in a High Arctic watershed. *ISME J.* 12, 2988–3000. doi: 10.1038/s41396-018-0236-4

Copland, L., and Mueller, D. R. (eds) (2017). *Arctic ice shelves and ice islands*. New York, NY: Springer. doi: 10.1007/978-94-024-1101-0

Cruaud, P., Vigneron, A., Fradette, M. S., Charette, S. J., Rodriguez, M. J., Dorea, C. C., et al. (2017). Open the Sterivex (TM) casing: An easy and effective way to improve DNA extraction yields. *Limnol. Oceanogr. Methods* 15, 1015–1020. doi: 10.1002/lom3.10221

Drevnick, P. E., Muir, D. C. G., Lamborg, C. H., Horgan, M. J., Canfield, D. E., Boyle, J. F., et al. (2010). Increased accumulation of sulfur in lake sediments of the High Arctic. *Environ. Sci. Technol.* 44, 8415–8421. doi: 10.1021/es101991p

Dyke, A. S. (2004). "An outline of North American deglaciation with emphasis on central and northern Canada," in *Quaternary Glaciation extent and chronology: Part II: North America*, eds J. Ehlers and P. L. Gibbard (Amsterdam: Elsevier), 373–424. doi: 10.1016/S1571-0866(04)80209-4

Erguder, T. H., Boon, N., Wittebolle, L., Marzorati, M., and Verstraete, W. (2009). Environmental factors shaping the ecological niches of ammonia-oxidizing archaea. *FEMS Microbiol. Rev.* 33, 855–869. doi: 10.1111/j.1574-6976.2009.00179.x

Ezzedine, J. A., Jacas, L., Desdevises, Y., and Jacquet, S. (2020). *Bdellovibrio* and like organisms in Lake Geneva: An unseen elephant in the room? *Front. Microbiol.* 11:98. doi: 10.3389/fmicb.2020.00098

Falkowski, P. G., Fenchel, T., and Delong, E. F. (2008). The microbial engines that drive Earth's biogeochemical cycles. *Science* 320, 1034–1039. doi: 10.1126/science.1153213

Fauteux, L., Cottrell, M. T., Kirchman, D. L., Borrego, C. M., Garcia-Chaves, M. C., and Del Giorgio, P. A. (2015). Patterns in abundance, cell size and pigment content of aerobic anoxygenic phototrophic bacteria along environmental gradients in northern lakes. *PLoS One* 10:e0124035. doi: 10.1371/journal.pone.0124035

Fuchsman, C. A., Staley, J. T., Oakley, B. B., Kirkpatrick, J. B., and Murray, J. W. (2012). Free-living and aggregate associated Planctomycetes in the Black Sea. *FEMS Microbiol. Ecol.* 80, 402–416. doi: 10.1111/j.1574-6941.2012.01306.x

Fuhrman, J. A., and Noble, R. T. (1995). Viruses and protists cause similar bacterial mortality in coastal seawater. *Limnol. Oceanogr.* 40, 1236–1242. doi: 10.4319/lo.1995.40.7.1236

Hatzenpichler, R. (2012). Diversity, physiology, and niche differentiation of ammonia-oxidizing archaea. *Appl. Environ. Microbiol.* 78, 7501–7510. doi: 10.1128/AEM.01960-12

Haukka, K., Kolmonen, E., Hyder, R., Hietala, J., Vakkilainen, K., Kairesalo, T., et al. (2006). Effect of nutrient loading on bacterioplankton community composition in lake mesocosms. *Microb. Ecol.* 5, 137–146. doi: 10.1007/s00248-005-0049-7

He, S. M., Stevens, S. L. R., Chan, L. K., Bertilsson, S., del Rio, T. G., Tringe, S. G., et al. (2017). Ecophysiology of freshwater Verrucomicrobia inferred from metagenome-assembled genomes. *mSphere* 2, e277–e217. doi: 10.1128/mSphere.00277-17

Hill, T. C. J., Walsh, K. A., Harris, J. A., and Moffett, B. F. (2003). Using ecological diversity measures with bacterial communities. *FEMS Microbiol. Ecol.* 43, 1–11. doi: 10.1111/j.1574-6941.2003.tb01040.x

Hollibaugh, J. T. (2017). Oxygen and the activity and distribution of marine Thaumarchaeota. *Environ. Microbiol. Rep.* 9, 186–188. doi: 10.1111/1758-2229.12534

Imhoff, J. F., Rahn, T., Künzel, S., and Neulinger, S. C. (2018). Photosynthesis is widely distributed among *Proteobacteria* as demonstrated by the phylogeny of PufLM reaction center proteins. *Front. Microbiol.* 8:2679. doi: 10.3389/fmicb.2017.02679

Johnke, J., Baron, M., de Leeuw, M., Kushmaro, A., Jurkevitch, E., Harms, H., et al. (2017). A generalist protist predator enables coexistence in multitrophic predator-prey systems containing a phage and the bacterial predator *Bdellovibrio*. *Front. Ecol. Evol.* 5:124. doi: 10.3389/fevo.2017.00124

Juottonen, H., Fontaine, L., Wurzbacher, C., Drakare, S., Peura, S., and Eiler, A. (2020). Archaea in boreal Swedish lakes are diverse, dominated by Woesearchaeota and follow deterministic community assembly. *Environ. Microbiol.* 22, 3158–3171. doi: 10.1111/1462-2920.15058

Karlov, D. S., Marie, D., Sumbatyan, D. A., Chuvochina, M. S., Kulichevskaya, I. S., Alekhina, I. A., et al. (2017). Microbial communities within the water column of freshwater Lake Radok, East Antarctica: Predominant 16S rDNA phylotypes and bacterial cultures. *Polar Biol.* 40, 823–836. doi: 10.1007/s00300-016-2008-9

Karr, E. A., Sattley, W. M., Rice, M. R., Jung, D. O., Madigan, M. T., and Achenbach, L. A. (2005). Diversity and distribution of sulfate-reducing bacteria in permanently frozen Lake Fryxell, McMurdo Dry Valleys, Antarctica. *Appl. Environ. Microbiol.* 71, 6353–6359. doi: 10.1128/AEM.71.10.6353-6359.2005

Kersters, K., De Vos, P., Gillis, M., Swings, J., Van Damme, P., and Stackebrandt, E. (2006). "Introduction to the *Proteobacteria*," in *The Prokaryotes*, eds M. M. Dworkin, S. Falkow, E. Rosenberg, K.-H. Schleifer, and E. Stackebrandt (New York, NY: Springer), 3–37. doi: 10.1007/0-387-30745-1_1

Klanten, Y., Triglav, K., Marois, C., and Antoniades, D. (2021). Under-ice limnology of coastal valley lakes at the edge of the Arctic Ocean. *Arct. Sci.* 7, 813–831. doi: 10.1139/as-2020-0038

Kolmonen, E., Haukka, K., Rantala-Ylinen, A., Rajaniemi-Wacklin, P., Lepisto, L., and Sivonen, K. (2011). Bacterioplankton community composition in 67 Finnish lakes differs according to trophic status. *Aquat. Microb. Ecol.* 62, 241–250. doi: 10.3354/ame01461

Kolmonen, E., Sivonen, K., Rapala, J., and Haukka, K. (2004). Diversity of cyanobacteria and heterotrophic bacteria in cyanobacterial blooms in Lake Joutikas, Finland. *Aquat. Microb. Ecol.* 36, 201–211. doi: 10.3354/ame036201

Lindstrom, E. S. (2000). Bacterioplankton community composition in five lakes differing in trophic status and humic content. *Microb. Ecol.* 40, 104–113. doi: 10.1007/s002480000036

Lindstrom, E. S., Vrede, K., and Leskinen, E. (2004). Response of a member of the Verrucomicrobia, among the dominating bacteria in a hypolimnion, to increased phosphorus availability. *J. Plankton Res.* 26, 241–246. doi: 10.1093/plankt/fbh010

Lizotte, M. P. (2008). "Phytoplankton and primary production," in *Polar lakes and rivers: Limnology of Arctic and Antarctic aquatic ecosystems*, eds W. F. Vincent and J. Laybourn-Parry (Oxford: Oxford University Press), 157–178.

Luoto, T. P., Rantala, M. V., Kivila, E. H., Nevalainen, L., and Ojala, A. E. K. (2019). Biogeochemical cycling and ecological thresholds in a High Arctic lake (Svalbard). *Aquat. Sci.* 81, 34. doi: 10.1007/s00027-019-0630-7

Magnuson, J. J., Robertson, D. M., Benson, B. J., Wynne, R. H., Livingstone, D. M., Arai, T., et al. (2000). Historical trends in lake and river ice cover in the Northern Hemisphere. *Science* 289, 1743–1746. doi: 10.1126/science.289.5485.1743

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

Meredith, M., Sommerkorn, M., Cassotta, S., Derksen, C., Ekaykin, A., Hollowed, A., et al. (2019). "Polar regions," in *IPCC Special report on the ocean and cryosphere in a changing climate*, eds D. C. Roberts, V. Masson-Delmotte, P. Zhai, M. Tignor, E. Poloczanska, et al. (Geneva: IPCC), 203–320.

Mohit, V., Culley, A., Lovejoy, C., Bouchard, F., and Vincent, W. F. (2017). Hidden biofilms in a far northern lake and implications for the changing Arctic. *NPJ Biofilms Microbiomes* 3:17. doi: 10.1038/s41522-017-0024-3

Moore, G. W. K., Schweiger, A., Zhang, J., and Steele, M. (2019). Spatiotemporal variability of sea ice in the Arctic's Last Ice Area. *Geophys. Res. Lett.* 46, 11237–11243. doi: 10.1029/2019GL083722

Mu, D. S., Wang, S., Liang, Q. Y., Du, Z. Z., Tian, R. M., Ouyang, Y., et al. (2020). Bradymonabacteria, a novel bacterial predator group with versatile survival strategies in saline environments. *Microbiome* 8:126. doi: 10.1186/s40168-020-00902-0

Mueller, D. R., Van Hove, P., Antoniades, D., Jeffries, M. O., and Vincent, W. F. (2009). High Arctic lakes as sentinel ecosystems: Cascading regime shifts in climate, ice cover, and mixing. *Limnol. Oceanogr.* 54, 2371–2385. doi: 10.4319/lo.2009.54.6_part_2.2371

Müller, K., and Wickham, H. (2020). *tibble: Simple data frames. R package version 3.0.3*. Vienna: R Core Team.

Newton, R. J., Jones, S. E., Eiler, A., McMahon, K. D., and Bertilsson, S. (2011). A guide to the natural history of freshwater lake bacteria. *Microbiol. Mol. Biol. Rev.* 75, 14–49. doi: 10.1128/MMBR.00028-10

Ntougias, S., Polkowska, Z., Nikolaki, S., Dionyssopoulou, E., Stathopoulou, P., Doudoumis, V., et al. (2016). Bacterial community structures in freshwater polar environments of Svalbard. *Microbes Environ.* 31, 401–409. doi: 10.1264/jsme2.ME16074

Ogle, D., Wheeler, P., and Dinno, A. (2020). *FSA: Fisheries stock analysis. R package version 0.8.30*. Vienna: R Core Team.

Oikonomou, A., Pachiadaki, M., and Stoeck, T. (2014). Protistan grazing in a meromictic freshwater lake with anoxic bottom water. *FEMS Microbiol. Ecol.* 87, 691–703. doi: 10.1111/1574-6941.12257

Okazaki, Y., and Nakano, S. I. (2016). Vertical partitioning of freshwater bacterioplankton community in a deep mesotrophic lake with a fully oxygenated hypolimnion (Lake Biwa, Japan). *Environ. Microbiol. Rep.* 8, 780–788. doi: 10.1111/1758-2229.12439

Okazaki, Y., Fujinaga, S., Tanaka, A., Kohzu, A., Oyagi, H., and Nakano, S. (2017). Ubiquity and quantitative significance of bacterioplankton lineages inhabiting the oxygenated hypolimnion of deep freshwater lakes. *ISME J.* 11, 2279–2293. doi: 10.1038/ismej.2017.89

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., et al. (2019). *vegan: Community ecology package. R package version 2.5.6*. Vienna: R Core Team.

Paix, B., Ezzedine, J. A., and Jacquet, S. (2019). Diversity, dynamics, and distribution of *Bdellovibrio* and like organisms in perialpine lakes. *Appl. Environ. Microbiol.* 85, e2494–e2418. doi: 10.1128/AEM.02494-18

Parada, A. E., Needham, D. M., and Fuhrman, J. A. (2016). Every base matters: Assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environ. Microbiol.* 18, 1403–1414. doi: 10.1111/1462-2920.13023

Perez, J., Moraleda-Munoz, A., Marcos-Torres, F. J., and Munoz-Dorado, J. (2016). Bacterial predation: 75 years and counting! *Environ. Microbiol.* 18, 766–779. doi: 10.1111/1462-2920.13171

Pester, M., Knorr, K. H., Friedrich, M. W., Wagner, M., and Loy, A. (2012). Sulfate-reducing microorganisms in wetlands–fameless actors in carbon cycling and climate change. *Front. Microbiol.* 3:72. doi: 10.3389/fmicb.2012.00072

Picazo, A., Rochera, C., Villaescusa, J. A., Miralles-Lorenzo, J., Velázquez, D., Quesada, A., et al. (2019). Bacterioplankton community composition along environmental gradients in lakes from Byers peninsula (Maritime Antarctica) as determined by next-generation sequencing. *Front. Microbiol.* 10:908. doi: 10.3389/fmicb.2019.00908

Picazo, A., Villaescusa, J. A., Rochera, C., Miralles-Lorenzo, J., Quesada, A., and Camacho, A. (2021). Functional metabolic diversity of bacterioplankton in maritime Antarctic lakes. *Microorganisms* 9:2077. doi: 10.3390/microorganisms9102077

Pienitz, R., Doran, P. T., and Lamoureux, S. F. (2008). "Origin and geomorphology of lakes in the polar regions," in *Polar lakes and rivers: Limnology of Arctic and Antarctic aquatic ecosystems*, eds W. F. Vincent and J. Laybourn-Parry (Oxford: Oxford University Press), 25–42. doi: 10.1093/acprof:oso/9780199213887.003.0002

Pouliot, J., Galand, P. E., Lovejoy, C., and Vincent, W. F. (2009). Vertical structure of archaeal communities and the distribution of ammonia monooxygenase A gene variants in two high Arctic lakes. *Environ. Microbiol.* 11, 687–699. doi: 10.1111/j.1462-2920.2008.01846.x

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2013). The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res.* 41, 590–596. doi: 10.1093/nar/gks1219

R Core Team (2020). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.

Ram, A. S. P., Palesse, S., Colombet, J., Thouvenot, A., and Sime-Ngando, T. (2014). The relative importance of viral lysis and nanoflagellate grazing for prokaryote mortality in temperate lakes. *Freshwater Biol.* 59, 300–311. doi: 10.1111/fwb.12265

Rochera, C., Quesada, A., Toro, M., Rico, E., and Camacho, A. (2017). Plankton assembly in an ultra-oligotrophic Antarctic lake over the summer transition from the ice-cover to ice-free period: A size spectra approach. *Polar Sci.* 11, 72–82. doi: 10.1016/j.polar.2017.01.001

Rochera, C., Toro, M., Rico, E., Fernández-Valiente, E., Villaescusa, J. A., Picazo, A., et al. (2013). Structure of planktonic microbial communities along a trophic gradient in lakes of Byers Peninsula, South Shetland Islands. *Antarct. Sci.* 25, 277–287. doi: 10.1017/S0954102012000971

Ruiz-González, C., Niño-García, J. P., and Giorgio, P. A. (2015). Terrestrial origin of bacterial communities in complex boreal freshwater networks. *Ecol. Lett.* 18, 1198–1206. doi: 10.1111/ele.12499

Sawstrom, C., Lisle, J., Anesio, A. M., Priscu, J. C., and Laybourn-Parry, J. (2008). Bacteriophage in polar inland waters. *Extremophiles* 12, 167–175. doi: 10.1007/s00792-007-0134-6

Schlesner, H., Rensmann, C., Tindall, B. J., Gade, D., Rabus, R., Pfeiffer, S., et al. (2004). Taxonomic heterogeneity within the Planctomycetales as derived by DNA-DNA hybridization, description of Rhodopirellula baltica gen. nov., sp nov., transfer of *Pirellula marina* to the genus Blastopirellula gen. nov as Blastopirellula marina comb. nov and emended description of the genus Pirellula. *Int. J. Syst. Evol. Microbiol.* 54, 1567–1580. doi: 10.1099/ijs.0.63113-0

Schutte, U. M. E., Cadieux, S. B., Hemmerich, C., Pratt, L. M., and White, J. R. (2016). Unanticipated geochemical and microbial community structure under seasonal ice cover in a dilute, dimictic Arctic lake. *Front. Microbiol.* 7:1035. doi: 10.3389/fmicb.2016.01035

Segata, N., Izard, J., Waldron, L., Gevers, D., Miropolsky, L., Garrett, W. S., et al. (2011). Metagenomic biomarker discovery and explanation. *Genome Biol.* 12:R60. doi: 10.1186/gb-2011-12-6-r60

Shade, A., Jones, S. E., and McMahon, K. D. (2008). The influence of habitat heterogeneity on freshwater bacterial community composition and dynamics. *Environ. Microbiol.* 10, 1057–1067. doi: 10.1111/j.1462-2920.2007.01527.x

Shugar, D. H., Burr, A., Haritashya, U. K., Kargel, J. S., Watson, C. S., Kennedy, M. C., et al. (2020). Rapid worldwide growth of glacial lakes since 1990. *Nat. Clim. Change* 10, 939–945. doi: 10.1038/s41558-020-0855-4

Smith, L. C., Sheng, Y., MacDonald, G. M., and Hinzman, L. D. (2005). Disappearing Arctic lakes. *Science* 308, 1429–1429. doi: 10.1126/science.1108142

Somers, D. J., Strock, K. E., and Saros, J. E. (2020). Environmental controls on microbial diversity in Arctic lakes of West Greenland. *Microb. Ecol.* 80, 60–72. doi: 10.1007/s00248-019-01474-9

Staehr, P. A., Baastrup-Spohr, L., Sand-Jensen, K., and Stedmon, C. (2012). Lake metabolism scales with lake morphometry and catchment conditions. *Aquat. Sci.* 74, 155–169. doi: 10.1007/s00027-011-0207-6

Stahl, D. A., and de la Torre, J. R. (2012). "Physiology and diversity of ammonia-oxidizing archaea," in *Annual review of microbiology*, eds S. Gottesman, C. S. Harwood, and O. Schneewind (Palo Alto, USA: Annual Reviews), 83–101. doi: 10.1146/annurev-micro-092611-150128

Storesund, J. E., Lanzen, A., Nordmann, E. L., Armo, H. R., Lage, O. M., and Ovreas, L. (2020). Planctomycetes as a vital constituent of the microbial communities inhabiting different layers of the meromictic Lake Saelenvannet (Norway). *Microorganisms* 8:1150. doi: 10.3390/microorganisms8081150

Tan, Z. L., and Zhuang, Q. L. (2015). Arctic lakes are continuous methane sources to the atmosphere under warming conditions. *Environ. Res. Lett.* 10, 054016. doi: 10.1088/1748-9326/10/5/054016

Van der Gucht, K., Sabbe, K., De Meester, L., Vloemans, N., Zwart, G., Gillis, M., et al. (2001). Contrasting bacterioplankton community composition and seasonal dynamics in two neighbouring hypertrophic freshwater lakes. *Environ. Microbiol.* 3, 680–690. doi: 10.1046/j.1462-2920.2001.00242.x

Van der Gucht, K., Vandekerckhove, T., Vloemans, N., Cousin, S., Muylaert, K., Sabbe, K., et al. (2005). Characterization of bacterial communities in four freshwater lakes differing in nutrient load and food web structure. *FEMS Microbiol. Ecol.* 53, 205–220. doi: 10.1016/j.femsec.2004.12.006

Veillette, J., Martineau, M.-J., Antoniades, D., Sarrazin, D., and Vincent, W. F. (2011). Effects of loss of perennial lake ice on mixing and phytoplankton dynamics: Insights from High Arctic Canada. *Ann. Glaciol.* 51, 56–70. doi: 10.3189/172756411795931921

Vigneron, A., Cruaud, P., Culley, A. I., Couture, R. M., Lovejoy, C., and Vincent, W. F. (2021). Genomic evidence for sulfur intermediates as new biogeochemical hubs in a model aquatic microbial ecosystem. *Microbiome* 9:46. doi: 10.1186/s40168-021-00999-x

Vigneron, A., Cruaud, P., Mohit, V., Martineau, M.-J., Culley, A. I., Lovejoy, C., et al. (2018). Multiple strategies for light-harvesting, photoprotection and carbon flow in high latitude microbial mats. *Front. Microbiol.* 9:2881. doi: 10.3389/fmicb.2018.02881

Vincent, W. F. (2010). Microbial ecosystem responses to rapid climate change in the Arctic. *ISME J.* 4, 1089–1090. doi: 10.1038/ismej.2010.108

Vincent, W. F. (2018). *Lakes: A very short introduction*. Oxford, UK: Oxford University Press. doi: 10.1093/actrade/9780198766735.001.0001

Vincent, W. F. (2020). "Arctic climate change: Local impacts, global consequences, and policy implications," in *Palgrave handbook of Arctic policy and politics*, eds K. Coates and C. Holroyd (London, UK: Palgrave Macmillan), 507–526. doi: 10.1007/978-3-030-20557-7_31

Vincent, W. F., and Mueller, D. (2020). Witnessing ice habitat collapse in the Arctic. *Science* 370, 1031–1032. doi: 10.1126/science.abe4491

Ward, L. M., Hemp, J., Shih, P. M., McGlynn, S. E., and Fischer, W. W. (2018). Evolution of phototrophy in the Chloroflexi phylum driven by horizontal gene transfer. *Front. Microbiol.* 9:260. doi: 10.3389/fmicb.2018.00260

Wertz, J. T., Kim, E., Breznak, J. A., Schmidt, T. M., and Rodrigues, J. L. M. (2012). Genomic and physiological characterization of the Verrucomicrobia isolate *Diplosphaera colitermitum* gen. nov., sp nov., reveals microaerophily and nitrogen fixation genes. *Appl. Environ. Microbiol.* 78, 1544–1555. doi: 10.1128/AEM.06466-11

White, A., and Copland, L. (2018). Area change of glaciers across Northern Ellesmere Island, Nunavut, between ˜ 1999 and ˜ 2015. *J. Glaciol.* 64, 609–623. doi: 10.1017/jog.2018.49

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. New York, NY: Springer-Verlag. doi: 10.1007/978-3-319-24277-4

Wickham, H. (2020). *tidyr: Tidy messy data. R package version 1.1.2*. New York, NY: Springer-Verlag.

Wickham, H., François, R., Henry, L., and Müller, K. (2020). *dplyr: A grammar of data manipulation. R package version 1.0.2*. Vienna: R Core Team.

Williamson, C. E., Dodds, W., Kratz, T. K., and Palmer, M. A. (2008). Lakes and streams as sentinels of environmental change in terrestrial and atmospheric processes. *Front. Ecol. Environ.* 6, 247–254. doi: 10.1890/070140

Wommack, K. E., and Colwell, R. R. (2000). Virioplankton: Viruses in aquatic ecosystems. *Microbiol. Mol. Biol. Rev.* 64, 69–114. doi: 10.1128/MMBR.64.1.69-114.2000

Wrona, F. J., Johansson, M., Culp, J. M., Jenkins, A., Mård Karlsson, J., Myers-Smith, I. H., et al. (2016). Transitions in Arctic ecosystems: Ecological implications of a changing freshwater system. *J. Geophys. Res.: Biogeosci.* 121, 650–674. doi: 10.1002/2015JG003133

Yadav, A. N., Yadav, N., Kour, D., Kumar, A., Yadav, K., Kumar, A., et al. (2019). "Bacterial community composition in lakes," in *Freshwater microbiology: Perspectives of bacterial dynamics in lake ecosystems*, eds S. A. Bandh, S. Shafi, and N. Shameem (Amsterdam: Elsevier), 1–71. doi: 10.1016/B978-0-12-817495-1.00001-3

Yilmaz, P., Parfrey, L. W., Yarza, P., Gerken, J., Pruesse, E., Quast, C., et al. (2014). The SILVA and "All-species Living Tree Project (LTP)" taxonomic frameworks. *Nucleic Acids Res.* 42, 643–648. doi: 10.1093/nar/gkt1209

Zwart, G., Crump, B. C., Agterveld, M., Hagen, F., and Han, S. K. (2002). Typical freshwater bacteria: An analysis of available 16S rRNA gene sequences from plankton of lakes and rivers. *Aquat. Microb. Ecol.* 28, 141–155. doi: 10.3354/ame028141

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

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 11 February 2022
doi: 10.3389/fmicb.2021.786094

Freshwater Microbial Eukaryotic Core Communities, Open-Water and Under-Ice Specialists in Southern Victoria Island Lakes (Ekaluktutiak, NU, Canada)

*Marianne Potvin[1], Milla Rautio[2,3,4] and Connie Lovejoy[1]**

[1] *Département de Biologie, Québec Océan, and Institut Intégrative et des Systèmes (IBIS), Université Laval, Quebec, QC, Canada,* [2] *Département des Sciences Fondamentales, Université du Québec à Chicoutimi, Saguenay, QC, Canada,* [3] *Groupe de Recherche Interuniversitaire de Limnologie (GRIL), Montreal, QC, Canada,* [4] *Center D'Études Nordiques (CEN), Quebec, QC, Canada*

Across much of the Arctic, lakes and ponds dominate the landscape. Starting in late September, the lakes are covered in ice, with ice persisting well into June or early July. In summer, the lakes are highly productive, supporting waterfowl and fish populations. However, little is known about the diversity and ecology of microscopic life in the lakes that influence biogeochemical cycles and contribute to ecosystem services. Even less is known about the prevalence of species that are characteristic of the seasons or whether some species persist year-round under both ice cover and summer open-water conditions. To begin to address these knowledge gaps, we sampled 10 morphometrically diverse lakes in the region of Ekaluktutiak (Cambridge Bay), on southern Victoria Island (NU, Canada). We focused on Greiner Lake, the lakes connected to it, isolated ponds, and two nearby larger lakes outside the Greiner watershed. The largest lakes sampled were Tahiryuaq (Ferguson Lake) and the nearby Spawning Lake, which support commercial sea-run Arctic char (*Salvelinus alpinus*) fisheries. Samples for nucleic acids were collected from the lakes along with limnological metadata. Microbial eukaryotes were identified with high-throughput amplicon sequencing targeting the V4 region of the 18S rRNA gene. Ciliates, dinoflagellates, chrysophytes, and cryptophytes dominated the lake assemblages. A Bray–Curtis dissimilarity matrix separated communities into under-ice and open-water clusters, with additional separation by superficial lake area. In all, 133 operational taxonomic units (OTUs) occurred either in all under-ice or all open-water samples and were considered "core" microbial species or ecotypes. These were further characterized as seasonal indicators. Ten of the OTUs were characteristic of all lakes and all seasons sampled. Eight of these were cryptophytes, suggesting diverse functional capacity within the lineage. The core open-water indicators were mostly chrysophytes, with a few ciliates and uncharacterized Cercozoa, suggesting that summer communities are mixotrophic with contributions by heterotrophic taxa. The core under-ice indicators included a dozen ciliates along with chrysophytes, cryptomonads, and dinoflagellates, indicating a more heterotrophic community augmented by mixotrophic taxa in winter.

Keywords: Arctic, chrysophytes, cryptophytes, season, ciliates

OPEN ACCESS

Edited by:
Julia Kleinteich, Bundesanstalt für Gewässerkunde (BFG), Germany

Reviewed by:
Veljo Kisand, University of Tartu, Estonia
Sophie Charvet, American Museum of Natural History, United States

***Correspondence:**
Connie Lovejoy
connie.lovejoy@bio.ulaval.ca

Specialty section:
This article was submitted to Extreme Microbiology, a section of the journal Frontiers in Microbiology

Received: *29 September 2021*
Accepted: *15 December 2021*
Published: *11 February 2022*

Citation:
Potvin M, Rautio M and Lovejoy C (2022) Freshwater Microbial Eukaryotic Core Communities, Open-Water and Under-Ice Specialists in Southern Victoria Island Lakes (Ekaluktutiak, NU, Canada). Front. Microbiol. 12:786094. doi: 10.3389/fmicb.2021.786094

INTRODUCTION

The Arctic is home to emblematic endemic species that have evolved in response to seasonal isolation and an extreme environment, where light and temperatures are limiting for much of the year. Across the Arctic, lakes and ponds dominate the landscape. The thousands of lakes of southern Victoria Island, Nunavut, are an integral component of local indigenous communities that depend upon them for valuable ecosystem services, providing potable water, food security, and cultural grounding (White et al., 2007). This cultural and ecological intertwining is reflected in the local name, Ekaluktutiak, for the main hamlet (Cambridge Bay) and the surrounding region of southern Victoria Island, which means good fishing place and refers to the abundance of both lake trout (*Salvelinus namaycush*) and Arctic char (*Salvelinus alpinus*) (Grosbois et al., 2022). The aquatic macrofauna in the lakes (zooplankton and emergent insects) that support the higher food webs are in turn largely dependent on small single-celled microbial eukaryotes (phytoplankton and other protists) that live in the lakes and ponds. These protists are abundant over the short summer, when surface waters are exposed to 24 h of light and reach temperatures into the teens well above freezing (Rautio et al., 2011a). By late September, the lakes are covered in ice and the underlying waters near the freezing point. The lake ice typically persists well into June or early July (Imbeau et al., 2021). However, despite dark and cold conditions, life continues throughout the winter (Hampton et al., 2017; Schneider et al., 2017; Grosbois and Rautio, 2018).

The Arctic is experiencing rapid warming, and the duration and phenology of the ice-free season is becoming less predictable, with the observed rates of climate-driven change in the Arctic exceeding existing climate model scenarios [Arctic Monitoring and Assessment programe [AMAP], 2017; Meredith et al., 2019]. The lakes are becoming increasingly subjected to longer ice-free periods as northern Canada warms at up to 5 times the rate of the global average (Meredith et al., 2019). The ecological function or health and the biodiversity of Arctic freshwaters are likely to be impacted by the changing conditions (Wrona et al., 2006). However, little is known about the microbial species that underpin the lake ecosystems, which contribute disproportionately to Arctic and sub-Arctic biodiversity (Rautio et al., 2011a; Charvet et al., 2012a).

Seasonal comparisons between ice-covered and summer open-water conditions in Arctic freshwaters are rare (Vigneron et al., 2019), impeding basic understanding of ecosystem function and possible vulnerability. The lack of data on the seasonality of microbial eukaryotes extends to Victoria Island, Nunavut, particularly at the level of species or putative ecotypes, which can now be distinguished using molecular techniques. As part of a larger study on the overall connectivity of Ekaluktutiak (Cambridge Bay) lakes, we carried out a survey of 10 diverse lakes with the aim of identifying ubiquitous protist species that would constitute the microbial eukaryote contribution to the core microbiome of the region. The area is accessible from the Canadian High Arctic Research Station (CHARS), which is set to become a monitoring site [Environmental Research Area (ERA), designated the CHARS-ERA] with a mandate to document ongoing change in the Arctic. Knowledge of the makeup of this core community can be used to monitor change and identify major perturbations in the future.

Specifically, we targeted a range of morphometrically diverse lakes in the CHARS-ERA Ekaluktutiak area. This region includes Greiner Lake, which is of particular importance to the local community of Cambridge Bay. For this study, sampling sites were from the same watershed and included lakes connected to Greiner Lake and hydrologically isolated ponds. In addition, two larger lakes outside of the Greiner watershed, Tahiryuaq (Ferguson Lake) and Spawning Lake, were also sampled to increase limnological diversity. These larger lakes are extensively used by sea-run Arctic char (Moore et al., 2016; Harris et al., 2020). We then further compared under-ice microbial eukaryotic plankton with open-water plankton in most of the same lakes, including Greiner Lake.

Water samples for nucleic acids were collected along with limnological metadata with the aim to identify relevant physical and chemical variables associated with total community compositions in the different seasons and lakes. We then focused on commonalities among the lakes by distinguishing a "core" community of microbial eukaryotes in the lakes with a view toward laying the groundwork for future studies. The concept of core in this case was firstly based on universal occurrence in all lakes sampled either under open-water or ice-covered conditions. In addition to presence/absence, the core community needed to meet the other criteria suggested by Shade and Handelsman (2012), which are persistence, connectivity, community composition, and phylogeny. The connectivity was based on geography, with the core taxa being found in lakes that were all within 40 km of Greiner Lake. Persistence was based on these core taxa occurring either under ice or in open waters of Greiner Lake over 3 years. The other criteria of Shade and Handelsman (2012), community composition and phylogeny, were implicit in our approach using molecular identification. To this end, the microbial eukaryotic communities were identified using high-throughput amplicon sequencing targeting the V4 region of the 18S rRNA gene (rDNA) and 18S rRNA (rRNA). The core operational taxonomic units (OTUs with 98% similarity) consisted of OTUs found in all open-water or all under-ice samples, a subset of which were OTUs that were found in all the sampled lakes irrespective of season. We then identified the seasonal specialists and indicators for open-water (summer) and under-ice (winter and spring) conditions.

MATERIALS AND METHODS

Sampling and Laboratory Analysis

The sampling sites were in the CHARS-ERA located on Victoria Island, which is at the western entrance of the Northwest Passage through the Canadian Arctic Archipelago (**Figure 1**). We sampled a total of 10 lakes over multiple years from June 2015 to August 2018. The lakes were classified by size based on surface area. ERA1–ERA4 were defined as small lakes with areas of 0.75 km^2 (ERA4), 0.19 km^2 (ERA1), and 30–40 m^2 (ERA2 and

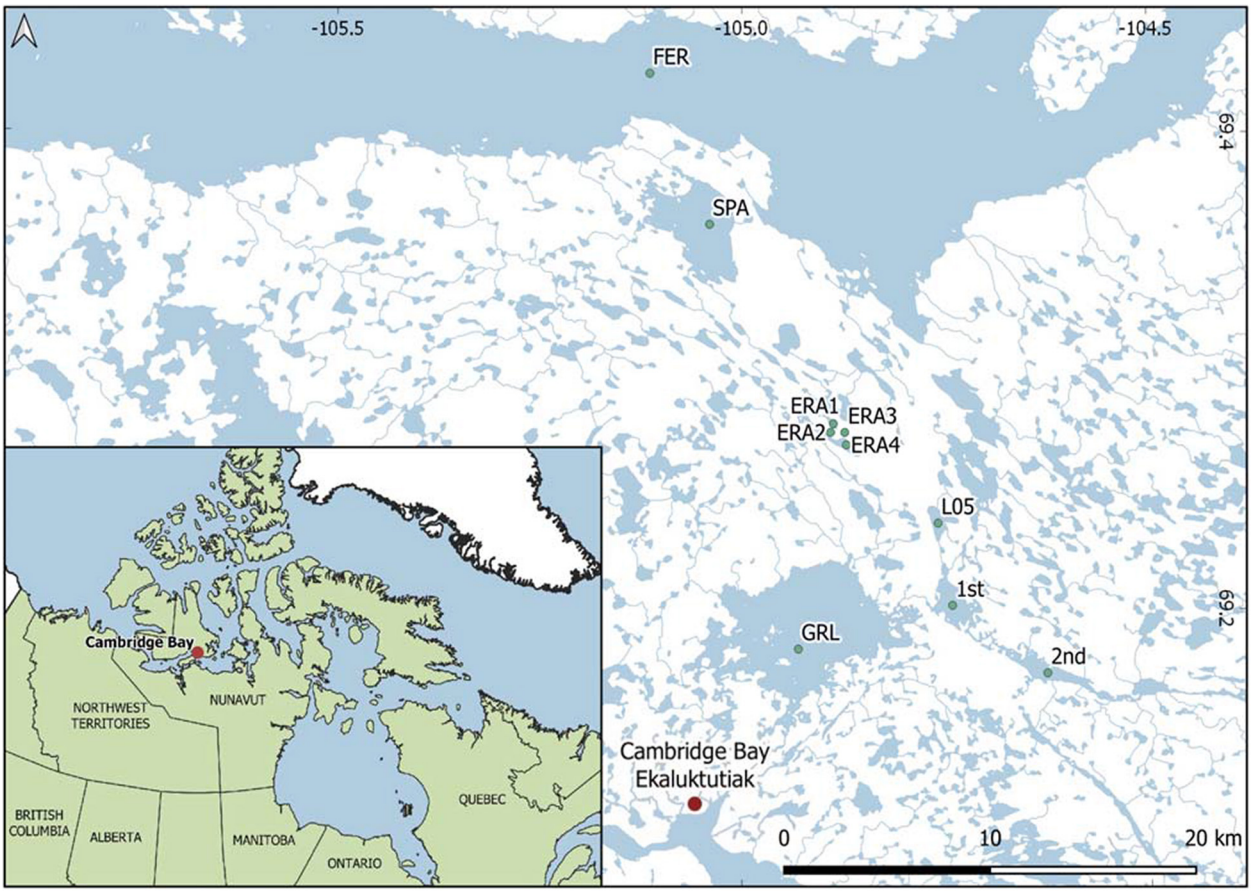

FIGURE 1 | Map of the study area showing the lakes sampled. Lakes with names are Tahiryuaq, previously known as Ferguson (*FER*), Spawning (*SPA*), Greiner (*GRL*), and Inuhuktok (*L05*). Map was created using Q-GIS 3.14 based on Landsat 8 Operational Land Imager (OLI) imagery and *in situ* GPS data.

ERA3). The lakes designated as 1st, 2nd, and L05 (locally known as Inuhuktok) were medium lakes (3.16, 2.68, and 1.11 km^2, respectively). Large lakes were Greiner (36.9 km^2), Spawning (12.5 km^2), and Tahiryuaq (Ferguson Lake, 588 km^2). Eight of the lakes were from the Greiner watershed, while Spawning and Tahiryuaq were nearby, but in a separate watershed. Spawning Lake drains into Tahiryuaq, which is directly linked to the sea *via* the Ekalluk River. The river flows into Wellington Bay, which is contiguous with Dease Strait and the Arctic Ocean. All the lakes in this region are ice-free for about 3 months of the year, between early to mid-July and the end of September. The open-water samples were collected in September (2015 and 2016) and August (2017 and 2018). The under-ice samples were collected in June 2015, October and November 2017, and April 2018. Samples were collected from the center or the deepest portion of the lake based on soundings for the small lakes and available bathymetry for the larger lakes. Lake water was accessed either by drilling a 254 mm hole through the ice in winter with a Jiffy ice augur or by an inflatable boat or float plane during the open-water season. Water for all lakes was collected using a 2-L Limnos water sampler (Limnos Ltd., Komorow, Poland), which was closed just below the surface. Water from multiple casts was collected and mixed in a larger cleaned container that had been rinsed 3 times prior to filling. In 2015, additional nucleic acid samples from 4 and 8 m were collected from Greiner Lake following the same protocol used for surface samples.

Temperature and conductivity were measured using a Ruskin RBR Concerto probe (Ottawa, ON, Canada). Water for ancillary variables was collected from just below the surface in conjunction with nucleic acid sampling. For chlorophyll *a* (Chl *a*), 500 ml was transferred immediately into separate opaque polycarbonate (PC) bottles. Aliquots of sample water for total nitrogen (TN) and phosphorus (TP) were transferred into acid-washed glass bottles. All water samples were placed directly into a cooler and transported by float plane (Spawning and Ferguson Lakes), helicopter, snowmobile, or all-terrain vehicles to the CHARS campus for filtration and laboratory manipulations within 1–3 h of collection. The samples for Chl *a* were filtered onto Whatman GF/F glass fiber filters, wrapped in an aluminum foil, and frozen at −20°C until extraction in 90% ethanol and spectrofluorometric analysis (Nusch, 1980). Samples for dissolved organic carbon (DOC) were filtered through a pre-combusted Whatman GF/F and stored in acid-washed glass

bottles. TN, TP, and DOC samples were analyzed by Environment and Climate Change Canada at the National Laboratory for Environmental Testing (Burlington, ON, Canada) following internal protocols (Environment Canada, 2019).

Filtration and Laboratory Treatment of Nucleic Acids

Up to 2 L of water for nucleic acids was poured through a 50-μm net to remove zooplankton and then filtered successively through 3-μm pore size 47-mm PC membrane filters and 0.2-μm pore size Sterivex filter units (Millipore, Burlington, MA, United States) using a peristaltic pump system. The 47-mm filters were placed in 1.5-ml microfuge tubes filled with RNAlater; 1.8 ml of RNAlater was also added to the Sterivex units. The preserved samples were initially kept at −20°C until they were shipped south within 3 months, then at −80°C until nucleic acid extraction. All collected (47 mm and Sterivex) filters were extracted separately for nucleic acids using AllPrep DNA/RNA Mini Kit (QIAGEN, Hilden, Germany). RNA was converted to cDNA using the Applied Biosystems High-Capacity cDNA Reverse Transcription Kit (Thermo Fisher Scientific, Waltham, MA, United States). The DNA or cDNA (from RNA) was amplified following protocols and using the V4 primers given in Kalenitchenko et al. (2019). Illumina Mi-Seq sequencing was carried at the Université Laval sequencing facility (Plateforme d'analyze génomique, IBIS). Raw sequences are archived in NCBI Sequence Read Archive (SRA) under the Bioproject accession PRJNA623385 and in EMBL European Nucleotide Archive (ENA) under the accession number PRJEB24089.

Bioinformatics Analysis

Since the goal of this study was to identify common OTUs from all sampled lakes in the geographical area, we pooled and analyzed all sequencing results to obtain a single OTU table. These included all reads generated from both DNA and cDNA (RNA) templates and from the separate large (3-μm pore size filter) and small (0.22-μm pore size filter) fractions. The two size fractions were collected and sequenced separately to reduce primer bias as the small fraction enriches for smaller cells with low rRNA gene copy numbers that tend to be swamped by larger cells when not size fractionated. Sequence analysis followed the method of Kalenitchenko et al. (2019), with minor modifications. Briefly, read pairs were merged using BBMerge v37.36 (Bushnell et al., 2017), followed by quality filtering with vsearch (maxEE parameter of 0.5) (Rognes et al., 2016). Unique sequences were selected to decrease the computational need for the chimera checking with USEARCH [unoise3, min size = 4 (zotus)] and OTU clustering in USEARCH (Edgar, 2010). A similarity threshold of 98% was used to define OTUs, and the centroid of each OTU was selected for downstream sequence analysis. We then assigned taxonomy using the Wang method in mothur v1.39 (Schloss, 2020), and the PR^2 database v4.11.0 (Guillou et al., 2013) was used as a reference database. In addition, we performed the same analysis using the Silva v132 database (Quast et al., 2013). An OTU table was then constructed, which contains the number of copies of each unique sequence of OTUs from all separate sample libraries. The final OTU table was then filtered to include only microbial eukaryotes, excluding fungi (essentially chytrids, Cryptomycota and Microsporidiomycota). OTUs matching bacteria, archaea, and multicellular organisms (metazoan and land plants) were also removed. Data from a single geographic location (lake) and unique sampling date were combined to obtain single-site/single-date libraries that were then treated as sample units for downstream analysis (**Supplementary Table 1**).

Phylogenetic Analysis

PR^2 and SILVA taxonomic assignments for individual OTUs concurred for most of the OTUs. When there was disagreement, we carried out individual BLAST searches of the OTUs and assigned them to the closest species-level match, which, in most cases, corresponded to the PR^2 assignment. Because of the high level of ambiguity in the case of cryptophytes and chrysophytes, we carried out separate alignments using Evolutionary Placement Algorithm (Stamatakis, 2014). We first constructed reference trees containing nearly full-length 18S rRNA gene reference sequences for the cryptophytes and chrysophytes mined from GenBank. Sequences were aligned with the multiple sequence alignment program (MAFFT) using default parameters (Katoh and Standley, 2013). The resulting alignments were trimmed to remove gaps at the beginning and the end of the alignments. The representative sequences of the selected OTUs were mapped onto the reference trees in MAFFT-add (–addfragments, Auto) using the evolutionary placement algorithm within RAxML (GTRGAMMA). The sequences of the representative OTUs were placed on the tree at the node that had the highest likelihood result. Trees were visualized and edited using FigTree v1.4.3. For all the microbial eukaryote OTUs, when more than one OTU (98% similarity) was associated with a single species, separate OTUs were considered as distinct genetic populations since several separated into open-water or under-ice distributions, suggesting ecotypes.

Statistical Analysis

Statistical analyses were performed using the "vegan" package in R (R Core Team, 2017). We used Hellinger transformed raw (not rarefied) and the complete OTU table to construct a matrix for the Bray–Curtis dissimilarity analysis. A resulting dendrogram was plotted using Ward clustering. Groups defined by the Bray–Curtis analysis were used to run analysis of similarities (ANOSIM) of the samples and similarity percentage (SIMPER) tests.

To address the environmental context behind whole community clustering, we used the vegan package in R to link environmental data (**Table 1**) with the eukaryote communities assessed by amplicon sequencing. We used the Hellinger-transformed OTU table, raw environmental data, and the Bray–Curtis dissimilarity distance to perform distance-based redundancy analysis (db-RDA) on the complete and on the seasonal subsets of samples (open water and under ice).

TABLE 1 | Sample environmental data.

Sample	Ice (cm)	DOC (mg L^{-1})	TN (μg L^{-1})	TP (μg L^{-1})	Temp. (°C)	Cond (μ S cm^{-1})	Chl *a* (μg L^{-1})	DL (h)	Size	Group
GRL_23AP18	155	5.6	600	11	1.86	608	0.39	17.23	L	Ice1
L05_24AP18	170	3.8	381	8.3	0.79	382	0.27	17.39	M	Ice1
1ST_26AP18	180	5.2	511	13.1	0.71	496	4.20	17.71	M	Ice1
2ND_26AP18	175	5.1	537	15.5	0.54	440	2.46	17.71	M	Ice1
GRL_10JN15	205	5.7	nd	6.1	1.80	675	1.07	24.00	L	Ice1
GRL_15JN15	195	4.9	nd	7.5	nd	nd	0.86	24.00	L	Ice1
GRL_12JN15	200	nd	nd	nd	nd	nd	Nd	24.00	L	Ice1
ERA4_24AP18	167	8.7	736	8.5	1.14	878	0.10	17.39	S	Ice2
L05_02NO17	40	3.7	333	nd	0.56	318	0.54	6.50	M	Ice2
ERA4_03NO17	44	6.6	661	8.4	nd	nd	0.84	6.35	S	Ice2
2ND_01NO17	36	4.4	456	13.6	0.83	269	1.64	6.66	M	Ice2
1ST_31OC17	40	4.5	350	17.9	0.45	317	2.81	6.81	M	Ice2
GRL_30OC17	25	3.9	384	16.8	2.10	325	0.43	6.96	L	Ice2
ERA1_04NO17	42	8.1	887	13.9	2.80	1,049	2.80	6.19	S	Ice2
ERA1_24AP18	170	15.5	1,493	15.3	0.51	2,195	5.52	17.39	S	Ice2
ERA1_10AU17	0	6.7	560	<5.0	10.80	820	1.60	18.60	S	Open3
ERA3_10AU17	0	7	570	<5.0	10.60	770	3.26	18.60	S	Open3
ERA1_12AU18	0	6.9	493	9.2	9.60	524	2.11	18.30	S	Open3
ERA4_10AU17	0	5.9	425	<5.0	11.10	501	1.15	18.60	S	Open3
ERA4_12AU18	0	6	436	5.7	9.10	818	1.16	18.30	S	Open3
1ST_09AU18	0	4.2	331	5.4	10.00	262	2.57	18.82	M	Open3
ERA2_10AU17	0	15.2	1,400	8	10.80	648	1.28	18.6	S	Open3
GRL_04SE16	0	nd	nd	nd	5.40	344	Nd	14.06	L	Open4
SPA_13AU17	0	3.2	271	<5.0	12.40	295	0.85	18.10	L	Open4
GRL_07AU18	0	6.1	832	8.4	9.50	270	3.14	19.17	L	Open4
FER_13AU17	0	2.1	168	<5.0	8.50	245	1.80	18.10	L	Open4
GRL_09SE15	0	3.2	372	5.2	12.40	266	1.29	14.17	L	Open4
GRL_23AU16	0	3.5	427	6.4	10.80	330	Nd	16.49	L	Open4
GRL_09AU17	0	3.7	298	<5.0	10.80	278	1.71	18.77	L	Open4

*Sample name consists of the lake abbreviation from **Figure 1** and the day, month, and year of sampling (JN refers to June).*
When values were under the limit of detection, half of the detection limit value was used for statistics.
*Lake order follows the Bray–Curtis group clustering from **Figure 2**, with (group) Ice1, Ice2, Open3, and Open4 indicated.*
The detection limit for TP was 0.5 μg L^{-1}.
DOC, dissolved organic carbon; TP, total phosphorus; TN, total nitrogen; Ice, ice thickness; Temp, temperature; Cond, conductivity; Chl a, chlorophyll a; DL, day length; Size, lake size category; L, large; M, medium; S, small (as defined in text); nd, no data available.

Defining the Microbial Eukaryote Core Community

The eukaryotic core OTUs comprise those in all "under-ice" samples and those in all "open-water" samples. We also noted a "year-round" category that contained OTUs that were present in all of the samples in the study. Following the definition of prevalence given in Shade and Handelsman (2012), the under-ice category contained OTUs that were present in all the colder season samples (prevalence under ice = 100%). The open-water category (prevalence open water = 100%) contained all the OTUs that were present in all the open-water samples. We note that the open-water and under-ice core OTUs were rarely exclusive to a particular season, with under-ice core OTUs found in the open-water samples and *vice versa*.

Indicator species were investigated using the IndVal command in R (labdsv package). For the IndVal test, the raw OTU table was converted into proportions. As IndVal compares two groups at a time, the comparison groups were based on the two main clusters from the Bray–Curtis dissimilarity analysis. The significance of the indicator OTUs was set at a *p*-value threshold of 0.05 from 10,000 iterations. The same Bray–Curtis sample clusters were compared to identify OTUs as specialists or generalists with the multinomial species classification method (CLAM). This test classifies OTUs as generalists, specialists of a group, or too rare to be defined.

RESULTS

Diversity

Overall, our analysis included 29 one-site/one-date samples from the 10 lakes (**Figure 1**). A total of 2,717,541 reads were retained after quality and taxonomic filtering, with an average of 93,708 reads per sample (**Supplementary Table 1**). The filtered and collapsed final OTU table used for all the community analyses (29 samples) resulted in 1,435 OTUs at 98% similarity. The "top

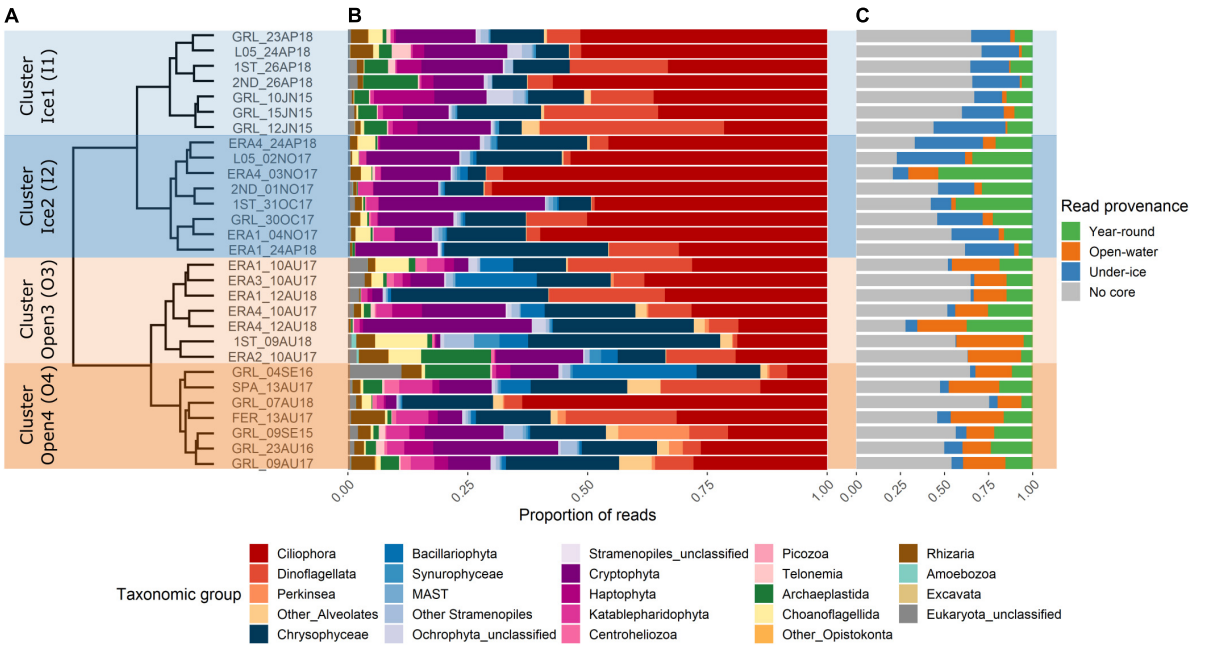

FIGURE 2 | Overall eukaryote diversity in Ekaluktutiak (Cambridge Bay region) revealed with high-throughput sequencing. **(A)** Dendrogram of Bray–Curtis dissimilarity from the Hellinger-transformed operational taxonomic unit (OTU) table containing all protist OTUs. The samples clustered primarily as under ice (Ice1 and Ice2) and open water (Open3 and Open4). **(B)** Relative abundance of reads classified by higher-level taxonomy. **(C)** Relative abundance of reads from the samples within the OTU categories following the initial core definitions. OTUs present in all samples are shown in *green*, OTUs with 100% prevalence in under-ice samples are in *blue*, OTUs with 100% prevalence in open-water samples are in *orange*, and the prevalence of reads from OTUs that were not defined as core is in *gray*.

five" OTUs, defined as those that were most abundant in terms of read counts in each sample, varied with some overlaps among lakes (**Supplementary Figure 1**). The *Fasta* files for these and the other OTUs mentioned in the main text and figures are provided in as a **Supplementary Data File 1**. The five most abundant OTUs in each sample (**Supplementary Figure 1**) accounted for 19.9–56.2% of the reads in their respective samples. Many of the OTUs in the top five were dominant in more than one lake, with a total of 70 distinct top five OTUs at the 98% similarity level. Alveolates accounted for 39 of these, with 29 ciliates, 9 dinoflagellates, and 1 Perkinsida (**Supplementary Figure 1**). The other predominant top five OTUs included chrysophytes and cryptophytes, along with one Thraustochytriaceae, one Katablepharida, and two Chlorophyta (**Supplementary Figure 1**).

At the level of major groups, Ciliophora, Dinoflagellata, Chrysophyceae, and Cryptophyceae were predominant in all samples (**Figures 2A,B**). However, there were major differences at the OTU level, and the Bray–Curtis dissimilarity analysis revealed that samples clustered first by season, with under-ice samples collected in October, November, April, and June separated from open-water samples collected in August and September (**Figure 2**). The relative proportions of reads belonging to the different core categories (**Figure 2C** under ice and open water) followed the clustering pattern of the whole communities. All communities included generalist "all-year" OTUs and non-core OTUs (**Figure 2C**).

The under-ice samples were more generally characterized by higher proportions of dinoflagellate and ciliate reads, with approx. 40–70% of the reads per sample. This under-ice cluster was further separated into two distinct groups (**Figure 2**, left hand labels), designated as Ice1 and Ice2. The Ice1 samples collected in spring (April and June) included samples from Greiner Lake (GRL) and from the medium-sized lakes, with slightly higher proportions of Archaeplastida (green algae) and Haptophyta reads; on the other hand, the Ice2 cluster had samples that were mostly from October and November, but also the April samples from smaller lakes. This Ice2 cluster consisted of slightly higher proportions of choanoflagellates and katablepharids, which are heterotrophic flagellates.

The open-water main cluster from the Bray–Curtis dendrogram contained all the samples that were collected from August and September. These were also separated into two sub-clusters: Open3 and Open4 (**Figure 2**). Open3 included the larger lakes (Greiner, Tahiryuaq, and Spawning), while the Open4 sub-cluster consisted of the small- and intermediate-sized lakes (1st lake and ERA1–ERA4). Open-water samples from L05 and 2nd lake were not available.

The dissimilarity of the four clusters was also significant according to the ANOSIM results. A first statistical test was performed by computing ANOSIM comparing the seasonal clusters (open water *vs.* under ice), which showed significant dissimilarity ($R = 0.7672$, $p = 0.001$). The same statistic was then performed considering the 4 different clusters from the Bray–Curtis cluster analysis, which showed significant dissimilarity among the four sub-clusters ($R = 0.8008$, $p = 0.001$). In summary, dissimilarity of the eukaryote community was

primarily determined by the sampling season (open water *vs*. under ice), and there were also significant differences associated with the lake size classes for both seasons.

Linking Communities to Environmental Data

The db-RDA of all samples and available abiotic environmental data clearly showed the strong seasonality of the region and the response by the eukaryotic communities along the first axis (40.6% of the variance, significant $p = 0.001$) (**Figure 3A**), which was associated with gradients in temperature and ice cover. Greater ice thickness and higher TP concentrations were associated with the under-ice samples, whereas higher temperatures and higher TN/TP ratios (**Figure 3A**) were associated with the open-water samples. The TN/TP ratio was mainly driven by the higher TP concentrations under the ice compared to the open-water samples (**Table 1**). The sub-clusters in **Figure 2**, which were associated with lake size separated along the second axis, were consistent with the Bray–Curtis clustering. Overall, 14.5% of the variance was explained by the second axis, with DOC having the greatest contribution. Spearman's rank correlation showed that the abiotic variables of the first axis—ice cover, temperature, and TP—were co-correlated. On the other hand, the explanatory variables of the second axis—DOC, conductivity, and TN—were co-correlated and negatively correlated with lake area.

To gain further insight, we repeated the analysis on the under-ice and open-water samples separately.

The db-RDA on the under-ice dataset (**Figure 3B**) showed that ice thickness was the most influential environmental variable along the first axis that separated the autumn and most of the spring under-ice samples. Spearman's rank correlations showed a correlation between ice and TP. Conductivity, DOC, and TN were significantly correlated with each other.

The db-RDA carried out using only open-water samples (**Figure 3C**) was not significant. However, as seen in **Figure 2**, the larger lakes tended to separate from the smaller lakes along the first axis. Spearman's rank correlations indicated significant negative correlations between lake area and conductivity, DOC, and TN. Significant positive correlations were seen among the same three variables (conductivity, DOC, and TN) (**Figure 3C**). The significance of individual variables was further examined by applying ANOVA to the open-water dataset, which indicated that conductivity ($p = 0.055$) was associated with lake size during the open-water season.

Core Eukaryote Microbial Taxa

We next identified potential core microbial eukaryotes within the geographic area of Greiner Lake. The OTUs that occurred in all (100%) of the under-ice or all (100%) of the open-water samples were defined as "core" OTUs. Using this definition, 133 putative core OTUs were identified (**Figure 4**). The core OTUs were primarily ciliates and dinoflagellates, but with high proportions of Chrysophyceae and Cryptophyceae. We verified the finer-level taxonomic assignment of these two groups using the Evolutionary Placement Algorithm (EPA). The 46 core Chrysophyceae OTUs were placed at 30 nodes. Globally, where there was a species- or genus-level taxonomic assignation by PR^2, the EPA agreed (**Supplementary Figure 2**). For example, OTUs identified as *Dinobryon* were placed at nodes G, H, and I with verified *Dinobryon* spp. Similarly, OTUs classified as the genus *Kephyrion* were placed at nodes J and K with verified *Kephyrion* 18S rRNA sequences. Since *Ochromonas* is polyphyletic, OTUs that were affiliated to any *Ochromonas* in the NCBI taxonomy were named after the clade in which they appeared, in this case clade C-a (top part of the C clade) and clade C-b (bottom part of the C clade) (**Figure 4** and **Supplementary Figure 2**). Similarly, the EPA placement separated cryptomonad species (**Supplementary Figure 3**) into various clades. The cryptophyte OTUs were placed into three separate clades, with the first clade consisting of an unresolved mix of named species of three genera: *Plagioselmis*, *Teleaulax*, and *Geminigera*. Since our OTUs were placed on a *Plagioselmis nannoplanctica* node, we designated these as *P. nannoplanctica*, which is described as a freshwater cryptophyte with the type species from Sweden (Novarino et al., 1994). The other two clades of cryptophytes were related to *Cryptomonas* species—*Cryptomonas ovata* and *Cryptomonas reflexa*—and a clade referred to in PR^2 as "Basal Cryptophyceae" (**Supplementary Figure 3**). In addition, OTU_700, which occurred in all samples, was a cryptophyte nucleomorph and was significantly correlated with the ensemble of *P. nannoplanctica* OTUs (cor. > 0.7, $p < 0.0001$), suggesting that OTU 700 is the *P. nannoplanctica* nucleomorph.

Seasonal Operational Taxonomic Units

There were 133 putative core OTUs, of which 25 OTUs were found in every sample and were provisionally considered "year-round" (**Figure 4**). After removing putative generalists, 55 "under-ice" and 53 "open-water" OTUs remained (**Figure 5**). The groups were provisionally designated generalists, under-ice, and open-water specialists (**Figure 5**, inner ring).

To provide statistical support for this classification and to identify under-ice "winter" or open-water "summer" season specialists and indicators, two independent tests were carried out. These tests involved analysis of all 1,435 microbial eukaryote OTUs from all samples. For this, we determined seasonal indicator OTUs using IndVal. Since some OTUs were assigned to the same species, we considered the OTUs with 98% similarity level as ecotypes. Out of the total, 582 OTUs were classified as significant ($p \leq 0.05$) season indicators of one condition or the other. Of these, 337 were characteristic of the open-water samples and 245 were associated with under-ice samples (**Figure 6A**). Out of the 133 putative core OTUs, the IndVal tests placed 48 as open-water and 44 as under-ice indicators (**Figure 5**, outer ring). When the initial generalist classification was confirmed by the CLAM test (see below), we marked the OTU as a generalist indicator (**Figure 5**, green slices).

The CLAM test was used to classify the 1,435 OTUs as either specialists for a given season, generalists, or too rare to be classified. From this, 192 were classified as generalists, 421 OTUs were under-ice specialists, 713 OTUs were open-water specialists, and 109 OTUs were too rare to be classified (**Figure 6B**). Examination of the 133 OTUs classified as the

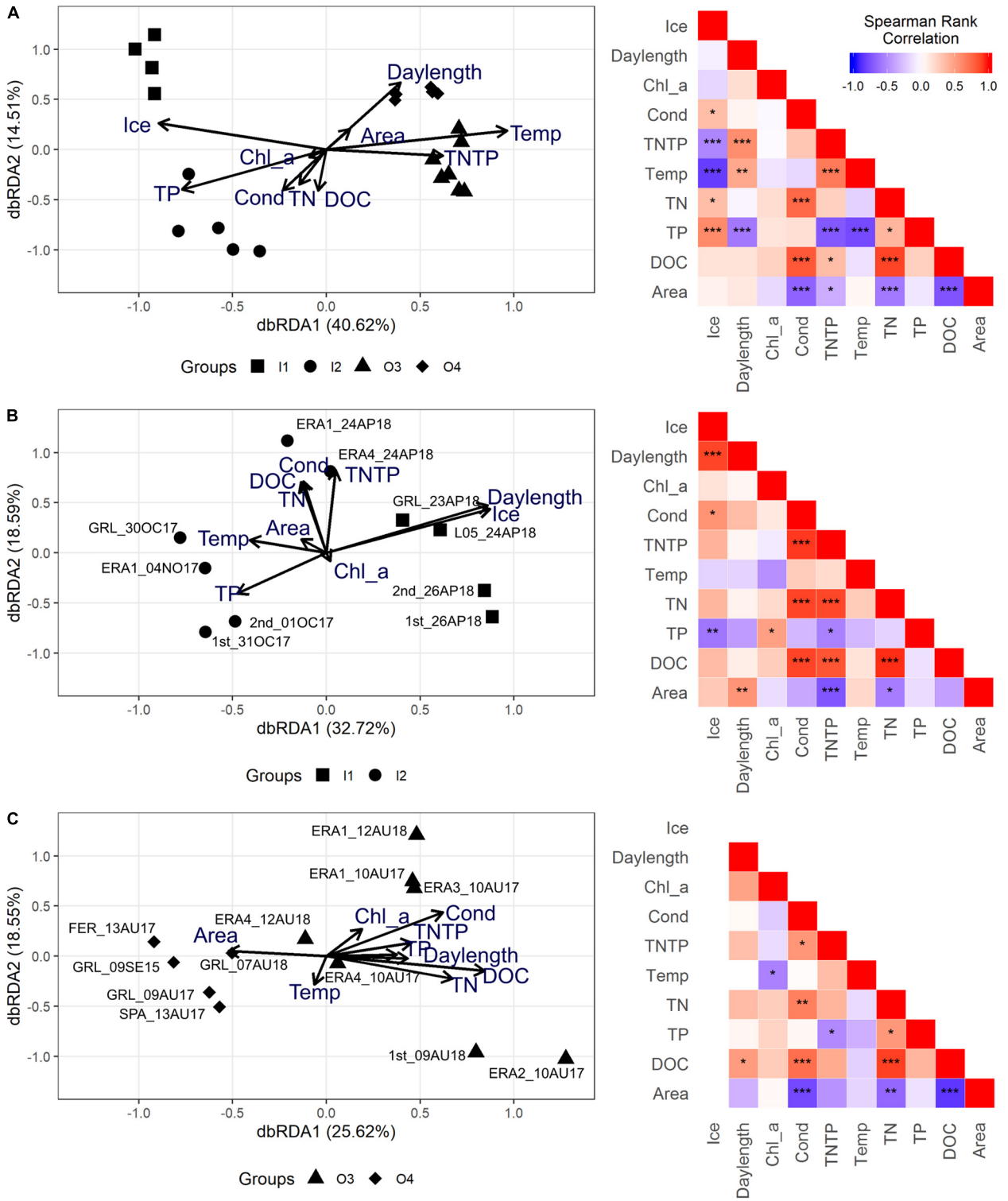

FIGURE 3 | Bray–Curtis distance of the Hellinger-transformed operational taxonomic unit (OTU) table analyzed with different environmental variables. On the *right* are the associated Spearman's rank correlation matrices of the environmental variables of corresponding samples. **(A–C)** Distance-based redundancy analysis (db-RDA) and Spearman's rank correlation on the complete dataset **(A)**, the open-water sample subset **(B)**, and the under-ice sample subset **(C)**. The date of sampling corresponds to the last part of the sample name. Spearman's rank correlation: ***$p < 0.01$, **$p < 0.05$, and *$p < 0.1$.

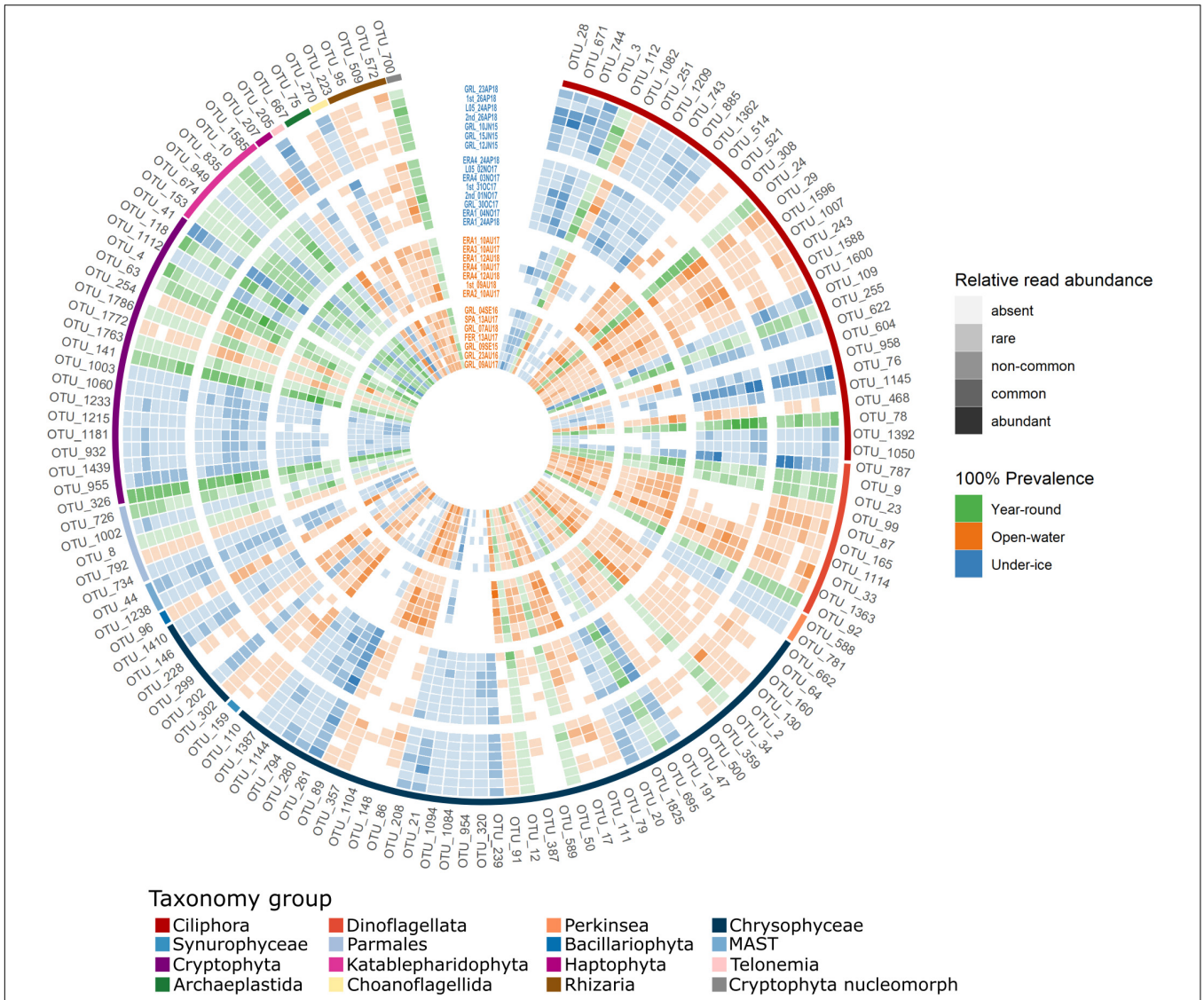

FIGURE 4 | Heatmap of operational taxonomic units (OTUs) with 100% prevalence showing the relative read abundances of the core OTUs within all samples. Samples are arranged following the order of clustering in **Figure 2**, from the exterior to the center. Initial classification of OTUs is by prevalence. Those detected in all samples are shown in *shades of green*, open-water samples are in *shades of orange*, and under-ice samples in *shades of blue*. Intensity of the colors was based on the relative read abundance categories: absent (0), rare (<0.1%), non-common (0.1–1%), common (1–10%), and abundant (>10%). Taxonomy assignment is at the most precise taxonomic rank possible (see text). *Outside ring color* indicates the major taxonomic group of the OTUs.

putative core community revealed that 36 were generalists, 49 were open-water specialists, and 48 were under-ice specialists (**Figure 5**, middle ring).

A majority of the putative core community OTUs (63%) showed consistency in their initial classification as generalists or as open-water or under-ice specialists with the CLAM classification and were also indicator species of a particular season. There were only 12 OTUs (9%) where the IndVal and CLAM classifications were not consistent. CLAM analysis resulted in the displacement of 26 under-ice and open-water OTUs into the generalist category. Conversely, 15 members of the "all seasons" category were classified as specialists of a season (**Figure 5**). Only one of the OTUs (OTU_1007), which was a *Halteria*, switched between open water and under ice in the initial *vs.* the CLAM categories.

We then assessed the overlap of the "core taxa" and the taxa with the highest read proportions (top OTUs). We compared the 133 core OTUs (**Figure 4**) and the 70 top five OTUs (**Supplementary Figure 1**) from all lakes (**Supplementary Figure 4**). Only 27 OTUs were shared by the core and in the top five OTUs. At the major group level, ciliates were the most overrepresented in the top five compared to the core, while chrysophytes were the most overrepresented in the core compared to the top five OTUs. Some top OTUs were associated with the Bray–Curtis clustering according to a SIMPER analysis, with the 10 OTUs that best explained the

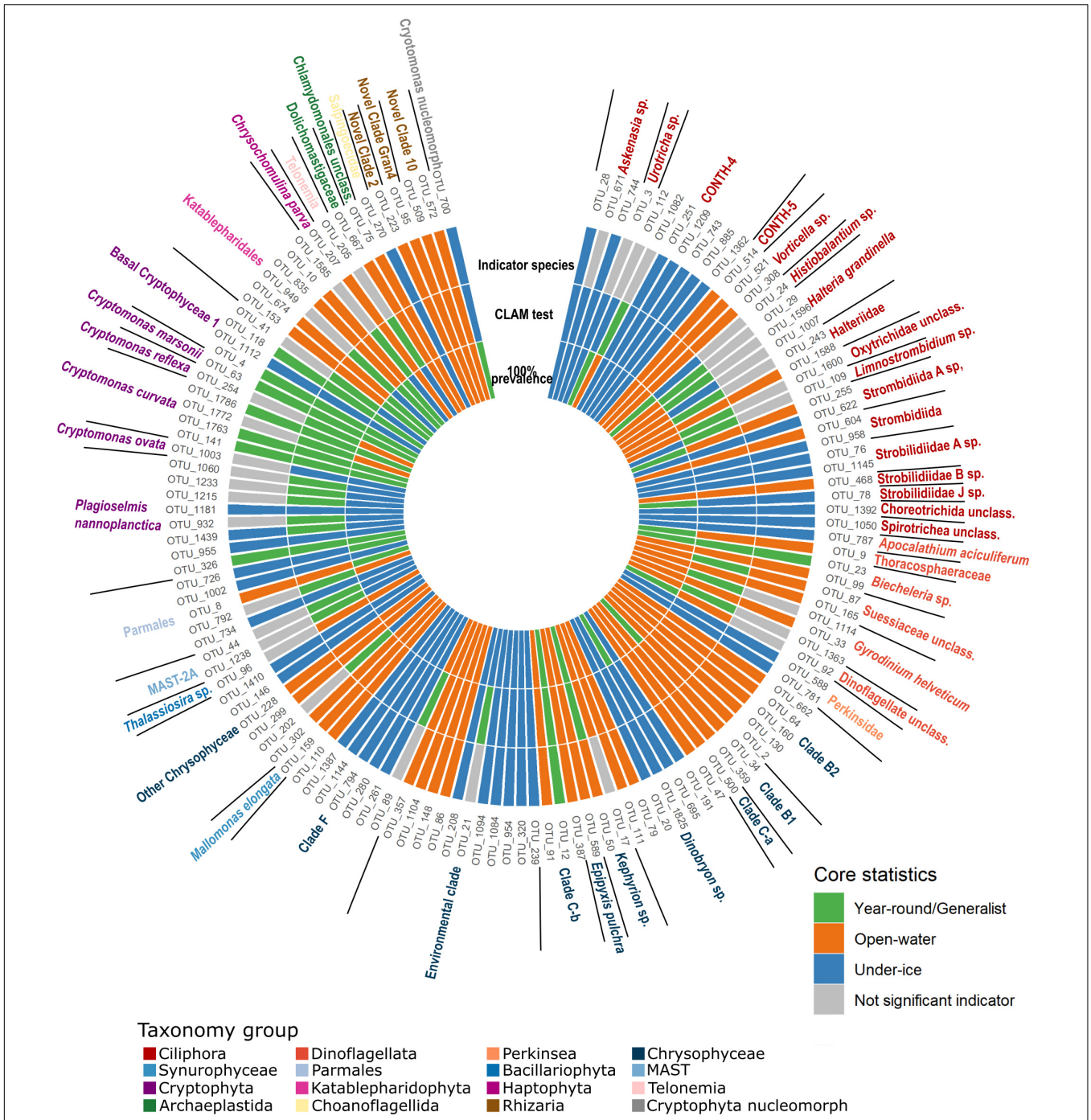

FIGURE 5 | Core species categorizations. *Innermost ring* indicates initial classification from 100% prevalence data in all under ice, all open water, or both (from **Figure 4**). *Middle ring* shows the categorization of operational taxonomic units (OTUs) as under ice or open water or all year using the multinomial species classification method (CLAM test). *Outermost ring* shows the results of IndVal analysis showing the positive indicator OTUs for under-ice, open-water, or all-year conditions. Generalist OTUs, where the initial and CLAM tests agreed, are also marked in *green in the outer ring*. For all three rings, positive results are color coded, with year-round generalists in *green*, open-water OTUs in *orange*, and under-ice OTUs in *blue*. *Gray in the outermost ring* indicates OTUs that were not significant InVal indicators of any of the three conditions and differed between the CLAM and initial classifications.

difference between the open-water and under-ice clusters all among the top five OTUs (**Supplementary Figure 1**). Eight of the top OTUs separating the two clusters were also core species, as defined in our initial classification (**Figures 4**, **5**). Eight were also classified as either open-water or under-ice indicators using IndVal (**Supplementary Table 2**). The SIMPER analysis was then

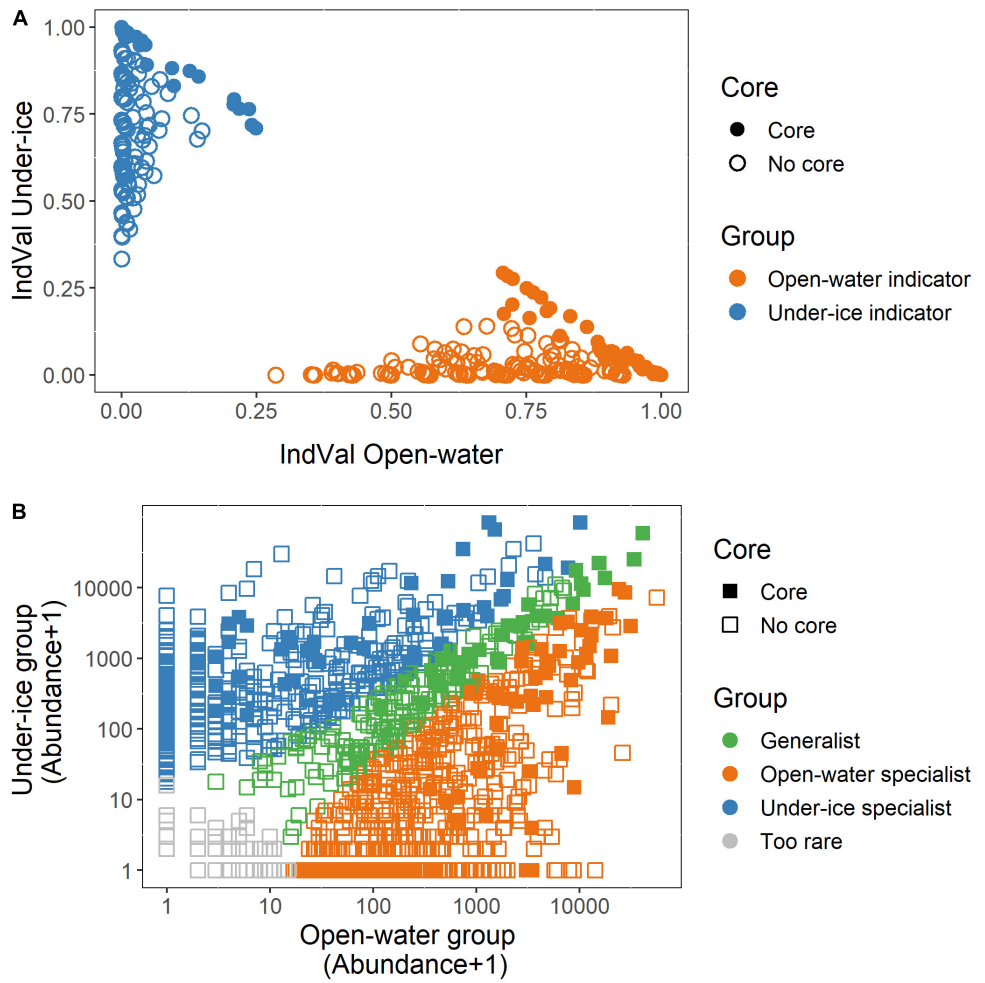

FIGURE 6 | Open-water and under-ice statistical tests and classification. Tests were carried out on the entire dataset. *Filled symbols* are operational taxonomic units (OTUs) that were identified in the core analysis (**Figures 4**, **5**) and *open symbols* are those not in the core. **(A)** Indicator (IndVal) OTUs comparing under ice (*blue symbols*) and open water (*orange symbols*). **(B)** Multinomial species classification method (CLAM) test signifying the open-water (*orange*), under-ice (*blue*), or year-round (*green*) specialists. OTUs that were too rare to classify are shown in *light gray*.

applied to the separate under-ice and open-water sub-clusters. Overall, the 10 best explanatory OTUs from each comparison (under ice *vs.* open water, Ice1 *vs.* Ice2, and Open3 *vs.* Open4) accounted for 21.8, 30, and 21.7% of the difference between the respective compared groups (**Supplementary Table 2**).

DISCUSSION

General Microbial Diversity

Given the direction of human-induced climate change, freshwater biodiversity is expected to be impacted by abiotic disturbance regimes and physical habitat modifications. The changes may disfavor existing resident taxa and favor invasions of taxa previously excluded by Arctic conditions (Wrona et al., 2006, 2016). However, basic knowledge of the core and characteristic taxa is required before short- or long-term impacts from ongoing global warming or changes in the precipitation patterns linked to recent anthropogenic climate perturbations can be assessed. Here, we identified the most common and widespread resident taxa living under ice and in open waters of small and large lakes of the Ekaluktutiak region of southern Victoria Island. Among the 133 OTUs that were found in either all open-water or all under-ice lakes, none were exclusive to a particular season (**Figure 4**), suggesting that the pool of phylotypes that always persist in most lakes may have been underestimated using our strict criteria of 100% prevalence in all lakes on all dates. Despite the ubiquity of many OTUS, the relative abundances of OTUs from the pool differed by lake and day of sampling (**Figure 4**). Targeted in-depth studies of these lakes are needed to fully understand successional patterns.

At the community level, the analysis showed the marked seasonality and the differences between the under-ice and open-water communities. In addition to the strong influence of the number of hours of daylight and temperature driving seasonality, communities were influenced by specific environmental

conditions associated with lake ecology. Species succession is known for phytoplankton and has been reported for zooplankton in other Arctic lakes (Rautio et al., 2011b), but less is known about heterotrophic microbial eukaryotes in winter and summer communities. Annual production largely depends on inorganic nutrient supply in the summer, but in ice-covered lakes, organic matter produced in the summer is required to sustain biological activity during winter (Lizotte, 2009; Hampton et al., 2017; Teufel et al., 2017; de Melo et al., 2019). Dissolved organic matter (DOM) is processed by the winter bacteria that support heterotrophic flagellates, dinoflagellates, and ciliates (Lizotte, 2009; Vigneron et al., 2019), which are food for small zooplankton and which maintain resident fish populations (Grosbois et al., 2020), ensuring a healthy ecosystem.

To reveal the trends that might have been masked by the season, environmental factors associated with the open-water and under-ice communities were analyzed separately. The under-ice communities were separated along the principal axis by day length, ice thickness, and TP (**Figure 3**). The effects of day length and ice thickness were consistent with the succession of microbial eukaryotes over the winter, from freeze-up in October to when ice melts in June. The effect of TP could be associated with both biological and abiotic conditions over winter. TP is a key lake health indicator since lakes are often phosphorus-limited, but excess phosphorus can result in eutrophication and suitable conditions for harmful algal bloom (HAB) events. Recently, a naturally eutrophic lake with high TP levels was reported from near the four ERA lakes sampled here (Ayala-Borda et al., 2021). The high TP load was thought to be due to the phosphorus-rich bedrock in the immediate area combined with high snow accumulation due to local lake morphology and orientation, which resulted in the increased water tracks in early spring maintaining summer phytoplankton production. Here, we found that the TP concentrations were greater under the ice, in agreement with another recent study reporting higher winter inventories of TP (Imbeau et al., 2021). The generally lower TP concentrations in samples from August and September indicate that TP was drawn down over the growing season, and the low Chl *a* levels are in keeping with most of the lakes in the region being more oligotrophic.

In contrast to under-ice conditions, during the open-water season DOC and conductivity had the greatest effect on the communities. Both DOC and conductivity were influenced by lake area, with smaller lakes more affected by near-shore allochthonous DOC (Abnizova et al., 2014) and conductivity, which increases in the summer due to evaporation (Young and Abnizova, 2011; Andresen and Lougheed, 2015). The moderately lower conductivity in Greiner Lake in open- *vs.* under-ice conditions (**Table 1**) would be consistent with salt ejection during ice formation in the winter (Bluteau et al., 2017) and lake flushing during the spring thaw.

The SIMPER analysis indicated that the seasonal and within-season community clustering was associated with OTUs with high relative read abundance, with rare species having less individual weight. In particular, diverse almost lake-specific ciliates were dominant (**Supplementary Figures 1**, **4** and **Supplementary Table 2**). Although the high 18S rRNA gene copy numbers (Blazewicz et al., 2013) in ciliates and the large numbers of ribosomes (from cDNA) in these larger cells could explain some of their high relative read abundance, the diversity of ciliates in the top five OTUs of lakes could be consistent with the ephemeral blooms of individual species. The sporadic occurrences of different ciliate OTUs suggest that ciliate species could be sensitive indicators of local lake conditions, as previously suggested for the Antarctic Dry Valley communities (Xu et al., 2014). Most ciliates likely form cysts (Verni and Rosati, 2011), including Spirotrichea and species such as the *Halteria grandinella* (Foissner et al., 2007) found here. Encystment facilitates the dispersion of species (Verni and Rosati, 2011) and provides inoculum for population recruitment when conditions are suitable for a given species. Once a "bloom" is terminated, the local lake would be a repository for the locally favored ciliate taxa.

All chrysophytes are believed to produce cysts (stomatocysts or statospores) as a response to stresses (Holen, 2014). In the Ekaluktutiak region, chrysophyte OTUs tended to separate the open-water *vs.* under-ice samples, especially in the smaller ERA lakes. Our finding that different species of chrysophytes were seen under different conditions reinforces the use of chrysophytes as indicators of lake status. For example, chrysophyte cysts, which are species-specific, are used in paleolimnological studies as indicators of lake conductivity and area (Pla and Anderson, 2005).

Mixotrophy in Arctic Lakes

The prevalence of mixotrophic taxa over both seasons and all lakes was striking. Mixotrophic microbial eukaryotes range from being primarily heterotrophic and using phototrophy as an energy supplement to being primarily phototrophic and using phagotrophy to supplement growth under low light and darkness (Lie et al., 2017). Mixotrophy is a useful strategy under low light conditions in the High Arctic (Charvet et al., 2014). Mixotrophs can also be classified as facultative or obligate by requiring prey to obtain essential nutrients for growth (Lie et al., 2018). In the core microbiome of Ekaluktutiak, we retrieved multiple mixotrophs from a range of phylogenetic groups and potential diverse behaviors. For example, several chloroplast-containing *Ochromonas* sp. are principally phagotrophic (Andersson et al., 1989; Lie et al., 2018), which may be the case for the core *Ochromonas* related to another Arctic ice-associated *Ochromonas* (CCMP2298). CCMP2298 is probably an obligate phagotroph as it lacks inorganic nitrogen transporters and would only be able to obtain nitrogen from prey (Terrado et al., 2015). Other chrysophytes such as *Dinobryon* can supplement their growth with prey ingestion under low light conditions (Caron et al., 1993), consistent with being found during the ice-covered season. *Uroglena americana* is an example of an obligate phagotroph using prey to obtain growth factors (Sanders, 2011; Kazamia et al., 2012).

The cryptophytes *P. nannoplanctica* and *Cryptomonas curvata*, both of which were found in the under-ice core, ingest bacteria to acquire nutrients (P and N) when they are depleted in the environment (Izaguirre et al., 2012), but also use

phagotrophy to maintain populations under low light regimes (Mitra et al., 2016). Their predominance in the under-ice core suggests the latter as a primary driver in our Arctic lakes. Another way of being mixotrophic is to acquire phototrophy by endosymbiosis and kleptoplasty (Stoecker et al., 2009). *Halteria*, *Askenasia*, *Vorticella*, *Limnostrombidium*, and *Histiobalantium* (**Figure 5**) are all reported to be kleptoplastidic (Stoecker et al., 2009; Esteban et al., 2010). Overall, mixotrophic taxa, whether obligate or optional, appeared to be favored, at least during the times of year we were able to sample. The high zooplankton populations in the lakes at both times of year (Grosbois et al., 2022) would also benefit since, at least in North American boreal lakes, zooplankton show a preference for mixotrophic species (Hansson et al., 2019).

All-Year Generalists

Most of the generalist core taxa were cryptophytes, with one chrysophyte, the two lineages that are typical of northern temperate (Cruaud et al., 2019) and polar freshwater environments (Charvet et al., 2012b). The majority of the generalist core OTUs belonged to mixotrophic–photosynthetic cryptophyte lineages. Their ability to survive under ice and in open water could be aided by a capacity to adjust their pigment profiles for high or low light conditions, as reported for the marine Arctic cryptomonad *Baffinella frigidus* (Daugbjerg et al., 2018). However, this remains to be verified for these freshwater species, and the adaptability to light could be a more general feature of cryptomonads (Greenwold et al., 2019). Two other cryptomonads were in the basal group of Cryptophyceae (**Supplementary Figure 3**). The nearest known relatives of this environmental cryptomonad clade are the non-photosynthetic Goniomonadales suggesting that the OTUs in this study may be purely heterotrophic. The two other year-round core taxa were from probable mixotrophic lineages: clade Cb Chrysophyceae (**Supplementary Figure 3**) and the dinoflagellate *Biecheleria* (Suessiaceae). Freshwater members of *Biecheleria* have been previously classified as *Woloszynskia pseudopalustris* (Moestrup et al., 2009). *Biecheleria* are typical of most photosynthetic woloszynskioid dinoflagellates and most likely mixotrophic, making the majority of the year-round "generalist" species able to photosynthesize when light is available and switch to heterotrophy under low light or darkness.

Under Ice

The under-ice specialists were more trophically diverse compared to the year-round generalist core group. The phylogenetically diverse taxa included phagotrophic mixotrophs, parasites, and strict heterotrophs. Specifically, within the under-ice specialist category were 17 likely mixotrophic stramenopiles, mostly within the Chrysophyceae. Chrysophyceae included mixotrophic genera belonging to *Dinobryon* spp. and OTUs at the nodes of the freshwater genera *Chromophyton* and *Chrysosaccu*s spp. Putative heterotrophic chrysophytes included *Paraphysomonas* spp. and other unclassified Chrysophyceae (**Supplementary Figure 2**), with unknown trophic status. Several Parmales, an order that is synonymous with Bolidophyceae, which are primarily marine but reported from Lake Baikal (Annenkova et al., 2020), were also among the under-ice indicators. The flagellated stages of Parmales are thought to be mixotrophic (Frias-Lopez et al., 2009). Two cryptophyte under-ice OTUs were assigned to *P. nannoplanctica*, which is presumably mixotrophic. Recent taxonomic revisions have suggested that *Plagioselmis* may be part of a di-morphospecies duo related to *Teleaulax* (Altenburger et al., 2020), which have complex life stages and serve as photosynthetic partners in some ciliates, suggesting potential for even more complex trophic roles of some under-ice taxa.

Alveolates that were under-ice specialists included parasitic Perkinsideae and ciliates classified within the well-supported 18S rRNA CONThree (CONTH) clade. The CONTH clade is named after the three major ciliate groups that were initially placed in the clade: Colpodea, Oligohymenophorea, and nassophoreans. CONTH now includes Phyllopharyngea and prosomateans (see UniEuk taxonomy[1]). The PR2 database places the environmental clade CONTH-4 into the CONTH clade. CONTH-4 appears to be related to the parasitic *Cryptocaryon irritans*, which is a Prostomatea that infects fish (Wright and Colorni, 2002; Yambot et al., 2003).

Heterotrophic ciliates under ice were diverse and included *Askenasia*, other Prostomatea, and various Spirotrichea. Also in the under-ice group were several dinoflagellates including a putative mixotrophic dinoflagellate classified as *Apocalathium aciculiferum*. The only other representative of a major group consistently categorized as under ice was a choanoflagellate in the Salpingoecidae family, which are bacterivores. The diverse mixotrophic and heterotrophic microorganisms would be essential for maintaining active overwintering zooplankton populations (Rautio et al., 2011b).

Open Water

Open-water communities included more photosynthetic species, with two green algal OTUs, Chlamydomonadales sp. and Dolichomastigaceae, and a diatom. Other photosynthetic core OTUs were from mixotrophic lineages: *Dinobryon* spp., an undescribed ochromonad, and two other chrysophytes, *Uroglena* and *Mallomonas elongata*, which are potential bloom-forming species. We also found a second *Biecheleria* OTU, other unclassified Suessiaceae, and the mixotrophic haptophyte *Chrysochromulina parva*. Majority of the heterotrophic open-water core OTUs were alveolates; among these were ciliates classified in PR^2 in the group CONTH-5, also distantly related to the fish parasite (Yambot et al., 2003), the genus *Vorticella* and other ciliate grazers from the class Spirotrichea. One putatively heterotroph dinoflagellate, *Gyrodinium helveticum*, was also indicative of open water along with the presumed non-photosynthetic chrysophyte *Paraphysomonas*, Katablepharidales, and Cercozoa.

Seasonal Ecotypes

This analysis highlighted closely related OTUs with different seasonal preferences. Among these were clades that were almost always open-water specialists, such as chrysophyte clades B1 and

[1]https://eukmap.unieuk.org/taxonomy/1163

B2 (**Figure 5**) and others that were almost always associated with under-ice conditions, such as the ciliates classified as *Askenasia*, *Urotricha*, and the CONTH-4 group. Other OTUs that were nominally the same species or genus were found in both the open-water and under-ice core groups, sometimes with a generalist representative, suggesting ecotype diversity and specialization in the same or closely related species. For example, *Dinobryon* spp. included both open-water and under-ice OTUs, and the closely related clade F chrysophytes were predominantly under ice, but with a single more open-water-trending OTU. Other species showed some seasonal preference, but were also found year-round in all lakes. In particular, the cryptophyte *P. nannoplanctica* was diverse with 9 OTUs found in all lakes, but with a tendency to occur in winter. In contrast, the dinoflagellate *Biecheleria* and the chrysophytes in clade C-b were trending toward a summer preference, but with one OTU found year-round (**Figure 5** and **Supplementary Figure 2**) for each. Similar seasonal diversity has been reported in the diatom *Ditylum brightwellii* (Koester et al., 2010) and geographic diversity described for strains of *Prymnesium parvum* (Lysgaard et al., 2018), which suggests that considerable genetic or potentially epigenetic diversity (Ait-Mohamed et al., 2020) within species is widespread. This environmental microdiversity suggests caution when interpreting laboratory-based experiments on single strains of diverse lineages.

Implications for Human Activities

Some OTUs retrieved in the core could potentially have harmful impacts or have impacts on freshwater services. For example, OTUs related to *U. americana*, a colonial flagellate, were retrieved in the open-water samples. This species causes a distinct unpleasant odor in drinking water (Juttner, 1995), and *U. americana* blooms, although not toxic, impacts local human populations. The presence of potential HABs emphasizes a need for monitoring lakes that are used by local communities. Other HABs are harmful to the health of the lake ecosystem. For example, the dinoflagellate *A. aciculiferum*, previously known as *Peridinium aciculiferum* (Craveiro et al., 2017), was reported to form high biomass under the ice in Swedish lakes and can produce a toxin under P limitation. This toxin can be lethal for other phytoplankton and is linked to poor recruitment of larvae of the salmonid whitefish *Coregonus albula* (Rengefors and Legrand, 2001). In this context, the relative abundance of reads classified as *A. aciculiferum* (OTU_787) was high in the under-ice samples collected in June 2015 (GRL_10JU15, GRL_15JU15, and GRL_12JU15) and the end of October 2017 (GRL_30OC17) from Greiner Lake (**Figure 4**), suggesting potential for a bloom. The effect on other phytoplankton or whitefish that live in the lake has not been studied, but would warrant further investigation.

CONCLUSION

We investigated both under-ice and open-water communities from shallow ponds to the 588-km^2 Tahiryuaq, which is an important habitat for sea-run Arctic char and local commercial fisheries. Core microbial taxa were identified, with year-round occurrences of presumably mixotrophic Cryptophyceae, a chrysophyte, and a dinoflagellate. Seasonally defined core taxa included parasites (Perkinsidae and the CONTH ciliates), heterotrophs (other ciliates, dinoflagellates, colorless chrysophytes, katablepharids, Cercozoa, and choanoflagellates) and mixotrophs (cryptophytes, chrysophytes, dinoflagellates, and kleptoplastidic ciliates). The core OTUs were not synonymous with the five most abundant OTUs from each sample. Ciliates made up nearly a quarter of the open-water and under-ice core OTUs, but accounted for over two-thirds of the reads of the top five most abundant OTUs of the individual samples, suggesting the potential utility of monitoring ciliate ecotypes for changes in lake ecology. Our survey of the dominant and most common species or ecotypes presently inhabiting these Arctic water bodies provides context for future studies.

Impact Statement

Microbial eukaryotes support the upper trophic levels providing food to zooplankton, which are food for fish, a primary staple and cultural icon for local Inuvialuit. This study establishes the first overview of the microbial eukaryote freshwater communities on Victoria Island, in small and large lakes that are threatened by ongoing climate warming.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

CL and MR conceived of the study and collected samples. MP and CL wrote the manuscript with input from MR. MP carried out laboratory work, bioinformatic analysis, and drafted figures. MR provided ancillary data and map. All authors contributed to the article and approved the submitted version.

FUNDING

The study was funded by Polar Knowledge Canada, the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery, Northern Supplement and Strategic Projects grants to CL and MR, and ArcticNet, a Canadian Center of Research and Excellence. Additional support was by way of the Fonds de recherche du Québec—Nature et Technologies (FRQNT) to Québec Océan and Center Études Nordiques. The analyses were conducted using Compute Canada facilities.

ACKNOWLEDGMENTS

We would like to thank the logistic support from the Polar Continental Shelf Program (PCSP). We would also like to thank

the Ekaluktutiak Hunters and Trappers Association for support and Johann Wagner, Elise Imbeau, Aili Pedersen, and Jasmine Tiktalek for field assistance. We thank Brian Boyle and the IBIS Platform staff for sharing their knowledge and expertise, Paola Ayala Borda for help with the lab analyses, and Juliette Lapeyre for producing the map. Finally, we acknowledge the Lovejoy Laboratory former and present members and collaborators for their kind input and brainstorming: Marie-Amélie Blais, Perrine Cruaud, Nastasia Freyria, Loïc Jacquemot, Dimitri Kalenitchenko, and Adrien Vigneron.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.786094/full#supplementary-material

REFERENCES

Abnizova, A., Young, K. L., and Lafreniere, M. J. (2014). Pond hydrology and dissolved carbon dynamics at Polar Bear Pass wetland, Bathurst Island, Nunavut, Canada. *Ecohydrology* 7, 73–90. doi: 10.1002/eco.1323

Ait-Mohamed, O., Vanclov, A., Joli, N., Liang, Y., Zhao, X., Genovesio, A., et al. (2020). PhaeoNet: a holistic RNAseq-based portrait of transcriptional coordination in the model diatom *Phaeodactylum tricornutum*. *Front. Plant Sci.* 11:590949. doi: 10.3389/fpls.2020.590949

Altenburger, A., Blossom, H. E., Garcia-Cuetos, L., Jakobsen, H. H., Carstensen, J., Lundholm, N., et al. (2020). Dimorphism in cryptophytes-the case of *Teleaulax amphioxeia/Plagioselmis prolonga* and its ecological implications. *Sci. Adv.* 6:eabb1611. doi: 10.1126/sciadv.abb1611

Andersson, A., Falk, S., Samuelsson, G., and Hagstrom, A. (1989). Nutritional characteristics of a mixotrophic nanoflagellate, *Ochromonas* sp. *Microb. Ecol.* 17, 251–262. doi: 10.1007/bf02012838

Andresen, C. G., and Lougheed, V. L. (2015). Disappearing Arctic tundra ponds: fine-scale analysis of surface hydrology in drained thaw lake basins over a 65 year period (1948-2013). *J. Geophys. Res. Biogeosci.* 120, 466–479. doi: 10.1002/2014jg002778

Annenkova, N. V., Giner, C. R., and Logares, R. (2020). Tracing the origin of planktonic protists in an ancient lake. *Microorganisms* 8:19. doi: 10.3390/microorganisms8040543

Arctic Monitoring and Assessment programe [AMAP] (2017). *Snow, Water, Ice, Permafrost in the Arctic (SWIPA)*. Norway: Arctic Monitoring and Assessment programe (AMAP), 269.

Ayala-Borda, P., Lovejoy, C., Power, M., and Rautio, M. (2021). Evidence of eutrophication in Arctic Lakes. *Arct. Sci.* 7, 859–871. doi: 10.1139/AS-2020-0033

Blazewicz, S. J., Barnard, R. L., Daly, R. A., and Firestone, M. K. (2013). Evaluating rRNA as an indicator of microbial activity in environmental communities: limitations and uses. *ISME J.* 7, 2061–2068. doi: 10.1038/ismej.2013.102

Bluteau, C. E., Pieters, R., and Lawrence, G. A. (2017). The effects of salt exclusion during ice formation on circulation in lakes. *Environ. Fluid Mech.* 17, 579–590. doi: 10.1007/s10652-016-9508-6

Bushnell, B., Rood, J., and Singer, E. (2017). BBMerge - Accurate paired shotgun read merging via overlap. *PLoS One* 12:e0185056. doi: 10.1371/journal.pone.0185056

Caron, D. A., Sanders, R. W., Lim, E. L., Marrase, C., Amaral, L. A., Whitney, S., et al. (1993). Light-dependent phagotrophy in the fresh-water mixotrophic chrysophyte *Dinobryon cylindricum*. *Microb. Ecol.* 25, 93–111. doi: 10.1007/BF00182132

Charvet, S., Vincent, W. F., Comeau, A., and Lovejoy, C. (2012a). Pyrosequencing analysis of the protist communities in a High Arctic meromictic lake: DNA preservation and change. *Front. Microbiol.* 3:422. doi: 10.3389/fmicb.2012.00422

Charvet, S., Vincent, W. F., and Lovejoy, C. (2012b). Chrysophytes and other protists in High Arctic lakes: molecular gene surveys, pigment signatures and microscopy. *Polar Biol.* 35, 733–748. doi: 10.1007/s00300-011-1118-7

Charvet, S., Vincent, W. F., and Lovejoy, C. (2014). Effects of light and prey availability on Arctic freshwater protist communities examined by high-throughput DNA and RNA sequencing. *FEMS Microbiol. Ecol.* 88, 550–564. doi: 10.1111/1574-6941.12324

Craveiro, S. C., Daugbjerg, N., Moestrup, O., and Calado, A. J. (2017). Studies on *Peridinium aciculiferum* and *Peridinium malmogiense* (=*Scrippsiella hangoei*): comparison with *Chimonodinium lomnickii* and description of *Apocalathium* gen. nov. (Dinophyceae). *Phycologia* 56, 21–35. doi: 10.2216/16-20.1

Cruaud, P., Vigneron, A., Fradette, M. S., Dorea, C. C., Culley, A. I., Rodriguez, M. J., et al. (2019). Annual protist community dynamics in a freshwater ecosystem undergoing contrasted climatic conditions: the Saint-Charles River (Canada). *Front. Microbiol.* 10:2359. doi: 10.3389/fmicb.2019.02359

Daugbjerg, N., Norlin, A., and Lovejoy, C. (2018). *Bafinella frigidus* gen et sp nov (Baffinellaceae fam. nov., Cryptophyceae) from Baffin Bay: morpholgy, pigment profile, phylogeny and growth rate response to three abiotic factors. *J. Phycol.* 54, 665–680. doi: 10.1111/jpy.12766

de Melo, T. X., Dias, J. D., Simoes, N. R., and Bonecker, C. C. (2019). Effects of nutrient enrichment on primary and secondary productivity in a subtropical floodplain system: an experimental approach. *Hydrobiologia* 827, 171–181. doi: 10.1007/s10750-018-3763-0

Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26, 2460–2461. doi: 10.1093/bioinformatics/btq461

Environment Canada (2019). *NLET Method Descriptions. Version 1*. Burlington: Environmental Science and Technology Laboratories.

Esteban, G. F., Fenchel, T., and Finlay, B. J. (2010). Mixotrophy in ciliates. *Protist* 161, 621–641. doi: 10.1016/j.protis.2010.08.002

Foissner, W., Muller, H., and Agatha, S. (2007). A comparative fine structural and phylogenetic analysis of resting cysts in oligotrich and hypotrich Spirotrichea (Ciliophora). *Eur. J. Protistol.* 43, 295–314. doi: 10.1016/j.ejop.2007.06.001

Frias-Lopez, J., Thompson, A., Waldbauer, J., and Chisholm, S. W. (2009). Use of stable isotope-labelled cells to identify active grazers of picocyanobacteria in ocean surface waters. *Environ. Microbiol.* 11, 512–525. doi: 10.1111/j.1462-2920.2008.01793.x

Greenwold, M. J., Cunningham, B. R., Lachenmyer, E. M., Pullman, J. M., Richardson, T. L., and Dudycha, J. L. (2019). Diversification of light capture ability was accompanied by the evolution of phycobiliproteins in crptophyte algae. *Proc. R. Soc. B Biol. Sci.* 286:20190655. doi: 10.1098/rspb.2019.0655

Grosbois, G., Power, M., Evans, M., Koehler, G., and Rautio, M. (2022). Content, composition and transfer of polyunsaturated fatty acids in an Arctic lake food web. *Ecosphere* 13:e03881. doi: 10.1002/ecs2.3881

Grosbois, G., and Rautio, M. (2018). Active and colorful life under lake ice. *Ecology* 99, 752–754. doi: 10.1002/ecy.2074

Grosbois, G., Vachon, D., del Giorgio, P. A., and Rautio, M. (2020). Efficiency of crustacean zooplankton in transferring allochthonous carbon in a boreal lake. *Ecology* 101:15. doi: 10.1002/ecy.3013

Guillou, L., Bachar, D., Audic, S., Bass, D., Berney, C., Bittner, L., et al. (2013). The Protist Ribosomal Reference database (PR2): a catalog of unicellular eukaryote Small Sub-Unit rRNA sequences with curated taxonomy. *Nucleic Acids Res.* 41, D597–D604. doi: 10.1093/nar/gks1160

Hampton, S. E., Galloway, A. W. E., Powers, S. M., Ozersky, T., Woo, K. H., Batt, R. D., et al. (2017). Ecology under lake ice. *Ecol. Lett.* 20, 98–111. doi: 10.1111/ele.12699

Hansson, T. H., Grossart, H. P., del Giorgio, P. A., St-Gelais, N. F., and Beisner, B. E. (2019). Environmental drivers of mixotrophs in boreal lakes. *Limnol. Oceanogr.* 64, 1688–1705. doi: 10.1002/lno.11144

Harris, L., Yurkowski, D., Gilbert, M., Else, B., Duke, P., Ahmed, M., et al. (2020). Depth and temperature preference of anadromous Arctic Char *Salvelinus alpinus* in the Kitikmeot Sea, a shallow and low-salinity area of the Canadian Arctic. *Mar. Ecol. Prog. Ser.* 634, 175–197. doi: 10.3354/meps13195

Holen, D. A. (2014). Chrysophyte stomatocyst production in laboratory culture and descriptions of seven cyst morphotypes. *Phycologia* 53, 426–432. doi: 10.2216/14-001.1

Imbeau, E., Vincent, W. F., Wauthy, M., Cusson, M., and Rautio, M. (2021). Hidden stores of organic matter in northern lake ice: selective retention of terrestrial particles, phytoplankton and labile carbon. *J. Geophys. Res. Biogeosci.* 126:e2020JG006233. doi: 10.1029/2020jg006233

Izaguirre, I., Sinistro, R., Schiaffino, M. R., Sanchez, M. L., Unrein, F., and Massana, R. (2012). Grazing rates of protists in wetlands under contrasting light conditions due to floating plants. *Aquat. Microb. Ecol.* 65, 221–232. doi: 10.3354/ame01547

Juttner, F. (1995). Physiology and biochemistry of odorous compounds from freshwater cyanobacteria and algae. *Water Sci. Technol.* 31, 69–78. doi: 10.1016/0273-1223(95)00458-y

Kalenitchenko, D., Joli, N., Potvin, M., Tremblay, J. -É, and Lovejoy, C. (2019). Biodiversity and species change in the Arctic Ocean: a view through the lens of Nares Strait. *Front. Mar. Sci.* 6:479. doi: 10.3389/fmars.2019.00479

Katoh, K., and Standley, D. M. (2013). MAFFT Multiple Sequence Alignment Software Version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. doi: 10.1093/molbev/mst010

Kazamia, E., Czesnick, H., Thi, T. V. N., Croft, M. T., Sherwood, E., Sasso, S., et al. (2012). Mutualistic interactions between vitamin B12-dependent algae and heterotrophic bacteria exhibit regulation. *Environ. Microbiol.* 14, 1466–1476. doi: 10.1111/j.1462-2920.2012.02733.x

Koester, J. A., Swalwell, J. E., von Dassow, P., and Armbrust, E. V. (2010). Genome size differentiates co-occurring populations of the planktonic diatom *Ditylum brightwellii* (Bacillariophyta). *BMC Evol. Biol.* 10:1. doi: 10.1186/1471-2148-10-1

Lie, A. A. Y., Liu, Z. F., Terrado, R., Tatters, A. O., Heidelberg, K. B., and Caron, D. A. (2017). Effect of light and prey availability on gene expression of the mixotrophic chrysophyte, *Ochromonas* sp. *BMC Genomics* 18:163. doi: 10.1186/s12864-017-3549-1

Lie, A. A. Y., Liu, Z. F., Terrado, R., Tatters, A. O., Heidelberg, K. B., and Caron, D. A. (2018). A tale of two mixotrophic chrysophytes: insights into the metabolisms of two *Ochromonas* species (Chrysophyceae) through a comparison of gene expression. *PLoS One* 13:e0192439. doi: 10.1371/journal.pone.0192439

Lizotte, M. P. (2009). "Phytoplankton and primary production," in *Polar Lakes and Rivers: Limnology of Arctic and Antarctic Aquatic Ecosystems*, eds W. F. Vincent and J. Laybourn-Parry (Oxford: Oxford University Press), doi: 10.1093/acprof:oso/9780199213887.001.0001

Lysgaard, M. L., Eckford-Soper, L., and Daugbjerg, N. (2018). Growth rates of three geographically separated strains of the ichthyotoxic *Prymnesium parvum* (*Prymnesiophyceae*) in response to six different pH levels. *Estuar. Coast. Shelf Sci.* 204, 98–102. doi: 10.1016/j.ecss.2018.02.030

Meredith, M. M., Sommerkorn, S., Cassotta, C., Derksen, A., Ekaykin, A., Hollowed, G., et al. (2019). "3.4 Actic snow, freshwater ice and permafrost: Changes consequences and impact," in *IPCC Special Report on the Ocean and Cryosphere in a Changing Climate*, eds H.-O. Pörtner, D. C. Roberts, V. Masson-Delomotte, P. Zhai, M. Tignor, E. Polozanska, et al. (Geneva: IPCC).

Mitra, A., Flynn, K. J., Tillmann, U., Raven, J. A., Caron, D., Stoecker, D. K., et al. (2016). Defining planktonic protist functional groups on mechanisms for energy and nutrient acquisition: incorporation of diverse mixotrophic strategies. *Protist* 167, 106–120. doi: 10.1016/j.protis.2016.01.003

Moestrup, O., Lindberg, K., and Daugbjerg, N. (2009). Studies on woloszynskioid dinoflagellates IV: the genus *Biecheleria* gen. nov. *Phycol. Res.* 57, 203–220. doi: 10.1111/j.1440-1835.2009.00540.x

Moore, J. S., Harris, L. N., Kessel, S. T., Bernatchez, L., Tallman, R. F., and Fisk, A. T. (2016). Preference for nearshore and estuarine habitats in anadromous Arctic char (*Salvelinus alpinus*) from the Canadian high Arctic (Victoria Island, Nunavut) revealed by acoustic telemetry. *Can. J. Fish. Aquat. Sci.* 73, 1434–1445. doi: 10.1139/cjfas-2015-0436

Novarino, G., Lucas, I. A. N., and Morrall, S. (1994). Observations on the genus *Plagioselmis* (*Cryptophyceae*). *Cryptogam. Algol.* 15, 87–107.

Nusch, E. A. (1980). Comparison of different methods for chlorophyll and phaeopicgments determination. *Arch. Hydrobiol.* 14, 14–35.

Pla, S., and Anderson, N. J. (2005). Environmental factors correlated with chrysophyte cyst assemblages in low Arctic lakes of southwest Greenland. *J. Phycol.* 41, 957–974. doi: 10.1111/j.1529-8817.2005.00131.x

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* 41, D590–D596. doi: 10.1093/nar/gks1219

R Core Team (2017). *R: A Language and environment for statistical computing*. Vienna: R foundation for statistical computing. https://www.R-project.org

Rautio, M., Dufresne, F., Laurion, I., Bonilla, S., Vincent, W. F., and Christoffersen, K. S. (2011a). Shallow freshwater ecosystems of the circumpolar Arctic. *Ecoscience* 18, 204–222. doi: 10.2980/18-3-3463

Rautio, M., Mariash, H., and Forsstrom, L. (2011b). Seasonal shifts between autochthonous and allochthonous carbon contributions to zooplankton diets in a subarctic lake. *Limnol. Oceanogr.* 56, 1513–1524. doi: 10.4319/lo.2011.56.4.1513

Rengefors, K., and Legrand, C. (2001). Toxicity in *Peridinium aciculiferum* - an adaptive strategy to outcompete other winter phytoplankton?. *Limnol. Oceanogr.* 46, 1990–1997. doi: 10.4319/lo.2001.46.8.1990

Rognes, T., Flouri, T., Nichols, B., Quince, C., and Mahe, F. (2016). VSEARCH: a versatile open source tool for metagenomics. *PeerJ* 4:e2584. doi: 10.7717/peerj.2584

Sanders, R. W. (2011). Alternative nutritional strategies in protists: symposium introduction and a review of freshwater protists that combine photosynthesis and heterotrophy. *J. Eukaryot. Microbiol.* 58, 181–184. doi: 10.1111/j.1550-7408.2011.00543.x

Schloss, P. D. (2020). Reintroducing mothur: 10 Years Later. *Appl. Environ. Microbiol.* 86, e02343–19. doi: 10.1128/aem.02343-19

Schneider, T., Grosbois, G., Vincent, W. F., and Rautio, M. (2017). Saving for the future: pre-winter uptake of algal lipids supports copepod egg production in spring. *Freshw. Biol.* 62, 1063–1072. doi: 10.1111/fwb.12925

Shade, A., and Handelsman, J. (2012). Beyond the Venn diagram: the hunt for a core microbiome. *Environ. Microbiol.* 14, 4–12. doi: 10.1111/j.1462-2920.2011.02585.x

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313. doi: 10.1093/bioinformatics/btu033

Stoecker, D. K., Johnson, M. D., de Vargas, C., and Not, F. (2009). Acquired phototrophy in aquatic protists. *Aquat. Microb. Ecol.* 57, 279–310. doi: 10.3354/ame01340

Terrado, R., Monier, A., Edgar, R., and Lovejoy, C. (2015). Diversity of nitrogen assimilation pathways among microbial photosynthetic eukaryotes. *J. Phycol.* 51, 490–506. doi: 10.1111/jpy.12292

Teufel, A. G., Li, W., Kiss, A. J., and Morgan-Kiss, R. M. (2017). Impact of nitrogen and phosphorus on phytoplankton production and bacterial community structure in two stratified Antarctic lakes: a bioassay approach. *Polar Biol.* 40, 1007–1022. doi: 10.1007/s00300-016-2025-8

Verni, F., and Rosati, G. (2011). Resting cysts: a survival strategy in Protozoa Ciliophora. *Ital. J. Zool.* 78, 134–145. doi: 10.1080/11250003.2011.560579

Vigneron, A., Lovejoy, C., Cruaud, P., Kalenitchenko, D., Culley, A., and Vincent, W. F. (2019). Contrasting winter versus summer microbial communities and metabolic functions in a permafrost thaw lake. *Front. Microbiol.* 10:1656. doi: 10.3389/fmicb.2019.01656

White, D. M., Gerlach, S. C., Loring, P., Tidwell, A. C., and Chambers, M. C. (2007). Food and water security in a changing arctic climate. *Environ. Res. Lett.* 2:4. doi: 10.1088/1748-9326/2/4/045018

Wright, A.-D. G., and Colorni, A. (2002). Taxonomic re-assignment of *Cryptocaryon irritans*, a marine fish parasite. *Eur. J. Prostitol.* 37, 375–378.

Wrona, F. J., Johansson, M., Culp, J. M., Jenkins, A., Mard, J., Myers-Smith, I. H., et al. (2016). Transitions in Arctic ecosystems: ecological implications of a changing hydrological regime. *J. Geophys. Res. Biogeosci.* 121, 650–674. doi: 10.1002/2015jg003133

Wrona, F. J., Prowse, T. D., Reist, J. D., Hobbie, J. E., Levesque, L. M. J., and Vincent, W. F. (2006). Climate impacts on Arctic freshwater ecosystems and fisheries: background, rationale and approach of the Arctic Climate Impact Assessment (ACIA). *Ambio* 35, 326–329. doi: 10.1579/0044-7447(2006)35[326:cioafe]2.0.co;2

Xu, Y., Vick-Majors, T., Morgan-Kiss, R., Priscu, J. C., and Amaral-Zettler, L. (2014). Ciliate diversity, community structure, and novel taxa in lakes of

the McMurdo Dry Valleys, Antarctica. *Biol. Bull.* 227, 175–190. doi: 10.1086/BBLv227n2p175

Yambot, A. V., Song, Y. L., and Sung, H. H. (2003). Characterization of *Cryptocaryon irritans*, a parasite isolated from marine fishes in Taiwan. *Dis. Aquat. Org.* 54, 147–156. doi: 10.3354/dao054147

Young, K. L., and Abnizova, A. (2011). Hydrologic thresholds of ponds in a polar desert wetland environment, Somerset Island, Nunavut, Canada. *Wetlands* 31, 535–549. doi: 10.1007/s13157-011-0172-9

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

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 04 March 2022
doi: 10.3389/fmicb.2022.805694

Antarctic Glacial Meltwater Impacts the Diversity of Fungal Parasites Associated With Benthic Diatoms in Shallow Coastal Zones

*Doris Ilicic[1], Jason Woodhouse[1], Ulf Karsten[2], Jonas Zimmermann[3], Thomas Wichard[4], Maria Liliana Quartino[5], Gabriela Laura Campana[5,6], Alexandra Livenets[1], Silke Van den Wyngaert[7] and Hans-Peter Grossart[1,8]**

[1] Department of Experimental Limnology, Leibniz Institute of Freshwater Ecology and Inland Fisheries, Neuglobsow, Germany, [2] Institute of Biological Sciences, Applied Ecology and Phycology, University of Rostock, Rostock, Germany, [3] Botanic Garden and Botanical Museum Berlin-Dahlem, Freie Universität Berlin, Berlin, Germany, [4] Institute for Inorganic and Analytical Chemistry, Friedrich-Schiller-University Jena, Jena, Germany, [5] Department of Coastal Biology, Argentinean Antarctic Institute, Buenos Aires, Argentina, [6] Department of Basic Sciences, National University of Luján, Luján, Buenos Aires, Argentina, [7] Department of Biology, University of Turku, Turku, Finland, [8] Institute of Biochemistry and Biology, University of Potsdam, Potsdam, Germany

Aquatic ecosystems are frequently overlooked as fungal habitats, although there is increasing evidence that their diversity and ecological importance are greater than previously considered. Aquatic fungi are critical and abundant components of nutrient cycling and food web dynamics, e.g., exerting top-down control on phytoplankton communities and forming symbioses with many marine microorganisms. However, their relevance for microphytobenthic communities is almost unexplored. In the light of global warming, polar regions face extreme changes in abiotic factors with a severe impact on biodiversity and ecosystem functioning. Therefore, this study aimed to describe, for the first time, fungal diversity in Antarctic benthic habitats along the salinity gradient and to determine the co-occurrence of fungal parasites with their algal hosts, which were dominated by benthic diatoms. Our results reveal that Ascomycota and Chytridiomycota are the most abundant fungal taxa in these habitats. We show that also in Antarctic waters, salinity has a major impact on shaping not just fungal but rather the whole eukaryotic community composition, with a diversity of aquatic fungi increasing as salinity decreases. Moreover, we determined correlations between putative fungal parasites and potential benthic diatom hosts, highlighting the need for further systematic analysis of fungal diversity along with studies on taxonomy and ecological roles of Chytridiomycota.

Keywords: Antarctica, aquatic fungi, Chytridiomycota, phytoplankton host, salinity gradient, Illumina amplicon sequencing, Carlini Station

OPEN ACCESS

Edited by:
David Velazquez, Autonomous University of Madrid, Spain

Reviewed by:
Jun Gong, Sun Yat-sen University, China
Marlis Reich, University of Bremen, Germany

***Correspondence:**
Hans-Peter Grossart
hgrossart@igb-berlin.de

Specialty section:
This article was submitted to Extreme Microbiology, a section of the journal Frontiers in Microbiology

Received: *30 October 2021*
Accepted: *12 January 2022*
Published: *04 March 2022*

Citation:
Ilicic D, Woodhouse J, Karsten U, Zimmermann J, Wichard T, Quartino ML, Campana GL, Livenets A, Van den Wyngaert S and Grossart H-P (2022) Antarctic Glacial Meltwater Impacts the Diversity of Fungal Parasites Associated With Benthic Diatoms in Shallow Coastal Zones. Front. Microbiol. 13:805694. doi: 10.3389/fmicb.2022.805694

INTRODUCTION

Fungi are morphologically, phylogenetically, and functionally diverse. They constitute a well-founded component of terrestrial ecology due to more than 100 years of research that has highlighted their role in biogeochemical cycling and promoting biodiversity (Peay et al., 2016). Aquatic ecosystems, in particular the oceans, are frequently overlooked as fungal habitats with a

Abbreviations: EukSSU, dataset containing whole eukaryotic community; FunLSU, dataset containing only fungal community.

systematic analysis of fungal diversity and their ecological roles still in their infancy. However, there is increasing evidence that fungal diversity in aquatic ecosystems is greater and more important than previously considered (Shearer et al., 2007; Amend et al., 2019; Grossart et al., 2019). Despite sampling efforts of fungi in aquatic habitats being low compared to terrestrial environments (Rojas-Jimenez et al., 2017), molecular analyses of environmental DNA samples reveal a great diversity of novel fungal sequences, the so-called Dark Matter Fungi (Grossart et al., 2016). Although recent advances in DNA-sequencing technology have revealed that fungi are highly abundant in marine environments (Comeau et al., 2016; Taylor and Cunliffe, 2016; Tisthammer et al., 2016; Hassett et al., 2019a; Banos et al., 2020), their ecological functions and interactions with other microorganisms remain largely unexplored and missing from current general concepts (Amend et al., 2019; Richards et al., 2021).

Based on recent studies, marine fungal communities are dominated by members of the phylum Ascomycota (Tisthammer et al., 2016; Amend et al., 2019; Hassett et al., 2019a) and Chytridiomycota (Comeau et al., 2016; Hassett and Gradinger, 2016; Hassett et al., 2017; Richards et al., 2021). Chytridiomycota, frequently referred to as chytrids, are often recognized as zoosporic virulent parasites on phytoplankton which play significant roles in controlling population sizes of their hosts (Ibelings et al., 2004; Kagami et al., 2004; Rasconi et al., 2011) and altering food web structure by transferring carbon and energy between trophic levels (Rasconi et al., 2012, 2020; Sime-Ngando and Twiss, 2012; Klawonn et al., 2021b). The majority of quantitative studies of zoosporic fungi to date have been laboratory-based, with only a small number of field studies conducted to assess their importance into the broader environment. In addition, zoosporic fungi are difficult to identify using solely morphological characteristics. Although direct examination of the diversity of zoosporic fungi using high-throughput sequencing has been increasingly applied (Monchy et al., 2011; Tedersoo et al., 2015, 2018; Seto et al., 2017; Li et al., 2021; Richards et al., 2021), data are still lacking for polar regions.

Microphytobenthos (MPB) comprise phototrophic benthic microalgae living in intertidal and shallow subtidal sediments and hard substrates, among which benthic diatoms are typically the dominant organisms and are known as preferential hosts for fungal parasites such as chytrids (Scholz et al., 2016a). Although the amount of research on fungal parasites infecting pelagic forms is growing, data about microphytobenthic taxa are sparse (Wulff et al., 2009; Scholz et al., 2016b). Generally, benthic microalgae are among the main contributors to primary production in coastal zones, particularly those poor in nutrients and organic matter. They can contribute up to 42% of the marine benthic primary production (Woelfel et al., 2010). Specifically, according to Cahoon (1999), the global production of benthic microalgae ranges from 8.9 to 14.4 Gt C m^{-2} $year^{-1}$, representing approximately 20% of the global ocean production. Hence, they are important suppliers of organic carbon to grazers and sediment feeding macro- and meiofauna (Middelburg et al., 2000; Oakes et al., 2010). Except for sediments, benthic microalgae are also a major and ecologically important components of epilithic biofilms (Cahoon, 1999). They exert further important ecological roles in marine shallow water environments as they occur at the sediment-water interface and thus directly influence, for example, vertical nutrient exchange processes (Sundbäck et al., 2000). Dissolved organic carbon is regularly released by benthic diatom excretion of extrapolymeric substances (EPS), which act as stabilizing compounds against hydrodynamic sediment erosion as well as bacterial substrate (de Brouwer et al., 2005; Aslam et al., 2012).

Marine environments, especially polar regions, which have historically experienced minimum human disturbance, are a vast reservoir of microbial diversity. Their clean air, water and ice are of great importance to science for understanding different aspects of the Earth's environment. Hence, they have been drawing more attention and as such, they have not escaped the negative impacts of human activity. Their unique marine ecosystems are affected on local and regional scales by overfishing, pollution, introduction of invasive species and exploitation of mineral reserves, oil and gas (Aronson et al., 2011). Recently, in the light of global warming, the specific focus is on the loss of sea ice (Stammerjohn et al., 2008). During summer and winter, Antarctic sea ice does not show a significant overall trend (Parkinson and DiGirolamo, 2021). Nevertheless, south and west of the Antarctic Peninsula, one region has shown a persistent decline (Grosfeld et al., 2013). In February 2020 weather stations recorded the hottest temperature on record for Antarctica, reaching 18.3°C (64.9°F) (Rocha Francelino et al., 2021). This heatwave was the third major melt event of the 2019–2020 summer, following warm spells in November 2019 and January 2020. Consequently, polar aquatic ecosystems face local changes that include higher water temperatures and altered patterns in light penetration, dust deposition, sediment load and changes in salinity due to intensified glacier melt and subsequent terrestrial runoff (Eraso and Dominguez, 2007; Hernández et al., 2019). Similar climate-change-driven changes in abiotic factors were shown to affect microbial community structure, e.g., chytrid community composition in Arctic waters (Kilias et al., 2020), but for Antarctica, such data are scarce. Slow-moving or sedentary benthic communities in shallow nearshore areas are most vulnerable to these processes, with their species richness and diversity strongly impacted. Moon et al. (2015) discuss the impact of glacial retreat on epibenthic megafaunal assemblages on the West Antarctic Peninsula (WAP) and note significant differences in assemblages related to distance from the glacier, substrate grain size, and organic content. Furthermore, Braeckman et al. (2021) showed that benthic communities along the WAP shift from net autotrophy to net heterotrophy when affected by glacial meltwater. Nevertheless, more information is needed to gain a comprehensive understanding of the effects of glacial meltwater on benthic, especially fungal, communities. Although a growing amount of research focuses on fungal parasites in aquatic ecosystems, polar regions remain mainly unexplored. Rojas-Jimenez et al. (2017) reported early diverging lineages within Chytridiomycota and Cryptomycota as dominant among fungal sequences in the Antarctic Dry Valley lakes. Yet, in Antarctic marine systems, further research needs to be done

to gain more detailed insights into their diversity and abundance changes and their ecological role under current climate change.

Despite the potentially important effect of fungal parasites on benthic diatoms, to our knowledge, no published experimental or field studies have been conducted on their diversity in Antarctic coastal waters. Therefore, in this study, we aim to qualitatively describe fungal community composition in Antarctic benthic habitats. Our main focus being on the diversity of parasitic fungi and the co-occurrence with their benthic microalgal hosts along the salinity gradient in Potter Cove, Antarctic Peninsula, which is impacted by a summer induced glacial meltwater runoff. We hypothesized that members of the fungal phylum Chytridiomycota predominate fungal taxa in Antarctic shallow coastal waters and exhibit a still unexplored biodiversity with many undescribed species.

MATERIALS AND METHODS

Sites and Sample Collection

The study was conducted at the German-Argentine Dallmann Laboratory in the Argentine Scientific Station "Dr. Carlini," which is located at Potter Cove, King George Island/25 de Mayo I, Antarctic Peninsula (62°14′S, 58°31′W). This area combines zones of glacier fronts, rocky shores and extensive seabed overlaid with sand or sediment. Sampling took place in January–February 2020 in several locations around Carlini Station and the Antarctic Specially Protected Area (ASPA) 132 (**Figure 1**). The sites included marine shallow water locations and open water locations down to 60 m depth, limnic locations that included lakes and ponds, and brackish meltwater runoff (**Table 1**). The area provided diverse sediment and biofilm habitats for the sampling of benthic diatoms. Sediment surface samples were taken using Plexiglas sediment corer, and biofilm samples were taken by scratching off the surface of at least five stones with a sterilized knife. Sampling of the ocean bottom was done with the help of Argentinian army divers. Each sample was divided into two parts and prepared the following way: (i) the part for DNA analysis was fixed with 70% ethanol, and the samples were kept at –20°C; (ii) the part for microscopy was fixed with Lugol's solution.

DNA Extraction and Sequencing

DNA from the microorganisms in ethanol fixed sediment and biofilm samples was extracted according to a modified protocol described by Nercessian et al. (2005). Briefly, cell lysis was achieved using small (0.1–1 mm) zirconia-silica beads that were suspended in cetyltrimethyl ammoniumbromide (CTAB), to which anion surfactants sodium dodecyl sulfate and N-Lauroylsarcosin, proteinase K and phenol–chloroform–isoamylalcohol were added. DNA purification was facilitated by the addition of chloroform–isoamylalcohol and polyethylene glycol (PEG). Finally, DNA was precipitated at 4°C, washed with ethanol, air-dried and dissolved in ultra-pure water. The detailed protocol is available in **Supplementary Material**. PCR, library preparation and sequencing was undertaken at mrDNA laboratories (Shallow Waters, Texas, United States). The V8 region of the 18S rRNA gene was amplified using primers 18S-82F (5′- GAAACTGCGAATGGCTC-3′) and Ek-516R (5′- ACCAGACTTGCCCTCC-3′) (Hassett et al., 2019b) for molecular characterization of diatoms, followed by library preparation (2 × 300 bp) and sequencing on a MiSeq (Illumina) platform. This dataset will be referred to as EukSSU further in the text. For molecular characterization of fungi we amplified the LSU D1 region of rRNA using primers ITS4ngsF (5′-GCA TAT CAA TAA GCG SAG GAA-3′) and LF402R (5′-TTC CCT TTY ARC AAT TTC AC-3′) (Tedersoo et al., 2015). This dataset will be referred to as FunLSU further in the text.

Taxonomic Identification and Amplicon Sequence Variant Generation

Sequences were processed in R (ver. 4.0.3) using the DADA2 pipeline (ver. 1.8) (Callahan et al., 2016). Primers were removed from demultiplexed reads using cutadapt ver. 3.5 (Martin, 2011). The primer-free sequences were then filtered and trimmed to remove low-quality sequences. The DADA2 algorithm was used to infer amplicon sequence variants (ASVs). Paired-end reads were merged to obtain full denoised sequences. Chimeric sequences were removed, with the exception of 18S rRNA chimeric sequences, which were identified and removed using DECIPHER online tool ver. 11.4 (Wright et al., 2012). 18S rRNA gene sequences were run against the SILVA SSU ver. 132 database (Quast et al., 2013) for identification of all eukaryotic taxa and LSU sequences were run against LSU database using RDP Classifier and Fungal LSU training set 11 (Wang et al., 2007) for identification of fungal taxa.

Data Analysis

Data processing, visualizations and statistical analysis were performed in R. Prior to any analysis, singletons were removed from both EukSSU and FunLSU datasets. Differences in community structure in respect to water types between samples (beta diversity) were calculated using a Bray-Curtis dissimilarity measure, using *phyloseq* package (McMurdie and Holmes, 2013), and visualized through non-metric multidimensional scaling (NMDS) ordination. Permutational multivariate analysis of variance (PERMANOVA) was used to test the effect of different water types on community structure using the "Adonis" function in the R package *vegan* (Oksanen et al., 2019). The statistical significance was calculated with a *post hoc* pairwise *t*-test using "pairwise adonis" function (Martinez Arbizu, 2020) with 999 permutations. *P*-values of the pairwise *t*-test were adjusted with the Bonferroni method. To calculate alpha diversity values, *microbiome* package was used (Lahti and Shetty, 2017). The graphical representation of results was realized using the R package *ggplot2* (Wickham, 2016). Ternary plots were made using *ggtern* package (Hamilton and Ferry, 2018). Map of sampling points was generated in software QGIS 3.16.14 (QGIS Development Team, 2021) using a geospatial data package and visualization platform Quantarctica (Matsuoka et al., 2021).

Network Analysis

To infer whether the presence of suitable hosts is correlated with the presence of fungal parasites in collected samples, we performed a co-occurrence network analysis using the SparCC

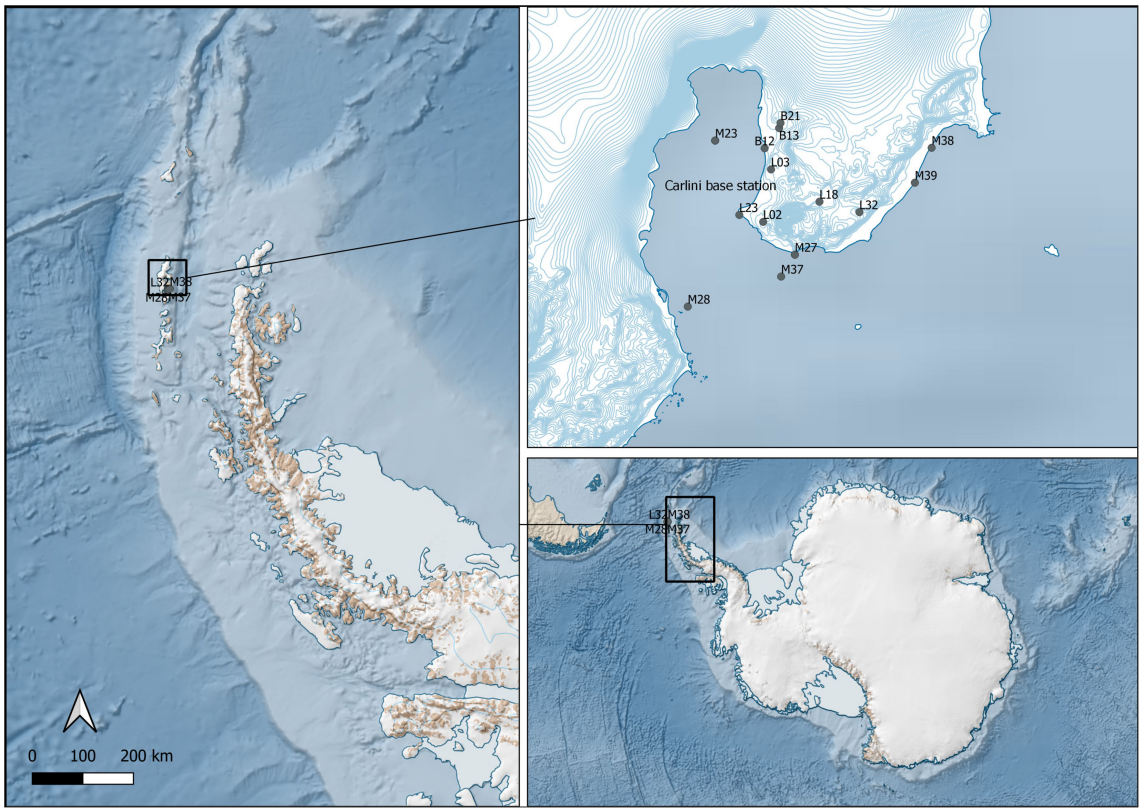

FIGURE 1 | Sampling sites in the Potter Cove area (South Shetland Islands, western Antarctic Peninsula).

TABLE 1 | Sample points characteristics.

Sample	Location	Latitude	Longitude	Habitat	Sample type	Depth	Habitat specification
B21	Punta Elefante	62°14′ 08.0″ S	58° 38′ 54.5″ W	Brackish	Biofilm	0	Meltwater
B12	Casa de bomba	62°14′ 07.78″ S	58° 39′ 27.91″ W	Brackish	Sediment	0	Meltwater
B13	Casa de bomba	62°14′ 08.76″ S	58° 39′ 30.18″ W	Brackish	Sediment	0	Meltwater
L03	Peñón I	62°14′ 16.39″ S	58° 39′ 44.4″ W	Limnic	Biofilm	0	Pond
L18	Punta Stranger	62°14′ 46.77″ S	58° 39′ 47.05″ W	Limnic	Biofilm	0	Lake
L23	Punta Elefante	62°14′ 14.6″ S	58° 40′ 46.45″ W	Limnic	Biofilm	0	Lake
L02	Peñón I	62°14′ 27.15″ S	58° 40′ 39.23″ W	Limnic	Water	0	Pond
L32	Refugio Albatros	62° 15′ 07.44″ S	58° 39′ 33.89″ W	Limnic	Water	0	Meltwater
M26	Peñón I	62°14′ 50.22″ S	58° 40′ 52.18″ W	Marine	Biofilm	0	Shore
M27	Peñón I	62°14′ 50.22″ S	58° 40′ 52.18″ W	Marine	Biofilm	2	Ocean bottom
M28	Peñón de Pesca	62°14′ 16.5″ S	58° 42′ 44.2″ W	Marine	Biofilm	5	Ocean bottom
M37	Peñón I	62°14′ 50.01″ S	58° 40′ 81.13″ W	Marine	Biofilm	0	Shore
M38	Refugio Elefante	62°15′ 22.16″ S	58° 37′ 50.1″ W	Marine	Biofilm	0	Shore
M39	Refugio Elefante	62°15′ 24.16″ S	58° 38′ 33.31″ W	Marine	Biofilm	0	Shore
M14	A4	62°13′ 43.61″ S	58° 39′ 49.36″ W	Marine	Sediment	15	Ocean bottom
M15	A4	62°13′ 43.61″ S	58° 39′ 49.36″ W	Marine	Sediment	15	Ocean bottom
M23	A4	62°13′ 43.61″ S	58° 39′ 49.36″ W	Marine	Sediment	20	Ocean bottom

algorithm (Friedman and Alm, 2012) implemented in FastSpar (Watts et al., 2019). Prior to the network analysis, ASVs with < 10% prevalence and occurring in less than three samples were removed from both EukSSU and FunLSU datasets. Data was further processed in R using *WGCNA* package (Langfelder and Horvath, 2008). We maintained the independence of both datasets and converted the correlation coefficients obtained from FastSpar into a topological overlap matrix (TOM) as the first step to identify modules of co-occurring taxa across our samples. Using parallel minimum (pmin) we calculated a consensus

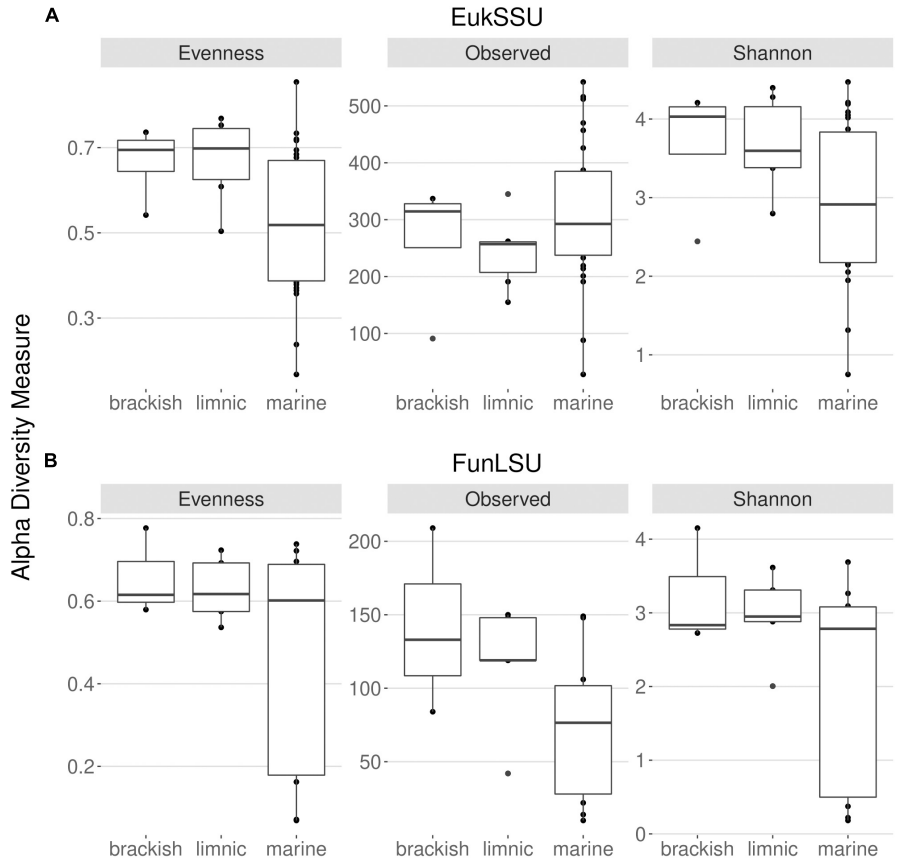

FIGURE 2 | Distribution of the alpha diversity estimators according to sampled water types in **(A)** 18S and **(B)** LSU data. This includes Pielou's evenness, observed ASV richness and Shannon diversity index.

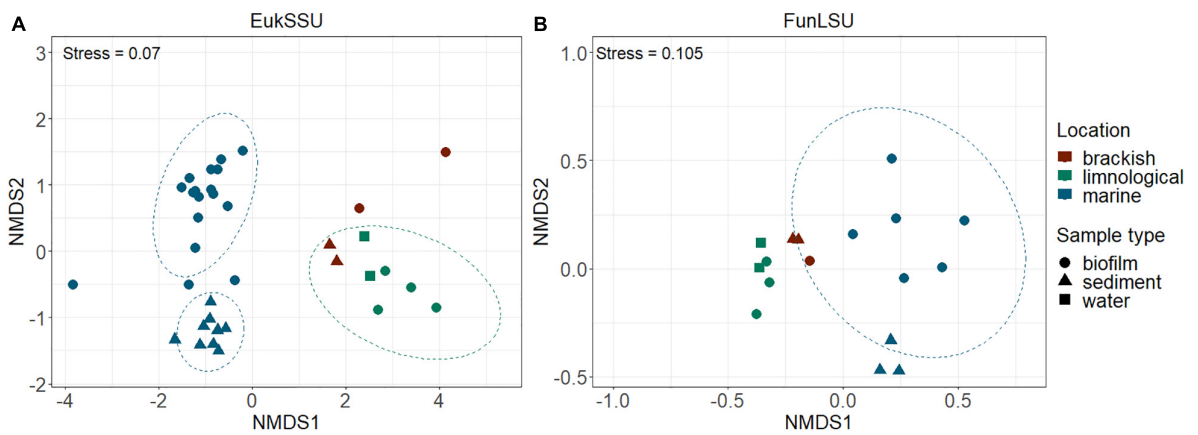

FIGURE 3 | NMDS multivariate clustering of communities according to location and sample type. In both **(A)** EukSSU and **(B)** FunLSU datasets marine samples separated from limnic (*post hoc* pairwise *t*-test, $p = 0.003$) and brackish samples (*post hoc* pairwise *t*-test, $p = 0.006$) as did biofilm from sediment samples (*post hoc* pairwise *t*-test, $p = 0.003$).

TOM from the EukSSU and FunLSU TOMs which was then used as an input for hierarchical average linkage clustering. To identify consensus modules we used cutreeDynamic with a "deepSplit" of two and "minModuleSize" of five. During this step ASVs are assigned to certain modules depending on their TOM-based topology. Further co-occurrence between identified EukSSU and FunLSU modules was determined by extracting their respective Eigen values and calculating Spearman

rank correlation coefficients. Co-occurrence between individual EukSSU and FunLSU ASVs was calculated by multiplying for each ASV pair, the between module Spearman correlation, by the module membership (range 0–1) for each ASV. In practice, this ensures that the two ASVs with the highest module membership from each respective module are most strongly correlated, and corr values decrease with decreasing module membership. We visualized for selected module pairs, weighted correlation coefficients (> 0.5) between ASVs using Cytoscape 3.8.2 software.

Microscopy

CalcoFluor White (CFW) staining approach along with epifluorescence microscopy, as suggested by Rasconi et al. (2009), was used to detect and identify fungal parasites attached to an algal hosts. CFW is a non-specific, fluorescent dye that binds to chitin in fungal cell walls but also cellulose, that is often present in cell wall of some algae and fungi-like organisms (Kagami et al., 2007; Priest et al., 2021). Pre-treatment of sediment samples was done using ultrasound (3 × 10 s) for the mechanical disruption of diatom-sand agglomerates (Scholz et al., 2014). Staining was performed as described in Klawonn et al. (2021a). Five microliter of CFW (1 mg/ml) was added to 1 ml sample and incubated for 15 min at room temperature. Samples were then transferred in Utermöhl counting chambers, after which they were incubated for 10 min, allowing the cells to sink to the bottom of the chamber. For further visualization, Nikon Eclipse Ti2 inverted microscope was used and the samples were screened for fungal parasites on 600x magnification.

RESULTS

With this study it was intended to highlight the prevalence of fungi in Antarctic benthic environments, determine the impact of glacial meltwater on benthic diatoms and their associated fungal communities and to identify correlation in their presence. To this end, freshwater, brackish and marine benthic environments were sampled and two datasets obtained: EukSSU containing the whole eukaryotic community and FunLSU containing fungal community. When examining the EukSSU community composition, a total of 3,790 sequence variants (ASVs) were identified of which 145 were classified as fungi. The relative proportion of these 145 fungal ASVs varied between different water types, from 0.97% in marine, 7.4% in brackish, to 8.9% in limnic habitats. LSU primers were not strictly fungi-specific and out of 1,204 identified ASVs, of which 727 were fungi, the remaining identified as Metazoa (262), Stramenopiles (131), Amoebozoa (35), Chlorophyta (33), Rhodophyta (12), and Alveolata (4). Before any downstream analysis, all non-fungal sequences were removed from FunLSU dataset. Nevertheless, relative proportions of fungal ASVs in FunLSU confirmed differences observed in EukSSU and 51.3% of sequences were classified as fungi in marine, 73.2% in brackish and 69.3% in limnic environments. Calculating alpha diversity measures, we note that eukaryotic and fungal diversity were lower in marine samples compared to either brackish or limnic samples (**Figure 2**). For eukaryotes, the dominance of a few taxa was the largest impact on diversity in marine environments, with Pielou's evenness higher for eukaryotic communities in brackish and limnic samples (**Figure 2A**). In contrast, fungal communities had more even species composition amongst all three water types, with marine samples being less diverse compared to brackish or limnic samples (**Figure 2B**). Due to artificial variation introduced through subsampling, the omission of valid data through loss of sequence counts, or exclusion of samples with small library sizes, sequences were not rarefied prior to calculating alpha diversity (Cameron et al., 2021). As Willis (2019) discussed, environments can be identical with respect to one alpha diversity metric, but the different abundance structures will induce different biases when rarefied.

Fungal Community Composition

To determine community composition, we carried out a non-NMDS analysis and it confirmed a distinct separation based on sample habitat (marine vs. brackish vs. limnic) (PERMANOVA, $p = 0.001$) in both EukSSU (**Figure 3A**) and FunLSU (**Figure 3B**) datasets. Marine samples were separated from limnic (*post hoc* pairwise *t*-test, $p = 0.003$) and brackish samples ($p = 0.006$), but no significant difference was observed between limnic and brackish sites. We detected a difference based on sample type (biofilm vs. sediment vs. water) (PERMANOVA, $p = 0.001$). *Post hoc* tests confirmed a difference between biofilm and sediment samples ($p = 0.003$). Summary of all the statistical tests performed can be found in **Supplementary Table 1** for EukSSU dataset and **Supplementary Table 2** for FunLSU dataset.

Ascomycota and Chytridiomycota were the most abundant taxa (FunLSU dataset) in marine Antarctic benthic habitats, with no apparent distinction based on habitat or sample type (**Figure 4A**). Ascomycota represented 49% of the fungal reads and 25.4% of the fungal ASVs, while Chytridiomycota represented 22.4% of the fungal reads and 38.9% of the ASVs. In addition, a high diversity of the phylum Chytridiomycota was observed (**Figure 4B**), with Spizellomycetales and Monoblepharidales being the most abundant. In the EukSSU dataset, it was noted that fungi were more dominant in freshwater and brackish environments (8.9 and 7.4%), relative to marine environments (0.97%) (**Figure 5**). This was consistent with an increased diversity of fungi, as observed in both the EukSSU and FunSSU datasets.

Co-occurrence of Fungal Parasites and Their Hosts

Using microscopy, we observed Calcofluor White stained sporangia attached to benthic algae which indicated active fungal infections (**Supplementary Figure 1**). Furthermore, we performed a co-occurrence network analysis to identify correlations between fungi and benthic algae to identify putative hosts. Due to the compositional nature of both the EukSSU and FunLSU datasets, it was not possible to estimate a direct linear or sparse correlation between the two datasets. Rather, we calculated sparse correlations amongst ASVs within each dataset, clustered ASVs into modules, extracted Eigen values for each module and used these values to identify correlated modules.

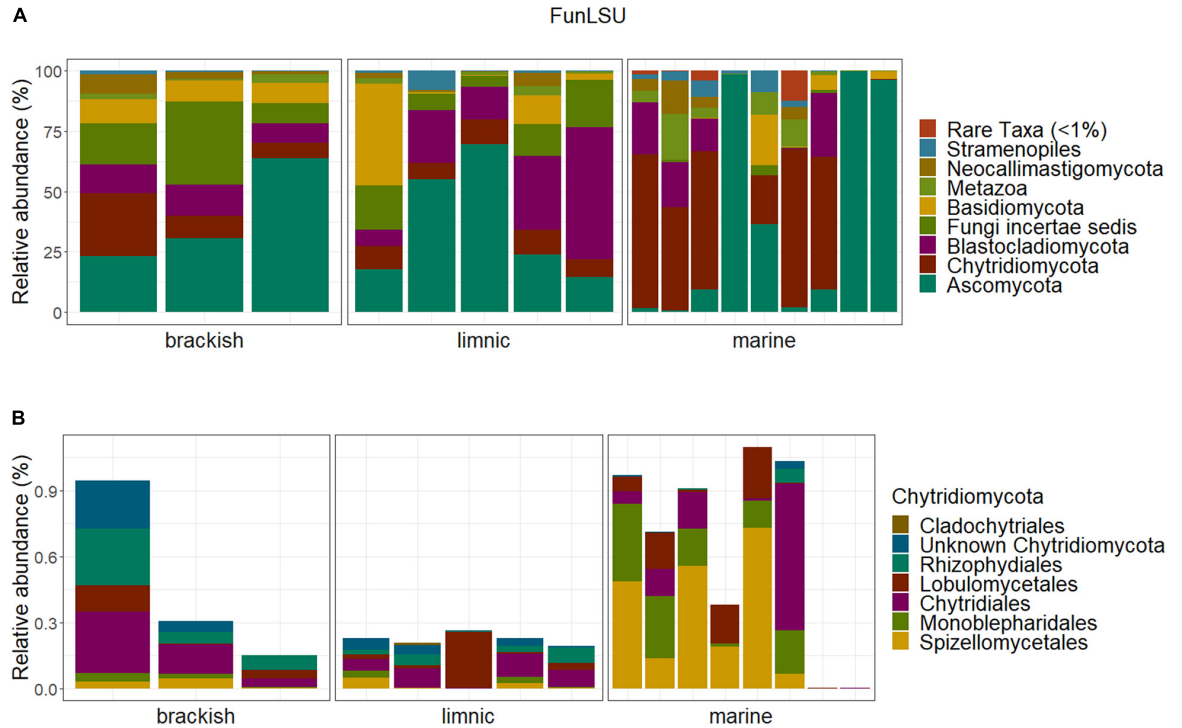

FIGURE 4 | Community composition in the Antarctic benthic habitats, considering the abundance distribution in composition in sampled water types **(A)** of the whole fungal community **(B)** and the phylum Chytridiomycota relative to the total fungal community.

In total, we identified seven modules in FunLSU and eight modules in EukSSU dataset. Modules largely reflected differences between habitats. FunLSU modules typically always contained at least one Chytrid ASV (**Supplementary Figure 2**), with Mesochytrium, Maunachytrium, Betamyces, and Spizellomyces being most abundant. For four out of the six FunLSU modules containing chytrids, we observed a correlation between that module and a EukSSU module containing a high abundance of diatoms (**Supplementary Figure 3**). For two FunLSU modules that contained chytrids, we did not observe a correlation with a EukSSU module, although all EukSSU modules contained diatoms (**Supplementary Figure 4**).

To better illustrate the correlation between diatoms and zoosporic fungi, we visualized weighted correlation coefficients amongst ASVs obtained from two EukSSU and FunSSU module pairs. In EukSSU3 and FunLSU5 module pair (**Figure 6A**) we visualized only those correlations that had a correlation value higher than 0.2. Both modules contained ASVs enriched from marine benthic environments. Specifically, we observed *Fragilariopsis cylindrus*, a pennate, benthic sea-ice diatom found both in Arctic and Antarctic waters (Helmcke and Krieger, 1953), three different *Navicula* spp. (*N. phyllepta*, *N. rhynchocephala*, *N. perminuta*) and *Pleurosigma intermedium*, all marine benthic diatoms that could serve as potential hosts for parasitic chytrids. Modules EukSSU4 and FunLSU2 (**Figure 6B**) contained ASVs enriched in brackish/limnic benthic habitats. In benthic environments, we observed a diverse number of benthic diatoms associated with some Chytridiomycota. In this case, we observed Betamyces (Rhizophydiales), chytrids mostly recognized as parasites on different phytoplankton species (Christaki et al., 2017), and Maunachytrium, for which there is no solid evidence for a specific lifestyle. Most abundant in both networks were ASVs classified as unknown Chytridiomycota, and they were highly correlated with certain diatom ASVs.

DISCUSSION

Climate warming in Antarctic environments has been associated with glacier retreat and increased ice melting (Lee et al., 2017) which, in turn, change the vertical structure of the water column, especially in Antarctic shallow coastal environments. During the summer months (December—March), meltwaters on the Western Antarctic Peninsula, where this study was undertaken, occur regularly. It has been shown that meltwater transports high particle loads to the coastal water thereby reducing incident light conditions and salinity that affect microbial community dynamics, leading to changes in species composition (Schloss et al., 2008; Antoni et al., 2020). In this study, we showed different proportions of fungal sequences between different water types. Hence, we assume a shaping impact of salinity on both the eukaryotic and fungal diversity in Potter Cove. Brackish sites, receiving strong inputs of fresh meltwater, exhibited a higher eukaryotic and fungal diversity than that of marine sites. Moreover, many of the same fungal taxa/ASVs found in freshwater environments were also determined in brackish habitats, suggesting fresh meltwater runoff, including terrestrial soil particles, might be

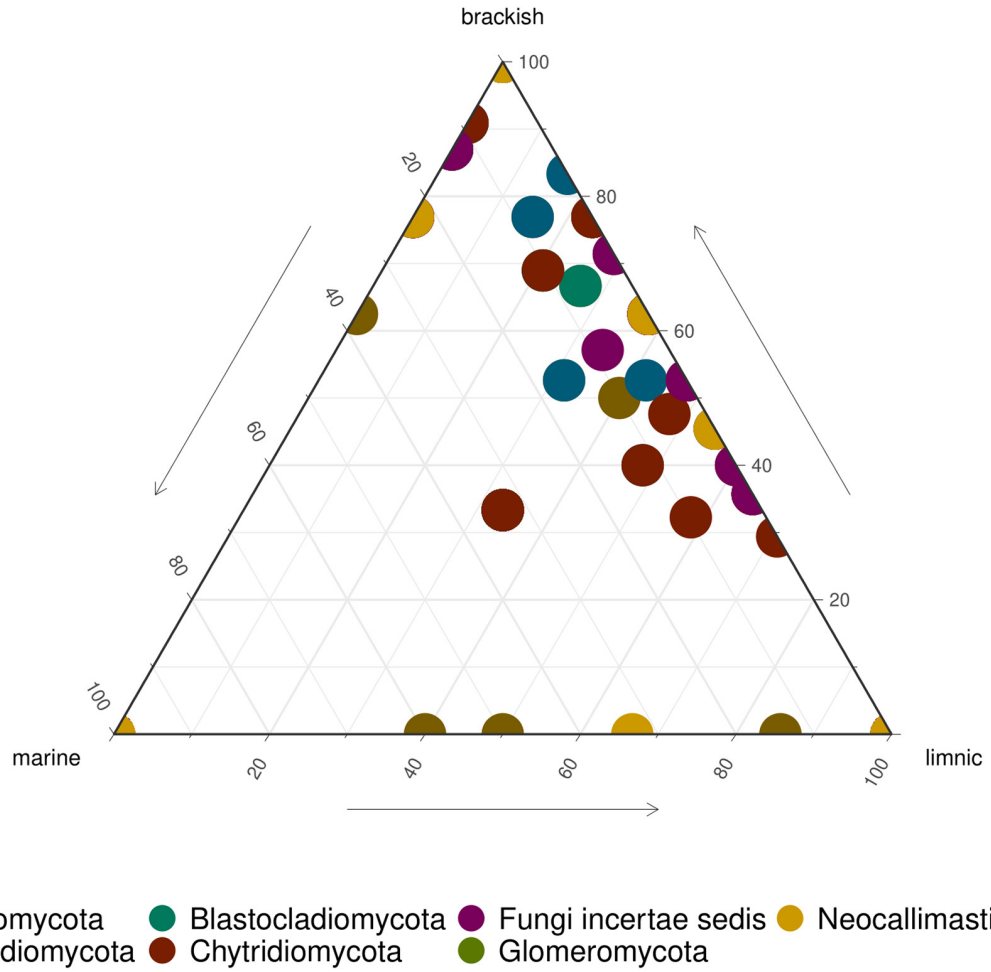

FIGURE 5 | Ternary plot showing relative abundance of fungal ASVs in LSU data. Each point represents an ASV. The position represents the relative abundance of the ASV with respect to each environment.

a direct source of this increased diversity. Furthermore, it is tempting to speculate that the increased richness of fungi and the decreased dominance of individual algal hosts in brackish/freshwater environments, as shown with alpha diversity measures, might be linked. We argue that this observation can be directly linked with fungal parasites playing an active role in maintaining the diversity of algae (van Donk, 1989; Sime-Ngando and Twiss, 2012). Alternatively, this might be indirect in the sense that fertilization (Schloss et al., 2008) opens up new niches for algae, with fungal diversity, of both saprotrophic and/or parasitic fungi, promoted by an increase in the diversity of autochthonous and allochthonous organic matter (Wang et al., 2017). Finally, we conclude that the increased evenness in brackish/freshwater habitats compared to marine habitats may indicate that glacial meltwaters reduce selective pressure and increase niche opportunities for algae, possibly due to increased nutrients as observed by Death et al. (2014). In contrast, observed differences in fungal diversity may have arisen from a decreased richness in marine habitats relative to brackish/freshwater.

Our results are supported by the study of Kilias et al. (2020), who showed that chytrid fungi are primarily encountered at sites influenced by sea ice melt. The evolutionary history of Chytridiomycota suggests that they originated under brackish-/freshwater-like conditions and thus evolved successively in parallel with the host organisms or existing food (Berbee et al., 2017). It remains an open question whether marine Chytridiomycota proliferate due to lower competition in high salinity water or whether specialization of some taxa has occurred over an evolutionary time scale (Yang et al., 2021). Moreover, a substantial proportion of fungal ASVs in our FunLSU dataset were classified as terrestrial fungi. For example, Spizellomycetales, the most abundant genera among chytrids, are described as soil-inhabiting fungi (Barr, 1980). As a result, we hypothesize that spores and fungal hyphae may arrive in Antarctic coastal waters via glacial/terrestrial runoff from ice melting followed by deposition due to sedimentation alongside other particles. In particular, larger areas of bare soil at higher temperatures lead to an increased deposition of soil particles on the melting glaciers, staining them dark and even decreasing

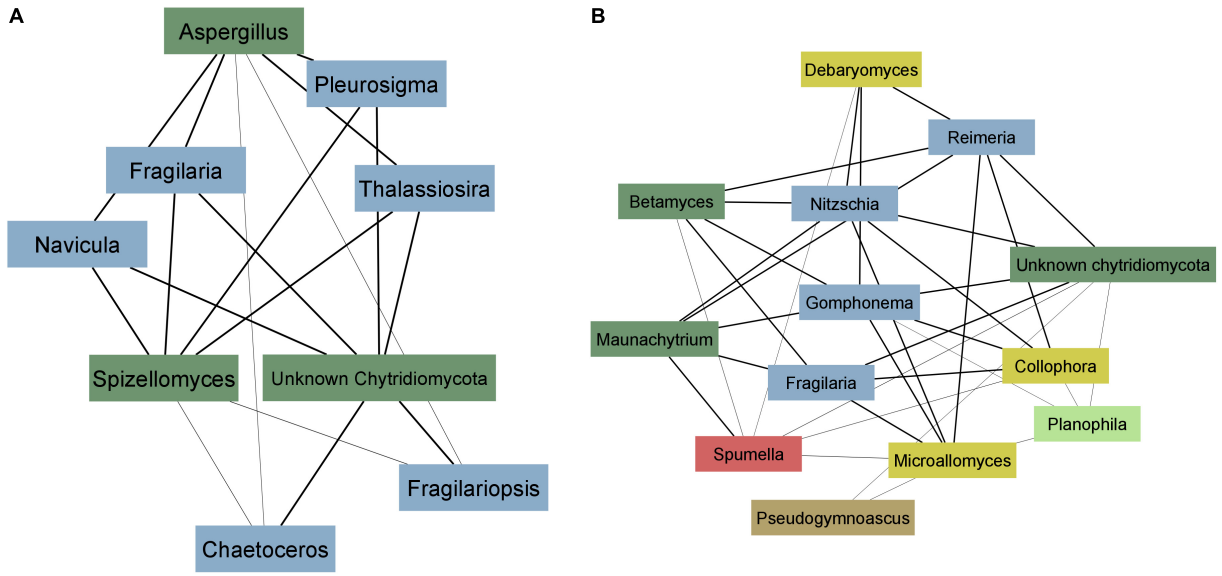

FIGURE 6 | Co-occurrence network analysis. **(A)** EukSSU3 and FunLSU5 modules **(B)** EukSSU4 and FunLSU2 module. Blue—Baccilariophyceae; dark green—Chytridiomycota; yellow—other Fungi; brown—Opisthokonta, red—Stramenopiles; light green—Chlorophyta.

their albedo and increasing their melting (Schmitt et al., 2015). This notion is supported by the presence of Monoblepharidales, which are described as typical freshwater fungi (Sparrow, 1933), and constitute the dominant genera among fungal taxa in the sampled marine environments. Such sediment accumulation driven by increased glacier retreat may decrease the rate of microphytobenthic (MPB) primary production. As earliers discussed by Hoffmann et al. (2019), MPB might survive the consequences but their contribution to the overall primary production as a carbon resource may decline. Consequently, food competition of the benthic heterotrophic community will increase, with unpredictable consequences for biomass, density, structure and diversity of the benthic community and food web structure.

Using microscopy and performing co-occurrence network analysis, we found preliminary evidence suggesting that zoosporic fungi in Antarctic waters are involved in parasitic relationships with benthic diatoms. In summary, in most cases, parasitic chytrids were associated with the presence of Bacillariophyceae, but some diatom taxa occurred in some cases independent of parasitic chytrids. Furthermore, unclassified Chytridiomycota and Spizellomyces, which are known as terrestrial parasites on other chytrid species and soil nematodes (Lozupone and Klein, 2002), were non-specifically associated with multiple potential diatom hosts. It is tempting to speculate that these chytrids exhibit a parasitic lifestyle in aquatic habitats with broad host tolerance, considering that these ASVs represent a large hidden diversity. Whilst network approaches may provide hints for interactions between fungi and diatoms, these are ultimately correlations. Cultivation and targeted single cell approaches that identify host-parasite pairs and better discriminate between parasitic and saprotrophic interactions are still needed (van den Wyngaert et al., 2017). Despite this, our study highlights that diatom-parasite interactions are likely in polar benthic environments and may have significant ecological roles.

CONCLUSION

Our work suggests that salinity has a major impact in shaping eukaryotic microbial communities in shallow coastal waters of Antarctica. In particular, we determined a significant difference in eukaryotic and fungal diversity between sampled habitats, with both eukaryotic and fungal diversity higher in brackish/freshwater environments. On the one hand, considering the potential parasitic lifestyle of dominant fungal taxa, we may conclude that fungal parasites play an active role in maintaining the diversity of benthic algae. On the other hand, increased ice melting during summer months may cause nutrient loading in brackish waters resulting in increased diversity. It remains an open question whether the positive effect on fungal diversity also has a positive or a negative effect on the ecosystem on a broad-scale, considering parasitic interactions. The factor of inputs of meltwater is also supported by the fact that freshwater and brackish environments shared many of the same fungal taxa/ASVs and that a substantial proportion of fungal ASVs in the FunLSU dataset were classified as terrestrial fungi. In terms of parasitic lifestyle, while many of the chytrid ASVs were found to be non-specifically associated with multiple potential diatom hosts, drawing mechanistic conclusion about host-parasite interactions is difficult because the most abundant ASVs in our networks were classified as unknown Chytridiomycota. Systematic analysis of fungal diversity in Antarctic regions is still in its infancy, and the taxonomy and ecological functions of Chytridiomycota species remain largely unexplored.

Our findings are a significant step toward illuminating the diversity of fungal communities and occurrence of, what it may be, fungal parasites in the Antarctic region. Benthic diatoms play a key role in polar food webs. Hence, studying chytrid diversity and their interactions with the algal hosts will provide us with a comprehensive understanding of polar ecosystems. The effects of salinity changes on fungi-microalgae interactions and microalgal communities, as well as its consequences for the coastal ecosystem of Potter Cove, are therefore subject to future studies.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are publicly available. This data can be found here: (https://www.ebi.ac.uk/ena/browser/view/PRJEB49266/PRJEB49266).

AUTHOR CONTRIBUTIONS

AL, H-PG, MLQ, GLC, JZ, UK, and TW organized an expedition to Antarctica and collected samples. AL performed microscopy and extracted DNA from collected samples. DI and JW analyzed the data. DI wrote the manuscript. SVdW assisted with her knowledge during the analysis. H-PG, JW, MLQ, GLC, JZ, UK, TW, and SVdW edited and provided feedback on the original draft. All authors approved the submitted version.

FUNDING

The study was performed in the frame of the Priority Programme of the German Science Foundation (DFG) entitled "Antarktisforschung mit vergleichenden Untersuchungen in arktischen Eisgebieten" (SPP 1158) in the subproject "Polar parasites" (GR1540/33-1) given to H-PG and UK (KA899/39-1). The DFG-SPP 1158 subproject "Benthic diatom diversity and biogeography" ZI 1628/2-1 is given to JZ and UK (KA899/39-1). GLC was supported by grants from IAA-DNA (PIH18), PADI Foundation (47918/2020), MINCYT-BMBF Programme (AL/17/06-01DN18024), and Universidad Nacional de Luján (Disp CDD-CB 296/18, 343/19, and 69/21). GLC and MLQ were supported by grants from ANPCyT-DNA (PICTO 2010-0116 and PICT 2017-2691). This study presents an outcome of Coastcarb (Funding ID 872690, H2020-MSCA-RISE-2019). Publication is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project Number 491466077.

ACKNOWLEDGMENTS

We would like to thank the team of the Argentinian Antarctic Research Station "Carlini" for their support and logistics as well as AWI Logistics department. We also appreciate the support of the MIBI group at IGB Stechlin for their frequent scientific input and discussions. We acknowledge the support of the Deutsche Forschungsgemeinschaft and Open Access Publishing Fund of University of Potsdam.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2022.805694/full#supplementary-material

REFERENCES

Amend, A., Burgaud, G., Cunliffe, M., Edgcomb, V. P., Ettinger, C. L., Gutiérrez, M. H., et al. (2019). Fungi in the marine environment: open questions and unsolved problems. *mBio* 10:e01189-18. doi: 10.1128/mBio.01189-18

Antoni, J. S., Almandoz, G. O., Ferrario, M. E., Hernando, M. P., Varela, D. E., Rozema, P. D., et al. (2020). Response of a natural Antarctic phytoplankton assemblage to changes in temperature and salinity. *J. Exp. Mar. Biol. Ecol.* 532:151444. doi: 10.1016/j.jembe.2020.151444

Aronson, R. B., Thatje, S., McClintock, J. B., and Hughes, K. A. (2011). Anthropogenic impacts on marine ecosystems in Antarctica. *Ann. N. Y. Acad. Sci.* 1223, 82–107. doi: 10.1111/J.1749-6632.2010.05926.X

Aslam, S. N., Cresswell-Maynard, T., Thomas, D. N., and Underwood, G. J. C. (2012). Production and characterization of the intra- and extracellular carbohydrates and polymeric substances (EPS) of three sea-ice diatom species, and evidence for a cryoprotective role for EPS. *J. Phycol.* 48, 1494–1509. doi: 10.1111/JPY.12004

Banos, S., Gysi, D. M., Richter-Heitmann, T., Glöckner, F. O., Boersma, M., Wiltshire, K. H., et al. (2020). Seasonal dynamics of pelagic mycoplanktonic communities: interplay of taxon abundance. Temporal occurrence, and biotic interactions. *Front. Microbiol.* 11:1305. doi: 10.3389/fmicb.2020.01305

Barr, D. J. S. (1980). An outline for the reclassification of the Chytridiales, and for a new order, the Spizellomycetales. *Can. J. Bot.* 58, 2380–2394. doi: 10.1139/b80-276

Berbee, M. L., James, T. Y., and Strullu-Derrien, C. (2017). Early diverging fungi: diversity and impact at the dawn of terrestrial life. *Annu. Rev. Microbiol.* 71, 41–60. doi: 10.1146/annurev-micro-030117-020324

Braeckman, U., Pasotti, F., Hoffmann, R., Vázquez, S., Wulff, A., Schloss, I. R., et al. (2021). Glacial melt disturbance shifts community metabolism of an Antarctic seafloor ecosystem from net autotrophy to heterotrophy. *Commun. Biol.* 4:148. doi: 10.1038/s42003-021-01673-6

Cahoon, L. B. (1999). The role of benthic microalgae in neritic ecosystems. *Oceanogr. Mar. Biol. Annu. Rev.* 37, 55–94. doi: 10.1201/9781482298550-4

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

Cameron, E. S., Schmidt, P. J., Tremblay, B. J.-M., Emelko, M. B., and Müller, K. M. (2021). To rarefy or not to rarefy: enhancing diversity analysis of microbial communities through next-generation sequencing and rarefying repeatedly. *bioRxiv*[Preprint]. doi: 10.1101/2020.09.09.290049

Christaki, U., Genitsaris, S., Monchy, S., Li, L. L., Rachik, S., Breton, E., et al. (2017). Parasitic eukaryotes in a meso-eutrophic coastal system with marked Phaeocystis globosa blooms. *Front. Mar. Sci.* 4:416. doi: 10.3389/fmars.2017.00416

Comeau, A. M., Vincent, W. F., Bernier, L., and Lovejoy, C. (2016). Novel chytrid lineages dominate fungal sequences in diverse marine and freshwater habitats. *Sci. Rep.* 6:30120. doi: 10.1038/srep30120

de Brouwer, J. F. C., Wolfstein, K., Ruddy, G. K., Jones, T. E. R., and Stal, L. J. (2005). Biogenic stabilization of intertidal sediments: the importance of extracellular

polymeric substances produced by benthic diatoms. *Microb. Ecol.* 49, 501–512. doi: 10.1007/S00248-004-0020-Z

Death, R., Wadham, J. L., Monteiro, F., le Brocq, A. M., Tranter, M., Ridgwell, A., et al. (2014). Antarctic ice sheet fertilises the Southern Ocean. *Biogeosciences* 11, 2635–2643. doi: 10.5194/BG-11-2635-2014

Eraso, A., and Dominguez, C. (2007). Physicochemical characteristics of the subglacier discharge in Potter Cove, King George Island, Antarctica. *Karst Cryokarst Stud. Faculty Earth Sci.* 45, 111–122.

Friedman, J., and Alm, E. J. (2012). Inferring correlation networks from genomic survey data. *PLoS Comput. Biol.* 8:e1002687. doi: 10.1371/JOURNAL.PCBI.1002687

Grosfeld, K., Treffeisen, R., Asseng, J., Bartsch, A., Bräuer, B., Fritzsch, B., et al. (2013). Online sea-ice knowledge and data platform. *Polarforschung* 85, 143–155. doi: 10.2312/polfor.2016.011

Grossart, H. P., van den Wyngaert, S., Kagami, M., Wurzbacher, C., Cunliffe, M., and Rojas-Jimenez, K. (2019). Fungi in aquatic ecosystems. *Nat. Rev. Microbiol.* 17, 339–354. doi: 10.1038/s41579-019-0175-8

Grossart, H. P., Wurzbacher, C., James, T. Y., and Kagami, M. (2016). Discovery of dark matter fungi in aquatic ecosystems demands a reappraisal of the phylogeny and ecology of zoosporic fungi. *Fungal Ecol.* 19, 28–38. doi: 10.1016/j.funeco.2015.06.004

Hamilton, N. E., and Ferry, M. (2018). Ggtern: ternary diagrams using ggplot2. *J. Stat. Softw.* 87, 1–17. doi: 10.18637/jss.v087.c03

Hassett, B. T., Borrego, E. J., Vonnahme, T. R., Rämä, T., Kolomiets, M. V., and Gradinger, R. (2019a). Arctic marine fungi: biomass, functional genes, and putative ecological roles. *ISME J.* 13, 1484–1496. doi: 10.1038/s41396-019-0368-1

Hassett, B. T., Ducluzeau, A. L. L., Collins, R. E., and Gradinger, R. (2017). Spatial distribution of aquatic marine fungi across the western Arctic and sub-arctic. *Environ. Microbiol.* 19, 475–484. doi: 10.1111/1462-2920.13371

Hassett, B. T., and Gradinger, R. (2016). Chytrids dominate arctic marine fungal communities. *Environ. Microbiol.* 18, 2001–2009. doi: 10.1111/1462-2920.13216

Hassett, B. T., Thines, M., Buaya, A., Ploch, S., and Gradinger, R. (2019b). A glimpse into the biogeography, seasonality, and ecological functions of arctic marine Oomycota. *IMA Fungus* 10:6. doi: 10.1186/s43008-019-0006-6

Helmcke, J. G., and Krieger, W. (1953). *Diatomeenschalen im elektronenmikroskopischen Bild. | Citations - Diatoms of North America*. Available online at: https://diatoms.org/citations/helmcke_jg_and_krieger_w-1953-diatomeenschalen_im_elektronenmikroskopischen (accessed August 5, 2021).

Hernández, E. A., Lopez, J. L., Piquet, A. M. T., mac Cormack, W. P., and Buma, A. G. J. (2019). Changes in salinity and temperature drive marine bacterial communities' structure at Potter Cove, Antarctica. *Polar Biol.* 42, 2177–2191. doi: 10.1007/S00300-019-02590-5/FIGURES/6

Hoffmann, R., Al-Handal, A. Y., Wulff, A., Deregibus, D., Zacher, K., Quartino, M. L., et al. (2019). Implications of glacial melt-related processes on the potential primary production of a microphytobenthic community in potter cove (Antarctica). *Front. Mar. Sci.* 6:655. doi: 10.3389/FMARS.2019.00655/BIBTEX

Ibelings, B. W., de Bruin, A., Kagami, M., Rijkeboer, M., Brehm, M., and van Donk, E. (2004). Host parasite interactions between freshwater phytoplankton and chytrid fungi (Chytridiomycota). *J. Phycol.* 40, 437–453. doi: 10.1111/j.1529-8817.2004.03117.x

Kagami, M., Arnout De Bruin, A. E., Bas, A. E., Ibelings, W., Ellen, A. E., and Donk, V. (2007). Parasitic chytrids: their effects on phytoplankton communities and food-web dynamics. *Hydrobiologia* 578, 113–129. doi: 10.1007/s10750-006-0438-z

Kagami, M., van Donk, E., de Bruin, A., Rijkeboer, M., and Ibelings, B. W. (2004). *Daphnia* can protect diatoms from fungal parasitism. *Limnol. Oceanogr.* 49, 680–685. doi: 10.4319/lo.2004.49.3.0680

Kilias, E. S., Junges, L., Šupraha, L., Leonard, G., Metfies, K., and Richards, T. A. (2020). Chytrid fungi distribution and co-occurrence with diatoms correlate with sea ice melt in the Arctic Ocean. *Commun. Biol.* 3:183. doi: 10.1038/s42003-020-0891-7

Klawonn, I., van den Wyngaert, S., Parada, A. E., Arandia-Gorostidi, N., Whitehouse, M. J., Grossart, H.-P., et al. (2021b). Characterizing the "fungal shunt": parasitic fungi on diatoms affect carbon flow and bacterial communities in aquatic microbial food webs. *Proc. Natl. Acad. Sci. U.S.A.* 118:e2102225118. doi: 10.1073/PNAS.2102225118

Klawonn, I., Dunker, S., Kagami, M., Grossart, H. P., and van den Wyngaert, S. (2021a). Intercomparison of two fluorescent dyes to visualize parasitic fungi (Chytridiomycota) on Phytoplankton. *Microb. Ecol.* doi: 10.1007/s00248-021-01893-7

Lahti, L., and Shetty, S. (2017). *Tools for Microbiome Analysis in R. Version*. Available online at: https://microbiome.github.io/ (accessed January 20, 2022).

Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9:559. doi: 10.1186/1471-2105-9-559

Lee, J. R., Raymond, B., Bracegirdle, T. J., Chadès, I., Fuller, R. A., Shaw, J. D., et al. (2017). Climate change drives expansion of Antarctic ice-free habitat. *Nature* 547, 49–54. doi: 10.1038/nature22996

Li, Y., Steenwyk, J. L., Chang, Y., Wang, Y., James, T. Y., Stajich, J. E., et al. (2021). A genome-scale phylogeny of the kingdom Fungi. *Curr. Biol.* 31, 1653–1665.e5. doi: 10.1016/j.cub.2021.01.074

Lozupone, C. A., and Klein, D. A. (2002). Molecular and cultural assessment of chytrid and Spizellomyces populations in grassland soils. *Mycologia* 94, 411–420. doi: 10.1080/15572536.2003.11833206

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.J.* 17:10. doi: 10.14806/ej.17.1.200

Martinez Arbizu, P. (2020). pairwiseAdonis: Pairwise Multilevel Comparison Using Adonis. R Package Version 0.4.

Matsuoka, K., Skoglund, A., Roth, G., de Pomereu, J., Griffiths, H., Headland, R., et al. (2021). Quantarctica, an integrated mapping environment for Antarctica, the Southern Ocean, and sub-Antarctic islands. *Environ. Modell. Softw.* 140:105015. doi: 10.1016/j.envsoft.2021.105015

McMurdie, P. J., and Holmes, S. (2013). Phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* 8:e61217. doi: 10.1371/journal.pone.0061217

Middelburg, J. J., Barranguet, C., Boschker, H. T. S., Herman, P. M. J., Moens, T., and Heip, C. H. R. (2000). The fate of intertidal microphytobenthos carbon: an in situ 13C-labeling study. *Limnol. Oceanogr.* 45, 1224–1234. doi: 10.4319/LO.2000.45.6.1224

Monchy, S., Sanciu, G., Jobard, M., Rasconi, S., Gerphagnon, M., Chabé, M., et al. (2011). Exploring and quantifying fungal diversity in freshwater lake ecosystems using rDNA cloning/sequencing and SSU tag pyrosequencing. *Environ. Microbiol.* 13, 1433–1453. doi: 10.1111/j.1462-2920.2011.02444.x

Moon, H. W., Wan Hussin, W. M. R., Kim, H. C., and Ahn, I. Y. (2015). The impacts of climate change on Antarctic nearshore mega-epifaunal benthic assemblages in a glacial fjord on King George Island: responses and implications. *Ecol. Indic.* 57, 280–292. doi: 10.1016/J.ECOLIND.2015.04.031

Nercessian, O., Noyes, E., Kalyuzhnaya, M. G., Lidstrom, M. E., and Chistoserdova, L. (2005). Bacterial populations active in metabolism of C1 compounds in the sediment of Lake Washington, a freshwater lake. *Appl. Environ. Microbiol.* 71, 6885–6899. doi: 10.1128/AEM.71.11.6885-6899.2005

Oakes, J. M., Connolly, R. M., and Revill, A. T. (2010). Isotope enrichment in mangrove forests separates microphytobenthos and detritus as carbon sources for animals. *Limnol. Oceanogr.* 55, 393–402. doi: 10.4319/LO.2010.55.1.0393

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., et al. (2019) *Vegan: Community Ecology Package, R package version 2.5-6*. Available online at: https://CRAN.R-project.org/package=vegan (accessed January 20, 2022).

Parkinson, C. L., and DiGirolamo, N. E. (2021). Sea ice extents continue to set new records: Arctic, Antarctic, and global results. *Remote Sens. Environ.* 267:112753. doi: 10.1016/J.RSE.2021.112753

Peay, K. G., Kennedy, P. G., and Talbot, J. M. (2016). Dimensions of biodiversity in the Earth mycobiome. *Nat. Rev. Microbiol.* 14, 434–447. doi: 10.1038/NRMICRO.2016.59

Priest, T., Fuchs, B., Amann, R., and Reich, M. (2021). Diversity and biomass dynamics of unicellular marine fungi during a spring phytoplankton bloom. *Environ. Microbiol.* 23, 448–463. doi: 10.1111/1462-2920.15331

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* 41, D590–D596. doi: 10.1093/NAR/GKS1219

Rasconi, S., Jobard, M., Jouve, L., and Sime-Ngando, T. (2009). Use of calcofluor white for detection, identification, and quantification of phytoplanktonic fungal parasites. *Appl. Environ. Microbiol.* 75, 2545–2553. doi: 10.1128/AEM.02211-08

Rasconi, S., Jobard, M., and Sime-Ngando, T. (2011). Parasitic fungi of phytoplankton: ecological roles and implications for microbial food webs. *Aquat. Microb. Ecol.* 62, 123–137. doi: 10.3354/AME01448

Rasconi, S., Niquil, N., and Sime-Ngando, T. (2012). Phytoplankton chytridiomycosis: community structure and infectivity of fungal parasites in aquatic ecosystems. *Environ. Microbiol.* 14, 2151–2170. doi: 10.1111/J.1462-2920.2011.02690.X

Rasconi, S., Ptacnik, R., Danner, S., van den Wyngaert, S., Rohrlack, T., Pilecky, M., et al. (2020). Parasitic chytrids upgrade and convey primary produced carbon during inedible algae proliferation. *Protist* 171:125768. doi: 10.1016/j.protis.2020.125768

Richards, T. A., Leonard, G., Mahé, F., del Campo, J., Romac, S., Jones, M. D. M., et al. (2021). Molecular diversity and distribution of marine fungi across 130 European environmental samples. *Proc. R. Soc. B* 282:20152243. doi: 10.1098/rspb.2015.2243

Rocha Francelino, M., Schaefer, C., de Los Milagros Skansi, M., Colwell, S., Bromwich, D. H., Jones, P., et al. (2021). WMO evaluation of two extreme high temperatures occurring in February 2020 for the Antarctic Peninsula region. *Bull. Am. Meteorol. Soci.* 102, 1–20. doi: 10.1175/BAMS-D-21-0040.1

Rojas-Jimenez, K., Wurzbacher, C., Bourne, E. C., Chiuchiolo, A., Priscu, J. C., and Grossart, H. P. (2017). Early diverging lineages within Cryptomycota and Chytridiomycota dominate the fungal communities in ice-covered lakes of the McMurdo Dry Valleys, Antarctica. *Sci. Rep.* 7:15348. doi: 10.1038/s41598-017-15598-w

Schloss, I., Ferreyra, G., González, Ó, Atencio, A., Fuentes, V., Tosonotto, G., et al. (2008). Long term hydrographic conditions and climate trends in Potter Cove. The Potter Cove coastal ecosystem, Antarctica. Synopsis of research performed 1999-2006 at the Dallmann Laboratory and Jubany Station, King George Island (Isla 25 de Mayo). *Ber. Polarforsch. Meeres.*

Schmitt, C. G., All, J. D., Schwarz, J. P., Arnott, W. P., Cole, R. J., Lapham, E., et al. (2015). Measurements of light-absorbing particles on the glaciers in the Cordillera Blanca, Peru. *Cryosphere* 9, 331–340. doi: 10.5194/TC-9-331-2015

Scholz, B., Guillou, L., Marano, A. V., Neuhauser, S., Sullivan, B. K., Karsten, U., et al. (2016a). Zoosporic parasites infecting marine diatoms — A black box that needs to be opened. *Fungal Ecol.* 19:59. doi: 10.1016/J.FUNECO.2015.09.002

Scholz, B., Küpper, F. C., Vyverman, W., and Karsten, U. (2016b). Effects of eukaryotic pathogens (Chytridiomycota and Oomycota) on marine benthic diatom communities in the Solthörn tidal flat (southern North Sea, Germany). *Eur. J. Phycol.* 51, 253–269. doi: 10.1080/09670262.2015.1134814

Scholz, B., Küpper, F. C., Vyverman, W., and Karsten, U. (2014). Eukaryotic pathogens (Chytridiomycota and Oomycota) infecting marine microphytobenthic diatoms-a methodological comparison. *J. Phycol.* 50, 1009–1019. doi: 10.1111/jpy.12230

Seto, K., Kagami, M., and Degawa, Y. (2017). Phylogenetic position of parasitic chytrids on diatoms: characterization of a novel clade in Chytridiomycota. *J. Eukaryot. Microbiol.* 64, 383–393. doi: 10.1111/jeu.12373

Shearer, C. A., Descals, E., Kohlmeyer, B., Kohlmeyer, J., Marvanová, L., Padgett, D., et al. (2007). Fungal biodiversity in aquatic habitats. *Biodivers. Conserv.* 16, 49–67. doi: 10.1007/s10531-006-9120-z

Sime-Ngando, T., and Twiss, M. R. (2012). Phytoplankton chytridiomycosis: fungal parasites of phytoplankton and their imprints on the food web dynamics. *Front. Microbiol.* 3:361. doi: 10.3389/fmicb.2012.00361

Sparrow, F. K. (1933). The Monoblepharidales. Available online at: https://about.jstor.org/terms (accessed May 26, 2021).

Stammerjohn, S. E., Martinson, D. G., Smith, R. C., and Iannuzzi, R. A. (2008). Sea ice in the western Antarctic Peninsula region: spatio-temporal variability from ecological and climate change perspectives. *Deep Sea Res. II Top. Stud. Oceanogr.* 55, 2041–2058. doi: 10.1016/j.dsr2.2008.04.026

Sundbäck, K., Miles, A., and Göransson, E. (2000). Nitrogen fluxes, denitrification and the role of microphytobenthos in microtidal shallow-water sediments: an annual study. *Mar. Ecol. Prog. Ser.* 200, 59–76. doi: 10.3354/MEPS200059

Taylor, J. D., and Cunliffe, M. (2016). Multi-year assessment of coastal planktonic fungi reveals environmental drivers of diversity and abundance. *ISME J.* 10, 2118–2128. doi: 10.1038/ismej.2016.24

QGIS Development Team (2021). *QGIS Geographic Information System*. Available online at: https://www.qgis.org (accessed December 7, 2021).

Tedersoo, L., Anslan, S., Bahram, M., Põlme, S., Riit, T., Liiv, I., et al. (2015). Shotgun metagenomes and multiple primer pair-barcode combinations of amplicons reveal biases in metabarcoding analyses of fungi. *Mycokeys* 10, 1–43. doi: 10.3897/mycokeys.10.4852

Tedersoo, L., Tooming-Klunderud, A., and Anslan, S. (2018). PacBio metabarcoding of Fungi and other eukaryotes: errors, biases and perspectives. *New Phytol.* 217, 1370–1385. doi: 10.1111/nph.14776

Tisthammer, K. H., Cobian, G. M., and Amend, A. S. (2016). Global biogeography of marine fungi is shaped by the environment. *Fungal Ecol.* 19, 39–46. doi: 10.1016/j.funeco.2015.09.003

van den Wyngaert, S., Seto, K., Rojas-Jimenez, K., Kagami, M., and Grossart, H.-P. (2017). A new parasitic chytrid, Staurastromyces oculus (Rhizophydiales, Staurastromycetaceae fam. nov.), infecting the freshwater desmid Staurastrum sp. *Protist* 168, 392–407. doi: 10.1016/J.PROTIS.2017.05.001

van Donk, E. (1989). The role of fungal parasites in phyto-plankton succession.. *Plankton Ecol.* ed. U. Sommer (Berlin: Springer), 171–194. doi: 10.1007/978-3-642-74890-5_5

Wang, Q., Garrity, G. M., Tiedje, J. M., and Cole, J. R. (2007). Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* 73, 5261–5267. doi: 10.1128/AEM.00062-07

Wang, Y., Guo, X., Zheng, P., Zou, S., Li, G., and Gong, J. (2017). Distinct seasonality of chytrid-dominated benthic fungal communities in the neritic oceans (Bohai Sea and North Yellow Sea). *Fungal Ecol.* 30, 55–66. doi: 10.1016/j.funeco.2017.08.008

Watts, S. C., Ritchie, S. C., Inouye, M., and Holt, K. E. (2019). FastSpar: rapid and scalable correlation estimation for compositional data. *Bioinformatics* 35, 1064–1066. doi: 10.1093/BIOINFORMATICS/BTY734

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. New York, NY: Springer-Verlag, doi: 10.1007/978-3-319-24277-4

Willis, A. D. (2019). Rarefaction, alpha diversity, and statistics. *Front. Microbiol.* 10:2407. doi: 10.3389/FMICB.2019.02407/BIBTEX

Woelfel, J., Schumann, R., Peine, F., Flohr, A., Kruss, A., Tegowski, J., et al. (2010). Microphytobenthos of Arctic Kongsfjorden (Svalbard, Norway): biomass and potential primary production along the shore line. *Polar Biol.* 33, 1239–1253. doi: 10.1007/s00300-010-0813-0

Wright, E. S., Yilmaz, L. S., and Noguera, D. R. (2012). DECIPHER, a search-based approach to chimera identification for 16S rRNA sequences. *Appl. Environ. Microbiol.* 78, 717–725. doi: 10.1128/AEM.06516-11

Wulff, A., Iken, K., Quartino, M. L., Al-Handal, A., Wiencke, C., and Clayton, M. N. (2009). Biodiversity, biogeography and zonation of marine benthic micro- and macroalgae in the Arctic and Antarctic. *Bot. Mar.* 52, 491–507. doi: 10.1515/BOT.2009.072

Yang, Y., Banos, S., Gerdts, G., Wichels, A., and Reich, M. (2021). Mycoplankton biome structure and assemblage processes differ along a transect from the Elbe River down to the river plume and the adjacent marine waters. *Front. Microbiol.* 12:640469. doi: 10.3389/fmicb.2021.640469

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

The handling editor declared a past co-authorship with one of the authors UK.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 24 March 2022
doi: 10.3389/fmicb.2022.787146

Microbial Community Changes in 26,500-Year-Old Thawing Permafrost

Maria Scheel[1], Athanasios Zervas[2], Carsten S. Jacobsen[2] and Torben R. Christensen[1,3]*

[1] Department of Ecoscience, Arctic Research Centre, Aarhus University, Roskilde, Denmark, [2] Department of Environmental Science, Aarhus University, Roskilde, Denmark, [3] Oulanka Research Station, Oulu University, Oulu, Finland

Northern permafrost soils store more than half of the global soil carbon. Frozen for at least two consecutive years, but often for millennia, permafrost temperatures have increased drastically in the last decades. The resulting thermal erosion leads not only to gradual thaw, resulting in an increase of seasonally thawing soil thickness, but also to abrupt thaw events, such as sudden collapses of the soil surface. These could affect 20% of the permafrost zone and half of its organic carbon, increasing accessibility for deeper rooting vegetation and microbial decomposition into greenhouse gases. Knowledge gaps include the impact of permafrost thaw on the soil microfauna as well as key taxa to change the microbial mineralization of ancient permafrost carbon stocks during erosion. Here, we present the first sequencing study of an abrupt permafrost erosion microbiome in Northeast Greenland, where a thermal erosion gully collapsed in the summer of 2018, leading to the thawing of 26,500-year-old permafrost material. We investigated which soil parameters (pH, soil carbon content, age and moisture, organic and mineral horizons, and permafrost layers) most significantly drove changes of taxonomic diversity and the abundance of soil microorganisms in two consecutive years of intense erosion. Sequencing of the prokaryotic 16S rRNA and fungal ITS2 gene regions at finely scaled depth increments revealed decreasing alpha diversity with depth, soil age, and pH. The most significant drivers of variation were found in the soil age, horizons, and permafrost layer for prokaryotic and fungal beta diversity. Permafrost was mainly dominated by Proteobacteria and Firmicutes, with *Polaromonas* identified as the most abundant taxon. Thawed permafrost samples indicated increased abundance of several copiotrophic phyla, such as Bacteroidia, suggesting alterations of carbon utilization pathways within eroding permafrost.

Keywords: permafrost erosion, abrupt thaw, 16S, fungi, Greenland, amplicon sequencing, soil microbiome, biodiversity

OPEN ACCESS

Edited by:
Jérôme Comte,
Université du Québec, Canada

Reviewed by:
Bärbel Ulrike Fösel,
Helmholtz Center München,
Helmholtz Association of German Research Centres (HZ), Germany
Christine M. Foreman,
Montana State University, United States
Sophie Crevecoeur,
Environment and Climate Change, Canada

****Correspondence:***
Maria Scheel
maria.scheel@ecos.au.dk

Specialty section:
This article was submitted to Extreme Microbiology, a section of the journal Frontiers in Microbiology

Received: *30 September 2021*
Accepted: *09 February 2022*
Published: *24 March 2022*

Citation:
Scheel M, Zervas A, Jacobsen CS and Christensen TR (2022) Microbial Community Changes in 26,500-Year-Old Thawing Permafrost. Front. Microbiol. 13:787146. doi: 10.3389/fmicb.2022.787146

INTRODUCTION

Ambient temperatures in the Arctic have increased by 3.1°C between 1971 and 2019, equaling roughly three times that of the global average temperature increase (AMAP, 2021). This trend is predicted to continue over the next 30 years with a local increase of as much as 3–6°C (IPCC, 2019). These complex consequences are projected to intensify in future scenarios (IPCC, 2021), including the recession of glacial and ice sheet extent, a decline in sea ice, and the thawing of permafrost (AMAP, 2021). These soils consist of a seasonally thawing active layer and the

underlying permafrost table, which has remained permanently frozen for at least two consecutive years. Permafrost-affected soils take up a quarter of the global soil surface while storing more than half of the global soil carbon (Tarnocai et al., 2009; Schuur et al., 2015). The Northern hemisphere permafrost zone currently contains roughly twice the amount of carbon as the atmosphere is estimated to store (Schuur and Mack, 2018). These geological reservoirs are at risk as Arctic permafrost soils have warmed by 2–3°C since the 1970s and by 0.3°C between 2007 and 2016, recorded within the coldest permafrost sites, such as found in Greenland (IPCC, 2019; AMAP, 2021). An increase of 3.3–10°C above the 1985–2014 average temperature is expected by 2100 (AMAP, 2021). Thermally induced gradual active layer deepening makes formerly frozen carbon stocks available for both deeper rooting vegetation and microbial decomposition (Hayes et al., 2014; AMAP, 2019). In contrast, abrupt thaw, often triggered by melting ground ice, can impact deeper permafrost carbon by means of deep incisions into the soil profile. Currently, these processes affect up to 20% of all permafrost-affected areas, although by 2100, up to 60% could be impacted (Turetsky et al., 2020). A global increase of 1.5–2°C is projected to make 28–53% of all permafrost organic carbon bioavailable (IPCC, 2019). Turetsky et al. (2020) modeled that 613–802 and 624 Tg CO_2 equivalents per year loss to the atmosphere could be accounted to gradual and abrupt permafrost thaw, respectively, until 2100 under the RCP8.5 scenario (IPCC, 2019). Permafrost carbon has accumulated over geological timescales and, once decomposed into gases, its degradation is an irreversible climate change indicator (AMAP, 2021).

When studying the connection between climate change and Arctic ecosystems (Cavicchioli et al., 2019), the importance of soil microorganisms is often underestimated. They are ecological sentinels of changing environmental conditions due to their fast growth rates and diverse metabolic potential (Edwards et al., 2020). In previous studies, DNA has been successfully extracted from up to 1 million-year-old permafrost (Willerslev et al., 2004) as well as 600,000-year-old DNA from viable permafrost organisms (Johnson et al., 2007). Microbial activity has been measured in permafrost samples as cold as −39°C (Panikov et al., 2006). Cryophilic organisms are adapted physiologically to sub-zero temperatures, but also to desiccation, high salinity, nutrient scarcity, and ground radiation, by adapting membrane fluidity, forming spores, and metabolizing complex organic sources (Collins and Margesin, 2019). As permafrost soils thaw, the existing cold-adapted microbial community, containing bacteria, archaea, viruses, and micro-eukaryotes such as fungi, is strongly affected by warmer temperatures, higher soil water content from melted ground ice, and increased nutrient availability from deeper rooting vegetation (Margesin, 2009). Not only do increasing ambient temperatures impose changing ecological conditions on the soil microbiome, but metabolic rates increase, and microorganisms can decompose diverse carbon substrates into climate-active gases (Jansson and Taş, 2014; Mackelprang et al., 2016; Schuur and Mack, 2018). As recently reviewed, microbial decomposition rates of ancient permafrost carbon are one aspect of the essential linkage between climate change and environmental microbiology and can greatly influence the carbon flux between permafrost and the atmosphere, as well as the linkage with the global nitrogen cycle (Cavicchioli et al., 2019).

Interest in Arctic permafrost has recently increased, rooted partly in the speed of irreversible erosion processes and knowledge gaps in dominant taxa and their metabolic potential on decomposing ancient permafrost matter during microbial succession in thawing permafrost soils (Edwards et al., 2020). Research in this area has been conducted but is still rare and focused on a few sites and mainly carried out using laboratory incubations and cultivation studies (Mackelprang et al., 2016; Metcalfe et al., 2018). In this paper, we investigated *in situ* which physicochemical soil parameters drive the microbial community composition and diversity changes with vertical abrupt thermal erosion as well as after 1 year of additional thawing. These insights will facilitate the development of quantitative approaches and builds a taxonomic baseline for future genetic research in this field, such as comparative metatranscriptomics. Furthermore, the rarely observed impact of abrupt loss of ancient microbiomes and the following ecological succession enhances our understanding of ecological cascades under climate change.

MATERIALS AND METHODS

Sampling and Site Description

The Zackenberg Valley, Northeast Greenland (74°30′N, 20°30′W), is a wide lowland valley dominated by continuous permafrost. According to Elberling et al. (2010), annual air temperatures averaged −9.5°C between 1996 and 2007 and summer temperatures varied between 3 and 7°C from 1996 to 2007 (Hansen et al., 2008; Hollesen et al., 2011). The vegetation in the valley consists primarily of wet hummocky fens and low shrub and graminoid species (Elberling et al., 2008), while the lowlands east of the Zackenberg river are *Cassiope tetragona* and grass heathland (Bay, 1998). Permafrost up to 1 m deep had an average temperature of −2°C in summer and −14°C in winter between 1997 and 2006 (Christiansen et al., 2008). The average maximum active layer thickness ranged from 40 cm up to 2 m in depth and increased by 0.8–1.5 cm per year between 1996 and 2012 (Elberling et al., 2013). Long-term soil temperature data from 1995 to 2020 from the monitoring site SAL-1 at depths of 0, 20, 40, 60, 80, and 100 cm were retrieved from the Greenland Ecosystem Monitoring database[1,2] and averaged for August and September.

After an intense snowmelt event in the 2018 summer season, a formerly existing minor thermokarst, a depression of the eroding permafrost soil surface, collapsed 400 m north-east of the Zackenberg Research Station. The incision into the soil profile was about 1 m deep and created an erosion gully that led toward the Zackenberg River. An ice lens at 40–60 cm depth was visible, which had melted in 2019 before sampling took place (**Supplementary Figure 1**). The thaw depth was measured with a thin steel pole pushed into the soil vertically three times until resistance indicated the frozen permafrost table had been reached

[1]https://data.g-e-m.dk/

[2]https://doi.org/10.17897/89G6-QD54

TABLE 1 | Physicochemical soil parameters were measured at the sampling site and included radiocarbon dating (^{14}C), which was measured only in 2019 and was calculated in fM (*) for samples younger than 1963 or BP (**) for samples older than 1945.

Depth [cm]	^{14}C [*fM **BP]	H_2O [%]		SOM [%]		pH		Horizon	Layer	
		2019	2020	2019	2020	2019	2020		2019	2020
0–10	*1.04	7.94 ± 1.03	28.80 ± 3.38	16.01 ± 4.22	8.73 ± 1.82	3.87 ± 0	4.22 ± 0.03	O1	AL	
10–20	*1.13	11.77 ± 0.93	22.52 ± 0.73	7.09 ± 1.71	5.39 ± 0.83	4.46 ± 0.05	4.02 ± 0.05	M1		
20–30	*1.16	11.83 ± 0.34	22.05 ± 3.30	5.75 ± 0.21	10.19 ± 3.22	4.25 ± 0.02	4.29 ± 0.01			
30–40	*1.20	13.13 ± 3.12	26.57 ± 0.83	5.85 ± 2.23	13.68 ± 0.44	4.73 ± 0.09	4.25 ± 0.01			
40–50	**2,635	28.27 ± 4.92	7.73 ± 1.58	24.42 ± 4.78	2.59 ± 0.68	5.08 ± 0.05	4.63 ± 0.03	O2	TZ	TZ
50–60	**3,770	23.49 ± 1.92	15.61 ± 6.28	15.76 ± 4.50	2.93 ± 0.30	5.45 ± 0.10	4.13 ± 0.02			
60–70	**26,500	14.14 ± 8.43	8.72 ± 2.53	3.78 ± 0.09	1.52 ± 0.16	6.54 ± 0.05	4.86 ± 0.05	M2		
70–80	**22,100	16.31 ± 1.09	6.35 ± 0.60	3.99 ± 0.25	1.10 ± 0.24	7.01 ± 0.01	4.48 ± 0		PF	
80–90	**26,200	1.65 ± 4.43	7.96 ± 0.61	1.78 ± 0.30	1.00 ± 0.20	6.61 ± 0	4.57 ± 0.01			
90–100	NA	NA	7.80 ± 0.59	NA	1.04 ± 0.26	NA	4.91 ± 0.01			PF

In both years, weight-based relative soil moisture (H_2O) and soil organic matter (SOM) were measured based on loss on ignition and the pH was determined in triplicates and standard deviation is indicated (±). Layers included the active layer (AL, 0-40 cm) as long-term seasonally thawing material above the former ice lens, the transition zone (TZ, 2019: 40-70 cm, 2020: 40-90 cm) as the newly thawed material since the collapse, and permafrost (PF, 2019: > 70 cm, 2020: > 90 cm) as the still frozen ground. Based on SOM and ^{14}C, four separate horizons were visible, namely the recent surface organic horizon (O1, 0-10 cm), the recent underlying mineral horizon (M1, 10-40 cm), the older buried organic horizon (O2, 40-60 cm), and an ancient mineral horizon (M2, > 60 cm).

(Christiansen, 1999). Based on these measurements, the layers were defined as the active layer (the long-term seasonally thawing top 40 cm), the transition zone of newly thawed permafrost material (40–70 cm in 2019 and 40–90 cm in 2020), and permafrost as the continuously frozen layer below these (below 70 cm in 2019 and 90 cm in 2020) as visualized in **Supplementary Figure 2**. In 2019 and 2020, three biological replicates were taken aseptically for every 10 cm interval until a depth of 90 or 100 cm, respectively. Due to the long transport chain without stable frozen transport technology and the bias, inconsistent freeze-thawing processes were shown to inflict on cold-adapted microbiomes (Lim et al., 2020), all samples were stored at 4°C at the Research Station before controlled, cooled transportation to Denmark, then stored at 4°C until further laboratory analysis took place as reported earlier (Gittel et al., 2014b).

Loss on Ignition and Radiocarbon Dating

For radiocarbon dating, field-wet soil per 10-cm horizon for 2019 was sifted in technical triplicated with a 0.5-mm sieve in deionized water to retrieve macro plant residues for further analysis, while the roots were removed under a stereomicroscope. To counter the low biomass, the triplicates were pooled per depth, before chemical pre-treatment with HCl and NaOH. The samples were graphitized before measuring the ^{14}C isotope activity using an accelerator mass spectrometer (Radiocarbon Dating Laboratory, Lund University, Lund, Sweden). The obtained peaks were compared to established ^{14}C calibration curves, considering anthropogenic atmospheric ^{14}C activity changes. The results were reported as age in ^{14}C years in BP (before present = AD 1950) from 1650 to 1950 using the terrestrial calibration curve IntCal13 (Reimer et al., 2013) and for younger datings than 1963 in fM (fraction modern) using the Levin post-Bomb calibration (Levin and Kromer, 2004).

Loss on ignition was performed in technical triplicate for each 10-cm depth interval per year on wet soil, by air-drying at 70°C for 48 h, then weighing to determine the weight-based relative soil water (H_2O) loss. The samples were burned at 450°C for 2 h in ceramic cups and weighed to obtain the weight-based relative organic carbon content (SOM). To measure the pH, 10 ml of air-dried soil in triplicates were added to 50 ml of 1 M KCl in Falcon Tubes in technical triplicate. After shaking for 1 h and resting for 1.5 h, the pH was measured with a Mettler Toledo FiveEasy Plus™ pH Meter (Mettler Toledo GmbH, Gieβ en, Germany).

Based on the radiocarbon age and organic matter content of the soil, we defined each a surface (1) and buried (2) organic (O) and mineral (M) horizon, resulting in two organic (O1, O2) and mineral (M1, M2) horizons (**Table 1** and **Supplementary Figure 2**). In contrast, layers were defined based on the thermal state of the soil with a local active layer (AL) designated as the seasonally thawing material until 40 cm. The transition zone (TZ) was designated as former permafrost material which had thawed since the collapse to 70 cm in 2019 and 90 cm in 2020. The underlying permafrost layer (PF) remained frozen, indicated by visible ice crystals in the material (**Supplementary Figure 3**).

DNA Extraction and Sequencing

DNA was extracted from 2 g field-wet soil per year and 10-cm depth interval for each biological replicate no later than 67 days after sampling, using the DNAeasy® PowerSoil® Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. Given the low biomass and DNA concentration, these extraction triplicates were pooled for sequencing in order to maximize coverage. The resulting extracts were quantified with a Qubit® 2.0 Fluorometer (Thermo Fisher Scientific, Life Technologies, Roskilde, Denmark). For prokaryotic DNA, the 16S rRNA V3-V4 region was amplified with the primer set F341 (5′-CCTAYGGGRBGCASCAG-3′) and R806 (5′-GGACTACNNGGGTATCTAAT-3′) (Takahashi et al., 2014), while for fungal sequences the ITS2 region was targeted with the primers ITS1F2 (5′-GAACCWGCGGARGGATCA-3′) and ITS2 (5′-GCTGCGTTCTTCATCGATGC-3′) (Gaylarde et al., 2017). For the PCR, negative controls confirmed absence of

contamination. The final PCR products were cleaned with magnetic beads (MagBio Genomics Inc., Gaithersburg, Maryland, US) and fresh 80% ethanol, and stored in 1x TE buffer at −20°C. The quality and size of the amplicons were verified by running an agarose gel and screening with Tapestation 4150 (Agilent Technologies, Santa Clara, California, US). Paired Illumina MiSeq 2 × 300 bp sequencing was performed at Teknologisk Institut, Taastrup, Denmark.

Community and Diversity Analysis

All of the following analyses were performed using the QIIME 2 pipeline and plugins version 2019.10.0 (Bolyen et al., 2019). Raw sequence reads were filtered by removing the multiplexing and gene region primer sequences with Cutadapt (Martin, 2011). The 50–60 cm 16S samples from 2020 and 0–10 cm and 60–70 cm ITS sequences from 2019 were removed due to an insufficient quality of reads after filtering. Noisy and redundant sequences were removed, to reduce sequencing errors and chimeras with dada2 denoise (Callahan et al., 2017). Unique filtered sequences were defined as amplicon sequencing variants (ASVs) and used for downstream analysis (**Supplementary Table 1**).

The prokaryotic ASVs were classified with the closest taxonomic affiliation with the classifier sklearn (Pedregosa et al., 2011) against the SILVA database (version 138_99) and the UNITE database version 8.2 (version 2020-02-20, Quast et al., 2013) with global and 97% singletons for fungi (Nilsson et al., 2019). The phylogeny was created with fasttree on classified sequences (Version 2.1.10, Price et al., 2010). For each sample, alpha phylogenetic diversity (PD) was determined with Faith's PD index. Several other diversity indices were calculated but not used for downstream analysis (**Supplementary Table 1**). Beta diversity, as the variation of diversity across several samples, was calculated as Bray–Curtis dissimilarities (BC) as a non-phylogenetic index, and weighted Unifrac as pairwise phylogenetic distances.

Statistical Significance Tests

On ASV level, phylogenetic alpha diversity correlation was tested with environmental metadata for variation, including pairwise Kruskal–Wallis covariation (non-parametric Mann–Whitney *U*-test) in QIIME2. Permutational multivariate analysis of variance (PERMANOVA) on 999 permutations was performed using the "adonis" function to determine significant drivers of community composition. A Principal coordinate analysis (PCoA) was performed on the weighted Unifrac distances and plotted as biplots, utilizing the Emperor QIIME2 plugin. Additional analyses were conducted using the "vegan" packages (Oksanen et al., 2019) in the R studio environment and R version 4.1.2 (R Core Team, 2021; R Studio Team, 2021). Correlation-based indicator species analysis (Dufrêne and Legendre, 1997) was performed on phyla abundances for each categorical parameter using the "indval" function. Non-Metric Multidimensional Scaling (NMDS) biplots with fitted environmental parameters were performed on BC dissimilarities based on the relative abundance of phyla.

RESULTS

Physicochemical Soil Analyses

At our study site, the weight-based relative soil organic matter (SOM) content indicated a mainly organic top 10 cm horizon (O1). It was underlain by the silt and clay mineral horizon (M1), which reached until the active layer limit at 30–40 cm depth with an average SOM content < 10%. Both were dated to be younger than 58 years old. Formerly separated from these by ground ice, the 20-cm deep organic horizon (O2) from 50–60 cm depth indicated up to 3,841 years old underlying material and after the melting of the ice lens in 2019 highest relative, weight-based soil water content (H_2O) of 28.27%. The deepest horizon (M2) from 60 cm depth and deeper consisted of sand and pebbles and up to 26,571-year-old material. This horizon included the upper permafrost table at 70 cm in 2019 and 90 cm in 2020. In 2019, the pH ranged from 3.87 of the top 10 cm to 7.01 at 70–80 cm depth, while the pH range in 2020 was stable throughout the profile from 4.01 to 4.91 (**Table 1**).

Prokaryotic Community Composition and Abundance

Amplicon sequencing revealed a frequency of 10,256,910 counts across 18 samples for raw 16S sequences. After filtering, 25,471 unique reads were defined as prokaryotic amplicon sequence variants (ASVs, for sequencing statistics see **Supplementary Table 1**). Overall, four ASVs remained taxonomically unclassified and 75 ASVs were archaeal (0.19% relative abundance). Ubiquitous phyla included Proteobacteria (29%), Acidobacteriota (19%), Verrucomicrobiota (13%), Bacteroidota (11%), Actinobacteriota (8%), Chloroflexi (6%), Gemmatimonadota (4%), and Patescibacteria (3%), as well as rarer taxa such as Firmicutes (0.7%), Bdellovibrionata (0.3%), and Cyanobacteria (0.1%), relative abundance across all samples indicated in brackets (**Figure 1**). While no single ASV was shared by all samples, the 11 most abundant ASVs with overall > 1% of all counts depicted 20% of total relative abundance (**Figure 2**). The overall most dominant identified genera belonged to the order Burkholderiales (Comamonadaceae family: *Polaromonas* and *Rhodoferax,* Gallionellaceae family: *Sideroxydans*, and Hydrogenophilaceae family: *Thiobacillus*), while Chthoniobacterales were dominated by *Candidatus Udaeobacter*, and Pyrinomonadales by *RB41*.

Specifically in permafrost samples (2019: 70–90 cm, 2020: 90–100 cm), Proteobacteria increased with depth, with the order Burkholderiales accounting for up to 60% of reads in the permafrost samples. Within this order, the genus *Polaromonas* was the most abundant ASV in the deepest permafrost sample of 2019 (80–90 cm), where also Firmicutes mainly occurred. *Rhodoferax* had the highest counts in the deepest and only permafrost sample of 2020 (90–100 cm). *Sideroxydans* occurred most in permafrost samples in 2019 (60–90 cm). In 2019, *Thiobacillus* dominated the community most between 60 and 80 cm (**Figure 2**). Actinobacteriota abundance increased with depth but decreased from 2019 to 2020.

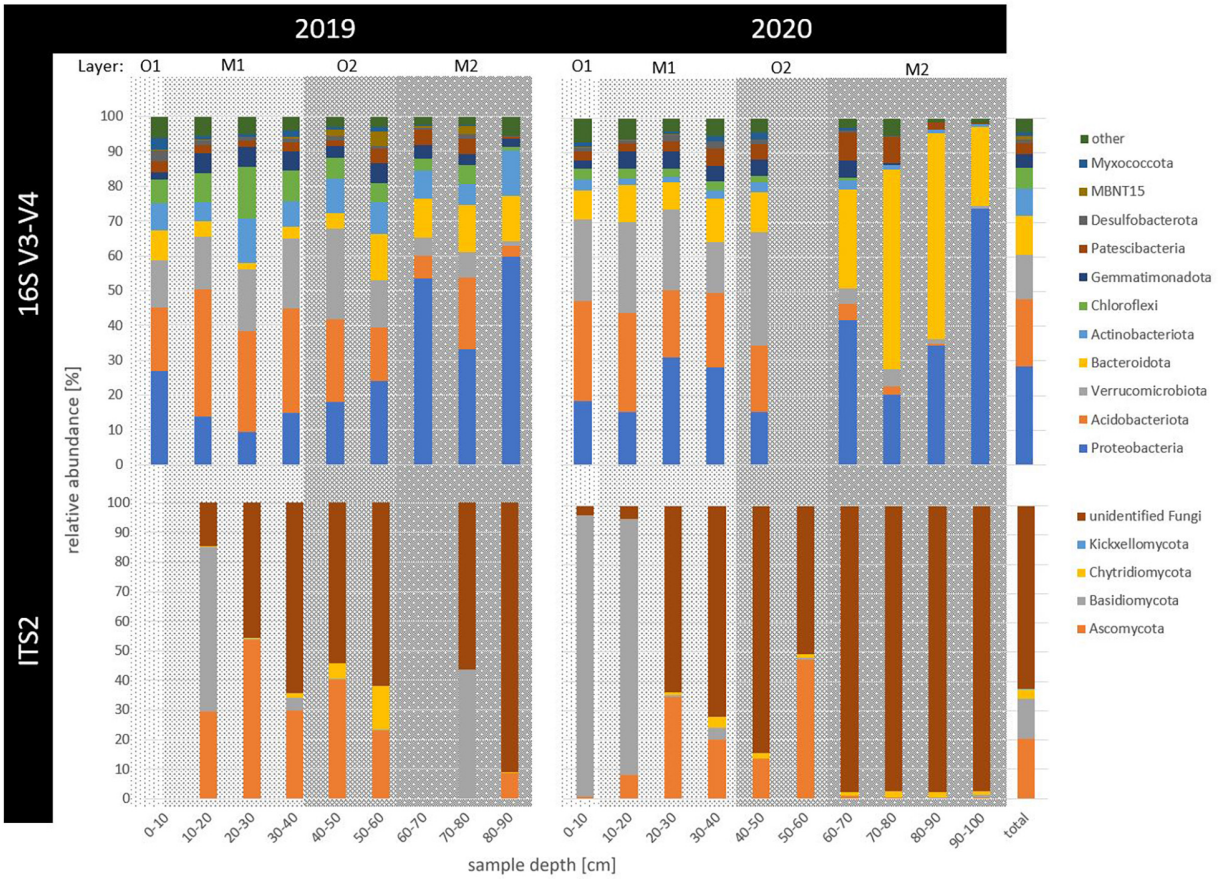

FIGURE 1 | Relative abundance of prokaryotic (16S V3-V4) and fungal (ITS2) taxonomic phyla across 2 years of sampling (2019 and 2020) and in up to 100 cm deep permafrost (sample depth in cm). Phyla with less than 1% of relative abundance, including archaeal counts, were summarized as "other." Horizons were indicated as gray background with the surface organic (O1) in light gray, buried mineral (M1) in medium gray, buried organic (O2) in darker and deepest, ancient mineral horizon (M2) in the darkest gray shading.

Within the freshly thawing permafrost material of the transition zone (2019: 40–70 cm, 2020: 40–90 cm), relative abundance of Bacteroidota increased between 2019 and 2020, driven mainly by the Sphingobacteriales order in 60–90 cm deep samples. The phyla MBNT15 and Patescibacteria also increased in the transition zone samples of 40–80 cm in 2019 and 60–80 cm in 2020, respectively. Verrucomicrobiota were present throughout the thawed samples of active layer and transition zone. In both years, the genus *Cand. Udaeobacter* was found both in the thawed mineral top and buried organic horizons until 60 cm depth (M1 and O2). Here, Acidobacteriota were also abundant and most represented by the orders 11–24 and Pyrinomonadales, which were most abundant between 10 and 50 cm depth, decreasing toward 2020. For the latter order, the genus *RB41* was among the highest counting ASVs in thawed soils in 2019, only seconded by *Cand. Udaeobacter*. Similarly, Chloroflexi also decreased in relative abundance with depth and time. Gemmatimonadota, Desulfobacterota, and Myxococcota abundance was rather constant throughout the samples, while the latter two had their maximal abundance in the surface organic horizon.

The indval analysis indicated the correlation of phyla with categorical environmental parameters (**Supplementary Figure 6**). Significantly more taxa were dominant in 2019, and only a few increased toward 2020, including Bacteroidota, Bdellovibrionota, and Fibrobacterota. These also correlated with increased abundance in the transition zone, together with MBNT15 and Patescibacteria, compared to other layers. While most taxa dominated the active layer (AL), Firmicutes indicated the strongest signal for significant abundance in permafrost samples. In contrast, each horizon had numerous relatively abundant taxa, without clearly outstanding trends (although Fibrobacterota and Firmicutes had their highest values in the deepest mineral horizon compared to other horizons; **Supplementary Figure 6**).

Archaeal relative abundance was marginal and even absent in permafrost samples. Phyla accounting for at least 1% of all archaeal counts included Crenarchaeota (59%), Nanoarchaeota (25%), Halobacterota (9%), Euryarchaeota (5%), and Thermoplasmatota (2% relative abundance within only archaeal ASVs across all samples). The two highest counting archaeal ASVs were assigned as Bathyarchaeia

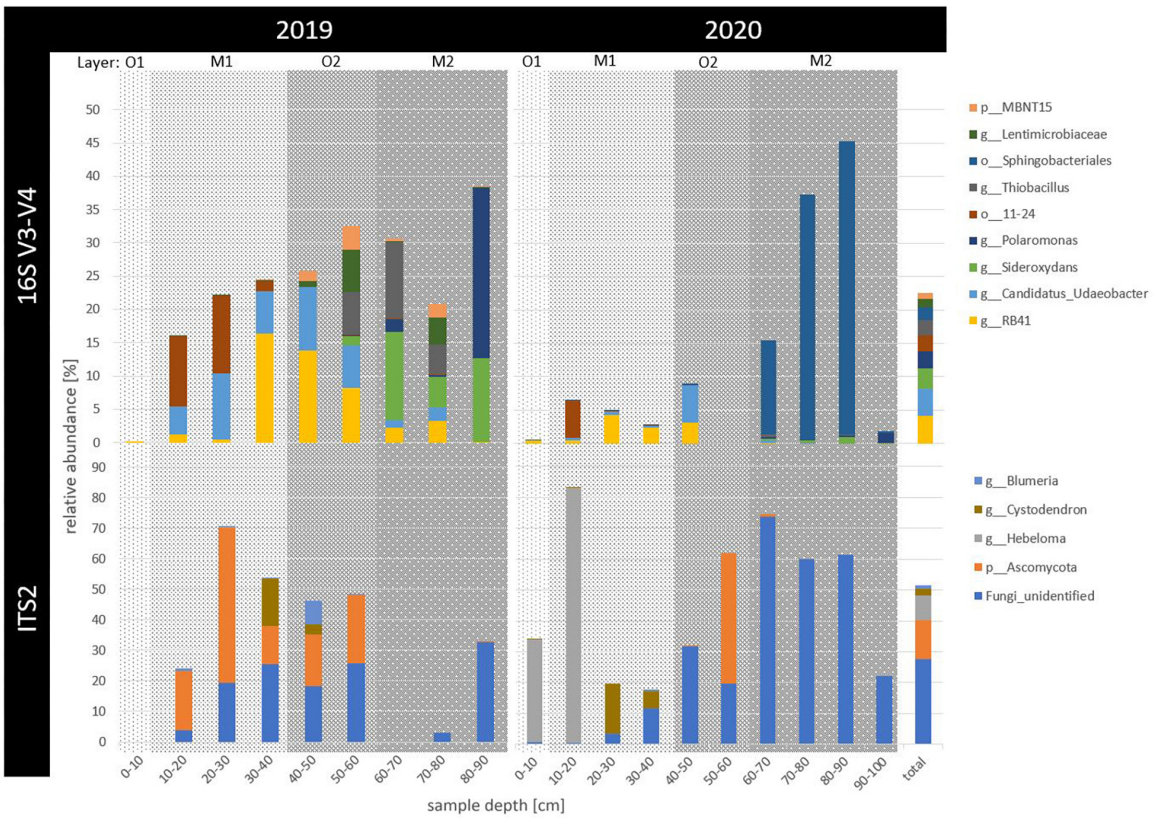

FIGURE 2 | Relative abundance of the most abundant prokaryotic (16S V3-V4) and fungal (ITS2) ASVs across 2 years of sampling (2019 and 2020) and in up to 100 cm deep permafrost (sample depth in cm). Only ASVs with more than 1% relative abundance across all samples, including archaeal counts, were summarized to the highest taxonomic level possible. That taxonomic level is indicated as phylum (p_), order (o_), or genus (g_). Horizons were indicated as gray background with the surface organic (O1) in light gray, buried mineral (M1) in medium gray, buried organic (O2) in darker and deepest, ancient mineral horizon (M2) in the darkest gray shading.

on 39% of all archaeal counts. The phyla Nanoarchaeota, Euryarchaeota, and Thermoplasmata each were driven mainly by one order, Woesearchaeales (25%), and the methanogenic Methanobacteriales (5%) and Methanomassiliicoccales (2%), respectively. Most of the Halobacterota abundance was also driven by other methanogenic taxa, such as the orders Methanomicrobiales and Methanosarcinales (7 and 2% total relative abundance).

Fungal Community Composition and Abundance

Across 17 samples, raw ITS2 sequences had 1,543,256 counts, revealing 1,624 fungal ASVs after filtering, of which 63% remained unclassified. Only four phyla were classified as Ascomycota (21%), Basidiomycota (14%), Chytridiomycota (3%), and Kickxellomycota (< 0.005% relative abundance across all samples). No phylum appeared across all samples (**Figure 1**). For Basidiomycota, the genus *Hebeloma* was the most abundant in the first 20 cm in both years, accounting for 83% of the counts in the 10–20 cm deep sample in 2020 and 9% of all fungal counts. Ascomycota were particularly dominated between 0 and 50 cm depth by the genera *Cystodendron*, *Blumeria,* and *Neobulgaria*, accounting for 7% of all fungal counts (**Figure 2**). Most fungi within the permafrost samples remained unclassified.

Microbial Diversity Along a Spatiotemporal Gradient

Sequence statistics and alpha diversity indices were calculated and are summarized in **Supplementary Table 1**. As all alpha diversity indices indicated similar trends, we utilized Faith's PD for all downstream analyses. The prokaryotic alpha diversity (overall Faith's PD = 181 ± 91) consistently decreased with depth in 2019, while in 2020, very high diversity was maintained throughout the whole active layer and peaked in 40–50 cm deep samples before decreasing with depth (**Figure 3**). In contrast, fungal phylogenetic alpha diversity was comparatively lower (overall Faith's PD = 54 ± 24), peaking in both years at 30–40 cm depth.

Multiple Drivers of Prokaryotic Diversity

Alpha diversity was significantly negatively correlated with soil age (ANOVA, $P = 0.0002$) and pH (ANOVA, $P = 0.0057$). Pairwise covariation revealed no significant differences between years, while diversity across all layers was significantly different

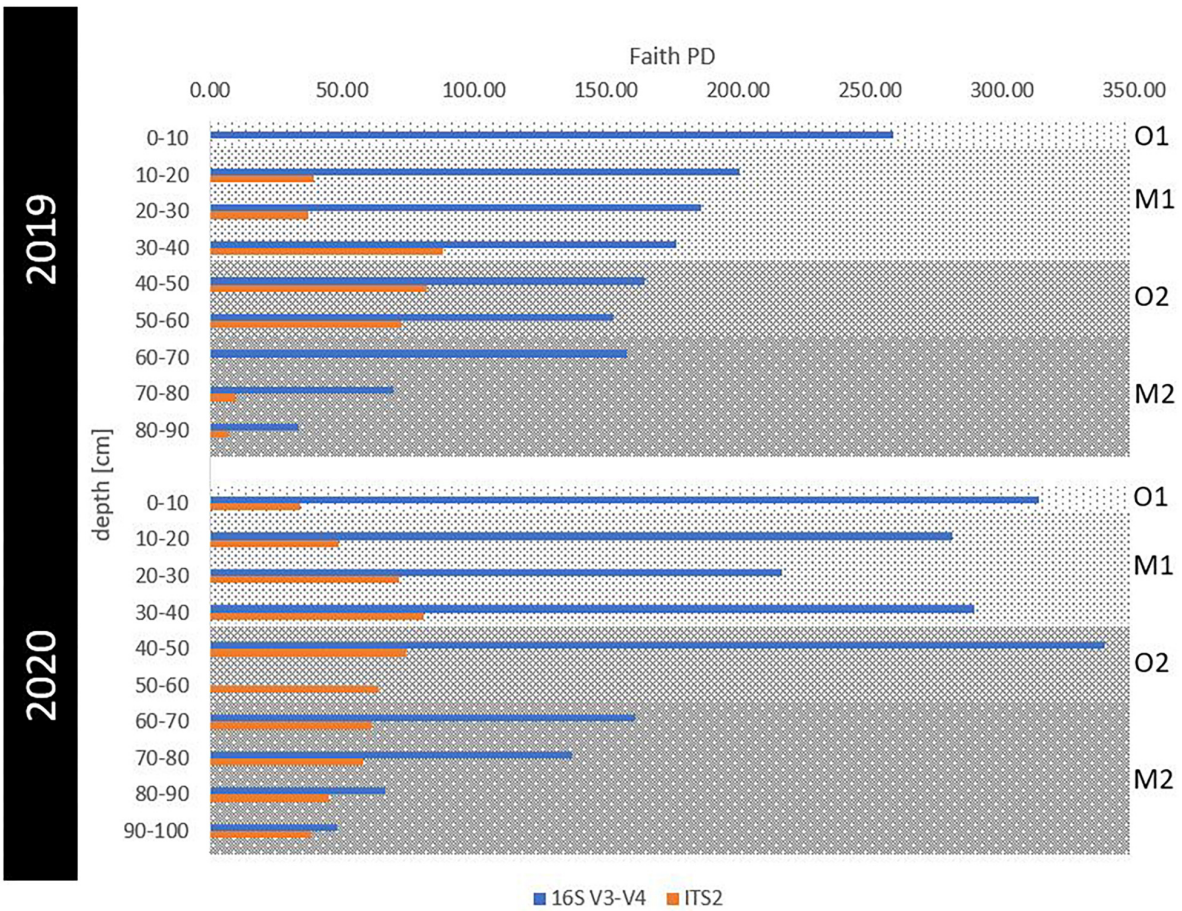

FIGURE 3 | Phylogenetic alpha diversity for prokaryotic and fungal communities in eroding permafrost samples. The Faith's PD index was visualized for 16S (blue) and ITS (orange) sequences per 10-cm depth intervals for 2019 and 2020. Horizons were indicated as gray background with the surface organic (O1) in light gray, buried mineral (M1) in medium gray, buried organic (O2) in darker and deepest, ancient mineral horizon (M2) in the darkest gray shading.

($P < 0.05$) (**Table 2**). As for the horizons, the deepest mineral horizon community was significantly different from all other horizons ($P < 0.05$). The beta diversity was significantly correlated with age (PERMANOVA, $P = 0.001$), horizon (PERMANOVA, $P = 0.001$), layer (PERMANOVA, $P = 0.005$), and year (PERMANOVA, $P = 0.026$). Soil moisture and organic matter content had no significant correlations (**Table 2**). PCoA ordination revealed that 49, 17, and 9% of all variation could be explained by the first three axes, totaling 75% (**Figure 4**). Samples from the top 0–60 cm horizons were visually separated from samples of the deepest horizon M2, the latter also considerably older (> 22,000 years).

Trends in Fungal Diversity Changes

Fungal alpha diversity was also negatively correlated with age (ANOVA, $P = 0.0403$) and pH (ANOVA, $P = 0.0305$). Significant differences for alpha diversity stemmed from the pairwise covariation between the transition zone and permafrost samples ($P = 0.0167$) and similarly between the two deepest horizons O2 and M2 ($P = 0.0143$, **Table 2**). Beta diversity was significantly correlated with changes of age (PERMANOVA, $P = 0.007$), layer (PERMANOVA, $P = 0.017$), horizon (PERMANOVA, $P = 0.021$), and year (PERMANOVA, $P = 0.04$). The variability of the fungal community could be explained by the first three PCoA axes with 32, 21, and 16%, respectively, although no visual separation by any environmental factor became apparent (**Supplementary Figure 5**).

DISCUSSION

Investigating microbial community responses to abrupt permafrost thawing processes is crucial to advance our understanding of ecological responses to global warming, yet it is greatly understudied (Edwards et al., 2020). There is an urgent need for linking microorganism abundances in response to increasing bioavailability with their potential to metabolize ancient permafrost carbon stocks (Turetsky et al., 2020). The findings presented enhance our understanding of cryophilic microbial biodiversity under *in situ* thermal stress within the first 2 years of an abruptly eroding thermal erosion gully in Northeast Greenland. We showed that both permafrost prokaryotes and

TABLE 2 | Significance tests between prokaryotic (16S) and fungal (ITS) diversity indices and soil parameters.

Parameters	Covariation	16S		ITS	
		α	β	α	β
^{14}C		−0.77***	0.67***	−0.50**	0.28**
pH		−0.62**	0.05	−0.53*	0.04
Year		1.03	2.92*	0.15	2.29*
Layer		10.5**	3.66**	6.36*	2.14*
	AL–PF	6.00*		3.75	
	AL–TZ	5.36*		0.69	
	TZ–PF	4.69*		5.73*	
Horizon		11.8**	4.1***	8.51*	1.9*
	M1–M2	9.00**		3.49	
	M1–O1	2.14		2.33	
	M1–O2	0.32		1.29	
	M2–O1	4.00*		0.09	
	M2–O2	4.27*		6.00*	
	O1–O2	0.33		2	
H_2O		0.46	0.11	0.37	0.14
SOM		0.34	0.1	0.37	0.14

*Significance between alpha (α) or beta (β) diversity with soil parameters was tested with linear two-way ANOVA on 999 permutations. For pairwise covariation between layers (AL, TZ, and PF) and horizons (O1, M1, O2, M2), two-sided Kruskal–Wallis test produced the statistic H for alpha diversity values. Significance with indicated with *P < 0.05, **P < 0.01, ***P < 0.005, insignificant pairs were omitted.*

fungi respond strongly to vertical variation within the quickly eroding permafrost habitat. Most significantly, stratification of the microbial community along organic and mineral horizons was evident, and the thaw state and the age of the soil explained the changes in diversity within the community better than other deterministic soil parameters, such as pH, organic matter content, and moisture.

Changes in Prokaryotic Community Are Driven by Inhabited Horizons

Community Composition

We found Proteobacteria, Verrucomicrobiota, Bacteroidota, Actinobacteriota, Gemmatimonadota, and Acidobacteriota to be ubiquitous across all samples, in agreement with recently reviewed abundances across several Arctic sites (Malard et al., 2019). To our knowledge, only two amplicon sequencing studies have investigated the permafrost soil community in Northeast Greenland (Ganzert et al., 2014; Gittel et al., 2014b). Despite such closeness to their sampling site, we observed an overall slightly higher relative abundance of Verrucomicrobiota, Bacteroidota, and Patescibacteria as well as fewer Actinobacteriota. This might be explained by the erosion characteristics of our site, as certain abundant taxa in our study particularly occurred in the freshly thawed transition zone. Defining horizons as consisting of an organic top, underlying mineral buried topsoil, and deeper buried horizons, two studies also found a significant correlation between the horizon and microbial diversity (Frank-Fahle et al., 2014; Gittel et al., 2014b). Surprisingly, per horizon, our prokaryotic alpha diversity measurements were 1.5–2 times higher compared to their data. We hypothesized this to be simply based on varying methodological approaches, as we used ASVs instead of OTUs, increasing the number of unique sequences, which only a few permafrost erosion studies did previously (Doherty et al., 2020; Holm et al., 2020; Kirkwood et al., 2021).

The Permafrost Layer

At our site, several Proteobacteria classes were abundant throughout the entire profile, in agreement with previous studies (Frank-Fahle et al., 2014; Gittel et al., 2014a; Deng et al., 2015; Hultman et al., 2015; Müller et al., 2018), and explained by the metabolic diversity of those classes (Koyama et al., 2014). Abundance of Proteobacteria increased toward permafrost depths, mainly governed by the copiotrophic family

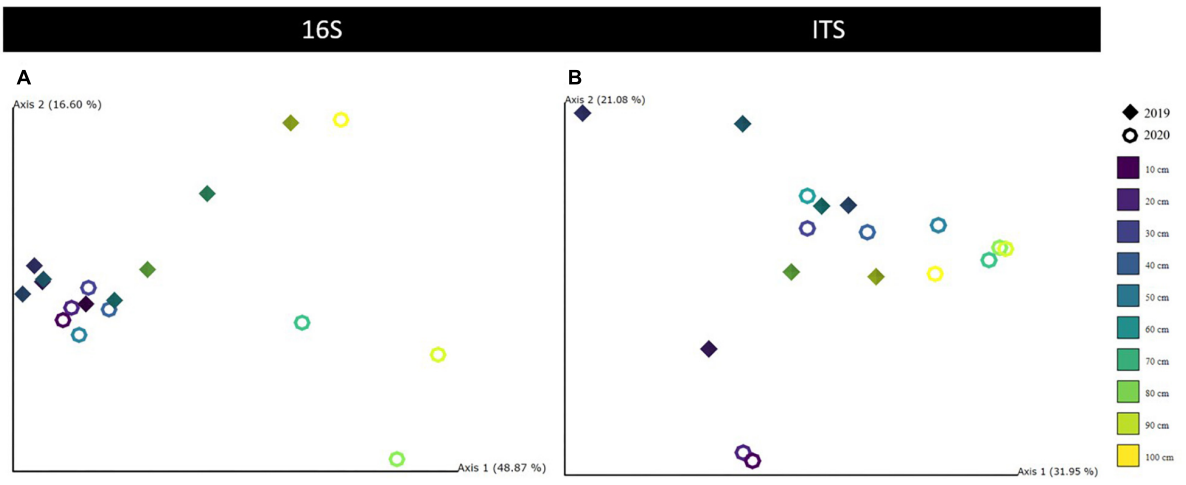

FIGURE 4 | Principal coordinate analysis (PCoA) results were visualized with QIIME2 Emperor plots for 16S **(A)** and ITS **(B)** Unifrac distances along the first two axes, each depicting the variability explained in %. Diamonds are used for 2019 and rings for 2020 measurements. Different colors indicate the 10-cm depth horizons. All other soil parameters were tested and summarized in **Supplementary Figure 2** (16S) and **Supplementary Figure 3** (ITS).

Comamonadaceae formerly documented as spore formers that grew rapidly with a fresh input of labile nutrients (Fierer et al., 2007). This may explain the high relative abundance of the genera *Polaromonas*, *Sideroxydans*, *Thiobacillus,* and *Rhodoferax* in our freshly thawing permafrost samples. *Polaromonas* exclusively occurred in permafrost samples, in agreement with former studies that found this aerobic psychrophile in glacial ice and sediment environments (Darcy et al., 2011). The Gallionellaceae and Comamonadaceae families had formerly been found to dominate in several millennia-old permafrost samples (Willerslev et al., 2004; Altshuler et al., 2019; Burkert et al., 2019), confirming these taxa as part of our ancient permafrost community. In the 2019 permafrost, we found a significant abundance of *Sideroxydans*, a chemolithotroph, formerly found in peatland soils (Lipson et al., 2010; Margesin and Collins, 2019). In anaerobic subglacial lake water of Antarctica, *Polaromonas* and *Sideroxydans* were found to co-occur (Margesin, 2017).

Anaerobic degraders in permafrost were documented to include Actinobacteriota, Firmicutes, and Proteobacteria (Willerslev et al., 2004; Yergeau et al., 2010; Jansson and Taş, 2014). While in our study, Proteobacteria clearly dominated deeper mineral soils, Actinobacteriota were found mainly in 2019, but then especially in the deepest, still intact permafrost sample. Only here, also Firmicutes occurred, where they had the strongest trend as indicator species in our study. They have been found to be extremophiles persisting in intact permafrost soils (Frank-Fahle et al., 2014; Gittel et al., 2014a,b; Deng et al., 2015; Monteux et al., 2018; Tripathi et al., 2018; Burkert et al., 2019) due to their endospore-formation. We note that Firmicutes spores might be harder to extract DNA from and thereby sequence, and their higher relative abundance potentially does not reflect *in situ* intact permafrost abundance, but a response to the storage conditions.

Often, Chloroflexi taxa endemic to permafrost are not classified in databases (Jansson and Taş, 2014). We found Chloroflexi present throughout the entire soil profile, in contrast to other studies that found this taxon to dominate in permafrost samples (Frank-Fahle et al., 2014; Tuorto et al., 2014; Hultman et al., 2015). Only one other study confirmed the occurrence of Chloroflexi at surface depths (Gittel et al., 2014b). As this study was the only one within the same study area, the local biogeography could explain this trend (Gittel et al., 2014b).

The Thawed Soils of Transition Zone and Active Layer

Previous studies have shown the most significant changes in the microbial community to take place during long-term thawing (Elberling et al., 2013; Schostag et al., 2019). The transition from permafrost to thawed soil was among the most significant drivers of the differences in community composition in our study. This trend was typified by the increased abundance of Bacteroidota, in agreement with the literature (Frank-Fahle et al., 2014; Coolen and Orsi, 2015; Deng et al., 2015; Burkert et al., 2019), also found at the upper permafrost limit and during thawing (Müller et al., 2018; Tripathi et al., 2018). Bacteroidia were strongly represented by the Sphingobacteriales order, which include psychrotolerant non-spore-forming soil genera and lichen symbionts (Cernava et al., 2017).

In agreement with previous studies, we found Actinobacteriota (Rivkina et al., 2007; Frank-Fahle et al., 2014; Tuorto et al., 2014; Hultman et al., 2015; Mackelprang et al., 2017), Gemmatimonadota (Gittel et al., 2014a; Schostag et al., 2015; Müller et al., 2018; Altshuler et al., 2019), and Verrucomicrobiota (Gittel et al., 2014a; Deng et al., 2015; Schostag et al., 2015; Müller et al., 2018; Tripathi et al., 2018; Altshuler et al., 2019) throughout the thawed active and transition zone samples. Verrucomicrobiota are often among the first responders to seasonal thawing (Malard and Pearce, 2018) and although previously seen mainly in surface samples (Deng et al., 2015), they were the most abundant until 60 cm depth in our study. For this phylum, a relative abundance of over 20% per sample in 2020 was rarely observed previously (Schostag et al., 2015), but confirmed by the one earlier study performed in this valley (Gittel et al., 2014b). Notably, the primers used in this study have been known to indicate a slightly higher Verrucomicrobiota abundance (Takahashi et al., 2014).

Based on the utilized monitoring data, local fluctuation of surface temperature indicates almost a 60°C range. The annual freeze-up in fall supplies dead organic matter, which serves as nutrients in the next spring, when increasing temperatures and soil moisture increases after snowmelt enable microbial growth again (Jansson and Taş, 2014; Malard and Pearce, 2018). Taxa particularly adapted to this extreme and less thermally stable environment of the active layer include Acidobacteriota, Alphaproteobacteria, and Verrucomicrobiota (Deng et al., 2015; Altshuler et al., 2019). In our study, Acidobacteriota abundance decreased significantly with depth, which has been reported earlier (Frank-Fahle et al., 2014; Gittel et al., 2014b; Monteux et al., 2018; Tripathi et al., 2018; Altshuler et al., 2019), and depends on a low pH (Männistö et al., 2007; Ganzert et al., 2011). In line with this, we found that even the small changes of pH in 2019 were found significant as driver of alpha diversity in our study. Willms et al. (2021) found *Cand. Udaeobacter* as one of the most abundant soil taxa, especially in strongly acidic soils. As one of the most abundant taxa in our study, the overall acidic pH might support its high relative abundance. Malard et al. (2019) reviewed Blastocatellia abundance similar to those in our study in acidoneutral to alkaline soils, which could explain their higher abundance in our more alkaline M2 samples in 2019 as opposed to 2020. The main order in phylum in our study were Pyrinomonadales, which formerly have been observed as phototrophic lichen symbionts and are part of a rather specialist soil community. An acidoneutral soil pH, such as in most Greenland studies, was found to supply optimum microbial growth conditions (Rousk et al., 2010), which could explain the particularly high alpha diversity found across most of our samples.

Archaeal Abundance Within Degrading Permafrost

Archaeal abundance was marginal within the prokaryotic community composition, which is in agreement with previous findings (Yergeau et al., 2010; Rivkina et al., 2016; Mondav et al., 2017; Müller et al., 2018; Hough et al., 2020). Although we selected primers that were designed to create the least possible bias toward bacterial or archaeal taxa

(Takahashi et al., 2014), they still have previously been shown to be less efficient for archaea (Parada et al., 2016). The dominance of the anaerobic, metabolically flexible Bathyarchaeia has previously been documented (Lazar et al., 2016; Xiang et al., 2017). In our study, this taxon occurred at soil depths with maximum soil moisture, supplying saturated conditions. Woesearchaeota have been documented as syntrophic partners to methanogens (Liu et al., 2018; Tveit et al., 2020), which could justify their abundance throughout the soil column, where we found several methanogenic orders, such as Methanomicrobiales, Methanosarcinales, Methanomassiliicoccales, and Methanobacteriales. Especially Methanomicrobiales have been documented in thawing permafrost (Deng et al., 2015; Wei et al., 2018). Both Euryarchaeota classes Methanomicrobiales and Methanobacteriales have been found in Greenlandic active layer soils before (Malard and Pearce, 2018). The nitrifying Nitrososphaerales oxidize ammonia, possibly explaining their abundance within the active layer where deeply rooting vegetation supplies nitrogen compounds (DeLong et al., 2014).

Different Fungal Community in Ancient Permafrost

In our study, fungal diversity was generally lower than prokaryotic diversity and most driven by the difference of age and the thaw state of the soil. We could confirm representation of Ascomycota by Leotiomycetes and Basidiomycota by Agarimycetes in alignment with Gittel et al. (2014b), which might indicate a local pattern specific to Northeast Greenland, or even the Zackenberg valley. Ascomycota dominated our fungal community, in agreement with previous findings (Malard and Pearce, 2018; Margesin and Collins, 2019). The identified Ascomycete genera *Blumeria*, *Cystodendron*, and *Neobulgaria* of the Helotiales order are recognized as saprotrophic and plant pathogens, growing on decaying organic matter (Ekanayaka et al., 2019), as abundant in the buried organic horizon, where they occurred predominantly. Fungal abundance was commonly reported with decrease with depth in Arctic soils and is linked to a saprotrophic metabolism and the breakdown of labile carbon compounds (Gittel et al., 2014a).

Microbial Diversity Is Driven by the Thawing of Soil Horizons

Prokaryotic alpha diversity in our study was twice as high as reported in other studies (Chu et al., 2010; Tveit et al., 2013; Frank-Fahle et al., 2014; Gittel et al., 2014b; Schostag et al., 2019), which did not detect a deeper second peak, as we did at the 40–50 cm depth in 2020. Faith's PD values both for prokaryotic and fungal communities were comparable to one of the few permafrost studies that previously employed this index (Feng et al., 2020). A decrease of alpha diversity with depth and age has been previously shown (Tripathi et al., 2018; Burkert et al., 2019). Interestingly, utilizing relic permafrost DNA, if amplifiable, was earlier found to lead to up to 45% increased diversity estimates (Fierer, 2017). Age was a significant driver of biodiversity for both kingdoms. The clustering of samples from the youngest samples opposed to higher dispersal in samples older than 3,500 years has also been described before (Tripathi et al., 2018), and aligns with the significant age effect on diversity (Mackelprang et al., 2017; Burkert et al., 2019).

For prokaryotic diversity, each layer depicted a significantly different community, but clearly only the most significant one, the oldest horizon, was visually split in the community, as indicated by the PCoA plots. For fungi, diversity changes were driven significantly similarly driven by the age effect and the transition to the permafrost layer, indicating a strong role for the permafrost table depth of the soil rather than organic or mineral content. Furthermore, current work has suggested that physical barriers to microbial dispersal were the main driver of changes in the permafrost soil microbiomes (Bottos et al., 2018). We think the separation of the community in the study by horizons most closely depicts these barriers of habitat, hence its significance for explaining changes in community composition.

Previous studies confirmed that prokaryotic survival in permafrost is due to the occupation of viable microhabitats, such as brine channels (Gilichinsky et al., 2003) or the ability to adherence to silt or clay particles, whereas the fungal mycelium often form larger networks across rather than within the soil horizons (Jansson and Hofmockel, 2019). Jansson and Hofmockel (2019) found bacteria usually outcompeted fungi as the temperature increased, due to higher growth rates and competition for newly available nutrients. However, for all microorganisms, our results indicate the strongest community shift from frozen to thawed ground.

Copiotrophy in Thawing Permafrost

While nutrient load and lability of organic carbon decreases within intact permafrost soils, the more stress-resilient oligotrophic (slow-growing specialist) microorganisms are usually favored over copiotrophic (fast-growing and generalist) taxa (Fierer et al., 2007). But these ecosystems are facing increased abrupt erosion until the end of this century, both globally (Turetsky et al., 2020) as well as in the Zackenberg Valley (Westermann et al., 2015; Rasmussen et al., 2018). Christensen et al. (2020) have already documented the connection between increasing events of extreme precipitation and summer temperatures, and the permafrost erosion site studied here. Habtewold et al. (2021) recently confirmed a correlation between higher carbon released from warmed Canadian soils and increased abundance of copiotrophic taxa. A former study on Alpine permafrost soil found soil warming to trigger copiotrophy (Perez-Mon et al., 2022), indicating elevated mineralization potential of increasingly bioavailable soil carbon and fresh nutrient inputs after thawing. The taxa that dominated this vulnerable zone of freshly thawed ancient carbon stocks in our study included both the copiotrophic Bacteroidia and MBNT15 phyla, as well as the oligotrophic Chthoniobacterales and Pyrinomonadales (Ho et al., 2017). But Bacteroidia were found as one of the clearest thaw zone indicator phyla and also dominated the 2020 thaw samples with more than half of all counts. They have higher carbon mineralization potential and can adapt their population size better to the available resources. In comparison, their oligotrophic counterpart Acidobacteria might be better adapted to the thermally variable active layer due to their higher stress resilience (Fierer et al., 2007). Interestingly, the abundance

of copiotrophic taxa, such as the Comamonadaceae family and the often as copiotrophic classified Firmicutes even in our deepest permafrost samples (Ho et al., 2017), indicates that the short time period of storage at thawing temperatures initiated a microbial response to the abrupt thaw, both *in situ* and *in vivo*. This, clearly, indicates a need for controlled studies to verify and quantify the metabolic response of ancient permafrost taxa to thaw. Fierer et al. (2007) argued that knowledge about community composition and dominance of oligo- and copiotrophy is a crucial component for improving the prediction permafrost carbon loss due to microbial mineralization (Langille et al., 2013). Due to the quick response to readily available carbon, copiotrophic metabolism has been reviewed before to correlate with carbon release from soil ecosystems (Trivedi et al., 2013; Hurst, 2019).

Limitations and Future Perspectives

Malard et al. (2019) and Tripathi et al. (2019) explored the description of stochastic (random and dispersal) and deterministic (biotic and abiotic selection) drivers of community composition in permafrost soils. They found top-soils to be dominated by deterministic processes, while deeper soils were controlled by stochastic processes. As permafrost soils are highly heterogenic, an assessment of these processes would help to confirm the significance of changes in this vulnerable ecosystem on its microbiome and should be considered in future efforts. Still, we can interpret from our community that higher abundance of specialist taxa in the permafrost samples indicates an importance of deterministic processes to govern, while these soils are intact.

In Arctic remote sampling sites, both core sampling techniques to reduce contamination (Bang-Andreasen et al., 2017) and temperature stable sampling and transport oppose great challenges. Due to the remote sampling site and long transport route, we decided to store the samples at 4°C and extract the DNA as fast as possible. Greater biological perturbation could have been induced with freezing and thawing. While DNA sequencing is a proxy for the total microbial community composition and not just the viable one—including taxa that might have ceased or increased abundance in response to the storage induced thaw—viability assays, incorporation, or transcriptomic studies are needed to quantify the copiotrophic taxa trends visible in our samples (Mackelprang et al., 2021). Still, taxa abundance in the deepest and most intact permafrost sample in 2019, as opposed to other depths and years, indicated the potential of depicting close to *in situ* conditions in ancient permafrost before and while thawing.

CONCLUSION

While investigating the impact of abrupt permafrost thawing on the microbial community composition is usually analyzed under constraining laboratory incubations, we were the first to study an *in situ* temporal natural thaw gradient in Northeast Greenland and to show the swift adaptation potential of the ancient permafrost microbiome to abrupt erosion on a 2-year timescale and a high-resolution sampling scale, which has been rarely done previously. We found the soil age and horizons to be the most significant driver of microbial community. Furthermore, year of sampling after erosion and permafrost thaw state significantly drove microbial community composition changes with depth. Alpha diversity decreased with increasing depth, pH, and especially soil age. We were able to differentiate dominant prokaryotic taxa within the active layer, freshly thawed transition zone, and ancient permafrost samples. Finally, the transition zone, especially with increasing erosion, gave rise to copiotrophic taxa, such as Bacteroidota, indicating potential impact on the release of permafrost carbon through microbial carbon mineralization. While the universal primers we utilized were sufficient for the scope of our study, a combination with viability assays or transcriptomic studies of ancient taxa within the deepest permafrost samples could significantly improve results. These approaches are still rare across the Arctic, as recently pointed out by Mackelprang et al. (2021). They could provide a higher level of confidence about composition changes of ancient microbial taxa under thawing conditions. We recommend that future research efforts focus on the differentiation of living from dead ancient microbiomes, as well as transcriptomic approaches to elucidate how the complex process of permafrost thaw ends dormancy stages and initiates metabolic activity. These approaches could empirically elaborate the connection between microbial community changes and the globally changing climate.

DATA AVAILABILITY STATEMENT

The datasets produced and analyzed for this study can be found in the online repository: https://www.ncbi.nlm.nih.gov/bioproject/PRJNA769429, with the accession number PRJNA769429.

AUTHOR CONTRIBUTIONS

MS, TC, AZ, and CJ conceptualized and designed the study. MS and AZ organized the databases. MS performed the laboratory work, statistical analysis, and wrote the first draft of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

FUNDING

This work was entirely funded by the Faculty of Technical Sciences (Aarhus University) and included costs for fieldwork, laboratory materials, sequencing services, and publication fees.

ACKNOWLEDGMENTS

Data from the Greenland Ecosystem Monitoring Programme were provided by the Department of Ecoscience, Aarhus University, Denmark in collaboration with Department of Geosciences and Natural Resource Management, Copenhagen University, Denmark and Asiaq—Greenland Survey, Nuuk,

Greenland. We also thank the Zackenberg Research Station and its field assistants for smooth sample handling, Lisa Bröder, Julien Fouche, and Catherine Hirst for contributing to the discussion of soil horizons, and Tina Thane and Tanja Begovic for advice during laboratory work.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2022.787146/full#supplementary-material

REFERENCES

Altshuler, I., Hamel, J., Turney, S., Magnuson, E., Lévesque, R., Greer, C. W., et al. (2019). Species interactions and distinct microbial communities in high Arctic permafrost affected cryosols are associated with the CH4 and CO2 gas fluxes. *Environ. Microbiol.* 21, 3711–3727. doi: 10.1111/1462-2920.14715

AMAP (2019). *AMAP Climate Change Update 2019: an Update to Key Findings of Snow, Water, Ice and Permafrost in the Arctic (SWIPA) 2017*. Norway: Arctic Monitoring and Assessment Programme (AMAP), 12.

AMAP (2021). *Arctic Climate Change Update 2021: key Trends and Impacts. Summary for Policy-Makers*. Norway: Arctic Monitoring and Assessment Programme (AMAP). 16.

Bang-Andreasen, T., Schostag, M., Priemé, A., Elberling, B., and Jacobsen, C. S. (2017). Potential microbial contamination during sampling of permafrost soil assessed by tracers. . *Sci. Rep.* 7, 1–11. doi: 10.1038/srep43338

Bay, C. (1998). *Vegetation Mapping of the Zackenberg Valley, Northeast Greenland*. Copenhagen: Danish Polar Center & Botanical Museum, University of Copenhagen.

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., et al. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* 37, 852–857. doi: 10.1038/s41587-019-0209-9

Bottos, E. M., Kennedy, D. W., Romero, E. B., Fansler, S. J., Brown, J. M., Bramer, L. M., et al. (2018). Dispersal limitation and thermodynamic constraints govern spatial structure of permafrost microbial communities. *FEMS Microbiol. Ecol.* 94, 1–14. doi: 10.1093/femsec/fiy110

Burkert, A., Douglas, T. A., Waldrop, M. P., and Mackelprang, R. (2019). Changes in the Active, Dead, and Dormant Microbial Community Structure across a Pleistocene Permafrost Chronosequence. *Appl. Environ. Microbiol.* 85, 1–16. doi: 10.1128/AEM.02646-18

Callahan, B. J., McMurdie, P. J., and Holmes, S. P. (2017). Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME J.* 11, 2639–2643. doi: 10.1038/ismej.2017.119

Cavicchioli, R., Bakken, L. R., Baylis, M., Foreman, C. M., Karl, D. M., Koskella, B., et al. (2019). Scientists' warning to humanity: microorganisms and climate change. *Nat. Rev. Microbiol.* 17, 569–586. doi: 10.1038/s41579-019-0222-5

Cernava, T., Erlacher, A., Aschenbrenner, I. A., Krug, L., Lassek, C., Riedel, K., et al. (2017). Deciphering functional diversification within the lichen microbiota by meta-omics. *Microbiome* 5:82. doi: 10.1186/s40168-017-0303-5

Christensen, T. R., Lund, M., Skov, K., Abermann, J., Scheller, J., Scheel, M., et al. (2020). Multiple Ecosystem Effects of Extreme Weather Events in the Arctic. *Ecosystems* 24, 122–136. doi: 10.1007/s10021-020-00507-6

Christiansen, H. H. (1999). Active Layer Monitoring in two Greenlandic Permafrost Areas: zackenberg and Disko Island. *Danish J. Geogr.* 99, 117–121.

Christiansen, H. H., Sigsgaard, C., Humlum, O., Rasch, M., and Hansen, B. U. (2008). Permafrost and Periglacial Geomorphology at Zackenberg. *Adv. Ecol. Res.* 40, 151–174. doi: 10.1016/S0065-2504(07)00007-4

Chu, H., Fierer, N., Lauber, C. L., Caporaso, J. G., Knight, R., and Grogan, P. (2010). Soil bacterial diversity in the Arctic is not fundamentally different from that found in other biomes. *Environ. Microbiol.* 12, 2998–3006. doi: 10.1111/j.1462-2920.2010.02277.x

Coolen, M. J. L., and Orsi, W. D. (2015). The transcriptional response of microbial communities in thawing Alaskan permafrost soils. *Front. Microbiol.* 6:197. doi: 10.3389/fmicb.2015.00197

Collins, T., and Margesin, R. (2019). Psychrophilic lifestyles: mechanisms of adaptation and biotechnological tools. *Appl. Microbiol. Biotechnol.* 103, 2857–2871. doi: 10.1007/s00253-019-09659-5

Darcy, J. L., Lynch, R. C., King, A. J., Robeson, M. S., and Schmidt, S. K. (2011). Global distribution of Polaromonas phylotypes – evidence for a highly successful dispersal capacity. *PLoS One* 6:e23742. doi: 10.1371/journal.pone.0023742

DeLong, E. F., Lory, S., Stackebrandt, E., and Thompson, F. (eds) (2014). *The Prokaryotes – Other Major Lineages of Bacteria and the Archaea*. Germany: Springer. doi: 10.1007/978-3-642-38954-2_329

Deng, J., Yungfu, G., Zhang, J., Xu, K., Qin, Y., Yuan, M., et al. (2015). Shifts of tundra bacterial and archaeal communities along a permafrost thaw gradient in Alaska. *Mol. Ecol.* 24, 222–234. doi: 10.1111/mec.13015

Doherty, S. J., Barbato, R. A., Grandy, A. S., Thomas, W. K., Monteux, S., Dorrepaal, E., et al. (2020). The Transition From Stochastic to Deterministic Bacterial Community Assembly During Permafrost Thaw Succession. *Front. Microbiol.* 11:596589. doi: 10.3389/fmicb.2020.596589

Dufrêne, M., and Legendre, P. (1997). Species assemblages and indicator species: the need for a flexible asymmetrical approach. *Ecol. Monogr.* 67, 345–366. doi: 10.2307/2963459

Edwards, A., Cameron, K. A., Cook, J. M., Debbonaire, A. R., Furness, E., Hay, M. C., et al. (2020). Microbial genomics amidst the Arctic crisis. *Microb. Genom.* 6:e000375. doi: 10.1099/mgen.0.000375

Ekanayaka, A. H., Hyde, K. D., Gentekaki, E., McKenzie, E. H. C., Zhao, Q., Bulgakov, T. S., et al. (2019). Preliminary classification of Leotiomycetes. *Mycosphere* 10, 310–489. doi: 10.5943/mycosphere/10/1/7

Elberling, B., Christiansen, H. H., and Hansen, B. U. (2010). High nitrous oxide production from thawing permafrost. *Nat. Geosci.* 3, 332–335. doi: 10.1038/ngeo803

Elberling, B., Michelsen, A., Schädel, C., Schuur, E. A. G., Christiansen, H. H., Berg, L., et al. (2013). Long-term CO2 production following permafrost thaw. *Nat. Clim. Chang.* 3, 890–894. doi: 10.1038/nclimate1955

Elberling, B., Tamstorf, M. P., Michelsen, A., Arndal, M. F., Sigsgaard, C., Illeris, L., et al. (2008). Soil and Plant Community-Characteristics and Dynamics at Zackenberg. *Adv. Ecol. Res.* 40, 223–248. doi: 10.1016/S0065-2504(07)00010-4

Feng, J., Wang, C., Lei, J., Yang, Y., Yan, Q., Zhou, X., et al. (2020). Warming-induced permafrost thaw exacerbates tundra soil carbon decomposition mediated by microbial community. *Microbiome* 8:3. doi: 10.1186/s40168-019-0778-3

Fierer, N. (2017). Embracing the unknown: disentangling the complexities of the soil microbiome. *Nat. Rev. Microbiol.* 15, 579–590. doi: 10.1038/nrmicro.2017.87

Fierer, N., Bradford, M. A., and Jackson, R. B. (2007). Toward an ecological classification of soil bacteria. *Ecology* 88, 1354–1364. doi: 10.1007/s00209-018-2115-0

Frank-Fahle, B. A., Yergeau, É, Greer, C. W., Lantuit, H., and Wagner, D. (2014). Microbial functional potential and community composition in permafrost-affected soils of the NW Canadian Arctic. *PLoS One* 9:e84761. doi: 10.1371/journal.pone.0084761

Ganzert, L., Bajerski, F., and Wagner, D. (2014). Bacterial community composition and diversity of five different permafrost-affected soils of Northeast Greenland. *FEMS Microbiol. Ecol.* 89, 426–441. doi: 10.1111/1574-6941.12352

Ganzert, L., Lipski, A., Hubberten, H.-W., and Wagner, D. (2011). The impact of different soil parameters on the community structure of dominant bacteria from nine different soils located on Livingston Island, South Shetland Archipelago, Antarctica. *FEMS Microbiol. Ecol.* 76, 476–491. doi: 10.1111/j.1574-6941.2011.01068.x

Gaylarde, C., Ogawa, A., Beech, I., Kowalski, M., and Baptista-Neto, J. A. (2017). Analysis of dark crusts on the church of Nossa Senhora do Carmo in Rio de Janeiro, Brazil, using chemical, microscope and metabarcoding microbial identification techniques. *Int. Biodeterior. Biodegrad.* 117, 60–67. doi: 10.1016/j.ibiod.2016.11.028

Gilichinsky, D., Rivkina, E., Shcherbakova, V., Laurinavichuis, K., and Tiedje, J. (2003). Supercooled water brines within permafrost – An unknown ecological niche for microorganisms: a model for astrobiology. *Astrobiology* 3, 331–341. doi: 10.1089/153110703769016424

Gittel, A., Bárta, J., Kohoutová, I., Mikutta, R., Owens, S., Gilbert, J., et al. (2014a). Distinct microbial communities associated with buried soils in the Siberian tundra. *ISME J.* 8, 841–853. doi: 10.1038/ismej.2013.219

Gittel, A., Bárta, J., Kohoutová, I., Schnecker, J., Wild, B., Čapek, P., et al. (2014b). Site- and horizon-specific patterns of microbial community structure and enzyme activities in permafrost-affected soils of Greenland. *Front. Microbiol.* 5:541. doi: 10.3389/fmicb.2014.00541

Habtewold, J. Z., Helgason, B. L., Yanni, S. F., Janzen, H. H., Ellert, B. H., and Gregorich, E. G. (2021). Warming effects on the structure of bacterial and fungal communities in diverse soils. *Appl. Soil Ecol.* 163:103973. doi: 10.1016/j.apsoil.2021.103973

Hansen, B. U. L. F., Sigsgaard, C., Rasmussen, L., Cappelen, J., Hinkler, J., Mernild, S. H., et al. (2008). "Present – Day Climate at Zackenberg. *Adv. Ecol. Res.* 40, 111–49. doi: 10.1016/S0065-2504(07)00006-2

Hayes, D. J., Kicklighter, D. W., McGuire, A. D., Chen, M., Zhuang, Q., Yuan, F., et al. (2014). The impacts of recent permafrost thaw on land-atmosphere greenhouse gas exchange. *Environ. Res. Lett.* 9:045005. doi: 10.1088/1748-9326/9/4/045005

Ho, A., Di Lonardo, D. P., and Bodelier, P. L. E. (2017). Revisiting life strategy concepts in environmental microbial ecology. *FEMS Microbiol. Ecol.* 93, 1–14. doi: 10.1093/femsec/fix006

Hollesen, J., Elberling, B., and Jansson, P. E. (2011). Future active layer dynamics and carbon dioxide production from thawing permafrost layers in Northeast Greenland. *Glob. Chang. Biol.* 17, 911–926. doi: 10.1111/j.1365-2486.2010.02256.x

Holm, S., Walz, J., Horn, F., Yang, S., Grigoriev, M. N., Wagner, D., et al. (2020). Methanogenic response to long-term permafrost thaw is determined by paleoenvironment. *FEMS Microbiol. Ecol.* 96, 1–13. doi: 10.1093/femsec/fiaa021

Hough, M., McClure, A., Bolduc, B., Dorrepaal, E., Saleska, S., Klepac-Ceraj, V., et al. (2020). Biotic and Environmental Drivers of Plant Microbiomes Across a Permafrost Thaw Gradient. *Front. Microbiol.* 11:796. doi: 10.3389/fmicb.2020.00796

Hultman, J., Waldrop, M. P., Mackelprang, R., David, M. M., McFarland, J., Blazewicz, S. J., et al. (2015). Multi-omics of permafrost, active layer and thermokarst bog soil microbiomes. *Nature* 521, 208–212. doi: 10.1038/nature14238

Hurst, C. J. (ed.) (2019). *Understanding Terrestrial Microbial Communities*. Berlin: Springer. doi: 10.1007/978-3-10777-2

IPCC (2019). "Global Warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways" in *The Context of Strengthening the Global Response to the Threat of Climate Change, Sustainable Development, and Efforts to Eradicate Poverty*. eds V. Masson-Delmotte, P. Zhai, H.-O. Pörtner, D. Roberts, J. Skea, P. R. Shukla, et al. (Switzerland: IPCC).

IPCC (2021). "Summary for Policymakers" in *Climate Change 2021: the Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Edn. eds V. Masson-Delmotte, P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, et al. (Cambridge: Cambridge University Press).

Jansson, J. K., and Hofmockel, K. S. (2019). Soil microbiomes and climate change. *Nat. Rev. Microbiol.* 18, 35–46. doi: 10.1038/s41579-019-0265-7

Jansson, J. K., and Taş, N. (2014). The microbial ecology of permafrost. *Nat. Rev. Microbiol.* 12, 414–425. doi: 10.1038/nrmicro3262

Johnson, S. S., Hebsgaard, M. B., Christensen, T. R., Mastepanov, M., Nielsen, R., Munch, K., et al. (2007). Ancient bacteria show evidence of DNA repair. *PNAS* 105, 10631–10631. doi: 10.1073/pnas.0710637105

Kirkwood, J. A. H., Roy-Léveillée, P., Mykytczuk, N., Packalen, M., McLaughlin, J., Laframboise, A., et al. (2021). Soil Microbial Community Response to Permafrost Degradation in Palsa Fields of the Hudson Bay Lowlands: implications for Greenhouse Gas Production in a Warming Climate. *Global Biogeochem. Cycles* 35:e2021GB006954. doi: 10.1029/2021GB006954

Koyama, A., Wallenstein, M. D., Simpson, R. T., and Moore, J. C. (2014). Soil bacterial community composition altered by increased nutrient availability in Arctic tundra soils. *Front. Microbiol.* 5:516. doi: 10.3389/fmicb.2014.00516

Langille, M. G. I., Zaneveld, J., Caporaso, J. G., McDonald, D., Knights, D., Reyes, J. A., et al. (2013). Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat. Biotechnol.* 31, 814–821. doi: 10.1038/nbt.2676

Lazar, C. S., Baker, B. J., Seitz, K., Hyde, A. S., Dick, G. J., Hinrichs, K. U., et al. (2016). Genomic evidence for distinct carbon substrate preferences and ecological niches of *Bathyarchaeota* in estuarine sediments. *Environ. Microbiol.* 18, 1200–1211. doi: 10.1111/1462-2920.13142

Levin, I., and Kromer, B. (2004). The tropospheric 14CO2 level in mid-latitudes of the Northern Hemisphere (1959-2003). *Radiocarbon* 46, 1261–1272. doi: 10.2458/azu_js_rc.46.4181

Lim, P. P., Pearce, D. A., Convey, P., Lee, L. S., Chan, K. G., and Tan, G. Y. A. (2020). Effects of freeze-thaw cycles on High Arctic soil bacterial communities. *Polar Sci.* 23:100487. doi: 10.1016/j.polar.2019.100487

Lipson, D. A., Jha, M., Raab, T. K., and Oechel, W. C. (2010). Reduction of iron (III) and humic substances plays a major role in anaerobic respiration in an Arctic peat soil. *J. Geophys. Res.* 115, 1–13. doi: 10.1029/2009JG001147

Liu, X., Li, M., Castelle, C. J., Probst, A. J., Zhou, A., Pan, J., et al. (2018). Insights into the ecology, evolution, and metabolism of the widespread Woesearchaeotal lineages. *Microbiome* 6:102. doi: 10.1186/s40168-018-0488-2

Mackelprang, R., Burkert, A., Haw, M., Mahendrarajah, T., Conaway, C. H., Douglas, T. A., et al. (2017). Microbial survival strategies in ancient permafrost: insights from metagenomics. *ISME J.* 11, 2305–2318. doi: 10.1038/ismej.2017.93

Mackelprang, R., Saleska, S. R., Jacobsen, C. S., Jansson, J. K., and Taş, N. (2016). Permafrost Meta-Omics and Climate Change. *Annu. Rev. Earth Planet. Sci.* 44, 439–462. doi: 10.1146/annurev-earth-060614-105126

Mackelprang, R., Tas, N., and Waldrop, M. (2021). "2 Functional response of microbial communities to permafrost thaw" in *Microbial Life in the Cryosphere and Its Feedback on Global Change*. eds S. Liebner and L. Ganzert (Berlin: De Gruyter). 27–42. doi: 10.1515/9783110497083-002

Malard, L. A., and Pearce, D. A. (2018). Microbial diversity and biogeography in Arctic soils. *Environ. Microbiol. Rep.* 10, 611–625. doi: 10.1111/1758-2229.12680

Malard, L. A., Anwar, M. Z., Jacobsen, C. S., Pearce, A., Pearce, D. A., and Pearce, A. (2019). Biogeographical patterns in soil bacterial communities across the Arctic region. *FEMS Microbiol. Ecol.* 95:fiz128. doi: 10.1101/655431

Männistö, M. K., Tiirola, M., and Häggblom, M. M. (2007). Bacterial communities in Arctic fjelds of Finnish Lapland are stable but highly pH-dependent. *FEMS Microbiol. Ecol.* 59, 452–465. doi: 10.1111/j.1574-6941.2006.00232.x

Margesin, R. (2009). "Permafrost Soils" in *Soil Biology*. 4th Edn. ed. R. Margesin (Berlin: Springer-Verlag). doi: 10.1007/978-3-540-69371-0

Margesin, R. (2017). *Psychrophiles: from Biodiversity to Biotechnology*. 2nd edition. Berlin: Springer-Verlag. doi: 10.1007/978-3-319-57057-0.

Margesin, R., and Collins, T. (2019). Microbial ecology of the cryosphere (glacial and permafrost habitats): current knowledge. *Appl. Microbiol. Biotechnol.* 103, 2537–2549. doi: 10.1007/s00253-019-09631-3

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.J.* 17:10. doi: 10.14806/ej.17.1.200

Metcalfe, D. B., Hermans, T. D. G., Ahlstrand, J., Becker, M., Berggren, M., Björk, R. G., et al. (2018). Patchy field sampling biases understanding of climate change impacts across the Arctic. *Nat. Ecol. Evol.* 2, 1443–1448. doi: 10.1038/s41559-018-0612-5

Mondav, R., McCalley, C. K., Hodgkins, S. B., Frolking, S., Saleska, S. R., Rich, V. I., et al. (2017). Microbial network, phylogenetic diversity and community membership in the active layer across a permafrost thaw gradient. *Environ. Microbiol.* 19, 3201–3218. doi: 10.1111/1462-2920.13809

Monteux, S., Weedon, J. T., Gavazov, G. B. K., Blume-Werry, G., Gavazov, K., Jassey, V. E. J., et al. (2018). Long-term in situ permafrost thaw effects on bacterial communities and potential aerobic respiration. *ISME J.* 12, 2129–2141. doi: 10.1038/s41396-018-0176-z

Müller, O., Bang-Andreasen, T., White, R. A., Elberling, B., Taş, N., Kneafsey, T., et al. (2018). Disentangling the complexity of permafrost soil by using high resolution profiling of microbial community composition, key functions and respiration rates. *Environ. Microbiol.* 20, 4328–4342. doi: 10.1111/1462-2920.14348

Nilsson, R. H., Larsson, K., Taylor, A. F. S., Bengtsson-Palme, J., Jeppesen, T. S., Schigel, D., et al. (2019). The UNITE database for molecular identification of fungi: handling dark taxa and parallel taxonomic classifications. *Nucleic Acids Res.* 47, D259–D264. doi: 10.1093/nar/gky1022

Oksanen, J., Blanchet, G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., et al. (2019). *vegan: Community Ecology Package*. Available online at: https://CRAN.R-project.org/package=vegan

Panikov, N. S., Flanagan, P. W., Oechel, W. C., Mastepanov, M. A., and Christensen, T. R. (2006). Microbial activity in soils frozen to below −39°C. *Soil Biol. Biochem.* 38, 785–794. doi: 10.1016/j.soilbio.2005.07.004

Parada, A. E., Needham, D. M., and Fuhrman, J. A. (2016). Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environ. Microbiol.* 18, 1403–1414. doi: 10.1111/1462-2920.13023

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: machine Learning in Python. *J. Mach. Learn. Res.* 12, 2825–2830. doi: 10.1080/13696998.2019.1666854

Perez-Mon, C., Stierli, B., Plötze, M., and Frey, B. (2022). Fast and persistent responses of alpine permafrost microbial communities to in situ warming. *Sci. Total Environ.* 807:150720. doi: 10.1016/j.scitotenv.2021.150720

Price, M. N., Dehal, P. S., and Arkin, A. P. (2010). FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. *PLoS One* 5:e9490. doi: 10.1371/journal.pone.0009490

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Glo, F. O., et al. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* 41, 590–596. doi: 10.1093/nar/gks1219

R Core Team (2021). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

R Studio Team (2021). *RStudio: Integrated Development Environment for R*. Boston, MA: RStudio, Inc.

Rasmussen, L. H., Zhang, W., Hollesen, J., Cable, S., Christiansen, H. H., Jansson, P. E., et al. (2018). Modelling present and future permafrost thermal regimes in Northeast Greenland. *Cold Reg. Sci. Technol.* 146, 199–213. doi: 10.1016/j.coldregions.2017.10.011

Reimer, P. J., Edouard Bard, B., Alex Bayliss, B., Warren Beck, B. J., Paul Blackwell, B. G., Christopher Bronk, et al. (2013). Intcal13 and Marine13 Radiocarbon Age Calibration Curves 0–50,000 Years Cal Bp. *Radiocarbon* 55, 1869–1887. doi: 10.2458/azu_js_rc.55.16947

Rivkina, E., Petrovskaya, L., Vishnivetskaya, T., Krivushin, K., Shmakova, L., Tutukina, M., et al. (2016). Metagenomic analyses of the late Pleistocene permafrost – Additional tools for reconstruction of environmental conditions. *Biogeosciences* 13, 2207–2219. doi: 10.5194/bg-13-2207-2016

Rivkina, E., Shcherbakova, V., Laurinavichius, K., Petrovskaya, L., Krivushin, K., Kraev, G., et al. (2007). Biogeochemistry of methane and methanogenic archaea in permafrost. *FEMS Microbiol. Ecol.* 61, 1–15. doi: 10.1111/j.1574-6941.2007.00315.x

Rousk, J., Bååth, E., Brookes, P. C., Lauber, C. L., Lozupone, C., Caporaso, J. G., et al. (2010). Soil bacterial and fungal communities across a pH gradient in an arable soil. *ISME J.* 4, 1340–1351. doi: 10.1038/ismej.2010.58

Schostag, M., Priemé, A., Jacquiod, S., Russel, J., Ekelund, F., and Jacobsen, C. S. (2019). Bacterial and protozoan dynamics upon thawing and freezing of an active layer permafrost soil. *ISME J.* 13, 1345–1359. doi: 10.1038/s41396-019-0351-x

Schostag, M., Stibal, M., Jansson, J. K., Jacobsen, C. S., Stibal, M., Priemé, A., et al. (2015). Distinct summer and winter bacterial communities in the active layer of Svalbard permafrost revealed by DNA- and RNA-based analyses. *Front. Microbiol.* 6:399. doi: 10.3389/fmicb.2015.00399

Schuur, E. A. G., and Mack, M. C. (2018). Ecological Response to Permafrost Thaw and Consequences for Local and Global Ecosystem Services. *Annu. Rev. Ecol. Evol. Syst.* 49, 279–301. doi: 10.1146/annurev-ecolsys-121415-032349

Schuur, E. A. G., McGuire, A. D., Schädel, C., Grosse, G., Harden, J. W., Hayes, D. J., et al. (2015). Climate change and the permafrost carbon feedback. *Nature* 520, 171–179. doi: 10.1038/nature14338

Takahashi, S., Tomita, J., Nishioka, K., Hisada, T., and Nishijima, M. (2014). Development of a prokaryotic universal primer for simultaneous analysis of Bacteria and Archaea using next-generation sequencing. *PLoS One* 9:e105592. doi: 10.1371/journal.pone.0105592

Tarnocai, C., Canadell, J. G., Schuur, E. A. G. G., Kuhry, P., Mazhitova, G., Zimov, S., et al. (2009). Soil organic carbon pools in the northern circumpolar permafrost region. *Global Biogeochem. Cycles* 23, 1–11. doi: 10.1029/2008GB003327

Tripathi, B. M., Kim, H. M., Jung, J. Y., Nam, S., Hyeon, T. J., Kim, M., et al. (2019).). Distinct taxonomic and functional profiles of the microbiome associated with different soil horizons of a moist tussock tundra in Alaska. *Front. Microbiol.* 10:1442. doi: 10.3389/fmicb.2019.01442

Tripathi, B. M., Kim, M., Kim, Y., Byun, E., Yang, J. W., Ahn, J., et al. (2018). Variations in bacterial and archaeal communities along depth profiles of Alaskan soil cores. *Nat. Sci. Reports* 8:504. doi: 10.1038/s41598-017-18777-x

Trivedi, P., Anderson, I. C., and Singh, B. K. (2013). Microbial modulators of soil carbon storage: integrating genomic and metabolic knowledge for global prediction. *Trends Microbiol.* 21, 641–651. doi: 10.1016/j.tim.2013.09.005

Tuorto, S. J., Darias, P., McGuinness, L. R., Panikov, N., Zhang, T., Häggblom, M. M., et al. (2014). Bacterial genome replication at subzero temperatures in permafrost. *ISME J.* 8, 139–149. doi: 10.1038/ismej.2013.140

Turetsky, M., Abbott, B., Jones, M., Walter Anthony, K. M., Olefeldt, D., Schuur, E. A. G., et al. (2020). Carbon release through abrupt permafrost thaw. *Nat. Geosci.* 13, 138–143. doi: 10.1038/s41561-019-0526-0

Tveit, A. T., Kiss, A., Winkel, M., Horn, F., Hájek, T., Svenning, M. M., et al. (2020). Environmental patterns of brown moss- and Sphagnum-associated microbial communities. *Sci Rep.* 10:22412. doi: 10.1038/s41598-020-79773-2

Tveit, A. T., Schwacke, R., Svenning, M. M., and Urich, T. (2013). Organic carbon transformations in high-Arctic peat soils: key functions and microorganisms. *ISME J.* 7, 299–311. doi: 10.1038/ismej.2012.99

Wei, S., Cui, H., Zhu, Y., Lu, Z., Pang, S., Zhang, S., et al. (2018). Shifts of methanogenic communities in response to permafrost thaw results in rising methane emissions and soil property changes. *Extremophiles* 22, 447–459. doi: 10.1007/s00792-018-1007-x

Westermann, S., Elberling, B., Højlund Pedersen, S., Stendel, M., Hansen, B. U., and Liston, G. E. (2015). Future permafrost conditions along environmental gradients in Zackenberg. *Greenland. Cryosphere* 9, 719–735. doi: 10.5194/tc-9-719-2015

Willerslev, E., Hansen, A. J., Rønn, R., Brand, T. B., Barnes, I., Wiuf, C., et al. (2004). Long-term persistence of bacterial DNA. *Curr. Biol.* 14, 13–14. doi: 10.1016/j.cub.2003.12.012

Willms, I. M., Bolz, S. H., Yuan, J., Krafft, L., Schneider, D., Schöning, I., et al. (2021). The ubiquitous soil verrucomicrobial clade 'Candidatus Udaeobacter' shows preferences for acidic pH. *Environ. Microbiol. Rep.* 13, 878–883. doi: 10.1111/1758-2229.13006

Xiang, X., Wang, R., Wang, H., Gong, L., Man, B., and Xu, Y. (2017). Distribution of *Bathyarchaeota* communities across different terrestrial settings and their potential ecological functions. *Sci. Rep.* 7:45028. doi: 10.1038/srep45028

Yergeau, E., Hogues, H., Whyte, L. G., and Greer, C. W. (2010). The functional potential of high Arctic permafrost revealed by metagenomic sequencing, qPCR and microarray analyses. *ISME J.* 4, 1206–1214. doi: 10.1038/ismej.2010.41

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

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Microbiology

REVIEW
published: 25 March 2022
doi: 10.3389/fmicb.2022.786764

Diatoms and Their Microbiomes in Complex and Changing Polar Oceans

*Reuben Gilbertson[1], Emma Langan[1,2] and Thomas Mock[1]**

[1] School of Environmental Sciences, University of East Anglia, Norwich Research Park, Norwich, United Kingdom, [2] The Earlham Institute, Norwich Research Park, Norwich, United Kingdom

Diatoms, a key group of polar marine microbes, support highly productive ocean ecosystems. Like all life on earth, diatoms do not live in isolation, and they are therefore under constant biotic and abiotic pressures which directly influence their evolution through natural selection. Despite their importance in polar ecosystems, polar diatoms are understudied compared to temperate species. The observed rapid change in the polar climate, especially warming, has created increased research interest to discover the underlying causes and potential consequences on single species to entire ecosystems. Next-Generation Sequencing (NGS) technologies have greatly expanded our knowledge by revealing the molecular underpinnings of physiological adaptations to polar environmental conditions. Their genomes, transcriptomes, and proteomes together with the first eukaryotic meta-omics data of surface ocean polar microbiomes reflect the environmental pressures through adaptive responses such as the expansion of protein families over time as a consequence of selection. Polar regions and their microbiomes are inherently connected to climate cycles and their feedback loops. An integrated understanding built on "omics" resources centered around diatoms as key primary producers will enable us to reveal unifying concepts of microbial co-evolution and adaptation in polar oceans. This knowledge, which aims to relate past environmental changes to specific adaptations, will be required to improve climate prediction models for polar ecosystems because it provides a unifying framework of how interacting and co-evolving biological communities might respond to future environmental change.

Keywords: diatoms, psychrophile, genomics, meta-omics, polar winter, climate change, adaptive evolution, microbiomes

OPEN ACCESS

Edited by:
Anne D. Jungblut, Natural History Museum, United Kingdom

Reviewed by:
Mark Moore, University of Southampton, United Kingdom
Caroline Chénard, National Research Council Canada (NRC-CNRC), Canada

****Correspondence:***
Thomas Mock
t.mock@uea.ac.uk

Specialty section:
This article was submitted to Extreme Microbiology, a section of the journal Frontiers in Microbiology

Received: *30 September 2021*
Accepted: *23 February 2022*
Published: *25 March 2022*

Citation:
Gilbertson R, Langan E and Mock T (2022) Diatoms and Their Microbiomes in Complex and Changing Polar Oceans. Front. Microbiol. 13:786764. doi: 10.3389/fmicb.2022.786764

INTRODUCTION

Polar oceans are major drivers of the global carbon pump, circulation of nutrients and reflection of solar radiation (Murata and Takizawa, 2003; Sigman et al., 2010; MacGilchrist et al., 2014; Laufkötter et al., 2018; Wadham et al., 2019). Despite the global importance of both the geophysical and biological aspects of the polar oceans, they are critically understudied. In addition to

consistently low temperatures, polar oceans comprise a multitude of stressors such as extremes in solar irradiance, salinity and UV radiation (Boyd, 2002). While the Arctic and Antarctic have similar climates due to the high latitudes, each ecosystem is characterized by unique geography which shapes the local environment. The Arctic Ocean is surrounded by large shelf areas connected to land masses and annually undergoes dramatic changes in the extent of sea ice. Conversely, the Southern Ocean is a merry-go-round system characterized by strong latitudinal gradients of temperature (Oceanic fronts) isolating the Antarctic continent (Hansom and Gordon, 1998; Maksym, 2019). These fronts have recently been discovered to be responsible for ecotypic differentiation and speciation in the endemic and pelagic Southern Ocean diatom *Fragilariopsis kerguelensis* (Postel et al., 2020). Hence, they contribute to creating and maintaining diatom biodiversity in the Southern Ocean. Wherever there are fewer barriers such as in the Arctic Ocean, there is likely more exchange (e.g., gene flow) between microbial populations. Their terrestrial landscapes also differ. For example, higher vascular plants have colonized around 75% of the Arctic (Walker et al., 2005; Iversen et al., 2015), compared to a limited number of plant species across the Antarctic (Pointing et al., 2015; Singh et al., 2018). Despite these differences, most of the primary productivity and nutrient cycling for both ecosystems takes place in marine microbial communities inhabiting their associated oceans (Nelson et al., 1987; Smythe-Wright et al., 2010).

Polar microbiomes, mainly composed of diverse microalgae and their associated prokaryotes, are of particular importance in polar primary productivity and nutrient cycling and are the base of the highly productive polar food webs (Stretch et al., 1988; Harrison and Cota, 1991; Boyd, 2002; Bluhm and Gradinger, 2008; Aslam et al., 2012; Petrou et al., 2014; Hayward and Grigor, 2020). Diatoms are the most abundant and diverse group of eukaryotic polar phytoplankton and are key components of both pelagic and sea-ice habitats (Thomas and Dieckmann, 2002; Kooistra et al., 2007; Armbrust, 2009; Bracher et al., 2009; Tréguer et al., 2018; Trefault et al., 2021), as they thrive in seasonally mixed, cold and nutrient-rich water, characteristic of the polar oceans. Thus, polar oceans are their preferred environment where they outcompete many other phytoplankton groups, at least under past and current climatic conditions. Reasons for the dominance of diatoms in polar oceans include: (a) the ability to "boom and bust", which matches the response required to thrive under extreme seasonality, (b) an increased abundance of genetic elements that enhance an adaptive response (i.e., transposable elements), (c) differential allelic expression and (d) multiple sources of genes from endosymbiotic (EGT) and horizontal gene transfer (HGT), both of which contribute to metabolic plasticity and therefore facilitate specific adaptations (Martin et al., 1998; Sjöqvist and Kremp, 2016; Mock et al., 2017). Given their prevalence in polar phytoplankton microbiomes, initial research focused on understanding their physiological mechanisms for survival under extreme conditions. The advent of genomics tools and methods has influenced later research with a focus on identifying the underpinning genes and pathways (Sackett et al., 2013; Palenik, 2015).

Molecular dating of the emergence of adaptations and radiations in polar species coincide with major geological events, such as the opening of the Drake Passage (∼35 Mya), which resulted in the Antarctic Circumpolar Current, and subsequent isolation and freezing of Antarctica (Suto et al., 2012; Benoiston et al., 2017). Generally, between the Eocene and Oligocene (∼34 Ma), the Earth cooled, and changes in upwelling patterns created favorable conditions for diatoms, especially in the Southern Ocean, where the evidence of the rise and sustained dominance of diatoms in the water column is found in large siliceous ooze deposits in the area (Salamy and Zachos, 1999; Dutkiewicz et al., 2015; Benoiston et al., 2017). All microbial life was affected by this global cooling, including the divergence of polar clades of *Chlamydomonas sp.* from temperate lineages (Zhang et al., 2020), the Atlantic *Chaetoceros* Explosion (ACE), and global radiation in soil diatoms (Suto et al., 2012; Pinseel et al., 2020). It has been suggested that the rise in diatom abundance and subsequently primary production, specifically *Chaetoceros,* during this time enabled organisms in higher trophic levels such as zooplankton and marine mammals to thrive and diversify (Suto et al., 2012). These data highlight that major changes in marine microbial communities are inextricably linked to climate cycles and can have far-reaching implications on the rest of the food chain (Falkowski, 1998; Suto et al., 2012). Just as diversification events such as the ACE can be observed in fossil records, the genetic code of marine organisms can show how they adapted and evolved to survive changing climates. Understanding the evolution of diverse polar diatom genomes and their associated microbes can therefore provide insight into past and current climate conditions, potentially improving models based on either traits or functional types, especially due to the rapid nature of climate change in polar ecosystems (Kwok, 2018). It appears that warming polar oceans have so far been generally beneficial for polar phytoplankton populations as consistently increasing temperatures have extended the season in which polar phytoplankton can grow but have so far been small enough to prevent significant encroachment of invasive species (Arrigo et al., 2008; Erwin, 2009). Increasing temperatures are generally resulting in prolonged stratification and growing seasons, ocean acidification and reduced upwelling creating conditions which some phytoplankton are unable to adapt to, opening niches for invasive species (Vincent, 2010; Boyd et al., 2016). However, the polar oceans are continuing to warm and the wider effect this will have on diverse polar microbiomes in the future is not well known. Unlike diatom fossil records there are no historical databases built on omics data, without a reference of the current diversity of polar microbiomes and associated functional traits, our ability to predict changes is limited. To combat this, we suggest an increase in metaomics sequencing of environmental communities from the polar regions to provide a background of current populations to relate future changes to.

However, evolutionary genomics with polar microbes is still in its infancy but will be necessary to improve predictions of key species' responses to climate change (Waldvogel et al., 2020) which could have significant effects on the food-web structure and biogeochemical cycles of elements, as seen during past major geological events (Suto et al., 2012). Even intraspecific changes

in biodiversity will have knock-on effects on the carbon cycle, as the changing polar ocean likely selects for different strains (e.g., ecotypes) with different traits impacting food web dynamics and the cycling of elements especially if keystone groups such as diatoms are affected (Field et al., 1998; Wolf et al., 2019). Thus, forecasts for how this rapid climate change will impact polar ecosystems remain incomplete unless we improve our understanding of how the polar environment has shaped the evolution and biodiversity of polar organismal communities (Bindoff et al., 2007; Steig et al., 2009; Lee, 2014; Mock et al., 2016; Murphy et al., 2016; Verde et al., 2016; Brown et al., 2019).

Consequently, to address the uncertainty of the effect future climate change will have on biodiversity change and loss in polar marine ecosystems, integrative approaches are required. We think they should be based on sequencing data because they provide comprehensive insights into microbial functional diversity and how this diversity might change due to selection driven by climate change. However, our current molecular knowledge about polar microbes is very limited because of (a) lack of diverse polar model species for cell biology and therefore fundamental insights into their biochemical adaptation and molecular evolution, and (b) a very limited number of environmental sequencing initiatives to reveal genetic and genomic biodiversity of polar microbes. The aim of this review paper, therefore, is to reflect on what we have learned so far from the omics resources currently available from limited model psychrophilic microalgae, specifically *Fragilariopsis cylindrus,* and their associated microbiomes in the context of future advances exemplified by developments in neighboring fields such as plant sciences, microbiome research, and macroecology. In addition to highlighting some of the latest advances in polar diatom molecular biology, we provide suggestions as to how to build bridges to neighboring disciplines to fill gaps in our knowledge and therefore to advance our field for tackling the challenges mentioned above.

Using *Fragilariopsis cylindrus* as a Model to Reveal How Marine Phytoplankton Are Adapted to the Polar Climate

Although diatoms are the most species-rich group of algae, diatom research has immensely benefited from model species such as *Thalassiosira pseudonana* and *Phaeodactylum tricornutum* (Falciatore et al., 2020). However, due to limitations given by the nature of both models (e.g., small and streamlined genomes, limited evidence for sexual reproduction, mesophiles, in culture for many decades), additional diatom species have been developed into models. One of them is *F. cylindrus*, which is the first and only eukaryotic psychrophile so far that is genetically tractable (Faktorová et al., 2020) with a sequenced genome, and grows well under laboratory conditions (Mock et al., 2017). Hence, it has been used to study physiological and molecular adaptation to polar conditions. Significant seasonality in light and the overall low temperatures are amongst the strongest selecting agents shaping the evolution and adaptation of *F. cylindrus* and likely other polar diatoms. In this section of the review paper, we will focus on molecular data gained from the sequencing of polar diatoms, mainly *F. cylindrus*, and how this molecular data compares to that from other psychrophilic microbes, and mesophilic model diatom species.

The Dark Polar Winter

In most temperate and polar aquatic environments, especially coastal waters, diatoms are typically one of the first taxa groups to bloom with the changing seasons by taking advantage of renewed nutrients and increased photosynthetically active radiation (PAR) (Soreide et al., 2010; Kvernvik et al., 2018). Polar diatoms are no exception, due to mechanisms allowing them to survive the polar winter with the necessary cellular components ready and waiting for the return of spring (Peters and Thomas, 1996; van de Poll et al., 2020). However, the molecular mechanisms underpinning these adaptations have been a mystery. Polar phytoplankton have been documented to employ a diverse set of strategies, such as forming resting spores, heterotrophy and respiration of stored lipids (Bunt and Lee, 1972; Reeves et al., 2011; McMinn and Martin, 2013; Schaub et al., 2017). The traditional view of a pause in biological activity has recently been challenged as more data during the winter months is collected, indicating that there is stable biological activity (Berge et al., 2015). This is suggesting a diverse polar microbial community with multiple mechanisms of survival in prolonged darkness. The addition of detailed transcriptomics and proteomics data from *F. cylindrus* cultured in complete darkness has enabled researchers to explore genetic variation, and how it could explain the physiological activity in prolonged darkness. Within 24h of the onset of darkness, there was a significant reduction in the abundance of light-harvesting protein complexes (LHCs), which transfer light energy to the photosystem reaction center (Kennedy et al., 2019). Despite the sudden widespread downregulation in the abundance of LHCs, these complexes are maintained at a low basal level in the absence of light. This coincided with the upregulation of proteins affiliated with glycolysis, the TCA cycle and the Entner–Doudoroff pathway, suggesting a coordinated shift from photosynthesis to cellular respiration to maintain essential functions for survival (**Figure 1**; Kennedy et al., 2019). This large response in the proteome corroborates data on differential expression in the *F. cylindrus* transcriptome in response to prolonged darkness (Mock et al., 2017). Dynamic regulation of the proteome and transcriptome to significant changes in light intensity does not appear to be exclusive to *F. cylindrus*. For example, the model psychrophile Chlamydomonas raudensis and the natural community around it, also maintained photosystem complexes at a basal level when cultured in dialysis bags in Lake Bonney (McMurdo Dry Valleys, Antarctica) during the transition to polar night conditions (Morgan-Kiss et al., 2016). Despite this study focusing on a different ecosystem and a distantly related species from diatoms, the same molecular response was discovered, suggesting co-adaptation to the polar climate across distinct phytoplankton communities.

After months of darkness, the sea-ice thickness begins to decline in the spring and more light can reach diatom populations in the surface ocean eventually becoming sustained high light during the summer months. The *F. cylindrus* genome encodes significantly more LHCx genes, a subset of LHC

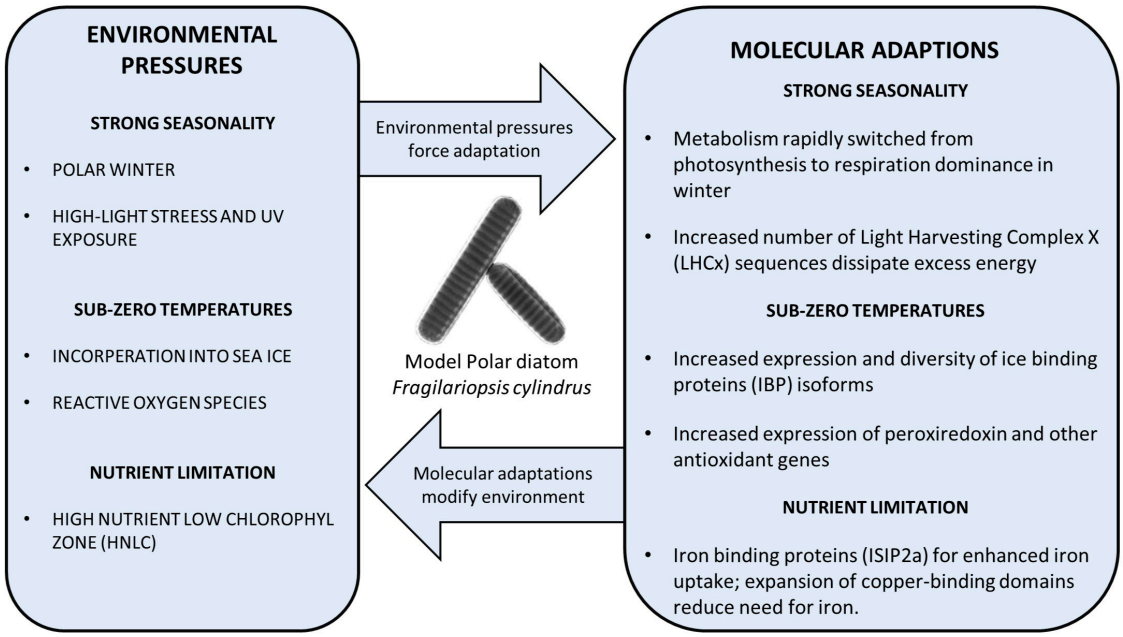

FIGURE 1 | Diagram showing examples of the adaptations of polar diatoms to a variety of polar environmental stressors, discovered using the model diatom *F. cylindrus*. Directional arrows display the pressure from the environment and the feedback/modification of the environment by polar diatoms.

sequences that are induced by high-light stress, than the model temperate diatoms *T. pseudonana* and *P. tricornutum*. The evolutionary expansion and upregulation of LHCx genes, known to reduce high-light stress, suggests their key role in the adaptive evolution to the polar light climate (**Figure 1**; Mock et al., 2017). Interestingly, the LHCx gene family is also expanded in the mesophilic alga *Aureococcus anophagefferens,* enabling this species to rapidly form blooms in high-light coastal environments (Gobler et al., 2011; Mock et al., 2017). Thus, to be prepared for high-light environments after a period of darkness, polar diatoms appear to be downregulating a subset of photosynthesis-associated genes accompanied by increasing cellular respiration. This combination of acclamatory metabolic processes enables them to survive the polar winter in an active state (Berge et al., 2015) and take full advantage of the spring and summer high-light conditions, especially in open water by upregulation of LHCx genes. These results show the usefulness of genomic data in researching how phytoplankton adapted to the physical polar environment, and the extent to which adaptations are possibly unique to specific taxa groups, giving more insight into the potential impact climate change will have on diverse polar phytoplankton communities.

Freezing

Ice-Binding Proteins

Ice-binding proteins (IBPs) are molecules that act on the interface between ice and water (Dolev et al., 2016). Their activity can reduce the freezing point and inhibit ice-crystal growth altogether. IBPs have been found in a diverse array of taxa across the world from insects to fish, with the majority found in species living in Arctic and Antarctic ecosystems such as oceans, frozen lakes and glaciers (DeVries and Wohlschlag, 1969; Raymond, 2014; Jung et al., 2016; Vance et al., 2019). The structure of IBPs is diverse throughout the domains of life, but the general function is conserved, suggesting multiple independent origins (Dolev et al., 2016).

Many diatom species are incorporated into the sea ice each year (Eicken, 1992; Thomas and Dieckmann, 2002). When frozen into forming sea ice, diatoms are enclosed into interconnected channels and pockets filled with brine of high salinity. To maintain an aqueous habitat inside sea ice, polar diatoms are known to produce IBPs and extracellular polymeric substances (EPS) to influence the physical state of their surrounding icy environment (Hoagland et al., 1993; Bayer-Giraldi et al., 2010; Lyon and Mock, 2014; Raymond, 2014; Arrigo et al., 2017; Aslam et al., 2018; Liang et al., 2019; Vance et al., 2019). There is evidence that the IBPs they secrete help maintain the aqueous state of the brine channel system to ensure access to nutrients from the seawater underneath through diffusive and convective transport processes (Raymond et al., 2009; Dolev et al., 2016). One interesting aspect of the functional conservation of IBPs is the Domain of Unknown Function 3494 (DUF3494), which has been identified in a large range of taxa in a variety of cold habitats. Despite significant sequence divergence, all homologs share the biophysical ability to reduce the growth of ice crystals (Vance et al., 2019; Raymond et al., 2021). Phylogenetic analysis of IBP sequences does not correlate with 18S-based phylogeny, which suggests that horizontal gene transfer (HGT) may have been a vector for the DUF3494's widespread presence (Keeling and Palmer, 2008; Bayer-Giraldi et al., 2010; Raymond and Kim, 2012; Mock et al., 2017; Raymond and Morgan-Kiss, 2017; Vance et al.,

2019). HGT in sea-ice communities is considered to be facilitated by the proximity of organisms in the narrow brine-channel system in combination with a strong selection pressure imposed by the harsh environmental conditions such as high salinities and subfreezing temperatures (Raymond et al., 2009; Raymond and Kim, 2012). Their important role in the adaptive evolution to the sea-ice habitat is supported by the fact that they usually expand (e.g., gene duplications) once they have been acquired *via* HGT (Mock et al., 2017; Raymond and Morgan-Kiss, 2017). Hence, many polar organisms have more than a single IBP gene encoded in their genomes such as *F. cylindrus* and *Chlamydomonas sp*. ICE-L.

The genome of *F. cylindrus* encodes 11 unique IBP isoforms, several of which are significantly upregulated under freezing temperatures and elevated salinity (**Figure 1**; Mock et al., 2017). Isoform 11 from *F. cylindrus* (*Fc*IBP11), despite having only moderate activity, can bind to multiple planes of ice crystals (Kondo et al., 2018). In addition to *Fc*IBPs containing the DUF3494, predicted to inhibit ice crystallization, several proteins such as *Fc*IBP-1, have transmembrane domains and are therefore hypothesized to protect the cell membrane from ice crystals (Mock et al., 2017). However, alternative roles are equally likely, such as sensing the formation of ice crystals and therefore initiating the appropriate physiological response to mitigate the impact of ice-crystal formation on the integrity of cellular structures.

Encoding multiple isoforms of antifreeze proteins to a variety of ice planes is evident in other polar eukaryotes, such as the psychrophilic yeast *Glaciozyma antarctica* (**Figure 2**) and the green algae *Coccomyxa subellipsoidea* (**Figure 2**; Blanc et al., 2012; Firdaus-Raih et al., 2018). The genome of the psychrophilic green alga *Chlamydomonas sp.* ICE-L, the first polar green algae sequenced, also contains expanded IBP gene families. All 12 IBPs in the genome of *Chlamydomonas sp.* contain the DUF3494, and again these sequences show a closer phylogenetic relationship to that of bacteria (Zhang et al., 2020). Polar diatoms and the wider microbiome produce EPS to work in synergy with IBPs by creating disorder in ice-crystal formation leading to retention of salinity, effectively reducing the freezing point (Ewert and Deming, 2011; Krembs et al., 2011; Mykytczuk et al., 2013; Aslam et al., 2018; Raymond-Bouchard et al., 2018; Torstensson et al., 2019). By having multiple IBP isoforms with a diverse range of functions and applications to different conditions, the cell appears to be able to sense the icy environment when it forms, which likely induces diverse protection mechanisms in which IBPs play a key role.

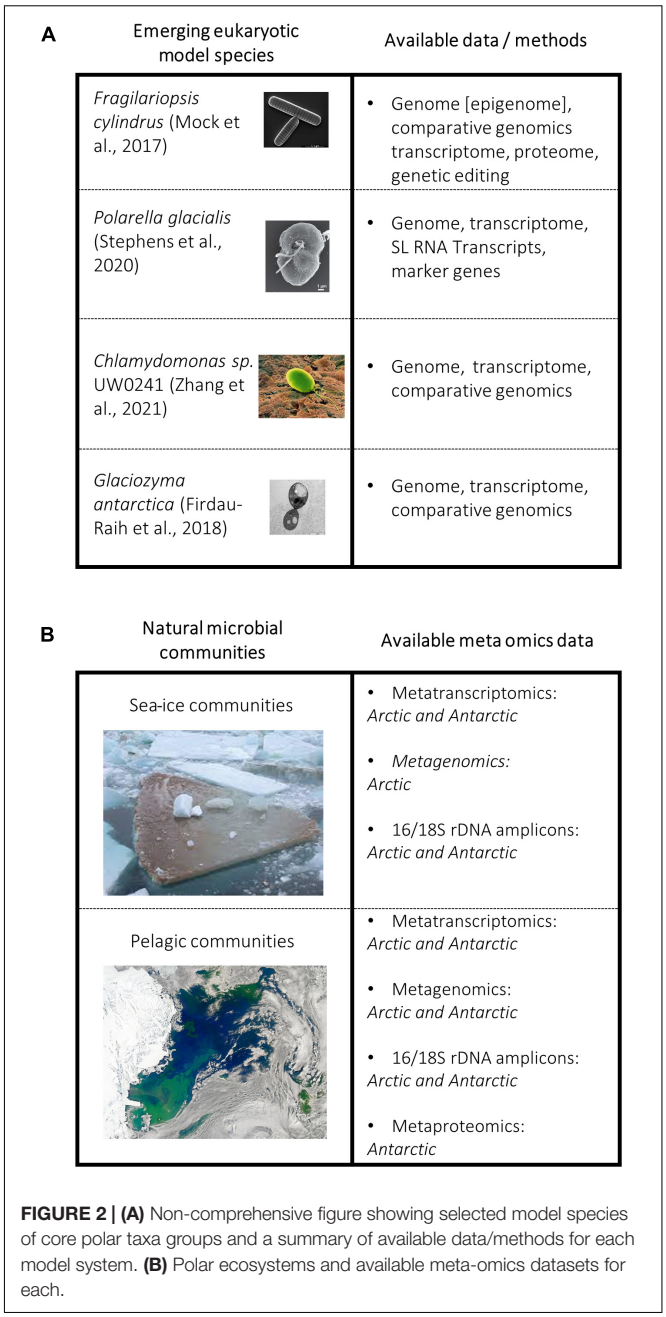

FIGURE 2 | (A) Non-comprehensive figure showing selected model species of core polar taxa groups and a summary of available data/methods for each model system. **(B)** Polar ecosystems and available meta-omics datasets for each.

Radical Oxygen Species and the Role of Trace Metals

Low temperatures generally reduce enzyme kinetics. Because of slow kinetics, excess free radicals accumulate within the cell causing dangerous levels of oxidative stress which can seriously damage cellular components, including DNA and membranes (Wingsle et al., 1999). As mentioned above, *F. cylindrus* displays a divergent expression of homologous alleles under varying environmental conditions. The most divergently expressed allele set across all tested conditions, including freezing temperatures, is glutathyamine s-transferase I (GST-1), a scavenger for free radical oxygen species (ROS). Thus, this enzyme is a good example that diverged expression of alleles might be driven by environmental pressures (Mock et al., 2017). Similarly, *F. kerguelensis* overexpressed transcripts for peroxiredoxin (PrxQ) to inhibit a diverse set of peroxides (Moreno et al., 2020). Similar responses across other psychrophiles include growth changes, upregulating expression of other antioxidant proteins such as specific catalases, overexpression of antifreeze proteins, and superoxide dismutases (Wong et al., 2019; Zhang et al., 2020). The latter converts superoxide radicals to more stable species (Miteva-Staleva et al., 2011; Scholz et al., 2014; Raymond-Bouchard et al., 2018). These data suggest that

multiple molecular strategies have evolved to cope with common environmental stressors in polar ecosystems.

The low concentrations of micronutrients, specifically iron, compared with macronutrients in the Southern Ocean results in a High Nutrient Low Chlorophyll region (HNLC) which limits phytoplankton growth (Martin et al., 1991; Pitchford and Brindley, 1999; Venables and Moore, 2010; Hassler et al., 2012). Thus, iron limitation can impose cellular stress. For example, iron limitation compounds the intracellular ROS stress due to a reduction in the capacity of the electron transport chain (Allen et al., 2008). Genomics, transcriptomics, and proteomics studies have discovered unique proteins which are responsive to micronutrient fluxes to mitigate this stress such as through FLDA2b in *F. kerguelensis* (Allen et al., 2008; Marchetti et al., 2009; Lommer et al., 2012; Bender et al., 2014; Marchetti, 2019; Moreno et al., 2020). Southern Ocean diatoms exposed to varying light intensities, and different iron concentrations, were found to have an increased number of genes associated with iron uptake compared to temperate diatoms (Mock et al., 2017; Moreno et al., 2018). Iron concentrating proteins in polar diatoms, such as iron starvation-induced protein 2a (ISIP2a) identified in *F. kerguelensis*, are significantly overexpressed in low iron conditions to increase uptake of iron from the environment (Marchetti et al., 2009; Moreno et al., 2020). In addition to increased affinity for iron, the *F. cylindrus* genome is enriched for copper-binding domains, outnumbering iron-binding domains (Mock et al., 2017). Most are of the plastocyanin/azurin-like family, possibly reducing the requirement for iron in cellular processes (Peers and Price, 2006; Mock et al., 2017). *F. kerguelensis* contains several plastocyanin isoforms which are significantly upregulated in response to low iron, specifically *PCYN-2b* (Moreno et al., 2020). The *F. cylindrus* genome also revealed an elevated number of zinc-binding domain-containing genes, with more than six times the number of homologous clusters compared to *P. tricornutum* and *T. pseudonana*; specifically, MYND Zinc Fingers. MYND domains in *F. cylindrus* are associated with a variety of accompanying domains with diverse functions, many of which are unknown (Laity et al., 2001; Mock et al., 2017). The evolutionary expansion of this family suggests their importance potentially in terms of regulating the activity of gene expression and/or the activity of enzymes (Mock et al., 2017). Their expansion may have been facilitated by the relatively high zinc concentrations in the Sothern Ocean (Croot et al., 2011; Mock et al., 2017). Most of the expansion in this family is predicted to have originated around 30 million years ago, which therefore coincides with the opening of the Drake passage (Mock et al., 2017). Thus, the coincidence between the evolutionary expansion of zinc-binding protein genes and the cooling of the Southern Ocean begs the question about the role of zinc in the adaptation of diatoms to environmental conditions in surface polar oceans.

The studies discussed so far in this review have given insight into the viability and usefulness of polar model organisms to reveal fundamental processes underpinning adaptation and evolution, however, relying on model organisms limits our ability to understand how biodiversity changes due to environmental change. For instance, the role of microbial interactions is neither well studied nor represented by using a monoculture of an individual model species (Wolf et al., 2019) because intra- and interspecific competition is based on complex dynamic populations and their communities shaping the evolution and adaptation of all microbes involved. For example, in the study by Wolf et al. (2019), different strains of the species *Thalassiosira hyaline* were grown in monoculture under predicted future climate conditions and then mixed with other strains to see if their physiology changed when cultivated altogether. Under elevated temperatures, the fastest-growing strain outcompeted the others when grown in mixed culture. Conversely, under ambient temperatures, its growth rate resembled that of the slowest growing strain (Wolf et al., 2019). This demonstrates that monoculture studies may not be representative of how individual species behave as part of complex communities in nature. Thus, understanding intraspecific variation between strains of the same species is important because differences in traits might shape the structure and function of food webs and therefore biogeochemical cycling of elements (Vance et al., 2019). Hence, differences in the co-occurrence of strains and species in complex microbial communities are important for considering how environmental conditions have shaped these communities (e.g., Martin et al., 2021). To address this question, meta-omics approaches have proven successful in revealing the intricacies of complex microbial communities and their activity.

POLAR META-OMICS

Metagenomes and Their Assembled Genomes

Metagenomics, metatranscriptomics, and metaproteomics are emerging as important tools for understanding the molecular basis of adaptation in polar microbes by allowing us to study organisms in their ecosystem and to capture their interactions in the natural environment (McCain et al., 2021). Metagenomics has been used to establish which diatom species are present in certain locations, providing a baseline for future monitoring systems. For example, the Tara Oceans project investigated the global distribution of diatoms and discovered unexpected species in polar oceans (Malviya et al., 2016; Carradec et al., 2018; Obiol et al., 2020). Similarly unexpected was that polar oceans are a particular hotspot of viral diversity with high levels of novel genes (Gregory et al., 2019). Thus, metagenomics helps to reveal the microbial biodiversity in polar oceans and how different it is when compared to metagenomes from non-polar oceans. Metagenomics has also been applied to provide insights into how polar environmental gradients (e.g., sea-ice water interface) influence microbial biodiversity. For instance, a metagenomics survey of microbial communities in the Canadian Arctic found that in sea-ice there was a higher concentration of algal genes and chlorophyll-a compared to seawater underneath, where prokaryotic genes were more dominant. Metagenomic data showed that this was largely due to diatoms. Differences in operational taxonomic units (OTUs), a group of closely related individuals, between sea-ice and sea-water samples indicated that sea-ice dwellers have

different strategies for adaptation than seawater-based microbes. Increased variability was found in the sea-ice communities, perhaps indicative of the adaptability of polar diatoms as a dominant and diverse group of microbes in the sea-ice habitat (Yergeau et al., 2017).

Sequencing of individual species such as *F. cylindrus*, *P. glacialis, and Chlamydomonas sp.* UW0241 (**Figure 2**), has advanced our fundamental understanding of polar algae (Bayer-Giraldi et al., 2010; Mock et al., 2017; Stephens et al., 2020; Zhang et al., 2021). However, their sequencing requires resources and takes time. Yet, they serve as a reference not only for fundamental research but also for analyzing phytoplankton-enriched metagenomes. The latter together with reference genomes therefore will allow for the assembly of novel algal genomes isolated from natural communities including their associated microbiomes. Insights from this work will help us to understand the similarities and differences between genomes from natural communities, rather than relying on model organisms which risks not representatively capturing the diversity of adaptive mechanisms underlying their complex intertwined co-evolution. For instance, targeted metagenomics has been used to establish the evolutionary history of picoprymnesiophyte populations in the North Atlantic, which make up approximately 25% of global pico-plankton biomass. Their dominance was found to be due to high gene density, and genes that are likely to be involved in defense and nutrient uptake (Cuvelier et al., 2010).

Metagenome-assembled genomes (MAGs) of prokaryotes have been produced using large quantities of metagenomic data from both the Arctic and Antarctic (Cao et al., 2020; Royo-Llonch et al., 2021). Increased availability of high-quality MAGs without the requirement for individual culturing and sequencing will allow researchers to study a broader, more representative range of polar microbiomes and compare between species to understand mechanisms for polar survival (Cao et al., 2020; Duncan et al., 2020). MAGs have been used to study polar bacterial and archaeal populations from the Tara Oceans dataset, providing the first compendium of Arctic prokaryotic MAGs including differences in gene enrichment between Arctic and Antarctic populations (Cao et al., 2020; Royo-Llonch et al., 2021).

Although most MAGs from polar oceans are still from prokaryotes, one of the first MAG-based datasets from natural Arctic and Atlantic microbial communities was used to reveal inter-kingdom species associations including microbial eukaryotes (Duncan et al., 2020). Besides the generation of first eukaryotic MAGs from a polar ocean (e.g., diatoms, prasinophytes), this dataset was used to identify metabolism enriched in MAG-based species associations. By identifying which protein families are enriched in selected species, comparisons of associations to a background set of phylogenetically related species known not to have associations shed first insights into shared metabolism, potentially underpinning their biotic interactions. By applying this approach, Duncan et al. (2020) found positive and negative associations between algal and bacterial MAGs including some that were only found in the Arctic (e.g., Micromonas MAG associated with a Gammaproteobacteria MAG). Combined with cell isolations from the same samples, these MAG-based species associations can be empirically tested if the microbial partners can be co-cultivated under laboratory conditions.

Metatranscriptomes and Metaproteomes

Investigation of the genes expressed across global oceans was carried out as part of the Tara Oceans project (Bork et al., 2015) and the Sea of Change project (Martin et al., 2021). The former has been used to create an ocean atlas of eukaryotic genes from temperate and tropical regions whereas the latter has been used to produce the first pole-to-pole catalog of expressed genes from microalgae and other microeukaryotes. Most of the expressed genes in the Tara Oceans Atlas were novel, indicating that eukaryotic ocean life is under-studied and therefore incompletely understood. Additionally, genes specific to the Southern Ocean were different to those found elsewhere, potentially indicating that adaptation to polar conditions results in a very different gene set to non-polar species (Carradec et al., 2018). The latter insights were corroborated and extended by the Sea of Change project, which revealed that global algal microbiomes can be largely separated into two main groups: polar and non-polar species associations. The same demarcation was found for expressed genes of the algal partners considering their biogeographic distribution from pole to pole based on a combination of sequence co-occurrence analysis and the geographical location of breakpoints in their beta diversity (Difference in species composition between neighboring assemblages) (Martin et al., 2021). Thus, it appears that there are ecosystem boundaries in sub-polar regions of the upper ocean in both hemispheres separating polar from non-polar algal microbiomes.

Although metaproteomics with natural polar microbial communities is still in its infancy, the first study on coastal Antarctic phytoplankton communities has been undertaken. This study by McCain et al. (2021) used proteome-level traits to identify ecological strategies for different taxa. For instance, haptophytes appear to have a lower regulatory cost than diatoms, which may explain the observed haptophyte-to-diatom bloom progression in the Ross Sea. As protein synthesis is the main energy sink in cells, the quantity of cellular proteins measured as part of complex metaproteomes relates to costs involved in their synthesis. Hence, the nature and quantity of specific proteins identified potentially provides insights into the costs of traits and trade-offs. Challenges arise, however, with incomplete taxon-specific proteomes as part of complex metaproteomes especially if reference data are missing.

Omics-Based Monitoring and Modeling to Improve Prediction of Species' Responses to Polar Warming

Genetic monitoring using metagenomics and metatranscriptomics of mixed communities to oversere changes in populations over time can be used to study adaptations

to changing environmental conditions, showing the rate of evolution over time and how adaptations are occurring (Hansen et al., 2012; Pearson et al., 2015). However, sequence-based monitoring of microbial communities and their populations needs to be carefully designed to identify the most appropriate sites for capturing changes in the biogeography of species and their populations including alterations of gene flow and climate-driven range-shift expansions in their distribution. Reoccurring surveys with sufficient spatial sampling in a given geographic area or multiple monitoring sites with frequent sampling will be needed for a rigorous assessment of how warming in polar oceans influences microbial populations and the biogeochemical cycles they drive. Furthermore, monitoring needs to cover complete seasonal cycles to allow us to tease apart long-term ecosystem change from seasonal effects, which might be challenging as the latter are much more pronounced at high latitudes. Although many different geographical sites have been used for sampling and monitoring polar microbial communities, two stand out: The Fram Observatory in the Arctic and long-term observation of plankton at the Western Antarctic Peninsula (Lin et al., 2021; Wietz et al., 2021). The Fram observatory is based on autonomous samplers, which can sample at discrete time intervals to cover a complete seasonal cycle. Preliminary results covering 12 months including the dark winter period have revealed that the strong seasonal dynamics including the variable sea-ice extent are the main drivers for the succession in changes of microbial biodiversity based on 16 and 18S rDNA data and subsequent analysis (e.g., alpha diversity). Furthermore, the persistent sea-ice cover appears to reduce the seasonality effect on changes in microbial biodiversity (Wietz et al., 2021). As the autonomous samplers can be redeployed, they provide a platform for long-term monitoring of microbial

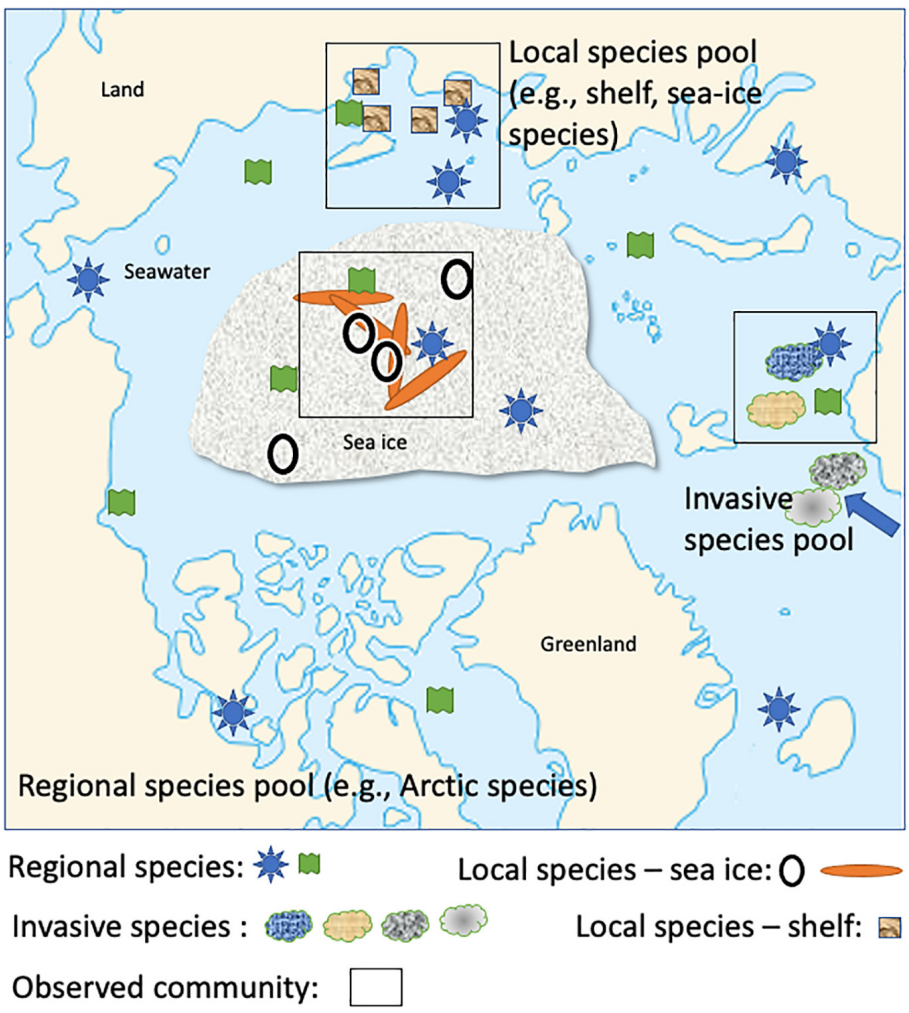

FIGURE 3 | Conceptual diagram of the assembly process influencing microbial communities at different spatiotemporal scales in the Arctic Ocean (e.g., regional, local). Geometric forms are arbitrary, but their composition represents how different microbial taxa (e.g., a specific geometric form such as a circle representing a taxon) contribute to complex microbiomes as the outcome of assembly processes at different spatiotemporal scales (Adapted from Ovaskainen et al., 2017). Drivers for the assembly process are likely a combination of abiotic and biotic conditions. For instance, the ephemeral nature of the sea-ice habitat in combination with strong gradients of temperature and salinity in brine channels can be assumed to significantly contribute to the assembly process of sea-ice microbiomes.

communities in the Arctic Ocean. A similar monitoring system is not yet, to our knowledge, available for the Southern Ocean, although there is a ship-based multi-year monitoring programme at the Western Antarctic Peninsula (Lin et al., 2021) based on similar methods (e.g., 16/18S rDNA). Modeling has also been shown that sea ice plays a significant role in structuring microbial communities and their productivity (Lin et al., 2021). However, increasing temperatures due to warming in this region are likely to result in declining microbial diversity (Tonelli et al., 2021).

Both of these monitoring programs have provided novel and highly needed insights into how the environmental conditions of the changing polar oceans shape their microbial inventory. To further understand how biotic and abiotic forces determine the abundance and distribution of microbial taxa in a polar ocean, future studies need to provide quantitative insights into how microbial traits (e.g., freeze-thaw resistance) underpin species interactions and how they are associated with habitat characteristics (e.g., sea-ice thickness, oceanic fronts, temperature gradients, snow-cover thickness). These biotic and abiotic forces likely encompass speciation, dispersal, and biotic interactions in a complex and highly dynamic polar environment.

The responses of microbial taxa to different habitat characteristics (environmental filters) and biotic interactions (biotic filters) vary depending on taxa-specific traits including their ability to acclimate, adapt, and their level of competitiveness (Ovaskainen et al., 2017). Thus, these traits will determine which taxa colonize habitats and dominate in seasonal successions (phenology) and therefore contribute to the community assembly process. To address this fundamental question, it is instrumental to apply explanatory models which can integrate data generated for quantitative insights into how microbial communities including their networks relate to environmental conditions and larger-scale ecosystem processes (e.g., seasonality of microbial diversity linked with habitat transformations such as freezing and melting of sea ice) (Ovaskainen et al., 2017; Tikhonov et al., 2020). One approach to tackle this challenge is hierarchical joint species/taxa distribution modeling (HJSDM) (Tikhonov et al., 2020). The heart of this approach is integrating species co-occurrence in co-variation with environmental conditions, and phylogeny of species and traits. Data required for this approach can be extracted from metagenomes, metatranscriptomes, and metaproteomes. For instance, MAGs will provide a genomic catalog (e.g., Cao et al., 2020; Royo-Llonch et al., 2021) of co-occurring microbial taxa in a phylogenetic but also spatio-temporal context if based on long-term sampling. Their genes, transcripts, and proteins provide trait information for revealing co-variation with environmental conditions. If sampling covers larger geographic regions, as done for the West Antarctic Peninsula, or the central Arctic Ocean as part of MOSAiC (Multidisciplinary drifting Observatory for the Study of Arctic Climate) (Wake, 2019), we will be able to delineate processes influencing the biodiversity of microbial communities at different spatio-temporal scales (**Figure 3**). Generally, a spatio-temporal context is required for separating drivers of seasonal differences vs climate-driven differences in microbial community composition and therefore to make robust predictions as to how warming impacts polar ecosystems.

DISCUSSION

The Earth's climate has been changing across the globe due to anthropogenic influences since the industrial revolution. There is no doubt of the importance of polar environments to regulate the global climate, with the Southern Ocean alone responsible for the uptake of around 40% of all anthropogenic CO_2 from the atmosphere (Khatiwala et al., 2009; Takahashi et al., 2009; DeVries, 2014; Long et al., 2021). As previously mentioned, polar ecosystems are experiencing more rapid changes, specifically warming, than any other biome on Earth. This alarming shift has been met with increased research over recent decades to understand the physical properties of these environments and the effect they have on food webs, with a focus on marine microbiomes, and diatoms in particular as one of the most important primary producers in polar oceans. Genome-enabled approaches have provided a step-change in our understanding of the adaptability of polar diatoms and allowed us to understand how they reacted to past climate events. However, there is still only one publicly available polar diatom genome, *F. cylindrus*, and with > 350 species of phytoplankton found in the Southern Ocean alone, using one genome to study such a diverse community is not viable. Hence, these studies need to expand as exemplified by the 100 Diatom Genomes Project[1] and the MOSAiC sequencing project at the United States Department of Energy Joint Genome Institute[2]. If monitoring programs build on these initiatives and adopt high-throughput omics approaches, we can generate the foundation for cataloging polar microbiomes through space and time. This information will help to integrate species association networks with traits to provide quantitative insights into how traits of species and their co-occurrence networks are associated with habitat characteristics. Based on this integrated approach, we potentially will be able to predict the future of microbial taxa and their diverse populations with greater confidence in warming polar oceans.

AUTHOR CONTRIBUTIONS

RG, EL, and TM wrote the article and equally contributed to the design of the figures. All authors contributed to the article and approved the submitted version.

FUNDING

RG and TM acknowledge funding from The Leverhulme Trust (RPG-2017-364) and the Natural Environment Research Council (NERC) (Grant Nos. NE/K004530/1 and NE/R000883/1). EL acknowledges funding from NEXUSS, a Natural Environmental Research Council (NERC) funded Doctoral Training Partnership (DTP). This work also partially funded from the School of Environmental Sciences at the University of East Anglia, Norwich, United Kingdom.

[1]https://jgi.doe.gov/csp-2021-100-diatom-genomes/

[2]https://jgi.doe.gov/csp-2020-arctic-ice-drift-experiment-mosaic/

REFERENCES

Allen, A. E., LaRoche, J., Maheswari, U., Lommer, M., Schauer, N., Lopez, P. J., et al. (2008). Whole-cell response of the pennate diatom *Phaeodactylum tricornutum* to iron starvation. *Proc. Natl. Acad. Sci. U.S.A.* 105, 10438–10443. doi: 10.1073/pnas.0711370105

Armbrust, E. V. (2009). The life of diatoms in the world's oceans. *Nature* 459, 185–192. doi: 10.1038/nature08057

Arrigo, K. R., van Dijken, G., and Pabi, S. (2008). Impact of a shrinking Arctic ice cover on marine primary production. *Geophys. Res. Lett.* 35:L19603. doi: 10.1029/2008GL035028

Arrigo, K. R., van Dijken, G. L., Alderkamp, A. C., Erickson, Z. K., Lewis, K. M., Lowry, K. E., et al. (2017). Early spring phytoplankton dynamics in the Western Antarctic Peninsula. *J. Geophys. Res. Oceans* 122, 9350–9369. doi: 10.1002/2017jc013281

Aslam, S. N., Cresswell-Maynard, T., Thomas, D. N., and Underwood, G. J. C. (2012). Production and characterization of the intra- and extracellular carbohydrates and polymeric substances (EPS) of three sea-ice diatom species, and evidence for a cryoprotective role for EPS. *J. Phycol.* 48, 1494–1509.

Aslam, S. N., Strauss, J., Thomas, D. N., Mock, T., and Underwood, G. J. C. (2018). Identifying metabolic pathways for production of extracellular polymeric substances by the diatom *Fragilariopsis cylindrus* inhabiting sea ice. *ISME J.* 12, 1237–1251. doi: 10.1038/s41396-017-0039-z

Bayer-Giraldi, M., Uhlig, C., John, U., Mock, T., and Valentin, K. (2010). Antifreeze proteins in polar sea ice diatoms: diversity and gene expression in the genus *Fragilariopsis*. *Environ. Microbiol.* 12, 1041–1052. doi: 10.1111/j.1462-2920.2009.02149.x

Bender, S. J., Durkin, C. A., Berthiaume, C. T., Morales, R. L., and Armbrust, E. V. (2014). Transcriptional responses of three model diatoms to nitrate limitation of growth. *Front. Mar. Sci.* 1:3. doi: 10.3389/fmars.2014.00003

Benoiston, A. S., Ibarbalz, F. M., Bittner, L., Guidi, L., Jahn, O., Dutkiewicz, S., et al. (2017). The evolution of diatoms and their biogeochemical functions. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 372:20160397. doi: 10.1098/rstb.2016.0397

Berge, J., Daase, M., Renaud, P. E., Ambrose, W. G., Darnis, G., Last, K. S., et al. (2015). Unexpected levels of biological activity during the polar night offer new perspectives on a Warming Arctic. *Curr. Biol.* 25, 2555–2561. doi: 10.1016/j.cub.2015.08.024

Bindoff, N. L., Willebrand, J., Artale, V., Cazenave, A., Gregory, J. M., Gulev, S., et al. (2007). "Observations: oceanic climate change and sea level," in *Climate Change 2007 The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, eds S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, et al. (Cambridge, MA: Cambridge University Press).

Blanc, G., Agarkova, I., Grimwood, J., Kuo, A., Brueggeman, A., Dunigan, D. D., et al. (2012). The genome of the polar eukaryotic microalga *Coccomyxa subellipsoidea* reveals traits of cold adaptation. *Genome Biol.* 13:R39. doi: 10.1186/gb-2012-13-5-r39

Bluhm, B. A., and Gradinger, R. (2008). Regional variability in food availability for arctic marine mammals. *Ecol. Appl.* 18, S77–S96. doi: 10.1890/06-0562.1

Bork, P., Bowler, C., De Vargas, C., Gorsky, G., Karsenti, E., and Wincker, P. (2015). Tara Oceans studies plankton at Planetary scale. *Science* 348:873.

Boyd, P. W. (2002). Review of environmental factors controlling phytoplankton processes in the Southern Ocean. *J. Phycol.* 38, 844–861. doi: 10.1046/j.1529-8817.2002.t01-1-01203.x

Boyd, P. W., Cornwall, C. E., Davison, A., Doney, S. C., Fourquez, M., Hurd, C. L., et al. (2016). Biological responses to environmental heterogeneity under future ocean conditions. *Glob. Change Biol.* 22, 2633–2650. doi: 10.1111/gcb.13287

Bracher, A., Vountas, M., Dinter, T., Burrows, J. P., Röttgers, R., and Peeken, I. (2009). Quantitative observation of cyanobacteria and diatoms from space using PhytoDOAS on SCIAMACHY data. *Biogeosciences* 6, 751–764. doi: 10.5194/bg-6-751-2009

Brown, M. S., Munro, D. R., Feehan, C. J., Sweeney, C., Ducklow, H. W., and Schofield, O. M. (2019). Enhanced oceanic CO2 uptake along the rapidly changing West Antarctic Peninsula. *Nat. Clim. Change* 9, 678–683. doi: 10.1038/s41558-019-0552-3

Bunt, J. S., and Lee, C. C. (1972). Data on the composition and dark survival of four sea-ice microalgae. *Limnol. Oceanogr.* 17, 458–461.

Cao, S., Zhang, W., Ding, W., Wang, M., Fan, S., Yang, B., et al. (2020). Structure and function of the Arctic and Antarctic marine microbiota as revealed by metagenomics. *Microbiome* 8, 1–12. doi: 10.1186/s40168-020-00826-9

Carradec, Q., Pelletier, E., Da Silva, C., Alberti, A., Seeleuthner, Y., Blanc-Mathieu, R., et al. (2018). A global ocean atlas of eukaryotic genes. *Nat. Commun.* 9, 1–13. doi: 10.1038/s41467-017-02342-1

Croot, P. L., Baars, O., and Streu, P. (2011). The distribution of dissolved zinc in the Atlantic sector of the Southern Ocean. *Deep Sea Res. II Top. Stud. Oceanogr.* 58, 2707–2719. doi: 10.1016/j.dsr2.2010.10.041

Cuvelier, M. L., Allen, A. E., Monier, A., McCrow, J. P., Messi'e, M., Tringe, S. G., et al. (2010). Targeted metagenomics and ecology of globally important uncultured eukaryotic phytoplankton. *Proc. Natl. Acad. Sci. U.S.A.* 107, 14679–14684. doi: 10.1073/pnas.1001665107

DeVries, A. L., and Wohlschlag, D. E. (1969). Freezing resistance in some Antarctic fishes. *Science* 163, 1073–1075. doi: 10.1126/science.163.3871.1073

DeVries, T. (2014). The oceanic anthropogenic CO2 sink: storage, air-sea fluxes, and transports over the industrial era. *Glob. Biogeochem. Cycles* 28, 631–647. doi: 10.1002/2013GB004739

Dolev, M. B., Braslavsky, I., and Davies, P. L. (2016). Ice-binding proteins and their function. *Annu. Rev. Biochem.* 85, 515–542. doi: 10.1146/annurev-biochem-060815-014546

Duncan, A., Barry, K., Daum, C., Eloe-Fadrosh, E., Roux, S., Tringe, S., et al. (2020). Metagenome-assembled genomes of phytoplankton communities across the Arctic Circle. *BioRxiv* [Preprint]. doi: 10.1101/2020.06.16.154583

Dutkiewicz, A., Müller, R. D., O'Callaghan, S., and Jónasson, H. (2015). Census of seafloor sediments in the world's ocean. *Geology* 43, 795–798. doi: 10.1130/g36883.1

Eicken, H. (1992). The role of sea ice in structuring Antarctic ecosystems. *Polar Biol.* 12, 3–13. doi: 10.1007/BF00239960

Erwin, D. H. (2009). Climate as a driver of evolutionary change. *Curr. Biol.* 19, R575–R583. doi: 10.1016/j.cub.2009.05.047

Ewert, M., and Deming, J. (2011). Selective retention in saline ice of extracellular polysaccharides produced by the cold-adapted marine bacterium *Colwellia psychrerythraea* strain 34H. *Ann. Glaciol.* 52, 111–117. doi: 10.3189/172756411795931868

Faktorová, D., Nisbet, R. E. R., Fernández Robledo, J. A., Casacuberta, E., Sudek, L., Allen, A. E., et al. (2020). Genetic tool development in marine protists: emerging model organisms for experimental cell biology. *Nat. Methods* 17, 481–494. doi: 10.1038/s41592-020-0796-x

Falciatore, A., Jaubert, M., Bouly, J.-P., Bailleul, B., and Mock, T. (2020). Diatom molecular research comes of age: model species for studying phytoplankton biology and diversity. *Plant Cell* 32, 547–572. doi: 10.1105/tpc.19.00158

Falkowski, P. G. (1998). Biogeochemical controls and feedbacks on ocean primary production. *Science* 281, 200–206. doi: 10.1126/science.281.5374.200

Field, C. B., Behrenfeld, M. J., Randerson, J. T., and Falkowski, P. (1998). Primary production of the biosphere: integrating terrestrial and oceanic components. *Science* 281:237. doi: 10.1126/science.281.5374.237

Firdaus-Raih, M., Hashim, N. H. F., Bharudin, I., Abu Bakar, M. F., Huang, K. K., Alias, H., et al. (2018). The *Glaciozyma antarctica* genome reveals an array of systems that provide sustained responses towards temperature variations in a persistently cold habitat. *PLoS One* 13:e0189947. doi: 10.1371/journal.pone.0189947

Gobler, C. J., Berry, D. L., Dyhrman, S. T., Wilhelm, S. W., Salamov, A., Lobanov, A. V., et al. (2011). Niche of harmful alga *Aureococcus anophagefferens* revealed through ecogenomics. *Proc. Natl. Acad. Sci. U.S.A.* 108, 4352–4357. doi: 10.1073/pnas.1016106108

Gregory, A. C., Zayed, A. A., Conceição-Neto, N., Temperton, B., Bolduc, B., Alberti, A., et al. (2019). Marine DNA viral macro-and microdiversity from pole to pole. *Cell* 177, 1109–1123. doi: 10.1016/j.cell.2019.03.040

Hansen, M. M., Olivieri, I., Waller, D. M., Nielsen, E. E., and Group, G. W. (2012). Monitoring adaptive genetic responses to environmental change. *Mol. Ecol.* 21, 1311–1329. doi: 10.1111/j.1365-294X.2011.05463.x

Hansom, J. D., and Gordon, J. E. (1998). *Antarctic Environments and Resources: A Geographical Perspective*, 1st Edn. Abingdon: Routledge.

Harrison, W. G., and Cota, G. F. (1991). Primary production in polar waters: relation to nutrient availability. *Polar Res.* 10, 87–104. doi: 10.3402/polar.v10i1.6730

Hassler, C. S., Schoemann, V., Boye, M., Tagliabue, A., Rozmarynowycz, M., and McKay, R. M. L. (2012). "Iron bioavailability in the Southern Ocean," in *Oceanography and Marine Biology: An Annual Review*, eds R. N. Gibson, R. J. A. Atkinson, J. D. M. Gordon, and R. N. Hughes (Boca Raton, FL: CRC Press-Taylor & Francis Group).

Hayward, A., and Grigor, J. (2020). The bottom of the Arctic's food web is of top importance. *Front. Young Minds* 8:122. doi: 10.3389/frym.2020.00122

Hoagland, K. D., Rosowski, J. R., Gretz, M. R., and Roemer, S. C. (1993). Diatom Extracellular polymeric substances: function, fine structure, chemistry, and physiology. *J. Phycol.* 29, 537–566. doi: 10.1111/j.0022-3646.1993.00537.x

Iversen, C. M., Sloan, V. L., Sullivan, P. F., Euskirchen, E. S., McGuire, A. D., Norby, R. J., et al. (2015). The unseen iceberg: plant roots in arctic tundra. *New Phytol.* 205, 34–58. doi: 10.1111/nph.13003

Jung, W., Campbell, R. L., Gwak, Y., Kim, J. I., Davies, P. L., and Jin, E. (2016). New cysteine-rich ice-binding protein secreted from antarctic microalga, *Chloromonas* sp. *PLoS One* 11:e0154056. doi: 10.1371/journal.pone.0154056

Keeling, P. J., and Palmer, J. D. (2008). Horizontal gene transfer in eukaryotic evolution. *Nat. Rev. Genet.* 9, 605–618. doi: 10.1038/nrg2386

Kennedy, F., Martin, A., Bowman, J. P., Wilson, R., and McMinn, A. (2019). Dark metabolism: a molecular insight into how the Antarctic sea-ice diatom *Fragilariopsis cylindrus* survives long-term darkness. *New Phytol.* 223, 675–691. doi: 10.1111/nph.15843

Khatiwala, S., Primeau, F., and Hall, T. (2009). Reconstruction of the history of anthropogenic CO2 concentrations in the ocean. *Nature* 462, 346–349. doi: 10.1038/nature08526

Kondo, H., Mochizuki, K., and Bayer-Giraldi, M. (2018). Multiple binding modes of a moderate ice-binding protein from a polar microalga. *Phys. Chem. Chem. Phys.* 20, 25295–25303. doi: 10.1039/c8cp04727h

Kooistra, W. H. C. F., Gersonde, R., Medlin, L. K., and Mann, D. G. (2007). "The origin and evolution of the diatoms: their adaptation to a planktonic existence," in *Evolution of Primary Producers in the Sea*, eds P. Falkowski and A. H. Knoll (Cambridge, MA: Academic Press), 207–249.

Krembs, C., Eicken, H., and Deming, J. W. (2011). Exopolymer alteration of physical properties of sea ice and implications for ice habitability and biogeochemistry in a warmer Arctic. *Proc. Natl. Acad. Sci. U.S.A.* 108, 3653–3658. doi: 10.1073/pnas.1100701108

Kvernvik, A. C., Hoppe, C. J. M., Lawrenz, E., Prasil, O., Greenacre, M., Wiktor, J. M., et al. (2018). Fast reactivation of photosynthesis in arctic phytoplankton during the polar night. *J. Phycol.* 54, 461–470. doi: 10.1111/jpy.12750

Kwok, R. (2018). Arctic sea ice thickness, volume, and multiyear ice coverage: losses and coupled variability (1958–2018). *Environ. Res. Lett.* 13:105005. doi: 10.1088/1748-9326/aae3ec

Laity, J. H., Lee, B. M., and Wright, P. E. (2001). Zinc finger proteins: new insights into structural and functional diversity. *Curr. Opin. Struct. Biol.* 11, 39–46. doi: 10.1016/s0959-440x(00)00167-6

Laufkötter, C., Stern, A. A., John, J. G., Stock, C. A., and Dunne, J. P. (2018). Glacial iron sources stimulate the southern ocean carbon cycle. *Geophys. Res. Lett.* 45, 377–413. doi: 10.1029/2018GL079797

Lee, S. (2014). A theory for polar amplification from a general circulation perspective. *Asia Pac. J. Atmos. Sci.* 50, 31–43. doi: 10.1007/s13143-014-0024-7

Liang, Y., Koester, J. A., Liefer, J. D., Irwin, A. J., and Finkel, Z. V. (2019). Molecular mechanisms of temperature acclimation and adaptation in marine diatoms. *ISME J.* 13, 2415–2425. doi: 10.1038/s41396-019-0441-9

Lin, Y., Moreno, C., Marchetti, A., Ducklow, H., Schofield, O., Delage, E., et al. (2021). Decline in plankton diversity and carbon flux with reduced sea ice extent along the Western Antarctic Peninsula. *Nat. Commun.* 12:4948. doi: 10.1038/s41467-021-25235-w

Lommer, M., Specht, M., Roy, A. S., Kraemer, L., Andreson, R., Gutowska, M. A., et al. (2012). Genome and low-iron response of an oceanic diatom adapted to chronic iron limitation. *Genome Biol.* 13:20. doi: 10.1186/gb-2012-13-7-r66

Long, M. C., Stephens, B. B., McKain, K., Sweeney, C., Keeling, R. F., Kort, E. A., et al. (2021). Strong Southern Ocean carbon uptake evident in airborne observations. *Science* 374, 1275–1280. doi: 10.1126/science.abi4355

Lyon, B. R., and Mock, T. (2014). Polar microalgae: new approaches towards understanding adaptations to an extreme and changing environment. *Biology* 3, 56–80. doi: 10.3390/biology3010056

MacGilchrist, G. A., Naveira Garabato, A. C., Tsubouchi, T., Bacon, S., Torres-Valdés, S., and Azetsu-Scott, K. (2014). The Arctic Ocean carbon sink. *Deep Sea Res. I Oceanogr. Res. Pap.* 86, 39–55. doi: 10.1016/j.dsr.2014.01.002

Maksym, T. (2019). Arctic and Antarctic Sea ice change: contrasts, commonalities, and causes. *Annu. Rev. Mar. Sci.* 11, 187–213. doi: 10.1146/annurev-marine-010816-060610

Malviya, S., Scalco, E., Audic, S., Vincent, F., Veluchamy, A., Poulain, J., et al. (2016). Insights into global diatom distribution and diversity in the world's ocean. *Proc. Natl. Acad. Sci. U.S.A.* 113, E1516–E1525. doi: 10.1073/pnas.1509523113

Marchetti, A. (2019). A global perspective on iron and plankton through the Tara Oceans Lens. *Glob. Biogeochem. Cycles* 33, 239–242. doi: 10.1029/2019gb006181

Marchetti, A., Parker, M. S., Moccia, L. P., Lin, E. O., Arrieta, A. L., Ribalet, F., et al. (2009). Ferritin is used for iron storage in bloom-forming marine pennate diatoms. *Nature* 457, 467–470. doi: 10.1038/nature07539

Martin, J. H., Gordon, M., and Fitzwater, S. E. (1991). The case for iron. *Limnol. Oceanogr.* 36, 1793–1802. doi: 10.4319/lo.1991.36.8.1793

Martin, K., Schmidt, K., Toseland, A., Boulton, C. A., Barry, K., Beszteri, B., et al. (2021). The biogeographic differentiation of algal microbiomes in the upper ocean from pole to pole. *Nat. Commun.* 12:5483. doi: 10.1038/s41467-021-25646-9

Martin, W., Stoebe, B., Goremykin, V., Hansmann, S., Hasegawa, M., and Kowallik, K. V. (1998). Gene transfer to the nucleus and the evolution of chloroplasts. *Nature* 393:162. doi: 10.1038/30234

McCain, J. S. P., Allen, A. E., and Bertrand, E. M. (2021). Proteomic traits vary across taxa in a coastal Antarctic phytoplankton bloom. *ISME J.* 16, 569–579. doi: 10.1038/s41396-021-01084-9

McMinn, A., and Martin, A. (2013). Dark survival in a warming world. *Proc. Biol. Sci.* 280:20122909. doi: 10.1098/rspb.2012.2909

Miteva-Staleva, J., Stefanova, T., Krumova, E., and Angelova, M. (2011). Growth-phase-related changes in reactive oxygen species generation as a cold stress response in Antarctic *Penicillium* strains. *Biotechnol. Biotechnol. Equ.* 25, 58–63. doi: 10.5504/bbeq.2011.0131

Mock, T., Daines, S. J., Geider, R., Collins, S., Metodiev, M., Millar, A. J., et al. (2016). Bridging the gap between omics and earth system science to better understand how environmental change impacts marine microbes. *Glob. Change Biol.* 22, 61–75. doi: 10.1111/gcb.12983

Mock, T., Otillar, R. P., Strauss, J., McMullan, M., Paajanen, P., Schmutz, J., et al. (2017). Evolutionary genomics of the cold-adapted diatom *Fragilariopsis cylindrus*. *Nature* 541, 536–540. doi: 10.1038/nature20803

Moreno, C. M., Gong, W., Cohen, N. R., DeLong, K., and Marchetti, A. (2020). Interactive effects of iron and light limitation on the molecular physiology of the Southern Ocean diatom *Fragilariopsis kerguelensis*. *Limnol. Oceanogr.* 65, 1511–1531. doi: 10.1002/lno.11404

Moreno, C. M., Lin, Y., Davies, S., Monbureau, E., Cassar, N., and Marchetti, A. (2018). Examination of gene repertoires and physiological responses to iron and light limitation in Southern Ocean diatoms. *Polar Biol.* 41, 679–696. doi: 10.1007/s00300-017-2228-7

Morgan-Kiss, R. M., Lizotte, M. P., Kong, W., and Priscu, J. C. (2016). Photoadaptation to the polar night by phytoplankton in a permanently ice-covered Antarctic lake. *Limnol. Oceanogr.* 61, 3–13. doi: 10.1002/lno.10107

Murata, A., and Takizawa, T. (2003). Summertime CO2 sinks in shelf and slope waters of the western Arctic Ocean. *Continent. Shelf Res.* 23, 753–776. doi: 10.1016/S0278-4343(03)00046-3

Murphy, E. J., Cavanagh, R. D., Drinkwater, K. F., Grant, S. M., Heymans, J. J., Hofmann, E. E., et al. (2016). Understanding the structure and functioning of polar pelagic ecosystems to predict the impacts of change. *Proc. Biol. Sci.* 283:1646. doi: 10.1098/rspb.2016.1646

Mykytczuk, N. C. S., Foote, S. J., Omelon, C. R., Southam, G., Greer, C. W., and Whyte, L. G. (2013). Bacterial growth at-15 degrees C; molecular insights from the permafrost bacterium *Planococcus halocryophilus* Or1. *ISME J.* 7, 1211–1226. doi: 10.1038/ismej.2013.8

Nelson, D. M., Smith, W. O. Jr., Gordon, L. I., and Huber, B. A. (1987). Spring distributions of density, nutrients, and phytoplankton biomass in the ice edge zone of the Weddell-Scotia Sea. *J. Geophys. Res. Oceans* 92, 7181–7190.

Obiol, A., Giner, C. R., Sánchez, P., Duarte, C. M., Acinas, S. G., and Massana, R. (2020). A metagenomic assessment of microbial eukaryotic diversity in the global ocean. *Mol. Ecol. Resour.* 20, 718–731. doi: 10.1111/1755-0998.13147

Ovaskainen, O., Tikhonov, G., Norberg, A., Guillaume Blanchet, F., Duan, L., Dunson, D., et al. (2017). How to make more out of community data? A

conceptual framework and its implementation as models and software. *Ecol. Lett.* 20, 561–576. doi: 10.1111/ele.12757

Palenik, B. (2015). Molecular mechanisms by which marine phytoplankton respond to their dynamic chemical environment. *Annu. Rev. Mar. Sci.* 7, 325–340. doi: 10.1146/annurev-marine-010814-015639

Pearson, G. A., Lago-Leston, A., Cánovas, F., Cox, C. J., Verret, F., Lasternas, S., et al. (2015). Metatranscriptomes reveal functional variation in diatom communities from the Antarctic Peninsula. *ISME J.* 9, 2275–2289. doi: 10.1038/ismej.2015.40

Peers, G., and Price, N. M. (2006). Copper-containing plastocyanin used for electron transport by an oceanic diatom. *Nature* 441, 341–344. doi: 10.1038/nature04630

Peters, E., and Thomas, D. N. (1996). Prolonged darkness and diatom mortality I: marine Antarctic species. *J. Exp. Mar. Biol. Ecol.* 207, 25–41. doi: 10.1016/S0022-0981(96)02520-8

Petrou, K., Trimborn, S., Rost, B., Ralph, P. J., and Hassler, C. S. (2014). The impact of iron limitation on the physiology of the Antarctic diatom Chaetoceros simplex. *Mar. Biol.* 161, 925–937. doi: 10.1007/s00227-014-2392-z

Pinseel, E., Janssens, S. B., Verleyen, E., Vanormelingen, P., Kohler, T. J., Biersma, E. M., et al. (2020). Global radiation in a rare biosphere soil diatom. *Nat. Commun.* 11:2382. doi: 10.1038/s41467-020-16181-0

Pitchford, J. W., and Brindley, J. (1999). Iron limitation, grazing pressure and oceanic high nutrient-low chlorophyll (HNLC) regions. *J. Plankton Res.* 21, 525–547. doi: 10.1093/plankt/21.3.525

Pointing, S., Buedel, B., Convey, P., Gillman, L., Koerner, C., Leuzinger, S., et al. (2015). Biogeography of photoautotrophs in the high polar biome. *Front. Plant Sci.* 6:692. doi: 10.3389/fpls.2015.00692

Postel, U., Glemser, B., Salazar Alekseyeva, K., Eggers, S. L., Groth, M., Glöckner, G., et al. (2020). Adaptive divergence across Southern Ocean gradients in the pelagic diatom *Fragilariopsis kerguelensis*. *Mol. Ecol.* 29, 4913–4924. doi: 10.1111/mec.15554

Raymond, J. A. (2014). The ice-binding proteins of a snow alga, *Chloromonas brevispina*: probable acquisition by horizontal gene transfer. *Extremophiles* 18, 987–994. doi: 10.1007/s00792-014-0668-3

Raymond, J. A., Janech, M. G., and Fritsen, C. H. (2009). Novel ice-binding proteins from a psychrophilic Antarctic alga (Chlamydomonadaceae, Chlorophyceae). *J. Phycol.* 45, 130–136. doi: 10.1111/j.1529-8817.2008.00623.x

Raymond, J. A., Janech, M. G., and Mangiagalli, M. (2021). Ice-binding proteins associated with an Antarctic *Cyanobacterium*, Nostoc sp. HG1. *Appl. Environ. Microbiol.* 87:e02499-20. doi: 10.1128/AEM.02499-20

Raymond, J. A., and Kim, H. J. (2012). Possible role of horizontal gene transfer in the colonization of Sea Ice by Algae. *PLoS One* 7:e35968. doi: 10.1371/journal.pone.0035968

Raymond, J. A., and Morgan-Kiss, R. (2017). Multiple ice-binding proteins of probable prokaryotic origin in an antarctic Lake Alga, *Chlamydomonas* sp ICE-MDV (Chlorophyceae). *J. Phycol.* 53, 848–854. doi: 10.1111/jpy.12550

Raymond-Bouchard, I., Tremblay, J., Altshuler, I., Greer, C. W., and Whyte, L. G. (2018). Comparative transcriptomics of cold growth and adaptive features of a eury- and steno-psychrophile. *Front. Microbiol.* 9:1565. doi: 10.3389/fmicb.2018.01565

Reeves, S., McMinn, A., and Martin, A. (2011). The effect of prolonged darkness on the growth, recovery and survival of Antarctic sea ice diatoms. *Polar Biol.* 34, 1019–1032. doi: 10.1007/s00300-011-0961-x

Royo-Llonch, M., Sánchez, P., Ruiz-González, C., Salazar, G., Pedrós-Alió, C., Sebastián, M., et al. (2021). Compendium of 530 metagenome-assembled bacterial and archaeal genomes from the polar Arctic Ocean. *Nat. Microbiol.* 6, 1561–1574. doi: 10.1038/s41564-021-00979-9

Sackett, O., Petrou, K., Reedy, B., De Grazia, A., Hill, R., Doblin, M., et al. (2013). Phenotypic plasticity of Southern Ocean diatoms: key to success in the sea ice habitat? *PLoS One* 8:e81185. doi: 10.1371/journal.pone.0081185

Salamy, K. A., and Zachos, J. C. (1999). Latest Eocene–Early Oligocene climate change and Southern Ocean fertility: inferences from sediment accumulation and stable isotope data. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 145, 61–77. doi: 10.1016/S0031-0182(98)00093-5

Schaub, I., Wagner, H., Graeve, M., and Karsten, U. (2017). Effects of prolonged darkness and temperature on the lipid metabolism in the benthic diatom *Navicula perminuta* from the Arctic Adventfjorden, Svalbard. *Polar Biol.* 40, 1425–1439. doi: 10.1007/s00300-016-2067-y

Scholz, B., Rua, A., and Liebezeit, G. (2014). Effects of UV radiation on five marine microphytobenthic Wadden sea diatoms isolated from the Solthorn tidal flat (Lower Saxony, southern North Sea) - Part I: growth and antioxidative defence strategies. *Eur. J. Phycol.* 49, 68–82. doi: 10.1080/09670262.2014.889214

Sigman, D. M., Hain, M. P., and Haug, G. H. (2010). The polar ocean and glacial cycles in atmospheric CO2 concentration. *Nature* 466, 47–55. doi: 10.1038/nature09149

Singh, J., Singh, R. P., and Khare, R. (2018). Influence of climate change on Antarctic flora. *Polar Sci.* 18, 94–101. doi: 10.1016/j.polar.2018.05.006

Sjöqvist, C. O., and Kremp, A. (2016). Genetic diversity affects ecological performance and stress response of marine diatom populations. *ISME J.* 10, 2755–2766. doi: 10.1038/ismej.2016.44

Smythe-Wright, D., Cunningham, S. A., Lampitt, R. A., Kent, E. C., King, B. A., Quartly, G. D., et al. (2010). "Sustained observations in the Atlantic and Southern Oceans," in *Proceedings of OceanObs'09: Sustained Ocean Observations and Information for Society*, eds J. Hall, D. E. Harrison, and D. Stammer (Paris: European Space Agency).

Soreide, J. E., Leu, E., Berge, J., Graeve, M., and Falk-Petersen, S. (2010). Timing of blooms, algal food quality and *Calanus glacialis* reproduction and growth in a changing Arctic. *Glob. Change Biol.* 16, 3154–3163. doi: 10.1111/j.1365-2486.2010.02175.x

Steig, E. J., Schneider, D. P., Rutherford, S. D., Mann, M. E., Comiso, J. C., and Shindell, D. T. (2009). Warming of the Antarctic ice-sheet surface since the 1957 International Geophysical Year. *Nature* 457, 459–462. doi: 10.1038/nature07669

Stephens, T. G., González-Pech, R. A., Cheng, Y., Mohamed, A. R., Burt, D. W., Bhattacharya, D., et al. (2020). Genomes of the dinoflagellate *Polarella glacialis* encode tandemly repeated single-exon genes with adaptive functions. *BMC Biol.* 18:56. doi: 10.1186/s12915-020-00782-8

Stretch, J. J., Hamner, P. P., Hamner, W. M., Michel, W. C., Cook, J., and Sullivan, C. W. (1988). Foraging behavior of Antarctic krill Euphausia-superba on sea ice microalgae. *Mar. Ecol. Prog. Ser.* 44, 131–139. doi: 10.3354/meps044131

Suto, I., Kawamura, K., Hagimoto, S., Teraishi, A., and Tanaka, Y. (2012). Changes in upwelling mechanisms drove the evolution of marine organisms. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 339, 39–51. doi: 10.1016/j.palaeo.2012.04.014

Takahashi, T., Sutherland, S. C., Wanninkhof, R., Sweeney, C., Feely, R. A., Chipman, D. W., et al. (2009). Climatological mean and decadal change in surface ocean pCO2, and net sea–air CO2 flux over the global oceans. *Deep Sea Res. II Top. Stud. Oceanogr.* 56, 554–577. doi: 10.1016/j.dsr2.2008.12.009

Thomas, D. N., and Dieckmann, G. S. (2002). Antarctic sea ice — a habitat for extremophiles. *Science* 295:641. doi: 10.1126/science.1063391

Tikhonov, G., Opedal, ØH., Abrego, N., Lehikoinen, A., de Jonge, M. M. J., Oksanen, J., et al. (2020). Joint species distribution modelling with the r-package Hmsc. *Methods Ecol. Evol.* 11, 442–447. doi: 10.1111/2041-210X.13345

Tonelli, M., Signori, C. N., Bendia, A., Neiva, J., Ferrero, B., Pellizari, V., et al. (2021). Climate projections for the southern ocean reveal impacts in the marine microbial communities following increases in sea surface temperature. *Front. Mar. Sci.* 8:636226. doi: 10.3389/fmars.2021.636226

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

Trefault, N., De la Iglesia, R., Moreno-Pino, M., Lopes dos Santos, A., Gérikas Ribeiro, C., Parada-Pozo, G., et al. (2021). Annual phytoplankton dynamics in coastal waters from Fildes Bay, Western Antarctic Peninsula. *Sci. Rep.* 11:1368. doi: 10.1038/s41598-020-80568-8

Tréguer, P., Bowler, C., Moriceau, B., Dutkiewicz, S., Gehlen, M., Aumont, O., et al. (2018). Influence of diatom diversity on the ocean biological carbon pump. *Nat. Geosci.* 11, 27–37. doi: 10.1038/s41561-017-0028-x

van de Poll, W. H., Abdullah, E., Visser, R. J. W., Fischer, P., and Buma, A. G. J. (2020). Taxon-specific dark survival of diatoms and flagellates affects Arctic phytoplankton composition during the polar night and early spring. *Limnol. Oceanogr.* 65, 903–914. doi: 10.1002/lno.11355

Vance, T. D. R., Bayer-Giraldi, M., Davies, P. L., and Mangiagalli, M. (2019). Ice-binding proteins and the 'domain of unknown function' 3494 family. *FEBS J.* 286, 855–873. doi: 10.1111/febs.14764

Venables, H., and Moore, M. (2010). Phytoplankton and light limitation in the Southern Ocean: learning from high-nutrient, high-chlorophyll areas. *J. Geophys. Res.* 115:C02015. doi: 10.1029/2009JC005361

Verde, C., Giordano, D., Bellas, C. M., di Prisco, G., and Anesio, A. M. (2016). Polar marine microorganisms and climate change. *Adv. Microb. Physiol.* 69, 187–215. doi: 10.1016/bs.ampbs.2016.07.002

Vincent, W. F. (2010). Microbial ecosystem responses to rapid climate change in the Arctic. *ISME J.* 4, 1087–1090. doi: 10.1038/ismej.2010.108

Wadham, J. L., Hawkings, J. R., Tarasov, L., Gregoire, L. J., Spencer, R. G. M., Gutjahr, M., et al. (2019). Ice sheets matter for the global carbon cycle. *Nat. Commun.* 10:3567. doi: 10.1038/s41467-019-11394-4

Wake, B. (2019). A drift in the Arctic. *Nat. Clim. Change* 9:733. doi: 10.1038/s41558-019-0597-3

Waldvogel, A.-M., Feldmeyer, B., Rolshausen, G., Exposito-Alonso, M., Rellstab, C., Kofler, R., et al. (2020). Evolutionary genomics can improve prediction of species' responses to climate change. *Evol. Lett.* 4, 4–18. doi: 10.1002/evl3.154

Walker, D. A., Raynolds, M. K., Daniëls, F. J. A., Einarsson, E., Elvebakk, A., Gould, W. A., et al. (2005). The Circumpolar Arctic vegetation map. *J. Veg. Sci.* 16, 267–282. doi: 10.1111/j.1654-1103.2005.tb02365.x

Wietz, M., Bienhold, C., Metfies, K., Torres-Valdés, S., von Appen, W. J., Salter, I., et al. (2021). The polar night shift: annual dynamics and drivers of microbial community structure in the arctic ocean. *bioRxiv* [Preprint]. doi: 10.1101/2021.04.08.436999w

Wingsle, G., Karpinski, S., and Hallgren, J. E. (1999). Low temperature, high light stress and antioxidant defence mechanisms in higher plants. *Phyton Ann. Rei Bot.* 39, 253–268.

Wolf, K. K. E., Romanelli, E., Rost, B., John, U., Collins, S., Weigand, H., et al. (2019). Company matters: the presence of other genotypes alters traits and intraspecific selection in an Arctic diatom under climate change. *Glob. Change Biol.* 25, 2869–2884. doi: 10.1111/gcb.14675

Wong, C., Boo, S. Y., Voo, C. L. Y., Zainuddin, N., and Najimudin, N. (2019). A comparative transcriptomic analysis provides insights into the cold-adaptation mechanisms of a psychrophilic yeast, *Glaciozyma antarctica* PI12. *Polar Biol.* 42, 541–553. doi: 10.1007/s00300-018-02443-7

Yergeau, E., Michel, C., Tremblay, J., Niemi, A., King, T. L., Wyglinski, J., et al. (2017). Metagenomic survey of the taxonomic and functional microbial communities of seawater and sea ice from the Canadian Arctic. *Sci. Rep.* 7:42242.

Zhang, X., Cvetkovska, M., Morgan-Kiss, R., Hüner, N. P. A., and Smith, D. R. (2021). Draft genome sequence of the Antarctic green alga *Chlamydomonas* sp, UWO241. *iScience* 24:102084. doi: 10.1016/j.isci.2021.102084

Zhang, Z., Qu, C., Zhang, K., He, Y., Zhao, X., Yang, L., et al. (2020). Adaptation to extreme antarctic environments revealed by the genome of a sea ice green alga. *Curr. Biol.* 30, 3330.e7–3341.e7. doi: 10.1016/j.cub.2020.06.029

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

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Microbiology

ORIGINAL RESEARCH
published: 09 May 2022
doi: 10.3389/fmicb.2022.782789

Bacterial Colonisation: From Airborne Dispersal to Integration Within the Soil Community

Lucie A. Malard[1,2]* *and David A. Pearce*[1,3]*

[1]Department of Applied Sciences, Faculty of Health and Life Sciences, Northumbria University, Newcastle-upon-Tyne, United Kingdom, [2]Department of Ecology and Evolution, University of Lausanne, Lausanne, Switzerland, [3]British Antarctic Survey, Natural Environment Research Council, Cambridge, United Kingdom

The deposition of airborne microorganisms into new ecosystems is the first stage of colonisation. However, how and under what circumstances deposited microorganisms might successfully colonise a new environment is still unclear. Using the Arctic snowpack as a model system, we investigated the colonisation potential of snow-derived bacteria deposited onto Arctic soils during and after snowmelt using laboratory-based microcosm experiments to mimic realistic environmental conditions. We tested different melting rate scenarios to evaluate the influence of increased precipitation as well as the influence of soil pH on the composition of bacterial communities and on the colonisation potential. We observed several candidate colonisations in all experiments; with a higher number of potentially successful colonisations in acidoneutral soils, at the average snowmelt rate measured in the Arctic. While the higher melt rate increased the total number of potentially invading bacteria, it did not promote colonisation (snow ASVs identified in the soil across multiple sampling days and still present on the last day). Instead, most potential colonists were not identified by the end of the experiments. On the other hand, soil pH appeared as a determinant factor impacting invasion and subsequent colonisation. In acidic and alkaline soils, bacterial persistence with time was lower than in acidoneutral soils, as was the number of potentially successful colonisations. This study demonstrated the occurrence of potentially successful colonisations of soil by invading bacteria. It suggests that local soil properties might have a greater influence on the colonisation outcome than increased precipitation or ecosystem disturbance.

Keywords: Arctic ecosystems, airborne dispersal, microbial colonisation, bacterial diversity, snow, soil

OPEN ACCESS

Edited by:
Anne D. Jungblut, Natural History Museum, United Kingdom

Reviewed by:
Henri M. P. Siljanen, University of Eastern Finland, Finland
Kevin C. Lee, Auckland University of Technology, New Zealand

***Correspondence:**
Lucie A. Malard
lucie.malard@unil.ch
David A. Pearce
david.pearce@northumbria.ac.uk

Specialty section:
This article was submitted to Extreme Microbiology, a section of the journal Frontiers in Microbiology

Received: *24 September 2021*
Accepted: *04 April 2022*
Published: *09 May 2022*

Citation:
Malard LA and Pearce DA (2022) Bacterial Colonisation: From Airborne Dispersal to Integration Within the Soil Community. Front. Microbiol. 13:782789. doi: 10.3389/fmicb.2022.782789

INTRODUCTION

Global dispersal of microorganisms has primarily been shown to occur through airborne transport *via* the aerosolisation of particles (Smets et al., 2016). Once airborne, microorganisms can travel thousands of kilometres (Smith et al., 2013; Barberán et al., 2014, 2015; Maki et al., 2019) and be deposited into the most remote places on Earth, from the Arctic (Harding et al., 2011; Cuthbertson et al., 2017; Šantl-Temkiv et al., 2018), to Antarctica (Pearce et al., 2009, 2010; Bottos et al., 2014; Archer et al., 2019). While microorganisms were once considered

to be ubiquitously distributed with limitless dispersal capability (De Wit and Bouvier, 2006), recent studies have refuted this hypothesis and shown that ecological drift and dispersal limitation significantly influence the distribution of microorganisms (Fierer and Jackson, 2006; Bahram et al., 2018; Delgado-Baquerizo et al., 2018; Malard et al., 2019a). Thus, airborne microorganisms have the potential to invade and colonise all ecosystems across the globe.

The first step of microbial colonisation is the deposition within the environment itself (Mallon et al., 2015a; Kinnunen et al., 2016). This deposition process is generally categorised as 'wet' or 'dry' deposition. Dry deposition occurs following adherence to buildings, plants, water or soil surfaces, while wet deposition is caused by precipitation events, such as rain or snow (Smets et al., 2016; Reche et al., 2018). While the constant deposition of microorganisms into new environments is accepted and has been demonstrated (Peter et al., 2014; Weil et al., 2017; Reche et al., 2018), whether the deposited microorganisms become established and colonise the new environment long term is still a subject of debate. For example, Fløistrup et al. (2018) demonstrated the quick colonisation of sterile soils by airborne microorganisms while Evans et al. (2019) suggested that dispersal *via* rainfall altered the soil microbial response to drought without actively demonstrating microbial colonisation.

The Arctic snowpack is an ideal model with which to study colonisation following wet deposition (by snowfall) since it covers the Arctic tundra for 8–10 months of the year while also isolating the soil from outside influence (Winther et al., 2003). The snowpack acts as an ephemeral transition ecosystem for airborne microorganisms. Indeed, although the snowpack is seeded by airborne microorganisms, its own unique microbial community develops over time (Harding et al., 2011; Larose et al., 2013; Els et al., 2019; Malard et al., 2019b). The snowpack community is well adapted to surrounding environmental conditions (Russell, 1990; Margesin and Miteva, 2011; Maccario et al., 2015), one of the key factors in successful colonisation (Mallon et al., 2015a; Kinnunen et al., 2016). Once the snow starts to melt, the run-off travels vertically under gravity on flat terrain, to reach the frozen soil layer and infiltrate the soil (Gray et al., 2001; Iwata et al., 2008; Larose et al., 2013). However, the percolation of meltwater is a complex process influenced by soil properties, such as temperature, moisture or the presence of ice, but also by the rate of snowmelt (Gray et al., 2001; Iwata et al., 2008). The snowpack is a major source of potential colonists as it supports between 10^1 and 10^4 microbial cells per mL (Amato et al., 2007; Zhang et al., 2010; Hauptmann et al., 2014; Cameron et al., 2015). Snowmelt also creates a peak in nutrient and solute availability in soils (Lipson and Schmidt, 2004; Edwards et al., 2006; Buckeridge and Grogan, 2010; Larose et al., 2013). This resource pulse may facilitate the colonisation of soils by snow microorganisms (Mallon et al., 2015b). Furthermore, ecosystem disturbance, such as the sudden addition of water (Fierer et al., 2003; Orwin et al., 2016), may promote successful colonisation (Liu et al., 2012; De Roy et al., 2013). Therefore, snowmelt may be an opportunity for snow microorganisms to establish in rich and diverse soil communities, which would otherwise be difficult to colonise (Jousset et al., 2011). Generally, studies of microbial invasion and colonisation investigate a single (or very few) invader taxa inoculated at a high density into manipulated communities (van Elsas et al., 2012; Acosta et al., 2015; Mallon et al., 2015b, 2018; Fløistrup et al., 2018; Li et al., 2019). In this study, we investigated changes in soil bacterial communities during snowmelt and evaluated the colonisation potential of snow bacteria deposited into Arctic soils during and after snowmelt using microcosm experiments. In order to examine these changes in a realistic scenario, the invading (snow) and invaded (soil) communities were not manipulated.

The first set of experiments evaluated the influence of increased precipitation on the colonisation potential by using different melt rate scenarios. We simulated a seasonal average and a fast-saturating snowmelt rate to test the hypothesis that increased colonisation would be observed in the soils influenced by a fast flow rate due to a higher number of microorganisms deposited as well as increased ecosystem disturbance (analogue to increased precipitation). Indeed, under the current climate trend, precipitation (either rain or snowfall) is expected to increase in the region (Bintanja and Selten, 2014; Bintanja and Andry, 2017), and the resulting deeper snowpack may melt faster due to higher temperatures (Wipf and Rixen, 2010). Whether such an increase in microorganisms and available water resulting from a deeper snowpack will influence the colonisation success remains to be determined.

The second set of experiments investigated the colonisation potential of microorganisms deposited into soil ecosystems with different pH ranges. In Arctic soils, as elsewhere, pH is the primary driver of microbial community structure and diversity (Fierer and Jackson, 2006; Männistö et al. 2007; Chu et al., 2010; Malard and Pearce, 2018; Malard et al., 2019a), demonstrating that climatic conditions are not the only abiotic factor invaders have to adapt to. Soil pH in the Arctic region is primarily acidoneutral, although the full pH spectrum can be found locally (Hengl et al., 2017; Malard et al., 2019a). We evaluated the influence of soil pH on the colonisation potential during average snowmelt rates to test the hypothesis that increased colonisation would be observed in acidoneutral soils compared to acidic and alkaline soils. In many ways, acidic and alkaline soils are considered harsh environments requiring a wide range of adaptations (Glenn and Dilworth, 1991; Horikoshi, 1999; Torsvik and Øvreås, 2008). Although microorganisms within the snow can be expected to be well adapted to climatic conditions (Russell, 1990; Margesin and Miteva, 2011; Maccario et al., 2015), whether they have the capacity to adapt to harsh physicochemical properties or to fast changing environmental conditions to colonise soils remains unknown.

MATERIALS AND METHODS

Approach

In all experiments, melted snow was input into soil columns and after percolation, the outflow was recovered (**Figure 1**). In the first set of experiments, only flow rate was manipulated

(at acidoneutral soil pH) to evaluate the influence of increased precipitation (using snowmelt rate as a proxy). In the second set of experiments, only soil pH was manipulated (at constant average flow rate) to evaluate the influence of soil pH on the colonisation potential of snow bacteria into soils.

The snowmelt period in the Arctic lasts anywhere between 7 and 30 days, with an average of 15–20 days (Foster, 1989; Winther et al., 2002). Here, melted snow was introduced to the system every day for 14 days to simulate this snowmelt period, followed by 15 days snow-free to simulate the post-melt season. All experiments were conducted at 4°C to simulate Arctic climatic conditions. However, they were kept in the dark, unlike summer Arctic conditions, to avoid the formation of cyanobacterial mats and to avoid a light regime change after storage in the dark. Indeed, differentiating the impact of light and water regime change would have required a different set of experiments. Bacterial communities within each component (snow, soil, outflow) were monitored with time using qPCR and 16S rRNA gene amplicon sequencing.

Sample Collection and Properties

The soil used for microcosm experiments was collected in Adventdalen (78°10′12″N, 16°3′0″E), Svalbard in July 2018, from the top 15 cm using ethanol-cleaned trowels and Whirl-pak bags (Nasco, WI, United States). The average temperature in July in Longyearbyen is 4°C. The soil was passed through a 1.5 cm sterilised sieve in a class II biological safety cabinet (ESCO, Singapore) to remove large organic matter and larger soil particles. pH, conductivity, moisture and total organic carbon were measured in the laboratory (**Supplementary Table S1**) using the same methods as Malard et al. (2019a). To avoid freezing the samples and the associated potential cell death (Soulides and Allison, 1961; Yanai et al., 2004), and because temperature was a constant factor in the microcosms, all materials were stored in the dark at 4°C after field sampling (DS) and for 100 days until the beginning of the experiments and were maintained at 4°C during the experiments. On the day of field sampling, 1 g of soil was frozen at −20°C for DNA extraction and later referred to as the day of field sampling (DS).

The snow used for microcosms was collected in Whirl-pak bags (Nasco) using ethanol-cleaned shovels, close to Mine 7 in Svalbard in July 2018. The snow was left to melt at room temperature, transferred to sterile containers and preserved at 4°C in the dark until transportation to the United Kingdom. On the day of field sampling (DS), 250 ml of melted snow was filtered through a 0.22 µm Whatman nitrocellulose filter (Merck, Darmstadt, Germany) using an ethanol-cleaned filtration unit (Nalgene Nunc International Corporation) and frozen at −20°C.

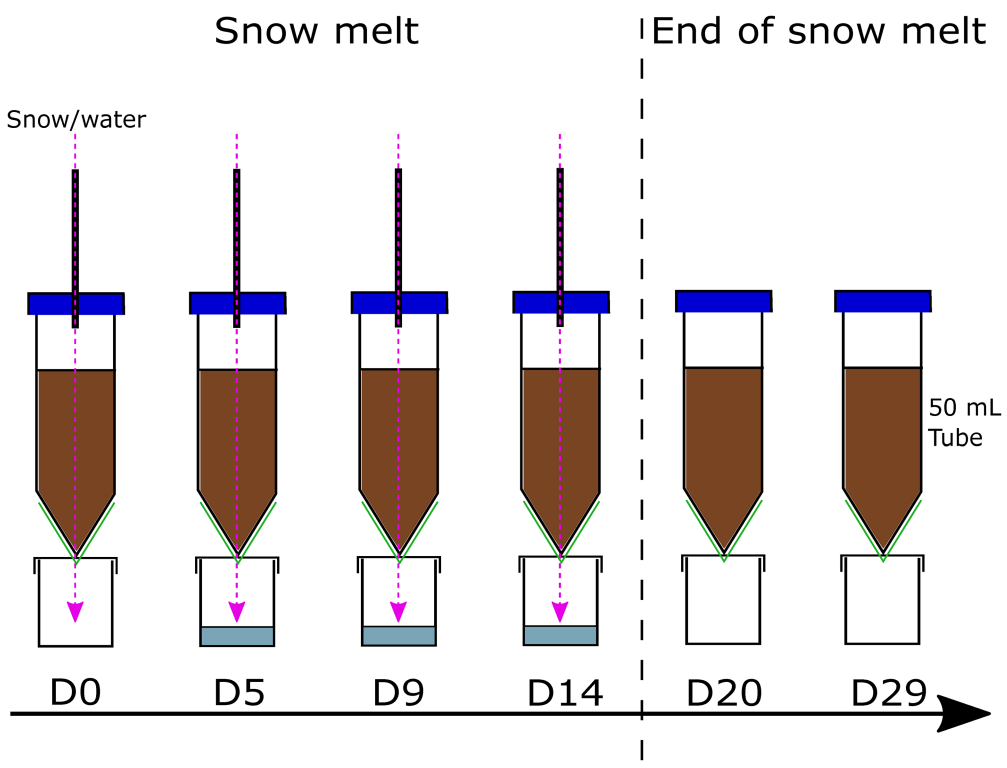

FIGURE 1 | Microcosm construction. The snow/sterile water was input using a programmable peristaltic dosing pump, through a hole in the centrifugal tube cap. To allow infiltration and evacuation of the excess liquid, a hole was made at the bottom of the tube, covered with a membrane (green) allowing the easy passage of liquid and cells while retaining the soil in the microcosm system. Snow/sterile water was input every day for 14 days and stopped thereafter; the hole in the centrifugal tube cap was covered with autoclaved aluminium to avoid contamination from outside the system.

Microcosm Construction

Prior to constructing the microcosms, a soil sample was frozen at −20°C and 250 ml of melted snow was filtered (using the same protocol as for DS) for DNA extraction, referred to as day 10 (D-10). Microcosms were constructed aseptically in a class II biological safety cabinet (ESCO) by adding 30 g of sieved soil in sterile 50 ml conical centrifugal tubes, packed at a density of approximately 0.96 g/cm^3, within the range of expected density relating to the soil organic carbon in Arctic soils (Hossain et al., 2015). Microcosms were left for 10 days to allow the soil bacterial communities acclimatise and adapt to the environmental conditions until the start of the experiment (D0). A multichannel peristaltic dosing pump (Jebao DP-4, DP-5 and DP-3S) was used for the delivery of melted snow (treatment) and sterile water (controls) into each microcosm. At the bottom of each tube, an evacuation hole was created using a sterile needle to let the excess liquid exit the system and avoid complete water saturation of the microcosm. A 0.45 µm Durapore membrane filter (Merck) was fitted at the bottom of the tube to let the water and microorganisms pass while stopping the soil from leaving the microcosm. The outflow, referred to as flow-through (FT), was collected in a sterile container, replaced on each sampling day (**Figure 1**). Triplicates of each experiment were run in parallel for a total of 12 microcosms per experiment. Soil samples were collected aseptically from the top of the column using sterile spatulas on day 0 (D0, first day of the experiment), day 5 (D5), day 9 (D9), day 14 (D14, last day of melted snow input), day 20 (D20) and day 29 (D29). The collected soil was frozen at −20°C until further processing. On each sampling day, 250 ml of melted snow was filtered (using the same protocol as for DS) and frozen at −20°C until further processing. For control microcosms, sterile, filtered MilliQ water was added to a clean and empty autoclaved container and autoclaved again. The sterility of the water was assessed on each sampling day by microscopy using Petroff-Hausser chambers and DNA extractions (**Supplementary Methods**). The flow-through output was filtered through a 0.22 µm nitrocellulose filter (Merck) on D5, D9 and D14 and frozen at −20°C until further processing.

On each sampling day, the pH of input snow/water and soil columns was measured using a 5 ml of snow/water or a 1:5 soil to water ratio and a Mettler-Toledo FE20 pH meter (Mettler-Toledo Instruments Co., Shanghai, China). The input snow pH (6.33 ± 0.15) and sterile water pH (5.82 ± 0.36) were acidoneutral. Overall, each microcosm remained within the pH category assigned throughout the duration of the experiments (**Supplementary Figure S1**).

Snowmelt Rate Experiments

In the Arctic, snow melts at an average rate of 9 mm water equivalent (we) per day (Marsh and Woo, 1984; Foster, 1989; Winther et al., 2002, 2003). This average flow rate (9 mm we/day) was equivalent to 6.4 ml/day for a 50 ml centrifugal tube, which formed the basis of the microcosms. In Svalbard, in cases of extremely rapid melt, rates up to 68 mm we/day have been recorded (summarised in Winther et al., 2002). At this fast rate, pilot experiments demonstrated the rapid saturation of the system, preventing the percolation of water through the soil. To simulate a fast melt rate at which the input water would saturate and percolate, a melting rate of 35 mm we/day was selected (equivalent to 24.7 ml/day for a 50 ml centrifugal tube). Using the peristaltic pump, a volume of 6.4 ml was introduced once a day for average rates and a volume of 6.2 ml was introduced four times a day, every 3 h (for 10 h) for the fast rate. Triplicates of each experiment were run as follows: average rates with melted snow (treatment 1), average rates with sterile water (control 1), fast rates with melted snow (treatment 2) and fast rates with sterile water (control 2).

Soil pH Experiments

On the day of microcosm set-up (D-10), the pH of the already acidoneutral soil (**Supplementary Table S1**) was adjusted to cover acidic and alkaline pH ranges. Around 0.35 g of aluminium sulphate [$Al_2(SO_4)_3$] was added to decrease soil pH and 0.44 g of calcium carbonate ($CaCO_3$) was added to increase soil pH. The optimum mass of each chemical to add to the soil was tested in a pilot experiment and in accordance with Nicol et al. (2008). A total of 12 microcosms were prepared, six acidic with pH = 3.65 ± 0.14 and six alkaline with pH = 7.97 ± 0.22 and left for 10 days at 4°C to let the soil bacterial communities acclimatise and adapt to the new environmental conditions. Using the average melt rate of 9 mm we/day, triplicates of each experiment were as follows: acidic soil with melted snow (treatment 1), acidic soil with sterile water (control 1), alkaline soil with melted snow (treatment 2) and alkaline soil with sterile water (control 2). The acidoneutral microcosms (pH = 5.31 ± 0.17) with average melt rate were run in the previous experiment (melt rate) and not repeated. Instead, the results were reused in different settings to compare the experiments.

DNA Extraction, Amplicon Sequencing, and Bioinformatic Processing

Soil DNA was extracted using the PowerSoil kit (Qiagen, Carlsbad, CA, United States) and following the manufacturers' protocol. Snow and flow-through (FT) DNA was extracted using the PowerWater kit (Qiagen) and following the manufacturers' protocol. Each extract was PCR amplified using the universal primers 515F-806R (Caporaso et al., 2010). Resulting amplicons were cleaned, normalised, pooled, sequenced on the Illumina MiSeq (as described in Malard et al., 2019a) and resulting amplicons were processed using the DADA2 pipeline v1.22 (Callahan et al., 2016). Forward and reverse read pairs were trimmed and filtered, with forward reads truncated at 230 base pairs (bp) and reverse reads at 200 bp, no ambiguous bases allowed, and each read required to have <2 expected errors based on their quality scores. Amplicon sequence variants (ASVs) were independently inferred from the forward and reverse reads of each sample using the run-specific error rates. Reads were dereplicated, pairs were merged, and chimeras were removed. Taxonomic assignment

was performed against the SILVA v128 database (Pruesse et al., 2007; Quast et al., 2012) using the implementation of the RDP (ribosomal database project) naive Bayesian classifier. A total of 6,715,429 reads (corresponding to ±30,114 reads/samples) were assigned against 20,583 ASVs.

16S rRNA qPCR

Quantitative real-time PCR (qPCR) was performed on a Bio-Rad CFX96 thermal cycler to quantify copy number of the bacterial 16S rRNA gene (Suzuki et al., 2000). PCR reactions were performed using the QuantiNova SYBR Green PCR kit (Qiagen) and 0.3 μM of universal bacterial 16S rRNA gene primers 1369F and 1492R (detailed protocol in **Supplementary Methods**). The results were normalised by the mass of soil or the volume of snow or FT filtered.

Statistical Analysis

All statistical analyses and visualisations were performed in the R environment using primarily a combination of the vegan (Dixon, 2003), phyloseq (McMurdie and Holmes, 2013) and ggplot2 (Wickham, 2016) packages. The decontam package (Davis et al., 2018) was used to identify potential contaminants using the prevalence function. The ASV table was also manually curated to discard ASVs present in the kit and MiSeq controls in higher abundance than in other samples, leaving 19,081 ASVs. The rarefaction curves saturated, suggesting that we reached the diversity plateau (**Supplementary Figure S2**). Bacterial richness and diversity indices (Kim et al., 2017) were calculated in phyloseq. Differences in 16S rRNA gene abundance and alpha diversity between sample type (snow, soil, FT), experiment (melt rate or pH), treatment (sterile water or snow) and sampling day and the interaction of all factors were assessed by a multi-factorial design using ANOVA and Tukey's Honest Significant Difference (HSD) tests with Bonferroni correction.

The ASV table was normalised to the relative abundance and used to evaluate changes in community composition. PERMANOVA were conducted using the adonis function with 999 permutations to identify significant differences in bacterial composition using the Bray–Curtis community dissimilarity and further observed using principal coordinate analysis (PCoA; Ramette, 2007; Paliy and Shankar, 2016). Betadisper (vegan) with 999 permutations was used to test whether the investigated groups were homogeneously dispersed.

To evaluate the colonisation potential, the ASVs identified in the control microcosms and pre-treated soils [day of field sampling (DS), set-up day (D-10) and first day of the experiment (D0)] were discarded from the treated ASV tables. Then, the ASV tables were filtered to keep the ASVs only identified in the snow (541 unique ASVs). Finally, only ASVs identified in the snow and present in the soils after the start of the experiment were conserved to obtain ASV tables of invaders and potential colonists for each experiment, leaving only 16 ASVs as invaders. All other 525 ASVs were either identified only in the FT or snow samples. Invaders included all taxa which were identified in the snow and soil but not detected in the soil prior to the start of the experiment. Potential colonists were invaders that were identified across multiple days throughout the experiments, regardless of whether they were identified in multiple replicates or not. As new members of a community can naturally undergo abundance fluctuations (van Elsas et al., 2012; Acosta et al., 2015; Mallon et al., 2018), ASVs identified across multiple days and still identified on D29 were considered potentially successful colonists, regardless of whether they had increasing or decreasing relative abundances. Differences between the number of invaders, colonists and potentially successful colonists by experiment were tested using ANOVA and Tukey's HSD tests with Bonferroni correction.

RESULTS

The Influence of Melt Rate

Overall, the differences in gene copy numbers were significant between sample types (**Table 1**). Soil and flow-through (FT) samples had significantly more gene copies than snow samples (**Table 1**; **Supplementary Figure S3A**). On average, 4.02×10^1 16S rRNA gene copies were measured per ml of melted snow and significantly changed by sampling day (**Table 1**; **Supplementary Figure S3A**). In soils, the number of gene copies per g of soil decreased during storage but increased again with the start of the experiment and the addition of water. The number of gene copies was higher in the fast flow than in the average flow experiments and the number of gene copies was consistently higher in the controls than in the treated soils (**Table 1**; **Supplementary Figure S3A**). On day 29, all samples returned close to the starting number of gene copies (D0), potentially indicating the stabilisation of the soil community 15 days after the end of the melt. In the flow-through samples, higher gene copies were observed in the fast flow than in the average flow experiments (**Table 1**; **Supplementary Figure S3A**). In contrast with the soil, the number of gene copies in the flow-through was higher in the treated than in the control samples. Although results were not significant, this is consistent with the addition of microorganisms from the melted snow. In all FT samples, the number of gene copies was highest on day 5 and decreased sharply on all other days. This spike on day 5 was expected as it may reflect the removal of unattached microorganisms, dead cells and relic DNA from the soil, evacuated with the addition of water.

Differences in alpha diversity were significant by sample type (**Figures 2A,B**; **Table 1**). In the snow, differences from the day of field sampling until the end of the experiments were not significant (**Figures 2A,B**; **Table 1**). In the soils, as with gene copies, there was a decrease in alpha diversity measures with storage (pre-treated), which increased at the beginning of the experiment with the addition of water (**Figures 2A,B**; **Table 1**). Alpha diversity was significantly higher in the average flow rate. However, there was no significant difference between control and treated samples and the day of sampling had limited influence on soil alpha diversity measures (**Figures 2A,B**; **Table 1**). In the flow-through, the number of ASVs was significantly higher in the fast flow experiments but decreased with time, highlighting the spike

TABLE 1 | Results of the ANOVA tests on gene copy numbers (qPCR), richness and Shannon diversity and results of the adonis tests on Bray–Curtis community dissimilarity).

Experiment	Sample type	Formula	Gene copies (qPCR)			Richness			Shannon			Community composition		
			F	*p*	Sign	*F*	*p*	Sign	*F*	*p*	Sign	R^2	*p*	Sign
Melt rate	All	Type	$F_{2-302} = 734$	2×10^{-16}	***	$F_{2-105} = 72$	2×10^{-16}	***	$F_{2-105} = 198$	2×10^{-16}	***	0.36	0.001	***
Melt rate	Snow	Day	$F_{4-9} = 15.2$	0.0005	***	$F_{4-5} = 3.24$	0.11		$F_{4-5} = 0.96$	0.50		0.55	0.008	**
Melt rate	Soil	Rate	$F_{2-180} = 11.6$	1.9×10^{-5}	***	$F_{4-103} = 9.24$	2×10^{-6}	***	$F_{3-103} = 11.5$	9×10^{-8}	***	0.24	0.001	***
Melt rate	Soil	Treat	$F_{1-180} = 6$	0.003	**	$F_{4-103} = 4.89$	0.001	**	$F_{4-103} = 11.4$	1×10^{-7}	***	0.13	0.004	**
Melt rate	Soil	Rate* Treat	$F_{1-178} = 1.89$	0.17		$F_{1-101} = 0.015$	0.90		$F_{1-101} = 0.03$	0.87		0.02	0.072	
Melt rate	Soil	Rate*Day	$F_{4-170} = 2.82$	0.03	*	$F_{6-90} = 0.66$	0.69		$F_{6-90} = 0.89$	0.51		0.06	0.11	
Melt rate	Soil	Treat *Day	$F_{4-170} = 1$	0.40		$F_{6-90} = 0.39$	0.89		$F_{6-90} = 0.69$	0.66		0.05	0.97	
Melt rate	Soil	Rate* Treat*Day	$F_{4-160} = 0.63$	0.64		$F_{4-80} = 0.25$	0.91		$F_{4-80} = 0.016$	1		0.04	0.67	
Melt rate	FT	Rate	$F_{1-106} = 5.53$	0.02	*	$F_{4-103} = 9.24$	2×10^{-6}	***	$F_{4-103} = 11.5$	2×10^{-8}	***	0.05	0.005	**
Melt rate	FT	Treat	$F_{1-106} = 0.69$	0.41		$F_{4-103} = 4.88$	0.001	**	$F_{4-103} = 11.4$	2×10^{-7}	***	0.03	0.41	
Melt rate	FT	Rate* Treat	$F_{1-104} = 0.18$	0.67		$F_{1-101} = 0.015$	0.90		$F_{1-101} = 0.03$	0.87		0.03	0.25	
Melt rate	FT	Rate*Day	$F_{2-102} = 10.9$	4.9×10^{-5}	***	$F_{6-90} = 0.66$	0.69		$F_{6-90} = 0.89$	0.51		0.08	0.005	**
Melt rate	FT	Treat*Day	$F_{2-102} = 1.58$	0.21		$F_{6-90} = 0.39$	0.89		$F_{6-90} = 0.69$	0.66		0.05	0.66	
Melt rate	FT	Rate* Treat*Day	$F_{2-96} = 0.91$	0.41		$F_{4-80} = 0.25$	0.91		$F_{4-80} = 0.02$	1		0.05	0.40	
Soil pH	All	Type	$F_{2-472} = 161$	2×10^{-16}	***	$F_{2-165} = 14.3$	2×10^{-6}	***	$F_{2-165} = 125.6$	2×10^{-16}	***	0.36	0.001	***
Soil pH	Snow	Day	$F_{4-8} = 25.3$	0.0001	***	$F_{4-5} = 0.86$	0.55		$F_{4-5} = 0.55$	0.71		0.49	0.051	
Soil pH	Soil	pH	$F_{3-296} = 152$	2×10^{-16}	***	$F_{3-101} = 19.02$	7.5×10^{-10}	***	$F_{3-101} = 22$	5×10^{-11}	***	0.32	0.001	***
Soil pH	Soil	Treat	$F_{2-296} = 1.61$	0.20		$F_{2-102} = 2.13$	0.12		$F_{2-102} = 2.96$	0.06		0.03	0.17	
Soil pH	Soil	pH* Treat	$F_{2-293} = 0.27$	0.77		$F_{2-98} = 5.15$	0.0075	**	$F_{2-98} = 6.6$	0.002	**	0.04	0.005	**
Soil pH	Soil	pH*Day	$F_{10-280} = 3.19$	0.0007	***	$F_{9-85} = 4.84$	3.2×10^{-5}	***	$F_{9-85} = 3.4$	0.001	**	0.13	0.001	***
Soil pH	Soil	Treat*Day	$F_{5-286} = 0.44$	0.082		$F_{5-90} = 0.80$	0.55		$F_{5-90} = 0.39$	0.86		0.53	1	
Soil pH	Soil	pH* Treat*Day	$F_{10-262} = 2.27$	0.015	*	$F_{9-68} = 0.99$	0.46		$F_{9-68} = 0.68$	0.73		0.05	0.53	
Soil pH	FT	pH	$F_{2-159} = 6.21$	0.003	**	$F_{2-50} = 1.56$	0.22		$F_{2-50} = 0.88$	0.42		0.07	0.001	***
Soil pH	FT	Treat	$F_{1-160} = 0.32$	0.57		$F_{1-51} = 0.11$	0.74		$F_{1-51} = 1.67$	0.20		0.02	0.70	
Soil pH	FT	pH*Treat	$F_{4-156} = 0.30$	0.75		$F_{2-47} = 0.5$	0.61		$F_{2-47} = 0.11$	0.90		0.03	0.79	
Soil pH	FT	pH*Day	$F_{4-153} = 7.91$	7×10^{-6}	***	$F_{4-44} = 1.23$	0.31		$F_{4-44} = 1.23$	0.30		0.09	0.003	**
Soil pH	FT	Treat*Day	$F_{4-144} = 0.36$	0.70		$F_{2-47} = 0.96$	0.39		$F_{2-47} = 2.98$	0.06		0.04	0.55	
Soil pH	FT	pH* Treat*Day	$F_{4-144} = 0.35$	0.84		$F_{4-35} = 2.06$	0.11		$F_{4-35} = 5.12$	0.002	**	0.08	0.13	

The experiment column refers to either the different melt rate or soil pH microcosms. The sample type column differentiates between snow, soil and flow-through (FT). The formula column indicates the parameters used in the models. Type = sample type/rate = fast or slow/treatment (treat) = snow or sterile water input/day = day of sampling. Value of p, asterisk indicates significant differences (Sign).
**p < 0.05; **p < 0.01; ***p < 0.001.*

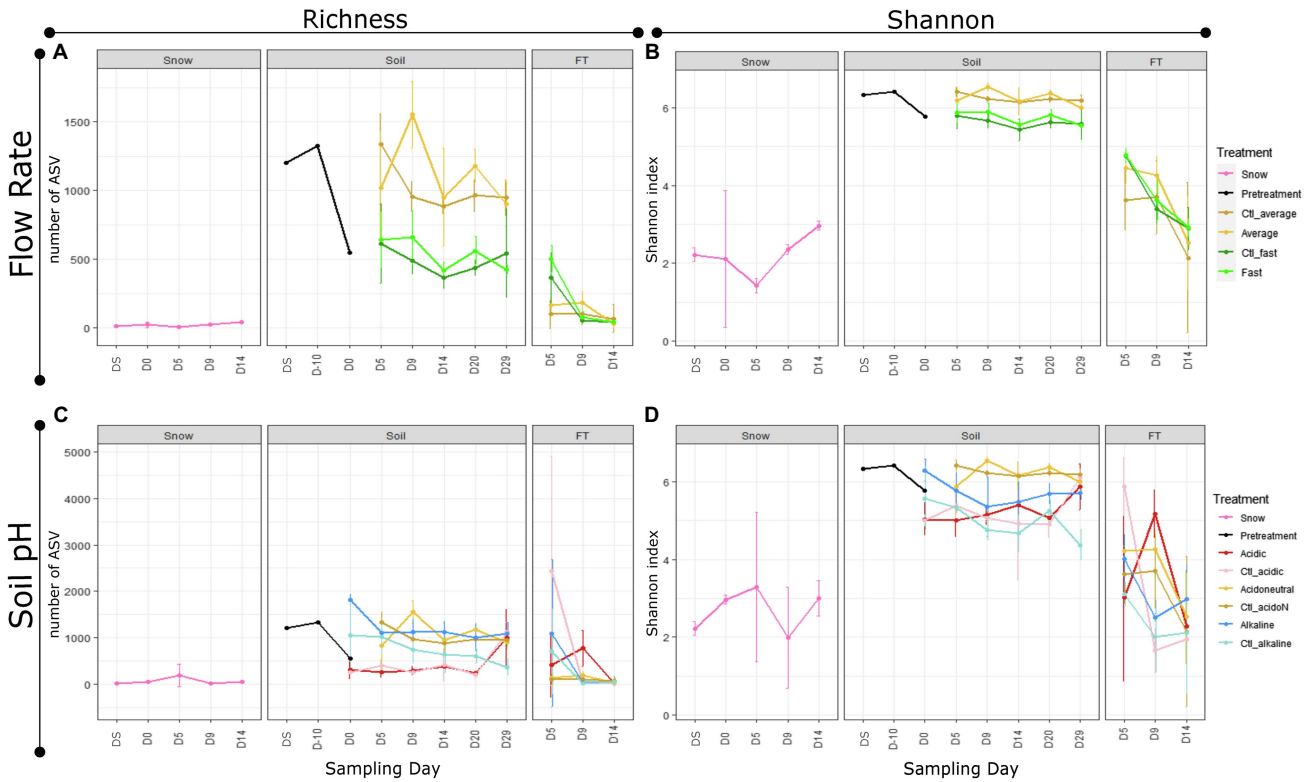

FIGURE 2 | Alpha diversity by experiment, treatment, sample type and day of sampling. **(A)** Observed number of amplicon sequence variants (ASVs) (Richness) and **(B)** diversity with the Shannon index in different flow rate experiments. **(C)** Observed number of ASVs (Richness) and **(D)** diversity with the Shannon index in different soil pH experiments. The shades of yellow correspond to the average and acidoneutral controls (Ctl) and treatments, resulting from the same microcosms.

at D5, also observed with gene copy numbers (**Figures 2A,B**; **Table 1**).

Overall, differences in bacterial community composition by sample type were observed using PCoA based on the Bray–Curtis dissimilarity (**Figure 3A**; **Table 1**), although these differences were likely also influenced by within-sample dispersion (betadisper $F=151$, $p<0.001$). Snow communities changed significantly between the day of field sampling and the end of the experiment (**Table 1**) and were primarily composed of Proteobacteria and Bacteroidetes (**Figure 3B**; **Supplementary Figure S4A**). Large variations in the composition of FT samples masked the potential variation in soil samples (**Figure 3A**) and therefore, the PCoA of soil communities was assessed separately (**Supplementary Figure S5**). The soil community on the day of field sampling changed with storage (**Figure 3A**; **Supplementary Figure S5**). Soil communities clustered separately between flow rates, but as for alpha diversity, controls and treated samples were similar (**Table 1**; **Supplementary Figure S5**). We should note that group dispersion was heterogeneous (betadisper $F=7.4$, $p<0.001$). Soil communities were primarily composed of Acidobacteria, Actinobacteria, Bacteroidetes, Chloroflexi, Gemmatimonadetes, Planctomycetes, Proteobacteria and Verrucomicrobia (**Figure 3B**; **Supplementary Figure S4A**). The communities identified in the flow-through presented clear variations, primarily separated by flow rate and sampling day (**Figure 3**). All FT samples were dominated by Proteobacteria but differences in communities with flow rate were observed. For instance, all fast FT samples had large proportions of Acidobacteria, Bacteroidetes and Proteobacteria and were relatively stable while in the average rate, communities significantly changed on D14 with the depletion of Acidobacteria (**Figure 3B**; **Supplementary Figure S4A**).

The Influence of Soil pH

Here, we assessed the influence of soil pH on the colonisation potential of snow microorganisms. The results from the average flow rate were used as the acidoneutral samples for comparison. As in the melt rate experiments, the differences in gene copy numbers were significant between sample types, with soil and FT samples harbouring significantly more gene copies than snow samples, which changed significantly by sampling day (**Table 1**; **Supplementary Figure S3B**). In soils, the number of gene copies decreased during storage between the day of field sampling (DS) and D-10. After pH manipulation (D-10) and until D0, the number of gene copies increased in alkaline soils but decreased in acidic soils. During the experiment, the gene copy number remained lowest in acidic and highest in alkaline soils with no significant differences between controls

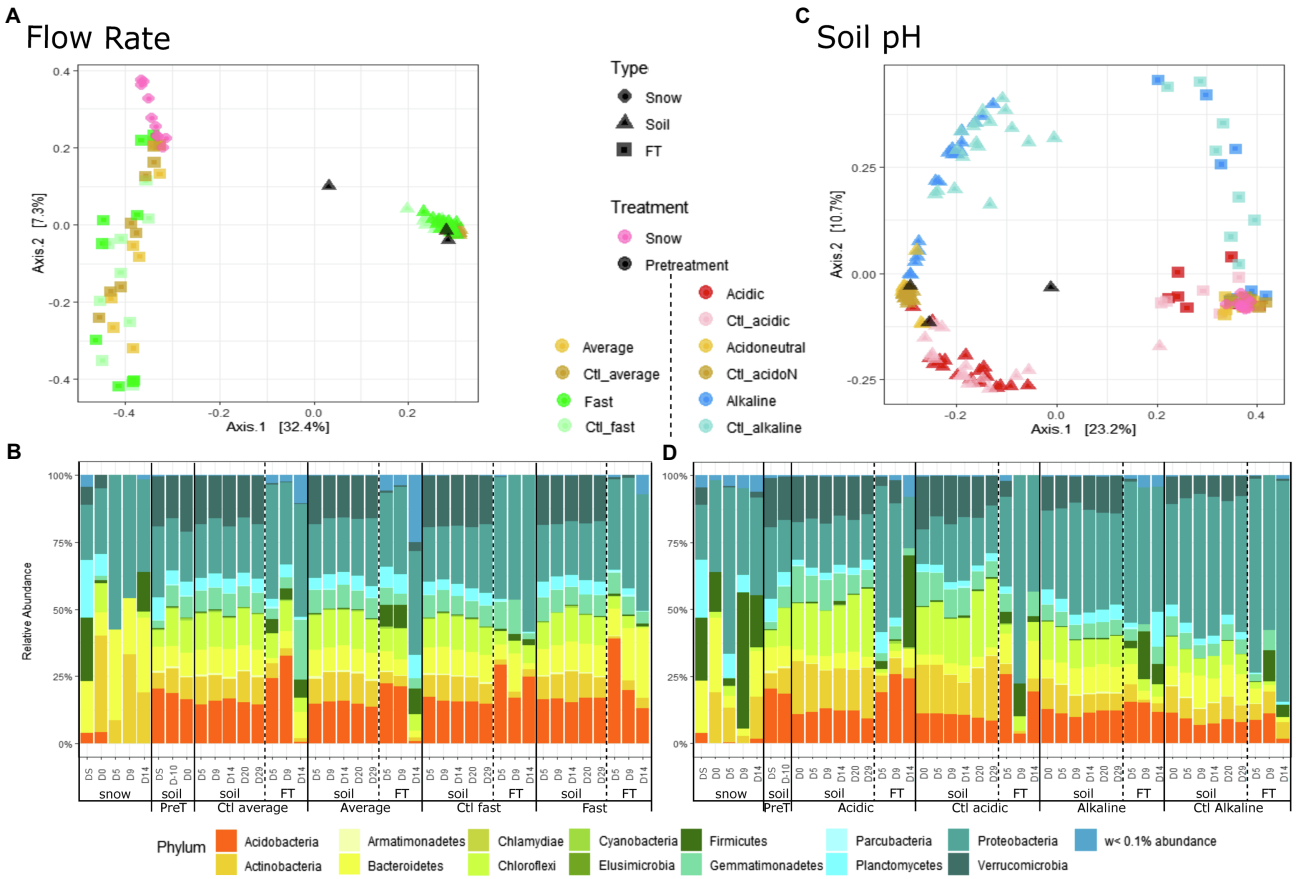

FIGURE 3 | Principal coordinate analysis (PCoA) and community composition at the phylum level, by experiment, treatment, sample type and day of sampling. **(A)** PCoA and **(B)** community composition of flow rate microcosms. **(C)** PCoA and **(D)** community composition of soil pH microcosms. Pretreatment (PreT) includes the day of field sampling (DS), the day of microcosm set-up (D-10) and D0 for the flow rate experiment as the soil was not yet treated in any way. 'w<0.1% abundance' regroups all phyla with less than 0.1% relative abundance.

and treated soils (**Table 1**; **Supplementary Figure S3B**). In the flow-through, higher gene copies were quantified in the alkaline soils, which also presented the highest variability (**Supplementary Figure S3B**; **Table 1**). As observed in the melt rate experiments, the number of gene copies peaked on D5 and decreased with time.

Differences in alpha diversity were significant by sample type (**Figures 2C,D**; **Table 1**). In the snow, alpha diversity did not change significantly from the day of sampling until the end of the experiments (**Figures 2C,D**; **Table 1**). In the soils, as with gene copies, there was a decrease in alpha diversity measures with storage (pre-treated). However, alpha diversity changed following pH manipulation showing an increase in alkaline soils and a decrease in acidic soils (**Figures 2A,B**). During the experiment, both soil pH and snow/water treatment were identified as significant variables influenced by the day of sampling (**Table 1**). Interestingly, on day 29, all microcosms harboured similar alpha diversity levels, except the alkaline controls (**Figures 2C,D**), suggesting stabilisation of the communities. In the flow-through, alpha diversity decreased with time and the interactions between soil pH, treatment and sampling day were significant (**Figures 2C,D**; **Table 1**).

Overall, differences in bacterial community composition by sample type were observed [groups were homogeneously dispersed (betadisper $F = 0.93$, $p = 0.40$)] (**Figure 3C**; **Table 1**). In the PCoA, the snow samples clustered closely together, indicating similar communities over time (**Figure 3C**; **Table 1**) dominated by Proteobacteria and Firmicutes (**Supplementary Figure S4B**). Acidic, acidoneutral and alkaline soil samples all formed different clusters indicating important differences in community composition (**Figure 3C**). The horseshoe effect observed in the PCoA likely reflects the pH gradient and highlights the change in community composition across this gradient (Morton et al., 2017), although these differences were likely also influenced by within-sample dispersion (betadisper $F = 8.40$, $p < 0.001$). This difference in soil community composition was further observed in **Figure 3D**, illustrating the community composition at the phylum level. We observed a shift in community composition during storage but also after pH manipulation, between D-10 and D0. Acidic communities presented an increase in Actinobacteria and

Chloroflexi while an increase in Proteobacteria was observed in alkaline communities (**Supplementary Figure S4B**). The addition of water after the start of the experiments further enhanced this increase in Proteobacteria (**Figure 3D**; **Supplementary Figure S4B**). Acidic soil communities were dominated by Actinobacteria, Chloroflexi, Proteobacteria and Actinobacteria while Alkaline soil communities were dominated by Proteobacteria (**Figure 3D**; **Supplementary Figure S4B**). Flow-through samples we dominated by Proteobacteria **Supplementary Figure S4B**) and displayed high variability in community composition (**Figures 3C,D**; **Table 1**).

Colonisation Potential

To assess whether colonisation of the soil by microorganisms in the snow occurred, only the ASVs identified in the snow, absent from the controls and pre-treated soils (DS, D-10 and D0) but identified in the soils from D5 were selected for further analysis. Only 16 ASVs fulfilled these conditions and were considered potential invaders. All were classified as Proteobacteria or Bacteroidetes, primarily Alphaproteobacteria, Betaproteobacteria, Flavobacteriia and Sphingobacteriia at the class level (**Figure 4**; **Supplementary Table S2**). In some cases, we identified taxa on D20 or D29 that were not previously identified and appeared new. It is likely that these taxa were previously present but had abundances below the detection threshold.

In the fast flow experiments, a total of 13 ASVs were identified as invaders depositing in the soils (**Figure 4A**; **Supplementary Table S2**). Of these 13 invaders, seven were identified across multiple days and were considered potential colonists. However, only two persisted until D29 suggesting potentially successful colonists. In the average flow experiments (acidoneutral experiments), nine ASVs were identified as invaders (**Figure 4B**; **Supplementary Table S2**). Of these nine ASVs, eight were identified across multiple days and considered potential colonists. Five ASVs persisted until D29 and were considered potentially successful colonists. Differences in invaders, colonists and potentially successful colonists were not statistically significant between flow rates. In acidic soils, only five ASVs were identified as invaders (**Figure 4C**; **Supplementary Table S2**). Of these five invaders, four were identified across multiple days and considered potential colonists; however, only two persisted until D29. In alkaline soils, nine ASVs were identified as invaders (**Figure 4D**). Of these nine invaders, seven were identified across multiple days and considered potential colonists but

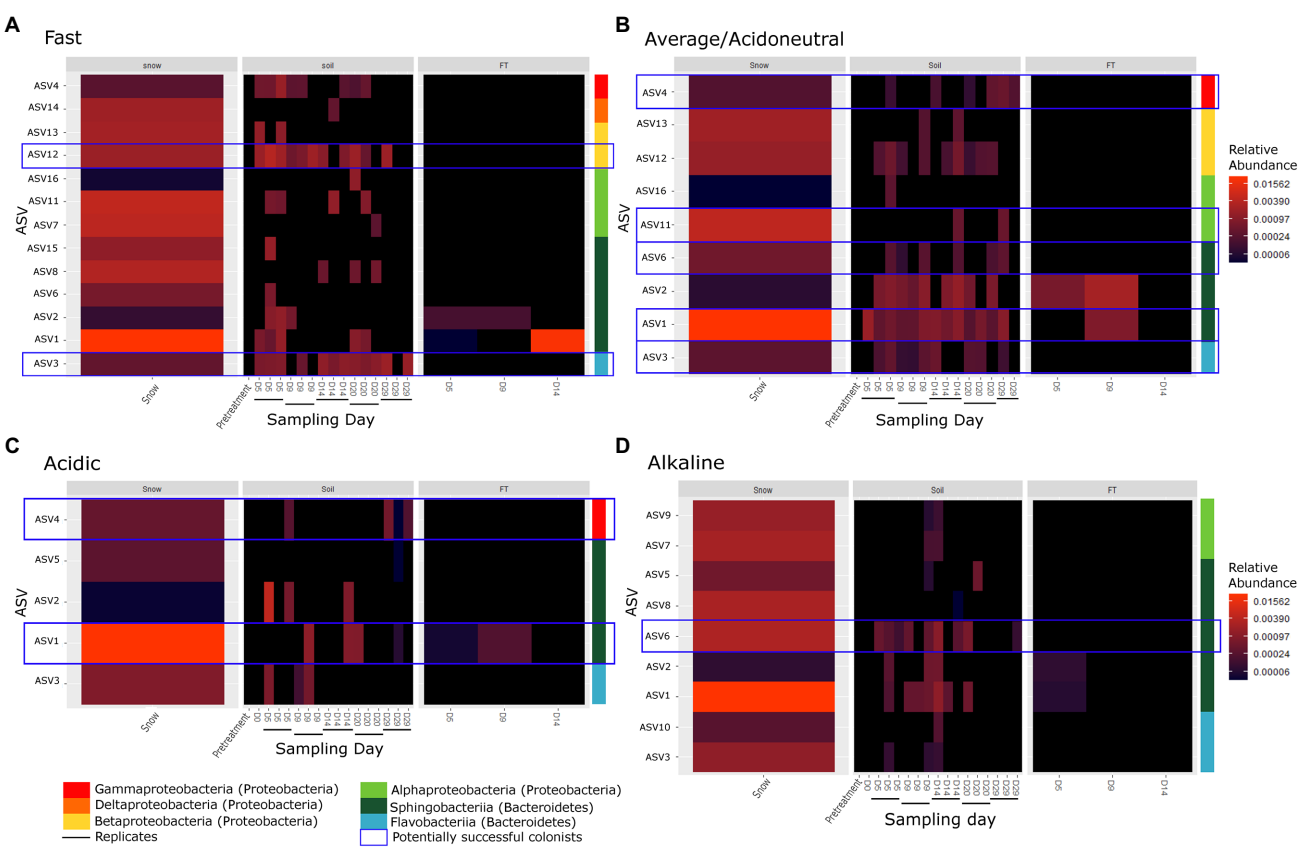

FIGURE 4 | Heatmap of invaders of **(A)** the fast flow, **(B)** the average flow/acidoneutral, **(C)** the acidic, and **(D)** the alkaline experiments where each row corresponds to an ASV. The coloured bar indicates the taxonomy associated to each ASV at the class level. Blue boxes highlight potentially successful colonists identified across multiple days and still present on D29. Black horizontal bars indicate replicate microcosms.

only one persisted until D29. Overall, the persistence of taxa until the end of the experiment was highest at the average flow in acidoneutral soils, with five potentially successful colonisation events. The higher flow rate increased the number of invaders but the persistence across time decreased. In acidic soils at the average flow rate, the number of invaders and persistence with time was much lower. In alkaline soils, the number of invaders was equal to that observed in acidoneutral soils but persistence with time was strongly reduced. Statistically, the number of invaders and colonists was significantly higher in acidoneutral soils (ANOVA, $p = 0.02$ and $p = 0.009$ respectively), and potentially successful colonists were only marginally more successful (ANOVA, $p = 0.056$).

DISCUSSION

In this study, we assessed the influence of increased precipitation (*via* the melt rate) and soil pH on the colonisation potential of snow microorganisms. Overall, we rejected the first hypothesis that increased precipitation (higher flow rate) would promote successful colonisations due to the higher inoculum density as well as increased ecosystem disturbance. However, we did support the second hypothesis by observing increased colonisation in acidoneutral soils compared to acidic and alkaline soils.

In both sets of experiments, the number of gene copies was consistently higher in the control than in the treated soils. This may be a signal of invasion, that is, species in the controls were not experiencing interactions with outside invaders, did not have to partition resources and could therefore grow to higher abundances. This signal of invasion was further observed for the richness and diversity, consistently higher in the treated soils, in line with the addition of snow microorganisms.

Of the two flow rate scenarios tested, the fast flow rate increased the input of microorganisms to the system, as would be expected from increased precipitation, and was expected to cause greater disturbance than the average melt rate. Ecosystem disturbance is considered a key factor for successful colonisation by newly deposited microbial invaders, with increasing disturbance enhancing the chances of successful colonisation (Liu et al., 2012; De Roy et al., 2013). We expected the addition of water at the higher flow rate to disturb soil communities (observed *via* changes in alpha and beta diversity) and increase the colonisation potential. Instead, the strongest impact on communities was observed at the average rate. Richness and diversity increased, and the soil community composition shifted away from the composition on the first day of the experiment (**Supplementary Figure S5**). Furthermore, melted snow has been shown to add nutrients to the ecosystem, resulting in a nutrient pulse of carbon, nitrogen and phosphorus (Schmidt and Lipson, 2004; Edwards et al., 2006; Buckeridge and Grogan, 2010; Larose et al., 2013). As nutrient pulses increase the chances of successful colonisation (Li and Stevens, 2012; Van Nevel et al., 2013; Mallon et al., 2015b), we expected to observe differences between the controls and treated samples. Instead, controls and treated soils within each experiment remained similar and only differences between experiments (average and fast rates) were observed. The shift in communities between average and fast flow rate suggested physical selection of microorganisms. For instance, fast growing r-strategists or EPS (extracellular polymeric substance) producing bacteria might be selected in the fast flow microcosms to ensure attachment, aggregation and growth in the soil system (Ho et al., 2017; Costa et al., 2018). Overall, following the increased input of microorganisms, we observed more invaders but low persistence with time and only two potentially successful colonists. At the average rate, the number of invaders was lower but persistence with time increased and five potentially successful colonists were identified, suggesting that a slower rate may give more time for invading microorganisms to compete with indigenous microorganisms and successfully colonise the ecosystem instead of being pushed out of the system by a higher flow rate.

As pH has previously been identified as a key driver of bacterial community structure in global (Fierer and Jackson, 2006; Delgado-Baquerizo et al., 2018) and in Arctic soils (Malard et al., 2019a), and dispersal is an important process structuring bacterial communities, the influence of soil pH on the colonisation success was investigated. The soil pH appeared to have a stronger influence on soil bacterial communities than flow rate. Not only were richness and diversity lower than in acidoneutral soils, but community composition was clearly different, forming three distinct clusters in the PCoA. Contrary to the flow rate experiments, communities between control and treated soils were also different, showing that the input of snow or sterile water had different consequences, potentially due to the input of nutrients from the snow and the different bacterial communities in acidic and alkaline soils. Acidic soils contained the lowest number of invaders deposited, as only two were considered potentially successful colonists. In alkaline soils, as in acidoneutral soils, an equal number of invaders were identified. However, persistence with time was much lower and only one potentially successful colonist was identified. The low colonisation observed in acidic and alkaline samples and successful colonisation in acidoneutral soils supports the second hypothesis that acidoneutral soils would promote successful colonisations. Furthermore, it suggests that while the flow rate is an important parameter in promoting successful colonisation, soil pH may be an even more important factor limiting colonisation.

Overall, all potentially successful colonists were identified in multiple replicates across different time points within each experiment, as well as in different experiments (**Figure 4**; **Supplementary Table S2**), suggesting active and selective colonisation of the soil by these taxa. For instance, ASV3 is a colonist at both flow rates, ASV1 and 4 are colonists of both acidoneutral and acidic soils while ASV6 colonised acidoneutral and alkaline soils. The potentially successful colonists

(Proteobacteria and Bacteroidetes) were primarily classified as r-strategists (as in Fierer et al., 2007 and Ho et al., 2017), with high growth rates and generally considered more likely to successfully colonise (Andrews and Harris, 1986; Litchman, 2010).

The colonisation and persistence with time observed in acidoneutral soils was in agreement with the consensus that acidoneutral soils may decrease the adaptative pressure on microbial communities (Voroney and Heck, 2014; Morris and Blackwood, 2015). Arctic acidoneutral soils harbour more generalist taxa (Malard et al., 2019a), generally considered more prone to dispersal (Pandit et al., 2009; Graham and Stegen, 2017; Sriswasdi et al., 2017). Therefore, the higher successful colonisation of acidoneutral soils further supported the role of dispersal in shaping bacterial communities in these ecosystems.

On the other hand, the lower number of invading taxa in acidic soils was not surprising considering that most bacterial taxa decrease in abundance with decreasing pH (Chu et al., 2010; Rousk et al., 2010), including the invading taxa Bacteroidetes. Potentially successful colonists of acidic and alkaline soils colonised acidoneutral soils, however, they did not colonise each other. The low rate of colonisation of acidic and alkaline soils, both generally characterised as harsh (Fierer and Jackson, 2006; Rousk et al., 2010), may highlight the need for effective adaptations (Beales, 2004; Ward et al., 2009) lacking in the deposited microorganisms. They may also be worse competitors than the indigenous soil communities, generally dominated by specialist taxa (Malard et al., 2019a) and acclimated to the soil pH. However, in this study, we opted to use the same soil bacterial community as a baseline to evaluate the colonisation potential. Therefore, we manipulated the soil pH by the addition of aluminium sulphate and calcium carbonate (Nicol et al., 2008). While we let the communities acclimatise to the new soil pH, they may be less adapted than indigenous communities in naturally acidic and alkaline soils. The addition of sulphate coupled with the addition of water could have formed sulphuric acid, especially at lower flow rates where the soil was exposed to air. Toxic metals, such as manganese, aluminium and arsenic, may have been released (Robson and Abbott, 1989; Kunhikrishnan et al., 2016). On the other hand, calcium carbonate is used to increase soil pH and reduce the possible toxicity of manganese and aluminium ions (Muhlbachova and Tlustos, 2006; Kunhikrishnan et al., 2016; Guo et al., 2019). In all manipulated pH microcosms, we observed a shift in bacterial communities following these adjustments (between D-10 and D0). Future studies could repeat these experiments using soils within the targeted pH range and the associated indigenous communities to establish if the colonisation potential is also lower in these natural systems. It is also interesting to note that the snow pH was acidoneutral (6.33 ± 0.15) and therefore, snow microorganisms may have already been well adapted for acidoneutral soil conditions. As snow pH is generally acidoneutral (Kol and Taylor, 1942; Drake and Moore, 1980; Newton, 1982; Li et al., 2007; Ali et al., 2010; Zhang et al., 2013), this scenario is rather realistic. However, we cannot exclude the possibility that acidic or alkaline precipitation may promote colonisation in soils within the same pH range.

Overall, local environmental conditions may be more important determinant factors influencing the outcome of colonisation than increased precipitation or faster melt events. Furthermore, we did not observe the inverse relationship between microbial diversity and invasion outcome previously identified or theorised (van Elsas et al., 2012; Mallon et al., 2015a; Kinnunen et al., 2016; Vila et al., 2019). Here, more potential colonisation events were observed in the […]. Instead, only invaders presenting the right adaptations to colonise a free niche (Malard et al., 2021) or outcompete indigenous microorganisms may be able to successfully colonise in the medium to long term. Following the deposition, the invaders have to adapt, compete for resources and then grow and spread in the ecosystem to colonise (Mallon et al., 2015a; Kinnunen et al., 2016). Here, we demonstrated that microorganisms were successfully deposited but only few taxa had the potential to successfully colonise the soils. After 15 days, the majority had disappeared leading to likely failed colonisations; although, even failed colonisations may influence indigenous microbial communities (Mallon et al., 2018).

CONCLUSION

This study used the Arctic snowpack as a model system to investigate microbial colonisation of snow bacteria deposited into Arctic soils. First, we tested the impact of increased precipitation (inoculum density) and the subsequent faster snowmelt rate (ecosystem disturbance) on the colonisation outcome. We identified more invaders but decreased persistence with time and a lower number of successful colonisation events. We also evaluated the influence on soil pH on this colonisation potential. Here, persistence with time decreased and a lower number of successful colonisation events were recorded in acidic and alkaline soils compared with acidoneutral soils. Overall, we demonstrated that soil bacterial communities could change significantly with snowmelt and that microorganisms were successfully deposited in soils following snowmelt events. However, we showed that only few taxa successfully colonised and established in these soil communities. Results suggest that local soil properties might have a greater influence on the colonisation outcome than increased precipitation or ecosystem disturbance.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: https://www.ncbi.nlm.nih.gov/, PRJNA564428.

AUTHOR CONTRIBUTIONS

DP secured the funding. LM and DP conceived and designed the study. LM carried the experimental work, laboratory work, bioinformatics processing, statistical analysis, and drafted the manuscript. DP revised and approved the final version. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by a grant from the European Commission's Marie Sklowdowska Curie Actions program under project number 675546.

ACKNOWLEDGMENTS

The authors thank Lewis Cuthbertson and Khadija Jabeen for the occasional care of the experiment. MiSeq sequencing of the 16S rRNA gene was performed by the NU-OMICS sequencing service (Northumbria University). Samples were collected under the Research in Svalbard (RIS) project number 10700.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2022.782789/full#supplementary-material

REFERENCES

Acosta, F., Zamor, R. M., Najar, F. Z., Roe, B. A., and Hambright, K. D. (2015). Dynamics of an experimental microbial invasion. *Proc. Natl. Acad. Sci.* 112, 11594–11599. doi: 10.1073/pnas.1505204112

Ali, K., Sonbawane, S., Chate, D., Siingh, D., Rao, P., Safai, P., et al. (2010). Chemistry of snow and lake water in Antarctic region. *J. Earth Syst. Sci.* 119, 753–762. doi: 10.1007/s12040-010-0063-0

Amato, P., Hennebelle, R., Magand, O., Sancelme, M., Delort, A.-M., Barbante, C., et al. (2007). Bacterial characterization of the snow cover at Spitzberg, Svalbard. *FEMS Microbiol. Ecol.* 59, 255–264. doi: 10.1111/j.1574-6941.2006.00198.x

Andrews, J. H., and Harris, R. F. (1986). "R-and K-selection and microbial ecology," in *Advances in Microbial Ecology*, ed. K. E. Nelson (Boston, MA: Springer), 99–147.

Archer, S. D., Lee, K. C., Caruso, T., Maki, T., Lee, C. K., Cary, S. C., et al. (2019). Airborne microbial transport limitation to isolated Antarctic soil habitats. *Nat. Microbiol.* 4, 925–932. doi: 10.1038/s41564-019-0370-4

Bahram, M., Hildebrand, F., Forslund, S. K., Anderson, J. L., Soudzilovskaia, N. A., Bodegom, P. M., et al. (2018). Structure and function of the global topsoil microbiome. *Nature* 560, 233–237. doi: 10.1038/s41586-018-0386-6

Barberán, A., Henley, J., Fierer, N., and Casamayor, E. O. (2014). Structure, inter-annual recurrence, and global-scale connectivity of airborne microbial communities. *Sci. Total Environ.* 487, 187–195. doi: 10.1016/j.scitotenv.2014.04.030

Barberán, A., Ladau, J., Leff, J. W., Pollard, K. S., Menninger, H. L., Dunn, R. R., et al. (2015). Continental-scale distributions of dust-associated bacteria and fungi. *Proc. Natl. Acad. Sci.* 112, 5756–5761. doi: 10.1073/pnas.1420815112

Beales, N. (2004). Adaptation of microorganisms to cold temperatures, weak acid preservatives, low pH, and osmotic stress: a review. *Compr. Rev. Food Sci. Food Saf.* 3, 1–20. doi: 10.1111/j.1541-4337.2004.tb00057.x

Bintanja, R., and Andry, O. (2017). Towards a rain-dominated Arctic. *Nat. Clim. Chang.* 7, 263–267. doi: 10.1038/nclimate3240

Bintanja, R., and Selten, F. (2014). Future increases in Arctic precipitation linked to local evaporation and sea-ice retreat. *Nature* 509, 479–482. doi: 10.1038/nature13259

Bottos, E. M., Woo, A. C., Zawar-Reza, P., Pointing, S. B., and Cary, S. C. (2014). Airborne bacterial populations above desert soils of the McMurdo dry valleys, Antarctica. *Microb. Ecol.* 67, 120–128. doi: 10.1007/s00248-013-0296-y

Buckeridge, K. M., and Grogan, P. (2010). Deepened snow increases late thaw biogeochemical pulses in Mesic low arctic tundra. *Biogeochemistry* 101, 105–121. doi: 10.1007/s10533-010-9426-5

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

Cameron, K. A., Hagedorn, B., Dieser, M., Christner, B. C., Choquette, K., Sletten, R., et al. (2015). Diversity and potential sources of microbiota associated with snow on western portions of the Greenland ice sheet. *Environ. Microbiol.* 17, 594–609. doi: 10.1111/1462-2920.12446

Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Lozupone, C. A., Turnbaugh, P. J., et al. (2010). Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc. Natl. Acad. Sci.* 108(Supplement 1), 4516–4522. doi: 10.1073/pnas.1000080107

Chu, H., Fierer, N., Lauber, C. L., Caporaso, J. G., Knight, R., and Grogan, P. (2010). Soil bacterial diversity in the Arctic is not fundamentally different from that found in other biomes. *Environ. Microbiol.* 12, 2998–3006. doi: 10.1111/j.1462-2920.2010.02277.x

Costa, O. Y., Raaijmakers, J. M., and Kuramae, E. E. (2018). Microbial extracellular polymeric substances: ecological function and impact on soil aggregation. *Front. Microbiol.* 9:1636. doi: 10.3389/fmicb.2018.01636

Cuthbertson, L., Amores-Arrocha, H., Malard, L. A., Els, N., Sattler, B., and Pearce, D. A. (2017). Characterisation of Arctic bacterial communities in the air above Svalbard. *Biology* 6:29. doi: 10.3390/biology6020029

Davis, N. M., Proctor, D. M., Holmes, S. P., Relman, D. A., and Callahan, B. J. (2018). Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome* 6:226. doi: 10.1186/s40168-018-0605-2

De Roy, K., Marzorati, M., Negroni, A., Thas, O., Balloi, A., Fava, F., et al. (2013). Environmental conditions and community evenness determine the outcome of biological invasion. *Nat. Commun.* 4, 1–5. doi: 10.1038/ncomms2392

De Wit, R., and Bouvier, T. (2006). 'Everything is everywhere, but, the environment selects'; what did baas Becking and Beijerinck really say? *Environ. Microbiol.* 8, 755–758. doi: 10.1111/j.1462-2920.2006.01017.x

Delgado-Baquerizo, M., Oliverio, A. M., Brewer, T. E., Benavent-González, A., Eldridge, D. J., Bardgett, R. D., et al. (2018). A global atlas of the dominant bacteria found in soil. *Science* 359, 320–325. doi: 10.1126/science.aap9516

Dixon, P. (2003). VEGAN, a package of R functions for community ecology. *J. Veg. Sci.* 14, 927–930. doi: 10.1111/j.1654-1103.2003.tb02228.x

Drake, J. J., and Moore, T. (1980). Snow pH and dust loading at Schefferville, Quebec. *Can. Geogr.* 24, 286–291. doi: 10.1111/j.1541-0064.1980.tb00343.x

Edwards, K. A., McCulloch, J., Kershaw, G. P., and Jefferies, R. L. (2006). Soil microbial and nutrient dynamics in a wet Arctic sedge meadow in late winter and early spring. *Soil Biol. Biochem.* 38, 2843–2851. doi: 10.1016/j.soilbio.2006.04.042

Els, N., Larose, C., Baumann-Stanzer, K., Tignat-Perrier, R., Keuschnig, C., Vogel, T. M., et al. (2019). Microbial composition in seasonal time series of free tropospheric air and precipitation reveals community separation. *Aerobiologia* 35, 671–701. doi: 10.1007/s10453-019-09606-x

Evans, S., Bell-Dereske, L., Dougherty, K., and Kittredge, H. (2019). Dispersal alters soil microbial community response to drought. *Environ. Microbiol.* 22, 905–916. doi: 10.1111/1462-2920.14707

Fierer, N., Bradford, M. A., and Jackson, R. B. (2007). Toward an ecological classification of soil bacteria. *Ecology* 88, 1354–1364. doi: 10.1890/05-1839

Fierer, N., and Jackson, R. B. (2006). The diversity and biogeography of soil bacterial communities. *Proc. Natl. Acad. Sci. U. S. A.* 103, 626–631. doi: 10.1073/pnas.0507535103

Fierer, N., Schimel, J., and Holden, P. (2003). Influence of drying–rewetting frequency on soil bacterial community structure. *Microb. Ecol.* 45, 63–71. doi: 10.1007/s00248-002-1007-2

Fløistrup, K. M., Olsen, M. N., Rasmussen, T. G., Ekelund, F., and Altenburger, A. (2018). Recruitment of airborne microorganisms on sterilized soil at different heights above ground. *Appl. Soil Ecol.* 126, 85–87. doi: 10.1016/j.apsoil.2018.02.011

Foster, J. (1989). The significance of the date of snow disappearance on the Arctic tundra as a possible indicator of climate change. *Arct. Alp. Res.* 21, 60–70. doi: 10.2307/1551517

Glenn, A. R., and Dilworth, M. J. (1991). "Soil acidity and the microbial population: survival and growth of bacteria in low pH," in *Plant-Soil Interactions at Low pH*, eds. R. J. Wright, V. C. Baligar, and P. Murrmann (Dordrecht: Springer), 567–579.

Graham, E. B., and Stegen, J. C. (2017). Dispersal-based microbial community assembly decreases biogeochemical function. *PRO* 5:65. doi: 10.3390/pr5040065

Gray, D., Toth, B., Zhao, L., Pomeroy, J., and Granger, R. (2001). Estimating areal snowmelt infiltration into frozen soils. *Hydrol. Process.* 15, 3095–3111. doi: 10.1002/hyp.320

Guo, A., Ding, L., Tang, Z., Zhao, Z., and Duan, G. (2019). Microbial response to $CaCO_3$ application in an acid soil in southern China. *J. Environ. Sci.* 79, 321–329. doi: 10.1016/j.jes.2018.12.007

Harding, T., Jungblut, A. D., Lovejoy, C., and Vincent, W. F. (2011). Microbes in high Arctic snow and implications for the cold biosphere. *Appl. Environ. Microbiol.* 77, 3234–3243. doi: 10.1128/AEM.02611-10

Hauptmann, A. L., Stibal, M., Bælum, J., Sicheritz-Pontén, T., Brunak, S., Bowman, J. S., et al. (2014). Bacterial diversity in snow on north pole ice floes. *Extremophiles* 18, 945–951. doi: 10.1007/s00792-014-0660-y

Hengl, T., Mendes de Jesus, J., Heuvelink, G. B., Ruiperez Gonzalez, M., Kilibarda, M., Blagotić, A., et al. (2017). SoilGrids250m: global gridded soil information based on machine learning. *PLoS One* 12:e0169748. doi: 10.1371/journal.pone.0169748

Ho, A., Di Lonardo, D. P., and Bodelier, P. L. (2017). Revisiting life strategy concepts in environmental microbial ecology. *FEMS Microbiol. Ecol.* 93:fix006. doi: 10.1093/femsec/fix006

Horikoshi, K. (1999). Alkaliphiles: some applications of their products for biotechnology. *Microbiol. Mol. Biol. Rev.* 63, 735–750. doi: 10.1128/MMBR.63.4.735-750.1999

Hossain, M., Chen, W., and Zhang, Y. (2015). Bulk density of mineral and organic soils in the Canada's arctic and sub-arctic. *Inf. Process. Agric.* 2, 183–190. doi: 10.1016/j.inpa.2015.09.001

Iwata, Y., Hayashi, M., and Hirota, T. (2008). Comparison of snowmelt infiltration under different soil-freezing conditions influenced by snow cover. *Vadose Zone J.* 7, 79–86. doi: 10.2136/vzj2007.0089

Jousset, A., Schulz, W., Scheu, S., and Eisenhauer, N. (2011). Intraspecific genotypic richness and relatedness predict the invasibility of microbial communities. *ISME J.* 5, 1108–1114. doi: 10.1038/ismej.2011.9

Kim, B.-R., Shin, J., Guevarra, R. B., Lee, J. H., Kim, D. W., Seol, K.-H., et al. (2017). Deciphering diversity indices for a better understanding of microbial communities. *J. Microbiol. Biotechnol.* 27, 2089–2093. doi: 10.4014/jmb.1709.09027

Kinnunen, M., Dechesne, A., Proctor, C., Hammes, F., Johnson, D., Quintela-Baluja, M., et al. (2016). A conceptual framework for invasion in microbial communities. *ISME J.* 10, 2773–2779. doi: 10.1038/ismej.2016.75

Kol, E., and Taylor, W. R. (1942). "The snow and ice algae of alaska," in *Smithsonian Miscellaneous Collections*. Washington DC: Smithsonian Institution.

Kunhikrishnan, A., Thangarajan, R., Bolan, N., Xu, Y., Mandal, S., Gleeson, D., et al. (2016). "Functional relationships of soil acidification, liming, and greenhouse gas flux," in *Advances in Agronomy*, *Vol. 139*. ed. D. L. Sparks (Oxford, UK: Elsevier), 1–71.

Larose, C., Dommergue, A., and Vogel, T. M. (2013). The dynamic arctic snow pack: an unexplored environment for microbial diversity and activity. *Biology* 2, 317–330. doi: 10.3390/biology2010317

Li, X., Li, Z., Ding, Y., Liu, S., Zhao, Z., Luo, L., et al. (2007). Seasonal variations of pH and electrical conductivity in a snow-firn pack on glacier no. 1, eastern Tianshan, China. *Cold Reg. Sci. Technol.* 48, 55–63. doi: 10.1016/j.coldregions.2006.09.006

Li, W., and Stevens, M. H. H. (2012). Fluctuating resource availability increases invasibility in microbial microcosms. *Oikos* 121, 435–441. doi: 10.1111/j.1600-0706.2011.19762.x

Li, S.-p., Tan, J., Yang, X., Ma, C., and Jiang, L. (2019). Niche and fitness differences determine invasion success and impact in laboratory bacterial communities. *ISME J.* 13, 402–412. doi: 10.1038/s41396-018-0283-x

Lipson, D. A., and Schmidt, S. K. (2004). Seasonal changes in an alpine soil bacterial community in the Colorado Rocky Mountains. *Appl. Environ. Microbiol.* 70, 2867–2879. doi: 10.1128/AEM.70.5.2867-2879.2004

Litchman, E. (2010). Invisible invaders: non-pathogenic invasive microbes in aquatic and terrestrial ecosystems. *Ecol. Lett.* 13, 1560–1572. doi: 10.1111/j.1461-0248.2010.01544.x

Liu, M., Bjørnlund, L., Rønn, R., Christensen, S., and Ekelund, F. (2012). Disturbance promotes non-indigenous bacterial invasion in soil microcosms: analysis of the roles of resource availability and community structure. *PLoS One* 7. doi: 10.1371/journal.pone.0045306

Maccario, L., Sanguino, L., Vogel, T. M., and Larose, C. (2015). Snow and ice ecosystems: not so extreme. *Res. Microbiol.* 166, 782–795. doi: 10.1016/j.resmic.2015.09.002

Maki, T., Lee, K. C., Kawai, K., Onishi, K., Hong, C. S., Kurosaki, Y., et al. (2019). Aeolian dispersal of bacteria associated with desert dust and anthropogenic particles over continental and oceanic surfaces. *J. Geophys. Res. Atmos.* 124, 5579–5588. doi: 10.1029/2018JD029597

Malard, L. A., Anwar, M. Z., Jacobsen, C. S., and Pearce, D. A. (2019a). Biogeographical patterns in soil bacterial communities across the Arctic region. *FEMS Microbiol. Ecol.* 95. doi: 10.1093/femsec/fiz128

Malard, L.A., Mod, H.K., Guex, N., Broennimann, O., Yashiro, E., Lara, E., et al. (2021). Comparative Analysis of Diversity and Environmental Niches of Soil Bacterial, Archaeal, Fungal and Protist Communities Reveal Niche Divergences Along Environmental Gradients in the Alps. Available at: Research Square. doi: 10.21203/rs.3.rs-609984/v2

Malard, L. A., and Pearce, D. A. (2018). Microbial diversity and biogeography in Arctic soils. *Environ. Microbiol. Rep.* 10, 611–625. doi: 10.1111/1758-2229.12680

Malard, L. A., Šabacká, M., Magiopoulos, I., Mowlem, M., Hodson, A., Tranter, M., et al. (2019b). Spatial variability of Antarctic surface snow bacterial communities. *Front. Microbiol.* 10:461. doi: 10.3389/fmicb.2019.00461

Mallon, C. A., Le Roux, X., Van Doorn, G., Dini-Andreote, F., Poly, F., and Salles, J. (2018). The impact of failure: unsuccessful bacterial invasions steer the soil microbial community away from the invader's niche. *ISME J.* 12, 728–741. doi: 10.1038/s41396-017-0003-y

Mallon, C. A., Poly, F., Le Roux, X., Marring, I., van Elsas, J. D., and Salles, J. F. (2015b). Resource pulses can alleviate the biodiversity–invasion relationship in soil microbial communities. *Ecology* 96, 915–926. doi: 10.1890/14-1001.1

Mallon, C. A., Van Elsas, J. D., and Salles, J. F. (2015a). Microbial invasions: the process, patterns, and mechanisms. *Trends Microbiol.* 23, 719–729. doi: 10.1016/j.tim.2015.07.013

Männistö, M. K., Tiirola, M., and Häggblom, M. M. (2007). Bacterial communities in Arctic fjelds of Finnish Lapland are stable but highly pH-dependent. *FEMS Microbiol. Ecol.* 59, 452–465. doi: 10.1111/j.1574-6941.2006.00232.x

Margesin, R., and Miteva, V. (2011). Diversity and ecology of psychrophilic microorganisms. *Res. Microbiol.* 162, 346–361. doi: 10.1016/j.resmic.2010.12.004

Marsh, P., and Woo, M. K. (1984). Wetting front advance and freezing of meltwater within a snow cover: 1. Observations in the Canadian Arctic. *Water Resour. Res.* 20, 1853–1864. doi: 10.1029/WR020i012p01853

McMurdie, P. J., and Holmes, S. (2013). phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* 8:e61217. doi: 10.1371/journal.pone.0061217

Morris, S. J., and Blackwood, C. B. (2015). "The ecology of the soil biota and their function," in *Soil Microbiology, Ecology and Biochemistry*, ed. E. A. Paul (Oxford UK: Academic Press). 273–309.

Morton, J. T., Toran, L., Edlund, A., Metcalf, J. L., Lauber, C., and Knight, R. (2017). Uncovering the horseshoe effect in microbial analyses. *mSystems* 2. doi: 10.1128/mSystems.00166-16

Muhlbachova, G., and Tlustos, P. (2006). Effects of liming on the microbial biomass and its activities in soils long-term contaminated by toxic elements. *Plant Soil Environ.* 52, 345–352. doi: 10.17221/3451-PSE

Newton, A. (1982). Red-coloured snow algae in Svalbard—some environmental factors determining the distribution of Chlamydomonas nivalis (Chlorophyta volvocales). *Polar Biol.* 1, 167–172. doi: 10.1007/BF00287003

Nicol, G. W., Leininger, S., Schleper, C., and Prosser, J. I. (2008). The influence of soil pH on the diversity, abundance and transcriptional activity of ammonia oxidizing archaea and bacteria. *Environ. Microbiol.* 10, 2966–2978. doi: 10.1111/j.1462-2920.2008.01701.x

Orwin, K. H., Dickie, I. A., Wood, J. R., Bonner, K. I., and Holdaway, R. J. (2016). Soil microbial community structure explains the resistance of respiration to a dry–rewet cycle, but not soil functioning under static conditions. *Funct. Ecol.* 30, 1430–1439. doi: 10.1111/1365-2435.12610

Paliy, O., and Shankar, V. (2016). Application of multivariate statistical techniques in microbial ecology. *Mol. Ecol.* 25, 1032–1057. doi: 10.1111/mec.13536

Pandit, S. N., Kolasa, J., and Cottenie, K. (2009). Contrasts between habitat generalists and specialists: an empirical extension to the basic metacommunity framework. *Ecology* 90, 2253–2262. doi: 10.1890/08-0851.1

Pearce, D. A., Bridge, P. D., Hughes, K. A., Sattler, B., Psenner, R., and Russell, N. J. (2009). Microorganisms in the atmosphere over Antarctica. *FEMS Microbiol. Ecol.* 69, 143–157. doi: 10.1111/j.1574-6941.2009.00706.x

Pearce, D. A., Hughes, K., Lachlan-Cope, T., Harangozo, S., and Jones, A. E. (2010). Biodiversity of air-borne microorganisms at Halley station, Antarctica. *Extremophiles* 14, 145–159. doi: 10.1007/s00792-009-0293-8

Peter, H., Hörtnagl, P., Reche, I., and Sommaruga, R. (2014). Bacterial diversity and composition during rain events with and without S aharan dust influence reaching a high mountain lake in the A lps. *Environ. Microbiol. Rep.* 6, 618–624. doi: 10.1111/1758-2229.12175

Pruesse, E., Quast, C., Knittel, K., Fuchs, B. M., Ludwig, W., Peplies, J., et al. (2007). SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res.* 35, 7188–7196. doi: 10.1093/nar/gkm864

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2012). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* 41, D590–D596. doi: 10.1093/nar/gks1219

Ramette, A. (2007). Multivariate analyses in microbial ecology. *FEMS Microbiol. Ecol.* 62, 142–160. doi: 10.1111/j.1574-6941.2007.00375.x

Reche, I., D'Orta, G., Mladenov, N., Winget, D. M., and Suttle, C. A. (2018). Deposition rates of viruses and bacteria above the atmospheric boundary layer. *ISME J.* 12, 1154–1162. doi: 10.1038/s41396-017-0042-4

Robson, A., and Abbott, L. (1989). "The effect of soil acidity on microbial activity in soils," in *Soil Acidity and Plant Growth*, ed. A. Robson (Oxford, UK: Academic Press) 139–165.

Rousk, J., Bååth, E., Brookes, P. C., Lauber, C. L., Lozupone, C., Caporaso, J. G., et al. (2010). Soil bacterial and fungal communities across a pH gradient in an arable soil. *ISME J.* 4, 1340–1351. doi: 10.1038/ismej.2010.58

Russell, N. (1990). Cold adaptation of microorganisms. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 326, 595–611. doi: 10.1098/rstb.1990.0034

Šantl-Temkiv, T., Gosewinkel, U., Starnawski, P., Lever, M., and Finster, K. (2018). Aeolian dispersal of bacteria in Southwest Greenland: their sources, abundance, diversity and physiological states. *FEMS Microbiol. Ecol.* 94:fiy031. doi: 10.1093/femsec/fiy031

Schmidt, S., and Lipson, D. (2004). Microbial growth under the snow: implications for nutrient and allelochemical availability in temperate soils. *Plant Soil* 259, 1–7. doi: 10.1023/B:PLSO.0000020933.32473.7e

Smets, W., Moretti, S., Denys, S., and Lebeer, S. (2016). Airborne bacteria in the atmosphere: presence, purpose, and potential. *Atmos. Environ.* 139, 214–221. doi: 10.1016/j.atmosenv.2016.05.038

Smith, D. J., Timonen, H. J., Jaffe, D. A., Griffin, D. W., Birmele, M. N., Perry, K. D., et al. (2013). Intercontinental dispersal of bacteria and archaea by transpacific winds. *Appl. Environ. Microbiol.* 79, 1134–1139. doi: 10.1128/AEM.03029-12

Soulides, D., and Allison, F. (1961). Effect of drying and freezing soils on carbon dioxide production, available mineral nutrients, aggregation, and bacterial population. *Soil Sci.* 91, 291–298. doi: 10.1097/00010694-196105000-00001

Sriswasdi, S., Yang, C.-C., and Iwasaki, W. (2017). Generalist species drive microbial dispersion and evolution. *Nat. Commun.* 8, 1–8. doi: 10.1038/s41467-017-01265-1

Suzuki, M. T., Taylor, L. T., and DeLong, E. F. (2000). Quantitative analysis of small-subunit rRNA genes in mixed microbial populations via 5′-nuclease assays. *Appl. Environ. Microbiol.* 66, 4605–4614. doi: 10.1128/AEM.66.11.4605-4614.2000

Torsvik, V., and Øvreås, L. (2008). "Microbial diversity, life strategies, and adaptation to life in extreme soils," in *Microbiology of Extreme Soils*. eds. P. Dion, and C. Shekhar Nautiyal (Berlin, Heidelberg: Springer), 15–43.

van Elsas, J. D., Chiurazzi, M., Mallon, C. A., Elhottovā, D., Krištůfek, V., and Salles, J. F. (2012). Microbial diversity determines the invasion of soil by a bacterial pathogen. *Proc. Natl. Acad. Sci.* 109, 1159–1164. doi: 10.1073/pnas.1109326109

Van Nevel, S., De Roy, K., and Boon, N. (2013). Bacterial invasion potential in water is determined by nutrient availability and the indigenous community. *FEMS Microbiol. Ecol.* 85, 593–603. doi: 10.1111/1574-6941.12145

Vila, J. C., Jones, M. L., Patel, M., Bell, T., and Rosindell, J. (2019). Uncovering the rules of microbial community invasions. *Nat. Ecol. Evol.* 3, 1162–1171. doi: 10.1038/s41559-019-0952-9

Voroney, R., and Heck, R. (2014). "The soil habitat," in *Soil Microbiology, Ecology, and Biochemistry*. ed. E. A. Paul (Oxford, UK: Academic Press), 15–39.

Ward, N. L., Challacombe, J. F., Janssen, P. H., Henrissat, B., Coutinho, P. M., Wu, M., et al. (2009). Three genomes from the phylum Acidobacteria provide insight into the lifestyles of these microorganisms in soils. *Appl. Environ. Microbiol.* 75, 2046–2056. doi: 10.1128/AEM.02294-08

Weil, T., De Filippo, C., Albanese, D., Donati, C., Pindo, M., Pavarini, L., et al. (2017). Legal immigrants: invasion of alien microbial communities during winter occurring desert dust storms. *Microbiome* 5:32. doi: 10.1186/s40168-017-0249-7

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis.* New York, NY: Springer.

Winther, J.-G., Bruland, O., Sand, K., Gerland, S., Marechal, D., Ivanov, B., et al. (2003). Snow research in Svalbard—an overview. *Polar Res.* 22, 125–144. doi: 10.3402/polar.v22i2.6451

Winther, J.-G., Godtliebsen, F., Gerland, S., and Isachsen, P. E. (2002). Surface albedo in Ny-Ålesund, Svalbard: variability and trends during 1981–1997. *Glob. Planet. Chang.* 32, 127–139. doi: 10.1016/S0921-8181(01)00103-5

Wipf, S., and Rixen, C. (2010). A review of snow manipulation experiments in Arctic and alpine tundra ecosystems. *Polar Res.* 29, 95–109. doi: 10.1111/j.1751-8369.2010.00153.x

Yanai, Y., Toyota, K., and Okazaki, M. (2004). Effects of successive soil freeze-thaw cycles on soil microbial biomass and organic matter decomposition potential of soils. *Soil Sci. Plant Nutr.* 50, 821–829. doi: 10.1080/00380768.2004.10408542

Zhang, Y., Xiu, G., Wu, X., Moore, C. W., Wang, J., Cai, J., et al. (2013). Characterization of mercury concentrations in snow and potential sources, Shanghai, China. *Sci. Total Environ.* 449, 434–442. doi: 10.1016/j.scitotenv.2013.01.088

Zhang, S., Yang, G., Wang, Y., and Hou, S. (2010). Abundance and community of snow bacteria from three glaciers in the Tibetan plateau. *J. Environ. Sci.* 22, 1418–1424. doi: 10.1016/S1001-0742(09)60269-2

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

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership