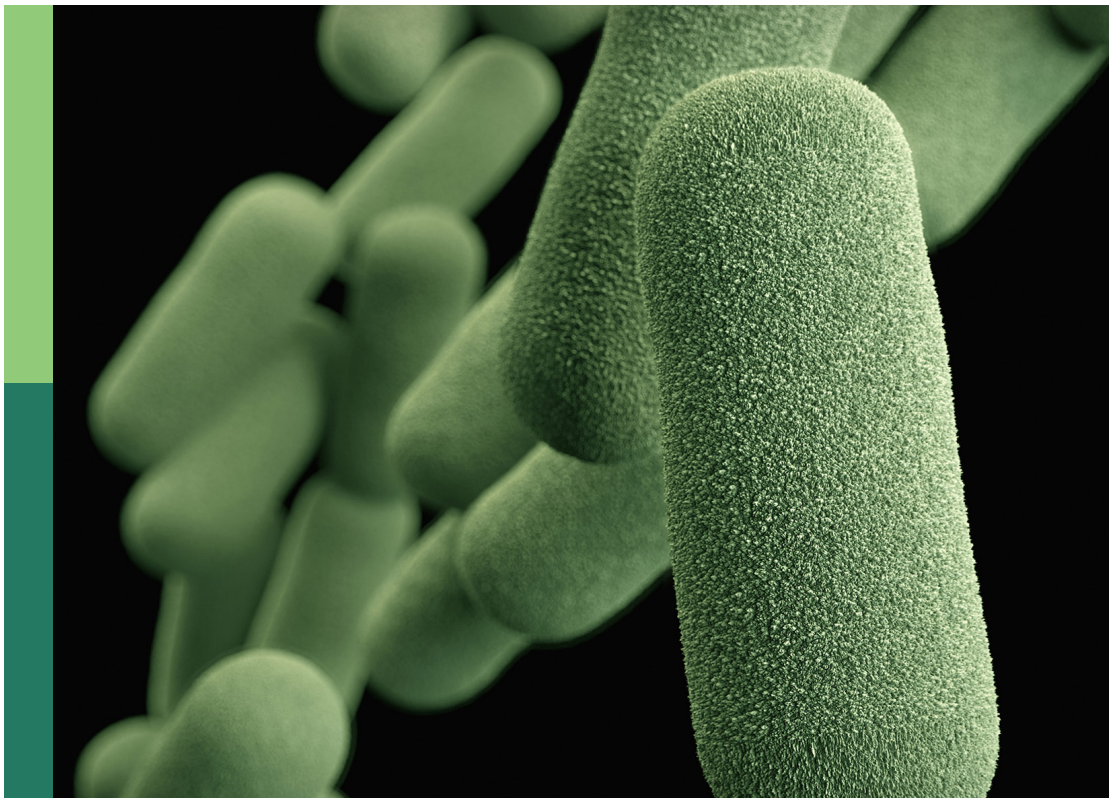
frontiers | Research Topics

Ecology, evolution, and biodiversity of microbiomes and viromes from extreme environments

Edited by
Janina Rahlff, Gareth Trubl, Ramesh K. Goel and Lucie Malard

Published in
Frontiers in Microbiomes



ISSN 1664-8714
ISBN 978-2-8325-6322-9
DOI 10.3389/978-2-8325-6322-9

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

Ecology, evolution, and biodiversity of microbiomes and viromes from extreme environments

Topic editors

Janina Rahlff — Friedrich Schiller University Jena, Germany
Gareth Trubl — Lawrence Livermore National Laboratory (DOE), United States
Ramesh K. Goel — The University of Utah, United States
Lucie Malard — Université de Genève, Switzerland

Citation

Rahlff, J., Trubl, G., Goel, R. K., Malard, L., eds. (2025). *Ecology, evolution, and biodiversity of microbiomes and viromes from extreme environments*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-6322-9

Table of contents

frontiers | Frontiers in Microbiomes

TYPE Editorial
PUBLISHED 24 April 2025
DOI 10.3389/frmbi.2025.1604002

OPEN ACCESS

EDITED AND REVIEWED BY
Helen L. Hayden,
RMIT University, Australia

*CORRESPONDENCE
Janina Rahlff
janina.rahlff@lnu.se

RECEIVED 01 April 2025
ACCEPTED 08 April 2025
PUBLISHED 24 April 2025

CITATION
Trubl G, Malard L and Rahlff J (2025) Editorial: Ecology, evolution, and biodiversity of microbiomes and viromes from extreme environments. *Front. Microbiomes* 4:1604002. doi: 10.3389/frmbi.2025.1604002

Editorial: Ecology, evolution, and biodiversity of microbiomes and viromes from extreme environments

Gareth Trubl[1], Lucie Malard[2] and Janina Rahlff[3,4,5]*

[1]Physical and Life Sciences Directorate, Lawrence Livermore National Laboratory, Livermore, CA, United States, [2]Department F.-A. Forel for Environmental and Aquatic Sciences, Faculty of Sciences, University of Geneva, Geneva, Switzerland, [3]Centre for Ecology and Evolution in Microbial Model Systems (EEMiS), Department of Biology and Environmental Science, Linnaeus University, Kalmar, Sweden, [4]Aero-Aquatic Virus Research Group, Faculty of Mathematics and Computer Science, Friedrich Schiller University Jena, Jena, Germany, [5]Leibniz Institute on Aging - Fritz Lipmann Institute (FLI), Jena, Germany

KEYWORDS

microbial communities, extreme environments, metagenomics, viromics, adaptation, biotechnology

Editorial on the Research Topic
Ecology, evolution, and biodiversity of microbiomes and viromes from extreme environments

Ecology, evolution, and biodiversity of microbiomes and viromes in extreme environments are key areas of research that explore how microbial communities adapt, survive, and thrive under harsh conditions. The studies published in our Research Topic advance our understanding of microbial and viral diversity, evolutionary processes, and the ecological roles of these communities, with implications for biotechnology, climate resilience, and even astrobiology.

Microbial life in extreme environments is shaped by the interplay between environmental pressures and adaptive strategies. Deep-sea hydrothermal vents serve as natural laboratories for studying microbial evolution and adaptation, offering analogs for primordial biochemical processes. A study by Alian et al. on the Lost City Hydrothermal Field near the Mid-Atlantic Ridge investigated microbial communities thriving within its actively venting carbonate chimneys. Their findings revealed that bacterial communities—including *Desulfotomaculum, Sulfurovum, Thiomicrorhabdus*, and *Serpentinicella*—exhibit high levels of microdiversity, driven by mineral composition rather than just vent fluid chemistry. Shotgun metagenomic analyses further showed that genes involved in carbon, methane, nitrogen, and sulfur cycling are closely linked to specific mineralogical characteristics, highlighting the importance of microbe-mineral interactions in shaping microbial community structure and function.

Moving from the ocean depths to the terrestrial extreme of hypersaline environments, Ionescu et al. explored microbial biofilms in Dead Sea underwater springs, where bacteria must rapidly adjust to unpredictable fluctuations in salinity, pH, and oxygen. Metagenomic analysis of an enrichment culture identified key taxa—*Prosthecochloris, Flexistipes, Izemoplasma, Halomonas*, and *Halanaerobium*—that employ both the "salt-in" and

"salt-out" osmoregulation strategies, allowing scalable adaptation similar to halophilic archaea. Genomic data indicate that while *Izemoplasma* relies on compatible solutes, others use mechanosensitive channels for rapid ionic adjustments. The findings challenge the conventional dichotomy of osmoregulation strategies and suggest a flexible mixed approach, expanding our understanding of microbial resilience in fluctuating environments.

Furthermore, the study by Tammert et al. explored the impact of increasing salinity on freshwater sediment and water bacterial communities, an increasing concern from industrial salt discharges and climate change-induced salinization. They showed that bacterial communities, especially from sediments, are highly resistant to increasing salinity but that the community adapts when faced with long-term exposure. For instance, the abundance of *Hydrogenophaga* increased as salinity increased. This shift indicates that although the community withstands salinity changes, individual species may be differentially affected.

Beyond extreme salinity, extreme hydrocarbon environments also harbor diverse microbial life with unique metabolic adaptations. Beilig et al. advanced our understanding of anaerobic oil degradation by employing reverse stable isotope labeling to quantify microbial mineralization rates in light oil reservoirs. Oil reservoirs contain microbial communities capable of degrading crude oil over geologic timescales, yet the actual rates of degradation remain poorly constrained. By incubating formation water from a newly drilled well, they estimated a degradation rate of 15.2 mmol CO_2/mol CH_2 per year—closely aligning with modeled reservoir degradation predictions. Their study highlights the role of sulfate-reducing and fermentative microbial communities, with Desulfobacterota, Thermotogota, and Bacteroidota emerging as key taxa, while also demonstrating the potential of reverse stable isotope labeling as a powerful tool to measure microbial degradation in complex hydrocarbon environments.

Moving into the cold, nutrient-poor conditions of icy landscapes, Faber et al. investigated the meltwater microbial communities on a glacial ice surface in Alaska. The metagenomic analysis revealed unique communities enriched in Gammaproteobacteria, Dothideomycetes, Microbotryomyetes and algae, with three times more fungal genes, associated homeostasis, cellular regulation, and stress responses. These findings contribute to understanding microbial life in glacial ecosystems and provide insights into their adaptation to extreme, cold environments.

Continuing our journey in the cold Swiss Alps, Busi et al. examined the stability of benthic biofilms in proglacial streams, focusing on interactions between bacterial and fungal communities. The authors showed that co-occurrence networks were relatively stable and that the identified bacterial and eukaryotic keystone taxa were critical for network stability. Interestingly, low-abundance taxa played significant roles in stabilizing microbial communities. These findings suggest that microbial functions of keystone taxa may be essential to community stability.

Still in the Swiss Alps, Peter et al. investigated the viral communities in sediment-derived benthic biofilm communities from proglacial streams, revealing that viral community composition closely mirrors bacterial diversity, with high host specificity and potential functional roles in microdiverse bacterial clades through auxiliary metabolic genes (AMGs), such as those involved in the metabolism of cofactors and vitamins. The work demonstrates bacteria-phage interactions and ecological dynamics in streams influenced by glacial meltwaters, which are considered extreme environments due to their harsh conditions, such as near-freezing temperatures and low nutrient availability (ultra-oligotrophy). They highlight that viruses also play a critical role in extreme ecosystems by shaping microbial diversity, influencing host metabolism, and driving biogeochemical cycles.

Finally, Oliveira et al. conducted research on the Great Amazon Reef System (GARS), revealing how sponges in this unique, low-light, sediment-rich environment rely on their microbiomes to support biogeochemical cycles and host nutrition. The analysis of metagenome-assembled genomes (MAGs) from GARS sponges, particularly the candidate phylum Latescibacterota, uncovered metabolic pathways related to nutrient cycling, pollutant degradation, and bioactive compound production. These findings emphasize the role of microbial symbionts in supporting host survival in extreme environments and highlight their promising potential for bioremediation and biotechnological applications.

Together, these studies provide a comprehensive look at how microbial communities and their viral counterparts adapt to extreme conditions, from the deep sea and hypersaline springs to hydrocarbon reservoirs, glacial meltwaters, and mesophotic reefs. Despite the diversity of these environments, common themes emerge: microbial adaptation is driven by selective pressures that shape metabolic flexibility, symbiotic relationships, and evolutionary innovation. Whether through microbe-mineral interactions in hydrothermal vents, scalable osmoregulation in fluctuating salinities, or virus-host dynamics in nutrient-limited streams, these studies collectively advance our understanding of life's resilience in extreme habitats. By integrating genomic, geochemical, and isotopic approaches, this Research Topic provides novel insights into microbial and viral evolution, with broader implications for Earth's biogeochemical cycles, ecosystem stability, and even the search for life beyond our planet.

Author contributions

GT: Writing – original draft, Writing – review & editing. LM: Writing – original draft, Writing – review & editing. JR: Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. GT was supported by LLNL's U.S. Department of Energy, Office of Biological and Environmental Research, Genomic Science Program "Microbes Persist" Scientific Focus Area (SCW1632) and work conducted at

LLNL was conducted under the auspices of the US Department of Energy under Contract DE-AC52-07NA27344. JR was supported by the Swedish Research Council, Starting Grant ID 2023-03310_VR.

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 author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

frontiers | Frontiers in Microbiomes

TYPE Original Research
PUBLISHED 17 August 2023
DOI 10.3389/frmbi.2023.1206961

OPEN ACCESS

EDITED BY
Janina Rahlff,
Linnaeus University, Sweden

REVIEWED BY
Maryam Rezaei Somee,
Linnaeus University, Sweden
Andrei S. Steindorff,
Berkeley Lab (DOE), United States

*CORRESPONDENCE
Ricardo H. Kruger
kruger@unb.br

RECEIVED 16 April 2023
ACCEPTED 31 July 2023
PUBLISHED 17 August 2023

CITATION
Oliveira RS, Pinto OHB, Quirino BF, de Freitas MAM, Thompson FL, Thompson C and Kruger RH (2023) Genome-resolved metagenomic analysis of Great Amazon Reef System sponge-associated Latescibacterota bacteria and their potential contributions to the host sponge and reef.
Front. Microbiomes 2:1206961.
doi: 10.3389/frmbi.2023.1206961

Genome-resolved metagenomic analysis of Great Amazon Reef System sponge-associated Latescibacterota bacteria and their potential contributions to the host sponge and reef

Rafael S. Oliveira[1], Otávio H. B. Pinto[1], Betania F. Quirino[2,3,4], Mayanne A. M. de Freitas[5], Fabiano Lopes Thompson[5], Cristiane Thompson[5] and Ricardo H. Kruger[1]*

[1]Laboratory of Enzymology, Department of Cell Biology, Institute of Biological Sciences, University of Brasília, Brasília, Brazil, [2]Genetics and Biotechnology Laboratory, Embrapa-Agroenergy, Brasília, Brazil, [3]Genomic Sciences and Biotechnology Graduate Program, Catholic University of Brasília, Brasília, Brazil, [4]Microbial Biology Graduate Program, University of Brasília, Brasília, Brazil, [5]Department of Genetics, Institute of Biology, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil

The Great Amazon Reef System (GARS) is an extensive biogenic reef influenced by a plume layer of sediments. This creates an extreme environment where light is reduced, thus affecting physicochemical properties as well as living organisms such as sponges and their microbiomes. The sponge's microbiome has numerous ecological roles, like participation in biogeochemical cycles and host nutrition, helping the sponge thrive and contributing to the ecosystem. Also, sponges and sponge-associated microorganisms are rich sources of bioactive compounds, and their products are applied in different areas, including textile, pharmaceutical, and food industries. In this context, metagenome-assembled genomes (MAG), obtained from GARS sponges microbiota, were analyzed to predict their ecological function and were prospected for biotechnological features. Thus, in this work, tissues of GARS sponges were collected, their metagenomes were sequenced and assembled, and 1,054 MAGs were recovered. Ten of those MAGs were selected based on their taxonomic classification in the candidate phylum Latescibacterota and this group's abundance in GARS sponges. The workflow consisted of MAG's quality definition, taxonomic classification, metabolic reconstruction, and search for bioactive compounds. Metabolic reconstruction from medium to high-quality MAGs revealed genes related to degradation and synthesis pathways, indicating functions that may be performed by GARS sponge-associated Latescibacterota. Heterotrophy, a recurring attribute in Latescibacterota that might be crucial for GARS sponge holobiont nutrition, was verified by the presence of genes related to respiration and fermentation. Also, the analyzed bacteria may contribute to the host's survival in multiple ways, including host protection via defense systems; aid in nutrient consumption by breaking complex substrates and producing essential nutrients like vitamins and certain amino acids; and detoxification of mercury,

arsenic, ammonia, and hydrogen sulfide. Additionally, genes linked to persistent organic pollutant degradation, including glyphosate, and biogeochemical cycles reactions, such as ammonification, sulfate reduction, thiosulfate disproportionation, phosphorus remineralization, and complex organic matter degradation, were identified, suggesting the participation of these Latescibacterota in bioremediation and nutrient cycling. Finally, the investigated MAGs contain genes for numerous bioactive compounds, including industrial enzymes, secondary metabolites, and biologically active peptides, which may have biotechnological value.

KEYWORDS

great amazon reef system (GARS), host-associated, metagenome-assembled genomes (MAGs), sponge microbiome, turbid reef

1 Introduction

For many years, elevated concentrations of suspended sediments and low light levels were considered threats to marine invertebrates' growth. These conditions might affect sponges by reducing their growth or disturbing their physiology, morphology, and microbial community, resulting in increased bleaching and mortality (Pineda et al., 2017; Zweifler et al., 2021). However, there is currently enough evidence to support that marine invertebrate communities can live in natural turbid reefs and benefit from these circumstances by becoming more resilient to stress and bleaching. To thrive in an environment with high sediment load, turbid sponge communities use strategies to rapidly adapt to these conditions, such as reducing channel openings, raising mucus production and photoefficiency, or using heterotrophy to compensate for reduced photosynthesis (Pineda et al., 2017; Burt et al., 2020; Zweifler et al., 2021). An example of a turbid reef is the Great Amazon Reef System (GARS).

The GARS is an extensive biogenic reef located near the Amazon River mouth, between the Brazil-Guyana border and the Brazilian state of Maranhão (latitude: 5° N to 1° S; longitude: 44° W to 51° W), with an estimated area of 9,500 km^2 and a depth of 220 m (Moura et al., 2016; Francini-Filho et al., 2018). The GARS is composed of multiple habitats that shelter a diversity of species, including algae, sponges, corals, crustaceans, and fish (Moura et al., 2016). The region is greatly influenced by the Amazon River, which discharges up to 300×10^3 m^3/s of water rich in suspended matter. The massive amount of water and suspended sediments discharged alters physicochemical properties of the place, including pH, salinity, dissolved nutrients, and lighting (Mahiques et al., 2019). Additionally, this discharge generates a plume containing dissolved and particulate nutrients, sediments, and microbes. The GARS plume attenuates light, changes nitrogen, and organic matter availability, and may provide nutrients for local sponges (Moura et al., 2016; Francini-Filho et al., 2018; Menezes et al., 2022).

Sponges are sessile aquatic filter feeder invertebrates that can be considered a holobiont formed by the host sponge and its associated organisms (Hentschel et al., 2012; Pita et al., 2018). These associated organisms are found inside or outside cells, usually in the mesohyl, where a diverse microbial community lives (Taylor et al., 2007; Bibi et al., 2017). This microbiome, which may correspond to 40% of the sponge's total volume, includes viruses, bacteria, archaea, and unicellular eukaryotes (Webster and Taylor, 2012). Sponge microbial communities are highly diverse and comprise numerous species of bacteria, including cultivated bacteria, uncultivated bacteria, and members of candidate phyla (KIran et al., 2018). Interactions between microorganisms and the host sponge can be either beneficial or detrimental to the participants. Microorganisms can help the sponge acquire nutrients and protect it against predators and pathogens. On the other hand, the symbionts get a supply of nutrients, obtained during water filtration, as well as shelter from predators and light (Taylor et al., 2007; Selvin et al., 2010). Conversely, microorganisms can harm the sponge by directly acting as a parasite or pathogen, causing damage to its skeletal structure, or indirectly via biofilm formation that promotes surface encrustation, resulting in channel blockage and displacement. On the other side of the interaction, sponges produce inhibitory and selective substances that shape their microbiome and might consume symbionts for nutrition (Taylor et al., 2007; Thacker and Freeman, 2012). Furthermore, sponge-associated microbes are involved in several ecological processes. They can participate in biogeochemical cycles, including carbon, nitrogen, sulfur, and phosphorus, and impact ecological relations, affecting predation and competitiveness (KIran et al., 2018; Pita et al., 2018).

Sponges and sponge-associated microorganisms are rich sources of biologically active compounds. They produce various classes of secondary metabolites such as peptides, polyketides, terpenes, alkaloids, pigments, sterols, and fatty acids, which exhibit relevant properties like antimicrobial, antiparasitic, antitumor, anticancer, anti-inflammatory, antioxidant, and cytotoxic (Bramhachari et al., 2016; Bibi et al., 2017; Brinkmann et al., 2017). The sponge holobiont is also a valuable source of hydrolytic enzymes that transform organic matter into nutrients. Many of these biocatalysts, including amylase, protease, cellulase, lipase, and esterase, are applied in the textile, pharmaceutical, and food industries. Sponge-derived enzymes occasionally show unique

characteristics like thermal, acid, alkaline, and salinity tolerance (Santos-Gandelman et al., 2014). The sponge holobiont has immense biotechnological potential, but due to difficulties in large-scale production, few sponge-derived compounds have progressed to clinical stage or have been approved for use (Brinkmann et al., 2017). Culture-independent methods, such as metagenomics and metatranscriptomics, can be applied to access the biotechnological potential of hard-to-isolate microorganisms and obtain genes to efficiently produce a desired molecule (Pallela and Kim, 2011). Another application of sponge symbionts concerns bioremediation, which involves the conversion of toxic substances, such as oils and heavy metals, into less damaging forms by microorganisms (Abatenh et al., 2017). The sponge's microbiome can also indicate the presence of contaminants and be used as a pollution bioindicator via changes in community composition or the detection of resistance genes (Selvin et al., 2010; Santos-Gandelman et al., 2014).

The candidate phylum Latescibacterota is a group of uncultured bacteria composed of heterotrophic microorganisms. This candidate phylum has been identified in the microbial communities of waters and animals, including marine invertebrates (Vavourakis et al., 2018; Engelberts et al., 2020). In this work, 10 Latescibacterota metagenome-assembled genomes, obtained from bacteria associated with GARS sponges, were investigated to predict their ecological role and potential contributions to the host and the reef, helping to explain how these sponges thrive in a light-reduced habitat. Additionally, knowing that the GARS is a unique location with peculiar characteristics and sponge microbiomes are rich sources of bioactive compounds, biotechnological features were searched for in the MAGs.

2 Materials and methods

2.1 Sample collection and DNA extraction

Sponge samples were collected near the mouth of the Amazon River (latitude: 1.299817; longitude: -46.778867) on September 27, 2014. Samples from sponges of the *Agelas* and *Geodia* genera were collected, including the species *Agelas dispar*, *Agelas clathrodes*, *Agelas clathrodes*, *Geodia* sp., *Geodia* cf. *corticostylifera*, and *Geodia neptuni*. The main criterion for sponge selection was the availability of at least three individuals per genus. After being sampled on the RV, the sponges were immediately stored in liquid nitrogen. Each sponge's tissue was dried and dissected with a scalpel to remove associated macroscopic organisms. Approximately 1 g of sponge tissue was frozen in liquid nitrogen and subsequently pulverized. DNA from a single individual of a selected sponge species was extracted and purified with 4 M guanidine hydrochloride, 50 mM Tris-HCl (pH 8.0), 0.05 M EDTA, 0.5% sodium N-lauroylsarcosine, and 1% β-mercaptoethanol. This was followed by phenol/chloroform extraction, isopropanol precipitation, and resuspension in 50 µL of ultrapure water (Trindade-Silva et al., 2012).

2.2 Metagenome sequencing and assembly

A DNA library was constructed for each sample with Illumina® TruSeq Nano DNA Preparation Kit. Sequencing was performed on the Novaseq system (Illumina®), generating 150 bp paired-end reads at 20 GB sequence coverage (Supplementary Data Set S1). Sickle v1.33 (https://github.com/najoshi/sickle) trimmed the reads, while BBtools v35 (https://sourceforge.net/projects/bbmap/) identified and removed phiX sequences and Illumina adapters. A total of six metagenomes were generated, one for each sponge species. The metagenomes were assembled with metaSPAdes v3.13.0 (Nurk et al., 2017), removing scaffolds shorter than 1 kb. All programs were used with default settings.

2.3 MAGs recovery and selection

Metagenome-assembled genomes (MAGs) were recovered using MaxBin2 v2.2.14 with 40 markers for bacteria and archaea, and 107 markers for bacteria only (Wu et al., 2016). DASTool v1.1.2 aggregated the MAGs (Sieber et al., 2018). uBin v0.9.14 (Bornemann et al., 2023) selected MAGs based on GC content, coverage, and taxonomy. Dereplication was executed with dRep v3.2.2 (Olm et al., 2017) with a threshold of 95%. Default parameters were used in all of these programs. A total of 1,054 MAGs were recovered from the sponges' microbiome. To exclude low-quality MAGs, coverage and contamination were calculated, respectively, by CoverM v0.6.1 (https://github.com/wwood/CoverM) and CheckM v1.0.13 (Parks et al., 2015). Genomes with <70% completeness or >10% contamination were discarded.

The remaining 205 MAGs underwent additional analysis. These MAGs were annotated with Prokka v1.14.5 (Seemann, 2014) and taxonomically assigned using the SILVA database v138 of 16S rRNA genes (Quast et al., 2013). MAGs with inconclusive 16S rRNA classification results were classified using phylogenetic marker genes trees. Based on these preliminary classifications, 10 MAGs were selected for further investigation as they were all assigned to the same under-described group, Latescibacterota. This phylum was selected because it is abundantly found in GARS sponges with low microbial abundance (Menezes et al., 2022), and its members may represent up to 2% of the GARS sponges' metagenomes (Pinto et al., 2022) (Supplementary Data Set S2). Moreover, Latescibacterota members cannot be cultivated yet; therefore, one method of learning more about this candidate is through analysis of their genomes.

2.4 Genomes description

The quality of selected genomes was more accurately assessed based on criteria established by Bowers et al. (2017), including completeness, contamination, presence of 5S, 16S, and 23S rRNA genes, and the quantity of tRNA. CheckM, CoverM, Prokka, and GenomeQC web server (Manchanda et al., 2020) were used to obtain these genomic metrics. Additional assembly statistics that

help define MAG quality, such as the number of scaffolds, assembly size, N50, L50, NG50, LG50, and GC content, were acquired with GenomeQC.

Furthermore, eukaryotic-like proteins (ELP), which contain domains important for symbiosis, were searched using Pannzer2 and CD-Search web servers in default settings. Pannzer2 uses the Uniprot database, and these proteins were manually identified using the description of annotated genes (Törönen and Holm, 2022). CD-Search utilizes a specific conserved domain database (CDD v3.20), and ELPs were manually curated according to the domain name (Marchler-Bauer and Bryant, 2004).

2.5 Taxonomic assignment and genome comparison

GTDB-Tk (Chaumeil et al., 2020; version 2.3.0) assigned the selected MAGs to a taxonomic group using GTDB (Parks et al., 2020; version 214). A codon tree was generated using resources present in PATRIC v3.6.12 (Davis et al., 2020). The "Similar Genome Finder" function found genomes similar to the MAGs, while the "Phylogenetic Tree" function identified homologous groups of single-copy genes, aligned those sequences, and concatenated the alignments. The tree was constructed using maximum likelihood estimation with RAxML (Stamatakis, 2014). FigTree v1.4.4 (https://github.com/rambaut/figtree) was used to visualize the tree, making aesthetic changes for improved visualization. The codon tree was composed of Latescibacterota members, members from a similar phylum (Gemmatimonadetes), and an outgroup of Firmicutes that was used to root the tree.

Additionally, average nucleotide index (ANI) and digital DNA-DNA hybridization (dDDH) were used to compare the 10 MAGs and determine their relatedness. While ANI was calculated on JspeciesWS v.3.9.1 (Richter et al., 2016), dDDH was calculated on GGDC v2.1 (Meier-Kolthoff et al., 2013). Default settings were applied to both web servers.

2.6 Metabolic reconstruction

Metabolic reconstruction was performed by verifying the existence of genes involved in reactions of specific metabolic pathways in KEGG v97.0 (Kanehisa et al., 2017) and Metacyc v26.0 (Caspi et al., 2020) databases. BlastKOALA v2.2 (Kanehisa et al., 2016) and KEGG Mapper v5.0 (Kanehisa and Sato, 2020) were used to reannotate genomes and access KEGG data. The software Pathway Tools v23.0 was used to reconstruct pathways present in Metacyc (Karp et al., 2021). Graphics and tables were generated with LibreOffice Calc and LibreOffice Writer, while the R package "pheatmap" was utilized to create heat maps.

2.7 Biotechnological prospection

Genes that encode industrial enzymes were searched for in annotated genomes. The enzymes were selected based on previously reported applications in scientific papers, using the BRENDA database (Jeske et al., 2019) to aid in the selection (Supplementary Data Set S3). AntiSMASH v6.0 (Blin et al., 2021) was used to identify clusters of genes related to the production of secondary metabolites. DbCAN2 web server (Zhang et al., 2018) searched for carbohydrate-active enzymes (CAZyme) present in CAZy database v08062022 (Drula et al., 2022) and dbCAN HMMdb v11 using three tools: Hmmer, Diamond, and eCAMI. Only the genes identified by these three methods were considered valid.

Additionally, biologically active peptides (BAP) were searched for. Peptide sequences were extracted from annotated genomes using FaBox tools v1.61 (Villesen, 2007). Due to software limitations, only sequences with <100 amino acid residues were retrieved. PeptideRanker calculated the chance that a BAP would be present in a protein sequence. Sequences with a probability >0.5 were selected for functional prediction. AntiCP 2.0 (Agrawal et al., 2021), Antifp server (Agrawal et al., 2018), antiAngioPred server (Ramaprasad et al., 2015), and Toxinpred server (Gupta et al., 2013) predicted, respectively, anticancer, antifungal, anti-inflammatory, and toxin activities. Antimicrobial activity was also predicted using antiBP2 server (Lata et al., 2010) and CAMPr3 (Waghu et al., 2016). A predicted function was considered valid when the calculated probability value was >0.5. All programs were run with default settings.

3 Results

3.1 Genomes description

According to the criteria defined by Bowers et al. (2017), a high-quality genome draft must have 5S, 16S, and 23S rRNA genes, at least 18 tRNA genes, completeness >90%, and contamination <5%. Medium-quality drafts only need completeness ≥50% and contamination <10%. Finally, low-quality drafts have completeness <50% and contamination >10%. Therefore, only three of the selected metagenome-assembled genomes (MAGs) are considered high-quality genome drafts (Supplementary Data Set S4). The others are medium-quality genomes due to contamination above 5%, completeness less than 90%, or the absence of certain rRNA genes. Additional metrics used to define MAG quality were also obtained and are described in Supplementary Data Set S4.

Both Pannzer2 and CD-Search detected several copies of eukaryotic-like proteins (ELPs) genes in all MAGs (Figure 1). The following domains were found in CD-Search: ankyrin repeat (ANK); leucine-rich repeat (LRR); NCL-1, HT2A, and Lin-41 repeat (NHL); SEL1-like repeat (SEL1); tetratricopeptide repeat (TPR); and WD40 repeat (WD40). Pannzer2 identified the same domains found by CD-Search, with the addition of fibronectin type III domain (FN3), pyrroloquinoline quinone repeat (PQQ), and NipSnap family. For both programs, the tetratricopeptide repeat (TRP) was the predominant ELP.

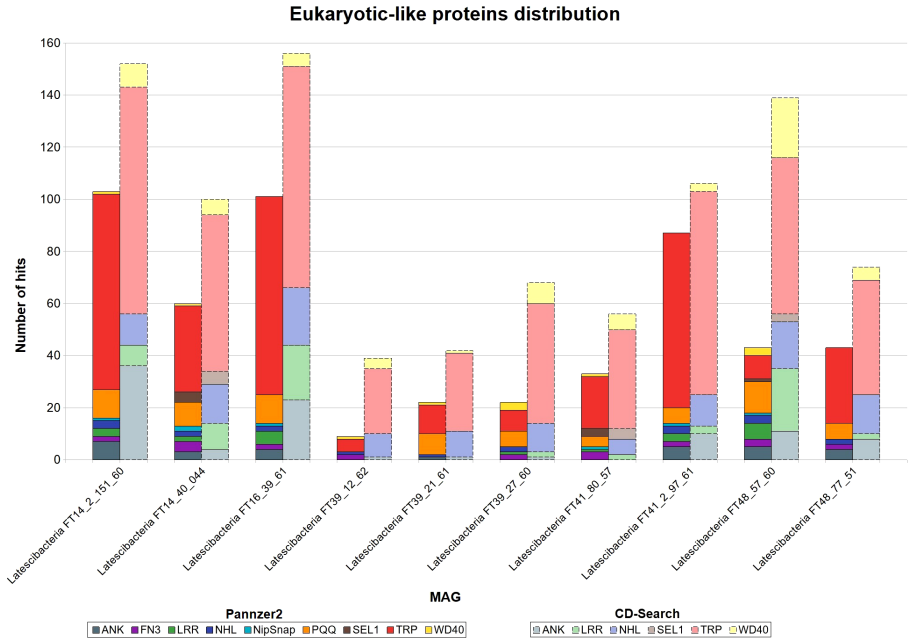


FIGURE 1
Eukaryotic-like proteins distribution in 10 GARS sponge-associated Latstescibacterota MAGs. Results of eukaryotic-like proteins (ELP) identification in each GARS sponge-associated Latescibacterota MAGs, using the programs CD-Search and Pannzer2. Straight bars depict Pannzer2 results while dashed bars depict CD-Search results. ANK, Ankyrin repeat; FN3, Fibronectin type-III repeat; LRR, Leucine-rich repeat; NHL, NCL-1, HT2A and Lin-41 repeat; NipSnap, NipSnap family; PQQ, Pyrrolo-quinoline quinone; SEL1, Sel1-like repeat; TRP, Tetratricopeptide repeat; WD40, WD40 repeat.

3.2 Taxonomic classification

All 10 MAGs showing inconclusive taxonomic affiliation based on 16S rRNA genes were assigned to the candidate phylum Latescibacterota by GTDB-Tk (Supplementary Data Set S4). The codon tree shows the relation between the selected MAGs, other individuals from the candidate phylum Latescibacterota, and members from the close-related phylum Gemmatimonades (Figure 2).

In addition, ANI and dDDH were used to assess how similar the genomes were to one another. ANI values between 95% and 96% or dDDH values ≥70% indicate that the pair of analyzed genomes probably belong to the same species (Chun et al., 2018). Thus, the pair of MAGs Latescibacteria FT14_2_151_60/Latescibacteria FT41_2_97_61 (ANI >96%; dDDH =73.6%) and Latescibacteria FT41_80_57/Latescibacteria FT14_40_044 (ANI >95%; dDDH =72.3%) are likely from the same species (Table 1).

3.3 Metabolic reconstruction

Supplementary Data Set S5 contains a list of all the genes and potentially expressed proteins that were identified during metabolic reconstruction, including transporters and enzymes. All Latescibacterota MAGs contained complete heterotrophic aerobic processes like the tricarboxylic acid (TCA) cycle, the electron transport chain (complexes I–V), and ATP production (Figure 3). As for anaerobic processes, a gene related to the dissimilatory sulfate reduction to hydrogen sulfide (*sat*) was present in seven MAGs, whereas genes related to dissimilatory reduction of nitrate to ammonium (*narI* and *nirD*) were present in six MAGs. Genes for pyruvate fermentation to acetoin (*alsS* and *budC*), ethanol (*adh*), or lactate (*ldh*) were found in all MAGs. Glycolysis and the pentose phosphate pathway were among the complete catabolic processes observed in three MAGs. Moreover, all of the analyzed Latescibacterota have the gene that encodes the enzyme carbon monoxide dehydrogenase, linked to energy generation from carbon monoxide. Furthermore, the gene for hydrogenase, an enzyme used to generate energy from hydrogen, was found in the Latescibacteria FT14_40_044 MAG. Finally, genes for carbon fixation pathway enzymes were found in all MAGs, though all those pathways are incomplete and genes for key enzymes are missing (Figure 3).

In addition, metabolic reconstruction results indicate that GARS sponge-associated Latescibacterota may degrade a variety of organic and inorganic molecules, including carbohydrates, lipids, amino acids, and other sulfur, phosphorous, and nitrogen compounds (Supplementary Data Set S5). The findings also suggest that they can synthesize important molecules such as carbohydrates, lipids, phospholipids, essential amino acids, vitamins, and secondary metabolites (Figure 3). Complete pathways for the production of phosphatidylethanolamine, myo-inositol, vitamins B6 and B7, and the essential amino acids isoleucine, leucine, lysine, phenylalanine, threonine, thryptophan, and valine were identified.

Furthermore, the MAGs contain genes for enzymes involved in specific steps of the nitrogen, sulfur, and phosphorus cycles. Regarding the nitrogen cycle, genes for hydroxylamine oxidation to nitrite (*hao*), nitrate reduction to nitrite (*narI*), and conversion of nitrite to ammonia (*nirD*) were detected (Figure 4A). These genes indicate participation in nitrification, interconversion of nitrate and

10.3389/frmbi.2023.1206961

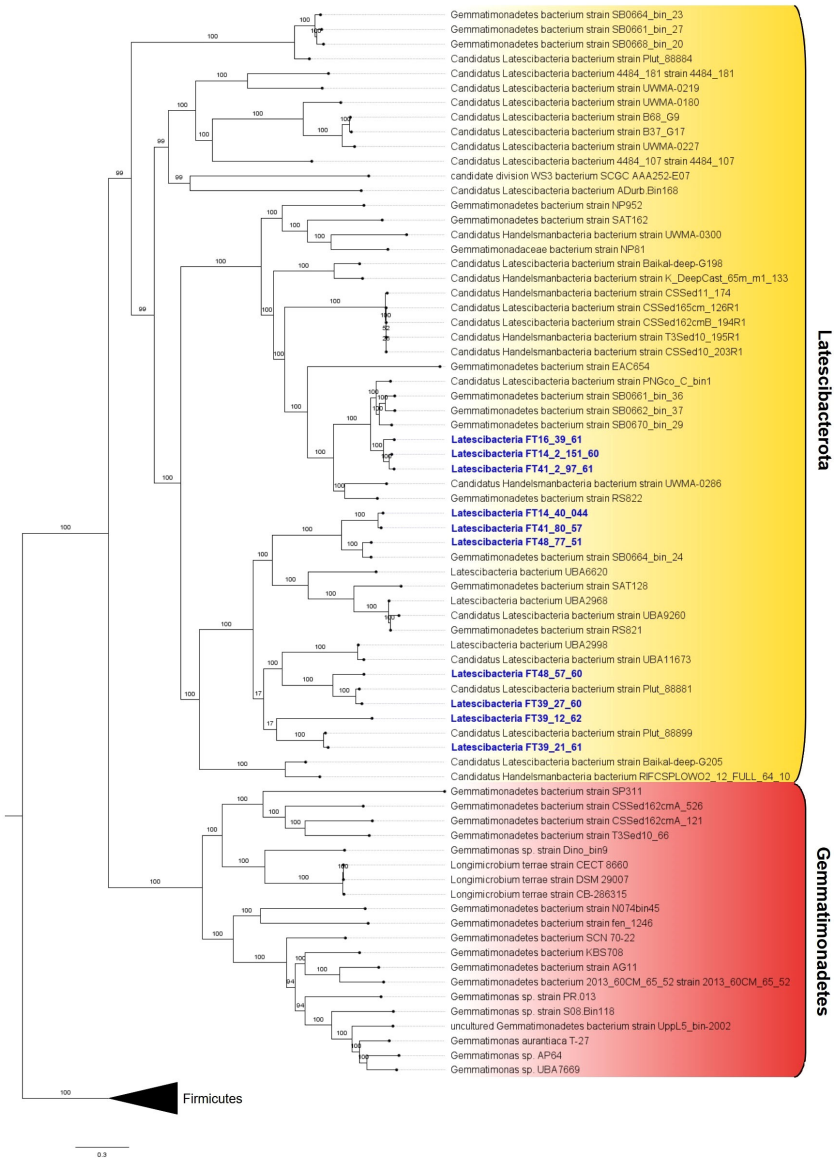


FIGURE 2
Phylogenetic tree with 10 GARS sponge-associated Latstescibacterota MAGs. Codon tree showing the 10 analyzed GARS sponge-associated MAGs (in blue) in relation to Latescibacterota candidate phylum members (in yellow background) and the similar phylum Gemmatimonadetes (in red background). The tree was rooted in the outgroup composed of Firmicutes members.

nitrite, and ammonification. Two genes required for the incorporation of ammonia into amino acids during glutamine synthesis (*glnA* and *gltB*) were observed as well. Concerning the sulfur cycle, the results suggest participation in assimilatory sulfate reduction (*sat*, *cysC*, *cysH*, and *sir*), dissimilatory sulfate reduction (*sat*), thiosulfate disproportionation (*glpE*), and incorporation of sulfide into cysteine (*cysK*) (Figure 4B). As for the phosphorus cycle, metabolic reconstruction indicates a contribution to phosphorus remineralization due to the presence of genes related to organic phosphorus compound degradation, such as aminoethyl phosphonate (*phnW* and *phnY*), methyl phosphonate (*phnPP*), and phosphonoacetate (*phnA*). Additionally, all genes for the polyphosphate cycle (*ppx*, *ppk*, and *ppk2*), a pathway involved in phosphorus storage, were verified in two MAGs (Figure 3).

Moreover, MAGs analysis shows that GARS sponge-associated Latescibacterota might be involved in heavy metals metabolism (Figure 3). All genes for arsenic compound bioremediation (*arsC* and *gap*) and narrow mercury resistance (*merA* and *merP)* were detected in eight and five MAGs, respectively. These microorganisms might also participate in the degradation of harmful organic compounds such as naphthalene, 4-chloronitrobenzene, 1,2-dichloroethane, fluoroacetate, and polybrominated diphenyl ether (PBDE), as verified respectively by the identification of genes *nagAB*, *cnbZ*, *dhmA*, *dehH2*, and *bphA*. Additionally, all MAGs contain the gene for glycine oxidase (*thiO*), which acts in the bioremediation of the pesticide glyphosate, whereas two MAGs have the gene for 2,4-dichlorophenol 6-monooxygenase (*tdfA*), which is involved in the degradation of

TABLE 1 Nucleotide-level genomic similarity comparison between candidate phylum Latescibacterota MAGs by Average nucleotide index (ANI) and digital DNA-DNA hybridization (dDDH) .

		Latescibacteria FT14_2_151_60	Latescibacteria FT14_40_044	Latescibacteria FT16_39_61	Latescibacteria FT39_12_62	Latescibacteria FT39_21_61	Latescibacteria FT39_27_60	Latescibacteria FT41_80_57	Latescibacteria FT41_2_97_61	Latescibacteria FT48_57_60	Latescibacteria FT48_77_51
ANI	Latescibacteria FT14_2_151_60	*	67.4%	90.71%	64.72%	65.32%	64.84%	66.73%	**96.01%**	66.13%	66.32%
	Latescibacteria FT14_40_044	67.56%	*	68.23%	65.69%	67.62%	66.33%	**95.62%**	66.99%	66.83%	77.32%
	Latescibacteria FT16_39_61	90.84%	68.04%	*	64.71%	65.41%	64.95%	67.03%	90.97%	66.08%	66.64%
	Latescibacteria FT39_12_62	64.89%	65.54%	64.71%	*	69.65%	67.64%	65.84%	64.83%	67.71%	64.8%
	Latescibacteria FT39_21_61	65.47%	67.43%	65.7%	69.66%	*	70.04%	67.38%	65.58%	70.48%	66.98%
	Latescibacteria FT39_27_60	66.34%	66.25%	65.4%	67.53%	70.12%	*	66.39%	65.37%	81.03%	66.1%
	Latescibacteria FT41_80_57	66.47%	**96.44%**	66.95%	65.78%	67.43%	66.57%	*	66.12%	66.76%	77.88%
	Latescibacteria FT41_2_97_61	**96.89%**	67.12%	91.69%	64.7%	65.46%	64.88%	66.23%	*	65.99%	66.15%
	Latescibacteria FT48_57_60	65.86%	66.75%	65.94%	67.57%	70.36%	80.58%	66.64%	65.79%	*	66.65%
	Latescibacteria FT48_77_51	66.07%	77.85%	66.51%	65.23%	67.31%	65.94%	77.74%	66.01%	66.63%	*
dDDH	Latescibacteria FT14_2_151_60	*	25.5%	43.2%	21.7%	23.2%	16.9%	24.6%	**73.6%**	22.7%	21%
	Latescibacteria FT14_40_044	25.5%	*	24.8%	17.2%	21.6%	15.3%	**72.3%**	27.9%	18.2%	22.2%
	Latescibacteria FT16_39_61	43.2%	24.8%	*	22.4%	24.8%	21.4%	28%	45.7%	19.2%	24.7%
	Latescibacteria FT39_12_62	21.7%	17.2%	22.4%	*	18.9%	16.6%	28.9%	29.9%	18.1%	23.7%
	Latescibacteria FT39_21_61	23.2%	21.6%	24.8%	18.9%	*	19.2%	19.4%	22.4%	18.4%	28.3%
	Latescibacteria FT39_27_60	16.9%	15.3%	21.4%	16.6%	19.2%	*	15.7%	15.7%	24.8%	16.4%
	Latescibacteria FT41_80_57	24.6%	**72.3%**	28%	28.9%	19.4%	15.7%	*	23.6%	15.4%	22.2%

(Continued)

10.3389/frmbi.2023.1206961

10.3389/frmbi.2023.1206961

TABLE 1 Continued

	Latescibacteria FT14_2_151_60	Latescibacteria FT14_40_044	Latescibacteria FT16_39_61	Latescibacteria FT39_12_62	Latescibacteria FT39_21_61	Latescibacteria FT39_27_60	Latescibacteria FT41_80_57	Latescibacteria FT41_2_97_61	Latescibacteria FT48_57_60	Latescibacteria FT48_77_51
Latescibacteria FT41_2_97_61	**73.6%**	27.9%	45.7%	29.9%	22.4%	15.7%	23.6%	*	19.7%	27.9%
Latescibacteria FT48_57_60	22.7%	18.2%	19.2%	18.1%	18.4%	24.8%	15.4%	19.7%	*	21.9%
Latescibacteria FT48_77_51	21%	22.2%	24.7%	23.7%	28.3%	16.4%	22.2%	27.9%	21.9%	*

Values above the cutoff (≥95% for ANI; ≥70% for dDDH) are highlighted in black background.
*, 100% match.

the herbicides 2,4-dichlorophenoxyacetic acid (2,4-D) and 2-methyl-4-chlorophenoxyacetic acid (MCPA).

Importantly, the following prokaryotic defense mechanisms were observed in the GARS sponge-associated Latescibacterota MAGs: clustered regularly interspaced short palindromic repeats (CRISPR) and CRISPR-associated protein (Cas) system (types I and III); restriction-modification (R-M) system (types I, II, III, and IV); toxin-antitoxin system (type II); and DNA phosphothiolation system (Figure 3). Five MAGs also contain the catalase (katG) and superoxide dismutase (sodA) genes, which encode enzymes that degrade reactive oxygen species (ROS).

Lastly, genes that encode enzymes related to the production of antibiotics were present in all Latescibacterota sponge-associated MAGs (Supplementary Data Set S5). However, all pathways were incomplete. The enzymes found were associated with the production of aurachin (*auaCH*), anthracycline (*dauE*), bacilisin (*bacCG*), beta-lactams (*axeA*, *acyII*, *cefD*, *G7AC*, *PENDE*), fosfomycin (*fom1*), kanamycin (*kanJ*), lividomycin (*livQ*), macrolide (*pikAII*), novobiocin (*novN*), pentalenolactone (*ptlD*), prodigiosin (*pigC*), pyocyanin (*phzF*), pyrrolnitrin (*prnD*), tetracycline (*oxyT*), and chlortetracycline (*ctcP*). Antibiotic resistance genes were also found in the 10 MAGs, indicating resistance to multiple drugs, including aminoglycosides, macrolides, vancomycin, penicillin, and triclosan (Supplementary Data Set S5).

3.4 Biotechnological prospection

The GARS sponge-associated Latescibacterota MAGs contain genes that encode industrial enzymes, including lipase, phospholipase, esterase, protease, aminopeptidase, amidase, dioxygenase, monooxygenase, dehalogenase, nitrile hydratase, alginate-lyase, alcohol dehydrogenase, and glycerol dehydrogenase (Table 2).

Gene clusters involved in secondary metabolites production were identified by AntiSMASH in all GARS sponge-associated Latescibacterota MAGs, with the MAG Latescibacteria FT41_80_57 as an exception (Figure 5). The clusters are linked to the production of terpene, non-ribosomal peptide synthase-like fragment (NRPS-like), linear azole(in)e-containing peptide (LAP), polyketide synthase type I (T1PKS), ranthipeptide, cluster containing RiPP recognition element (RRE), and other unspecified ribosomally synthesized and post-translationally modified peptide product (RiPP-like).

A total of 124 genes for CAZymes were annotated by DbCAN2 in the 10 MAGs (Figure 6). The following modules were identified: glycoside hydrolase (GH); glycosyltransferase (GT); polysaccharide lyase (PL); carbohydrate esterase (CE); and carbohydrate-binding modules (CBM). The auxiliary activity module (AA) was absent. It is worth noting that 13 sequences also have a signal peptide.

A total of 425 peptides were identified in all sponge-associated Latescibacterota MAGs. Of those peptides, 166 were categorized as biologically active peptides (BAP) by PeptideRanker (Figure 7A). In 126 of those BAPs, there was no biological function predicted. Nevertheless, antimicrobial, antifungal, anti-inflammatory, and

10.3389/frmbi.2023.1206961

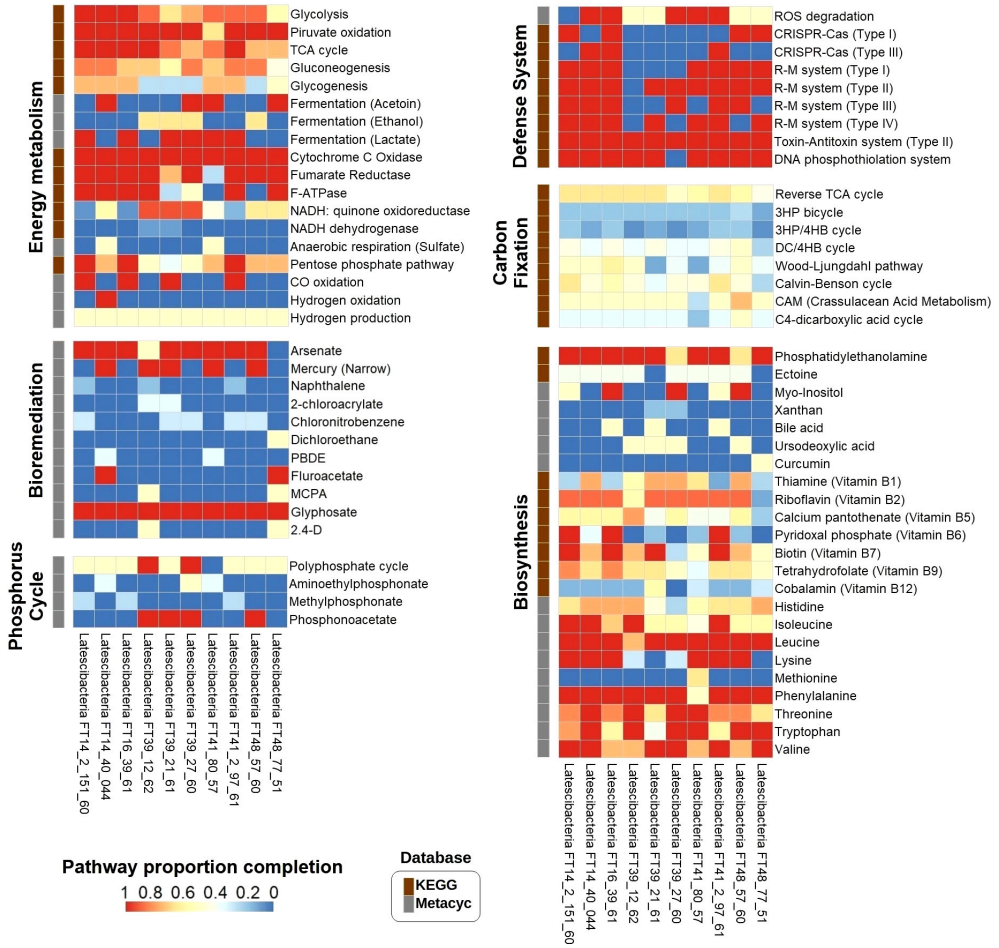


FIGURE 3
Proportion completion of pathways in 10 GARS sponge-associated Latstescibacterota MAGs. Heat map showing the proportion completion of different pathways in each GARS sponge-associated Latstescibacterota MAG based on the presence of specific genes involved in a certain pathway. The described pathways are related to energy metabolism, carbon fixation, bioremediation, biosynthesis, phosphorus cycle, and defense mechanisms. Information about the pathways were obtained from KEGG (brown squares) or Metacyc (grey squares) databases. 3HP, 3-hydroxypropionate; 4HB, 4-hydroxybutyrate; DC, Dicarboxylate; PBDE, Polybrominated diphenyl ether; MCPA, 2-methyl-4-chlorophenoxyaceticacid; 2,4-D, 2,4-dichlorophenoxyacetic acid.

anticancer activities were predicted in 40 BAPs (Figure 7B). For eight peptides, more than one type of activity was predicted. In contrast, Toxinpred did not detect toxin activity in any peptide.

4 Discussion

4.1 Contributions to the sponge and symbiosis

GARS MAGs provide valuable insights into the functional potential of uncultured Latescibacterota associated with the sponge holobiont. These MAGs encode enzymes that degrade complex substrates such as proteases and lipases, releasing nutrients for easier absorption by the sponge (Selvin et al., 2010; Pita et al., 2018). These bacteria also possess transporters that may transfer nutrients and other compounds between the host and symbionts (Fan et al., 2012; Menezes et al., 2022). Additionally, they can synthesize myo-inositol, which serves as an osmolyte or part of cell structures, and phosphatidylethanolamine, which is a structural component of sponges and bacterial membranes (Murzyn et al., 2005; Roberts, 2006; Genin et al., 2008). Moreover, these microorganisms could supply the host's demand for essential nutrients like vitamins and some amino acids (Pita et al., 2018; Robbins et al., 2021).

Another role potentially performed by the GARS sponge-associated Latescibacterota is host defense. Their MAGs contain genes involved in the production of antibiotics, biologically active peptides, and other secondary metabolites with antibiotic activity, which can prevent pathogen colonization (Selvin et al., 2010). Secondary metabolites may also contribute to chemical defense, predator evasion, and physiological improvement, influencing the sponge's susceptibility to predation (Pita et al., 2018). Additionally, the 10 MAGs have genes related to stress proteins, defensive enzymes like phospholipases, and defense systems such as CRISPR-Cas, R-M system, toxin-antitoxin system, and DNA phosphotiolation. These genes, which are generally enriched in sponge symbionts, may protect the sponge against harmful

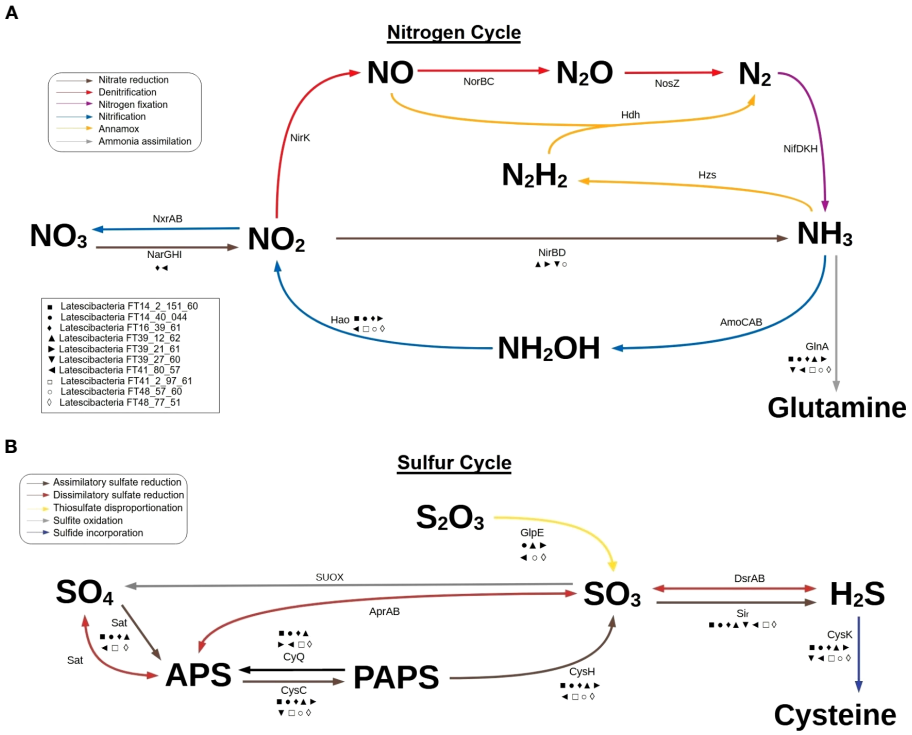


FIGURE 4
Participation of 10 GARS sponge-associated Latescibacterota MAGs in nitrogen cycle and sulfur cycle. Predicted involvement of GARS sponge-associated bacteria from the candidate phylum Latescibacterota in biogeochemical cycles based on the presence of specific genes in MAGs. **(A)** Representation of the nitrogen cycle highlighting the possible participation of GARS Latescibacterota in nitrate reduction, nitrification, and ammonia assimilation. **(B)** Representation of the sulfur cycle, emphasizing the possible participation of GARS Latescibacterota in assimilatory sulfate reduction, dissimilatory sulfate reduction, thiosulfate disproportionation, and sulfide incorporation. Each of the 10 MAGs analyzed is represented by a symbol, as shown in the box. The presence of a MAG symbol below an arrow indicates that the gene involved in the reaction is present in the MAG.

substances and foreign organisms (Selvin et al., 2010; Horn et al., 2016). Moreover, five MAGs contain genes for antioxidant enzymes, which may improve the host's defense against reactive oxygen species by increasing antioxidant activity and neutralizing capability (Regoli et al., 2000).

Furthermore, GARS sponge-associated Latescibacterota could aid in the elimination or mitigation of toxic substances. They may oxidize and assimilate ammonia, one of the main wastes generated by sponges, preventing its accumulation during moments of low water pumping and the consequential death of host tissues (Webster and Taylor, 2012; Pita et al., 2018). These bacteria may also regulate the concentration of hydrogen sulfide, which can be toxic in aquatic environments, by incorporating sulfide into cysteine, similar to the process observed in plants (Bagarinao, 1992; Birke et al., 2012). In addition, GARS sponge-associated Latescibacterota potentially convert mercury and arsenic into less toxic forms, safeguarding the host from their harmful effects. This protection could be crucial as heavy metal exposure impacts the sponge's microbial community, physiology, and overall health (Selvin et al., 2010; Webster and Taylor, 2012).

To establish a symbiotic relationship, sponges can recognize and differentiate symbionts from consumable microbes through immune signaling (Thomas et al., 2010). ELPs, which are usually enriched in sponges and their symbionts, may play a role in this recognition as they are linked to phagocytosis evasion (Pita et al., 2018). Different ELPs were identified in the GARS sponge-associated Latescibacterota, consistent with findings in other sponge-associated microorganisms (Reynolds and Thomas, 2016; Díez-Vives et al., 2017; Frank, 2019).

4.2 Contributions to the reef

Besides aiding their host, GARS sponge-associated Latescibacterota can contribute to the reef by participating in nitrogen, sulfur, and phosphorus cycles. Concerning the nitrogen cycle, genes related to the interconversion of nitrate and nitrite, ammonification, nitrification, and incorporation of ammonia into amino acids were identified. Nitrifying bacteria and ammonia-oxidizing bacteria are frequently found in sponges from different environments (Pajares and Ramos, 2019). About the sulfur cycle, the analyzed Latescibacterota might be involved in assimilatory sulfate reduction, dissimilatory sulfate reduction, and thiosulfate disproportionation. However, this last reaction may serve a different purpose as it is catalyzed by rhodanese, an enzyme used to convert hydrogen cyanide into less toxic forms (Cipollone et al., 2007). Nonetheless, these Latescibacterota are likely sulfate-reducing bacteria, a type of microorganism regularly found in association with sponges (Jensen et al., 2017; Lavy et al., 2018). Regarding the phosphorus cycle, the described Latescibacterota may participate in

TABLE 2 Enzyme-coding genes with biotechnological potencial identified in candidate phylum Latescibacterota MAGs from GARS sponges' microbiome.

Enzyme	EC Number	Latescibacteria FT14_2_151_60	Latescibacteria FT14_40_044	Latescibacteria FT16_39_61	Latescibacteria FT39_12_62	Latescibacteria FT39_21_61	Latescibacteria FT39_27_60	Latescibacteria FT41_80_57	Latescibacteria FT41_2_97_61	Latescibacteria FT48_57_60	Latescibacteria FT48_77_51
Alcohol dehydrogenase	EC1.1.1.2										
Glycerol dehydrogenase	EC1.1.1.6										
Catalase	EC1.11.1.6										
Hydrogenase	EC1.12.-.-										
Dioxygenase	EC1.13.12.-										
Monooxygenase	EC1.14.13.-										
Esterase	EC3.1.1.-										
Lipase	EC3.1.1.3										
Phospholipase	EC3.1.1.32										
Alkaline phosphatase	EC3.1.3.1										
Epoxide hydrolase	EC3.3.2.9										
Protease	EC3.4.-.-										
Glutaminase	EC3.5.1.2										
Amidase	EC3.5.1.4										
Dehalogenase	EC3.8.1.-										
Nitrile hydratase	EC4.2.1.84										
Alginate lyase	EC4.2.2.-										

Green squares indicate the presence of a gene that encodes a certain enzyme in a MAG. Red squares indicate gene's absence.

10.3389/frmbi.2023.1206961

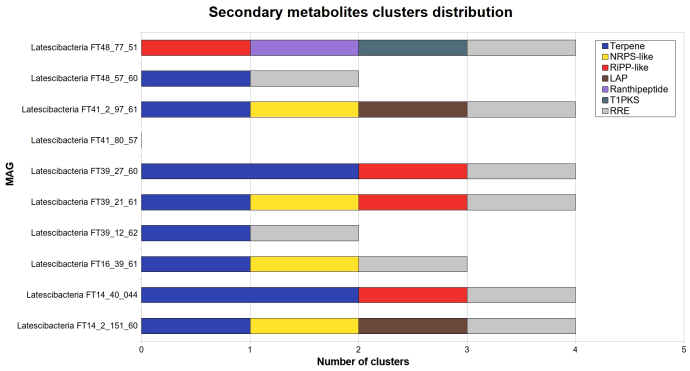


FIGURE 5
Secondary metabolites clusters distribution in 10 GARS sponge-associated Latescibacterota MAGs. Identification of clusters of genes related to secondary metabolites production in GARS sponge-associated Latescibacterota MAGs, according to AntiSMASH. These clusters were absent in Latescibacteria FT41_80_57. The following types of secondary metabolites clusters were identified: Non-ribosomal peptide synthetase-like fragment (NRPS-like); other unspecified ribosomally synthesized and post-translationally modified peptides (RiPP); linear azol(in)e-containing peptides (LAP); type I polyketide synthase(T1PKS); and RiPP recognition element containing cluster (RRE).

phosphorus storage through its sequestration and deposit in polyphosphate granules, which are produced during the polyphosphate cycle. This process affects phosphorus availability as well as marine microbial metabolism and diversity (Dyhrman et al., 2007; Zhang et al., 2015; Pita et al., 2018). There is also possible involvement in remineralization via phosphonate degradation, a pathway commonly found in invertebrates' microbiomes (Villarreal-Chiu et al., 2012; Podell et al., 2020).

The investigated microorganisms might also contribute to bioremediation. GARS sponge-associated Latescibacterota MAGs contain genes related to the degradation of pollutants such as naphthalene, 4-chloronitrobenzene, 1,2-dichloroethane, organohalides, fluoroacetate, PBDE, glyphosate, 2,4-D, and MCPA. However, none of these contaminants directly impact sponges. Conversely, PBDE is produced by a few species of sponges (Agarwal et al., 2017). Nonetheless, some of these pollutants, including naphthalene, glyphosate, and certain types of PBDE, are known persistent organic pollutants in the ocean, disturbing other aquatic organisms (Ipen, 2018; Vagi et al., 2021). Moreover, all MAGs contain genes for bioremediation enzymes like oxygenases, nitrile hydratase, and dehalogenase, which can break, transform, or remove toxic molecules such as aromatics, nitriles, and organohalogens (Karigar and Rao, 2011; Cheng et al., 2020; Oyewusi et al., 2020).

Furthermore, metabolic reconstruction of GARS sponge-associated Latescibacterota revealed an energy metabolism typical of heterotrophic bacteria, with coexisting aerobic and anaerobic processes. Heterotrophy is a feature often found in bacteria associated with sponges, typically through aerobiosis. However, facultative anaerobes and sulfate-reducing bacteria can be found in sponges' microbiomes, possibly due to pumping fluctuations that create moments of anaerobiosis (Osinga et al., 2001; Taylor et al.,

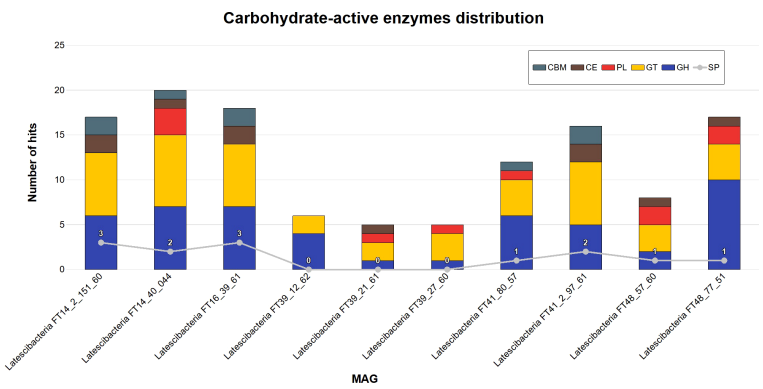


FIGURE 6
Carbohydrate-active enzymes distribution in 10 GARS sponge-associated Latescibacterota MAGs. Five distinct classes of carbohydrate-active enzymes (CAZyme) were identified in GARS sponge-associated Latescibacterota MAGs by DbCAN2: Carbohydrate-binding modules (CBM); Carbohydrate esterases (CE); Polysaccharide lyases (PL); GlycosylTransferases (GT); Glycoside hydrolase (GH). The auxiliary activities class was absent. 13 sequences also contained a signal peptide (SP).

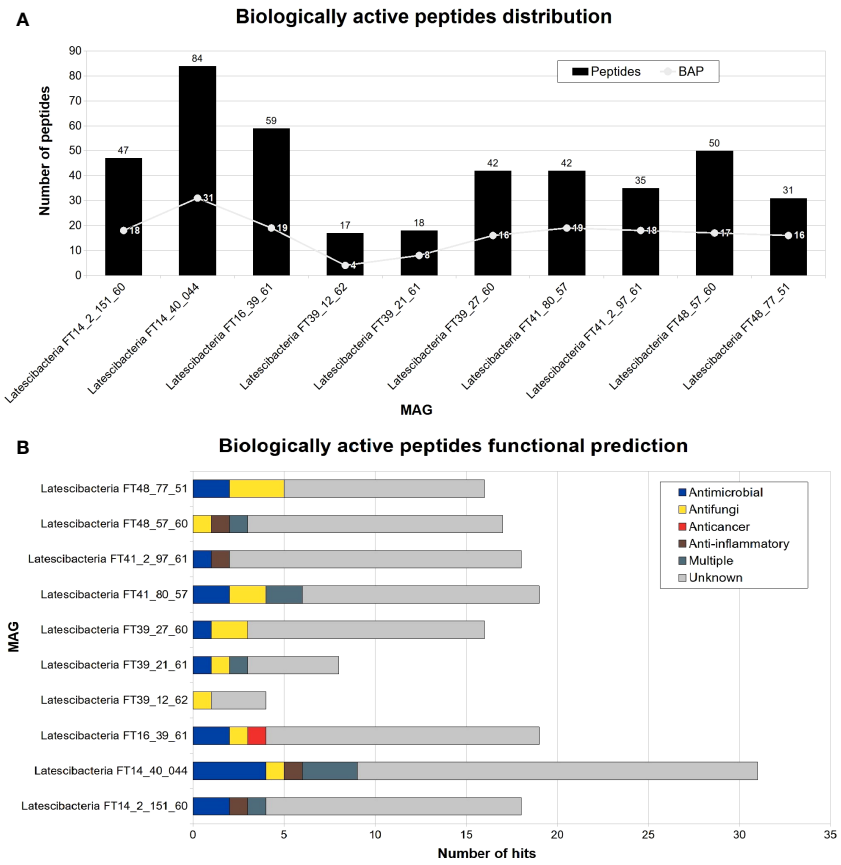


FIGURE 7
Peptides distribution and functional prediction in 10 GARS sponge-associated Latescibacterota MAGs. Number of biologically active peptides (BAP) predicted in each GARS sponge-associated Latescibacterota MAGs. **(A)** Total number of peptides (i.e., amino acid sequences with less than 100 residues) and number of peptides with a predicted biological function. **(B)** Function of predicted biologically active peptides.

2007). In the GARS, heterotrophic processes may be important for sponges, particularly the ones with low microbial abundance, as they can obtain a carbon supply from organic particles in the Amazon River plume (Calegario et al., 2021; Menezes et al., 2022). Carbon fixation is another process occurring in the GARS turbid areas, albeit at a minor level (Pinto et al., 2022). These pathways are frequently seen in sponge symbionts and the candidate phylum Latescibacterota (Feng and Li, 2019; Tran et al., 2021), but were absent or incomplete in the 10 MAGs. Carbon monoxide dehydrogenase, an enzyme frequently present in litohererotrophics, was detected in all MAGs, which may allow the use of carbon monoxide as a carbon source or for energy conservation (Robb and Techtmann, 2018; Burgsdorf et al., 2022).

4.3 MAG's quality and limitations

It is worth mentioning that the MAGs' quality varied from medium to high. The assembly quality is essential for annotation and may affect the quantity and quality of identified genes (Florea et al., 2011), which could explain the predominance of incomplete pathways in the studied MAGs. Another hypothesis that could explain this incompleteness of metabolic pathways is the presence of alternative genes encoding unknown enzymes that may be involved in the missing reactions (Lannes et al., 2019). In some pathways, cooperation with other organisms in the community may be necessary to complete metabolic processes. Through the metabolic division of labor, different populations execute complementary reactions of a particular pathway, as seen in nitrification (Embree et al., 2015; Tsoi et al., 2018). Microorganisms can also interact with the host through cosynthesis or cometabolism. In sponges, integration between symbiont and host genes may be required to complete a pathway, or the product of one organism's reactions may act as an input in the other's pathway (Selvin et al., 2010). It is important to highlight that inferences were made using DNA sequences and did not include information about gene expression. Having a specific gene in a MAG does not guarantee that a particular process is being executed. Instead, it suggests potential for its execution.

4.4 Biotechnological prospection

In addition to their ecological role, GARS sponge-associated Latescibacterota may serve as a significant source of natural products as they potentially produce relevant metabolites. For

example, these bacteria may produce myo-inositol, which is used in the food, pharmaceutical, and cosmetic industries as a beverage additive as well as to produce uronic acid and other inositol derivatives (You et al., 2020). Moreover, nine MAGs contain clusters of genes involved in the generation of secondary metabolites, including terpenes, ranthipeptides, LAPs, T1PKS, and NRPS-like proteins, as well as clusters with RRE, a conserved domain that interacts with peptides and is used to identify novel compounds (Burkhart et al., 2015). Sponge symbionts are a known source of NRPS and PKS (Pimentel-Elardo et al., 2012; Trindade-Silva et al., 2013). Except for ranthipeptides, whose function remains unknown, these metabolites may have exploitable activities. For instance, terpenes are employed as a repellent, pigment, flavoring agent, and medical component for infections and cardiovascular diseases (Schempp et al., 2018; Cox-Georgian et al., 2019).

Additionally, genes for commercial enzymes that are applied in many areas were identified in GARS sponge-associated Latescibacterota MAGs. For instance, proteases are used in the food industry to process meat, milk, and beverages, as well as in the production of pharmaceuticals, agrochemicals, and textiles (Adrio and Demain, 2014; Raveendran et al., 2018). Esterases, phospholipases, and aminopeptidases are also used in the food industry to enhance food flavor and fragrance (Raveendran et al., 2018; Nandan and Nampoothiri, 2020). Moreover, five classes of CAZymes, enzymes that act on the synthesis, breakdown, or modification of complex carbohydrates and other glycoconjugates, were found in the 10 MAGs (Cantarel et al., 2009). CAZymes from marine organisms are used in several areas, including bioenergy, food industry, and chemical industry, and may possess unique properties (Oliveira et al., 2020; Contesini et al., 2021).

Furthermore, GARS sponge-associated Latescibacterota showed great potential for generating bioactive peptides, containing possible peptides with antimicrobial, antifungal, anticancer, and anti-inflammatory properties predicted. BAPs are a group of peptides with less than 50 amino acid residues that exhibit a biological function (Macedo et al., 2021; Akbarian et al., 2022). These peptides also have significant pharmaceutical and nutraceutical potential, with commercialized peptide products derived from sponges such as the antibacterial polydiscamide A and the anticancer phakellistatin (Brinkmann et al., 2017; Macedo et al., 2021).

In conclusion, the 10 sponge-associated Latescibacterota may play a significant role in the sponge's success in the GARS. These symbionts may participate in the host's nutrition and protection against pathogens, heavy metals, and other harmful chemicals. Also, the presence of genes related to biogeochemical cycles and bioremediation indicates that these Latescibacterota contribute to nutrient flux in the reef and mitigation of damage from detrimental compounds. Additionally, the predominance of heterotrophic processes is in agreement with a recent hypothesis that GARS organisms might feed from sediments and other particles present in the Amazon River discharge (Menezes et al., 2022). However, it is essential to acquire transcriptional data to confirm the expression of found genes and endorse the execution of discussed functions. Finally, GARS sponge-associated Latescibacterota have great biotechnological value. Their genomes contain genes for industrial enzymes, secondary metabolites, and biologically active compounds that could be explored in future works.

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/genbank/, The sequence datasets generated for this study were deposited at NCBI under Bio-Project number PRJNA795684.

Ethics statement

Ethical review and approval was not required for the study on animals in accordance with the local legislation and institutional requirements.

Author contributions

RO, OP, and RK designed the study. Data analyses were performed by OP and RO. Data collection was performed by FT. The manuscript was drafted by RO and revised by all authors. BQ, MF, FT and CT critically reviewed and substantially edited the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This study was financially supported by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), which also provided financial support for the first author, Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ), and Fundação de Apoio à Pesquisa do Distrito Federal (FAP-DF).

Acknowledgments

The authors thank Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ), and Fundação de Apoio à Pesquisa do Distrito Federal (FAP-DF) for financial support.

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.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frmbi.2023.1206961/full#supplementary-material

References

Abatenh, E., Gizaw, B., Tsegaye, Z., and Wassie, M. (2017). The role of microorganisms in bioremediation- A review. *Open J. Environ. Biol.* 2, 38–46. doi: 10.17352/ojeb

Adrio, J. L., and Demain, A. L. (2014). Microbial enzymes: tools for biotechnological processes. *Biomolecules* 4, 117–139. doi: 10.3390/biom4010117

Agarwal, V., Blanton, J. M., Podell, S., Taton, A., Schorn, M. A., Busch, J., et al. (2017). Metagenomic discovery of polybrominated diphenyl ether biosynthesis by marine sponges. *Nat. Chem. Biol.* 13, 537–543. doi: 10.1038/nchembio.2330

Agrawal, P., Bhagat, D., Mahalwal, M., Sharma, N., and Raghava, G. P. S. (2021). AntiCP 2.0: An updated model for predicting anticancer peptides. *Brief. Bioinform.* 22, 3. doi: 10.1093/bib/bbaa153

Agrawal, P., Bhalla, S., Chaudhary, K., Kumar, R., Sharma, M., and Raghava, G. P. S. (2018). In silico approach for prediction of antifungal peptides. *Front. Microbiol.* 9. doi: 10.3389/fmicb.2018.00323

Akbarian, M., Khani, A., Eghbalpour, S., and Uversky, V. N. (2022). Bioactive peptides: synthesis, sources, applications, and proposed mechanisms of action. *Int. J. Mol. Sci.* 23, 1445. doi: 10.3390/ijms23031445

Bagarinao, T. (1992). Sulfide as an environmental factor and toxicant: tolerance and adaptations in a aquatic organism. *Aquat. Toxicol.* 24, 21–62. doi: 10.1016/0166-445X(92)90015-F

Bibi, F., Faheem, M., Azhar, E. I., Yasir, M., Alvi, S. A., Kamal, M. A., et al. (2017). Bacteria from marine sponges: A source of new drugs. *Curr. Drug Metab.* 17, 11–15. doi: 10.2174/1389200217666161013090610

Birke, H., Haas, F. H., De Kok, L. J., Balk, J., Wirtz, M., and Hell, R. (2012). Cysteine biosynthesis, in concert with a novel mechanism, contributes to sulfide detoxification in mitochondria of Arabidopsis thaliana. *Biochem. J.* 445, 275–283. doi: 10.1042/BJ20120038

Blin, K., Shaw, S., Kloosterman, A. M., Charlop-Powers, Z., Van Wezel, G. P., Medema, M. H., et al. (2021). AntiSMASH 6.0: Improving cluster detection and comparison capabilities. *Nucleic Acids Res.* 49, 29–35. doi: 10.1093/nar/gkab335

Bornemann, T. L. V., Esser, S. P., Stach, T. L., Burg, T., and Probst, A. J. (2023). uBin: A manual refining tool for genomes from metagenomes. *Environ. Microbiol.* 25 (6), 1077–1083. doi: 10.1111/1462-2920.16351

Bowers, R. M., Kyrpides, N. C., Stepanauskas, R., Harmon-Smith, M., Doud, D., Reddy, T. B. K., et al. (2017). Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. *Nat. Biotechnol.* 35, 725–731. doi: 10.1038/nbt.3893

Bramhachari, P. V., Mutyala, S., Bhatnagar, I., and Pallela, R. (2016). "Novel insights on the symbiotic interactions of marine sponge-associated microorganisms: Marine microbial biotechnology perspective," in *Marine Sponges: Chemicobiological and Biomedical Applications*. Eds. R. Pallela and H. Ehrlich (India: Springer), 69–95. doi: 10.1007/978-81-322-2794-6_6

Brinkmann, C. M., Marker, A., and Kurtböke, D. I. (2017). An overview on marine sponge-symbiotic bacteria as unexhausted sources for natural product discovery. *Diversity* 9, 40. doi: 10.3390/d9040040

Burgsdorf, I., Sizikov, S., Squatrito, V., Britstein, M., Slaby, B. M., Cerrano, C., et al. (2022). Lineage-specific energy and carbon metabolism of sponge symbionts and contributions to the host carbon pool. *ISME J.* 16, 1163–1175. doi: 10.1038/s41396-021-01165-9

Burkhart, B. J., Hudson, G. A., Dunbar, K. L., and Mitchell, D. A. (2015). A prevalent peptide-binding domain guides ribosomal natural product biosynthesis. *Nat. Chem. Biol.* 11, 564–570. doi: 10.1038/nchembio.1856

Burt, J. A., Camp, E. F., Enochs, I. C., Johansen, J. L., Morgan, K. M., Riegl, B., et al. (2020). Insights from extreme coral reefs in a changing world. *Coral Reefs* 39, 495–507. doi: 10.1007/s00338-020-01966-y

Calegario, G., Freitas, L., Appolinario, L. R., Venas, T., Arruda, T., Otsuki, K., et al. (2021). Conserved rhodolith microbiomes across environmental gradients of the Great Amazon Reef. *Sci. Total Environ.* 760, 143411. doi: 10.1016/j.scitotenv.2020.143411

Cantarel, B. I., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V., and Henrissat, B. (2009). The Carbohydrate-Active EnZymes database (CAZy): An expert resource for glycogenomics. *Nucleic Acids Res.* 37, D233–D238. doi: 10.1093/nar/gkn663

Caspi, R., Billington, R., Keseler, I. M., Kothari, A., Krummenacker, M., Midford, P. E., et al (2020). The MetaCyc database of metabolic pathways and enzymes - a 2019 update. *Nucleic Acids Res* 48, D445–D453. doi: 10.1093/nar/gkz862

Chaumeil, P. A., Mussig, A. J., Hugenholtz, P., and Parks, D. H. (2020). GTDB-Tk: A toolkit to classify genomes with the genome taxonomy database. *Bioinformatics* 36, 1925–1927. doi: 10.1093/bioinformatics/btz848

Cheng, Z., Xia, Y., and Zhou, Z. (2020). Recent advances and promises in nitrile hydratase: from mechanism to industrial applications. *Front. Bioeng. Biotechnol.* 8. doi: 10.3389/fbioe.2020.00352

Chun, J., Oren, A., Ventosa, A., Christensen, H., Arahal, D. R., Costa, M. S., et al (2018). Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. *Int. J. Syst. Evol. Microbiol.* 68, 461–466. doi: 10.1099/ijsem.0.002516

Cipollone, R., Frangipani, E., Tiburzi, F., Imperi, F., Ascenzi, P., and Visca, P. (2007). Involvement of Pseudomonas aeruginosa rhodanese in protection from cyanide toxicity. *Appl. Environ. Microbiol.* 73, 390–398. doi: 10.1128/AEM.02143-06

Contesini, F. J., Frandsen, R. J. N., and Damasio, A. (2021). Editorial: CAZymes in biorefinery: from genes to application. *Front. Bioeng. Biotechnol.* 9. doi: 10.3389/fbioe.2021.622817

Cox-Georgian, D., Ramados, N., Dona, C., and Basu, C. (2019). "Therapeutic and Medicinal Uses of Terpenes," in Medicinal Plants. Eds N. Joshee, S. Dhekney and P. Parajuli (Switzerland: Springer, Cham), 333–359. doi: 10.1007/978-3-030-31269-5_15

Davis, J. J., Wattam, A. R., Aziz, R. K., Brettin, T., Butler, R., Butler, R. M., et al. (2020). The PATRIC Bioinformatics Resource Center: Expanding data and analysis capabilities. *Nucleic Acids Res.* 48, D606–D612. doi: 10.1093/nar/gkz943

Díez-Vives, C., Moitinho-Silva, L., Nielsen, S., Reynolds, D., and Thomas, T. (2017). Expression of eukaryotic-like protein in the microbiome of sponges. *Mol. Ecol.* 26, 1432–1451. doi: 10.1111/mec.14003

Drula, E., Garron, M. L., Dogan, S., Lombard, V., Henrissat, B., and Terrapon, N. (2022). The carbohydrate-active enzyme database: Functions and literature. *Nucleic Acids Res.* 50, D571–D577. doi: 10.1093/nar/gkab1045

Dyhrman, S. T., Ammerman, J. W., and Van Mooy, B. J. S. (2007). Microbes and the marine phosphorus cycle. *Oceanography* 20, 110–116. doi: 10.5670/oceanog.2007.54

Embree, M., Liu, J. K., Al-Bassam, M. M., and Zengler, K. (2015). Networks of energetic and metabolic interactions define dynamics in microbial communities. *Proc. Natl. Acad. Sci. U. S. A.* 112, 15450–15455. doi: 10.1073/pnas.1506034112

Engelberts, J. P., Robbins, S. J., de Goeij, J. M., Aranda, M., Bell, S. C., and Webster, N. S. (2020). Characterization of a sponge microbiome using an integrative genome-centric approach. *ISME J.* 14, 1100–1110. doi: 10.1038/s41396-020-0591-9

Fan, L., Reynolds, D., Liu, M., Stark, M., Kjelleberg, S., Webster, N. S., et al. (2012). Functional equivalence and evolutionary convergence in complex communities of microbial sponge symbionts. *Proc. Natl. Acad. Sci. U. S. A.* 109, E1878–E1887. doi: 10.1073/pnas.1203287109

Feng, G., and Li, Z. (2019). "Carbon and nitrogen metabolism of sponge microbiome," in *Symbiotic Microbiomes of Coral Reefs Sponges and Corals*. Ed. Z. Li (Netherlands: Springer), 145–169. doi: 10.1007/978-94-024-1612-1_9

Florea, L., Souvorov, A., Kalbfleisch, T. S., and Salzberg, S. L. (2011). Genome assembly has a major impact on gene content: A comparison of annotation in two bos Taurus assemblies. *PloS One* 6, e0021400. doi: 10.1371/journal.pone.0021400

Francini-Filho, R. B., Asp, N. E., Siegle, E., Hocevar, J., Lowyck, K., D'Avila, N., et al. (2018). Perspectives on the Great Amazon Reef: Extension, biodiversity, and threats. *Front. Mar. Sci.* 5. doi: 10.3389/fmars.2018.00142

Frank, A. C. (2019). Molecular host mimicry and manipulation in bacterial symbionts. *FEMS Microbiol. Lett.* 366, 4. doi: 10.1093/femsle/fnz038

Genin, E., Wielgosz-Collin, G., Njinkoué, J. M., Velosaotsy, N. E., Kornprobst, J. M., Gouygou,, et al. (2008). New trends in phospholipid class composition of marine sponges. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* 150, 427–431. doi: 10.1016/j.cbpb.2008.04.012

Gupta, S., Kapoor, P., Chaudhary, K., Gautam, A., Kumar, R., and Raghava, G. P. S. (2013). In silico approach for predicting toxicity of peptides and proteins. *PloS One* 8, e0073957. doi: 10.1371/journal.pone.0073957

Hentschel, U., Piel, J., Degnan, S. M., and Taylor, M. W. (2012). Genomic insights into the marine sponge microbiome. *Nat. Rev. Microbiol.* 10, 641–654. doi: 10.1038/nrmicro2839

Horn, H., Slaby, B. M., Jahn, M. T., Bayer, K., Moitinho-Silva, L., Förster, F., et al. (2016). An Enrichment of CRISPR and other defense-related features in marine sponge-associated microbial metagenomes. *Front. Microbiol.* 7. doi: 10.3389/fmicb.2016.01751

IPEN (2018) *Ocean Pollutants Guide: Toxic Threats To Human Health And Marine Life, Technical report*. Available at: http://www.ipen.org.

Jensen, S., Fortunato, S. A. V., Hoffmann, F., Hoem, S., Rapp, H. T., Øvreås, L., et al. (2017). The relative abundance and transcriptional activity of marine sponge-associated microorganisms emphasizing groups involved in sulfur cycle. *Microb. Ecol.* 73, 668–676. doi: 10.1007/s00248-016-0836-3

Jeske, L., Placzek, S., Schomburg, I., Chang, A., and Schomburg, D. (2019). BRENDA in 2019: A European ELIXIR core data resource. *Nucleic Acids Res.* 47, D542–D549. doi: 10.1093/nar/gky1048

Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y., and Morishima, K. (2017). KEGG: New perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* 45, D353–D361. doi: 10.1093/nar/gkw1092

Kanehisa, M., and Sato, Y. (2020). KEGG Mapper for inferring cellular functions from protein sequences. *Protein Sci.* 29, 28–35. doi: 10.1002/pro.3711

Kanehisa, M., Sato, Y., and Morishima, K. (2016). BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *J. Mol. Biol.* 428, 726–731. doi: 10.1016/j.jmb.2015.11.006

Karigar, C. S., and Rao, S. S. (2011). Role of microbial enzymes in the bioremediation of pollutants: A review. *Enzyme Res.* 2011, 805187. doi: 10.4061/2011/805187

Karp, P. D., Midford, P. E., Billington, R., Kothari, A., Krummenacker, M., Latendresse, M., et al. (2021). Pathway Tools version 23.0 update: Software for pathway/genome informatics and systems biology. *Brief. Bioinform.* 22, 109–126. doi: 10.1093/bib/bbz104

KIran, G. S., Sekar, S., Ramasamy, P., Thinesh, T., Hassan, S., Lipton, A. N., et al. (2018). Marine sponge microbial association: Towards disclosing unique symbiotic interactions. *Mar. Environ. Res.* 140, 169–179. doi: 10.1016/j.marenvres.2018.04.017

Lannes, R., Olsson-Francis, K., Lopez, P., and Bapteste, E. (2019). Carbon fixation bymarine ultrasmall prokaryotes. *Genome Biol. Evol.* 11, 1166–1177. doi: 10.1093/gbe/evz050

Lata, S., Mishra, N. K., and Raghava, G. P. S. (2010). AntiBP2: Improved version of antibacterial peptide prediction. *BMC Bioinform.* 11, S19. doi: 10.1186/1471-2105-11-S1-S19

Lavy, A., Keren, R., Yu, K., Thomas, B. C., Alvarez-Cohen, L., Banfield, J. F., et al. (2018). A novel Chromatiales bacterium is a potential sulfide oxidizer in multiple orders of marine sponges. *Environ. Microbiol.* 20, 800–814. doi: 10.1111/1462-2920.14013

Macedo, M. W. F. S., Cunha, N. B., Carneiro, J. A., Costa, R. A., Alencar, S. A., Cardoso, M. H., et al. (2021). Marine organisms as a rich source of biologically active peptides. *Front. Mar. Sci.* 8. doi: 10.3389/fmars.2021.667764

Mahiques, M. M., Siegle, E., Francini-Filho, R. B., Thompson, F. L., Rezende, C. E., Gomes, J. D., et al. (2019). Insights on the evolution of the living Great Amazon Reef System, equatorial West Atlantic. *Sci. Rep.* 9, 13699. doi: 10.1038/s41598-019-50245-6

Manchanda, N., Portwood, J. L. II, Woodhouse, M. R., Seetharam, A. S., Lawrence-Dill, C. J., Andorf, C. M., et al. (2020). GenomeQC: a quality assessment tool for genome assemblies and gene structure annotations. *BMC Genomics* 21, 1–9. doi: 10.1186/s12864-020-6568-2

Marchler-Bauer, A., and Bryant, S. H. (2004). CD-Search: Protein domain annotations on the fly. *Nucleic Acids Res.* 32, W327–W331. doi: 10.1093/nar/gkh454

Meier-Kolthoff, J. P., Auch, A. F., Klenk, H., and Goker, M. (2013). Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinform.* 14, 60. doi: 10.1186/1471-2105-14-60

Menezes, T. A., Freitas, M. A. M., Lima, M. S., Soares, A. C., Leal, C., Busch, M. S., et al. (2022). Fluxes of the Amazon River plume nutrients and microbes into marine sponges. *Sci. Total Environ.* 847, 157474. doi: 10.1016/j.scitotenv.2022.157474

Moura, R. L., Amado-Filho, G. M., Moraes, F. C., Brasileiro, P. S., Salomon, P. S., Mahiques, M. M., et al. (2016). An extensive reef system at the Amazon River mouth. *Sci. Adv.* 2, e1501252. doi: 10.1126/sciadv.1501252

Murzyn, K., Róg, T., and Pasenkiewicz-Gierula, M. (2005). Phosphatidylethanolamine-phosphatidylglycerol bilayer as a model of the inner bacterial membrane. *Biophys. J.* 88, 1091–1103. doi: 10.1529/biophysj.104.048835

Nandan, A., and Nampoothiri, K. M. (2020). Therapeutic and biotechnological applications of substrate specific microbial aminopeptidases. *Appl. Microbiol. Biotechnol.* 104, 5243–5257. doi: 10.1007/s00253-020-10641-9

Nurk, S., Meleshko, D., Korobeynikov, A., and Pevzner, P. A. (2017). MetaSPAdes: A new versatile metagenomic assembler. *Genome Res.* 27, 824–834. doi: 10.1101/gr.213959.116

Oliveira, B. F. R., Carr, C. M., Dobson, A. D. W., and Laport, M. S. (2020). Harnessing the sponge microbiome for industrial biocatalysts. *Appl. Microbiol. Biotechnol.* 104, 8131–8154. doi: 10.1007/s00253-020-10817-3

Olm, M. R., Brown, C. T., Brooks, B., and Banfield, J. F. (2017). DRep: A tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. *ISME J.* 11, 2864–2868. doi: 10.1038/ismej.2017.126

Osinga, R., Armstrong, E., Burgess, J. B., Hoffmann, F., Reitner, J., and Schumann-Kindel, G. (2001). Sponge-microbe associations and their importance for sponge bioprocess engineering. *Hydrobiologia* 461, 55–62. doi: 10.1023/A:1012717200362

Oyewusi, H. A., Wahab, R. A., and Huyop, F. (2020). Dehalogenase-producing halophiles and their potential role in bioremediation. *Mar. pollut. Bull.* 160, 111603. doi: 10.1016/j.marpolbul.2020.111603

Pajares, S., and Ramos, R. (2019). Processes and microorganisms involved in the marine nitrogen cycle: knowledge and gaps. *Front. Mar. Sci.* 6. doi: 10.3389/fmars.2019.00739

Pallela, R., and Kim, K. (2011). Ecobiotechnology of marine sponges and their symbionts-review and present status. *J. Mar. Biosci. Biotechnol.* 5, 15–25. doi: 10.15433/KSMB.2011.5.4.015

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

Pimentel-Elardo, S. M., Grozdanov, L., Proksch, S., and Hentschel, U. (2012). Diversity of nonribosomal peptide synthetase genes in the microbial metagenomes of marine sponges. *Mar. Drugs* 10, 1192–1202. doi: 10.3390/md10061192

Pineda, M. C., Strehlow, B., Sternel, M., Duckworth, A., Jones, R., and Webster, N. S. (2017). Effects of suspended sediments on the sponge holobiont with implications for dredging management. *Sci. Rep.* 7, 4925. doi: 10.1038/s41598-017-05241-z

Pinto, O. H. B., Bornemann, T. L. V., Oliveira, R. S., Frederico, T. D., Quirino, B. F., Probst, A. J., et al. (2022). Plume layer influences the amazon reef sponge microbiome primary producers. *Front. Mar. Sci.* 9. doi: 10.3389/fmars.2022.867234

Pita, L., Rix, L., Slaby, B. M., Franke, A., and Hentschel, U. (2018). The sponge holobiont in a changing ocean: from microbes to ecosystems. *Microbiome* 6, 46. doi: 10.1186/s40168-018-0428-1

Podell, S., Blanton, J. M., Oliver, A., Schorn, M. A., Agarwal, V., Biggs, J. S., et al. (2020). A genomic view of trophic and metabolic diversity in clade-specific Lamellodysidea sponge microbiomes. *Microbiome* 8, 97. doi: 10.1186/s40168-020-00877-y

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

Ramaprasad, A. S. E., Singh, S., Gajendra, R. P. S., and Venkatesan, S. (2015). AntiAngioPred: A server for prediction of anti-angiogenic peptides. *PloS One* 10, e0136990. doi: 10.1371/journal.pone.0136990

Raveendran, S., Parameswaran, B., Ummalyma, S. B., Abraham, A., Mathew, A. K., Madhavan, A., et al. (2018). Applications of microbial enzymes in food industry. *Food Technol. Biotechnol.* 56, 16–30. doi: 10.17113/ftb.56.01.18.5491

Regoli, F., Cerrano, C., Chierici, E., Bompadre, S., and Bavestrello, G. (2000). Susceptibility to oxidative stress of the Mediterranean demosponge Petrosia ficiformis: role of endosymbionts and solar irradiance. *Mar. Biol.* 137, 453–461. doi: 10.1007/s002270000369

Reynolds, D., and Thomas, T. (2016). Evolution and function of eukaryotic-like proteins from sponge symbionts. *Mol. Ecol.* 25, 5242–5253. doi: 10.1111/mec.13812

Richter, M., Rosselló-Móra, R., Glöckner, F. O., and Peplies, J. (2016). JSpeciesWS: A web server for prokaryotic species circumscription based on pairwise genome comparison. *Bioinformatics* 32, 929–931. doi: 10.1093/bioinformatics/btv681

Robb, F. T., and Techtmann, S. M. (2018). Life on the fringe: Microbial adaptation to growth on carbon monoxide. *F1000Research* 7, 1981. doi: 10.12688/f1000research.16059.1

Robbins, S. J., Song, W., Engelberts, J. P., Glasl, B., Slaby, B. M., Boyd, J., et al. (2021). A genomic view of the microbiome of coral reef demosponges. *ISME J.* 15, 1641–1654. doi: 10.1038/s41564-019-0532-4

Roberts, M. F. (2006). "Inositol in bacteria and archaea," in *Biology of Inositols and Phosphoinositides*. Eds. A. L. Majumder and B. B. Biswas (New York, NY: Springer), 103–133. doi: 10.1007/0-387-27600-9

Santos-Gandelman, J., Giambiagi-deMarval, M., Oelemann, W., and Laport, M. (2014). Biotechnological potential of sponge-associated bacteria. *Curr. Pharm. Biotechnol.* 15, 143–155. doi: 10.2174/1389201015666140711115033

Schempp, F. M., Drummond, L., Buchhaupt, M., and Schrader, J. (2018). Microbial cell factories for the production of terpenoid flavor and fragrance compounds. *J. Agric. Food Chem.* 66, 2247–2258. doi: 10.1021/acs.jafc.7b00473

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

Selvin, J., Ninawe, A. S., KIran, G. S., and Lipton, A. P. (2010). Sponge-microbial interactions: Ecological implications and bioprospecting avenues. *Crit. Rev. Microbiol.* 36, 82–90. doi: 10.3109/10408410903397340

Sieber, C. M. K., Probst, A. J., Sharrar, A., Thomas, B. C., Hess, M., Tringe, S. G., et al. (2018). Recovery of genomes from metagenomes via a dereplication, aggregation and scoring strategy. *Nat. Microbiol.* 3, 836–843. doi: 10.1038/s41564-018-0171-1

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

Taylor, M. W., Radax, R., Steger, D., and Wagner, M. (2007). Sponge-associated microorganisms: evolution, ecology, and biotechnological potential. *Microbiol. Mol. Biol. Rev.* 71, 295–347. doi: 10.1128/mmbr.00040-06

Thacker, R. W., and Freeman, C. J. (2012). Sponge-microbe symbioses. Recent advances and new directions. *Adv. Mar. Biol.* 62, 57–111. doi: 10.1016/B978-0-12-394283-8.00002-3

Thomas, T., Rusch, D., DeMaere, M. Z., Yung, P. Y., Lewis, M., Halpern, A., et al. (2010). Functional genomic signatures of sponge bacteria reveal unique and shared features of symbiosis. *ISME J.* 4, 1557–1567. doi: 10.1038/ismej.2010.74

Törönen, P., and Holm, L. (2022). PANNZER - A practical tool for protein function prediction. *Protein Sci.* 31 (1), 118–128. doi: 10.1002/pro.4193

Tran, P. Q., Bachand, S. C., McIntyre, P. B., Kraemer, B. M., Vadeboncoeur, Y., Kimirei, I. A., et al. (2021). Depth-discrete metagenomics reveals the roles of microbes in biogeochemical cycling in the tropical freshwater Lake Tanganyika. *ISME J.* 15, 1971–1986. doi: 10.1038/s41396-021-00898-x

Trindade-Silva, A. E., Rua, C. P. J., Andrade, B. G. N., Vicente, A. C. P., Silva, G. G. Z., Berlinck, R. G. S., et al. (2013). Polyketide synthase gene diversity within the microbiome of the sponge Arenosclera brasiliensis, endemic to the southern Atlantic Ocean. *Appl. Environ. Microbiol.* 79, 1598–1605. doi: 10.1128/AEM.03354-12

Trindade-Silva, A. E., Rua, C., Silva, G. G. Z., Dutilh, B. E., Moreira, A. P. B., Edwards, R. A., et al. (2012). Taxonomic and functional microbial signatures of the endemic marine sponge arenosclera brasiliensis. *PloS One* 7, e39905. doi: 10.1371/journal.pone.0039905

Tsoi, R., Wu, F., Zhang, C., Bewick, S., Karig, D., and You, L. (2018). Metabolic division of labor in microbial systems. *Proc. Natl. Acad. Sci. U. S. A.* 115, 2526–2531. doi: 10.1073/pnas.1716888115

Vagi, M. C., Petsas, A. S., and Kostopoulou, M. N. (2021). Potential effects of persistent organic contaminants on marine biota: A review on recent research. *Water* 13, 2488. doi: 10.3390/w13182488

Vavourakis, C. D., Andrei, A. S., Mehrshad, M., Ghai, R., Sorokin, D. Y., and Muyzer, G. (2018). A metagenomics roadmap to the uncultured genome diversity in hypersaline soda lake sediments. *Microbiome* 6, 168. doi: 10.1186/s40168-018-0548-7

Villarreal-Chiu, J. F., Quinn, J. P., and McGrath, J. W. (2012). The genes and enzymes of phosphonate metabolism by bacteria, and their distribution in the marine environment. *Front. Microbiol.* 3. doi: 10.3389/fmicb.2012.00019

Villesen, P. (2007). FaBox: An online toolbox for FASTA sequences. *Mol. Ecol. Notes* 7, 965–968. doi: 10.1111/j.1471-8286.2007.01821.x

Waghu, F. H., Barai, R. S., Gurung, P., and Idicula-Thomas, S. (2016). CAMPR3: A database on sequences, structures and signatures of antimicrobial peptides. *Nucleic Acids Res.* 44, D1094–D1097. doi: 10.1093/nar/gkv1051

Webster, N. S., and Taylor, M. W. (2012). Marine sponges and their microbial symbionts: Love and other relationships. *Environ. Microbiol.* 14, 335–346. doi: 10.1111/j.1462-2920.2011.02460.x

Wu, Y. W., Simmons, B. A., and Singer, S. W. (2016). MaxBin 2.0: An automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics* 32, 605–607. doi: 10.1093/bioinformatics/btv638

You, R., Wang, L., Shi, C., Chen, H., Zhang, S., Hu, M., et al. (2020). Efficient production of myo-inositol in Escherichia coli through metabolic engineering. *Microb. Cell Fact.* 19, 109. doi: 10.1186/s12934-020-01366-5

Zhang, F., Blasiak, L. C., Karolin, J. O., Powell, R. J., Geddes, C. D., Hill, R. T., et al. (2015). Phosphorus sequestration in the form of polyphosphate by microbial symbionts in marine sponges. *Proc. Natl. Acad. Sci. U. S. A.* 112, 4381–4386. doi: 10.1073/pnas.1423768112

Zhang, H., Yohe, T., Huang, L., Entwistle, S., Wu, P., Yang, Z., et al. (2018). DbCAN2: A meta server for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res.* 46, W95–W101. doi: 10.1093/nar/gky418

Zweifler, A., O'Leary, M., Morgan, K., and Browne, N. K. (2021). Turbid coral reefs: Past, present and future—a review. *Diversity* 13, 251. doi: 10.3390/d13060251

frontiers | Frontiers in Microbiomes

TYPE Original Research
PUBLISHED 13 November 2023
DOI 10.3389/frmbi.2023.1232571

OPEN ACCESS

EDITED BY
Lucie Malard,
Université de Lausanne, Switzerland

REVIEWED BY
Dandan Izabel-Shen,
Helmholtz Institute for Functional Marine Biodiversity (HIFMB), Germany
Angel Rain-Franco,
University of Zurich, Switzerland

*CORRESPONDENCE
Daniel P. R. Herlemann
daniel.herlemann@io-warnemuende.de

RECEIVED 31 May 2023
ACCEPTED 11 October 2023
PUBLISHED 13 November 2023

CITATION
Tammert H, Kivistik C, Kisand V, Käiro K and Herlemann DPR (2023) Resistance of freshwater sediment bacterial communities to salinity disturbance and the implication for industrial salt discharge and climate change-based salinization.
Front. Microbiomes 2:1232571.
doi: 10.3389/frmbi.2023.1232571

Resistance of freshwater sediment bacterial communities to salinity disturbance and the implication for industrial salt discharge and climate change-based salinization

Helen Tammert[1,2], Carmen Kivistik[1], Veljo Kisand[1,2], Kairi Käiro[1] and Daniel P. R. Herlemann[1,3]*

[1]Centre for Limnology, Chair of Hydrobiology and Fisheries, Estonian University of Life Sciences, Tartu, Estonia, [2]Institute of Technology, University of Tartu, Tartu, Estonia, [3]Department of Biological Oceanography, Leibniz Institute for Baltic Sea Research Warnemünde, Rostock, Germany

The impact of salinization on freshwater ecosystems became apparent during the 2022 ecological disaster in the Oder River, located in Poland and Germany, which was caused by salt discharge from mining activities. How bacterial communities respond to salinization caused by industrial salt discharge, or climate change-driven events, depends on the sensitivity of these complex bacterial communities. To investigate the sensitivity of bacterial communities to pulse salinization, we performed an experiment in the salinity range from 0.2 to 6.0. In addition, we sampled similar salinities in the littoral zone of the Baltic Sea where the bacterial communities are permanently exposed to the aforementioned salinities. To simulate a major disturbance, we included an ampicillin/streptomycin treatment in the experiment. Although the addition of antibiotics and increase in salinity had a significant impact on the water bacterial richness and community composition, only antibiotics affected the sediment bacterial community in the experiment. In contrast, sediment bacterial communities from the Baltic Sea littoral zone clustered according to salinity. Hence, sediment bacterial communities are more resistant to pulse changes in salinity than water bacteria but are able to adapt to a permanent change without loss in species richness. Our results indicate that moderate pulse salinization events such as industrial salt discharge or heavy storms will cause changes in the water bacterial communities with unknown consequences for ecosystem functioning. Sediment bacterial communities, however, will probably be unaffected in their ecosystem functions depending on the disturbance strength. Long-term disturbances, such as sea level rise or constant salt discharge, will cause permanent changes in the sediment bacterial community composition.

KEYWORDS

Baltic Sea, sea level rise, littoral, salinization, pulse disturbance, experiment, industrial discharge management

Introduction

Climate change is responsible for weather extremes that can cause salinity pulse shocks to freshwater environments during storm events or long-term salinization due to sea level rise (IPCC, 2022). Moreover, changes in evaporation and precipitation result in the salinization of freshwater in inland waters (Jeppesen et al., 2020). Another source of salinization is the often neglected industrial pollution caused by mining activities, which are the main sources of river salinization in Europe (Bäthe and Coring, 2011). The discharge of hypersaline effluents into freshwater ecosystems can increase the salinity levels in a matter of a few hours or days to brackish conditions (Johnson et al., 2010), which in turn affects the ecosystem structure and functioning (Sauer et al., 2016). Hypersaline discharge was also the key factor for the 2022 ecological disaster of the Oder River, located in Poland and Germany, which caused the death of hundreds of tons of fish along a 500-km stretch of the river due to the sudden growth of toxic brackish water algae (Free et al., 2023). The increase in salinity causes extreme conditions for freshwater bacterial communities, as microorganisms face particular barriers when transitioning to another salinity level (Logares et al., 2009). The salt concentration of the environment has been reported to be one of the main factors shaping the distribution of aquatic bacterial communities (Crump et al., 2004; Herlemann et al., 2011; Fortunato et al., 2012). Despite the sensitivity of water bacterial community composition to changes in salinity, a constant species richness has been found in different salinities for water bacteria (Herlemann et al., 2011) and sediment bacteria in the Baltic Sea (Klier et al., 2018). This is in contrast to observations of eukaryotic richness where lower numbers of organisms in intermediate salinities (i.e., 6–10) for eukaryotes were found (Remane, 1934; Olli et al., 2023).

One of the fundamental goals of microbial ecology is to understand to what extent environmental disturbance is accompanied by changes in bacterial richness and community diversity. Disturbances that lead to changes in the community composition (beta diversity) and richness (alpha diversity) may alter the functioning of the community and affect the ecosystem processes (Berga et al., 2017). The recovery of the community after compositional reduction or change depends on the strength and the duration of the disturbance, as well as the dispersal of the taxa and recruitment from the existing seedbank. In conclusion, the reaction of the community members to the disturbances is a combination of past and present events (Hillebrand and Blenckner, 2002; Andersson et al., 2014; Renes et al., 2020; Philippot et al., 2021). In the littoral environment, microbial communities both in the water and in the sediment upper layer are exposed to unstable and inhomogeneous sets of biological and physical conditions that cause a complex mixture of growth-influencing factors. The diverse range of ecological niches has created seedbanks for dynamic and flexible communities, with the potential to inhabit a high number of rare species (Shade et al., 2012). The response of community members to disturbance depends on the frequency and intensity of the disturbance (Berga et al., 2012). Small-scale and possibly recurring changes in the environment, such as inter-seasonal chemical–biological fluctuations, are expected to cause minor modulations in the bacterial diversity and the presence of niche specialists favored by certain abiotic or biotic factors (Andersson et al., 2014). Strong pulse disturbances (e.g., industrial salt discharge or extreme weather events) and continuous long-term pressures (e.g., sea level rise, evaporation, precipitation, and industrial salt leakage) can lead to more severe disturbances in the bacterial community. In general, bacterial communities have shown to be sensitive to disturbance and are usually unable to recover to their original state (Renes et al., 2020). The flexibility of bacterial physiology and the energetic cost of adaptation mechanisms to withstanding disturbance can determine the resilience and resistance level of bacterial community members. For aquatic environments, the vast majority of the disturbance studies have shown sensitivity to temperature, salinity, and acidification disturbance of bacterial communities (Andersson et al., 2014; Renes et al., 2020; Seidel et al., 2023).

In this study, we investigate the responses of littoral freshwater sediment and water bacterial communities to a pulse increase in salinity in manipulation experiments. The results are compared with the bacterial community composition in natural brackish sites with similar salinities. Our hypotheses were that an increase in salinity (i) significantly alters the bacterial community composition and brackish-tolerant bacteria dominate in the water and sediment, and (ii) does not affect the bacterial richness of freshwater sediment and water bacterial communities.

Materials and methods

The top 0- to 2-cm layer of sandy sediment, at a water depth of 0.5 m, and corresponding water were collected for the freshwater salinization experiment from the littoral zone of Lake Võrtsjärv, Estonia. The sediment was sieved through 0.5-mm mesh filter to remove larger debris, homogenized, and then distributed to 10-L aquaria in a 3- to 4-cm-thick layer. To the sediment, freshly 85-µm mesh-filtered Lake Võrtsjärv water was added. The aquaria also contained snails (*Ampullaceana balthica*), pebbles, and small stones with biofilm. The snails and their microbiome were analyzed in the study of Kivistik et al. (2022). In total, 13 × 10 L aquaria were divided into four different treatments: three reference aquaria without further manipulation (REF); three aquaria with an addition of 5 mg/L of ampicillin and 5 mg/L of streptomycin (AB); three aquaria where the salinity of the water was increased to 3 (SAL3, oligohaline); and four aquaria where the salinity of the water was increased to 6 (SAL6, mesohaline). For salinization, we used commercially available Reef Salt (AquaMedic) containing 1,000 mg/L sodium, 1,200 mg/L magnesium, 420 mg/L calcium, 350 mg/L potassium, 19,700 mg/L chloride, 2,200 mg/L sulfate, 180 mg/L carbonate, and 16 mg/L strontium. The experimental aquaria were constantly supplied by air and held at 16.1°C –17.5°C for 8 days. The water and sediment samples were taken on days 1, 3, 6, and 8 of the experiment. For water samples, 100 mL of water from each aquarium, for a total of 55 samples, was filtered through 0.2-µm membrane filters (Whatman) and frozen at –80°C. In addition, 3 g of sediment was sampled from each aquarium, for a total of 53 samples, shock frozen in liquid nitrogen, and stored at –80°C. The

temperature, oxygen, pH, and salinity were measured every day with a YSI ProDSS multiparameter meter.

The natural site water and sediment samples were collected in the Baltic Sea coastal area in Estonia on 17 and 18 June 2019. The triplicate water and surface sediment samples were collected at the natural sites from a 0.5- to 1-m depth ($n = 42$) at similar salinities as in the experiment. The *in situ* freshwater sampling sites were Selja Pond (FP), and the freshwater (FW) rivers Selja River (SR) and Kunda River (KR). The sampling site with a salinity of 3 (oligohaline; SAL3) was Selja Bay (SB). The sampling sites with a salinity of 6 (mesohaline; SAL6) were Nõva (NÕ), Ristna West (RW), and Ristna East (RE) (Table 1).

DNA extraction and sequence processing

The DNA from the experimental water samples and the DNA and RNA of the sediment and natural site water samples were extracted according to the modified protocols from Lueders et al. (2004) and Weinbauer et al. (2002). For the water filters dichlorodimethylsilane-treated glass beads (three beads with a diameter of 3 mm, and 0.5 g of beads with diameter 0.5 mm) were added to 2-mL tubes. Cell lysis was performed in 750 µL of 120 mM $NaPO_4$ buffer (pH 8) and 250 µL of TNS solution [500 mM Tris-HCl pH 8.0, 100 mM NaCl, and 10% sodium dodecyl sulfate (SDS) (weight to volume (wt/vol)], and the samples were bead beaten for 3 min at 2,000 rpm (revolutions per minute) using a Mikro-dismembrator U (B. Braun Biotech International, Melsungen, Germany). After a 1-h incubation at 65°C, we applied a second bead beating for 3 min at 2,000 rpm, which was followed by centrifugation at 14,000 rpm for 5 min. The supernatant was transferred to a new 2-mL tube and a mixture of phenol : chloroform : isoamyl alcohol (25 : 24 : 1) at pH 8 was then added and carefully mixed. The phases were separated by centrifugation at 14,000 rpm for 5 min and the upper aqueous phase was placed in a new 2-mL tube. For purification of both DNA and RNA samples, 1 volume of chloroform : isoamyl (24 : 1) was added and mixed carefully. After centrifugation at 14,000 rpm for 12 min, nucleic acids in the upper aqueous phase of sediment samples were divided equally between two new 1.5-mL tubes: one for DNA and another for RNA. For the removal of the RNA from the DNA sample, 2 µL of RNase (100 mg/mL; QIAGEN, Venlo, the Netherlands) was added and incubated at 37°C for 30 min. The nucleic acids were precipitated by incubating at room temperature for 15 min with a 0.7 volume of cold isopropanol. The samples were then centrifuged at 14,000 rpm for 20 min. The resulting pellet was washed with 250 µL of 95% ethanol, centrifuged at 14,000 rpm for 5 min, and then dried at 50°C after ethanol removal (≈ 5–15 min). The pellet was resuspended in a 50 µL of elution buffer (10 mM Tris-HCl, 0.5 mM EDTA, pH 9.0; QIAGEN).

To remove the DNA from the RNA samples, DNase treatment was performed by using the TURBO DNA-free™ Kit (Invitrogen, Thermo Fisher Scientific, Waltham MA, USA), in accordance with the manufacturer's protocol. Both the amount and quality of the nucleic acids were measured with a NanoDrop™ UV–Vis spectrophotometer (Thermo Fisher Scientific, Waltham MA, USA). The iScript™ Select cDNA synthesis kit (Bio-Rad, Hercules, CA, USA) was used to transcribe the RNA of the sediment samples into cDNA, in accordance with the manufacturer's protocol.

For the bacterial community analysis, a two-step PCR protocol was used. In the first step, DNA and cDNA were PCR amplified using the primers Bakt_341F and Bakt_805R (Herlemann et al., 2011), as described by Kivistik et al. (2020). In the second step, sample-specific tags were added to the first step's PCR products. The amplicons were purified with PCR Kleen™ (Bio-Rad, Hercules, CA, USA). The sequencing was conducted, at FIMM at the University of Helsinki, Finland, with Illumina sequencing using PE250 chemistry.

The sequences were quality checked using Trimmomatic (V0.36) (Bolger et al., 2014) to remove any Illumina-specific sequences and regions with low-sequence quality (i.e., with an average quality score < Q20). The PCR primer sequences were removed using the default values in Cutadapt (V2.3) (Martin, 2011). The reads were paired (16-bp overlap, with a minimum length of 300 bp) and quality trimmed using the VSEARCH tool (Rognes et al., 2016). The sequences were then taxonomically assigned using the SILVA next-generation sequencing (NGS) pipeline (Glöckner et al., 2017), which was based on the SILVA

TABLE 1 Environmental variables in natural sampling sites.

Sampling site	Salinity	Salinity class	Temperature (°C)	DO (mg/L)	pH	Latitude (°) Longitude (°)
Kunda River (KR)	0.3	Freshwater (FW)	17.6	10.23	8.55	59.546306 N 26.650722 E
Selja River (SR)	0.3	Freshwater (FW)	19.9	10.85	8.72	59.555118 N 26.343352 E
Selja Pond (FP)	0.3	Freshwater (FW)	23.8	5.8	8.11	59.509762 N 26.538798 E
Selja Bay, Baltic Sea (SB)	2.5	Oligohaline (SAL3)	19	12.45	8.93	59.554670 N 26.340361 E
Nõva site, Baltic Sea (NÕ)	6.3	Mesohaline (SAL6)	16.5	7.46	7.71	59.554553 N 26.339870 E
Ristna East, Baltic Sea (RE)	6.5	Mesohaline (SAL6)	20.7	14.05	8.87	59.270542 N 23.738648 E
Ristna West, Baltic Sea (RW)	6.5	Mesohaline (SAL6)	19.9	7.62	8.03	59.271705 N 23.734715 E

release version 138.1 (released 2020). SILVA NGS was used to perform additional quality checks according to SINA-based alignments (Pruesse et al., 2012) with a curated seed database in which the PCR artifacts or non-small subunit (SSU) reads are excluded. The longest read served as a reference for taxonomic classification using a BLAST (version 2.2.30+) search against the SILVA SSURef dataset. The classification of the reference sequence of each cluster (98% sequence identity) was then mapped to all members of the respective cluster and to their replicates. Samples with less than 1,000 reads were removed. Non/bacterial sequences, such as chloroplastic, mitochondrial, eukaryotic, and archaean, were excluded because the primer set employed in the analysis had only a very limited coverage of these groups. This resulted in 7,422,343 sequences for 197 samples. The raw reads were deposited at the National Center for Biotechnology Information (NCBI)'s Sequence Read Archive (SRA) under bioproject number PRJNA724976, accession number SAMN18865857-SAMN18866052.

Statistical analysis

For alpha diversity measures, the data were rarefied with bootstrapping using Explicet (Robertson et al., 2013) and expressed as Chao1 and Shannon index values. Sampling events with at least three parallel samples were tested by analysis of variance (ANOVA) (natural sites) and repeated measures ANOVA (experiment) combined with *post-hoc* Tukey's pairwise test to calculate the significant differences between the taxonomic richness using the PAST software package (version 4.12; Hammer et al., 2001). For beta diversity, the data were cumulative sum scaling (CSS) normalized by the R package metagenomeSeq (Paulson et al., 2013). The differences in the bacterial community composition were tested by multivariate permutational analysis of variance (PERMANOVA) in natural samples and by repeated measures PERMANOVA in the experiment. ANCOVA was used to compare the slopes in the ordination axis–salinity plots using PAST. Community composition was visualized by principal coordinate analysis (PCoA). PERMANOVA and PCoA were both based on the Bray–Curtis dissimilarity, as implemented in the PAST software package version 4.12 (Hammer et al., 2001) using read abundances normalized with CSS. The Molbiol Tools' online Multiple List Comparator (https://molbiotools.com/) was used to analyze the similarity of the bacterial communities (Jaccard index) and for visualization of similarity as Venn diagrams.

Results

Bacterial richness

The sediment bacterial Chao1 richness estimate showed insignificant changes in the reference aquaria (REF: Chao1 1184, SE ± 40.0) and salinity 3 (SAL3: Chao1 1142, SE ± 63.7) treatment during the 8 days of the experiment (Figure 1A). The salinity 6 (SAL6: Chao1 1156, SE ± 61.6) and antibiotic treatment (AB: Chao1 979, SE ± 76.4) reduced the sediment bacterial richness in the DNA-based analysis (Figures 1A, 2A) after day 6 of the experiment (Tukey's test, $p < 0.05$). One of the salinity treatments was significantly different from the reference (Figure 2A). A similar pattern was observed for the cDNA-based analysis, but too few samples were available for statistical analysis. In contrast, the bacterial richness in the water samples showed an intense dynamic by decreasing for day 6 but recovering by the end of the experiment (Figure 1B). In the salinity 6 treatment, water bacterial Chao1 richness estimates decreased significantly (Tukey's test, $p < 0.05$) during the experiment from 935 (SE ± 49.1) to 387 (SE ± 63.0) and recovered at day 8 to a lower level (Chao1 705, SE ± 51.7), compared with the beginning of the experiment (Tukey's test, $p < 0.05$). In the antibiotic treatment, a similar pattern to the water bacterial community was observed: the Chao1 richness estimate decreased significantly from 1,144 (SE ± 98.4) on day 1 to 294 (SE ± 15.0) on day 6 (Tukey's test, $p < 0.05$), and recovered for day 8. The Shannon index values showed a similar pattern to the Chao1 richness estimates for water and sediment bacterial communities (Supplementary Figures 1A, B).

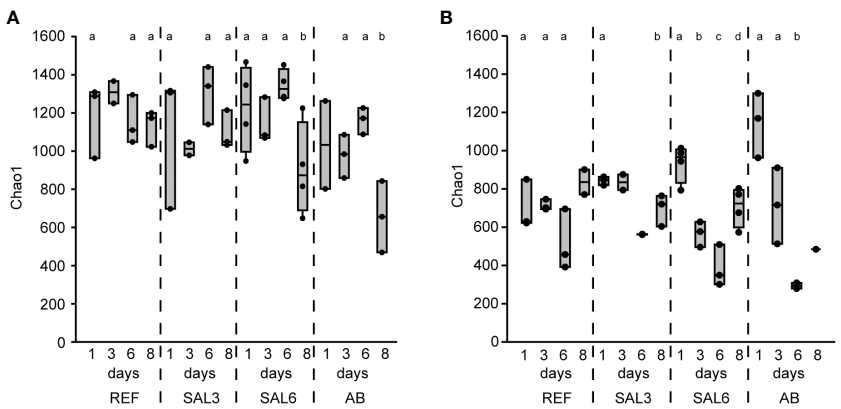


FIGURE 1
The bacterial richness based on DNA represented by the Chao1 richness estimates of the reference and treatments during the time course of the experiment. **(A)** Sediment; and **(B)** water. REF, reference aquaria; SAL3, salinity increased to 3; SAL6, salinity increased to 6; AB, antibiotic treatment. The non-capital letters (a and b) above the box plots indicate statistical significance within one group, which are separated by the dashed lines.

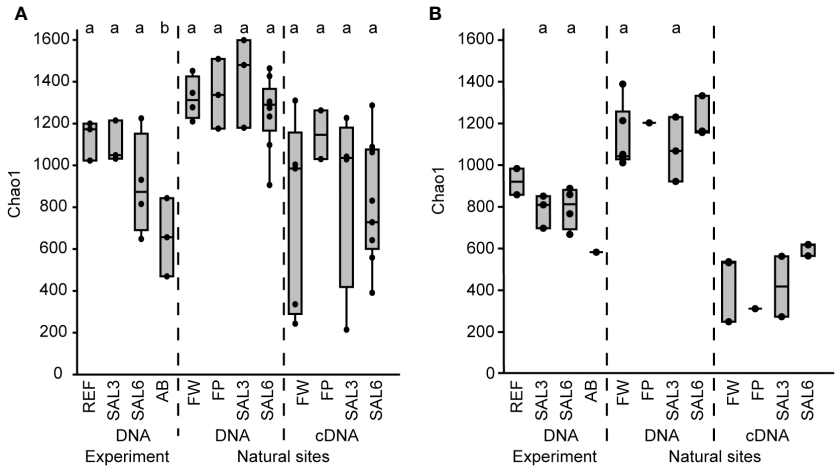


FIGURE 2
The bacterial Chao1 richness estimates based on the DNA and cDNA on the last day (day 8) of the experiment and in the natural sites. **(A)** Sediment; and **(B)** water. Experiment: REF, reference aquaria; SAL3, salinity increased to 3; SAL6, salinity increased to 6; AB, antibiotic treatment. Natural sites: FW, freshwater; FP, freshwater pond; SAL3, salinity 3; SAL6, salinity 6. The non-capital letters (a, b, and c) above the box plots indicate statistical significance within one group, which are separated by the dashed lines.

The differences in the sediment bacterial richness in natural site rivers and the coast of the Baltic Sea (i.e., FW, FP, SAL3, and SAL6 sites) and also between the experiment (i.e., REF, SAL3, and SAL6) and the natural sites were insignificant for the DNA-based analysis (Figure 2A). The Chao1 richness estimate and Shannon index revealed similar patterns, whereas the cDNA of the experiment did not contain enough samples for statistical analysis (Supplementary Figures 2A, 3A). The comparison of experiment day 8 and the natural sediment bacterial communities with the same salinity revealed that half of the operational taxonomic units (OTUs) were shared, whereas 35% were specific to natural SAL3 and 42% to SAL6 sites (Supplementary Figure 4). In the experiment fewer specific OTUs were observed than in the natural sites (16% in SAL3 and 12% in SAL6). A significantly higher Chao1 richness estimate was observed for only sediment DNA than with the cDNA in the natural sites for SAL6 (Tukey's test, $p < 0.05$; Figure 2A).

The water bacterial community in the natural sites had little variation in the taxonomic richness, the Chao1 estimate was 1,045 (SE ± 63.6) in freshwater rivers, 994 (SE ± 92.7) in the salinity 3 site, 1,144 (SE ± 59.9) in the salinity 6 sites, and 1,129 in the freshwater pond (Figure 2B). The cDNA-based Chao1 richness estimate was lower in all sites: 336 (SE ± 98.8) in FW, 314 (SE ± 150.5) in SAL3, 503 (SE ± 18.7) in the SAL6 sites, and 203 in the FP. The Shannon index of the water samples was comparable between DNA and cDNA (Supplementary Figure 3B).

In the water, the majority of OTUs were specific to natural sites, and 42% were shared between experiment day 8 and the natural sites for SAL3 and SAL6 (Supplementary Figure 5). Similar to the low number of specific OTUs observed in the sediment of the experiment, a low number of specific OTUs were also found in the water of the experiment (12% SAL3 and 9% SAL6) compared with the natural sites.

Bacterial community composition

The sediment and water bacterial community composition at the phylum/class level was dominated by Gammaproteobacteria (20.4%–33.6%), Alphaproteobacteria (9.3%–24.5%), and Bacteroidetes (8.8%–22.7%) in both the experiment and natural sites (Supplementary Figure 6).

On the finest phylogenetic level (OTU), the sediment bacterial community composition in the treatments showed high variability on PCoA (Figure 3A). The AB treatment differed significantly from the salinity treatments and reference (repeated measures PERMANOVA, $p = 0.007$, $R^2 = 49.0\%$). The SAL6 treatment did not show significant changes in the bacterial community composition during the first 6 days of the experiment. At the end of the experiment (i.e., day 8), the bacterial communities were separated into two clusters on the PCoA plot. Both clusters were represented by two replicates out of the four, with differences from the reference for one set (repeated measures PERMANOVA test, $p = 0.009$, $R^2 = 92.5\%$). The SAL3 treatment was not different from the REF. The cDNA-based analysis showed no difference in the active fraction of the bacterial community in any treatments (PERMANOVA, $p > 0.05$). On the basis of sediment DNA, sequences from *Hydrogenophaga*, *Flavobacterium*, unclassified Saprospiraceae, unclassified Anaerolineaceae, vadinHA17, unclassified *Chthoniobacter*, *Chthoniobacter*, unclassified Comamonadaceae, *Thiobacillus*, *Crenotrix*, SC-I-84, unclassified Steroidobacteriaceae, *Terriomonas*, and *Leptothrix* were among the most abundant bacteria in all the samples on any day of the experiment (Figure 4A). In the antibiotic treatment, the abundance of several OTUs decreased.

Artificial salinization and the addition of antibiotics caused a significant shift in the water bacterial community composition on

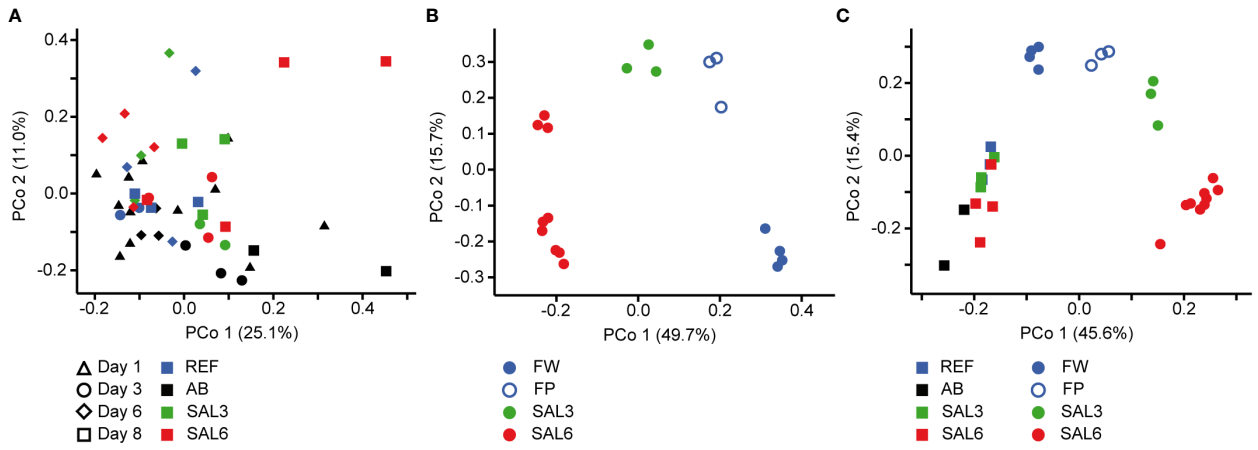


FIGURE 3
Principal coordinate analysis of the bacterial community composition in the sediment based on the DNA **(A)** during the experiment, **(B)** in the natural sites, and **(C)** on the last day of the experiment (day 8) and in the natural sites with the same salinity. Experiment: REF, reference aquaria; SAL3, salinity increased to 3; SAL6; salinity increased to 6; AB, antibiotic treatment. Natural sites: FW, freshwater; FP, freshwater pond; SAL3, salinity 3; SAL6, salinity 6.

the finest phylogenetic level (repeated measures PERMANOVA test, $p = 0.0001$, $R^2 = 73.7\%$; Figure 5A). In the REF and the SAL3 treatments, smaller shifts in the bacterial community composition occurred in the experiment, with the SAL6 and the antibiotic (AB) treatment showing stronger treatment effects, especially at day 6.

The bacterial community composition in the SAL6 treatment was characterized by a high abundance of a flavobacterial OTU (Figure 4B). This OTU was also present in high abundances in other treatments and dominated the natural sites but with lower abundances than the experiment samples. In the salinity 6 treatment, hydrogenophagal OTUs became most abundant, whereas in the SAL3 aquaria a smaller increase was observed. The chitinobacterial LD29 was also abundantly found in the treatment, by having the highest abundance in SAL6 on day 6. The genera *Limnohabitans* and *Fluvicola* were more abundant in the REF and SAL3 treatments, whereas in the SAL6 and AB treatments they had

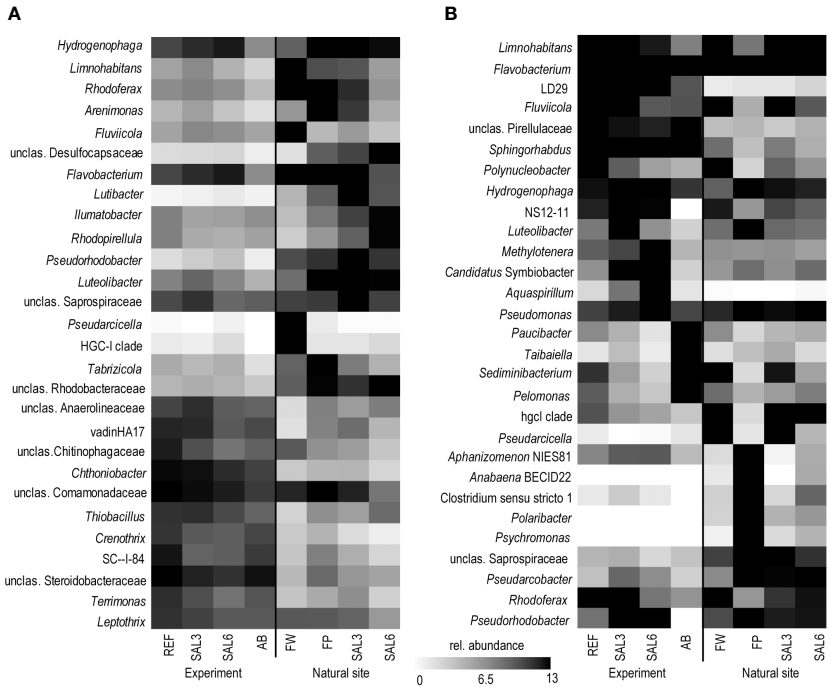


FIGURE 4
Most abundant OTUs on the last day of the experiment and in the natural sites based on DNA. **(A)** Sediment; and **(B)** water. Experiment: REF, reference aquaria; SAL3, salinity increased to 3; SAL6, salinity increased to 6; AB, antibiotic treatment. Natural sites: FW, freshwater; FP, freshwater pond; SAL3, salinity 3; SAL6, salinity 6. OTU, operational taxonomic unit.

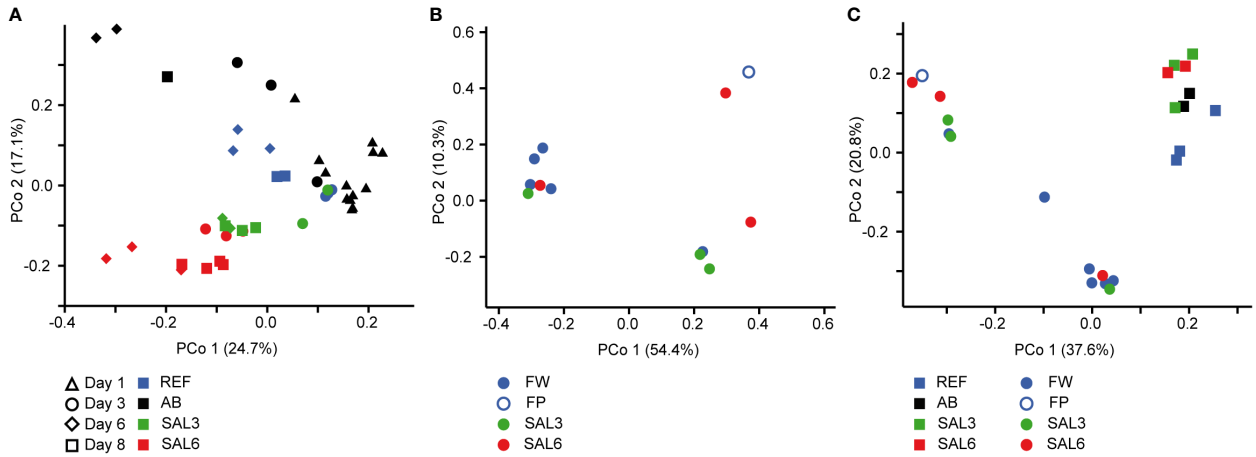


FIGURE 5
Principal coordinate analysis of the bacterial community composition in the water based on the DNA **(A)** during the experiment, **(B)** in the natural sites, and **(C)** on the last day of the experiment (day 8) and in the natural sites with the same salinity. Experiment: REF, reference aquaria; SAL3, salinity increased to 3; SAL6, salinity increased to 6; AB, antibiotic treatment. Natural sites: FW, freshwater; FP, freshwater pond; SAL3, salinity 3; SAL6, salinity 6.

a reduced prevalence. In addition, in both salinity treatments *Candidatus* Symbiobacter was among the most abundant bacteria. In the AB manipulation, *Sediminibacterium*, *Taibaiella*, *Paucibacter*, and *Sphingorhabdus* were resistant to the AB treatment, whereas the numbers of other bacteria were reduced compared with the reference and other treatments. The increase in Chao1 richness estimate in the water bacterial community between day 6 and day 8 (Figure 1) was connected with the appearance of highly diverse and low-abundance OTUs (Supplementary Figure 7), of which approximately 50% differed between each of the treatments (Supplementary Table 1).

Visualization by PCoA of the sediment and water bacterial community, including only day 1 and day 8, revealed a difference between day 1 and day 8 bacterial community composition on the first coordinate, and salinity on the second coordinate, especially for the water (Supplementary Figures 8A, B). The plotting of the second coordinate against salinity showed a steeper slope (water day 8: $y = -0.045x + 0.0328$; sediment day 8: $y = -0.035x + 0.032$) for day 8 water bacterial community composition than the sediment bacterial community composition (Supplementary Figure 9). The water bacterial community ANCOVA slopes were not equal ($p < 0.0001$, $F = 33.8$) indicating a stronger salinity effect in day 8 water bacterial community composition than in the sediment bacterial community.

The sediment bacterial communities had a statistically different composition in natural sites with different salinity (PERMANOVA, $R^2 = 44\%$, $p < 0.0001$; Figure 3B). The sediment samples from the temporal freshwater pond (FP) had the highest compositional similarity with the SAL3 site. The Jaccard similarity index value of the freshwater pond bacterial community composition, compared with other freshwater sites, was 61%, with SAL3 and SAL6 sites 66% and 58%, respectively. The pond sediment bacterial community was statistically different from salinity 6 sites (PERMANOVA, $R^2 = 32.8\%$, $p < 0.001$). The bacterial community in the Nõva sampling site (SAL6) was statistically different from other SAL6 sites (PERMANOVA, $R^2 = 67\%$, $p < 0.05$) indicating compositional variability in the littoral area with the same salinity. The cDNA-based analysis showed overlap in the sediment bacterial communities in FW, FP, and SAL3 sites leading to a statistically significant differentiation between sites having freshwater (FW and FP) and SAL6 sites (PERMANOVA, $R^2 = 66\%$, $p < 0.001$). In the natural sediment sites and in the experiment, *Flavobacterium*, *Limnohabitans*, and *Fluvicola* were the most abundant genera in the FW sites, but not in the pond, and their numbers decreased in the higher salinity (Figure 4A). *Pseudarcicella* and the HGC-I clade were also abundant in FW and had very low abundances in the FP, SAL3, and SAL6 sites. The freshwater pond was differentiated from the other sites by the higher numbers of *Tabrizicola* and *Arenimonas* and the shared abundant genera, such as *Hydrogenophaga*, *Luteolibacter*, and unclassified Rhodobacteraceae with the saline sites. In addition, salinity 3 had more *Pseudorhodobacter*, *Hydrogenophaga*, and unclassified Saprospiraceae in abundance, whereas in the salinity 6 sites *Ilumatobacter*, unclassified Desulfocapsaceae, and *Rhodopirellula* predominated. When the natural and treatment sites were compared, sediment bacterial communities from similar salinity sites did not group together. The PERMANOVA test validated statistically significant differences for SAL6 natural and experimental bacterial communities (Figure 3C; PERMANOVA $R^2 = 34\%$, $p < 0.05$).

The DNA- and cDNA-based bacterial community composition in the water sampled on the natural sites did not differ significantly (PERMANOVA $p > 0.05$) between the different salinities. The comparison of natural water sites and experiment bacterial communities with the respective salinity resulted with no clustering by salinity groups (Figure 5C).

The genera *Flavobacterium*, *Limnohabitans*, and *Fluvicola* were the most abundant OTUs in the natural sites (Figure 4B). The most abundant bacteria in the FW and SAL3 sites were very similar (Figure 5B). A very different bacterial community was found in the freshwater pond. The dominating genera were *Aphanizomenon*

NIES81, *Anabaena* BECID22, *Clostridium* sensu stricto 1, *Polaribacter*, and *Hydrogenophaga*.

Discussion

Salinity is among the most important factors in structuring bacterial communities (Lozupone and Knight, 2007). Current climate change and anthropogenic activities cause freshwater areas to become saline. This can occur as a pulse disturbance during an extreme weather event (IPCC, 2022) or as a result of anthropogenic activities, including industrial salt discharge (Cañedo-Argüelles et al., 2013). Alternatively, slower salinization occurs by sea level rise (IPCC, 2022), increased evaporation, and lower precipitation (Jeppesen et al., 2020), or via permanent industrial salt discharge (Bäthe and Coring, 2011). The aim of this study was to compare salinization effects on sediment and water bacterial communities in experimental setups and in natural sites to investigate if salinity disturbance alters the bacterial community composition and richness. Salinization occurs by mixing freshwater with saltier water, which results in brackish water typically characterized by a salinity range 3–26. We focused on the salinity range 0.2–6.0, which covers the first stage in freshwater salinization. Comparable to our experimental setup, a previous study used salinities of 3.3, 5.7, and 8.7 to mimic industrial salt discharges (Cañedo-Argüelles et al., 2013). The 2022 ecological disaster, caused by industrial salt discharge in the Oder River was also within the same salinity range (Free et al., 2023).

The first hypothesis, that an increase in salinity alters the bacterial community composition, was only partially proven in this study. An increase in salinity did not cause significant changes in the sediment bacterial community composition in the experiment (Figure 3A). However, for the water bacteria, a significant effect was observed (Figure 5A). This suggests that sediment bacterial communities are more resistant to these moderate pulse salinity disturbances than water bacterial communities. However, a potential effect on the sediment bacterial community due to increased salinity was observed for day 8 in the salinity 6 treatment. Here a change appeared in half of the replicated aquaria, suggesting that salinity 6 could be a tipping point where the sediment bacterial community becomes sensitive. This is supported by a drop in the Chao1 richness estimate after 8 days at salinity 6 (Figure 1A). In the well-established sediment bacterial communities in the Baltic Sea, the composition differed between salinity 4 and 8 (Klier et al., 2018). Accordingly, pulse salinity changes exceeding a salinity of 6 could impact sediment bacterial communities. More studies with a higher resolution in this salinity range are needed to evaluate a potential tipping point for sediment bacterial communities. As expected, the addition of antibiotics 1,000× stronger than the typical concentrations found in the environment (usually in the low ng/L range; Larsson, 2014) caused a significant disturbance in the sediment bacterial community composition and richness. Therefore, intense disturbance also results in changes in the sediment bacterial community. In contrast to the resistance against a pulse increase in salinity in the experiment, the littoral sediment bacterial community composition in rivers and coastal salinity 3 and 6 sites clustered according to salinity (Figure 3B). This is similar to previous studies that have found a significant separation of offshore sediment bacterial communities according to salinity (Yang et al., 2016; Klier et al., 2018). Hence, sediment bacterial communities are adapted to specific salinities during permanent exposure.

A significant salinization effect and a strong disturbance of antibiotics to the water bacterial community was observed in the experiment (Figure 5A) showing a sensitivity to different types of pulse disturbances. The responses of the water bacterial communities to the alteration of salinity (Langenheder et al., 2003) and antibiotics (Eckert et al., 2019) have been shown previously. The missing clustering of water bacterial community composition in natural sites according to salinity (Figure 5B) suggests that the littoral bacterial communities were strongly influenced by local factors, including macrophytes (Duarte et al., 2005), nutrient concentration, and terrestrial effects (Figueroa et al., 2021). This supports the hypothesis that water bacterial communities were sensitive to complex mixtures of environmental factors that were stronger than the effect of salinity alone. However, as we focused on the effect of salinity, we can only speculate about the impact of the other parameters. In addition, historical site-specific events and temporal changes in salinity could have altered the development of the bacterial community in the water (Renes et al., 2020; Philippot et al., 2021). In contrast to the complex mixture of changing factors in the environment, the variability of these factors was minimized and similar in all aquaria in the experiment thus allowing us to focus on the effect of salinity to bacterial communities alone.

The second hypothesis about the reduction of bacterial richness during salinity disturbance can only partially be proven as the water bacterial richness changed significantly in the experiment but sediment bacterial richness did not. A decrease in the bacterial richness is a typical response of aquatic bacterial communities to pulse disturbance (Allison, 2004; Downing and Leibold, 2010; Renes et al., 2020). However, a decrease in the water bacterial richness was observed in parallel in all aquaria, including in the reference. This was a reaction to the confined artificial aquaria conditions where there was an altered nutrient availability (Ionescu et al., 2015; Herlemann et al., 2017; Herlemann et al., 2019) and different grazing patterns of bacteria by ciliates and heterotrophic nanoflagellates (Ionescu et al., 2015; Herlemann et al., 2019). A significantly stronger dynamic of bacterial richness was observed for the more intense manipulation (i.e., SAL6 and AB) supporting a sensitivity of the water bacterial richness to pulse disturbance in addition to confinement stress. Disturbance intensity is a key determining how water bacterial communities respond to disturbance (Berga et al., 2012; Philippot et al., 2021; Kivistik et al., 2022). In salinization experiments of rock pools (salinity range 3–12), bacterial communities changed at an increasing magnitude with increasing salinities (Berga et al., 2017). In contrast to pulse stress, bacterial richness has been shown to be constant at different salinity levels in permanent brackish water environments such as in the Baltic Sea (Herlemann et al., 2011; Herlemann et al., 2016). Similar results were found in our study for natural sites with different salinity (Figure 2). Hence, our results

suggest that water bacterial communities are reacting to strong pulse disturbances with decreased species diversity and changes in the bacterial community composition. A longer exposure to different salinity caused the adaptation of the bacterial community resulting in comparable richness. However, a reduced water bacterial richness in permanent exposure to the salinity can be found in hypersaline lakes (Ji et al., 2019) suggesting that this is only applicable to the brackish environment.

The OTUs abundantly found in the sediment and water, including *Flavobacterium*, *Hydrogenophaga*, and *Limnohabitans*, were also abundant in the experiment. They are typical aquatic bacteria, of which the *Flavobacterium* species are known members of freshwater but also the Baltic Sea (Klier et al., 2018). Water bacteria of the genus *Limnohabitans* are capable of utilizing algal exudates and can survive well in artificial environments (Pérez and Sommaruga, 2006; Jezberová et al., 2017). The abundance of *Hydrogenophaga*, a genus containing many hydrogen-oxidizing bacteria, increased with increased salinity. In contrast, the abundance of the HGC-I OTU was high in natural water sites, except the pond, but was significantly lower in the experiment, as this bacterial group is known to be sensitive to artificial environments (Herlemann et al., 2019).

The increase in salinity did not change the sediment bacterial community in the experiment severely. The bacteria abundant during the experiment (unclassified Comamonadaceae, unclassified Steroidobacteriaceae, *Chthoniobacter*, Bacteroidetes vadinHA17, and *Thiobacillus*) were also among the most abundant sediment bacteria in the natural samples (Figure 4A). Aerobic Comamonadaceae have been found in lower salinity concentrations in previous studies (Aguirre et al., 2017). The *Chthoniobacter* genus contains only free-living, aerobic chemoheterotrophic, and metabolizing organic carbon species found in soil and freshwater sediments (Sangwan et al., 2004; MGnify, 2019). The majority of the available vadinHA17 sequences in the SILVA NR database have been obtained from sediments and anaerobic bioreactors where proteinaceous substrates are abundant and these bacteria are most likely to be capable of protein and amino acids degradation (Mei et al., 2020). The data of salinity tolerance for the vadinHA17 group suggest an adaptation to brackish environments (Klier et al., 2018). This is in accordance with our results where the abundance did not decrease up to a salinity of 6. *Thiobacillus* species have been found to dominate surface sediments in freshwater lake–river systems (Duan et al., 2020). In a salinity of 6, the *Chloroflexi* group KD4–96 also became abundant. This group has been found in deep-sea sediments, but also in the Bothnian Sea with a similar salinity to that which was used in our experiment (Rasigraf et al., 2020). Several of the most abundant OTUs that tolerated the increase in salinity were also found in the antibiotic treatment (e.g., *Hydrogenophaga*, *Limnohabitans*, *Arenimonas*, *Luteoliobacter*) but in lower abundances. The differences in salinity of natural site sediments were enough to facilitate different dominating bacteria. In a salinity of 3, the genera *Hydrogenophaga*, *Flavobacterium*, and *Pseudorhodobacter*, and unclassified Saprospiraceae predominated. The salinity 6 samples from the envrionment and the experiment contained bacteria found in brackish water sediments with both aerobic and anaerobic metabolisms, such as *Ilumatobacter* (aerobic heterotroph) (Fang et al., 2015), unclassified Desulfocapsaceae (anaerobic respiration) (Galushko and Kuever, 2021), and *Rhodopirellula* (chemoheterotrophic aerobes) (Sreya et al., 2023).

In addition to *Limnohabitans*, *Flavobacterium*, and HGC-I, the verrucomicrobial lineage LD29 were highly abundant in the experiment water but had very low abundance in the natural sites. Previously these bacteria have been found in a similar salinity range in the Baltic Sea (Herlemann et al., 2013) but have strong seasonal dynamics (Lindh et al., 2015) and were therefore possibly missed in our natural samples. In the water bacterial community, antibiotic treatment supported the dominance of *Sphingorhabdus*, *Sediminibacterium*, *Taibaiella*, and *Paucibacter*, which are probably more resistant to the added antibiotics. Hence, the Chao1 richness recovery in the antibiotic treatment is probably connected with the ability of the bacteria to proliferate in the presence of the antibiotics used.

Despite the overlaps of the OTUs between the experimental and natural sites, there was a significant difference in the overall bacterial community composition (Figures 3C, 5C). An important difference between the natural site and experiment are the abundances of the bacteria but also the less abundant OTUs.

An exceptional salinity-driven bacterial community was found in the natural pond situated on the shoreline of the Baltic Sea. It is located in close proximity to the outflow from a freshwater river. The pond contained freshwater during sampling, whereas the investigated sediment bacterial community had a high similarity to the brackish coastal bacterial communities (Figure 3B). The pond water bacterial community composition clustered randomly. We hypothesize that the pond was filled by freshwater due to the natural shift of the river outflow in the spring (i.e., 4 months previous to the sampling), but the sediment was originally part of the brackish Baltic Sea shoreline. Therefore, the brackish sediment bacterial community was overlaid with freshwater but contained its natural brackish sediment bacteria. As a result, the pond had a brackish sediment bacterial community despite being filled with freshwater. This supports the resistance of sediment bacterial communities to changes in salinity possibly for several months in the environment. The freshwater pond sediment was different from the other sites by the increased abundance of *Tabrizicola*, *Arenimonas*, and the facultative anaerobe *Rhodoferax*. The genus *Tabrizicola* consists mostly of chemotrophic bacteria but also has members with anoxygenic photosynthetic capabilities that have been isolated from freshwater or saline lake sediment (Liu et al., 2019; Tarhriz et al., 2019). The genus *Arenimonas*, which is abundant in the sediment of the freshwater pond and salinity 3 site, has been previously isolated from soil (Han et al., 2020) but also from seashore sand (Kwon et al., 2007). Most known members of the genus *Rhodoferax* are anoxygenic photoheterotrophs or organomixotrophs with capability to reduce Fe(III) (Finneran et al., 2003). The pond sediment shared abundant genera, such as *Hydrogenophaga* and *Luteolibacter*, and unclassified Rhodobacteraceae with saline site sediments, reflecting its history of being in contact with the Baltic Sea. Hence a brackish photoheterotrophic bacterial community seems to be characteristic for this pond.

During the last decades, a number of studies have investigated the effects of disturbance on the bacterial community composition with the conclusion that bacterial communities are in most cases

sensitive to disturbances (Allison and Martiny, 2008; Shade et al., 2012; Griffiths and Philippot, 2013; Lindh and Pinhassi, 2018). Disturbance may change species richness and bacterial community composition, which may in turn have consequences for the stability of ecosystem processes (Langenheder et al., 2006; Peter et al., 2011; Reed and Martiny, 2013; Gibbons et al., 2016; Berga et al., 2017; Renes et al., 2020). The effect of disturbance depends on its intensity, disturbance type, frequency, and length, as well as the species tolerance capacity (Shade et al., 2012; Gibbons et al., 2016; Eckert et al., 2019). Compositional resistance of water bacterial communities has been shown to decrease with increasing salinity levels (Berga et al., 2017). Contrary to this, our study showed that the sediment bacterial community was resistant and buffered pulse disturbance in the salinity range tested. A permanent exposure to natural brackish conditions caused a selection of sediment bacteria without loss in species richness. This represents a long-term adaptation of the brackish sediment bacterial communities that perform the ecosystem processes. A resistance of sediment bacteria to changes in salinity has been shown previously (Reed and Martiny, 2013). The reason for the resistance of bacterial communities to salinification has been connected with fluctuating environments (Berga et al., 2017). However, since the origin of the water and sediment used in our experiment was a freshwater lake that has not been in contact with saltwater since the last ice age, adaptation to salt fluctuations cannot explain the observed resistance. Sediment bacteria are dispersal limited due to the physical barrier of the sediments. This dispersal limitation could be a reason for the higher resistance of sediment bacterial communities toward salinity disturbance. The importance of dispersal to bacterial communities in aquatic environments has been shown in several studies (Jones and McMahon, 2009; Lindström and Östman, 2011; Martiny et al., 2011; Székely et al., 2013). In the environment, dispersal can supply bacteria that are better adapted to a disturbed environment. The example of the pond shows that several months of being exposed to freshwater was insufficient for a shift in the brackish sediment bacterial community composition. The laboratory experiments, in contrast, excluded effects of natural dispersal that were present in the environment (Lennon and Jones, 2011). Together with the limited time, dispersal limitation of the sediment resulted in the resistance of the sediment bacterial community to salinity pulse changes. Moreover, a similar bacterial richness in the sediment based on DNA and cDNA suggests a limited seedbank of bacteria available that could respond to changes in salinity. These dormant bacteria are a valuable seedbank, which consists of individuals capable of being resuscitated after or during disturbance (Lennon and Jones, 2011). The pulse salinity disturbances had a significant effect on the water bacterial richness and community composition in the experiments. Also, for the water bacterial community, the laboratory experiment excludes the effects of natural dispersal. The cDNA-to-DNA ratio was much lower in the water than in the sediment, suggesting that a large water bacterial seedbank provides a source for shifts in the bacterial community. Finally, neither the sediment nor the water bacterial community in the experiment resembled the natural brackish bacterial communities after changes in salinity (Figures 3C, 5C). This supports the important role of a seedbank during disturbance and also suggests a biogeography for bacterial communities.

We conclude that sediment bacterial communities are more resistant to pulse salinization than the water bacterial communities. The resistance of sediment bacteria may be caused by dispersal limitations due to the physical structure of the sediment and a smaller seedbank in our experiment. However, a longer exposure to salinity and dispersal, as experienced in the environment, results in an adaptation of the sediment bacterial communities. Moreover, stronger effects such as antibiotics indicate that other stressors than salinity may evade the resistance of sediment bacterial communities during pulse disturbance. For the management of industrial salt disposal and climate change, this implies a resistance of sediment bacterial communities to mild short-term salinization events, possibly without significant effects on ecosystem functions. For water bacterial communities significant responses to changes in salinity are expected.

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) are as follows: https://www.ncbi.nlm.nih.gov/, bioproject PRJNA724976 accession numbers SAMN18865857-SAMN18866052.

Author contributions

All authors performed the experiments and sampling. HT and CK did DNA extraction and preparation for sequencing. HT, VK, and DH did statistical and bioinformatics analysis. HT submitted sequences. HT wrote the original draft. HT, DH, CK, and VK contributed to writing, reviewing, and editing of the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This study was funded by the European Regional Development Fund/Estonian Research Council project “Mobilitas Plus Top Researcher grant MOBTT24”. The study was also supported by the Estonian University of Life Sciences projects P190250PKKH and P200028PKKH and the Leibniz Institute for Baltic Sea Research Warnemünde in the new research area Shallow water processes and Transitions to the Baltic Scale.

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 author(s) DH and VK declared that they were an editorial board members of Frontiers at the time of submission. This had no impact on the peer review process and the final decision.

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.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frmbi.2023.1232571/full#supplementary-material

References

Aguirre, M., Abad, D., Albaina, A., Cralle, L., Goñi-Urriza, M. S., Estonba, A., et al. (2017). Unraveling the environmental and anthropogenic drivers of bacterial community changes in the estuary of Bilbao and its tributaries. *PloS One* 12, e0178755. doi: 10.1371/journal.pone.0178755

Allison, G. (2004). The influence of species diversity and stress intensity on community resistance and resilience. *Ecol. Monogr.* 74, 117–134. doi: 10.1890/02-0681

Allison, S. D., and Martiny, J. B. H. (2008). Resistance, resilience, and redundancy in microbial communities. *P Natl. Acad. Sci. U.S.A.* 105, 11512–11519. doi: 10.1073/pnas.0801925105

Andersson, M. G., Berga, M., Lindström, E. S., and Langenheder, S. (2014). The spatial structure of bacterial communities is influenced by historical environmental conditions. *Ecology* 95, 1134–1140. doi: 10.1890/13-1300.1

Bäthe, J., and Coring, E. (2011). Biological effects of anthropogenic salt-load on the aquatic Fauna: A synthesis of 17 years of biological survey on the rivers Werra and Weser. *Limnologica* 41, 125–133. doi: 10.1016/j.limno.2010.07.005

Berga, M., Székely, A. J., and Langenheder, S. (2012). Effects of disturbance intensity and frequency on bacterial community composition and function. *PloS One* 7, e36959. doi: 10.1371/journal.pone.0036959

Berga, M., Zha, Y., Székely, A. J., and Langenheder, S. (2017). Functional and compositional stability of bacterial metacommunities in response to salinity changes. *Front. Microbiol.* 8, 948. doi: 10.3389/fmicb.2017.00948

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

Cañedo-Argüelles, M., Kefford, B. J., Piscart, C., Prat, N., Schäfer, R. B., and Schulz, C.-J. (2013). Salinisation of rivers: An urgent ecological issue. *Environ. pollut.* 173, 157–167. doi: 10.1016/j.envpol.2012.10.011

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

Downing, A. L., and Leibold, M. A. (2010). Species richness facilitates ecosystem resilience in aquatic food webs. *Freshw. Biol.* 55, 2123–2137. doi: 10.1111/j.1365-2427.2010.02472.x

Duan, J. L., Sun, J. W., Ji, M. M., Ma, Y., Cui, Z. T., Tian, R. K., et al. (2020). Indicatory bacteria and chemical composition related to sulfur distribution in the river-lake systems. *Microbiol. Res.* 236, 126453. doi: 10.1016/j.micres.2020.126453

Duarte, C. M., Holmer, M., and Marbà, N. (2005). "Plant–microbe interactions in seagrass meadows." in Interactsions between macro- and microorganisms in marine sediments. Eds. E. Kristensen, R. R. Haese and J. E. Kostka. (Washington, DC: American Geophysical Union), 31–60. doi: 10.1029/CE060p0031

Eckert, E. M., Quero, G. M., Di Cesare, A., Manfredini, G., Mapelli, F., Borin, S., et al. (2019). Antibiotic disturbance affects aquatic microbial community composition and food web interactions but not community resilience. *Mol. Ecol.* 28, 1170–1182. doi: 10.1111/mec.15033

Fang, L., Chen, L., Liu, Y., Tao, W., Zhang, Z., Liu, H., et al. (2015). Planktonic and sedimentary bacterial diversity of Lake Sayram in summer. *Microbiologyopen* 4, 814–825. doi: 10.1002/mbo3.281

Figueroa, D., Capo, E., Lindh, M. V., Rowe, O. F., Paczkowska, J., Pinhassi, J., et al. (2021). Terrestrial dissolved organic matter inflow drives temporal dynamics of the bacterial community of a subarctic estuary (northern Baltic Sea). *Environ. Microbiol.* 23, 4200–4213. doi: 10.1111/1462-2920.15597

Finneran, K. T., Johnsen, C. V., and Lovley, D. R. (2003). *Rhodoferax ferrireducens* sp. nov., a psychrotolerant, facultatively anaerobic bacterium that oxidizes acetate with the reduction of Fe(III). *Int. J. Syst. Evol. Microbiol.* 53, 669–673. doi: 10.1099/ijs.0.02298-0

Fortunato, C. S., Herfort, L., Zuber, P., Baptista, A. M., and Crump, B. C. (2012). Spatial variability overwhelms seasonal patterns in bacterioplankton communities across a river to ocean gradient. *Isme J.* 6, 554–563. doi: 10.1038/ismej.2011.135

Free, G., Bund, W., Gawlik, B., Wijk, L., Wood, M., Guagnini, E., et al. (2023). *An EU analysis of the ecological disaster in the Oder River of 2022*. (Luxembourg: Publications Office of the European Union). doi: 10.2760/067386

Galushko, A., and Kuever, J. (2021). "*Desulfocapsaceae,*" *in Bergey's manual of systematics of archaea and bacteria*. Eds M. E. Trujillo, S. Dedysh, P. DeVos, B. Hedlund, P. Kämpfer, F. A. Rainey, et al (John Wiley & Sons, Inc., in association with Bergey's Manual Trust), 1–6. doi: 10.1002/9781118960608.fbm00332

Gibbons, S. M., Scholz, M., Hutchison, A. L., Dinner, A. R., Gilbert, J. A., and Coleman, M. L. (2016). Disturbance regimes predictably alter diversity in an ecologically complex bacterial system. *mBio* 7. doi: 10.1128/mBio.01372-16

Glöckner, F. O., Yilmaz, P., Quast, C., Gerken, J., Beccati, A., Ciuprina, A., et al. (2017). 25 years of serving the community with ribosomal RNA gene reference databases and tools. *Journal of biotechnology* 261, 169-176.

Griffiths, B. S., and Philippot, L. (2013). Insights into the resistance and resilience of the soil microbial community. *FEMS Microbiol. Rev.* 37, 112–129. doi: 10.1111/j.1574-6976.2012.00343.x

Hammer, Ø., Harper, D. A., and Ryan, P. D. (2001). PAST: Paleontological statistics software package for education and data analysis. *Palaeontol electron* 4, 9.

Han, D. M., Chun, B. H., Kim, H. M., Khan, S. A., and Jeon, C. O. (2020). *Arenimonas terrae* sp. nov., isolated from orchard soil. *Int. J. Syst. Evol. Microbiol.* 70, 537–542. doi: 10.1099/ijsem.0.003785

Herlemann, D. P. R., 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 jl* 5, 1571–1579. doi: 10.1038/ismej.2011.41

Herlemann, D. P. R., Lundin, D., Andersson, A. F., Labrenz, M., and Jürgens, K. (2016). Phylogenetic signals of salinity and season in bacterial community composition across the salinity gradient of the Baltic Sea. *Front. Microbiol.* 7. doi: 10.3389/fmicb.2016.01883

Herlemann, D. P. R., Lundin, D., Labrenz, M., Jürgens, K., Zheng, Z., Aspeborg, H., et al. (2013). Metagenomic *de novo* assembly of an aquatic representative of the verrucomicrobial Class *Spartobacteria*. *mBio* 4, e00569–e00512. doi: 10.1128/mBio.00569-12

Herlemann, D. P. R., Manecki, M., Dittmar, T., and Jürgens, K. (2017). Differential responses of marine, mesohaline and oligohaline bacterial communities to the addition of terrigenous carbon. *Environ. Microbiol.* 19, 3098–3117. doi: 10.1111/1462-2920.13784

Herlemann, D. P. R., Markert, S., Meeske, C., Andersson, A. F., De Bruijn, I., Hentschker, C., et al. (2019). Individual physiological adaptations enable selected bacterial taxa to prevail during long-term incubations. *Appl. Environ. Microbiol.* 85, e00825–19. doi: 10.1128/AEM.00825-19

Hillebrand, H., and Blenckner, T. (2002). Regional and local impact on species diversity – from pattern to processes. *Oecologia* 132, 479–491. doi: 10.1007/s00442-002-0988-3

Ionescu, D., Bizic-Ionescu, M., Khalili, A., Malekmohammadi, R., Morad, M. R., De Beer, D., et al. (2015). A new tool for long-term studies of POM-bacteria interactions: overcoming the century-old Bottle Effect. *Sci. Rep-UK* 5, 14706. doi: 10.1038/srep14706

IPCC (2022). *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, UK: Cambridge University Press, and New York, NY, USA). doi: 10.1017/9781009157926

Jeppesen, E., Beklioğlu, M., Özkan, K., and Akyürek, Z. (2020). Salinization increase due to climate change will have substantial negative effects on inland waters: A call for multifaceted research at the local and global scale. *Innovation* 1. doi: 10.1016/j.xinn.2020.100030

Jezberová, J., Jezbera, J., Znachor, P., Nedoma, J., Kasalický, V., and Šimek, K. (2017). The limnohabitans genus harbors generalistic and opportunistic subtypes: evidence

from spatiotemporal succession in a canyon-shaped reservoir. *Appl. Environ. Microbiol.* 83. doi: 10.1128/AEM.01530-17

Ji, M., Kong, W., Yue, L., Wang, J., Deng, Y., and Zhu, L. (2019). Salinity reduces bacterial diversity, but increases network complexity in Tibetan Plateau lakes. *FEMS Microbiol. Ecol.* 95. doi: 10.1093/femsec/fiz190

Johnson, B. R., Haas, A., and Fritz, K. M. (2010). Use of spatially explicit physicochemical data to measure downstream impacts of headwater stream disturbance. *Water Resour. Res.* 46, W09526. doi: 10.1029/2009WR008417

Jones, S. E., and McMahon, K. D. (2009). Species-sorting may explain an apparent minimal effect of immigration on freshwater bacterial community dynamics. *Environ. Microbiol.* 11, 905–913. doi: 10.1111/j.1462-2920.2008.01814.x

Kivistik, C., Käiro, K., Tammert, H., Sokolova, I. M., Kisand, V., and Herlemann, D. P. R. (2022). Distinct stages of the intestinal bacterial community of *Ampullaceana balthica* after salinization. *Front. Microbiol.* 13. doi: 10.3389/fmicb.2022.767334

Kivistik, C., Knobloch, J., Käiro, K., Tammert, H., Kisand, V., Hildebrandt, J.-P., et al. (2020). Impact of salinity on the gastrointestinal bacterial community of *Theodoxus fluviatilis*. *Front. Microbiol.* 11. doi: 10.3389/fmicb.2020.00683

Klier, J., Dellwig, O., Leipe, T., Jürgens, K., and Herlemann, D. P. R. (2018). Benthic bacterial community composition in the oligohaline-marine transition of surface sediments in the Baltic Sea based on rRNA analysis. *Front. Microbiol.* 9. doi: 10.3389/fmicb.2018.00236

Kwon, S. W., Kim, B. Y., Weon, H. Y., Baek, Y. K., and Go, S. J. (2007). *Arenimonas donghaensis* gen. nov., sp. nov., isolated from seashore sand. *Int. J. Syst. Evol. Microbiol.* 57, 954–958. doi: 10.1099/ijs.0.64457-0

Langenheder, S., Kisand, V., Wikner, J., and Tranvik, L. J. (2003). Salinity as a structuring factor for the composition and performance of bacterioplankton degrading riverine DOC. *FEMS Microbiol. Ecol.* 45, 189–202. doi: 10.1016/S0168-6496(03)00149-1

Langenheder, S., Lindström, E. S., and Tranvik, L. J. (2006). Structure and function of bacterial communities emerging from different sources under identical conditions. *Appl. Environ. Microbiol.* 72, 212–220. doi: 10.1128/AEM.72.1.212-220.2006

Larsson, D. G. (2014). Antibiotics in the environment. *Ups J. Med. Sci.* 119, 108–112. doi: 10.3109/03009734.2014.896438

Lennon, J. T., and Jones, S. E. (2011). Microbial seed banks: the ecological and evolutionary implications of dormancy. *Nat. Rev. Microbiol.* 9, 119–130. doi: 10.1038/nrmicro2504

Lindh, M. V., and Pinhassi, J. (2018). Sensitivity of bacterioplankton to environmental disturbance: A review of Baltic Sea field studies and experiments. *Front. Microbiol.* 5. doi: 10.3389/fmars.2018.00361

Lindh, M. V., Sjöstedt, J., Andersson, A. F., Baltar, F., Hugerth, L. W., Lundin, D., et al. (2015). Disentangling seasonal bacterioplankton population dynamics by high-frequency sampling. *Environ. Microbiol.* 17, 2459–2476. doi: 10.1111/1462-2920.12720

Lindström, E. S., and Östman, Ö. (2011). The importance of dispersal for bacterial community composition and functioning. *PloS One* 6, e25883. doi: 10.1371/journal.pone.0025883

Liu, Z.-X., Dorji, P., Liu, H.-C., Li, A.-H., and Zhou, Y.-G. (2019). *Tabrizicola sediminis* sp. nov., one aerobic anoxygenic photoheterotrophic bacteria from sediment of saline lake. *Int. J. Syst. Evol. Microbiol.* 69, 2565–2570. doi: 10.1099/ijsem.0.003542

Logares, R., Bråte, J., Bertilsson, S., Clasen, J. L., Shalchian-Tabrizi, K., and Rengefors, K. (2009). Infrequent marine-freshwater transitions in the microbial world. *Trends Microbiol.* 17, 414–422. doi: 10.1016/j.tim.2009.05.010

Lozupone, C. A., and Knight, R. (2007). Global patterns in bacterial diversity. *P Natl. Acad. Sci. U.S.A.* 104.27, 11436–11440. doi: 10.1073/pnas.0611525104

Lueders, T., Manefield, M., and Friedrich, M. W. (2004). Enhanced sensitivity of DNA- and rRNA-based stable isotope probing by fractionation and quantitative analysis of isopycnic centrifugation gradients. *Environ. Microbiol.* 6, 73–78. doi: 10.1046/j.1462-2920.2003.00536.x

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* 17 (1), 10–12. doi: 10.14806/ej.17.1.200

Martiny, J. B., Eisen, J. A., Penn, K., Allison, S. D., and Horner-Devine, M. C. (2011). Drivers of bacterial beta-diversity depend on spatial scale. *P Natl. Acad. Sci. U.S.A.* 108, 7850–7854. doi: 10.1073/pnas.1016308108

Mei, R., Nobu, M. K., Narihiro, T., and Liu, W. T. (2020). Metagenomic and Metatranscriptomic analyses revealed uncultured bacteroidales populations as the dominant proteolytic amino acid degraders in anaerobic digesters. *Front. Microbiol.* 11, 593006. doi: 10.3389/fmicb.2020.593006

MGnify (2019). Microbial community in PAH-contaminated riverine sediments. Sampling event dataset. doi: 10.15468/jv4atf

Olli, K., Tamminen, T., and Ptacnik, R. (2023). Predictable shifts in diversity and ecosystem function in phytoplankton communities along coastal salinity continua. *Limnol. Oceanogr.* 8 (1), 173–180. doi: 10.1002/lol2.10242

Paulson, J. N., Stine, O. C., Bravo, H. C., and Pop, M. (2013). Differential abundance analysis for microbial marker-gene surveys. *Nat. Methods* 10, 1200–1202. doi: 10.1038/nmeth.2658

Pérez, M. T., and Sommaruga, R. (2006). Differential effect of algal- and soil-derived dissolved organic matter on alpine lake bacterial community composition and activity. *Limnol. Oceanogr.* 51, 2527–2537. doi: 10.4319/lo.2006.51.6.2527

Peter, H., Beier, S., Bertilsson, S., Lindström, E. S., Langenheder, S., and Tranvik, L. J. (2011). Function-specific response to depletion of microbial diversity. *Isme J.* 5, 351–361. doi: 10.1038/ismej.2010.119

Philippot, L., Griffiths, B. S., and Langenheder, S. (2021). Microbial community resilience across ecosystems and multiple disturbances. *Microbiol. Mol. Biol. Rev.* 85. doi: 10.1128/MMBR.00026-20

Pruesse, E., Peplies, J., and Glöckner, F. O. (2012). SINA: accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics* 28 (14), 1823–1829. doi: 10.1093/bioinformatics/bts252

Rasigraf, O., Van Helmond, N., Frank, J., Lenstra, W. K., Egger, M., Slomp, C. P., et al. (2020). Microbial community composition and functional potential in Bothnian Sea sediments is linked to Fe and S dynamics and the quality of organic matter. *Limnol. Oceanogr.* 65, S113–S133. doi: 10.1002/lno.11371

Reed, H. E., and Martiny, J. B. H. (2013). Microbial composition affects the functioning of estuarine sediments. *Isme J.* 7, 868–879. doi: 10.1038/ismej.2012.154

Remane, A. (1934). *Die Brackwasserfauna: Mit besonderer Berücksichtigung der Ostsee* (Zoologischer Anzeiger).

Renes, S. E., Sjöstedt, J., Fetzer, I., and Langenheder, S. (2020). Disturbance history can increase functional stability in the face of both repeated disturbances of the same type and novel disturbances. *Sci. Rep-UK* 10, 11333. doi: 10.1038/s41598-020-68104-0

Robertson, C. E., Harris, J. K., Wagner, B. D., Granger, D., Browne, K., Tatem, B., et al. (2013). Explicet: graphical user interface software for metadata-driven management, analysis and visualization of microbiome data. *Bioinformatics* 29, 3100–3101. doi: 10.1093/bioinformatics/btt526

Rognes, T., Flouri, T., Nichols, B., Quince, C., and Mahé, F. (2016). VSEARCH: a versatile open source tool for metagenomics. *PeerJ* 4, e2584.

Sangwan, P., Chen, X., Hugenholtz, P., and Janssen, P. H. (2004). *Chthoniobacter flavus* gen. nov., sp. nov., the first pure-culture representative of subdivision two, Spartobacteria classis nov., of the phylum Verrucomicrobia. *Appl. Environ. Microbiol.* 70, 5875–5881. doi: 10.1128/AEM.70.10.5875-5881.2004

Sauer, F. G., Bundschuh, M., Zubrod, J. P., Schäfer, R. B., Thompson, K., and Kefford, B. J. (2016). Effects of salinity on leaf breakdown: Dryland salinity versus salinity from a coal mine. *Aquat Toxicol.* 177, 425–432. doi: 10.1016/j.aquatox.2016.06.014

Seidel, L., Broman, E., Nilsson, E., Ståhle, M., Ketzer, M., Pérez-Martínez, C., et al. (2023). Climate change-related warming reduces thermal sensitivity and modifies metabolic activity of coastal benthic bacterial communities. *Isme J.* 17, 855–869. doi: 10.1038/s41396-023-01395-z

Shade, A., Peter, H., Allison, S., Baho, D., Berga, M., Buergmann, H., et al. (2012). Fundamentals of microbial community resistance and resilience. *Front. Microbiol.* 3. doi: 10.3389/fmicb.2012.00417

Sreya, P., Suresh, G., Rai, A., Ria, B., Vighnesh, L., Agre, V. C., et al. (2023). Revisiting the taxonomy of the genus *Rhodopirellula* with the proposal for reclassification of the genus to *Rhodopirellula* sensu stricto, *Aporhodopirellula* gen. nov., *Allorhodopirellula* gen. nov. and *Neorhodopirellula* gen. nov. *A Van Leeuw J. Microb.* 116, 243–264. doi: 10.1007/s10482-022-01801-0

Székely, A. J., Berga, M., and Langenheder, S. (2013). Mechanisms determining the fate of dispersed bacterial communities in new environments. *Isme J.* 7, 61–71. doi: 10.1038/ismej.2012.80

Tarhriz, V., Hirose, S., Fukushima, S.-I., Hejazi, M. A., Imhoff, J. F., Thiel, V., et al. (2019). Emended description of the genus *Tabrizicola* and the species *Tabrizicola aquatica* as aerobic anoxygenic phototrophic bacteria. *A Van Leeuw J. Microb.* 112, 1169–1175. doi: 10.1007/s10482-019-01249-9

Weinbauer, M. G., Fritz, I., Wenderoth, D. F., and Höfle, M. G. (2002). Simultaneous extraction from bacterioplankton of total RNA and DNA suitable for quantitative structure and function analyses. *Appl. Environ. Microbiol.* 68, 1082–1087. doi: 10.1128/AEM.68.3.1082-1087.2002

Yang, J., Ma, L. A., Jiang, H., Wu, G., and Dong, H. (2016). Salinity shapes microbial diversity and community structure in surface sediments of the Qinghai-Tibetan Lakes. *Sci. Rep-UK* 6, 25078. doi: 10.1038/srep25078

frontiers | Frontiers in Microbiomes

TYPE Original Research
PUBLISHED 08 January 2024
DOI 10.3389/frmbi.2023.1329925

OPEN ACCESS

EDITED BY
Gareth Trubl,
Lawrence Livermore National Laboratory (DOE), United States

REVIEWED BY
Shengwei Hou,
Southern University of Science and Technology, China
Chandni Sidhu,
Max Planck Society, Germany

*CORRESPONDENCE
Danny Ionescu
danny.ionescu@igb-berlin.de

RECEIVED 30 October 2023
ACCEPTED 14 December 2023
PUBLISHED 08 January 2024

CITATION
Ionescu D, Zoccarato L, Cabello-Yeves PJ and Tikochinski Y (2024) Extreme fluctuations in ambient salinity select for bacteria with a hybrid "salt-in"/" salt-out" osmoregulation strategy. *Front. Microbiomes* 2:1329925. doi: 10.3389/frmbi.2023.1329925

Extreme fluctuations in ambient salinity select for bacteria with a hybrid "salt-in"/"salt-out" osmoregulation strategy

Danny Ionescu[1]*, Luca Zoccarato[2,3], Pedro J. Cabello-Yeves[4,5,6] and Yaron Tikochinski[7]

[1]Plankton and Microbial Ecology, Leibniz Institute of Freshwater Ecology and Inland Fisheries (IGB), Stechlin, Germany, [2]Institute of Computational Biology, University of Natural Resources and Life Sciences 24 (BOKU), Vienna, Austria, [3]Core Facility Bioinformatics, University of Natural Resources and Life Sciences (BOKU), Vienna, Austria, [4]Evolutionary Genomics Group, Departamento de Producción Vegetal y Microbiología, Universidad Miguel Hernández, Alicante, Spain, [5]Cavanilles Institute of Biodiversity and Evolutionary Biology, University of Valencia, Valencia, Spain, [6]School of Life Sciences, University of Warwick, Coventry, United Kingdom, [7]The School of Marine Sciences, The Ruppin Academic Center, Michmoret, Israel

Abundant microbial biofilms inhabit underwater freshwater springs of the Dead Sea. Unlike the harsh (i.e., over 35% total dissolved salts) yet stable environment of the basin, the flow rate of the springs changes with random amplitude and duration, resulting in drastic shifts in salinity, pH, and oxygen concentrations. This requires the organisms to continuously adapt to new environmental conditions. Osmotic regulation is energetically expensive; therefore, the response of the biofilm organisms to rapid and drastic changes in salinity is interesting. For this purpose, we studied the metagenome of an enrichment culture obtained from a green biofilm-covered rock positioned in a spring. We obtained metagenome-assembled genomes (MAGs) of *Prosthecochloris* sp. (*Chlorobiales*), *Flexistipes* sp. (*Deferribacterales*), *Izemoplasma* (*Izemoplasmatales*), *Halomonas* sp. (*Oceanospirillales*), and *Halanaerobium* (*Halanaerobiales*). The MAGs contain genes for both the energetically cheaper "salt-in" and more expensive "salt-out" strategies. We suggest that the dynamic response of these bacteria utilizes both osmoregulation strategies, similar to halophilic archaea. We hypothesize that the frequent, abrupt, and variable-in-intensity shifts in salinity, typical of the Dead Sea spring system, select for microorganisms with scalable adaptation strategies.

KEYWORDS

osmoregulation, Dead Sea, underwater springs, metagenome-assembled genome (MAG), salinity fluctuation

Introduction

Halophilic bacteria and Archaea thrive in a wide array of salinities ranging from seawater to saturated brines. The environments where the sodium chloride (NaCl) concentration is close to saturation and where extreme halophiles thrive are common globally (Atanasova et al., 2013) and harbor relatively low biodiversity. Such environments provide a suitable environment for studying biological, ecological, and evolutionary questions regarding the limits of microbial life on Earth (Ma et al., 2010). Moreover, halophilic bacteria are of special interest for biotechnological industries, which require bioprocessing under high salt concentrations and non-sterile conditions, in an energy- and resource-saving way, for the production of cheaper materials and fuels (Moreno et al., 2016; Chen and Jiang, 2018). Additionally, mechanosensitive channels used for osmoregulation by bacteria are relevant models in the field of mechanobiology, applied to the study of diseases such as cardiopathies or muscular dystrophy (Martinac et al., 2014).

To account for the external osmotic pressure halophile microorganisms have adopted two main strategies. The first, commonly used by hyper-halophiles, was named the "salt-in" strategy as it involves a high intracellular concentration of salts, mainly potassium. The second strategy, "salt-out", is more common among moderate halophiles, as well as marine microbes, and relies on the accumulation of small organic compounds, generally referred to as compatible solutes, e.g., glycine betaine and trehalose (Kunte et al., 2002; Scanlan et al., 2009; Kunte, 2009).

The type of osmoregulatory mechanism used by an organism has various implications. On the one hand, the salt-in strategy is energetically less expensive than the salt-out alternative, in which compatible solutes need to be synthesized or acquired from the environment if available (Oren, 2008). On the other hand, the salt-in strategy requires the presence of an acidic proteome capable of functioning in the presence of molar concentration of intracellular salts (Deole et al., 2013), whereas the salt-out strategy does not require any special adaptations.

Traditionally it was believed that the type of osmoregulation employed by an organism follows a phylogenetic separation by which extreme halophilic Archaea make use of the salt-in strategy (Becker et al., 2014), while moderate and slight halophilic bacteria and eukaryotes use the salt-out alternative (Roessler and Muller, 2001; Oren, 2005). Nevertheless, an increasing number of halophilic bacteria have been shown to utilize the salt-in strategy, the first of which was *Salinibacter ruber* (Oren, 2013a). Subsequently, two sister proteobacterial taxa, *Halorhodospira halochloris* and *Halorhodospira halophila*, were found to utilize different osmoregulation strategies, one using the salt-out and the other the salt-in strategy, respectively (Oren, 2013b). Interestingly, the suggestion that *H. halochloris* does not have an acidic proteome based on isoelectric gel analysis was later disproved from its genome sequence, which showed that the transition of *H. halochloris* to the salt-out strategy may have happened late in the evolution of this genus (Deole et al., 2013).

The Dead Sea is a hypersaline lake with a concentration of ca. 350 g L^{-1} total dissolved salts consisting mainly of 2 M Mg^{2+}, 1.5 M Na^{+}, and 0.5 M Ca^{2+} and a matching ca. 7 M Cl^{-} (Oren, 2010). The microbial life in the Dead Sea has been described in a series of papers in the first half of the 20th century covering both prokaryotes and eukaryotes (Wilkansky, 1936; Elazari-Volcani, 1940; Elazari-Volcani, 1943; Elazari-Volcani, 1944). Unlike other aquatic environments, including hypersaline systems, due to the high concentration of divalent cations and possibly due to the low concentrations of available phosphorus, bacterial concentrations in the lake's water are low, ranging between 10^4 and 10^5 cells mL^{-1} under normal conditions (Ionescu et al., 2012). Occasionally, during strong winter seasons the input of freshwater increases, leading to a dilution of the upper water layer, which in turn results in a bloom of the green alga *Dunaliella* sp. A consequence of this is that the number of prokaryotes in the water column increases as well (Kaplan and Friedmann, 1970; Oren, 1985; Oren and Gurevich, 1995).

Submarine groundwater discharge has been documented in the Dead Sea for over a decade (Ionescu et al., 2012; Mallast et al., 2013). Nevertheless, the effects it has on the biota of the Dead Sea near the water outlets was only recently investigated (Ionescu et al., 2012). Extensive biofilms on the sediments and stones at the freshwater/saltwater interface have been discovered (Figure 1). Interestingly, the water flow in these underwater springs is characterized by extreme fluctuations that occur at a seemingly random amplitude and frequency (Häusler et al., 2014a, Figure S1). The ambient water salinity is directly affected by the freshwater flow from the springs whose velocity dictates the distance from the sediment at which the Dead Sea water mixes in. The higher the flow, the further away from the sediment and the biofilms the mixing is, and *vice versa*. Thus, great fluctuations in the flow result in extreme and rapid fluctuations in the ambient biofilm salinity, as well as other physical–chemical parameters such as O_2 concentration and pH (Figure S1).

In this study we used a metagenomics approach to hypothesize the osmoregulation strategy employed by the biofilm microorganisms living close to the groundwater outlets and to hypothesize whether or not such a strategy allows them to survive these salinity fluctuations. Despite being, on average, exposed to moderate salinities, we consider it unlikely that microorganisms constantly exposed to either freshwater or saturated brine in the freshwater springs of the Dead Sea can be osmotically balanced by exclusively using the salt-out strategy.

Materials and methods

Cultivation

The green biofilm samples were collected by scuba diving during 2011 (Ionescu et al., 2012). The cultures aimed at green sulfur bacteria isolation and enrichment were set using Dead Sea water diluted to 10% of the original salinity. The diluted brine was supplemented with 25 ml L^{-1} of PF-I solution (6.6 g/L of each KCl, $MgCl_2{\cdot}6H_2O$, NH_4Cl, and KH_2PO_4), 2.5 ml L^{-1} of PF-II trace element solution (Per L: 1 g of Na_2-EDTA, 48 mg of $CoSO_4{\cdot}7H_2O$, 400 mg of $FeSO_4{\cdot}7H_2O$, 20 mg of $ZnSO_4{\cdot}7H_2O$, 3 mg of $CuSO_4{\cdot}5H_2O$, 4 mg of $NiCl_2{\cdot}6H_2O$, 6 mg of $Na_2MoO_4{\cdot}2H_2O$,

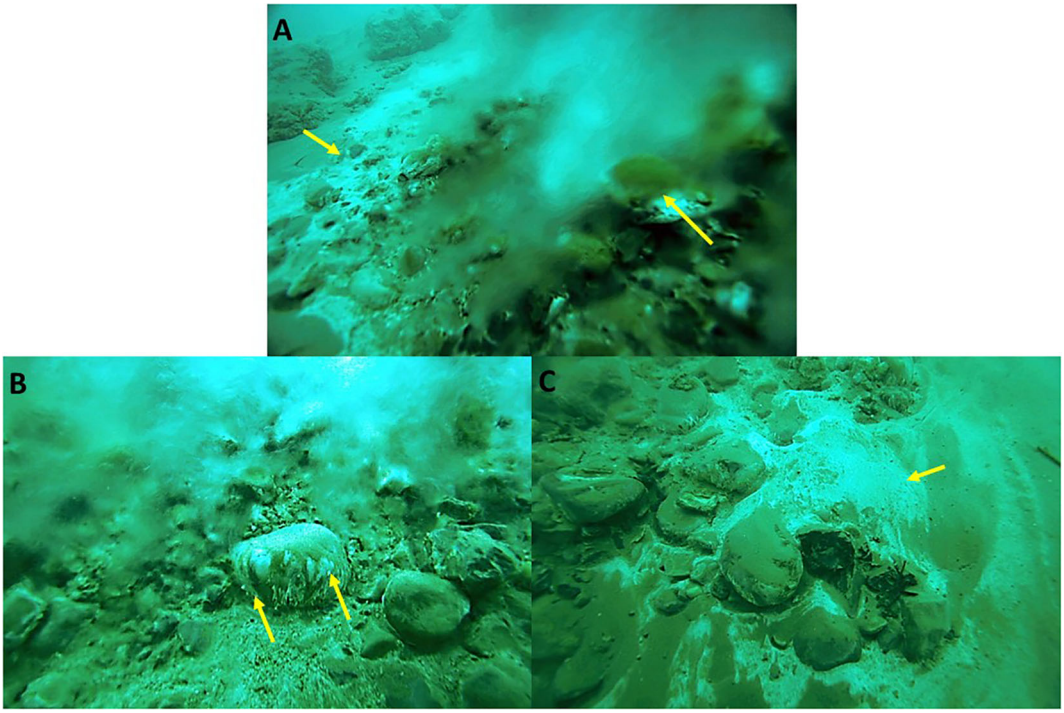


FIGURE 1
White and green biofilms as they appear in the Dead Sea underwater springs (examples are marked by arrows). Often, white biofilms are found at the peripheries of the main outlet of springs, whereas stones covered in green and white biofilms are found in the main flow **(A)**. A close-up of a stone shows that the biofilms are located on the underside of the stone where they benefit from an average lower salinity (Häusler et al., 2014a) **(B)**. A white biofilm on the sediment photographed at a time where no flow could be observed, demonstrating the vast fluctuations in the flow regime **(C)**.

6 mg of $MnCl_2{\cdot}4H_2O$, and 60 mg of H_3BO_3), 2.5 ml L^{-1} of 1% $NaHCO_3$, 2 mM sulfide (using $Na_2S{\cdot}9H_2O$), 1 ml L^{-1} 10% Na-acetate, and 0.25 ml L^{-1} of 0.002% vitamin B_{12}. No growth was observed when artificial Dead Sea water was used as a basis for media instead of the diluted lake water. The cultures were incubated at 30°C in dim light (10 µmol photons m^{-2} s^{-1}).

Nucleic acid extraction and sequencing

The DNA was extracted using a phenol–chloroform-based method, as described by Ionescu et al. (2012) using a basic (ca. pH 8) phenol solution. The DNA samples were sequenced on an Illumina HiSeq 2000 as 2 × 150 bp paired-end reads at MrDNA (Shallowater, TX, USA).

Genome data analysis

All sequences were cleaned from adapter sequences and low-quality bases (Q < 15) using the Trimmomatic software (V 0.32) (Bolger et al., 2014). The DNA sequences were assembled using SPADES (V 3.14) (Bankevich et al., 2012), after which supervised binning was done using MetaBAT2 (Kang et al., 2019). The raw reads from the DNA sample were mapped using BBMap (V 38.0) to the contigs associated with each bin, after which the reads from each bin were assembled and binned individually. The process was repeated until the scaffold length was maximized and the contigs with obvious taxonomic mismatch were no longer present. The bin-refining module of Anvi'o (Eren et al., 2015) was used to refine some of the bins; however, having only one sample, the process led to minimal or no improvement. The scaffolds and the contigs of each bin were annotated separately using the Prokka pipeline (Seemann, 2014), the eggNOG database using eMapper (Huerta-Cepas et al., 2019), and RAST (Aziz et al., 2008; Brettin et al., 2015). The taxonomic annotation of the metagenome-assembled genomes (MAGs) was finally assigned using GTDB-Tk (V.2.3 DB version 207) (Chaumeil et al., 2022). The predicted proteins from each genome were further used for isoelectric point prediction using the IPC 2.0 software (Kozlowski, 2016).The genome-based phylogenetic trees of the obtained genomes and reference organisms were constructed using FastTree (Price et al., 2010) (v. 2.1.11; double precision; default parameters with gamma20 rescaling) from the multilocus alignment of the single-copy marker genes generated during the genome annotation using the GTDB-Tk tool (Chaumeil et al., 2022). In addition, the principal coordinate analysis (PCO) plots comparing SEED protein presence/absence between all the available genomes within the same genera to our retrieved MAGs were constructed. The pangenome comparisons were conducted using Anvi'o (Eren et al., 2015).

Data accessibility

The raw sequence data from the enrichment cultures and the MAGs were deposited into the National Center for Biotechnology Information (NCBI)'s Sequence Read Archive (SRA) (metagenomes) and GenBank® databases under Bioproject PRJNA1033173 (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1033173). The MAG sequences are also available as part of the Supplementary Material of this paper to allow accessibility during NCBI internal processing.

Results

Characterization of the freshwater submarine springs

The freshwater springs in the Dead Sea were investigated and sampled during routine scuba dives between 2009 and 2014. White and green biofilms were observed throughout the sampling period (Figure 1) and various aspects of them have been studied and are described elsewhere (Ionescu et al., 2012; Häusler et al., 2014a; Häusler et al., 2014b; Häusler et al., 2014c).

These biofilms are exposed to great osmotic fluctuations, as changes in the groundwater inflows (Figures S1A, S1AB) are common and occur on a timescale of seconds to minutes, with extreme cases of complete ceasing of the freshwater flow being observed as well. Such fluctuations result in short- and long-term changes in salinity (Figure S1C) and oxygen concentration (Figure S1C) as well as in pH, temperature, and sulfide concentration. The underlying mechanism responsible for the inconsistent flow has not yet been elucidated and may vary between different springs, as the chemical analysis of the outflowing water shows different sources and different flow paths (Ionescu et al., 2012).

Phylogenomic analysis

We obtained six MAGs from the metagenomic library of the studied enrichment culture (Figure 2). These genomes were identified as *Izemoplasma* sp. (Figure 2A), *Prosthecochloris* sp. (Figure 3B), two different species of *Halanaerobiales* (Figure 2C), *Flexistipes* sp. (Figure 2D), and *Halomonas* sp. (Figure 2E). Assembly statistics and completeness are reported in Table 1.

The Dead Sea *Izmoplasma* genome clusters together with two genomes obtained from deep sea cold seeps: *Candidatus Izimaplasma* sp. HR2 from the Pacific Ocean hydrate ridge and *Xianfuyuplasma coldseepsis* from the South China Sea (Figure 2A).

The *Prosthecochloris* sp. from the Dead Sea falls into a cluster with sequences of other members of the same genus (Figure 2B). *Prosthecochloris* sp. strain ZM 2 (*Prosthecochloris ethylica* in GTDB) was isolated from the meromictic Green Cape Lake, *Prosthecochloris* sp. HL130 was isolated from the $MgSO_4$-rich

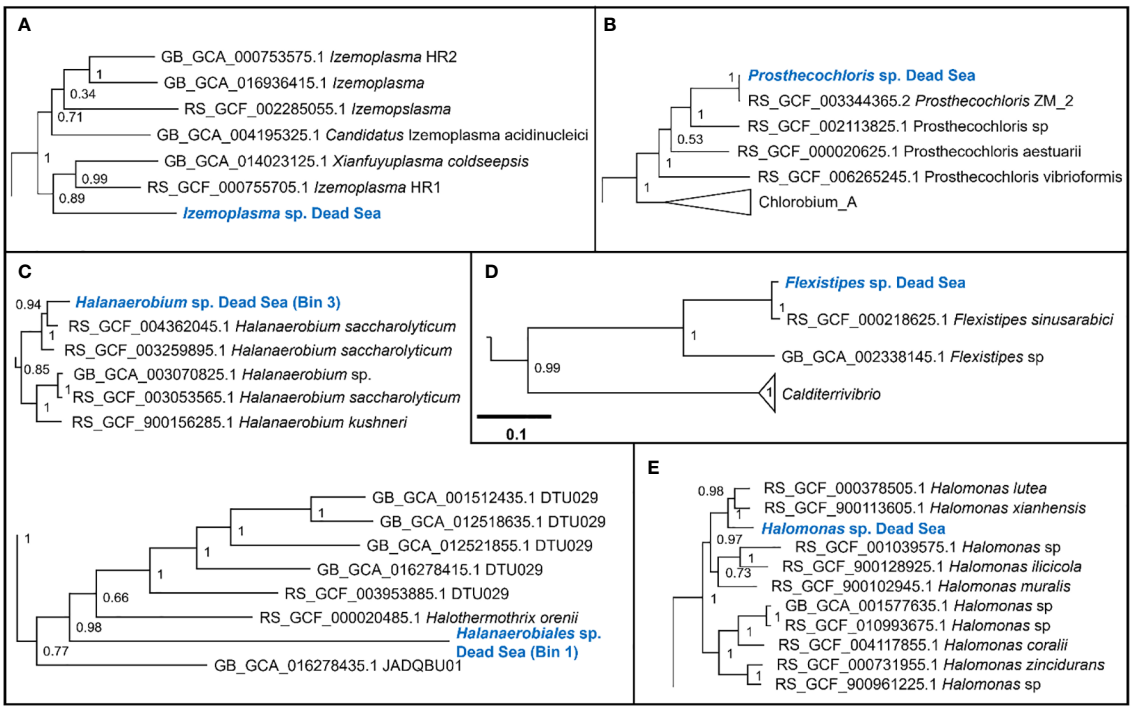


FIGURE 2
Sections of the whole-genome maximum likelihood phylogenetic trees showing the suggested taxonomic identity of the MAGs reconstructed from the Dead Sea cultures: **(A)** *Izemoplasma* sp.; **(B)** *Prosthecochloris* sp.; **(C)** *Halanaerobiales* spp.; **(D)** *Flexistipes* sp.; an **(E)** *Halomonas* sp. The trees were generated based on multisequence alignment, as generated by the GTDB-Tk tool using a bacterial marker set of 120 genes and 42 amino acids per marker (Chaumeil et al., 2022). The numbers next to the branches are the Shimodaira–Hasegawa support values (Shimodaira and Hasegawa, 1999; Guindon et al., 2010). The full trees are available in the Supplementary Figures 2-6.

10.3389/frmbi.2023.1329925

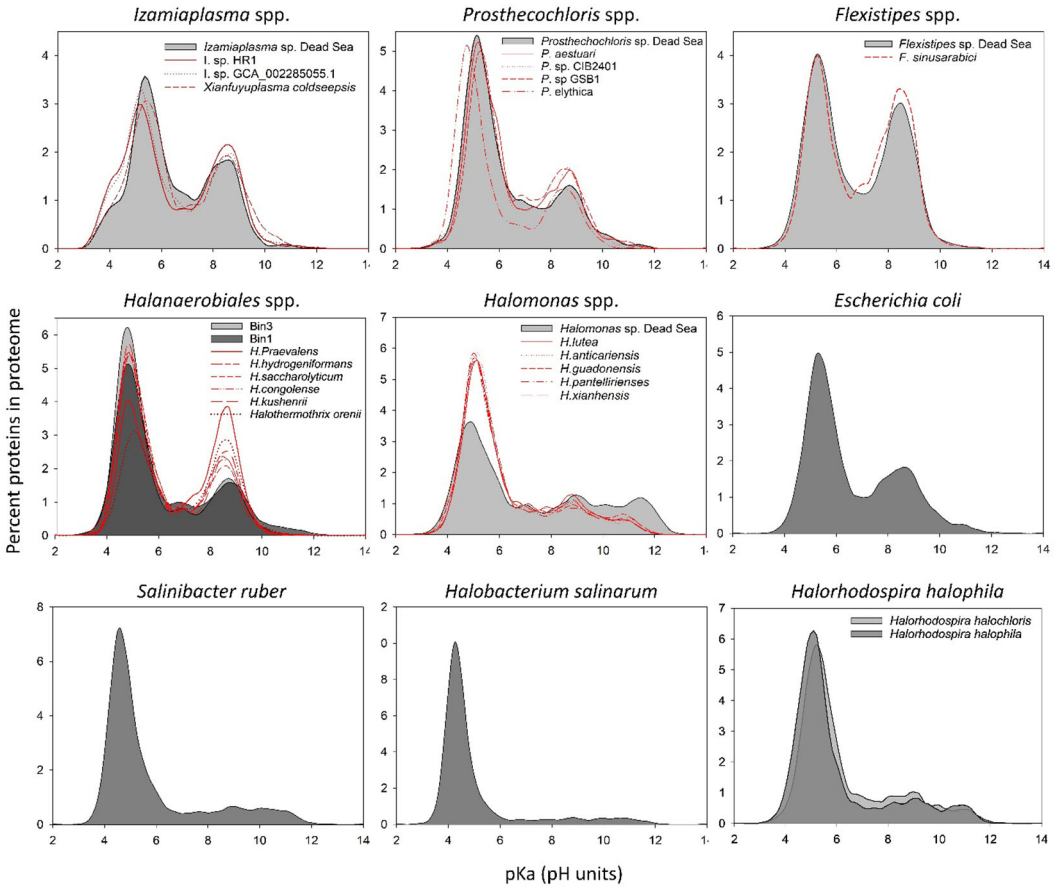


FIGURE 3
In silico whole-proteome isoelectric point profiles of the Dead Sea MAGs compared with several phylogenetically related organisms and model organisms.

hypersaline lake Hot Lake, and *Prosthecochloris aestuarii* was isolated from the saline Lake Sasyk Sivash (Gorlenko et al., 1970). A comparison of 16S rRNA genes extracted from the metagenomic data (not shown) further shows similarity to *Prosthecochloris* sp. L21 and *Prosthecochloris* sp. LA53, which were obtained from environments with drastically different salinities. *Prosthecochloris* sp. L21 was obtained from Lake 21 on the atoll Kiritimati, which has a salinity of ca. 200 ppt (parts per thousand) (Spring et al., 2015). In contrast, *Prosthecochloris* sp. LA53 was obtained from the Hawaiian archipelago, where the salinity is ca. 30 ppt (Donachie et al., 2004).

The Dead Sea *Flexistipes* clusters together with a genome obtained from the Atlantis II deep brine pool in the Red Sea. The next-closest genome is fairly distant and was obtained from a large genome recovery project of the GTDB team, with no information on the environment from which it was recovered.

Within the *Halanaerobiales*, one Dead Sea strain (Bin1) clusters close to two genomes obtained from deep subsurface microbiomes found in extracted gas shales. The second genome (Bin3) has no known close relatives with sequenced genomes. The nearest know genome is that of *Halothermothrix orenii*, which was isolated from the hypersaline lake Chott El Guettar in Tunisia (Cayol et al., 1994).

TABLE 1 Assembly statistics.

Name	Comp.	Contam.	Assembly length	Number of contigs	N50	Max Contig
Prosthecochloris sp.	99.99	0.01	2.33 Mbp	47	84 Kbp	215 Kbp
Flexistipes sp.	99.88	0.05	2.47 Mbp	31	113 Kbp	381 Kbp
Izemoplasma sp.	99.01	0.00	1.68 Mbp	10	517 Kbp	571 Kbp
Halanaerobium sp. (Bin3)	63.22	2.40	1.89 Mbp	192	10 Kbp	31 Kbp
Halanaerobiales sp. (Bin1)	50.02	0.80	1.39 Mbp	157	9 Kbp	29 Kbp
Halomonas sp.	49.76	0.56	2.40 Mbp	417	5 Kbp	28 Kbp

Comp: Completness; Cont: Contamination.

The last genome that was obtained from the Dead Sea biofilm enrichment culture was of a *Halomonas* strain. This genome clustered together with two other *Halomonas* strains obtained from saline environments: *Halomonas lutea* (Wang et al., 2008) and *Halomonas xianhensis* (Zhao et al., 2012). Both *H. lutea* and *H. xuanhensis* were characterized by their ability to grow in an extremely broad salinity range (< 1%–20% salt).

Analysis of isoelectric point profile of predicted proteomes from the metagenomic data

It has been suggested that changes at the level of protein amino acid composition and pIs (isoelectric points) represent a tool to predict the preferred habitat (i.e., salt adapted or salt free) of the different microorganisms (Cabello-Yeves and Rodriguez-Valera, 2019). To evaluate whether and which of the organism genomes that were isolated from the Dead Sea enrichment culture might have a high-salt-in-adapted proteome, we calculated the pIs of each predicted protein in the genomes (Figure 3). The Dead Sea *Prosthecochlois* sp., *Flexistipes* sp., and *Izemoplasma* sp. genomes were on par with data calculated from other available genomes. The pI profile generated from the *Prosthecochloris* genome is similar to that of non-halophilic organisms, such as *Escherichia coli*, and is slightly less acidic than its closest relative *P. ethylica*. In comparison, the Dead Sea *Flexistipes* sp. has more basic proteins than any of the other five genomes obtained. Both *Halanaerobium* sp. genomes have more acidic pI profiles than the related reference genomes. *Halomonas* reference genomes show that this genus is characterized by an overall acidic proteome, but it has more proteins with a high pI (i.e., > 10) than other organisms evaluated in this study. Nevertheless, the Dead Sea *Halomonas* genome appears less acidic than the references and contains more proteins with a pI above a pH of 10 (Figure 3). None of the bacteria enriched from the Dead Sea springs in this study has a comparable acidic proteome to those of *Halobacterium salinarum* and *Salinibacter ruber*, an archaeon and bacterium, respectively (Figure 3). *Halorhodospira halophila* and *H. halochloris* are two sister taxa that were reported to use the salt-in and salt-out strategies for osmoregulation, respectively (Deole et al., 2013; Oren, 2013a). Prior to genome sequencing, *H. halochloris* was thought not to have an acidic proteome, yet our analysis of the genome published by Singh et al. (Singh et al., 2014) shows otherwise (Figure 3).

Genomic comparison

An overall genome comparison of the Dead Sea MAGs and their closest relatives (Figure 4) reveals that based on the COGs (cluster of orthologous genes), the Dead Sea organisms share a large common core genome (See also Table S2-S4). A multi-environment SEED protein presence/absence PCO analysis for each taxon compared with their closest relatives to assess the uniqueness of Dead Sea MAGs compared with other habitats further confirms this (Figure S1). The smallest set of unique genes not shared with closely related genomes was observed in *Prostechochloris* sp. (141 gene clusters) (Figure 4B) and the largest in *Izamiaplasma* sp. (592 gene clusters) (Figure 4A), with *Flexistipes* harboring 216 unique gene clusters with 23, 28, and 63 unique COG functions for *Prosthecochloris* sp., *Flexistipes* sp., and *Izamiaplasma* sp., respectively. However, when inspecting functions there are fewer unique functions harbored by the Dead Sea organisms. It is important to state that, although the calculated completeness of the three MAGs presented in Figure 5 is above 99%, the lack of some functions may still be a result of an incomplete genome. For *Prosthecochloris*, no unique pathway was observed, yet it appears to be entirely missing the F-type ATPase (Table S1). Compared with its nearest relatives, the Dead Sea *Izemoplasma* sp. harbors the pathway to produce riboflavin (vitamin B_2), as well as the DAP (diaminopimelic acid) and the acetyl-DAP pathways for lysine synthesis (Table S1). The Dead Sea *Izemoplasma* sp. lacks several pathways when compared with some of its closest relatives; however, these are not exclusively present in all other genomes. When compared with the genomes of five different strains of *F. sinusarabici*, the Dead Sea *Flexistipes* sp. was found to lack the pathway for assimilatory sulfate reduction and biotin synthesis. In contrast, it was found to possess the commamox (complete oxidation of ammonia to nitrate) pathway, and almost exclusively possess the pathway for methionine degradation (Table S1).

In addition to comparing the presence and absence of KEGG (Kyoto Encyclopedia of Genes and Genomes) enzymes and modules, we compared the genomes for genes potentially involved in osmoregulation. Furthermore, while for a whole genome comparison we investigated only genomes with a high completeness, for the following analysis we investigated all six bins.

The Dead Sea *Izemoplasma* sp. genome, as well as other close relatives, were unique among the other obtained genomes in containing only glycine betaine transporters when considering the "salt-out" osmoregulation strategy (Figure 5). No significant dissimilarity between the Dead Sea genome and its closest relatives was observed with regard to osmoregulation genes.

Similarly to the *Izemoplasma* sp. genomes, the Dead Sea *Prosthecochloris* genome showed no remarkable differences to its closest relatives (Figure 5) with all species containing both genes for K^+ uptake as well as synthesis and uptake of different compatible solutes. The genome comparison suggests *Prosthecochloris* sp. can utilize several types of compatible solutes but not Ectoine, while only the Dead Sea isolate appears to be capable of taking up proline via a sodium/Proline symporter.

Two different *Halanaerobium* genomes were obtained (Figure 2). Osmoregulation genes differ between the two genomes themselves as well as between the Dead Sea isolates and available genomes of other *Halanaerobium* sp. Both Dead Sea isolates contain more, and novel, cation transport genes as than the reference genomes, with the *Halanaerobium* Bin 3 genome being richer in cation transport genes than the *Halanaerobium* Bin 1. The *Halanaerobium* Bin 3 genome also contains a higher number of genes related to compatible solutes than *Halanaerobium* Bin 1 and most references. Among the compared *Halanaerobiales*, the Dead Sea *Halanaerobium* Bin 1 appears to be the only one able to synthesize trehalose rather than take it up. Nevertheless, missing

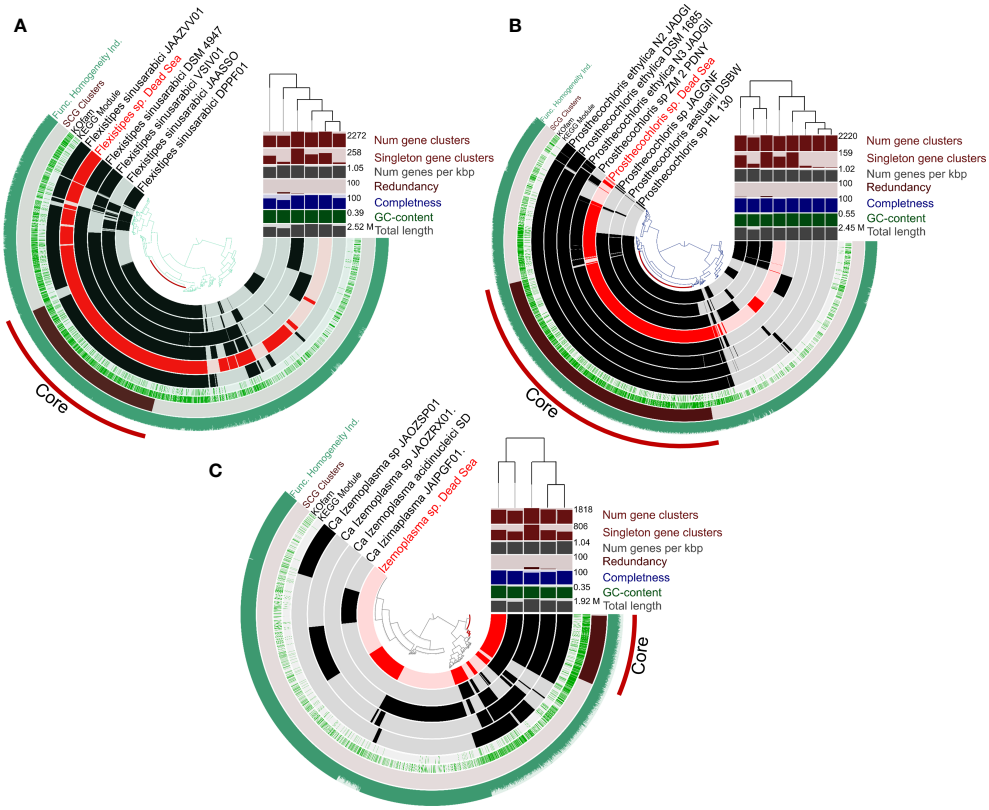


FIGURE 4
Circular representation of pangenomes of MAGs obtained from the Dead Sea springs enrichment culture and the most-closely available genomes in the database. Panels **(A-C)** show *Flexistipes* sp., *Prosthecochloris* sp., and *Isemoplasma* sp., respectively. The genomes of *Halanaerobiales* and *Halomonas* are not shown due to their relatively low level of completeness. The figures were generated using Anvi'o (Eren et al., 2015; Eren et al., 2020).

genes may be a result of the relatively low completeness of these MAGs.

Comparing the Dead Sea *Flexistipes* genome to *F. sinusarabici* from a deep Red Sea brine, most of the differences are between genes related to compatible solutes (Figure 4B). Both *Flexistipes* sp. can synthesize ectoine, the most common compatible solute (Shivanand and Mugeraya, 2011), yet the Dead Sea isolate harbors additional genes for uptake of glycine betaine, proline, and other osmoprotectants. The complete absence of genes related to trehalose synthesis or metabolism suggests this genus does not utilize trehalose as a compatible solute.

The Dead Sea *Halomonas* isolate is enriched in cation transport genes as compared with the reference genomes (Figure 4A), and also harbors more diverse genes for compatible solute transport. In contrast, it has fewer copies of some of the genes for trehalose and glycine betaine transport. Like the other *Halomonas* it can synthesize ectoine but it can take up more osmoprotectants than most of the other *Halomonas*.

Discussion

Halophilic organisms do not normally experience rapid, random, and drastic changes in salinity. However, the large biomass that is present in the Dead Sea underwater springs evidently shows that these organisms can cope with such atypical changes in salinity. In the case of Dead Sea underwater freshwater springs, it is evident by the large biomass present in the springs that these organisms cope with it. Additionally, previous studies have shown that these microbial biofilms are active. Sulfide oxidation was shown to take place in white biofilms (Häusler et al., 2014c) and oxygenic photosynthesis in brown and green biofilms found on sediments and rocks (Häusler et al., 2014b). In the latter case, a clear separation between diatoms and cyanobacterial biofilms was observed and was attributed to the organisms' ability to withstand rapid versus long-term changes in salinity. Diatoms were shown to respond better to short-term, rapid fluctuations while the cyanobacteria are more tolerant to long exposures to undiluted Dead Sea water (Häusler et al., 2014b).

Using an enrichment culture targeting a green sulfur bacterium previously identified in the springs as *Prosthecochloris* sp. (Ionescu et al., 2012), we looked into the genomic potential for osmoregulation of several co-occurring bacteria to elucidate whether they adopt the salt-in or salt-out strategy.

We propose that the organisms inhabiting the Dead Sea springs do not exclusively use the salt-in strategy. The salt-in approach in which osmoregulation is handled via the uptake and release of K^+ ions is energetically less expensive (Oren, 1999; Oren, 2011).

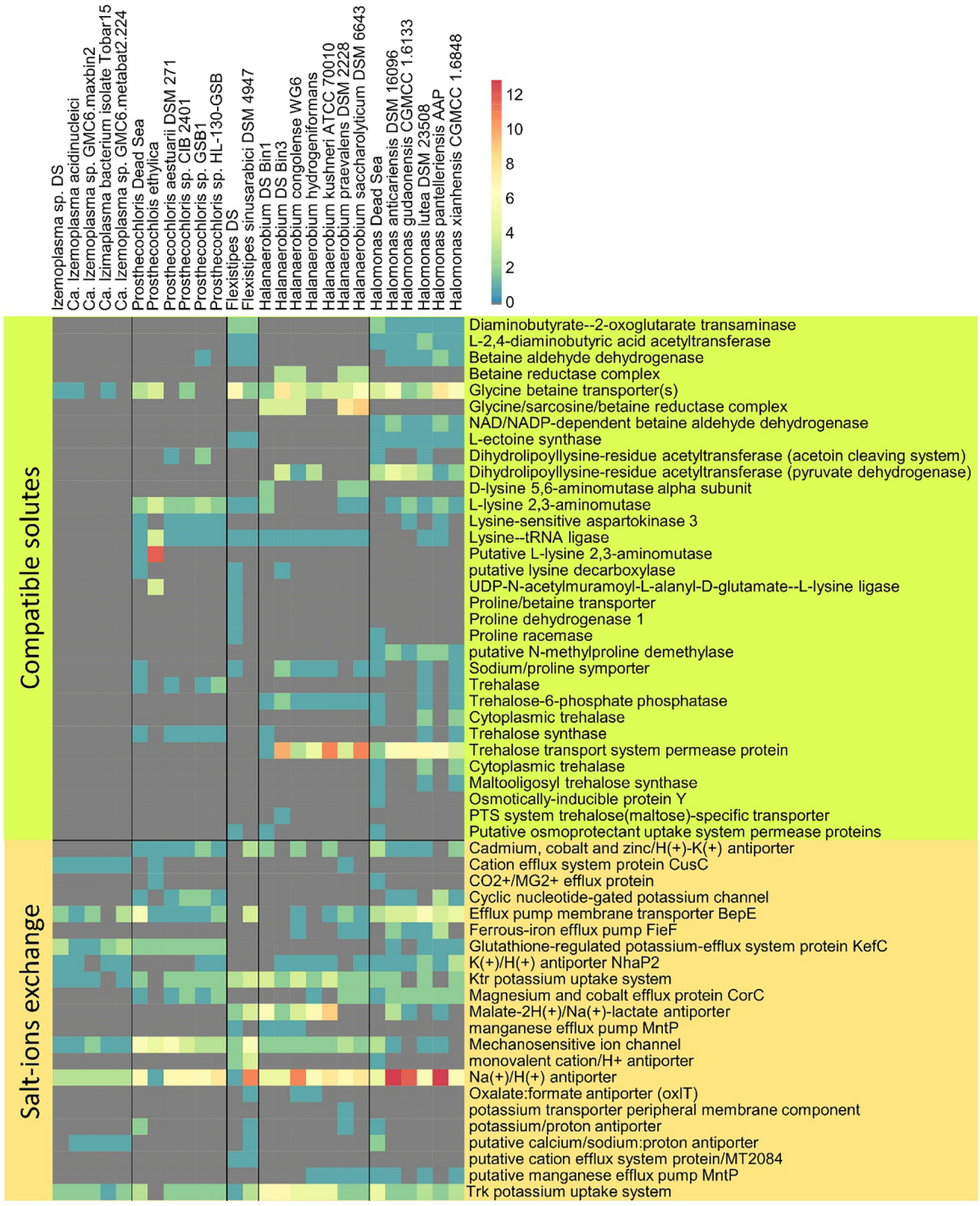


FIGURE 5
Osmotic response genes distribution across reference genomes and MAGs. The data are organized into compatible solute-related genes and genes involved in ion transport. The color scale indicates the number of genes identified per genome, related to a particular function (gray = 0).

Furthermore, the salt-in approach is expected to offer a more rapid response than the salt-out approach, as there is no need to synthesize or take up different organic compounds such as ectoine, glycine betaine, or trehalose. However, when compared with extreme halophiles, the proteome isoelectric point histograms of the Dead Sea MAGs did not resemble the extreme "salt-in"-using halophiles. Except for the two *Halanaerobiales* bins that appeared to harbor a more acidic proteome than their closest sequenced relatives, all bins had a similar profile to their phylogenetically close relatives. *Halorhodospira halophila* and *H. halochloris* are phylogenetically closely related or sister taxa. The former uses the salt-in strategy, whereas the latter uses the salt-out approach. Deole et al. (2013) have shown that *H. halophila* can withstand low-intracellular-K^+ concentrations despite having an acidic proteome,

thus showing that high-intracellular-salt concentrations are not essential for an acidic proteome to function. The acidic proteome of the sister taxa *H. halochloris*, which uses the salt-out strategy, further supports these conclusions and reveals that bacteria using the salt-out strategy may also have acidic proteomes. Nevertheless, to date, no salt-in utilizer has been identified with a non-acidic proteome. Therefore, it is unlikely that the Dead Sea spring organisms discussed here rely exclusively on this strategy.

The mechanistic of static versus dynamic osmoregulation in halophiles was seldom discussed, and when it was it particularly addressed Archaea (Becker et al., 2014). The last study suggested that halophilic Archaea utilizes both salt-in and salt-out strategies under fluctuating salinities, yet maintains an acidic proteome, as suited for their typical strategy. Our genome comparison analysis revealed that the isolates from the Dead Sea springs do not differ significantly from their closest relatives, and especially not with regard to known genes involved in osmoregulation. Overall, all the organisms appear to be able to make use of one (*Izemiplasma* sp.) or more compatible solutes, in parallel to the transport of ions, specifically K^+. Interestingly, all genomes contain multiple mechanosensitive channels. These likely play a crucial role during sudden drops in salinity (Martinac, 2001; Blount and Iscla, 2020). In such events, the high intracellular concentration of salts and or solutes may lead to the rapid intake of water and eventually to cell burst. Thus, harboring multiple, and in some cases different, mechanosensitive channels can aid in the rapid release of solutes to prevent irreversible damage to the cells (Martinac, 2001; Bremer and Krämer, 2019; Blount and Iscla, 2020). During rapid increases in salinity, in parallel to the immediate uptake of K^+ ions through one or more systems, all cells can start synthesizing solutes to increase the intracellular osmolarity while maintaining an overall low intracellular salt concentration (Bremer and Krämer, 2019).

The salinity optimum of the Dead Sea enrichment culture was determined based on the growth of the green sulfur bacterium *Prosthecochloris* at ca. 10% of Dead Sea water salinity. It may be that other members in the culture have a higher salinity optimum, as *Halanaerobium*, *Halomonas*, and *Flexistipes* are halophilic genera. Nevertheless, while both *Halanaerobium* genomes exhibit a more acidic proteome than other members of their genus, none of the bacteria studied here has a proteome as acidic as the salt-in users *Salinibacter ruber* or *Halobacterium salinarum*. The abundance of these organisms in the spring biofilms, alongside the monitored fluctuations in spring flow and salinity provide, strong evidence that these bacteria were coping well with the fluctuating environment.

Thus, we extend the conclusions presented by Becker et al. (2014) to include bacteria and suggest that organisms facing a dynamic osmo-environment make use of both the salt-in and salt-out strategies. At the same time, the proteome acidity of these organisms remains most compatible with the osmotic conditions most frequently encountered. This is further supported by the halophilic *Halorhodosphira halochloris*, which, despite losing the ability to accumulate high concentrations of potassium, maintains an acidic proteome typically suitable for the high-salinity, salt-in strategy.

The Dead Sea underwater spring biofilms provide a unique setting for the study of osmoregulation in bacteria, whereby these are subject to rapid and extreme changes in salinity in a natural way. Our study shows that in moderately halophilic, hyper-halotolerant bacteria, the "salt-in" and "salt-out" strategies are not mutually exclusive, as generally accepted. The response of microorganisms to extreme physical changes in such short periods in this ecosystem provides an opportunity for the real-time study of bacterial evolution on conditions likely similar to those during early life. Moreover, it is a source of information, physiological and genetic, for those looking to obtain strains that could adapt to such changing conditions in biotechnology.

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

DI: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Project administration, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. LZ: Data curation, Formal Analysis, Writing – original draft, Writing – review & editing. PC-Y: Data curation, Formal Analysis, Writing – review & editing. YT: Investigation, Resources, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. DI was supported by DFG project IO 98/3–1. PC-Y work was funded by a APOSTD/2019/009 Post-Doctoral Fellowship from Generalitat Valenciana and is currently supported by a Post-Doctoral Fellowship from the Fundación Alfonso Martín Escudero, Spain. The study was partly funded by the Max Planck Society.

Acknowledgments

We thank Neta Raab and Maria Arias for their help with some of the experimental and analytical aspects of this study.

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 author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

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.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frmbi.2023.1329925/full#supplementary-material

References

Atanasova, N. S., Pietilä, M. K., and Oksanen, H. M. (2013). Diverse antimicrobial interactions of halophilic archaea and bacteria extend over geographical distances and cross the domain barrier. *Microbiologyopen* 2, 811–825. doi: 10.1002/mbo3.115

Aziz, R. K., Bartels, D., Best, A. A., DeJongh, M., Disz, T., Edwards, R. A., et al. (2008). The RAST Server: rapid annotations using subsystems technology. *BMC Genomics* 9, 75. doi: 10.1186/1471-2164-9-75

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., et al. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 19, 455–477. doi: 10.1089/cmb.2012.0021

Becker, E. A., Seitzer, P. M., Tritt, A., Larsen, D., Krusor, M., Yao, A. I., et al. (2014). Phylogenetically driven sequencing of extremely halophilic archaea reveals strategies for static and dynamic osmo-response. *PloS Genet.* 10, e1004784. doi: 10.1371/journal.pgen.1004784

Blount, P., and Iscla, I. (2020). Life with bacterial mechanosensitive channels, from discovery to physiology to pharmacological target. *Microbiol. Mol. Biol. Rev.* 84, e00055–19. doi: 10.1128/MMBR.00055-19

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

Bremer, E., and Krämer, R. (2019). Responses of microorganisms to osmotic stress. *Annu. Rev. Microbiol.* 73, 313–334. doi: 10.1146/annurev-micro-020518-115504

Brettin, T., Davis, J. J., Disz, T., Edwards, R. A., Gerdes, S., Olsen, G. J., et al. (2015). RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Sci. Rep.* 5, 8365. doi: 10.1038/srep08365

Cabello-Yeves, P. J., and Rodriguez-Valera, F. (2019). Marine-freshwater prokaryotic transitions require extensive changes in the predicted proteome. *Microbiome* 7, 117. doi: 10.1186/s40168-019-0731-5

Cayol, J.-L., Olliver, B., Patel, B. K. C., Prensier, G., Gezennec, J., and Garcia, J.-L. (1994). Isolation and characterization of *Halothermothrix orenii* gen. nov., sp. nov., a calophilic, thermophilic, fermentative, strictly anaerobic bacterium. *Int. J. Syst. Bacteriol.* 44, 534–540. doi: 10.1099/00207713-44-3-534

Chaumeil, P.-A., Mussig, A. J., Hugenholtz, P., and Parks, D. H. (2022). GTDB-Tk v2: memory friendly classification with the genome taxonomy database. *Bioinformatics* 38, 5315–5316. doi: 10.1093/bioinformatics/btac672

Chen, G.-Q., and Jiang, X.-R. (2018). Next generation industrial biotechnology based on extremophilic bacteria. *Curr. Opin. Biotechnol.* 50, 94–100. doi: 10.1016/j.copbio.2017.11.016

Deole, R., Challacombe, J., Raiford, D. W., and Hoff, W. D. (2013). An extremely halophilic proteobacterium combines a highly acidic proteome with a low cytoplasmic potassium content. *J. Biol. Chem.* 288, 581–588. doi: 10.1074/jbc.M112.420505

Donachie, S. P., Hou, S., Lee, K. S., Riley, C. W., Pikina, A., Belisle, C., et al. (2004). The hawaiian archipelago: A microbial diversity hotspot. *Microb. Ecol.* 48, 509–520. doi: 10.1007/s00248-004-0217-1

Elazari-Volcani, B. (1940). Algae in the bed of the dead sea. *Nature* 145, 975–975. doi: 10.1038/145975a0

Elazari-Volcani, B. (1943). Bacteria in the bottom sediments of the Dead Sea. *Nature* 152, 301–302. doi: 10.1038/152301a0

Elazari-Volcani, B. (1944). A ciliate from the dead sea. *Nature* 154, 335–335. doi: 10.1038/154335a0

Eren, A. M., Esen, Ö.C., Quince, C., Vineis, J. H., Morrison, H. G., Sogin, M. L., et al. (2015). Anvi'o: an advanced analysis and visualization platform for 'omics data. *PeerJ* 3, e1319. doi: 10.7717/peerj.1319

Eren, A. M., Kiefl, E., Shaiber, A., Veseli, I., Miller, S. E., Schechter, M. S., et al. (2020). Community-led, integrated, reproducible multi-omics with anvi'o. *Nat. Microbiol.* 6, 3–6. doi: 10.1038/s41564-020-00834-3

Gorlenko, V. M. (1970). A new phototrophic green sulphur bacterium. Prosthecochloris aestuarii nov. gen. nov. spec. *Z Allg Mikrobiol* 10, 147–149.

Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of phyML 3.0. *Syst. Biol.* 59, 307–321. doi: 10.1093/sysbio/syq010

Häusler, S., Noriega-Ortega, B. E., Polerecky, L., Meyer, V., de Beer, D., and Ionescu, D. (2014a). Microenvironments of reduced salinity harbour biofilms in Dead Sea underwater springs. *Environ. Microbiol. Rep.* 6, 152–158. doi: 10.1111/1758-2229.12140

Häusler, S., Weber, M., de Beer, D., and Ionescu, D. (2014b). Spatial distribution of diatom and cyanobacterial mats in the Dead Sea is determined by response to rapid salinity fluctuations. *Extremophiles* 18, 1085–1094. doi: 10.1007/s00792-014-0686-1

Häusler, S., Weber, M., Siebert, C., Holtappels, M., Noriega-Ortega, B. E., De Beer, D., et al. (2014c). Sulfate reduction and sulfide oxidation in extremely steep salinity gradients formed by freshwater springs emerging into the Dead Sea. *FEMS Microbiol. Ecol.* 90, 956–969. doi: 10.1111/1574-6941.12449

Huerta-Cepas, J., Szklarczyk, D., Heller, D., Hernández-Plaza, A., Forslund, S. K., Cook, H., et al. (2019). eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res.* 47, D309–D314. doi: 10.1093/nar/gky1085

Ionescu, D., Siebert, C., Polerecky, L., Munwes, Y. Y., Lott, C., Häusler, S., et al. (2012). Microbial and chemical characterization of underwater fresh water springs in the dead sea. *PloS One* 7, e38319. doi: 10.1371/journal.pone.0038319

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

Kaplan, I. R., and Friedmann, A. (1970). Biological productivity in the dead sea part I. Microorganisms in the water column. *Isr. J. Chem.* 8, 513–528. doi: 10.1002/ijch.197000058

Kozlowski, L. P. (2016). IPC – isoelectric point calculator. *Biol. Direct* 11, 55. doi: 10.1186/s13062-016-0159-9

Kunte, H. J. (2009). Osmoregulation in halophilic bacteria. in C. Gerday and N. Glansdorff Eds. *Encyclopedia of life support sciences - Extremophiles Vol II, United Nations*, 263–277.

Kunte, H. J., Trüper, H. G., and Stan-Lotter, H. (2002). Halophilic Microorganisms In: G. Horneck and C. Baumstark-Khan (eds) (Berlin, Heidelberg: Springer). doi: 10.1007/978-3-642-59381-9_13

Ma, Y., Galinski, E. A., Grant, W. D., Oren, A., and Ventosa, A. (2010). Halophiles 2010: life in saline environments. *Appl. Environ. Microbiol.* 76, 6971–6981. doi: 10.1128/AEM.01868-10

Mallast, U., Siebert, C., Wagner, B., Sauter, M., Gloaguen, R., Geyer, S., et al. (2013). Localisation and temporal variability of groundwater discharge into the Dead Sea using thermal satellite data. *Environ. Earth Sci.* 69, 587–603. doi: 10.1007/s12665-013-2371-6

Martinac, B. (2001). Mechanosensitive channels in prokaryotes. *Cell. Physiol. Biochem.* 11, 61–76. doi: 10.1159/000047793

Martinac, B., Nomura, T., Chi, G., Petrov, E., Rohde, P. R., Battle, A. R., et al. (2014). Bacterial mechanosensitive channels: models for studying mechanosensory transduction. *Antioxid. Redox Signal* 20, 952–969. doi: 10.1089/ars.2013.5471

Moreno, M., de, L., Márquez, M. C., García, M. T., and Mellado, E. (2016). "Halophilic Bacteria and Archaea as Producers of Lipolytic Enzymes BT - Biotechnology of Extremophiles: Advances and Challenges,". Ed. P. H. Rampelotto (Cham: Springer International Publishing), 375–397. doi: 10.1007/978-3-319-13521-2_13

Oren, A. (1985). The rise and decline of a bloom of halobacteria in the Dead Sea. *Limnol. Oceanogr.* 30, 911–915. doi: 10.4319/lo.1985.30.4.0911

Oren, A. (1999). Bioenergetic aspects of halophilism. *Microbiol. Mol. Biol. Rev.* 63, 334–348. doi: 10.1128/MMBR.63.2.334-348.1999

Oren, A. (2005). A hundred years of Dunaliella research: 1905–2005. *Saline Syst.* 1, 2. doi: 10.1186/1746-1448-1-2

Oren, A. (2008). Microbial life at high salt concentrations: phylogenetic and metabolic diversity. *Saline Syst.* 4, 2. doi: 10.1186/1746-1448-4-2

Oren, A. (2010). The dying Dead Sea: The microbiology of an increasingly extreme environment. *Lakes Reserv* 15, 215–222. doi: 10.1111/j.1440-1770.2010.00435.x

Oren, A. (2011). Thermodynamic limits to microbial life at high salt concentrations. *Environ. Microbiol.* 13, 1908–1923. doi: 10.1111/j.1462-2920.2010.02365.x

Oren, (2013a). "Life at High Salt Concentrations," in *The Prokaryotes*. Eds. E. Rosenberg, E. F. DeLong, S. Lory, E. Stackebrandt and T. Fabiano (Berlin, Heidelberg: Springer), 421–440. doi: 10.1007/978-3-642-30123-0_57

Oren, A. (2013b). Life at high salt concentrations, intracellular KCl concentrations, and acidic proteomes. *Front. Microbiol.* 4. doi: 10.3389/fmicb.2013.00315

Oren, A., and Gurevich, P. (1995). Dynamics of a bloom of halophilic archaea in the Dead Sea. *Hydrobiologia* 315, 149–158. doi: 10.1007/BF00033627

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

Roessler, M., and Muller, V. (2001). Osmoadaptation in bacteria and archaea: common principles and differences. *Environ. Microbiol.* 3, 743–754. doi: 10.1046/j.1462-2920.2001.00252.x

Scanlan, D. J., Ostrowski, M., Mazard, S., Dufresne, A., Garczarek, L., Hess, W. R., et al. (2009). Ecological genomics of marine picocyanobacteria. *Microbiol. Mol. Biol. Rev.* 73, 249–299. doi: 10.1128/MMBR.00035-08

Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30(14), 2068–2069. doi: 10.1093/bioinformatics/btu153

Shimodaira, H., and Hasegawa, M. (1999). Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Mol. Biol. Evol.* 16, 1114–1116. doi: 10.1093/oxfordjournals.molbev.a026201

Shivanand, P., and Mugeraya, G. (2011). Halophilic bacteria and their compatible solutes-osmoregulation and potential applications. *Curr. Sci.* 100, 1516–1521.

Singh, K. S., Kirksey, J., Hoff, W. D., and Deole, R. (2014). Draft genome sequence of the extremely halophilic phototrophic purple sulfur bacterium halorhodospira halochloris. *J. Genomics* 2, 118–120. doi: 10.7150/jgen.9123

Spring, S., Brinkmann, N., Murrja, M., Spröer, C., Reitner, J., and Klenk, H.-P. (2015). High diversity of culturable prokaryotes in a lithifying hypersaline microbial mat. *Geomicrobiol. J.* 32, 332–346. doi: 10.1080/01490451.2014.913095

Wang, Y., Tang, S.-K., Lou, K., Mao, P.-H., Jin, X., Jiang, C.-L., et al. (2008). Halomonas lutea sp. nov., a moderately halophilic bacterium isolated from a salt lake. *Int. J. Syst. Evol. Microbiol.* 58, 2065–2069. doi: 10.1099/ijs.0.65436-0

Wilkansky, B. (1936). Life in the dead sea. *Nature* 138, 467–467. doi: 10.1038/138467a0

Zhao, B., Wang, H., Mao, X., Li, R., Zhang, Y.-J., Tang, S., et al. (2012). Halomonas xianhensis sp. nov., a moderately halophilic bacterium isolated from a saline soil contaminated with crude oil. *Int. J. Syst. Evol. Microbiol.* 62, 173–178. doi: 10.1099/ijs.0.025627-0

frontiers | Frontiers in Microbiomes

TYPE Original Research
PUBLISHED 11 January 2024
DOI 10.3389/frmbi.2023.1280809

OPEN ACCESS

EDITED BY
Angela Kent,
University of Illinois at Urbana-Champaign, United Sates

REVIEWED BY
Rey Mourot,
GFZ German Research Centre for Geosciences Potsdam, Germany, in collaboration with reviewer CK
Christoph Keuschnig,
GFZ German Research Centre for Geosciences, Germany
Nathan A. M. Chrismas,
Marine Biological Association of the United Kingdom, United Kingdom

*CORRESPONDENCE
Tom J. Battin
tom.battin@epfl.ch
Susheel Bhanu Busi
susbus@ceh.ac.uk

†These authors have contributed equally to this work

RECEIVED 21 August 2023
ACCEPTED 22 December 2023
PUBLISHED 11 January 2024

CITATION
Busi SB, Peter H, Brandani J, Kohler TJ, Fodelianakis S, Prameteftaki P, Bourquin M, Michoud G, Ezzat L, Lane S, Wilmes P and Battin TJ (2024) Cross-domain interactions confer stability to benthic biofilms in proglacial streams.
Front. Microbiomes 2:1280809.
doi: 10.3389/frmbi.2023.1280809

Cross-domain interactions confer stability to benthic biofilms in proglacial streams

Susheel Bhanu Busi[1,2*†], Hannes Peter[3†], Jade Brandani[3], Tyler J. Kohler[3,4], Stilianos Fodelianakis[3], Paraskevi Prameteftaki[3], Massimo Bourquin[3], Grégoire Michoud[3], Leïla Ezzat[3,5], Stuart Lane[6], Paul Wilmes[2,7] and Tom J. Battin[3*]

[1]UK Centre for Ecology & Hydrology (UKCEH), Wallingford, Oxfordshire, United Kingdom, [2]Systems Ecology Group, Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Esch-sur-Alzette, Luxembourg, [3]River Ecosystems Laboratory, Alpine and Polar Environmental Research Center, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland, [4]Department of Ecology, Faculty of Science, Charles University, Prague, Czechia, [5]MARBEC, Université de Montpellier, CNRS, Ifremer, IRD, Montpellier, France, [6]Institute of Earth Surface Dynamics (IDYST), University of Lausanne, Lausanne, Switzerland, [7]Department of Life Sciences and Medicine, Faculty of Science, Technology and Medicine, University of Luxembourg, Esch-sur-Alzette, Luxembourg

Cross-domain interactions are an integral part of the success of biofilms in natural environments but remain poorly understood. Here, we describe cross-domain interactions in stream biofilms draining proglacial floodplains in the Swiss Alps. These streams, as a consequence of the retreat of glaciers, are characterised by multiple environmental gradients and perturbations (e.g., changes in channel geomorphology, discharge) that depend on the time since deglaciation. We evaluate co-occurrence of bacteria and eukaryotic communities along streams and show that key community members have disproportionate effects on the stability of community networks. The topology of the networks, here quantified as the arrangement of the constituent nodes formed by specific taxa, was independent of stream type and their apparent environmental stability. However, network stability against fragmentation was higher in the streams draining proglacial terrain that was more recently deglaciated. We find that bacteria, eukaryotic photoautotrophs, and fungi are central to the stability of these networks, which fragment upon the removal of both pro- and eukaryotic taxa. Key taxa are not always abundant, suggesting an underlying functional component to their contributions. Thus, we show that there is a key role played by individual taxa in determining microbial community stability of glacier-fed streams.

KEYWORDS

glacier-fed streams, cross-domain interactions, networks, community fragmentation, microbiome

Introduction

Biofilms represent the dominant microbial lifestyle in streams and rivers (Battin et al., 2016). These matrix-enclosed microbial communities colonise sediment surfaces, alongside the surface of benthic rocks and plants, and regulate critical ecosystem processes (Battin et al., 2016). Stream biofilm communities are highly diverse, harbouring members from all domains of life, including prokaryotes and various microeukaryotes. This biodiversity enclosed within streams fosters interactions, such as those between bacterial heterotrophs, algae, and fungi. The bacterial and eukaryotic photoautotrophic interactions, the latter of which includes algae, are likely due to metabolic exchanges, whereby algal exudates serve as a source of organic matter for bacterial heterotrophs (Kaplan and Bott, 1989; Wagner et al., 2017). The fungi, meanwhile, may play important roles where fungal metabolites are utilised by distinct bacterial taxa such as Burkholderia (Stopnisek et al., 2016). Furthermore, they are also capable of parasitizing algae, in the case of Chytridiomycetes, to release carbon for bacterial utilisation (Klawonn et al., 2021).

In plants and animals, elevated biodiversity and related biotic interactions can contribute to community stability (McCann, 2000). Yet, studying biotic interactions within complex microbial communities is not a trivial task. For instance, signalling molecules have been suggested as a proxy for identifying interactions among microorganisms (Braga et al., 2016). However, such approaches have remained largely limited to model systems (Vetsigian et al., 2011; Lee and Zhang, 2015), e.g., communities of *Streptomyces, Pseudomonas* and *Vibrio* spp. Given the multitude of interacting taxa and the small spatial scales at which interactions occur, the direct observation and quantification of individual microorganisms in stream biofilms is not possible to this date. Instead, patterns of taxa co-occurrence across samples can be used to infer microbial interactions, niches, and key taxa (Faust and Raes, 2012; Berry and Widder, 2014). Despite limitations stemming from the notion that co-occurrence networks are non-empirical and derived from correlations, co-occurrence patterns are often useful for assessing ecological networks and obtaining insights into the organisation of microbial communities (Faust and Raes, 2012; Layeghifard et al., 2017). Co-occurrence networks further allow the exploration of emergent properties, such as the density of interactions, clusters of interacting taxa or the stability of networks against fragmentation. For example, studying bacterial co-occurrences across a dendritic stream network, Widder et al. found evidence for the role of spatial and hydrological processes in shaping co-occurrence network structure and stability (Widder et al., 2014). More recently, Ma et al. highlighted the interconnected patterns across microbiomes in various environments, emphasising the critical impact of microbial interactions on community assembly processes (Ma et al., 2020).

The environment of glacier-fed streams (GFS) is harsh. Low water temperature, coupled with high turbidity, oligotrophy, and snow- and ice-cover over extended periods of the year collectively contribute to this harshness. Unstable stream channels resulting from high turbulence and sediment loads further contribute to challenging conditions, making it difficult for benthic biofilms to establish (Roncoroni et al., 2023). Depending on local topography and geomorphology, GFS may develop into extensive floodplains, with dynamic braided channels exhibiting channel migration on a diel basis (Roncoroni et al., 2023). As distance increases downstream, harshness tends to decrease as channels consolidate and pioneering vegetation stabilises channels (Miller and Lane, 2019). Channel stabilisation increases the habitability of GFS, making these ecosystems more hospitable for benthic biofilms (Roncoroni et al., 2023). Furthermore, towards the edges of the proglacial floodplain, non-glacial tributary streams (TRIB), predominantly fed by groundwater and snowmelt, drain adjacent elevated terrasses that ultimately discharge into the glacier-fed mainstem (Brown et al., 2007). Given the lack of connectivity with the glacier, the TRIB environment is therefore more stable compared to GFS (Roncoroni et al., 2023), with potential consequences for biofilm structure and function (Freimann et al., 2013; Michoud et al., 2023). In fact, despite their close spatial proximity, GFS and TRIB streams host biofilms that differ in their biomass, composition, and diversity, as well as functional potential (Freimann et al., 2013; Brandani and Busi, 2022). It is predicted that as glaciers shrink, GFS will gradually transition into streams fed by groundwater and snowmelt, hence becoming more alike TRIB (Milner et al., 2017; Kohler et al., 2022). However, the consequences of this environmental change for microbial life remain poorly understood.

Brandani et al. recently reported on the complexity of GFS benthic biofilms in the proglacial floodplains of the Swiss Alps (Brandani and Busi, 2022). For instance, alongside bacteria, eukaryotic photoautotrophs such as Ochrophytes (algae) are a major component of these streams. Additionally, fungi such as Chytridiomycetes play key roles in microbial interactions, being involved in 12% of the overall interactions within these streams (Brandani and Busi, 2022). However, the role of taxa within these complex networks in influencing and maintaining community composition have not been resolved. Here, we investigated the properties of cross-domain co-occurrence networks of benthic biofilms in GFS and TRIB with different deglaciation histories in two proglacial floodplains in the Swiss Alps. We focused on putative interactions between bacteria, eukaryotic photoautotrophs, and parasitic fungi. Eukaryotic photoautotrophs are indeed important to stream biofilms as their exudates constitute an energy source for heterotrophic bacteria (e.g. during algal blooms (Kaplan and Bott, 1989), in the presence of light (Battin et al., 2003; Wagner et al., 2017)), while fungi can parasitize algae, thereby affecting microbial carbon flow (Klawonn et al., 2021). Therefore, we posit eukaryotic photoautotrophs and fungi as keystone taxa that take a central role, defined here based on network centrality measures, in maintaining network structure, and further hypothesise that the apparent stability of co-occurrence networks in GFS and TRIB changes along downstream and lateral gradients of deglaciation histories and environmental stability. To evaluate these roles, we assessed the stability of cross-domain co-occurrence networks upon removal of keystone taxa. We also investigated the variance in bacterial community composition that can be explained by specific bacterial and eukaryotic keystone taxa, subsequently contrasting this to the variance that can be explained by environmental differences among sites.

Materials and methods

Sample collection

Benthic biofilms were collected from stream sediments (0 - 5 cm depth) near to the glacier snout and extending to the floodplain's outlet in various TRIB and GFS stream reaches within the Otemma Glacier (Otemma; 45° 56' 08.4" N, 7° 24' 55.1" E) and Val Roseg Glacier (Val Roseg; 46° 24' 21.1" N, 9° 51' 55.1" E) floodplains in Switzerland. In each reach, we collected sandy sediments (0.25 - 3.15 mm) with flame-sterilised sieves and spatulas. Sediments samples were transferred to sterile vials, immediately frozen in the dark on dry ice in the field, and stored at -80°C in the laboratory until analysis (all of which were completed within a few months). Samples were collected during early (June/July) and late (August/September) summer (Brandani and Busi, 2022) in 2019. As shown previously, the two sample periods did not show differences in terms of community composition and structure (Brandani and Busi, 2022). For water, i.e. DOC and major ion measurements, samples were refrigerated, while nutrient samples were frozen at -20 C. Samples for the analysis of ions were sterile-filtered and kept at 4C until processing within a few weeks. Samples for the analysis of DOC were GF/F filtered, kept at 4C in the dark and analysed immediately upon return to the laboratory (i.e. within 48h after sampling). Study reaches were categorised into GFS or TRIB depending on their connectivity to glacier runoff based on visual field observations, drone-based imagery, and physicochemical characteristics (Brandani and Busi, 2022). The physicochemical parameters were established previously as part of the study characterising the diversity of the pro- and eukaryotic communities (Brandani and Busi, 2022). Overall, a total of 136 samples (GFS: 50; TRIB: 86) were collected across both floodplains. The sampling sites, indicating pre-2000 (UP) versus post-2000 (DOWN), from GFS and TRIB across Otemma and Val Roseg are depicted in Supplementary Figure 1. These included 68 samples each for the Otemma Glacier and Val Roseg Glacier floodplains, with the exact breakdown of these samples into GFS (Otemma:20; Val Roseg: 30) and TRIB (Otemma:48; Val Roseg: 38) including metadata listed in Supplementary Table 1.

Deglaciation histories

We identified past glacier extents from historic orthophotos and maps using SWISSIMAGE journey through time (Rickenbacher, 2020), and the GLIMS glacier inventory (Raup et al., 2007). These extents were compared with GLAMOS (Linsbauer et al., 2021) frontal variation measurements to verify glacial re-advances. The year of latest glaciation was thus interpolated for each sample site, which provided the longitudinal deglaciation history. We further split the reaches of the floodplain into those which were already deglaciated in the year 2000 (post-2000 or DOWN) and those still glaciated in 2000 (pre-2000 or UP; Supplementary Figure 1A). A lateral gradient (relative to the GFS mainstem) is also given by the TRIB streams that drain the adjacent terraces on the margins of the proglacial floodplains.

Benthic algal biomass

Benthic algal biomass was estimated as chlorophyll α (Supplementary Table 2) using a modified ethanol extraction protocol (Kohler et al., 2020), processed within a few months after collection. For this, sediment (ca. 2 g) samples were treated with 5 ml of 90% EtOH and then placed in a hot water bath (78°C, 10 min), followed by an incubation in the dark (4°C, 24 h). They were thereafter vortexed, centrifuged, and the supernatant read on a plate reader at 436/680 nm (excitation/emission). Chlorophyll α concentrations were inferred from a spinach standard and normalised by the sediment dry mass (DM).

Metabarcoding library preparation and sequencing

A previously established protocol (Busi et al., 2020) utilising phenol-chloroform was used for DNA extraction from benthic sediments (ca. 0.5 g). After initial processing, the DNA samples were diluted to a final concentration of ≤ 2-3 ng/µl. For the 16S rRNA gene metabarcoding analyses, we used the methodology previously described in Fodelianakis et al. (Fodelianakis et al., 2022), where the V3-V4 hypervariable region of the 16S rRNA gene were targeted with the 341F/785R primers. This was done in line with the 16S library preparation Illumina guidelines for the MiSeq system. The eukaryotic 18S rRNA gene metabarcoding library preparation was performed similarly but using the TAReuk454F-TAReukREV3 primers (Stoeck et al., 2010). Based on the MiSeq manufacturer's protocol, amplicon libraries were prepared where a second PCR was used to add dual indices to the purified amplicon PCR products. This allowed for extensive multiplexing of samples on a single sequencing lane of the MiSeq (Illumina) platform after quantification and normalisation. Samples were subsequently sequenced using a 300-base paired-end protocol in the Lausanne Genomic Technologies Facility (Switzerland).

Metabarcoding analyses

For the 16S and 18S rRNA metabarcoding data analyses, a combination of Trimmomatic v0.36 (Bolger et al., 2014) and QIIME2 v.2020.8 (Bolyen et al., 2019) were used with the latest SILVA database (Quast et al., 2013) v138.1 for taxonomic classification of the gene amplicons, i.e. 16S rRNA and 18S rRNA. From the 16S rRNA amplicon dataset, non-bacterial amplicon sequence variants (ASVs), i.e., archaea, chloroplasts, and mitochondria, were removed from all downstream analyses. The dataset was not rarefied for the analyses owing to the saturation of sequencing curves (Brandani and Busi, 2022) and also based on our previous work where rare taxa play important

roles in GFS community structure (Wilhelm et al., 2014). Additionally, the rationale behind discarding the archaeal reads was that the primers used were not designed, and are therefore not optimal, for detecting all lineages of archaea (Bahram et al., 2019). A total of 192 sample libraries were generated for the 16S rRNA sequencing and paired-end sequencing produced a total of 17,200,512 reads, with an average of 89,586 reads per sample. For the 18S rRNA amplicon dataset, a total of 157 amplicon sequence libraries from sediment samples were generated. Meanwhile, singletons and ASVs observed only once were discarded. The paired end 18S rRNA sequencing generated a total of 10,837,518 reads, with an average of 64,127 reads per sample. The 18S ASVs were further clustered into operational taxonomic units (OTUs) based on a 97% identity threshold using the *de novo* clustering method in *vsearch*, which has been implemented in QIIME2. This approach was used in order to avoid the overestimation of diversity driven by a high copy number of 18S in eukaryotic cells (Brandani and Busi, 2022). For the analyses, photoautotrophic eukaryotes along with fungi and protists were retained in the 18S rRNA amplicon dataset, given their contributions to community composition as highlighted in our previous work (Battin et al., 2016). Consequently, other non-phototrophic eukaryotes were discarded from the downstream analyses. The 18S rRNA dataset was also not rarefied, and any singletons/OTUs observed in only one sample were removed from downstream analyses, resulting in a dataset of 18S rRNA eukaryotic photoautotrophs and fungi with 429 OTUs. For our downstream analyses (i.e., co-occurrence), it was critical to include paired datasets. This was also a limitation of SpiecEasi (Kurtz et al., 2015), where only paired samples can be included to run pro- and eukaryotic network analyses. To address this, only 136 samples with paired 16S and 18S data were used for downstream analyses.

Co-occurrence networks

To study potential interactions between bacteria, eukaryotic photoautotrophs, and fungi, co-occurrence network analyses were performed with samples meeting specific criteria. These included: 1) the presence of both 16S and 18S sequence data for each sample, and 2) samples had to be categorised the same way across both samplings to ensure replicability (i.e., either designated as GFS or TRIB for both samplings as described by Brandani et al. (Brandani and Busi, 2022). Due to the dynamic nature of proglacial streams, GFS channels tend to meander and migrate, leaving some sites dry, under the influence of TRIB, or infiltrating TRIB streams. Hence, this approach was adopted to avoid potential confounders.

Subsequently, to reduce the noise and overall computational effort, any ASVs found in less than 5% of the samples were discarded from the 16S dataset for the co-occurrence networks. Prior to setting a threshold of 5%, the total number of ASVs retained in the datasets at 1, 5 and 10% thresholds was assessed to reduce noise and computational burden. From a total of 25,307 and 26,912 ASVs in Otemma and Val Roseg, respectively, only 5,268 and 5,216 ASVs were retained at 10%. To avoid exclusivity to only highly-prevalent taxa, we chose the 5% cutoff, where, 17,825 and 16,512 ASVs were respectively retained in Otemma and Val Roseg (Supplementary Table 3).

Co-occurrence networks between 16S and 18S (i.e., bacteria, eukaryotic photoautotrophs and fungi) were constructed using an ensemble of the distance matrices created from SparCC (Friedman and Alm, 2012), Spearman's correlation (Hauke and Kossowski, 2011), and SpiecEasi (Kurtz et al., 2015) where the networks were constructed using the Meinshausen and Bühlmann (mb) method (Meinshausen and Bühlmann, 2006). Networks were constructed across reaches, for pre-2000 and post-2000segments separately, and across GFS and TRIB for both Otemma and Val Roseg floodplains. Since our analyses are based on amplicon sequence data alone, we focused on the positive interactions across domains to assess potential mutualism within the microbiome. While reports suggest that negative interactions are indicative of co-exclusion mechanisms, especially in human microbiomes (Faust et al., 2012), the paucity of information available, especially in poorly characterised ecosystems may be insufficient to establish via amplicon sequencing data.

To detect communities in the network analyses, we used the Louvain clustering algorithm (Ghosh et al., 2018), removing clusters with less than five nodes. Herein, each community is defined as nodes within the graph with a higher probability of being connected to each other than to the rest of the network. Following this, we calculated network topology measures, including number of nodes and edges, number of clusters, diameter, edge-density, and modularity. These refer to the overall structure of the network such as the number of constituent entities (nodes), their connections (edges), overall length of connections (diameter), and independent cliques (clusters). The adjacency matrix was visualised using the *igraph* package (Csardi and Nepusz, 2006) in R v4.0.3 (R Core Team, 2021). Centrality measures, degree and betweenness, were also estimated per node, using the *igraph* v1.3.4 package. The fragmentation (*f*) of the network was determined as the percentage of the number of disconnected subgraphs over the overall nodes in each network (Widder et al., 2014). Fragmentation was estimated iteratively by the removal of each keystone taxa, i.e., top 10 nodes with both a high degree and a high betweenness in each graph. This information was further used for the subsequent generation of network topologies such as the number of clusters following the initial Louvain clustering of the network.

Community analyses

To explore the role of keystone taxa in structuring biofilm communities, we used constrained ordinations (db-RDA, R function vegan:capscale) using Bray-Curtis distances. In contrast to unconstrained ordination, constrained, or canonical ordination, resolves only the variation that can be explained by the constraints. We first employed a forward selection strategy (vegan:ordistep) to identify sets of non-redundant and significant ($p < 0.01$) bacterial and eukaryotic keystone taxa that we used as constraints for RDA. We subsequently used the explained variance in the constrained ordination as a measure of how important bacterial and eukaryotic

keystone taxa are for explaining bacterial community composition in the two floodplains. Prior to db-RDA, Wisconsin-double standardisation was applied to the ASV counts. The relative abundances of keystone taxa were inspected for multicollinearity and were provided as constraints for stepwise model creation (using 999 permutations). Model significance was evaluated for each RDA axis and explained variance of the constraints was extracted. To contrast variance in bacterial community composition that could be explained by keystone taxa, we performed a similar analysis using environmental parameters. For this, important environmental parameters including pH, water temperature, specific conductivity, dissolved oxygen (DO), turbidity and major ions and nutrients (obtained from Brandani and Busi, 2022) were first standardised (log-transformed) and then supplied to forward selection in db-RDA as described above.

Data analysis

All statistical analyses were performed in R v4.0.3. The *ggplot2* (Wickham, 2011) package was used for generating plots in R, while *patchwork* (https://github.com/thomasp85/patchwork) and Adobe Illustrator were used to arrange the figures as displayed.

Results

Cross-domain interactions underlie stream community structure in proglacial floodplains

In both proglacial floodplains, GFS and TRIB harbour diverse microbial communities which include bacteria, fungi, and phototrophic eukaryotes (Supplementary Figure 1). Previously, we reported on the overall community structure, diversity and abundances of the various microbial components in these streams (Brandani and Busi, 2022), including the role of benthic primary producers in shaping the microbial communities (Michoud et al., 2023). Microalgae such as Chlorophyta, Charophyta, Cryptomonadales, and Ochrophyta form the majority of the phototrophic eukaryotes, which collectively contribute a higher diversity in TRIB compared to GFS (Brandani and Busi, 2022). We previously observed the richness and Pielou's evenness indices were higher in TRIB compared to GFS. In light of this, we similarly found that bacterial communities were more diverse within the TRIB and dominated by Proteobacteria (Gamma- and Alphaproteobacteria), Bacteroidia, Cyanobacteria, and Verrucomicrobia. Based on covariation of taxa abundances across samples, we built co-occurrence networks. These networks were based on 1,090 nodes including bacteria, eukaryotic photoautotrophs, and fungi, with an average of 61,115 edges (interactions) connecting the nodes. The topological characteristics of the individual networks yielded similar metrics, reflected in density, modularity, assortativity and transitivity (Supplementary Table 1). In all networks, except Otemma post-2000, we identified three dense clusters of co-occurring taxa, one with a majority of eukaryotic photoautotrophs, another comprising mainly bacteria, and an intersecting third cluster composed of microbial eukaryotes including fungi and bacteria (Figure 1).

Next, we assessed the relative abundance of taxa present in the networks at the family level. Across both floodplains and stream types, we found that Acetobacteraceae were significantly overrepresented in networks constructed from pre-2000compared to post-2000reaches (two-way ANOVA, adj. $p < 0.05$, Supplementary Figures 2A, B, 3B). On the other hand, Comamonadaceae were significantly overrepresented in post-2000networks (two-way ANOVA, adj. $p < 0.05$), especially in TRIB (Supplementary Figures 2B and 3B). We also found that Chrysophyceae were overrepresented in pre-2000networks, while Diatomea decreased in pre-2000networks (two-way ANOVA, adj. $p < 0.05$) (Supplementary Figures 2C, D and 3C, D). Chytridiomycota, parasitic fungi infecting algae (Klawonn et al., 2021), were prevalent in both GFS and TRIB networks, but their abundance did not significantly differ across pre-2000or post-2000sites. However, Zoopagomycota, also parasitic fungi (Spatafora et al., 2016), were considerably enriched in post-2000 reaches across stream types and floodplains (Supplementary Figures 1E-F and 3E, F; adj. $p < 0.05$, Two-way ANOVA).

Apparent stability of co-occurrence networks

Based on our observations of differential abundance patterns across stream types and deglaciation gradients, we further assessed the contributions of the individual taxa to the overall network. For this, we first identified potential keystone taxa within each network by identifying the top 10 nodes with both a high degree and a high betweenness in each network (Supplementary Figures 4 and 5). For example, taxa classified as Dikarya, Phragmoplastophyta, Chlorophyceae, Cryptomycota, and Diatomea, along with an ASV classified as Burkholderiales, were determined to be keystone taxa in the GFS network at the pre-2000segment of the Otemma Glacier floodplain (Supplementary Figure 4A). Conversely, at the post-2000 segment of the same floodplain, Burkholderiales, Phragmoplastophyta, Xanthophyceae, Chrysophceae, and Dikarya, for instance, were identified as keystone taxa. Similarly, in the pre-2000segment of the Val Roseg Glacier floodplain, Dikarya, Phragmoplastophyta, Gemmatales, Burkholderiales, Cryptomycota, and Diatomea, for instance, were identified as keystone taxa contributing to the network topology (Supplementary Figures 5A, C). Finally, we found various bacteria (e.g., Rhodobacterales, Sphingomonadales) and fungi (e.g., Chytridiomycota) to be keystone taxa in the post-2000 reaches within the Val Roseg floodplain (Supplementary Figures 5B, 5D).

To further understand the role of the keystone taxa in community structure and their effect on apparent network stability, we first assessed network fragmentation upon their removal. For this, the numbers of clusters based on Louvain clustering were determined for each network, following which, a keystone was removed. The fragmentation (f) of the network was assessed before and after iterative removal of the top 10 keystone taxa. Interestingly, we found that in the Otemma Glacier floodplain (Figure 2A), the fragmentation of the networks constructed from

10.3389/frmbi.2023.1280809

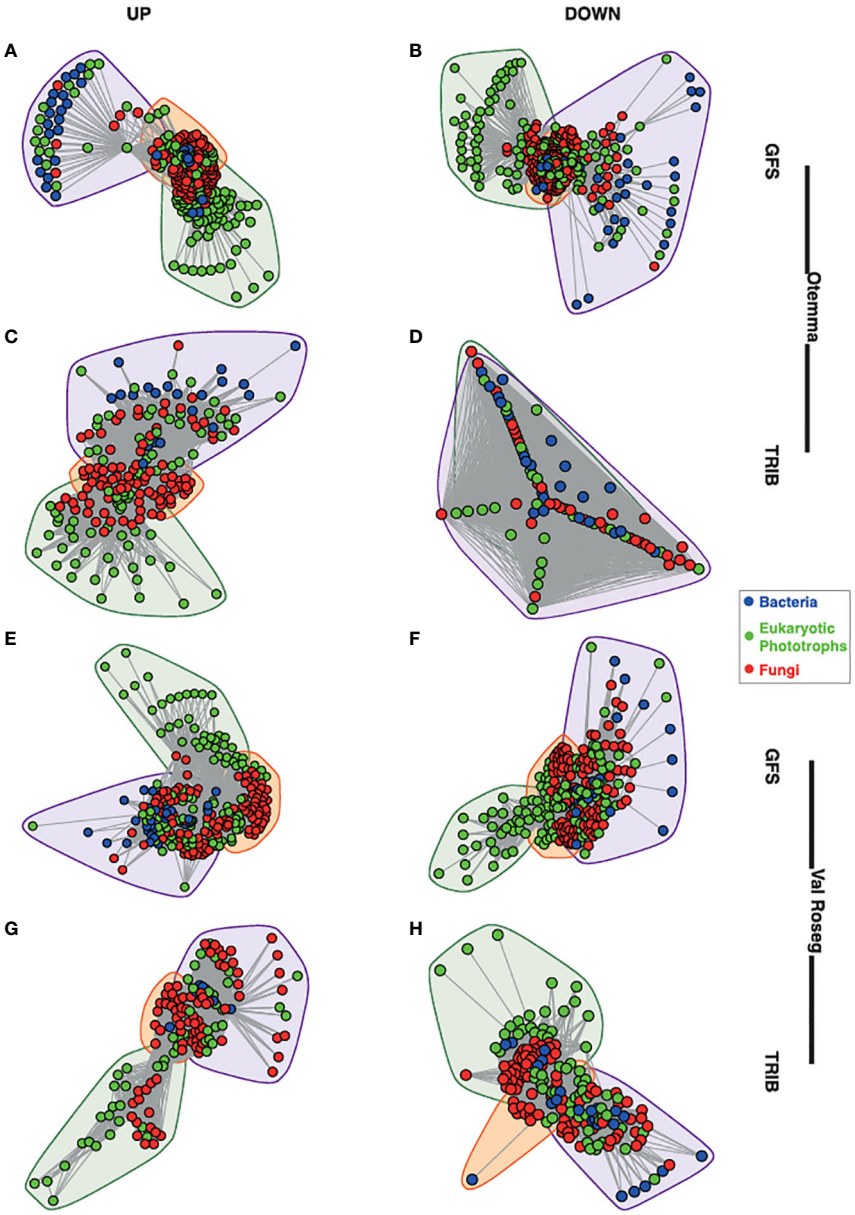


FIGURE 1
Network structure of glacier-fed streams and tributary streams. The overall structure of the cross-domain networks from the glacier-fed streams (GFS) and non-glacial tributaries (TRIB) are depicted. **(A)** GFS from the pre-2000reaches at Otemma, **(B)** GFS from the pre-2000reaches at Otemma, **(C)** TRIB from the post-2000 reaches at Otemma, **(D)** TRIB from the pre-2000reaches at Otemma. From the Val Roseg glacier, the network structures are depicted as follows: **(E)** GFS from the pre-2000reaches, **(F)** GFS from the post-2000 reaches, **(G)** TRIB from the pre-2000reaches, **(H)** TRIB from the post-2000 reaches. Each node represents a single amplicon sequence variant (ASV), and the lines represent the edges between them, while the colours indicate bacteria, eukaryotic photoautotrophs (indicated as phototrophs) and fungi. The convex hulls indicate clusters identified based on Louvain clustering of the overall network.

the GFS in the post-2000 reaches, increased upon removal of two to three keystone taxa, while the TRIB fragmentation increased upon removal of five keystone taxa. The pre-2000networks, however, appeared more stable, where fragmentation occurred only upon removal of five or eight keystone taxa. In Val Roseg, especially in TRIB (Figure 2B), the overall fragmentation of the microbial network was higher (f_{mean}=0.48) compared to GFS (f_{mean}=0.18) upon removal of four or five keystone taxa.

Finally, we unravelled the role of keystone taxa for biofilm community composition. Constrained ordinations revealed that both bacterial as well as eukaryotic phototrophic keystone taxa can explain a substantial fraction of bacterial community dissimilarity at the floodplain scale (Figure 3). Specifically, the relative abundance of bacterial keystone taxa explained 35.0% and 25.4% of variance in bacterial community similarity in Val Roseg and Otemma, respectively. While eukaryotic keystone taxa appeared particularly important for explaining network stability, they played a minor role in explaining bacterial community composition (i.e., 8.5% and 2.4% of explained variance in Val Roseg and Otemma, respectively). This is surprising, particularly

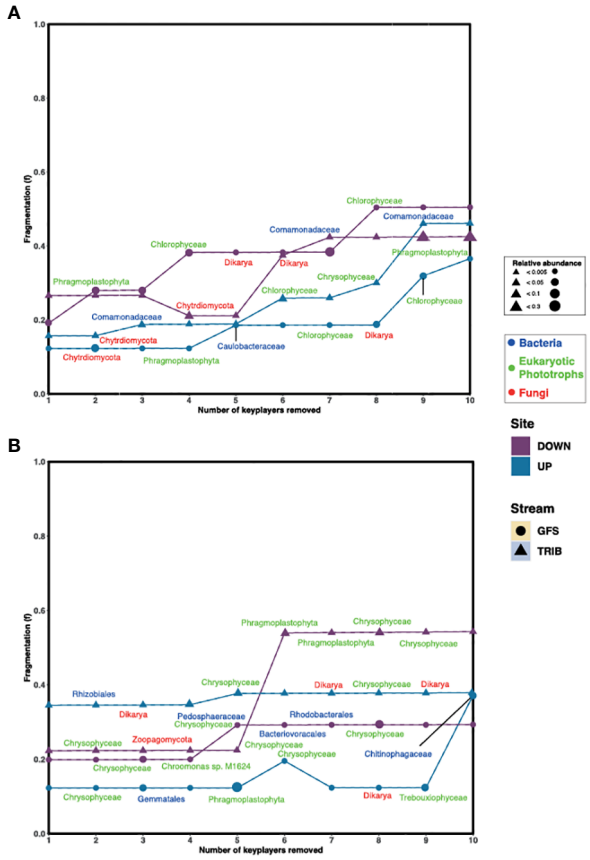


FIGURE 2
Keystone removal leads to fragmentation of the network. The change in fragmentation (*f*) for **(A)** Otemma and **(B)** Val Roseg are indicated in the line plots, where *f* was recalculated after each keystone (i.e., bacteria, eukaryotic photoautotroph or phototroph and fungi) was removed from the network. The size of the symbols indicates the relative abundance of the individual keystone taxa within the 16S or 18S data, respectively.

in relation to the variance in bacterial community composition that could be explained by environmental conditions, which accounted for a mere 16.5% and 14.5%, respectively. The retained environmental parameters, including streamwater temperature, nutrients and DOC concentration explain differences among TRIB and GFS bacterial communities.

Discussion

Biotic interactions are a salient property of microbial communities, with evidence of cross-domain interactions reported from various ecosystems, including freshwaters (Sun et al., 2022; Wijewardene et al., 2022), oceans (Rowan-Nash et al., 2019), glaciers (Brown and Jumpponen, 2015), and icesheets (Perini et al., 2023). To date, such interactions have not been studied in proglacial stream biofilms. Our findings suggest that biotic interactions, as inferred from co-occurrence patterns, play a pivotal role in influencing the apparent stability of stream biofilm communities along deglaciation and environmental gradients. Although previous reports showed structural and functional differences of the biofilm communities dwelling in different stream types within proglacial floodplains (Freimann et al., 2013; Brandani and Busi, 2022), we found that the overall network topology was similar between both proglacial floodplains, stream types, and deglaciation gradients. This contrasts with our expectation of successional imprints owing to deglaciation on co-occurrence networks. On the one hand, biotic interactions may be established very early on during community succession in the streams that drain recently deglaciated terrain. Indeed, our sampling design covered the successional timescale of the past 20 (pre-2000sites) and 80 (post-2000) years. Furthermore, functional redundancies across clades may also contribute to the apparent similarity of cross-domain interaction networks. Functionally redundant taxa may transiently occupy the same position in interaction networks and therefore result in similar network topologies. However, additional work will be necessary to relate network topology, taxa position and stability with functional characteristics to substantiate this notion.

Cross-domain networks have the potential to reveal key associations between microbial taxa (Williams et al., 2014). We found that biofilms in GFS and TRIB draining recently deglaciated terrain (i.e., pre-2000sites) had relatively more stable networks compared to the post-2000 sites. This finding suggests that bacterial keystone taxa are important for the apparent stability of the cross-domain interaction networks of biofilms dwelling in these nascent stream ecosystems. Furthermore, our results reveal that not all keystone taxa are typically among the most abundant community members, suggesting that low abundance taxa may also play important roles in stabilising microbial networks, corresponding to the notion of keystone species (Han et al., 2022). Our findings agree with observations from recent reports (Crump et al., 2009; Wilhelm et al., 2014; de Cena et al., 2021) highlighting the role of low-abundance taxa in ecosystem function and structure. For example, de Cena et al. recently hypothesised that low-abundance taxa, albeit in the human microbiome, act as keystone taxa, and might often be more metabolically influential within the community (de Cena et al., 2021). Similarly, Crump et al. (Crump et al., 2009) identified microbial keystone taxa, albeit neither rare nor abundant, that are central to ecosystem-level metabolic activity. Additionally, as Ren et al. (Ren et al., 2021) highlighted, there may be several causes including but not limited to environmental variables and/or the metacommunity structure of tributaries in Otemma. It is plausible that both metacommunity structure such as connectivity, directionality of dispersal, and also if these sites are rather discrete communities or part of a gradient/continuum may be contributing to these differences (Leibold and Mikkelson, 2002). It is also likely that the physical heterogeneity of the floodplains, i.e. tributaries in Otemma were larger compared to Val Roseg, may potentially lead to greater heterogeneity and thus increased fragmentation.

Work on multi-trophic food webs (Stouffer and Bascompte, 2011) and agroecosystems (Pocock et al., 2012) has demonstrated the fragility of ecological networks towards removal of key nodes. Our fragmentation analysis substantiates the notion of keystone

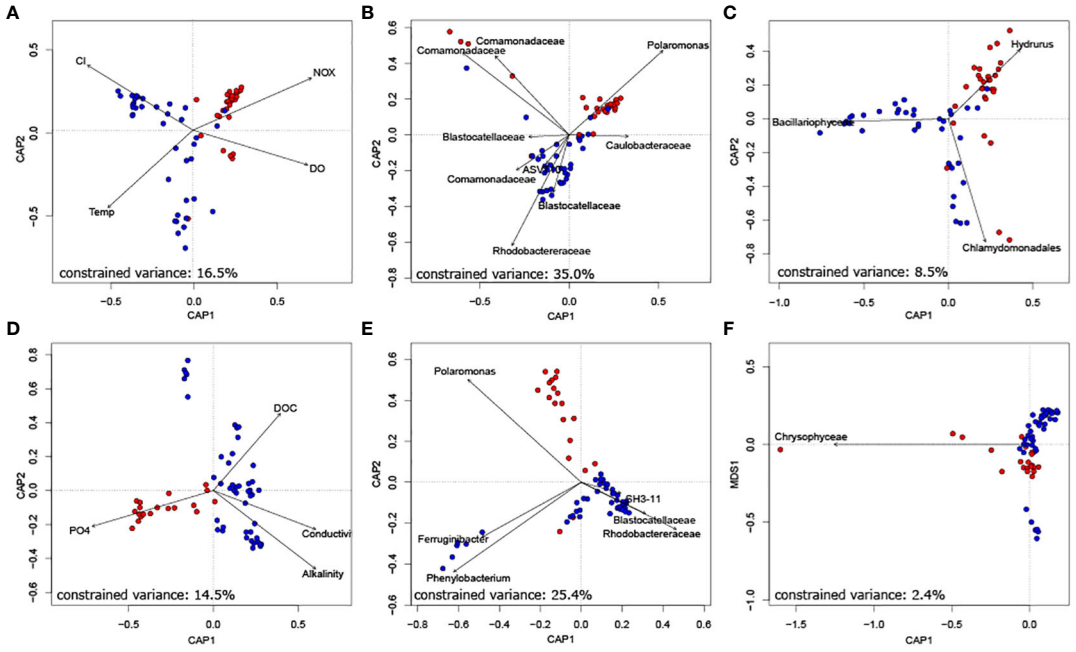


FIGURE 3
Bacterial keystone taxa well explain bacterial community composition. Constrained ordination of Val Roseg **(A–C)** and Otemma **(D–F)** floodplain samples revealed that bacterial keystone taxa **(B, E)**, as identified by their position in co-occurrence networks explained most of the variance in Bray-Curtis distance-based bacterial community composition. This outweighed the role of key environmental parameters **(A, D)** and of eukaryotic keystone taxa **(C, F)**.

taxa and their role for the stability of the cross-domain network. Interestingly, we identified several eukaryotes as keystone taxa, underscoring their relevance for biofilm structure and functioning. In GFS in Central Asia, Ren et al. (Ren and Gao, 2019) reported that fungi form integral components of cross-domain interactions networks, forming more clustered networks less susceptible to disturbances. As highlighted previously for stream biofilms (Wagner et al., 2015; Busi et al., 2022), eukaryotic algae serve as sources of organic matter thereby fuelling phototrophic-heterotrophic interactions. Simultaneously, parasitic fungi also foster the release of organic compounds from algae via the 'fungal shunt' (Klawonn et al., 2021). The prevalence of parasitic fungi has been noted previously in GFS (Kohler et al., 2022) and other cryospheric ecosystems (Anesio et al., 2017); our analyses further point to the importance of interactions among parasitic fungi and their algal host in proglacial stream biofilms. Along these lines, Mo et al. (Mo et al., 2021) recently suggested that interactions of microeukaryotes between them in the Lena River continental shelf were more stable compared to that of the estuary, potentially explained by variability in salinity. In contrast, Liu and Jiang reported that bacteria-bacteria interactions dominate co-occurrence networks in coastal sea waters of Antarctica (Liu and Jiang, 2020) and related this to competitive abilities of prokaryotes.

Taken together, the roles of bacterial and eukaryotic keystone taxa in ecological networks and their stability may very much be context dependent. We argue that, likely driven by the provisioning of organic matter to heterotrophs, eukaryotic algae and their fungal counterparts play central roles in biofilm interaction networks. However, we quantified the importance of pro- and eukaryotic keystone taxa to overall bacterial community structure and found that the relative abundance of bacterial keystone taxa could explain much of the bacterial community structure. This points towards a hierarchical structuring of interactions among eukaryotic and bacterial members of the biofilm. While eukaryotic primary producers may directly interact with only some bacterial keystone taxa, these bacterial keystone taxa themselves interact, likely via the exchange of secondary metabolites, with a much larger number of bacteria in the biofilm assemblage (Hessler et al., 2023). Such a hierarchical organisation of interactions is likely sensitive to changes in taxa at the base (i.e., the algal primary producers) whereas functional redundancies may dampen the impacts of taxa replacement. This may be particularly relevant in proglacial streams, where reduced light availability due to suspended particles and substrate instability typically inhibit algal growth. The current retreat of glaciers weakens these controls with potential effects on stream microbial communities.

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/, PRJNA808857 and https://doi.org/10.5281/zenodo.7524289, zenodo-7524289.

Author contributions

SBB: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Visualization, Validation, Writing – original draft. HP: Conceptualization, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review & editing. JB: Investigation, Writing – original draft, Writing – review & editing. TK: Writing – original draft, Writing – review & editing. SF: Data curation, Writing – review & editing. PP: Methodology, Writing – review & editing. MB: Data curation, Writing – review & editing. GM: Data curation, Writing – review & editing. LE: Data curation, Writing – review & editing. SL: Conceptualization, Funding acquisition, Resources, Writing – original draft, Writing – review & editing. PW: Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Writing – original draft, Writing – review & editing. TB: Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. Funding was provided by the Swiss National Science Foundation grant (CRSII5_180241) to TB, SL and PW.

Acknowledgments

We would like to acknowledge Kevin Casellini, Nicola Deluigi, Matteo Roncoroni, for their help sampling in the field. We would also like to thank Martina Schön for her help with collecting and interpolating glacial extent. The Hunting and Fishing Office of the Canton of the Grisons gave permission to fly the drone on the floodplain of the Tschierva Glacier.

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frmbi.2023.1280809/full#supplementary-material

SUPPLEMENTARY FIGURE 1
16S and 18S community profiles. **(A)** Bird's eye-view of the Otemma (left) and Val Roseg (right) floodplains depicting the glacier-fed stream mainstem (GFS) and the branching non-glacial tributaries (TRIB). The dashed line indicates the year 2000, where samples were classified as 'pre-2000 or UPor 'post-2000 or DOWN' site above and below, respectively. **(B)** Family-level profiles of the top 15 bacteria found in the floodplains across reaches and stream types (GFS and TRIB). **(C)** Relative abundance of the top 15 eukaryotic photoautotrophs.

SUPPLEMENTARY FIGURE 2
Taxa contributing to cross-domain interactions in Otemma. Relative abundance of bacteria found in the cross-domain networks of the **(A)** GFS and **(B)** TRIB in Otemma. **(C)** and **(D)** show the relative abundance of the eukaryotic photoautotrophs in the GFS and TRIB respectively, while **(E)** and **(F)** depict the relative abundance of the fungi in Otemma.

SUPPLEMENTARY FIGURE 3
Taxa contributing to cross-domain interactions in Val Roseg. Relative abundance of bacteria found in Val Roseg in the cross-domain networks of the **(A)** GFS and **(B)** TRIB. Phototroph relative abundances in the **(C)** GFS and **(D)** TRIB. (E) and (F) depict the relative abundance of the fungi in GFS and TRIB in Val Roseg.

SUPPLEMENTARY FIGURE 4
Keystonetaxa in Otemma. The keystone taxa for the GFS at the **(A)** pre-2000and **(B)** post-2000 reaches are highlighted based on their domain of origin. Keystone taxa in the TRIB at the **(C)** pre-2000and **(D)** post-2000reaches from the TRIB are simultaneously shown. The x-axis represents the overall betweenness of the individual taxa, whereas the y-axis indicates the degree centrality.

SUPPLEMENTARY FIGURE 5
Keystone taxa in Val Roseg. The keystone taxa for the GFS at the **(A)** pre-2000and **(B)** post-2000reaches in Val Roseg are highlighted. keystone taxa in the tributaries at the **(C)** pre-2000and **(D)** post-2000reaches from the TRIB are depicted in the scatter plots. The x-axis represents the overall betweenness of the individual taxa, whereas the y-axis indicates the degree centrality, i.e., number of connections per node.

SUPPLEMENTARY TABLE 1
Metadata and network topology. Glacier metadata, including the glacier from which samples were collected, pre-2000or post-2000 reaches, and type of stream, i.e., Glacier-fed (GFS) or tributaries (TRIB), are indicated along with network topology measures. The dashed line (- - - -) indicates the 'millennium cut' based on which samples were classified as 'pre-2000 or up' or 'post-2000 or down'. The solid lines represent the deglaciated history based on the Glacier Extent Database to determine the date since 'last glaciation'.

SUPPLEMENTARY TABLE 2
Chlorophyll-α measurements. Levels of chlorophyll-α measured at the site for each sample are listed along with metadata.

SUPPLEMENTARY TABLE 3
Filtering thresholds for network analyses. ASVs were filtered at various thresholds, i.e. 1%, 5% and 10%, to reduce computational burden without losing diversity. Table indicates ASVs retained at chosen thresholds.

SUPPLEMENTARY TABLE 4
Pro- and eukaryotic keystone taxa. Abundance information for all ASVs detected in the bacteria (16S) and eukaryote (18S) datasets are provided alongside their indication of Keystone taxa or otherwise.

References

Anesio, A. M., Lutz, S., Chrismas, N. A. M., and Benning, L. G. (2017). The microbiome of glaciers and ice sheets. *NPJ Biofilms Microbiomes* 3, 10. doi: 10.1038/s41522-017-0019-0

Bahram, M., Anslan, S., Hildebrand, F., Bork, P., and Tedersoo, L. (2019). Newly designed 16S rRNA metabarcoding primers amplify diverse and novel archaeal taxa from the environment. *Environ. Microbiol. Rep.* 11, 487–494. doi: 10.1111/1758-2229.12684

Battin, T. J., Besemer, K., Bengtsson, M. M., Romani, A. M., and Packmann, A. I. (2016). The ecology and biogeochemistry of stream biofilms. *Nat. Rev. Microbiol.* 14, 251–263. doi: 10.1038/nrmicro.2016.15

Battin, T. J., Kaplan, L. A., Newbold, J. D., Cheng, X., and Hansen, C. (2003). Effects of current velocity on the nascent architecture of stream microbial biofilms. *Appl. Environ. Microbiol.* 69, 5443–5452. doi: 10.1128/AEM.69.9.5443-5452.2003

Berry, D., and Widder, S. (2014). Deciphering microbial interactions and detecting keystone species with co-occurrence networks. *Front. Microbiol.* 5, 219. doi: 10.3389/fmicb.2014.00219

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

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

Braga, R. M., Dourado, M. N., and Araújo, W. L. (2016). Microbial interactions: ecology in a molecular perspective. *Braz. J. Microbiol.* 47 Suppl 1, 86–98. doi: 10.1016/j.bjm.2016.10.005

Brandani, P., and Busi, (2022). Spatial patterns of benthic biofilm diversity among streams draining proglacial floodplains. *Front. Microbiol.* 13, 948165. doi: 10.3389/fmicb.2022.948165

Brown, S. P., and Jumpponen, A. (2015). Phylogenetic diversity analyses reveal disparity between fungal and bacterial communities during microbial primary succession. *Soil Biol. Biochem.* 89, 52–60. doi: 10.1016/j.soilbio.2015.06.025

Brown, L. E., Milner, A. M., and Hannah, D. M. (2007). Groundwater influence on alpine stream ecosystems. *Freshw. Biol.* 52, 878–890. doi: 10.1111/j.1365-2427.2007.01739.x

Busi, S. B., Bourquin, M., Fodelianakis, S., Michoud, G., Kohler, T. J., Peter, H., et al. (2022). Genomic and metabolic adaptations of biofilms to ecological windows of opportunity in glacier-fed streams. *Nat. Commun.* 13, 2168. doi: 10.1038/s41467-022-29914-0

Busi, S. B., Prameteftaki, P., Brandani, J., Fodelianakis, S., Peter, H., Halder, R., et al. (2020). Optimised biomolecular extraction for metagenomic analysis of microbial biofilms from high-mountain streams. *PeerJ* 8, e9973. doi: 10.7717/peerj.9973

Crump, B. C., Peterson, B. J., Raymond, P. A., Amon, R. M. W., 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

Csardi, G., and Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal Complex Syst.* 1695, 1–9.

de Cena, J. A., Zhang, J., Deng, D., Damé-Teixeira, N., and Do, T. (2021). Low-abundant microorganisms: the human microbiome's dark matter, a scoping review. *Front. Cell. Infect. Microbiol.* 11, 689197. doi: 10.3389/fcimb.2021.689197

Faust, K., and Raes, J. (2012). Microbial interactions: from networks to models. *Nat. Rev. Microbiol.* 10, 538–550. doi: 10.1038/nrmicro2832

Faust, K., Sathirapongsasuti, J. F., Izard, J., Segata, N., Gevers, D., Raes, J., et al. (2012). Microbial co-occurrence relationships in the human microbiome. *PloS Comput. Biol.* 8, e1002606. doi: 10.1371/journal.pcbi.1002606

Fodelianakis, S., Washburne, A. D., Bourquin, M., Prameteftaki, P., Kohler, T. J., Styllas, M., et al. (2022). Microdiversity characterizes prevalent phylogenetic clades in the glacier-fed stream microbiome. *ISME J.* 16, 666–675. doi: 10.1038/s41396-021-01106-6

Freimann, R., Bürgmann, H., Findlay, S. E. G., and Robinson, C. T. (2013). Bacterial structures and ecosystem functions in glaciated floodplains: contemporary states and potential future shifts. *ISME J.* 7, 2361–2373. doi: 10.1038/ismej.2013.114

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

Ghosh, S., Halappanavar, M., Tumeo, A., Kalyanaraman, A., Lu, H., Chavarrià-Miranda, D., et al. (2018). "Distributed louvain algorithm for graph community detection," in *2018 IEEE International Parallel and Distributed Processing Symposium (IPDPS)*. 885–895.

Han, G., Luong, H., and Vaishnava, S. (2022). Low abundance members of the gut microbiome exhibit high immunogenicity. *Gut Microbes* 14, 2104086. doi: 10.1080/19490976.2022.2104086

Hauke, J., and Kossowski, T. (2011) *Comparison of values of Pearson's and Spearman's correlation coefficients on the same sets of data*. Available at: https://sciendo.com/downloadpdf/journals/quageo/30/2/article-p87.pdf?pdfJsInlineViewToken=1302953392&inlineView=true.

Hessler, T., Huddy, R. J., Sachdeva, R., Lei, S., Harrison, S. T. L., Diamond, S., et al. (2023). Vitamin interdependencies predicted by metagenomics-informed network analyses and validated in microbial community microcosms. *Nat. Commun.* 14, 4768. doi: 10.1038/s41467-023-40360-4

Kaplan, L. A., and Bott, T. L. (1989). Diel fluctuations in bacterial activity on streambed substrata during vernal algal blooms: Effects of temperature, water chemistry, and habitat. *Limnol. Oceanogr.* 34, 718–733. doi: 10.4319/lo.1989.34.4.0718

Klawonn, I., Van den Wyngaert, S., Parada, A. E., Arandia-Gorostidi, N., Whitehouse, M. J., Grossart, H.-P., et al. (2021). Characterizing the "fungal shunt": Parasitic fungi on diatoms affect carbon flow and bacterial communities in aquatic microbial food webs. *Proc. Natl. Acad. Sci.* 118, e2102225118. doi: 10.1073/pnas.2102225118

Kohler, T. J., Fodelianakis, S., Michoud, G., Ezzat, L., Bourquin, M., Peter, H., et al. (2022). Glacier shrinkage will accelerate downstream decomposition of organic matter and alters microbiome structure and function. *Glob. Change Biol.* doi: 10.1111/gcb.16169

Kohler, T. J., Peter, H., Fodelianakis, S., Prameteftaki, P., Styllas, M., Tolosano, M., et al. (2020). Patterns and drivers of extracellular enzyme activity in New Zealand glacier-fed streams. *Front. Microbiol.* 11, 591465. doi: 10.3389/fmicb.2020.591465

Kurtz, Z. D., Müller, C. L., Miraldi, E. R., Littman, D. R., Blaser, M. J., and Bonneau, R. A. (2015). Sparse and compositionally robust inference of microbial ecological networks. *PloS Comput. Biol.* 11, e1004226. doi: 10.1371/journal.pcbi.1004226

Layeghifard, M., Hwang, D. M., and Guttman, D. S. (2017). Disentangling interactions in the microbiome: A network perspective. *Trends Microbiol.* 25, 217–228. doi: 10.1016/j.tim.2016.11.008

Lee, J., and Zhang, L. (2015). The hierarchy quorum sensing network in Pseudomonas aeruginosa. *Protein Cell* 6, 26–41. doi: 10.1007/s13238-014-0100-x

Leibold, M. A., and Mikkelson, G. M. (2002). Coherence, species turnover, and boundary clumping: elements of meta-community structure. *Oikos* 97, 237–250. doi: 10.1034/j.1600-0706.2002.970210.x

Linsbauer, A., Huss, M., Hodel, E., Bauder, A., Fischer, M., Weidmann, Y., et al. (2021). The new Swiss Glacier Inventory SGI2016: From a topographical to a glaciological dataset. *Front. Earth Sci.* 9. doi: 10.3389/feart.2021.704189

Liu, Q., and Jiang, Y. (2020). Application of microbial network analysis to discriminate environmental heterogeneity in Fildes Peninsula, Antarctica. *Mar. pollut. Bull.* 156, 111244. doi: 10.1016/j.marpolbul.2020.111244

Ma, B., Wang, Y., Ye, S., Liu, S., Stirling, E., Gilbert, J. A., et al. (2020). Earth microbial co-occurrence network reveals interconnection pattern across microbiomes. *Microbiome* 8, 82. doi: 10.1186/s40168-020-00857-2

McCann, K. S. (2000). The diversity–stability debate. *Nature* 405, 228–233. doi: 10.1038/35012234

Meinshausen, N., and Bühlmann, P. (2006). Variable selection and high-dimensional graphs with the lasso. *Ann. Stat.* 34, 1436–1462. doi: 10.1214/009053606000000281

Michoud, P., Kohler, T. J., Peter, H., Brandani, J., Banu Busi, S., and Battin, T. J. (2023). Unexpected functional diversity of stream biofilms within and across proglacial floodplains despite close spatial proximity. *Limnol. Oceanogr.* doi: 10.1002/lno.12415

Miller, H. R., and Lane, S. N. (2019). Biogeomorphic feedbacks and the ecosystem engineering of recently deglaciated terrain. *Prog. Phys. Geography: Earth Environ.* 43, 24–45. doi: 10.1177/03091333188165

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

Mo, Y., Peng, F., Gao, X., Xiao, P., Logares, R., Jeppesen, E., et al. (2021). Low shifts in salinity determined assembly processes and network stability of microeukaryotic plankton communities in a subtropical urban reservoir. *Microbiome* 9, 128. doi: 10.1186/s40168-021-01079-w

Perini, L., Gostinčar, C., Likar, M., Frisvad, J. C., Kostanjšek, R., Nicholes, M., et al. (2023). Interactions of fungi and algae from the Greenland ice sheet. *Microb. Ecol.* 86, 282–296. doi: 10.1007/s00248-022-02033-5

Pocock, M. J. O., Evans, D. M., and Memmott, J. (2012). The robustness and restoration of a network of ecological networks. *Science* 335, 973–977. doi: 10.1126/science.1214915

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. *A language and environment for statistical computing* (Vienna, Austria: R Foundation for Statistical Computing). Available at: https://ci.nii.ac.jp/naid/20001692429/.

Raup, B., Racoviteanu, A., Khalsa, S. J. S., Helm, C., Armstrong, R., and Arnaud, Y. (2007). The GLIMS geospatial glacier database: A new tool for studying glacier change. *Glob. Planet. Change* 56, 101–110. doi: 10.1016/j.gloplacha.2006.07.018

Ren, Z., and Gao, H. (2019). Ecological networks reveal contrasting patterns of bacterial and fungal communities in glacier-fed streams in Central Asia. *PeerJ* 7, e7715. doi: 10.7717/peerj.7715

Ren, Z., Zhang, C., Li, X., Ma, K., Zhang, Z., Feng, K., et al. (2021). Bacterial communities present distinct co-occurrence networks in sediment and water of the thermokarst lakes in the yellow river source area. *Front. Microbiol.* 12, 716732. doi: 10.3389/fmicb.2021.716732

Rickenbacher, M. (2020). Journeys through time with the Swiss national map series. *Proc. 26th Int. Cartographic.*

Roncoroni, M., Mancini, D., Miesen, F., Müller, T., Gianini, M., Ouvry, B., et al. (2023). Decrypting the stream periphyton physical habitat of recently deglaciated floodplains. *Sci. Total Environ.* 867, 161374. doi: 10.1016/j.scitotenv.2022.161374

Rowan-Nash, A. D., Korry, B. J., Mylonakis, E., and Belenky, P. (2019). Cross-domain and viral interactions in the microbiome. *Microbiol. Mol. Biol. Rev.* 83. doi: 10.1128/MMBR.00044-18

Spatafora, J. W., Chang, Y., Benny, G. L., Lazarus, K., Smith, M. E., Berbee, M. L., et al. (2016). A phylum-level phylogenetic classification of zygomycete fungi based on genome-scale data. *Mycologia* 108, 1028–1046. doi: 10.3852/16-042

Stoeck, T., Bass, D., Nebel, M., Christen, R., Jones, M. D. M., Breiner, H.-W., et al. (2010). Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. *Mol. Ecol.* 19 Suppl 1, 21–31. doi: 10.1111/j.1365-294X.2009.04480.x

Stopnisek, N., Zühlke, D., Carlier, A., Barberán, A., Fierer, N., Becher, D., et al. (2016). Molecular mechanisms underlying the close association between soil Burkholderia and fungi. *ISME J.* 10, 253–264. doi: 10.1038/ismej.2015.73

Stouffer, D. B., and Bascompte, J. (2011). Compartmentalization increases food-web persistence. *Proc. Natl. Acad. Sci. U. S. A.* 108, 3648–3652. doi: 10.1073/pnas.1014353108

Sun, P., Wang, Y., Huang, X., Huang, B., and Wang, L. (2022). Water masses and their associated temperature and cross-domain biotic factors co-shape upwelling microbial communities. *Water Res.* 215, 118274. doi: 10.1016/j.watres.2022.118274

Vetsigian, K., Jajoo, R., and Kishony, R. (2011). Structure and evolution of Streptomyces interaction networks in soil and in silico. *PloS Biol.* 9, e1001184. doi: 10.1371/journal.pbio.1001184

Wagner, K., Bengtsson, M. M., Findlay, R. H., Battin, T. J., and Ulseth, A. J. (2017). High light intensity mediates a shift from allochthonous to autochthonous carbon use in phototrophic stream biofilms. *J. Geophysical Research: Biogeosciences* 122, 1806–1820. doi: 10.1002/2016JG003727

Wagner, K., Besemer, K., Burns, N. R., Battin, T. J., and Bengtsson, M. M. (2015). Light availability affects stream biofilm bacterial community composition and function, but not diversity. *Environ. Microbiol.* 17, 5036–5047. doi: 10.1111/1462-2920.12913

Wickham, H. (2011). ggplot2: ggplot2. *Wiley Interdiscip. Rev. Comput. Stat.* 3, 180–185.

Widder, S., Besemer, K., Singer, G. A., Ceola, S., Bertuzzo, E., Quince, C., et al. (2014). Fluvial network organization imprints on microbial co-occurrence networks. *Proc. Natl. Acad. Sci. U. S. A.* 111, 12799–12804. doi: 10.1073/pnas.1411723111

Wijewardene, L., Wu, N., Fohrer, N., and Riis, T. (2022). Epiphytic biofilms in freshwater and interactions with macrophytes: Current understanding and future directions. *Aquat. Bot.* 176, 103467. doi: 10.1016/j.aquabot.2021.103467

Wilhelm, L., Besemer, K., Fasching, C., Urich, T., Singer, G. A., Quince, C., et al. (2014). Rare but active taxa contribute to community dynamics of benthic biofilms in glacier-fed streams. *Environ. Microbiol.* 16, 2514–2524. doi: 10.1111/1462-2920.12392

Williams, R. J., Howe, A., and Hofmockel, K. S. (2014). Demonstrating microbial co-occurrence pattern analyses within and between ecosystems. *Front. Microbiol.* 5, 358. doi: 10.3389/fmicb.2014.00358

frontiers | Frontiers in Microbiomes

TYPE Original Research
PUBLISHED 15 January 2024
DOI 10.3389/frmbi.2023.1279550

OPEN ACCESS

EDITED BY
Janina Rahlff,
Friedrich Schiller University Jena, Germany

REVIEWED BY
Ovidiu-Nicolae Popa,
Heinrich Heine University Düsseldorf, Germany
Emelie Nilsson,
Linnaeus University, Sweden

*CORRESPONDENCE
Hannes Peter
hannes.peter@epfl.ch

RECEIVED 18 August 2023
ACCEPTED 22 December 2023
PUBLISHED 15 January 2024

CITATION
Peter H, Michoud G, Busi SB and Battin TJ (2024) The role of phages for microdiverse bacterial communities in proglacial stream biofilms.
Front. Microbiomes 2:1279550.
doi: 10.3389/frmbi.2023.1279550

The role of phages for microdiverse bacterial communities in proglacial stream biofilms

Hannes Peter[1]*, Grégoire Michoud[1], Susheel Bhanu Busi[2,3] and Tom J. Battin[1]

[1]River Ecosystems Laboratory, Alpine and Polar Environmental Research Center, Ecole Polytechnique Fédérale de Lausanne, Sion, Switzerland, [2]Systems Ecology Group, Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Esch-sur-Alzette, Luxembourg, [3]UK Centre for Ecology & Hydrology (UKCEH), Wallingford, Oxfordshire, United Kingdom

Viruses modulate the diversity and activity of microbial communities. However, little is known about their role for the structure of stream bacterial biofilm communities. Here, we present insights into the diversity and composition of viral communities in various streams draining three proglacial floodplains in Switzerland. Proglacial streams are characterized by extreme environmental conditions, including near-freezing temperatures and ultra-oligotrophy. These conditions select for few but well-adapted bacterial clades, which dominate biofilm communities and occupy niches via microdiversification. We used metagenomic sequencing to reveal a diverse biofilm viral assemblage in these streams. Across the different floodplains and streams, viral community composition was tightly coupled to that of the bacterial hosts, which was underscored by generally high host specificity. Combining predictions of phage-host interactions with auxiliary metabolic genes (AMGs), we identify specific AMGs shared by phages infecting microdiverse clade members. Our work provides a step towards a better understanding of the complex interactions among bacteria and phages in stream biofilm communities in general and streams influenced by glacier meltwaters and characterized by microdiversity in particular.

KEYWORDS

phage-host interactions, host range, auxiliary metabolic genes, environmental selection, coevolution

Introduction

Viruses infecting bacteria, also called bacteriophages, play important roles in modulating the diversity and activity of bacterial assemblages in every biome on Earth (Dion et al., 2020). Particularly phage-induced mortality is known to control bacterial abundance and biomass, community composition and diversity (Suttle, 1994; Suttle, 2007;

Roux et al., 2016). Metabolic reprogramming during lytic infections redirects host metabolism towards viral progeny production, with sizable impacts on ecosystem-scale biogeochemical cycles (reviewed by: Zimmerman et al., 2020). However, beyond altering host' metabolic pathways and inducing mortality, phages influence the fitness of their bacterial hosts in numerous ways. Lysogenic infections, for example, can impede host functioning by insertion of phage genetic material into functional genes (reviewed by: Obeng et al., 2016). And phage-encoded sigma factors can influence bacterial spore formation, affecting important bacterial life-history traits such as dormancy (Schwartz et al., 2022).

While the consequences of viral infections for individual bacterial cells are often detrimental, beneficial effects of phage-host interactions typically arise at the level of populations, communities or even holobionts (Shkoporov et al., 2022). For instance, phage-driven diversification (Koskella and Brockhurst, 2014), lysogenic conversion (Howard-Varona et al., 2017) or auxiliary metabolic genes [AMGs (Kieft et al., 2021b)] can increase phenotypic plasticity or augment the metabolic repertoire of microbial assemblages (reviewed by: Shkoporov et al., 2022). In biofilms, where densely packed communities of microbes interact, phage genomic material has been shown to enter even phylogenetically distant hosts directly or during conjugative transfers, ultimately increasing the immunological memory (Hwang et al., 2023).

Besides the density, diversity, and activity of phages and their hosts, the eco-evolutionary consequences of phage infections depend on host specificity. While a notion of a predominantly narrow phage host range prevailed in early work, recent advances, including single-cell and metagenomics studies of hydrothermal biofilms (Jarett et al., 2020; Hwang et al., 2023), suggest indeed a continuum of host ranges (reviewed by: de Jonge et al., 2019; Holtappels et al., 2023).

Here, we present insights into the interactions between bacterial biofilm communities and their phages in high-altitude stream ecosystems. Alpine streams are extreme environments, characterized by ultra-oligotrophy, low water temperature, exposure to ultraviolet (UV) radiation, and short ice-free seasons, yet represent hotspots of microbial biodiversity (Freimann et al., 2013; reviewed by: Hotaling et al., 2017; Brandani et al., 2022; Busi et al., 2022; Fodelianakis et al., 2022). Glacier influence adds additional constraints for life in alpine streams. Close to the glacier terminus, glacier-fed streams (GFS) are permanently ice cold and the erosive activity of glaciers liberates large quantities of mineral particles, which render GFS turbid and limit light for phototrophic primary production. Compared to GFS, groundwater-fed streams (GWS) that drain elevated terrain towards the edges of the floodplains, are warmer, their streambeds more consolidated and lower turbidity allows photosynthetic primary producers to dwell (Freimann et al., 2013; Brandani et al., 2022).

Previous work on biofilm communities in GFS and GWS revealed that they are diverse with members of all domains of life present (reviewed by: Battin et al., 2016; Brandani et al., 2022), distinct from the microbial community suspended in the water column (Ezzat et al., 2022), and shaped predominantly by environmental selection (Fodelianakis et al., 2022; Brandani et al., 2023). More specifically, this means that ecologically successful taxa in both GFS and GWS are recruited from similar clades, including taxa classified as *Polaromonas*, *Rhodoferax*, *Rhizobacter*, *Methylotenera*, and *Massilia*. These genera are microdiverse (Fodelianakis et al., 2022), indicating that they efficiently fill available niches which cannot be occupied by less-well adapted taxa. Microdiversity is an intrinsic property of many microbial communities (reviewed by: Larkin and Martiny, 2017), including marine viral communities (Needham et al., 2017; Gregory et al., 2019).

By focusing on microdiverse biofilm communities in proglacial streams, we aim to shed new light on the role of phage-host interactions in microbial communities in extreme environments. We hypothesize that narrow phage host ranges predominate in communities dominated by microdiverse clades because selective constraints which lead to microdiversification may also shape bacteria-phage coevolution. Moreover, we speculate that interactions between phages and microdiverse clade members in extreme environments may be predominantly beneficial, thus contributing to the ecological success of microdiverse clades in extreme environments. The rationale behind this is that reciprocal feedbacks between beneficial effects of phage-host interactions and the density of hosts could stabilize the process of microdiversification over evolutionary timescales – in contrast to a situation where density-dependent phage predation would lead to turnover in host communities (and thus suppress microdiversification). We argue that studying microbial communities in alpine stream biofilms is ideally suited to address such questions, because of the pronounced microdiversity and limited dispersal due to geographic and topographic isolation, which can increase phage-host encounter probabilities and hence lessen the coevolutionary strength of phage-host associations.

Materials and methods

Study sites

We here present data from an extensive survey of benthic biofilm communities sampled from proglacial streams in the Swiss Alps in summer 2019. Detailed descriptions of the sampling design and environmental parameters as well as analyses of spatial diversity patterns (Brandani et al., 2022), assembly processes (Brandani et al., 2023) and the functional diversity (Michoud et al., 2023) are available elsewhere. In this field survey, a total of 259 benthic sediment samples were collected from the Otemma Glacier (OTE, 45.95E, 7.45N), the Valsorey Glacier (SOY, 45.91E, 7.27N) and Val Roseg Glacier (VAR, 46.39E, 9.84N) floodplains. The sampling sites were distributed along gradients from the glacier snout to the outflows of the respective proglacial floodplains, covering terrain which has been deglaciated since ca. 36 (OTE), 65 (VAR), and 70 (SOY) years. Samples of the coarse sandy sediment fraction were collected using flame-sterilized shovels (0-5 cm depth) and fractionated using flame-sterilized sieves (Retsch, 0.25 and 3.15 mm). Subsamples were immediately flash-frozen on

dry ice and stored at -80°C until processing. Samples for bacterial cell counting were fixed using 1.8 mL of a mixture containing filter-sterilized paraformaldehyde (1% w/v) and 0.5% (w/w) glutaraldehyde, flash frozen and kept at -80°C (Brandani et al., 2022; Fodelianakis et al., 2022). After extraction from sediments using pyrophosphate, shaking and sonication, bacterial cells were stained using SYBR Green I (1 x final concentration; Thermo Fisher) for 15 minutes at 37°C and counted using flow cytometry (NovoCyte, ACEA) equipped with a 488 nm laser. Samples were analyzed with a reading time of 2 minutes at a flow rate of 14 µL min^{-1} and thresholds on the forward scatter channel set to 300 and the 530/30 nm fluorescence channel set to 1500. Cells were distinguished from background noise based on fluorescence signals on biplots of 530/30 and 725/40 nm using the NovoExpress (ACEA) software. Cell abundances were converted to cells per g dry sediment.

Here, we focus on the viral communities, which we assess using bulk metagenomic sequencing of a subset of samples. We obtained metagenomes for 47 samples from GFS (n=12) and GWS (n=35) from OTE (n_{GFS}=3, n_{GWS}=14), SOY (n_{GFS}=6, n_{GWS}=10) and VAR (n_{GFS}=3, n_{GWS}=11), respectively. Twenty-four of these samples were obtained in June–July, and 23 samples were collected in August–September 2019 (Supplementary Table S1).

DNA extraction, library preparation, and sequencing

DNA was extracted from 0.5 g of sediment using an optimized extraction protocol tailored to the low-biomass and mineral nature of these samples (Busi et al., 2020). Metabarcoding libraries were prepared using 2-3 ng $µL^{-1}$ input DNA and primers targeting the V3-V4 hypervariable region of the 16S rRNA gene [341f (5'-CCTACGGGNGGCWGCAG-3') and 785r (5'-GACTACHVGGGTATCTAATCC-3') (Klindworth et al., 2013)]. Amplification was performed on a Biometra Trio (Biometra) using the KAPA HiFi DNA Polymerase (Hot Start and Ready Mix formulation) in a 25 µL-amplification reaction containing 1 x PCR buffer, 1 µM of each primer, 0.48 µg $µL^{-1}$ bovine serum albumin (BSA), and 1.0 µL of template DNA. After an initial denaturation step at 95°C for 3min, 25 cycles of 94°C for 30 s, 55°C for 30 s, and 72°C for 30 s were followed by a final extension at 72°C for 5 min. Sequencing libraries were prepared using dual indices (Illumina). Prior to paired-end sequencing on a single MiSeq (Illumina) lane, library DNA concentrations were quantified, normalized, and pooled. Libraries were then sequenced at the Lausanne Genomic Technologies Facility (Switzerland), producing 300 bp paired end reads.

Metagenome libraries were prepared using the NEBNext Ultra II FS library kit, with 50 ng of input DNA. Briefly, DNA was enzymatically fragmented for 12.5 min and amplified using 6 PCR cycles. Qubit (Invitrogen) and Bioanalyzer (Agilent) were used to check the quality and insert size of the libraries (450 bp). The metagenomes were sequenced at the Functional Genomics Centre Zürich using an S4 flowcell on a NovaSeq (Illumina) platform, producing 150 bp paired end reads.

Bioinformatic processing

Amplicon sequences were quality checked using Trimmomatic v0.36 (Bolger et al., 2014) and assigned to exact Amplicon Sequence Variants (ASVs) using DADA2 v1.10.0 (Callahan et al., 2016) implemented in QIIME2 v2020.8 (Bolyen et al., 2019) with default parameters. Singleton ASVs were removed and taxonomy was assigned using the naïve Bayesian classifier implemented in QIIME2 and the SILVA v138.1 reference database. Non-bacterial ASVs (i.e. chloroplasts, mitochondria, and archaea) were discarded. We built *de novo* a phylogenetic tree using RAxML v8.2.12 (Stamatakis, 2014) and the GTRCAT substitution model and the rapid bootstrapping option. Phylogenetic distances among ASVs were calculated using function *cophenetic* implemented in R v4.1 (R Core Team, 2021).

Metagenomic sequence reads were preprocessed using trim_galore v0.6.6 (Krueger et al., 2021), which uses fastqc v0.11.9 (Andrews, 2010) for quality control and cutadapt v3.4 (Martin, 2011) for adapter removal. Trimmed reads were assembled separately for each sample with megahit v1.2.9 (Li et al., 2015) using default parameters and a minimum contig length of 1,000 bp. Prokaryotic Metagenome-Assembled Genomes (MAGs) were binned using MetaBAT2 v2.15 (Kang et al., 2019) and CONCOCT v1.1 (Alneberg et al., 2014). Reads were mapped to the assemblies with CoverM v0.6.1 (https://github.com/wwood/CoverM). DAS Tool v1.1.2 (Sieber et al., 2018) was used to obtain a non-redundant set of MAGs for each sample. All MAGs were then dereplicated with dRep v3.2.2 (Olm et al., 2017), considering a completeness level greater than 75% and a contamination level less than 10% as determined by checkM v1.1.3 (Parks et al., 2015). Taxonomy of MAGs was assigned using GTDB-Tk v1.7.0 (Chaumeil et al., 2019) based on release 202 obtained from https://data.ace.uq.edu.au/public/gtdb/data/releases/ (Quast et al., 2013). DefenseFinder v1.0.9 (Tesson et al., 2022) using default settings was used to identify antiviral systems on MAGs.

Putative viral contigs (n=57168) were identified from the assemblies using VIBRANT v1.2.1 (Kieft et al., 2020) with default parameters. VIBRANT also provides information about auxiliary metabolic genes on viral contigs based on KEGG annotations (i.e. including "metabolic pathways" and "sulfur relay system"). Complete viral contigs were then identified using CheckV v0.8.1 (Nayfach et al., 2021) and set aside (n=801). The remaining, non-complete viral contigs were binned into viral metagenome-assembled genomes (vMAGs) with PHAMB v1.0.1 (Johansen et al., 2022) which uses VAMB v3.0.2 (Nissen et al., 2021) and DeepVirFinder v1.0 (Ren et al., 2020) to obtain high-quality viral genomes. Subsequently, complete viral genomes and high-quality vMAGs were dereplicated using vRhyme v1.1.0 (Kieft et al., 2022) taking the longest bin and 97% identity to obtain a final set of vMAGs. CheckV was then used to assess vMAG completeness and quality. To obtain vMAG abundances, reads were mapped to the concatenated vMAG sequences using CoverM v0.6.1 in contig mode, using 100% minimum read identity, and the "trimmed_mean" method, which removes 5% of the highest and lowest coverages, respectively. The number of reads mapped to vMAGs are reported in Supplementary Table S1. An additional

detection threshold of 5% of the contig length covered was applied and vMAG abundances below this threshold were set to zero. Coverage was then normalized (i.e. to range between 0 and 1) and treated as relative abundances in downstream analyses. BACPHLIP v0.9.6 (Hockenberry and Wilke, 2021) was used to assign temperate or virulent lifestyles to complete vMAGs. Viral taxonomy was assigned using kaiju v1.9.0 (Menzel et al., 2016) and the *viruses* database (only viruses from the NCBI RefSeq database, accessed 07/05/2022). PHIST v1.1.0 (Zielezinski et al., 2022), an alignment-free, k-mer frequency-based tool was used to predict phage-host interactions (based on $p_{adjusted}$ <0.05). A snakemake workflow for our bioinformatic analyses can be found here: https://github.com/hpeter0803/viromes/tree/ensemble_paper.

Statistical analyses

The statistical software language R v4.1 was used to prepare figures and to perform all statistical analyses. Specifically, viral taxonomic composition was visualized using the heat_tree function of the *metacoder* v*0.3.6* (Foster et al., 2017) R package. Differential abundance analysis for vMAGs was performed using DESeq2 v3.15 (Love et al., 2014). Multivariate analyses of bacterial (16S rRNA gene amplicon) and viral (vMAG) community composition included non-metric multidimensional scaling ordination (*metaMDS*), procrustes correlation (*procrustes* and *protest*), and permutational multivariate analysis of variance using distance matrices (PERMANOVA, *adonis2*) using Bray-Curtis dissimilarities and were performed using the R package *vegan v 2.6-4* (Oksanen et al., 2022) with default settings. Prior to these analyses, Wisconsin double standardization was applied to relative abundances and AMG counts were Hellinger transformed prior to Principal Component Analysis. Phage-host interactions were visualized using Cytoscape v3.10.0 (Shannon et al., 2003).

Results and discussion

Bacterial communities in proglacial streams are characterized by microdiversity

After removal of singletons, we retained 35,170 bacterial ASVs across the three different floodplains and the different stream types. The bacterial communities were taxonomically dominated by members of Gamma- and Alphaproteobacteria (17.4 and 8.2% of ASVs, respectively), Bacteroidia (12.5%), Planctomycetes (7.8%), and Verrucomicrobiae (6.4%). This diversity was further classified into 925 different genera, of which members of *Candidatus Nomurabacteria* (1.8% of ASVs), *Flavobacterium* (1.7%), and *Candidatus Saccharimonadales* (1.4%) accounted for most ASV richness.

Previous work has identified environmental factors common to both GFS and GWS, such as low water temperature (average 6.5 ± 4.6°C) and ultra-oligotrophy to shape proglacial stream biofilm communities (Brandani et al., 2022). Yet, important environmental differences between GFS and GWS, such as turbidity and diurnal variation in stream flow select for different ASVs which account for the compositional (Freimann et al., 2013; Brandani et al., 2022; Brandani et al., 2023) and functional differences (Michoud et al., 2023) between stream types. Leveraging phylogenetic and ecological signatures, we previously related environmental selection to microdiversity, particularly among members of the Gammaproteobacterial family *Comamonadacea* in proglacial streams in New Zealand (Fodelianakis et al., 2022) and the European Alps (Brandani et al., 2023). Here, we re-iterate this and use phylogenetic distances between ASVs, prevalence and relative abundance to identify microdiverse bacterial clades. We found that the genera *Polaromonas*, *Rhodoferax*, *Rhizobacter*, *Methylotenera*, and *Massilia* featured consistently short phylogenetic distances among ASVs (Figure 1A) and are microdiverse. Short phylogenetic distances could be a consequence of the number of ASVs with the same taxonomic classification. Indeed, we found that genera with few ASVs had generally shorter phylogenetic distances than genera containing many ASVs. *Polaromonas* (n_{ASVs}=169), *Rhodoferax* (n_{ASVs}=386), *Rhizobacter* (n_{ASVs}=283), *Methylotenera* (n_{ASVs}=134), and *Massilia* (n_{ASVs}=121), however, contain numerous ASVs, comparable to other, non-microdiverse bacterial genera (Figure 1A). And while these five microdiverse genera account on average (± SD) for only 5.0 ± 2.3% of ASV richness detected in each sample, they contribute 14.5 ± 11.3% to relative abundance (Figures 1B, C).

We obtained 2732 high-quality bacterial MAGs and previous work unraveled the functional differences between MAGs found in GFS and GWS (Michoud et al., 2023). Here, we used MAGs to identify phage-host interactions and screened them for their anti-phage arsenal. Nevertheless, MAGs, while only resolving a subset of the ASVs-resolved microbial diversity, reflected the composition of bacterial communities (Supplementary Figure S1). Specifically, matching ASVs and MAGs with same genus-level classification, mean relative abundance of MAGs was well correlated with relative abundance of ASVs (Spearman's rho = 0.65, p<0.01).

Viral communities in proglacial stream biofilms are diverse

Generally, little is known about viral biodiversity in extreme environments, particularly in mountain streams and rivers (Payne et al., 2020; Bekliz et al., 2022; Busi et al., 2022). Here, we identified 1452 high-quality (n=870) or complete (n=582) vMAGs. vMAG genome size varied considerably, with complete vMAG genome size averaging 47.8 kbp (range 2.6 to 373.3 kbp) and high-quality vMAG size averaging 70.4 kbp (range: 5.0 to 622.1 kbp). Taxonomically, viral communities were dominated by viruses classified as *Caudoviricetes* (Figure 2). *Caudoviricetes* classified as *Ralstonia phage*, *Tsukubavirus*, and *Gordonia phage* were particularly abundant in GFS, whereas, unclassified viruses, *Unahavirus* and *Koutsourovirus* were relatively more abundant in GWS (Supplementary Figure S2A). Besides *Caudoviricetes* (n=930), *Malgrandaviricetes* (n=60), and *Megaviricetes* (n=6) constituted the viral communities.

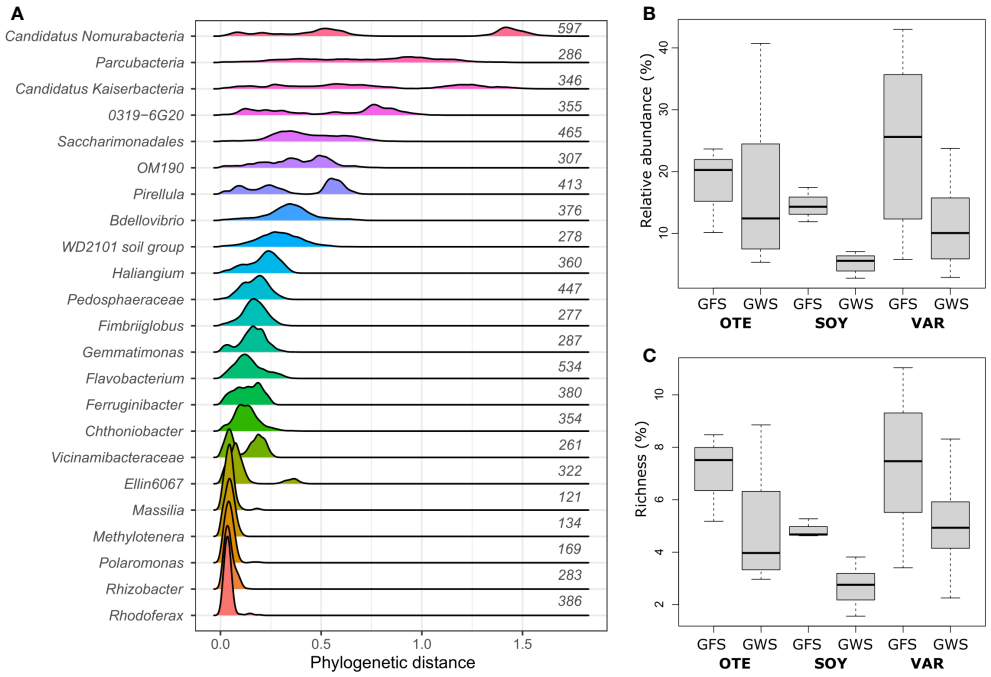


FIGURE 1
Microdiverse bacterial clades. Microdiverse bacterial genera in proglacial streams are abundant, prevalent and characterized by short pairwise phylogenetic distances among ASVs. Shown are the distributions of phylogenetic distances among ASVs with same genus-level taxonomy for the most abundant genera **(A)**. Note that e.g. *Parcubacteria*, *Saccharimonadales, Pedosphaeraceae*, and Vicinamibacteraceae refer to genus-level taxonomic assignments of further unclassified ASVs. The number of ASVs per genus is given to the right of each distribution. Microdiverse genera, i.e. *Rhodoferax*, *Rhizobacter*, *Polaromonas*, *Methylotenera*, and *Massilia*, were particularly abundant in GFS **(B)** while they contributed disproportionately less to overall community richness **(C)**.

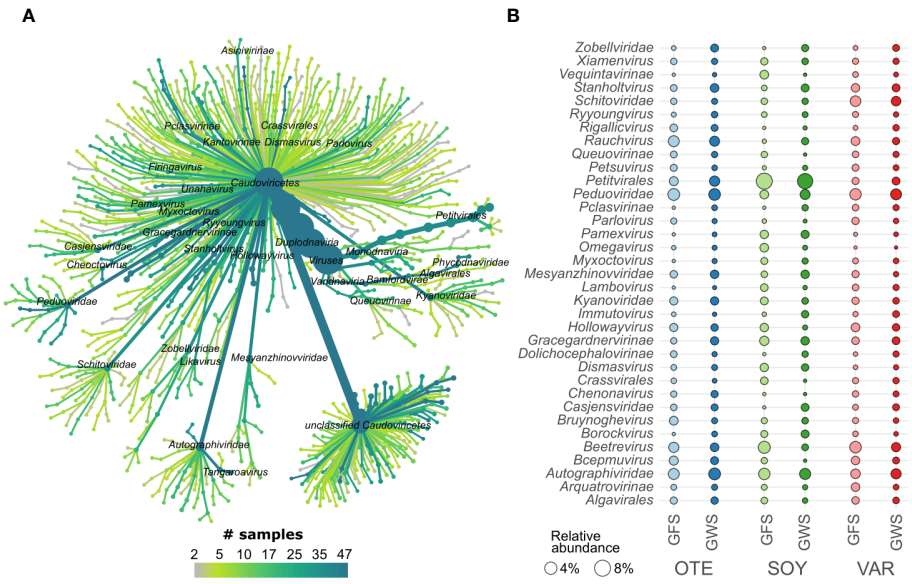


FIGURE 2
Viral taxonomic composition. vMAGs were classified mostly as members of *Caudoviricetes* **(A)**. Displayed is a tree-like representation of viral taxonomy. We found high prevalence (the number of samples in which a specific taxonomic group was detected) across several taxonomic groups. The mean relative abundance (circle size, panel **B)** of the most abundant vMAGs highlights differences across floodplains (e.g. *Autographiviridae, Peduoviridae*, and the eukaryotic viruses *Algalvirales*) and differences between GFS and GWS (e.g. *Petitvirales, Beetrevirus*, and *Kyanoviridae;* see Supplementary Figure S2A for more detail).

On average, 225.4 ± 74.7 vMAGs were detected in each sample with no significant differences in viral richness across stream types (ANOVA, F=0.85, p=0.38) or floodplains (ANOVA, F=1.35, p=0.26) (Supplementary Figure S3). Across the three floodplains, 831 vMAGs were present in GFS samples and 1350 vMAGs were present in GWS samples. While 15 vMAGs were present in all samples, 52 vMAGs were found in both stream types and across all three floodplains, forming the viral proglacial stream biofilm core community (Supplementary Figure S2B). These 52 core vMAGs accounted on average for 9.0% of vMAG relative abundance, with a notable overrepresentation in GFS of OTE, where they represented 33.5 ± 12.9% of vMAG relative abundance. A large number of vMAGs were exclusively found in GWS of SOY (n=264), VAR (n=150), and OTE (n=116) (Supplementary Figure S2B). In contrast, fewer vMAGs were exclusively present in GFS of VAR (n=45), SOY (n=31), and OTE (n=19).

Across all floodplains and stream types, most vMAGs were rare (i.e. low abundance) and only few vMAGs were abundant (i.e. reached more than 0.1% of relative abundance in any given sample), yet, these abundant vMAGs accounted on average for 96.8% of total viral relative abundance (Supplementary Figures S3A, C). Screening for lysogeny-associated protein domains (e.g. integrases and recombinases) on complete vMAG genomes (n=582), we found that 81.0% of them featured a virulent lifestyle, whereas 19.0% followed a temperate lifestyle (Supplementary Figure S4). Contrary to Killing-the-winner and Piggyback-the-winner model expectations (Thingstad and Lignell, 1997; Knowles et al., 2016; Silveira and Rohwer, 2016; reviewed by: Chen et al., 2021), we found no relationship between the ratio of temperate to virulent lifestyles (mean ratio: 0.41 ± 0.21) and bacterial abundance, despite a range of two orders of magnitude in bacterial abundance across our samples (Supplementary Figure S4). Additional work would be needed to better constrain these predictions (i.e. extending beyond complete vMAGs and accounting for novelty). Moreover, bulk metagenomics, as compared to viromics, may better capture actively replicating viruses but may offer only limited resolution to resolve non-infecting or rare viruses (i.e. viruses infecting rare hosts (Roux et al., 2021). Using metagenomics, low phage richness has previously been linked to Clustered Regularly Interspaced Short Palindromic Repeats (CRISPER)-Cas abundance (Meaden et al., 2022), suggesting a trade-off exists between this adaptive bacterial immune response and phage diversity. We screened the bacterial MAGs (n=2732) and identified a total of 18,840 antiviral defense genes on 1947 MAGs. Restriction-modification (n_{genes}=8,705), Cas (n_{genes}=4,483), abortive infection (n_{genes}=1,365), and the Septu defense system (n_{genes}=1,299) were the most commonly found defense types. MAGs contained up to 12 different antiviral defense system types (median number of defense types: 2) highlighting the general importance of antiviral defense for bacterial stream biofilm communities. Moreover, we found a significant relationship between viral diversity and the relative abundance of Cas-containing MAGs (Spearman's rho = 0.49, p=0.0006, Supplementary Figure S5). Taken together, the diverse viral community and versatile antiviral defense arsenal in proglacial stream biofilms suggests a coevolutionary history of phages and biofilm-dwelling bacteria.

Phage communities are coupled to bacterial host communities

Bacterial community composition based on Bray-Curtis similarity obtained from ASV relative abundance significantly differed between floodplains (PERMANOVA, R^2 = 0.08, p<0.001) and stream types (PERMANOVA, R^2 = 0.03, p<0.001), with a significant interaction between floodplain and stream type (PERMANOVA, R^2 = 0.05, p<0.001). Additional variation can be explained by the spatial positioning of samples (e.g. distance to the glacier snout (Brandani et al., 2022), but substantial variation in the bacterial host community remains inexplicable. This highlights the large spatial heterogeneity among nearby streams in proglacial floodplains and the importance of stochastic community assembly processes in nascent ecosystems (Brandani et al., 2023). Moreover, bacterial communities from the Valsorey Glacier catchment were compositionally more distinct from communities of Otemma and Val Roseg, whereas particularly GWS sites of Otemma and Val Roseg were more similar (Figure 3A). This compositional similarity was mirrored by the viral community (Figure 3B), with significant differences between floodplains (PERMANOVA, R^2 = 0.10, p<0.001) and stream types (PERMANOVA, R^2 = 0.04, p<0.001), and a significant interaction term (PERMANOVA, R^2 = 0.06, p<0.001). Superimposition of non-metric multidimensional scaling ordinations further substantiated the congruence between viral and bacterial communities (Procrustes correlation: 0.74, p=0.001, Supplementary Figure S6). The strength of these associations is comparable to earlier reports on the coupling of stream biofilm bacterial and viral communities (Bekliz et al., 2022) and suggests that host availability dominates over environmental conditions in shaping viral communities.

We used phage-host predictions (PHIST) to further resolve the coupling between bacterial and phage communities and predicted a total of 1435 phage-host interactions. Similar to other work (Luo et al., 2022), the majority of phage-host interactions (84.2%) involved a single phage-host pair and only six phages had more than 10 predicted interactions with different MAGs (maximum number of predicted interactions = 13). In line with generally high host specificity, vMAGs with multiple host interactions tended to interact with taxonomically similar MAGs (Figure 3C). For instance, 11 out of 13 interactions of a vMAG classified as *Lacusarxvirus* involved the alpha-proteobacterial CYK-10 genus. The other two interactions were predicted for MAGs classified as *Sphingomonas* (Alphaproteobacteria).

Phage-host interactions involving the microdiverse genera *Polaromonas*, *Rhizobacter*, *Rhodoferax*, *Methylotenera*, and *Massilia* were also dominated by single phage-host pairs (66 out of 110 interactions). Phages interacting with *Rhizobacter*, *Polaromonas*, and *Rhodoferax* also interacted with several other bacterial genera (Supplementary Figure S7). For instance, of the 38 vMAGs putatively interacting with *Rhizobacter*, 22 vMAGs exclusively interacted with *Rhizobacter* whereas 16 vMAGs also interacted with different bacterial genera. However, this was not consistent for all microdiverse bacterial clades: all of the 19 phages predicted to interact with the microdiverse genus *Methylotenera*, exclusively interacted with *Methylotenera* MAGs. Strikingly, only a

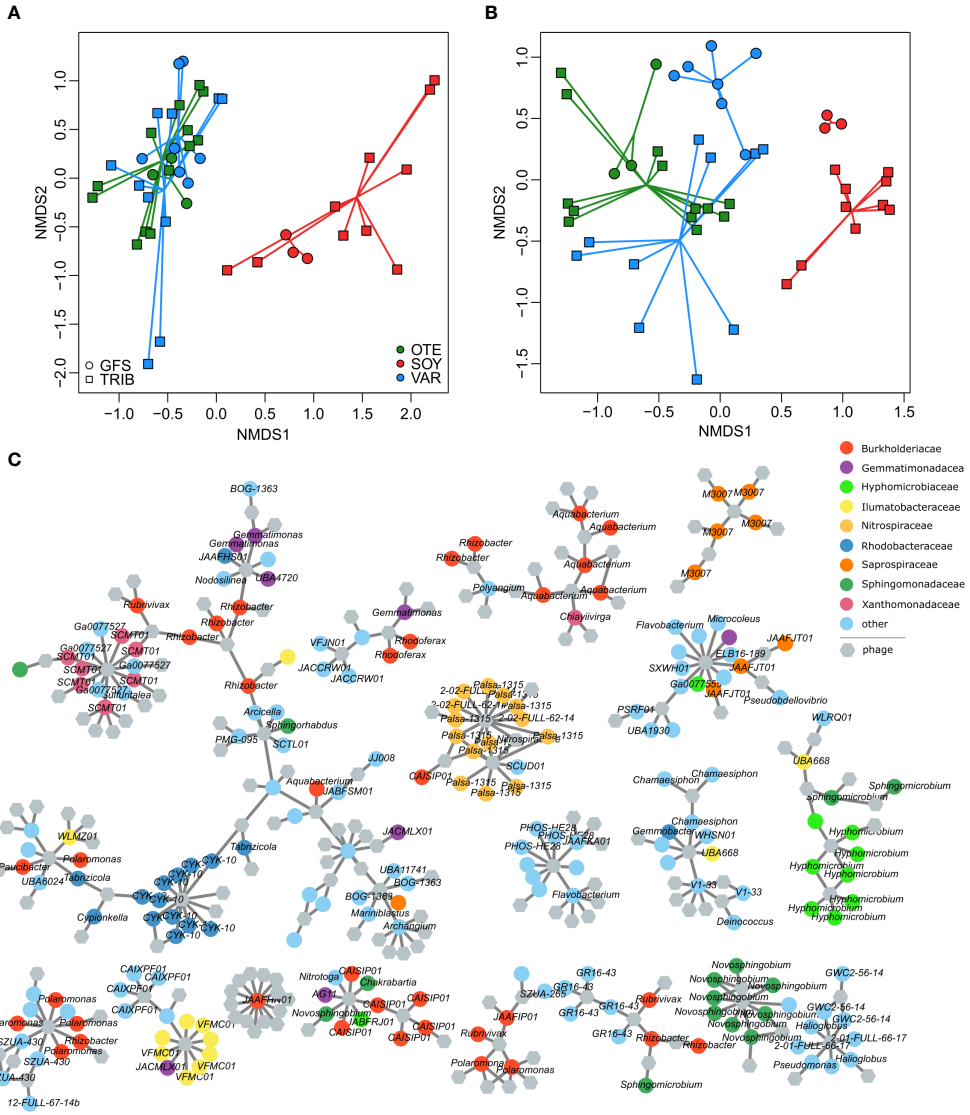


FIGURE 3
Phage-bacteria coupling. Non-metric multidimensional scaling ordinations based on vMAG relative abundance **(A)** and bacterial 16S rRNA gene ASV relative abundance **(B)** reveal the compositional dissimilarities among floodplains and stream types. This separation was more pronounced for the bacterial communities whereas viral communities were more similar (particularly samples of both GFS and GWS in OTE and VAR were similar). Resolving putative phage-host interactions (grey lines) among individual vMAGs (hexagons) and bacterial MAGs (colored circles) illustrate the generally high phage-host specificity (i.e. phages tend to interact with bacteria of similar taxonomy) **(C)**. The 17 largest sub-networks of the entire interaction network are shown.

single phage-host interaction was identified involving the microdiverse genus *Massilia*. Taken together, we discovered a continuum of potential phage host ranges but mostly found evidence for host specificity in proglacial streams.

While the limitations of *in silico* phage-host predictions need to be considered in this context (Brum and Sullivan, 2015; de Jonge et al., 2019), we report here a large number of putative phage-host interactions in stream biofilms, in line with reports from other extreme environments (Munson-McGee et al., 2018; Jarett et al., 2020). Moreover, the continuum of phage host range suggests that multiple eco-evolutionary processes shape bacteria-phage interactions in these communities (Weitz et al., 2013; de Jonge et al., 2019). On the one hand, trade-offs between virulence and host range (de Jonge et al., 2019) or resource limitation (Weitz et al., 2013) can explain narrow host ranges, whereas reduced dispersal and the dense packing of diverse bacteria have been associated with broader host ranges in biofilms (Hwang et al., 2023). In communities characterized by microdiversity, narrow host ranges may be beneficial for phages given the increased encounter probability with phylogenetically closely related members. Phages interacting with the microdiverse bacterial genus *Methylotenera* may, for instance, benefit from such narrow host range. On the other hand, broad host ranges may increase the likelihood of exchanging genetic material across distant bacterial populations. This may, for example, be the case for the microdiverse genus *Rhizobacter* in proglacial stream biofilms. We may thus hypothesize that beneficial effects of broad host ranges could contribute to the longer-term success of microdiverse bacterial clades. Microdiversity

implies that phylogenetically closely related taxa occupy available niches. Typically, such closely related taxa are expected to compete for similar resources (Larkin and Martiny, 2017) and phage-mediated exchange of genetic material, including auxiliary metabolic genes, may alleviate competition among microdiverse clade members.

Specific auxiliary metabolic genes are associated with microdiverse bacterial genera

In total, we detected 938 AMGs on 448 vMAGs, broadly reflecting the taxonomic composition of the entire vMAG community. In contrast to a recent analysis of Pearl River Estuary viromes (Luo et al., 2022), we did not detect differences in the number of AMGs between virulent and temperate phages. Specifically, 318 out of 1041 virulent vMAGs and 130 out of 411 temperate vMAGs harbored at least one AMG. When accounting for viral lifestyle distribution (i.e. 81.0% of vMAGs were virulent), both viral lifestyles represented a similar number of AMGs. Moreover, AMGs abundant among virulent phages were also abundant among temperate phages (Spearman's rho = 0.72, $p<0.0087$).

AMGs encoded for a large functional diversity, with in total 166 different KEGG orthologues (KOs) present. Particularly numerous were AMGs encoding *dcm* (159 AMGs) which in prokaryotes can be involved in DNA methylation in restriction-modification systems (Pingoud et al., 2014). However, in phages, so-called orphan methyltransferases (i.e., without endonucleases of the prokaryotic restriction-modification system) can contribute to the protection of phage DNA against digestion (Heyerhoff et al., 2022). Other common AMGs were *queE, queD, queE, and folE* genes (137 AMG), involved in the synthesis of the deazapurine nucleoside $preQ_O$, which also has been shown to protect viruses from restriction enzymes (Hutinet et al., 2019; Kieft et al., 2020). In contrast to pelagic systems (Kieft et al., 2021b; Heyerhoff et al., 2022) and despite the importance of cross-domain interactions in proglacial stream biofilms (Busi et al., 2022), however, AMGs encoding the core photosystem II proteins *psbA* and *psbD* were rare, with only 3 AMGs present across our entire dataset. The majority (34.4%) of AMGs encoded KOs involved in the metabolism of cofactors and vitamins including AMGs involved in folate biosynthesis, the metabolisms of nicotinate/nicotinamide, and porphyrine/chlorophyll. AMGs involved in amino acid metabolism (cysteine and methionine) accounted for another 20.7% of all AMGs. AMGs involved in carbohydrate metabolism (15.4%), glycan biosynthesis (8.4%), and sulfur metabolism (4.3%) were also commonly detected. Similar to marine sediments (Heyerhoff et al., 2022), *cysH*, which is involved in the synthesis of sulfite during assimilatory sulfate reduction, dominated AMGs involved in sulfur metabolism. *cysH* ranks among the globally conserved AMGs (Kieft et al., 2020) and has the potential to influence ecosystem-scale sulfur metabolism (Kieft et al., 2021a). This may be particularly relevant in GFS, where high turbidity and the absence of light fosters chemolithoautotrophic bacterial metabolisms. Chemolithoautotrophic bacteria can oxidize reduced sulfur to fix CO_2 and to produce biomass. In the absence of photosynthetic organisms in GFS, chemolithoautotrophic carbon production may thus be an important ecosystem function, and phage AMGs may play a pivotal role.

We did not detect significant differences in AMGs between viral communities from different floodplains or stream types (Figure 4A). Abundant AMG families, including the above-mentioned AMGs, tended to be present and abundant in most samples, whereas other AMG families were only sporadically present and often at low relative abundance. However, when coupling AMG with phage-host predictions (17.3% of AMGs were on vMAGs with a host prediction), we found that common and abundant AMGs were found on phages associated with specific bacterial community members (Figure 4B). For instance, 19 AMGs encoding folate biosynthesis were found on phages associated with 6 bacterial genera, including members of *Novosphingobium*, *Calothrix*, and *Rhodobacter*. Similarly, AMGs encoding for cysteine and methionine metabolism were particularly prevalent among vMAGs infecting members of the bacterial genera *Polaromonas*, *Microcoleus*, *Ferruginibacter*, and *Segetibacter*. Finally, using Principal Component Analysis with all AMGs for which phage-host predictions were available, we found distinct groups of bacterial genera to be host to phages with specific AMGs (Figure 4C). For example, the two microdiverse bacterial genera *Methylotenera* and *Rhizobacter*, together with *Lysobacter* and *Sandarakinorhabdus* formed a distinct group dominated by AMGs involved in porphyrine and chlorophyll metabolism. Another microdiverse bacterial genus, *Rhodoferax*, together with several other, non-microdiverse bacterial genera formed a distinct cluster based on AMGs involved in the metabolism of cysteine and methionine. AMGs involved in the biosynthesis of folate and lipopolysaccharides, explained clusters of bacterial genera involving members such as *Rhodobacter*, *Nitrospira*, *Novosphingobium*, and *Calothrix* as well as a cluster composed of *Methylibium*, *Pedosphaera*, and *Ideonella*. Interestingly, one of the most common and microdiverse bacterial genus in proglacial stream biofilms, *Polaromonas*, did not associate with any of the clusters. These findings highlight that similar AMGs may be found on phages associated with microdiverse and non-microdiverse clades.

Despite their compact genomes, phages regularly encode AMGs which can provide them a fitness advantage by augmenting or redirecting specific metabolic processes (Dion et al., 2020; Kieft et al., 2021b). This highlights the reciprocal benefits of phage-host interactions, which manifest during coevolution (Koskella and Brockhurst, 2014). The selective environmental constraints in proglacial streams have persisted over geological time scales, shaping microbial diversity and putatively influencing phage-bacteria coevolution. Host specificity of AMGs has been previously reported (Luo et al., 2022) and here we report specific AMGs linked with microdiverse (and non-microdiverse) clades in proglacial stream biofilms. Arguably, microdiverse clades and their phages have shared a long evolutionary time in a selective environment, refining the AMGs library and putatively contributing to the eco-evolutionary success of both, the phages and their prokaryotic hosts. Microdiversity is an intrinsic property of many microbial communities and horizontal gene transfer can

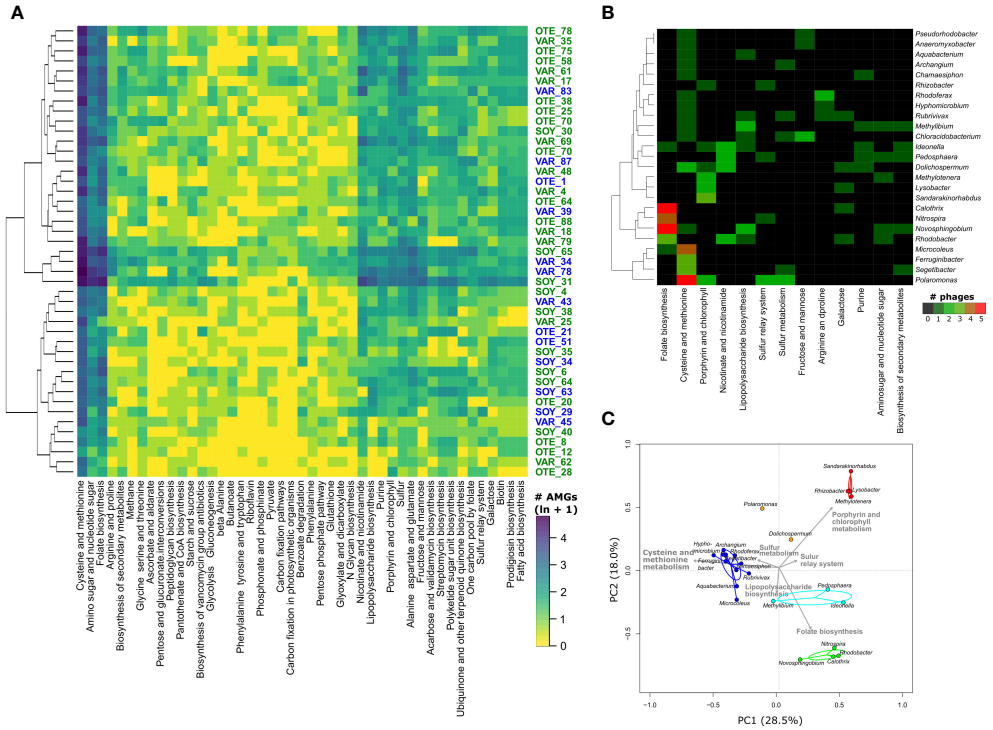


FIGURE 4
Auxiliary metabolic genes. A heatmap of the number of different AMGs found in the different samples highlights the prevalence of AMGs involved in the metabolism of cysteine and methionine, amino sugars and nucleotide sugar and folate biosynthesis **(A)**. The colors represent ln(x+1) transformed number of AMGs. Despite the pronounced compositional difference between GFS and GWS, no differences in AMG composition between GFS (blue sample IDs) and GWS (green sample IDs) were apparent. The dendrogram to the left of the heatmap shows the results of single-linkage clustering, hierarchically connecting samples based on similar AMG distributions. Coupling AMGs with phage-host interactions revealed AMGs shared by phages putatively infecting bacterial MAGs **(B)**. Shown are the 13 most common AMG-encoded pathways for 25 bacterial genera. Colors reflect the number of vMAGs. Considering all phage-host interactions and AMGs in a Principal Component Analysis, further highlighted the pronounced differences in phages carrying AMGs and infecting specific bacterial genera **(C)** Based on the number of AMGs on phages, different bacterial host genera (circles) cluster into distinct groups. Colored lines connect host genera to cluster centroids. Ellipses denote 95% confidence intervals for these clusters. Grey arrows denote KO pathways distinctive for the different clusters. For instance, the two microdiverse genera *Rhizobacter* and *Methylotenera* cluster with *Sandarakinorhabdus* and *Lysobacter* in terms of AMGs found on their respective phages. These AMGs are enriched in genes encoding porphyrin and chlorophyll metabolism. Note the distribution of microdiverse bacterial genera in different clusters and the absence of the microdiverse bacterial genus *Massilia*.

generate microdiversity (Larkin and Martiny, 2017). Phage-mediated HGT transfer may therefore contribute to microdiversification and narrow host range and AMG specificity suggests that in proglacial stream biofilms, virus-host interactions may foster 'Maestro Microbes' (Larkin and Martiny, 2017), which occupy niches by optimizing specific traits (in contrast to 'Renaissance Microbes' which acquire new traits (Larkin and Martiny, 2017).

Conclusions

Generally, little is known about viral biodiversity in extreme environments, particularly however in streams and rivers draining mountain regions (Payne et al., 2020; Bekliz et al., 2022; Busi et al., 2022). Here, we expand the current knowledge of phage diversity and its role in shaping microbial communities in streams and rivers. Biofilms are thought to represent a physical barrier for viruses (Abedon, 2011), and our work unravels viral communities, which are shaped by their bacterial host community composition. Extreme environmental conditions in proglacial streams lead to strong selective sorting of these host communities, ultimately resulting in communities dominated by few but well-adapted and microdiverse clades. Contrary to our expectation of generally narrow phage-host ranges in communities dominated by microdiverse bacterial clades, we found a continuum of host ranges, which we attribute to different eco-evolutionary dynamics, putatively associated with resource limitation in proglacial streams (explaining narrow host ranges) and the biofilm mode of life (putatively explaining broader host ranges). Phages infecting members of different microdiverse bacterial genera carry AMGs, which appear to be distinct and specific. Future work may further elucidate the role of non-microdiverse clades, which share similar AMGs and assess to which degree the coevolution between phages and their hosts shapes the AMG repertoire. This may be particularly relevant in extreme environments, where stark selective processes shape communities. On the other hand, microdiversity is a hallmark of many microbial communities, and beneficial phages-host interactions may play important roles in generating and maintaining microdiversity.

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/, PRJNA808857. Supplementary Table S1 lists individual BioSample accession numbers. Raw bioinformatic output is available at Figshare under doi 10.6084/m9.figshare.24511972.

Author contributions

HP: Conceptualization, Formal Analysis, Investigation, Methodology, Visualization, Writing – original draft. GM: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Writing – original draft. SB: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Writing – original draft. TB: Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Writing – original draft.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. Funding supporting this work was provided by the Swiss National Science Foundation (grant CRSII5_180241) to TJB.

Acknowledgments

We thank Jade Brandani, Nicola Deluigi, Kevin Casellini, and Paraskevi Pramateftaki for assistance in the field and laboratory. We further thank Stylianos Fodelianakis, Tyler Kohler, Massimo Bourquin, Leila Ezzat, Martin Boutroux, and two reviewers for fruitful discussions and feedback on the manuscript.

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 author(s) SB, GM declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frmbi.2023.1279550/full#supplementary-material

References

Abedon, S. T. (2011). Bacteriophages and biofilms: ecology, phage therapy, plaques. *Nova Science Publishers*.

Alneberg, J., Bjarnason, B. S., de Bruijn, I., Schirmer, M., Quick, J., Ijaz, U. Z., et al. (2014). Binning metagenomic contigs by coverage and composition. *Nat. Methods* 11, 1144–1146. doi: 10.1038/nmeth.3103

Andrews, S. (2010) *FastQC: A Quality Control Tool for High Throughput Sequence Data*. Available at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

Battin, T. J., Besemer, K., Bengtsson, M. M., Romani, A. M., and Packmann, A. I. (2016). The ecology and biogeochemistry of stream biofilms. *Nat. Rev. Microbiol.* 14, 251–263. doi: 10.1038/nrmicro.2016.15

Bekliz, M., Pramateftaki, P., Battin, T. J., and Peter, H. (2022). Viral diversity is linked to bacterial community composition in alpine stream biofilms. *ISME Commun* 2, 27. doi: 10.1038/s43705-022-00112-9

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

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

Brandani, J., Peter, H., Busi, S. B., Kohler, T. J., Fodelianakis, S., Ezzat, L., et al. (2022). Spatial patterns of benthic biofilm diversity among streams draining proglacial floodplains. *Front. Microbiol.* 13. doi: 10.3389/fmicb.2022.948165

Brandani, J., Peter, H., Fodelianakis, S., Kohler, T. J., Bourquin, M., Michoud, G., et al. (2023). Homogeneous environmental selection structures the bacterial communities of benthic biofilms in proglacial floodplain streams. *Appl. Environ. Microbiol.* 89, e02010–e02022. doi: 10.1128/aem.02010-22

Brum, J. R., and Sullivan, M. B. (2015). Rising to the challenge: accelerated pace of discovery transforms marine virology. *Nat. Rev. Microbiol.* 13, 147–159. doi: 10.1038/nrmicro3404

Busi, S. B., Bourquin, M., Fodelianakis, S., Michoud, G., Kohler, T. J., Peter, H., et al. (2022). Genomic and metabolic adaptations of biofilms to ecological windows of opportunity in glacier-fed streams. *Nat. Commun.* 13, 2168. doi: 10.1038/s41467-022-29914-0

Busi, S. B., Pramateftaki, P., Brandani, J., Fodelianakis, S., Peter, H., Halder, R., et al. (2020). Optimised biomolecular extraction for metagenomic analysis of microbial biofilms from high-mountain streams. *PeerJ* 8, e9973. doi: 10.7717/peerj.9973

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

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

Chen, X., Weinbauer, M. G., Jiao, N., and Zhang, R. (2021). Revisiting marine lytic and lysogenic virus-host interactions: Kill-the-Winner and Piggyback-the-Winner. *Sci. Bull.* 66, 871–874. doi: 10.1016/j.scib.2020.12.014

de Jonge, P. A., Nobrega, F. L., Brouns, S. J. J., and Dutilh, B. E. (2019). Molecular and evolutionary determinants of bacteriophage host range. *Trends Microbiol.* 27, 51–63. doi: 10.1016/j.tim.2018.08.006

Dion, M. B., Oechslin, F., and Moineau, S. (2020). Phage diversity, genomics and phylogeny. *Nat. Rev. Microbiol.* 18, 125–138. doi: 10.1038/s41579-019-0311-5

Ezzat, L., Fodelianakis, S., Kohler, T. J., Bourquin, M., Brandani, J., Busi, S. B., et al. (2022). Benthic biofilms in glacier-fed streams from Scandinavia to the Himalayas host

distinct bacterial communities compared with the streamwater. *Appl. Environ. Microbiol.* 88, e00421–e00422. doi: 10.1128/aem.00421-22

Fodelianakis, S., Washburne, A. D., Bourquin, M., Pramateftaki, P., Kohler, T. J., Styllas, M., et al. (2022). Microdiversity characterizes prevalent phylogenetic clades in the glacier-fed stream microbiome. *ISME J.* 16, 666–675. doi: 10.1038/s41396-021-01106-6

Foster, Z. S. L., Sharpton, T. J., and Grünwald, N. J. (2017). Metacoder: An R package for visualization and manipulation of community taxonomic diversity data. *PLoS Comput. Biol.* 13, e1005404. doi: 10.1371/journal.pcbi.1005404

Freimann, R., Bürgmann, H., Findlay, S. E., and Robinson, C. T. (2013). Bacterial structures and ecosystem functions in glaciated floodplains: contemporary states and potential future shifts. *ISME J.* 7, 2361–2373. doi: 10.1038/ismej.2013.114

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.e14. doi: 10.1016/j.cell.2019.03.040

Heyerhoff, B., Engelen, B., and Bunse, C. (2022). Auxiliary metabolic gene functions in pelagic and benthic viruses of the Baltic Sea. *Front. Microbiol.* 13. doi: 10.3389/fmicb.2022.863620

Hockenberry, A. J., and Wilke, C. O. (2021). BACPHLIP: predicting bacteriophage lifestyle from conserved protein domains. *PeerJ* 9, e11396. doi: 10.7717/peerj.11396

Holtappels, D., Alfenas-Zerbini, P., and Koskella, B. (2023). Drivers and consequences of bacteriophage host range. *FEMS Microbiol. Rev.* 47, fuad038. doi: 10.1093/femsre/fuad038

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

Howard-Varona, C., Hargreaves, K. R., Abedon, S. T., and Sullivan, M. B. (2017). Lysogeny in nature: mechanisms, impact and ecology of temperate phages. *ISME J.* 11, 1511–1520. doi: 10.1038/ismej.2017.16

Hutinet, G., Kot, W., Cui, L., Hillebrand, R., Balamkundu, S., Gnanakalai, S., et al. (2019). 7-Deazaguanine modifications protect phage DNA from host restriction systems. *Nat. Commun.* 10, 5442. doi: 10.1038/s41467-019-13384-y

Hwang, Y., Roux, S., Coclet, C., Krause, S. J. E., and Girguis, P. R. (2023). Viruses interact with hosts that span distantly related microbial domains in dense hydrothermal mats. *Nat. Microbiol.* 8, 946–957. doi: 10.1038/s41564-023-01347-5

Jarett, J. K., Džunková, M., Schulz, F., Roux, S., Paez-Espino, D., Eloe-Fadrosh, E., et al. (2020). Insights into the dynamics between viruses and their hosts in a hot spring microbial mat. *ISME J.* 14, 2527–2541. doi: 10.1038/s41396-020-0705-4

Johansen, J., Plichta, D. R., Nissen, J. N., Jespersen, M. L., Shah, S. A., Deng, L., et al. (2022). Genome binning of viral entities from bulk metagenomics data. *Nat. Commun.* 13, 965. doi: 10.1038/s41467-022-28581-5

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

Kieft, K., Adams, A., Salamzade, R., Kalan, L., and Anantharaman, K. (2022). vRhyme enables binning of viral genomes from metagenomes. *Nucleic Acids Res.* 50, e83. doi: 10.1093/nar/gkac341

Kieft, K., Breister, A. M., Huss, P., Linz, A. M., Zanetakos, E., Zhou, Z., et al. (2021a). Virus-associated organosulfur metabolism in human and environmental systems. *Cell Rep.* 36, 109471. doi: 10.1016/j.celrep.2021.109471

Kieft, K., Zhou, Z., and Anantharaman, K. (2020). VIBRANT: automated recovery, annotation and curation of microbial viruses, and evaluation of viral community function from genomic sequences. *Microbiome* 8, 90. doi: 10.1186/s40168-020-00867-0

Kieft, K., Zhou, Z., Anderson, R. E., Buchan, A., Campbell, B. J., Hallam, S. J., et al. (2021b). Ecology of inorganic sulfur auxiliary metabolism in widespread bacteriophages. *Nat. Commun.* 12, 3503. doi: 10.1038/s41467-021-23698-5

Klindworth, A., Pruesse, E., Schweer, T., Peplies, J., Quast, C., Horn, M., et al. (2013). Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res.* 41, e1. doi: 10.1093/nar/gks808

Knowles, B., Silveira, C. B., Bailey, B. A., Barott, K., Cantu, V. A., Cobián-Güemes, A. G., et al. (2016). Lytic to temperate switching of viral communities. *Nature* 531, 466–470. doi: 10.1038/nature17193

Koskella, B., and Brockhurst, M. A. (2014). Bacteria-phage coevolution as a driver of ecological and evolutionary processes in microbial communities. *FEMS Microbiol. Rev.* 38, 916–931. doi: 10.1111/1574-6976.12072

Krueger, F., James, F., Ewels, P., Afyounian, E., and Schuster-Boeckler, B. (2021). *TrimGalore: v0.6.7*. doi: 10.5281/zenodo.5127899

Larkin, A. A., and Martiny, A. C. (2017). Microdiversity shapes the traits, niche space, and biogeography of microbial taxa. *Environ. Microbiol. Rep.* 9, 55–70. doi: 10.1111/1758-2229.12523

Li, D., Liu, C.-M., Luo, R., Sadakane, K., and Lam, T.-W. (2015). MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* 31, 1674–1676. doi: 10.1093/bioinformatics/btv033

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, 550. doi: 10.1186/s13059-014-0550-8

Luo, X.-Q., Wang, P., Li, J.-L., Ahmad, M., Duan, L., Yin, L.-Z., et al. (2022). Viral community-wide auxiliary metabolic genes differ by lifestyles, habitats, and hosts. *Microbiome* 10, 190. doi: 10.1186/s40168-022-01384-y

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* 17, 10–12. doi: 10.14806/ej.17.1.200

Meaden, S., Biswas, A., Arkhipova, K., Morales, S. E., Dutilh, B. E., Westra, E. R., et al. (2022). High viral abundance and low diversity are associated with increased CRISPR-Cas prevalence across microbial ecosystems. *Curr. Biol.* 32, 220–227.e5. doi: 10.1016/j.cub.2021.10.038

Menzel, P., Ng, K. L., and Krogh, A. (2016). Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nat. Commun.* 7, 11257. doi: 10.1038/ncomms11257

Michoud, G., Kohler, T. J., Peter, H., Brandani, J., Busi, S. B., and Battin, T. J. (2023). Unexpected functional diversity of stream biofilms within and across proglacial floodplains despite close spatial proximity. *Limnology Oceanography* 68, 2183–2194. doi: 10.1002/lno.12415

Munson-McGee, J. H., Peng, S., Dewerff, S., Stepanauskas, R., Whitaker, R. J., Weitz, J. S., et al. (2018). A virus or more in (nearly) every cell: ubiquitous networks of virus-host interactions in extreme environments. *ISME J.* 12, 1706–1714. doi: 10.1038/s41396-018-0071-7

Nayfach, S., Camargo, A. P., Schulz, F., Eloe-Fadrosh, E., Roux, S., and Kyrpides, N. C. (2021). CheckV assesses the quality and completeness of metagenome-assembled viral genomes. *Nat. Biotechnol.* 39, 578–585. doi: 10.1038/s41587-020-00774-7

Needham, D. M., Sachdeva, R., and Fuhrman, J. A. (2017). Ecological dynamics and co-occurrence among marine phytoplankton, bacteria and myoviruses shows microdiversity matters. *ISME J.* 11, 1614–1629. doi: 10.1038/ismej.2017.29

Nissen, J. N., Johansen, J., Allesøe, R. L., Sønderby, C. K., Armenteros, J. J. A., Grønbech, C. H., et al. (2021). Improved metagenome binning and assembly using deep variational autoencoders. *Nat. Biotechnol.* 39, 555–560. doi: 10.1038/s41587-020-00777-4

Obeng, N., Pratama, A. A., and van Elsas, J. D. (2016). The significance of mutualistic phages for bacterial ecology and evolution. *Trends Microbiol.* 24, 440–449. doi: 10.1016/j.tim.2015.12.009

Oksanen, J., Simpson, G. L., Blanchet, G. F., Kindt, R., Legendre, P., et al. (2022) *vegan: Community ecology package. R package version 2.6-4*. Available at: https://CRAN.R-project.org/package=vegan.

Olm, M. R., Brown, C. T., Brooks, B., and Banfield, J. F. (2017). dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. *ISME J.* 11, 2864–2868. doi: 10.1038/ismej.2017.126

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

Payne, A. T., Davidson, A. J., Kan, J., Peipoch, M., Bier, R., and Williamson, K. (2020). Widespread cryptic viral infections in lotic biofilms. *Biofilm* 2, 100016. doi: 10.1016/j.bioflm.2019.100016

Pingoud, A., Wilson, G. G., and Wende, W. (2014). Type II restriction endonucleases–a historical perspective and more. *Nucleic Acids Res.* 42, 7489–7527. doi: 10.1093/nar/gku447

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 (2021). *R: A language and environment for statistical computing* (Vienna, Austria: R Foundation for Statistical Computing). Available at: https://www.R-project.org/.

Ren, J., Song, K., Deng, C., Ahlgren, N. A., Fuhrman, J. A., Li, Y., et al. (2020). Identifying viruses from metagenomic data using deep learning. *Quantitative Biol.* 8, 64–77. doi: 10.1007/s40484-019-0187-4

Roux, S., Brum, J. R., Dutilh, B. E., Sunagawa, S., Duhaime, M. B., Loy, A., et al. (2016). Ecogenomics and potential biogeochemical impacts of globally abundant ocean viruses. *Nature* 537, 689–693. doi: 10.1038/nature19366

Roux, S., Matthijnssens, J., and Dutilh, B. E. (2021). Metagenomics in virology. *Encyclopedia Virol.* Dennis H. Bamford, Mark Zuckerman, Encyclopedia of Virology (Fourth Edition), Academic Press. 133–140. doi: 10.1016/B978-0-12-809633-8.20957-6

Schwartz, D. A., Lehmkuhl, B. K., and Lennon, J. T. (2022). Phage-encoded sigma factors alter bacterial dormancy. *mSphere* 7, e00297–e00222. doi: 10.1128/msphere.00297-22

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13, 2498–2504. doi: 10.1101/gr.1239303

Shkoporov, A. N., Turkington, C. J., and Hill, C. (2022). Mutualistic interplay between bacteriophages and bacteria in the human gut. *Nat. Rev. Microbiol.* 20, 737–749. doi: 10.1038/s41579-022-00755-4

Sieber, C. M. K., Probst, A. J., Sharrar, A., Thomas, B. C., Hess, M., Tringe, S. G., et al. (2018). Recovery of genomes from metagenomes via a dereplication, aggregation and scoring strategy. *Nat. Microbiol.* 3, 836–843. doi: 10.1038/s41564-018-0171-1

Silveira, C. B., and Rohwer, F. L. (2016). Piggyback-the-Winner in host-associated microbial communities. *NPJ Biofilms Microbiomes* 2, 1–5. doi: 10.1038/npjbiofilms.2016.10

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

Suttle, C. (1994). The significance of viruses to mortality in aquatic microbial communities. *Microbial Ecol.* 28, 237–243. doi: 10.1007/BF00166813

Suttle, C. A. (2007). Marine viruses — major players in the global ecosystem. *Nat. Rev. Microbiol.* 5, 801–812. doi: 10.1038/nrmicro1750

Tesson, F., Hervé, A., Mordret, E., Touchon, M., d'Humières, C., Cury, J., et al. (2022). Systematic and quantitative view of the antiviral arsenal of prokaryotes. *Nat. Commun.* 13, 2561. doi: 10.1038/s41467-022-30269-9

Thingstad, T., and Lignell, R. (1997). Theoretical models for the control of bacterial growth rate, abundance, diversity and carbon demand. *Aquat. Microbial Ecol.* 13, 19–27. doi: 10.3354/ame013019

Weitz, J. S., Poisot, T., Meyer, J. R., Flores, C. O., Valverde, S., Sullivan, M. B., et al. (2013). Phage-bacteria infection networks. *Trends Microbiol.* 21, 82–91. doi: 10.1016/j.tim.2012.11.003

Zielezinski, A., Deorowicz, S., and Gudyś, A. (2022). PHIST: fast and accurate prediction of prokaryotic hosts from metagenomic viral sequences. *Bioinformatics* 38, 1447–1449. doi: 10.1093/bioinformatics/btab837

Zimmerman, A. E., Howard-Varona, C., Needham, D. M., John, S. G., Worden, A. Z., Sullivan, M. B., et al. (2020). Metabolic and biogeochemical consequences of viral infection in aquatic ecosystems. *Nat. Rev. Microbiol.* 18, 21–34. doi: 10.1038/s41579-019-0270-x

frontiers | Frontiers in Microbiomes

TYPE Original Research
PUBLISHED 07 March 2024
DOI 10.3389/frmbi.2024.1324967

OPEN ACCESS

EDITED BY
Gareth Trubl,
Lawrence Livermore National Laboratory (DOE), United States

REVIEWED BY
Boonfei Tan,
Singapore-MIT Alliance for Research and Technology (SMART), Singapore
Brooke Peyton Stemple,
Colorado State University, United States

*CORRESPONDENCE
Lisa Voskuhl
lisa.voskuhl@uni-due.de
Rainer U. Meckenstock
rainer.meckenstock@uni-due.de

†PRESENT ADDRESS
Mark Pannekens,
IWW Rheinisch-Westfälisches Institut für Wasserforschung gemeinnützige GmbH, Mülheim an der Ruhr, Germany

RECEIVED 20 October 2023
ACCEPTED 13 February 2024
PUBLISHED 07 March 2024

CITATION
Beilig S, Pannekens M, Voskuhl L and Meckenstock RU (2024) Assessing anaerobic microbial degradation rates of crude light oil with reverse stable isotope labelling and community analysis.
Front. Microbiomes 3:1324967.
doi: 10.3389/frmbi.2024.1324967

Assessing anaerobic microbial degradation rates of crude light oil with reverse stable isotope labelling and community analysis

Sebastian Beilig, Mark Pannekens†, Lisa Voskuhl* and Rainer U. Meckenstock*

Department of Chemistry, Environmental Microbiology and Biotechnology (EMB) - Aquatic Microbiology, University of Duisburg-Essen, Essen, Germany

Oil reservoirs represent extreme environments where anaerobic degradation profoundly influences oil composition and quality. Despite the common observation of biodegraded oil, the microbial degradation rates remain largely unknown. To address this knowledge gap, we conducted microcosm incubations with light oil as carbon source, original formation water and sulfate as electron acceptor, closely mimicking *in situ* conditions to assess oil degradation rates. Samples were taken from a newly drilled oil well to exclude contamination with injection water and allochthonous microorganisms. At the end of the incubations, microbial community analyses with 16S rRNA gene amplicon sequencing revealed the most prominent phyla as Desulfobacterota, Thermotogota, Bacteroidota, Bacillota (formerly Firmicutes), and Synergistota, collectively accounting for up to 44% of relative abundance. Ion chromatography and reverse stable isotope labeling were used to monitor sulfate reduction and CO_2 evolution respectively. We calculated an average degradation rate of 0.35 mmol CO_2 per year corresponding to 15.2 mmol CO_2/mol $CH_{2(oil)}$ per year. This resembles to approximately 200 years to degrade one gram of oil under the applied, presumably ideal conditions. Factoring in the available oil-water-contact (OWC) zone within the incubations yielded a degradation rate of 120 g CH_2 m^{-2} OWC per year, closely aligning with the modeled degradation rates typically observed in oil reservoirs. Moreover, our study highlighted the utility of the reverse stable isotope labeling (RSIL) approach for measuring complex substrate degradation at minute rates.

KEYWORDS

sulfate reduction, community composition, oil reservoir microbiology, 16S rRNA amplicon gene, microbial community, oil microbiome, RSIL

Introduction

Oil reservoirs are extreme habitats due to saturated hydrocarbon concentrations, low water activity and anoxic conditions (Pannekens et al., 2019). Nevertheless, microbial degradation is taking place in oil reservoirs where microorganisms favor the relatively easier-to-degrade light oil components such as alkanes over the more recalcitrant polycyclic aromatic hydrocarbons (PAHs) or asphaltenes (Larter et al., 2003; Mbadinga et al., 2011; Kayukova et al., 2016).

The microbial oil degradation alters the chemical composition of the oil in moderately temperate reservoirs leading to heavier oil or even bitumen with increased density (Head et al., 2003; Aitken et al., 2004; Youssef et al., 2009; Rajbongshi and Gogoi, 2021). Such processes take place over geological time scales, which makes it very challenging to assess the actual degradation processes and rates in oil reservoirs (Head et al., 2003; Larter et al., 2003; Aitken et al., 2004). Hence, microbial activities and rates resulting in hydrocarbon degradation in the subsurface remain still unknown. So far, there have been only a few studies that aimed at modeling *in situ* oil degradation rates at the oil water transition zone (Head et al., 2003; Larter et al., 2003) or directly determining the rate within a heavily degraded bitumen reservoir (Golby et al., 2012; Pannekens et al., 2021). Oil degradation rates certainly depend on several factors such as the reservoir temperature, oil column height, oil gravity (API) and chemical composition, diffusion coefficients in the porous media and maybe most importantly the oil-water distribution and interface area (Larter et al., 2006; Pannekens et al., 2019).

Biodegradation processes in oil reservoirs can be assessed by different methods such as oil-charging biodegradation models (Larter et al., 2003) or monitoring changes in the ratios of linear to branched alkanes (phytane or pristine) with gas chromatography – mass spectroscopy (GC-MS) (Kondyli and Schrader, 2019). Alternatively, the overall chemical composition of the crude oil can be analyzed through Fourier transform ion cyclotron mass spectrometry (FT-ICR-MS) analyses (Meckenstock et al., 2014; Cho et al., 2015). These qualitative measurements allow to assess the degradation state of oil samples from reservoirs but quantitative assessment of degradation rates is very difficult. Measurements of degradation rates have mostly been performed with laboratory incubations of oil samples where methane production or sulfate depletion were measured (Gieg et al., 2010). Methane can be very easily and sensitively measured with GC-analysis but sulfate measurements are often not accurate enough to monitor the minute metabolic rates that occur in anaerobic oil degradation. Alternative methods are thus needed to assess degradation rates of hydrocarbons in microcosm experiments.

We recently introduced the monitoring of hydrocarbon mineralization to carbon dioxide with reverse stable isotope labeling (RSIL) in order to measure the degradation of complex oil samples and obtain degradation rates (Schulte et al., 2019; Pannekens et al., 2021) which was not possible with classical methods such as analyzing oil compositions. For RSIL microcosms are set up with a bicarbonate buffer containing 10 atom percent of labeled $H^{13}CO_3^-$. Upon degradation of hydrocarbons, carbon dioxide with natural abundance of ^{13}C (1.01%) is released which dilutes the label in the bicarbonate buffer (Schulte et al., 2019; Pannekens et al., 2021). This change in the isotopic composition of the bicarbonate buffer can be very precisely measured with stable isotope analyzers for CO_2. Pannekens et al. (2021) successfully applied this method to assess the anaerobic mineralization of natural bitumen. However, the monitoring took almost three years, which in many cases is too long for studying environmental processes.

With the present study, we intended at verifying if RSIL can be used to measure degradation of light oil and if the sensitivity and accuracy are sufficient for assessing degradation rates during shorter incubations. Furthermore, we wanted to determine potential degradation rates of a light oil reservoir using indigenous microbial communities from the formation water or water droplets entrapped in the oil and we compare the resulting rates with previously determined oil degradation rates of bitumen (Meckenstock et al., 2014; Pannekens et al., 2020, 2021). To this end, RSIL was applied to validate the biological activity and degradation rates of light oil in combination with anoxic formation water. Oil and formation water samples were gathered from a newly drilled oil well located off shore of Trinidad (Trinidad and Tobago) to prevent contamination with allochthonous microorganisms and incubated in microcosms amended with 10 atom% ^{13}C-labeled bicarbonate buffer and sulfate as electron acceptor to follow the mineralization of oil to carbon dioxide.

Materials and methods

Oil sampling

Light oil and formation water from a newly drilled production well located in the Caribbean Sea close to Trinidad (Trinidad and Tobago) were acquired on 16.08.2018. Until sampling, the well remained free from water injection water, ensuring that the sample remained uncontaminated by production or sea water. Drilling fluids were not accessible. In total, the sample contained 16 liters light oil and four liters formation water. The formation water (402 mM Cl^-, 47.8 mM PO_4^{3-}, 38.2 µM Mg^{2+}, 48.8 µM Ca^{2+}) was decanted from the oil phase and transferred into glass bottles (2 L) under anoxic and sterile conditions. The head space was flushed with N_2/CO_2 (80/20) and sealed with butyl rubber stoppers (Ochs Laborbedarf, Bovenden, Germany).

The composition of the sampled oil was unknown. However, available data pertains to the eastern Venezuelan basin. An examination of formation water extracted from offshore oil fields, specifically Samaan and Poui, revealed the presence of short-chained aliphatic acids indicating a light oil reservoir (Fisher, 1987).

Geochemical formation water compositions depend on the oil reservoirs geology and environmental conditions. Dissolved inorganic carbon (DIC) and dissolved organic carbon (DOC) of the formation water were analyzed with a DOC/DIC-analyzer (TOC-L, Shimadzu, Duisburg) and revealed 505 mg DOC L^{-1} and

1373 mg DIC L^{-1}. The amount of DOC in the studied area is similar to those found in southern California (6 - 2930 mg L^{-1}) and three offshore reservoirs in the Eastern Venezuelan Basin (200 - 16514 mg L^{-1}) (Fisher, 1987; Mcmahon et al., 2018). However, the concentrations of DIC are higher in our study compared to the Eastern Venezuelan Basin with 4-765 mg L^{-1}.

For the reverse-isotope-labeling method it is important that the DIC in the formation water is lower than 10 mM. To remove the carbonate, 4 M phosphoric acid was added to the formation water until the pH reached 2.9 and the DIC evaporated as carbon dioxide, the acidic pH was maintained for 5 minutes. Then, the pH was directly adjusted to 7.5 with an anoxic NaOH solution (4 M) and the carbon dioxide-free formation water was used for microcosm setup. The phosphoric acid treatment also removed dissolved calcium by precipitation of $CaPO_4$ which was separated from the water by decanting. The acid treatment might have influenced the microbial community to a certain extent and may have led to a reduction in the observed initial activities. However, indigenous microbiota in subsurface oil reservoirs can be adapted to different pH conditions (Gao et al., 2016). Additionally, various mechanisms of acidic stress resistance (active proton pumps, cell internal buffer systems, changes in the membrane, chaperones) are known that might have protected the organisms from acidification resulting in the surviving community in our microcosms.

Microcosm setup

Microbial incubations were set up in serum bottles (120 mL) adding one drop (0.33 ± 0.09 g) of light oil to each bottle. Afterwards, 96 mL of anoxic, carbon dioxide-free formation water was added and the serum bottles were flushed with N_2/CO_2 (80/20) and sealed with butyl rubber stoppers (Ochs Laborbedarf, Bovenden, Germany). The water was buffered with ^{13}C-labeled bicarbonate (final concentration 10 mM with 10 atom % ^{13}C) made from a mixture of ^{13}C-labeled $NaHCO_3$ with an ^{13}C-atom percent $x(^{13}C)$ = 98% (Sigma-Aldrich, MO, USA) and regular $NaHCO_3$ $x(^{13}C)$ = 1.11% (Carl Roth, Germany). Finally, the terminal electron acceptor sulfate (20 mM) and Na_2S as reducing agent (0.5 mM) were added to remove any remaining traces of oxygen (Widdel and Pfennig, 1981). Immediate sulfate reduction proved that the bottles were all anoxic. Six replicate microcosms were prepared, whereof two were autoclaved to serve as negative heat-treated controls. The microcosms were incubated at 30°C in the dark in order to promote a wide range of microorganism without temperature restrictions. The overall duration of the experiment spanned 857 days, with monthly samplings conducted for the initial 122 days, followed by the final sampling point at day 857.

Assessing light oil mineralization

To assess degradation rates of light oil by carbon dioxide evolution, 1 mL of the aqueous phase of each microcosm was sampled with a N_2-flushed syringe through the stopper. The samples were transferred into 12 mL Labco Exertainer vials (Labco Limited, UK) that contained 50 µL of 85% phosphoric acid, closed with butyl rubber septa and screw caps, and flushed with carbon dioxide-free synthetic air (6.0 grade; Air Liquide, Germany) (Assayag et al., 2006; Dong et al., 2017, 2019; Schulte et al., 2019). The samples were analyzed using a Delta Ray CO_2 Isotope Ratio Infrared Spectrometer (Thermo Fisher Scientific, MA, USA) with Universal Reference Interface Connect according to Pannekens et al. (2021). To achieve the highest precision possible, the carbon dioxide concentration of the reference and sample gas entering the machine was set to 380 ppm.

After measuring each sample for 5 minutes, the resulting $^{13}C/^{12}C$ isotope ratio given in $\delta^{13}C$ values were averaged and transferred into isotope-amount fraction ($x(^{13}C)_{CO2}$) (Equation 1, Coplen, 2011).

$$isotope\ \ amount\ \ fraction\ \ (x(^{13}C)CO_2) = \frac{1}{1 + \frac{1}{(Delta\ \ Value\ \ Sample + 1) \times VPDB\ \ Standard\ \ Ratio}} \quad (1)$$

The isotope-amount fraction depicts the relative amount of ^{13}C in the gas phase originating from the DIC of the aqueous phase evaporated through the acid. The VPDB Standard ratio depicts the stable isotope ratio of the international Vienna PeeDee Belemnite carbon isotope standard given as $\delta^{13}C$ = 0.0111802.

The calculated isotope amount fraction was furthermore converted into atom percent by multiplying by 100 (atom percent = $x(^{13}C)_{CO2} \times 100$ [%]). This value represents the abundance of ^{13}C atoms within the sample in percent and is used to finally calculate the produced carbon dioxide according to Equation 2.

$$produced\ \ CO_2 = \frac{total\ \ CO_2\ \ in\ \ buffer[mM] \times (atom\ \ percent\ \ CO_2 - {}^{13}C\ \ content\ \ of\ \ buffer\ [\%])}{({}^{13}C\ \ content\ \ in\ \ substrate\ [\%] - atom\ \ percent\ \ CO_2)} \quad (2)$$

All calculations were performed using the actual total carbon dioxide concentration of 7.58 mM in the buffer and a ^{13}C content of 1.075% in the substrate oil (Pannekens et al., 2021).

The depletion of the terminal electron acceptor sulfate was monitored via ion chromatography (IC) as described in Voskuhl et al. (2021). The samples were pretreated by addition of KOH (0.1 M) and subsequent centrifugation to remove particles and heavy metals. Furthermore, a solid-phase-extraction using octadecyl-modified silica gel columns (Macherey-Nagel, Germany) was performed to remove organic compounds from the samples.

The growth of microorganisms in the aqueous phase was monitored via microscopic cell counting. Cells were stained with DAPI (4,6-diamidino-2-phenylindole) and counted via fluorescence microscopy (Zeiss AxioScope.A1, Oberkochen, Germany) according to Saby et al. (1997).

Microbial community composition

DNA extraction was performed with 1 mL of aqueous sample from each timepoint and microcosm using the DNeasy PowerLyser Power

Soil Kit (Qiagen GmbH, Germany) following the manufacturer's instructions. Slight variations were conducted according to Pannekens et al. (2021). 16S rRNA gene amplicon sequencing of the V3-V4 region was performed using the primer set S-D-Arch-0519-a-S-15/S-D-Bact-0785-b-A-18 (Klindworth et al., 2013). A touchdown PCR was carried out to avoid non-specific amplification (Korbie and Mattick, 2008). The PCR products were purified (MagSi NGSprep Plus magnetic beads) and cleaned following the Illumina guide ("16S Metagenomic Sequencing Library Preparation," 15044223). The last PCR step with Illumina indices and sequencing adapters was performed using the Nextera XT DNA Library Preparation Kit v2 Set D (FC-131-2004) by Illumina (München, Germany). After each DNA preparation step a 1% agarose-gel (w/v) electrophoresis was performed. The DNA concentration of each sample was normalized to 10 ng μL^{-1} and pooled into one ready-to-load library. Ready-to-load sequencing was outsourced to CeGaT GmbH (Tübingen, Germany) using a Illumnia MiSeq system with V2 500 cycle Flow Cell. Demultiplexing and adapter trimming was performed by the sequencing company. The resulting sequences were analyzed using the R-packages DADA2 version 1.24.0 and phyloseq version 1.40.0 (Mcmurdie and Holmes, 2013; Callahan et al., 2016). The forward and reverse reads were trimmed to 240 and 150 base pairs respectively before merging and chimera removal. Afterwards, the taxonomy was assigned using the SILVA database version 138.1 and the sequences were clustered into Amplicon Sequence Variants (ASVs) (Quast et al., 2013; Yilmaz et al., 2014; Callahan et al., 2016). In order to exclude rare organisms or sequencing errors, all amplicon sequence variants (ASVs) that either were only abundant in one of the two technical sequencing replicates or sequences with less than 10 reads per ASV were removed from the data set. Furthermore, all reads were rarified to the lowest detected read number of 14,212. Finally, the technical replicates of each timepoint were combined by calculating the mean of each ASV. Because many ASVs were received with the same taxonomy, we merged all of those with the "tax_glom" command always at the highest available taxonomic level, for visualization in (Figure 1). Raw sequences were deposited on the NCBI database (BioProject ID: PRJNA1030274).

Results

Mineralization of light oil

Microcosms with crude oil as carbon source and sulfate as electron acceptor were set up to determine biodegradation rates with reverse stable isotope labeling (RSIL) measuring the mineralization of oil to carbon dioxide. A major increase in carbon dioxide concentration was observed within the first 122 days of incubation with an average carbon dioxide production of 17.6 mM per year. The degradation rate in the four replicate microcosms was very reproducible within the first 122 days with 18.17 mM, 15.27 mM, 19.29 mM and 17.61 mM/year, respectively (Figure 2). The average production of carbon dioxide over the entire incubation period of 857 days was 8.31 ± 1.5 mM. Considering the volume of media used (100 mL), this resulted in the generation of 0.83 mmol of carbon dioxide after 857 days. Therefore, the overall degradation rate was 0.35 mmol CO_2 per year or 1.06 mmol CO_2 per gram of oil and year. This accounts to a rate of 15.2 mmol CO_2/mol CH_2(oil) per year or 120 g CH_2 m^{-2} OWC per year. However, the majority of degradation activity occurred within the first 122 days of incubation with a degradation rate of 0.63 mmol CO_2/mol CH_2(oil) per day. The CO_2 evolution demonstrated that 11.36 mg of oil were consumed throughout the entire incubation process. Given that we introduced 330 mg of oil, the electron and carbon source was not limiting.

The microorganisms in the system reduced the terminal electron acceptor sulfate, leading to a decrease from the initially added 20.6 mM to 15.2 mM over the 857-day incubation period indicating that the electron acceptor was also not limiting during the experiment. Considering the volume of media used, this corresponds to a total reduction of 0.54 mmol of sulfate and an average sulfate reduction rate of 0.23 mmol of sulfate per year.

However, the majority of sulfate degradation occurred within the first 122 days. Throughout the entire experiment, the pH values of the incubations remained stable at 7.5. The carbonate equilibrium at pH 7.5 implies that the majority of the carbon dioxide produced was completely dissolved as bicarbonate in the liquid phase, while a negligible amount was present in the gas phase.

An electron balance calculation was conducted based on the assumption that all the sulfate is reduced to hydrogen sulfide (5.4 mM). The crude oil used in our study is a complex blend of various hydrocarbons with an unknown oxidation state. Hence, we assumed a carbon oxidation state of -2 (CH_2). The electron balance revealed that 88% (7.3 mM) of the electrons from the generated carbon dioxide (8.3 mM) were used for the reduction of sulfate. The remaining 12% are mostly likely used for biomass production which is a common yield value for sulfate-reducing organisms.

Cell growth and community composition

Microbial growth in the aqueous phase was monitored during the exponential phase of carbon dioxide production. All microorganisms originated from the formation water or the oil itself. On average, the cell counts tripled from $4.0 \times 10^5 \pm 1.5 \times 10^5$ at the start of the experiment to $1.4 \times 10^6 \pm 1.9 \times 10^5$ cells mL^{-1} during the first 122 days (Figure 2). The cell counts in the autoclaved controls stayed mostly constant around $1.8 \times 10^5 \pm 1.8 \times 10^4$ cells mL^{-1} (Figure 2), probably because the cells were either dead or present as duration forms such as spores. High error bars at day 857 resulted from one autoclaved microcosm, where we could detect an increase in cell counts due to the proliferation of spore-forming microorganisms of the genus *Desulfotomaculum*. It remained unclear what triggered this germination and why it occurred in only one microcosm. The observed growth in the autoclaved controls can be attributed to several factors. Firstly, the autoclaving process was conducted in closed bottles to maintain anoxic conditions, necessary for the experiment. However, this method might have resulted in less efficient sterilization, as the microorganisms and water phase were not directly exposed to steam, leading to lower pressure compared to open bottles. Secondly, the closed system may have experienced reduced energy

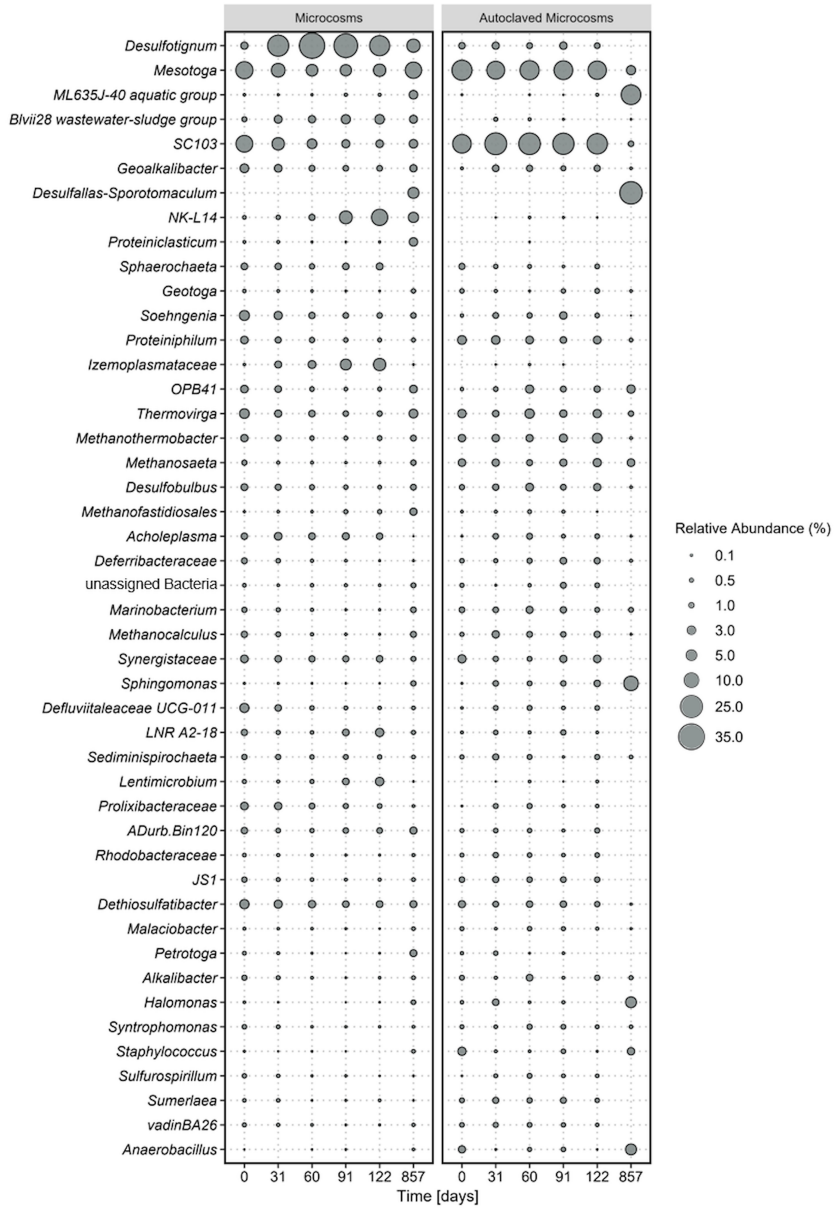


FIGURE 1
Average microbial community composition of most abundant taxa, calculated over four replicate microcosms (left panel) and two autoclaved controls (right panel) over 854 incubation days. Displayed taxa represent clustered ASVs on genus level. The size of each data point corresponds to the relative abundance in % at the corresponding time point.

transfer and heating, potentially resulting in temperatures below the optimal 121°C during autoclaving.

Shifts in the microbial communities where analyzed based on 16S rRNA gene amplicon sequences of the V3-V4 region that have been agglomerated based on the genus level. Out of a total of 1926 amplicon sequence variants (ASVs), the most abundant taxa were selected for further analyses. Bacterial ASVs were more abundant (> 87%) than archaea (< 13%).

The most dominant phyla were Desulfobacterota, (*Desulfotignum, Geoalkalibacter, Desulfobulbus, SEEP-SRB1, Syntrophotalea*) (Figure 1), followed by Thermotogota, Bacteroidota, Bacillota (former Firmicutes), and Synergistota. ASVs of *Clostridiales* were also detected in our original oil microbial communities although not displayed in Figure 1, as individual abundances were all below < 0.2% at the start of the experiment. The five most abundant phyla account for 9-44% of the total relative abundance in the microcosms. Within the phylum Desulfobacterota, members of the genus *Desulfotignum*, *Geoakalibacter* and *Desulfobulbus* dominated the community composition in the first 122 days reaching in total a maximum of 58% relative abundance after 60 days. After the first 60 days the relative abundance of the three genera constantly declined to a minimum of 20% after 587 days. In addition, a rapid rise in members of the phylum Cloacimonadota and especially the

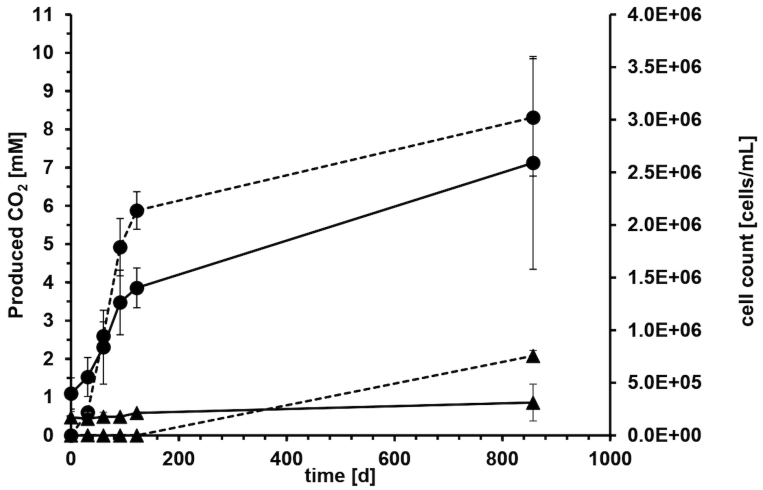


FIGURE 2
Average carbon dioxide evolution (dashed line) and cell counts (solid line) with crude oil as carbon and electron source and sulfate as electron acceptor of four microcosms over 857 days of incubation. Error bars depict standard deviations of four replicate incubations. Solid circles show life cultures (●) and solid triangles average values of heat-treated controls (▲) where the error bars indicate the spreading of duplicate incubations.

genera *NK-L14* (Speth et al., 2022) and *LNR A2-18* (Dutra et al., 2023) were observed during this time. The relative abundance started at 2% and stayed stable within the first 60 days of incubation. After that, it rapidly grew to a maximum relative abundance of 17% after 122 days. This increase of members of Cloacimonadota happened simultaneously to the decline of members of Desulfobacterota.

At the start of the experiment, Thermotogota (*Mesotoga, SC103, Geotoga, Petrotoga, Defluviitoga, Kosmotoga, Pseudothermotoga, Oceanotoga, Thermotoga, Thermosipho*) was the most abundant phylum with a relative abundance of over 31%. However, the relative abundance constantly decreased in the first 91 days to a minimum of 9% while stabilizing around 20% at the end of the incubation. Most prominent representatives of the phylum Thermotogota belonged to the genus *Mesotoga* (Nesbo et al., 2012, 2019) and *SC103* (Gao et al., 2023). The microbial community composition in the autoclaved controls remained consistent throughout the entire incubation period, which correlated with the constant cell counts. At the final sampling point, the increased abundance of two spore-forming organisms in the controls suggested that autoclaving the closed, anoxic bottles did not completely eradicate all spores. Instead, it appears that some spores managed to germinate after 122 days of incubation.

After 857 days (2.3 years) the spore-forming genus *Desulfotomaculum* had probably proliferated and dominated the community composition of this control bottle. The growth of the spore-forming bacteria was mainly observed in the autoclaved controls and to lesser extent in the active microcosms.

Discussion

Although oil reservoirs are extreme environments, microorganisms degrade the oil at the oil water contact (OWC) zone as well as in water phases within the oil leg (Head et al., 2003, 2010; Meckenstock et al., 2014; Pannekens et al., 2021). The objective of this study was to evaluate the potential of reverse stable isotope labeling (RSIL) to monitor small microbial oil degradation rates at the example of light oil from an oil reservoir located in Trinidad and Tobago. To this end, we collected reservoir fluids from a newly drilled production well including formation water and light oil without adding additional carbon sources. The oil well did not experience water injection ensuring that only the allochthonous microbiota was studied in the microcosm experiments. Using the RSIL, we were able to accurately and sensitively measure the rate of oil mineralization from a complex, naturally occurring light oil using the unaltered natural fluid composition.

In a previous publication, we conducted an extensive study on the degradation of heavy oil (bitumen) from the Pitch Lake in Trinidad and Tobago, which is a natural oil seep on the island of Trinidad (Pannekens et al., 2021). Since the incubation times in this experiment lasted over three years, we wanted to analyze here, if the RSIL method can be used to monitor the degradation of light oil in a reasonably shorter time frame. The bitumen degradation rate in the previous study was around 0.01 mmol produced carbon dioxide per gram of oil and year with a sulfate reduction rate of 0.46 mmol $year^{-1}$ (Pannekens et al., 2021). The degradation of light oil in the present study was 1.06 mmol carbon dioxide per gram of oil and year with 0.23 mmol sulfate reduction per year. Thus, the reduction rate of light oil was 100-fold faster compared to the bitumen experiment (Pannekens et al., 2021). A conceivable explanation is that the short-chain hydrocarbons in the light oil are utilized easier than the extensively degraded bitumen.

Available data on anaerobic microbial degradation rates of crude oil are scarce. Numerical modeling of oil biodegradation at the OWC-zone have resulted in oil degradation rates of 10^{-3}–10^{-4} kg petroleum m^{-2} OWC per year, dependent on reservoir temperature and oil composition. However, this rate was calculated from oil compositions based on gas chromatography measurements rather

than from microbiological incubation experiments (Head et al., 2003; Larter et al., 2006). In order to draw a comparison between our findings and the modeled results, we normalized our determined degradation rate of 15.2 mmol CO_2/mol CH_2 per year to the oil-water contact-area of the incubations (16.61 cm^2). The resulting degradation rate of 120 x 10^{-3} kg CH_2 m^{-2} OWC per year closely aligned with the previous modeled rates of 10^{-3}–10^{-4} kg petroleum m^{-2} OWC (Head et al., 2003; Larter et al., 2006). Nevertheless, one has to consider that the obtained rates in our study provide a degradation potential for the investigated sample under optimal incubation conditions and do not necessarily reflect real *in situ* degradation rates. Our microcosms provided very high concentrations of the favorable electron acceptor sulfate which is not present in pristine oil reservoirs and gets only introduced with injection water during secondary oil production. In oil reservoirs, mostly methanogenic degradation takes place which is orders of magnitude slower (Jones et al., 2008; Head et al., 2014). The microcosm incubations only contained an aqueous and an oil phase which diminished diffusion limitations that obviously occur in sediment or reservoir rock. Hence, the environmental conditions in the sediment or porous rock of the oil reservoir are distinct from those in the microcosm incubation. The obtained degradation rates primarily serve as an indicator of the presence and activity of microorganisms responsible for oil degradation in the original formation water and oil samples. Additionally, it is important that these degradation rates can be affected by various factors such as sampling of the environmental sample, storage conditions of the oil samples before incubation, and the experimental setup. For instance, the exposure to air before microcosm setup can rapidly lead to the demise of anaerobic microorganisms, potentially impacting the measured degradation rates. On the other hand, an exposure to oxygen will lead to an immediate growth of aerobic oil degraders, quickly changing the microbial community composition and the availability of easily degradable carbon sources. Nevertheless, our microcosm study can provide an estimate of the maximal degradation potential which provides a comparable parameter for biodegradation in different oil samples.

One way to qualitatively assess microbial degradation in oil reservoirs is by examining oil and gas samples for alterations in hydrocarbon composition, particularly the ratio of linear to branched alkanes. This analysis is primarily conducted to evaluate the weathering impact on crude oil and, to a lesser extent, to detect biodegradation, employing gas chromatography–mass spectrometry (GC–MS) techniques (Wang and Fingas, 1995; Munoz et al., 1997; Fernandez-Varela et al., 2009). However, this approach only offers qualitative insights into alterations in oil composition and does not provide information about degradation rates. Furthermore, it is conceivable that not all compounds from the intricate oil mixture are efficiently extracted during the sample preparation, which could result in their absence from the GC-MS spectra. This omission may potentially lead to overlooking degradation of non-identified compounds.

Since microbial degradation rates in oil reservoirs are extremely small, changes in oil composition or stable isotope ratios occur over geological time frames and cannot be monitored over time. Hence, such data have to be correlated to geochemical gradients or non-degraded reference samples to obtain indications of biodegradation but they can rarely provide degradation rates. However, gradients of compound compositions and nutrient supply capability have been used to model degradation rates in oil reservoirs (Larter et al., 2006; Kacewicz et al., 2010). There are two methods for assessing degradation rates through modeling. The first approach, known as the geochemical zero-order degradation model (referred to as "g0d"), entails the decomposition of hydrocarbons at a rate dependent on temperature. In this model, fresh oil containing known saturated hydrocarbon levels is introduced into a simulated geological trap at realistic rates (Larter et al., 2006). The second method, known as "Cyclops," involves a coupled oil charging and biodegradation model with advective-diffusive mixing within the reservoir. This approach is employed to estimate degradation fluxes by analyzing compositional gradients in biodegraded oil columns consisting of multiple components (Larter et al., 2006). However, despite their cost and time efficiency, these modeling approaches suffer from limited accuracy due to their reliance on assumptions and simplifications that fail to fully capture the complexity of the entire system. This drawback is particularly pronounced when dealing with oil reservoirs. Nevertheless, the degradation rates derived from modeling aligned with the degradation rates observed in this study through experimentation.

Compound Specific Isotope Analysis (CSIA) is a good method to assess microbial degradation of oil, by measuring the carbon stable isotope composition of hydrocarbons (Rooney et al., 1998; Schmidt et al., 2004; Meckenstock et al., 2004a). The fundamental concept underlying this approach is that stable isotope effects occurring during biodegradation induce alterations in the stable isotope composition of specific compounds, which CSIA can effectively measure. It is essential to recognize that CSIA provides only a snapshot of the current situation. To gauge the extent of biodegradation, one can compare the stable isotope compositions of selected compounds with samples from non-degraded locations within the reservoir. Unfortunately, due to a lack of information regarding the duration of this degradation process, CSIA encounters significant challenges when attempting to calculate degradation rates in oil reservoirs.

An established method to assess the biodegradation of organic compounds like crude oil is to measure the increase in carbon dioxide partial pressure in microcosm (Struijs and Stoltenkamp, 1990; Battersby, 1997). However, this method is usually of limited sensitivity when small amounts of substrate are oxidized as in the case of anaerobic crude oil degradation. Since the measurement of the total amount of CO_2 or the partial pressure is not very accurate, a small increase is hardly measurable. This problem can be overcome when microcosm incubations are amended with ^{13}C or ^{14}C-labeled substrates and the evolution of $^{13}CO_2$ or $^{14}CO_2$ is followed (Johnsen et al., 2013). Such measurements can be very accurate and sensitive but they are only applicable for specific labeled compounds that must be commercially available and cannot be used for complex mixtures or substrates of unknown composition such as crude oil.

To address this limitation, RSIL was developed as a robust and accurate method to assess the complete mineralization of all organic

compounds in the sample (Dong et al., 2017). The advantage lies in accurately tracking the degradation of any organic substrate, including complex mixtures such as crude oil, by measuring changes in the isotopic composition of the carbon dioxide in the microcosm over time. In this study, we demonstrate that RSIL can serve as an excellent method for evaluating biodegradation rates of light oil from production wells or other environmental samples with complex and unknown substrate compositions.

Microorganisms inhabiting oil reservoirs have been identified as mesophilic and thermophilic sulfate-reducing, fermentative, manganese- and iron-reducing, acetogenic bacteria and methanogenic archaea (Magot et al., 2000; Li et al., 2017b; Pannekens et al., 2019; Zhou et al., 2019a, 2023). Comparing three oil seeps located thousands of kilometers apart revealed similarities in the most prominent representatives, which belonged to the bacterial phyla Proteobacteria, Bacteroidetes, Bacillota (former Firmicutes), Synergistetes, Deferribacteres, Thermotogae, Chloroflexi, and Bacteroidia (Pannekens et al., 2020). To this end, numerous systematic studies exploring the composition and distribution of microbial communities across water-flooded oil reservoirs have been documented. The dominant bacterial taxa encompass Alpha-, Beta-, Delta-, Gammaproteobacteria, Actinobacteria, Thermotogae and Sphingomonas (Feng et al., 2011; Wang et al., 2012; Gao et al., 2016; Li et al., 2017a; Kim et al., 2018).

It is obvious that the conditions within our microcosms differed from those in the oil reservoirs, as the microbial community compositions changed rapidly upon incubation (Figure 1). Examination of the microbial communities over time revealed the dominance the phyla Desulfobacterota, Cloacimonadota, Bacteroidota, Pseudomonadota, and Thermotogota. Interestingly, the abundance of these phyla did not remain constant over time. The primary energy conservation pathways during the first 122 days were most likely fermentation and sulfate reduction. Initially, the experiment started with a community that was dominated by the obligate anaerobic fermenters of the genus Thermotoga (Lanzilli et al., 2020; Shao et al., 2020), indicating that fermentation was the dominating metabolism in the original crude oil. Genomic analysis of Thermotoga indicated that they likely colonized oil reservoirs from the subsurface, rather than being originally deposited during oil reservoir formation (Nesbo et al., 2015, 2019). Following the incubation of oil, heavy oil, or water from oil reservoirs, fermenters frequently emerge as the most prevalent microorganisms (Mbadinga et al., 2011; Zhou et al., 2019b; Pannekens et al., 2021). In most studies on oil reservoir microbiology, sulfate reduction was observed to take a subordinate role due to rapid depletion of sulfate (Jones et al., 2008) or potential inhibition of the sulfate-reducing bacteria (Voskuhl et al., 2022). Typically when sulfate concentrations exceed 50 μM, sulfate reduction becomes the prevailing degradation process (Jimenez et al., 2016). The primary introduction of sulfate into oil reservoirs occurs through water injection. The allochthonus microbial community in our experiments was not exposed to injection water before because it was obtained from a newly drilled well. Nevertheless, the community was able to quickly utilize the added sulfate in our experiments which indicates that an oil reservoir community can rapidly produce reservoir souring upon sulfate injection.

Over the incubation time, Desulfobacterota took over and became dominant in the community within the first 60 days of incubation, probably due to reduction of the added sulfate. Then, the abundance of Desulfobacterota declined steadily over time, while the abundance of Cloacimonadota increased, suggesting a potential competition between these two phyla. Desulfobacterota, Thermotogota, and Cloacimonadota have been identified as core phyla specifically in low-temperature oil reservoirs (Hidalgo et al., 2021; Gittins et al., 2023; Zhou et al., 2023). Core phyla are phyla that are prevalent and most prominent within various oil reservoirs around the world. We could identify these phyla now in yet another oil reservoir confirming their status as core phyla.

The genus *Desulfotignum* (Desulfobacterota) can dominate sulfate-reducing communities from oil reservoirs containing high sulfate concentrations. This well-known member of oil reservoir consortia was very abundant in oil fields located in the Neuquen Basin (Argentina), Akita (Japan) and Jinangsu (China) (Grigoryan et al., 2008; Li et al., 2016; Li et al., 2017a; Kamarisima et al., 2018). Its prevalence in these environments can be attributed to its ability to utilize aliphatic organic acids, which are commonly present in petroleum reservoirs (Hatton and Hanor, 1984; Fisher, 1987, Means and Hubbard, 1987). In all anaerobic degradation pathways of aromatic hydrocarbons aliphatic acids are built as metabolites which can also be taken as indicators of microbial degradation (Gieg and Toth, 2020). The metabolites have to be unique and indicative of the degraded parent compounds. This is hardly possible for the majority of the alkanes in oil since the potential product fatty acids are universal to life. Only when specific modifications are performed, e. g. the activation with formate via glycyl radical enzymes, specific signature metabolites are produced that indicate certain metabolic pathways (Mbadinga et al., 2011). In the case of polycyclic aromatic hydrocarbons (PAHs) degradation, very specific signature metabolites are produced (e. g. different naphthoic acids) that can be attributed to the anaerobic degradation of naphthalene or other PAHs (Meckenstock et al., 2004b).

One interesting aspect of *Desulfotignum* is its ability to also degrade aromatics such as toluene through the fumarate-addition pathway (Hasegawa et al., 2014). Furthermore, its capacity to generate propionate from succinate and lactate further enhances its metabolic versatility and adaptability to varying reservoir conditions (Yang et al., 2017). The prevalence of *Desulfotignum* in moderate temperate oil reservoirs suggests a specific adaptation to such regimes (Tian et al., 2020).

Members of the Cloacimonadota have been identified in anoxic environments worldwide. Aside from engineered systems like anaerobic digesters (Dyksma and Gallert, 2019), Cloacimonadota have also been detected in natural environments such as sulfidic water layers of the Black Sea, Ursu Lake in Romania, and the Thuwal cold seep brine pool in the Red Sea (Williams et al., 2021). Based on incubations and analysis of Black Sea samples, Cloacimonadota have been inferred as fermentative heterotrophic generalists capable of utilizing various carbon sources, including proteins. As demonstrated here, they are also present in oil reservoirs as well as the biggest natural bitumen lake, the Pitch Lake in Trinidad-Tobago (Voskuhl et al., 2021). Furthermore, they were found in the high temperature Qinghai Oil Field, which was

subject to water-flooding for more than 15 years (Hidalgo et al., 2021) or oil pipelines in Brazil (Dutra et al., 2023). Overall, the diversity of carbon utilization enzymes and metabolic pathways within the Cloacimonadota phylum suggests an anaerobic, acetogenic, and mixed fermentative lifestyle, potentially relying on flavin-based bifurcation of hydrogen during chemolithotrophic growth (Johnson and Hug, 2022).

Mesotoga are commonly found in mesothermic anaerobic environments rich in hydrocarbons. They are characterized by low or no hydrogen production as well as sulfur-compound reduction (Nesbo et al., 2019). Mesotoga are strictly anaerobic and their growth is slightly stimulated by thiosulfate, sulfite, and elemental sulfur (Nesbo et al., 2012). Sulfate reducing bacteria can potentially stimulate the growth of Mesotoga by sulfite production. Human activities, such as wastewater treatment and oil exploration, have expanded the habitats and dispersal of Mesotoga. It has been isolated from and detected in polluted marine sediments, low-temperature oil reservoirs, wastewater treatment plants, and anaerobic methanogenic environments, where it may participate in syntrophic acetate degradation (Nesbo et al., 2019).

Anthropogenic activities can significantly effect microbial communities in oil fields during production (Vigneron et al., 2017). In the initial stages of oil production, microbial communities were predominantly characterized by slow-growing anaerobes such as members of the *Thermotogales* and *Clostridiales*, presumably thriving on hydrocarbons and complex organic matter. During secondary oil production, the injection of seawater and nitrate induced significant alterations in microbial community composition with a shift towards fast-growing opportunistic species, encompassing members of the *Deferribacteres*, or Delta-, Epsilon-, and Gammaproteobacteria (Vigneron et al., 2017). These microorganisms exhibited a genomic profile indicating processes such as nitrate reduction and hydrogen sulfide (H_2S) or sulfur oxidation. However, a change of the microbial community composition is not necessarily caused by injected microorganisms since the microbial communities in the oil reservoir do not necessarily contain allochthonous microorganisms (Mbow et al., 2024). Besides, findings from Gao et al. in 2016 elucidated specific microbial taxa consistently present in oil fields. Among others, members of *Geoakalibakter, Marinobacter, Sphingomonas*, and Thermotoga were identified in five out of nine analyzed oil fields in China (Gao et al., 2016). In our study, we detected the presence of several *Desulfobacterota* ASVs which was also found by Vigneron et al. in oil reservoirs before water injection. Hence, our study confirms the prevalence and persistence of these microbial groups across diverse oil field ecosystems.

Concluding, our results demonstrate that the reverse stable isotope labeling (RSIL) method provides a valuable method to assess the degradation of light oil by indigenous microorganisms. The found microorganisms and observed changes in microbial community composition over time show that fermentation could be a main metabolic trait in the investigated oil reservoir.

Data availability statement

The datasets presented in this study can be found in online repositories. The sequence datasets generated for this study were deposited at NCBI under BioProject ID: PRJNA1030274.

Author contributions

SB: Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. LV: Data curation, Investigation, Methodology, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. MP: Conceptualization, Formal analysis, Methodology, Writing – review & editing. RM: Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Validation, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This study was supported by the European Research Council (ERC; grant number 666952) and the German Research Foundation DFG through grant ME2049/11-1.

Acknowledgments

The authors thank Dr. Daniel Köster, University of Duisburg-Essen for performing the DOC and DIC measurements and Dr. Verena Brauer, University of Duisburg-Essen for assisting with the bioinformatic analysis.

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

Aitken, C. M., Jones, D. M., and Larter, S. R. (2004). Anaerobic hydrocarbon biodegradation in deep subsurface oil reservoirs. *Nature* 431, 291–294. doi: 10.1038/nature02922

Assayag, N., Rive, K., Ader, M., Jezequel, D., and Agrinier, P. (2006). Improved method for isotopic and quantitative analysis of dissolved inorganic carbon in natural water samples. *Rapid Commun. Mass Spectrom.* 20, 2243–2251. doi: 10.1002/rcm.2585

Battersby, N. S. (1997). The ISO headspace CO_2 biodegradation test. *Chemosphere* 34, 1813–1822. doi: 10.1016/S0045-6535(96)00358-X

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

Cho, Y., Ahmed, A., Islam, A., and Kim, S. (2015). Developments in FT-ICR MS instrumentation, ionization techniques, and data interpretation methods for petroleomics. *Mass Spectrom Rev.* 34, 248–263. doi: 10.1002/mas.21438

Coplen, T. B. (2011). Guidelines and recommended terms for expression of stable-isotope-ratio and gas-ratio measurement results. *Rapid Commun. Mass Spectrom.* 25, 2538–2560. doi: 10.1002/rcm.5129

Dong, X., Backer, L. E., Rahmatullah, M., Schunk, D., Lens, G., and Meckenstock, R. U. (2019). Quantification of microbial degradation activities in biological activated carbon filters by reverse stable isotope labelling. *AMB Express* 9, 109. doi: 10.1186/s13568-019-0827-0

Dong, X., Jochmann, M. A., Elsner, M., Meyer, A. H., Backer, L. E., Rahmatullah, M., et al. (2017). Monitoring microbial mineralization using reverse stable isotope labeling analysis by mid-infrared laser spectroscopy. *Environ. Sci. Technol.* 51, 11876–11883. doi: 10.1021/acs.est.7b02909

Dutra, J., Garcia, G., Gomes, R., Cardoso, M., Cortes, A., Silva, T., et al. (2023). Effective biocorrosive control in oil industry facilities: 16S rRNA gene metabarcoding for monitoring microbial communities in produced water. *Microorganisms* 11, 846. doi: 10.3390/microorganisms11040846

Dyksma, S., and Gallert, C. (2019). *Candidatus* Syntrophosphaera thermopropionivorans: a novel player in syntrophic propionate oxidation during anaerobic digestion. *Environ. Microbio. Rep.* 11, 558–570. doi: 10.1111/1758-2229.12759

Feng, W.-W., Liu, J.-F., Gu, J.-D., and Mu, B.-Z. (2011). Nitrate-reducing community in production water of three oil reservoirs and their responses to different carbon sources revealed by nitrate-reductase encoding gene (napA). *Int. Biodeterioration Biodegradation* 65, 1081–1086. doi: 10.1016/j.ibiod.2011.05.009

Fernandez-Varela, R., Andrade, J. M., Muniategui, S., Prada, D., and Ramirez-Villalobos, F. (2009). The comparison of two heavy fuel oils in composition and weathering pattern, based on IR, GC-FID and GC-MS analyses: application to the Prestige wreackage. *Water Res.* 43, 1015–1026. doi: 10.1016/j.watres.2008.11.047

Fisher, J. B. (1987). Distribution and occurrence of aliphatic acid anions in deep subsurface waters. *Geochimica Cosmochimica Acta* 51, 2459–2468. doi: 10.1016/0016-7037(87)90297-3

Gao, P., Tian, H., Wang, Y., Li, Y., Li, Y., Xie, J., et al. (2016). Spatial isolation and environmental factors drive distinct bacterial and archaeal communities in different types of petroleum reservoirs in China. *Sci. Rep.* 6, 20174. doi: 10.1038/srep20174

Gao, Y., Cai, T., Yin, J., Li, H., Liu, X., Lu, X., et al. (2023). Insights into biodegradation behaviors of methanolic wastewater in up-flow anaerobic sludge bed (UASB) reactor coupled with *in-situ* bioelectrocatalysis. *Bioresour. Technol.* 376, 128835. doi: 10.1016/j.biortech.2023.128835

Gieg, L. M., Davidova, I. A., Duncan, K. E., and Suflita, J. M. (2010). Methanogenesis, sulfate reduction and crude oil biodegradation in hot Alaskan oilfields. *Environ. Microbio* 12, 3074–3086. doi: 10.1111/j.1462-2920.2010.02282.x

Gieg, L. M., and Toth, C. R. A. (2020). Signature metabolite analysis to determine in *situ* anaerobic hydrocarbon biodegradation. In: M. Boll *Anaerobic Utilization of Hydrocarbons, Oils, and Lipids. Handbook of Hydrocarbon and Lipid Microbiology* (Cham: Springer). doi: 10.1007/978-3-319-50391-2_19

Gittins, D. A., Bhatnagar, S., and Hubert, C. R. J. (2023). Environmental selection and biogeography shape the microbiome of subsurface petroleum reservoirs. *mSystems* 8, e0088422. doi: 10.1128/msystems.00884-22

Golby, S., Ceri, H., Gieg, L. M., Chatterjee, I., Marques, L. L., and Turner, R. J. (2012). Evaluation of microbial biofilm communities from an Alberta oil sands tailings pond. *FEMS Microbiol. Ecol.* 79, 240–250. doi: 10.1111/j.1574-6941.2011.01212.x

Grigoryan, A. A., Cornish, S. L., Buziak, B., Lin, S., Cavallaro, A., Arensdorf, J. J., et al. (2008). Competitive oxidation of volatile fatty acids by sulfate- and nitrate-reducing bacteria from an oil field in Argentina. *Appl. Environ. Microbio* 74, 4324–4335. doi: 10.1128/AEM.00419-08

Hasegawa, R., Toyama, K., Miyanaga, K., and Tanji, Y. (2014). Identification of crude-oil components and microorganisms that cause souring under anaerobic conditions. *Appl. Microbio Biotechno* 98, 1853–1861. doi: 10.1007/s00253-013-5107-3

Hatton, R., and Hanor, J. (1984). Dissolved volatile fatty acids in subsurface, hydropressure brines: a review of published literature on occurrence, genesis and thermodynamic properties. *Technical Report for Geopressured-Geothermal Activities in Louisiana: Final Geological Report for the Period 1 November 1981 to 31 October 1982. DOE Report No. DOE/NV/10174-3Technical Report for Geopressured-Geothermal Activities in Louisiana: Final Geological Report for the Period 1 November 1981 to 31 October 1982. DOE Report No. DOE/NV/10174-3.*

Head, I. M., Gray, N. D., and Larter, S. R. (2014). Life in the slow lane; biogeochemistry of biodegraded petroleum containing reservoirs and implications for energy recovery and carbon management. *Front. Microbio.* 5, 566. doi: 10.3389/fmicb.2014.00566

Head, I. M., Jones, D. M., and Larter, S. R. (2003). Biological activity in the deep subsurface and the origin of heavy oil. *Nature* 426, 344–352. doi: 10.1038/nature02134

Head, I., Larter, S., Gray, N., Sherry, A., Adams, J., Aitken, C., et al. (2010). "Hydrocarbon degradation in petroleum reservoirs," In: Timmis, K.N. (eds). *Handbook of hydrocarbon and lipid microbiology*. (Berlin, Heidelberg: Springer). doi: 10.1007/978-3-540-77587-4_232

Hidalgo, K. J., Sierra-Garcia, I. N., Zafra, G., and De Oliveira, V. M. (2021). Genome-resolved meta-analysis of the microbiome in oil reservoirs worldwide. *Microorganisms* 9, 1812. doi: 10.3390/microorganisms9091812

Jimenez, N., Richnow, H. H., Vogt, C., Treude, T., and Kruger, M. (2016). Methanogenic hydrocarbon degradation: evidence from field and laboratory studies. *J. Mol. Microbiol. Biotechnol.* 26, 227–242. doi: 10.1159/000441679

Johnsen, A. R., Binning, P. J., Aamand, J., Badawi, N., and Rosenbom, A. E. (2013). The gompertz function can coherently describe microbial mineralization of growth-sustaining pesticides. *Environ. Sci. Techno*. 47, 8508–8514. doi: 10.1021/es400861v

Johnson, L. A., and Hug, L. A. (2022). Cloacimonadota metabolisms include adaptations in engineered environments that are reflected in the evolutionary history of the phylum. *Environ. Microbio Rep.* 14, 520–529. doi: 10.1111/1758-2229.13061

Jones, D. M., Head, I. M., Gray, N. D., Adams, J. J., Rowan, A. K., Aitken, C. M., et al. (2008). Crude-oil biodegradation via methanogenesis in subsurface petroleum reservoirs. *Nature* 451, 176–180. doi: 10.1038/nature06484

Kacewicz, M., Peters, K. E., and Curry, D. J. (2010). 2009 napa AAPG hedberg research conference on basin and petroleum systems modeling. *AAPG Bull.* 94, 773–789. doi: 10.1306/10270909128

Kamarisima,, Hidaka, K., Miyanaga, K., and Tanji, Y. (2018). The presence of nitrate- and sulfate-reducing bacteria contributes to ineffectiveness souring control by nitrate injection. *Int. Biodeterioration Biodegradation* 129, 81–88. doi: 10.1016/j.ibiod.2018.01.007

Kayukova, G. P., Uspensky, B. V., Abdrafikova, I. M., and Musin, R. Z. (2016). Characteristic features of the hydrocarbon composition of Spiridonovskoe (Tatarstan) and Pitch Lake (Trinidad and Tobago) asphaltites. *Petroleum Chem.* 56, 572–579. doi: 10.1134/S0965544116070082

Kim, D. D., O'farrell, C., Toth, C. R. A., Montoya, O., Gieg, L. M., Kwon, T. H., et al. (2018). Microbial community analyses of produced waters from high-temperature oil reservoirs reveal unexpected similarity between geographically distant oil reservoirs. *Microb. Biotechnol.* 11, 788–796. doi: 10.1111/1751-7915.13281

Klindworth, A., Pruesse, E., Schweer, T., Peplies, J., Quast, C., Horn, M., et al. (2013). Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res.* 41, e1. doi: 10.1093/nar/gks808

Kondyli, A., and Schrader, W. (2019). High-resolution GC/MS studies of a light crude oil fraction. *J. Mass Spectrom* 54, 47–54. doi: 10.1002/jms.4306

Korbie, D. J., and Mattick, J. S. (2008). Touchdown PCR for increased specificity and sensitivity in PCR amplification. *Nat. Protoc.* 3, 1452–1456. doi: 10.1038/nprot.2008.133

Lanzilli, M., Esercizio, N., Vastano, M., Xu, Z., Nuzzo, G., Gallo, C., et al. (2020). Effect of cultivation parameters on fermentation and hydrogen production in the phylum *thermotogae*. *Int. J. Mol. Sci.* 22, 341. doi: 10.3390/ijms22010341

Larter, S., Huang, H., Adams, J., Bennett, B., Jokanola, O., Oldenburg, T., et al. (2006). The controls on the composition of biodegraded oils in the deep subsurface: Part II—Geological controls on subsurface biodegradation fluxes and constraints on reservoir-fluid property prediction. *AAPG Bull.* 90, 921–938. doi: 10.1306/01270605130

Larter, S., Wilhelins, A., Head, I., Koopmans, M., Aplin, A., Di Primio, R., et al. (2003). The controls on the composition of biodegraded oils in the deep subsurface - part 1: biodegradation rates in petroleum reservoirs. *Org. Geochem.* 34, 601–613. doi: 10.1016/S0146-6380(02)00240-1

Li, X.-X., Liu, J.-F., Yao, F., Wu, W.-L., Yang, S.-Z., Mbadinga, S. M., et al. (2016). Dominance of *Desulfotignum* in sulfate-reducing community in high sulfate production-water of high temperature and corrosive petroleum reservoirs. *Int. Biodeterioration Biodegradation* 114, 45–56. doi: 10.1016/j.ibiod.2016.05.018

Li, X. X., Liu, J. F., Zhou, L., Mbadinga, S. M., Yang, S. Z., Gu, J. D., et al. (2017a). Diversity and composition of sulfate-reducing microbial communities based on genomic DNA and RNA transcription in production water of high temperature and corrosive oil reservoir. *Front. Microbio.* 8, 1011. doi: 10.3389/fmicb.2017.01011

Li, X. X., Mbadinga, S. M., Liu, J. F., Zhou, L., Yang, S. Z., Gu, J. D., et al. (2017b). Microbiota and their affiliation with physiochemical characteristics of different subsurface petroleum reservoirs. *Int. Biodeterioration Biodegradation* 120, 170–185. doi: 10.1016/j.ibiod.2017.02.005

Magot, M., Ollivier, B., and Patel, B. K. (2000). Microbiology of petroleum reservoirs. *Antonie Van Leeuwenhoek* 77, 103–116. doi: 10.1023/A:1002434330514

Mbadinga, S. M., Wang, L.-Y., Zhou, L., Liu, J.-F., Gu, J.-D., and Mu, B.-Z. (2011). Microbial communities involved in anaerobic degradation of alkanes. *Int. Biodeterioration Biodegradation* 65, 1–13. doi: 10.1016/j.ibiod.2010.11.009

Mbow, F. T., Akbari, A., Dopffel, N., Schneider, K., Mukherjee, S., and Meckenstock, R. U. (2024). Insights into the effects of anthropogenic activities on oil reservoir microbiome and metabolic potential. *N. Biotechnol.* 79, 30–38. doi: 10.1016/j.nbt.2023.11.004

Mcmahon, P. B., Kulongoski, J. T., Vengosh, A., Cozzarelli, I. M., Landon, M. K., Kharaka, Y. K., et al. (2018). Regional patterns in the geochemistry of oil-field water, southern San Joaquin Valley, California, USA. *Appl. Geochem.* 98, 127–140. doi: 10.1016/j.apgeochem.2018.09.015

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

Means, J. L., and Hubbard, N. (1987). Short-chain aliphatic acid anions in deep subsurface brines: a review of their origin, occurrence, properties, and importance and new data on their distribution and geochemical implications in the Palo Duro Basin, Texas. *Org. Geochem.* 11, 177-191.

Meckenstock, R. U., Morasch, B., Griebler, C., and Richnow, H. H. (2004a). Stable isotope fractionation analysis as a tool to monitor biodegradation in contaminated acquifers. *J. Contam. Hydrol.* 75, 215–255. doi: 10.1016/j.jconhyd.2004.06.003

Meckenstock, R. U., Safinowski, M., and Griebler, C. (2004b). Anaerobic degradation of polycyclic aromatic hydrocarbons. *FEMS Microbiol. Ecol.* 49, 27–36. doi: 10.1016/j.femsec.2004.02.019

Meckenstock, R. U., Von Netzer, F., Stumpp, C., Lueders, T., Himmelberg, A. M., Hertkorn, N., et al. (2014). Water droplets in oil are microhabitats for microbial life. *Science* 345, 673–676. doi: 10.1126/science.1252215

Munoz, D., Doumenq, P., Guiliano, M., Jacquot, F., Scherrer, P., and Mille, G. (1997). New approach to study of spilled crude oils using high resolution GC-MS (SIM) and metastable reaction monitoring GC-MS-MS. *Talanta* 45, 1–12. doi: 10.1016/S0039-9140(96)02054-1

Nesbo, C. L., Bradnan, D. M., Adebusuyi, A., Dlutek, M., Petrus, A. K., Foght, J., et al. (2012). *Mesotoga prima* gen. nov., sp. nov., the first described mesophilic species of the Thermotogales. *Extremophiles* 16, 387–393. doi: 10.1007/s00792-012-0437-0

Nesbo, C. L., Charchuk, R., Pollo, S. M. J., Budwill, K., Kublanov, I. V., Haverkamp, T. H. A., et al. (2019). Genomic analysis of the mesophilic Thermotogae genus *Mesotoga* reveals phylogeographic structure and genomic determinants of its distinct metabolism. *Environ. Microbiol.* 21, 456–470. doi: 10.1111/1462-2920.14477

Nesbo, C. L., Dahle, H., Haverkamp, T. H., Birkeland, N. K., Sokolova, T., Kublanov, I., et al. (2015). Evidence for extensive gene flow and Thermotoga subpopulations in subsurface and marine environments. *ISME J.* 9, 1532–1542. doi: 10.1038/ismej.2014.238

Pannekens, M., Kroll, L., Muller, H., Mbow, F. T., and Meckenstock, R. U. (2019). Oil reservoirs, an exceptional habitat for microorganisms. *N. Biotechnol.* 49, 1–9. doi: 10.1016/j.nbt.2018.11.006

Pannekens, M., Voskuhl, L., Meier, A., Muller, H., Haque, S., Frosler, J., et al. (2020). Densely populated water droplets in heavy-oil seeps. *Appl. Environ. Microbio.* 86, e00164-20. doi: 10.1128/AEM.00164-20

Pannekens, M., Voskuhl, L., Mohammadian, S., Koster, D., Meier, A., Kohne, J. M., et al. (2021). Microbial degradation rates of natural bitumen. *Environ. Sci. Technol.* 55, 8700–8708. doi: 10.1021/acs.est.1c00596

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

Rajbongshi, A., and Gogoi, S. B. (2021). A review on anaerobic microorganisms isolated from oil reservoirs. *World J. Microbiol. Biotechnol.* 37, 111. doi: 10.1007/s11274-021-03080-9

Rooney, M. A., Vuletich, A. K., and Griffith, C. E. (1998). Compound-specific isotope analysis as a tool for characterizing mixed oils: an example from the West of Shetlands area. *Org. Geochem.* 29, 241–254. doi: 10.1016/S0146-6380(98)00136-3

Saby, S., Sibille, I., Mathieu, L., Paquin, J. L., and Block, J. C. (1997). Influence of water chlorination on the counting of bacteria with DAPI (4',6-diamidino-2-phenylindole). *Appl. Environ. Microbiol.* 63, 1564–1569. doi: 10.1128/aem.63.4.1564-1569.1997

Schmidt, T. C., Zwank, L., Elsner, M., Berg, M., Meckenstock, R. U., and Haderlein, S. B. (2004). Compound-specific stable isotope analysis of organic contaminants in natural environments: a critical review of the state of the art, prospects, and future challenges. *Anal. Bioanal Chem.* 378, 283–300. doi: 10.1007/s00216-003-2350-y

Schulte, S. M., Koster, D., Jochmann, M. A., and Meckenstock, R. U. (2019). Applying reverse stable isotope labeling analysis by mid-infrared laser spectroscopy to monitor BDOC in recycled wastewater. *Sci. Total Environ.* 665, 1064–1072. doi: 10.1016/j.scitotenv.2019.02.061

Shao, W., Wang, Q., Rupani, P. F., Krishnan, S., Ahmad, F., Rezania, S., et al. (2020). Biohydrogen production via thermophilic fermentation: A prospective application of Thermotoga species. *Energy* 197, 117199. doi: 10.1016/j.energy.2020.117199

Speth, D. R., Yu, F. B., Connon, S. A., Lim, S., Magyar, J. S., Peña, M. E., et al. (2022). Microbial community of recently discovered Auka vent field sheds light on vent biogeography and evolutionary history of thermophily. *The ISME Journal* 16 (7), 1750–1764. doi: 10.1101/2021.08.02.454472

Struijs, J., and Stoltenkamp, J. (1990). Headspace determination of evolved carbon dioxide in a biodegradability screening test. *Ecotoxicol Environ. Saf.* 19, 204–211. doi: 10.1016/0147-6513(90)90068-G

Tian, Y., Xue, S., and Ma, Y. (2020). Comparative analysis of bacterial community and functional species in oil reservoirs with different in *situ* temperatures. *Int. Microbio.* 23, 557–563. doi: 10.1007/s10123-020-00125-1

Vigneron, A., Alsop, E. B., Lomans, B. P., Kyrpides, N. C., Head, I. M., and Tsesmetzis, N. (2017). Succession in the petroleum reservoir microbiome through an oil field production lifecycle. *ISME J.* 11, 2141–2154. doi: 10.1038/ismej.2017.78

Voskuhl, L., Akbari, A., Muller, H., Pannekens, M., Brusilova, D., Dyksma, S., et al. (2021). Indigenous microbial communities in heavy oil show a threshold response to salinity. *FEMS Microbio. Eco.* 97 (12), fiab157. doi: 10.1093/femsec/fiab157

Voskuhl, L., Brusilova, D., Brauer, V. S., and Meckenstock, R. U. (2022). Inhibition of sulfate-reducing bacteria with formate. *FEMS Microbiol. Ecol.* 98 (1), fiac003. doi: 10.1093/femsec/fiac003

Wang, L. Y., Duan, R. Y., Liu, J. F., Yang, S. Z., Gu, J. D., and Mu, B. Z. (2012). Molecular analysis of the microbial community structures in water-flooding petroleum reservoirs with different temperatures. *Biogeosciences* 9, 4645–4659. doi: 10.5194/bg-9-4645-2012

Wang, Z., and Fingas, M. (1995). Differentiation of the source of spilled oil and monitoring of the oil weathering process using gas chromatography-mass spectrometry. *J. Chromatogr. A* 712, 321–343. doi: 10.1016/0021-9673(95)00546-Y

Widdel, F., and Pfennig, N. (1981). Studies on dissimilatory sulfate-reducing bacteria that decompose fatty acids. I. Isolation of new sulfate-reducing bacteria enriched with acetate from saline environments. Description of *Desulfobacter postgatei* gen. nov., sp. nov. *Arch. Microbio* 129, 395–400. doi: 10.1007/BF00406470

Williams, T. J., Allen, M. A., Berengut, J. F., and Cavicchioli, R. (2021). Shedding light on microbial "Dark matter": insights into novel cloacimonadota and omnitrophota from an antarctic lake. *Front. Microbio.* 12, 741077. doi: 10.3389/fmicb.2021.741077

Yang, T., Mbadinga, S. M., Zhou, L., Yang, S. Z., Liu, J. F., Gu, J. D., et al. (2017). Propionate metabolism and diversity of relevant functional genes by in silico analysis and detection in subsurface petroleum reservoirs. *World J. Microbio. Biotechno.* 33, 182. doi: 10.1007/s11274-017-2350-2

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, D643–D648. doi: 10.1093/nar/gkt1209

Youssef, N., Elshahed, M. S., and Mcinerney, M. J. (2009). Microbial processes in oil fields: culprits, problems, and opportunities. *Adv. Appl. Microbio.* 66, 141–251. doi: 10.1016/S0065-2164(08)00806-X

Zhou, Z. C., Liang, B., Wang, L. Y., Liu, J. F., Mu, B. Z., Shim, H., et al. (2019b). Identifying the core bacterial microbiome of hydrocarbon degradation and a shift of dominant methanogenesis pathways in the oil and aqueous phases of petroleum reservoirs of different temperatures from China. *Biogeosciences* 16, 4229–4241. doi: 10.5194/bg-16-4229-2019

Zhou, L., Wu, J., Ji, J. H., Gao, J., Liu, Y. F., Wang, B., et al. (2023). Characteristics of microbiota, core sulfate-reducing taxa and corrosion rates in production water from five petroleum reservoirs in China. *Sci. Total Environ.* 858, 159861. doi: 10.1016/j.scitotenv.2022.159861

Zhou, L., Zhou, Z., Lu, Y.-W., Ma, L., Bai, Y., Li, X.-X., et al. (2019a). The newly proposed TACK and DPANN archaea detected in the production waters from a high-temperature petroleum reservoir. *Int. Biodeterioration Biodegradation* 143, 104729. doi: 10.1016/j.ibiod.2019.104729

frontiers | Frontiers in Microbiomes

TYPE Original Research
PUBLISHED 13 November 2024
DOI 10.3389/frmbi.2024.1488744

OPEN ACCESS

EDITED BY
Gareth Trubl,
Lawrence Livermore National Laboratory (DOE), United States

REVIEWED BY
Yong Zheng,
Fujian Normal University, China
Rubens Tadeu Delgado Duarte,
Federal University of Santa Catarina, Brazil

*CORRESPONDENCE
Brent Christner
xner@ufl.edu

†PRESENT ADDRESS
Christina Davis,
McGill University, Department of Natural Resource Sciences, Sainte-Anne-de-Bellevue, QC, Canada

RECEIVED 30 August 2024
ACCEPTED 22 October 2024
PUBLISHED 13 November 2024

CITATION
Faber Q, Davis C and Christner B (2024) Metagenomic inference of microbial community composition and function in the weathering crust aquifer of a temperate glacier.
Front. Microbiomes 3:1488744.
doi: 10.3389/frmbi.2024.1488744

Metagenomic inference of microbial community composition and function in the weathering crust aquifer of a temperate glacier

Quincy Faber, Christina Davis† and Brent Christner*

Department of Microbiology and Cell Science, University of Florida, Gainesville, FL, United States

Bacterial, fungal, and algal communities that colonize aquatic systems on glacial ice surfaces mediate biogeochemical reactions that alter meltwater composition and affect meltwater production and storage. In this study, we sought to improve understanding of microbial communities inhabiting the shallow aquifer that forms seasonally within the ice surface of a glacier's ablation zone (i.e., the weathering crust aquifer). Using a metagenomic approach, we compared gene contents of microbial assemblages in the weathering crust aquifer (WCA) of the Matanuska Glacier (Alaska, USA) to those recovered from supraglacial features and englacial ice. High abundances of Pseudomonadota, Cyanobacteriota, Actinomycetota, and Bacteroidota were observed across all samples, while taxa in class Gammaproteobacteria were found at significantly higher abundances in the weathering crust aquifer. The weathering crust aquifer samples also contained higher abundances of Dothideomycetes and Microbotryomyetes; fungal classes commonly observed in snow and other icy ecosystems. Phylogenetic analysis of 18S rRNA and *rbcL* gene sequences indicated high abundances of algae in the WCA that are closely related (> 98% and > 93% identity, respectively) to taxa of *Ancylonema* (Streptophyta) and *Ochromonas* (Ochrophyta) reported from glacial ice surfaces in Svalbard and Antarctic sea ice. Many functional gene categories (e.g., homeostasis, cellular regulation, and stress responses) were enriched in samples from the weathering crust aquifer compared to those from proximal englacial and supraglacial habitats, providing evidence for ecological specialization in the communities. The identification of phagotrophic phytoflagellate taxa and genes involved in mixotrophy implies that combined phototrophic and heterotrophic production may assist with persistence in the low light, low energy, and ephemeral conditions of the weathering crust environment. The compositional and functional differences we have documented indicate distinct microbial distributions and functional processes occur in the weathering crust aquifer environment, and we discuss how deciphering these nuances is essential for developing a more complete understanding of ecosystem biogeochemistry in supraglacial hydrological systems.

KEYWORDS

metagenomics, ice algae, supraglacial, weathering crust aquifer, glacial ice

Introduction

The seasonal generation of meltwater on the surface of glaciers and ice sheets creates liquid water that allows solute transport, microbial metabolism, and biogeochemical transformations in snow- and ice-based aquatic ecosystems during the melt season. Examples of these include algal- and cyanobacterial-based communities that colonize ice (Yallop et al., 2012; Remias et al., 2012), snow (Hoham and Duval, 2001), cryoconite holes (Hodson et al., 2010), and supraglacial streams (Irvine-Fynn et al., 2012; Foreman et al., 2013). In the ablation zone (i.e., where annual melting leads to net loss of ice from the glacier) on western portions of the Greenland Ice Sheet, the growth of darkly pigmented algae generates large blooms (Williamson et al., 2020) that accelerate ice melt (Yallop et al., 2012; Remias et al., 2012). Cryoconite holes are water-filled depressions of dark sediment that have melted into the ice surface and contain communities composed of phototrophs, heterotrophic bacteria, and fungi, as well as ciliates, tardigrades, and rotifers (Cook et al., 2016a; Zawierucha et al., 2015). Streams flowing in channels on the glacier surface are also known to be biogeochemical hotspots, with microbial metabolism altering nutrient compositions of water during its transport through the supraglacial hydrological system (Scott et al., 2010). Research on the microbiomes and biogeochemistry of glacial environments has accelerated over the last decade due to the realization of the rapid changes currently underway (Constable et al., 2023) and an increased awareness of their contributions to global carbon cycling (Anesio et al., 2017; Wadham et al., 2019).

Water observed in standing pools or flowing in streams on a glacier's surface represents the portion of the meltwater that can be most rapidly exported to subglacial and proglacial environments (Varliero et al., 2023). This meltwater is sourced from water-bearing porous ice in the near-surface that forms a perched aquifer termed the weathering crust aquifer (WCA; Karlstrom et al., 2014; Cook et al., 2016b; Cooper et al., 2018). The WCA forms and persists in the ablation zone as seasonal increases in air temperature and incident shortwave radiation cause warming and internal melting to depths of 30 to 100 cm in the ice (Hoffman et al., 2014; Cooper et al., 2018; Irvine-Fynn et al., 2021). Given the presence of liquid water and light, photosynthesis is possible in the WCA (Hodson et al., 2013; Irvine-Fynn and Edwards, 2014; Christner et al., 2018). Flow through tortuous paths in the porous ice tends to increase residence time relative to water in other supraglacial environments (Stevens et al., 2018); conditions that may also enhance microbial interactions, biogeochemical processing, and production of new biomass. This contention is supported by data from western Greenland that estimate 37 kg km^{-2} of the cellular carbon produced in the WCA is exported to proglacial waters every melt season, representing a significant and overlooked contribution to global carbon cycling (Irvine-Fynn et al., 2021). The hydrological properties of the weathering crust decrease ice albedo (Tedstone et al., 2020) and enhance retention of nutrients that support algal growth (Holland et al., 2019); however, little is known about microbial nutrient processing in the WCA. Though previous studies have documented bacterial communities in WCA pore waters that differ from those associated with other supraglacial features (Christner et al., 2018; Rassner et al., 2024), it is not well understood if the WCA ecosystem is functionally distinct or contains similar metabolisms and taxa to those in proximal supraglacial microbial communities.

Given that environmental changes in glaciated regions have important social, scientific, and environmental impacts globally (e.g., Heckmann et al., 2016), there is a critical need to improve understanding on the microbiological sources and sinks of nutrients cycled in ice-based ecosystems and biogeochemical consequences of the meltwater discharged to proglacial environments. The ubiquity of algae and cyanobacteria in supraglacial communities (e.g., Anesio et al., 2009; Cook et al., 2012) has led to an assumption that autotrophic activity is driving carbon metabolism, yet studies of cryoconite holes have shown many to be net heterotrophic (Hodson et al., 2010). Algal blooms on glacier surfaces (Cook et al., 2012; Williamson et al., 2018) that coincide with higher abundances of presumably heterotrophic bacteria (Nicholes et al., 2019) and fungi (Perini et al., 2019) suggest these net autotrophic communities fuel supraglacial food webs. While certain fungi are known to be parasites of glacier algae (i.e., Chytridiomycota; Perini et al., 2022; Nakanishi et al., 2023), few studies have explored the diversity and role of fungal communities in supraglacial environments (Perini et al., 2022; Fiołka et al., 2021).

A previous study of the temperate Matanuska Glacier (Alaska, USA) during the melt season inferred internal melting from solar radiation to a maximum depth of ~2 m, estimated that ~10% of the surface ice volume was liquid water (200 L m^{-2}), and revealed meltwaters containing growing microbial populations with estimated doubling times of ~two weeks (Christner et al., 2018). Sequencing and analysis of amplified small subunit (SSU) rRNA genes showed that samples from the near surface, water-bearing ice were enriched with certain taxa within the Alphaproteobacteria, Bacteroidota, and Cyanobacteriota, and contained a bacterial community composition distinct from assemblages detected in supraglacial waters and englacial ice. Phylotypes closely related to common ice and snow algae (i.e., *Ancylonema nordenskioeldii*, *Mesotaenium* sp. AG-2009-1, *Ochromonas* sp. CMP1899, and *Chrysophyceae* sp. 176) were also identified. However, the SSU rRNA gene data are poorly suited for resolving subtle ecological variation in natural populations and predicting the metabolic potential of the communities. In this study, we tested the hypothesis that the physical and hydrogeochemical conditions in the WCA select for a community comprised of bacterial and eukaryotic taxa distinct from the varieties in proximal supraglacial features. To do this, we used a metagenomic approach to more thoroughly examine community composition and functional gene content of supraglacial communities associated with the Matanuska Glacier. The purpose of our study was to holistically examine the bacterial, archaeal, and eukaryotic communities of the WCA, examine the complement of genes they possess, make predictions of their functional characteristics, and compare them to proximal supraglacial habitats. Indeed, the bacterial, algal, and fungal assemblages in the WCA possess distinct compositions and structures, and we discuss how many of the functions identified are highly relevant to survival in the low temperature, light, and nutrient conditions of an ecotone bridging the supraglacial and englacial environment.

Materials and methods

Site description and sample collection

The Matanuska Glacier is a temperate valley glacier in the Chugach Mountains of south-central Alaska, USA. The samples analyzed in this study were collected during June and July of 2014 and 2015 at a field site (61°42′9.3" N, 147°37′23.2" W) on the glacier's western margin and ~8 km up glacier from the terminus. During the study, ten boreholes (ranging from 4 to 30 m in depth) were melted into the glacier that allowed meltwater primarily sourced from the englacial ice and the near-surface WCA to be sampled (Christner et al., 2018).

The range of borehole depths examined were chosen to collectively access the transition of weathered ice on the surface into the underlying sequence of englacial ice. To sample englacial ice, we selected depths exceeding 5 m because it provided high confidence that the ice had not been affected by surface effects (i.e. solar radiation and melting). Subsequent modeling of ice temperature and downward radiation data implied that liquid water from internal melting was unlikely at depths > 2 m (Christner et al., 2018). Extensive detail on the methods used to generate the boreholes and sample the meltwater have been described previously (Christner et al., 2018). Briefly, the near-surface ice (depth < 5 m) and englacial ice (depths of 5 to 15 m; Figure 1) were sampled by melting boreholes into the surface using equipment that was precleaned with 3% (v/v) hydrogen peroxide. This generated meltwater that was subsequently transferred to the surface using clean tubing and a peristaltic pump (Figures 1C, D). Two of the boreholes (BH6 and BH7) were melted using technology that allowed sampling at discrete depths (1.1 to 15.4 m; Supplementary Table S1). To sample the WCA, the water generated from melting boreholes of ~4 m depth was discarded, and following refill with water that had percolated laterally through the near surface ice (Figure 1B), the material that collected in the borehole was filter concentrated and preserved. A supraglacial stream ~100 m upstream of the primary sampling site was also sampled (Figure 1A). For most samples, 20 to 60 L of meltwater was filtered using a peristaltic pump (Supplementary Table S1); however, several samples were obtained from large volumes of water (~400 L) with a deployable borehole filtration system (Christner et al., 2014). Cells in the water samples were concentrated on 90 mm or 142 mm Supor 0.2 µm pore size filters (Pall 66549), preserved in 40 mM EDTA pH 8.0, 50 mM Tris pH 8.3, and 0.73 M sucrose, shipped frozen overnight from Alaska to the laboratory, and were stored at -80°C until analyzed.

DNA extraction and sequencing

DNA was extracted from cells concentrated on the filters using the PowerWater DNA Isolation Kit (MO BIO Laboratories, Inc.) and the modifications described by Christner et al. (2018). The extracted DNA was quantified with the dsDNA HS Assay Kit (ThermoFisher Scientific) and a Qubit 3 fluorometer. Samples with low genomic DNA yields (< 0.3 ng mL^{-1}; indicated in Supplementary Table S1) were amplified using the GenomiPhi V3 DNA Amplification kit (Cytiva). Following the manufacturer's protocol, 10 µL of the DNA sample was added to 10 µL of denaturing buffer, followed by addition to the Ready-To-Go GenomiPhi cake for amplification at 30°C using a Bio-Rad C1000 Touch thermal cycler. After 1.5 h, the reaction was terminated by heating to 65°C for 10 min and the samples were stored at 4°C for subsequent processing.

Library preparation and sequencing were performed at the ICBR Next Gen DNA Sequencing Core at the University of

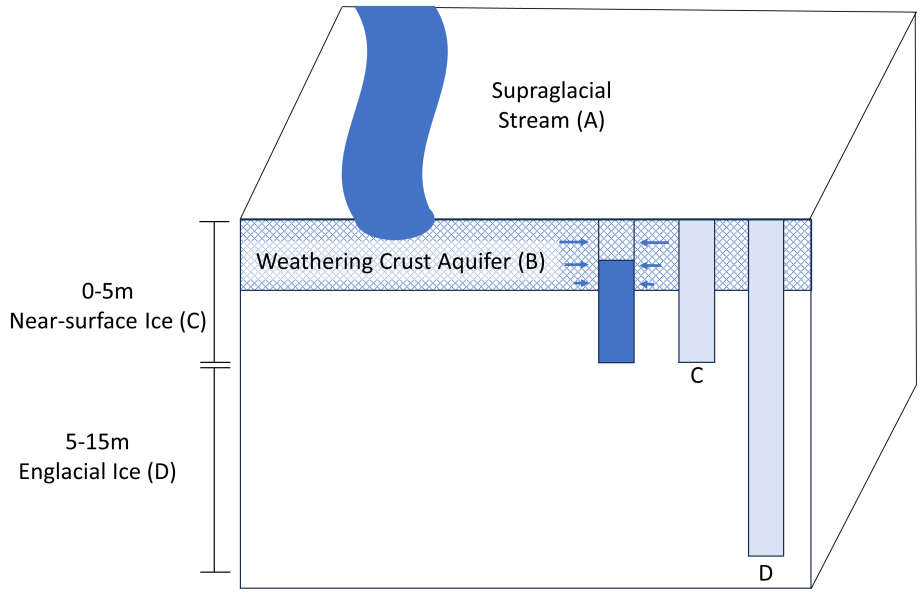


FIGURE 1
Schematic of the supraglacial and englacial features sampled in this study from the ablation zone of the Matanuska Glacier including the supraglacial stream **(A)**, weathering crust aquifer **(B)**, near-surface ice **(C)**, and englacial ice **(D)**.

Florida. Due to low DNA extraction yields for many of the samples, libraries were constructed using the Ovation Ultralow system v2 kit. Quality and concentration of the libraries were checked using an Agilent 4150 TapeStation system and Qubit fluorometer, respectively. Paired-end sequencing was conducted on the NextSeq500 2x150 MID platform (Illumina, Inc.).

Read processing and assembly

The sequencing reads were paired and low-quality reads were removed using Trimmomatic 0.39 (LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:36; Bolger et al., 2014). For analyses with unassembled sequences, paired-end reads were merged using BBmerge in BBMap Version 37.62 (Bushnell et al., 2017). *De novo* assembly of reads into contigs and scaffolds was performed using SPAdes 3.13.0 and default parameters (Bankevich et al., 2012). MetaBAT (V2.13; Kang et al., 2019) was used to create bins that were assessed for completeness using CheckM (V 1.1.2; Parks et al., 2015).

Taxonomic and functional analysis

The reads containing sequence derived from SSU rRNA genes were extracted using Parallel Meta 3.6. OTUs were assigned in QIIME2 V2023.7 using the SILVA 138 99% OTUs classifier trained on full-length sequences and taxonomy (Bokulich et al., 2018; Quast et al., 2013). Only samples containing > 100 mapped rRNA gene sequences were used for alpha and beta diversity analysis (Supplementary Table S3). Simpson's Diversity Index and Good's estimator of coverage were calculated for each sample in R using the vegan package and the QsRutils package, respectively. Reads containing the ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit gene (*rbcL*) and the internal transcribed spacer 2 region (ITS2) were identified using BlastN v2.9 (Camacho et al., 2009) with a custom database of all *rbcL* genes and fungal ITS2 sequences in the NCBI Reference Sequence database as of August 2020. QIIME 1.9.1 was used for reference-based OTU assignment and taxonomic identifications were conducted using the custom databases from RefSeq at the 97% sequence identity threshold (Caporaso et al., 2010). Nonmetric multidimensional scaling (NMDS) plots based on Bray-Curtis distance (iterations = 20) were visualized using the vegan package in R. Abundance tables based on the number of reads identified as SSU rRNA and *rbcL* genes, as well as the ITS2 region, were visualized using the ggplot2 package in R (v4.1.2; R Core Team 2021).

The sequences of 18S rRNA and *rbcL* genes were identified in the contigs using BlastN v2.9 (Camacho et al., 2009) and examined with the assistance of a database created using nucleotide sequences available in NCBI and with an E-value cutoff < 1e-100. To generate alignments of partial 18S rRNA genes for phylogenetic analysis, the sequences were aligned based on secondary structure using the SINA aligner (v1.2.12, Pruesse et al., 2012). For the *rbcL* sequences, the top five BLAST hits for each query sequence were aligned using MAFFT v7 (Kuraku et al., 2013). The alignments were used for phylogenetic analysis by maximum likelihood, using the Jukes-Cantor model with 100 bootstraps, and to generate distance matrices in MEGAX v11.0.10 (Kumar et al., 2018).

A kmer-based strategy was used to identify the contigs as prokaryotic, or eukaryotic with Eukrep (West et al., 2018). Functional predictions of bacterial and archaeal genes encoded by the contigs were conducted using Prodigal v2.6.3 (Hyatt et al., 2010) and eukaryotic genes were predicted using MetaEuk v6 (Levy Karin et al., 2020). Genes were aligned to the SWISS-PROT database using BlastX in DIAMOND v2.0.9.147 (Buchfink et al., 2021; Bairoch and Apweiler, 2000) and their taxonomy was assigned using MEGAN6 (Huson et al., 2016). All metagenomic reads were mapped to the predicted genes using Bowtie2 v2.4.5 (Langmead and Salzberg, 2012), duplicates were removed using SAMtools v1.18 (Li et al., 2009), and the "pileup.sh" script in BBMap v38.9 (Bushnell et al., 2017) was used to summarize the number of reads mapped to each gene. The outputs of these analyses were exported to R for statistical analysis and the number of reads per gene was normalized to gene length.

Statistical analysis

The Shapiro-Wilk test was used to test normality and correlations between alpha diversity and quantity of extracted DNA (i.e., prior to genome amplification of select samples) were tested using Spearman's rank order correlation analysis in R. To determine if community composition among samples based on the 16S and 18S rRNA gene data were significantly different ($\alpha = 0.05$), Adonis tests were run using Vegan in R. The results from MANOVA tests and *post hoc* analysis with Tukey's Honest Significant Difference test were used to determine statistical significance of relative abundances of taxa among sample types. To examine correlations in the relative abundances of bacteria, fungi, and algae, a Pearson correlation analysis was done in R using the "Hmisc" package and visualized using the "corrplot" package. ANOVA tests were used to compare the relative abundance of genes assigned to bacteria, algae, and fungi among sample types. The Wald test was used to determine statistically significant differences in the abundance of OTUs and functional genes with the DESEQ2 v1.40.1 tool (Love et al., 2014).

Results

Metagenomic sequencing output

Table 1 summarizes the sequencing results from 39 DNA samples extracted from a supraglacial stream (2 samples), meltwater from the WCA (9 samples), near-surface ice (15 samples from depths of 0 to 5m below the surface), and englacial ice (13 samples from depths of 5 to 15 m; Figure 1). A total of 174 million paired end (2 × 150 bp) reads were generated, which corresponded to an average of 4 million reads per sample (range of ~75,000 to 60 million reads per sample; Supplementary Table S1). Most sequence reads classify as bacterial or eukaryotic (Table 1), with < 0.01% of the

reads (< 0.1% of total genes) identified as archaeal (i.e. taxa within the Methanococci, Methanomicrobia, and Thermococcales). From these data, 10 million contigs were assembled with a maximum length of 372,000 bp (Supplementary Table S2). Attempts at *de novo* assembly for each sample failed to generate a large number of high-quality metagenome-assembled genomes; only 8 bins were obtained with completeness > 80% and contamination < 10%. Presumably, this was due to the short length of the contigs used for the assembly and low coverage for many of the libraries (Supplementary Table S3). Therefore, we restricted further analysis to the unassembled reads and contigs containing protein-encoding genes.

Community composition based on small subunit rRNA sequences in the metagenomes

OTUs were assigned in QIIME2 for the 32 samples that contained > 100 SSU rRNA gene sequences. For 16S rRNA genes, 5,252 bacterial OTUs and 0 archaeal OTUs were generated from a total of 104,932 total reads, whereas fewer eukaryotic 18S rRNA genes (69,421 reads) and OTUs (1,770) were observed (Supplementary Table S3). To assess this approach, we analyzed sequence data obtained using samples of the ZymoBIOMICS™ Microbial Community Standard (Catalog No. D6300) and found that the analysis in QIIME produced genus and phylum level predictions that agreed well with the species abundances in the standard (Supplementary Figure S1).

Simpson's complementary diversity (i.e., 1-D) indices for bacteria are high and range from 0.98 to 0.99, with the highest diversity observed in the stream samples (average of 0.988; Supplementary Table S3). The Good's coverage values of species richness for the 16S rRNA gene sequences were lowest in supraglacial samples (average of 0.34) and higher coverage was observed in the WCA (0.18 to 0.81), near-surface ice (0.18 to 0.90), and englacial ice (0.33 to 0.95) samples (Supplementary Table S3). This coverage is much lower than values > 0.96 based on amplicon data from these samples (Christner et al., 2018). From analysis of the 18S rRNA gene data, a lower diversity of eukaryotes was indicated by Simpson's complementary diversity values of 0.81 to 0.99. While Good's coverage for the 18S and 16S rRNA gene data ranged widely (0.125 to 0.95), the values are not statistically different (ANOVA, $p > 0.05$) among sample types. Good's coverage (Supplementary Table S3) is moderately positively correlated with the quantity of DNA used to prepare the libraries (Spearman's $r = 0.431$, $p < 0.01$, $n = 39$), consistent with low coverage being due to the lower amounts of extractable DNA in the samples. The trimmed sequence count was positively correlated with Good's coverage values for 16S and 18S rRNA genes ($r > 0.6$, $p < 0.001$; Supplementary Table S1), but did not significantly correlate to Simpson's complementary diversity. Because we analyzed non-overlapping rRNA gene sequences from the metagenomes, the number of unique OTUs and singletons is likely to be artificially inflated, also contributing to low Good's coverage values.

Bacterial beta diversity was visualized in NMDS plots of Bray-Curtis dissimilarity distances based on 16S rRNA genes extracted from metagenomic data (Figure 2A) and compared with data derived from amplicon sequencing (Christner et al., 2018; Supplementary Figure S2). The two-dimensional NMDS ordination for bacteria showed similar clustering of samples between methods and differentiated englacial samples from those of the supraglacial stream, WCA, and near-surface ice (Figure 2A), but it has a relatively high ordination stress value of 0.2209. Therefore, we also conducted a PCoA analysis of the 16S rRNA gene data (data not shown), which resulted in a similar clustering of samples. Bacterial assemblages in englacial ice are significantly different from the other samples (p value < 0.05, Adonis), as was also found for amplicon data by Christner et al. (2018). Most of the bacterial community samples from the WCA are most similar to those observed in bulk sampling of near-surface ice. This is not entirely unexpected because the meltwater sampled from the near-surface ice would contain a mixture of taxa originating from the WCA and from the melting of englacial ice. Bacterial community structure in the WCA was highly heterogenous, particularly in the samples collected during three consecutive days in 2015. The NMDS ordination for the 18S rRNA gene data resulted in a lower stress value of 0.1594 and showed compositions that were relatively

TABLE 1 Comparison of basic data outputs from the DNA sequencing and phylogenetic binning of reads from each sample type.

	Supraglacial Stream	Near-surface Ice (0-5m)	Weathering Crust Aquifer	Englacial Ice (5-15m)
Number of samples	2	15	9	13
Reads per sample	1295808	5291961 ± 14523424	3524928 ± 3100957	4385548 ± 5230017
Total number of reads	2591612	79379413	31724351	53792038
% Bacteria	89.2	80.7 ± 22.4	73.9 ± 23.2	86.2 ± 16.7
% Fungi	2.9	6.3 ± 12.5	14.6 ± 19	1 ± 0.5
% Algae	0.6	0.6 ± 0.4	1 ± 0.7	0.3 ± 0.2
% Other Eukarya	1.5	6.6 ± 10.8	4.5 ± 5.9	4.7 ± 6.3
Number of contigs	220861	308167 ± 661535	262737 ± 170151	199043 ± 179982

Values are averages ± the standard deviation of the mean.

heterogeneous within themselves and not well differentiated by environmental type. The only significant differences in composition among the eukaryotic communities were between sample years (p value < 0.05, Adonis; Figure 2B).

Pseudomonadota is the most abundant bacterial phylum in the samples (9.9 to 93.5% relative abundance; Figure 3A), with most taxa affiliated with class Betaproteobacteria, except for the englacial samples from 2014 in which Gammaproteobacteria were more numerous, similar to results from 16S rRNA gene amplicon sequencing (Christner et al., 2018). The 20 most abundant bacterial OTUs (Supplementary Table S4) are Pseudomonadota, and 12 of these classify within the genus *Pseudomonas*. Actinomycetota (0.7 to 24%), Bacteroidota (0.4 to 17%), and Patescibacteria (0 to 35%) were found at high abundances in the majority of samples, and Cyanobacteriota (0 to 25%) was found in several near-surface ice (BH07a_rep1 and BH003_rep2) and englacial ice (BH06b_rep1, BH06c_rep2, and BH07c_rep3) samples (Figure 3A). While many of the same bacterial phyla were shared and found in similar abundances among the stream, near-surface ice, WCA, and englacial ice samples, compositional differences were observed for each environmental type. For instance, there are significantly more taxa from the phylum Deinococcota (31 OTUs) in the supraglacial stream samples compared to those from the WCA (MANOVA, $p < 0.01$, $n = 2$). The number of gammaproteobacterial 16S rRNA gene reads (807 OTUs) varied significantly among samples (p value < 0.05, MANOVA), with higher numbers of taxa in the WCA than in englacial ice. DESEQ2 analysis showed that 15 and 8 OTUs classifying within the genera *Brevundimonas* and *Stenotrophomonas*, respectively, had significantly higher abundances in the WCA as compared to other samples (Supplementary Table S5).

Fungi and unicellular algae were the numerically dominant taxa in the eukaryotic communities, representing 1% to 15% and 0.3% to 1%, respectively, of the total sequencing reads (Table 1). The 20 most abundant OTUs are members of Ascomycota, Basidiomycota, Charophyta, and Ochrophyta divisions. Within the Basidiomycota (1 to 57% of the 18S rRNA sequences), OTUs in class Microbotryomycetes are most abundant, followed by Agaricomycetes, Malasseziomycetes, and Wallemiomycetes (Figure 3B). Ascomycota was the second most abundant fungal phylum (0 to 64% of the 18S rRNA sequences) and contained OTUs belonging to Dothideomycetes, Eurotiomycetes, Saccharomycetes, Sordariomycetes, and Tremellomycetes (Figure 3B). The distribution of fungal classes varies amongst individual samples, but abundance is not significantly different based on sample type. For example, Saccharomycetes was the most abundant group in the near-surface sample BH07a_rep2, with a relatively low abundances across the remainder of samples. Similarly, Sordariomycetes dominated two WCA samples (BH10b_rep1 and BH10b_rep2) but was found at lower abundance in the other WCA samples. The abundance of OTUs from the phylum Chytridiomycota was

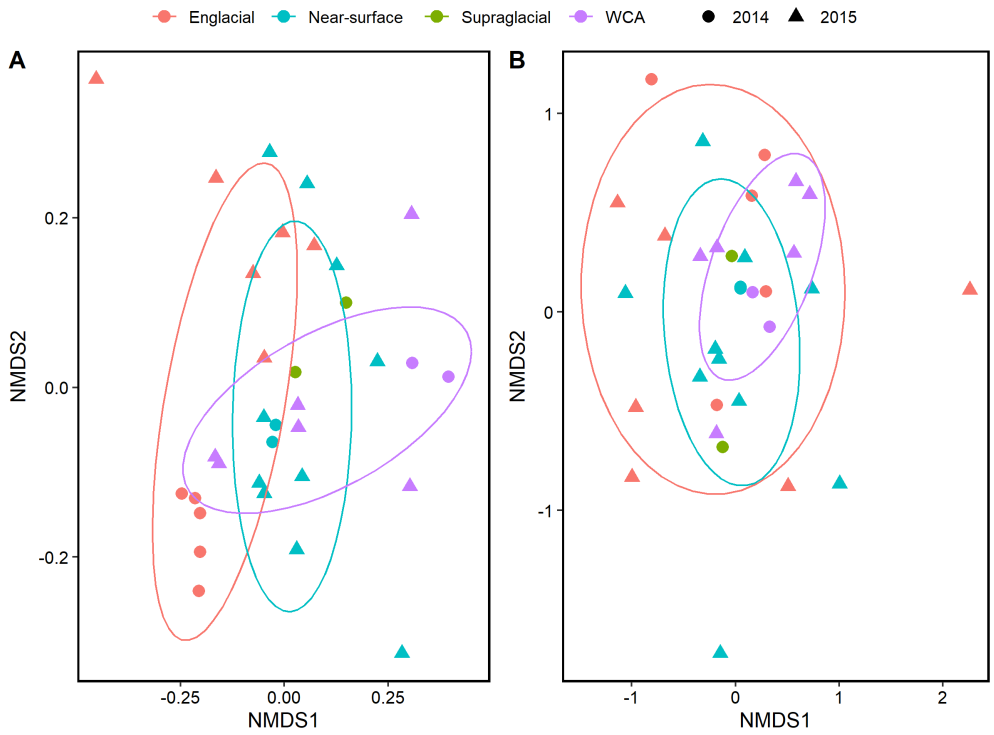


FIGURE 2
NMDS of Bray-Curtis distance among samples collected during 2014 and 2015 from the supraglacial stream, WCA, and ice from the near-surface and englacial zone. Similarity calculations and statistical significance were performed using Adonis with the vegan package in R. The ellipses indicate the clustering of each sample type with a confidence level of 0.70. **(A)** Analysis based on 16S rRNA genes (stress=0.2209). **(B)** Analysis based on 18S rRNA genes (stress=0.1594). The plots were generated using ggplot2 in R.

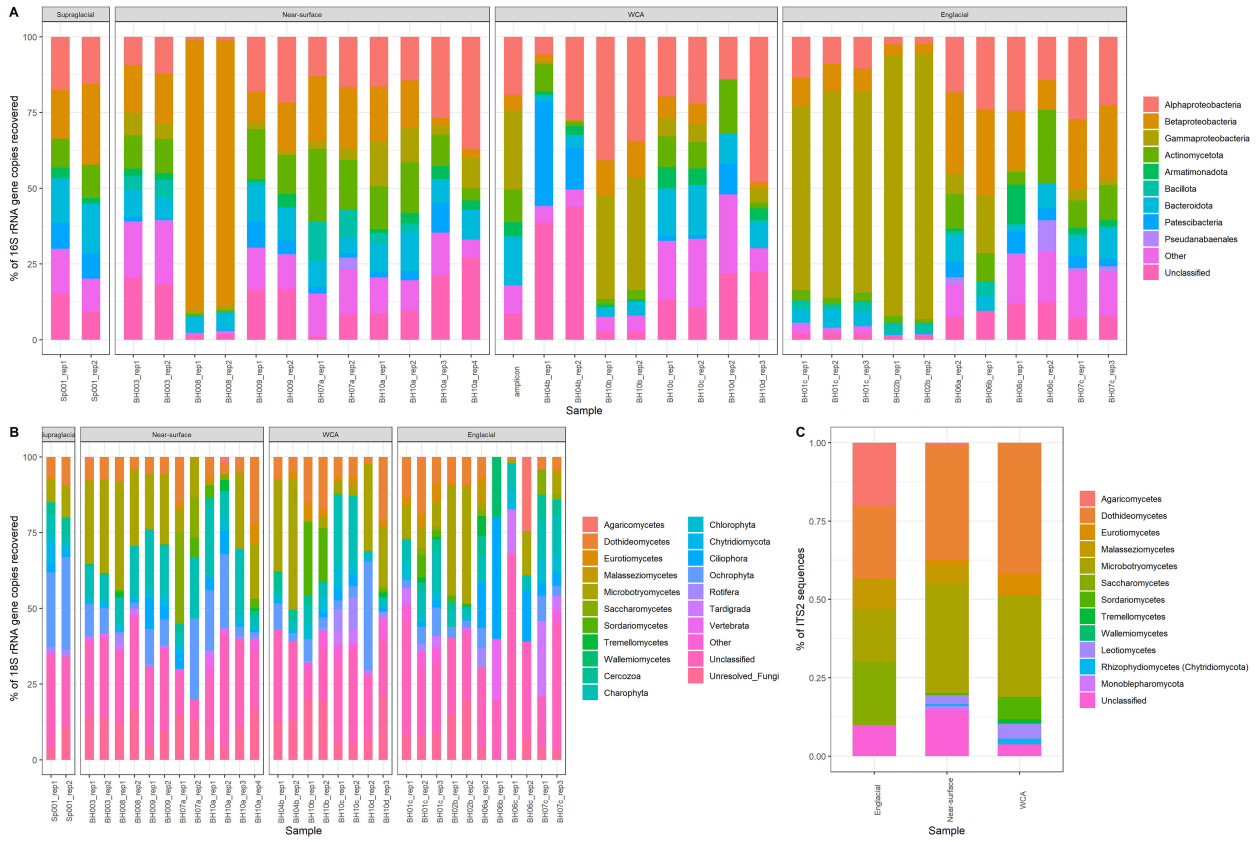


FIGURE 3
Community composition based on SSU rRNA genes and ITS2 sequences retrieved from metagenomes. **(A)** Relative abundances of bacteria phyla based on 16S rRNA genes, except for Pseudomonadota, which is shown at the class level, and Cyanobacteriota, which is shown at the order level. **(B)** Relative abundances of eukaryotic groups based on 18S rRNA genes. The "Other" category contains taxa with abundances <0.01%. **(C)** Relative abundance of fungal classes based on the internal transcribed spacer 2 (ITS2) sequence.

highest in the supraglacial sample Sp001-rep1 (6.8%). Tardigrada, and Rotifera phyla were present in some of the samples (0% to 25% and 0% to 10%, respectively). Vertebrata were also present in low abundances (0.36% average relative abundance), with BH10a_rep1 containing 4% relative abundance, and the lowest classification assigned to the OTUs is the class Mammalia.

Certain algae were highly abundant members of the eukaryotic communities, including green (Charophyta and Chlorophyta; 0%-24% and 0%-8%, respectively, of the 18S rRNA gene sequences) and golden (Ochrophyta; 0%-36%) algae. Samples from the supraglacial stream contained significantly more 18S rRNA gene reads for Ochrophyta (MANOVA, $p < 0.01$, $n = 2$; 39%-41%) than other sample types (near-surface: 0%-36%; WCA: 1%-41%; englacial ice: 0%-16%; Figure 3B). A comparison of bacterial and eukaryotic phylotypes in the WCA samples identified algal divisions with abundances highly positively correlated ($r > 0.7$, $p < 0.05$) to eight bacterial phyla or classes (Figure 4). Charophyta are positively correlated to Rotifera, Tardigrada, and five bacterial phyla (Armatimonadota, Bacteroidota, Gemmatimonadota, Myxococcota, and Planctomycetota). The relative abundance of Chlorophyta is negatively correlated to Basidiomycota, whereas Ochryophyta positively correlates to three bacterial phyla (Actinomycetota, Chloroflexota, *Candidatus* Eremiobacterota; Figure 4). The correlations between relative abundances of specific taxa to major algal groups was also supported by analyses conducted at lower taxonomic hierarchies. OTUs in Zygnematophyceae and Chrysophyceae positively correlate with those in the bacterial phyla Actinomycetota, Parcubacteria, and Pseudomonata (Table 2). In addition, Zygnematophyceae positively correlates with 4 OTUs in three bacterial phyla (Armatimonadota, Bacteroidota, and Cyanobacteriota), whereas Chrysophyceae negatively correlates to an OTU in the Armatimonadota phyla. There is also a positive correlation among OTUs belonging to the Zygnematophyceae and Chrysophyceae classes and several groups of fungi (i.e. Ascomycota and Basidiomycota; Table 2).

Analysis of 1000 bp 18S rRNA gene sequences on contigs recovered from the supraglacial stream, WCA, and near-surface ice samples provided information on the phylogeny of organisms within major algal divisions (Figures 5A, B). We identified various green and golden algal taxa (9 and 10 sequences, respectively) that are highly similar to those reported in previous studies of snow, glacial ice, and other icy ecosystems. For instance, eight of the 18S rRNA gene sequences are identical to species of *A. nordenskioeldii* that have been documented in supraglacial environments of Svalbard and Switzerland (Leya, 2004; Procházková et al., 2024; Figure 5A). A separate cluster of 18S rRNA gene sequences

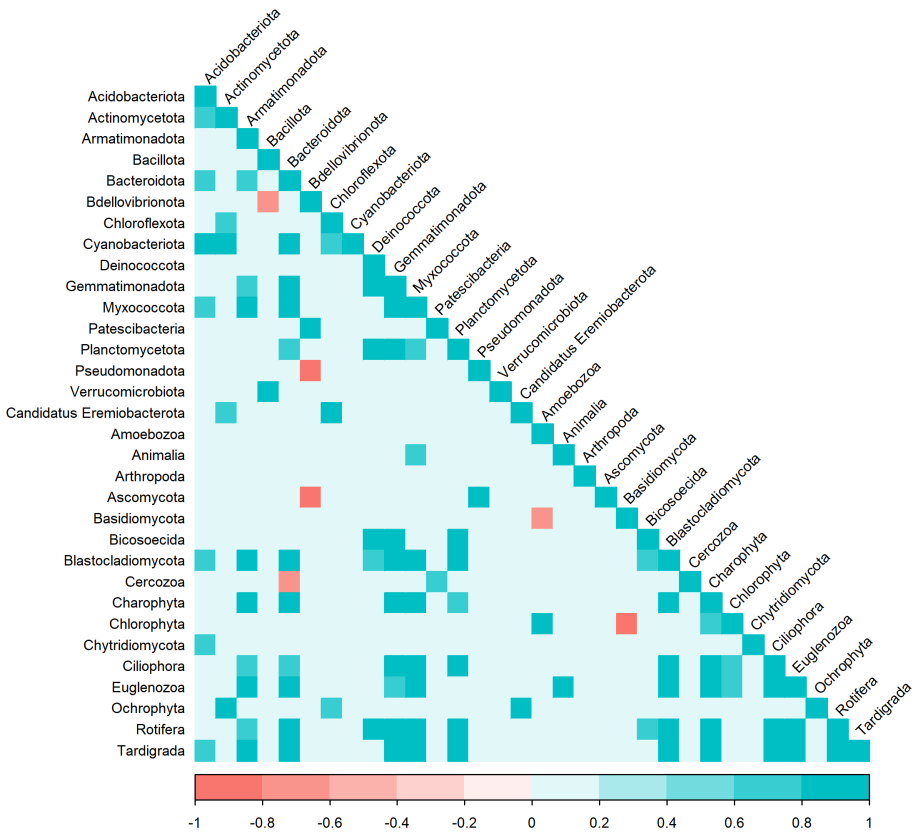


FIGURE 4
Correlation of bacterial (phylum and class) and eukaryotic (phylum and divisions) SSU rRNA phylotypes in WCA samples. The heat plot was generated based on Pearson correlation in the "Hmisc" package. Statistically different ($p < 0.05$) and highly correlated variables are shown for groups with greater than 0.5% abundance in the samples.

phylogenetically related to golden algae in the genus *Ochromonas* (Figure 5B) contains nine OTUs with 99.9% identity to *Ochromonas* CCMP 1899, a mixotrophic golden algal species isolated from Antarctic sea ice (Brodie and Lewis, 2007).

rbcL and ITS2 sequences in the metagenomes

Given the limited phylogenetic resolution provided by the 18S rRNA gene, further analysis of molecular diversity in the eukaryotic community was directed towards analyzing sequences of the internal transcribed spacer 2 (ITS2) and *rbcL* (large subunit of ribulose bisphosphate carboxylase) gene that were identified in the metagenome data. Analysis of 2,568 ITS2 sequences retrieved from the near-surface, WCA, and englacial samples identified 10 classes of fungi (Figure 5C) that largely agreed with compositions based on the 18S rRNA gene (15 classes identified; Figure 5B), but with notable differences. Both the 18S rRNA gene- and ITS2-based analyses showed that the divisions Ascomycota and Basidiomycota dominated, with the Dothideomycetes and Microbotryomycetes being the most abundant classes in these divisions, comprising 34% and 28%, respectively, of ITS2 sequences (Figures 5B, C). These methods also indicate that Chytridiomycota are in low abundance. However, when comparing relative abundances of 18S rRNA gene sequences to ITS2 sequences, Dothideomycetes was identified at lower abundances in the 18S rRNA gene sequences (15.4% of 18S rRNA gene fungal sequences), while Microbotryomycetes had a similar abundance (30.1% of 18S rRNA gene fungal sequences). In addition, Saccharomycetes, Agaricomycetes, and Malasseziomycetes taxa that were at relatively high abundances (5 to 7%) in the ITS2 sequences were inferred to have very low abundances in the 18S rRNA gene data (< 0.7% of fungal sequences; Figures 5B, C).

Analysis of 858 *rbcL* reads showed that most sequences (75.4%) were closely related to taxa in the phototrophic genera *Ancylonema*, *Cylindrocystis*, *Mesotaenium*, and *Ochromonas* (Supplementary Figure S3). *Ancylonema* and *Cylindrocystis* were most abundant in the WCA, making up half of the *rbcL* reads in these samples. *Ancylonema* was the most abundant genus in supraglacial, near-surface, and englacial samples based on the distribution of reads for *rbcL* (Supplementary Figure S3) and 18S rRNA (Figure 5B) genes. Analysis of *rbcL* sequences also confirmed the presence of taxa in phylum Ochrophyta (clade Stramenopile) that were identified from 18S rRNA gene analysis (Figure 5B). Of the *rbcL* genes in the supraglacial samples, the most abundant were Ochrophyta taxa (25%), which were at lower abundances in the other samples (6-10%; Supplementary Figure S3). Analysis of 387 bp *rbcL* gene sequences from contigs showed the samples contained seven *rbcL*

TABLE 2 Number of OTUs correlated to the relative abundance of major algal taxa in samples from the WCA.

Description	Correlation with Zygnematophyceae	Correlation with Chrysophyceae
Actinomycetia (Actinomycetota)	3+	1+
Armatimonadia (Armatimonadota)	2+	2-
Cytophagia (Bacteroidota)	3+	0
Cyanobacteriota-unclassified	1+	0
Parcubacteria-unclassified	1+	1+
Alphaproteobacteria (Pseudomonadota)	3+	1+
Betaproteobacteria (Pseudomonadota)	1+	0
Pseudomonadota-unclassified	1+	0
Ascomycota-unclassified	1+	0
Dothideomycetes (Ascomycota)	0	3+
Basidiomycota-unclassified	5+	4+
Microbotryomycetes (Basidiomycota)	10+	15+
Exobasidiomycetes (Basidiomycota)	1+	0
Chlorophyta-unclassified	1+	0
Chlorophyceae (Chlorophyta)	1+	0
Bdelloidea (Rotifera)	1+	0

Statistically significant ($p < 0.05$) and highly correlated results [$r > 0.6$ for positive associations (+) and $r < -0.6$ for negative associations (-)] are shown.

sequences with > 98% identity to that of *A. nordenskioeldii* and *A. alaskanum* (Figure 5C). Several *rbcL* sequences in the genus *Ochromonas* were also identified, and based on an operational taxonomic unit definition of > 98% identity for *rbcL* (An et al., 2018), the three sequences with 93% similarity to the *rbcL* gene of *Ochromonas* CCMP 1899 (Figure 5D) are likely to be a distinct species from the Antarctic sea ice isolate identified as their nearest neighbor by *rbcL* gene analysis.

Functional analysis of metagenomes

Functional annotation of assembled data using DIAMOND-Blastx analysis identified ~1.2 million bacterial and ~70,000 eukaryotic genes (Supplementary Table S2). Archaeal functional genes represented less than 0.1% of genes identified in the metagenomes and were excluded from the analysis. Based on sequence analysis conducted using InterPro, the dataset contains 150,796 bacterial and 11,413 eukaryotic protein-encoding genes. Most bacterial genes classify within the groups Actinomycetota, Alphaproteobacteria, Bacteroidota, Cyanobacteriota, and Gammaproteobacteria (Supplementary Figure S4A), largely consistent with compositions inferred from the 16S rRNA gene data (Figure 3A), with the exception of Patescibacteria, which was found at a much lower abundance (0.1%) compared to 16S rRNA gene data (3.1%). Compositional abundances differed greatly from those inferred from 18S rRNA gene analysis (Supplementary Figure S4B; Figure 3B). For instance, genes from Ochrophyta were found at very different abundances in the total gene data (18.5%) compared to the 18S rRNA gene data (7.9%), likely due to the lack of reference genomes available. However, Dikarya and Charophyta were found at similar abundances in both datasets. On average, ~90% of the genes identified from the samples were bacterial, but the fraction observed in the WCA samples is lower (76%) and significantly different ($p < 0.05$) from values in the near surface and englacial ice (Figure 6A). The lower fraction of bacterial genes in the WCA is offset by increased abundances of fungal genes (21% compared to <6% for other sites) that are significantly different from the englacial samples ($p < 0.01$; Figures 6B, C). Approximately 3.7% of the genes identified belong to the algal groups Charophyta, Ochrophyta, and Chlorophyta, and although slightly higher percentages of algal genes were observed in the WCA samples (8.5%), the values are not significantly different among locations (Figure 6B).

DESEQ2 was used to determine if functional genes and Gene Ontology terms (GO terms) were differentially distributed by environmental type (Supplementary Table S7). A total of 2,395 bacterial genes (1.6% of total genes) and 984 eukaryotic genes (8.6%) are differentially abundant in the WCA compared to one or more sample types (Wald test, $p < 0.05$). Bacterial gene functions in the WCA are most distinct from englacial ice, with 85 GO terms differentially abundant, compared to 6 and 13 GO terms when compared to supraglacial and near-surface ice samples, respectively. Eukaryotic gene functions also varied widely between the WCA and the near-surface and englacial ice, with 130 and 148 GO terms differentially abundant, respectively, and that span a wide range of biological processes (Supplementary Table S7). Many of the GO functions differentially abundant in the WCA are potentially related to adaptations that enhance environmental fitness and interactions within the ecosystem (Figure 7). The functions enriched in the WCA included genes associated with symbiotic interactions (GO:0085030), response to nutrients (GO:0007584), response to salt stress (GO:0009651), and processes that detoxify superoxide radicals (GO:0019430), whereas genes involved in substrate transport were depleted relative to englacial ice (Supplementary Table S7). Gene functions for response to UV (GO:0009411) and absence of light (GO:0009646) are also enriched in the WCA compared to near-surface and englacial ice (Figure 7B). Curiously, genes involved in response to DNA damage (GO:0006974) are depleted in the WCA compared to other sample types (Figure 7A). Gene content of assemblages in the WCA samples were also depleted for cellular response to water deprivation (GO:0042631) and nutrient levels (GO:0031669) when compared to englacial ice.

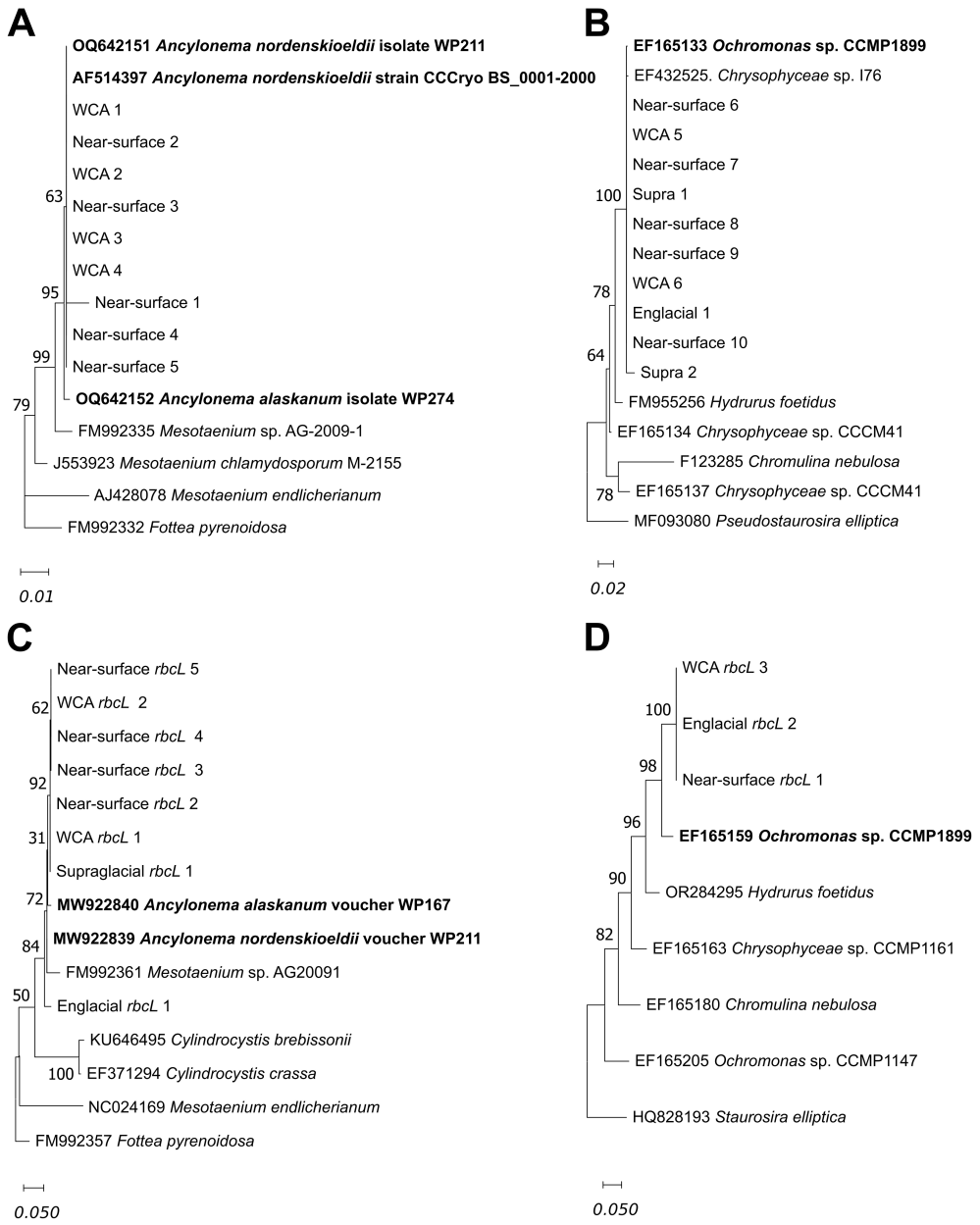


FIGURE 5
Phylogenetic analysis of partial 18S rRNA (1000 bp ranging from position 271 to 1270; *Fottea pyrenoidosa* FM992332) for Zygnematophyceae **(A)** and Chrysophyceae **(B)** and *rbcL* (387bp) gene sequences for Zygnematophyceae **(C)** and Chrysophyceae **(D)** by maximum likelihood using the Jukes-Cantor model with 100 bootstraps Branches corresponding to partitions reproduced in greater than 30% of bootstrap replicates are indicated. The phylogenetic trees for 18S rRNA and *rbcL* gene analysis were rooted using DNA sequences from *Fottea pyrenoidosa* for sequences belonging to Zygnematophyceae, and *Pseudostaurosira elliptica* for sequences belonging to Chrysophyceae. The scale bar represents the number of changes per nucleotide position. Names in bold are taxa previously reported from ice surfaces.

Genes encoding enzymes involved in carbon, energy, and nutrient metabolism provided insight on the major catabolic and anabolic pathways distributed among the communities (Supplementary Table S6). Samples containing genes annotated for photosystem I and II showed the potential for oxygenic photosynthetic pathways (GO:0015979). The presence of the Calvin–Bassham–Benson cycle was indicated by the presence of genes for ribulose bisphosphate carboxylase complex assembly (GO:0110102) and phosphofructokinase-1. Many sample types contained genes associated with lithotrophic metabolisms, including the oxidation of ferrous iron (GO:0019411) and reduced sulfur (GO:0019417; GO:0019418). Samples containing genes for carbohydrate metabolic processing (GO:0005975) coupled with respiration of O_2 and nitrate were observed, including aerobic respiration (GO:0009060), anaerobic respiration (GO:0009061), and denitrification [GO:0019333]. Metagenomes also encoded enzymes for fermentation via the Embden-Meyerhof (GO:0006096) and Entner-Doudoroff (GO:0009255) pathways. Genes involved in nitrogen fixation (GO:0009399) provided evidence for the potential for diazotrophy in the bacterial

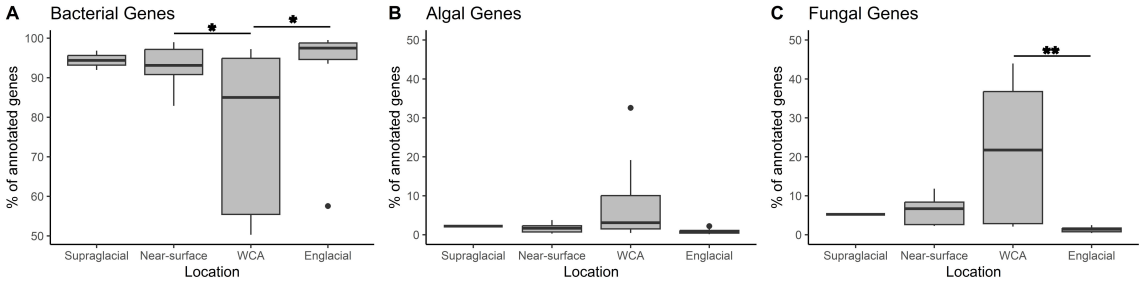


FIGURE 6
The percentage of genes assigned to bacterial **(A)**, algal **(B)**, and fungal **(C)** taxa for each sampling location. Asterisks show significant differences for *post hoc* analysis of ANOVA results using Tukey's Honest Significant Difference (*=<0.05, **=<0.01).

communities. Bacterial and fungal genes related to inorganic phosphorus mobilization were also widely distributed among samples, including quinoprotein glucose dehydrogenase A, exopolyphosphatase, inorganic pyrophosphatase, and alkaline phosphatase D (Supplementary Figure S5).

Genes involved in response to stresses associated with temperature and light were identified in the supraglacial and englacial assemblages. These included stress related adaptations such as extracellular polysaccharide biosynthetic processes (GO:0045226), response to oxidative (GO:0034599) and osmotic stress (GO:0071470), and carotenoid biosynthesis (GO:0016117; Supplementary Table S6). Additionally, genes involved in fungal sporulation and formation of a cellular spore (GO:0030435) were found, as well as genes known to be upregulated in mixotrophic algae (i.e., McKie-Krisberg et al., 2018) and involved in processes such as endocytosis (GO:0006897), phagocytosis (GO:0006909), and autophagy (GO:0006914; Supplementary Table S6).

Discussion

The interplay among physical, hydrological, and microbiological processes in the ablating surfaces of glaciers is recognized as an important research area due to their effects on altering ice properties, meltwater generation, supraglacial nutrient cycling, and the solute composition of runoff (Wadham et al., 2016; Musilova et al., 2017;

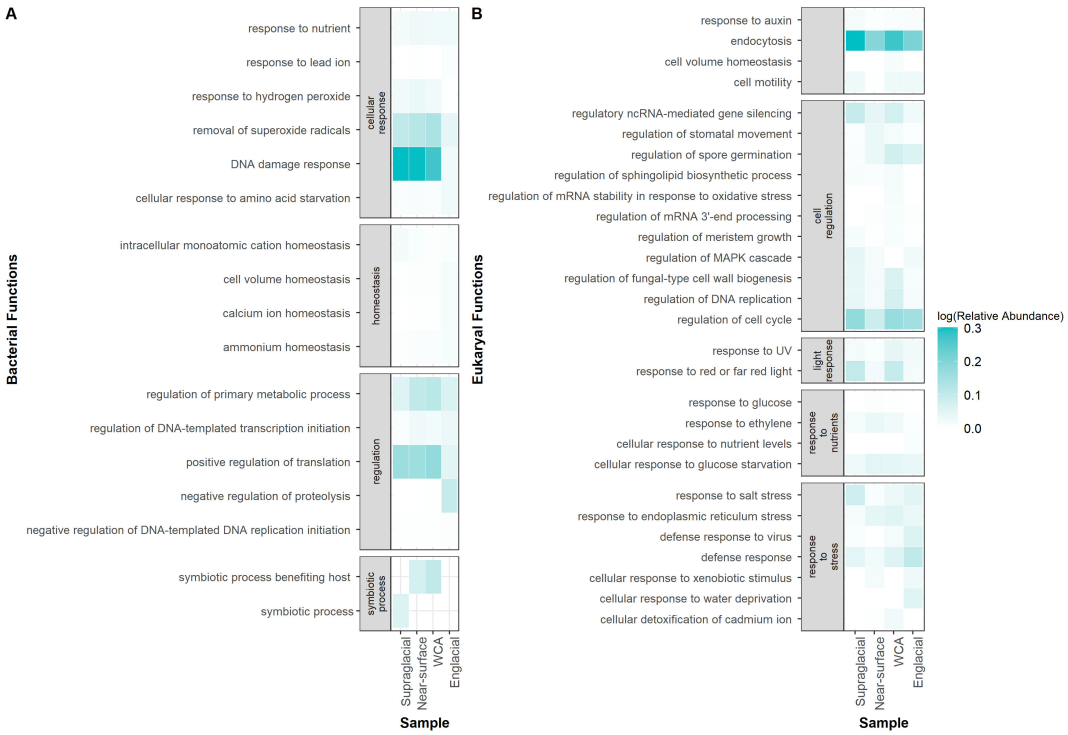


FIGURE 7
Comparison of GO functions related to stress adaptations significantly enriched or depleted in the WCA (Wald test, $p < 0.05$). Genes classified as bacterial **(A)** and eukaryotic **(B)** appear in separate panels. Relative abundance of genes was determined by mapping reads to genes, then normalizing by gene length.

Williamson et al., 2020). Recent studies of microbial communities inhabiting cryoconite holes, surficial ice, snow, meltwater lakes, and streams have highlighted the diversity of species and metabolisms in glacier and ice sheet microbiomes globally (e.g., Cameron et al., 2016; Anesio et al., 2017; Jaarsma et al., 2023). Bacterial communities in cryoconite holes and surface ice vary by geographical location, yet are largely comprised of taxa from the same phyla we have documented (i.e., Cyanobacteriota, Pseudomonadota, Actinomycetota, and Bacteroidota; Stibal et al., 2015; Cameron et al., 2016) in supraglacial features and englacial ice of a temperate Alaskan glacier (Figure 3A). Though few studies have examined microbial assemblages in the near-surface, water-saturated habitats of porous glacial ice, the unique distributions of bacterial and algal taxa reported in the weathering crust environment (Christner et al., 2018; Rassner et al., 2024) support a proposition for communities with distinct functions and biogeochemical roles in supraglacial meta-ecosystems. Hence, the aim of our study was to further understanding of microorganisms in the WCA by comparing their gene contents to communities associated with proximal features of the supraglacial ecosystem.

Low biomass and yields of extractable DNA (i.e., 20 samples with < 1 ng of DNA available for sequencing) produced a metagenomic dataset that assembled into few high-quality bins. While this prevented a broad comparison of metagenome-assembled genomes from the populations, the data available nevertheless allowed us to identify orthologous genes assembled from paired-end sequencing of samples from the WCA to infer molecular function. In comparison to the number of bacterial genes identified (~1,200,000), fewer eukaryotic genes were identified in the samples (~70,000; Supplementary Table S2). The 16S rRNA gene sequences recovered from the metagenomes indicate the bacterial communities of the WCA were distinct from samples in the englacial environment (Figures 2A, 3A), whereas analysis of 18S rRNA gene sequences showed that the eukaryotic community composition was similar amongst all sample types (Figures 2B, 3B). Despite the compositional similarities in the eukaryotic communities, we found higher numbers of fungal genes in the WCA in comparison to the other samples analyzed (Figure 6), implying that fungi are an important component of the WCA biodiversity and may be playing crucial roles in supraglacial ecosystem processes. Given that subtle differences in environmental conditions and resource availability have evolved ecologically specialized microbial variants over environmental gradients (e.g., Ward et al., 2006; Chase et al., 2018), we analyzed gene content to seek evidence that the supraglacial environment harbors populations with distinct ecologies related to exposure and access to meltwater. The parameters expected to affect distributions of microbes in the WCA are related to hydrological (i.e., retention of water, microbes, and nutrients; flowthrough velocities 4- to 6-orders of magnitude lower than surface streams; Irvine-Fynn et al., 2021) and physical properties (attenuated flux of solar radiation; Christner et al., 2018) of the porous ice. Considering differences in beta diversity among the bacterial communities (Figure 2A), it was surprising to find few differences in gene content (Figure 7A), suggesting that the supraglacial bacterial communities may have been functionally homogenous. This contrasted with the abundance of functionally annotated eukaryotic genes, which was significantly different among the WCA assemblages (Figure 7B). Eukaryotic communities of the WCA were functionally distinct from near-surface ice, as well as englacial ice, and their metagenomes are enriched in genes related to light stress, spore germination, and endocytosis (Supplementary Table S7B); properties that may assist WCA microbes in surviving the stresses associated with the low nutrient and temperature conditions on glacier surfaces.

The range of phototrophic and heterotrophic metabolisms we documented in the Matanuska Glacier's supraglacial environment are similar to those observed previously on glacial surfaces (Margesin and Collins, 2019; Supplementary Table S6). Supraglacial communities are exposed to high fluxes of incident solar radiation in the summer, complete darkness and low temperatures during winter, recurring freezing and thawing events, and oligotrophic conditions. Accordingly, genes with functions related to light-induced stress were enriched in the WCA, including those involved in the response to UV radiation exposure, red and far-red light, and oxidative stress (Figures 7A, B). Abundance of genes involved in endospore and fungal sporulation in the samples (Supplementary Table S6) implies that some species may form these environmentally durable structures to survive unfavorable conditions and disperse between supraglacial features. Cellular survival during freezing and thawing is enhanced by controlling water efflux and osmolyte accumulation (Doyle et al., 2011), and genes related to cell volume homeostasis, response to salt stress, and cellular response to water deprivation were enriched in the WCA samples (Figures 7A, B). Though supraglacial ecosystems typically contain low amounts of nutrients (Holland et al., 2019), a sufficient flux of PAR through the ice surface (Christner et al., 2018) coupled with the presence of phototrophic algae (Figure 5) is likely to be a key source of organic carbon in the WCA.

Understanding the interactions between photosynthetic primary producers and heterotrophic microbes is important for deciphering carbon cycling and nutrient exchange in the supraglacial environment. On ice surfaces, measured rates of net primary productivity exceed bacterial production (Williamson et al., 2018), implying that algal photosynthate is an essential organic carbon source for heterotrophic bacteria and fungi. We used species abundance correlations based on protein-encoding and rRNA genes to infer potential interactions between taxa identified in the metagenomes. The relative abundance of algal classes Zygnematophyceae and Chrysophyceae in the WCA were positively correlated to select bacterial taxa (i.e. Actinomycetota, Armatimonadota, Bacteroidota, Cyanobacteriota, and Pseudomonadota) (Figure 4; Table 2), similar to previous observations on ice surfaces (Nicholes et al., 2019). This result is supported by co-cultivation experiments that have demonstrated cross feeding between snow algae and bacteria (Krug et al., 2020) and that have indicated the algal community is also highly codependent on the bacteria they coexist with. For instance, most algal species lack the genes to produce B vitamins and must rely on the supply of this essential cofactor from bacteria (Croft et al., 2005; Helliwell et al., 2011). Members of the Bacteroidota phylum have previously been shown to be associated with Zygnematophyceae taxa and express genes for the production of thiamine, cobalamin, and biotin in the presence of these algae (Krohn-Molt et al., 2017).

Given that bacterial genes for cobalamin (B_{12}) biosynthesis are found across all sample types (Supplementary Table S6), mutualistic relationships between supraglacial algae and bacteria based on exchange of organic carbon and B_{12} is a plausible ecological scenario. Importantly, if these bacteria are involved in interactions that provide a nutrient essential for algal growth, then they would have key roles in the occurrence of algal blooms on ice surfaces (Yallop et al., 2012; Williamson et al., 2020) and dynamics of supraglacial carbon cycling.

Phosphorus availability has been shown to limit primary productivity in algal blooms on the Greenland Ice Sheet (McCutcheon et al., 2021), calling attention to the sources of this key nutrient in glacial meltwaters. The mineralization of organic phosphorus compounds by fungi (e.g., *Penicillium* (Eurotiomycetes); Figure 3B, Qiao et al., 2019) may represent an important source of inorganic phosphorus to the supraglacial environment. The presence of fungal genes encoding inorganic pyrophosphatase in the metagenomes, which hydrolyzes polyphosphate compounds (Supplementary Figure S5; Nannipieri et al., 2010), and bacteria known to be efficient at solubilizing phosphate from mineral sources (e.g., *Massilia*; Zheng et al., 2017) supports this contention. Positive correlations were noted between several algal and fungal taxa (OTUs in the Microbotryomycetes and unclassified Basidiomycota; Table 2), indicating that fungal taxa may support the growth of algae. Positive correlations between Alphaproteobacteria and several algal taxa (Table 2) also imply that bacteria may be important for providing phosphorus to algae, as several alphaproteobacterial phosphorus cycling genes were identified (Supplementary Figure S5). We also found positive associations between Zygnematophyceae and Cytophagia (Table 2), which is a group of bacteria known to be involved in phosphorus solubilization in soils and predicted to increase phosphate bioavailability (Wu et al., 2022). Fungi isolated from glacial ice have been shown to produce antibacterial secondary metabolites (de Menezes et al., 2020), and many of our samples contained genes encoding pathways for production of mycophenolic acid, polyketides, and alkaloids belonging to Eurotiomycetes and Dothideomycetes, which are abundant in WCA samples (Figure 3C; Supplementary Table S6). If derivatives of these compounds are produced *in situ*, these secondary metabolites could be involved in antagonistic interactions between supraglacial fungal and bacterial communities that affect community composition and carbon cycling.

Glacial ice surfaces are well known to be colonized by species of green algae from the class Zygnematophyceae that include *Cyclindrocystis brebbisonni*, *A. nordenskioeldii*, and *A. alaskanum* (Yallop et al., 2012; Procházková et al., 2021). *Ancylonema* has been documented on the surfaces of glaciers across the northern hemisphere (Procházková et al., 2021; Takeuchi et al., 2003, 2019; Yallop et al., 2012; Yoshimura et al., 1997), with fewer observations in the southern hemisphere (e.g., maritime Antarctica; Komárek and Komárek, 2001; Ling and Seppelt, 1990). However, since many of these studies used classical morphological approaches as the basis for identification, it is not possible to evaluate their relatedness to the taxa documented in this study. Given that blooms of darkly pigmented *Ancylonema* have been shown to reduce ice albedo (Yallop et al., 2012; Williamson et al., 2020), high abundances of these algae in water-saturated ice on the surface may accelerate melting as well (Cook et al., 2020). The near-surface ice and WCA samples contained high abundances of taxa with 100% identity to the 18S rRNA gene of isolate *A. nordenskioeldii* WP211 from Svalbard/Switzerland (Procházková et al., 2024; OQ642151) and > 99% identity to the *rbcL* genes of *A. nordenskioeldii* WP211 and *A. alaskanum* WP167 (MW922839 and MW922840; Procházková et al., 2021; Figures 5A, C) from Austria and Switzerland, respectively. These results indicate there is a high degree of relatedness in green algae separated by large geographical distances, suggesting the possibility of dispersal via atmospheric transport, similar to distribution patterns in certain red-snow algae (Segawa et al., 2018). Although aerial dissemination has been documented for many microorganisms, little is known about the environmental barriers limiting dispersal of ice algae on regional to global scales. Long distance dissemination via air transport and fall-out with dry or wet deposition is possible and supported by the high similarities of our data to phylotypes from distant locations. Future efforts aimed at recovering metagenomic assembled genomes or single amplified genomes from *Ancylonema* populations in supraglacial environments from different geographical locations represent fertile territory for testing the concept of ecological cosmopolitanism in these glacier algae. Due to their role in both carbon cycling and meltwater production in surface ice (Cook et al., 2020; Musilova et al., 2017), understanding the scale of ice algae dispersal is critical for predicting changes to glacial surfaces in the future.

Algae phylogenetically related to an *Ochromonas* species (CCMP1899) isolated from Antarctic sea ice (18S rRNA and *rbcL* genes with > 99.9 and 93% identity, respectively) were also identified (Figures 5B, D). While members from this group have been less commonly reported on glacier and ice sheet surfaces than *Ancylonema*, *Ochromonas* taxa are known to be associated with algal blooms in alpine snowpacks (Tanabe et al., 2011) and possess a mixotrophic nutrition based on nutrient and light availability (Lie et al., 2018). Considering their metabolic flexibility, the presence of taxa related to *Ochromonas* in the WCA and near-surface ice samples, and the identification of genes involved in autophagy and endocytosis (Figure 7; Supplementary Table S6), these microbes may be able to use sunlight or bacterial grazing for energy and nutrition. In the shaded, oligotrophic conditions characteristic of the WCA, the efficiency of a mixotrophic metabolism (e.g., Mitra et al., 2014) could provide golden algae with an advantageous physiological strategy when nutrient and energy sources are limited. In addition to enhancing growth rates and biomass production, mixotrophic microalgae may also be more resilient to the seasonal environmental stresses associated with the ice habitat. While the role of heterotrophic metabolism in algal adaptation to extended period of darkness remains a subject of debate (e.g., Lyon and Mock, 2014), shifts to mixotrophy have been documented in later winter sea ice communities (Bachy et al., 2011). Given the extended persistence of meltwater in the ice (~7.5 months per year; Christner et al., 2018), such a physiological adaptation could extend supraglacial carbon

cycling beyond the melt season into periods where decreased solar irradiance and temperature prevent photobiogeochemical processes from occurring on the uppermost ice surface.

As few studies have examined the mycobiomes of glaciers and ice sheets, there is an incomplete understanding of the diversity, ecology, and biogeochemical roles of the fungal communities ubiquitous in supraglacial environments. The fungal groups we identified (Figures 3B, C) have also been shown to be associated with surface ice (Microbotryomycetes, Basidiomycota) and snow (Dothideomycetes, Ascomycota and Agaricomycetes, Basidiomycota; Edwards et al., 2013; Brown et al., 2015; Duo Saito et al., 2018; Perini et al., 2019). Several studies have also documented a parasitic relationship between Chytridiomycota and glacier ice algae (Perini et al., 2022; Nakanishi et al., 2023; Kobayashi et al., 2023). The abundance of taxa related to the parasitic Chytridriomycota that infect snow and ice algae (Perini et al., 2022; Nakanishi et al., 2023; Kobayashi et al., 2023) was < 1.6% of total 18S rRNA gene reads (Figure 3B). These are distributions similar to those observed in other supraglacial environments (0.01 to 3.4% in surface ice; Perini et al., 2019) where their interactions with glacier algae are inferred to regulate population sizes and affect supraglacial food webs. Microbotryomycetes OTUs and the major algae taxa are positively correlated, and though their specific interactions with microalgae have not been reported, numerous species are known to be phytoparasites of plants (Kemler et al., 2020). In aquatic ecosystems, fungi typically outperform bacteria in decomposition (Gessner et al., 2007) and can often break down complex organic matter not metabolized by bacteria, such as lignin (Sun et al., 2020; Gu et al., 2022). Fungi likely play an important role as decomposers in supraglacial systems. This is supported by observations of high abundances of Microbotromycetes (Figure 3C), which are known to be saprotrophic (Oberwinkler, 2017), and fungal genes involved in lignin and pectin catabolism (Supplementary Table S6). Fungal interactions that control algal production could impact the biological darkening of ice, but it is also possible that some fungi could have a more direct role in melt water production through the production of pigments that can darken ice surfaces. The fungal metagenomes contain genes for the biosynthesis of melanin (GO:0042438) and carotenoids (GO:0016117; Supplementary Table S6) from classes that were highly abundant in supraglacial samples, including Dothideomycetes and Sordariomycetes (Figure 3C). These pigments protect cells from UV radiation and lead to increased heat absorption (Avalos and Carmen Limón, 2015; Cordero et al., 2018; Eisenman and Casadevall, 2012). Ice algal pigments have been studied extensively for their role in reducing ice albedo (Remias et al., 2012; Yallop et al., 2012; Williamson et al., 2020), but little is known on the role of fungal or bacterial pigments in ice darkening and meltwater generation in supraglacial environments.

Our analysis of data from two melt seasons verified the hypothesis that the physical and hydrogeochemical conditions in the Matanuska Glacier's WCA selects for a distinct microbial community. The WCA communities contained a distinct bacterial composition and threefold more fungal genes than those in proximal streams and glacial ice. Although the eukaryotic communities were taxonomically more uniform across the portion of the supraglacial meta-ecosystem we investigated, the gene content of fungi and algae in the WCA are functionally specialized and enriched with genes involved in endocytosis, light stress, and osmotic stress. Additionally, algal abundances in the WCA positively correlated with specific bacterial and fungal taxa that they may interact with to exchange nutrients and carbon sources. The roles of fungal communities in supraglacial biogeochemical processes have been largely ignored to date but are likely to have important implications to nutrient mineralization, production of secondary metabolites, and possibly, contribute to biological darkening of ice. Future studies that trace carbon flow, investigate species interactions, and examine gene expression within the WCA would be valuable for improving understanding of biogeochemical processing in supraglacial microbial communities. Such information would aid scientists with predicting changes in proglacial nutrient export as rates of global temperature and marginal ablation increase.

Data availability statement

The original contributions presented in the study are publicly available. This data can be found here: NCBI SRA, accession PRJNA430887.

Author contributions

QF: Conceptualization, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review & editing. CD: Investigation, Methodology, Writing – review & editing. BC: Conceptualization, Funding acquisition, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This research was supported by awards from the NASA ASTEP program (NNX11AJ89G) and NSF Office of Polar Programs (2000649). Partial support was also provided by the University of Florida through the Institute of Food and Agricultural Sciences and a Water Institute Graduate Fellowship.

Acknowledgments

The support and assistance provided by the following 2014 and 2015 VALKYRIE field team members was crucial to the success of this research: N. Bramall, E. Clark, P. Doran, C. Flesher, J. Harman, B. Hogan, H. Lavender, J. Moor, K. Myers, S. Neuhaus, E. Oliver, D. Rickel, D. Sampson, V. Siegel, W. Stone, S. Tulaczyk, and V. Yuan.

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.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frmbi.2024.1488744/full#supplementary-material

References

An, S. M., Choi, D. H., Lee, H., Lee, J. H., and Noh, J. H. (2018) Next-generation sequencing reveals the diversity of benthic diatoms in tidal flats. *Algae*. 33, 167–180.

Anesio, A. M., Hodson, A. J., Fritz, A., Psenner, R., and Sattler, B. (2009). High microbial activity on glaciers: Importance to the global carbon cycle. *Glob. Chang. Biol.* 15, 955–960. doi: 10.1111/j.1365-2486.2008.01758.x

Anesio, A. M., Lutz, S., Chrismas, N. A. M., and Benning, L. G. (2017). The microbiome of glaciers and ice sheets. *NPJ Biofilms. Microbiomes.* 3. doi: 10.1038/s41522-017-0019-0

Avalos, J., and Carmen Limón, M. (2015). Biological roles of fungal carotenoids. *Curr. Genet.* 61, 309–324. doi: 10.1007/s00294-014-0454-x

Bachy, C., López-García, P., Vereshchaka, A., and Moreira, D. (2011). Diversity and vertical distribution of microbial eukaryotes in the snow, sea ice and seawater near the North Pole at the end of the polar night. *Front. Microbiol.* 2. doi: 10.3389/fmicb.2011.00106

Bairoch, A., and Apweiler, R. (2000). The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. *Nucleic Acids Research*, 28(1), 45–48. doi: 10.1093/nar/28.1.45

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., et al. (2012). SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 19, 455–477. doi: 10.1089/cmb.2012.0021

Bokulich, N. A., Kaehler, B. D., Rideout, J. R., Dillon, M., Bolyen, E., Knight, R., et al. (2018). Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome* 6. doi: 10.1186/s40168-018-0470-z

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

Brodie, J., and Lewis, J. (2007). *Unravelling the algae: the past, present, and future of algal systematics* (Boca Raton: CRC).

Brown, S. P., Olson, B. J. S. C., and Jumpponen, A. (2015). Fungi and algae co-occur in snow: an issue of shared habitat or algal facilitation of heterotrophs? *Arct. Antarct. Alp. Res.* 47, 729–749. doi: 10.1657/AAAR0014-071

Buchfink, B., Reuter, K., and Drost, H. G. (2021). Sensitive protein alignments at tree-of-life scale using DIAMOND. *Nat. Methods* 18, 366–368. doi: 10.1038/s41592-021-01101-x

Bushnell, B., Rood, J., and Singer, E. (2017). BBMerge – Accurate paired shotgun read merging via overlap. *PloS One* 12. doi: 10.1371/journal.pone.0185056

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., et al. (2009). BLAST+: architecture and applications. *BMC Bioinf.* 10. doi: 10.1186/1471-2105-10-421

Cameron, K., Stibal, M., Zarsky, J. D., Gozderelier, E., Schostag, M., and Jacobsen, C. S. (2016). Supraglacial Bacterial Communities vary across the Greenland ice sheet. *FEMS Microbiol. Ecol.* 92, 1–11. doi: 10.1093/femsec/fiv164

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

Chase, A. B., Gomez-Lunar, Z., Lopez, A. E., Li, J., Allison, S. D., Martiny, A. C., et al. (2018). Emergence of soil bacterial ecotypes along a climate gradient. *Environ. Microbiol.* 20, 4112–4126. doi: 10.1111/1462-2920.14405

Christner, B. C., Lavender, H. F., Davis, C. L., Oliver, E. E., Neuhaus, S. U., Myers, K. F., et al. (2018). Microbial processes in the weathering crust aquifer of a temperate glacier. *Cryosphere* 12, 3653–3669. doi: 10.5194/tc-12-3653-2018

Christner, B. C., Priscu, J. C., Achberger, A. M., Barbante, C., Carter, S. P., Christianson, K., et al. (2014). A microbial ecosystem beneath the West Antarctic ice sheet. *Nature* 512, 310–313. doi: 10.1038/nature13667

Constable, A. J., Harper, S., Dawson, J., Holsman, K., Mustonen, T., Piepenburg, D., et al. (2023). "Polar regions," in *Climate Change 2022 – Impacts, Adaptation and Vulnerability* (Cambridge University Press, Cambridge, UK and New York, NY, USA), 2319–2368. doi: 10.1017/9781009325844.023

Cook, J., Edwards, A., Takeuchi, N., and Irvine-Fynn, T. (2016a). Cryoconite: The dark biological secret of the cryosphere. *Prog. Phys. Geogr.* 40, 66–111. doi: 10.1177/0309133315616574

Cook, J. M., Hodson, A. J., Anesio, A. M., Hanna, E., Yallop, M., Stibal, M., et al. (2012). An improved estimate of microbially mediated carbon fluxes from the Greenland ice sheet. *J. Glaciol.* 58, 1098–1108. doi: 10.3189/2012JoG12J001

Cook, J. M., Hodson, A. J., and Irvine-Fynn, T. D. L. (2016b). Supraglacial weathering crust dynamics inferred from cryoconite hole hydrology. *Hydrol. Process.* 30, 433–446. doi: 10.1002/hyp.10602

Cook, J. M., Tedstone, A. J., Williamson, C., Mccutcheon, J., Hodson, A. J., Dayal, A., et al. (2020). Glacier algae accelerate melt rates on the south-western Greenland Ice Sheet. *Cryosphere* 14, 309–330. doi: 10.5194/tc-14-309-2020

Cooper, M. G., Smith, L. C., Rennermalm, A. K., Mige, C., Pitcher, L. H., Ryan, J. C., et al. (2018). Meltwater storage in low-density near-surface bare ice in the Greenland ice sheet ablation zone. *Cryosphere* 12, 955–970. doi: 10.5194/tc-12-955-2018

Cordero, R. J. B., Robert, V., Cardinali, G., Arinze, E. S., Thon, S. M., and Casadevall, A. (2018). Impact of yeast pigmentation on heat capture and latitudinal distribution. *Curr. Biol.* 28, 2657–2664.e3. doi: 10.1016/j.cub.2018.06.034

Croft, M. T., Lawrence, A. D., Raux-Deery, E., Warren, M. J., and Smith, A. G. (2005). Algae acquire vitamin B12 through a symbiotic relationship with bacteria. *Nature* 438, 90–93. doi: 10.1038/nature04056

de Menezes, G. C. A., Porto, B. A., Amorim, S. S., Zani, C. L., de Almeida Alves, T. M., Junior, P. A. S., et al. (2020). Fungi in glacial ice of Antarctica: diversity, distribution and bioprospecting of bioactive compounds. *Extremophiles* 24, 367–376. doi: 10.1007/s00792-020-01161-5

Doyle, S., Dieser, M., Broemsen, E., and Christner, B. (2011). "General characteristics of cold-adapted microorganisms," in *Polar Microbiology* (ASM Press, Washington, DC, USA), 101–125. doi: 10.1128/9781555817183.ch5

Duo Saito, R. A., Connell, L., Rodriguez, R., Redman, R., Libkind, D., and de Garcia, V. (2018). Metabarcoding analysis of the fungal biodiversity associated with Castaño Overa Glacier – Mount Tronador, Patagonia, Argentina. *Fungal Ecol.* 36, 8–16. doi: 10.1016/j.funeco.2018.07.006

Edwards, A., Douglas, B., Anesio, A. M., Rassner, S. M., Irvine-Fynn, T. D. L., Sattler, B., et al. (2013). A distinctive fungal community inhabiting cryoconite holes on glaciers in Svalbard. *Fungal Ecol.* 6, 168–176. doi: 10.1016/j.funeco.2012.11.001

Eisenman, H. C., and Casadevall, A. (2012). Synthesis and assembly of fungal melanin. *Appl. Microbiol. Biotechnol.* 93, 931–940. doi: 10.1007/s00253-011-3777-2

Fiołka, M. J., Takeuchi, N., Sofińska-Chmiel, W., Wójcik-Mieszawska, S., Irvine-Fynn, T., and Edwards, A. (2021). Morphological and spectroscopic analysis of snow and glacier algae and their parasitic fungi on different glaciers of Svalbard. *Sci. Rep.* 11. doi: 10.1038/s41598-021-01211-8

Foreman, C. M., Cory, R. M., Morris, C. E., Sanclements, M. D., Smith, H. J., Lisle, J. T., et al. (2013). Microbial growth under humic-free conditions in a supraglacial stream system on the Cotton Glacier, Antarctica. *Environ. Res. Lett.* 8. doi: 10.1088/1748-9326/8/3/035022

Gessner, M. O., Gulis, V., Kuehn, K. A., Chauvet, E., and Suberkropp, K. (2007). "Fungal decomposers of plant litter in aquatic systems, " in *Environmental and microbial relationships* (Heidelberg: Springer), 301.

Gu, D., Xiang, X., Wu, Y., Zeng, J., and Lin, X. (2022). Synergy between fungi and bacteria promotes polycyclic aromatic hydrocarbon cometabolism in lignin-amended soil. *J. Hazard. Mater.* 425. doi: 10.1016/j.jhazmat.2021.127958

Heckmann, T., Mccoll, S., and Morche, D. (2016). Retreating ice: Research in pro-glacial areas matters. *Earth Surf. Process. Landf.* 41, 271–276. doi: 10.1002/esp.3858

Helliwell, K. E., Wheeler, G. L., Leptos, K. C., Goldstein, R. E., and Smith, A. G. (2011). Insights into the evolution of vitamin B 12 auxotrophy from sequenced algal genomes. *Mol. Biol. Evol.* 28, 2921–2933. doi: 10.1093/molbev/msr124

Hodson, A., Cameron, K., Bøggild, C., Irvine-Fynn, T., Langford, H., Pearce, D., et al. (2010). The structure, biological activity and biogeochemistry of cryoconite aggregates upon an Arctic valley glacier: Longyearbreen, Svalbard. *J. Glaciol.* 56, 349–362. Available at: https://www.cambridge.org/core.

Hodson, A., Paterson, H., Westwood, K., Cameron, K., and Laybourn-Parry, J. (2013). A blue-ice ecosystem on the margins of the East Antarctic ice sheet. *J. Glaciol.* 59, 255–268. doi: 10.3189/2013JoG12J052

Hoffman, M. J., Fountain, A. G., and Liston, G. E. (2014). Near-surface internal melting: A substantial mass loss on Antarctic Dry Valley glaciers. *J. Glaciol.* 60, 361–374. doi: 10.3189/2014JoG13J095

Hoham, R. W., and Duval, B. (2001). *Microbial Ecology of Snow and Freshwater Ice with Emphasis on Snow Algae*. (Cambridge: Cambridge University Press).

Holland, A. T., Williamson, C. J., Sgouridis, F., Tedstone, A. J., McCutcheon, J., Cook, J. M., et al. (2019). Dissolved organic nutrients dominate melting surface ice of the Dark Zone (Greenland Ice Sheet). *Biogeosciences* 16, 3283–3296. doi: 10.5194/bg-16-3283-2019

Huson, D. H., Beier, S., Flade, I., Górska, A., El-Hadidi, M., Mitra, S., et al. (2016). MEGAN community edition - interactive exploration and analysis of large-scale microbiome sequencing data. *PloS Comput. Biol.* 12. doi: 10.1371/journal.pcbi.1004957

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. doi: 10.1186/1471-2105-11-119

Irvine-Fynn, T. D. L., and Edwards, A. (2014). A frozen asset: The potential of flow cytometry in constraining the glacial biome. *Cytomet. Part A.* 85, 3–7. doi: 10.1002/cyto.a.22411

Irvine-Fynn, T. D. L., Edwards, A., Newton, S., Langford, H., Rassner, S. M., Telling, J., et al. (2012). Microbial cell budgets of an Arctic glacier surface quantified using flow cytometry. *Environ. Microbiol.* 14, 2998–3012. doi: 10.1111/j.1462-2920.2012.02876.x

Irvine-Fynn, T. D. L., Edwards, A., Stevens, I. T., Mitchell, A. C., Bunting, P., Box, J. E., et al. (2021). Storage and export of microbial biomass across the western Greenland Ice Sheet. *Nat. Commun.* 12, 3960. doi: 10.1038/s41467-021-24040-9

Jaarsma, A. H., Sipes, K., Zervas, A., Jiménez, F. C., Ellegaard-Jensen, L., Thøgersen, M. S., et al. (2023). Exploring microbial diversity in Greenland Ice Sheet supraglacial habitats through culturing-dependent and -independent approaches. *FEMS Microbiol. Ecol*, 99. doi: 10.1093/femsec/fiad119

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* 2019. doi: 10.7717/peerj.7359

Karlstrom, L., Zok, A., and Manga, M. (2014). Near-surface permeability in a supraglacial drainage basin on the Llewellyn Glacier, Juneau Icefield, British Columbia. *Cryosphere* 8, 537–546. doi: 10.5194/tc-8-537-2014

Kemler, M., Denchev, T. T., Denchev, C. M., Begerow, D., Piątek, M., and Lutz, M. (2020). Host preference and sorus location correlate with parasite phylogeny in the smut fungal genus Microbotryum (Basidiomycota, Microbotryales). *Mycol. Prog.* 19, 481–493. doi: 10.1007/s11557-020-01571-x

Kobayashi, K., Takeuchi, N., and Kagami, M. (2023). High prevalence of parasitic chytrids infection of glacier algae in cryoconite holes in Alaska. *Sci. Rep.* 13, 3973. doi: 10.1038/s41598-023-30721-w

Komárek, O., and Komárek, J. (2001). Contribution to the taxonomy and ecology of green cryosestic algae in the summer season 1995-96 at Kang George lsland, S. Shetland lslands. *Algae. Extreme. Environ.* 123, 121–140.

Krohn-Molt, I., Alawi, M., Förstner, K. U., Wiegandt, A., Burkhardt, L., Indenbirken, D., et al. (2017). Insights into Microalga and bacteria interactions of selected phycosphere biofilms using metagenomic, transcriptomic, and proteomic approaches. *Front. Microbiol.* 8. doi: 10.3389/fmicb.2017.01941

Krug, L., Erlacher, A., Markut, K., Berg, G., and Cernava, T. (2020). The microbiome of alpine snow algae shows a specific inter-kingdom connectivity and algae-bacteria interactions with supportive capacities. *ISME. J.* 14, 2197–2210. doi: 10.1038/s41396-020-0677-4

Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* 35, 1547–1549. doi: 10.1093/molbev/msy096

Kuraku, S., Zmasek, C. M., Nishimura, O., and Katoh, K. (2013). aLeaves facilitates on-demand exploration of metazoan gene family trees on MAFFT sequence alignment server with enhanced interactivity. *Nucleic Acids Res.* 41, 22–28. doi: 10.1093/nar/gkt389

Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359. doi: 10.1038/nmeth.1923

Levy Karin, E., Mirdita, M., and Söding, J. (2020). MetaEuk-sensitive, high-throughput gene discovery, and annotation for large-scale eukaryotic metagenomics. *Microbiome* 8. doi: 10.1186/s40168-020-00808-x

Leya, T. (2004). *Field studies and genetic investigations of the cryophilic snow algae of northwest Spitzbergen* (Berlin: Humboldt-Universitaet zu Berlin).

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352

Lie, A. A. Y., Liu, Z., 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. doi: 10.1371/journal.pone.0192439

Ling, H. U., and Seppelt, R. D. (1990). Snow algae of the Windmill Islands, continental Antarctica. Mesotaenium berggrenii (Zygnematales, Chlorophyta) the alga of grey snow. *Antarct. Sci.* 2, 143–148. doi: 10.1017/S0954102090000189

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. doi: 10.1186/s13059-014-0550-8

Lyon, B. R., and Mock, T. (2014). Polar microalgae: New approaches towards understanding adaptations to an extreme and changing environment. *Biol. (Basel).* 3, 56–80. doi: 10.3390/biology3010056

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

McCutcheon, J., Lutz, S., Williamson, C., Cook, J. M., Tedstone, A. J., Vanderstraeten, A., et al. (2021). Mineral phosphorus drives glacier algal blooms on the Greenland Ice Sheet. *Nat. Commun.* 12, 570. doi: 10.1038/s41467-020-20627-w

McKie-Krisberg, Z. M., Sanders, R. W., and Gast, R. J. (2018). Evaluation of mixotrophy-associated gene expression in two species of polar marine algae. *Front. Mar. Sci.* 5. 10.3389/fmars.2018.00273

Mitra, A., Flynn, K. J., Burkholder, J. M., Berge, T., Calbet, A., Raven, J. A., et al. (2014). The role of mixotrophic protists in the biological carbon pump. *Biogeosciences* 11, 995–1005. doi: 10.5194/bg-11-995-2014

Musilova, M., Tranter, M., Wadham, J., Telling, J., Tedstone, A., and Anesio, A. M (2017). Microbially driven export of labile organic carbon from the Greenland ice sheet. *Nat. Geosci.* 10, 360–365. doi: 10.1038/ngeo2920

Nakanishi, H., Seto, K., Takeuchi, N., and Kagami, M. (2023). Novel parasitic chytrids infecting snow algae in an alpine snow ecosystem in Japan. *Front. Microbiol.* 14. doi: 10.3389/fmicb.2023.1201230

Nannipieri, P., Giagnoni, L., Landi, L., and Renella, G. (2010). "Role of phosphatase enzymes in soil," in *Phosphorus in Action*. Eds. E. Bünemann, A. Oberson and E. Frossard (Springer Berlin Heidelberg, Berlin, Heidelberg), 215–243. doi: 10.1007/978-3-642-15271-9

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. doi: 10.3389/fmicb.2019.01366

Oberwinkler, F. (2017). Yeasts in pucciniomycotina. *Mycol. Prog.* 16, 831–856. doi: 10.1007/s11557-017-1327-8

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

Perini, L., Gostinčar, C., Anesio, A. M., Williamson, C., Tranter, M., and Gunde-Cimerman, N. (2019). Darkening of the Greenland ice sheet: Fungal abundance and diversity are associated with algal bloom. *Front. Microbiol.* 10. doi: 10.3389/fmicb.2019.00557

Perini, L., Gostinčar, C., Likar, M., Frisvad, J. C., Kostanjšek, R., Nicholes, M., et al. (2022). Interactions of fungi and algae from the Greenland ice sheet. *Microb. Ecol.* 86, 282–296. doi: 10.1007/s00248-022-02033-5

Procházková, L., Remias, D., Nedbalová, L., and Raymond, J. A. (2024). A DUF3494 ice-binding protein with a root cap domain in a streptophyte glacier ice alga. *Front. Plant Sci.* 14. doi: 10.3389/fpls.2023.1306511

Procházková, L., Řezanka, T., Nedbalová, L., and Remias, D. (2021). Unicellular versus filamentous: The glacial alga ancylonema alaskana comb. et stat. nov. and its ecophysiological relatedness to ancylonema nordenskioeldii (zygnematophyceae, streptophyta). *Microorganisms* 9. doi: 10.3390/microorganisms9051103

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

Qiao, H., Sun, X. R., Wu, X. Q., Li, G. E., Wang, Z., and Li, D. W. (2019). The phosphate-solubilizing ability of Penicillium guanacastense and its effects on the growth of Pinus massoniana in phosphate-limiting conditions. *Biol. Open* 8. doi: 10.1242/bio.046797

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. (2021). *R: A language and environment for statistical computing*. (Vienna, Austria: R Foundation for Statistical Computing). https://www.R-project.org/

Rassner, S. M. E., Cook, J. M., Mitchell, A. C., Stevens, I. T., Irvine-Fynn, T. D. L., Hodson, A. J., et al. (2024). The distinctive weathering crust habitat of a High Arctic glacier comprises discrete microbial micro-habitats. *Environ. Microbiol.* 26. doi: 10.1111/1462-2920.16617

Remias, D., Schwaiger, S., Aigner, S., Leya, T., Stuppner, H., and Lütz, C. (2012). Characterization of an UV- and VIS-absorbing, purpurogallin-derived secondary

pigment new to algae and highly abundant in Mesotaenium berggrenii (Zygnematophyceae, Chlorophyta), an extremophyte living on glaciers. *FEMS Microbiol. Ecol.* 79, 638–648. doi: 10.1111/j.1574-6941.2011.01245.x

Scott, D., Hood, E., and Nassry, M. (2010). In-stream uptake and retention of C, N and P in a supraglacial stream. *Ann. Glaciol.* 51, 80–86. doi: 10.3189/172756411795932065

Segawa, T., Matsuzaki, R., Takeuchi, N., Akiyoshi, A., Navarro, F., Sugiyama, S., et al. (2018). Bipolar dispersal of red-snow algae. *Nat. Commun.* 9. doi: 10.1038/s41467-018-05521-w

Stevens, I. T., Irvine-Fynn, T. D. L., Porter, P. R., Cook, J. M., Edwards, A., Smart, M., et al. (2018). Near-surface hydraulic conductivity of northern hemisphere glaciers. *Hydrol. Process.* 32, 850–865. doi: 10.1002/hyp.11439

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

Sun, Y., Liu, L., Zeng, J., Wu, Y., and Lin, X. (2020). Enhanced cometabolism of benzo(a)anthracene by the lignin monomer vanillate is related to structural and functional responses of the soil microbiome. *Soil Biol. Biochem.* 149. doi: 10.1016/j.soilbio.2020.107908

Takeuchi, N., Kohshima, S., and Segawa, T. (2003). Effect of cryoconite and snow algal communities on surface albedo on maritime glaciers in south Alaska. *Bull. Glaciol. Res.* 20, 21–27. Available at: https://www.researchgate.net/publication/267631131.

Takeuchi, N., Tanaka, S., Konno, Y., Irvine-Fynn, T. D. L., Rassner, S. M. E., and Edwards, A. (2019). Variations in phototroph communities on the ablating bare-ice surface of glaciers on brøggerhalvøya, svalbard. *Front. Earth Sci. (Lausanne).* 7. doi: 10.3389/feart.2019.00004

Tanabe, Y., Shitara, T., Kashino, Y., Hara, Y., and Kudoh, S. (2011). Utilizing the effective xanthophyll cycle for blooming of ochromonas smithii and O. itoi (chrysophyceae) on the snow surface. *PloS One* 6. doi: 10.1371/journal.pone.0014690

Tedstone, A. J., Cook, J. M., Williamson, C. J., Hofer, S., McCutcheon, J., Irvine-Fynn, T., et al. (2020). Algal growth and weathering crust state drive variability in western Greenland Ice Sheet ice albedo. *Cryosphere* 14, 521–538. doi: 10.5194/tc-14-521-2020

Varliero, G., Lebre, P. H., Frey, B., Fountain, A. G., Anesio, A. M., and Cowan, D. A. (2023). Glacial water: A dynamic microbial medium. *Microorganisms* 11. doi: 10.3390/microorganisms11051153

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

Wadham, J. L., Hawkings, J., Telling, J., Chandler, D., Alcock, J., O'donnell, E., et al. (2016). Sources, cycling and export of nitrogen on the Greenland Ice Sheet. *Biogeosciences* 13, 6339–6352. doi: 10.5194/bg-13-6339-2016

Ward, D. M., Bateson, M. M., Ferris, M. J., Kühl, M., Wieland, A., Koeppel, A., et al. (2006). Cyanobacterial ecotypes in the microbial mat community of Mushroom Spring (Yellowstone National Park, Wyoming) as species-like units linking microbial community composition, structure and function. *Philos. Trans. R. Soc. B.: Biol. Sci. (Royal. Society).* 361, 1997–2008. doi: 10.1098/rstb.2006.1919

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

Williamson, C. J., Anesio, A. M., Cook, J., Tedstone, A., Poniecka, E., Holland, A., et al. (2018). Ice algal bloom development on the surface of the Greenland Ice Sheet. *FEMS Microbiol. Ecol.* 94.

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.* 117, 5694–5705. doi: 10.1073/pnas.1918412117

Wu, X., Rensing, C., Han, D., Xiao, K.-Q., Dai, Y., Tang, Z., et al. (2022). Genome-resolved metagenomics reveals distinct phosphorus acquisition strategies between soil microbiomes. *mSystems* 7. doi: 10.1128/msystems.01107-21

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

Yoshimura, Y., Kohshima, S., and Ohtani, S. (1997). A community of snow algae on a himalayan glacier: change of algal biomass and community structure with altitude. *Arctic. Alpine. Res.* 29, 126–137. doi: 10.1080/00040851.1997.12003222

Zawierucha, K., Kolicka, M., Takeuchi, N., and Kaczmarek, L. (2015). What animals can live in cryoconite holes? A faunal review. *J. Zool.* 295, 159–169. doi: 10.1111/jzo.12195

Zheng, B. X., Bi, Q. F., Hao, X. L., Zhou, G. W., and Yang, X. R. (2017). Massilia phosphatilytica sp. nov., a phosphate solubilizing bacteria isolated from a long-term fertilized soil. *Int. J. Syst. Evol. Microbiol.* 67, 2514–2519. doi: 10.1099/ijsem.0.001916

frontiers | Frontiers in Microbiomes

TYPE Original Research
PUBLISHED 20 January 2025
DOI 10.3389/frmbi.2024.1401831

OPEN ACCESS

EDITED BY
Gareth Trubl,
Lawrence Livermore National Laboratory (DOE), United States

REVIEWED BY
David Emerson,
Bigelow Laboratory For Ocean Sciences, United States
Elaine Luo,
University of North Carolina at Charlotte, United States

*CORRESPONDENCE
Matthew O. Schrenk
schrenkm@msu.edu

RECEIVED 16 March 2024
ACCEPTED 19 December 2024
PUBLISHED 20 January 2025

CITATION
Alian OM, Brazelton WJ, Aquino KA, Twing KI, Pendleton HL, Früh-Green G, Lang SQ and Schrenk MO (2025) Microbial community differentiation in vent chimneys of the Lost City Hydrothermal Field reflects habitat heterogeneity.
Front. Microbiomes 3:1401831.
doi: 10.3389/frmbi.2024.1401831

Microbial community differentiation in vent chimneys of the Lost City Hydrothermal Field reflects habitat heterogeneity

Osama M. Alian[1], William J. Brazelton[2], Karmina A. Aquino[3], Katrina I. Twing[4], H. Lizethe Pendleton[2], Gretchen Früh-Green[5], Susan Q. Lang[6] and Matthew O. Schrenk[1,7]*

[1]Department of Microbiology, Genetics, and Immunology, Michigan State University, East Lansing, MI, United States, [2]School of Biological Sciences, University of Utah, Salt Lake City, UT, United States, [3]Department of Science and Technology - Philippine Nuclear Research Institute, Quezon City, Philippines, [4]Department of Microbiology, Weber State University, Ogden, UT, United States, [5]Department of Earth Sciences, Eidgenössiche Technische Hochschule Zürich, Zürich, Switzerland, [6]Department of Geology and Geophysics, Woods Hole Oceanographic Institution, Woods Hole, MA, United States, [7]Department of Earth and Environmental Science, Michigan State University, East Lansing, MI, United States

Oceanic hydrothermal vent systems represent some of the oldest habitats on Earth and serve as analogs for extraterrestrial environments. The Lost City Hydrothermal Field (LCHF) near the Mid-Atlantic Ridge is one such environment, and its large chimneys are unique in hosting actively venting hydrothermal fluids that are primarily controlled by serpentinization reactions in the subseafloor. Microbial communities within LCHF have been studied for insights into their functional adaptations to the warm, alkaline, and dissolved inorganic carbon-limited environment. Metagenomic and mineralogical data collected during a recent expedition to Lost City were analyzed to delineate associations between microbial populations and physical, chemical and biological characteristics of the chimneys. Bacterial 16S rRNA gene sequences show a high degree of putative microdiversity within the relatively dominant genera Desulfotomaculum, Sulfurovum, Thiomicrorhabdus, and Serpentinicella, which represent a large core of the overall LCHF vent bacterial community. This microdiversity relates to the compositional fraction of aragonite, brucite, and calcite minerals within chimney samples rather than just the composition of nearby vent fluids. Although many species are found in both chimneys and venting fluids, the overall microbial community structures in chimney biofilms remain distinct from the hydrothermal fluids that flow through them. Shotgun metagenomic analyses reveal differences among genes predicted to be involved in carbon, methane, nitrogen and sulfur cycling with respect to their correlations to the abundances of specific minerals. These data hint at microenvironmental

complexity lost within standard bulk analyses. The findings of this study underscore the need to more closely examine microbe-mineral interactions in natural environments, critically informing not just population-level distributions, but also the functional underpinnings of these extremophile microbial communities.

KEYWORDS

hydrothermal vent, geomicrobiology, correlations, microbe-rock interactions, extremophile, Lost City Hydrothermal Field

Introduction

Deep-sea hydrothermal vents are some of the most dynamic environments in Earth's biosphere. These unique ecosystems form when hydrothermal fluids, produced from processes below the seafloor, mix and react with cold, oxic seawater (Corliss et al., 1979; Früh-Green et al., 2022). Specific niches at deep-sea vents contain some of the harshest conditions known to support life, with extreme temperatures (>120°C), pressures (> 50 MPa) and pH (from 3-10) (Ding et al., 2005; Ludwig et al., 2006). Vent chimneys represent distinct habitats within marine hydrothermal ecosystems, serving as conduits for fluid flow as well as platforms for sustainable chemical disequilibria between the ocean and subseafloor (Sander and Koschinsky, 2011; Resing et al., 2015). The inhabitants of vent ecosystems redefine our understanding of biological resilience and the limits of habitability in extreme settings. Microbial communities found here often have unique taxonomic diversity, containing organisms with physiologies adapted to the distinctive physical-chemical conditions of the vents (Nakagawa and Takai, 2008; Anderson et al., 2015; Jebbar et al., 2015; Meier et al., 2017). Within the porous walls of hydrothermal chimneys, steep thermal and chemical gradients similar to those found in the subseafloor develop over distances as short as a few centimeters (Kristall et al., 2006). Although the microbiological features of vent chimneys have been previously studied at larger scales, limited research has been conducted on the characteristics of small-scale micro-environments within the chimney walls and how they influence microbial community organization and functions (Wang et al., 2009). Prior studies delineate microbial distributions at coarse scales, highlighting differences between the inner and outer portions of chimneys or between different vent habitats (Harmsen et al., 1997; Takai et al., 2001; Schrenk et al., 2003; Scott et al., 2017). However, the nuanced distribution and abundance of these communities across varying local conditions, exemplified by different mineralogy, remain to be fully understood.

Sulfide-hosted hydrothermal systems, driven by active magmatism, are among the best studied extreme environments since their discovery in the 1970s. These systems are teeming with uniquely adapted microbial and macrofaunal communities, some of which endure the highest temperatures known to life (Prieur, 1997; Gadanho and Sampaio, 2005; Wirth, 2017). A mix of seawater-derived nutrients and oxidants, combined with metals and volatiles from magmatic processes, sustain high microbial metabolic activity within these vibrant ecosystems (Kato et al., 2012; Zhong et al., 2022). In contrast, low to moderate temperature hydrothermal systems influenced by serpentinization, a process whereby water reacts with ultramafic rocks of the upper mantle, can be metal-deficient with scant dissolved inorganic carbon and high pH from the water-rock reactions (Ludwig et al., 2006). The Lost City Hydrothermal Field (LCHF; 30°N, Mid-Atlantic Ridge) is an example of this type of system, and vents alkaline, warm fluids. The chimney structures at LCHF consist primarily of three minerals; aragonite ($CaCO_3$), brucite ($Mg(OH)_2$) and calcite ($CaCO_3$). These minerals form as the emanating serpentinization-influenced fluids react with seawater (Chavagnac et al., 2013; Tutolo et al., 2018; Aquino and Früh-Green, 2024a; Aquino and Früh-Green, 2024b). Studies of LCHF chimneys indicate major controls of vent fluid and seawater mixing on the resulting mineralogy. Brucite-calcite is the preferred mineral assemblage that precipitates from the vent-dominated solutions found in interior chimney structures, with limited seawater mixing (Aquino et al., 2023; Aquino and Früh-Green, 2024b). In contrast, chimney exteriors in a seawater-dominated environment are primarily aragonite. These differences are further highlighted by stable isotope signatures suggesting higher formation temperatures in the brucite-calcite regions and $^{87}Sr/^{86}Sr$ ratios in aragonite that are much closer to seawater than primary calcite. As venting decreases and the chimneys become extinct, structures develop much larger fractions of secondary calcite infilling than either aragonite or brucite (Ludwig et al., 2011). Taken together, these minerals can be used as proxies for environmental conditions within chimney structures where brucite represents the hottest and newest chimneys, aragonite appears with increasing seawater intrusion, and the largest relative calcite fractions indicate terminal stage vents.

The fragile nature of LCHF chimneys, sampling challenges associated with this remote site, and difficulty culturing representative microbial populations have limited our ability to investigate the microbiology of the chimneys. Prior work has shown that the chimneys host dense microbial populations, reaching 10^9 cells per gram, primarily as biofilms (Brazelton et al., 2011). Although the serpentinization-influenced hydrothermal fluids at LCHF provide key energy sources such as methane, hydrogen and formate, they are largely deficient in trace metals that are critical to

enzymatic functions (Brazelton et al., 2006; Ludwig et al., 2006; Brazelton et al., 2022). The presence of dense microbial biofilms within these chimneys hint at sufficient nutrient availability to sustain metabolism in either the fluids or the chimney minerals, or an adaptation to utilize alternate metabolic cofactors.

Metal limitations can greatly impact microbial physiology and the functional organization of microbial communities. Metal cofactors are required for many metabolic processes, including methane and nitrogen cycling, and most basic cell functions, such as electron transport. They are especially critical to key steps in anaerobic metabolism that under thermophilic conditions, can require nearly double (hundreds of ppm and ppb levels) the trace metal amounts for mesophilic conditions (Takashima et al., 2011). Lost City vent fluids and carbonates contain generally lower metal concentrations than seawater (Kelley et al., 2005; Ludwig et al., 2006; Früh-Green et al., 2022). When chimneys are analyzed in bulk, constituent trace elements are generally minimal or below instrument detection depending on the element of interest, with difficulties in detectability arising from mineral matrix effects (Ludwig et al., 2006). In other systems, metal concentrations can vary over time with changes in the underlying vent fluid chemistry, temperature and pH (Schijf et al., 2015; Zwicker et al., 2022). Marine microorganisms inhabit a wide range of metal-scarce environments and have evolved sophisticated adaptive strategies for metal uptake, or evolved to use alternative cofactors altogether (Vasudevan et al., 2002; Samuel et al., 2012; Boyd et al., 2015; Fadel et al., 2017; Mus et al., 2019). These communities may also engage in cooperative strategies, sharing resources and products in a complex biofilm community (Liang et al., 2017; Skennerton et al., 2017). Effective utilization of these trace metals likely depends on the existence of localized microenvironments and mixing regions at the mineral-microbe interface (McCollom and Shock, 1997; Bianchi et al., 2018).

This study examines the diversity and differentiation of microbial populations within the Lost City chimneys and how they relate to physical-chemical features and chimney evolution. We highlight how specific taxa and genes are associated with the key indicator minerals aragonite, brucite and calcite, which are used as proxies for localized conditions. Our findings shed light on the interplay between geochemistry and biochemistry within these structures, identifying small scale environmental differences that may impact microbial community structure within the Lost City chimneys and delineating some of the key metabolic pathways prevalent in those environments.

Materials and methods

Sample collection

Hydrothermal chimney and fluid samples were collected using the R/V Atlantis and ROV Jason II during expedition AT42-01 in September 2018. Hydrothermal fluids and seawater were sampled with the Hydrothermal Organic Geochemistry sampler deployed on ROV Jason II or with Niskin bottles mounted to a CTD rosette, as previously described (Lang and Benitez-Nelson, 2021; Lang et al., 2021; Brazelton et al., 2022). Water samples were also collected with Niskin bottles mounted to a CTD rosette and Niskin bottles mounted to two seafloor drills during IODP Expedition 357 (Früh-Green et al., 2018; Motamedi et al., 2020). Hydrothermal fluid samples collected with ROV Jason II were filtered *in situ* through 0.22 μm Sterivex filter cartridges (Millipore, Billerica, Massachusetts, USA). Niskin water samples were filtered shipboard through 0.22 μm Sterivex filter cartridges using an inline peristaltic pump. Sterivex filter cartridges were stored frozen at –80°C until analysis in the shore-based laboratory, as previously reported (Motamedi et al., 2020; Brazelton et al., 2022). Ambient lab air samples were collected from the R/V Atlantis shipboard laboratory and at the University of Utah during DNA extraction procedures, as we have reported for other studies (Motamedi et al., 2020; Brazelton et al., 2022).

Chimney samples were collected by slurp sampling from the ROV or via grab sampling and placed in onboard boxes unique to each item. Samples were processed aseptically with sterile tools shipboard shortly after collection. Subsamples of various sizes and masses were collected shipboard and either stored in sterile Falcon tubes or Whirlpaks to preserve integrity. Each individual sample is associated with a unique longform Jason dive ID to track any associated parallel analyses performed. Collections were catalogued to record location information, depth, pH and temperature where measured and any notes during sampling. Samples destined for genomics work were immediately frozen at -80°C on the ship and transferred by dry ice to the University of Utah for DNA extraction.

X-ray diffraction mineralogy analysis

The mineralogy of the samples was determined using a Bruker AXS D8 Advance Powder X-ray Diffractomer (Bruker Corporation, Billerica, USA) at the Institute of Geochemistry and Petrology, ETH Zurich. The XRD was equipped with a LynxEye detector and a CuKα source. The analyses were carried out using a scan range of 5-90° 2θ, a step size of 0.02, and a scan time of 0.8 seconds per step. The mineral phases were identified using the ICDD PDF-2 database (International Center for Diffraction Data, USA) and the ICDD Sieve+ automatic peak search software. Weight proportions of the identified minerals were obtained using a full-profile Rietveld refinement method via Siroquant version 3.0 (Seitronics, Australia). Detailed mineralogical studies were carried out by Aquino et al. and are reported in a separate publication (Aquino and Früh-Green, 2024a; Aquino and Früh-Green, 2024b)

DNA extraction and sequencing

Extractions of DNA from chimney biofilm samples were conducted with the FastDNA SPIN kit (MP Biomedical) according to the manufacturer's instructions, followed by washing in Amicon Ultra (Millipore) filter units with 65°C Tris-EDTA buffer. Extractions of DNA from hydrothermal fluid samples and seawater samples were conducted as described previously (Motamedi et al., 2020; Brazelton et al., 2022) and the full extraction protocol is available on protocols.io (DOI: dx.doi.org/10.17504/protocols.io.5qpvoym5xg4o/v2). All DNA preparations were purified with magnetic beads prior to sequencing (Rohland and Reich, 2012).

Sequencing of 16S rRNA gene amplicons was conducted by the Michigan State University genomics core facility. The V4 region of the 16S rRNA gene was amplified with universal dual-indexed Illumina fusion primers (515F/806R) as described elsewhere (Wang et al., 2009). Amplicon concentrations were normalized and pooled using an Invitrogen SequalPrep DNA Normalization Plate. After library quality control (QC) and quantitation, the pool was loaded on an Illumina MiSeq v2 flow cell and sequenced using a standard 500 cycle reagent kit. Base calling was performed by Illumina Real Time Analysis (RTA) software v1.18.54. Output of RTA was demultiplexed and converted to FASTq files using Illumina Bcl2fastq v1.8.4.

Chimney biofilm metagenomic libraries were constructed with size-selected, sonicated DNA fragments of 500-700 bp with the NEBnext Ultra DNA II library kit for Illumina (E7645S). Paired-end sequencing (2 × 150 bp) of metagenomic libraries was conducted at the University of Utah High-Throughput Genomics Core Facility at the Huntsman Cancer Institute with an Illumina NovaSeq 6000 platform.

16S rRNA gene amplicon sequences were processed with cutadapt v. 1.15 and DADA2 v. 1.10.1, as in previous studies (Motamedi et al., 2020; Brazelton et al., 2022). This protocol includes quality trimming and filtering of reads, removal of chimeras, and inference of amplicon sequence variants (ASVs). Taxonomic classification of all ASVs was performed with DADA2 using the SILVA reference alignment (SSURefv132) and taxonomy outline. Sequences from the same sequencing run were analyzed together with DADA2, and the resulting ASVs from separate sequencing runs were merged with phyloseq (McMurdie and Holmes, 2013). The merged set of ASVs included sequences from chimney biofilms (present study), hydrothermal fluids (Brazelton et al., 2022), seawater (Motamedi et al., 2020; Lang et al., 2021; Brazelton et al., 2022), seafloor rock drill cores (Motamedi et al., 2020), ambient lab air [present study and (Motamedi et al., 2020; Brazelton et al., 2022)], and extraction blanks [present study and (Motamedi et al., 2020; Brazelton et al., 2022)]. Potential contaminants were removed from the merged set of ASVs with the decontam package using both the "prevalence" and "frequency" modes, as previously reported (Brazelton et al., 2022). A total of 53 ASVs representing 1,856 sequence counts were removed from chimney biofilm samples with this approach. In addition, ASVs with taxonomic assignments matching those identified as common contaminants of subsurface sequence datasets (Sheik et al., 2018) were removed from downstream analyses, resulting in the removal of an additional 422 ASVs. Finally, ASVs from chloroplasts and mitochondria were filtered from the data set based on taxonomic assignment.

Metagenomic sequencing assembly and annotation

Metagenomic analyses were conducted as described previously for a similar study (Brazelton et al., 2022) and summarized here. Adapters and low-quality bases were removed from metagenomes with bbduk and seq-qc (Bushnell et al., 2017; Thornton et al., 2020). A pooled, co-assembly of all 16 chimney biofilm metagenomes was conducted with Megahit v1.1.1 (Li et al., 2016). Genes were predicted with Prodigal v2.6.3 in meta mode (Hyatt et al., 2010). Predicted protein sequences were queried against the KEGG release 83.2 prokaryotes database with Diamond v0.9.14 (Buchfink et al., 2021). Unassembled sequences were mapped to the pooled Megahit assembly with Bowtie2 v2.3.2 (Langmead and Salzberg, 2012). Binning of metagenome-assembled genomes (MAGs) from the pooled assembly was conducted with BinSanity using the Binsanity-lc workflow v0.2.6.2 and a minimum contig size of 3 kb (Graham et al., 2017). MAGs were assigned taxonomic classifications with GTDB-Tk v1.5.1 (reference data version r202 (Chaumeil et al., 2019);). Completeness and contamination of MAGs were estimated with CheckM v1.0.5 (Parks et al., 2015). The coverage for each predicted protein and KEGG annotation feature was calculated as transcripts (or fragments) per million (TPM) with count_features v1.3.0, part of our seq-annot package (Thornton et al., 2020; Brazelton et al., 2022). TPM is a proportional unit (multiplied by one million for convenience) that is normalized to the length of each predicted protein sequence as well as to the total library size. The coverage of each MAG was calculated as the weighted sum of the normalized, proportional coverages (in TPM) of its member contigs. The contig coverages were obtained by mapping all unassembled reads against each MAG with Bowtie2 and then calculating the average coverage per position of each contig with the genomecov command (with the option -pc) in bedtools v2.30.0 (Quinlan and Hall, 2010). Normalized coverages in units of TPM were calculated by dividing the average coverage per position by the total number of read pairs for that library. Metagenomic bins used for this study are summarized in Supplementary Table 1.

Statistical analyses of amplicon and metagenomic data

16S rRNA data was imported as ASVs and analyzed using the R studio environment version 4.2.2, primarily utilizing the Phyloseq package (McMurdie and Holmes, 2013; R Core Team, 2022). NMDS analyses were done within Phyloseq using the ordinate() function on Bray-Curtis dissimilarities of ASV relative abundances. ASV counts were normalized to the median sequence depth and transformed to proportional abundance. Source-tracking to identify potential sources and sinks of sequences was carried out using the R package FEAST (Shenhav et al., 2019). Samples of seawater collected during the 2018 Lost City expedition, as well as IODP Expedition 357, were classified as putative sources, while plume, biofilm and vent fluids were categorized as sinks (Lang et al., 2021). FEAST outputs assigned a fractional contribution between labeled sources and sinks, categorizing unknown contributions as their own category. Pearson correlation tests were carried out in the R environment using the ltm package and the rcor.test() function for the ASV-mineralogy matrix and the genes-mineralogy matrix (Rizopoulos, 2007). A correction for multiple tests was carried out using the R package qvalue and qvalue() function (Storey, 2002; Storey and Tibshirani, 2003). Significant correlation scores ($p < 0.05$

and q < 0.05) were retained as the statistically significant output for further analysis. Accordingly, three groups of correlations were analyzed within these constraints: aragonite (ARA), brucite (BRU) and calcite (CAL), based on what mineral they were strongly correlated to.

Metagenome Assembled Genomes (MAGs) with over 90% completion and less than 10% contamination were tested for their mineralogical Pearson correlations in a pairwise matrix as before by using their calculated sequence coverage (in TPM). Bin metabolic potential was assessed by the presence and completeness of KEGG metabolic modules for carbon fixation, nitrogen cycling, methane cycling and sulfur cycling, in addition to a selected subset of metal cycling or environmental processing genes (Kanehisa and Goto, 2000).

Results

Chimney mineralogy and community characteristics

Chimney samples were collected from seven sites across the LCHF at depths between 729 and 875 meters below sea level (Table 1). Aragonite and brucite were the dominant minerals, but sample lithologies were usually a combination of all three minerals. Temperature, pH and chemistry of associated hydrothermal fluids were concurrently measured during the same expedition, indicating temperatures at the site up to 96°C, *in situ* pH up to 10, and trends in chemistry consistent with the mixture of H_2-rich hydrothermal fluids with ambient seawater (Aquino et al., 2022; Brazelton et al., 2022).

Examining chimney microbial communities through their ASV composition reveals only moderate clustering of samples collected from the same chimney, such as in Marker 3 samples (Supplementary Figure 1). This heterogeneity is in line with previous observations of physical variability in venting dynamics and mineral composition even across short distances within one chimney structure (Schrenk et al., 2004; Kelley et al., 2005; Ludwig et al., 2006). Microbial species richness is also variable among samples from the same chimney, with MKR3 and MKR8 samples having the lowest richness estimates (Supplementary Figure 2). Nearly 50% of all sequence counts in the dataset are dominated by just 50 ASVs constituting 14 genera (Supplementary Figure 3; Table 2). Omitting unassigned taxa, the most abundant genera show high numbers of unique ASVs, the largest being *Thiomicrorhabdus* with 111 ASVs, indicating a large contribution of sequence diversity from relatively few taxa, consistent with earlier studies (Brazelton et al., 2011).

One possible explanation for the dissimilarity within and between sites is the variable contribution from the ambient seawater microbial community. In this scenario, the least active chimneys would be more impacted by a seawater microbial imprint. To test this, source-sink dynamics were analyzed with the chimney and seawater samples collected in and around the Atlantis Massif (Figure 1). Seawater samples gathered away from the venting environment were considered potential microbial sources while chimney biofilms, hydrothermal fluids, and hydrothermal plumes mixing with seawater were classified as possible microbial sinks. Hydrothermal plume samples and vent fluids showed observable contributions from seawater, as expected, but chimney biofilms showed little if any contributions from seawater microbes. Only 3 of the 24 chimney samples (LC02923, LC02958, and LC02981) contained assignable seawater contribution fractions ranging from 0.02 – 0.20, one of which was an inactive vein, indicating overall, the chimney microbial community has little input from the surrounding seawater (Supplementary Table 2). Two of these samples have high brucite contents, and all three samples have very little calcite which is inconsistent with a simple interpretation that the microbial communities of less active chimneys become more seawater dominated. Long inactive chimneys were not included in this sampling campaign.

Correlations between mineralogy, taxa, and genes

A significant inverse correlation was noted between aragonite and brucite mass fractions (Figure 2). Of 10,416 microbial taxa in the ASV dataset, 86 showed significant correlations with these minerals. Taxon distributions that surpassed statistical significance (p<0.05, q<0.05) were grouped into ARA, BRU and CAL categories based on their associated mineral. Detailed information about these groups is available in Supplementary Table 3 and as relative abundances in Figure 3. Several ASVs comprised significant fractions of the samples, upwards of 15% of total relative abundance, highlighting their centrality within their respective communities.

Clusters with a positive aragonite correlation (ARA) included 6 ASVs, 3 Proteobacteria and one each of Acidobacteria, Actinobacteria, and Bacteroidetes. The highest correlated ASV of this group (R = 0.75) belonged to Acidobacteria *Thermoanaerobaculaceae* Subgroup 10. Clusters with a positive brucite (BRU) correlation score contained 4 ASVs, a *Lewinella*, a Chloroflexi, and a Patescibacteria order MSBL5, and an unclassified bacterial sequence. The highest positive correlation score to brucite was *Lewinella* and an unknown ASV (ASV 32052 in our dataset, R = 0.66). Clusters positively correlated to calcite (CAL) contained 68 ASVs, the largest number of all three groups, and the highest correlation score of all with a *Sulfurovum* ASV (R = 0.82). 21 of the scored sequences belonged to Proteobacteria. 4 archaeal sequences were identified, 3 of which were *ANME-1b* sequences, and 1 archaeal *Candidatus Nitrosopumilus*. Interestingly, one ASV of the recently identified *Candidatus Desulforudis* had a significant association with calcite. In addition, 11 unknown bacterial sequences were moderately to very strongly correlated with calcite.

Automated annotation of individual protein sequences predicted from the LCHF chimney metagenomes identified 7,452 different genes classified by the KEGG orthology. Our analysis focused on those related to carbon, methane, nitrogen and sulfur metabolisms, as well as selected trace metal cycling components. A total of 12 genes with three carbon fixation pathways showed distinct correlations to underlying mineralogy (Table 3). The Calvin-Benson-Bassham cycle showed positive correlations to brucite only in the steps for the conversion of glycerate-3-

TABLE 1 Samples included in this study, their location, label used for analysis, and mineralogical content for aragonite, brucite and calcite in weight percent.

Sample	Jason Sample ID	Location	Location Label	Type	Depth (m)	Latitude	Longitude	Aragonite (wt%)	Brucite (wt%)	Calcite (wt%)
LC02867	J.1107.16Sept.1708	Marker 3	MKR3	Active Chimney	729.67	30.1205243	-42.1243348	30	66	3
LC02870 A/C	J.1107.16Sept.1723	Marker 3	MKR3	Active Chimney	729.69	30.1205071	-42.1243435	20	78	2
LC02876	J.1107.16Sept.1714	Marker 3	MKR3	Active Chimney	729.7	30.1205091	-42.1243436	35	61	4
LC02881	J.1107.16Sept.1713	Marker 3	MKR3	Active Chimney	729.68	30.1205119	-42.1243414	50	49	1
LC02923	J.1108.17Sept.2301	Marker 6	MKR6	Active Chimney	797.97	30.1245919	-42.1193596	65	34	1
LC02928	J.1108.17Sept.2002	Calypso	CAL	Active Chimney	797.95	30.1245946	-42.1193284	23	54	22
LC02934	J.1108.17.Sept.2223	Venting Wall	VW	Venting Wall	795.19	30.1246195	-42.1193754	94	6	1
LC02938	J.1108.17Sept.1338	Calypso	CAL	Active Chimney	797.44	30.1242902	-42.1193738	26	71	2
LC02954	J.1109.Sep19.0756	Sombrero	SOM	Active Chimney	761.71	30.1240556	-42.1195502	96	4	1
LC02958	J.1109.Sep19.1033	Carbonate Veins	CV	Inactive Vein	741.1	30.1249291	-42.1188567	13	59	28
LC02964	J.1109.Sep19.1011	Carbonate Veins	CV	Inactive Vein	740.89	30.1249447	-42.1188485	29	42	29
LC02967	J.1109.Sep19.1016	Carbonate Veins	CV	Inactive Vein	740.9	30.1249298	-42.1188394	47	24	29
LC02981	J.1110.20Sep.0611	Marker 8	MKR8	Active Chimney	801.21	30.1204738	-42.1249175	30	68	1
LC02985	J.1110.20Sep.0609	Marker 8	MKR8	Active Chimney	801.21	30.1204738	-42.1249175	34	63	2
LC02990	J.1110.20Sep.0817	Marker 8 Vein	MKR8V	Inactive Vein	871.44	30.1203248	-42.1257789	14	61	25
LC02993 B/D	J.1110.20Sep.0838	Marker 8 Vein	MKR8V	Inactive Vein	874.96	30.1202751	-42.1256963	38	18	33
LC02998	J.1111.21Sep.0938	Marker 6 Wall	MKR6W	Inactive Vein	788.23	30.1177429	-42.1247982	94	1	6
LC03001	J.1111.20Sep.2240	Sombrero	SOM	Active Chimney	762.11	30.119726	-42.124407	62	32	6
LC03005	J.1111.20Sep.2214	Sombrero	SOM	Active Chimney	762.08	30.1197175	-42.1243882	68	24	9
LC03008	J.1112.22sSep.0113	Marker 6	MKR6	Active Chimney	777.05	30.1207253	-42.1241496	87	12	2
LC03012	J.1112.22Sep.0110	Marker 6	MKR6	Active Chimney	777.16	30.1207231	-42.1241433	51	49	1

10.3389/frmbi.2024.1401831

TABLE 2 Representative table of the most abundant genera sequenced from the vents in order of decreasing total percent relative abundance, and including the number of ASVs found in our dataset.

Genus	# Unique ASVs	% Abundance
Thiomicrorhabdus	111	8.9
Desulfotomaculum	50	8.3
Roseobacter_clade_NAC11-7_lineage	17	7.9
Sulfurovum	74	7.1
Marine_Methylotrophic_Group_2	37	3.5
Sulfurospirillum	15	2.8
Cocleimonas	33	2.6
Serpentinicella	20	2.3
pItb-vmat-59	24	2
Moritella	4	1.7
Marine_Methylotrophic_Group_3	19	1.7
Sva0996_marine_group	70	1.4
IheB2-23	47	1.3
Desulfobulbus	9	0.9
Actibacter	18	0.8
Methanosalsum	7	0.7
Sedimenticola	4	0.4
Lutimonas	6	0.4
Woeseia	85	0.4
Ruegeria	2	0.4
Desulfurivibrio	3	0.4
Filomicrobium	46	0.4
Litoreibacter	3	0.3
Cenarchaeum	2	0.3
Photobacterium	4	0.3

phosphate to 1,3-biphopshoglycerate and from glyceraldehyde-3P to glycerone-P through *pgk* (K00927), and *tpiA* (K01803) enzyme genes (Figure 4).

The TCA and rTCA cycle genes showed some negative correlations to aragonite and brucite. In the TCA cycle, *gltA* (K01647), *sucA* (K00164), *sdhB*/*frdB* (K00240), and *sdhC*/*frdC* (K00241) positively correlated to aragonite.

The Wood-Ljungdahl pathway showed no positive correlations, but *fdhA* (K05299) and *acsB* (K14138) had negative correlations with aragonite. Two subunits of CODH/ACS, *cdhC* (K00193) and *cdhD* (K00194), were positively correlated with brucite. *cdhE* (K00197) displayed a strong negative aragonite correlation. Acetyl-CoA synthetase (*acs:* K01895), which is associated with acetate metabolism in multiple pathways, including acetoclastic methanogenesis, was also positively correlated with brucite (Figure 5).

Methane oxidation showed no positive correlations, but *pmoC-amoC* showed a negative brucite correlation. Formaldehyde assimilation to the serine pathway contained mixed correlations, with *AGXT* (K00830), *hprA* (K00018), and *mtkA* (K14067) correlated with aragonite, and *mtkB* (K08692) negatively correlated to brucite. Conversion of glycerate-2-phosphate to phosphoenolpyruvate mediated by the *echA* (K01689) gene was correlated with brucite.

Correlations between minerals and genes involved in nitrogen cycling included those associated with nitrification and denitrification (Figure 6; Table 4). Ammonia conversion to hydroxylamine using *pmoC-amoC* (K10946) showed a negative brucite correlation, and *pmoB-amoB* (K10945) showed a strong negative brucite correlation. Denitrification from nitrate to nitrogen was aragonite-correlated, with *nirK* (K00368) showing significant negative brucite correlation.

Mineral-gene correlations related to sulfur metabolism included an aragonite correlation with *cysD* (K00957) and a strongly negative brucite correlation with *cysNC* (K00955), both of which are associated with assimilatory sulfate reduction, among other sulfur-metabolizing pathways (Figure 7; Table 5). A strong brucite correlation was found for *cysK* (K01738) with an equally strong negative correlation to aragonite, involved in utilization of sulfide in forming or maintaining the cellular cysteine pool. *hdrA2* (K03388), which is involved in sulfur redox reactions in many different organisms including methanogens and sulfur-cycling bacteria, was correlated with calcite. Components necessary for the SOX system of thiosulfate oxidation and dissimilatory processes showed no notable correlations.

A selection of genes related to trace metal transport systems were examined for any mineral associations, however all significant correlation scores were related to tungsten transport (Table 6). All genes exhibited negative correlations to aragonite, except for *wtpB* (K15496), which showed a negative correlation to brucite.

Metagenome assembled genomes and genes correlated with minerals

The sequence coverages (i.e. relative abundances) of MAGs with greater than 90% completion were tested for correlations with mineralogy and of these, 6 MAGs had significant correlations ($p<0.05$) but only one correlation (Desulfobulbales_Bin048 correlated to calcite) exhibited a q-value score better than our significance threshold ($q < 0.05$) (Supplementary Table 4; Figure 8). The correlations with non-significant q-value scores included aragonite correlations with two MAGs classified as Chromatiales and Thiotrichales and calcite correlations with four MAGs classified as Desulfobulbales, Caulobacterales, and Methylococcales. We will refer to these as aragonite-correlated and calcite-correlated MAGs despite the high q-value scores to survey potential overall differences between the MAGs.

The two aragonite-correlated MAGs contained genes related to sulfur metabolism, thiosulfate oxidation, and copper/zinc transport. Chromatiales_Bin108 contained genes for sulfur metabolism, molybdenum transport and zinc transport. Only one gene

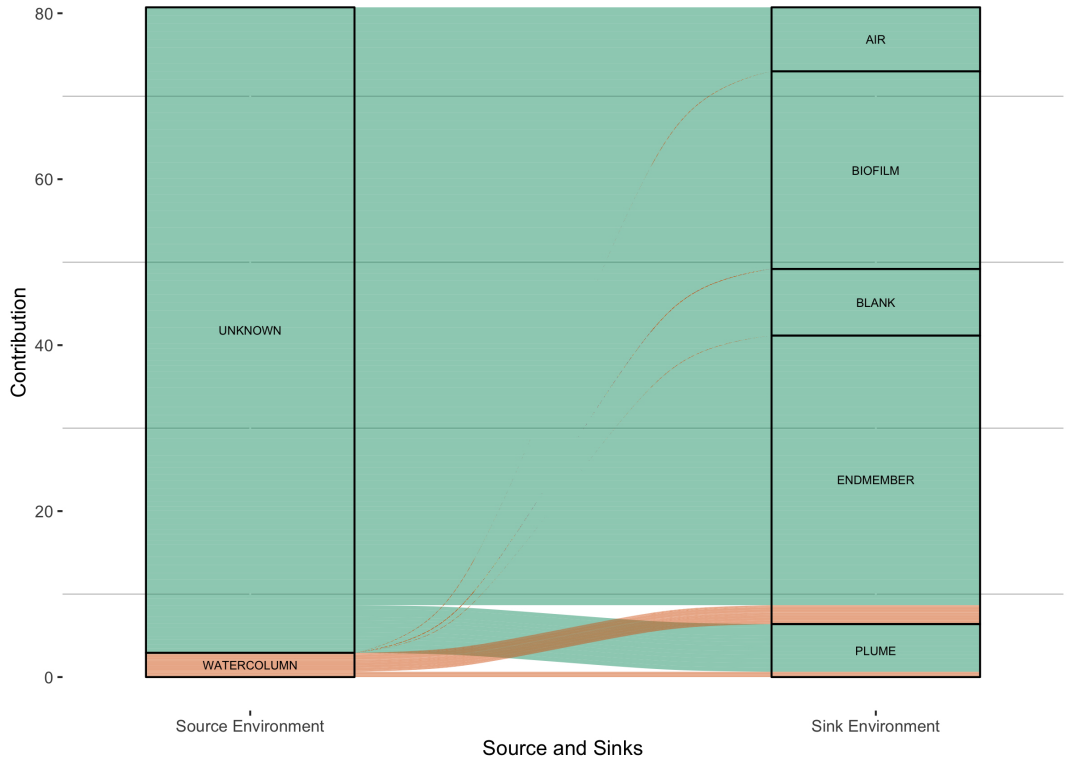


FIGURE 1
Simplified visualization of FEAST calculated source contributions of microbial community members within the LCHF local region. Sources and sinks were primarily fluid sampled historically from across the Atlantis Massif and vent end member fluids from the 2018 expedition. The left column are possible source categories into the column on the right. The majority of data cannot be assigned a source and is considered unknown, while the seawater samples collected from the water column at and around Lost City contribute notable portions into vent/endmember fluids and plume samples. Biofilm samples represent the crushed and DNA extracted carbonate samples from the 2018 expedition. Blank and Air samples were used during sample processing and DNA extraction and sequenced accordingly. Vent endmember samples were fluid samples taken directly from the venting chimneys at Lost City. Plume samples were putative Lost City plumes sampled via CTD and represent an intermixing between sea and vent environments. The data here is represented as the square root of the contribution fraction in order to enhance the visualization of small contributions against the largely unknown portions. The original unedited data can be seen in Supplementary Table 2.

potentially associated with assimilatory sulfate reduction (*cysC*) and three genes for thiosulfate oxidation (*soxB*, *soxY* and *soxZ*) were present in this MAG. The combination of genes found likely represents an oxidation pathway using the *Dsr*/*Apr* genes, supported by the absence of *dsrC*, and the taxonomy of organisms with oxidative and reductive pathways associated with *dsrAB*. Two nickel/peptide permease transporters and one nickel/peptide ATP-binding protein were also found. Thiotrichales_Bin148 showed a

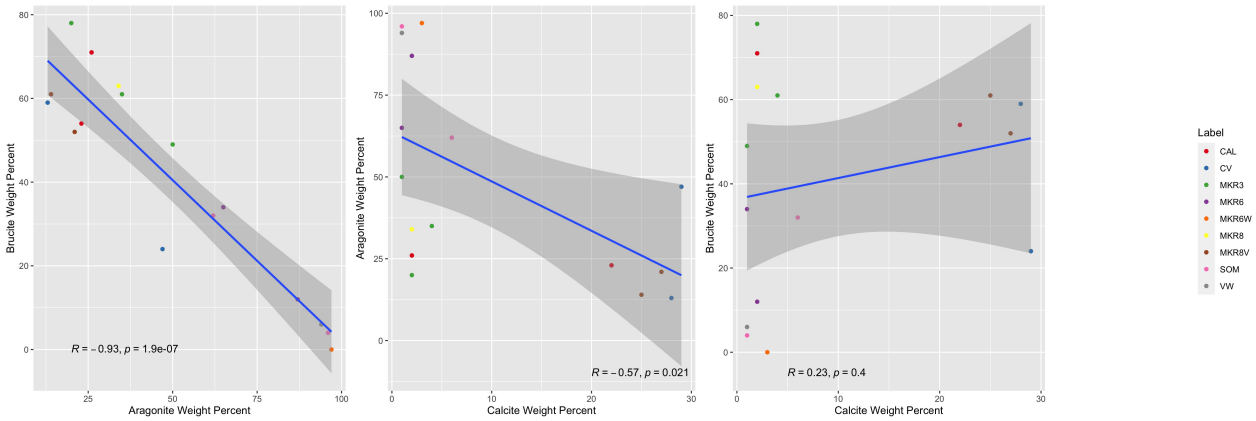


FIGURE 2
Linear correlation between the different mineralogy weight percents of samples, colored by location. From left to right: brucite weight percent to aragonite weight percent, calcite weight percent to aragonite weight percent, and brucite weight percent to calcite weight percent.

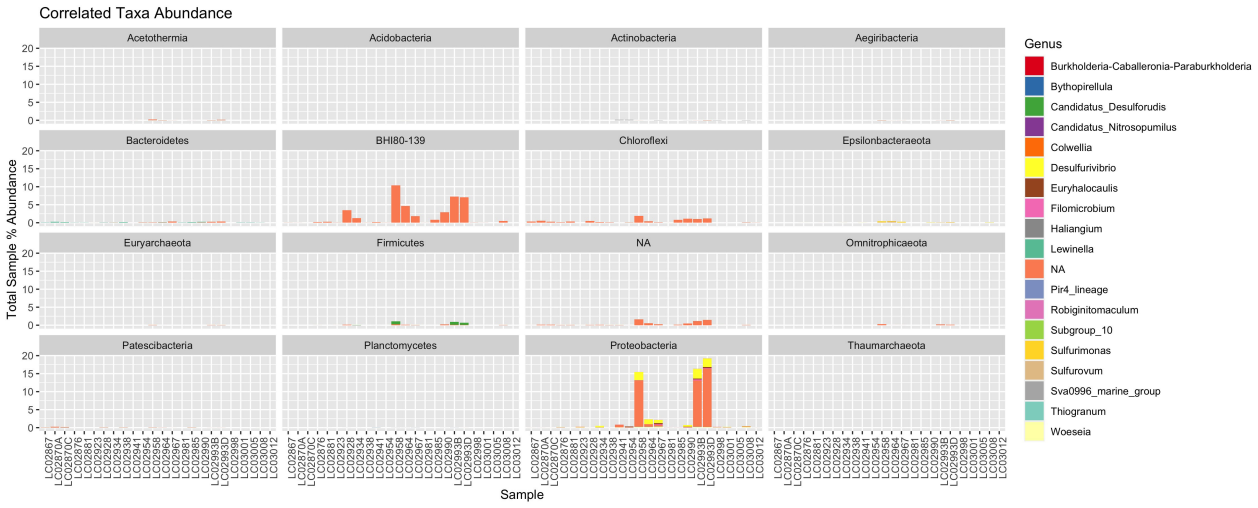


FIGURE 3
A stacked bar plot showing the abundance of correlated ASV genera in each sample, faceted by their respective phyla. Some of the phyla constitute over 15% of the total sample reads, making them significantly abundant community members.

similar pattern, with genes associated with sulfur and thiosulfate cycling, as well as complete modules for copper transport, heme export and tungstate transport. In both MAGs, there was an absence of nitrogen cycling and Wood-Ljungdahl genes.

The four calcite-correlated MAGs were classified as Desulfobulbales, Caulobacterales, and Methylococcales. Calcite-correlated MAGs contained more genes related to assimilatory sulfate reduction, denitrification, carbon cycling, and cobalt or nickel transporters. The Desulfobulbales_Bin048 contained complete modules for dissimilatory sulfate metabolism and zinc transport. A partial Wood-Ljungdahl pathway was found with the presence of *acsB, acsE, cooS, fdhA* and *cdhE*. Caulobacterales_Bin084 contained key genes for denitrification (Figure 8), as well as two sulfite reductase genes (*cysD* and *cysNC*) and some genes for heme export and nickel/peptide transport. The two Methylococcales MAGs had a similar collection of assimilatory sulfate reduction genes and several of the metal transport modules. However, Methylococcales_Bin126 contained two of the SOX system genes, *soxY* and *soxZ*, and several cobalt transport genes not found in its sister bin.

Discussion

Our study used mineralogy as a proxy for varying microenvironmental conditions to investigate changes in associated microbial communities living within hydrothermal chimneys at the LCHF. The three dominant minerals at the site, aragonite, brucite and calcite, represent physical snapshots of *in situ* temperature, pH and venting conditions. Younger chimneys with active and robust flow are dominated by brucite and primary calcite and are expected to be stable with hydrothermal venting at relatively higher temperatures, while lower temperature chimneys are composed of greater aragonite concentrations (Ludwig et al., 2006; Aquino and Früh-Green, 2024a). Inactive chimneys are dominated by secondary calcite as temperatures drop below 15°C (Ludwig et al., 2006). The presence of an inverse relationship in samples between aragonite and brucite (Figure 5) highlights this distinct relationship, though the exact parameters of vent evolution are beyond the scope of this particular study and have been investigated by other recent studies (Aquino et al., 2022; Aquino et al., 2023; Aquino and Früh-Green, 2024a).

This pattern provides a measurable physical environmental gradient against which to analyze correlations of different microbial taxa and genes. While correlation does not imply causation, examining microbial distributions through the lens of mineral proxies provides a window into the complex chimney environments. Community analysis through taxon abundance can be confounded by multiple environmental changes and microbial turnover over long timescales, creating noise that complicates identification of significant species or consortia. Bulk analysis by nature takes advantage of a processed sample of inherently mixed and slightly different environments and microbial residents. In a site like Lost City, no two samples are exactly alike, and we are faced with the challenge in defining the strategies of a large number of distinct species utilizing different substrates under variably extremophilic conditions (Hutchinson, 1961; Aquino and Früh-Green, 2024b). The first problem, longevity of the environment, means a slow evolution of conditions over timescales longer than the biological processes, ranging from decades to tens of thousands of years, as demonstrated by the ages of active and inactive Lost City chimneys (Ludwig et al., 2006). The second problem concerns microbial turnover, resulting in a slowly evolving environment conducive to microbial coevolution and a sufficient timespan for extensive mixing and community succession. This complexity hinders the accurate identification of primary colonizers or keystone species as the community diversifies with time. This is further complicated metabolically when considering highly variable rates of microbial activity, dormancy, intercellular cooperation, inhibition, or horizontal gene transfer imparting a mosaic of

10.3389/frmbi.2024.1401831

TABLE 3 Selected significantly correlated ($p < 0.05$, $q < 0.05$) carbon cycling genes.

KEGG ID	Mineral Correlation(s)	Correlation Score	Gene	Metal Cofactor
Carbon Cycling				
K00927	Brucite	0.73	*pgk*	Mg
K01803	Brucite	0.68	*tpiA*	–
K14138	Aragonite	-0.72	*acsB*	Ni
K05299	Aragonite	-0.71	*fdhA*	Fe, Mo
K00164	Aragonite	0.72	*sucA*	–
	Brucite	-0.70		
K00240	Aragonite	0.69	*sdhB/frdB*	Fe
K00241	Aragonite	0.70	*sdhC/frdC*	–
	Brucite	-0.68		
K01647	Aragonite	0.67	*gltA*	Fe
	Brucite	-0.68		
K15232	Aragonite	-0.68	*ccsA*	–
	Brucite	0.68		
K01677	Aragonite	-0.72	*fumA*	Fe
	Brucite	0.72		
K01678	Aragonite	-0.74	*fumB*	Fe
	Brucite	0.72		
K00177	Aragonite	-0.67	*korC*	–
K01895	Aragonite	-0.72	*acs*	Mg, Mn
	Brucite	0.71		
K00193	Aragonite	-0.73	*cdhC*	Fe, Ni
	Brucite	0.73		
K00194	Aragonite	-0.76	*cdhD*	Fe
	Brucite	0.70		
K00197	Aragonite	-0.75	*cdhE*	Fe
K03388	Aragonite	-0.70	*hdrA2*	Fe
	Calcite	0.74		
K10946	Brucite	-0.72	*pmoC-amoC*	–
K00830	Aragonite	0.68	*AGXT*	–
	Brucite	-0.72		
K01689	Brucite	0.67	*echA*	Mg
K08692	Brucite	-0.72	*mtkB*	Mg
K14067	Aragonite	0.71	*mtkA*	Mg
	Brucite	-0.75		
K01622	Aragonite	-0.70	*K01622*	Zn

Each gene characterized by its Kegg ID displays the Pearson correlation score and if found, its metal cofactor.

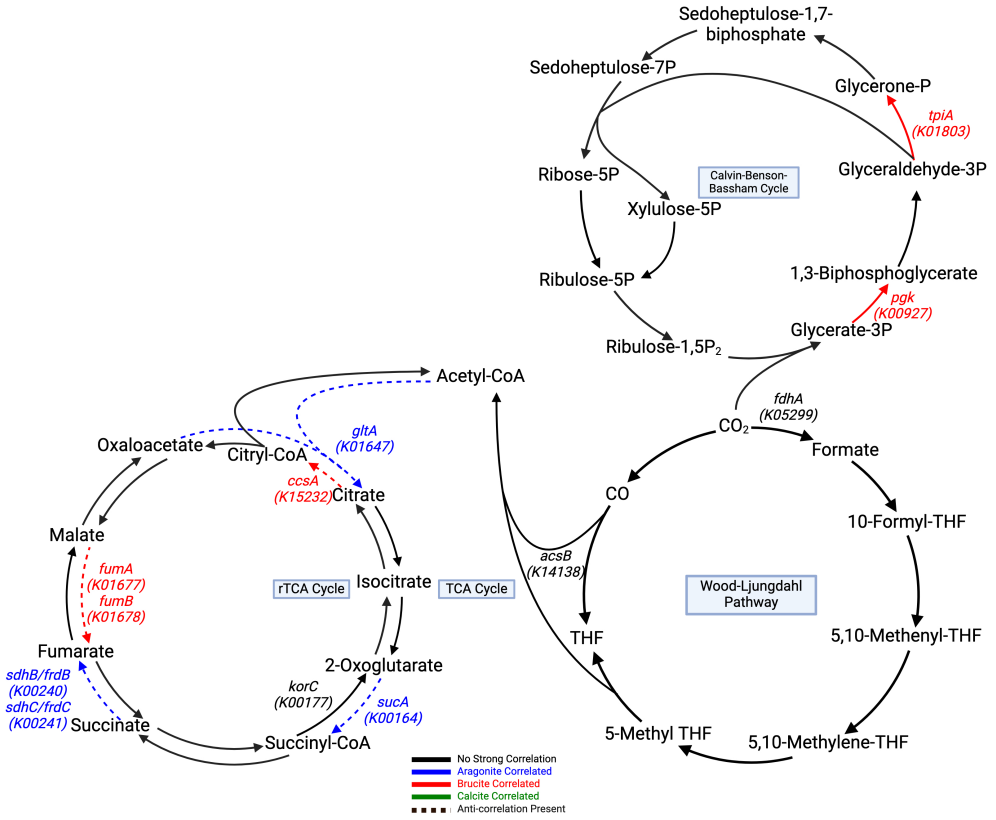


FIGURE 4
Carbon fixation pathways assembled from correlated genes with a Kegg ID color coded to their respective positive mineral correlation. A dashed line represents an anti-correlation present, as specified in Table 3. Genes with negative correlations are highlighted in black. Calvin-Benson-Bassham Cycle, and rTCA show generalized brucite association, while the Wood-Ljungdahl pathway shows negative aragonite correlation at key steps. The TCA cycle contains pathway steps correlated with aragonite, but negatively correlated to brucite, while the rTCA cycle pathway steps show brucite correlated genes that are also negatively correlated to aragonite. All remaining pathway genes were found but showed no significant correlations. Created with BioRender.com.

functions (Price and Sowers, 2004; Edwards et al., 2012; Fuchsman et al., 2017; Colman et al., 2018; Verma et al., 2022). The resultant challenge then becomes delineating a finite set of variables suitable across the relevant environmental age that some numbers of microbes are responsive to, allowing differentiation between many of the commonly found cosmopolitan microbes and more environmentally restricted individuals. Our results conservatively focus on the most significantly correlated genes and microbes to explore the most potentially responsive relationships within a highly complex biogeochemical system. Potentially, there are many co-associations within our dataset to further explore, and this analysis represents a new perspective for consideration of the slowly changing environmental conditions at Lost City.

Chimney microbial communities generally differ from surrounding fluids

Source tracking analysis shows that the microbial communities detected within the chimneys are not heavily influenced by surrounding seawater communities. This differs from vent fluid and plume samples showing greater exchange between seawater and subsurface fluid sources. Furthermore, the microbial community structures of chimney biofilms are both highly variable and distinct from those of venting fluids (Supplementary Figure 4). While aragonite remains relatively stable through extinction of hydrothermal venting, increased calcite represents the final stage of the vent lifecycle due to its formation at much lower temperatures (Ludwig et al., 2006; Aquino and Früh-Green, 2024b). Higher levels of microbial diversity and richness are observed in vents with high calcite concentration which could be related to many variables including more moderate conditions, and an increased availability of substrates and cofactors either through seawater input or their increased solubility due to more moderate pH and lower ambient temperatures. Additionally, species feeding off organic matter produced by chemolithoautotrophs may facilitate colonization of the mineral substrate by diverse consumers, akin to successional processes documented in other environments (Nemergut et al., 2007; Tarlera et al., 2008). A strong *Candidatus_Nitrosopumilus* correlation to calcite may be indicative of this gradient as they include known autotrophic ammonia oxidizers that physiologically vary accordingly between bathypelagic and hadal zones of the ocean (Qin et al., 2017; Rios-Del Toro et al., 2018; Zhong et al., 2020). The identification of correlated *Sulfurovum* ASVs firmly within the calcite associated category may be evidence of at least some taxa adapted to

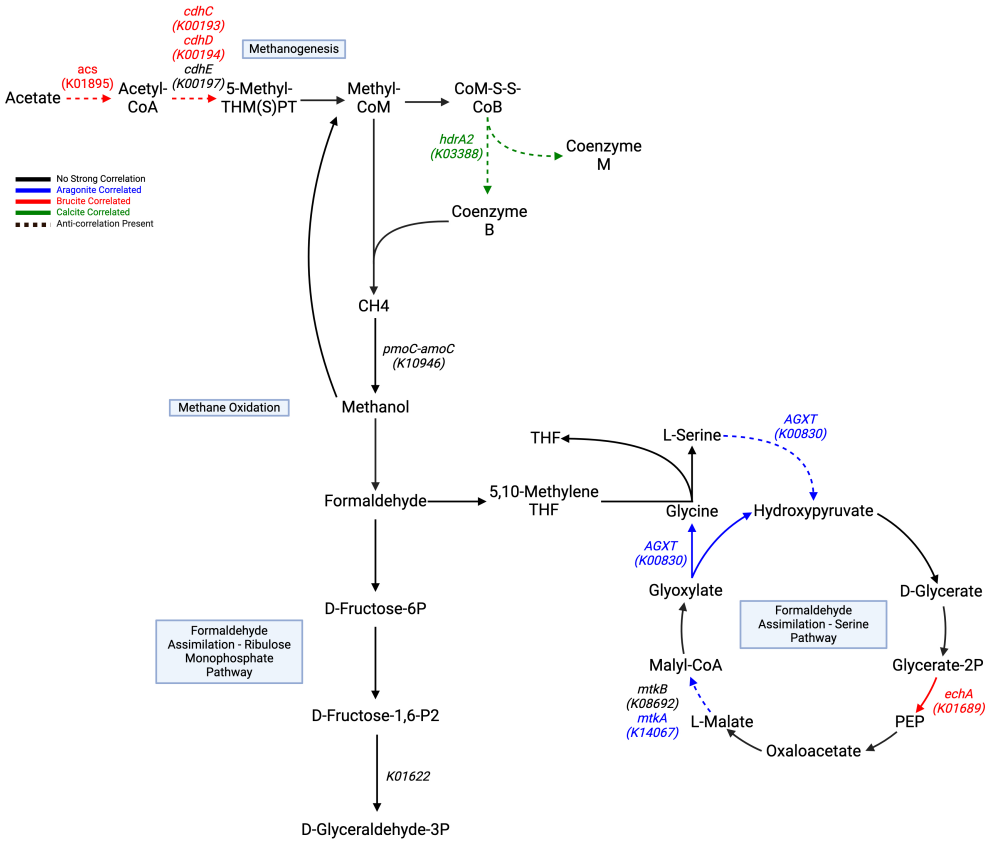


FIGURE 5
Different directions of methane/carbon cycling show trending correlations to different minerals, as highlighted in Table 3. Methanogenesis component genes correlate highly to brucite and calcite and showing negative correlations to aragonite. Methane oxidation and formaldehyde assimilation through the serine pathway are broadly correlate to aragonite, with negative correlations to brucite. Formaldehyde assimilation through the ribulose phosphate pathway shows no distinct positive correlation, but one strongly negative aragonite correlation for K01622. Acetate/*acs* are included to illustrate its utilization for acetoclastic methanogenesis. All remaining pathway genes were found but showed no significant correlations. Created with BioRender.com.

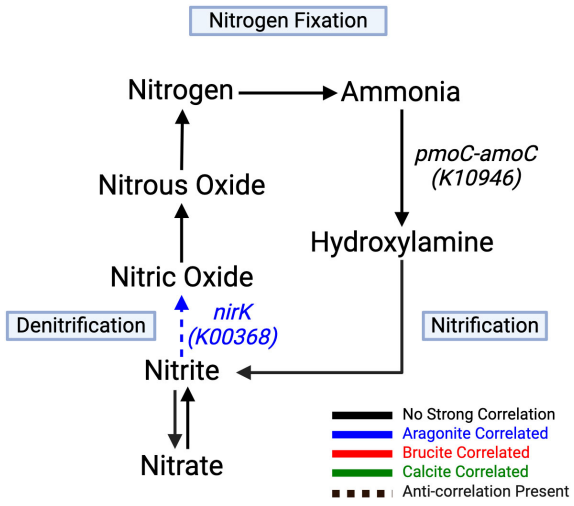


FIGURE 6
Nitrification and denitrification arms strongly correlate to aragonite. Conversion of nitrate to nitrite by the napA (K02567) gene product shows strong negative correlation to brucite. All remaining pathway genes were found but showed no significant correlations. Created with BioRender.com.

hydrothermal conditions, including exposure to less anoxic seawater. Previous studies examining Lost City *Sulfurovum* suggest a mixotrophic lifestyle or the capability of inhabiting a transition zone between the anoxic interior and exterior of the vent (McGonigle et al., 2020). Our results further support this hypothesis given the increased abundance of this taxon towards calcite enriched regions and the overall distribution through the vent may indeed show it colonizing an intermediate region. *Candidatus Desulforudis*, initially discovered kilometers below the continental surface and subsequently found in another terrestrial serpentinite-hosted site, suggest potential interconnections between subsurface environments and the possibility of an ecological niche along the venting pathway or a subsurface-driven dispersal mechanism (Sabuda et al., 2020; Gittins et al., 2022). *Sulfurovum* and multiple other hydrothermally associated taxa have been frequently identified in marine sediments to varying extents further highlighting some shared environmental similarity between different locales where they are found (Mino et al., 2014; Anderson et al., 2017; Mori et al., 2018; Sun et al., 2020).

Microbial activities have also been known to be nucleation sites for carbonate precipitation and have been hypothesized to contribute to the carbonate formation in hypersaline lakes, and

TABLE 4 Selected significantly correlated ($p < 0.05$, $q < 0.05$) nitrogen cycling genes.

KEGG ID	Mineral Correlation(s)	Correlation Score	Gene	Metal Cofactor
Nitrogen Cycling				
K10946	Brucite	-0.72	*pmoC-amoC*	–
K00368	Aragonite	0.76	*nirK*	Cu
	Brucite	-0.73		

Each gene characterized by its Kegg ID displays the Pearson correlation score and if found, its metal cofactor.

low temperature alkaline hydrothermal vents such as in Prony Bay (Dupraz et al., 2004; Ries et al., 2008; Pisapia et al., 2017). There is a high likelihood that carbon metabolism increases carbonate mineralization either through carbonic anhydrase (catalyzing bidirectional formation of bicarbonate from CO_2 and water) or urease (catalyzing urea breakdown to CO_2 and ammonia) activity, with additional evidence previously showing interconversion between formate and CO_2 (Wei et al., 2015; Krause et al., 2018; Lang et al., 2018; Sundaram and Thakur, 2018). This possible interlinkage between CO_2 cycling and carbonate precipitation could explain the microbial correlations with aragonite and calcite. The most abundant taxa, *Thiomicrorhabdus* (111 ASVs), is known to contain carbonic anhydrase activity, and although it is not correlated with specific minerals in our dataset, it is widely dispersed and occupies much of the carbonate-bearing substrate within the site (Brazelton and Baross, 2010; Brazelton et al., 2010). The same can be found for *Desulfotomaculum* (50 unique ASVs), *Sulfurovum* (74 ASVs) and *Woeseia* (85 ASVs), and if the pattern holds true, would place them within intermediate regions of the vent at Lost City and associated with a wide habitat range on the seafloor.

The large number of ASVs associated with the most abundant taxa points to a microdiversity arising from long lived and relatively stable populations within the site. The higher resolution provided with ASV analysis highlights the possibility that large ASV populations contain slightly different ecotypes across an evolving habitat range as a persistence and resilience strategy of the taxa, similar to environments with seasonal cycles (Larkin and Martiny, 2017; García-García et al., 2019; Fodelianakis et al., 2022). Though Lost City does not experience significant seasonal changes, the venting environment does change over the lifetime of the vent and the microdiversity could be a byproduct of this slow change at the microenvironmental scale over time (Martiny et al., 2023).

Brucite relationships as a window into the most extreme metabolic adaptations

The very limited number of taxa correlated with brucite, including the lack of any correlated high-quality MAGs, may be indicative of the more extreme environmental conditions associated with brucite stability. Aragonite and brucite show an inverse correlation to each other (Figure 2), and some taxa mirror this distribution creating the appearance of distinct habitat ranges for the correlated taxa. *Lewinella*, for example, is prevalent and diverse in anaerobic and saline environments, including bioreactors (Mourgela et al., 2023). An ASV of the order MSBL5 correlated with brucite and is notable because of its identification in hypersaline and other serpentinizing environments, such as the Prony Bay Hydrothermal Field (Suzuki et al., 2013; Postec et al., 2015; Frouin et al., 2018; Lecoeuvre et al., 2020). Several other MSBL5 ASVs were associated with calcite, and when considering the formation of primary calcite with brucite early in vent formation, and association with other serpentinizing sites there is a possibility that this order's niche may be within the interior of the vent brucite/calcite regions. A similar pattern can be seen for

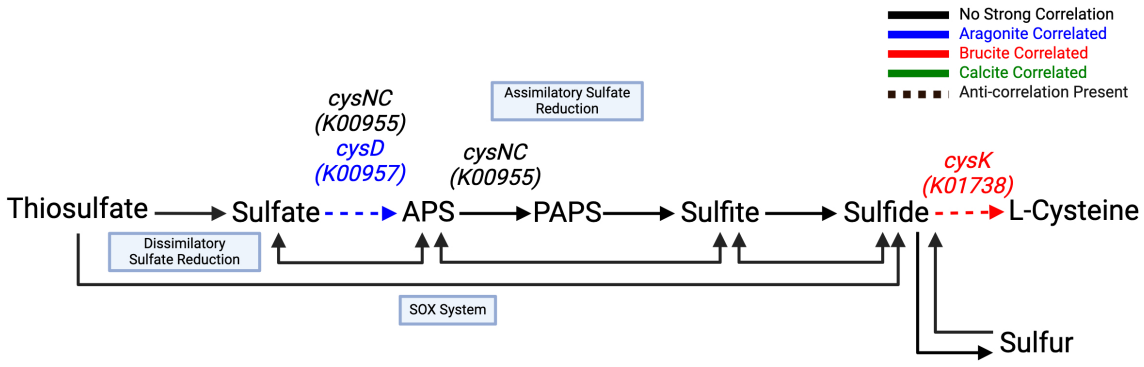


FIGURE 7
Sulfur cycling through assimilatory sulfate reduction is favored in the aragonite environment. Utilization of sulfide for cysteine biosynthesis is highly correlated to brucite via the cysK (K01738) gene product. All remaining pathway genes were found but showed no significant correlations. Created with BioRender.com.

TABLE 5 Selected significantly correlated ($p < 0.05$, $q < 0.05$) sulfur cycling genes.

KEGG ID	Mineral Correlation(s)	Correlation Score	Gene	Metal Cofactor
Sulfur Cycling				
K00955	Brucite	-0.76	*cysNC*	Mg
K00957	Aragonite	0.70	*cysD*	Mg
	Brucite	-0.82		
K01738	Aragonite	-0.69	*cysK*	–
	Brucite	0.74		

Each gene characterized by its Kegg ID displays the Pearson correlation score and if found, its metal cofactor.

ASVs of class Parcubacteria, known to exist in hyperalkaline and hadal environments, for which there is a single brucite associated taxa, and two calcite associated ones (León-Zayas et al., 2017; Suzuki et al., 2017).

Gene correlations highlight metabolic differentiation within vent chimneys and may inform community differentiation

Within vent chimneys, gene correlations shed light on the metabolic differentiation among resident microbes. By mapping correlated genes within their respective pathways, we highlight the dominant metabolic strategies of these microbial populations attributable to their environment and location within LCHF. Our amplicon sequence data and observed high number of ASVs within different taxa suggests specialized populations adapted to various niches within the Lost City environmental range. Substrate availability also plays a pivotal role, influencing either an increase in gene abundance within specialized microbial populations over time or reflecting inherent limitation *in situ*. Limited and essential substrates necessitate the heightened production of uptake and utilization machinery, leading to increased gene abundances associated with these processes.

While we identified few correlated components of the TCA and rTCA cycles (Figure 4), the former exhibited more correlations with aragonite. Initial use of acetyl-CoA by *gltA* correlated with aragonite, but also strongly negatively correlated with brucite. The rate-limiting step of the TCA cycle involves the conversion of isocitrate to 2-oxoglutarate/α-ketoglutarate, and its conversion to succinyl-CoA correlates with aragonite but negatively with brucite. The positive correlations observed for *sdhB/sdhC* (succinate dehydrogenase complexes) may represent an adaptation for maintaining TCA cycle products, particularly fumarate, under stressful conditions which would correspond with moderate hydrothermal conditions in high aragonite domains (Gaupp et al., 2010). Here, we presume that similar correlations in a common and critical cycle as this may indicate the overall trend, and why their placement with the strong negative and positive correlations places them in an aragonite region. Though only two correlated products were identified, the rTCA cycle may show the reverse relationship. Two genes positively correlated with brucite and negatively correlated with aragonite were *fumA* and *fumB*, which encode for conversion of malate to fumarate, further underlining convergence on fumarate. Two genes independently correlated to brucite may be driven by the specific stability of a final two subunit product over a larger single enzyme, but it is unclear why. The large subunit of citryl-CoA synthetase encoded by *ccsA*, highly correlating with brucite, participates in the specific ATP-dependent conversion of citrate to citryl-CoA, and has resemblance to succinyl-CoA synthetase with a high degree of ATP sensitivity (Aoshima et al., 2004). An inhibitory effect by succinyl-CoA synthetase may cause discontinuation of the rTCA cycle and would explain why an increased abundance of *ccsA* would help drive the conversion process forward against a possible inhibitory threshold. The mixtures of negative and positive correlations here could be suggestive of key cycle components being brucite associated.

Because of the prevalence of archaea associated with methanogenesis and the anaerobic oxidation of methane in LCHF fluids and chimneys we were particularly interested in testing for mineral correlations with genes involved in methane metabolism and the Wood-Ljungdahl pathway (Schrenk et al., 2004; Brazelton et al., 2006; Brazelton et al., 2022). Some of the

TABLE 6 Selected significantly correlated ($p < 0.05$, $q < 0.05$) metal transport genes.

KEGG ID	Mineral Correlation(s)	Correlation Score	Gene	Metal Cofactor
Metals				
K05773	Aragonite	-0.72	*tupB*	W
K06857	Aragonite	-0.73	*tupC*	W
K15496	Brucite	-0.70	*wtpB*	Mo/W
K15497	Aragonite	-0.67	*wtpC*	Mo/W

Each gene characterized by its Kegg ID displays the Pearson correlation score and if found, its metal cofactor.

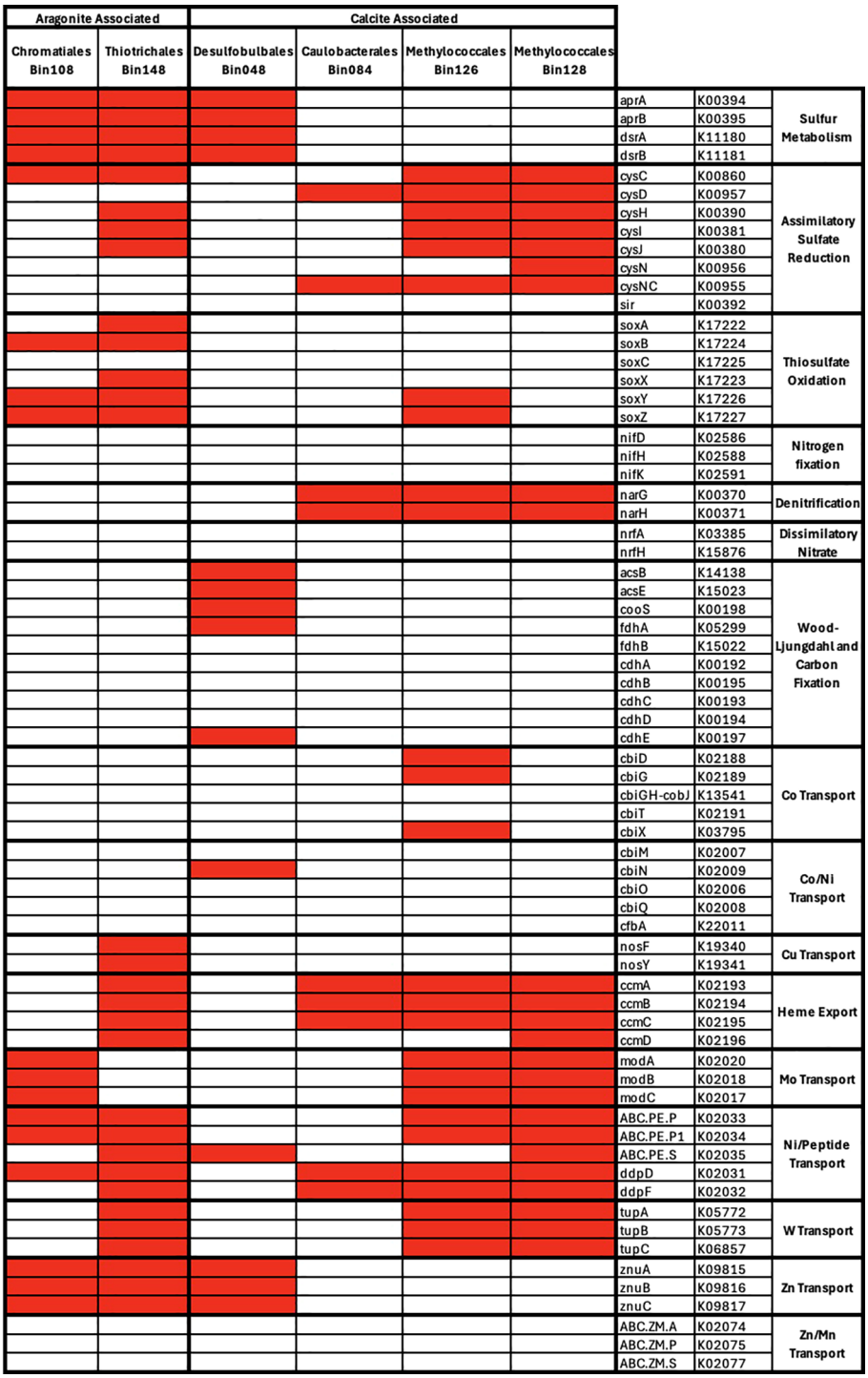


FIGURE 8
Presence and absence of selected genes related to metabolic and metal cycling in bins correlated with aragonite and calcite. Bins 90% complete with <10% contamination were selected and presented here at the taxonomic order level.

subunits of the CODH/ACS enzyme (*cdhC*, *cdhD*, and *cdhE*) exhibited correlations with aragonite or brucite, but overall, there were no consistent mineralogical correlations with the Wood-Ljungdahl pathway. One likely explanation for this result is the presence of Wood-Ljungdahl genes in several different archaeal and bacterial species in LCHF fluids and chimneys (McGonigle et al., 2020; Brazelton et al., 2022).

Multiple bacterial species associated with aerobic methane oxidation have been identified in LCHF chimneys and fluids (Brazelton et al., 2006; Brazelton et al., 2022). Indeed, our results show that brucite has a strong negative correlation with *pmoC-amoC*, consistent with aerobic methane oxidation occurring at cooler, more seawater-dominated zones of the chimney, leaving methanotrophy as a regional function isolated from the primarily reducing conditions at

the vent core. Furthermore, the ribulose monophosphate pathway to D-fructose-6P associated with type I methanotrophs showed no clear correlations, while type II methanotrophs utilizing the serine pathway of formaldehyde assimilation showed strong negative correlations to brucite (Bowman, 2006).

Trace metal availability and sensitivity can influence key metabolic cofactors

Many of the key enzymes in carbon fixation and sulfur, nitrogen, and methane cycling rely on metal cofactors for the catalytic centers of their products. Lost City chimneys, however, are considered iron-poor with concentrations below a 50 ppm detection limit, and other metals (Mn, Ni, Co, Cu, and Mo) are present in single digit ppm concentrations or below detection levels in active vents (Ludwig et al., 2006). Therefore, we were interested in potential connections between metal transporters and specific minerals. We identified no correlations with iron transporters or related genes, possibly indicating no significant differences in iron availability at the site. However, genes related to tungsten or tungsten/molybdenum transport show a strong correlation with brucite. Where those metal-associated genes show relationships could indicate a localized metal requirement and a significant energetic expenditure for their uptake. Previous studies examining the distribution of trace metals between biofilm and chimney mineral matrix showed 83% of iron detected was concentrated within the biofilms, and the remainder in the chimney matrix, in addition to high levels of tungsten accumulation in hydrothermal environment (Demina et al., 2015; Coimbra et al., 2017). The correlations point to a possibly increased utilization of tungsten and/or molybdenum within brucite regions (Ludwig et al., 2006).

Conclusions

Our research focused on using mineral abundances as a physical indicator of microenvironmental conditions within Lost City chimneys. This approach allowed us to pinpoint microbial taxa and crucial metabolic components that are responsive to these mineralogical variations. Initially, vent chimneys are rich in brucite, reflecting the hot, anoxic fluids from the serpentinizing system below. In this harsh setting, only a few organisms persist, possibly relying on mutually beneficial relationships to make the most of the available methane, hydrogen, and limited dissolved inorganic carbon. As the vents mature, they develop a combination of aragonite and brucite, with brucite diminishing as temperature decreases. This shift leads to a less exclusive environment, supporting a wider variety of microbes and metabolic processes, as evidenced by changes in carbon fixation, and methane, nitrogen, and sulfur cycling. In the absence of cultured specimens, this kind of high-throughput data analysis helps us to dissect how these microbial communities are structured and to identify their key constraints. Lost City-type environments are believed to have been widespread on early Earth. Studying the structure of microbial communities in these environments can reveal how they, and their potentially ancient metabolic strategies, evolved with the changing chemistry of the system. The origins of protometabolism on Earth may have occurred in conditions like those at Lost City (Russell and Hall, 1997; Sojo et al., 2016; Kitadai et al., 2019). This implies that the advantages and limitations of this environmental setting have persisted for billions of years, with contemporary organisms highly adapted to them. By further pinpointing, measuring, and defining the conditions these microbes persist within, we can gain a clearer understanding of how microbial life endures in other harsh environments.

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/, PRJNA1072998 https://www.ncbi.nlm.nih.gov/, PRJNA1074139.

Author contributions

OA: Conceptualization, Formal analysis, Investigation, Writing – original draft, Writing – review & editing. WB: Investigation, Methodology, Resources, Supervision, Validation, Writing – review & editing. KA: Formal analysis, Investigation, Methodology, Writing – review & editing. KT: Writing – review & editing. HP: Writing – review & editing. GF-G: Funding acquisition, Resources, Writing – review & editing, Methodology. SL: Funding acquisition, Resources, Writing – review & editing. MS: Funding acquisition, Project administration, Resources, Supervision, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. Support for this work was provided by NSF Awards OCE-1536702 / 1536405 / 153962, the Swiss National Science Foundation project 200021-163187, Joint Institute for the Study of the Atmosphere and the Ocean (JISAO) / NOAA Cooperative Agreement NA15OAR4320063 Contribution 2021-1132, the Deep Carbon Observatory 2015-14084, and the NASA Astrobiology Institute NNA15BB02A. Additional support was provided by the Michigan Space Grant Consortium and the Philippine Department of Science and Technology - Science Education Institute (DOST-SEI).

Acknowledgments

We extend our gratitude for the careful cooperation, advice and help of the captain and crew of the R/V Atlantis, ROV Jason and the scientific party aboard for expedition AT42-01. We are especially grateful for the technical support provided by Dr. Julia McGonigle for the preparation and processing of samples for sequencing.

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.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frmbi.2024.1401831/full#supplementary-material

References

Anderson, R. E., Reveillaud, J., Reddington, E., Delmont, T. O., Eren, A. M., McDermott, J. M., et al. (2017). Genomic variation in microbial populations inhabiting the marine subseafloor at deep-sea hydrothermal vents. *Nat. Commun.* 8, 1114.

Anderson, R. E., Sogin, M. L., and Baross, J. A. (2015). Biogeography and ecology of the rare and abundant microbial lineages in deep-sea hydrothermal vents. *FEMS Microbiol. Ecol.* 91, 1–11.

Aoshima, M., Ishii, M., and Igarashi, Y. (2004). A novel enzyme, citryl-CoA synthetase, catalysing the first step of the citrate cleavage reaction in Hydrogenobacter thermophilus TK-6. *Mol. Microbiol.* 52, 751–761.

Aquino, K. A., Früh-Green, G. L., Bernasconi, S. M., Bontognali, T. R. R., Foubert, A., and Lang, S. Q. (2024a). Controls on mineral formation in high pH fluids from the Lost City hydrothermal field. *Geochem. Explor. Environ. Analy.* 25, e2023GC011010. doi: 10.1029/2023GC011010

Aquino, K. A., Früh-Green, G. L., Bernasconi, S. M., Rickli, J., Lang, S. Q., and Lilley, M. D. (2024b). Fluid mixing and spatial geochemical variability in the Lost City hydrothermal field chimneys. *Geochem. Explor. Environ. Analy.* 25, e2023GC011011. doi: 10.1029/2023GC011011

Aquino, K. A., Fruh-Green, G. L., Bernasconi, S. M., Foubert, A., and Lang, S. Q. (2023). Others, Linking mineral formation and microbes in high pH fluids: Constraints from the Lost City hydrothermal field. ESS Open Archive . April 30, 2023. doi: 10.22541/essoar.168286823.37608718

Aquino, K. A., Früh-Green, G. L., Rickli, J., Bernasconi, S. M., Lang, S. Q., Lilley, M. D., et al. (2022). Multi-stage evolution of the Lost City hydrothermal vent fluids. *Geochim. Cosmochim. Acta.* 332, 239-262. doi: 10.1016/j.gca.2022.06.027

Bianchi, D., Weber, T. S., Kiko, R., and Deutsch, C. (2018). Global niche of marine anaerobic metabolisms expanded by particle microenvironments. *Nat. Geosci.* 11, 263–268.

Bowman, J. (2006). The Methanotroph s—The Fam ilies Methylococcaceae and Methylocy staceae. *Prokaryote s* 5, 266–289.

Boyd, E. S., Costas, A. M. G., Hamilton, T. L., Mus, F., and Peters, J. W. (2015). Evolution of molybdenum nitrogenase during the transition from anaerobic to aerobic metabolism. *J. Bacteriol.* 197, 1690–1699.

Brazelton, W. J., and Baross, J. A. (2010). Metagenomic comparison of two Thiomicrospira lineages inhabiting contrasting deep-sea hydrothermal environments. *PloS One* 5, e13530.

Brazelton, W. J., Ludwig, K. A., Sogin, M. L., Andreishcheva, E. N., Kelley, D. S., Shen, C.-C., et al. (2010). Archaea and bacteria with surprising microdiversity show shifts in dominance over 1,000-year time scales in hydrothermal chimneys. *Proc. Natl. Acad. Sci. U. S. A.* 107, 1612–1617.

Brazelton, W. J., McGonigle, J. M., Motamedi, S., Pendleton, H. L., Twing, K. I., Miller, B. C., et al. (2022). Metabolic strategies shared by basement residents of the lost city hydrothermal field. *Appl. Environ. Microbiol.* 88, e0092922.

Brazelton, W. J., Mehta, M. P., Kelley, D. S., and Baross, J. A. (2011). Physiological differentiation within a single-species biofilm fueled by serpentinization. *MBio* 2, 1–9.

Brazelton, W. J., Schrenk, M. O., Kelley, D. S., and Baross, J. A. (2006). Methane- and sulfur-metabolizing microbial communities dominate the Lost City hydrothermal field ecosystem. *Appl. Environ. Microbiol.* 72, 6257–6270.

Buchfink, B., Reuter, K., and Drost, H.-G. (2021). Sensitive protein alignments at tree-of-life scale using DIAMOND. *Nat. Methods* 18, 366–368.

Bushnell, B., Rood, J., and Singer, E. (2017). BBMerge – Accurate paired shotgun read merging via overlap. *PloS One* 12, e0185056.

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.

Chavagnac, V., Ceuleneer, G., Monnin, C., Lansac, B., Hoareau, G., and Boulart, C. (2013). Mineralogical assemblages forming at hyperalkaline warm springs hosted on ultramafic rocks: A case study of Oman and Ligurian ophiolites. *Geochem. Geophys. Geosyst.* 14, 2474–2495.

Coimbra, C., Farias, P., Branco, R., and Morais, P. V. (2017). Tungsten accumulation by highly tolerant marine hydrothermal Sulfitobacter dubius strains carrying a tupBCA cluster. *Syst. Appl. Microbiol.* 40, 388–395.

Colman, D. R., Poudel, S., Hamilton, T. L., Havig, J. R., Selensky, M. J., Shock, E. L., et al. (2018). Geobiological feedbacks and the evolution of thermoacidophiles. *ISME J.* 12, 225–236.

Corliss, J. B., Dymond, J., Gordon, L. I., Edmond, J. M., von Herzen, R. P., Ballard, R. D., et al. (1979). Submarine thermal sprirngs on the galapagos rift. *Science* 203, 1073–1083.

Demina, L. L., Lein, A. Y., Galkin, S. V., and Lisitzin, A. P. (2015). Features of trace metal distribution in the components of the ecosystem of the Lost City hydrothermal vent field (North Atlantic). *Dokl. Earth Sci.* 53, 224–240. doi: 10.1134/S1028334X1512017X

Ding, K., Seyfried, W. E., Zhang, Z., Tivey, M. K., Von Damm, K. L., and Bradley, A. M. (2005). The in *situ* pH of hydrothermal fluids at mid-ocean ridges. *Earth Planet. Sci. Lett.* 237, 167–174.

Dupraz, C., Visscher, P. T., Baumgartner, L. K., and Reid, R. P. (2004). Microbe–mineral interactions: early carbonate precipitation in a hypersaline lake (Eleuthera Island, Bahamas). *Sedimentology* 51, 745–765.

Edwards, K. J., Becker, K., and Colwell, F. (2012). The deep, dark energy biosphere: intraterrestrial life on earth. *Ann. Rev. Earth Planet. Sci.* 40 (1), 551-568. doi: 10.1146/annurev-earth-042711-105500

Fadel, M., Hassanein, N. M., Elshafei, M. M., Mostafa, A. H., Ahmed, M. A., and Khater, H. M. (2017). Biosorption of manganese from groundwater by biomass of Saccharomyces cerevisiae. *HBRC J.* 13, 106–113.

Fodelianakis, S., Washburne, A. D., Bourquin, M., Pramateftaki, P., Kohler, T. J., Styllas, M., et al. (2022). Microdiversity characterizes prevalent phylogenetic clades in the glacier-fed stream microbiome. *ISME J.* 16, 666–675.

Frouin, E., Bes, M., Ollivier, B., Quéméneur, M., Postec, A., Debroas, D., et al. (2018). Diversity of rare and abundant prokaryotic phylotypes in the prony hydrothermal field and comparison with other serpentinite-hosted ecosystems. *Front. Microbiol.* 9, 102.

Früh-Green, G. L., Kelley, D. S., Lilley, M. D., Cannat, M., Chavagnac, V., and Baross, J. A. (2022). Diversity of magmatism, hydrothermal processes and microbial interactions at mid-ocean ridges. *Nat. Rev. Earth Environ.* 3, 852–871.

Früh-Green, G. L., Orcutt, B. N., Rouméjon, S., Lilley, M. D., Morono, Y., Cotterill, C., et al. (2018). Magmatism, serpentinization and life: Insights through drilling the Atlantis Massif (IODP Expedition 357). *Lithos* 323, 137–155.

Fuchsman, C. A., Collins, R. E., Rocap, G., and Brazelton, W. J. (2017). Effect of the environment on horizontal gene transfer between bacteria and archaea. *PeerJ* 5, e3865.

Gadanho, M., and Sampaio, J. P. (2005). Occurrence and diversity of yeasts in the mid-atlantic ridge hydrothermal fields near the Azores Archipelago. *Microb. Ecol.* 50, 408–417.

García-García, N., Tamames, J., Linz, A. M., Pedrós-Alió, C., and Puente-Sánchez, F. (2019). Microdiversity ensures the maintenance of functional microbial communities under changing environmental conditions. *ISME J.* 13, 2969–2983.

Gaupp, R., Schlag, S., Liebeke, M., Lalk, M., and Götz, F. (2010). Advantage of upregulation of succinate dehydrogenase in Staphylococcus aureus biofilms. *J. Bacteriol.* 192, 2385–2394.

Gittins, D. A., Desiage, P.-A., Morrison, N., Rattray, J. E., Bhatnagar, S., Chakraborty, A., et al. (2022). Geological processes mediate a microbial dispersal loop in the deep biosphere. *Sci. Adv.* 8, eabn3485.

Graham, E. D., Heidelberg, J. F., and Tully, B. J. (2017). BinSanity: unsupervised clustering of environmental microbial assemblies using coverage and affinity propagation. *PeerJ* 5, e3035.

Harmsen, H., Prieur, D., and Jeanthon, C. (1997). Distribution of microorganisms in deep-sea hydrothermal vent chimneys investigated by whole-cell hybridization and enrichment culture of thermophilic subpopulations. *Appl. Environ. Microbiol.* 63, 2876–2883.

Hutchinson, G. E. (1961). The paradox of the plankton. *Am. Nat.* 95, 137–145.

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 Bioinf.* 11, 119.

Jebbar, M., Franzetti, B., Girard, E., and Oger, P. (2015). Microbial diversity and adaptation to high hydrostatic pressure in deep-sea hydrothermal vents prokaryotes. *Extremophiles* 19, 721–740.

Kanehisa, M., and Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 28, 27–30.

Kato, S., Nakamura, K., Toki, T., Ishibashi, J.-I., Tsunogai, U., Hirota, A., et al. (2012). Iron-based microbial ecosystem on and below the seafloor: a case study of hydrothermal fields of the southern mariana trough. *Front. Microbiol.* 3, 89.

Kelley, D. S., Karson, J. A., Früh-Green, G. L., Yoerger, D. R., Shank, T. M., Butterfield, D. A., et al. (2005). A serpentinite-hosted ecosystem: the Lost City hydrothermal field. *Science* 307, 1428–1434.

Kitadai, N., Nakamura, R., Yamamoto, M., Takai, K., Yoshida, N., and Oono, Y. (2019). Metals likely promoted protometabolism in early ocean alkaline hydrothermal systems. *Sci. Adv.* 5, eaav7848.

Krause, S., Liebetrau, V., Löscher, C. R., Böhm, F., Gorb, S., Eisenhauer, A., et al. (2018). Marine ammonification and carbonic anhydrase activity induce rapid calcium carbonate precipitation. *Geochim. Cosmochim. Acta* 243, 116–132.

Kristall, B., Kelley, D. S., and Hannington, M. D. (2006). Growth history of a diffusely venting sulfide structure from the Juan de Fuca Ridge: a petrological and geochemical study. *Geochem. Explor. Environ. Analy.* 7, Q07001. doi: 10.1029/2005GC001166

Lang, S. Q., and Benitez-Nelson, B. (2021). Hydrothermal Organic Geochemistry (HOG) sampler for deployment on deep-sea submersibles. *Deep Sea Res. Part 1 Oceanogr. Res. Pap.* 173, 103529.

Lang, S. Q., Früh-Green, G. L., Bernasconi, S. M., Brazelton, W. J., Schrenk, M. O., and McGonigle, J. M. (2018). Deeply-sourced formate fuels sulfate reducers but not methanogens at Lost City hydrothermal field. *Sci. Rep.* 8, 755.

Lang, S. Q., Lilley, M. D., Baumberger, T., Früh-Green, G. L., Walker, S. L., Brazelton, W. J., et al. (2021). Extensive decentralized hydrogen export from the Atlantis Massif. *Geology* 49, 851–856.

Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359.

Larkin, A. A., and Martiny, A. C. (2017). Microdiversity shapes the traits, niche space, and biogeography of microbial taxa. *Environ. Microbiol. Rep.* 9, 55–70.

Lecoeuvre, A., Ménez, B., Cannat, M., Chavagnac, V., and Gérard, E. (2020). Microbial ecology of the newly discovered serpentinite-hosted Old City hydrothermal field (southwest Indian ridge). *ISME J.* 15 (3), 818–832. doi: 10.1038/s41396-020-00816-7

León-Zayas, R., Peoples, L., Biddle, J. F., Podell, S., Novotny, M., Cameron, J., et al. (2017). The metabolic potential of the single cell genomes obtained from the Challenger Deep, Mariana Trench within the candidate superphylum Parcubacteria (OD1). *Environ. Microbiol.* 19, 2769–2784.

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.

Liang, J., Bai, Y., Men, Y., and Qu, J. (2017). Microbe-microbe interactions trigger Mn(II)-oxidizing gene expression. *ISME J.* 11, 67–77.

Ludwig, K. A., Kelley, D. S., Butterfield, D. A., Nelson, B. K., and Früh-Green, G. (2006). Formation and evolution of carbonate chimneys at the Lost City Hydrothermal Field. *Geochim. Cosmochim. Acta* 70, 3625–3645.

Ludwig, K. A., Shen, C.-C., Kelley, D. S., Cheng, H., and Edwards, R. L. (2011). U–Th systematics and 230Th ages of carbonate chimneys at the Lost City Hydrothermal Field. *Geochim. Cosmochim. Acta* 75, 1869–1888.

Martiny, J. B. H., Martiny, A. C., Brodie, E., Chase, A. B., Rodríguez-Verdugo, A., Treseder, K. K., et al. (2023). Investigating the eco-evolutionary response of microbiomes to environmental change. *Ecol. Lett.* 26 Suppl, 1, S81–1, S90.

McCollom, T. M., and Shock, E. L. (1997). Geochemical constraints on chemolithoautotrophic metabolism by microorganisms in seafloor hydrothermal systems. *Geochim. Cosmochim. Acta* 61, 4375–4391.

McGonigle, J. M., Lang, S. Q., and Brazelton, W. J. (2020). Genomic evidence for formate metabolism by chloroflexi as the key to unlocking deep carbon in lost city microbial ecosystems, appl. *Environ. Microbiol.* p, 831230.

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.

Meier, D. V., Pjevac, P., Bach, W., Hourdez, S., Girguis, P. R., Vidoudez, C., et al. (2017). Niche partitioning of diverse sulfur-oxidizing bacteria at hydrothermal vents. *ISME J.* 11, 1545–1558.

Mino, S., Kudo, H., Arai, T., Sawabe, T., Takai, K., and Nakagawa, S. (2014). Sulfurovum aggregans sp. nov., a hydrogen-oxidizing, thiosulfate-reducing chemolithoautotroph within the Epsilonproteobacteria isolated from a deep-sea hydrothermal vent chimney, and an emended description of the genus Sulfurovum. *Int. J. Syst. Evol. Microbiol.* 64, 3195–3201.

Mori, K., Yamaguchi, K., and Hanada, S. (2018). Sulfurovum denitrificans sp. nov., an obligately chemolithoautotrophic sulfur-oxidizing epsilonproteobacterium isolated from a hydrothermal field. *Int. J. Syst. Evol. Microbiol.* 68, 2183–2187.

Motamedi, S., Orcutt, B. N., Früh-Green, G. L., Twing, K. I., Pendleton, H. L., and Brazelton, W. J. (2020). Microbial residents of the atlantis massif's shallow serpentinite subsurface. *Appl. Environ. Microbiol.* 86 (11), e00356-20.

Mourgela, R. N., Kioukis, A., Pourjam, M., and Lagkouvardos, I. (2023). Large-scale integration of amplicon data reveals massive diversity within saprospirales, mostly originating from saline environments. *Microorganisms* 11, 1767. doi: 10.3390/microorganisms11071767

Mus, F., Colman, D. R., Peters, J. W., and Boyd, E. S. (2019). Geobiological feedbacks, oxygen, and the evolution of nitrogenase. *Free Radic. Biol. Med.* 140, 250–259.

Nakagawa, S., and Takai, K. (2008). Deep-sea vent chemoautotrophs: diversity, biochemistry and ecological significance. *FEMS Microbiol. Ecol.* 65, 1–14.

Nemergut, D. R., Anderson, S. P., Cleveland, C. C., Martin, A. P., Miller, A. E., Seimon, A., et al. (2007). Microbial community succession in an unvegetated, recently deglaciated soil. *Microb. Ecol.* 53, 110–122.

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.

Pisapia, C., Gérard, E., Gérard, M., Lecourt, L., Lang, S. Q., Pelletier, B., et al. (2017). Mineralizing filamentous bacteria from the prony bay hydrothermal field give new insights into the functioning of serpentinization-based subseafloor ecosystems. *Front. Microbiol.* 8, 57.

Postec, A., Quéméneur, M., Bes, M., Mei, N., Benaïssa, F., Payri, C., et al. (2015). Microbial diversity in a submarine carbonate edifice from the serpentinizing hydrothermal system of the Prony Bay (New Caledonia) over a 6-year period. *Front. Microbiol.* 6, 857.

Price, P. B., and Sowers, T. (2004). Temperature dependence of metabolic rates for microbial growth, maintenance, and survival. *Proc. Natl. Acad. Sci. U. S. A.* 101, 4631–4636.

Prieur, D. (1997). Microbiology of deep-sea hydrothermal vents. *Trends Biotechnol.* 15, 242–244.

Qin, W., Heal, K. R., Ramdasi, R., Kobelt, J. N., Martens-Habbena, W., Bertagnolli, A. D., et al. (2017). Nitrosopumilus maritimus gen. nov., sp. nov., Nitrosopumilus cobalaminigenes sp. nov., Nitrosopumilus oxyclinae sp. nov., and Nitrosopumilus ureiphilus sp. nov., four marine ammonia-oxidizing archaea of the phylum Thaumarchaeota. *Int. J. Syst. Evol. Microbiol.* 67, 5067–5079.

Quinlan, A. R., and Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841–842.

R Core Team (2022). *R: A language and environment for statistical computing* (Vienna, Austria: R Foundation for Statistical Computing). Available at: https://www.R-project.org/.

Resing, J. A., Sedwick, P. N., German, C. R., Jenkins, W. J., Moffett, J. W., Sohst, B. M., et al. (2015). Basin-scale transport of hydrothermal dissolved metals across the South Pacific Ocean. *Nature* 523, 200–203.

Ries, J. B., Anderson, M. A., and Hill, R. T. (2008). Seawater Mg/Ca controls polymorph mineralogy of microbial CaCO3: a potential proxy for calcite-aragonite seas in Precambrian time. *Geobiology* 6, 106–119.

Rios-Del Toro, E. E., Valenzuela, E. I., Ramírez, J. E., López-Lozano, N. E., and Cervantes, F. J. (2018). Anaerobic ammonium oxidation linked to microbial reduction of natural organic matter in marine sediments. *Environ. Sci. Technol. Lett.* 5, 571–577.

Rizopoulos, D. (2007). ltm: an R package for latent variable modeling and item response analysis. *J. Stat. Software* 17, 1–25.

Rohland, N., and Reich, D. (2012). Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. *Genome Res.* 22, 939–946.

Russell, M. J., and Hall, A. J. (1997). The emergence of life from iron monosulphide bubbles at a submarine hydrothermal redox and pH front. *J. Geol. Soc London* 154, 377–402.

Sabuda, M. C., Brazelton, W. J., Putman, L. I., McCollom, T. M., Hoehler, T. M., Kubo, M. D. Y., et al. (2020). A dynamic microbial sulfur cycle in a serpentinizing continental ophiolite. *Environ. Microbiol.* 22 (6), 2329-2345. doi: 10.1111/1462-2920.15006

Samuel, J., Paul, M. L., Pulimi, M., Nirmala, M. J., Chandrasekaran, N., and Mukherjee, A. (2012). Hexavalent chromium bioremoval through adaptation and consortia development from sukinda chromite mine isolates. *Ind. Eng. Chem. Res.* 51, 3740–3749.

Sander, S. G., and Koschinsky, A. (2011). Metal flux from hydrothermal vents increased by organic complexation. *Nat. Geosci.* 4, 145–150.

Schijf, J., Christenson, E. A., and Byrne, R. H. (2015). YREE scavenging in seawater: A new look at an old model. *Mar. Chem.* 177, 460–471.

Schrenk, M. O., Kelley, D. S., and Bolton, S. A. (2004). Low archaeal diversity linked to subseafloor geochemical processes at the Lost City Hydrothermal Field, Mid-Atlantic Ridge. *Environmentalist*. 6 (10), 1086-1095.

Schrenk, M. O., Kelley, D. S., Delaney, J. R., and Baross, J. A. (2003). Incidence and diversity of microorganisms within the walls of an active deep-sea sulfide chimney. *Appl. Environ. Microbiol.* 69, 3580–3592.

Scott, J. J., Glazer, B. T., and Emerson, D. (2017). Bringing microbial diversity into focus: high-resolution analysis of iron mats from the Lō'ihi Seamount. *Environ. Microbiol.* 19, 301–316.

Sheik, C. S., Reese, B. K., Twing, K. I., Sylvan, J. B., Grim, S. L., Schrenk, M. O., et al. (2018). Identification and removal of contaminant sequences from ribosomal gene databases: lessons from the census of deep life. *Front. Microbiol.* 9, 840.

Shenhav, L., Thompson, M., Joseph, T. A., Briscoe, L., Furman, O., Bogumil, D., et al. (2019). FEAST: fast expectation-maximization for microbial source tracking. *Nat. Methods* 16, 627–632.

Skennerton, C. T., Chourey, K., Iyer, R., Hettich, R. L., Tyson, G. W., and Orphan, V. J. (2017). Methane-Fueled Syntrophy through Extracellular Electron Transfer: Uncovering the Genomic Traits Conserved within Diverse Bacterial Partners of Anaerobic Methanotrophic Archaea. *MBio* 8, e00530–e00517.

Sojo, V., Herschy, B., Whicher, A., Camprubí, E., and Lane, N. (2016). The origin of life in alkaline hydrothermal vents. *Astrobiology* 16, 181–197.

Storey, J. D. (2002). A direct approach to false discovery rates. *J. R. Stat. Soc Ser. B Stat. Methodol.* 64, 479–498.

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

Sun, Q.-L., Zhang, J., Wang, M.-X., Cao, L., Du, Z.-F., Sun, Y.-Y., et al. (2020). High-throughput sequencing reveals a potentially novel sulfurovum species dominating the microbial communities of the seawater–sediment interface of a deep-sea cold seep in south China sea. *Microorganisms* 8, 687.

Sundaram, S., and Thakur, I. S. (2018). Induction of calcite precipitation through heightened production of extracellular carbonic anhydrase by CO2 sequestering bacteria. *Bioresour Technol.* 253, 368–371.

Suzuki, S., Ishii, S., Hoshino, T., Rietze, A., Tenney, A., Morrill, P. L., et al. (2017). Unusual metabolic diversity of hyperalkaliphilic microbial communities associated with subterranean serpentinization at The Cedars. *ISME J.* 11, 2584.

Suzuki, S., Ishii, S., Wu, A., Cheung, A., Tenney, A., Wanger, G., et al. (2013). Microbial diversity in The Cedars, an ultrabasic, ultrareducing, and low salinity serpentinizing ecosystem. *Proc. Natl. Acad. Sci. U. S. A.* 110, 15336–15341.

Takai, K., Komatsu, T., Inagaki, F., and Horikoshi, K. (2001). Distribution of archaea in a black smoker chimney structure. *Appl. Environ. Microbiol.* 67, 3618–3629.

Takashima, M., Shimada, K., and Speece, R. E. (2011). Minimum requirements for trace metals (iron, nickel, cobalt, and zinc) in thermophilic and mesophilic methane fermentation from glucose. *Water Environ. Res.* 83, 339–346.

Tarlera, S., Jangid, K., Ivester, A. H., Whitman, W. B., and Williams, M. A. (2008). Microbial community succession and bacterial diversity in soils during 77,000 years of ecosystem development. *FEMS Microbiol. Ecol.* 64, 129–140.

Thornton, C. N., Tanner, W. D., VanDerslice, J. A., and Brazelton, W. J. (2020). Localized effect of treated wastewater effluent on the resistome of an urban watershed. *Gigascience* 9 (11), giaa125.

Tutolo, B. M., Luhmann, A. J., Tosca, N. J., and Seyfried, W. E. (2018). Serpentinization as a reactive transport process: The brucite silicification reaction. *Earth Planet. Sci. Lett.* 484, 385–395.

Vasudevan, P., Padmavathy, V., and Dhingra, S. C. (2002). Biosorption of monovalent and divalent ions on baker's yeast. *Bioresour. Technol.* 82, 285–289.

Verma, D., Kumar, V., and Satyanarayana, T. (2022). Genomic attributes of thermophilic and hyperthermophilic bacteria and archaea. *World J. Microbiol. Biotechnol.* 38, 135.

Wang, F., Zhou, H., Meng, J., Peng, X., Jiang, L., Sun, P., et al. (2009). GeoChip-based analysis of metabolic diversity of microbial communities at the Juan de Fuca Ridge hydrothermal vent. *Proc. Natl. Acad. Sci. U. S. A.* 106, 4840–4845.

Wei, S., Cui, H., Jiang, Z., Liu, H., He, H., and Fang, N. (2015). Biomineralization processes of calcite induced by bacteria isolated from marine sediments. *Braz. J. Microbiol.* 46, 455–464.

Wirth, R. (2017). Colonization of black smokers by hyperthermophilic microorganisms. *Trends Microbiol.* 25, 92–99.

Zhong, H., Lehtovirta-Morley, L., Liu, J., Zheng, Y., Lin, H., Song, D., et al. (2020). Novel insights into the Thaumarchaeota in the deepest oceans: their metabolism and potential adaptation mechanisms. *Microbiome* 8, 78.

Zhong, Y.-W., Zhou, P., Cheng, H., Zhou, Y.-D., Pan, J., Xu, L., et al. (2022). Metagenomic features characterized with microbial iron oxidoreduction and mineral interaction in southwest Indian ridge. *Microbiol. Spectr.* 10, e0061422.

Zwicker, J., Smrzka, D., Vadillo, I., Jiménez-Gavilán, P., Giampouras, M., Peckmann, J., et al. (2022). Trace and rare earth element distribution in hyperalkaline serpentinite-hosted spring waters and associated authigenic carbonates from the Ronda peridotite. *Appl. Geochemistry* 147, 105492.

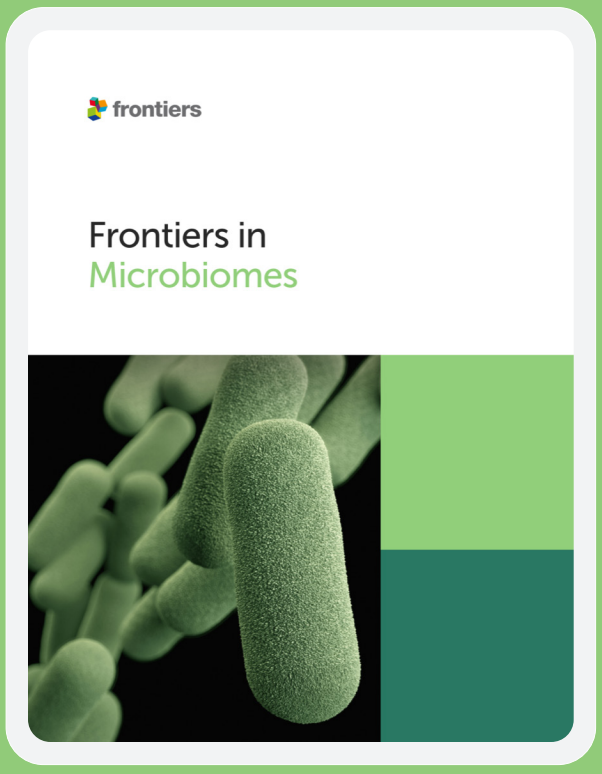
Frontiers in Microbiomes

Explores impactful connections between microbiomes and the world around us

Advances our understanding of how microbiomes generate positive or negative outcomes for their hosts and environments.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact