frontiers | Research Topics

Rising Stars in aquatic microbiology 2022

Edited by
Hongbin Liu and Tony Gutierrez

Published in
Frontiers in Microbiology
Frontiers in Marine Science

ISSN 1664-8714
ISBN 978-2-8325-3370-3
DOI 10.3389/978-2-8325-3370-3

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

Rising stars in aquatic microbiology: 2022

Topic editors

Hongbin Liu — Hong Kong University of Science and Technology, Hong Kong, SAR China

Tony Gutierrez — Heriot-Watt University, United Kingdom

Citation

Liu, H., Gutierrez, T., eds. (2023). *Rising stars in aquatic microbiology: 2022*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-3370-3

Table of contents

frontiers | Frontiers in Microbiology

TYPE Editorial
PUBLISHED 15 August 2023
DOI 10.3389/fmicb.2023.1265720

OPEN ACCESS

EDITED AND REVIEWED BY
Michael Rappe,
University of Hawaii at Manoa, United States

*CORRESPONDENCE
Tony Gutierrez
✉ tony.gutierrez@hw.ac.uk

RECEIVED 23 July 2023
ACCEPTED 31 July 2023
PUBLISHED 15 August 2023

CITATION
Gutierrez T and Liu H (2023) Editorial: Rising stars in aquatic microbiology: 2022. *Front. Microbiol.* 14:1265720. doi: 10.3389/fmicb.2023.1265720

Editorial: Rising stars in aquatic microbiology: 2022

Tony Gutierrez[1]* and Hongbin Liu[2]

[1]Institute of Mechanical Process and Energy Engineering (IMPEE), School of Engineering and Physical Sciences, Heriot-Watt University, Edinburgh, United Kingdom, [2]Department of Ocean Science, Hong Kong University of Science and Technology, Kowloon, Hong Kong SAR, China

KEYWORDS

environmental microbiology, microbial diversity, microbial ecology, ecosystem function analysis, metagenomics

Editorial on the Research Topic
Rising stars in aquatic microbiology: 2022

In this Research Topic, we recognize the work of future leaders in the field of Aquatic Microbiology by presenting a collection of articles that showcases the high-quality work of "rising stars" that are in the early stages of their careers and on a trajectory to potentially becoming internationally recognized leaders in their field. We aim to highlight research by these upcoming leading scientists of the future across the entire breadth of Aquatic Microbiology, with advances in theory, experimental design, and methodology with applications for compelling problems. We trust that this Research Topic will give a hint of who to follow in the coming years.

In this Research Topic, five articles highlight new findings on the functions and seasonal and ecosystem effects of microbial communities in aquatic systems. To begin, the article by Chen et al. investigates the seasonal patterns, ecological function, and assembly processes of bacterioplankton communities in the Danjiangkou Reservoir (DJR), China. Using high-throughput sequencing, PICRUSt2, and other methods, this study reveals the composition, function, interaction, and assembly of bacterioplankton communities in the DJR, providing a reference for the protection of water quality and the ecological functions of DJR bacterioplankton. In the article by Summers et al., the authors assess how different stone materials that are used in seawall construction (ranging from aluminosilicates to limestone and concrete) affect biofilm formation in tropical marine waters. Their results show bacterial biofilm colonization and succession are similar across different types of stone materials, suggesting that marine biofilms converged over time to a generic marine biofilm, rather than the underlying stone substrata type playing a significant role in driving community composition. The article by Brenes-Guillén et al. reports new insights into the prokaryotic diversity and community structure, in conjunction with physicochemistry along vertical gradients, during stratification and mixing periods in Lake Cote, Costa Rica, and provides an example of lake hydrodynamic impacts on microbial communities and their function in Central American lakes with implications for other shallow, upland, and oligotrophic lake systems. Zhao et al. report on the biogeographical patterns and assembly processes of the abundant and rare bacteria from river sediment at high altitudes (Lhasa River, China) and their potential association with antibiotic resistance genes (ARGs).

Their results reveal that abundant bacteria shaped by deterministic processes have a high abundance of potential ARGs in the plateau river sediment. The article by Puigcorbé et al. shows how the vertical connectivity of the prokaryotic assemblages associated with particles of three different sizes at two East Antarctic polynyas with different surface productivity is linked to the magnitude of the particle export fluxes measured using thorium-234 (^{234}Th) as particle tracer. Whilst the results support recent studies evidencing links between surface productivity and deep prokaryotic communities, this study provides the first evidence of sinking particles acting as vectors of microbial diversity to depth in Antarctic polynyas and highlights the direct influence of particle export in shaping the prokaryotic communities of mesopelagic waters.

A couplet of review articles (Tian et al.; Singh et al.) focus on the microbiology of wastewater treatment plants (WWTPs). The article by Tian et al. presents a critical overview of the existing knowledge of bioaerosol emissions from WWTPs, including their nature, magnitude, and size distribution, and highlights shortcomings associated with existing sampling and analysis methods. The mini-review by Singh et al. provides an appraisal of the different types and sources of coliphage and their fate and behavior in source waters and engineered drinking water treatment systems, and provides exciting future prospects for the use of coliphages in aquatic microbiology based on current scientific evidence and practical needs.

Body size is an important ecological trait, but it has been poorly explored in microbial communities. The article by Wu and Liu examines the effect of cell size on coastal eukaryotic communities across a size continuum of 0.2–3 (pico-), 3–20 (nano-), and 20–200 mm (micro-size), and their findings suggest that the cell size of microbial eukaryotes is a phylogenetically conserved trait which is tightly associated with biogeographic patterns.

In the study by Xia et al., comparative genomics was used to reveal a new strain of flavobacteria, *Tamlana* sp. S12, which can potentially degrade complex algal polysaccharides, thus expanding our knowledge on how marine *Flavobacteriaceae* adapt to marine algal polysaccharide environments.

In summary, these articles provide a snapshot of what is to come with respect to more high-quality work from these "rising stars" across the wide breadth of Aquatic Microbiology. We hope that this Research Topic will also inspire the younger "clan" of early career researchers to what may be possible in uncovering more about our enigmatic and incredibly profound world of microbes.

Author contributions

TG: Conceptualization, Funding acquisition, Validation, Writing—original draft, Writing—review and editing. HL: Validation, Writing—review and editing.

Funding

Preparation of this manuscript was made possible by a NERC Global Challenges Research Fund (NE/V006088/1) to TG.

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

frontiers | Frontiers in Microbiology

ORIGINAL RESEARCH
published: 15 June 2022
doi: 10.3389/fmicb.2022.884765

The Seasonal Patterns, Ecological Function and Assembly Processes of Bacterioplankton Communities in the Danjiangkou Reservoir, China

Zhao-Jin Chen[1], Yong-Qi Liu[1], Yu-Ying Li[1], Li-An Lin[1], Bao-Hai Zheng[1], Ming-Fei Ji[1], B. Larry Li[2] and Xue-Mei Han[3]**

[1] International Joint Laboratory of Watershed Ecological Security and Collaborative Innovation Center of Water Security for Water Source Region of Middle Route Project of South-North Water Diversion in Henan Province, School of Water Resource and Environmental Engineering, Nanyang Normal University, Nanyang, China, [2] Ecological Complexity and Modelling Laboratory, Department of Botany and Plant Sciences, University of California, Riverside, Riverside, CA, United States, [3] Ministry of Education Key Laboratory for Ecology of Tropical Islands, College of Life Sciences, Hainan Normal University, Haikou, China

OPEN ACCESS

Edited by:
Hongbin Liu,
Hong Kong University of Science and Technology, Hong Kong SAR, China

Reviewed by:
Lunguang Yao,
Nanyang Normal University, China
Shengwei Hou,
Southern University of Science and Technology, China
Henglun Shen,
Zaozhuang University, China

***Correspondence:**
Yu-Ying Li
lyying200508@163.com
Xue-Mei Han
hanxuemei916@163.com

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

Received: 27 February 2022
Accepted: 24 May 2022
Published: 15 June 2022

Citation:
Chen Z-J, Liu Y-Q, Li Y-Y, Lin L-A, Zheng B-H, Ji M-F, Li BL and Han X-M (2022) The Seasonal Patterns, Ecological Function and Assembly Processes of Bacterioplankton Communities in the Danjiangkou Reservoir, China. Front. Microbiol. 13:884765. doi: 10.3389/fmicb.2022.884765

As the water source for the Middle Route Project of the South-to-North Water Diversion Project (MR-SNWD) of China, the Danjiangkou Reservoir (DJR) is in the process of ecosystem reassembly, but the composition, function, and assembly mechanisms of bacterioplankton communities are not yet clear. In this study, the composition, distribution characteristics and influencing factors of bacterioplankton communities were analyzed by high-throughput sequencing (HTS); PICRUSt2 was used to predict community function; a molecular ecological network was used to analyze bacterioplankton interactions; and the assembly process of bacterioplankton communities was estimated with a neutral model. The results indicated that the communities, function and interaction of bacterioplankton in the DJR had significant annual and seasonal variations and that the seasonal differences were greater than that the annual differences. Excessive nitrogen (N) and phosphorus (P) nutrients in the DJR are the most important factors affecting water quality in the reservoir, N and P nutrients are the main factors affecting bacterial communities. Season is the most important factor affecting bacterioplankton N and P cycle functions. Ecological network analysis indicated that the average clustering coefficient and average connectivity of the spring samples were lower than those of the autumn samples, while the number of modules for the spring samples was higher than that for the autumn samples. The neutral model explained 66.3%, 63.0%, 63.0%, and 70.9% of the bacterioplankton community variations in samples in the spring of 2018, the autumn of 2018, the spring of 2019, and the autumn of 2019, respectively. Stochastic processes dominate bacterioplankton community assembly in the DJR. This study revealed the composition, function, interaction, and assembly of bacterioplankton communities in the DJR, providing a reference for the protection of water quality and the ecological functions of DJR bacterioplankton.

Keywords: Danjiangkou Reservoir, bacterioplankton communities, seasonal variations, ecological network analysis, neutral model

INTRODUCTION

The Danjiangkou Reservoir (DJR) is the water source for the MR-SNWD of China (Liu et al., 2018). To meet the demand of MR-SNWD water transfer, the water level in the DJR was raised from 157 m to 170 m by increasing the height of the dam in 2013, and the inundation area was increased by 302.5 km^2 (Shu et al., 2017). The water flow direction in the DJR changed dramatically, from the original outflow from the Danjiangkou Dam to the diversion of water at the head gate of the Taocha Channel to the north, eventually reaching Beijing and Tianjin. Therefore, the DJR ecosystem is in the process of reconstruction, providing an excellent location for the study of bacterioplankton composition, function, interaction and community assembly in large reservoirs (Pan et al., 2018).

Bacterioplankton, as the main component of the aquatic community, plays an important role in the matter cycling in freshwater bodies and drive chemical element cycles in the entire ecosystem (Ducklow et al., 1986; Grossart, 2010). In addition, bacterioplankton are sensitive to water quality and environment changes: changes in factors such as the content of nutrients in different forms and the physicochemical indicators of the freshwater body will influence bacterioplankton community structure (Morris and Lewis, 1992; Mohapatra et al., 2020). Therefore, analyzing variations in bacterioplankton community structure in a reservoir and the difference in the degree of response to the physical and chemical properties can well reflect the water environment of the reservoir and can be used as an important indicator for analyzing the health of the reservoir ecosystem. Currently, few studies have been carried out on the composition of the bacterioplankton communities in the DJR and the associated influencing factors. In our previous study, HTS was employed for bacterioplankton community composition and influencing factors in the DJR, found that the bacterioplankton community were composed of 27 phyla and 336 genera and that TN, pH, COD, and COD_{Mn} can significantly affect bacterioplankton community composition (Chen Z. J. et al., 2020). Sun et al. (2021) studied the composition of bacterioplankton in the upstream river of the DJR and reported that environmental parameters such as pH, TN, Cond, and NH_4^+–N significantly affected the composition of the bacterioplankton communities. The dry season at the DJR occurs from February to July, and the wet season occurs from August to January of the following year. The physical and chemical properties of the water body (temperature, nutrients, etc.) vary greatly during the different periods. DJR bacterioplankton may exhibit yearly and seasonal variation, but currently, there is no comprehensive comparative study on different years and seasons. The DJR area carries relatively large nitrogen and phosphorus loads, of which TN, the most important factor affecting water quality, significantly exceeds the standard (Shu et al., 2017; Chen et al., 2018). Bacterioplankton is a principal contributor to the N and P cycles. As noted by Zhang L. et al. (2021) in their study of bacterial communities in the MR-SNWD main water diversion canal, attention will be focusing on microbial communities that help remove N and P in the water body. Furthermore, microbial communities play an important role in ecological processes through direct and indirect interactions. Ecological network analysis technology has been applied in studies of bacterioplankton interactions (Kara et al., 2013; Chen Z. J. et al., 2020; Zhang J. et al., 2021). However, there is still little information regarding the interaction pattern among species of bacterioplankton in the DJR. The ecological network of the DJR bacterioplankton community was constructed using bioinformatics analysis, and key species were identified. These data are of great significance for analyzing and predicting the survival patterns of bacterioplankton communities in the DJR ecosystem. The mechanisms of biome assembly have always been one of the core issues of ecology, among which the neutral theory has become a research hotspot due to its simplicity and predictive ability and has been widely applied in terrestrial ecosystems such as forests and grasslands; neutral processes can play an important role in shaping biomes (Zhou and Ning, 2017; Zhang et al., 2018). The neutral theory has also been applied to bacterioplankton community assembly in aquatic ecosystems. Chen et al. (2019) showed that stochastic processes dominate microeukaryotic community assembly in the Tingjiang River. Zhang L. et al. (2021) showed that the bacterial communities in the MR-SNWD main canal were mainly shaped by a deterministic process and that stochasticity dominated microeukaryotic community assembly. Currently, there are no reports on bacterioplankton community assembly in the DJR, thus requiring attention.

Because the DJR is related to the safe operation of the MR-SNWD, studies on the composition, function and assembly of bacterioplankton communities are of great ecological significance. Based on the current research status for this topic, we pose the following research questions: (1) Are there yearly and seasonal variations in the composition and function (especially the N and P cycles) of the bacterioplankton communities in the DJR? (2) How do bacterioplankton communities interact with each other in the DJR, and do the ecological networks and core microbiome change between different years and seasons? (3) How does bacterioplankton community assembly occur in the DJR? Is the assembly determined by deterministic or stochastic processes?

MATERIALS AND METHODS

Study Area and Sample Collection

According to geographic location, we set up 11 ecological sites in the DJR: Dashiqiao (DS), Zhangying (ZY), Heijizui (HJ), Kuxin (KX), Songgang (SG), Taizishan (TZ), Qushou (QS), Ganqu (GQ), Bashang (BS), Baxia (BX), Langhekou (LH) (**Figure 1**). The dry season at the DJR occurs from February to July, and the wet season occurs from August to January of the following year. We collected samples in the spring of 2018 (May, 2018S), the autumn of 2018 (October, 2018A), the spring of 2019 (May, 2019S), and the autumn of 2019 (October, 2019A). At each site, three replicate water samples were collected from surface water (0–50 cm). The water samples were processed for subsequent DNA extraction of

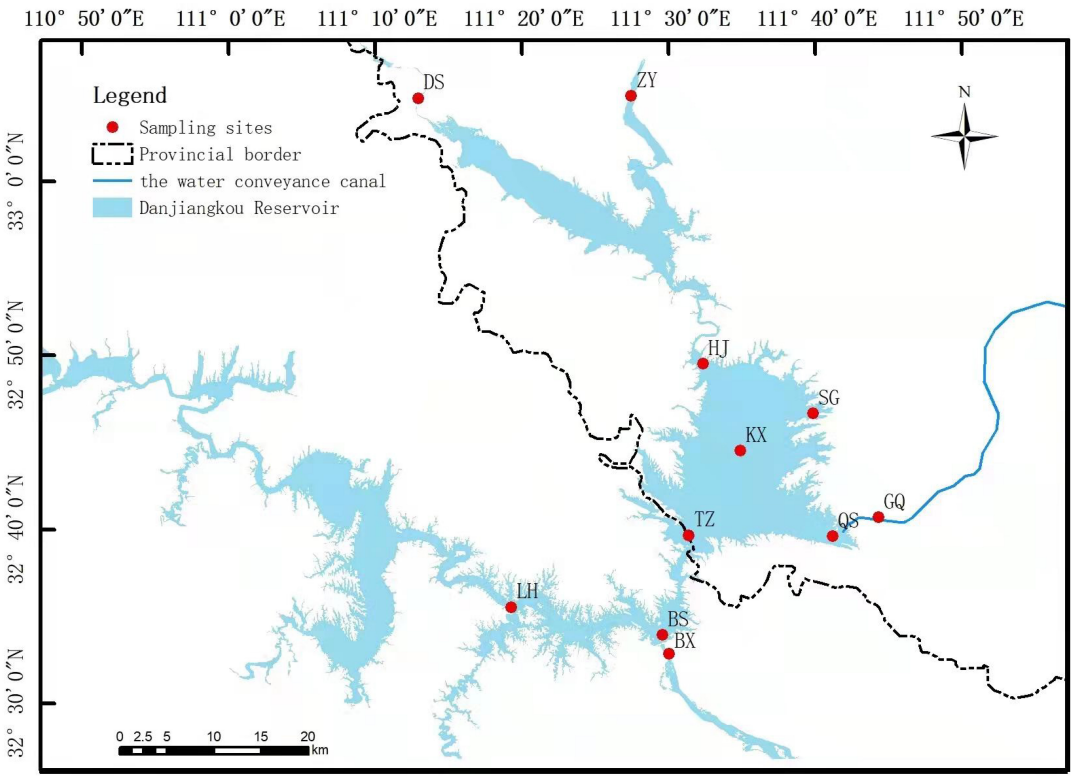


FIGURE 1 | Location of study area and distribution of sampling sites.

bacterioplankton and physicochemical analysis of water quality. 0.22-μm filter paper was used to collect bacterioplankton by filter 1 L of water samples, and the filter were stored in a −80°C refrigerator.

Physicochemical variables were measured according to the environmental quality standard for surface water of China (GB3838-2002). Water temperature (T), pH and dissolved oxygen (DO) were measured *in situ* using the YSI 6920 (YSI Inc., Yellow Springs, Ohio, United States). Secchi depth (SD) was determined with a 30-cm-diameter Secchi disk. Water samples for chemical analysis were transported to the laboratory within 24 h, stored in a refrigerator at 4 °C, and analyzed within one week after sample collection. The permanganate index (COD_{Mn}) was calculated using the potassium permanganate index method, and the chemical oxygen demand (COD) was measured by the potassium dichromate method. The total phosphorus (TP) was determined with acidified molybdate to form reduced phosphorus-molybdenum blue and measured spectrophotometrically. Total nitrogen (TN) was assayed *via* alkaline persulfate digestion and UV spectrophotometry, whereas ammonia nitrogen (NH_4^+–N) was measured using Nessler's reagent spectrophotometric method. Chlorophyll a (Chl a) concentrations were estimated spectrophotometrically after extraction in 90% ethanol.

The trophic status of the Danjiangkou Reservoir area was assessed by measuring the parameters TN, TP, COD_{Mn}, Chl a, and SD according to the improved Carlson's trophic level index (TLI) (Wang et al., 2002; Chen Z. J. et al., 2020).

DNA Extraction and Sequencing

DNA extraction from each filter was performed with the E.Z.N.A.® Water DNA Kit (OMEGA, United States) following the instructions given by the supplier. The DNA samples were sent to Shanghai Majorbio Bio-Pharm Technology Co., Ltd. and sequenced (2 × 300) on an Illumina MiSeq platform. For the amplification of the 16S rRNA gene, the specific primers 338F (ACTCCTACGGGAGGCAGCA) and 806R (GGACTACHVGGGTWTCTAAT) were used for high-throughput pyrosequencing. Samples were amplified on a T100 thermal cycler (Bio-Rad Laboratories). After pyrosequencing, the raw data was filtered according to barcode and primer sequences using the software of Trimmomatic v0.39 and FLASH v1.2.11. Then the high-quality sequences were processed using the using QIIME 2 (Bolyen et al., 2019). Non-repeating sequences were extracted from the optimized sequences using UPARSE v7.0.1090 (DeSantis et al., 2006). The bacterial sequences were identified and clustered into OTUs (97% similarity) by using UCLUST (version 7.1[1]) method (Edgar, 2010). The high-thoughput sequencing data were deposited in the MG-RAST[2] under accession number of mgp101864.

[1]http://drive5.com/uparse/

[2]http://www.mg-rast.org/

Bioinformatic Analysis

We performed beta diversity analysis online using the free online platform of Majorbio Cloud Platform.[3] The Unweighted pair-group method with arithmetic mean (UPGMA), partial least squares-discriminant analysis (PLS-DA), non-metric multidimensional scaling (NMDS), redundancy analysis (RDA), canonical correspondence analysis (CCA), variation partition analysis (VPA), correlation test and mantel test were conducted on this platform (Ren et al., 2022). Phylogenetic molecular ecological networks (pMENs) analysis was performed using the Molecular Ecological Network Analyses Pipeline (MENAP)[4] (Zhou J. Z. et al., 2011; Deng et al., 2012). The metagenomes predicted by the Phylogenetic investigation of communities by reconstruction of unobserved states (PICRUSt2) algorithm were classified into clusters of orthologous groups (COGs) (Douglas et al., 2020). In this study, PICRUSt2 was used to explore the functional profiles of the bacterial communities according to the online protocol. Heat map were generated from the gene copy number of the functional genes using the TBtools software (Chen C. et al., 2020). To determine the potential importance of stochastic processes on community assembly, we used a neutral community model (NCM) to predict the relationship between OTU detection frequency and their relative abundance across the wider metacommunity. In this model, Nm is an estimate of dispersal between communities. The parameter Nm determines the correlation between occurrence frequency and regional relative abundance, with N describing the metacommunity size and m being the immigration rate. The parameter R^2 represents the overall fit to the neutral model. Calculation of 95% confidence intervals around all fitting statistics was done by bootstrapping with 1000 bootstrap replicates.

RESULTS

Physical and Chemical Properties of Water

Except for TN and COD_{Mn}, the water quality of the DJR is generally good based on China's *Environmental Quality Standards for Surface Water* (GB38382-2002) and, overall, meets the requirements of class I water standards. In 2018S, 2018A, 2019S and 2019A, the TN contents were high, with average contents of 1.75, 1.80, 1.63 and 1.96 mg/L, respectively, concentrations that were higher than the water quality standards for class IV surface water, with trends of higher concentrations in autumn than in spring (**Supplementary Table 1**). The COD_{Mn} was similar to TN, with average concentrations of 2.60, 2.72, 2.72, and 2.96 mg/L in 2018S, 2018A, 2019S, and 2019A, respectively, concentrations that met the class II surface water standards, with trends of higher concentration in autumn than in spring.

[3]www.majorbio.com

[4]http://ieg4.rccc.ou.edu/mena/login.cgi

Bacterioplankton Composition and Yearly and Seasonal Variations

High-throughput sequencing (HTS) results indicated that the DJR bacterioplankton comprised the phyla Proteobacteria, Actinobacteria, Bacteroidetes, Cyanobacteria, Firmicutes, Verrucomicrobia, and Armatimonadetes, of which Proteobacteria, Actinobacteria and Bacteroidetes accounted for 71.78%∼96.98% of the total population (**Figure 2**). At the genus level, *CL500-29_marine_group*, *Acinetobacter*, *hgcI_clade*, *Limnohabitans*, *Cyanobium_PCC-6307*, *Flavobacterium*, *Brevundimonas*, *Sediminibacterium*, and *Exiguobacterium* accounted for 27.38% ∼ 96.60% of the bacterioplankton population (**Supplementary Figure 1**).

A dilution curve was used to evaluate the sequencing depth of the samples. The results indicated greater than 20,000 sample bands, with a dilution curve that tended to be flat (**Figure 3A**). UPGMA clustering tree analysis and PLS-DA were used to analyze the bacterioplankton community differences among different samples. In the UPGMA clustering tree, 2018S and 2019S were clustered in the lower part and separated from the 2018A and 2019A samples (**Figure 3B**). The PLS-DA analysis results were similar to the UPGMA clustering tree results. In the PLS-DA plot, 2018S and 2019S were distributed on the right side of the plot, and 2018A and 2019A were distributed on the left side of the plot, indicating that compared with different years, different seasons had a greater impact on the bacterioplankton communities and was the most important factor (**Figure 3C**). Additionally, comparing different seasons, the distances in the PLS-DA plots for 2018A and 2019A were greater than those for 2018S and 2019S, indicating that the bacterioplankton communities in the DJR varied in different seasons and that the variations in autumn were larger than those in spring. Samples from different seasons were clustered together in the UPGMA and PLS-DA plots, but the DS and ZY samples were poorly clustered. Based on the distribution of the sampling points, these two points were located on the tributaries of the reservoir, and thus, the bacterioplankton communities were different from those in the reservoir. Adonis and ANOISM were used to test the differences in the overall composition of the bacterioplankton communities (**Supplementary Table 2**). The results indicated that the differences in the composition of the bacterioplankton communities in different years and in different seasons were significant ($P < 0.05$). The number of OTUs in the samples at different time points was analyzed using a Venn diagram. The results indicated that the overall number of bacterial OTUs in the autumn was higher than that in the spring and higher in 2019 than in 2018 (**Figure 3D**).

Influencing Factors of Bacterioplankton Communities

First, the environmental factors with a variance inflation factor (VIF) > 10 were screened and removed by VIF analysis, and the screened environmental factors were used for RDA or CCA. The effects of environmental factors in different years and seasons on the bacterioplankton communities were analyzed by RDA or CCA. ORP, TP, TN, NH_4^+-N and SD

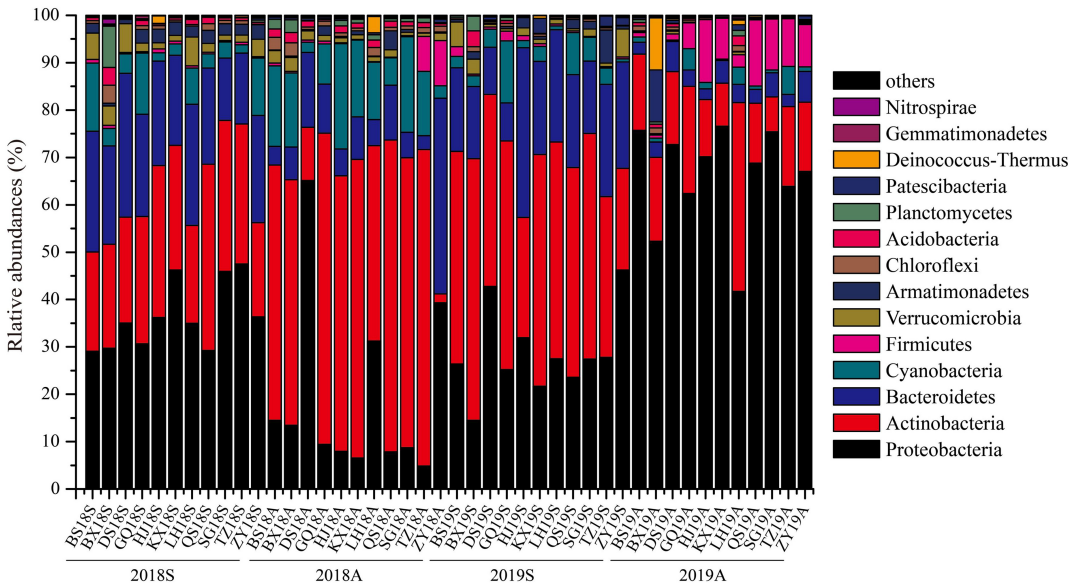


FIGURE 2 | Relative abundance of bacterioplankton sequences at the phylum.

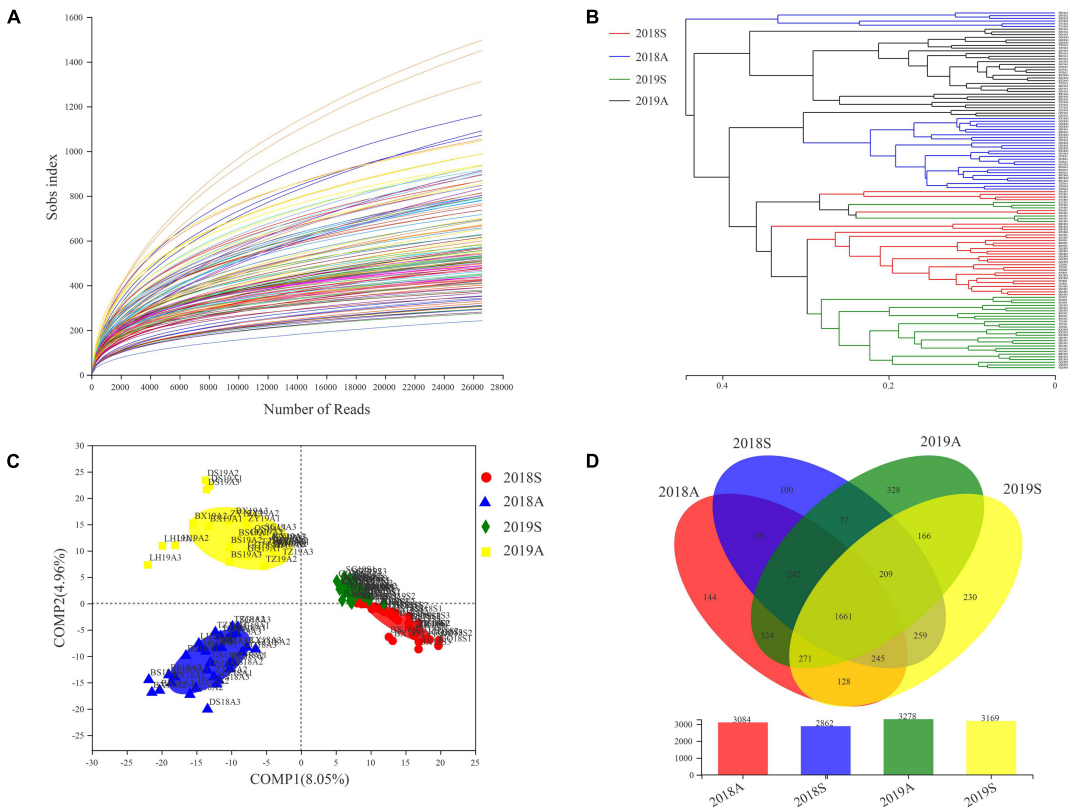


FIGURE 3 | Dilution curve **(A)**, UPGMA-tree **(B)**, PLS-DA **(C)** and Venn diagram **(D)** base on pyrosequencing of bacterioplankton communities.

were the significant factors affecting the 2018S bacterioplankton communities (**Figure 4A**). pH, COD_{Mn}, TN and Chla were the significant factors affecting the 2018A bacterioplankton communities (**Figure 4B**). T, pH, DO, COD, COD_{Mn}, TN, NH_4^+-N and Chla were the significant factors affecting the 2019S bacterioplankton communities (**Figure 4C**). Cond, ORP, COD_{Mn}, NH_4^+-N, and NO_3^--N were the significant factors affecting the 2019A bacterioplankton communities (**Figure 4D**).

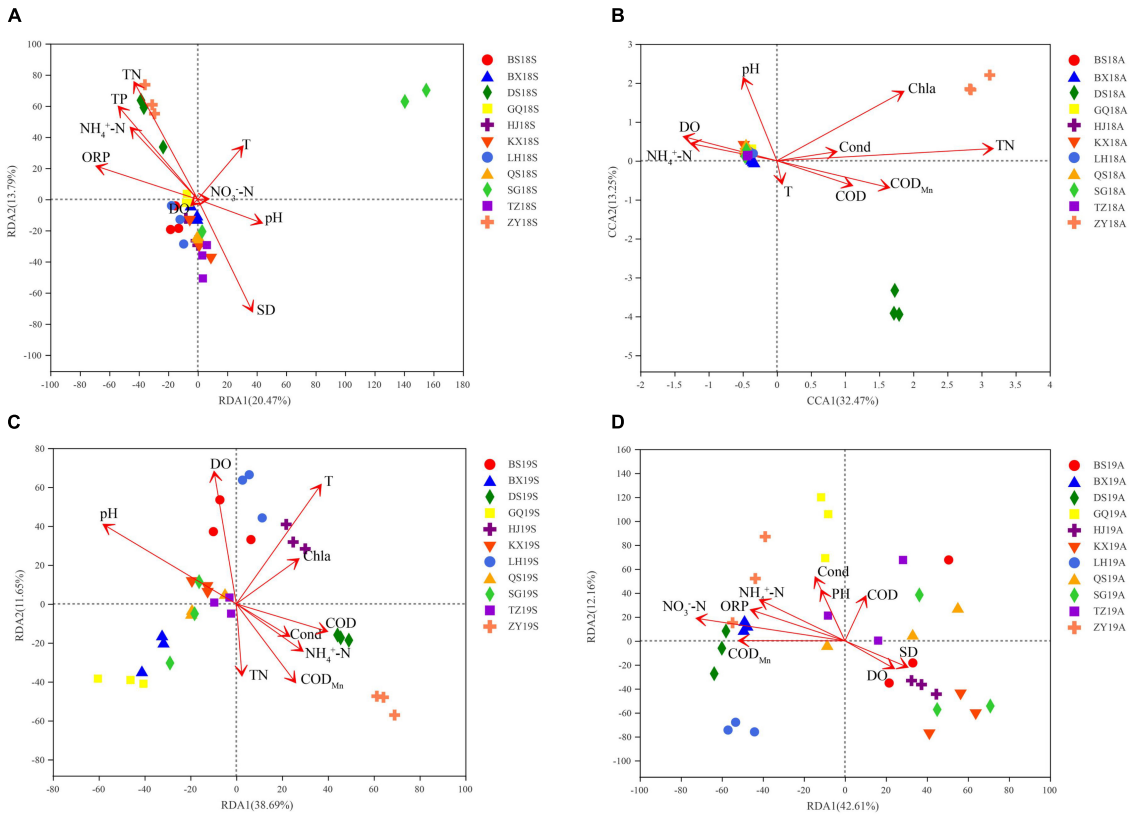


FIGURE 4 | RDA, CCA of 2018S **(A)**, 2018A **(B)**, 2019S **(C)**, 2019A **(D)** bacterioplankton communities of and physico-chemical water quality parameters.

The above results suggest that the composition and degree of impact of the significant environmental factors vary between different years and seasons and that the relationship between the DJR bacterioplankton communities and environmental factors exhibits annual and seasonal variations. Excessive N and P nutrients in the DJR are the most important factors affecting water quality in the reservoir. N and P nutrients are also the main factors affecting the bacterial communities. The effects of N and P nutrients (TN, NH_4^+-N, NO_3^--N, TP) and other environmental factors on the changes in bacterial community composition were analyzed by VPA. N and P nutrients (TN, NH_4^+-N, NO_3^--N, TP) explained 16.12%, 10.59%, 32.46%, and 24.93% of the variation in bacterial community composition in 2018S, 2018A, 2019S, and 2019A, respectively, indicating a strong interaction of N and P nutrients in shaping the microbial communities in the reservoir.

Bacterioplankton Functions

To assess the functions of bacterioplankton at different site in the DJR, PICRUSt2 software was used to perform microbiota predictions and analyses. The prediction results, based on the COG database, included a total of 24 functional gene families, of which six functional gene families, such as amino acid transport and metabolism and translation, Translation, ribosomal structure and biogenesis, Energy production and conversion, Cell wall/membrane/envelope biogenesis, Inorganic ion transport and metabolism, Carbohydrate transport and metabolism were the main functional gene families, accounting for 40.29%–45.76%. We analyzed the functional genes related to the N and P cycles in bacterioplankton samples from the DJR for different years and seasons. The results indicated that such genes were predominantly involved in nitrogen fixation (K02588 *nifH*), nitrification (K10535 *hao*), denitrification (K00368 *nirK*, K04561 *norB*, and K00376 *nosZ*), nitrogen assimilation reduction and dissimilarity reduction (K02575 *nasA*, K00367 *narB*, K02567 *napA*, K00366 *nirA*, K00362 *nirB*, and K03385 *nrfA*) and other related nitrogen cycle function genes. The results of cluster analysis of the copy number of nitrogen cycle genes indicated two separate groups in spring and autumn, indicating that season was the most important factor affecting the bacterioplankton N cycle (**Figure 5A**). The 2018 and 2019 samples were separated from each other in different seasons, indicating that year was also an important influencing factor. The predicted key genes of the P cycle were K00655 *plsC,* K01507 *ppa,* K02036 *pstB,* K02037 *pstC,* K02038 *pstA,* K00324 *pntA,* K06217 *phoH,* K03820 *lnt,* and K07636 *phoR.* The results of the gene copy number cluster analysis were similar to those for the N cycle, i.e., season and year were the main influencing factors (**Figure 5B**).

Ecological Network Analysis

The relative abundance of bacterioplankton was used to construct a bacterioplankton molecular ecological network for the DJR

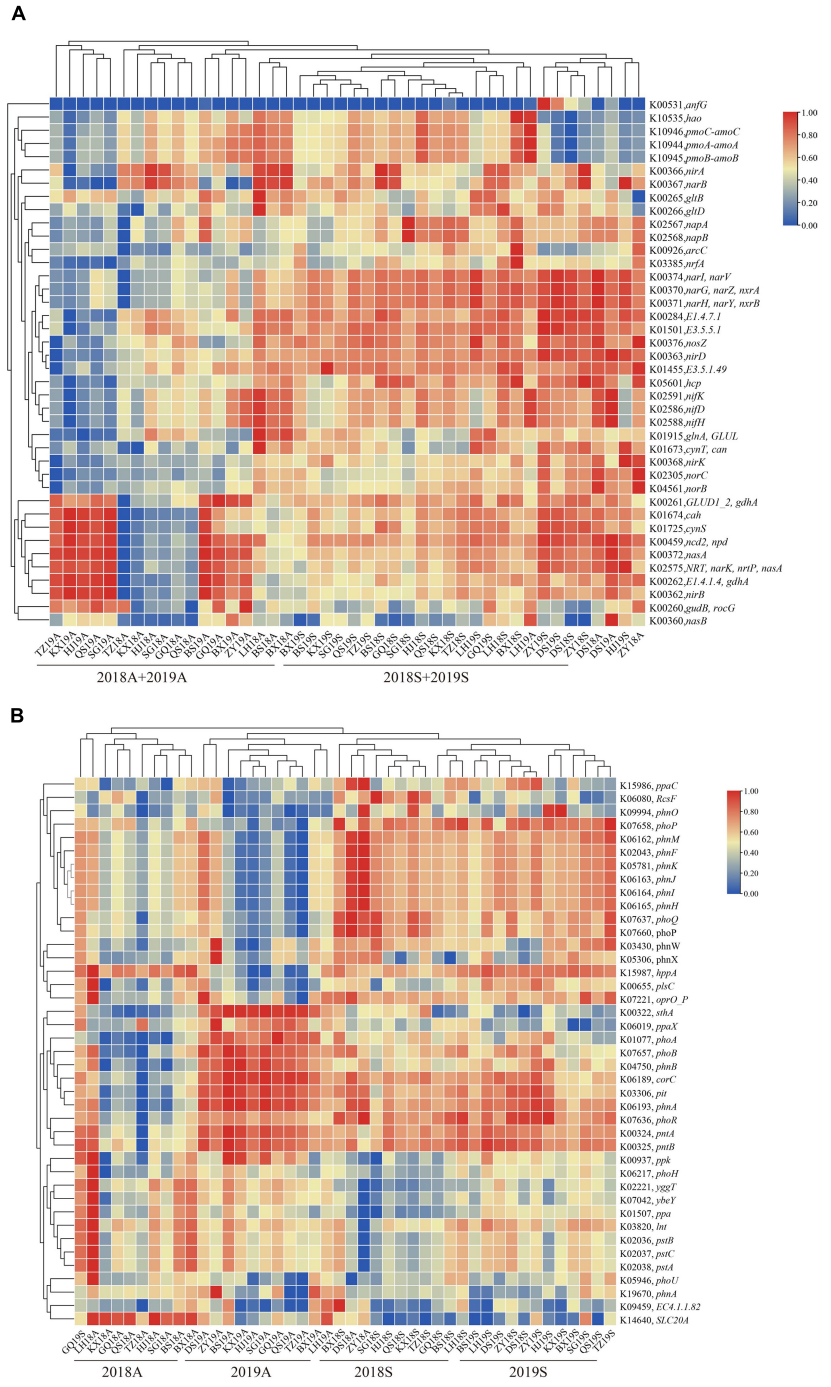


FIGURE 5 | Heatmap showing the hierarchical clustering of the predicted functional genes related to the N **(A)** and P **(B)** cycles in bacterioplankton sample.

bacterioplankton in different years and seasons based on the RMT method (**Figure 6**). Based on the analysis of the network properties, the number of nodes in the 2019 sample was higher than that in the 2018 sample, and it was higher in the spring than in the autumn. The main trend for the total number of links was that the number was lower in spring than in autumn. Additionally, the average clustering coefficient and average connectivity of the molecular ecological network for the spring samples were lower than those for the autumn samples, and the number of modules were higher for the spring samples than for the autumn samples.

In addition, molecular ecological networks have been widely used in the identification of core microbiomes. Nodes with $Zi \geq 2.5$ or $Pi \geq 0.62$ were defined as core microbiomes. The

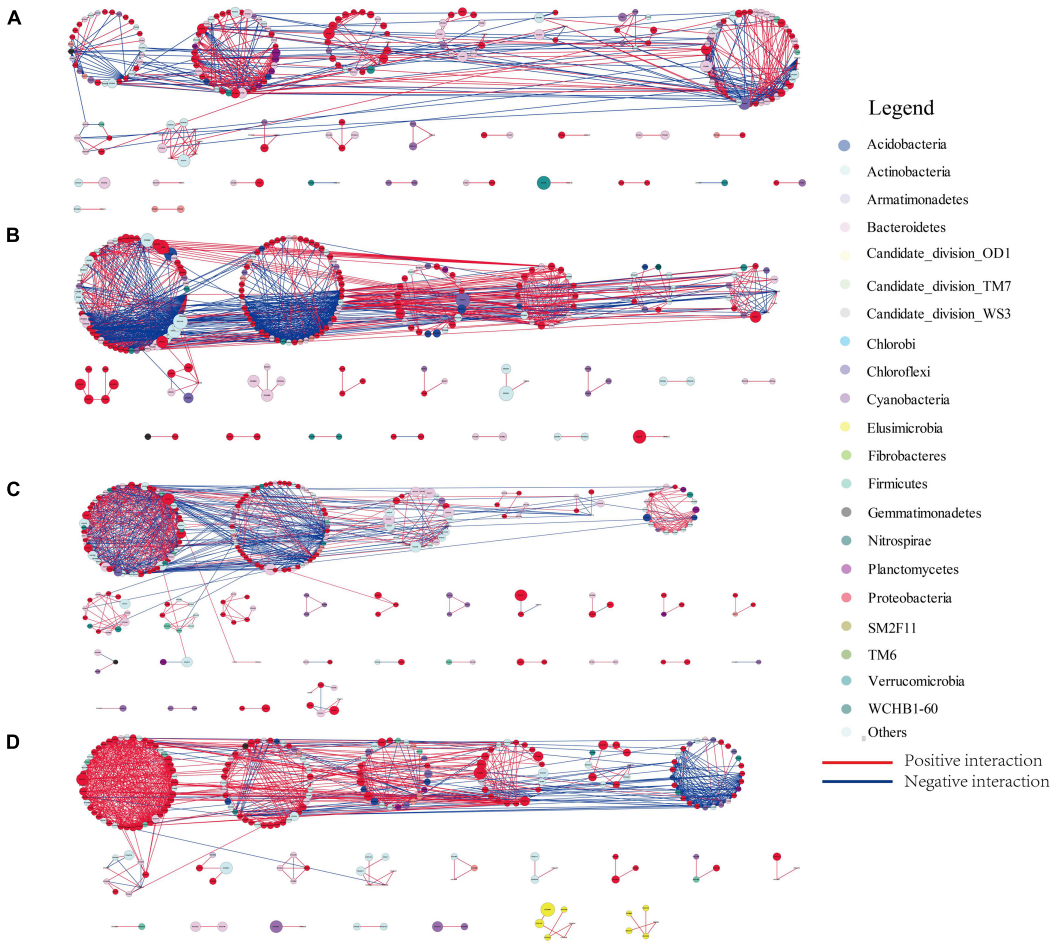


FIGURE 6 | Overview of the 2018S **(A)**, 2018A **(B)**, 2019S **(C)**, 2019A **(D)** bacterioplankton networks. Each node represents an OTU. The colored circles indicate the OTUs affiliated with particular phyla (color code on the right).

number and composition of core microbiomes in the molecular ecological network of bacterioplankton in different years and seasons in the DJR were different. There were 8 core microbiomes (OTU number) for the 2018S sample: OTU2563 and OTU3223 of Actinobacteria, OTU1049, OTU3534, and OTU3798 of Bacteroidetes, OTU2938 of Cyanobacteria, OTU2425 of Gemmatimonadetes, and OTU1780 of Proteobacteria (**Figure 7A**). There were 5 core microbiomes (OTU number) for the 2018A sample: OTU1765 of Actinobacteria, OTU89 of Bacteroidetes, OTU12 of Proteobacteria, OTU3233 of Verrucomicrobia (**Figure 7B**). There were 6 core microbiomes (OTU number) for the 2019S sample: OTU3748, OTU4243 and OTU3770 of Bacteroidetes, OTU1931, OTU2817 and OTU3044 of Proteobacteria (**Figure 7C**). There were 6 core microbiomes (OTU number) for the 2019A sample: OTU2082, OTU3717 of Actinobacteria, OTU2641 of Firmicutes, OTU1976 of Patescibacteria, OTU2034, OTU960 of Proteobacteria (**Figure 7D**). The OTU classification information for these key bacteria is provided in **Supplementary Table 3**. Spearman correlation was used to analyze the relationship between environmental factors and key bacteria. The results showed that T, Cond, COD_{Mn}, TP, NH_4^+-N, NO_3^--N and SD were the main factors affecting core microbiomes.

Community Assembly

Neutral model analysis of the bacterioplankton communities in the DJR showed that the neutral model explained 66.3%, 63.0%, 63.0%, and 70.9% of the bacterioplankton community variations in 2018S (**Figure 8A**), 2018A (**Figure 8B**), 2019S (**Figure 8C**), and 2019A (**Figure 8D**), respectively, and that stochastic processes played a leading role.

DISCUSSION

The Composition of and Yearly and Seasonal Variation in the Bacterioplankton Communities in the Danjiangkou Reservoir

High-throughput sequencing (HTS) has been widely used in the study of bacterial community composition in aquatic

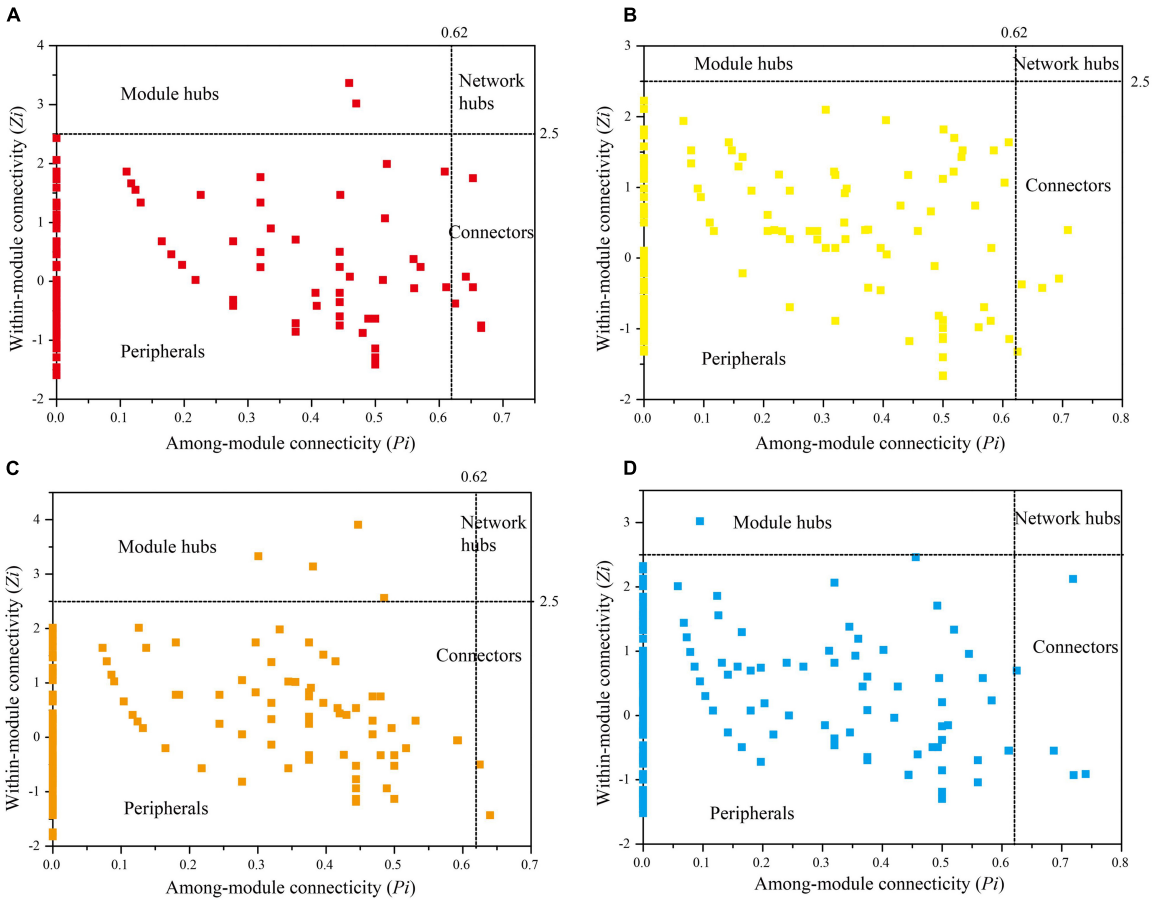


FIGURE 7 | *Zi-Pi* plot of 2018S **(A)**, 2018A **(B)**, 2019S **(C)**, 2019A **(D)** bacterioplankton networks.

ecosystems, giving people the opportunity to conduct more in-depth and comprehensive studies on the structure and function of bacterial communities in freshwater (Grossart, 2010). Liu et al. (2015) analyzed the bacterioplankton composition of 42 lakes and reservoirs in China. The results indicated that Actinobacteria, Bacteroidetes, and Cyanobacteria were the dominant populations. Sun et al. (2021) analyzed Han River bacterioplankton upstream of the DJR. The results showed that the dominant genera were *Flavobacterium*, *Planktophila*, and *Siphonobacter*. Our assessment of the bacterioplankton communities in the DJR showed that they were mainly composed of bacterial phyla common in water bodies, such as Proteobacteria, Actinobacteria, and Bacteroidetes. At the genus level, the communities were mainly composed of *CL500-29_marine_group*, *Acinetobacter*, *hgcI_clade*, and *Limnohabitans*. The community composition was similar to the composition of DJR bacterioplankton in May 2016 but differed from the composition of bacterioplankton in the main water diversion canal in the Han River, upstream of the DJR (Luo et al., 2019; Chen Z. J. et al., 2020; Zhang L. et al., 2021).

Many studies have shown that bacterioplankton exhibit temporal and spatial variation characteristics under the influence of physicochemical properties such as T, pH, nutrients, and DO in the water body and hydrological factors such as flow velocity, flow rate, and water level (Pearce, 2005; Qian-Qian et al., 2021). The dry season at the DJR occurs from February to July, and the wet season occurs from August to January of the following year. The water level of, nutrients in, and temperature of the reservoir vary greatly during different periods (Dong et al., 2020). Based on our analysis, the physicochemical properties of the water body were different in different periods, and T, TN, and COD_{Mn} exhibited trends of being higher in the autumn than in the spring (**Supplementary Table 1**). UPGMA clustering and PLS-DA of bacterioplankton community composition revealed that bacterioplankton communities were significantly different in different seasons and years and that the influence of season was the most important factor. This finding is consistent with the results reported by Luo et al. (2019), who showed that the seasonal variation in the composition of the bacterioplankton communities in the main canal was greater than that of the spatial distribution variation.

Seasonal variation in bacterioplankton can be reflected not only by community composition but also through the interaction and composition and function of core microbiomes. Jiao et al. (2020) conducted a network analysis of planktonic bacterial communities and showed seasonal variations in bacterioplankton

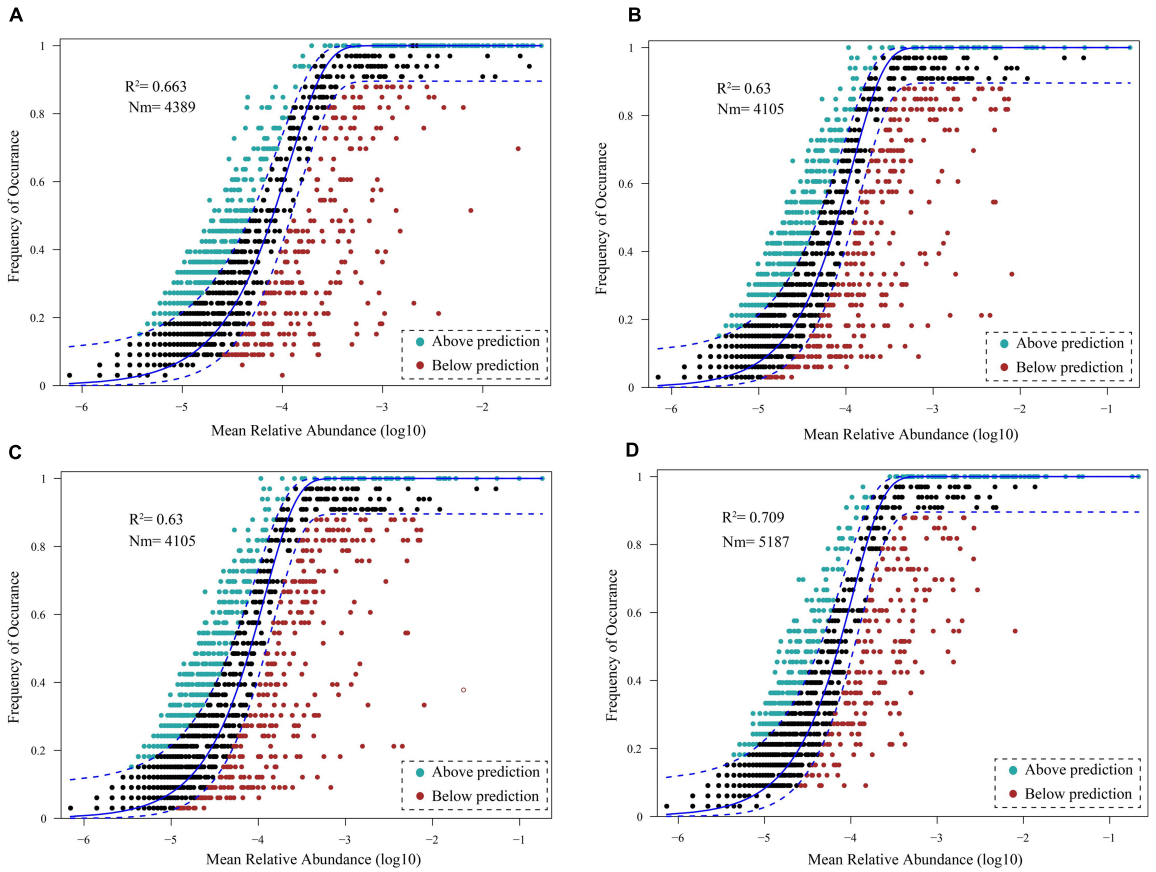


FIGURE 8 | Fit of the neutral community model (NCM) of 2018S **(A)**, 2018A **(B)**, 2019S **(C)**, 2019A **(D)** bacterioplankton community assembly.

network attributes. Molecular ecological network analysis of DJR bacterioplankton showed that network attributes such as nodes, links, average clustering coefficient, average connectivity, and modules all exhibited seasonal variations. These network attributes are used to indicate the size, complexity, efficiency of transferring matter, energy and information between species, and sensitivity to the external environment of the network (Zhou J. et al., 2011; Chen et al., 2021). The links, average clustering coefficient, and average connectivity for the autumn samples were all higher than those for the spring samples, indicating a more complex network association and higher susceptibility to interference from the external environment. Similarly, an analysis of the bacterioplankton network for the four seasons at Lake Taihu indicated that the network in autumn was the most complex and the network in spring was the simplest (Lin et al., 2019). The reason may be due to the massive reproduction of phytoplankton in autumn. The large amount of soluble organic matter, including hydrocarbons and organic acids, produced by phytoplankton photosynthesis provides abundant nutrients for bacteria, allowing bacterioplankton networks to be more complex. Another important function of network analysis is the identification of core microbiomes, which play important roles in the stabilization of ecological service functions. The results showed that the number of core microbes in 2018S was higher than that in 2018A. The number of core microbiomes in 2019 was the same, but the composition of core microbiomes was significantly different. The relative abundance of bacterial communities to which some key nodes belong to in the Danjiangkou bacterial network was relatively low (< 1%), indicating that low-abundance bacteria in the water body played an important role in bacterial network assembly and suggesting that further attention should be paid to the role of low-abundance bacteria (Liu et al., 2015; Peter et al., 2018). The molecular network analysis used to investigate the links between numerous taxa, however, these links are often difficult to provide any evidence of such interactions using biochemical or other standard microbiological tests referring to living microbes. Thus, it is required follow up experimental validation to confirm true bacterial interactions in Danjiangkou Reservoir.

N and P Nutrients Are the Main Factors Driving the Communities and Function of Bacterioplankton in the Danjiangkou Reservoir

The DJR is an important drinking water source in China, and water quality is always stable at classes II and above. A claim supported by the monitoring results from this experiment

(**Supplementary Table 1**). Due to the impact of non-point source pollution, such as agricultural production around the DJR and the input of main and tributary streams, such as the Hanjiang River and the Danjiang River, the excessive TN content in the DJR has become a major threat to water quality in the water body. Dong et al. (2020) conducted TN analysis of the DJR for 2016-2020 and showed that the TN content ranged from 0.928 to 1.020 mg/L. The results of this experiment showed that the TN content in the DJR exceeded 1.00 mg/L, with the highest content being 2.862 mg/L, meeting the standard for class IV surface water and showing a trend of higher concentrations in autumn than in spring. An explanation for thus finding is that autumn is the wet season, and N in the soil of the upstream basin enters the water body with rain. As the main factor in freshwater ecosystems, N and P nutrients significantly affect the composition of bacterioplankton communities (Kisand et al., 2001; Maria Montserrat et al., 2002). An RDA of the bacterioplankton communities and environmental factors in the DJR in May 2016 showed that TN was an important factor affecting the distribution of bacterioplankton (Chen Z. J. et al., 2020). In this study, RDA and CCA indicated that N and P nutrients significantly affected the composition of the bacterioplankton communities, and subsequent VPA analysis indicated that N and P nutrients were the most important factors affecting the bacterioplankton communities. Previous long-term monitoring of the DJR showed that it was in a mesotrophic state and that N and P contents in the DJR were lower than those in other lakes in China (Chen Z. J. et al., 2020). Many previous studies have shown that in lakes the nutrient concentration may be the limiting factor for bacterial growth and is the main factor affecting the composition of bacterioplankton communities (Vrede, 2005). Spearman correlation analysis showed that NH_4^+-N, NO_3^--N, TP and other nutrients were also the main factors affecting core microbiomes, indicating that they can also affect bacterioplankton interactions.

Microbes are also the main driving force of the N and P cycles in the DJR. Dang et al. (2021) found nitrogen-cycle bacteria such as *Rhodoferax*, *Polaromonas*, *Limnohabitans*, *Pararheinheimera*, *Desulfobulbus*, and *Pseudopelobacter* as well as 51 nitrogen functional genes in the DJR. Studies have shown that PICRUSt can accurately predict the presence and abundance of functional genes (Hartman et al., 2017; LeBrun and Kang, 2018; Ribeiro et al., 2018). We analyzed the functional genes related to the N and P cycles in the DJR in different years and seasons. PICRUSt2 predicted 40 N functional genes and 41 P functional genes, findings that were consistent with the functional genes of bacterioplankton for N cycling in Pearl River Estuary predicted by Zhu et al. (2018) and the functional genes of bacterioplankton for P cycling in Poyang Lake predicted by Ren et al. (2019). These results indicated that the bacterioplankton communities in the DJR had abundant N and P cycle functional genes. Cluster analysis of N cycle gene copy number indicated that similar to the composition of the bacterial communities, season was the most important factor affecting the bacterioplankton N cycle. The above results indicate that N and P nutrients are the main factors driving the communities and function of bacterioplankton in the DJR. Similar to other lakes and reservoirs, T, pH, COD, COD_{Mn}, Cond, and Chla are also important factors affecting the composition of the bacterioplankton communities in the DJR (Liu et al., 2015; Chen Z. J. et al., 2020; Dong et al., 2020; Qian-Qian et al., 2021; Sun et al., 2021).

Both Stochastic Processes and Deterministic Processes Dominate Bacterioplankton Community Assembly in the Danjiangkou Reservoir

Currently, there is no study on bacterioplankton community assembly in the DJR. Relevant studies are not only conducive to predicting variations in community composition but also play a potential role in bacterioplankton ecological function and diversity protection (Nemergut et al., 2013). Niche theory and neutral theory are the main models that explain the formation and maintenance of biodiversity, that is, the community assembly. The niche theory proposes that biological communities are regulated by environmental selection and biological interaction, which are deterministic processes; the neutral theory proposes that stochastic processes, including the birth, death, migration, and diffusion of species, shape biological communities (Zhou and Ning, 2017; Zhang et al., 2018; Chen et al., 2019). Researchers have tried to strengthen the understanding of the influence of stochastic processes on bacterioplankton community assembly based on the neutral theory. In this study, Environmental factors such as T, pH, DO, COD, COD_{Mn}, TN, NH_4^+-N and Chla were significantly related to variations in the bacterioplankton community. They demonstrated that environmental factors have impact on bacterial community assembly. But, the high proportion of unexplained variation in bacterioplankton communities indicated the potential importance of neutral or stochastic processes for community assembly. Our analysis of the neutral model of bacterioplankton communities in the DJR showed that the neutral model explained 66.3%, 63.0%, 63.0%, and 70.9% of the bacterioplankton community variation in 2018S, 2018A, 2019S, and 2019A, respectively. The neutral community model fitted well for the bacterioplankton community with a moderate fitted value (R^2 = 0.630∼0.709). The fitted value indicated that stochastic process played only a moderate role in the community assembly process by comparing with other studies (Chen et al., 2019; Zhang Z. F. et al., 2021; Zhang et al., 2022). Due to the heightening of the dam and the increase in the water level, the aquatic ecosystem in the DJR is in a process of reconstruction. Additionally, due to the demand for water storage and water transfer, the hydrological and water quality physicochemical properties of the reservoir during the dry season and the wet season vary greatly each year, potentially affecting neutral (dispersal-related) process in the DJR and making it mainly a stochastic process.

CONCLUSION

The water quality in the DJR is important because it is associated with the safety of drinking water for hundreds of millions

of people along the MR-SNWD in China. In view of the important role of bacterioplankton in aquatic ecosystems, the study of bacterioplankton community composition, function and community assembly has important ecological and economic significance. We analyzed the bacterioplankton community composition and distribution characteristics in different years and seasons and found that bacterioplankton had annual and seasonal variations. Additionally, the main factors affecting the bacterial communities were analyzed through RDA and CCA, and it was found that N and P nutrients were the main driving factors. Subsequently, the function of bacterioplankton was predicted by PICRUSt2, and it was found that the N and P cycle functions of bacterioplankton had significant seasonality. In addition, ecological network analysis revealed that the links, average clustering coefficient, and average connectivity of the autumn samples were all higher than those of the spring samples, indicating that the network was more complex and more susceptible to interference from the external environment. The analysis of the neutral model showed that stochastic processes dominated bacterioplankton community assembly in the DJR. This study systematically studied the composition, function, interaction, and assembly of the bacterioplankton communities in the DJR as well as the influencing factors, providing a reference for the protection of water quality and the ecological functions of DJR bacterioplankton.

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: MG-RAST metagenomics analysis server, accession number: mgp101864 (https://www.mg-rast.org/linkin.cgi?project=mgp101864).

AUTHOR CONTRIBUTIONS

Z-JC: conceptualization, investigation, and writing – original draft. Y-QL, L-AL, and B-HZ: investigation. Y-YL: conceptualization, funding acquisition, and writing – review and editing. M-FJ: methodology and investigation. BL: conceptualization and writing – review and editing. X-MH: conceptualization, methodology, and writing – original draft. All authors contributed to the article and approved the submitted version.

FUNDING

This research was supported by the National Natural Science Foundation of China (Grant Nos. 51879130, U2004145, 41601332), higher discipline innovation and talent introduction base of Henan Province (No. CXJD2019001), and the Key Scientific and Technological Project of Henan Province (Grant No. 212102310844).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2022.884765/full#supplementary-material

REFERENCES

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.

Chen, W., Ren, K., Isabwe, A., Chen, H., Liu, M., and Yang, J. (2019). Stochastic processes shape microeukaryotic community assembly in a subtropical river across wet and dry seasons. *Microbiome* 7:138.

Chen, Z.-J., Tian, W., Li, Y.-J., Sun, L.-N., Chen, Y., Zhang, H., et al. (2021). Responses of rhizosphere bacterial communities, their functions and their network interactions to Cd stress under phytostabilization by *Miscanthus* spp. *Environ. Pollut.* 287:117663. doi: 10.1016/j.envpol.2021.117663

Chen, C., Chen, H., Zhang, Y., Thomas, H. R., Frank, M. H., He, Y., et al. (2020). TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Mol. Plant* 13, 1194–1202. doi: 10.1016/j.molp.2020.06.009

Chen, Z.-J., Xu, G., Ding, C.-Y., Zheng, B.-H., Chen, Y., Han, H., et al. (2020). Illumina MiSeq sequencing and network analysis the distribution and co-occurrence of bacterioplankton in Danjiangkou Reservoir. *China. Arch. Microbiol.* 202, 859–873. doi: 10.1007/s00203-019-01798-7

Chen, Z., Yuan, J., Sun, F., Zhang, F., Chen, Y., Ding, C., et al. (2018). Planktonic fungal community structures and their relationship to water quality in the Danjiangkou Reservoir. *China. Sci. Rep.* 8:10596. doi: 10.1038/s41598-018-28903-y

Dang, C., Liu, S., Chen, Q., Sun, W., Zhong, H., Hu, J., et al. (2021). Response of microbial nitrogen transformation processes to antibiotic stress in a drinking water reservoir. *Sci. Total Environ.* 797:149119.

Deng, Y., Jiang, Y.-H., Yang, Y., He, Z., Luo, F., and Zhou, J. (2012). Molecular ecological network analyses. *BMC Bioinform.* 13:113.

DeSantis, T. Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E. L., Keller, K., et al. (2006). Greengenes, a Chimera-Checked 16S rRNA gene database and workbench compatible with ARB. *Appl. Environ. Microbiol.* 72, 5069–5072. doi: 10.1128/AEM.03006-05

Dong, G., Hu, Z., Liu, X., Fu, Y., and Zhang, W. (2020). Spatio-Temporal variation of total nitrogen and ammonia nitrogen in the water source of the middle route of the south-to-north water diversion project. *Water* 12:2615.

Douglas, G. M., Maffei, V. J., Zaneveld, J. R., Yurgel, S. N., Brown, J. R., Taylor, C. M., et al. (2020). PICRUSt2 for prediction of metagenome functions. *Nat. Biotechnol.* 38, 685–688. doi: 10.1038/s41587-020-0548-6

Ducklow, H. W., Purdie, D. A., Williams, P. J. L., and Davies, J. M. (1986). Bacterioplankton: a sink for carbon in a coastal marine plankton community. *Science* 232, 865–867. doi: 10.1126/science.232.4752.865

Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26, 2460–2461. doi: 10.1093/bioinformatics/btq461

Grossart, H.-P. (2010). Ecological consequences of bacterioplankton lifestyles: changes in concepts are needed. *Environ. Microbiol. Rep.* 2, 706–714. doi: 10.1111/j.1758-2229.2010.00179.x

Hartman, W. H., Ye, R., Horwath, W. R., and Tringe, S. G. (2017). A genomic perspective on stoichiometric regulation of soil carbon cycling. *ISME J.* 11, 2652–2665. doi: 10.1038/ismej.2017.115

Jiao, C., Zhao, D., Zeng, J., Guo, L., and Yu, Z. (2020). Disentangling the seasonal co-occurrence patterns and ecological stochasticity of planktonic and benthic bacterial communities within multiple lakes. *Sci. Total Environ.* 740:140010. doi: 10.1016/j.scitotenv.2020.140010

Kara, E. L., Hanson, P. C., Hu, Y. H., Winslow, L., and McMahon, K. D. (2013). A decade of seasonal dynamics and co-occurrences within freshwater

bacterioplankton communities from eutrophic Lake Mendota, WI, USA. *ISME J.* 7, 680–684. doi: 10.1038/ismej.2012.118

Kisand, V., Tuvikene, L., and Nõges, T. (2001). Role of phosphorus and nitrogen for bacteria and phytopankton development in a large shallow lake. *Hydrobiologia* 457, 187–197.

LeBrun, E., and Kang, S. (2018). A comparison of computationally predicted functional metagenomes and microarray analysis for microbial P cycle genes in a unique basalt-soil forest. *F1000Res.* 7:179. doi: 10.12688/f1000research.13841.1

Lin, Y., Zhao, D., Zeng, J., Cao, X., and Jiao, C. (2019). Network analysis reveals seasonal patterns of bacterial community networks in lake taihu under aquaculture conditions. *Water* 11:1868.

Liu, H., Yin, J., and Feng, L. (2018). The dynamic changes in the storage of the danjiangkou reservoir and the influence of the south-north water transfer project. *Sci. Rep.* 8:8710. doi: 10.1038/s41598-018-26788-5

Liu, L., Yang, J., Yu, Z., and Wilkinson, D. M. (2015). The biogeography of abundant and rare bacterioplankton in the lakes and reservoirs of China. *ISME J.* 9, 2068–2077. doi: 10.1038/ismej.2015.29

Luo, Z., Li, S., Hou, K., and Ji, G. (2019). Spatial and seasonal bacterioplankton community dynamics in the main channel of the middle route of south-to-north water diversion project. *Res. Microbiol.* 170, 24–34. doi: 10.1016/j.resmic.2018.08.004

Maria Montserrat, S., Francesc, P., Josep, M. G., Carlos, PÂs-AÂ, Celia, M. Â, and Dolors, V. Â (2002). Seasonal and spatial variations in the nutrient limitation of bacterioplankton growth in the northwestern Mediterranean. *Aquat. Microb. Ecol.* 27, 47–56.

Mohapatra, M., Behera, P., Kim, J. Y., and Rastogi, G. (2020). Seasonal and spatial dynamics of bacterioplankton communities in a brackish water coastal lagoon. *Sci. Total Environ.* 705:134729. doi: 10.1016/j.scitotenv.2019.134729

Morris, J. R. D. P., and Lewis, W. M. (1992). Nutrient limitation of bacterioplankton growth in Lake Dillon. Colorado. *Limnol. Oceanogra.* 37, 1179–1192.

Nemergut, D. R., Schmidt, S. K., Fukami, T., O'Neill, S. P., Bilinski, T. M., Stanish, L. F., et al. (2013). Patterns and processes of microbial community assembly. *Microbiol. Mol. Biol. Rev.* 77, 342–356.

Pan, Y., Guo, S., Li, Y., Yin, W., Qi, P., Shi, J., et al. (2018). Effects of water level increase on phytoplankton assemblages in a drinking water reservoir. *Water* 10:256.

Pearce, D. A. (2005). The structure and stability of the bacterioplankton community in Antarctic freshwater lakes, subject to extremely rapid environmental change. *FEMS Microbiol. Ecol.* 53, 61–72. doi: 10.1016/j.femsec.2005.01.002

Peter, H., Jeppesen, E., De Meester, L., and Sommaruga, R. (2018). Changes in bacterioplankton community structure during early lake ontogeny resulting from the retreat of the Greenland Ice Sheet. *ISME J.* 12, 544–555. doi: 10.1038/ismej.2017.191

Qian-Qian, Z., Sheng-Long, J., Ke-Mao, L., Zhen-Bing, W., Hong-Tao, G., Jin-Wen, H., et al. (2021). Community structure of bacterioplankton and its relationship with environmental factors in the upper reaches of the Heihe River in Qinghai Plateau. *Environ. Microbiol.* 23, 1210–1221. doi: 10.1111/1462-2920.15358

Ren, Z., Qu, X., Zhang, M., Yu, Y., and Peng, W. (2019). Distinct Bacterial Communities in Wet and Dry Seasons During a Seasonal Water Level Fluctuation in the Largest Freshwater Lake (Poyang Lake) in China. *Front. Microbiol.* 10:1167. doi: 10.3389/fmicb.2019.01167

Ribeiro, H., de Sousa, T., Santos, J. P., Sousa, A. G. G., Teixeira, C., Monteiro, M. R., et al. (2018). Potential of dissimilatory nitrate reduction pathways in polycyclic aromatic hydrocarbon degradation. *Chemosphere* 199, 54–67. doi: 10.1016/j.chemosphere.2018.01.171

Ren, Y., Yu, G., Shi, C., Liu, L., Guo, Q., Han, C., et al. (2022). Majorbio cloud: a one-stop, comprehensive bioinformatic platform for multiomics analyses. *iMeta* 1:e12. doi: 10.1002/imt2.12

Shu, X., Zhang, K., Zhang, Q., and Wang, W. (2017). Response of soil physico-chemical properties to restoration approaches and submergence in the water level fluctuation zone of the Danjiangkou Reservoir, China. *Ecotoxicol. Environ. Saf.* 145, 119–125. doi: 10.1016/j.ecoenv.2017.07.023

Sun, H., Pan, B., He, H., Zhao, G., Jiang, X., Han, X., et al. (2021). Characterization of the bacterioplankton community and the influencing factors in the upper reaches of the Han River basin. *Environ. Sci. Pollut. Res.* 28, 61748–61759. doi: 10.1007/s11356-021-14906-2

Vrede, K. (2005). Nutrient and temperature limitation of bacterioplankton growth in temperate lakes. *Microb. Ecol.* 49, 245–256. doi: 10.1007/s00248-004-0259-4

Wang, M. C., Liu, X. Q., and Zhang, J. H. (2002). Evaluate method and classification standard on lake eutrophication. *Environ. Monit. China* 18, 47–49.

Zhang, J., Chen, Y., Huo, Y., Guo, J., Wan, L., Lu, Z., et al. (2021). Eutrophication increases deterministic processes and heterogeneity of co-occurrence networks of bacterioplankton metacommunity assembly at a regional scale in tropical coastal reservoirs. *Water Res.* 202:117460. doi: 10.1016/j.watres.2021.117460

Zhang, L., Yin, W., Wang, C., Zhang, A., Zhang, H., Zhang, T., et al. (2021). Untangling microbiota diversity and assembly patterns in the world's largest water diversion canal. *Water Res.* 204:117617. doi: 10.1016/j.watres.2021.117617

Zhang, T., Xu, S., Yan, R., Wang, R., Gao, Y., Kong, M., et al. (2022). Similar geographic patterns but distinct assembly processes of abundant and rare bacterioplankton communities in river networks of the Taihu Basin. *Water Res.* 211:118057. doi: 10.1016/j.watres.2022.118057

Zhang, W., Pan, Y., Yang, J., Chen, H., Holohan, B., Vaudrey, J., et al. (2018). The diversity and biogeography of abundant and rare intertidal marine microeukaryotes explained by environment and dispersal limitation. *Environ. Microbiol.* 20, 462–476. doi: 10.1111/1462-2920.13916

Zhang, Z.-F., Pan, J., Pan, Y.-P., Li, M., and Gilbert, J. A. (2021). Biogeography, Assembly Patterns, Driving Factors, and Interactions of Archaeal Community in Mangrove Sediments. *mSystems* 6, e1381–e1320. doi: 10.1128/mSystems.01381-20

Zhou, J., Deng, Y., Luo, F., He, Z., and Yang, Y. (2011). Phylogenetic molecular ecological network of soil microbial communities in response to elevated CO 2. *MBio* 2:e00122-11.

Zhou, J., and Ning, D. (2017). Stochastic Community Assembly: Does It Matter in Microbial Ecology? *Microbiol. Mol. Biol. Rev.* 81:e00002–e17. doi: 10.1128/MMBR.00002-17

Zhou, J. Z., Deng, Y., Luo, F., He, Z. L., and Yang, Y. F. (2011). Phylogenetic molecular ecological network of soil microbial communities in response to elevated CO2. *mBio* 2, e122–e111.

Zhu, J., Hong, Y., Zada, S., Hu, Z., and Wang, H. (2018). Spatial variability and co-acclimation of phytoplankton and bacterioplankton communities in the pearl river estuary, China. *Front. Microbiol.* 9:2503. doi: 10.3389/fmicb.2018.02503

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

The reviewer LY declared a shared affiliation with the authors Z-JC, Y-QL, Y-YL, L-AL, B-HZ, and M-FJ to the handling editor at the time of review.

frontiers | Frontiers in Microbiology

TYPE Original Research
PUBLISHED 25 July 2022
DOI 10.3389/fmicb.2022.928877

OPEN ACCESS

EDITED BY
Hongbin Liu,
Hong Kong University of Science and Technology, Hong Kong SAR, China

REVIEWED BY
Weipeng Zhang,
Ocean University of China, China
Rashmi Rathour,
Jawaharlal Nehru University, India

*CORRESPONDENCE
Stephen Summers
ssummers@ntu.edu.sg
William R. Birch
w-birch@imre.a-star.edu.sg

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

RECEIVED 26 April 2022
ACCEPTED 07 July 2022
PUBLISHED 25 July 2022

CITATION
Summers S, Pek YS, Vinod DP, McDougald D, Todd PA, Birch WR and Rice SA (2022) Bacterial biofilm colonization and succession in tropical marine waters are similar across different types of stone materials used in seawall construction. *Front. Microbiol.* 13:928877. doi: 10.3389/fmicb.2022.928877

Bacterial biofilm colonization and succession in tropical marine waters are similar across different types of stone materials used in seawall construction

Stephen Summers[1]*, Y. Shona Pek[2], Deepthi P. Vinod[3], Diane McDougald[4], Peter A. Todd[5], William R. Birch[2]* and Scott A. Rice[1,6]

[1]Singapore Centre for Environmental Life Sciences Engineering, Nanyang Technological University, Singapore, Singapore, [2]Institute of Materials Research and Engineering, Agency for Science, Technology and Research (A*STAR), Singapore, Singapore, [3]School of Bioscience and Technology, Vellore Institute of Technology, Vellore, India, [4]Australian Institute for Microbiology and Infection, The University of Technology Sydney, Sydney, NSW, Australia, [5]Department of Biological Sciences, National University of Singapore, Singapore, Singapore, [6]Agriculture and Food, Microbiomes for One Systems Health, Commonwealth Scientific and Industrial Research Organisation (CSIRO), Canberra, ACT, Australia

Seawalls are important in protecting coastlines from currents, erosion, sea-level rise, and flooding. They are, however, associated with reduced biodiversity, due to their steep orientation, lack of microhabitats, and the materials used in their construction. Hence, there is considerable interest in modifying seawalls to enhance the settlement and diversity of marine organisms, as microbial biofilms play a critical role facilitating algal and invertebrate colonization. We assessed how different stone materials, ranging from aluminosilicates to limestone and concrete, affect biofilm formation. Metagenomic assessment of marine microbial communities indicated no significant impact of material on microbial diversity, irrespective of the diverse surface chemistry and topography. Based on KEGG pathway analysis, surface properties appeared to influence the community composition and function during the initial stages of biofilm development, but this effect disappeared by Day 31. We conclude that marine biofilms converged over time to a generic marine biofilm, rather than the underlying stone substrata type playing a significant role in driving community composition.

KEYWORDS

seawall, biofilm, microbial diversity, marine microbes, coastal protection

Introduction

A warming planet has led to thermal expansion of the oceans and melting of ice sheets, thus raising sea levels around the world (Warrick and Oerlemans, 1990; Douglas, 1991; Golledge et al., 2019) and accelerating coastal erosion and flooding. Low lying islands and coastal cities are particularly vulnerable (Bruun, 1962; Nicholls and Cazenave, 2010). As human populations continue to migrate toward the coast, the risk to property and life will increase (Temmerman et al., 2013). The installation of artificial defense structures such as seawalls has become the primary mitigation strategy for preventing some of the more extreme effects of sea level rise (Firth et al., 2014; Todd et al., 2019). These can be constructed from any material that can withstand the constant action of waves and currents, but concrete and natural stone materials are commonly used (Dugan et al., 2011). For many countries, large scale installations represent a significant amount of the coastline. For example, Korea's artificial Saemangeum dyke is > 33 km in length (Lam et al., 2009), Japan is planning over 400 km of seawalls to protect from future tsunami disasters (Yamashita, 2020) and in Singapore, > 63% of the coast is some form of seawall, either vertical concrete or granite rip-rap (Lai et al., 2015). However, seawall proliferation comes at a considerable cost to coastal ecosystems. Heatherington and Bishop (2012) reported that seawalls affect the physical and hydrodynamic processes of a natural coastal region, such as sediment and larval transportation. Others have shown that they support lower species diversity and different communities (Ming et al., 2010; Lai et al., 2015) and can encourage invasive species (Simkanin et al., 2012). Critically, the pre-existing habitat is often obliterated by the installation of such large and intrusive infrastructure (Firth et al., 2014).

While there exists a strong ecological case for moving away from seawalls to ecosystem-based coastal defenses such as mangroves and saltmarshes, these require substantial shoreline space that is generally not available near cities (Temmerman et al., 2013). It is inevitable that seawalls will continue to be constructed. However, there remains the potential to improve their design to make them more viable as a habitat for marine life (Loke et al., 2015; Morris et al., 2019). One solution that is being extensively researched is to increase the topographical complexity of the seawall to provide niche availability for a wider range of species (Loke et al., 2019). It is thought that anthropogenic seawall installations are lacking in complexity and thus have a reduce niche environment (Lawrence et al., 2021). However, more recent investigations have highlighted that complexity alone is not universally positive with respect to increasing biodiversity, as this can be affected by local stressors, such as predation and should be seen only as an element in improving seawalls (Chee et al., 2021; Strain et al., 2021). Another, much less studied approach, is to identify construction materials that may encourage colonization by marine biota (Dodds et al., 2022; Hartanto et al., 2022). Before any invertebrates or other eukaryotic organisms settle on the seawall surface, it is known that microbial colonizers begin to form "conditioning" biofilms. These biofilms are a vital first step in the recruitment of other organisms (Freckelton et al., 2017), as the biofilms and species therein have been shown to have a role in regulating the settlement of sessile organisms onto the substratum (Hadfield, 2011; Freckelton et al., 2022). However, it is unclear whether manipulating the seawall construction material has any substantial impact on the biofilms that initially form on the seawall surface. Studies comparing seawall microbial biofilms to those found on natural rocky shorelines have shown differences in the relative abundance of biofilm inhabitants between the two environments (Tan et al., 2015). However, settlement control plates present at both sites (seawall and rocky shore) did not differ in biofilm settlement. This suggests that the locations are not driving any observed differences, but instead it is the chemical and/or physical properties of those sites that drive differences in the settled and formed microbial communities.

The current study selected natural stones comprising aluminosilicates and calcium carbonate, which present different surface chemistry. As substrates, they also differ in their resistance to erosion. Granite is an aluminosilicate that resists chemical erosion, which has led to its widespread implementation as a durable mechanical barrier on seawalls. In contrast, limestone is composed from calcium carbonate, which is more readily eroded in a mildly acidic environment (Revelle and Emery, 1957), thus giving rise to natural macrostructures with higher complexity (Higgins, 1980). Sandstone differs by offering a high degree of porosity and it has been reported to promote the growth of foliose algae (Green et al., 2012). Settlement on natural stones and the ensuing erosion generated by microbial species have been studied and shown to influence the microbial biofilms associated with the stones (Cockell et al., 2011; Summers et al., 2016). Differences have been observed between microorganisms weathering limestone, granite, and concrete (Coombes et al., 2011), with boring also being influenced by environmental factors (Cockell and Herrera, 2008). The surface chemistry of aluminosilicates presents silanol groups that have high affinity to water (Birch et al., 2008), while calcium carbonate substrates present moieties that are reported to bind to organic molecules (Suess, 1970; Carter, 1978; Zeller et al., 2020).

Based on these observations, we hypothesized that differing construction materials would recruit different microbial biofilm communities. If this hypothesis is correct, then construction material could be a key component not only for seawall strength and durability, but also for any biodiversity enhancement strategy, such as the addition of tiles or cladding. Therefore, to address the role of substrata in marine biofilm community composition, we examined several construction materials for the formation of natural tropical marine biofilms, to elucidate

what impacts these varied materials may have on the initial colonization of an anthropogenic sea defense structures.

Materials and methods

Substrata used as proxies for potential seawall materials

To determine the effect that stone type has on marine biofilm formation on a tropical seawall, eight different substrata to monitor biofilm community formation were selected to represent salient potential construction materials, encompassing aluminosilicates and calcium carbonates, respectively. These were: granite (GRANITE), limestone (LIMESTN), marble (MARBLE), pink sandstone (SANDSP), violet sandstone (SANDSV), pH neutralized concrete (CONCTN), rough textured concrete (CONCRO), and smooth textured concrete (CONCTR). These were chosen to represent salient natural stone categories, with differing surface structure, mineral, and chemical compositions. They were selected with uniform surface appearance, on a cm-scale, purchased from a local stone mason (Hot Spring Stone Pte. Ltd., Singapore). The surfaces were saw-cut without polishing, thus mitigating surface roughness variations and avoiding alteration of their surface chemistry, respectively. Stone samples were used as 5 cm × 5 cm × 2.5 cm coupons and concrete coupons were prepared by pouring concrete premix into set molds of the same dimensions. Concrete grain composition was 80% sand and 20% cement, resulting in a mortar comparable to the standards employed in masonry. The intrinsic alkalinity of as-formed concrete, with pH ≈ 13, can be mitigated by carbonation. This was achieved by exposing the concrete coupons to a saturated CO_2 environment, generated within a sealed vessel by sublimating dry ice, for 1 month (Hsiung et al., 2020). This process confers pH neutrality to the surface region, and progressively penetrates into the bulk volume of the concrete.

The stone surfaces were examined by scanning electron microscopy (SEM) using a JEOL JSM-7600F field emission scanning electron microscope operating at an accelerating voltage of 15 kV. Each was sputter coated with gold using a JEOL JFC-1200 fine coater and its surface morphology was imaged, together with an elemental cartography. Surface elemental and compositional analysis was performed with an Oxford Energy Dispersive X-Ray Analysis (EDX) system. Internal surface area, pore volume, and overall porosity measurements were obtained via Brunauer–Emmett–Teller (BET) nitrogen adsorption, using a Micromeritics ASAP 2020 system. Pore distribution was obtained from mercury intrusion, using a MicroActive Autopore V9600 mercury porosimeter. Average surface roughness (Ra) was measured over a 5 mm length using a TENCOR P-10 Surface Profiler at a scan speed of 50 μms^{-1}, with a sampling rate of 100 Hz. Thermal conductivity was measured with a NETZSCH HFM 436 heat flow meter.

Natural biofilm generation— Amplicon sequencing experiment

To develop natural marine biofilms on the different stone types, each coupon was sterilized in an autoclave (121°C for 15 min), submerged vertically in one of two outdoor flow-through aquaria to mitigate sedimentation and were supplied with a constantly renewed supply of sand filtered natural seawater for 1 month. Each aquarium held 300 L of seawater and was fed at a flow rate of 15 L min^{-1}, resulting in three water exchanges per hour. This exchange rate was chosen to mitigate any effects from leachates from the stone coupons during the incubation. Temperature and LUX intensity logs of the water and air (<5 cm above water line) were recorded each hour for 31 days, using Hobo data loggers (Onset, 8k data logger pendant).

To monitor biofilm development over the 1-month period, sacrificial coupons of all eight substrata types were removed at six time points (after 3 h and 1, 3, 7, 14, and 31 days) for examination. The experiment was divided over two aquaria, fed from the same seawater supply, with two replicates of each treatment were present in each tank, for a total of four replicates overall for each stone type at each time point (192 coupons in total). The position of each coupon within each aquarium was randomized, using the "sample" function in R base, to mitigate any positional bias within the aquaria with respect to biofilm growth. To determine the variability of the microbial community in the sand filtered seawater, 50 mL of seawater was collected at each time point to act as a background control. These aliquots were filtered onto a 0.22 μm polycarbonate filter and subsequently stored at −20°C.

Natural biofilm generation—Metagenome sequencing experiment

This experiment was repeated 12 months later using a similar experimental design to that employed for amplicon sequencing. The second run did not include concrete coupons and sacrificial sampling occurred weekly (weeks 1, 2, 3, and 4). One reason for omitting the concrete coupons was the emergence from our trials with chemically altered concretes that their differences appear to have little effect on biological diversity at the macro scale (Hsiung et al., 2020). Glass microscope slides were also included during this experiment to allow comparison with other studies which employ glass substrata as controls. All analyses involving the control glass slides were standardized to account for the difference in area between the glass and stone substrata. In addition, total nitrogen

(TN), total organic carbon (TOC), and total carbon (TC) from the aquaria were measured each week for the duration of the experiment (TOC-L analyzer, Shimadzu, Singapore), from these data the total inorganic carbon (TIC) was calculated. The water chemistry was recorded each week using a YSI ProQuatro meter (YSI; Ohio, United States). The primary purpose of this second experiment was to enable metagenomic sequencing of the biofilms. All treatments were performed in triplicate for this experiment.

Biofilm collection and nucleic acid extraction

To assess the biodiversity of the biofilms across time, the stone coupons were removed from the aquaria at the time points indicated above. The stones were removed from the aquaria and the biofilm sampled from a single 5 × 5 cm face. A sterile cotton swab was used to collect the biofilm material formed on the surface. These were immediately labeled and frozen at −20°C, before being transported to the laboratory in a liquid nitrogen cooled dry shipper and then stored at −20°C until nucleic acid extraction was performed.

To extract and isolate the nucleic acids, individual cotton swab heads were removed and placed into a Lysis Matrix E tube, before extraction with the FastDNA soil extract kit (MP Biomedical, Singapore), which was implemented according to the manufacturer's instructions. The resulting total nucleic acid extractions were quantified using a NanoDrop spectrophotometer, before being stored at −20°C for downstream analysis. These concentration data were employed as proxies for the general biomass of the biofilms and cannot distinguish between the main taxonomic groups (Bohórquez et al., 2017). For the background controls, the seawater was filtered and each filter was cut into half and placed in a sterile microcentrifuge tube, before being frozen in liquid N_2 and then milled down to a powder, which was then subjected to the same process of extraction as the cotton swabs. All extractions were stored at −20°C, for downstream analysis.

Nucleic acid amplification and MiSeq amplicon library preparation

Two sets of primers were used to target the 16S rRNA (515f-806r; V4 region bacteria and archaea) (Caporaso et al., 2011) and the 18S rRNA (Euk 1392f—Euk Br; V9 region) (Medlin et al., 1988; Lane, 1991) to generate libraries for amplicon sequencing. For each reaction for the 16S rRNA gene, 0.5 μL of each primer (10 μM) was added to 10 μL of Jumpstart Taq Ready Mix (#P2893-100RXN, Merck, Singapore). Additionally, 1 μL of gDNA template was added to each reaction and then made up to 20 μL using sdH_2O. For each reaction for the 18S rRNA gene, 1.0 μL of each primer (10 μM) was added to 12 μL of Jumpstart Taq Ready Mix. Additionally, 1 μL of gDNA template was added to each reaction and made up to 25 μL using sdH_2O. The amplification condition for the 16S rRNA libraries were 94°C for 3 min initial denaturation, followed by 35 cycles of 94°C (45 s), 53°C (60 s), and 72°C (90 s); concluding with 72°C for 10 min. Amplification of the 18S rRNA libraries was achieved by 94°C for 3 min initial denaturation, followed by 35 cycles of 94°C (45 s), 57°C (60 s), and 72°C (90 s), concluding with 72°C for 10 min. Upon amplification, sequencing libraries were generated using the TruSeq kit v2.0 and paired-end MiSeq sequencing was conducted.

Amplicon sequence quality, assembly and identification

Quality filtration of the raw sequences was performed using the DaDa2 (Callahan et al., 2016) protocol, within an R environment (R Core Team, 2014). Briefly, all sequences were observed for QC Phred scores above 25 and bases trimmed from the start (10 bp) and end (50 bp) of the fragments. Subsequently, these sequences were assembled and processed using DaDa2 to obtain chimeric free amplicon single nucleotide variants (SNVs). Prokaryotic and eukaryotic sequences were identified through comparison to the Silva v.132 Ribosomal database. Each identified SNV required a minimum similarity of 80% in order to be classified as a particular taxon, otherwise it was labeled as unclassified. Rarefaction plots indicating sampling effort are available (**Supplementary Figure 1**).

Metagenome sequencing library

Total nucleic acid extractions from the second experiment were quantified using a Qubit HS (high sensitivity) dsDNA kit. Sequencing library preparation was achieved using the Swift Biosciences Accel-NGS 2S Plus DNA Kit, as per the manufacturer's instructions. Using a Covaris S220 ultrasonicator, fragments were sheared to ∼450 bp, prior to adding barcodes using the Swift Biosciences 2S dual indexing kit. The subsequent libraries were checked for fragment length using a Bioanalyser DNA 7500 chip and quantification was performed using a picrogreen fluorescence assay. Libraries were then standardized to 4 nM as verified by qPCR (Kapa Biosystems Library Quantification kit, Applied Biosciences). An equimolar pool of all samples was generated and sequenced on Illumina HiSeq 2500 rapid runs (10–11 pM: V2 rapid sequencing chemistry), yielding reads of 251 bp paired-end sequences.

Metagenomic sequence analyses

Shotgun sequencing reads were assessed for quality and adapters were removed using the Trim Galore package

(v0.6.4_dev). Low quality ends were removed from sequences with a phred score lower than 20. Paired reads were merged using Pear (v0.9.6), using default settings. The resulting sequences were trimmed to between 200 and 300 bases before being randomly sub-sampled, which resulted in 750,000 reads per sample. All sub-sampled reads were subjected to taxonomic identification by comparison to the Maxikraken2 database (Loman, 2020), using the Kraken2 package (v2.0.8_beta) and visualized as Sankey plots using R (Pavian Package). Independent of the taxonomic identification, the raw reads were *de novo* assembled using spades (v3.13.0), thus generating contigs. Gene annotation of the contigs was conducted using Prokka (1.13) and these were examined for functional identification using Microbeannotator (v2.0.5) which annotated the proteins against the KOfam and RefSeq databases, before being searches and compared again the KEGG database (release 551) using the KOfamscan package (Kanehisa et al., 2016; Kanehisa and Sato, 2019). From these metagenomic sequencing reads, all rRNA reads were identified and classified using sortmerna package (v4.3.4) and Kraken2 (v2.0.8_beta) package, both equipped with the Silva v.132 database used for the amplicon experiments. Rarefaction plots indicating sampling effort of all rRNAs are available (**Supplementary Figure 3**).

Statistical analysis of results

Data were processed and visualized in R Core Team (2014) using the vegan (Oksanen et al., 2011), complexheatmap (Gu et al., 2016), and ggplots (Wickham, 2016). Briefly, all absolute values of abundance were standardized to relative abundance for SNVs. Statistical analysis was performed using Permanova (Anderson, 2001) (adonis; vegan) to determine community differences and univariate differences between data were assessed using ANOVA or ANCOVA analyses (aov; vegan). Potential functional differences were assessed using ANOSIM to determine if identified complete KEGG pathways differed between treatment types. *Post hoc* analyses were conducted using the Tukey test (Tukey, 1949) of difference, to determine variables that differ significantly ($p \leq 0.05$). Alpha-diversity was assessed using Shannon-Weiner indices (Shannon, 1948). All data was arcsine transformed prior to statistical assessment, to ensure the assumptions needed for parametric analyses were met where appropriate.

Data availability

Data generated during this study is available within the **Supplementary Information**. All sequences are available from the NCBI, Sequence Read Archive (BioProject: PRJNA698754).

Results

Biofilm physico-chemical environment

Three of the five natural stone substrata examined were classed as aluminosilicates (Al_2SiO_5; SANDSP, SANDSV, and GRANITE), while the remaining were calcium carbonates (LIMESTN and MARBLE). The elemental composition of LIMESTN was predominately Ca, C, and O, with trace amounts of Mg (0.02%) and Ni (0.1%) (**Table 1**). While broadly similar to LIMESTN, MARBLE contained a lower fraction of Ca, together with higher percentage (11.5%) of Mg and trace amounts of Co and Al. SANDSP, SANDSV, and GRANITE contained Si and Al, as anticipated for aluminosilicates, with higher fractions of Mg in SANDSV (3.64%) and GRANITE (1.72%). In GRANITE, Fe generally appeared to be co-localized with Mg. Al and Na were similarly co-located, but in complementary locations (**Supplementary Figure 3**). A similar trend was observed for Al and Na, in SANDSP. SANDSV presented comparatively small, localized regions of Ca, which are presumed to indicate the presence of incorporated calcium carbonate fragments. Si and O, associated with silica moieties, were exhibited uniformly over the natural stone aluminosilicates. In contrast, CONCRO samples showed Si in locations complementary to Ca.

While natural stone coupons were cut with identical tooling, their surfaces exhibited different roughness. This is attributed to differences in hardness and their exposed porosity. The structure of LIMESTN presents a fractional pore volume over 11% (**Table 2**), with a larger fraction in the sub-micron range ≈280–500 nm and a smaller proportion over ≈50–250 μm. This differs substantially from MARBLE, which presents a 10-fold lower overall porosity (≈ 1%) that occurs primarily around 250 nm and over a larger size range that is similar to LIMESTN ≈65–270 μm. These ranges were estimated from the width at half of the maximum of pore size (**Supplementary Figure 2**). SANDSP and SANDSV both present fractional pore volumes of the same order as limestone, with a significant proportion occurring in the range 150–500 nm. They contrast in their larger pore dimensions, with SANDSP exhibiting a narrow peak around 7.9 μm, while SANDSV offers a broad distribution, from ≈30 to 300 μm. GRANITE offers a comparatively closed structure, with a fractional pore volume that is similar to MARBLE. This appears as a broad distribution above 150 nm and larger size pores with a peak at 4.5 μm. Thermal conductivity was similar across the natural stones.

The environmental parameters of the aquaria hosting the stone coupons during biofilm growth showed some variability for the duration of the first experiment. Air and water data logs for temperature and LUX intensity indicated measurable diurnal fluctuations of both parameters, with light levels largely consistent and water temperature levels dropping over the course of the experiment (**Supplementary Figure 5**). For example, there was a temperature range of 5.68°C (max:

TABLE 1 Surface elemental composition of the treatment stone types obtained from SEM/EDX.

Substratum	Ca (%)	C (%)	Al (%)	Si (%)	O (%)	Fe (%)	Mg (%)	Na (%)	K (%)	Ti (%)	Ni (%)	Co (%)	S (%)
LIMESTN	41.62 (1.27)	13.06 (0.68)			45.18 (0.57)		0.02 (0.05)				0.10 (0.14)		
MARBLE	21.52 (1.25)	17.18 (0.38)	0.04 (0.09)		49.56 (1.04)		11.5 (0.45)					0.12 (0.08)	
SANDSP	0.32 (0.05)	8.96 (0.37)	3.28 (0.26)	35.06 (0.65)	48.32 (0.65)	0.56 (0.06)	0.30 (0.10)	0.04 (0.06)	3.08 (0.34)	0.06 (0.13)			
SANDSV	1.38 (0.08)	8.28 (0.85)	6.62 (0.23)	24.00 (0.50)	45.02 (0.91)	5.66 (0.23)	3.64 (0.39)	3.48 (0.19)	1.58 (0.18)	0.28 (0.05)	0.04 (0.06)	0.02 (0.05)	
GRANITE	6.72 (1.08)	9.90 (1.59)	10.26 (1.35)	22.42 (2.31)	37.76 (4.13)	7.82 (5.43)	1.72 (1.42)	1.88 (0.37)	0.40 (0.24)	0.62 (1.39)	0.10 (0.14)	0.06 (0.06)	0.06 (0.06)
CONCRO	14.2 (5.44)	10.7 (3.18)	1.4 (0.07)	20.6 (8.49)	51.4 (5.73)	0.7 (0.21)	0.3 (0.11)		0.2 (0.07)				0.4 (0.28)
CONCTR	33.16 (1.11)	13.46 (1.63)	0.82 (0.18)	3.38 (1.94)	45.05 (3.62)	0.72 (0.04)	0.74 (0.42)	0.22 (0.20)	0.42 (0.33)			0.03 (0.06)	0.2 (0.28)
CONCTN	25.47 (0.80)	20.30 (2.31)	0.90 (0.26)	1.7 (0.3)	48.67 (2.84)	0.63 (0.25)	0.73 (0.38)	0.27 (0.12)	0.23 (0.15)			0.10 (0.01)	0.3 (0.1)

Each element has been listed as the w/w percentage fraction of the sum of the detected elements. Abbreviations are used in subsequent tables and figures to indicate each stone type. Values presented indicate the measured mean with standard deviations in parentheses.

TABLE 2 Physical measurements of the five stone substrata employed in this study.

Substratum	Thermal Conductivity ($W\ m^{-1}\ K^{-1}$) @ 30°C	Internal surface area (m^2/g)	Internal pore volume (mm^3/g)	Overall porosity (%)	Surface roughness (R_a, μm)
LIMESTN	0.44	0.099	8.27	6.95	2.21
MARBLE	0.45	0.0025	0.79	0.43	1.79
SANDSP	0.40	0.98	61.87	31.91	8.07
SANDSV	0.41	2.02	6.83	19.29	6.56
GRANITE	0.45	0.13	0.96	6.28	3.30
CONCRO		4.05	19.38	4.44	25.15
CONCTN		3.79	20.12	4.63	6.81
CONCTR		3.34	21.29	4.86	11.84

Note concrete coupons were not able to be measure for thermal conductivity using our laboratory protocols/equipment.

31.78°C, min: 26.10°C) in the air temperature over the course of the experiment, and the daily mean air temperature varied by 3.10°C. The water temperature showed an overall range of 2.39°C (max: 30.05°C, min: 27.67°C). The water temperature had a maximum diurnal variability of up to 1.20°C per day. However, the mean diurnal temperature change was 0.67°C. Over the course of the experiment the mean daily temperature of the water lowered by 1.85°C.

For the metagenomic experiment, the water samples tested throughout the experiment indicate some nutrient enrichment of inorganic carbon (Table 3). In addition, there was a reduction in TIC for weeks 2 and 3. TN did not follow the same pattern of fluctuations as the carbon.

Biofilm analyses

The amplicon experiment (with time points at 3 h and 1, 3, 7, 14, and 31 days) did not yield sufficient nucleic acids to provide a confident assessment of biomass. However, the metagenomic experiment yielded measurable nucleic acid concentrations at each of the four time points (Figure 1). Data indicate that there was no significant difference in biomass among substratum types [ANOVA, $F_{(5,120)} = 1.551$, $p = 0.179$] or time points [ANOVA, $F_{(3,120)} = 0.551$, $p = 0.649$]. However, while a visual inspection of these data suggested that the biomass on glass was lower than that of calcium carbonate-based substrata (MARBLE and LIMESTN), pairwise analyses revealed no significant differences (Tukey HSD, $p = 0.121$ and $p = 0.251$, respectively).

Amplicon sequences

From the 192 samples processed, 190 produced suitable sequencing libraries, with two SANDSP coupons (days 3 and 31) failing to amplify. For the remaining 190 samples from the first experiment, there was no significant influence of substratum

TABLE 3 Carbon and Nitrogen measure taken of the water present in the aquaria at the time of each sampling point.

	T_0	Week 1 (ppm)	Week 2 (ppm)	Week 3 (ppm)	Week 4 (ppm)
TC	14.82	14.89	18.06	17.68	17.24
TIC	13.44	13.74	17.76	17.71	16.13
TOC	1.38	1.15	0.30	0.03	1.11
TN	0.13	0.08	0.10	0.07	0.15
Conductivity(ms/cm)	30.67 (0.42)	25.92 (0.55)	36.56 (0.31)	33.87 (0.37)	38.08 (0.16)
Salinity (ppt)	27.59 (0.24)	30.66 (0.46)	27.73 (0.06)	29.44 (0.05)	29.39 (0.05)
DO (mg/L)	6.08 (0.27)	8.01 (0.48)	2.99 (0.28)	4.45 (0.53)	3.53 (0.65)
ORP (mV)	126.7 (8.65)	135.53 (9.06)	183.38 (8.3)	166.13 (4.97)	154.23 (3.36)

Probe measurements of conductivity, salinity, DO, and ORP taken at the time of sampling. Standard deviation available in parenthesis. All data is in parts per million and written to 2 d.p. unless specified.

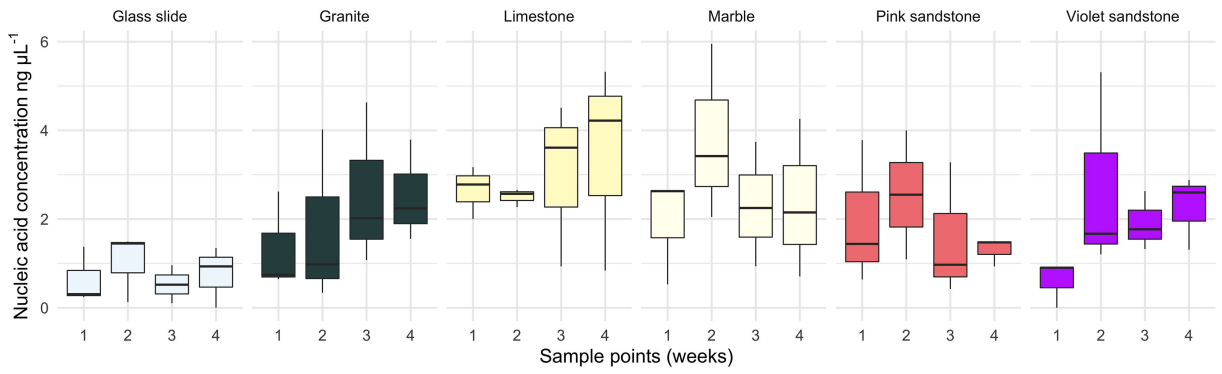


FIGURE 1
Box and whisker plot of the biofilm biomass estimates based on total nucleic acid concentration extracted from each time point grouped by substratum. Measured values were obtained using a Qubit fluorescence HS dsDNA kit.

on the Prokaryotic composition based on 16S rRNA marker genes [PERMANOVA, $F_{(7,181)} = 1.203$, $p = 0.066$; Figure 2). However, the age of the biofilm resulted in significant differences in bacterial composition [PERMANOVA, $F_{(1,189)} = 13.782$, $p = 0.001$]. Pairwise investigation of these data indicates that no single substratum was significantly different in overall microbial community composition to any other substratum. Conversely, biofilm age had a significant effect on the microbial composition for all time point pairwise comparisons, with the exception of 3 h vs. 14 days and 7 days vs. 14 days which showed no statistical difference in microbial communities at these times.

Univariate analysis of the effect of biofilm age and substratum on the biofilm's α diversity also showed no significant influence of substratum [ANOVA, $F_{(7,182)} = 0.746$, $p = 0.634$]. It was apparent that α diversity increased over time [ANOVA, $F_{(1,188)} = 31.87$, $p < 0.001$], thought to result in the similarities in β diversity observed for all samples at the end of the experiment (31 days). Analyses of the seawater controls indicated that there was no change in Bray-Curtis variability over the course of the experiment and neither did microbial diversity differ [ANOVA, $F_{(1,6)} = 1.603$, $p = 0.29$], indicating that no measurable outside influences impacted the experimental design.

The overall bacterial composition of the biofilms was dominated by a low abundance of single nucleotide variants (SNVs; Supplementary Figure 6). Most of the identified SNVs contributed < 1% of the overall relative abundance for any given sample. However, a few key taxa were dominant (high mean relative abundance) across all treatments. These included SNVs identified as *Colwellia* (5.3%), *Escherichia-Shigella* (4.3%), *Neptuniibacter* (3.2%), *Thalassotalea* (3.0%), and *Rhodobacteraceae* (2.9%); with *Thalassotalea* being the highest individual fraction (>80%) observed for one of the SANDSP replicates. While the overall bacterial community composition was not significantly different among substrata, 87 of the 3,631 individual SNVs identified in this study were significantly different among substratum type. These differences appeared to be primarily dominated by larger relative abundances of the SNVs on CONCRO (Supplementary Figure 7).

Examination of the 18S rRNA marker genes revealed no difference in Eukaryotic communities associated with the substratum material [PERMANOVA, $F_{(9-95)} = 0.9125$, $p = 0.691$]. There was, however, a significant difference

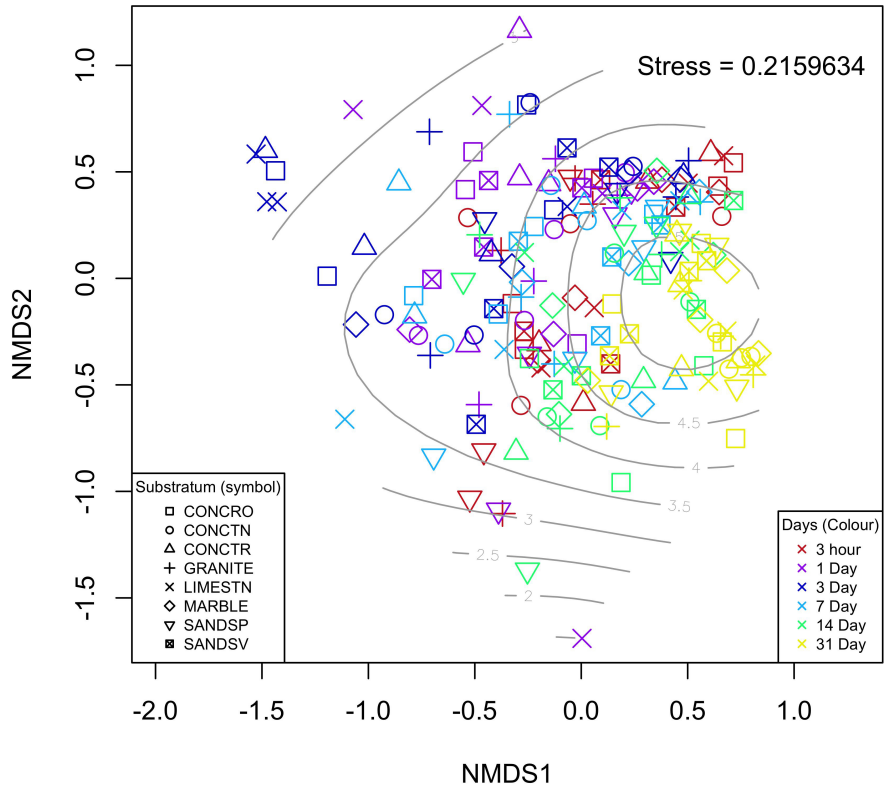


FIGURE 2
nMDS ordination of the individual samples used for 16S rRNA gene SNV analyses. Each substratum is indicated by a different symbol, whereas biofilm age is represented by color of the points on the plot. Light gray contour lines are indicative of Shannon-wiener diversity indices for each sample.

over time within the Eukaryotic communities [PERMANOVA, $F_{(7-95)} = 5.906$, $p = 0.001$]. The 18S rRNA gene dataset yielded 4,419 specific SNVs across all treatments. From the 20 most abundant SNVs, it was clear that time was a contributing factor for their distribution (**Supplementary Figure 8**). From this subset (61.3% of total) of Eukaryotic SNVs, invertebrates contributed significantly to the SNVs, with *Copepoda* being responsible for > 27% of the total SNVs. Representatives of the *Choradates*, *Vertebrata* (4.53%), and *Asciidians* (4.38%) were also present. In addition, there was evidence of fungi (*Basidiomycota*; 3.48%), although the oligonucleotides used in this study were not optimized for fungal identification. In addition, the dominance of *Copepoda* 18S rRNA gene markers in the Eukaryotic fraction, is likely due to a high number of gene copy numbers within the genome his (Wyngaard et al., 1995).

Metagenomic sequence analyses

The metagenomic experiment revealed that the metagenome and amplicon data differed with respect to the dominant taxa in the Prokaryotic and Eukaryotic groups. This difference was likely due to amplification bias from the oligos used in amplicon sequencing or the gene copy number from the rRNA gene markers monitored. For example, amplicon analyses of the Prokaryotic fraction of the biofilm indicated a dominance of *Colwellia* spp., while the metagenomic data showed that *Rhodobacterales* was the most dominant of the contigs. This was also evident in the eukaryotic fraction, with *Copeopoda* being abundant in amplicon sequences, yet absent from the common metagenome sequences. All rRNA identified from the metagenome sequencing was compared to the rRNA data from the amplicon sequencing experiments. From these data we concluded that the most abundant organisms found in the amplicon experiments were also present in the metagenome experiment, with only relative abundances differing (**Supplementary Figure 9**). This abundance difference was expected due to different library preparation techniques employed for the two experiments. Assessment of the alpha diversity of these *in silico* obtained rRNA genes was performed by ANCOVA and we observed that the substrata had no measurable impact [ANCOVA, $F_{(5)} = 0.98$, $p > 0.05$] on diversity, whereas the age of the biofilm did [ANCOVA, $F_{(3)} = 3.23$, $p = 0.028$]. In addition to the differences between contigs and amplicon assessments, the use of contigs enabled us to determine the relative contribution of Eukaryotic and

Prokaryotic fractions. While Eukaryotic contigs were observed, their contribution was minimal compared to Prokaryotic organisms (Bacteria ∼ 98%, Archaea ∼ 0.2%, Eukaryotes ∼ 0.8%, and Viruses ∼ 0.07%; **Supplementary Data**—taxa abundance tables).

Despite these subtle differences in biofilm community identifications, the metagenome experiment once again showed that biofilm age was a driving factor in biofilm composition [PERMANOVA, $F_{(5-71)} = 3.45$, $p < 0.001$], as was previously observed in the amplicon experiment. In contrast to the amplicon experiment, we found a significant effect of substratum on biofilm communities. However, this was only between the glass control substratum and the other stone types—there were no significant differences amongst the experimental stone types and biofilm community (**Figure 3**).

Functional assessment of contigs

From the contigs there was a total of 3.9 M identified and hypothetical proteins annotated. When pooled for each treatment type across the four time points measured, these yielded 160 complete identified KEGG pathways. Of these, 81 KEGG pathways occurred within biofilms from all substrata. As with the taxonomic comparison, functional diversity also showed that the age of the biofilm had a statistically significant impact on the KEGG modules identified as complete (ANOSIM, $R = 0.14$, $p = 0.01$). However, substratum type had no significant effect (ANOSIM, $R = 0.03$, $p > 0.05$). Some KEGG modules which were common, but not shared across all samples, included Lipid A biosynthesis (M00866), Galactose degradation (M00632), and Thiosulfate oxidation (M00595). Some of the less common KEGG modules identified include Thyroid hormone biosynthesis (M00043), Glycosaminoglycan biosynthesis (M00058), and Lactosylceramide biosynthesis (M00066). As glass was the most dissimilar, significant differences in enrichment were observed for this substratum. Biofilms on glass slides were enriched for aromatic amino acid degradation while genes associated with assimilatory nitrate reduction (i.e., M00531) exhibited reduced abundance. As the biofilms increased in age, we also observed changes in the KEGG module enrichment, with initial (week 1) biofilms being enriched for salicylate (i.e., M00638). No enrichment of any KEGG module was observed for weeks 1 and 3. Four modules were enriched for weeks 1 and 4: homoprotocatechuate degradation, methane oxidation, ammonia nitrification, and comammox (i.e., M00533, M00174, M00528, and M00804, respectively). A full list of the identified modules, together with their level of completeness, can be found in **Supplementary Figure 7**.

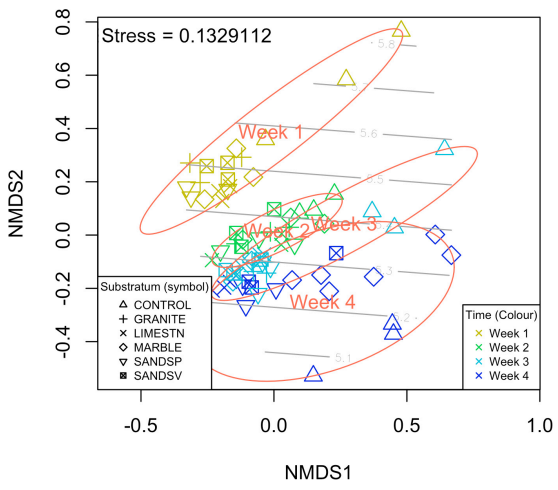


FIGURE 3
nMDS plot shows that that α and β diversity of the meta-contigs clearly separate over the *y*-axis suggesting that biofilm age is a factor of interest. Colors signify biofilm age: Gold = Week1, Green = Week2, light blue = Week 3, and dark blue = Week 4. Light gray contour lines indicate the Shannon Wiener diversity (H′).

Discussion

The construction of seawalls is often the go-to defense against erosion and flooding, especially in coastal cities, hence finding ways to mitigate their negative impacts on species' biodiversity is critical to ecological functioning and sustainable development (Firth et al., 2014; Cole et al., 2015). The current study aims to determine whether seawall construction materials affect the initial biofilm formation, with a view to identifying stone types that can provide ecological uplift. Our results show that biofilms changed, in both taxonomic composition and functional potential, over the initial 4 weeks of growth. However, these were not significantly affected by the stone type in either of the two experiments; findings that align with a recent seawall study examining colonization by macrobiota on similar materials (Hartanto et al., 2022). In contrast, smooth glass supported a different biofilm community, as compared with the natural stones. It was also observed that some SNVs were significantly more abundant on rough concrete surfaces.

The calcium carbonate-based substrata supported a consistently high biomass throughout the experiment, while the glass slides induced growth of a consistently lower biomass biofilm. This is congruent with studies describing stronger adsorption of organic matter on calcium carbonate than on silica (Carter, 1978; Zeller et al., 2020). Moreover, glass slides had an intrinsic surface roughness approximately three orders of magnitude lower (R_a ∼ 3–5 nm) than that of the stones used in this study. This difference in roughness may limit the ability of biofilms to bind strongly onto the smoother surfaces (Encinas et al., 2020). This argument is supported by the increased relative abundance of some SNVs, as found on the rougher CONCRO samples (**Supplementary Figure 2**).

Indeed, the latter substratum harbored many of the SNVs that differed significantly among substrata. However, some studies show that larger organisms such as diatoms have the opposite relationship with surface roughness, with rough surfaces yielding lower diversity and abundance (Sweat and Johnson, 2013). Previous research from other aquatic environments, such as rivers, have also shown that surface roughness is a driver of biofilm composition and diversity (Schneck et al., 2011), as well as augmenting the biomass of benthic algae (Murdock and Dodds, 2007). Furthermore, sandstone's substantial fraction of micron-scale porosity is expected to facilitate the penetration of filaments and enhance water retention, leading to its propensity for supporting the growth of foliose algae (Green et al., 2012).

Porosity measurements indicate pore sizes that are compatible with promoting microbial adhesion, namely above 50 nm and up to several hundred nanometers (Feng et al., 2015). Granite presented almost no pores with sub-μm size and a narrow size distribution around 4.5 μm, together with a low (≈1%) fractional porosity. LIMESTN, SANDSP, and SANDSV incorporate significant porosity, ranging over sizes from μm to sub-mm. Their interior may thus potentially harbor micro-organisms, but these were not harvested and analyzed in the present study.

Biofilm diversity (α and β) did not expose significant differences for the tested substrata, thus implying the absence of surface-induced biofilm variations. While it is probable that surface roughness and the settling organism's size and shape may contribute to driving the recruitment process of these initial colonizers (Lu et al., 2016; Huang et al., 2018), the homogeneity in marine microbial settlement contrasts with the diverse surface chemistry of stone surfaces, as broadly classified into calcium carbonate, aluminosilicates, and concrete (**Table 1**). Moreover, Huggett et al. (2009) found that marine biofilm taxonomy converged on similar substrata, but that invertebrate settlement was impacted by wettability (amongst other factors). Therefore, it is likely that surface roughness, surface chemistry, and sub-surface porosity in relation to the settling organism's size and shape are key characteristics driving the recruitment process of these initial colonizers (Lu et al., 2016; Huang et al., 2018).

The overall diversity (α and β) of the biofilms appears not to have been driven by the stone materials, but rather regulated by the age of the biofilm. In total, we observed 863 16S SNVs (2.4% of total) that where significantly different the time periods sampled. Of the SNVs most enriched were candidates from *Gemmatimonadetes*, *Calditrichaceae*, *Entotheonella*, *Microtrichales*, *Nitrsococcaceae*, *Lentisphaerae*, and four uncultured bacteria. While the average water temperature did decline throughout the experiment, this was by only ∼1.84°C (mean daily average). There was no significant change in LUX over the duration of this experiment. Data showed (**Figure 2**) that, as the biofilms age, α diversity indices increased and the dissimilarly between biofilms on each substratum (β diversity) reduced, in the case of amplicon sequences. This increase in α diversity and decrease in β diversity is also observed in similar on-going field experiments in the region (data unpublished). However, the α diversity of the metagenome samples only showed minor fluctuations. This is most likely due to the lack of high frequency sampling in the first weeks, such as was employed for the amplicon experiments. This indicates a more homogenous biofilm with increased convergence in identified SNVs and identified contigs. For the contigs the α diversity was most similar at weeks 2 and 3, which matched the reduction in TOC measures. This suggests that recruitment from the water column may be a factor in regulating the α diversity. The β diversity changes on the glass controls (open triangles in **Figure 3**) were subtly different from other substrata of an equal age, supporting the observation that the (sub-nm) roughness of the glass may be more important than the chemical composition of the glass when compared to the silicate stones (SANDSP and SANDSV).

In addition to taxonomic diversity, we examined functional traits. Given the complexity in microbial systems, and that functional diversity does not have a global definition (Escalas et al., 2019), we opted for trait-based richness as a measure of functional potential. We showed that the age of the biofilm was a driver for changes in functional trait richness, although *post hoc* analyses show that it was only samples from weeks 1 to 3 that differed significantly. Both the functional and taxonomic differences observed could also be influenced by lower TOC levels at week 3, which may have altered recruitment. However, our data do not determine this conclusively (**Supplementary Figure 10**).

By observing enrichment of functional trait richness, the differences between the time points was associated with enrichment of specific KEGG modules only for weeks 1 and 4. This enrichment could have implications with respect to the function of the biofilm community and the efficacy by which invertebrate larvae settle on the biofilms. For example, the presence of methanotrophs in weeks 1 and 4 may reduce the settlement of some coral larvae as it has been shown that the presence of methanotrophs deter larval settlement (Sneed et al., 2015). Nevertheless, Petersen and Dubilier (2009) showed that methanotrophy can encourage symbiosis between invertebrates and bacteria, suggesting that not all invertebrates would be repelled by these bacteria. A similar scenario is also likely for coral settlement with regards to nitrogen fixation, as N_2 is a major factor in coral health (Rädecker et al., 2015) and any changes in the nitrogen cycling ability of the biofilm may impact the settlement of coral larvae. Therefore, the enrichments observed for weeks 1 and 4, as well as the reduction observed on the glass substrata, could impact what types of invertebrate larvae choose to settle on these surface biofilms. Furthermore, the increase in biomass over time observed for some of the substrata may further enhance invertebrate settlement as it has been shown that increases in settlement occur with high

bacterial cell numbers (Lema et al., 2019). However, it is unclear from these data whether this was an influence of biomass alone or a change in function that would likely result from a higher cell density within the biofilm.

Conclusion

The initial aim of this study was to determine what effect different seawall building material would have on the taxonomic and functional diversity of a typical marine biofilm during early settlement. We discovered, however, that differences in material surface chemistry and topography had no measurable lasting impact on the biofilms. Only glass, which was orders of magnitude smoother, and the rougher concrete substratum, had any significant influence on the biofilm composition, and only during the initial stages. The overall observation was that the marine biofilms converged over time to a generic marine biofilm, rather than the underlying stone substrata type playing a significant role in driving distinct 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 in the article/**Supplementary Material**.

Author contributions

SS, YP, DM, PT, WB, and SR equally contributed to the design of this project. SS, YP, and DV generated all data and performed analyses. All authors contributed to the writing of this article.

Funding

This research was funded by the National Research Foundation, Prime Minister's Office, Singapore under its Marine Science Research and Development Programme (Award No. MSRDP-P05). We acknowledge financial support from the Singapore Centre for Environmental Life Sciences Engineering, whose research is supported by the National Research Foundation Singapore, Ministry of Education, Nanyang Technological University and National University of Singapore, under its Research Centre of Excellence Programme. We also acknowledge financial support from the Institute of Materials Research and Engineering (IMRE) under the Agency of Science, Technology and Research (A*STAR).

Acknowledgments

We would like to thank the staff at St Johns Island National Marine Laboratory, Singapore for their help in facilitating these experiments.

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/fmicb.2022.928877/full#supplementary-material

References

Anderson, M. J. (2001). A new method for non-parametric multivariate analysis of variance. *Austral Ecol.* 26, 32–46. doi: 10.1046/j.1442-9993.2001.01070.x

Birch, W., Carré, A., and Mittal, K. L. (2008). "Wettability techniques to monitor the cleanliness of surfaces," in *Developments in Surface Contamination and Cleaning*, eds R. Kohli and K. L. Mittal (Netherlands: Elsevier), 693–723. doi: 10.1016/B978-081551555-5.50015-0

Bohórquez, J., Mcgenity, T. J., Papaspyrou, S., García-Robledo, E., Corzo, A., and Underwood, G. J. (2017). Different types of diatom-derived extracellular polymeric substances drive changes in heterotrophic bacterial communities from intertidal sediments. *Front. Microbiol.* 8:245. doi: 10.3389/fmicb.2017.00245

Bruun, P. (1962). Sea-level rise as a cause of shore erosion. *J. Waterways Harbors Division* 88, 117–132. doi: 10.3390/antiox11020244

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. doi: 10.1038/nmeth.3869

Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Lozupone, C. A., Turnbaugh, P. J., et al. (2011). Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc. Natl. Acad. Sci. U. S. A.* 108, 4516–4522. doi: 10.1073/pnas.1000080107

Carter, P. W. (1978). Adsorption of amino acid-containing organic matter by calcite and quartz. *Geochim. Cosmochim. Acta* 42, 1239–1242. doi: 10.1016/0016-7037(78)90117-5

Chee, S. Y., Yee, J. C., Cheah, C. B., Evans, A. J., Firth, L. B., Hawkins, S. J., et al. (2021). Habitat complexity affects the structure but not the diversity of sessile communities on tropical coastal infrastructure. *Front. Ecol. Evol.* 9:673227. doi: 10.3389/fevo.2021.673227

Cockell, C. S., and Herrera, A. (2008). Why are some microorganisms boring? *Trends Microbiol.* 16, 101–106. doi: 10.1016/j.tim.2007.12.007

Cockell, C. S., Kelly, L. C., Summers, S., and Marteinsson, V. (2011). Following the kinetics: iron-oxidizing microbial mats in cold icelandic volcanic habitats and their rock-associated carbonaceous signature. *Astrobiology* 11, 679–694. doi: 10.1089/ast.2011.0606

Cole, M., Lindeque, P., Fileman, E., Halsband, C., and Galloway, T. S. (2015). The Impact of Polystyrene Microplastics on Feeding, Function and Fecundity in the Marine Copepod *Calanus helgolandicus*. *Environ. Sci. Technol.* 49, 1130–1137. doi: 10.1021/es504525u

Coombes, M. A., Naylor, L. A., Thompson, R. C., Roast, S. D., Gómez-Pujol, L., and Fairhurst, R. J. (2011). Colonization and weathering of engineering materials by marine microorganisms: an SEM study. *Earth Surf. Process. Landf.* 36, 582–593. doi: 10.1002/esp.2076

Dodds, K. C., Schaefer, N., Bishop, M. J., Nakagawa, S., Brooks, P. R., Knights, A. M., et al. (2022). Material type influences the abundance but not richness of colonising organisms on marine structures. *J. Environ. Manag.* 307:114549. doi: 0.016/j.jenvman.2022.114549

Douglas, B. C. (1991). Global sea level rise. *J. Geophys. Res.* 96, 6981–6992. doi: 10.1029/91JC00064

Dugan, J., Airoldi, L., Chapman, M., Walker, S., Schlacher, T., Wolanski, E., et al. (2011). 8.02-Estuarine and coastal structures: environmental effects, a focus on shore and nearshore structures. *Treat. Estuar. Coast. Sci.* 8, 17–41. doi: 10.1016/B978-0-12-374711-2.00802-0

Encinas, N., Yang, C. Y., Geyer, F., Kaltbeitzel, A., Baumli, P., Reinholz, J., et al. (2020). Submicrometer-Sized Roughness Suppresses Bacteria Adhesion. *ACS Appl. Mater. Interfaces* 12, 21192–21200. doi: 10.1021/acsami.9b22621

Escalas, A., Hale, L., Voordeckers, J. W., Yang, Y., Firestone, M. K., Alvarez-Cohen, L., et al. (2019). Microbial functional diversity: from concepts to applications. *Ecol. Evol.* 9, 12000–12016. doi: 10.1002/ece3.5670

Feng, G., Cheng, Y., Wang, S. Y., Borca-Tasciuc, D. A., Worobo, R. W., and Moraru, C. I. (2015). Bacterial attachment and biofilm formation on surfaces are reduced by small-diameter nanoscale pores: how small is small enough? *NPJ Biofilms Microbiomes* 1:15022. doi: 10.1038/npjbiofilms.2015.22

Firth, L. B., Thompson, R. C., Bohn, K., Abbiati, M., Airoldi, L., Bouma, T. J., et al. (2014). Between a rock and a hard place: environmental and engineering considerations when designing coastal defence structures. *Coast. Eng.* 87, 122–135. doi: 10.1016/j.coastaleng.2013.10.015

Freckelton, M. L., Nedved, B. T., and Hadfield, M. G. (2017). Induction of Invertebrate Larval Settlement; Different Bacteria, Different Mechanisms? *Sci. Rep.* 7:42557. doi: 10.1038/srep42557

Freckelton, M. L., Nedved, B. T., Cai, Y. S., Cao, S., Turano, H., Alegado, R. A., et al. (2022). Bacterial lipopolysaccharide induces settlement and metamorphosis in a marine larva. *Proc. Natl. Acad. Sci. U. S. A.* 119:e2200795119. doi: 10.1073/pnas.2200795119

Golledge, N. R., Keller, E. D., Gomez, N., Naughten, K. A., Bernales, J., Trusel, L. D., et al. (2019). Global environmental consequences of twenty-first-century ice-sheet melt. *Nature* 566, 65–72. doi: 10.1038/s41586-019-0889-9

Green, D., Chapman, M., and Blockley, D. (2012). Ecological consequences of the type of rock used in the construction of artificial boulder-fields. *Ecol. Eng.* 46, 1–10. doi: 10.1016/j.ecoleng.2012.04.030

Gu, Z., Eils, R., and Schlesner, M. (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* 32, 2847–2849. doi: 10.1093/bioinformatics/btw313

Hadfield, M. G. (2011). Biofilms and marine invertebrate larvae: what bacteria produce that larvae use to choose settlement sites. *Ann. Rev. Mar. Sci.* 3, 453–470. doi: 10.1146/annurev-marine-120709-142753

Hartanto, R. S., Loke, L. H., Heery, E. C., Hsiung, A. R., Goh, M. W., Pek, Y. S., et al. (2022). Material type weakly affects algal colonisation but not macrofaunal community in an artificial intertidal habitat. *Ecol. Eng.* 176:106514. doi: 10.1016/j.ecoleng.2021.106514

Heatherington, C., and Bishop, M. J. (2012). Spatial variation in the structure of mangrove forests with respect to seawalls. *Mar. Freshw. Res.* 63, 926–933. doi: 10.1071/MF12119

Higgins, C. G. (1980). Nips, notches, and the solution of coastal limestone: an overview of the problem with examples from Greece. *Estuar. Coast. Mar. Sci.* 10, 15–30. doi: 10.1016/S0302-3524(80)80046-6

Hsiung, A. R., Tan, W. T., Loke, L. H., Firth, L. B., Heery, E. C., Ducker, J., et al. (2020). Little evidence that lowering the pH of concrete supports greater biodiversity on tropical and temperate seawalls. *Mar. Ecol. Prog. Ser.* 656, 193–205.

Huang, Y., Zheng, Y., Li, J., Liao, Q., Fu, Q., Xia, A., et al. (2018). Enhancing microalgae biofilm formation and growth by fabricating microgrooves onto the substrate surface. *Bioresour. Technol.* 261, 36–43. doi: 10.1016/j.biortech.2018.03.139

Huggett, M. J., Nedved, B. T., and Hadfield, M. G. (2009). Effects of initial surface wettability on biofilm formation and subsequent settlement of Hydroides elegans. *Biofouling* 25, 387–399. doi: 10.1080/08927010902823238

Kanehisa, M., and Sato, Y. (2019). 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

Lai, S., Loke, L. H., Hilton, M. J., Bouma, T. J., and Todd, P. A. (2015). The effects of urbanisation on coastal habitats and the potential for ecological engineering: a Singapore case study. *Ocean Coast. Manag.* 103, 78–85. doi: 10.1016/j.ocecoaman.2014.11.006

Lam, N. W., Huang, R., and Chan, B. K. (2009). Variations in intertidal assemblages and zonation patterns between vertical artificial seawalls and natural rocky shores: a case study from Victoria Harbour, Hong Kong. *Zool. Stud.* 48, 184–195.

Lane, D. (1991). "16S/23S rRNA sequencing," in *Nucleic Acid Techniques in Bacterial Systematics*, eds E. Stackebrandt and M. Goodfellow (Chichester, UK: John Wiley and Sons), 115–175.

Lawrence, P. J., Evans, A. J., Jackson-Bué, T., Brooks, P. R., Crowe, T. P., Dozier, A. E., et al. (2021). Artificial shorelines lack natural structural complexity across scales. *Proc. R. Soc. B* 288:20210329. doi: 10.1098/rspb.2021.0329

Lema, K. A., Constancias, F., Rice, S. A., and Hadfield, M. G. (2019). High bacterial diversity in nearshore and oceanic biofilms and their influence on larval settlement by Hydroides elegans (*Polychaeta*). *Environ. Microbiol.* 21, 3472–3488. doi: 10.1111/1462-2920.14697

Loke, L. H., Heery, E. C., Lai, S., Bouma, T. J., and Todd, P. A. (2019). Area-independent effects of water-retaining features on intertidal biodiversity on eco-engineered seawalls in the tropics. *Front. Mar. Sci.* 6:16. doi: 10.3389/fmars.2019.00016

Loke, L. H., Ladle, R. J., Bouma, T. J., and Todd, P. A. (2015). Creating complex habitats for restoration and reconciliation. *Ecol. Eng.* 77, 307–313. doi: 10.1016/j.ecoleng.2015.01.037

Loman, N. (2020). *MaxiKraken2_1903_140GB Mock Community Database*. Available online at: https://lomanlab.github.io/mockcommunity/mc_databases.html (accessed April 2021).

Lu, N., Zhang, W., Weng, Y., Chen, X., Cheng, Y., and Zhou, P. (2016). Fabrication of PDMS surfaces with micro patterns and the effect of pattern sizes on bacteria adhesion. *Food Control.* 68, 344–351. doi: 10.1016/j.foodcont.2016.04.014

Medlin, L., Elwood, H. J., Stickel, S., and Sogin, M. L. (1988). The characterization of enzymatically amplified eukaryotic 16S-like rRNA-coding regions. *Gene* 71, 491–499. doi: 10.1016/0378-1119(88)90066-2

Ming, C. L., Lionel, N. C. S., Jeremy, C. S. M., and Angie, S. L. (2010). Natural coral colonization of a marina seawall in Singapore. *J. Coast. Dev. ISSN* 14, 11–17.

Morris, R. L., Heery, E. C., Loke, L. H., Lau, E., Strain, E., Airoldi, L., et al. (2019). Design options, implementation issues and evaluating success of ecologically engineered shorelines. *Oceanogr. Mar. Biol.* 57, 169–228. doi: 10.1201/9780429026379-4

Murdock, J. N., and Dodds, W. K. (2007). Linking benthic algal biomass to stream substratum topography 1. *J. Phycol.* 43, 449–460. doi: 10.1111/j.1529-8817.2007.00357.x

Nicholls, R. J., and Cazenave, A. (2010). Sea-level rise and its impact on coastal zones. *Science* 328, 1517–1520. doi: 10.1126/science.1185782

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'hara, R., et al. (2011). *Vegan: Community Ecology Package. R Package Version*. Available online at: http://CRAN.Rproject.org/package=vegan (accessed April 2020).

Petersen, J. M., and Dubilier, N. (2009). Methanotrophic symbioses in marine invertebrates. *Environ. Microbiol. Rep.* 1, 319–335. doi: 10.1111/j.1758-2229.2009.00081.x

R Core Team (2014). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Rädecker, N., Pogoreutz, C., Voolstra, C. R., Wiedenmann, J., and Wild, C. (2015). Nitrogen cycling in corals: the key to understanding holobiont functioning? *Trends Microbiol.* 23, 490–497. doi: 10.1016/j.tim.2015.03.008

Revelle, R., and Emery, K. O. (1957). *Chemical Erosion of Beach Rock and Exposed Reef Rock: Bikini and Nearby Atolls*. Marshall Islands: US Government Printing Office.

Schneck, F., Schwarzbold, A., and Melo, A. S. (2011). Substrate roughness affects stream benthic algal diversity, assemblage composition, and nestedness. *J. North Am. Benthol. Soc.* 30, 1049–1056. doi: 10.1899/11-044.1

Shannon, C. E. (1948). A mathematical theory of communication. *Bell Syst. Tech. J.* 27, 379–423. doi: 10.1002/j.1538-7305.1948.tb01338.x

Simkanin, C., Davidson, I. C., Dower, J. F., Jamieson, G., and Therriault, T. W. (2012). Anthropogenic structures and the infiltration of natural benthos by invasive ascidians. *Mar. Ecol.* 33, 499–511. doi: 10.1111/j.1439-0485.2012.00516.x

Sneed, J. M., Ritson-Williams, R., and Paul, V. J. (2015). Crustose coralline algal species host distinct bacterial assemblages on their surfaces. *ISME J.* 9, 2527–2536. doi: 10.1038/ismej.2015.67

Strain, E. M., Steinberg, P. D., Vozzo, M., Johnston, E. L., Abbiati, M., Aguilera, M. A., et al. (2021). A global analysis of complexity–biodiversity relationships on marine artificial structures. *Glob. Ecol. Biogeogr.* 30, 140–153. doi: 10.1111/geb.13202

Suess, E. (1970). Interaction of organic compounds with calcium carbonate—I. Association phenomena and geochemical implications. *Geochim. Cosmochim. Acta* 34, 157–168. doi: 10.1016/0016-7037(70)90003-7

Summers, S., Thomson, B. C., Whiteley, A. S., and Cockell, C. S. (2016). Mesophilic mineral-weathering bacteria inhabit the critical-zone of a perennially cold basaltic environment. *Geomicrobiol. J.* 33, 52–62. doi: 10.1080/01490451.2015.1039672

Sweat, L. H., and Johnson, K. B. (2013). The effects of fine-scale substratum roughness on diatom community structure in estuarine biofilms. *Biofouling* 29, 879–890. doi: 10.1080/08927014.2013.811492

Tan, E. L. Y., Mayer-Pinto, M., Johnston, E. L., and Dafforn, K. A. (2015). Differences in Intertidal Microbial Assemblages on Urban Structures and Natural Rocky Reef. *Front. Microbiol.* 6:1276. doi: 10.3389/fmicb.2015.01276

Temmerman, S., Meire, P., Bouma, T. J., Herman, P. M., Ysebaert, T., and De Vriend, H. J. (2013). Ecosystem-based coastal defence in the face of global change. *Nature* 504, 79–83. doi: 10.1038/nature12859

Todd, P. A., Heery, E. C., Loke, L. H., Thurstan, R. H., Kotze, D. J., and Swan, C. (2019). Towards an urban marine ecology: characterizing the drivers, patterns and processes of marine ecosystems in coastal cities. *Oikos* 128, 1215–1242. doi: 10.1111/oik.05946

Tukey, J. W. (1949). Comparing individual means in the analysis of variance. *Biometrics* 5, 99–114. doi: 10.2307/3001913

Warrick, R., and Oerlemans, J. (1990). "Sea Level Rise," in *Climate Change: The IPCC Scientific Assessment*, eds J. T. Houghton, G. J. Jenkins, and J. J. Ephraums (Sweden: Intergovernmental Panel on Climate Change), 257–281.

Wickham, H. (2016). *Ggplot2: Elegant Graphics for Data Analysis*. New York: Springer. doi: 10.1007/978-3-319-24277-4

Wyngaard, G., Mclaren, I., White, M., and Sévigny, J.-M. (1995). Unusually high numbers of ribosomal RNA genes in copepods (*Arthropoda*: *Crustacea*) and their relationship to genome size. *Genome* 38, 97–104. doi: 10.1139/g95-012

Yamashita, H. (2020). Living together with seawalls: risks and reflexive modernization in Japan. *Environ. Sociol.* 6, 166–181. doi: 10.1080/23251042.2019.1709680

Zeller, M. A., Van Dam, B. R., Lopes, C., and Kominoski, J. S. (2020). Carbonate-associated organic matter is a putative FDOM source in a subtropical seagrass meadow. *Front. Mar. Sci.* 7:580284. doi: 10.3389/fmars.2020.580284

frontiers | Frontiers in Microbiology

TYPE Mini Review
PUBLISHED 26 July 2022
DOI 10.3389/fmicb.2022.941532

OPEN ACCESS

EDITED BY
Tony Gutierrez,
Heriot-Watt University,
United Kingdom

REVIEWED BY
Karen Fong,
Agriculture and Agri-Food Canada (AAFC), Canada
Masoud Yavarmanesh,
Ferdowsi University of Mashhad, Iran
Felipe Molina,
University of Extremadura, Spain

*CORRESPONDENCE
Francis Hassard
francis.hassard@cranfield.ac.uk

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

RECEIVED 11 May 2022
ACCEPTED 30 June 2022
PUBLISHED 26 July 2022

CITATION
Singh S, Pitchers R and Hassard F (2022) Coliphages as viral indicators of sanitary significance for drinking water. *Front. Microbiol.* 13:941532. doi: 10.3389/fmicb.2022.941532

Coliphages as viral indicators of sanitary significance for drinking water

Suniti Singh[1], Robert Pitchers[2] and Francis Hassard[1,3]*

[1]Cranfield Water Science Institute, Cranfield University, Bedford, United Kingdom, [2]Water Research Centre, Swindon, United Kingdom, [3]Institute for Nanotechnology and Water Sustainability, University of South Africa, Johannesburg, South Africa

Coliphages are virus that infect coliform bacteria and are used in aquatic systems for risk assessment for human enteric viruses. This mini-review appraises the types and sources of coliphage and their fate and behavior in source waters and engineered drinking water treatment systems. Somatic (cell wall infection) and F^+ (male specific) coliphages are abundant in drinking water sources and are used as indicators of fecal contamination. Coliphage abundances do not consistently correlate to human enteric virus abundance, but they suitably reflect the risks of exposure to human enteric viruses. Coliphages have highly variable surface characteristics with respect to morphology, size, charge, isoelectric point, and hydrophobicity which together interact to govern partitioning and removal characteristics during water treatment. The groups somatic and F^+ coliphages are valuable for investigating the virus elimination during water treatment steps and as indicators for viral water quality assessment. Strain level analyses (e.g., Qβ or GA-like) provide more information about specific sources of viral pollution but are impractical for routine monitoring. Consistent links between rapid online monitoring tools (e.g., turbidity, particle counters, and flow cytometry) and phages in drinking water have yet to be established but are recommended as a future area of research activity. This could enable the real-time monitoring of virus and improve the process understanding during transient operational events. Exciting future prospects for the use of coliphages in aquatic microbiology are also discussed based on current scientific evidence and practical needs.

KEYWORDS

coliphage, somatic coliphage, drinking water quality, online monitoring, drinking water treatment, F^+ coliphage

Coliphage characteristics

Phages are the most widely distributed and abundant biological forms on Earth estimated at $\sim 10^{31}$ particles in the biosphere (Hendrix et al., 1999; Comeau et al., 2008; Mushegian, 2020). Phages are part of a complex microbial ecosystem and exist either as free-floating infectious particles in environmental matrices or within a bacteria or associated directly/indirectly to particles (Clokie et al., 2011; Zimmerman et al., 2020). Phages are obligate parasites of prokaryotes and replicated by members of two domains of cellular life—bacteria and archaea, with some evidence of interactions with eukaryotic organisms

through effects to their microbiome or to their bacterial pathogens (Sime-Ngando, 2014; Putra and Lyrawati, 2020). The coliphages are a specific group of bacteriophages that infect coliforms (including *Escherichia coli*) and other closely related bacteria that are present in human and animal gut microbiomes.

The coliphages are a diverse group of phage from several families and therefore consensus genomic sequences do not exist, which prevents design of universal primer sets limiting the use of qPCR for quantification of coliphage. Instead, they are classified into diverse taxonomic groups based on their *replication mechanism*, *mode of infecting hosts*, *morphological characteristics*, and *genomic content* (Table 1). Coliphages can be divided in two groups, virulent or temperate, which depends on their *replication mechanism* in host cells (Grabow, 2001). Virulent phages follow an obligate lytic cycle, whereas temperate phages undergo lysogenic cycle from where they can switch to a "lytic" or "chronic" cycle (Figure 1A). Coliphages infect host cells by first adsorbing to the host cell and injecting genetic material (DNA or RNA). Next, the coliphage nucleic material circularizes and enters lytic (virulent phage), lysogenic (temperate phage), or chronic cycle (temperate phage) (Figure 1A). In a lytic cycle, coliphages replicate in their host cells to synthesize new coliphage DNA or RNA and proteins which are assembled into phage virions. Subsequently, large numbers of coliphage virions can be released to the environment during each lytic cycle *via* lysis-protein-mediated rupture of the host cell wall. There are similarities between the lytic cycle and the lysogenic cycle. In the lysogenic cycle, coliphage attaches to the host cell and injects its genetic material into the bacteria host. During these processes, the coliphage DNA stably integrates into the chromosome of host bacterium forming a *prophage*. This entity is not infectious or lethal and replicates alongside and within the host DNA, increasing its titer with each prokaryotic replication cycle. Intermittently, an environmental cue will trigger the *prophage* to excise from the bacterial chromosome to become lytic. This marks a shift in the physiological state of host cells from lysogenic to lytic: a process which is termed *induction*. After *induction*, host cells are lysed, producing new phage virions (Figure 1A). The lysis of host cells expels intracellular components and cell debris into the surrounding environment and contributes to organic loading in waters. Thus, lytic phages (alongside higher organism grazing) have a key role in driving biogeochemical nutrient cycling and availability by structuring the population dynamics of their hosts in most aquatic systems (Clokie et al., 2011; Sime-Ngando, 2014). Another aspect of phage biology is the so-called chronic mode. This mode is exhibited by some archaeal phages but has not been widely reported in coliform bacteria. In chronic mode, the cell grows at a slower rate as new virions are continuously produced and excreted at low levels, despite the infected host cell remaining intact and viable (Howard-Varona et al., 2017).

Activation of the lytic-lysogenic switch occurs either spontaneously or requires stimulation by natural or anthropogenic inductors. Induction can be triggered by exposure of bacterial host (containing prophage) to stressed conditions, such as UV, low nutrient conditions, a change to the bacterial trophic status (i.e., reduced competitors due to antibiotic), or replication inhibitors (mitomycin C); or through inactivation of a phage repressor. This process inevitably contributes to the coliphage loads in the surrounding environmental matrices. This aspect of (coli)phage biology offers a potential challenge to its utility as a viral indicator—as its abundance can be linked to its environment, e.g., solar irradiation in catchment or UV dose in treatment processes, rather than the fecal load exclusively. The impact of lysogeny on coliphage loading in an aquatic system is reviewed below.

Lysogeny and its contribution to phage abundance in aquatic matrices have been studied extensively. For example, the induction of lysogenic bacteria (lysogens) from tropical marine waters by mitomycin method contributed to 4–27% of the total marine phage concentrations (Ashy and Agustí, 2020). Exposure to surface ambient ultraviolet radiation levels induces lysogeny in bacteria within lake waters (Maranger et al., 2002). Solar radiation or hydrogen peroxide stress induces lysogeny and accounts for up to 29–63 and 47–53%, respectively, of phages produced by induction (Weinbauer and Suttle, 1999); however, the lysogenic contribution is low in marine (0.8–11%) and riverine bacteria (1%) (Muniesa and Jofre, 2007), suggesting total phage concentrations are governed by other factors than *in situ* lysogenic contribution, a finding which has been replicated in raw sewage and river waters (Casjens, 2003; Jofre, 2009). Therefore, the majority of coliphage loading, i.e., *plaque forming units* (PFU), are contributed from lytic coliphages generated within the gut of animals and to a lesser extent in the environment.

The PFU is an indicative measure of infective phage numbers. A single PFU can manifest from simultaneous infection initiated by a single- or multiple-phage particles (containing numerous copies of phage). Therefore, theoretically, one PFU is not always equivalent to one infective phage particle. However, from a practical perspective, the plaque numbers correlate well to absolute numbers of phage quantified via other methods (e.g., PCR-based approaches; Rose et al., 1997). However, what is less well-established is the influence of water physicochemistry on coliphage enumeration. The presence of natural organic and inorganic compounds can result in phage complexes to form. Key water treatment processes such as coagulation/flocculation aggregate particles and remove substantial amounts of virus from water. However, the aggregative effect could artificially reduce the quantification of phages (as multiple infective phages could present single plaque). In addition, Matsushita et al. (2011) identified that PFU infectivity assays were particularly sensitive to coagulants which impacted both the infectivity and the aggregative properties of coliphages compared to qPCR methods. Determining whether water treatments (e.g., coagulation) can impact viability

TABLE 1 Characteristics of major groups of bacteriophages and their representative phage species.

Order	Family	Phage genus	Representative species	Type	Genogroup based on serotypes	Nucleic acid	Virion shape	Virion diameter (nm)	Tail	Measured pI
Levivirales	*Fiersviridae (earlier Leviviridae)*	*Levivirus*	*Escherichia virus MS2*	F-specific RNA phage	Genogroup I	ssRNA, linear genome	icosahedral	26	No tail	2–4, 9.04[A]
			Escherichia virus BZ13		Genogroup II	ssRNA, linear genome	icosahedral	26	No tail	2.1–2.3
		Allolevirus	*Escherichia virus Qbeta (Qβ)*		genogroup III	ssRNA, linear genome	icosahedral	26	No tail	2–4, 5.3, 2–7, 1.9
			Enterobacteria phage SP		Genogroup IV	ssRNA, linear genome	icosahedral	26	No tail	2.1–2.6, 6.37[A]
Tubulavirales	*Inoviridae*	*Inovirus*	*Enterobacteria virus M13*	F-specific DNA phage	Not applicable	ss DNA, circular genome	filamentous	6, (1,000–2,000 nm long)	No tail, flexible filaments	4.05, 7.33[#]
Kalamavirales	*Tectiviridae*	*Alphatectivirus*	*Salmonella virus PRD1*	Somatic	Not applicable	ds DNA, linear genome	icosahedral	62	Pseudotail	3–4, 6.82[A]
Caudovirales	*Myoviridae*	*T4virus*	*Enterobacteria virus T4*			ds DNA, linear genome	elongated icosahedral	90, (200 nm long)	Long contractile tail	2, 4–5, 6.53[A]
	Siphoviridae	*Lambdavirus*	*Enterobacteria virus lambda*			ds DNA, linear genome	icosahedral	62	Long non-contractile tail	3.8, 7.04[#]
	Podoviridae	*T7 virus*	*Enterobacteria virus T7*			ds DNA, linear genome	icosahedral	55	Short non-contractile tail	6.98[A]
Petitvirales	*Microviridae*	*phix174*	*Escherichia virus phiX174*			ss DNA, circular genome	icosahedral	25	No tail	6–7.4, 7.66[A]

[#] Average calculated pI for phage proteome.

10.3389/fmicb.2022.941532

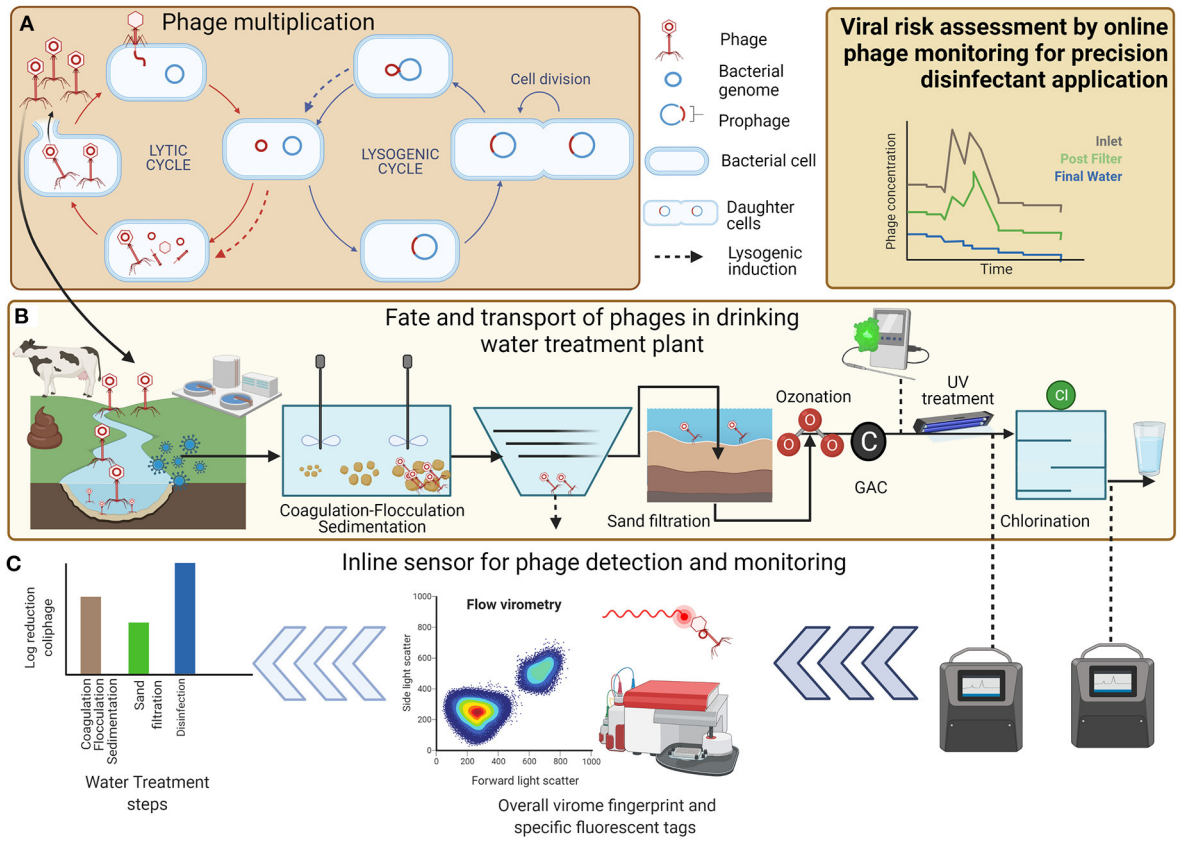


FIGURE 1
Schematic representing coliphages as viral indicators of sanitary significance for drinking water. Each panel represents a different example of this field. **(A)** Phage multiplication in environmental water matrices, **(B)** fate and transport of phages in drinking water treatment plant, and **(C)** inline sensors for phage detection and monitoring for disinfection optimization and risk assessment.

(infectivity) via plaque assay or just result in false low numbers due to aggregation is critical, considering the role of infectious virus dose on human enteric/respiratory disease response. This is particularly important for understudied areas of water treatment with respect to virus and their indicators.

Coliphages can also be classified based on their "*mode of infection a host cell*" into somatic and F^+ coliphages (also called male-specific or F-specific coliphages). Generally, individual coliphage strain infects specific *E. coli* strains by attaching to lipopolysaccharide or protein receptors in the cell wall and may lyse the host cell in within 20 min of infection (De Paepe and Taddei, 2006; Stone et al., 2019). They produce plaques of diverse size and morphology and can be useful for distinguishing different types of phages. Coliphage-host infection dynamics is strain specific (Molina et al., 2020), and as such model somatic phages are often used. One commonly applied coliphage during laboratory and pilot-scale experiments is phiX174 (ΦX174) which is host specific to *E. coli* ATCC 13706 (strain C) and PC 0886. In contrast, some other somatic coliphages may multiply in other hosts such as the *Enterobacteriaceae*. Of these, the two most commonly found species are *Shigellla* sp. and *Klebsiella pneumoniae* (Leclerc et al., 2000; Goodridge et al., 2003; Muniesa et al., 2003). Four phage families, *Myoviridae*, *Siphoviridae*, *Podoviridae*, and *Microviridae*, include strains which are considered to be somatic coliphages. The *Microviridae* group infects a diverse range of hosts such as *Enterobacteria*, *Bdellovibrio*, *Chlamydia*, and *Siroplasma*. For the *Myoviridae* group, principal hosts are *Enterobacteria*, *Bacillus*, and *Halobacterium*. For the *Siphoviridae* group, *Enterobacteria*, *Mycobacterium*, and *Lactococcus* are the major host groups. For the *Podoviridae* group, *Enterobacteria* and *Bacillus* are the main host groups (Lee, 2009). The tailed bacteriophages (order *Caudovirales*, including families *Myoviridae*, *Siphoviridae*, *Podoviridae*, and *Microviridae*) constitute 96% of all known phages (Ackermann, 1998). Of these, there are at least 150 "species" of *Caudovirales* phages (International Committee on Taxonomy of Viruses [ICTV]) which infect the genus *Escherichia*. Sequencing technologies are driving new phage discoveries which in turn is shedding light on the ecology of these entities (Korf et al.,

2019) and the development of phage analysis tools, e.g., VIrus Classification and Tree building Online Resource (VICTOR) is aiding efforts for rapid and systematic phage classification (Meier-Kolthoff and Göker, 2017).

F^+ coliphages are either single-stranded (ss) DNA or RNA viruses that are infectious to bacteria possessing fertility (F)-plasmid and infect their host through the F-pili (Jones et al., 2017). Of the F^+ coliphages, the single-stranded RNA phages (FRNAPH) have been subclassified further into four serologically and phylogenetically distinct genogroups (designated GI, GII, GIII, and GIV). These FRNAPH genogroups differ in their tail morphologies (Table 1). Subsequently, several types of quantitative (reverse) PCR assays have been developed for rapid, sensitive, and specific detection of each F+RNA coliphage genogroup by targeting different conserved genes including the RNA replicases, maturase, capsid (coat protein), β-chain, or assembly genes (Jofre et al., 2016). The specific coliphage genogroups can assist in discrimination of the sources of fecal contamination via apportioning the load originating from humans and other animals—offering application for pollution tracking in catchments. FRNAPH strains MS2 (GI), GA (GII), Qβ (GIII), and SP (GIV) are widely used as representatives of their genogroups. There is some concern about the specificity of these genogroups to *E. coli* as the F-plasmid is transferable from/to *E. coli* and other related *Enterobacteria*. The transfer of the PCR target regions used in detecting FRNAPH genogroups (RNA replicases, maturase, capsid, or assembly genes) through the F-plasmid may confound quantification of FRNAPH genogroups, especially as a single bacterium is likely to contain numerous copies of the plasmid. For example, F^+ coliphages may multiply in coliforms, including not only *E. coli* but also *Salmonella*, *Shigella*, *Bacteroides*, *Caulobacter*, *Pseudomonas*, and *Acinetobacter* providing an appropriate pili expressed on the bacterial surface (Leclerc et al., 2000; Cann, 2001; Virolle et al., 2020). The fragments of genetic material may be transferred along with F-plasmid, resulting in horizontal gene transfer, thereby transmitting genes (pathogenicity, metabolic properties, or antimicrobial resistance linked) among the bacteria across the water distribution network (Maganha de Almeida Kumlien et al., 2021). Few studies have reported the use of strain level monitoring in microbial source tracking (Hartard et al., 2016; Fauvel et al., 2017; Lee et al., 2019; Hata et al., 2021) as differentiating the F^+ strains remains challenging and labor intensive.

Fate and transport of coliphages in water used for human consumption

Coliphages are found in aquatic systems which are impacted by fecal contamination, including most surface and ground water drinking water sources, demonstrating region- and season-specific variation in abundance (Nappier et al., 2019). Coliphages are shed in high numbers in feces, with numbers dependent on the type of coliphage, animal host, and size and frequency of defecation events. There is some evidence of coliphages replicating in bacteria naturally present within surface water environments (Figure 1; Hassard et al., 2016), although low specific densities of host or phage are unlikely to permit coliphage replication in most aquatic environments (Muniesa and Jofre, 2004). Therefore, replication of coliphage in natural waters does not contribute a detectable increase in the numbers of somatic coliphages (Jofre et al., 2016). For some coliphages, e.g., F^+ coliphages, the F+ pili is not readily produced by the host at $<25^{\circ}C$; therefore, phage attachment and replication are unlikely in temperate water sources (Franke et al., 2009). Thus, the fate and behavior of coliphages are useful to assess the ability of water treatment to eliminate human enteric viruses and assess degree of contamination of surface and groundwaters (Jofre et al., 2016). Coliphages are considered in ambient water quality regulations (Anonymous, ISO 10705-2, 2001; Anonymous, ISO 10705-1, 2001; WHO, 2017) and the European Commission included somatic coliphages for the revised EU drinking water directive (2020), highlighting the importance of the method for understanding the risk profile of different treatment works. In addition, coliphage could be used to assess viral water quality impacts due to natural (i.e., runoff) or anthropogenic (i.e., sewer overflows) forces. An approach to appraise the risk is based on source water viral risk characterization, e.g., the presence of somatic coliphage exceeding 50 PFU/100 ml raw water would trigger the use of this microbial water quality parameter. This will be analyzed interstage in the water treatment works to demonstrate log removal to ensure that the risk of a breakthrough of pathogenic viruses is controlled. A meta-analysis revealed that the total coliphage median density in untreated wastewater was about 80,000 PFU/100 ml compared to around 30 PFU/100 ml for lowland river source waters used for drinking water sources (Nappier et al., 2019). The justification for use of somatic coliphage is that it is the most readily detected and at greater density justifying its selection as a useful indicator of viral risk (Jebri et al., 2017). Coliphages have been readily employed for log removal credit and assessing viral risk from raw water quality changes and transient operational or process events within water treatment works (Figures 1B,C). It is noteworthy to mention that the total coliphage concentrations do not always correlate to the total concentrations of infectious viruses whose abundance is ephemeral and often linked to outbreaks within catchments. However, the presence of specific indicators (e.g., coliphages) has been linked to viral water quality and likelihood of the presence of human enteric viruses.

Tools for monitoring phage and viral transience and loading

Emergent technologies may facilitate detection of coliphages using molecular methods (PCR-based), direct enumeration using optical properties or fluorescent dyes, and indirect methods using enzymatic or immunological reactions. Among these, PCR-based methods have been applied extensively to different water matrices for coliphage detection. Primer sets have been developed for detecting type strains representing the families in somatic or F^+ coliphages (Table 1). Real-time PCR has been used for the four somatic coliphage families, represented by type strain T4 for *Myoviridae*, type strain phiX174 for *Microviridae*, and type strain lambda (λ) phage for *Siphoviridae* and type strain T7 for *Podoviridae* (Lee, 2009). Despite the initial categorization of the genogroups of FRNAPH based on serological and physicochemical properties, development of genogroup-specific primers enables a more specific detection of the FRNAPH genogroups (Ogorzaly and Gantzer, 2006). Molecular methods such as PCR and sequencing are fast and reliable but complex and costly especially for routine water quality assessment. One alternative is the adenosine triphosphate (ATP) assessment, which is a rapid, sensitive, and is amenable to automated inline/real-time sensing of the microbiological activity in water samples (Ochromowicz and Hoekstra, 2005; Vang et al., 2014). Continuous sampling combined with ATP measurements displays potential for microbial drinking water quality assessments. The ability of the ATP assay to detect microbial ingress or poor biostability is influenced by both the ATP load from the contaminant event and the ATP concentration in the specific drinking waters (Vang et al., 2014). While ATP is an indicator of total microbial activity, its application or direct correlations with coliphage concentrations is not reported. This raises the prospect of whether online technologies (such as online ATP or particle count) could better inform/trigger coliphage sampling campaigns to assess pathways for health risks.

Another approach is direct quantification and this method relies on virus and their genomes being labeled using either specific or non-specific fluorescent dyes which permits quantification of virus-like particles (VLP) and viruses in environmental water samples using quantification tools, e.g., flow cytometry (Figure 1C) (Gaudin and Barteneva, 2015; Huang et al., 2016; Safford and Bischel, 2019; Olivenza et al., 2020). The tools have been adapted and applied to characterize phages including coliphages (Wilhartitz et al., 2013; Roudnew et al., 2014). For example, imaging flow cytometry was used to quantify somatic coliphages phiX174 and *E. coli* phage Qβ were detected with high specificity (within a concentration range of 10^4-10^7 PFU/ml) providing results in 1 h with results comparable to the double-layer agar technique (Yang et al., 2019). Rajnovic and Mas (2020) evaluated the sensitivity of a resazurin-reporter sensor in *E. coli* DSMZ 613/T4 phage model using a fluorimeter. Viable bacterial host cells reduce resazurin to a fluorescent compound resorufin which can be enumerated and phage-induced changes of activity of the bacterial host are measured using resazurin as a reporter. Green fluorescent protein (gfp)-based bacterial tagging used epigenetic biosensing to provide sensitive detection of infectious liposaccharide-binding phages (1.6 phages per 100 ml) (Olivenza et al., 2020). Other fluorescent tracer compounds have been developed to demonstrate log removal values (LRV) and have been applied in different scenarios. An example is TRASAR®, an organic fluorescent dye that carries a net negative charge, thus mimicking some of the charge repulsion experienced by negatively charged virus particles and can be detected at level of 10 μg/L with proprietary online sensor. LRVs for MS2 coliphage, TRASAR®, and conductivity using intact membranes averaged 6.2, 4.3, and 1.7, respectively, suggesting the tracer compound acts as a conservative surrogate of MS2 LRVs (Steinle-Darling et al., 2016). Nanoparticles or nanoplastics (tagged and/or coated) have been readily applied in assess fate and transport studies of virus in bench of lab scale sand columns. Good association between nanoparticles and two pathogenic virus suggested that physical tracers could be employed in fundamental and applied studies of virus or phage (Pang et al., 2014). Applicability of fluorescent dyes or nanoparticles to represent viruses or phages in environmental matrices requires further investigation especially bench scale or pilot studies to better understand the behavior and removal in treatment processes (Pulido-Reyes et al., 2022).

Online monitoring tools: Limitations, future outlook, and recommendations

There are several limitations associated with detection of phages using fluorescence-based technologies, particularly in real-environmental water matrices. The detection limit of flow cytometry is ∼150 kbp (Dlusskaya et al., 2021), whereas majority of bacteriophages (87%) with known genomes in NCBI database have a smaller genome size than this. Therefore, most bacteriophages and small enteric viruses can not be directly quantified using flow cytometry with the current limits of sensitivity—assuming the phage will be planktonic and not particle associated. Pathogenic viruses are also present and can pose health risks at concentrations well below the detection limits of 100 pathogens/ml, and as such the direct FCM-based assessment of waterborne pathogens is challenging (Ramírez-Castillo et al., 2015). Coliphages, however, are suitable process indicators and

surrogates of human viral pathogens present at higher concentrations but are labor intensive to measure. We posit that there could be a plausible link between data generated from online tools, e.g., turbidity, online flow cytometry, or fluorescence (e.g., tryptophan-like), a strategy to understand the removal performance for virus and their indicators (Figure 1; Rockey et al., 2019). Biosensors offer another approach for monitoring of contaminants/indicators in water matrices (Gautam et al., 2012). Biosensor development using FACS, flow cytometry, imaging cytometry, OD-based techniques, resazurin reduction, gfp tagging, or ATP measurements is a promising technology. However, a thorough investigation of the matrix effects from environmental matrices on phage detection techniques is warranted. To date, application of "online" FCM has not been demonstrated for automated, real-time detection of waterborne viruses and requires further development (Safford and Bischel, 2019).

There are other online monitors such as turbidity and particle-size distribution which are being used in a diagnostic capacity by drinking water treatment plants (DWTPs) operators for monitoring raw and treated water quality. Interesting applications showed that elevated water flow levels linked to rainfall or snowmelt events increased the viral and phage concentrations in a catchment (Fauvel et al., 2016; Sylvestre et al., 2021). Infectious F-specific RNA bacteriophage concentrations increased during rainfall events (Fauvel et al., 2016), whereas viral concentrations increased by ∼0.5-log during two snowmelt/rainfall episodes and ∼1.0-log following a planned wastewater discharge upstream of the drinking water intake (Sylvestre et al., 2021). Another approach is to utilize enzyme-based sensors specific to coliforms. These provide rapid sensitive online monitoring, providing a proxy for culture-based assays. Online enzyme sensors such as β-D-glucuronidase provided estimates of increased viral loads, for example, transient peaks in raw water fecal contamination were identified at two urban DWTPs (Sylvestre et al., 2021). However, the enzyme sensors monitor bacterial indicators, and although their application for viral or phage quantification may not be accurate, the automated enzymatic sensors could be practically used to trigger intensive sampling campaigns. Catchment-level demographic and hydroclimatic parameters (agricultural or residential catchment, sediment property, precipitation frequency, and load) drive the physico-chemical properties such as turbidity, particulate, and water flow, and thereby affect the coliphage detection. Moreover, inhibitors present in water matrices are varied and impact the LOD achieved during enumeration. Hence, standardization of the enzyme sensors in different catchments is recommended prior to the implementation of monitoring program.

Focus to develop the fluorescent detection methods and their assessment criteria for coliphage detection offers opportunities to setup a reliable online monitoring system. We recommend matrix-based rigorous background controls for quantitative evidence of the sensitivity and accuracy as applied to aquatic environmental matrices and standardization of biosensing methods using representative coliphages [somatic, phiX174, and F-specific coliphage (MS2)]. There is an additional need to validate that potential virus and phage populations identified are indeed viruses rather than bacterial debris or other small particles. To achieve this, development of reliable positive controls, and internal standards is needed through evaluations using mixture of fluorescent-labeled bacteriophages or VLP of various sizes, genomes, and capsid characteristics. Method standardization is a herculean task that needs to be undertaken to take fluorescent methods to the next stage of real-time coliphage monitoring. If realized, this will help operators optimize DWTP for viral removal and assess potential pathways for risks for public health.

Conclusions

- Coliphages have a host range broader than *E. coli* and have highly variable surface characteristics which govern partitioning in aquatics systems and influence removal during water treatment.
- General coliphage groups, either somatic coliphages or F^+ coliphages, are valuable for investigating the virus elimination of water treatment.
- Current poor process level understanding of phage particle interactions, phage reproduction cycles within source water, and water treatment combined with methodological challenges is limiting the utility of coliphage as a sanitary tool in aquatic environments.
- The combined use of chemical and biological analyses alongside inline sensors could enable phage and thereby virus monitoring and enable the revision and development of more accurate environmental risk assessment.
- Finally, standardization of rapid online monitoring tools (e.g., turbidity and flow cytometry) for coliphages is recommended as a future area of activity.

Author contributions

SS, RP, and FH wrote the manuscript and undertook the literature survey. All authors had substantial contributions to the conception and design of the work and approved the final version to be published.

Funding

This work was funded by the UK Water Industry Research Limited under funding for the Independent Microbial Advisory Service at WRc and Cranfield University.

Acknowledgments

The contributions of specific colleagues, institutions, or agencies that aided the efforts of the authors. The figure was created with BioRender.com.

Conflict of interest

The Water Research Centre received funding from the United Kingdom Water Industry Research. The funder was not involved in the study design, collection, analysis, interpretation of data, the writing of this article, and the decision to submit it for publication.

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

Ackermann, H.-W. (1998). Tailed bacteriophages: the order Caudovirales. *Adv. Virus Res.* 51, 135–197. doi: 10.1016/S0065-3527(08)60785-X

Anonymous, ISO 10705-1 (2001). *ISO 10705-1: Water Quality. Detection and Enumeration of F-Specific RNA Bacteriophages-Part 1*. Geneva: International Organisation for Standardisation.

Anonymous, ISO 10705-2 (2001). *ISO 10705-2: Water Quality. Detection and Enumeration of Bacteriophages-Part 2: Enumeration of Somatic Coliphages*. Geneva: International Organisation for Standardisation.

Ashy, R. A., and Agustí, S. (2020). Low host abundance and high temperature determine switching from lytic to lysogenic cycles in planktonic microbial communities in a tropical sea (Red Sea). *Viruses* 12:761. doi: 10.3390/v12070761

Cann, A. J. (2001). *Principles of Molecular Virology (Standard Edition)*. Geneva: International Standardisation Organisation; Academic Press.

Casjens, S. (2003). Prophages and bacterial genomics: what have we learned so far? *Mol. Microbiol.* 49, 277–300. doi: 10.1046/j.1365-2958.2003.03580.x

Clokie, M. R. J., Millard, A. D., Letarov, A. V., and Heaphy, S. (2011). Phages in nature. *Bacteriophage* 1, 31–45. doi: 10.4161/bact.1.1.14942

Comeau, A. M., Hatfull, G. F., Krisch, H. M., Lindell, D., Mann, N. H., and Prangishvili, D. (2008). Exploring the prokaryotic virosphere. *Res. Microbiol.* 159, 306–313. doi: 10.1016/j.resmic.2008.05.001

De Paepe, M., and Taddei, F. (2006). Viruses' life history: towards a mechanistic basis of a trade-off between survival and reproduction among phages. *PLoS Biol.* 4, 1248–1256. doi: 10.1371/journal.pbio.0040193

Dlusskaya, E., Dey, R., Pollard, P. C., and Ashbolt, N. J. (2021). Outer limits of flow cytometry to quantify viruses in water. *ACS EST Water* 1, 1127–1135. doi: 10.1021/acsestwater.0c00113

Fauvel, B., Cauchie, H. M., Gantzer, C., and Ogorzaly, L. (2016). Contribution of hydrological data to the understanding of the spatio-temporal dynamics of F-specific RNA bacteriophages in river water during rainfall-runoff events. *Water Res.* 94, 328–340. doi: 10.1016/j.watres.2016.02.057

Fauvel, B., Ogorzaly, L., Cauchie, H. M., and Gantzer, C. (2017). Interactions of infectious F-specific RNA bacteriophages with suspended matter and sediment: towards an understanding of FRNAPH distribution in a river water system. *Sci. Total Environ.* 574, 960–968. doi: 10.1016/j.scitotenv.2016.09.115

Franke, C., Rechenburg, A., Baumanns, S., Willkomm, M., Christoffels, E., Exner, M., et al. (2009). The emission potential of different land use patterns for the occurrence of coliphages in surface water. *Int. J. Hyg. Environ. Health* 212, 338–345. doi: 10.1016/j.ijheh.2008.07.003

Gaudin, R., and Barteneva, N. S. (2015). Sorting of small infectious virus particles by flow virometry reveals distinct infectivity profiles. *Nat. Commun.* 6:6022. doi: 10.1038/ncomms7022

Gautam, P., Suniti, S., Prachi, K.umari, A., Deepa, M., and Brijesh Nair, A. N. (2012). A review on recent advances in biosensors for detection of water contamination. *Agris On-line Pap. Econ. Informatics* 2, 1564–74. doi: 10.6088/ijes.00202030041

Goodridge, L., Gallaccio, A., and Griffiths, M. W. (2003). Morphological, host range, and genetic characterization of two coliphages. *Appl. Environ. Microbiol.* 69, 5364–5371. doi: 10.1128/AEM.69.9.5364-5371.2003

Grabow, W. O. K. (2001). Bacteriophages: Update on application as models for viruses in water. *Water Sa.* 27, 251–268. doi: 10.4314/wsa.v27i2.4999

Hartard, C., Banas, S., Loutreul, J., Rincé, A., Benoit, F., Boudaud, N., et al. (2016). Relevance of F-Specific RNA bacteriophages in assessing human norovirus risk in shellfish and environmental waters. *Appl. Environ. Microbiol.* 82, 5709–5719. doi: 10.1128/AEM.01528-16

Hassard, F., Gwyther, C. L., Farkas, K., Andrews, A., Jones, V., Cox, B., et al. (2016). Abundance and distribution of enteric bacteria and viruses in coastal and estuarine sediments—a review. *Front. Microbiol.* 7:1692. doi: 10.3389/fmicb.2016.01692

Hata, A., Shirasaka, Y., Ihara, M., Yamashita, N., and Tanaka, H. (2021). Spatial and temporal distributions of enteric viruses and indicators in a lake receiving municipal wastewater treatment plant discharge. *Sci. Total Environ.* 780:146607. doi: 10.1016/j.scitotenv.2021.146607

Hendrix, R. W., Smith, M. C. M., Burns, R. N., Ford, M. E., and Hatfull, G. F. (1999). Evolutionary relationships among diverse bacteriophages and prophages: all the world's a phage. *Proc. Natl. Acad. Sci. U.S.A.* 96, 2192–2197. doi: 10.1073/pnas.96.5.2192

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

Huang, X., Zhao, Z., Hernandez, D., and Jiang, S. C. (2016). Near real-time flow cytometry monitoring of bacterial and viral removal efficiencies during water reclamation processes. *Water* 8:464. doi: 10.3390/w8100464

Jebri, S., Muniesa, M., and Jofre, J. (2017). "Part Two. Indicators and Microbial Source Tracking Markers General and Indicators of Fecal Pollution," in *Water and Sanitation for the 21st Century: Health and Microbiological Aspects of Excreta and Wastewater Management (Global Water Pathogen Project)*, eds J.B. Rose and B. Jiménez-Cisneros (East Lansing, MI: Michigan State University), 1–43.

Jofre, J. (2009). Is the replication of somatic coliphages in water environments significant? *J. Appl. Microbiol.* 106, 1059–1069. doi: 10.1111/j.1365-2672.2008.03957.x

Jofre, J., Lucena, F., Blanch, A. R., and Muniesa, M. (2016). Coliphages as model organisms in the characterization and management of water resources. *Water* 8:199. doi: 10.3390/w8050199

Jones, T. H., Brassard, J., Topp, E., Wilkes, G., and Lapend, D. R. (2017). Waterborne viruses and f-specific coliphages in mixed-use watersheds : microbial associations, host specificities, and affinities with environmental/land use factors. *Appl. Environ. Microbiol.* 83, 1–22. doi: 10.1128/AEM.02763-16

Korf, I. H. E., Meier-Kolthoff, J. P., Adriaenssens, E. M., Kropinski, A. M., Nimtz, M., Rohde, M., et al. (2019). Still something to discover: novel

insights into *Escherichia coli* phage diversity and taxonomy. *Viruses* 11:454. doi: 10.3390/v11050454

Leclerc, H., Edberg, S., Pierzo, V., and Delattre, J. M. (2000). Bacteriophages as indicators of enteric viruses and public health risk in groundwaters. *J. Appl. Microbiol.* 88, 5–21. doi: 10.1046/j.1365-2672.2000.00949.x

Lee, H. S. (2009). Somatic coliphage families as potential indicators of enteric viruses in water and methods for their detection. *J. Agric. Food Chem.* [Epub ahead of print]. doi: 10.1016/j.chemosphere.2009.02.022

Lee, S., Suwa, M., and Shigemura, H. (2019). Metagenomic analysis of infectious f-specific rna bacteriophage strains in wastewater treatment and disinfection processes. *Pathogens* 8:217. doi: 10.3390/pathogens8040217

Maganha de Almeida Kumlien, A. C., Borrego, C. M., and Balcázar, J. L. (2021). Antimicrobial resistance and bacteriophages: an overlooked intersection in water disinfection. *Trends Microbiol.* 29, 517–527. doi: 10.1016/j.tim.2020.12.011

Maranger, R., Del Giorgio, P. A., and Bird, D. F. (2002). Accumulation of damaged bacteria and viruses in lake water exposed to solar radiation. *Aquat. Microb. Ecol.* 28, 213–227. doi: 10.3354/ame028213

Matsushita, T., Shirasaki, N., Matsui, Y., and Ohno, K. (2011). Virus inactivation during coagulation with aluminum coagulants. *Chemosphere* 85, 571–576. doi: 10.1016/j.chemosphere.2011.06.083

Meier-Kolthoff, J. P., and Göker, M. (2017). VICTOR: genome-based phylogeny and classification of prokaryotic viruses. *Bioinformatics* 33, 3396–3404. doi: 10.1093/bioinformatics/btx440

Molina, F., Simancas, A., Tabla, R., Gómez, A., Roa, I., and Rebollo, J. E. (2020). Diversity and local coadaptation of *Escherichia coli* and coliphages from small ruminants. *Front. Microbiol.* 11:564522. doi: 10.3389/fmicb.2020.564522

Muniesa, M., and Jofre, J. (2004). Factors influencing the replication of somatic coliphages in the water environment. *Antonie van Leeuwenhoek* 86, 65–76. doi: 10.1023/B:ANTO.0000024909.75523.be

Muniesa, M., and Jofre, J. (2007). The contribution of induction of temperate phages to the numbers of free somatic coliphages in waters is not significant. *FEMS Microbiol. Lett.* 270, 272–276. doi: 10.1111/j.1574-6968.2007.00676.x

Muniesa, M., Mocé-Llivina, L., Katayama, H., and Jofre, J. (2003). Bacterial host strains that support replication of somatic coliphages. *Antonie van Leeuwenhoek* 83, 305–315. doi: 10.1023/a:1023384714481

Mushegian, A. R. (2020). Are there 10 31 virus particles on earth, or more, or fewer ? *J. Bacteriol.* 202, e00052–20. doi: 10.1128/JB.00052-20

Nappier, S. P., Hong, T., Ichida, A., Goldstone, A., and Eftim, S. E. (2019). Occurrence of coliphage in raw wastewater and in ambient water: a meta-analysis. *Water Res.* 153, 263–273. doi: 10.1016/j.watres.2018.12.058

Ochromowicz, K., and Hoekstra, E. J. (2005). *ATP as an indicator of microbiological activity in tap* water.

Ogorzaly, L., and Gantzer, C. (2006). Development of real-time RT-PCR methods for specific detection of F-specific RNA bacteriophage genogroups: application to urban raw wastewater. *J. Virol. Methods* 138, 131–139. doi: 10.1016/j.jviromet.2006.08.004

Olivenza, D. R., Casadesús, J., and Ansaldi, M. (2020). Epigenetic biosensors for bacteriophage detection and phage receptor discrimination. *Environ. Microbiol.* 22, 3126–3142. doi: 10.1111/1462-2920.15050

Pang, L., Farkas, K., Bennett, G., Varsani, A., Easingwood, R., Tilley, R., et al. (2014). Mimicking filtration and transport of rotavirus and adenovirus in sand media using DNA-labeled, protein-coated silica nanoparticles. *Water Res.* 62, 167–179. doi: 10.1016/j.watres.2014.05.055

Pulido-Reyes, G., Magherini, L., Bianco, C., Sethi, R., von Gunten, U., Kaegi, R., et al. (2022). Nanoplastics removal during drinking water treatment: laboratory- and pilot-scale experiments and modeling. *J. Hazard. Mater.* 436:129011. doi: 10.1016/j.jhazmat.2022.129011

Putra, R. D., and Lyrawati, D. (2020). Interactions between Bacteriophages and Eukaryotic Cells. *Scientifica* 2020:3589316. doi: 10.1155/2020/3589316

Rajnovic, D., and Mas, J. (2020). Fluorometric detection of phages in liquid media: application to turbid samples. *Anal. Chim. Acta* 1111, 23–30. doi: 10.1016/j.aca.2020.03.016

Ramírez-Castillo, F. Y., Loera-Muro, A., Jacques, M., Garneau, P., Avelar-González, F. J., Harel, J., et al. (2015). Waterborne pathogens: detection methods and challenges. *Pathogens* 4, 307–334. doi: 10.3390/pathogens4020307

Rockey, N., Bischel, H. N., Kohn, T., Pecson, B., and Wigginton, K. R. (2019). The utility of flow cytometry for potable reuse. *Curr. Opin. Biotechnol.* 57, 42–49. doi: 10.1016/j.copbio.2018.12.009

Roudnew, B., Lavery, T. J., Seymour, J. R., Jeffries, T. C., and Mitchell, J. G. (2014). Variability in bacteria and virus-like particle abundances during purging of unconfined aquifers. *Groundwater* 52, 118–124. doi: 10.1111/gwat.12044

Safford, H. R., and Bischel, H. N. (2019). Flow cytometry applications in water treatment, distribution, and reuse: a review. *Water Res.* 151, 110–133. doi: 10.1016/j.watres.2018.12.016

Sime-Ngando, T. (2014). Environmental bacteriophages: viruses of microbes in aquatic ecosystems. *Front. Microbiol.* 5:355. doi: 10.3389/fmicb.2014.00355

Steinle-Darling, E., Salveson, A., Sutherland, J., Dickenson, E., Hokanson, D., Trussell, S., et al. (2016). *Direct Potable Reuse Monitoring: Testing Water Quality in a Municipal Wastewater Effluent Treated to Drinking Water Standards Volume 1 of 2*. Autin, TX: Texas Water Development Board.

Stone, E., Campbell, K., Grant, I., and McAuliffe, O. (2019). Understanding and exploiting phage–host interactions. *Viruses* 11:567. doi: 10.3390/v11060567

Sylvestre, É., Prévost, M., Burnet, J. B., Pang, X., Qiu, Y., Smeets, P., et al. (2021). Demonstrating the reduction of enteric viruses by drinking water treatment during snowmelt episodes in urban areas. *Water Res. X* 11:100091. doi: 10.1016/j.wroa.2021.100091

Vang, Ó. K., Corfitzen, C. B., Smith, C., and Albrechtsen, H. J. (2014). Evaluation of ATP measurements to detect microbial ingress by wastewater and surface water in drinking water. *Water Res.* 4, 309–320. doi: 10.1016/j.watres.2014.07.015

Virolle, C., Goldlust, K., Djermoun, S., Bigot, S., and Lesterlin, C. (2020). Plasmid transfer by conjugation in gram-negative bacteria: From the cellular to the community level. *Genes* 11:1239. doi: 10.3390/genes11111239

Weinbauer, M. G., and Suttle, C. A. (1999). Lysogeny and prophage induction in coastal and offshore bacterial communities. *Aquat. Microb. Ecol.* 18, 217–225. doi: 10.3354/ame018217

WHO (2017). Drinking water parameter cooperation project - quality of water intended for human consumption. *World Heal. Organ.* 1–228. Available online at: http://ec.europa.eu/environment/water/water-drink/pdf/20171215_EC_project_report_final_corrected.pdf. (accessed March 1, 2022).

Wilhartitz, I. C., Kirschner, A. K. T., Brussaard, C. P. D., Fischer, U. R., Wieltschnig, C., Stadler, H., et al. (2013). Dynamics of natural prokaryotes, viruses, and heterotrophic nanoflagellates in alpine karstic groundwater. *Microbiologyopen* 2, 633–643. doi: 10.1002/mbo3.98

Yang, J., Liu, S., Sui, Z., Wang, J., and Niu, C. (2019). Establishing novel bacteriophage detection method based on imagining flow cytometry. *IOP Conf. Ser. Earth Environ. Sci.* 252:042116. doi: 10.1088/1755-1315/252/4/042116

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 Microbiology

TYPE Original Research
PUBLISHED 03 October 2022
DOI 10.3389/fmicb.2022.941897

OPEN ACCESS

EDITED BY
Tony Gutierrez,
Heriot-Watt University,
United Kingdom

REVIEWED BY
Edward Hall,
Colorado State University,
United States
Wenjie Wan,
Wuhan Botanical Garden (CAS), China

*CORRESPONDENCE
Francis Hassard
francis.hassard@cranfield.ac.uk

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

RECEIVED 11 May 2022
ACCEPTED 29 August 2022
PUBLISHED 03 October 2022

CITATION
Brenes-Guillén L, Vidaurre-Barahona D, Avilés-Vargas L, Castro-Gutierrez V, Gómez-Ramírez E, González-Sánchez K, Mora-López M, Umaña-Villalobos G, Uribe-Lorío L and Hassard F (2022) First insights into the prokaryotic community structure of Lake Cote, Costa Rica: Influence on nutrient cycling. *Front. Microbiol.* 13:941897. doi: 10.3389/fmicb.2022.941897

First insights into the prokaryotic community structure of Lake Cote, Costa Rica: Influence on nutrient cycling

Laura Brenes-Guillén[1], Daniela Vidaurre-Barahona[1], Lidia Avilés-Vargas[2], Victor Castro-Gutierrez[3], Eddy Gómez-Ramírez[2], Kaylen González-Sánchez[2], Marielos Mora-López[1], Gerardo Umaña-Villalobos[2], Lorena Uribe-Lorío[1] and Francis Hassard[4,5]*

[1]Cellular and Molecular Biology Research Center, University of Costa Rica, San José, Costa Rica, [2]Research Center in Sciences of the Sea and Limnology, University of Costa Rica, San José, Costa Rica, [3]Environmental Pollution Research Center, University of Costa Rica, San José, Costa Rica, [4]Cranfield Water Science Institute, Cranfield University, Cranfield, United Kingdom, [5]Institute for Nanotechnology and Water Sustainability, University of South Africa, Johannesburg, South Africa

Prokaryotic diversity in lakes has been studied for many years mainly focusing on community structure and how the bacterial assemblages are driven by physicochemical conditions such as temperature, oxygen, and nutrients. However, little is known about how the composition and function of the prokaryotic community changes upon lake stratification. To elucidate this, we studied Lake Cote in Costa Rica determining prokaryotic diversity and community structure in conjunction with physicochemistry along vertical gradients during stratification and mixing periods. Of the parameters measured, ammonium, oxygen, and temperature, in that order, were the main determinants driving the variability in the prokaryotic community structure of the lake. Distinct stratification of Lake Cote occurred (March 2018) and the community diversity was compared to a period of complete mixing (March 2019). The microbial community analysis indicated that stratification significantly altered the bacterial composition in the epi-meta- and hypolimnion. During stratification, the Deltaproteobacteria, Chloroflexi, Bacteroidetes, Nitrospirae, and Euryarchaeota were dominant in the hypolimnion yet largely absent in surface layers. Among these taxa, strict or facultative anaerobic bacteria were likely contributing to the lake nitrogen biogeochemical cycling, consistent with measurements of inorganic nitrogen measurements and microbial functional abundance predictions. In general, during both sampling events, a higher abundance of Alphaproteobacteria, Betaproteobacteria, Actinobacteria, and Cyanobacteria was found in the oxygenated layers. Lake Cote had a unique bacterial diversity, with 80% of Amplicon Sequence Variant (ASV) recovered similar to unclassified/uncultured

strains and exhibits archetypal shallow lake physicochemical but not microbial fluctuations worthy of further investigation. This study provides an example of lake hydrodynamics impacts to microbial community and their function in Central American lakes with implications for other shallow, upland, and oligotrophic lake systems.

KEYWORDS

volcanic lake, stratification, hypolimnion, epilimnion, oligotrophic, nutrient cycling, microbiome, 16S

Introduction

Microorganisms, such as bacteria and archaea are vital for the function of freshwater environments and their roles include the: (i) conversion of abiotic resources into biomass, (ii) transformation, maintenance and continuity of nutrient resources and complex organic compounds through biogeochemical cycling in aquatic systems (Hassard et al., 2016; Khachikyan et al., 2019; Grossart et al., 2020). For instance, photoautotrophs are the main primary producers of the microbial food web, and they support lake productivity by providing carbon to higher trophic levels (de Kluijver et al., 2012). Nutrients such as nitrogen are transformed by diazotrophos into biologically more bioavailable forms such as ammonium or nitrate (Fernandez et al., 2020; Ehrenfels et al., 2021). On the other hand, the increase of nutrients such as phosphorus might drive the proliferation of algae and cyanobacteria, which can have a detrimental effect on the lake community, ecological function, and its services to people (Michalak et al., 2013). Natural variability in environmental conditions gives rise to inherent heterogeneity, which in turn drives the dynamics, functional profile, and composition of the microbial communities in these systems (Yadav et al., 2019). Therefore, it is of significance to decipher distribution patterns and factors affecting community composition of aquatic prokaryotes.

Previous studies have found that in shallow and deep lakes, Proteobacteria are the most abundant phylum, generally followed by Chloroflexi, Cyanobacteria, Actinobacteria, and Bacteriodetes (Salmaso, 2019; Tran et al., 2021). Changes in the bacterial composition along the vertical profile have been primarily related to thermal stratification, which leads to strong gradients in oxygen and nutrient concentrations (e.g., N and P) (Diao et al., 2017; Tran et al., 2021; Díaz-Torres et al., 2022). The differences in the bacterial composition in the hypoxic layers are evident at lower taxonomic levels (Salmaso, 2019). Next generation sequencing technologies have been used to understand the genetic resources, distribution patterns, and the regulation of the microbial communities in freshwater lakes. However, it remains unclear how composition and function of prokaryotic community change during sequential periods of mixing and stratification (Aguilar et al., 2022).

Tropical lakes are important engines of local socioeconomic growth and stability (Lewis, 2002; Osborne, 2004; Menezes et al., 2019). Lakes provide many ecosystems and human beings with services such as water supply, water recharge, temperature regulation, recreation, and commercial activities. Lakes especially in Africa and in other tropical regions are especially vulnerable due to problems such as water eutrophication, harmful algal blooms, heavy metal, and agrochemical pollution, and global warming, the latter affecting their water levels and mixing regimes (Hecky et al., 2010; Simiyu et al., 2022).

Stratification is a well-established phenomenon in all but the most oligotrophic systems. In our study, we choose one representative lake to resolve the ecological event of stratification and mixing and impacts to microbiome. Lake Cote is a discontinuous polymictic shallow lake (depth ∼10 m), located 650 m above sea level and exhibits an intermittent thermocline present at 6 m depth. Stratification occurs during calm meteorological conditions which act to inhibit wind driven mixing of the lake. Previous studies in the lake have focused primarily on the physiochemistry e.g., nutrient, chlorophyll, dissolved oxygen, temperature, and conductivity (Umaña et al., 1999; Sibaja-Cordero and Umaña-Villalobos, 2008; Umaña, 2014). Additionally, taxonomic analyses of its macrofauna, benthic meiofauna, and plankton have been undertaken (Umaña-Villalobos and Aviles-Vargas, 2020).

We posit that stratification of Lake Cote is a driver for anoxic biogeochemical cycling to be established during periods of low mixing, and the diversity, and function of the microbial community will effectively reflect these physicochemical characteristics. In addition, certain nutrients (e.g., N) and oxygen concentrations may be suitable for the presence of specific and novel microorganisms. We aimed to (i) investigate changes in water physicochemical properties, (ii) decipher shifts in relative abundances of bacterial taxa and functions of the bacterial community, and (iii) elucidate drivers affecting bacterial community composition during mixing and stratification periods in shallow upland lake.

Materials and methods

Site characteristics

Lake Cote (10°34′54°N and 84°54′42°W) is located at 650 m above sea level, at the southern end of the Guanacaste volcanic range. It is the largest natural lake in Costa Rica, a size exceeding 1.99 km^2 and a maximum depth of 12 m. The lake has two tributaries, the Pierna de la Laguna river, and Zoncho stream, and used to have an outlet through the river Cote to the Caribe slope. Lake Cote is now dammed with a direct outflow to the Arenal Reservoir through a tunnel that was built during the 1980s. Two sampling events were conducted in March 2018 and March 2019, in each 12 water samples were obtained from six depths (0, 2, 4, 6, 8, and 10 m) within the deepest zone of the lake.

Sample collection and physicochemical analyses

We collected 12 water column samples from various depths in Lake Cote during March 2018 and March 2019 using 2 L alpha bottles (WildcoTM) while aboard an Apex rubber inflatable boat. The 12 water samples were taken from six depths (0, 2, 4, 6, 8, and 10 m) including the surface (0–2 m depth), the epilimnion (2–8 m), and the hypolimnion (8–10 m). The samples were stored immediately upon collection at 4°C within a cooler on melting ice, then they were storage in the same cooler for 24 h. The dissolved nutrients were quantified using a FIA LACHAT QuickChem8500 Flow Injection Analysis System, our method was modified from Standard Methods for the Examination of Water and Wastewater (American Public Health Association [APHA], 1998). Ammonium, nitrite, nitrate, and orthophosphate Soluble Reactive Phosphorus (SRP) concentrations were quantified based on 4500-NH3, 4500-NO3, 4500-NO2, and 4500-P protocols, respectively. Chlorophyll-a was extracted using 90% acetone and measured in a Thermo^{TM} UV 700 spectrophotometer Shimadzu (PharmaSpec UV-1700), at the wavelengths established in an empirical equation (Parsons et al., 1984). Alkalinity was measured using Hach's Alkalinity Test Model AL-TA. The vertical profile of temperature and dissolved oxygen (0–10 m depth) was measured *in situ* using a YSI meter (Models Pro20). Light attenuation (turbidity) was estimated using a Secchi disc.

Nucleic acid extraction

At each sampling depth we filtered 1 L of water *in situ* through a sterile 0.2 μm pore size cellulose nitrate filter (MilliporeTM 114). Filters were placed in 50 mL of sucrose lysis buffer (0.75 M Sucrose, 40 mM EDTA, pH 8, 50 mM TRIS-HCl, pH 8), and were stored initially on melting ice before being transferred within 48 h to the laboratory and subsequently stored at -70°C prior to DNA extractions. For the DNA extractions the filters were thawed on ice, agitated for 10 min (DS LAB Orbital Shaker, 100 rpms) and 1 mL of the solution was processed using the Nucleospin Soil Genomic DNA extraction kit (Macherey NagelTM) according to manufacturer's instructions. Integrity of the DNA was examined by electrophoresis in 1.0% agarose gels and quantified with a NanoDropTM ND-1000. DNA extracts were stored at -20°C.

Ilumina MiSeq sequencing, noise removal, and taxonomic assignment

DNA samples were sent to Macrogen Inc. (Seoul, South Korea) for amplification of the 16S rRNA gene using the MiSeq Platform (Illumina Inc.) forward Primer: 5'-TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG-5' and reverse primer 5'-GTC TCG TGG GCT CGG AGA TGT GTA TAA GAG ACA G-3' for amplification of the hypervariable regions V3–V4 (Mori et al., 2014), and approximately 300 bp long tags were obtained. Raw sequence data obtained from the Illumina sequencing platform were processed using QIIME2 (version 2018.11) (Bolyen et al., 2019) and its plugins. Further, the "dada2" plugin (Callahan et al., 2016) was applied to truncate reads (–p-trim-left 0, –p-trunc-len 280). We applied the vsearch uchime_ref method to identify chimeric feature sequences. Mitochondria and chloroplast sequences were removed. The KEGG Orthology (KO) predictions based on 16S data was obtained using PICRUSt2 plugin from QIIME 2 (Douglas et al., 2020). Taxonomy was assigned to the Amplicon Sequence Variant (ASV) against a SILVA 16S/18S rDNA non-redundant reference dataset (SSURef 132 NR) (Quast et al., 2012) with the feature-classifier classify-sklearn. Some taxonomic assignments of the ASVs were manually checked by comparing them with sequences in the database using a combination of initial BLASTN-based searches and an extension of the EzTaxon cured database (Yoon et al., 2017). We used the criteria published by Chun et al. (2010) for the taxonomic assignment of each read (x = similarity): species ($x \geq 97\%$), genus (97%; $x \geq 94\%$), family (94%; $x \geq 90\%$), order (90%; $x \geq 85\%$), class (85%; $x \geq 80\%$), and phylum (80%; $x \geq 75\%$). Reference GenBank sequences were used to illustrate the relationship of sequences to representative taxa. Nucleotide sequence accession numbers of partial 16S rRNA gene sequences were deposited to GenBank under the following accession numbers: MT458149-MT458168, MT351156-MT351180, MT348498-MT348459, MT26904-MT269071, MT241880-MT241845, MT039245-MT039296, and MK912176-MK912244.

Statistical analyses

All the analyses were performed using PRIMER 7/PERMANOVA+ (Clarke and Gorley, 2015). Non-metric multidimensional scaling (NMDS) was used to illustrate the dissimilarities among microbial community assemblages (after square root transformation of absolute abundances of the generated ASVs) using Bray-Curtis distances while the Euclidian distance was used for normalized environmental variables. The most abundant ASVs (contributing at least 2% to the overall abundance) were used to create a shade-plot graphs, and along with samples, were arranged and clustered according to their similarity (Bray Curtis similarity for samples and Whitaker index of association for species). The influence of year of sampling and depth on the microbial community composition (or predicted function) was assessed using permutational analysis of variance (PERMANOVA). To analyze the correlation between microbial community structure (or predicted function) and environmental data, we used a distance-based linear model (DistLM) with a stepwise variable selection approach based on the Akaike selection criterion (AIC). To represent the relationship between physicochemical variables and the microbial community (or predicted function) of the samples we used distance-based redundancy analysis (dbRDA). The Wilcoxon rank-sum test used to compare physicochemical properties between 2018 and 2019 was calculated using SPSS (IBM Corp, 2021).

Results

Physicochemical data

The vertical profiles of the physicochemical parameters examined in this study (e.g., temperature, dissolved oxygen, and nutrients) are shown in **Figure 1**. We defined the epilimnion between 0 and 6 m, metalimnion at 8 m, and hypolimnion at 10 m in 2018. There was significant difference ($p < 0.05$) in temperature, oxygen, nitrate, nitrite, ammonium, alkalinity, and chlorophyll-a between 2018 and 2019. The temperature and oxygen profiles measured in 2018 showed a moderate temperature decrease (~0.5°C) through the water column and a pronounced two-fold decrease of oxygen occurring at ~8 m. The peak temperature and oxygen concentrations occurred close to the surface (24.6°C and 7 mg/L). In the 2019 samples, temperature, and oxygen remained constant at most depths, at 24.0°C and 7.5 mg/L, throughout the water column confirming the mixed status of the lake during this period.

Alkalinity and chlorophyll-a showed higher values in 2019 than 2018 and the maximum values occurred at different depths. In 2019 the chlorophyll-a maximum was detected at ~4 m and was ~1.5-fold higher (7.7 mg/m^3) than the 2018 chlorophyll-a peak, which occurred in the metalimnion (8 m; 5.5 mg/m^3). During the 2018 sampling event, the ammonium and nitrate values exceed those measured in 2019. A peak in ammonium was detected at 6 m in 2018 and was between one-fold higher and three-fold higher than in 2019, depending on the depth. Nitrate was ~5-fold higher in 2018 than in 2019. The elevated values of phosphorus and nitrite occurred in 2019 were not observed in 2018.

Principal component analysis (PCA) described 72% of the total variance in prokaryotic diversity (based on relative abundance) along two principal dimensions (PCA1 and PCA2; **Supplementary Figure 1**). PERMANOVA analyses revealed that the year of sampling had a significant effect on microbial community structure ($p = 0.002$), whereas sampling depth did not ($p = 0.422$). Temperature did not decrease by more than ~0.5°C from the surface although the decrease was slightly greater in 2018. Samples in 2018 were enriched in nitrate and ammonium, while nitrite, chlorophyll-a, and phosphorus were present in greater concentrations in 2019 (**Figure 1**).

Community structure patterns

We identified a total of 3,220 ASVs affiliated with 52 prokaryotic phyla, with 80% of ASVs similar to unclassified/uncultured strains (**Figure 2**). Proteobacteria, Actinobacteria, Cyanobacteria, Verrucomicrobia, and Nitrospirae comprised >60% of the reads. Other phyla were present in smaller abundances (1 to <8%). Archaeal ASVs were found only a depth of 10 m 2018 (<1%). Community structure among depths (0, 2, 4, 6, 8, and 10 m) was compared during 2018 (the lake was stratified), and 2019 (the lake was mixed) using the Bray-Curtis similarity index (**Figure 3**). The analysis revealed distinct microbial community with sampling event indicating a time component to the observed community. However, the sample obtained at 10 m during 2018 was different from the rest in its respective collection year (at 30% similarity) with most of this variability driven by PC2. Samples from the epilimnion (0–2 m) also clustered differently from the rest. In 2019, the bacterial community reflected the mixed status of the water column (60% similarity). The NMDS analysis confirmed the separation between the hypolimnion (2018) and the remaining samples (**Figure 2**). A discrete separation between epilimnion and the metalimnion can be observed. The samples corresponding to lake mixing cluster together tightly. Stratification produced low dissolved oxygen zones within the lake at the hypolimnion and a well-oxygenated surface at the epilimnion, both showing a distinct microbial community.

The shade-plot analysis shows the first 30 statistically significant ASVs with similar patterns of abundance across the samples, clustered using the Whitaker index of association (**Figure 3**). The community was dominated by uncultured microorganisms and the specific identity these sequences derived are available (**Supplementary Table 1**). There were

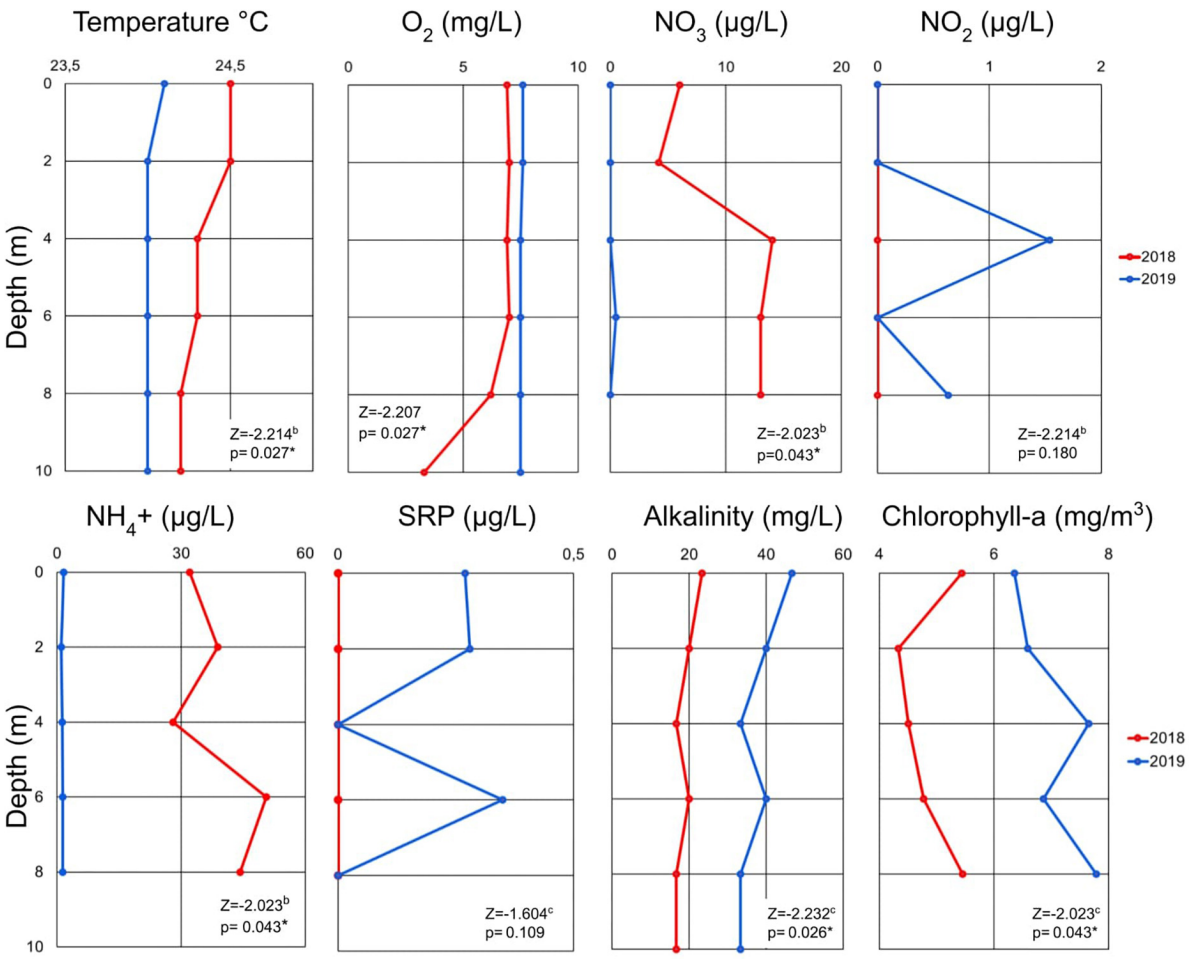


FIGURE 1
Vertical profile of physicochemical variables in Cote Lake for the two sampling dates in 2018 and 2019 according to depth. Wilcoxon sum rank test values for the comparison between 2018 and 2019 variables are showed in the table. The asterisk shows significant differences ($p < 0.05$) (SRP: Soluble Reactive Phosphorus; The 10 m depth data collected in 2018 excluded due to possible contamination because the alpha bottle touched the sediment).

some taxa which showed a ubiquitous distribution during both sampling events (e.g., 2018 and 2019), whilst rare and uncharacterized microorganisms were found mainly associated to the hypolimnion suggesting this environment presented a different ecological niche for bacteria. To understand the drivers of this variability, we performed a DistLM correlation analysis between biotic and environmental data which indicated that, sequentially, ammonium (p = 0.0026, 33.7%), oxygen (p = 0.0070, 21.3%), and temperature (p = 0.0112, 12.0%) were the main predictors of the variability in microbial structure of the samples and clear trends emerged during periods of stratification versus mixed water column. This is represented in the distance-based redundancy analysis (dbRDA) in which the environmental descriptors explained 67.4% of the community variation (Figure 4).

Analyses of hypothesized functions for nitrogen cycling using PICRUSt2 (Figure 5) showed that during stratification, those related to nitrogen fixation, nitrification, some determinants of denitrification, and dissimilatory nitrate reduction were more prevalent in the hypolimnion than in the metalimnion and epilimnion. This coincided with the increase in ammonium concentrations with depth, which were higher in the hypolimnion, presumably due to ammonification, ammonium diffusion from the sediment or dissimilatory nitrate reduction to ammonium, while nitrates were more abundant at the metalimnion. During the mixed lake condition, the predicted function related to nitrogen cycling showed smaller differences with depth than what could be observed during stratification, which coincided with the homogeneous ammonium concentrations detected throughout the water column (Figure 1).

To understand the influence of the environmental variables in the hypothesized function of the microbial communities, once again, a DistLM analysis was performed, which showed

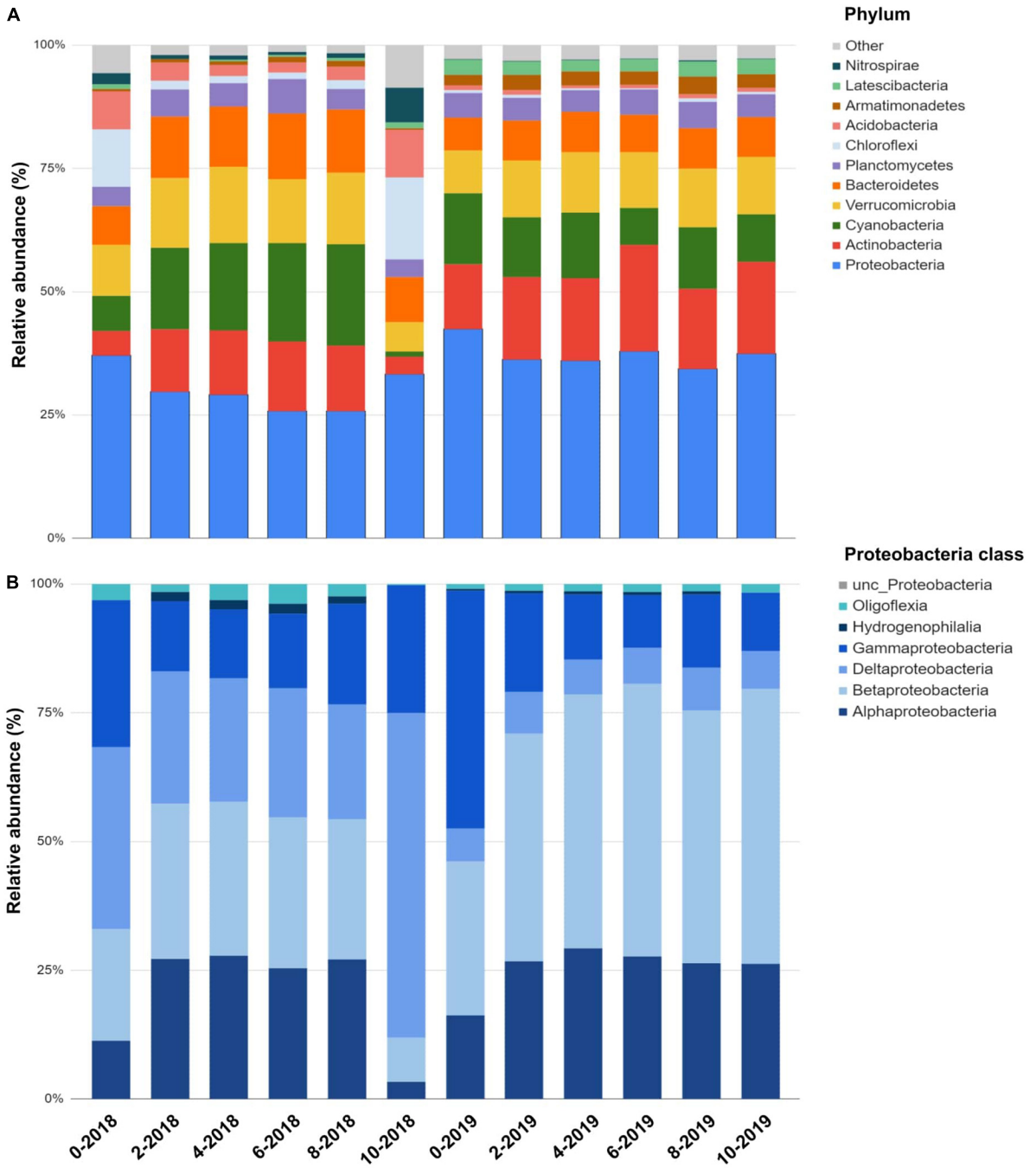


FIGURE 2
Microbial community structure in Lake Cote for the two sampling dates in 2018 and 2019 according to depth. **(A)** Profile according to phylum level, the category "others" includes Spirochaetes, Firmicutes, Nistropinae, Patescibactera, Omnitrophicaeota, Rokubacteria, Kiritimatiellaoeta, TA06, and Zixibacteria. **(B)** Proteobacteria classes relative abundance along the Lake Cote water column in 2018 and 2019. Deltaproteobacteria was dominant in the hypolimnion during lake stratification.

that, sequentially, temperature (p = 0.0003, 42.4%), depth (p = 0.0156, 14.0%), and oxygen (p = 0.0356, 12.0%) were the main predictors of the variability in predicted microbial function. This is represented in the dbRDA in which the environmental descriptors explained 79.6% of the community variation (Figure 6). PERMANOVA analyzes revealed that the year of sampling had a significant effect on hypothesized overall microbial community function (p = 0.0104), whereas, sampling depth did not (p = 0.4515).

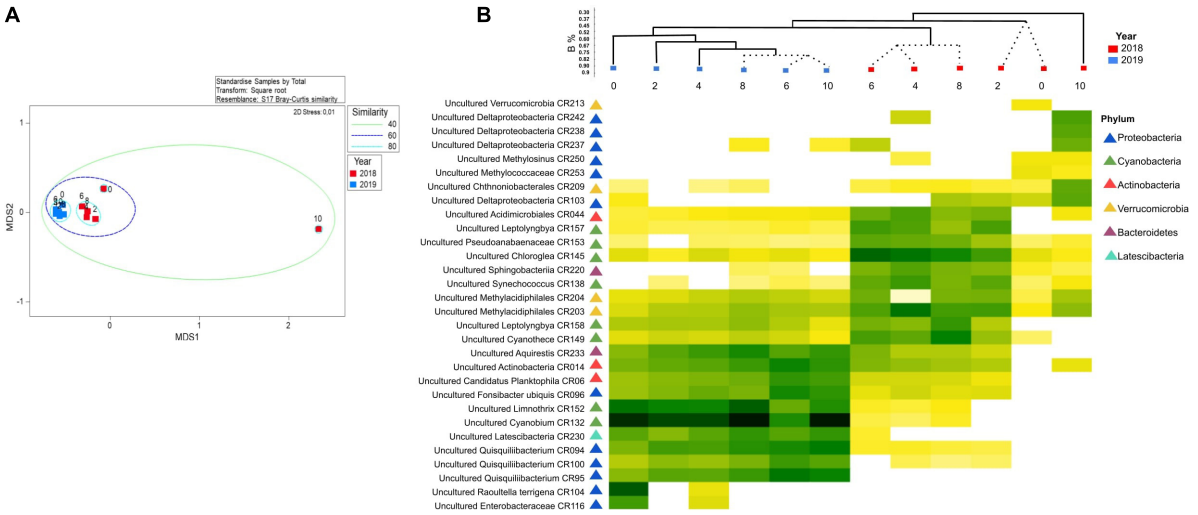


FIGURE 3
(A) Non-metric multidimensional scaling (NMDS) of bacterial community structure for the Lake Cote. Similarity circles were based on Bray-Curtis coefficient. **(B)** Shade-plot of standardized, square root-transformed abundance of 30 ASVs, selected by index of association. Samples were arranged according to LINKTREE results, where dotted lines indicate no significant relation between the samples. Dendrogram using Bray-Curtis similarity coefficient showing the similarity among the samples.

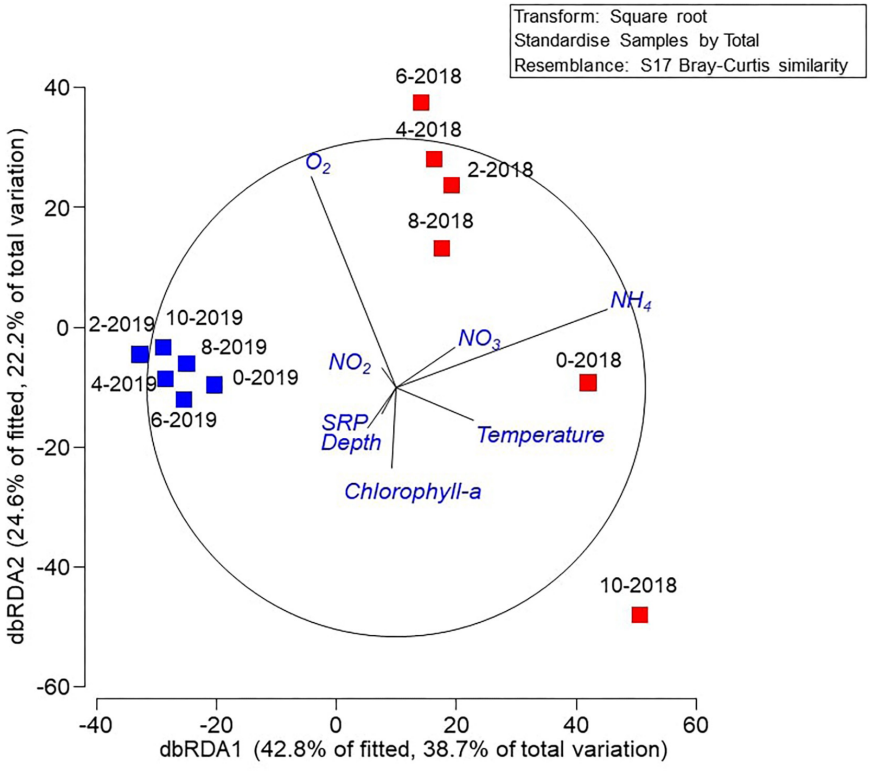


FIGURE 4
Distance-based redundancy analysis (dbRDA) plot of the DistLM based on geochemical parameters fitted to the variation in ASV abundance. Vectors indicate direction of the parameter effect in the ordination plot.

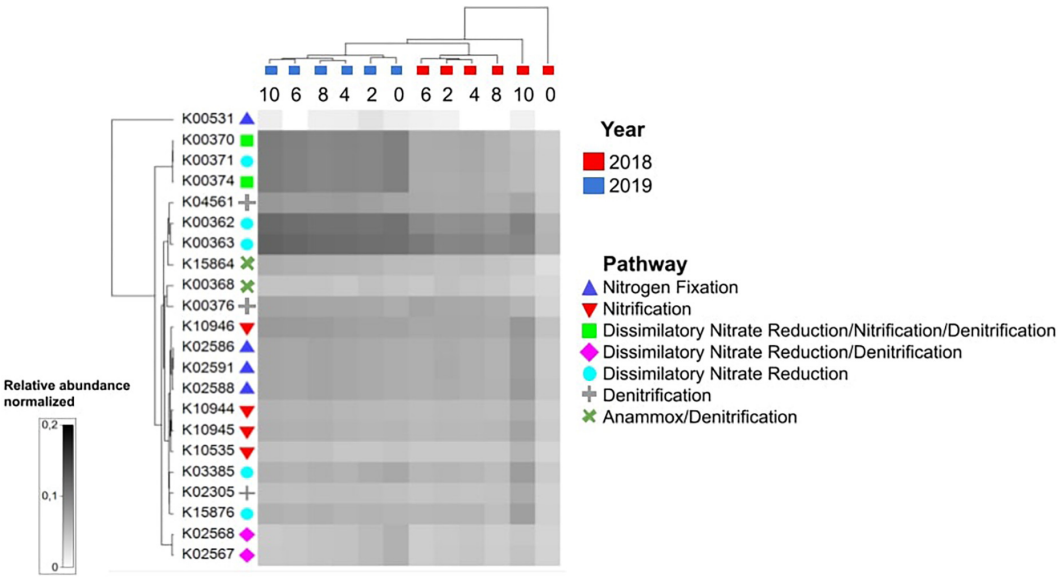


FIGURE 5
PICRUSt2 functional prediction analysis for nitrogen cycling functions across different temporal and spatial scales. During stratification, KOs related to nitrogen metabolism were more prevalent in the hypolimnion than in the other strata.

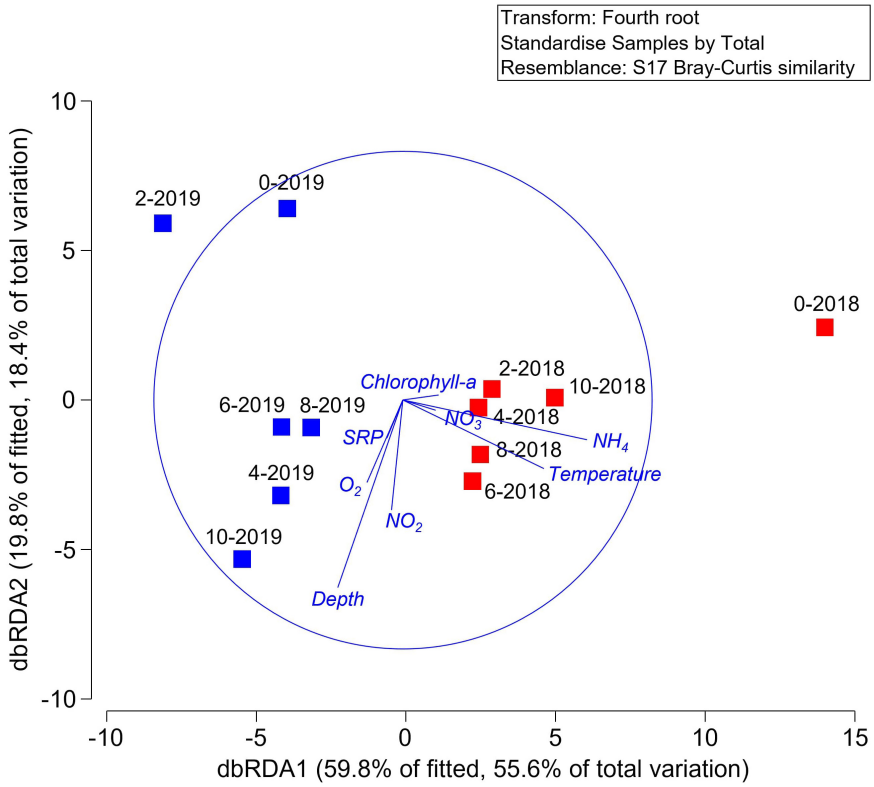


FIGURE 6
Distance-based redundancy analysis (dbRDA) plot of the DistLM based on geochemical parameters fitted to the variation in functional abundance prediction using PICRUSt2. Vectors indicate direction of the parameter effect in the ordination plot.

Diversity of microbial communities in Lake Cote during lake stratification

Proteobacteria comprised up to 42% of the bacterial community in the water column of Lake Cote in 2018 (**Figure 2**). In the hypolimnion, Deltaproteobacteria (Desulfobacterales) was the most abundant class within this phylum (63%) (**Figure 2**). ASVs classified within the Nitrospinaceae family were the most abundant in the hypolimnion and were 96.7% similar to the uncultured Nitrospirales strain in the sediment of Lake Dongping in China (Song et al., 2012), which is a eutrophic, temperate, and shallow lake (mean depth ∼1–2 m). After Deltaproteobacteria, Gammaproteobacteria followed in abundance. Within this class, the family Methylococcaceae was the most abundant.

Acidobacteria were more abundant during lake stratification in the epilimnion and hypolimnion (**Figure 2**). We found 35 ASVs classified into GP3, GP4, GP6, GP7, GP17, GP18, GP22, and GP23 subdivisions. Six ASVs were affiliated to the Bryobacteraceae family. Within Acidobacteria, GP6, GP7, GP17, and GP18 subdivisions were predominant in the water column, being more abundant at 0 m depth. ASVs affiliated to these subdivisions were similar to uncultured strains from soils, water columns of lakes, wetlands, and sediments (**Supplementary Table 1**). Based on the classification by Mehrshad et al. (2018), seven Chloroflexi classes were detected in Lake Cote, the majority of the ASVs affiliated to non-cultured strains from the benthic sediments of lakes (**Supplementary Table 1**). The families Candidatus Profundisolitariaceae fam. nov. within the CL500-11 cluster, Anaerolineae, Dehalococcoidia, Ktedonobacteraceae, and Roseiflexaceae were most abundant in the anoxic layer. Verrucomicrobia comprised between 6 and 15% of the sequences. ASVs were identified as the classes Verrucomicrobiae, Spartobacteria, Methylacidiphilaceae, Pedosphaera, and Opitutae (**Supplementary Table 1**). The microorganisms within the Nitrospirae phylum were found throughout the water column in both years. In 2018, all the ASVs within Nitrospirae were identical (96–98%) to *Magnetobacterium bavaricum* isolated from sediments in Lake Chiemsee, Germany (Spring et al., 1993; **Supplementary Table 1**).

Diversity of microbial communities in Lake Cote during lake mixing

In 2019, Alphaproteobacteria and Betaproteobacteria were the most abundant classes within Proteobacteria and had a ubiquitous distribution. Their relative abundance increased during lake mixing in 2019 (up to 20%, **Figure 2**). ASVs classified within the Alphaproteobacteria class as *Novosphingobium* sp., *Hyphomicrobium* sp., and *Ca Fonsibacter* sp. (LD12 group) were the most abundant in the lake. Reads similar to *Rhodoferax*, *Acidovorax*, *Limnohabitans*, and *Quisquiliibacterium* represent over 80% of sequences for the class Betaproteobacteria. Their abundance increased in the 2019 sampling (**Figure 3**). All of them are aerobic and mesophilic bacteria. *Acinetobacter*, *Enterobacter*, *Citrobacter*, *Escherichia*, *Proteus*, *Kluyvera*, and *Yokenella* sequences were found in 2019 in the upper lake samples (0–2 m). These microorganisms are often associated with human activities such as the presence of homes or farms located around the lake.

Common freshwater taxa, such as members of Actinobacteria, were found throughout the water column in both 2018 and 2019, however they comprised between 13.3 and 21.6% of the total bacterial community throughout the water column when the lake was mixed. All ASVs were similar to uncultured Actinobacteria strains found in freshwater reservoirs, soil, or marine environments (**Supplementary Table 1**). The most abundant Actinobacteria related ASVs were affiliated to sequences within the acI lineage such as Conexibacteraceae, Rubrobacterales, Gaeillales, Coriobacteriaceae, Ilumatobacteracea, Acidimicrobiales, Iamiaceae, Microtrichaceae, Actinomarinaceae, Sporichthyaceae, and Micrococcales. *Ca. Planktophila* spp, *Ca. Nanopelagicus* spp, *Ca. Planktoluna* spp, and members of the CL500-29 clade were the most prevalent genotypes. All ASVs in this study were similar to uncultured Actinobacteria strains found in freshwater reservoirs, soil, or marine environments (**Supplementary Table 1**).

Discussion

Impact of stratification on microbial community

The abundance and distribution of bacteria in lake environments can be influenced by numerous factors such as substrate and oxygen availability (Hassard et al., 2017). During periods of mixing there was a high relative abundance of aerobic heterotrophic bacteria present in the water column. Anaerobic microorganisms were prevalent in the bottom of the lake during stratification periods and may play an important role in carbon, nitrogen, sulfur, and other nutrient cycles (Tran et al., 2021). The community structure of Lake Cote was influenced by the physicochemical conditions, mainly ammonium, oxygen, and temperature. Yadav et al. (2019) indicated that bacterial communities differ less within a lake than between lakes, however ecological niche separation of the microbial taxa and microbial pattern distribution are influenced by the resource availability, physicochemical conditions, and other microbial factors such as grazing by meiofauna and ingress from surface water run-off for example.

The richness of the community based on the Simpson index was higher in the hypolimnion in 2018 than in 2019

(D = 109 and D = 35, respectively). In anoxic zones, low oxygen levels and high concentrations of some key nutrients (i.e., ammonium) may have driven an increase in community diversity, which would explain why it was higher in the hypolimnion compared to the epilimnion. This is consistent with previous results in stratified and mixed shallow lakes, which showed a higher number of unique OTUs in the hypolimnion versus the epilimnion (Schmidt et al., 2016). In Lake Cote, the top of the oxygen minimum zone was very shallow and had a brief period of stratification, however we found a notable increase in microbial diversity in this transitional hypoxic layer, where the co-occurrence of aerobic and anaerobic processes may explain the high diversity (Fernandes et al., 2020).

As expected, Proteobacteria was the most abundant phylum in the water column in line with other studies in lake ecosystems (Yadav et al., 2019). This taxon is very heterogeneous from a metabolic and ecological viewpoint. Proteobacterial community composition varied in our study, probably driven by a combination of environmental factors, including oxygen, nitrate, and ammonium showed significant positive effect on the diversity of this phylum. The metabolic and taxonomic diversity could be enhanced by high concentration of the nitrogen species (e.g., ammonium) and perhaps the long stratification periods.

The effects of nutrient status on bacterial diversity in lakes can have diametrically opposite conclusions depending on the balance of specific substrates and metabolic capabilities of the microbial communities (Liu et al., 2017). The exact role these variables can play regarding bacterial diversity is a complex one and is not completely understood. For instance, bacterial richness was positively influenced by total nitrogen, organic carbon, and total phosphorus in oligotrophic lakes in Sweden (Logue et al., 2012). Wang et al. (2022) found that in less eutrophicated lakes, microbial diversity was higher than in those with more eutrophication. On the other hand, Dai et al. (2016) found that bacterial diversity was negatively correlated with trophic level in plateau freshwater lakes.

Role of stratification to microbial driven nutrient cycling

The identification of the microbial diversity within Lake Cote and the physicochemical data collected here provide early findings that can hint toward potentially microbially-driven biogeochemical processes. The role of microorganisms in the cycling of key nutrients (e.g., N) at lower depths is important for the biogeochemistry of these lakes especially during stratification (Tran et al., 2021).

It has been suggested that members of some of the Proteobacteria phylum (Alphaproteobacteria, Betaproteobacteria, and Deltaproteobacteria), the most abundant taxa found in this study, play an important role in the carbon, nitrogen, and sulfur biogeochemical cycling in lake environments. For instance, ASVs belonging to the Alphaproteobacteria LD12 group (also known as SAR11 clade IIIb) which were abundant throughout the water column, and are ubiquitous in other freshwater environments in Russia (Cabello-Yeves et al., 2018), China (Oh et al., 2014), and USA (Tsementzi et al., 2019) and were shown to play a role in carbon fixation and oxidation. The non-LD12 group, mostly abundant in Lake Cote at depths with predominantly anoxic conditions, has been implicated in nitrogen fixation (Tran et al., 2021). Other Alphaproteobacteria present in this system (e.g., *Hyphomicrobium* spp.) have the capacity for complete denitrification (Martineau et al., 2015).

Nitrospina, another inhabitant of the hypolimnion, which belongs to the Deltaproteobacteria class, is a nitrite oxidizing bacterium with the capacity to oxidize nitrites to nitrates (Lücker and Daims, 2014). Even though no correlation was found between its abundance and nitrate concentration in the present study, its abundance did correlate with ammonium concentration. Based on genome annotation, Nitrospinae species also possess an Amt-type transporter for taking up ammonium (Ngugi et al., 2016), suggesting that Nitrospina could have a role in ammonium oxidation (COMMAMOX) process. This strain was found throughout the water column and its abundance suggests the capacity of active oxidation of nitrite to nitrate.

Other bacteria usually involved in nitrogen metabolism found in the present study was *Nitrospira*, a member of the phylum Nitrospirota. Its abundance was negatively correlated with total phosphorus; a similar result was obtained in Lake Dongping, China, although both lakes have very different characteristics, the former being eutrophic and having high concentrations of nutrients (>20 mg/kg NH_4-N, >3 mg/kg NO_3-N) derived from industrial and wastewater inputs, whereas, Lake Cote is oligotrophic, upland and less impacted by human activity.

Actinobacteria, Bacteroidetes, and Verrucomicrobia, also prevalent in Lake Cote, may be involved in nitrate oxidation and nitrate ammonification. Actinobacteria abundance could be affected by the amount of carbon generated by cyanobacteria, algae, decaying plant material, and by-products from fungi, crustaceans, and diatoms (Lewin et al., 2016). As most Actinobacteria are obligate aerobes, some of them have to survive long periods under hypoxic conditions in a latent, non-replicative state (den Hengst and Buttner, 2008). This phylum is considered ubiquitous in freshwater ecosystems. These lineages of freshwater Actinobacteria are cosmopolitan, dominant members of lake bacterial communities and have an important role in carbon and nitrogen cycling (Ghylin et al., 2014; Lewin et al., 2016). Although the acI lineage is widely distributed in freshwater ecosystems, it is difficult to use culture-dependent methods for molecular and phylogenetic identification of this phylum, however, some cultivation efforts have been carried out (Kang et al., 2017).

Diazotrophs are other key players in supplying new nitrogen to the aquatic ecosystems. Aerobic and anaerobic microbial nitrogen transformation processes were reported by Klawonn et al. (2016) within cyanobacterial aggregates in aerated surface waters in the Baltic Sea. Cyanobacterial ASVs belonging to Prochlorococcus and Chlorogloea, a nitrogen fixer (Fay, 1965) were the most abundant in the hypolimnion. High densities of diazotrophic cyanobacteria were associated with a low nitrate supply and remained phosphate in the surface waters (Ehrenfels et al., 2021). Members of Deltaproteobacteria, Chloroflexi, and Euryarchaeota are also nitrogen-fixing bacteria, and *Nitrospira* has been identified as capable of performing complete COMAMMOX (Tran et al., 2021).

Microbial diversity of Central American lakes

The vast majority of the ASVs observed during stratification in Lake Cote were identical to uncultured strains from different environments such as lakes and sediments. Genomic databases have been compiled from sequences obtained from culture-independent analysis; it has become common strategy to designate new phyla or higher taxa without designating lower rank (Chuvochina et al., 2019). The high abundance of uncultured taxa within the majority of phyla found in Lake Cote reflects the fact that the culture of bacteria from soil or freshwater ecosystems in tropical habitats is scarce; laboratory equipment provisions in these regions need to be enhanced to support the establishment of culture collections of rare bacteria. Uncultured microorganisms are a source of secondary metabolites, new niches, and novel lineages of bacteria and archaea, that could be important for understanding the ecology and evolution of microorganisms in freshwater ecosystems. Further studies are being performed to obtain novel cultured bacteria from Lake Cote using different approaches to enable genome comparative analysis and strain characterization.

Temporal studies provide a better resolution of the community structure than point studies and allow us to observe microbial compositional changes. Tropical lake microorganisms are genetically and metabolically diverse and ecologically important, although some studies have concluded that climate change might bring increased stratification and anoxia to freshwater lakes. Heterotrophic and autotrophic microbial diversity studies have been used as an indicator for the ecological conditions and suitability of the water quality, and as source of novel enzymes, and metabolites. However, there are few studies in the tropical area and the information is scarce. The use of high-throughput sequencing brings many advantages to the study of microorganism taxonomy, abundance, and roles in communities, but culture-dependent approaches should be developed to provide morphological and molecular descriptions and a better understanding of their physiology. These studies are very important for understanding the full metabolic potential of the lake ecosystem because not all energy is processed through phytoplankton photosynthesis and community respiration, and bacteria play key roles in the lake energy, material, and nutrient balances.

Conclusion

Central American freshwater lakes are vital ecosystems as engines for regional biogeochemical cycling, and aquatic biodiversity. Here, the prokaryotic diversity of Lake Cote was determined and for the first time, the role for nutrient cycling was posited for such systems. Of the parameters measured, the ammonium, oxygen, and temperature, in that order, were the main measured determinants driving the variability in the microbial community structure of the lake. During periods of distinct stratification, the Deltaproteobacteria, Chloroflexi, Bacteroidetes, Nitrospirae, and Euryarchaeota were dominant in the hypolimnion yet largely absent in surface layers. Anaerobic bacteria were likely contributing to the lake nitrogen biogeochemical cycling, confirmed *via* inorganic nitrogen measurements and microbial functional abundance predictions during stratification. The prokaryotic diversity of Lake Cote has been poorly studied previously, which was reflected in the finding that 80% of ASV recovered similar to unclassified/uncultured strains. Shallow lake physicochemical and microbial fluctuations are therefore worthy of further investigation. Future work should establish whether predicted functional abundance match actual functional characteristics elucidated using chemical, transcriptomic or metabolic assessments. Furthermore, microbial assemblages of the tropical lakes should be investigated to determine their responses to current and expected climate change, which is predicted to increment the stratification and hypoxia in the freshwater lakes.

Data availability statement

The original contributions presented in the study are publicly available. This data can be found here: https://www.ncbi.nlm.nih.gov/bioproject/PRJNA880862.

Author contributions

LB-G, GU-V, and LU-L designed the study. LB-G, LU-L, GU-V, and VC-G wrote the original manuscript. DV-B, LA-V, EG-R, KG-S, MM-L, GU-V, and LU-L edited the manuscript. FH advised on the analysis. FH and VC-G edited and commented on the final manuscript. All authors agreed to be accountable for the content of the work and approved the submitted version.

Funding

This work was supported by Vicerrectoria de Investigacion, University of Costa Rica project VI-801-B8238.

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/fmicb.2022.941897/full#supplementary-material

SUPPLEMENTARY FIGURE 1
Principal component analysis (PCA) of the samples, 72% of the total variance along two principal dimensions.

SUPPLEMENTARY TABLE 1
Abundance of amplicon sequence variants (ASVs) analyzed in Lake Cote and closely related sequences in the GenBank database.

References

Aguilar, P., Vila, I., and Sommaruga, R. (2022). Bacterioplankton zonation does exist in high elevation, polymictic Lakes. *Front. Microbiol.* 13:764566. doi: 10.3389/fmicb.2022.764566

American Public Health Association [APHA] (1998). *Standard Methods for the Examination of Water and Wastewater*, 20th Edn. Washington, DC: American Water Works Association, Water Environment Federation.

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

Cabello-Yeves, P. J., Zemskaya, T. I., Rosselli, R., Coutinho, F. H., Zakharenko, A. S., Blinov, V. V., et al. (2018). Genomes of novel microbial lineages assembled from the sub-ice waters of Lake Baikal. *Appl. Environ. Microbiol.* 84:e02132-17. doi: 10.1128/AEM.02132-17

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

Chun, J., Kim, K. Y., Lee, J. H., and Choi, Y. (2010). The analysis of oral microbial communities of wild-type and toll-like receptor 2-deficient mice using a 454 GS FLX *Titanium pyrosequencer*. *BMC Microbiol.* 10:101. doi: 10.1186/1471-2180-10-101

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

Clarke, K. R., and Gorley, R. N. (2015). *Getting Started with PRIMER v7*. Plymouth: PRIMER-E.

Dai, Y., Yang, Y., Wu, Z., Feng, Q., Xie, S., and Liu, Y. (2016). Spatiotemporal variation of planktonic and sediment bacterial assemblages in two plateau freshwater lakes at different trophic status. *Appl Microb Biot.* 100, 4161–4175. doi: 10.1007/s00253-015-7253-2

de Kluijver, A., Yu, J., Houtekamer, M., Middelburg, J. J., and Liu, Z. (2012). Cyanobacteria as a carbon source for zooplankton in eutrophic Lake Taihu, China, measured by 13C labeling and fatty acid biomarkers. *Limnol. Oceanography* 57, 1245–1254. doi: 10.4319/lo.2012.57.4.1245

den Hengst, C. D., and Buttner, M. J. (2008). Redox control in actinobacteria. *Biochim. Biophys. Acta (BBA)-General Subjects* 1780, 1201–1216. doi: 10.1016/j.bbagen.2008.01.008

Diao, M., Sinnige, R., Kalbitz, K., Huisman, J., and Muyzer, G. (2017). Succession of bacterial communities in a seasonally stratified lake with an anoxic and sulfidic hypolimnion. *Front. Microbiol.* 8:2511. doi: 10.3389/fmicb.2017.02511

Díaz-Torres, O., Lugo-Melchor, O. Y., de Anda, J., Pacheco, A., Yebra-Montes, C., Gradilla-Hernández, M. S., et al. (2022). Bacterial dynamics and their influence on the biogeochemical cycles in a subtropical hypereutrophic lake during the rainy season. *Front. Microbiol.* 13:832477. doi: 10.3389/fmicb.2022.832477

Douglas, G. M., Maffei, V. J., Zaneveld, J. R., Yurgel, S. N., Brown, J. R., Taylor, C. M., et al. (2020). PICRUSt2 for prediction of metagenome functions. *Nat. Biotechnol.* 38, 685–688. doi: 10.1038/s41587-020-0548-6

Ehrenfels, B., Bartosiewicz, M., Mbonde, A. S., Baumann, K. B., Dinkel, C., Junker, J., et al. (2021). Diazotrophic cyanobacteria are associated with a low nitrate resupply to surface waters in Lake Tanganyika. *Front. Environ. Sci.* 9:716765. doi: 10.3389/fenvs.2021.716765

Fay, P. (1965). Heterotrophy and nitrogen fixation in *Chlorogloea fritschii*. *J. Gen. Microbiol.* 39, 11–20. doi: 10.1099/00221287-39-1-11

Fernandes, G. L., Shenoy, B. D., and Damare, S. R. (2020). Diversity of bacterial community in the oxygen minimum zones of arabian sea and bay of Bengal as deduced by illumina sequencing. *Front. Microbiol.* 10:3153. doi: 10.3389/fmicb.2019.03153

Fernandez, L., Peura, S., Eiler, A., Linz, A. M., McMahon, K. D., and Bertilsson, S. (2020). Diazotroph genomes and their seasonal dynamics in a stratified humic bog lake. *Front. Microbiol.* 11:1500. doi: 10.3389/fmicb.2020.01500

Ghylin, T. W., Garcia, S. L., Moya, F., Oyserman, B. O., Schwientek, P., Forest, K. T., et al. (2014). Comparative single-cell genomics reveals potential ecological niches for the freshwater acI Actinobacteria lineage. *ISME J.* 8, 2503–2516. doi: 10.1038/ismej.2014.135

Grossart, H. P., Massana, R., McMahon, K. D., and Walsh, D. A. (2020). Linking metagenomics to aquatic microbial ecology and biogeochemical cycles. *Limnol. Oceanography* 65, S2–S20. doi: 10.1002/lno.11382

Hassard, F., Andrews, A., Jones, D. L., Parsons, L., Jones, V., Cox, B. A., et al. (2017). Physicochemical factors influence the abundance and culturability of human enteric pathogens and fecal indicator organisms in estuarine water and sediment. *Front. Microbiol.* 8:1996. doi: 10.3389/fmicb.2017.01996

Hassard, F., Gwyther, C. L., Farkas, K., Andrews, A., Jones, V., Cox, B., et al. (2016). Abundance and distribution of enteric bacteria and viruses in coastal and estuarine sediments – a review. *Front. Microbiol.* 7:1692. doi: 10.3389/fmicb.2016.01692

Hecky, R. E., Mugidde, R., Ramlal, P. S., Talbot, M. R., and Kling, G. W. (2010). Multiple stressors cause rapid ecosystem change in Lake Victoria. *Freshwater Biol.* 55, 19–42. doi: 10.1111/j.1365-2427.2009.02374.x

IBM Corp (2021). *IBM SPSS Statistics for Windows, Version 28.0*. Armonk, NY: IBM Corp.

Kang, I., Kim, S., Islam, M. R., and Cho, J.-C. (2017). The first complete genome sequences of the acI lineage, the most abundant freshwater Actinobacteria,

obtained by whole-genome-amplification of dilution-to-extinction cultures. *Sci. Rep.* 7:42252. doi: 10.1038/srep42252

Khachikyan, A., Milucka, J., Littmann, S., Ahmerkamp, S., Meador, T., Könneke, M., et al. (2019). Direct cell mass measurements expand the role of small microorganisms in nature. *Appl. Environ. Microbiol.* 85:e00493-19. doi: 10.1128/AEM.00493-19

Klawonn, I., Nahar, N., Walve, J., Andersson, B., Olofsson, M., Svedén, J., et al. (2016). Cell-specific nitrogen-and carbon-fixation of cyanobacteria in a temperate marine system (Baltic Sea). *Environ. Microbiol.* 18, 4596–4609. doi: 10.1111/1462-2920.13557

Lewin, G. R., Carlos, C., Chevrette, M. G., Horn, H. A., McDonald, B. R., Stankey, R. J., et al. (2016). Evolution and ecology of Actinobacteria and their bioenergy applications. *Annu. Rev. Microbiol.* 70, 235–254. doi: 10.1146/annurev-micro-102215-095748

Lewis, W. M. (2002). Basis for the protection and management of tropical lakes. *Lakes Reservoirs: Res. Manag.* 5, 35–48. doi: 10.1046/j.1440-1770.2000.00091.x

Liu, K. S., Liu, Y. Q., Jiao, N. Z., Xu, B. Q., Gu, Z. Q., Xing, T. T., et al. (2017). Bacterial community composition and diversity in Kalakuli, an alpine glacial-fed lake in Muztagh Ata of the westernmost Tibetan plateau. *FEMS Microbiol. Ecol.* 93, 1–9. doi: 10.1093/femsec/fix085

Logue, J. B., Langenheder, S., and Andersson, A. F. (2012). Freshwater bacterioplankton richness in oligotrophic lakes depends on nutrient availability rather than on species–area relationships. *ISME J.* 6, 1127–1136. doi: 10.1038/ismej.2011.184

Lücker, S., and Daims, H. (2014). "The family nitrospinaceae," in *The Prokaryotes*, eds E. Rosenberg, E. F. DeLong, S. Lory, E. Stackebrandt, and F. Thompson (Berlin: Springer). doi: 10.1007/978-3-642-39044-9_402

Martineau, C., Mauffrey, F., and Villemur, R. (2015). Comparative analysis of denitrifying activities of Hyphomicrobium nitrativorans, *Hyphomicrobium denitrificans*, and *Hyphomicrobium zavarzinii*. *Appl. Environ. Microbiol.* 81, 5003–5014. doi: 10.1128/AEM.00848-15

Mehrshad, M., Salcher, M. M., Okazaki, Y., Nakano, S. I., Šimek, K., Andrei, A. S., et al. (2018). Hidden in plain sight—highly abundant and diverse planktonic freshwater Chloroflexi. *Microbiome* 6:176. doi: 10.1186/s40168-018-0563-8

Menezes, R. F., Attayde, J. L., Kosten, S., Lacerot, G., Coimbra e Souza, L., Costa, L. S., et al. (2019). Differences in food webs and trophic states of Brazilian tropical humid and semi-arid shallow lakes: implications of climate change. *Hydrobiologia* 829, 95–111. doi: 10.1007/s10750-018-3626-8

Michalak, A. M., Anderson, E. J., Beletsky, D., Boland, S., Bosch, N. S., Bridgeman, T. B., et al. (2013). Record-setting algal bloom in Lake Erie caused by agricultural and meteorological trends consistent with expected future conditions. *Proc. Natl. Acad. Sci. U S A* 110, 6448–6452. doi: 10.1073/pnas.1216006110

Mori, H., Maruyama, F., Kato, H., Toyoda, A., Dozono, A., Ohtsubo, Y., et al. (2014). Design and experimental application of a novel non-degenerate universal primer set that amplifies prokaryotic 16S rRNA genes with a low possibility to amplify eukaryotic rRNA genes. *DNA Res.* 21, 217–227. doi: 10.1093/dnares/dst052

Ngugi, D., Blom, J., Stepanauskas, R., and Sting. (2016). Diversification and niche adaptations of Nitrospina-like bacteria in the polyextreme interfaces of Red Sea brines. *ISME J.* 10, 1383–1399. doi: 10.1038/ismej.2015.214

Oh, S., Zhang, R., Wu, Q. L., and Liu, W. T. (2014). Draft genome sequence of a novel SAR11 clade species abundant in a Tibetan Lake. *Genome Announcements* 2:e01137-14. doi: 10.1128/genomeA.01137-14

Osborne, P. L. (2004). "Eutrophication of shallow tropical lakes," in *The Lakes Handbook: Lake Restoration and Rehabilitation*, eds P. E. O'Sullivan and C. S. Reynolds (Oxford: Blackwell Science Ltd).

Parsons, T. R., Maita, Y., and Lally, C. M. (1984). *A Manual of Chemical and Biological Methods for Seawater Analysis*. Oxford: Pergamon Press.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2012). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* 41, D590–D596. doi: 10.1093/nar/gks1219

Salmaso, N. (2019). Effects of habitat partitioning on the distribution of bacterioplankton in deep lakes. *Front. Microbiol.* 10:2257. doi: 10.3389/fmicb.2019.02257

Schmidt, M. L., White, J. D., and Denef, V. J. (2016). Phylogenetic conservation of freshwater lake habitat preference varies between abundant bacterioplankton phyla. *Environ. Microbiol.* 18, 1212–1226. doi: 10.1111/1462-2920.13143

Sibaja-Cordero, J. A., and Umaña-Villalobos, G. (2008). Invertebrados bentónicos del Lago Cote. *Costa Rica. Rev. Biol. Trop.* 56, 205–213.

Simiyu, B. M., Amukhuma, L., Sitoki, L., Okello, W., and Kurmayer, R. (2022). Interannual variability of water quality conditions in the Nyanza Gulf of Lake Victoria, Kenya. *J. Gt. Lakes Res.* 48, 97–109. doi: 10.1016/j.jglr.2021.10.017

Song, H., Li, Z., Du, B., Wang, G., and Ding, Y. (2012). Bacterial communities in sediments of the shallow Lake Dongping in China. *J. Appl. Microbiol.* 112, 79–89. doi: 10.1111/j.1365-2672.2011.05187.x

Spring, S., Amann, R., Ludwig, W., Schleifer, K. H., van Gemerden, H., and Petersen, N. (1993). Dominating role of an unusual magnetotactic bacterium in the microaerobic zone of a freshwater sediment. *Appl. Environ. Microbiol.* 59, 2397–2403. doi: 10.1128/aem.59.8.2397-2403.1993

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

Tsementzi, D., Rodriguez-R, L. M., Ruiz-Perez, C. A., Meziti, A., Hatt, J. K., and Konstantinidis, K. T. (2019). Ecogenomic characterization of widespread, closely-related SAR11 clades of the freshwater genus "Candidatus Fonsibacter" and proposal of Ca. *Fonsibacter lacus* sp. nov. *Systematic Appl. Microbiol.* 42, 495–505. doi: 10.1016/j.syapm.2019.03.007

Umaña, G. (2014). Ten years of limnological monitoring of a modified natural lake in the tropics: Cote Lake, Costa Rica. *Rev. Biol. Trop.* 62, 576–578.

Umaña, G., Haberyan, K., and Horn, S. (1999). Limnology in Costa Rica. *Limnol. Dev. Countries* 2, 33–62.

Umaña-Villalobos, G., and Aviles-Vargas, L. (2020). Plankton variations in Lake Cote, Costa Rica, from 2002 to 2018. *Hydrobiologia* 847, 4177–4190. doi: 10.1007/s10750-020-04387-8

Wang, Y., Guo, M., Li, X., Liu, G., Hua, Y., Zhao, J., et al. (2022). Shifts in microbial communities in shallow lakes depending on trophic states: feasibility as an evaluation index for eutrophication. *Ecol. Indicators* 136:108691. doi: 10.1016/j.ecolind.2022.108691

Yadav, A. N., Yadav, N., Kour, D., Kumar, A., Yadav, K., Kumar, A., et al. (2019). "Bacterial community composition in lakes," in *Freshwater Microbiology* eds, S. A. Bandh, S. Shafi, and N. Shameem (Cambridge, MA: Academic Press). doi: 10.1016/B978-0-12-817495-1.00001-3

Yoon, S. H., Ha, S. M., Kwon, S., Lim, J., Kim, Y., Seo, H., et al. (2017). Introducing EzBioCloud: a taxonomically united database of 16S rRNA and whole genome assemblies. *Int. J. Systematic Evol. Microbiol.* 67, 1613–1617. doi: 10.1099/ijsem.0.001755

frontiers | Frontiers in Marine Science

TYPE Original Research
PUBLISHED 11 October 2022
DOI 10.3389/fmars.2022.933256

OPEN ACCESS

EDITED BY
Jie Li,
South China Sea Institute of Oceanology (CAS), China

REVIEWED BY
Gang Li,
South China Sea Institute of Oceanology (CAS), China
Jinjun Kan,
Stroud Water Research Center, United States

*CORRESPONDENCE
Wenxue Wu
wwu@hainanu.edu.cn

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

RECEIVED 30 April 2022
ACCEPTED 22 September 2022
PUBLISHED 11 October 2022

CITATION
Wu W and Liu H (2022) Cell size is a key ecological trait associated with biogeographic patterns of microbial eukaryotes in coastal waters. *Front. Mar. Sci.* 9:933256. doi: 10.3389/fmars.2022.933256

Cell size is a key ecological trait associated with biogeographic patterns of microbial eukaryotes in coastal waters

Wenxue Wu[1,2]* and Hongbin Liu[2,3]

[1]State Key Laboratory of Marine Resource Utilization in South China Sea, Hainan University, Haikou, China, [2]Southern Marine Science and Engineering Guangdong Laboratory, Zhuhai, China, [3]Department of Ocean Science, The Hong Kong University of Science and Technology, Kowloon, Hong Kong SAR, China

Body size is an important ecological trait, but it has been poorly explored in microbial communities. Here, we examined the effect of cell size on coastal eukaryotic communities across a size continuum of 0.2–3 (pico-), 3–20 (nano-), and 20–200 μm (micro-sized), which were characterized *via* high-throughput sequencing based on the V4 region of the 18S rRNA gene. We found that, at the alpha diversity level, there was a decreasing trend across the pico-, nano-, and micro-sized eukaryotic communities regarding both amplicon sequence variant (ASV) richness and Shannon index. At the beta diversity level, the three categories were significantly different, and these were accompanied by a relatively high local contribution to beta diversity in contrasting freshwater and seawater locations. The community variations observed for the microbial eukaryotes could largely be explained by the environmental effects which decreased between the pico- (40.5%), nano- (37.3%), and micro-sized (27.3%) fractions. These environmental effects were mainly contributed by several ASV modules showing opposing responses to environmental conditions. This might partly indicate the coalescence of the freshwater and seawater groups of microbial eukaryotes. In summary, our findings suggest that the cell size of microbial eukaryotes is a phylogenetically conserved trait, which is tightly associated with biogeographic patterns.

KEYWORDS

marine microbial eukaryote, 18S rRNA gene, size continuum, environmental effect, Pearl River Estuary

Introduction

Body size is perhaps the most fundamental ecological trait as it is associated with many biological properties. For all marine life, from bacteria to whales, size has several profound effects, for example, on the body temperature, resource encounter strategy (e.g., photosynthesis), mobility, sensing mode (e.g., echolocation), and life history strategy

(Andersen et al., 2016). A lot of statistical characteristics, such as biodiversity and abundance across marine taxa, have been basically scaled with body size (Reuman et al., 2014) which represents the most commonly measured trait in ecological studies. As a consequence, size-fractionated categories at the population or community levels often exhibit distinct distributions across space. For example, large cells usually dominate phytoplankton assemblages in estuaries in terms of biomass and production, whereas small cells dominate the open oceans (Cloern, 2018). This is because cell size represents a determinant of phytoplankton metabolism (Marañón, 2015), which is strongly affected by contrasting temperatures (Mousing et al., 2014; López-Urrutia and Morán, 2015) and the supply of resources (Finkel et al., 2010; Marañón et al., 2015) across the marine realms. The phytoplankton size structure principally consists of several categories, which can shift both spatially and temporally in the global oceans (Acevedo-Trejos et al., 2013). Body size is known to be more than simply a trait that controls the physiological characteristics, and it is also relevant for community-level phenomena linking environmental conditions and compositional variations. However, the effect of differences in size is still poorly understood in marine microbial communities, partly due to the complexity of characterizing the various microorganisms.

In recent years, marine microbial eukaryotes (mainly unicellular eukaryotes) (Caron et al., 2009), have been intensively assessed *via* high-throughput sequencing, and this has generated important information about their taxonomic diversity (Adl et al., 2019) coupled with their diverse ecological functions. For example, research indicates that Syndiniales (Dinoflagellata, Alveolata) are mainly intracellular parasites or symbionts (Guillou et al., 2008), and Dinophyceae (Dinoflagellata, Alveolata) are of a mixotrophic nature, feeding on other microbial eukaryotes such as Bacillariophyta (Ochrophyta, Stramenopiles; also known as diatoms) (Grattepanche et al., 2011). Bacillariophyta are major primary producers (Malviya et al., 2016), while MAST (MArine STramenopiles) are bacterivorous heterotrophs (Massana et al., 2014). Mamiellophyceae, Chlorophyceae, and Trebouxiophyceae (Chlorophyta, Archaeplastida) are representative autotrophic microalgae in coastal waters (Tragin and Vaulot, 2018), and Basidiomycota (Fungi, Opisthokonta) are described as being an amoebophagous group (Corsaro et al., 2018; Grossart et al., 2019). These complex trophic roles of microbial eukaryotes support the suggestion that they represent pivotal players in marine biogeochemical processes (Worden et al., 2015).

Marine microbial eukaryotes exhibit a wide spectrum of cell sizes, which span more than five orders of magnitude (Caron et al., 2012), and they show distinct biogeographic patterns among the different size-based categories (Sommeria-Klein et al., 2021). At the alpha diversity level, the smaller-sized fractions in general have higher estimates of diversity indices. For example, the smallest size fraction (i.e., pico-sized) often shows the highest richness (de Vargas et al., 2015). However, it is still unclear whether there is a size-dependent trend in alpha diversity across a fine size scale. At the beta diversity level, due to the vast differences in richness and taxonomy, size-fractionated microbial eukaryotes are (as expected) distinct from each other (Massana et al., 2015; Xu et al., 2020). Microbial eukaryotes of different sizes are supposed to have different traits (at least regarding the cell size), efficiently indicating niche differentiations, which can be further realized by environmental adaptions (Caron et al., 2012). Their community variations (beta diversity) are anticipated to be driven by environmental effects by different amounts (Clarke and Deagle, 2020). Therefore, size-based comparisons offer a more comprehensive understanding of marine microbial eukaryotes than considering them as a biological pool.

In this study, we focused on size-fractionated microbial eukaryotes in the Pear River Estuary, located in the northern South China Sea. This estuary connects rivers, tidal wetlands, and continental shelves comprising a coastal ocean with unique and changing biogeochemical cycles (Bauer et al., 2013). In particular, the Pear River Estuary harbors typical coastal waters, which exhibit large environmental gradients within a small geographical area. For example, salinity is one of the strongest barriers for organisms within the eukaryotic tree of life to cross (Jamy et al., 2022). This likely explains the big changes in biogeographic patterns of microbial eukaryotes that occur in the coastal waters. The data we acquired in this new estuarine region survey, when combined with those from previous studies conducted in the upper Pearl River Delta (Zou et al., 2021), Pearl River-South China Sea Continuum (Wu and Liu, 2018), and offshore South China Sea (Wang et al., 2021), will together provide a more complete picture of the microbial eukaryotic communities in this region. Moreover, size-fractionated microbial eukaryotes have previously been investigated in some isolated coastal locations (Massana et al., 2015; Genitsaris et al., 2016; Elferink et al., 2020); however, they have not been well documented *via* a fine sampling design over an entire estuary. We collected microbial eukaryotic communities during the summer season across a size continuum of 0.2–3 μm (pico-sized eukaryotes, PE), 3–20 μm (nano-sized eukaryotes, NE), and 20–200 μm (micro-sized eukaryotes, ME). We characterized these three categories using high-throughput sequencing of the V4 region of the 18S rRNA gene, and as mentioned above, we tested whether there was a decreasing trend of alpha diversity at this fine scale (i.e., from PE, NE to ME) and whether the strength of the environmental effects driving beta diversity was strong and varied across the size continuum. We conclude that this study contributes to the long-standing task of understanding the biodiversity of marine eukaryotes from a trait-based perspective.

Materials and methods

Sample collection and environmental factor

Sampling was carried out on 10–21 July 2017 in a summer cruise supported by the OCEAN-HK project. Field seawater samples were collected from the surface (i.e., at a depth of 1 m) and bottom (i.e., depths of 6–48 m at 1–6.1 m above the sediment) layers at 19 stations (Figure 1), using Niskin bottles mounted on an SBE 32 carousel water sampler (Sea-Bird Electronics, Bellevue, WA, USA). For the molecular analyses, seawater (350–1000 ml) was immediately pre-filtered through a 200-µm mesh and then sequentially through 20 µm, 3 µm, and 0.2 µm pore-sized polycarbonate membranes (47 mm diameter, Millipore, Carrigtwohill, Cork, Ireland). Cells retained on the 20 µm, 3 µm, and 0.2 µm filters represented the collections of ME, NE, and PE, respectively. The filters were kept at –20°C onboard and then frozen at –80°C in the laboratory until required for analysis.

Temperature, salinity, and turbidity were measured at all 19 sites using a conductivity-temperature-depth (CTD) profiler. Chlorophyll *a* was extracted in 90% acetone and quantified fluorometrically with a Turner Designs fluorometer (Turner Designs, Sunnyvale, CA, USA). The concentration of dissolved oxygen was determined by the Winkler titration method, and nutrients (including NO_2, NO_3, PO_4 and SiO_3), dissolved organic carbon, and total suspended matter were all measured using standard protocols, as described in previous studies (Dai et al., 2009; Han et al., 2012). More details of the environmental factors measured are also as described previously (Wu et al., 2021).

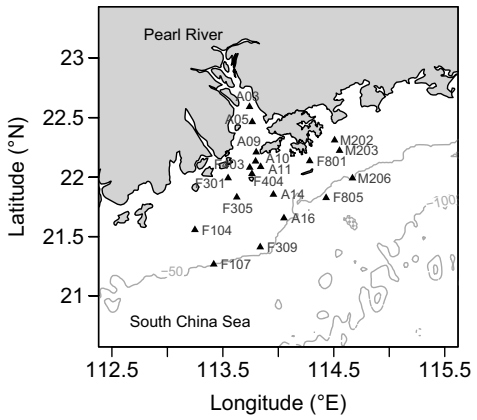


FIGURE 1
Map showing the sampling sites (indicated by filled triangles) in the Pearl River Estuary located in the northern South China Sea.

DNA extraction and amplification

Each filter was cut into small pieces using sterile scissors and then each piece was transferred to a 2 ml tube provided in the FastDNA Spin Kit (MP Biomedicals, Solon, OH, USA). The pieces of filter were vortexed in a mixture of 978 µl sodium phosphate buffer and 122 µl MT buffer, at a speed of 3,500 oscl min^{-1} using a Mini-Beadbeater-24 (Biospec Products, Bartlesville, OK, USA). DNA was then extracted from the filter pieces following the FastDNA Spin Kit manufacturer's protocol.

The hypervariable V4 region of the 18S rRNA gene was amplified using the following primer pair: TAReuk454FWD1 (5′-CCAGCA(G/C)C(C/T)GCGGTAATTCC-3′) and TAReukREV3 (5′-ACTTTCGTTCTTGAT(C/T)(A/G)A-3′), with a dual-index strategy (Stoeck et al., 2010). Each polymerase chain reaction (PCR) mixture (50 µl) contained 1× PCR buffer, 1.5 mM $MgCl_2$, 0.2 mM dNTP mix, 0.5 µM each primer, 2 U Invitrogen Platinum *Taq* DNA polymerase (Life Technologies, Carlsbad, CA, USA), and 2.5 µl of template DNA. In general, the reaction was run under the following conditions: 95°C for 5 min, and then 35 cycles at 94°C for 30 s, 47°C for 45 s, and 72°C for 1 min, with a final extension step at 72°C for 5 min. However, since many of the ME samples failed to be amplified with 35 cycles, 40 cycles were used for the PCR amplification of them. We assessed the data from 14 samples that were successfully amplified by both 35 and 40 cycles (Supplementary Table 1). We showed that there were no significant differences ($p > 0.05$) at the alpha diversity level by the nonparametric Kruskal-Wallis *H* test (Kruskal and Wallis, 1952) (Supplementary Figure 1), or at the beta diversity level by analysis of similarity (ANOSIM) when comparing these 35- and 40-cycle groups. In addition, four ME samples were not successfully amplified even when using the modified 40-cycle protocol, and so these were replaced by filters immersed in RNA*later*, which had been simultaneously collected in the same locations and water depths (Supplementary Table 1). All the products were sequenced on a HiSeq 2500 platform (Illumina, San Diego, CA, USA) with a (2 × 250 bp) paired-end strategy.

Sequence processing

Sequence data were de-multiplexed using the QIIME 2 (v. 2019.1.0) bioinformatics platform (Bolyen et al., 2019) and trimmed using cutadapt (Martin, 2011). Sequences were then quality filtered and dereplicated with DADA2 (Callahan et al., 2016) implemented in the QIIME 2 pipeline. Amplicon sequence variants (ASVs) were identified, and their taxonomies were assigned using the BLAST+ consensus classifier (Camacho et al., 2009; Bokulich et al., 2018) against the PR^2 database (v. 4.12) (Guillou et al., 2013). Singletons, metazoa, and sequences unsuccessfully assigned at least at rank 2 in PR^2 (e.g., Alveolata,

Stramenopiles, Archaeplastida and Opisthokonta) were removed from the subsequent analyses. The ASV representatives were aligned using MAFFT (Katoh and Standley, 2013) and a phylogenetic tree was built using FastTree (Price et al., 2010).

Alpha diversity

The numbers of ASVs found across the PE, NE, and ME were illustrated with a Venn diagram, based on the full ASV-sample table. To compare alpha diversity among the three categories, the ASVs in each sample were counted using a rarified ASV-sample table (4,982 sequences per sample). The Shannon index (Shannon, 1948) was also calculated for each sample using this rarified table. For both the observed ASV numbers and Shannon indices, the nonparametric Kruskal-Wallis *H* test was used to examine differences among the PE, NE, and ME.

Fraction-inferred size analysis

To differentiate between the ASV proportions in the PE, NE, and ME, we defined a fraction-inferred size (FIS) index. For the FIS calculation, the full (non-rarified) ASV table was proportionally normalized and rare ASVs (average proportion <0.001% in all 114 communities) were not included. For a given ASV, we recorded its average proportions in the PE, NE, and ME, for which we gave a size weight of 1, 2, and 3, respectively. The FIS was calculated as the sum of the fraction classification-weighted proportions. The observed FIS values ranged between 1 and 3, indicating a fraction-inferred size from smaller to larger. The phylogenetic placements of the FIS index were combined through a FastTree using the ggtree package (Yu et al., 2017) in R (R Core Team, 2018). The phylogenetic signal was further analyzed using Pagel's λ index (Pagel, 1999) with the phytools package (method = 'lambda') (Revell, 2012). The phylogenetic signal measure (λ) ranged between 0 and 1, indicating the extent by which a trait's correlations among closed related ASVs matched the motion model of evolution. The *p*-value for λ was based on a likelihood ratio test. Given a significant phylogenetic signal, the FIS index could be assumed to indicate a set of synthesized traits associated with size (Salazar et al., 2015; Mestre et al., 2018).

Beta diversity

For beta diversity among the PE, NE, and ME, we used the unweighted and weighted UniFrac index (Lozupone and Knight, 2005). The UniFrac dissimilarities were then analyzed by principal coordinate analysis (PCoA), and ANOSIM was used to test the significance of community dissimilarities among the three categories (999 permutations).

To identify where community dissimilarities occurred intensively, we calculated the local contribution to beta diversity (LCBD) index (Legendre and De Cáceres, 2013), using the adespatial package (Dray et al., 2022). We performed a principal component analysis (PCA) using standardized environmental factors, and the first axis (PC 1) was used as a proxy of the overall environments. Salinity, as the most representative variable (Wu et al., 2021), was positively and significantly correlated with the PC 1 measurements (Spearman's rank correlation coefficient $r = 0.98$, $p < 0.05$). The chlorophyll *a* concentration was not included when performing the PCA, because this (biotic) factor did not belong to typical physical and chemical parameters. The relationship between the LCBD values and PC 1 was tested using generalized additive models implemented in the mgcv package (Wood, 2011).

Variation partitioning

To examine pure environmental effects apart from spatial autocorrelation, we performed variation partitioning using the vegan package (Oksanen et al., 2018). It is well known that proportional (abundance-weighted) metrics mirror more environmental condition information than unweighted ones, and so the weighted UniFrac distance was used to account for community variations in this analysis. The environmental variables (excluding chlorophyll *a*) were standardized, and spatial distances were transformed to principal coordinates of neighbor matrices (PCNM) using the *pcnm* function. The relative importance of the environmental and spatial factors was determined using the *varpart* function, and significance was tested using distance-based redundancy analyses (999 permutations).

In addition, the permutational multivariate analysis of variance (PERMANOVA) was carried out using the *adonis* function (999 permutations) implemented in the vegan package to further identify the relationship between a particular environmental variable and community variations. Consistent with the variance partitioning analysis, the weighted UniFrac distances and standardized environmental variables were used.

Co-occurrence network analysis

To identify clusters contributing to community variations, we constructed co-occurrence networks using the igraph package (Csárdi and Nepusz, 2006). Prior to network construction, ASVs present in at least five sites were included to ensure the regional representativeness of the resulting co-occurrence relationships. In brief, robust co-occurrence relationships of pairwise ASVs ($p < 0.01$, $r > 0.8$, Spearman's

rank correlation coefficient) were retained to build the networks. In the resulting networks, each node represented an ASV, each edge represented the significantly positive or negative coexistence of two nodes, and the number of connections for each node represented its degree. Modules (i.e., clusters) in networks were inferred using a fast-greedy algorithm (Clauset et al., 2004). The co-occurrence networks were visualized using the interactive Gephi platform (Bastian et al., 2009).

The environmental effects within network modules were further examined. Nodes in the same module were expected to show similar or dissimilar responses to environmental conditions given significantly positive or negative co-occurrence relationships, respectively. Although it is generally accepted that the co-occurrence network can partly reveal biotic interactions among nodes (Fuhrman et al., 2015), our examinations primarily focused on the co-occurrence patterns themselves rather than biotic interactions due to several cautionary challenges (Röttjers and Faust, 2018). We then paid more attention to modules showing >10% contributions to the total degrees. We calculated the Spearman's rank correlations between the PC 1 (inferred from PCA) and regional patterns of each ASV in a module. It was anticipated that significant ($p < 0.05$) and non-significant ($p > 0.05$) correlations would be detected in ASVs linked to environmental factors either directly (e.g., autotrophs) or indirectly (e.g., parasitic groups), respectively. Consequently, the former would uncover a series of ASVs responsible for environmental effects inferred from the variation partitioning analyses above.

Results

Library characterization and alpha diversity

A total of 5,816,445 sequences (4,982–79,400 per sample) were retrieved from 114 microbial eukaryotic communities. Regarding the relative abundances, Alveolate, Stramenopiles, Archaeplastida, and Opisthokonta represented the four most abundant supergroups with an average abundance of 36.9%, 26.1%, 22.5%, and 9.7%, respectively. The PE, NE, and ME categories showed opposite trends of regional relative abundances (mean ± SD) of Alveolate (i.e., PE, 45.8 ± 20.2%; NE, 35.3 ± 17.7%; and ME, 29.5 ± 15.3%) and Stramenopiles (i.e., PE, 9.3 ± 4.8%; NE, 25.6 ± 14.1%; and ME, 43.5 ± 23.4%) (Figure 2A). With regard to the average relative abundances roughly at the class level (i.e., rank 4 in the PR^2 database), Syndiniales, Dinophyceae, Bacillariophyta, Chlorophyceae, Trebouxiophyceae, and Basidiomycota were the six groups with an average relative abundance >5% in the 114 communities (Supplementary Table 2).

The total sequences grouped into 7,011 ASVs, and the PE, NE, and ME accounted for decreasing unique proportions of 30.1%, 18.7%, and 10.8%, respectively (Figure 2B). Based on the ASV table rarified to 4,982 sequences per sample, we found that the ASV richness (i.e., the number of observed ASVs) significantly decreased from the PE-to-NE-to-ME ($p < 0.05$) (Figure 2C). However, there were no significant differences in the Shannon index between the PE and NE ($p > 0.05$) (Figure 2D), even though the PE had a higher mean than the NE.

FIS

The phylogenetic signal was strong ($\lambda = 0.85$) and significant ($p < 0.001$), indicating that cell size was a phylogenetically constrained trait of microbial eukaryotes. For this reason, ASVs with close phylogenetic placements showed approximate FIS calculations and a similar presence in the PE, NE, and ME (Figure 3A). The distribution of the FIS index displayed a high frequency density of 1, 2, and 3 (Figure 3B), which indicated the presence of unique ASVs in the PE, NE, and ME, respectively. At the rough phylum or class level, some clusters showed consistently distinct FIS values (Figure 3C). For example, Perkinsea mainly consisted of taxa of low FIS values, whereas Streptophyta generally included ASVs of high FIS.

Beta diversity and relative importance of environmental effects

The PE, NE, and ME were significantly different at the beta diversity level, according to the ANOSIM test ($p < 0.001$). However, these three categories were not separate from each other in the PCoA plots based on either unweighted or weighted UniFrac dissimilarities (Figure 4).

Moreover, the LCBD values in general showed a U-shaped curve against PC 1 which indicated an environmental transition from low- to high-salinity (Figure 5). Therefore, the two ends of the freshwater- and seawater-like habitats were the main contributors to the community variations in the Pear River Estuary.

Significant environmental effects on beta diversity were also detected. Based on weighted UniFrac dissimilarities, a purely relative importance of 40.5%, 37.3%, and 27.3% was observed for the PE, NE, and ME, respectively (Figure 6). In parallel, the spatial effects were much lower, and even non-significant in the ME fraction.

Co-occurrence pattern

Microbial eukaryotes mostly exhibited positive relationships in co-occurrence patterns (Figures 7A–C), and there were three, three, and one modules showing >10% contributions to total degrees in the PE, NE, and ME networks, respectively. The ASVs

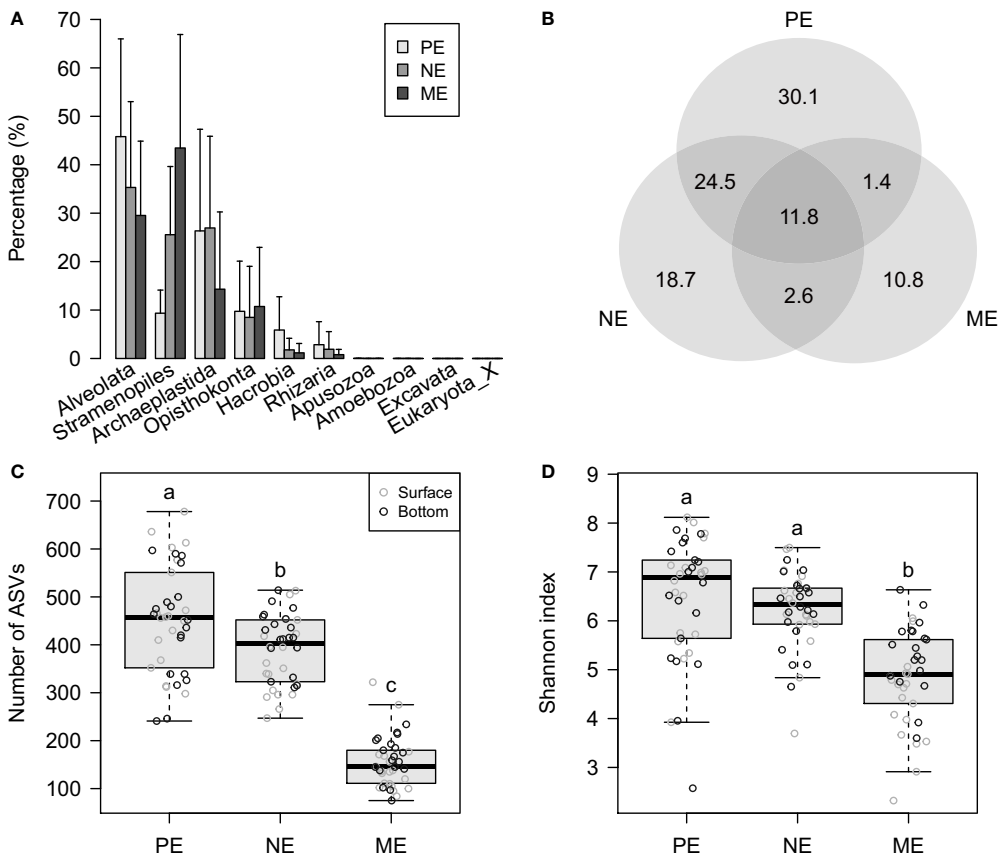


FIGURE 2
(A) Bar graph showing the percentage mean (and standard deviation) relative abundance of the various supergroups, based on the full ASV-sample table. **(B)** Venn diagram showing the percentage of ASVs that are unique in, and shared among, the PE, NE, and ME categories. **(C, D)** Boxplots showing the **(C)** observed ASV numbers and **(D)** Shannon indices using the rarified ASV-sample table. In **(C)** and **(D)**, the letters above each boxplot indicate significant differences at $p < 0.05$.

belonging to each module showed variable taxonomic compositions (Figures 7D–F) and relationships related to environmental conditions (Figures 7G–I). Remarkably, the PE and NE each displayed one module composed of ASVs showing a significant linear regression relationship between degrees and correlation coefficients.

Discussion

This study shows that microbial eukaryotes across the size continuum were significantly different in multiple aspects. Overall, our results support the two expectations that (1) the alpha diversity of microbial eukaryotes in general, decreased as the cell size increased, and (2) the strength of the environmental effects on the beta diversity was strong and varied across the size continuum. These results support the historical perspective that organism size is a useful trait for understanding biogeographic patterns, even for microbial communities at a fine size scale.

Size dependence of alpha diversity

Microbial eukaryotes have an expected trend in size-fractionated alpha diversity. This pattern is well supported by the overall ASV numbers of each fraction (Figure 2B) and the ASV richness estimations of the rarified communities (Figure 2C), which is consistent with previous global surveys (de Vargas et al., 2015). Meanwhile, differences between the PE and NE were not significant when using a proportional measure (i.e., the Shannon index) (Figure 2D). Proportional measures of alpha diversity often carry information on environmental adaptions of populations in a given community (i.e., evenness), which suggests that this size-dependent pattern of

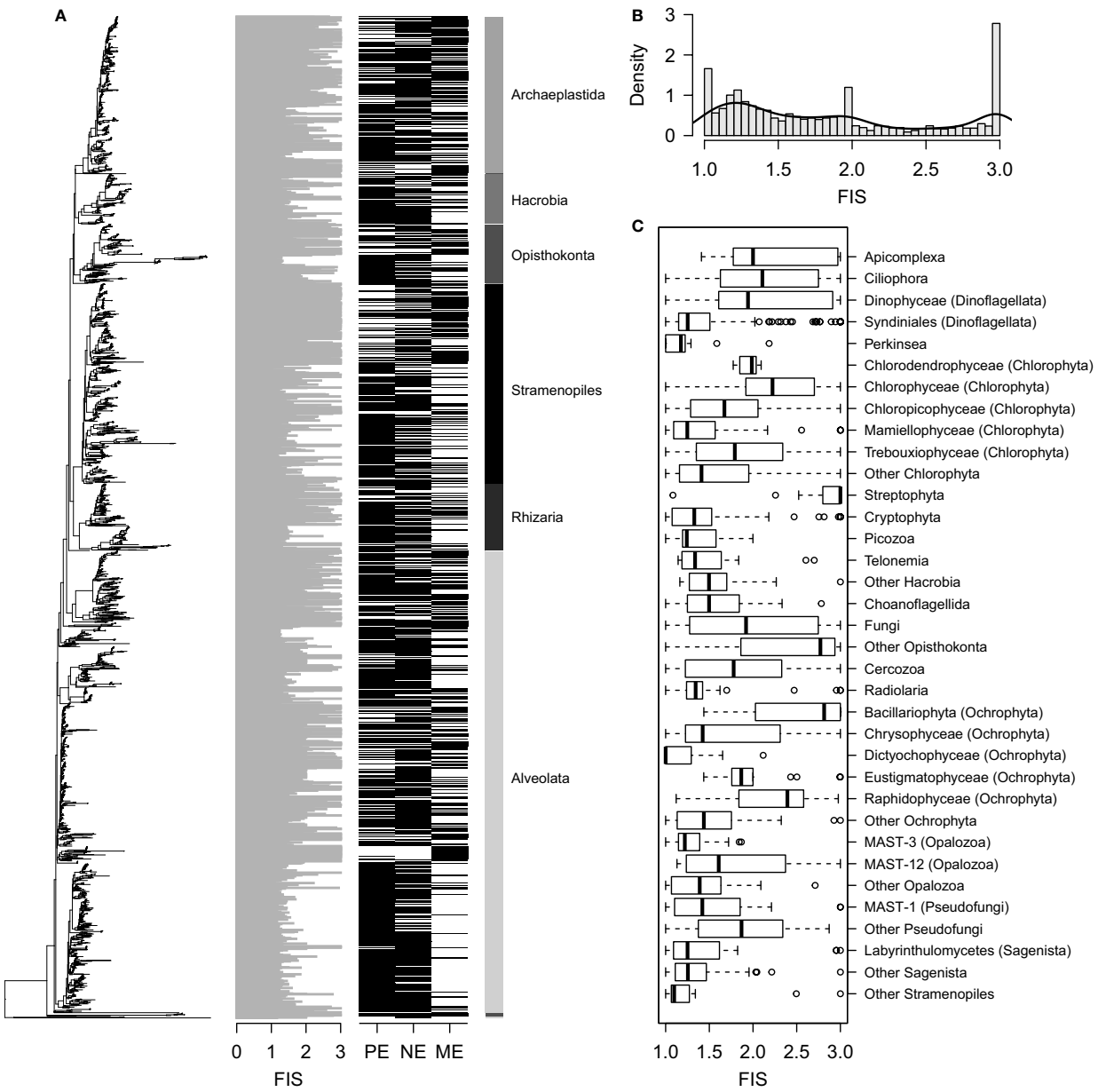


FIGURE 3
(A) Phylogenetic placement of the fraction-inferred size (FIS) index combined with a heatmap showing the ASV occurrence (indicated by the black bars) in each category. Taxonomic assignments at the supergroup level are added on the right side of the heatmap. **(B)** Histogram showing the distribution of the FIS index. The line indicates kernel density estimates for the FIS frequency. **(C)** Boxplots indicating the FIS index for the main lineages (generally collapsed at ranks 3 and 4 of the PR^2 database). One Dinoflagellata ASV (FIS = 1.3), which could not be further assigned to a lower taxonomic level (e.g., Syndiniales and Dinophyceae), and the Amoebozoa and Apusozoa supergroups, consisting of just one (FIS = 1.2) and two (FIS = 1.2 and 1.6) ASVs, respectively, were not included in the boxplots.

alpha diversity can be shaped by environmental conditions. We note that the results from this study were inconsistent with those from some previous reports. For example, Logares et al. (2014) found that the nano-sized fraction had higher richness than the pico- and micro-meso-sized fractions. This discrepancy might be partly due to the difference in the targeted regions of the 18S rRNA gene (i.e., V4 *versus* V9), which resulted in variable diversity estimations (Dunthorn et al., 2012; Decelle et al., 2014; Tragin et al., 2018). Alternatively, the size-dependent pattern relies on the spatial scale of the sampling design,

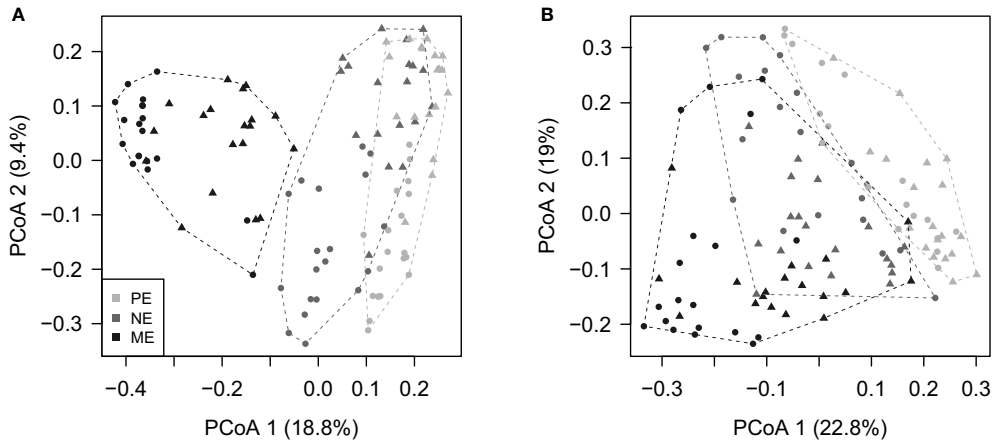


FIGURE 4
Principal coordinate analysis (PCoA) plots based on **(A)** unweighted and **(B)** weighted UniFrac distances showing the beta diversity patterns of the PE, NE, and ME categories. The dashed lines surrounding the symbols (i.e., the filled circles indicating surface samples, and filled triangles indicating bottom samples) delineate the convex hull for each category and are drawn in the same color.

which varies from isolated sites (Logares et al., 2014), an entire estuary (i.e., this study) to global oceans (de Vargas et al., 2015). More importantly, the pattern we observed is consistent with other previous investigations, which mostly address the marine diversity spectrum of large organisms (Reuman et al., 2014). Differences in diversity are mainly due to resource partitioning, through niche differentiation. Indeed, natural environments are expected to support more organisms of smaller sizes than of larger sizes, generating a decreasing trend in alpha diversity as the size of the organisms increase. Therefore, the diversity spectrum can perhaps be extended to microbial eukaryotes based on our observations.

The cell size of microbial eukaryotes is suggested to be associated with a set of organismal characteristics related to niche and fitness, and this suggestion was supported by our finding that cell size is a phylogenetically conserved trait (Figure 3A). It is well known that phylogeny is tightly connected to functional traits in microorganisms (Martiny et al., 2013; Goberna and Verdú, 2016). Although niches cannot be simply inferred through cell size alone, differences in size can predict variations in the niche and fitness of different microbial eukaryotes (Gallego et al., 2019). For example, Syndiniales are one of the most highly represented groups in the world oceans (de Vargas et al., 2015), and in general they exhibit small FIS values (Figure 3C), mainly due to their parasitic and symbiotic life strategies. In contrast, in the Pearl River Estuary, the Bacillariophyta class is known to be a highly represented group (Wu and Liu, 2018), and this group exhibited high FIS values. These findings agree with previous reports that the larger-sized Bacillariophyta can dominate the estuarine phytoplankton (Cloern, 2018), despite their variable cell sizes across the world oceans (Malviya et al., 2016). Together with the dominance of Syndiniales and Bacillariophyta (Supplementary Table 2), our data suggest that the co-existence of most of the microbial eukaryotes might be putatively inferred from both their size and related characteristics, which contributes to the size-dependent diversity pattern.

There are some limitations in our assessments of alpha diversity when relying on pore-sized filters and 18S rDNA gene sequencing. For example, we cannot rule out some biases caused by body fragments from larger-sized organisms (Liu et al., 2017). In addition, extracellular DNA putatively adhering to small particles might also be a source of some larger-sized taxa in the smaller-sized fractions (Sørensen et al., 2013). In contrast, grazing might contribute to the existence of smaller taxa in the larger-sized categories. Indeed, some herbivorous eukaryotes can exert grazing pressures on phytoplankton independent of cell size (Cabrerizo and Marañón, 2021), which leads to biases in the size classifications. Despite these limitations, 59.6% of the total ASVs were still restricted to a single category (Figure 2B), suggesting that the protocols we used to separate the whole microbial eukaryotes into the PE, NE, and ME were effective (Logares et al., 2014). Moreover, our size dependence of alpha diversity was based on the DNA signatures found in complex estuarine waters. This pattern might be affected by including dormant taxa and mixing organisms from freshwater, seawater, and sediments in the estuarine area. Employing RNA signatures (only including active taxa) and/or surveying other waters might result in a changed pattern with an absence of the size dependence of alpha diversity found in this study (Xu et al., 2020).

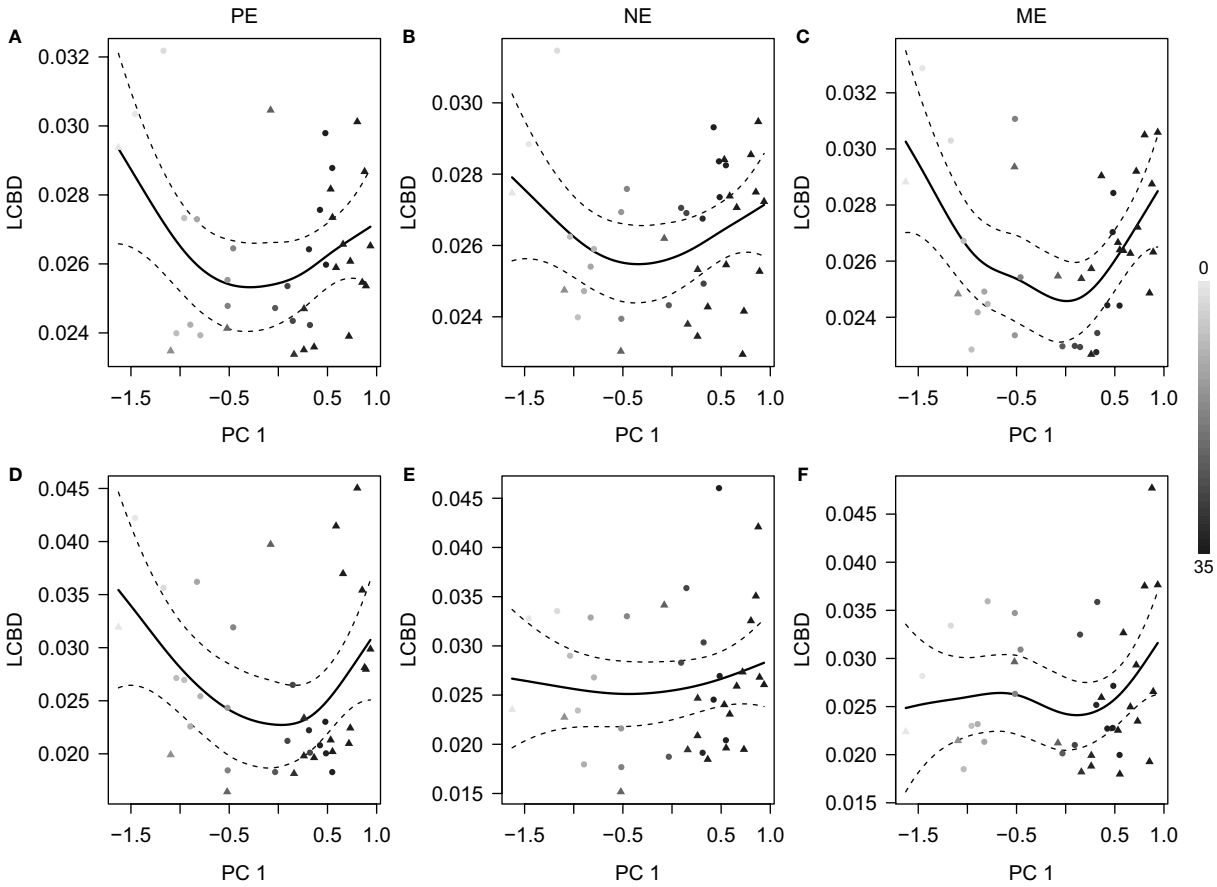


FIGURE 5
Relationship between the local contribution to beta diversity (LCBD) estimates and the first axis of the principal component analysis of environmental factors (PC 1). Panels **(A–C)** and **(D–F)** are results based on the unweighted and weighted UniFrac distances, respectively, for the **(A, D)** PE, **(B, E)** NE, and **(C, F)** ME categories. The fitted smoothers (solid lines with 95% confidence intervals) were predicted using generalized additive models. The shades of symbols (filled circles, surface samples; filled triangles, bottom samples) indicate the local salinity at each station.

Environmental effect on beta diversity

The community variations of microbial eukaryotes are largely driven by environmental effects, which decreased from the smaller- to larger-sized categories (Figure 6). The main possibility for the difference in environmental effects might be that smaller-sized microbial eukaryotes are more responsive to environmental changes. For example, a recent study conducted in the coastal waters off the Ría de Vigo in northwestern Spain showed that pico- and nano-sized eukaryotes (particularly of pigmented groups) responded profoundly to the addition of inorganic and organic nutrients (Hernández-Ruiz et al., 2020). From the size spectrum perspective, smaller microbial eukaryotes are mainly found in the lower trophic levels, and so they might respond to hydrographic conditions more efficiently. Moreover, it is supposed that as well as higher richness, smaller microbial eukaryotes also have higher cell densities (White et al., 2007). The smaller categories with more diverse and abundant assemblages might also have higher potentials to track environmental changes, which contributes to higher environmental effects.

The co-occurrence patterns further differentiate the responses of the size-fractionated microbial eukaryotes to the environmental contexts (Figures 7A–C). Indeed, many ASVs showed significant correlations with environments, suggesting that these are the main contributors of the environmental effects on the community variations observed (Figures 7G–I). However, some modules were of remarkably different taxonomic compositions. For example, modules composed mainly of ASVs with non-significant environmental correlations, had large contributions of Alveolata (Figures 7D–F), indicating that these modules might be controlled mainly by biotic interactions by parasitic or symbiotic taxa. Interestingly, the ASVs in M1 of the PE and M2 of the NE showed linear

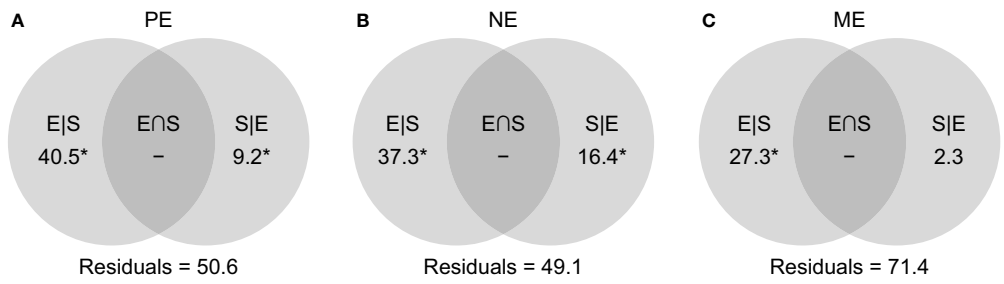


FIGURE 6
Venn diagrams showing the proportions of variations (in %) explained by shared (E∩S), as well as the unique effects of environmental (E|S) and spatial factors (S|E) for the **(A)** PE, **(B)** NE, and **(C)** ME categories. Negative values are indicated by '–', and significant effects ($p < 0.05$) are marked with an asterisk.

regression relationships between the degrees and coefficients (Figures 7G, H). Indeed, the three ASVs with the largest degrees (two ASVs for PE and one ASV for NE) are closely affiliated to freshwater Chlorophyta, supported by 100% sequence similarities with hits in a BLAST (NCBI) search. The negative coefficients mostly resulted from intensive filtering towards offshore environments. This phenomenon supports our cautiousness to not simply interpret the co-occurrence networks as biotic interactions (Blanchet et al., 2020), which can suffer from rapidly changing environments. We suggest that network analyses are informative for understanding environmental effects based on module-level characteristics.

In addition to the environmental effects, we acknowledge that spatial effects are also important in shaping the beta diversity of microbial eukaryotes. The spatial effects are particularly crucial for large geographical scales when ocean connectivity (rather than environmental differences) might be the key player (Villarino et al., 2018). Moreover, a considerable number of components could not be attributed to either environmental or spatial effects (Figure 6). These might result from unmeasured environmental variables or biotic factors such as *via* predator–prey interactions. For example, the microbial eukaryotes encompass numerous bacterivores, including mixotrophic phytoplankton (Unrein et al., 2014). These bacterivores are themselves controlled by grazing pressures commonly observed in the PE, NE, and ME (Wikner and Hagstrom, 1988; Tamigneaux et al., 1997; Caron et al., 1999). Notably, biotic interactions (e.g., grazing) can also be regulated by environmental factors (e.g., temperature) (Cabrerizo and Marañón, 2021), which suggests that there are mixed effects between biotic interactions and environmental conditions. Taking these ideas together, we suggest that environmental, spatial, and biotic factors exhibit combined and complicated effects on the beta diversity of the PE, NE, and ME.

It should be noted that the environmental effects we observed represent a single temporal scenario of the community variations across the Pearl River Estuary. A characteristic feature of the Pearl River Estuary is the occurrence of seasonal discharges of freshwater into the sea; this reaches a peak in the summer, which is when we collected our samples. As shown by the shapes of the LCBD graphs (Figure 5), the turnover of the inshore and offshore communities is considerable. This represents the outcome of the environmental gradients caused mainly by the Pearl River runoff, in comparison to other coastal waters that lack large freshwater discharges (Genitsaris et al., 2016). We suggest that more temporal investigations in this region are required to achieve a more complete view of the microbial eukaryotes. Moreover, we suggest that our results represent a view that is restricted to estuarine ecosystems. In particular, the large environmental gradients might weaken the between-fraction differences, as indicated by the overlap observed in the PCoA plots (Figure 4). A similar situation occurred in open oceans (Duret et al., 2015), which indicates that size-fractionated microbial eukaryotic communities can be strongly rearranged by water depth. More efforts are also required to survey diverse habitats in addition to estuaries.

Conclusions

This study provides new insights into the microbial eukaryotes across a size continuum of pico-, nano-, and micro-sized fractions. We showed that Alveolate, Stramenopiles, Archaeplastida, and Opisthokonta represent the four most abundant supergroups. In addition, we found that there was a decreasing trend of alpha diversity with increasing cell size. The phylogenetic signal was strong and significant for the fraction-inferred sizes, indicating that cell size was a phylogenetically conserved trait of microbial eukaryotes. Moreover, the three categories were significantly different at the beta diversity level, although they were not distinct in ordination analyses. In the Pear

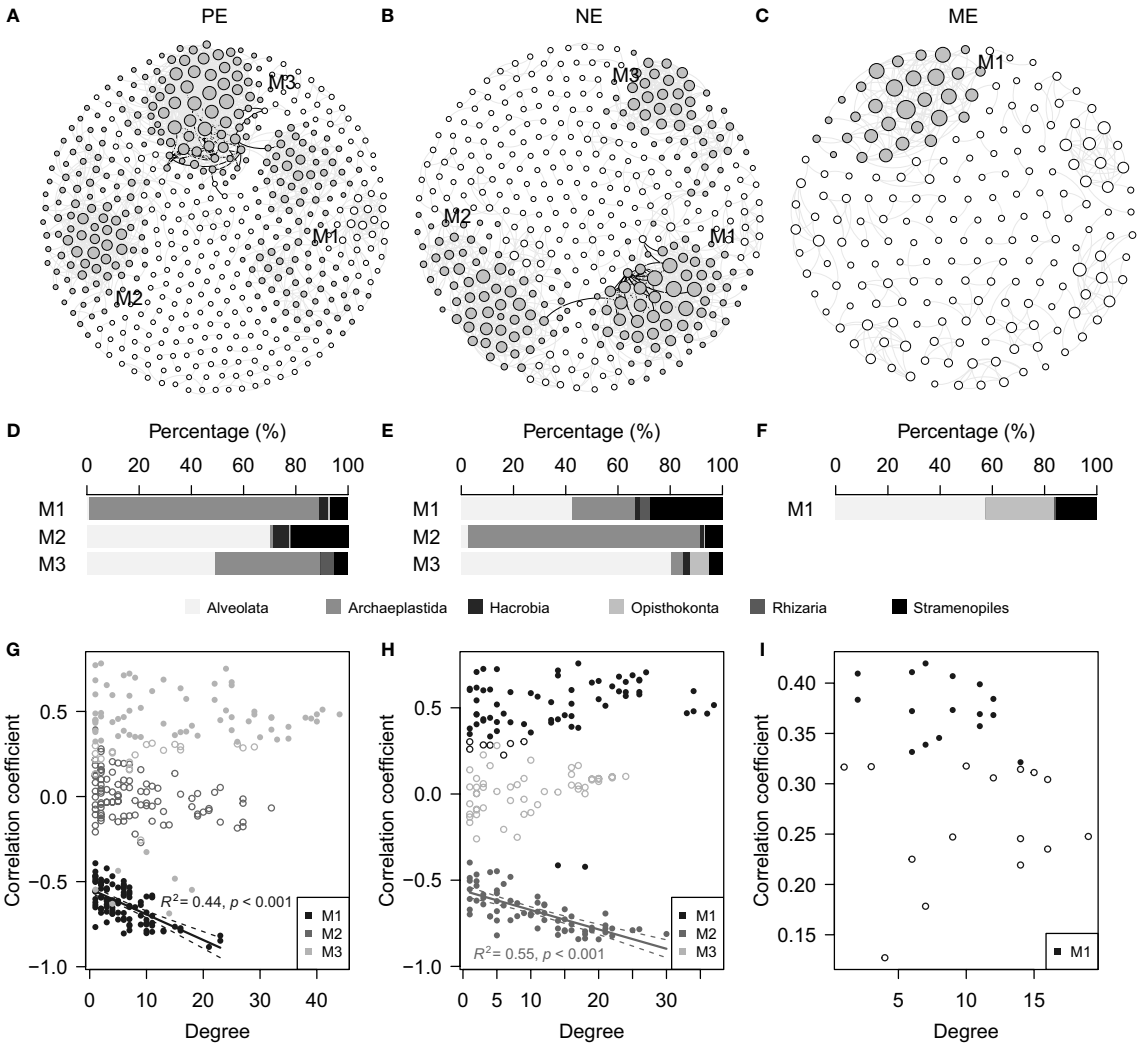


FIGURE 7
Co-occurrence networks showing the **(A)** PE, **(B)** NE and **(C)** ME categories. Major modules with >10% contributions to the total node degrees are marked with filled circles (3 modules in each of the PE and NE, and 1 module in the ME). Some of the nodes in **(A)** were manually moved to show a clear boundary among the M1 and M3 modules. **(D–F)** Horizontal stacked barcharts displaying the taxonomic compositions of the major modules shown in **(A–C)**, respectively. **(G–I)** Scatter plots illustrating the relationships between degrees and correlation coefficients inferred from Spearman's rank correlations between the ASV relative abundance and PC 1. Significant and non-significant Spearman's rank correlations are indicated by filled and open circles, respectively. Significant linear regressions (p < 0.05), which include **(G)** M1 in PE and **(H)** M2 in NE are indicated by solid lines with a 95% confidence interval.

River Estuary, the two ends of the freshwater- and seawater-like habitats were the major contributors of the community variations observed. In addition, the community variations of the microbial eukaryotes could mainly be explained by the environmental effects, which decreased from the smaller- to larger-sized categories. The co-occurrence patterns of each category further identified a few modules with opposite responses to the environmental context, suggesting that they were largely responsible for the community variations. In conclusion, in addition to cell size, it is important to conduct more trait-based examinations to achieve a better understanding of marine planktons.

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: NCBI BioProject: PRJNA836731

Author contributions

WW and HL conceived and designed the study. WW performed the experiments, analyzed the data and wrote the manuscript. WW and HL contributed to manuscript revision and approved the submitted version.

Funding

This work was supported by the National Natural Science Foundation of China (No. 42176143) and the Theme-based Research Scheme of the Hong Kong Research Grants Council (No. T21/602/16).

Acknowledgments

The authors would like to thank Zhimeng Xu (The Hong Kong University of Science and Technology) for his help with field sampling and Prof. Jianping Gan (The Hong Kong University of Science and Technology) and Prof. Minhan Dai (Xiamen University) for providing the environmental variables.

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/fmars.2022.933256/full#supplementary-material

References

Acevedo-Trejos, E., Brandt, G., Merico, A., and Smith, S. L. (2013). Biogeographical patterns of phytoplankton community size structure in the oceans. *Glob. Ecol. Biogeogr.* 22, 1060–1070. doi: 10.1111/geb.12071

Adl, S. M., Bass, D., Lane, C. E., Lukeš, J., Schoch, C. L., Smirnov, A., et al. (2019). Revisions to the classification, nomenclature, and diversity of eukaryotes. *J. Eukaryot. Microbiol.* 66, 4–119. doi: 10.1111/jeu.12691

Andersen, K. H., Berge, T., Gonçalves, R. J., Hartvig, M., Heuschele, J., Hylander, S., et al. (2016). Characteristic sizes of life in the oceans, from bacteria to whales. *Ann. Rev. Mar. Sci.* 8, 217–241. doi: 10.1146/annurev-marine-122414-034144

Bastian, M., Heymann, S., and Jacomy, M. (2009). Gephi: an open source software for exploring and manipulating networks. *Proc. Int. AAAI Conf. Web Soc. Media*. 3, 361-362

Bauer, J. E., Cai, W.-J., Raymond, P. A., Bianchi, T. S., Hopkinson, C. S., and Regnier, P. A. G. (2013). The changing carbon cycle of the coastal ocean. *Nature* 504, 61–70. doi: 10.1038/nature12857

Blanchet, F. G., Cazelles, K., and Gravel, D. (2020). Co-Occurrence is not evidence of ecological interactions. *Ecol. Lett.* 23, 1050–1063. doi: 10.1111/ele.13525

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, 90. doi: 10.1186/s40168-018-0470-z

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

Cabrerizo, M. J., and Marañón, E. (2021). Grazing pressure is independent of prey size in a generalist herbivorous protist: insights from experimental temperature gradients. *Microb. Ecol.* 81, 553–562. doi: 10.1007/s00248-020-01578-7

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

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., et al. (2009). BLAST+: architecture and applications. *BMC Bioinf.* 10, 421. doi: 10.1186/1471-2105-10-421

Caron, D. A., Countway, P. D., Jones, A. C., Kim, D. Y., and Schnetzer, A. (2012). Marine protistan diversity. *Ann. Rev. Mar. Sci.* 4, 467–493. doi: 10.1146/annurev-marine-120709-142802

Caron, D. A., Peele, E. R., Lim, E. L., and Dennett, M. R. (1999). Picoplankton and nanoplankton and their trophic coupling in surface waters of the Sargasso Sea south of Bermuda. *Limnol. Oceanogr.* 44, 259–272. doi: 10.4319/lo.1999.44.2.0259

Caron, D. A., Worden, A. Z., Countway, P. D., Demir, E., and Heidelberg, K. B. (2009). Protists are microbes too: a perspective. *ISME J.* 3, 4–12. doi: 10.1038/ismej.2008.101

Clarke, L. J., and Deagle, B. E. (2020). Eukaryote plankton assemblages in the southern kerguelen axis region: ecological drivers differ between size fractions. *Deep-Sea Res. II* 174, 104538. doi: 10.1016/j.dsr2.2018.12.003

Clauset, A., Newman, M. E. J., and Moore, C. (2004). Finding community structure in very large networks. *Phys. Rev. E* 70, 066111. doi: 10.1103/physreve.70.066111

Cloern, J. E. (2018). Why large cells dominate estuarine phytoplankton. *Limnol. Oceanogr.* 63, S392–S409. doi: 10.1002/lno.10749

Corsaro, D., Köhsler, M., Wylezich, C., Venditti, D., Walochnik, J., and Michel, R. (2018). New insights from molecular phylogenetics of amoebophagous fungi (Zoopagomycota, zoopagales). *Parasitol. Res.* 117, 157–167. doi: 10.1007/s00436-017-5685-6

Csárdi, G., and Nepusz, T. (2006). "The igraph software package for complex network research." *Inter Journal.* Complex Systems, 1695. Available at: http://igraph.org.

Dai, M., Meng, F., Tang, T., Kao, S.-J., Lin, J., Chen, J., et al. (2009). Excess total organic carbon in the intermediate water of the south China Sea and its export to the north pacific. *Geochem. Geophys. Geosyst.* 10, Q12002. doi: 10.1029/2009gc002752

Decelle, J., Romac, S., Sasaki, E., Not, F., and Mahé, F. (2014). Intracellular diversity of the V4 and V9 regions of the 18S rRNA in marine protists (radiolarians) assessed by high-throughput sequencing. *PloS One* 9, e104297. doi: 10.1371/journal.pone.0104297

de Vargas, C., Audic, S., Henry, N., Decelle, J., Mahé, F., Logares, R., et al. (2015). Eukaryotic plankton diversity in the sunlit ocean. *Science* 348, 1261605. doi: 10.1126/science.1261605

Dray, S., Bauman, D., Blanchet, G., Borcard, D., Clappe, S., Guenard, G., et al. (2022). Adespatial: Multivariate multiscale spatial analysis. R package version 0.3-19. Available at: https://CRAN.R-project.org/package=adespatial.

Dunthorn, M., Klier, J., Bunge, J., and Stoeck, T. (2012). Comparing the hypervariable V4 and V9 regions of the small subunit rDNA for assessment of ciliate environmental diversity. *J. Eukaryot. Microbiol.* 59, 185–187. doi: 10.1111/j.1550-7408.2011.00602.x

Duret, M. T., Pachiadaki, M. G., Stewart, F. J., Sarode, N., Christaki, U., Monchy, S., et al. (2015). Size-fractionated diversity of eukaryotic microbial communities in the Eastern tropical north pacific oxygen minimum zone. *FEMS Microbiol. Ecol.* 91, fiv037. doi: 10.1093/femsec/fiv037

Elferink, S., Wohlrab, S., Neuhaus, S., Cembella, A., Harms, L., and John, U. (2020). Comparative metabarcoding and metatranscriptomic analysis of microeukaryotes within coastal surface waters of west Greenland and northwest Iceland. *Front. Mar. Sci.* 7. doi: 10.3389/fmars.2020.00439

Finkel, Z. V., Beardall, J., Flynn, K. J., Quigg, A., Rees, T. A. V., and Raven, J. A. (2010). Phytoplankton in a changing world: cell size and elemental stoichiometry. *J. Plankton Res.* 32, 119–137. doi: 10.1093/plankt/fbp098

Fuhrman, J. A., Cram, J. A., and Needham, D. M. (2015). Marine microbial community dynamics and their ecological interpretation. *Nat. Rev. Microbiol.* 13, 133–146. doi: 10.1038/nrmicro3417

Gallego, I., Venail, P., and Ibelings, B. W. (2019). Size differences predict niche and relative fitness differences between phytoplankton species but not their coexistence. *ISME J.* 13, 1133–1143. doi: 10.1038/s41396-018-0330-7

Genitsaris, S., Monchy, S., Breton, E., Lecuyer, E., and Christaki, U. (2016). Small-scale variability of protistan planktonic communities relative to environmental pressures and biotic interactions at two adjacent coastal stations. *Mar. Ecol. Prog. Ser.* 548, 61–75. doi: 10.3354/meps11647

Goberna, M., and Verdú, M. (2016). Predicting microbial traits with phylogenies. *ISME J.* 10, 959–967. doi: 10.1038/ismej.2015.171

Grattepanche, J. D., Breton, E., Brylinski, J. M., Lecuyer, E., and Christaki, U. (2011). Succession of primary producers and micrograzers in a coastal ecosystem dominated by *Phaeocystis globosa* blooms. *J. Plankton Res.* 33, 37–50. doi: 10.1093/plankt/fbq097

Grossart, H.-P., Van den Wyngaert, S., Kagami, M., Wurzbacher, C., Cunliffe, M., and Rojas-Jimenez, K. (2019). Fungi in aquatic ecosystems. *Nat. Rev. Microbiol.* 17, 339–354. doi: 10.1038/s41579-019-0175-8

Guillou, L., Bachar, D., Audic, S., Bass, D., Berney, C., Bittner, L., et al. (2013). The protist ribosomal reference database (PR^2): a catalog of unicellular eukaryote small sub-unit rRNA sequences with curated taxonomy. *Nucleic Acids Res.* 41, D597–D604. doi: 10.1093/nar/gks1160

Guillou, L., Viprey, M., Chambouvet, A., Welsh, R. M., Kirkham, A. R., Massana, R., et al. (2008). Widespread occurrence and genetic diversity of marine parasitoids belonging to *Syndiniales* (Alveolata). *Environ. Microbiol.* 10, 3349–3365. doi: 10.1111/j.1462-2920.2008.01731.x

Han, A., Dai, M., Kao, S.-J., Gan, J., Li, Q., Wang, L., et al. (2012). Nutrient dynamics and biological consumption in a large continental shelf system under the influence of both a river plume and coastal upwelling. *Limnol. Oceanogr.* 57, 486–502. doi: 10.4319/lo.2012.57.2.0486

Hernández-Ruiz, M., Barber-Lluch, E., Prieto, A., Logares, R., and Teira, E. (2020). Response of pico-nano-eukaryotes to inorganic and organic nutrient additions. *Estuar. Coast. Shelf Sci.* 235, 106565. doi: 10.1016/j.ecss.2019.106565

Jamy, M., Biwer, C., Vaulot, D., Obiol, A., Jing, H., Peura, S., et al. (2022). Global patterns and rates of habitat transitions across the eukaryotic tree of life. *Nat. Ecol. Evol.* 6, 1458-1470. doi: 10.1038/s41559-022-01838-4

Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. doi: 10.1093/molbev/mst010

Kruskal, W. H., and Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *J. Am. Stat. Assoc.* 47, 583–621. doi: 10.2307/2280779

Legendre, P., and De Cáceres, M. (2013). Beta diversity as the variance of community data: dissimilarity coefficients and partitioning. *Ecol. Lett.* 16, 951–963. doi: 10.1111/ele.12141

Liu, L., Liu, M., Wilkinson, D. M., Chen, H., Yu, X., and Yang, J. (2017). DNA Metabarcoding reveals that 200-μm-size-fractionated filtering is unable to discriminate between planktonic microbial and large eukaryotes. *Mol. Ecol. Resour.* 17, 991–1002. doi: 10.1111/1755-0998.12652

Logares, R., Audic, S., Bass, D., Bittner, L., Boutte, C., Christen, R., et al. (2014). Patterns of rare and abundant marine microbial eukaryotes. *Curr. Biol.* 24, 813–821. doi: 10.1016/j.cub.2014.02.050

López-Urrutia, Á., and Morán, X. A. G. (2015). Temperature affects the size-structure of phytoplankton communities in the ocean. *Limnol. Oceanogr.* 60, 733–738. doi: 10.1002/lno.10049

Lozupone, C., and Knight, R. (2005). UniFrac: a new phylogenetic method for comparing microbial communities. *Appl. Environ. Microbiol.* 71, 8228–8235. doi: 10.1128/aem.71.12.8228-8235.2005

Malviya, S., Scalco, E., Audic, S., Vincent, F., Veluchamy, A., Poulain, J., et al. (2016). Insights into global diatom distribution and diversity in the world's ocean. *Proc. Natl. Acad. Sci. U. S. A.* 113, E1516–E1525. doi: 10.1073/pnas.1509523113

Marañón, E. (2015). Cell size as a key determinant of phytoplankton metabolism and community structure. *Ann. Rev. Mar. Sci.* 7, 241–264. doi: 10.1146/annurev-marine-010814-015955

Marañón, E., Cermeño, P., Latasa, M., and Tadonléké, R. D. (2015). Resource supply alone explains the variability of marine phytoplankton size structure. *Limnol. Oceanogr.* 60, 1848–1854. doi: 10.1002/lno.10138

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

Martiny, A. C., Treseder, K., and Pusch, G. (2013). Phylogenetic conservatism of functional traits in microorganisms. *ISME J.* 7, 830–838. doi: 10.1038/ismej.2012.160

Massana, R., del Campo, J., Sieracki, M. E., Audic, S., and Logares, R. (2014). Exploring the uncultured microeukaryote majority in the oceans: reevaluation of ribogroups within stramenopiles. *ISME J.* 8, 854–866. doi: 10.1038/ismej.2013.204

Massana, R., Gobet, A., Audic, S., Bass, D., Bittner, L., Boutte, C., et al. (2015). Marine protist diversity in European coastal waters and sediments as revealed by high-throughput sequencing. *Environ. Microbiol.* 17, 4035–4049. doi: 10.1111/1462-2920.12955

Mestre, M., Ruiz-González, C., Logares, R., Duarte, C. M., Gasol, J. M., and Sala, M. M. (2018). Sinking particles promote vertical connectivity in the ocean microbiome. *Proc. Natl. Acad. Sci. U. S. A.* 115, E6799–E6807. doi: 10.1073/pnas.1802470115

Mousing, E. A., Ellegaard, M., and Richardson, K. (2014). Global patterns in phytoplankton community size structure-evidence for a direct temperature effect. *Mar. Ecol. Prog. Ser.* 497, 25–38. doi: 10.3354/meps10583

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., et al. (2018) *Vegan: community ecology package. r package version 2.5-3*. Available at: http://CRAN.R-project.org/package=vegan.

Pagel, M. (1999). Inferring the historical patterns of biological evolution. *Nature* 401, 877–884. doi: 10.1038/44766

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

R Core Team (2018). *R: A language and environment for statistical computing* (Vienna, Austria: R Foundation for Statistical Computing). Available at: http://www.R-project.org/.

Reuman, D. C., Gislason, H., Barnes, C., Mélin, F., and Jennings, S. (2014). The marine diversity spectrum. *J. Anim. Ecol.* 83, 963–979. doi: 10.1111/1365-2656.12194

Revell, L. J. (2012). Phytools: an r package for phylogenetic comparative biology (and other things). *Methods Ecol. Evol.* 3, 217–223. doi: 10.1111/j.2041-210x.2011.00169.x

Röttjers, L., and Faust, K. (2018). From hairballs to hypotheses–biological insights from microbial networks. *FEMS Microbiol. Rev.* 42, 761–780. doi: 10.1093/femsre/fuy030

Sørensen, N., Daugbjerg, N., and Richardson, K. (2013). Choice of pore size can introduce artefacts when filtering picoeukaryotes for molecular biodiversity studies. *Microb. Ecol.* 65, 964–968. doi: 10.1007/s00248-012-0174-z

Salazar, G., Cornejo-Castillo, F. M., Borrull, E., Díez-Vives, C., Lara, E., Vaqué, D., et al. (2015). Particle-association lifestyle is a phylogenetically conserved trait in bathypelagic prokaryotes. *Mol. Ecol.* 24, 5692–5706. doi: 10.1111/mec.13419

Shannon, C. E. (1948). A mathematical theory of communication. *Bell Syst. Tech. J.* 27, 379–423, 623–656. doi: 10.1002/j.1538-7305.1948.tb01338.x

Sommeria-Klein, G., Watteaux, R., Ibarbalz, F. M., Pierella Karlusich, J. J., Iudicone, D., Bowler, C., et al. (2021). Global drivers of eukaryotic plankton biogeography in the sunlit ocean. *Science* 374, 594–599. doi: 10.1126/science.abb3717

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, 21–31. doi: 10.1111/j.1365-294x.2009.04480.x

Tamigneaux, E., Mingelbier, M., Klein, B., and Legendre, L. (1997). Grazing by protists and seasonal changes in the size structure of protozooplankton and phytoplankton in a temperate nearshore environment (western gulf of st. Lawrence, Canada). *Mar. Ecol. Prog. Ser.* 146, 231–247. doi: 10.3354/meps146231

Tragin, M., and Vaulot, D. (2018). Green microalgae in marine coastal waters: the ocean sampling day (OSD) dataset. *Sci. Rep.* 8, 14020. doi: 10.1038/s41598-018-32338-w

Tragin, M., Zingone, A., and Vaulot, D. (2018). Comparison of coastal phytoplankton composition estimated from the V4 and V9 regions of the 18S rRNA gene with a focus on photosynthetic groups and especially chlorophyta. *Environ. Microbiol.* 20, 506–520. doi: 10.1111/1462-2920.13952

Unrein, F., Gasol, J. M., Not, F., Forn, I., and Massana, R. (2014). Mixotrophic haptophytes are key bacterial grazers in oligotrophic coastal waters. *ISME J.* 8, 164–176. doi: 10.1038/ismej.2013.132

Villarino, E., Watson, J. R., Jönsson, B., Gasol, J. M., Salazar, G., Acinas, S. G., et al. (2018). Large-Scale ocean connectivity and planktonic body size. *Nat. Commun.* 9, 142. doi: 10.1038/s41467-017-02535-8

Wang, F., Huang, B., Xie, Y., Cai, S., Wang, X., and Mu, J. (2021). Diversity, composition, and activities of nano- and pico-eukaryotes in the northern south China Sea with influences of kuroshio intrusion. *Front. Mar. Sci.* 8. doi: 10.3389/fmars.2021.658233

White, E. P., Ernest, S. K. M., Kerkhoff, A. J., and Enquist, B. J. (2007). Relationships between body size and abundance in ecology. *Trends Ecol. Evol.* 22, 323–330. doi: 10.1016/j.tree.2007.03.007

Wikner, J., and Hagstrom, A. (1988). Evidence for a tightly coupled nanoplanktonic predator-prey link regulating the bacterivores in the marine environment. *Mar. Ecol. Prog. Ser.* 50, 137–145. doi: 10.3354/meps050137

Wood, S. N. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *J. R. Stat. Soc Ser. B* 73, 3–36. doi: 10.1111/j.1467-9868.2010.00749.x

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

Wu, W., and Liu, H. (2018). Disentangling protist communities identified from DNA and RNA surveys in the pearl river–south China Sea continuum during the wet and dry seasons. *Mol. Ecol.* 27, 4627–4640. doi: 10.1111/mec.14867

Wu, W., Xu, Z., Dai, M., Gan, J., and Liu, H. (2021). Homogeneous selection shapes free-living and particle-associated bacterial communities in subtropical coastal waters. *Divers. Distrib.* 27, 1904–1917. doi: 10.1111/ddi.13193

Xu, D., Kong, H., Yang, E., Li, X., Jiao, N., Warren, A., et al. (2020). Contrasting community composition of active microbial eukaryotes in melt ponds and sea water of the Arctic ocean revealed by high throughput sequencing. *Front. Microbiol.* 11. doi: 10.3389/fmicb.2020.01170

Yu, G., Smith, D. K., Zhu, H., Guan, Y., and Lam, T. T.-Y. (2017). Ggtree: an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods Ecol. Evol.* 8, 28–36. doi: 10.1111/2041-210x.12628

Zou, K., Wang, R., Xu, S., Li, Z., Liu, L., Li, M., et al. (2021). Changes in protist communities in drainages across the pearl river delta under anthropogenic influence. *Water Res.* 200, 117294. doi: 10.1016/j.watres.2021.117294

frontiers | Frontiers in Microbiology

TYPE Original Research
PUBLISHED 04 November 2022
DOI 10.3389/fmicb.2022.977037

OPEN ACCESS

EDITED BY
Hongbin Liu,
Hong Kong University of Science and Technology, Hong Kong SAR, China

REVIEWED BY
Yuying Li,
Nanyang Normal University, China
Anyi Hu,
Institute of Urban Environment (CAS), China

*CORRESPONDENCE
Weihong Zhang
zhangweihong@wbgcas.cn

†These authors have contributed equally to this work

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

RECEIVED 24 June 2022
ACCEPTED 10 October 2022
PUBLISHED 04 November 2022

CITATION
Zhao Y, Lin H, Liu Y, Jiang Y and Zhang W (2022) Abundant bacteria shaped by deterministic processes have a high abundance of potential antibiotic resistance genes in a plateau river sediment.
Front. Microbiol. 13:977037.
doi: 10.3389/fmicb.2022.977037

Abundant bacteria shaped by deterministic processes have a high abundance of potential antibiotic resistance genes in a plateau river sediment

Yuhong Zhao[1†], Hui Lin[2†], Yi Liu[3,4], Ying Jiang[3,4] and Weihong Zhang[3,4*]

[1]Tibet Agricultural and Animal Husbandry University, Nyingchi, China, [2]State Key Laboratory for Managing Biotic and Chemical Threats to the Quality and Safety of Agro-products, Institute of Environment Resource, Soil and Fertilizers, Zhejiang Academy of Agricultural Sciences, Hangzhou, China, [3]Key Laboratory of Aquatic Botany and Watershed Ecology, Wuhan Botanical Garden, Chinese Academy of Sciences, Wuhan, China, [4]University of Chinese Academy of Sciences, Beijing, China

Recent research on abundant and rare bacteria has expanded our understanding of bacterial community assembly. However, the relationships of abundant and rare bacteria with antibiotic resistance genes (ARGs) remain largely unclear. Here, we investigated the biogeographical patterns and assembly processes of the abundant and rare bacteria from river sediment at high altitudes (Lhasa River, China) and their potential association with the ARGs. The results showed that the abundant bacteria were dominated by *Proteobacteria* (55.4%) and *Cyanobacteria* (13.9%), while the *Proteobacteria* (33.6%) and *Bacteroidetes* (18.8%) were the main components of rare bacteria. Rare bacteria with a large taxonomic pool can provide function insurance in bacterial communities. Spatial distribution of persistent abundant and rare bacteria also exhibited striking differences. Strong selection of environmental heterogeneity may lead to deterministic processes, which were the main assembly processes of abundant bacteria. In contrast, the assembly processes of rare bacteria affected by latitude were dominated by stochastic processes. Abundant bacteria had the highest abundance of metabolic pathways of potential drug resistance in all predicted functional genes and a high abundance of potential ARGs. There was a strong potential connection between these ARGs and mobile genetic elements, which could increase the ecological risk of abundant taxa and human disease. These results provide insights into sedimental bacterial communities and ARGs in river ecosystems.

KEYWORDS

antibiotic resistance gene, potential hosts, plateau river sediment, community assembly, abundant bacteria

Introduction

Rivers play a vital and irreplaceable role in the process of human civilization and global biogeochemical cycling. Due to development of human society and rapid economic growth, river pollution has become a critical challenge (Shao et al., 2006; Wang et al., 2019). Antibiotic resistance genes (ARGs) and antibiotic-resistant bacteria (ARB) are recognized as emerging contaminants (Cosgrove, 2006; Amarasiri et al., 2020). Environmental pollution due to factors such as heavy metals may accelerate the enrichment and evolution of ARB and ARGs and increase the risk of transmission of the environmental resistome to humans (Yang et al., 2018). Bacterial community composition is a vital factor affecting the distribution of ARGs. For example, *Cyanobacteria* blooms promote the diversity of ARGs in aquatic ecosystems (Zhang et al., 2020). The change in the bacterial community promotes the improvement of ARGs in the chlorination process of drinking water (Jia et al., 2015). These studies on the correlation between bacterial community and ARGs were conducted at the overall level of the community. However, microbial communities in nature are comprised of a large number of species, while few of these species are abundant, and a large number of species are often called the "rare biosphere" (Sogin et al., 2006; Jiao et al., 2017). To date, we still know little about how spatial variation in ARG composition relates to bacterial taxonomic composition (i.e., abundant bacteria or rare bacteria) in a river continuum.

Abundant and rare bacteria in sediments are major participants in the biogeochemical cycle of rivers (Jia et al., 2015; Goldman et al., 2017). It is the core goal of community ecology to reveal the basic mechanism of the generation and maintenance of river microbial community diversity, and some interesting patterns have been discovered. For example, the physicochemical properties (such as pH, heavy metals content, and nutritional status) and spatial distribution (such as horizontal geographic distribution and vertical altitude distribution) were important drivers of the unique biogeographic patterns of microbial communities (Chen et al., 2020b; Wang Y. et al., 2020; Wang et al., 2021). However, there has been little consistency in the studies so far due to the heterogeneity of the river ecosystem.

In recent years, pollutants from industry and life have entered the water of the Lhasa River and caused a certain degree of pollution to the water quality. The changes in microbial community diversity and structure can indirectly or/and directly affect the aquatic ecological function, which is a comprehensive and sensitive index of environmental quality in the aquatic ecosystem (Long et al., 2021). The main objective of this study was to examine the biogeographical patterns and assembly processes of the abundant and rare bacteria in sediment and the potential association with the ARGs in the sediment of the Lhasa River. Therefore, 16S rRNA gene sequencing and qPCR reaction were used to analyze the sediments of the Lhasa River to determine the adaptation mechanism of microorganisms and resistance genes in the sediments. We hope this study could provide new insights into sedimental bacterial communities and ARGs in river ecosystems.

Materials and methods

Study sites and sample collection

Lhasa River (90.08–93.33°E, 29.33–31.25°N) known as the "Water Tower of Asia" is located in the Qinghai-Tibetan Plateau and is one of the highest rivers in the world (Qu et al., 2019). The Lhasa River basin is about 568 km long from east to west, and the altitude is between 3,570 and 5,200 m above sea level. More than 70% of the population of the Lhasa River Basin is concentrated from Mozhugongka county to Qushui county. Therefore, we set up 10 sampling sites along the Lhasa River from the Mozhugongka to the Qushui county with detailed geographic information on the sampling sites (Supplementary Table S1; Supplementary materials). Surface sediment (0–5 cm) was collected from each site in September 2019 using a stainless-steel core sampler. Three sub-samples were collected from each site, mixed as one sample, kept in a car refrigerator, transported to the laboratory, and stored at −80°C before 16S rRNA gene sequencing. The contents of Cr, Co, Cu, Zn, As, Cd, Hg, and Pb in sediment were detected by inductively coupled plasma mass spectrometer (ICP-MS, X Series 2, Germany). Detailed data and measurement methods are shown in Supplementary Table S2. Detailed information about sediment physicochemical properties (temperature, pH, salinity, and conductivity) and nutrients [total nitrogen (TN) and total carbon (TC)] are shown in Supplementary Table S3.

16S rRNA gene sequencing

Genomic DNA of the bacterial community from each site was extracted using a bacterial DNA Extraction Kit (Tiangen Biotech, Inc., Beijing, China) according to the manufacturer's protocols (Zhang et al., 2021a). The DNA served as a template for PCR amplification of the V4 region of 16S rRNA using the primer set 515F/806R (Caporaso et al., 2011; Walters et al., 2016). The sequencing library was set up when the amplicons of 16S rRNA were purified, and Ion S5™XL of Thermofisher was used in the sequencing. The raw fastq data were quality-filtered by low-quality parts and chimeric sequences to get clean reads (Martin, 2011; Rognes et al., 2016). The clean reads were clustered into operational taxonomic units (OTUs) at the 97% similarity level using Uparse (Edgar, 2013). Since this study only focused on bacteria, we deleted all OTUs that did not belong to bacteria. The MUSCLE method and the SSU rRNA database of silva132 were used for the annotation species analysis (Wang et al., 2007; Quast et al., 2013). We followed a previously reported method (Wemheuer et al., 2020) and applied Tax4Fun to reveal the functional and redundancy index (FRI) of the sequenced bacterial genome.

Analysis of ARGs in the sediment from the Lhasa River

A total of 23 ARGs and the 16S rRNA gene were selected to investigate the distribution of ARGs in the sediment from the Lhasa River. Herein, the representative ARGs in the environment and clinically important ARGs were taken into account based on the potential ecological risks and threats to human health (Zhang et al., 2021a; Jiang et al., 2021b). The 23 ARGs including seven major classes of antibiotic-related ARGs, which were the colistin (*mcr*-1, *mcr*-3, and *mcr*-7), beta-lactam ($bla_{\text{CTX-M-32}}$, $bla_{\text{NMD-1}}$, bla_{CMY}, $bla_{\text{CTX-M}}$, and bla_{TEM}), aminoglycosides (*aad*A, *str*B, and *arm*A), macrolide (*ere*A, *ere*B, and *mph*A), quinolones (*qnr*A, *qnr*B, and *qnr*S), sulfonamides (*sul*1, *sul*2, and *sul*3), and tetracycline (*tet*A, *tet*M, and *tet*X) resistance genes, respectively. Besides, the transposase gene (*tnp*A) and class 1 integron-integrase gene (*int*I1) were selected to investigate the transfer or propagation of ARGs in the Lhasa River sediment. Detailed information on the primers and their corresponding target genes was given in Supplementary Table S4. The qPCR reaction in a 10 μl reaction volume was performed according to the denaturation at 95°C for 30 s, followed by the thermal cycles of qPCR consisting of 40 cycles of 10 s at 95°C, 30 s annealing at 55°C, and 1 min extension at 72°C. The relative abundances of ARGs and mobile genetic elements (MGEs) were calculated using the $2^{-\Delta CT}$ method [Equations (1) and (2); Zhang et al., 2021a].

$$\Delta C_T = C_{T(\text{ARG})} - C_{T(16\text{SrRNA gene})} \quad (1)$$

The relative abundance of

$$\text{ARG} = 2^{-\Delta C_T} \quad (2)$$

Statistical analysis

Previous studies have generally defined OTUs at the regional level with average relative abundances >0.10% as "abundant," those with average relative abundances <0.01% as "rare" and those in between as "intermediate" (Jiao and Lu, 2020a; Wan et al., 2021a; Zhang et al., 2021b; Zhang Y. et al., 2021). However, this definition is not suitable for our data because the total abundance of these abundant bacterial OTUs in some samples was lower than 50% accounting for this sample's total reads. Similarly, some previous studies also defined OTUs at the regional level with average relative abundances >0.05% as "abundant" (Jiao et al., 2017; Hou et al., 2020). Thus, across all sediment, the average relative abundance of OTUs above 0.05% was defined as abundant bacteria, while the average relative abundance of OTUs below 0.01% was regarded as rare bacteria. The remainder OTUs (0.01–0.05%) were deemed as "intermediate." The community similarity (1-Bray–Curtis distance) and phylogenetic similarity (1-βNMTD) of abundant and rare bacteria were calculated based on taxonomic distance and phylogenetic distance, respectively. Then, the distance-decay relationship (DDR) was used to reveal the responses of community similarity and phylogenetic similarity to horizontal (geographic distance) and vertical (altitude distance) spatial distribution and environmental heterogeneity (Bray-Curtis distance). The network was constructed by Spearman correlation and visualized *via* Gephi software (0.9.1; Gephi, WebAtlas, France). We identify the contribution of different assembly processes of abundant and rare bacteria in the Lhasa River sediments *via* applying a null model analysis by Stegen et al. (2013).

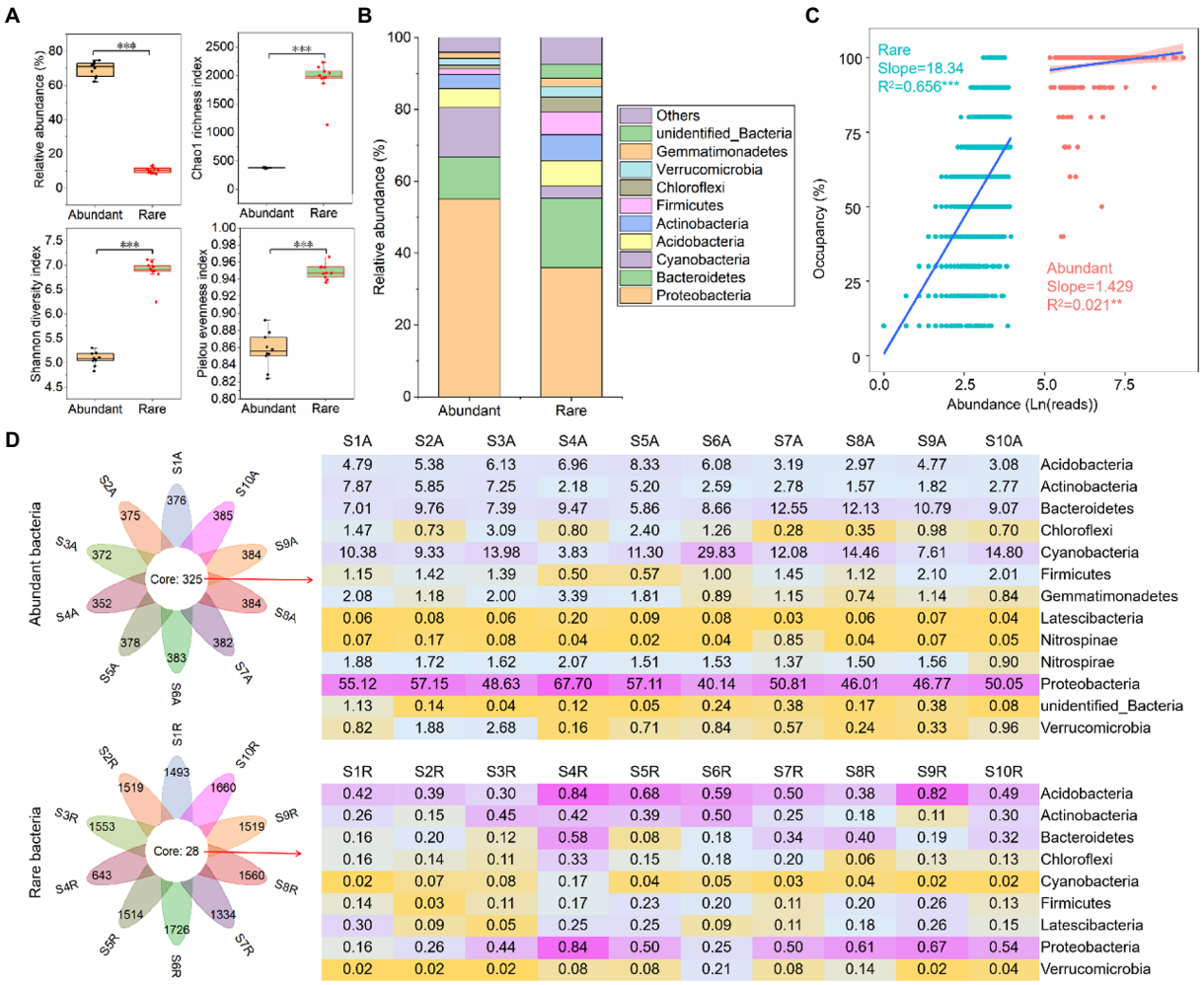
Results

Composition and distribution of abundant and rare bacteria

The relative abundance of abundant bacteria (mean = 69.6%) was higher than rare ones (10.5%; Figure 1A). Conversely, the Chao1 richness (381.6), Shannon diversity (5.08), and Pielou evenness (0.86) of abundant bacteria were lower than the rare ones (1937.5, 6.88, and 0.95, respectively; Figure 1A). At the bacterial phylum level, abundant bacteria were dominated by *Proteobacteria* (55.1%), *Cyanobacteria* (13.8%), and *Bacteroidetes* (11.6%), while the *Proteobacteria* (36.1%), *Bacteroidetes* (19.3%), and *Actinobacteria* (7.17%) were the main components of rare bacteria (Figure 1B). Abundance-occupancy relationships showed that rare bacteria possessed stronger positive correlations than abundant bacteria (Figure 1C). Meanwhile, abundant bacterial taxa had a wider distribution than the rare bacterial taxa. The petal diagram showed that abundant bacteria had 325 OTUs that persisted in all sediments, while rare bacteria only had 28 OTUs (Figure 1D). Even these persistent abundant and rare bacteria had obvious differences in spatial distribution.

The community similarity (Figure 2A) and phylogenetic similarity (Figure 2E) of abundant bacteria were higher than rare bacteria, indicating that the rare bacteria had more taxonomic and phylogenetic variation than the abundant bacteria. Furthermore, the community similarity for abundant and rare bacteria had significantly positive correlations with their corresponding phylogenetic similarity, and the correlations of rare bacteria were stronger than that of abundant bacteria (Supplementary Figure S1), indicating that the phylogeny of these abundant and rare bacteria had different sensitivities to environmental changes.

The DDR showed that the community similarity of both abundant and rare bacteria significantly decreased with the increased geographical distance (Figure 2B). Interestingly, the effect of geographical distance on the community composition of rare bacteria ($R^2 = 0.15$) was greater than that of abundant bacteria ($R^2 = 0.12$), whereas the community composition of abundant bacteria (Slope = −0.083) had more community turnover with increased of geographical distance. Besides, the composition of rare bacteria was also significantly affected by altitude in biogeographic patterns, and the community similarity significantly



	S1A	S2A	S3A	S4A	S5A	S6A	S7A	S8A	S9A	S10A	
	4.79	5.38	6.13	6.96	8.33	6.08	3.19	2.97	4.77	3.08	Acidobacteria
	7.87	5.85	7.25	2.18	5.20	2.59	2.78	1.57	1.82	2.77	Actinobacteria
	7.01	9.76	7.39	9.47	5.86	8.66	12.55	12.13	10.79	9.07	Bacteroidetes
	1.47	0.73	3.09	0.80	2.40	1.26	0.28	0.35	0.98	0.70	Chloroflexi
	10.38	9.33	13.98	3.83	11.30	29.83	12.08	14.46	7.61	14.80	Cyanobacteria
	1.15	1.42	1.39	0.50	0.57	1.00	1.45	1.12	2.10	2.01	Firmicutes
	2.08	1.18	2.00	3.39	1.81	0.89	1.15	0.74	1.14	0.84	Gemmatimonadetes
	0.06	0.08	0.06	0.20	0.09	0.08	0.03	0.06	0.07	0.04	Latescibacteria
	0.07	0.17	0.08	0.04	0.02	0.04	0.85	0.04	0.07	0.05	Nitrospinae
	1.88	1.72	1.62	2.07	1.51	1.53	1.37	1.50	1.56	0.90	Nitrospirae
	55.12	57.15	48.63	67.70	57.11	40.14	50.81	46.01	46.77	50.05	Proteobacteria
	1.13	0.14	0.04	0.12	0.05	0.24	0.38	0.17	0.38	0.08	unidentified_Bacteria
	0.82	1.88	2.68	0.16	0.71	0.84	0.57	0.24	0.33	0.96	Verrucomicrobia

	S1R	S2R	S3R	S4R	S5R	S6R	S7R	S8R	S9R	S10R	
	0.42	0.39	0.30	0.84	0.68	0.59	0.50	0.38	0.82	0.49	Acidobacteria
	0.26	0.15	0.45	0.42	0.39	0.50	0.25	0.18	0.11	0.30	Actinobacteria
	0.16	0.20	0.12	0.58	0.08	0.18	0.34	0.40	0.19	0.32	Bacteroidetes
	0.16	0.14	0.11	0.33	0.15	0.18	0.20	0.06	0.13	0.13	Chloroflexi
	0.02	0.07	0.08	0.17	0.04	0.05	0.03	0.04	0.02	0.02	Cyanobacteria
	0.14	0.03	0.11	0.17	0.23	0.20	0.11	0.20	0.26	0.13	Firmicutes
	0.30	0.09	0.05	0.25	0.25	0.09	0.11	0.18	0.26	0.15	Latescibacteria
	0.16	0.26	0.44	0.84	0.50	0.25	0.50	0.61	0.67	0.54	Proteobacteria
	0.02	0.02	0.02	0.08	0.08	0.21	0.08	0.14	0.02	0.04	Verrucomicrobia

FIGURE 1
Alpha diversity and composition of abundant and rare bacteria in sediment from the Lhasa River. **(A)** Alpha diversity of abundant and rare bacteria in sediment. **(B)** The composition of abundant and rare bacteria in sediment. **(C)** Abundance–occupancy relationship of abundant and rare bacteria in sediment. **(D)** The number and composition of shared operational taxonomic units (OTUs) in sediment. Asterisks denote significance (**p<0.01; ***p<0.001).

decreased with the increased altitude distance (Figure 2C). Similarly, the effect of environmental heterogeneity on rare bacteria (R^2 = 0.35) was greater than that of abundant bacteria (R^2 = 0.18). Rare bacteria (Slope = −0.609) had more community turnover with the increase in environmental change (Figure 2D). Specifically, the taxonomic composition of rare bacteria was significantly affected by heavy metals, such as Cu, Zn, Cd, and As, whereas only the composition of abundant bacteria was significantly affected by Cu (Supplementary Table S5). Especially, Cu had more influence on the taxonomic composition of rare bacteria than abundant bacteria (Supplementary Table S5). However, the phylogenetic similarity for abundant and rare bacteria did not decrease significantly with the increased geographical distance and altitude distance (Figures 2F,G). Only the phylogenetic similarity of rare bacteria was significantly affected by environmental heterogeneity (Figure 2H). The different responses of abundant and rare bacteria to geographic and environmental factors in the taxonomic and phylogenetic composition may indicate the presence of had distinct community ecological assembly processes.

Community assembly processes of abundant and rare bacteria in the sediment

Although the niche width of abundant bacteria (mean = 5.58) was higher than rare bacteria (2.80), the niche of the abundant bacteria showed higher differentiation (Figure 3A). Results from the null model showed that the differentiating was the dominant process for both abundant (99.8%) and rare bacteria (88.9%) assembly, while the homogenizing process (4.44%) had little impact on rare bacteria assembly (Figure 3B). Additionally, the stochastic process (62.2%) was the main assembly pattern of rare

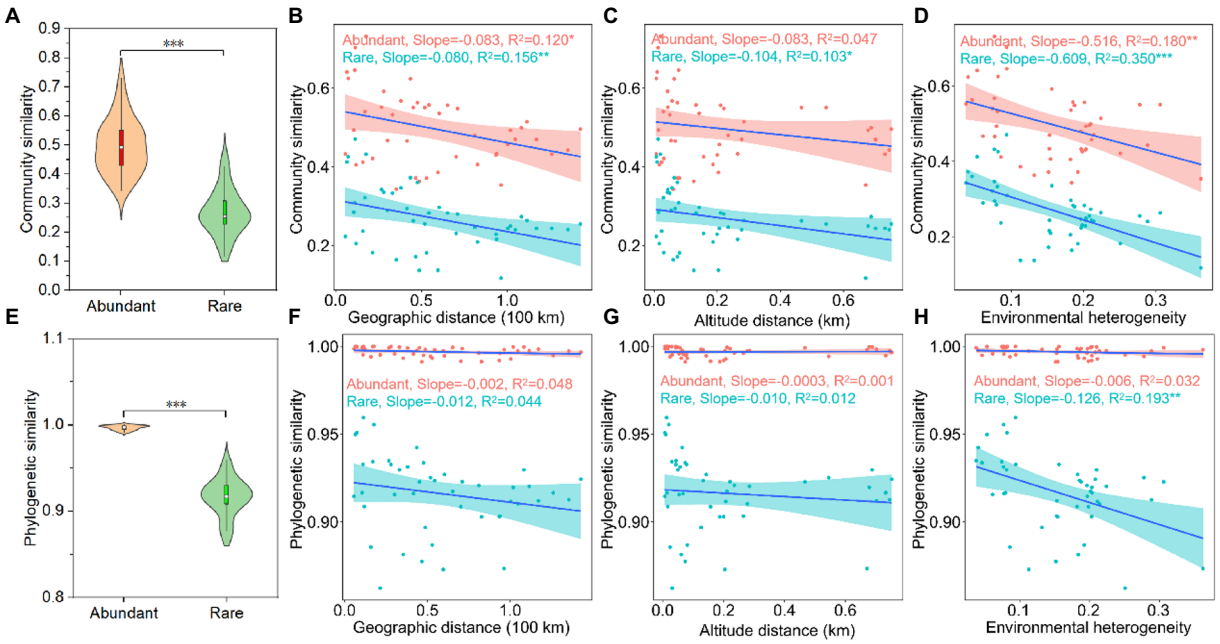


FIGURE 2
Beta-diversity patterns of taxonomic and phylogenetic for both abundant and rare bacteria in sediment from the Lhasa River. **(A)** Community similarity (1—Bray-Curtis distance) of abundant and rare bacteria. **(B–D)** Relationship of community similarity for both abundant and rare bacteria with the geographical distance, altitude distance, and environmental heterogeneity, respectively. **(E)** Phylogenetic similarity (1—βMNTD) of abundant and rare bacteria. **(F–H)** Relationship of phylogenetic similarity for both abundant and rare bacteria with the geographical distance, altitude distance, and environmental heterogeneity, respectively. Asterisks denote significance (*p<0.05; **p<0.01; and ***p<0.001).

bacteria in the sediment of the Lhasa River, followed by the deterministic process (37.8%). The deterministic process (53.3%) dominated the assembly of abundant bacteria, followed by the stochastic process (46.7%). The results show that the contribution of the stochastic and deterministic process for the assembly of abundant and rare bacteria in the sediment of the Lhasa River was different. Mantel tests suggested that the βNTI values of both abundant and rare bacteria had a significant correlation with the geospatial factor (latitude; Supplementary Table S6), indicating that the community assembly of abundant and rare bacteria may be affected by latitude. Furthermore, the βNTI values of abundant bacteria significantly correlated with pH, conductivity, and heavy metal (Cd; Supplementary Table S6). The results of the null model further suggested that the variable selection (53.3%) was the dominant assembly process of abundant bacteria, whereas the dispersal limitation (51.1%) was the dominant assembly process of rare bacteria (Figure 3C). These results suggested that the assembly of abundant bacteria was susceptible to the environmental selection, while the assembly of rare bacteria was susceptible to geospatial factors.

Co-occurrence patterns of abundant and rare bacteria

A metacommunity network was conducted based on the strong ($|r|>0.8$) and significant ($p<0.01$) Spearman correlations to explore the co-occurrence patterns of the sedimental microbial communities of the Lhasa River (Figure 4A). The network consisted of 2,699 nodes linked by 40,344 edges. Degree, Betweenness centrality, Closeness centrality, and Eigen centrality of the network within abundant bacteria were significantly higher than within rare bacteria (Figure 4B), indicating that the abundant bacteria played an important role in maintaining community structure. Abundant bacteria interacted more with other bacterial taxa than within themselves (Figure 4A). Although the number of positive correlation edges (59.1%) was higher than that of negative correlation edges (40.9%) in the co-occurrence network of whole bacteria in sediment, the proportion of negative correlation was different within and between the different bacterial taxa. For example, the proportion of negative correlation within bacterial taxa was lower than between these bacterial taxa, suggesting there may be stronger competition between different bacterial taxa than that within these bacterial taxa. Further, the proportion of negative correlation within rare bacteria (40.1%) was higher than within abundant bacteria (38.9%), indicating that there may be stronger competition within rare bacteria than within abundant bacteria.

Potential function analysis of abundant and rare bacteria

When compared to rare bacteria, abundant bacteria not only had the highest abundance of metabolic pathways of potential

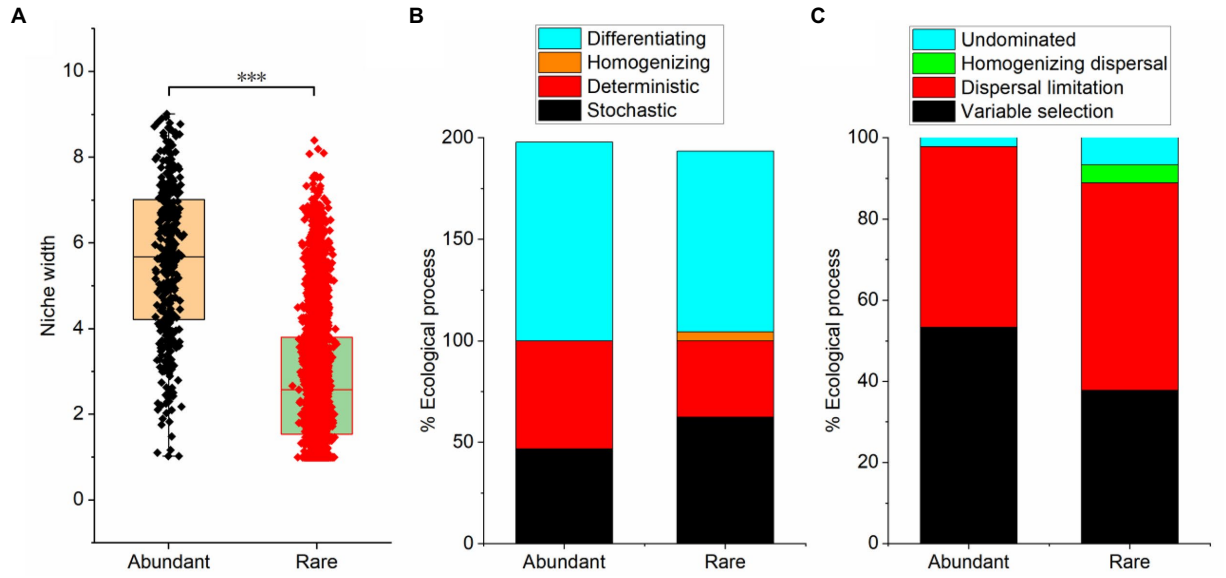


FIGURE 3
Niche width **(A)** and community assembly processes **(B,C)** of abundant and rare bacteria in sediment from the Lhasa River. Stochastic=Dispersal limitation+Homogenizing dispersal+Undominated processes; Deterministic=Variable selection+Homogeneous selection; Homogenizing=Homogeneous selection+Homogenizing dispersal; and Differentiating=Variable selection+Dispersal limitation.

drug resistance (such as antimicrobial and antineoplastic resistance), chemical structure transformation maps, cell growth and death, cell motility, and xenobiotics biodegradation and metabolism but also had the more potential pathogenic potential of human disease (such as infectious diseases, etc.) in all predicted functional genes (Figure 5A). However, abundant taxa had weak global and overview maps (such as biosynthesis of antibiotics and carbon metabolism), carbohydrate metabolism, metabolism of terpenoids and polyketides, nucleotide metabolism, glycan biosynthesis and metabolism, and biosynthesis of other secondary metabolites. The function redundancy index (FRI) of abundant bacteria (122) was lower than that of rare bacteria (8,878; Figure 5B), indicating that the probability of potential function loss of rare bacteria after the disturbance was lower than that of abundant bacteria.

Composition and distribution of ARGs

A total of 20 ARGs were detected in sediment samples of the Lhasa River, which including colistin (*mcr*-1, *mcr*-3, and *mcr*-7), beta-lactam ($bla_{\text{CTX-M-32}}$, bla_{CMY}, $bla_{\text{CTX-M}}$, and bla_{TEM}), aminoglycosides (*aad*A and *str*B), macrolide (*ere*A, *ere*B, and *mph*A), quinolones (*qnr*A, *qnr*B, and *qnr*S), sulfonamides (*sul*1, *sul*2, and *sul*3), and tetracycline (*tet*M and *tet*X) resistance genes, respectively (Figure 6A). However, $bla_{\text{NMD-1}}$, *arm*A, and *tet*A were not detected in any sediment sample. The total relative abundance of ARGs ranged from 4.60×10^{-3} to 1.72 copies per 16S rRNA, indicating that the ARGs were widely distributed in the sediment of the Lhasa River. The bla_{TEM} was the only ARG detected in all sediments and was the most abundant ARGs (mean relative abundance was 1.92×10^{-1} copies per 16S rRNA), followed by the *tet*M, *sul*1, and *aad*A. Besides, *aad*A, *str*B, *sul*1, *sul*2, and *tet*M were also detected in all sediments. Among the 10 sediment samples, the total relative abundance of ARGs at S6 was significantly higher than those of other locations. The total relative abundance of ARGs downstream from Lhasa (from S6 to S10) was significantly higher than those upriver from Lhasa (from S1 to S5), suggesting that human activities may promote the accumulation of ARGs in the sediments of the Lhasa River (Figure 6B).

Potential hosts and co-occurrence patterns of ARGs in sediment

Network analysis showed that more members of rare bacteria (36.5%) were the potential host of ARGs (Figure 7A). An ARG may have more potential hosts, such as the potential hosts of *mcr*-7 belonging to both abundant and rare bacteria. However, network topology features showed that abundant bacteria rather than rare bacteria had stronger connectivity and centrality, indicating that abundant bacteria may be the potential host of more ARGs (Supplementary Figure S2). Besides, the relative abundance of abundant bacteria (15.5%) in the whole bacterial community was higher than that of rare bacteria (0.57%; Supplementary Figure S3). This also suggested that the abundant bacteria were the main potential host of ARGs. The relative abundance of ARGs and their potential hosts downstream was higher than upstream, suggesting that urbanization may promote the occurrence of ARGs and their potential hosts.

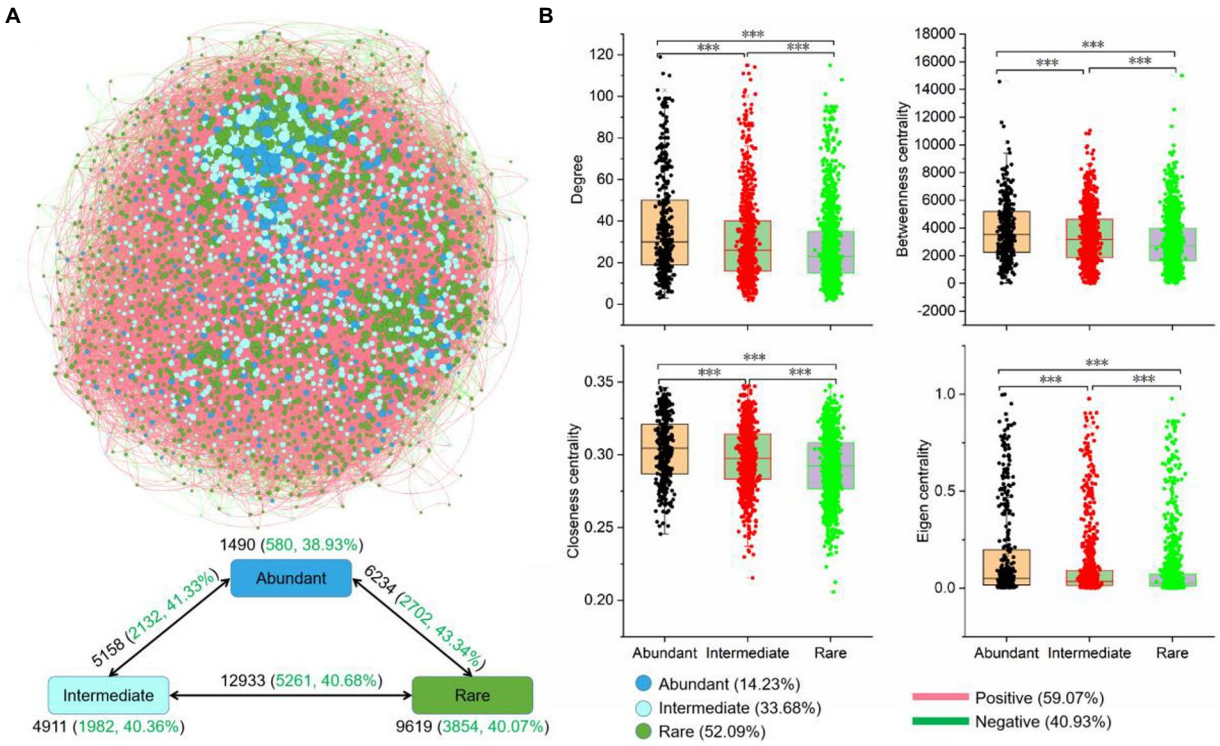


FIGURE 4
Co-occurrence network of abundant and rare bacteria in the sediment of the Lhasa River. **(A)** The network analysis showed the intra-associations within each bacterial taxa and inter-association between different bacterial taxa. OTUs occurred in more than half of samples were used for network analysis. A connection based on a strong ($|r|>0.8$) and significant ($p<0.01$) correlation *via* Spearman. The size of each node is proportional to the degree. Numbers outside and inside parentheses represent total edge numbers and negative edge numbers and their ratio, respectively. **(B)** Comparison of node-level topological features among three different bacterial taxa. Asterisks denote significance (***$p<0.001$).

Furthermore, the network results showed coexistence patterns among some ARGs, such as bla_{TEM} had a significant correlation with the *str*B, *mph*A, *sul*3, and *tet*M. More important, *transposase gene tnpA* had a significant correlation with the *aad*A, *str*B, *ere*A, *ere*B, *qnr*S, *sul*1, *sul*2, and *tet*X. *int*I1 had a significant correlation with the *aad*A, *str*B, *ere*B, *qnr*S, *sul*1, *sul*2, and *tet*X. These results suggested that some ARGs (*aad*A, *str*B, *ere*B, *qnr*S, *sul*1, *sul*2, and *tet*X) in the sediment of the Lhasa River may co-exist on MGEs, which may increase the risk of transmission of these ARGs in aquatic ecosystems.

Discussion

Bacterial communities are the foundation of every ecosystem on earth. They are often composed of abundant bacterial taxa with fewer species and rare bacteria with more species (Pedrós-Alió, 2012; Hou et al., 2020; Zhang et al., 2021b). Research on abundant and rare bacteria has expanded our understanding of bacterial community structure, but the relationships of abundant and rare bacteria with ARGs remain largely unclear. Revealing the dominant host bacteria (e.g., abundant or rare bacteria) of ARGs and assembly processes provide a good hold on the potential risks of ARGs. Therefore, we investigated the composition of abundant and rare bacteria and their relationship with ARGs. We also characterized the ecological assembly mechanism of sedimental abundant and rare bacteria in the Lhasa River, China.

Rare taxa can serve as function insurance of sediment bacterial communities

In this study, rare bacteria with low relative abundance were found to have high richness and diversity (Figure 1A), which was consistent with findings of the studies on the sediment of Erhai Lake (Zhang et al., 2021b) and Hangzhou Bay (Dai et al., 2016). Although rare bacteria did not dominate the taxonomic community, they may still play an important role in maintaining the bacterial community's stability in the Lhasa River sediment due to their large taxonomic pool. Because the more members of the rare bacteria, the stronger the buffering effect of their functional composition on environmental variation (Schindler et al., 2010). Previous studies showed that functional redundancy could protect microbial communities by maintaining ecosystem function homeostasis (Liang et al., 2020). Our study found that rare bacteria had stronger functional redundancy than abundant ones (Figure 5B). Besides, rare bacteria with high functional redundancy show a stronger adaptation to anthropogenic

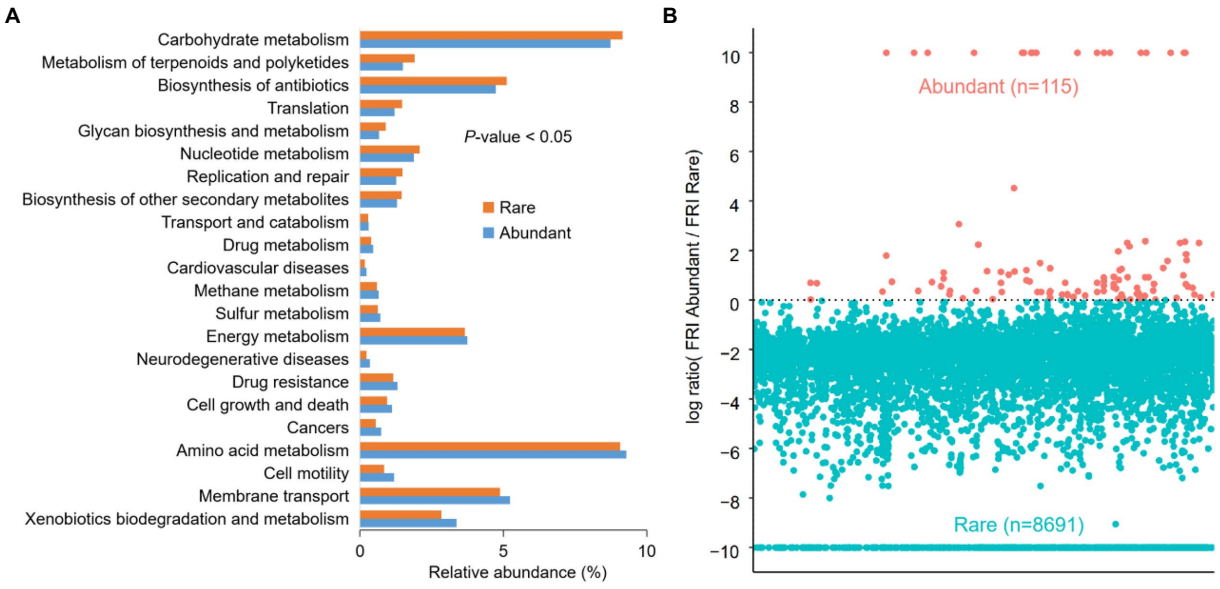


FIGURE 5
Comparison of functional differences **(A)** and functional redundancy **(B)** between rich and rare groups in sediments of the Lhasa River.

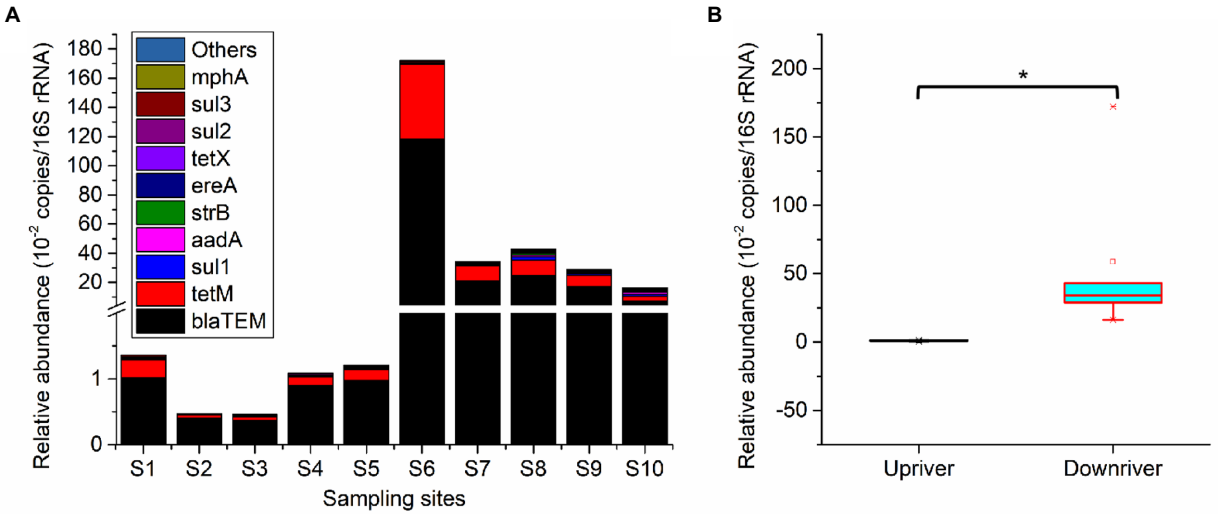


FIGURE 6
Main composition of the sedimental ARGs of the Lhasa River. Upriver including the sediment samples from sampling site S1 to S5, and downriver including the sediment samples from sampling site S6 to S10. "*" showed a significant difference at the 0.05 level.

disturbances (Wan et al., 2021b). Thus, rare bacteria can serve as an insurance source for the function of sediment bacterial communities in the Lhasa River during external disturbance.

Biogeographical patterns of abundant and rare bacteria in the sediment

Our study found obvious differences in diversity, taxonomy composition, and phylogenetic composition between abundant and rare bacteria. Even the persistent abundant and rare bacteria showed different biogeographical patterns. The number of spatial persistence existing OTUs in abundant bacteria was outdistanced than that in rare ones, which was consistent with the spatial persistence existing pattern of rare and abundant bacteria in the sediment from Erhai Lake (Zhang et al., 2021b). Furthermore, this study found that the community similarity of abundant bacteria stands out from rare bacteria (Figure 2A), suggesting the species composition of rare bacteria was more susceptible to geographical and environmental filters. Some studies found that the stronger

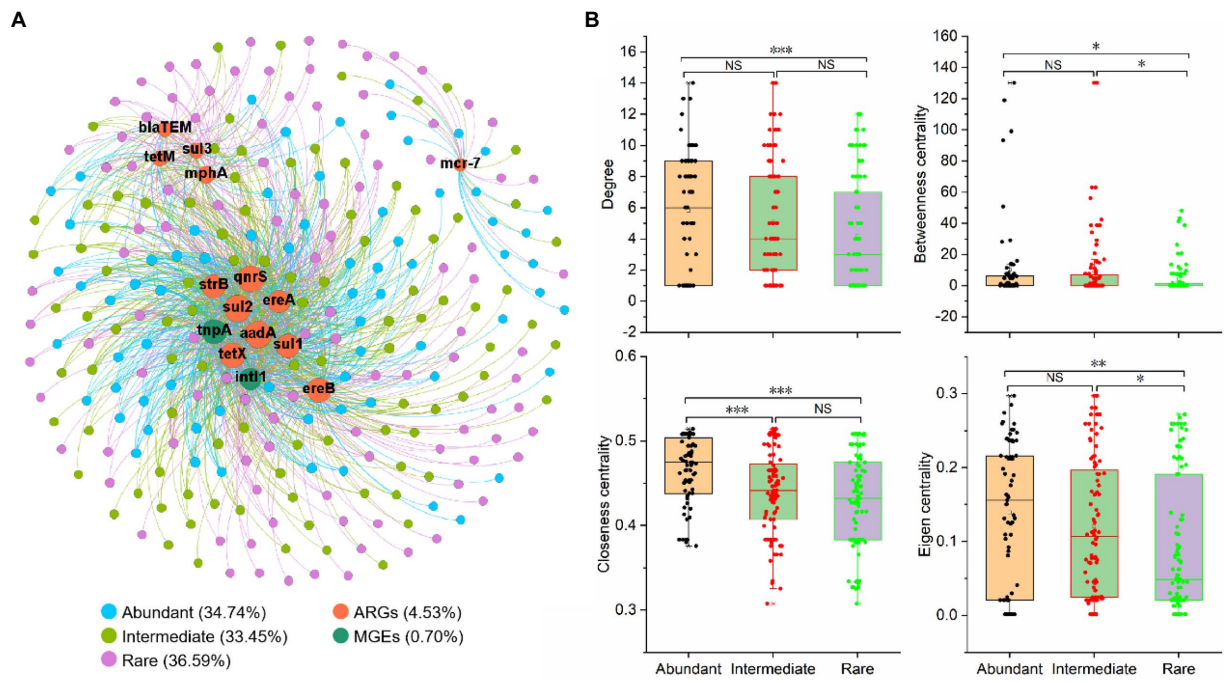


FIGURE 7
Co-occurrence patterns of ARGs and their potential hosts in the sediment of the Lhasa River. **(A)** Network analysis showed the co-occurrence patterns of ARGs and their potential hosts. The percentages were of these taxa OTUs or genes that accounted for total OTUs or genes in networks. OTUs and ARGs occurred in more than half of the samples were used for network analysis. A connection based on a strong (r>0.8) and significant (p<0.01) correlation *via* Spearman. **(B)** Major network topological properties of co-occurrence patterns of ARGs and their potential hosts. Asterisks denote significance (NS, $p \geq 0.05$; *p<0.05; **p<0.01; and ***p<0.001).

spatial variation within the microbial community, the more susceptible it is to environmental change (Wan et al., 2021b; Zhang et al., 2021b). It also may be foreshadowed that rare bacteria in sediments from the Lhasa River were more susceptible to environmental changes.

Geographic distance and environmental heterogeneity are abiotic factors that govern bacterial community assemblage (Gao et al., 2019; Langenheder and Lindström, 2019). This study found that abundant and rare bacteria had complicated responses to geographical and environmental differences. The spatial turnover of the bacterial community has been reported to be related to dispersal limitations (Wang et al., 2013; Lewthwaite et al., 2017). Our study found that rare and abundant bacteria had a more significant variation in horizontal spatial distribution because their community similarity decreased with the increase in geographical distance, which was similar to the biogeographical pattern of river microorganisms reported previously (Chen et al., 2020a,b; Wang Y. et al., 2020). This study also found that altitude was another important spatial factor that affects the community turnover of rare bacteria in the sediment. The R^2 value of DDR showed that the geographical distance, altitude distance, and environmental heterogeneity had greater effects on rare bacteria than on abundant bacteria, showing that rare bacteria rather than abundant bacteria were more susceptible to environmental changes. The slope of DDR showed that the effects of altitude and environmental heterogeneity on the spatial turnover of rare bacteria surpass abundant bacteria. These results portend that geographic and environmental factors together shaped the unique biogeographic pattern of sediment abundant and rare bacteria in the Lhasa River. This also means the greater impact of geographic and environmental factors on rare bacteria resulting in the community similarity of rare bacteria far below abundant bacteria.

Potential associations of bacterial communities

Association among the microbe-microbe is an essential biotic factor in the assembly processes of microbial communities except for the abiotic factor (geographic and environmental selection; Nemergut et al., 2013; Zhang et al., 2021b). In the community ecological assembly processes, network analysis could provide new insights into the associations within individual bacterial taxon and linkages between different bacterial taxa (Xue et al., 2018; Zhang et al., 2021b). The nodes in networks with high connectivity may play a crucial role in protecting the structural stability of the bacterial community (Xue et al., 2018). This study found that the connection within the abundant bacteria significantly overtopped rare bacteria, indicating that abundant bacteria may play an irreplaceable role in maintaining bacterial community structure. Furthermore, the positive interaction links in the network are mainly considered cooperative relationships among microbial members, while the

negative interaction links are mainly thought of as competitive relationships among microbial members (Faust et al., 2012). Cooperation among bacterial members helps improve the resilience of bacterial communities to respond to changing environments (Xue et al., 2018). Cooperation among the abundant bacteria was more than among rare bacteria, which may be an important reason for the widespread of abundant bacteria in the sediment of the Lhasa River.

Deterministic process was the dominant assembly process in abundant bacterial taxa

Traditional niche theory generally agrees that deterministic process mediated community structure is governed by species interaction (e.g., competition and mutualisms, etc.) and environmental variables (e.g., pH and temperature, etc.; Fargione et al., 2003; Zhou and Ning, 2017), whereas neutral theory assumes that community structure is shaped by limited dispersal and random fluctuations in species abundance (e.g., birth, death, and extinction, etc.; Chave, 2004; Zhou and Ning, 2017). Although deterministic and stochastic were generally accepted that occur simultaneously in the community assembly processes, their relative contribution to regulating community structure and biogeography is debatable (Zhou and Ning, 2017). This study showed that both deterministic and stochastic processes occur during the assembly of abundant and rare bacteria, which was consistent with previous studies (Zhou and Ning, 2017; Jiao and Lu, 2020b; Wan et al., 2021b; Zhang et al., 2021b). Among them, the deterministic process was the dominant assembly mechanism of abundant bacterial taxa, while the stochastic process was the dominant assembly mechanism of rare bacteria. One possible reason was that the high diversity of rare bacteria species allows them to occupy various ecological niches, while more rare species occurs spatial turnover in biogeographic distribution, which leads to the strong influence of stochastic processes on the assembly of rare bacterial taxa (Hou et al., 2020). Similarly, more persistent species from the abundant bacterial taxa were detected in the Lhasa River sediment, which may be one reason why the assembly processes of the abundant bacteria were more inclined to deterministic processes. Furthermore, environmental and spatial variables also seemed to control the biogeographic patterns and assembly of abundant and rare bacteria. Null model results showed that the variable selection was the dominant assembly process of abundant bacteria, followed by dispersal limitation. Conversely, the dispersal limitation was the main assembly process of rare bacteria, followed by variable selection. Moreover, the Mantel tests of geospatial and environmental factors against βNTI values also suggested that the βNTI values of abundant bacteria had a significant correlation with more geospatial and environmental factors (such as latitude, pH, conductivity, TN and TC ratio, and Cd), whereas βNTI of rare bacterial taxa only had a significant correlation with latitude. This may be a decisive reason why the abundant bacteria were more influenced by variable selection than rare bacteria.

Urbanization increased the occurrence of ARGs in the Lhasa River sediment

Antibiotic resistance genes were widely distributed in the sediment of the Lhasa River, among which bla_{TEM}, *aad*A, *str*B, *sul*1, *sul*2, and *tet*M were detected with 100%. Notably, bla_{TEM} was detected in all sediment samples with the highest abundance, consistent with bla_{TEM} in the surface sediments of Danjiangkou Reservoir (Jiang et al., 2021b). The bla_{TEM} was the clinically relevant ARG, which can be used as an indicator gene for ARG contamination caused by human activities (Narciso-da-Rocha et al., 2014; Stange et al., 2019). This also indicates that sediments of the Lhasa River were contaminated by ARGs related to human activities. Human activities have increased correspondingly with the decrease of the altitude of the Lhasa River, which also leads to the increase in the abundance of ARGs in the sediments. In contrast to a global survey that found that urbanization was strongly associated with lower rates of antibiotic resistance (Collignon et al., 2018), studies have reported that urbanization could promote the development of bacterial resistance to antibiotics in rivers (Peng et al., 2020; Liu et al., 2021). This study found that urbanization promoted the enrichment of ARGs, which was consistent with the results found in the Yarlung Tsangpo River (Liu et al., 2021). Therefore, more attention should be paid to the pollution of ARGs caused by urbanization on the watershed scale.

Abundant bacterial taxa were the main potential hosts of ARGs

Bacterial antibiotic resistance is one of the most serious global threats to environmental safety and human health (Ashbolt et al., 2013; Roca et al., 2015; Zainab et al., 2020). *Cyanobacteria* were found to be a reservoir and source of ARGs (Wang Z. et al., 2020), which contributes to the diversity increase of ARGs in the aquatic ecosystem (Zhang et al., 2020). In our study, the relative abundance of *Cyanobacteria* was second only to that of *Proteobacteria* among the abundant bacterial taxa, but not in rare bacterial taxa. Previous studies also found that *Proteobacteria*, *Bacteroidetes*, and *Actinobacteria* were often antibiotic producers or have the ability to transform/metabolize (Manaia, 2017). *Proteobacteria*, *Bacteroidetes*, and *Actinobacteria* were the dominant phyla in both abundant and rare bacteria. Furthermore, abundant bacteria account for the highest proportion in whole bacterial communities, indicating abundant bacteria may be the main hosts of ARGs in the sediments from the Lhasa River. Function prediction results show that abundant bacteria not only had a strong pathogenic potential for human diseases but also had a strong potential for drug resistance. Environment ARGs could threaten human health by increasing pathogenic ARB, leading to inefficient or ineffective use of therapeutic antibiotics for humans (Pan et al., 2020; Jiang et al., 2021a). To date, we still know little about how spatial variation in ARGs composition relates to bacterial taxonomic composition (i.e., abundant bacteria or rare

bacteria) in a river continuum. Therefore, we further explored the relationship of the ARGs with abundant bacteria and rare bacteria based on understanding the biogeographic patterns of ARGs.

The co-occurrence network is also a widely used as an important tool to explore the interaction between ARGs and their potential hosts (Chen et al., 2016; Peng et al., 2020). Important nodes in a network can be identified by network central location and high connectivity (Peng et al., 2020). In this study, the network analysis showed that the connectedness between the abundant bacteria and ARGs was higher than the rare ones (Figure 7B), indicating that the abundant bacteria may be the potential hosts of more ARGs. The relative abundance of these potential hosts belonging to abundant bacteria was also higher than that of rare bacteria. This may also be an important reason for the strong potential drug resistance in the abundant bacterial taxa. More importantly, ARGs in the environment lead to the rapid increase in the spread and number of ARB through horizontal gene transfer, which will also make antibiotic resistance an important and unavoidable global health problem affecting human health (Watts et al., 2017; Manyi-Loh et al., 2018; Kraemer et al., 2019; Yadav and Kapley, 2021). This study found that abundant bacteria not only had strong potential drug resistance but also have a high abundance of potential ARGs. There was a strong and significant correlation between these ARGs and MGEs, which could increase the ecological risk of abundant taxa and the potential for human disease.

Some potential limitations merit further discussion. Our analyses were focused on the abundant and rare bacterial taxa level, and we did not know the exact host bacteria for each ARG at the species level. Network results on the co-occurrence patterns between ARGs and bacterial taxa indicated the possible host information of ARGs. Therefore, further studies are needed to verify the ARG bacterial hosts at the species level. Abundant bacteria had a high abundance of metabolic pathways of potential drug resistance in all predicted functional genes may be due to their higher relative abundance in the whole community, but it also does not mean that rich taxa contain more ARGs than rare taxa. Therefore, further studies are needed to verify the ARG bacterial hosts at the species level.

Conclusion

In this study, we investigated biogeographical patterns and assembly mechanisms of rare and abundant bacteria and revealed the potential association among the ARGs with abundant and rare bacteria from the Lhasa River sediment. The different and complex responses of abundant and rare bacteria to geospatial and environmental changes may be influenced mainly by deterministic and stochastic processes, respectively. Rare taxa can serve as function insurance of bacterial communities in the Lhasa River sediment. This shall provide novel insights to explain the assembly and biogeographical patterns of abundant and rare bacteria in the sediment. To our knowledge, this study was the first time to reveal that the abundant bacteria have a high abundance of potential ARGs in Plateau Rivers, with strong pathogenic potential for human diseases. In particular, abundant bacteria with potential ARGs were also maybe the main potential hosts for the presence of MGEs, which may increase the ecological risks of abundant bacteria. These results provide new insights into understand the ARGs' association with abundant and rare bacteria in plateau river sediment. Given the importance of ARGs to the health of aquatic ecosystems, the findings of this study should be validated experimentally at the bacterial species level in more diverse freshwater and marine ecosystems.

Data availability statement

The data presented in the study are deposited in the Sequence Read Archive (SRA) repository (https://submit.ncbi.nlm.nih.gov/subs/sra/), accession number PRJNA681935.

Author contributions

YZ and WZ: conceptualization. YZ and YJ: methodology. YZ and HL: writing–original draft and validation. YL and HL: visualization. YL and YJ: software. WZ: writing–reviewing and editing. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the National Natural Science Foundation of China (Grant numbers 31860151 and 32171648).

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/fmicb.2022.977037/full#supplementary-material

References

Amarasiri, M., Sano, D., and Suzuki, S. (2020). Understanding human health risks caused by antibiotic resistant bacteria (ARB) and antibiotic resistance genes (ARG) in water environments: current knowledge and questions to be answered. *Crit. Rev. Environ. Sci. Technol.* 50, 2016–2059. doi: 10.1080/10643389.2019.1692611

Ashbolt, N., Amézquita, T., Backhaus, P., Borriello, K., Brandt, P., Collignon, A., et al. (2013). Human health risk assessment (HHRA) for environmental development and transfer of antibiotic resistance. *Environ. Health Perspect.* 121, 993–1001. doi: 10.1289/ehp.1206316

Caporaso, J., Lauber, C., Walters, W., Berg-Lyons, D., Lozupone, C., Turnbaugh, P., et al. (2011). Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc. Natl. Acad. Sci. U. S. A.* 108, 4516–4522. doi: 10.1073/pnas.1000080107

Chave, J. (2004). Neutral theory and community ecology. *Ecol. Lett.* 7, 241–253. doi: 10.1111/j.1461-0248.2003.00566.x

Chen, Q., An, X., Li, H., Su, J., Ma, Y., and Zhu, Y. (2016). Long-term field application of sewage sludge increases the abundance of antibiotic resistance genes in soil. *Environ. Int.* 645, 1230–1237. doi: 10.1016/j.scitotenv.2018.07.260

Chen, J., Wang, P., Wang, C., Wang, X., Miao, L., Liu, S., et al. (2020a). Fungal community demonstrates stronger dispersal limitation and less network connectivity than bacterial community in sediments along a large river. *Environ. Microbiol.* 22, 832–849. doi: 10.1111/1462-2920.14795

Chen, J., Wang, P., Wang, C., Wang, X., Miao, L., Liu, S., et al. (2020b). Distinct assembly mechanisms underlie similar biogeographic patterns of rare and abundant bacterioplankton in cascade reservoirs of a large river. *Front. Microbiol.* 11:158. doi: 10.3389/fmicb.2020.00158

Collignon, P., Beggs, J. J., Walsh, T. R., Gandra, S., and Laxminarayan, R. (2018). Anthropological and socioeconomic factors contributing to global antimicrobial resistance: a univariate and multivariable analysis. *Lancet Planet. Health* 2, e398–e405. doi: 10.1016/S2542-5196(18)30186-4

Cosgrove, S. E. (2006). The relationship between antimicrobial resistance and patient outcomes: mortality, length of hospital stay, and health care costs. *Clin. Infect. Dis.* 42, S82–S89. doi: 10.1086/499406

Dai, T., Zhang, Y., Tang, Y., Bai, Y., Tao, Y., Huang, B., et al. (2016). Identifying the key taxonomic categories that characterize microbial community diversity using full-scale classification: a case study of microbial communities in the sediments of Hangzhou Bay. *FEMS Microbiol. Ecol.* 92:fiw 150. doi: 10.1093/femsec/fiw150

Edgar, R. (2013). Uparse: Highly accurate otu sequences from microbial amplicon reads. *Nat. Methods*, 10, 996–998. doi: 10.1038/nmeth.2604

Fargione, J., Brown, C. S., and Tilman, D. (2003). Community assembly and invasion: an experimental test of neutral versus niche processes. *Proc. Natl. Acad. Sci. U. S. A.* 100, 8916–8920. doi: 10.1073/pnas.1033107100

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

Gao, Q., Yang, Y., Feng, J., Tian, R., Guo, X., Ning, D., et al. (2019). The spatial scale dependence of diazotrophic and bacterial community assembly in paddy soil. *Glob. Ecol. Biogeogr.* 28, 1093–1105. doi: 10.1111/geb.12917

Goldman, A. E., Graham, E. B., Crump, A. R., Kennedy, D. W., Romero, E. B., Anderson, C. G., et al. (2017). Biogeochemical cycling at the aquatic–terrestrial interface is linked to parafluvial hyporheic zone inundation history. *Biogeosciences* 14, 4229–4241. doi: 10.5194/bg-14-4229-2017

Hou, J., Wu, L., Liu, W., Ge, Y., Mu, T., Zhou, T., et al. (2020). Biogeography and diversity patterns of abundant and rare bacterial communities in rice paddy soils across China. *Sci. Total Environ.* 730:139116. doi: 10.1016/j.scitotenv.2020.139116

Jia, S., Shi, P., Hu, Q., Li, B., Zhang, T., and Zhang, X. (2015). Bacterial community shift drives antibiotic resistance promotion during drinking water chlorination. *Environ. Sci. Technol.* 49, 12271–12279. doi: 10.1021/acs.est.5b03521

Jiang, C., Diao, X., Wang, H., and Ma, S. (2021a). Diverse and abundant antibiotic resistance genes in mangrove area and their relationship with bacterial communities-a study in Hainan Island. *Environ. Pollut.* 276:116704. doi: 10.1016/j.envpol.2021.116704

Jiang, C., Pan, X., Grossart, H. P., Lin, L., Shi, J., and Yang, Y. (2021b). Vertical and horizontal distributions of clinical antibiotic resistance genes and bacterial communities in Danjiangkou reservoir, China. *Environ. Sci. Pollut. Res.* 28, 61163–61175. doi: 10.1007/s11356-021-15069-w

Jiao, S., and Lu, Y. (2020a). Abundant fungi adapt to broader environmental gradients than rare fungi in agricultural fields. *Glob. Chang. Biol.* 26, 4506–4520. doi: 10.1111/gcb.15130

Jiao, S., and Lu, Y. (2020b). Soil pH and temperature regulate assembly processes of abundant and rare bacterial communities in agricultural ecosystems. *Environ. Microbiol.* 22, 1052–1065. doi: 10.1111/1462-2920.14815

Jiao, S., Luo, Y., Lu, M., Xiao, X., Lin, Y., Chen, W., et al. (2017). Distinct succession patterns of abundant and rare bacteria in temporal microcosms with pollutants. *Environ. Pollut.* 225, 497–505. doi: 10.1016/j.envpol.2017.03.015

Kraemer, S. A., Ramachandran, A., and Perron, G. G. (2019). Antibiotic pollution in the environment: from microbial ecology to public policy. *Microorganisms* 7, 1–24. doi: 10.3390/microorganisms7060180

Langenheder, S., and Lindström, E. S. (2019). Factors influencing aquatic and terrestrial bacterial community assembly. *Environ. Microbiol. Rep.* 11, 306–315. doi: 10.1111/1758-2229.12731

Lewthwaite, J. M. M., Debinski, D. M., and Kerr, J. T. (2017). High community turnover and dispersal limitation relative to rapid climate change. *Glob. Ecol. Biogeogr.* 26, 459–471. doi: 10.1111/geb.12553

Liang, Y., Xiao, X., Nuccio, E. E., Yuan, M., Zhang, N., Xue, K., et al. (2020). Differentiation strategies of soil rare and abundant microbial taxa in response to changing climatic regimes. *Environ. Microbiol.* 22, 1327–1340. doi: 10.1111/1462-2920.14945

Liu, S., Wang, P., Wang, C., Wang, X., and Chen, J. (2021). Anthropogenic disturbances on antibiotic resistome along the Yarlung Tsangpo River on the Tibetan plateau: ecological dissemination mechanisms of antibiotic resistance genes to bacterial pathogens. *Water Res.* 202:117447. doi: 10.1016/j.watres.2021.117447

Long, S., Tong, H., Zhang, X., Jia, S., Chen, M., and Liu, C. (2021). Heavy metal tolerance genes associated with contaminated sediments from an e-waste recycling river in southern China. *Front. Microbiol.* 12:1134. doi: 10.3389/fmicb.2021.665090

Manaia, C. M. (2017). Assessing the risk of antibiotic resistance transmission from the environment to humans: non-direct proportionality between abundance and risk. *Trends Microbiol.* 25, 173–181. doi: 10.1016/j.tim.2016.11.014

Manyi-Loh, C., Mamphweli, S., Meyer, E., and Okoh, A. (2018). Antibiotic use in agriculture and its consequential resistance in environmental sources: potential public health implications. *Molecules* 23:795. doi: 10.3390/molecules23040795

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

Narciso-da-Rocha, C., Varela, A. R., Schwartz, T., Nunes, O. C., and Manaia, C. M. (2014). bla_{TEM} and van a as indicator genes of antibiotic resistance contamination in a hospital–urban wastewater treatment plant system. *J. Glob. Antimicrob. Resist.* 2, 309–315. doi: 10.1016/j.jgar.2014.10.001

Nemergut, D. R., Schmidt, S. K., Fukami, T., O'Neill, S. P., Bilinski, T. M., Stanish, L. F., et al. (2013). Patterns and processes of microbial community assembly. *Microbiol. Mol. Biol. Rev.* 77, 342–356. doi: 10.1128/MMBR.00051-12

Pan, X., Lin, L., Zhang, W., Dong, L., and Yang, Y. (2020). Metagenome sequencing to unveil the resistome in a deep subtropical lake on the Yunnan-Guizhou plateau, China. *Environ. Pollut.* 263:114470. doi: 10.1016/j.envpol.2020.114470

Pedrós-Alió, C. (2012). The rare bacterial biosphere. *Annu. Rev. Mar. Sci.* 4, 449–466. doi: 10.1146/annurev-marine-120710-100948

Peng, F., Guo, Y., Isabwe, A., Chen, H., Wang, Y., Zhang, Y., et al. (2020). Urbanization drives riverine bacterial antibiotic resistome more than taxonomic community at watershed scale. *Environ. Int.* 137:105524. doi: 10.1016/j.envint.2020.105524

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 webbased tools. *Nucleic Acids Res.* 41, D590–D596. doi: 10.1093/nar/gks1219

Qu, B., Zhang, Y., Kang, S., and Sillanpaa, M. (2019). Water quality in the Tibetan plateau: major ions and trace elements in rivers of the "water tower of Asia". *Sci. Total Environ.* 649, 571–581. doi: 10.1016/j.scitotenv.2018.08.316

Roca, I., Akova, F., Baquero, J., Carlet, M., Cavaleri, S., Coenen, J., et al. (2015). The global threat of antimicrobial resistance: science for intervention. *New Microbes and New Infections*. 6, 22–29. doi: 10.1016/j.nmni.2015.02.007

Rognes, T., Flouri, T., Nichols, B., Quince, C., and Mahe, F. (2016). VSEARCH: a versatile open source tool for metagenomics. *Peer, J.* 4:e2584. doi: 10.7717/peerj.2584

Schindler, D. E., Hilborn, R., Chasco, B., Boatright, C. P., Quinn, T. P., Rogers, L. A., et al. (2010). Population diversity and the portfolio effect in an exploited species. *Nature* 465, 609–612. doi: 10.1038/nature09060

Shao, M., Tang, X., Zhang, Y., and Li, W. (2006). City clusters in China: air and surface water pollution. *Front. Ecol. Environ.* 4, 353–361. doi: 10.1890/1540-9295(2006)004

Sogin, M. L., Morrison, H. G., Huber, J. A., Welch, D. M., Huse, S. M., Neal, P. R., et al. (2006). Microbial diversity in the deep sea and the underexplored "rare biosphere". *Proc. Natl. Acad. Sci. U. S. A.* 103, 12115–12120. doi: 10.1073/pnas.0605127103

Stange, C., Yin, D., Xu, T., Guo, X., Schafer, C., and Tiehm, A. (2019). Distribution of clinically relevant antibiotic resistance genes in Lake tai. *Sci. Total Environ.* 655, 337–346. doi: 10.1016/j.scitotenv.2018.11.211

Stegen, J. C., Lin, X., Fredrickson, J. K., Chen, X., Kennedy, D. W., Murray, C. J., et al. (2013). Quantifying community assembly processes and identifying features that impose them. *ISME J.* 7, 2069–2079. doi: 10.1038/ismej.2013.93

Walters, W., Hyde, E., Berg-Lyons, D., Ackermann, G., Humphrey, G., Parada, A., et al. (2016). Improved bacterial 16S rRNA gene (V4 and V4-5) and fungal internal transcribed spacer marker gene primers for microbial community surveys. *mSystems* 1, e00009–15. doi: 10.1128/mSystems.00009-15

Wan, W., Gadd, G. M., Yang, Y., Yuan, W., Gu, J., Ye, L., et al. (2021a). Environmental adaptation is stronger for abundant rather than rare microorganisms in wetland soils from the Qinghai-Tibet plateau. *Mol. Ecol.* 30, 2390–2403. doi: 10.1111/mec.15882

Wan, W., Grossart, H. P., He, D., Yuan, W., and Yang, Y. (2021b). Stronger environmental adaptation of rare rather than abundant bacterioplankton in response to dredging in eutrophic Lake Nanhu (Wuhan, China). *Water Res.* 190:116751. doi: 10.1016/j.watres.2020.116751

Wang, Z., Chen, Q., Zhang, J., Guan, T., Chen, Y., and Shi, W. (2020). Critical roles of cyanobacteria as reservoir and source for antibiotic resistance genes. *Environ. Int.* 144:106034. doi: 10.1016/j.envint.2020.106034

Wang, E., Li, Q., Hu, H., Peng, F., Zhang, P., and Li, J. (2019). Spatial characteristics and influencing factors of river pollution in China. *Water Environ. Res.* 91, 351–363. doi: 10.1002/wer.1044

Wang, Q., Garrity, G., Tiedje, J., and Cole, J. R. (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* 73, 5261–5267. doi: 10.1128/AEM.00062-07

Wang, S., Wang, X., Guo, H., Fan, W., Lv, H., and Duan, R. (2013). Distinguishing the importance between habitat specialization and dispersal limitation on species turnover. *Ecol. Evol.* 3, 3545–3553. doi: 10.1002/ece3.745

Wang, X., Wang, P., Wang, C., Chen, J., Hu, B., Liu, S., et al. (2021). Distinct strategies of abundant and rare bacterioplankton in river-reservoir system: evidence from a 2800 km plateau river. *Environ. Res.* 199:111418. doi: 10.1016/j.envres.2021.111418

Wang, Y., Ye, F., Wu, S., Wu, J., Yan, J., Xu, K., et al. (2020). Biogeographic pattern of bacterioplanktonic community and potential function in the Yangtze River: roles of abundant and rare taxa. *Sci. Total Environ.* 747:141335. doi: 10.1016/j.scitotenv.2020.141335

Watts, J., Schreier, H., Lanska, L., and Hale, M. (2017). The rising tide of antimicrobial resistance in aquaculture: sources, sinks and solutions. *Mar. Drugs* 15. doi: 10.3390/md15060158

Wemheuer, F., Taylor, J., Daniel, R., Johnston, E., Meinicke, P., Thomas, T., et al. (2020). Tax 4Fun2: prediction of habitat-specific functional profiles and functional redundancy based on 16S rRNA gene sequences. *Environ. Microb.* 15, 1–12. doi: 10.1186/s40793-020-00358-7

Xue, Y., Chen, H., Yang, J. R., Liu, M., Huang, B., and Yang, J. (2018). Distinct patterns and processes of abundant and rare eukaryotic plankton communities following a reservoir cyanobacterial bloom. *ISME J.* 12, 2263–2277. doi: 10.1038/s41396-018-0159-0

Yadav, S., and Kapley, A. (2021). Antibiotic resistance: global health crisis and metagenomics. *Biotechnol. Rep.* 29:e00604. doi: 10.1016/j.btre.2021.e00604

Yang, Y., Song, W., Lin, H., Wang, W., Du, L., and Xing, W. (2018). Antibiotics and antibiotic resistance genes in global lakes: a review and meta-analysis. *Environ. Int.* 116, 60–73. doi: 10.1016/j.envint.2018.04.011

Zainab, S., Junaid, N., Xu, N., and Malik, R. (2020). Antibiotics and antibiotic resistant genes (ARGs) in groundwater: A global review on dissemination, sources, interactions, environmental and human health risks. *Water Res.* 187:116455. doi: 10.1016/j.watres.2020.116455

Zhang, W., Suyamud, B., Lohwacharin, J., and Yang, Y. (2021a). Large-scale pattern of resistance genes and bacterial community in the tap water along the middle and low reaches of the Yangtze River. *Ecotox. Environ. Safe* 208:111517. doi: 10.1016/j.ecoenv.2020.111517

Zhang, W., Wan, W., Lin, H., Pan, X., Lin, L., and Yang, Y. (2021b). Nitrogen rather than phosphorus driving the biogeographic patterns of abundant bacterial taxa in a eutrophic plateau lake. *Sci. Total Environ.* 806:150947. doi: 10.1016/j.citotenv.2021.150947

Zhang, Y., Yao, P., Sun, C., Li, S., Shi, X., Zhang, X., et al. (2021). Vertical diversity and association pattern of total, abundant and rare microbial communities in deep-sea sediments. *Mol. Ecol.* 30, 2800–2816. doi: 10.1111/mec.15937

Zhang, Q., Zhang, Z., Lu, T., Peijnenburg, W., Gillings, M., Yang, X., et al. (2020). Cyanobacterial blooms contribute to the diversity of antibiotic-resistance genes in aquatic ecosystems. *Commun. Biol.* 3, 737–710. doi: 10.1038/s42003-020-01468-1

Zhou, J., and Ning, D. (2017). Stochastic community assembly: does it matter in microbial ecology? *Microbiol. Mol. Biol. Rev.* 81, e00002–e00017. doi: 10.1128/MMBR.00002-17

frontiers | Frontiers in Microbiology

TYPE Mini Review
PUBLISHED 10 November 2022
DOI 10.3389/fmicb.2022.958514

OPEN ACCESS

EDITED BY
Hongbin Liu,
Hong Kong University of Science and Technology, Hong Kong SAR, China

REVIEWED BY
Adèle Lazuka,
Veolia, France
Hamza Mbareche,
Laval University,
Canada

*CORRESPONDENCE
Zaheer Ahmad Nasir
z.a.nasar@cranfield.ac.uk

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

RECEIVED 31 May 2022
ACCEPTED 11 October 2022
PUBLISHED 10 November 2022

CITATION
Tian J, Yan C, Alcega SG, Hassard F, Tyrrel S, Coulon F and Nasir ZA (2022) Detection and characterization of bioaerosol emissions from wastewater treatment plants: Challenges and opportunities. *Front. Microbiol.* 13:958514. doi: 10.3389/fmicb.2022.958514

Detection and characterization of bioaerosol emissions from wastewater treatment plants: Challenges and opportunities

Jianghan Tian[1], Cheng Yan[2,3], Sonia Garcia Alcega[4], Francis Hassard[2,5], Sean Tyrrel[2], Frederic Coulon[2] and Zaheer Ahmad Nasir[2]*

[1]School of Chemistry, University of Bristol, Bristol, United Kingdom, [2]School of Water, Energy and Environment, Cranfield University, Cranfield, United Kingdom, [3]School of Environmental Studies, China University of Geosciences, Wuhan, China, [4]School of Physical Sciences, The Open University, Walton Hall, Milton Keynes, United Kingdom, [5]Institute for Nanotechnology and Water Sustainability, University of South Africa, Johannesburg, South Africa

Rapid population growth and urbanization process have led to increasing demand for wastewater treatment capacity resulting in a non-negligible increase of wastewater treatment plants (WWTPs) in several cities around the world. Bioaerosol emissions from WWTPs may pose adverse health risks to the sewage workers and nearby residents, which raises increasing public health concerns. However, there are still significant knowledge gaps on the interplay between process-based bioaerosol characteristics and exposures and the quantification of health risk which limit our ability to design effective risk assessment and management strategies. This review provides a critical overview of the existing knowledge of bioaerosol emissions from WWTPs including their nature, magnitude and size distribution, and highlights the shortcoming associated with existing sampling and analysis methods. The recent advancements made for rapid detection of bioaerosols are then discussed, especially the emerging real time detection methods to highlight the directions for future research needs to advance the knowledge on bioaerosol emissions from WWTPs.

KEYWORDS

bioaerosol, wastewater treatment plant, real time detection, emission characteristics, *in situ* measurement

Introduction

Rapid population growth and urbanization increased industrial process and other economic activities are all contributing to significant increase of wastewater generation, which is associated with a growing number of wastewater treatment plants (WWTPs) in cities. These WWTPs are often in close proximity to residential areas (Mateo-Sagasta et al., 2015; IWA and OFID, 2018; OECD, 2019). Wastewater treatment generally undergoes

pre-treatment, secondary treatment and advance treatment, while sludge is normally treated by dewatering process followed by stabilization prior to being used or disposed of in a WWTP (Korzeniewska, 2011; Mateo-Sagasta et al., 2015). Activated sludge and bio-membrane methods within WWTP infrastructures have been widely applied to treat wastewater, which take advantage of the aeration process during treatment.

Raw wastewater contains various microorganisms, such as bacteria, virus, fungi, and some of them are pathogenic. During the aeration process, these microorganisms can aerosolize from wastewater to the air forming bioaerosols. These can impact on human health by way of inhalation or ingestion, leading to a range of health effects (allergenicity, toxicity, and infectivity) especially in sewage workers and nearby communities (Gerardi and Zimmerman, 2005; Fracchia et al., 2006; Korzeniewska et al., 2008; Kim et al., 2018) as well directly impacting regional air quality (Upadhyay et al., 2013). Moreover, operations involving mechanical agitation can accelerate the dispersal of bioaerosols (Papke and Ward, 2004). Thus, WWTPs are considered to be an important source of bioaerosol emissions (Michalkiewicz, 2019). Figure 1 illustrates the bioaerosol dynamics from different typical sources at WWTPs and describes how various factors affect aerosolization, transport, dispersal, decay, exposure and potential health effects from emissions at wastewater treatment plants. This information is vital to understanding the processes and factors affecting the WWTPs aerobiology to inform the development and implementation of control measures.

A wide range of microbial species are found within bioaerosol emissions from WWTPs, including heterotrophic/mesophilic bacteria, fungi, and pathogenic organisms (Korzeniewska, 2011), as well as their metabolic products such as endotoxins (1,3)-β-D-glucan molecules and mycotoxins (Garcia-Alcega et al., 2017; Mbareche et al., 2019). Moreover, norovirus and adenovirus are significant viral species among bioaerosol emissions from WWTPs (Masclaux et al., 2014). Furthermore, actinomycetes and fecal coliform bacteria have also been investigated (Vítězová et al., 2012; Li et al., 2013). Since the start of COVID-19 pandemic, there has been growing interest in understanding the fate of SARS-CoV-2 virus in WWTPs and potential transmission through fecal-oral and fecal-inhalation routes (Foladori et al., 2020; Heller et al., 2020), and most importantly, the WWTP associated bioaerosols was considered to hold the potential as the early indicator for future pandemics (Singh et al., 2021).

Air samples collected around mechanical and biological treatment sites in WWTPs have shown microorganisms less than 2 μm in size (Bauer et al., 2002; Kowalski et al., 2017; Hsiao et al., 2020). These fine size fractions of bioaerosols have longer airborne residence time enabling their long-distance transport from their sources and can potentially reach the alveolar region of the respiratory tract. Specifically, pathogenic bacteria, fungi and yeast such as Citrobacter, Enterobacter, Klebsiella, Serratia, Pantoea were identified as being in inhalable range (Filipkowska et al., 2002, 2007, 2008; Korzeniewska et al., 2007, 2008, 2009).

Since the late 1970s, it was reported that several typical symptoms were frequently found on the WWTP's employees, named 'sewage worker's syndrome' (Rylander et al., 1976; Clark, 1987; Hung et al., 2010; Stellacci et al., 2010). The main characteristics of those symptoms are malaise, fatigue, weakness, headache, dizziness, acute rhinitis, fever, respiratory diseases (Hung et al., 2010; Stellacci et al., 2010). Studies have suggested that there was a strong connection between 'sewage worker's syndrome' and bioaerosol emissions (Cyprowski and Krajewski, 2003; Patentalakis et al., 2008). Nonetheless, further investigation will enhance the evidence base to strengthen the association between casual agents of sewage worker's exposure and ill health outcomes (Van Hooste et al., 2010; Mbareche et al., 2022). Consequently, concurrent with a growing number of WWTPs, there have been increasing concerns over public and environmental health risks associated with bioaerosol emissions from WWTPs, and the risk assessment of bioaerosols from WWTPs became significant to regulations of workplace safety and community public health (Peccia et al., 2008). However, the evidence base on nature and magnitude of bioaerosol emissions from WWTPs and potential public health impacts remains inconclusive primarily due to methodological constraints associated with diverse sampling and analysis methods,

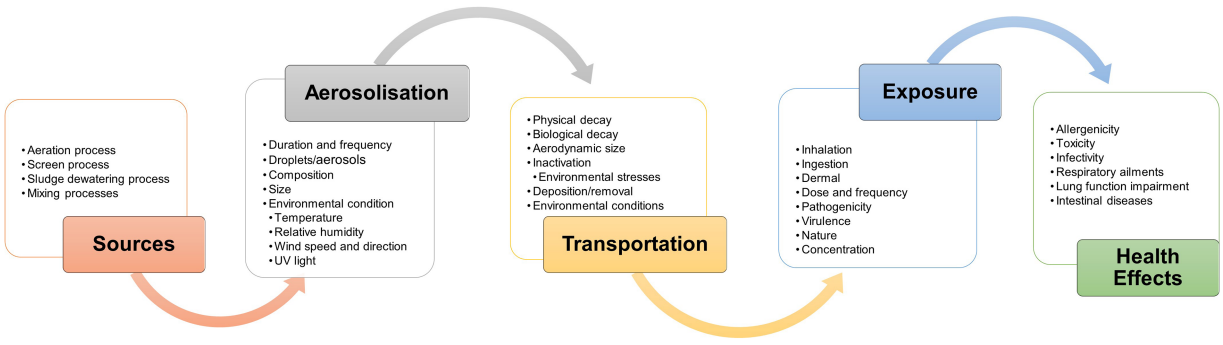


FIGURE 1
Factors influencing dynamics of bioaerosol and potential health effects from emissions at wastewater treatment plants.

insufficient dose–response data, and varied health endpoints (Tian et al., 2020; Mbareche et al., 2022). Hence, both operators and regulators are facing considerable challenges to devise proportionate risk-based operation and policies to permit efficient management of potential risks to human and environmental health.

This paper aims to review the existing state of knowledge on emission characteristics of bioaerosols from WWTPs, with a view to highlight the shortcoming associated with existing methods, discuss the recent advancements and identify future directions for research to advance the knowledge on bioaerosols emissions from WWTPs.

State of the art

Overall, the scientific literature on bioaerosol emissions from WWTPs is limited. A review of the literature was conducted to examine the progress on detection and characterization methods of bioaerosols in WWTPs, using major citation databases (Scopus, Web of Science, Google scholar). The studies were grouped according to different sampling methods and analytical approach. Table 1 provides a brief overview of representative studies conducted to understand the nature, magnitude, and size distributions of different analytes during varied sampling duration.

Sampling and analytical methods

To assess the risk of bioaerosols from WWTPs, different sampling and analytical methods have been applied for the qualitative and quantitative studies of bioaerosols. Major sampling methods involve bioaerosols collection through impaction, impingement, filtration, and cyclone followed by a range of post collection analysis methods. Broadly, those methods can be divided into culture-based and culture-independent methods, and the advantages and disadvantages of different sampling and analytical approaches have been intensively discussed in literatures (Mainelis, 2020; Kathiriya et al., 2021). Majority of studies on emission characteristics of bioaerosols were based on cultivation methods (Brandi et al., 2000; Bauer et al., 2002; Karra and Katsivela, 2007; Vítězová et al., 2012; Li et al., 2013; Niazi et al., 2015; Katsivela et al., 2017; Kowalski et al., 2017; Michalkiewicz, 2019). Typical processes of culture-based methods entail three stages including (i) sampling, (ii) incubation and (iii) enumeration. During the sampling stage, viable airborne microorganisms are collected either by impaction, filtration or impingement and transferred onto the culture medium (e.g., agar, depends on the targeted microorganisms; Korzeniewska, 2011; Mainelis, 2020). After sample collection, colonies of bacteria and fungi are incubated on a defined solid media and temperature for a period ranging between 2 and 7 days. The concentration of bioaerosols is then determined by counting colony formed and expressed as colony-forming units (CFUs) per 1 m^3 (Korzeniewska, 2011; Mainelis, 2020). The culture-based methods for bioaerosols are relatively sensitive and widely used in bioaerosols quantification (Douwes et al., 2003), however, they have some disadvantages such as low repeatability, relying on the incubation conditions and poor time resolution. Further to this and most importantly, there is a vast number of viable but non-culturable (VBNC) microorganisms in the environment, and therefore culture-based methods are leading to a significant underestimation of the actual viable bioaerosol concentrations in air samples. In order to improve quantitative and qualitative analysis of bioaerosols, several culture-independent methods have been developed, such as staining method, immunoassay method, molecular method (e.g., polymerase chain reaction, PCR; Carducci et al., 2000; Ranalli et al., 2000; Orsini et al., 2002; Thorn et al., 2002; Pascual et al., 2003; Cyprowski et al., 2011; Masclaux et al., 2014; Mbareche et al., 2017; Ferguson et al., 2019; Liu et al., 2020). These new methods, specifically that utilize DNA/RNA based approaches along with Next Generation Sequencing (NGS) technologies showed great capability in improving the understanding of identities, distribution, abundance, diversity and function of airborne microbial communities in WWTPs. (Liang et al., 2012; Brisebois et al., 2018; Han et al., 2019; Corpuz et al., 2020; Kabir et al., 2020; Han et al., 2020b; Bhardwaj et al., 2021; Kathiriya et al., 2021; Mbareche et al., 2022). To conclude, the existing evidence base on bioaerosol emission from WWTPs stems from a range of sampling and post-collection analysis methods ranging from culture-based methods to immunoassay and advanced molecular methods. Whilst these have advanced knowledge of bioaerosol emission from WWTPs, harmonization of sampling methods, sampling design and analytic methods is lacking. A decision tree framework will help to evaluate the relevance and utility of different sampling and analytical methods to a specific endpoint.

Nature and magnitude of bioaerosol emissions

In a study by Niazi et al. (2015), the results showed that *Bacillus*, *Staphylococcus* spp., and *Micrococcus* spp. were the most frequently observed bacteria types in the bioaerosols emitted from WWTPs, while the dominant fungi species were *Cladosporium* spp. and *Penicillium* spp. Similarly, Michalkiewicz et al. (2011) pointed out that *Corynebacterium*, *Bacillus* spp., *Staphylococcus* spp., *Pseudomonas aeruginosa* and *Micrococcus* spp. were the prevalent bacteria in their study and these potentially pathogenic infectious bacteria can pose a serious hazard to onsite workers an nearby communities. Another study done by Breza-Boruta and Paluszak (2007) showed that *Pseudomonas* were the predominant bacteria species. The most occurring species recorded by Kowalski et al. (2017) were Gram-positive cocci and non-sporing Gram-positive rods. Moreover, some by-products of airborne microorganisms' agents such as (1–3)-β-D glucans and bacterial endotoxin are being measured because of their toxic potency, immunological and allergic reactions through exposure (Thorn et al., 2002; Cyprowski et al.,

TABLE 1 Emission characteristics of bioaerosol emissions from wastewater treatment plants (WWTPs).

References	Place	Sampling site	Main analytes	Method	Instrument	Sampling duration	Analytical approach	Key findings
Bauer et al. (2002)	Vienna, Austria	Two WWTPs (an activated sludge plant and a fixed-film reactor)	Cultivable bacteria and fungi	Filtration	(Glass manifold) sterile cellulose nitrate filters (Sartorius, 47 mm diameter, 0.45 mm pore size)	About 5 min each	Cultivation	1. Activated sludge plant have a higher concentration of bacterial and fungal aerosols compared to the fixed-film reactor; 2. Majority of particles were smaller than 2.0 μm (respirable range); 3. It was suggested that the surface area of the aeration tank should be decreased.
Brandi et al. (2000)	Northwest Adriatic coast, Italy	Near the wastewater aeration tanks	Bacteria and fungi	Impaction and impingement	The Andersen Six-Stage Viable Particle Sampler, SAS (Surface Air System) impactor and All Glass Impinger	20 min each	Cultivation	1. Mechanical agitation of the sludge generates aerosols with high concentrations of bacteria and fungi; 2. Aerobic digestion with a submerged microbubble system seems to pose little risk from airborne transmission of pathogenic bacteria and fungi to waste-treatment workers and local residents.
Karra and Katsivela (2007)	Crete, Greece	Screens, aerated grit chambers (indoor), primary settling tanks, (partially covered) primary and secondary settling tanks, aeration tanks, chlorination and sludge processors	Mesophilic heterotrophic bacteria (total coliforms, fecal coliforms and Enterococci) and fungi	Impaction	Air sampler MAS 100 (Merck, Germany)	(Did not mention)	Cultivation	1. The highest concentration of bioaerosols were observed at aerated grit chambers (pre-treatment); 2. A gradual decrease of bioaerosol emissions was observed during the advanced wastewater treatment from the pre-treatment to the primary, secondary and tertiary treatment.
Katsivela et al. (2017)	Chania (Crete), Greece	A WWTP (aerated grit removal and primary sedimentation processes)	Bacteria and fungi	Impaction	Andersen six-stage viable particle sampler (Thermo ESM Andersen Instruments GmbH, Germany)	(Did not mention)	Cultivation	1. Proper wastewater treatment is efficient to the reduction of bioaerosol emissions thus decrease the possibility of associated health risks; 2. Findings on significant positive linear relationships between hydrogen sulphide and size-fractionated bacterial aerosol concentrations (and mass concentrations of different particulate matter fractions) may support the possible drift of some heterotrophic bacteria and particulate matter from the wastewater to the atmosphere by the release of the hydrogen sulphide gas.

(Continued)

10.3389/fmicb.2022.958514

TABLE 1 (Continued)

References	Place	Sampling site	Main analytes	Method	Instrument	Sampling duration	Analytical approach	Key findings
Kowalski et al. (2017)	Upper Silesia, Poland	Pre-treatment, biological treatment and activated sludge post-processing stages at five different WWTP	Viable airborne bacteria and fungi	Impaction	6-stage Andersen impactor	8 min each for Anderson impactor	Cultivation	1. The highest concentration of bacterial aerosol was found in the activated sludge post-processing and mechanical purifying stages, which was probably caused by the close environment of the sampling (indoor); 2. Dominant size of airborne bacteria: 2.1–3.3 μm (clarifiers and the sludge post-processing); 3.3–4.7 μm (mechanical treatment and aeration tanks).
Li et al. (2013)	Xi'an City, China	Aerated grit chamber, oxidation ditch, secondary settling tank, sludge dewatering house	Airborne viable bacteria, fungi and actinomycetes	Impaction	Andersen six-stage impactor	10 min each for the impactor (three repetitions each time)	Cultivation	1. The area of mechanical agitation of sludge (aeration and sludge thickening operations) is the largest emission source of bacteria and actinomycete aerosols; 2. Majority of bacteria, fungi and actinomycete aerosols were in respirable size range (less than 3.3 μm), which have high potential to cause adverse health effects.
Michalkiewicz (2019)	Wielkopolska Region, Poland	9 WWTPs	Bacteria and fungi	Sedimentation and impaction methods	MAS 100 Eco type air sampler (Merck)	(Did not mention)	Cultivation	1. Coliform bacteria (an important factor to be monitored) reflects the level of air pollution with bioaerosols from sewage; 2. The higher humidity favored the occurrence of a higher abundance of microorganisms.
Niazi et al. (2015)	Tehran, Iran	Four operational units and one background site	Bacteria and fungi	Impaction	QuickTake 30 sample pump and Bio Stage single-stage cascade impactor (SKC, United States)	2 min each, 240 samples in total	Cultivation. Identification: biochemical test for bacteria; microscopic method for fungi	1. Highest air pollution happened in warm season; 2. Significant relationship between environmental parameters and concentrations of bacterial and fungal were observed.
Vítězová et al. (2012)	Czech Republic	I: Near the screen (sand catcher); II: in the middle of activation tank; III: near area for dewatered sludge	Heterotrophic bacteria, fungi, actinomycetes and fecal coliform bacteria	Impaction	MAS-100 (Merck, Germany)	(Did not mention)	Cultivation	The vicinity of sand catcher and activation tanks are at highest risk due to the presence of higher concentration of pathogenic mesophilic bacteria (includes fecal coliform bacteria).

(Continued)

10.3389/fmicb.2022.958514

TABLE 1 (Continued)

References	Place	Sampling site	Main analytes	Method	Instrument	Sampling duration	Analytical approach	Key findings
Brisebois et al. (2018)	Eastern Canada	Screening, primary and secondary screening, biofiltration, grit/fats, oils and greases (FOGs) removal unit, and secondary decantation at four WWTPs	Human pathogen viruses	Impaction and filtration	Coriolis®μ (Bertin Technologies, Montigny-le-Bretonneux, France), Marple Personal Cascade Impactor (Thermo Fisher Scientific, Waltham, United States) and SASS 2300 sampler (Research International, Washington, United States)	6-h shift for Coriolis®μ, 5 h for Marple impactor, 100 min for SASS 2300 sampler	qPCR (molecular method) and viral metagenomic approach	1. Four viral pathogens of particular concern were tested; 2. Human DNA viruses were in much greater relative abundance than viral RNA from bioaerosols from WWTPs.
Carducci et al. (2000)	Livorno, Italy	1. An activated sludge plant; 2. An anaerobic sludge treatment plant; 3. A sewage washing station	Bacteria and virus	Impaction	SAS impactor	(Did not mention)	Cultivation and RT-PCR (molecular method)	1. Highest bioaerosol contamination was found at sewage washing station (high risk); 2. Enteric virus were frequently presented in the bioaerosol emissions, and isolated enteroviruses of human origin represent a significant health hazard, as they can be transmitted by the respiratory route and are known to cause often-serious extra-intestinal pathologies.
Han et al. (2020b)	Beijing, China	A municipal A^2/O WWTP	Total bacteria, bacteria with *Enterobacteriaceae*, *Staphylococcus aureus*, and *Pseudomonas aeruginosa*.	Impaction	Anderson six sampler	(Did not mention)	Cultivation and MiSeq high-throughput sequencing	1. The bioaerosol generated in the bubble bottom aeration is dominated by <3.3 μm in size, while the brush surface aeration is the opposite; 2. Capping or sealing of major treatment sections can effectively reduce the emission of bioaerosols from WWTP to the surrounding environment.
Han et al. (2019)	Beijing, China	The aerobic tank of a municipal WWTP, which uses an anaerobic/anoxic/oxic process	Submicron aerosols (SA)		A particulate matter sampler (TH-PM1.0–100, TianHong, Wuhan, China)	24 h	MiSeq sequencing, high-throughput sequencing-based metagenomic analysis	1. The unimodal SA size distribution ranges largely from 68 to 350 nm; 2. Antibiotic resistant genes in SA may lead to high risk to human.

(Continued)

10.3389/fmicb.2022.958514

TABLE 1 (Continued)

References	Place	Sampling site	Main analytes	Method	Instrument	Sampling duration	Analytical approach	Key findings
Liu et al. (2020)	Shijiazhuang City, China	A municipal WWTP and a pharmaceutical WWTP	Pathogenic and inhalable bacterial aerosol (PM2.5, PM10)	Filtration	Medium-flow continuous filtration air particulate samplers	44 h	Miseq sequencing and PCR technique (molecular/ DNA-based approach)	1. Indoor sludge dewatering facilities were the significant sources of bacterial aerosol emissions, and the process was dominated by aerosolization rather than dispersion; 2. Outdoor aerosols have higher similarity to ambient aerosols than wastewater, revealed that dispersion is more important than aerosolization process.
Masclaux et al. (2014)	Zürich, Switzerland	31 WWTPs	Airborne virus (adenovirus, norovirus and the hepatitis E virus)	(Gelatine) filtration	Gelatine filters embedded in standard cassettes (SKC, Inc. Eighty-Four, United States)	At least 1 h	qPCR	1. Adenovirus was present in 100% of summer WWTP samples and 97% of winter samples; 2. Concentrations of potentially pathogenic viral particles in WWTP air are non-negligible and could partly explain the work-related gastrointestinal symptoms often reported in employees in this sector.
Mbareche et al. (2022)	Quebec, Canada	Screening, grit/FOGs removal, settling tank, and biofiltration at eight WWTPs	Endotoxins, total cultivable and gram-negative bacteria and pathogens indicators	Impaction	Six-stage Anderson Impactor (Thermo Fisher Scientific, Waltham, MA, United States), SASS 3100 (Research International, Inc., Monroe, WA, United States)	(Did not mention)	qPCR	1. The screening, grit/FOGs removal and biofiltration were the most bioaerosol-loaded sites; 2. Water temperature is an important factor for microbial activity – increasing the rate of air changes per hour in summer would be beneficial to reduce the concentration of bioaerosols.
Orsini et al. (2002)	Italy	WWTP in a university hospital	Viable bacteria	Impaction	Impacting sampler (surface air system (SAS), PBI)	(Did not mention)	Cultivation, bio-chemical test and AP-PCR (molecular method)	An integrated molecular approach can contribute in studying bacteria air-diffusion and personnel contamination.
Pascual et al. (2003)	Spain	WWTP	Molds and yeasts, total and fecal coliforms	Impaction	Air Sampler MAS 100 impactor (Merck)	(Did not mention)	Cultivation and PCR (molecular method)	1. Heavy bioaerosol pollution was observed at pre-treatment and primary settlers; 2. Wind speed and daily inflow at the WWTP were two important factors.

(Continued)

10.3389/fmicb.2022.958514

TABLE 1 (Continued)

References	Place	Sampling site	Main analytes	Method	Instrument	Sampling duration	Analytical approach	Key findings
Ranalli et al. (2000)	Como, Italy	WWTP	Bacteria	Sedimentation and impaction	SAS (Surface Air Systems, PBI) viable sampler, Sartorius MD8	60 min	Cultivation and PCR (molecular method)	The highest bacterial emission rates (total aerobic heterotrophic and fecal coliform bacteria) were registered during the preliminary mechanical treatments.
Thorn et al. (2002)	Sweden	Three WWTP in different cities (sampling sites including pump stations, sludge hall, flocculoreaction room, indoor sedimentation basins etc.)	Airborne bacterial endotoxin	Filtration	Isopore filters (ATTP 0.8um; Millipore, Cambridge, MA). For personal samplers: Gil-Air 3 SC, Gilian personal air sampler, Gilian Instrument Corp., NJ, United States	(Did not mention)	Immunoassay	1. The highest endotoxin concentrations were found around agitation practices of wastewater; 2. The highest endotoxin values were found at worksites located indoors.
Cyprowski et al. (2011)	Poland	A WWTP in a metal industry	(1,3)-β-glucans	Filtration	25-mm-glass-fiber filters GF/F (Whatman, International Ltd., Maidstone, UK) and Button Aerosol Sampler (SKC Inc., Eighty-Four, PA, United States)	(Did not mention)	Immunoassay	Soluble and insoluble (1,3)-β-glucan fractions were suggested to be analyzed in occupational environment.
Hsiao et al. (2020)	Taiwan, China	Aeration area of an urban wastewater treatment plant	Fluorescent particles	*In-situ*	UV-APS	Quadruplicate over 60 min	Fluorescence-based method	1. On a number basis, the full PSD (particle size distribution) exhibits a unimodal distribution and is dominated by nanoparticles (contribute 99% of total concentration); 2. On a volume basis, the distribution exhibits dual peaks, and the mode sizes are 0.5–0.7 μm and 2–3 μm; 3. The volume concentration of submicron particles is 2 times that of supermicron particles; 4. Submicron particles are major contributors to both the number and volume concentration of particles in the WWTP. 5. Most fluorescent particles may be bacterial aggregates or fungal species.

(Continued)

10.3389/fmicb.2022.958514

TABLE 1 (Continued)

References	Place	Sampling site	Main analytes	Method	Instrument	Sampling duration	Analytical approach	Key findings
Li et al. (2016)	Beijing, China	7 intra-plant sites within a WWTP	1. Viable bacteria and fungi 2.Fluorescent particles	1. Impaction 2. *In-situ*	1. Reuter Centrifugal Sampler High Flow (RCS) (Biotest, Inc.) 2. UV-APS (ultraviolet aerodynamic particle sizer)	30 min	Fluorescence-based method and PCR	1. The sludge thickening basin and the screen room were the dominant emission sources of bacterial and fungal aerosols; 2. The UV-APS results showed that for most sampling site, the peak of the number size distribution (or number concentration) is in the range of 3–4 μm.
Tian et al. (2020)	Cranfield, UK	A WWTP within a university campus	Fluorescent particles	*In-situ*	Spectral Intensity Bioaerosol Sensor (SIBS)	Around 3–3.5 h (five repeated measurement)	Fluorescence-based method	1. Particle number concentrations were highly variable; 2. The on-site activity affected the particle size distribution and number concentrations; 3. Predominant particles were in fine size scale (less than 1 μm).

2011). Additionally, viruses were also investigated as the pathogenic viral particles in WWTP air that can partly explain the work-related symptoms (Masclaux et al., 2014; Brisebois et al., 2018; Corpuz et al., 2020). Additionally, mesophilic heterotrophic bacteria (total coliforms, fecal coliforms, and enterococci), *Escherichia coli* (or *E.coli*) and staphylococci were monitored in the surrounding air at different stages of wastewater treatment since they are the indicators of fecal pollutants in the wastewater (Michalkiewicz, 2019).

Bioaerosol emissions were found to exist in every stage of wastewater treatment, and highly variable during different stages of wastewater treatment (Li et al., 2013; Michalkiewicz, 2019; Han et al., 2020a), with a concentration from 10^2 to 10^4 CFU/m^3 for viable bacterial and fungal aerosols. Moreover, mechanical agitation, aeration tanks and pre-treatment section were generally considered to be the highest concentrations of bioaerosols (Li et al., 2016; Kowalski et al., 2017; Michalkiewicz, 2019). Though many studies have been done to establish the risk assessment procedure at WWTP such as quantitative microbiological risk assessment (QMRA), however, there is no uniform standards for assessing concentrations of airborne microorganisms on the workplace especially for WWTPs. At present, there is only limited regulations about the critical values people were exposed to bioaerosol concentrations. In the United Kingdom Environment Agency (2017), published a guidance about the environmental monitoring strategy of bioaerosols at regulated facilities proposing the use of culture based methods and sample collection by impaction and filtration. Michalkiewicz (2019) pointed out that there are no limit values for microorganisms and endotoxins in the air on worksite in Poland. Overall, the bioaerosol emissions from WWTPs are diverse (including bacteria, fungi, viruses, and secondary metabolites), and concentrations are highly variable depending upon multiple operational and meteorological variables affecting the timing, intensity, spatial extent and duration of emissions.

Size distribution of bioaerosol emissions

Dominant size ranges of airborne bacteria and fungi ranged between 2.1 and 3.3 μm at clarifiers and sludge post-processing stages, and between 3.3 and 4.7 μm near mechanical treatment and aeration tanks (Kowalski et al., 2017). Li et al. (2013) revealed that majority of bacteria, fungi and actinomycete aerosols were in respirable size range (less than 3.3 μm), which have high potential to cause adverse health effects. Recent work by Hsiao et al. (2020) further indicated that submicron particles are major contributors to both the number and volume concentration of particles in the WWTP.

Bioaerosol particle morphological characteristics such as size distribution, surface area and asymmetry factor (AF, i.e., shape/aspect ratio, or sphere/rod) are also central to understanding emissions and downwind dispersal from source of particles (Vestlund et al., 2014). These morphological

characteristics, especially size distribution, highly affect bioaerosol particles behavior and are important factors in predicting their dispersal (Madelin and Johnson, 1992). For example, the deposition rates for bioaerosols and non-biological particles are a function of particle size, rather than the nature of the particle (Kaye et al., 2000; Ho, 2002; Könemann et al., 2019). However, the particle size distributions of bioaerosol emissions are affected by multiple complex mechanisms as the particles will disintegrate into smaller fragments or single spores due to the disturbance/agitation activities of release mechanisms or during the sampling campaign (Madelin and Johnson, 1992), In addition, the emission characteristics of bioaerosols vary with time and process. For instance, the dominant bioaerosol particle size generated in the bubble bottom aeration is dominated by particles less than 3.3 μm in size while the brush surface aeration is particles larger than 3.3 μm (Han et al., 2020b). Whilst the knowledge on size distribution of bioaerosol emissions from different operational stages of WWTPs is limited, the available evidence suggests the dominance of respirable size fraction. This has implications for transport, dispersion, exposure, and resultant health impacts.

Factors responsible for bioaerosol emissions characteristics

Numerous studies indicated that the concentration of bioaerosols in WWTPs were influenced by the sampling location (Breza-Boruta and Paluszak, 2007), type of wastewater, aeration method, climatic conditions, wastewater treatment equipment, sunlight, wind speed, and relative humidity (Karra and Katsivela, 2007; Michalkiewicz et al., 2011). Specifically, the concentration of fungal aerosols was largely found at pre-treatment, primary treatment and grit chamber stages (Pascual et al., 2003; Kim et al., 2010; Niazi et al., 2015). Michalkiewicz et al. (2011) also reported that several key factors influenced bioaerosols emission from WWTP including: (a) turbulence and tremor in wastewater, (b) wind speed and direction and wind effect level, and (c) rainfall. In contrast to mechanical aeration, the diffusion aeration system undergoes less turbulence (Niazi et al., 2015). There was a significant relationship between environmental parameters and concentrations of bacterial and fungal bioaerosols. Among the different meteorological conditions recorded by Niazi et al. (2015), significant correlations were found between bacteria concentrations and temperature, and fungal concentrations and relative humidity in air. Similarly, Jones and Harrison (2004) reported that temperature and water availability will affect the size and growth of the bioaerosol source material. For instance, the water availability is critical to stimulate the release of fungal spores. Similarly, wind speed can affect the bioaerosol concentration through atmospheric mixing and removal of biological materials from surfaces. In general, a range of wastewater treatment processes along with meteorological and topographical conditions affect the emissions, fate and behavior of bioaerosols.

Key barriers to advance bioaerosols detection and characterization

As summarized in Table 1, a range of sampling/collection and analysis methods have been applied for the identification and quantification of bioaerosols. Each method has its own advantages and disadvantages. Depending on different analytes and health endpoint, appropriate sampler and analysis approach should be chosen to improve the efficiency of bioaerosols collection for subsequent analysis (Haddrell and Thomas, 2017). However, there are major limitations that impeded our understanding of bioaerosol emissions from WWTPs. For both culture-based and culture-independent methods, the collection efficiency and bio efficiency of different sampling methods are variable and relatively low (Pöhlker et al., 2012). This can largely affect the viability, cultivability, size, and representativeness of sampled particles, which leads to the underestimation of the bioaerosols and limit our understanding of bioaerosol emissions. Besides, those methods can only provide snapshot data with low temporal resolution, which is difficult to capture the true nature and magnitude of bioaerosol emissions from WWTPs (Mainelis, 2020; Šantl-Temkiv et al., 2020). Additionally, poor repeatability (affected by the culture environment, e.g., temperature, relative humidity), selectivity to certain species, and labor intensiveness are also the disadvantages of cultivation method (Griffiths and DeCosemo, 1994; Heidelberg et al., 1997; Chi and Li, 2007; Korzeniewska, 2011).

To summarize, those existing methods are limited in providing the information on the size, fate and behavior of bioaerosol particles. Therefore, there is a need for improving the understanding of the temporal variation of the nature, magnitude and size distribution regarding to different processes in WWTPs, which can greatly contribute to the risk analysis modelling thus improving public health applications and management.

Real time detection and characterization of bioaerosol emissions

In recent years, significant technological advancements have been made to develop rapid detection and characterization methods for bioaerosols, such as flow cytometry in conjunction with fluorescent technique (FCM/FL), laser induced breakdown spectroscopy (LIBS), laser/light-induced fluorescence (LIF), biochemistry and molecular biology analysis, aerosol mass spectrometry (MS) focusing on physical, chemical, and biological characterization of bioaerosols (Chen and Li, 2007; Steele et al., 2008; Ghosh et al., 2015; Druckenmüller et al., 2017; Fennelly

et al., 2017; Nasir et al., 2018; Zhang et al., 2019; Negron et al., 2020; Šantl-Temkiv et al., 2020; Bhardwaj et al., 2021). Typically, each method offers different information on the complex mixture of atmospheric bioaerosols. For instance, LIBS and MS provides elemental composition of the particles in comparison to direct analysis of biochemical composition by different biochemistry and molecular biology analysis. Whilst LIBS and MS can rapidly record the elemental composition of single particles, detection and discrimination of biological materials is limited. Similarly, biochemical analysis can offer high sensitivity and selectivity but these are labor intensive and provide data with low temporal resolution. However, among these techniques, fluorescence spectroscopy has shown great potential to detect and broadly classify bioaerosols non-destructively in real time. Instruments based on LIF and/or elastic scattering have shown their capability and utility to detect and characterize bioaerosols in real-time in a range of ambient environments and sources (Nasir et al., 2019).

Briefly, the LIF based instruments interrogates the characteristic intrinsic fluorescence emission of particles and record the size, shape, and fluorescence spectra of single particles with high time resolution. These have been deployed into both laboratory and field studies. For example, Ultraviolet Aerodynamic Particle Sizer (UV-APS), Wideband Integrated Bioaerosol Sensor (WIBS) series (Healy et al., 2014; Oconnor et al., 2014; Perring et al., 2015; Crawford et al., 2016), Rapid-E (Šikoparija, 2020), Swisens Poleno (Sauvageat et al., 2020) and Spectral Intensity Bioaerosol Sensor (SIBS), have been employed in the bioaerosol monitoring in different environments (Nasir et al., 2018, 2019).

The LIF based devices present spectral resolution in the form of excitation-emission matrix (EEM; Pöhlker et al., 2012). There are three-excitation wavelength bands, which are the most commonly, used for distinguishing bioaerosol particles (Huffman and Santarpia, 2017). The excitation at approximately 255–285 nm band has been utilized to distinguish certain amino acid residues (Pöhlker et al., 2012). The excitation at approximately 340–370 nm band has been shown to promote fluorescence from the ubiquitous biological coenzyme biofluorophore NADH (Kaye and Hirst, 2011). Some LIF devices also use 450 nm to excite riboflavin and a variety of flavoproteins in bioaerosol particles (Pan, 2015).

The use of dual-wavelength excitation (e.g., 285 or 280 nm, and 370 nm) and the measurement of fluorescence in broad emission detection bands along with size and shape of single particles in real time is most prevailing approach. However, it is difficult to discriminate between different types of bioaerosols based on broad fluorescence emission detection bands (Pöhlker et al., 2012). The detection systems comprised of dual wavelength excitation and generating highly resolved spectral information of single particles in real time have been developed to overcome this challenge. For instance, SIBS is able to provide highly resolved spectral information (measuring fluorescence emission spectra in 16 wavelength bands) as compared to broad emission bands in the WIBS and other LIF based devices (Nasir et al., 2018, 2019; Könemann et al., 2019). Table 2 presents a comparison of parameters for different LIF based bioaerosol detection systems.

The application of these in field measurements in WWTPs are rare. Li et al. (2016) and Hsiao et al. (2020) both employed UV-APS for the detection of fluorescent particles and reported that the peak concentration of fluorescent particles was centered at 3–4 μm, which suggests the contribution from bacterial aggregates or fungal spores. Whilst, Tian et al. (2020) found the predominant size range is 0.5–1 μm by utilizing a SIBS. This is probably due to the differences in the excitation and emission bands. The UV-APS has 355 nm excitation wavelength and records emission at 420–575 nm to detect fluorescent particles in comparison to dual wavelength excitation and multi-channel emission measurements of SIBS. The highly resolved fluorescence spectra provided by SIBS may have the potential to elucidate the contribution of bioaerosols to total particles, the impact of various processes specific activities and the biological materials associated with airborne particles at WWTPs. Such investigations of bioaerosol emissions characteristics from WWTPs are of great value to better understand the size distribution and composition of emission profiles. This is critical for advancing and improving modelling methods to simulate dispersion of bioaerosols from WWTPs and the resultant health and environmental impacts.

To summarize, the high temporal resolution of size and number concentration data could greatly enhance the understanding of fate and behavior of bioaerosols, and likely deposition in the human respiratory tract. In combination with dose response studies, it has potential to better unravel the health effects of bioaerosols. Concurrently, high time resolution data could improve the understanding of transient emission dynamics, diurnal and annual cyclical variability of bioaerosol emissions, thus informing management strategies for bioaerosol emissions.

Challenges for real time detection and characterization of bioaerosols

Whilst LIF based instruments have the potential to instantaneously detect airborne biological materials and provide an overall contribution of bioaerosols to total particles, there are still challenges with regards to confidence in their ability to discriminate biological and non-biological particles and categories biological particles. To begin with, fluorescent particles are not equivalent to viable bioaerosols, though live unculturable microorganisms can be detected, however the dead but morphologically intact microorganisms can also be quantified by fluorescence-based methods (Burdsall et al., 2021). Similarly, fluorescence from interfering non-biological compounds also make the discrimination of particles challenging. Hence, the development of more effective fluorescence threshold strategies will enable to filter the interfering particles and maximize the proxy for bioaerosols. There is also a need for developing standard reference materials for fluorescence measurements and intercomparison of LIF based measurements.

TABLE 2 Comparison of parameters for different LIF based bioaerosol detection systems.

	SIBS	WIBS-5	WIBS-4A	Rapid-E
Measured parameters	1. Particle size 2. Asphericity 3. Fluorescence spectra	1. Particle size 2. Asphericity 3. Integrated fluorescence in three channels	1. Particle size 2. Asphericity 3. Integrated fluorescence in three channels	1. Particle size 2. Particle shape (scattering images) 3.Fluorescence spectra
Particle size range	~ 0.5–30 μm	~ 0.5–30 μm	~ 0.5–31 μm	0.5–100 μm
Fluorescence excitation (λ_{ex})	285 and 370 nm	280 and 370 nm	280 and 370 nm	320 nm
Fluorescence emission bands (λ_{em})	λ_{mean} = 302–721 nm (16-channels)	310–400 nm and 420–650 nm	310–400 and 420–650 nm	350–800 nm (32 channels)

Following on, to enhance selectivity and discrimination of bioaerosol emissions, the highly resolved fluorescence intensity measurements (such as with SIBS) can help to discriminating between different biological particle types depending on their biofluorophore signatures due to significantly resolved spectral resolution. However, the key limitation is the information provided by resolved emission spectra requires meaningful interpretation (Könemann et al., 2019; Nasir et al., 2019). The assignment of fluorescence to specific biological fluorophores within atmospheric particles is challenging due to the complexity of the molecular environment (Pan, 2015) and the overlap of mixed signals from different fluorophores (e.g., mineral dust, polycyclic aromatic hydrocarbons, humic-like interferences; Pöhlker et al., 2012). Advanced data analysis ecosystems need to be developed to improve the discrimination and processing speed of analysis along with lab studies to improve certainty and validation of assigning spectral signatures to atmospherically relevant biological fluorophores.

Conclusion and future directions

Overall, bioaerosol emissions from WWTPs presents a multicomponent/heterogeneity in nature, and highly variable in magnitude and particle size distribution with reference to different wastewater treatment unit and processes, which is of increasing health concern of nearby residential and occupational settings along with the urbanization process and population growth.

The advancements in real-time detection methods have shown the potential to overcome the methodological constraints enabling to detect and characterize the temporal variability of bioaerosols. However, long-term field investigations will help to better understand the dynamics of bioaerosol emissions from WWTPs and efficiently resolve public health issues relating to wastewater and bioaerosols. Depending on the treatment methods and type of wastewater, the categories and level of bioaerosol emissions could vary greatly. Hence, it is vital to conduct laboratory studies with biofluorophores, biological and non-biological particles in conjunction with other complementary methods (such as fluorimetry, gas chromatography, molecular methods) as well as *in situ* measurement to develop and validate library/network, database, and selective assignment of spectral responses to bioaerosol classes, for better discriminating bioaerosol particles and identifying the hazards with regards to various emission sources. In addition, studies on inter-technique comparability of fluorescence-based detection with advanced molecular methods will help to evaluate their reliability and utility to inform bioaerosol risk management at WWTPs. Simultaneously significant progress is required in the development of data analytics for optimizing fluorescence thresholds and discrimination between different biological particles.

Moreover, long-term *in situ* measurements will help to probe the temporal variation in the concentration and size distribution of bioaerosols to inform regulations concerning with occupational exposure and take measures to effectively mitigate bioaerosol emissions (e.g., avoid exposing to the time period with higher concentrations of emissions). A holistic system approach, involving multiple disciplines such as environmental science, engineering, microbiology, aerosol science, toxicology and epidemiology will pave the way to advance the knowledge on the factors and mechanisms influencing bioaerosol emissions and resultant health impacts. Process-based quantitative microbial risk assessment (QMRA) and dose response studies will also advance the evidence base on the exposure risk of bioaerosols released from WWTPs. Additionally, studies focusing on investigating the interaction of bioaerosols with abiotic components including other air pollutants will help to better understand the transformation, and the governing influences on viability, toxicity, and infectivity during airborne transport. This will inform the studies on exposure assessment and mechanism of toxicity enhancing the evidence base to better understand the dose–response relationships.

Advancement in detection and characterization of bioaerosols is critical in gaining insights into physicochemical and biological characteristics of emissions from WWTPs and elucidating their impacts in the context of public health (allergenicity, toxicity, and infectivity). This new knowledge will underpin the development of proportionate risk assessment and management policies and strategies to protect public health while ensuring the development of wastewater treatment infrastructure.

Data availability statement

No new data were generated or analysed during this study.

Author contributions

JT: investigation, formal analysis, methodology, visualization, and writing—original draft. CY: investigation, methodology, and writing—review and editing. SA: review and editing. FH: methodology and writing—review and editing. ST: funding acquisition and writing—review and editing. FC: funding acquisition, supervision, and writing—review and editing. ZN: conceptualization, supervision, methodology, administration, resources, and writing—review and editing. All authors contributed to the article and approved the submitted version.

Funding

This paper received the financial support from the UK Research and Innovation (UKRI) Natural Environment Research Council (NERC) through the Environmental Microbiology and Human Health Programme (Grant references NE/M01163/1 and NE/M010961/1) and the UKRI Strategic Priorities Fund (SPF) Clean Air Programme (Grant reference NE/V002171/1). CY's academic visit at Cranfield University was supported by the National Natural Science Foundation of China (Grant reference 51608497).

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

Bauer, H., Fuerhacker, M., Zibuschka, F., Schmid, H., and Puxbaum, H. (2002). Bacteria and fungi in aerosols generated by two different types of wastewater treatment plants. *Water Res.* 36, 3965–3970. doi: 10.1016/S0043-1354(02)00121-5

Bhardwaj, J., Hong, S., Jang, J., Han, C. H., Lee, J., and Jang, J. (2021). Recent advancements in the measurement of pathogenic airborne viruses. *J. Hazard. Mater.* 420:126574. doi: 10.1016/j.jhazmat.2021.126574

Brandi, G., Sisti, M., and Amagliani, G. (2000). Evaluation of the environmental impact of microbial aerosols generated by wastewater treatment plants utilizing different aeration systems. *J. Appl. Microbiol.* 88, 845–852. doi: 10.1046/j.1365-2672.2000.01024.x

Breza-Boruta, B., and Paluszak, Z. (2007). Influence of water treatment plant on microbiological composition of air bioaerosol. *Polish J. Environ. Stud.* 16, 663–670.

Brisebois, E., Veillette, M., Dion-Dupont, V., Lavoie, J., Corbeil, J., Culley, A., et al. (2018). Human viral pathogens are pervasive in wastewater treatment center aerosols. *J. Environ. Sci. (China)* 67, 45–53. doi: 10.1016/j.jes.2017.07.015

Burdsall, A. C., Xing, Y., Cooper, C. W., and Harper, W. F. (2021). Bioaerosol emissions from activated sludge basins: characterization, release, and attenuation. *Sci. Total Environ.* 753:141852. doi: 10.1016/j.scitotenv.2020.141852

Carducci, A., Tozzi, E., Rubulotta, E., Casini, B., Cantiani, L., Rovini, E., et al. (2000). Assessing airborne biological hazard from urban wastewater treatment. *Water Res.* 34, 1173–1178. doi: 10.1016/S0043-1354(99)00264-X

Chen, P. S., and Li, C. S. (2007). Real-time monitoring for bioaerosols - flow cytometry. *Analyst* 132, 14–16. doi: 10.1039/b603611m

Chi, M. C., and Li, C. S. (2007). Fluorochrome in monitoring atmospheric bioaerosols and correlations with meteorological factors and air pollutants. *Aerosol Sci. Technol.* 41, 672–678. doi: 10.1080/02786820701383181

Clark, C. S. (1987). Potential and actual biological related health risks of wastewater industry employment. *J. Water Pollut. Control Fed.* 59, 999–1008.

Corpuz, M. V. A., Buonerba, A., Vigliotta, G., Zarra, T., Ballesteros, F., Campiglia, P., et al. (2020). Viruses in wastewater: occurrence, abundance and detection methods. *Sci. Total Environ.* 745:140910. doi: 10.1016/j.scitotenv.2020.140910

Crawford, I., Lloyd, G., Herrmann, E., Hoyle, C. R., Bower, K. N., Connolly, P. J., et al. (2016). Observations of fluorescent aerosol-cloud interactions in the free troposphere at the high-altitude Research Station Jungfraujoch. *Atmos. Chem. Phys.* 16, 2273–2284. doi: 10.5194/acp-16-2273-2016

Cyprowski, M., and Krajewski, J. A. (2003). Harmful agents in municipal wastewater treatment plants. *Med. Pr.* 54, 73–80.

Cyprowski, M., Sowiak, M., and Szadkowska-Stańczyk, I. (2011). B (1→3)-Glucan aerosols in different occupational environments. *Aerobiologia (Bologna).* 27, 345–351. doi: 10.1007/s10453-011-9201-7

Douwes, J., Thorne, P., Pearce, N., and Heederik, D. (2003). Bioaerosol health effects and exposure assessment: Progress and prospects. *Ann. Occup. Hyg.* 47, 187–200. doi: 10.1093/annhyg/meg032

Druckenmüller, K., Gärtner, A., Jäckel, U., Klug, K., Schiffels, J., Günther, K., et al. (2017). Development of a methodological approach for the characterization of bioaerosols in exhaust air from pig fattening farms with MALDI-TOF mass spectrometry. *Int. J. Hyg. Environ. Health* 220, 974–983. doi: 10.1016/j.ijheh.2017.05.003

Environment Agency (2017). Environmental monitoring of bioaerosols at regulated facilities (M9). Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/730226/M9_Environmental_monitoring_of_bioaerosols_at_regulated_facilities.pdf (Accessed March, 2022).

Fennelly, M. J., Sewell, G., Prentice, M. B., O'Connor, D. J., and Sodeau, J. R. (2017). Review: the use of real-time fluorescence instrumentation to monitor ambient primary biological aerosol particles (PBAP). *Atmosphere (Basel).* 9, 1–39. doi: 10.3390/atmos9010001

Ferguson, R. M. W., Garcia-Alcega, S., Coulon, F., Dumbrell, A. J., Whitby, C., and Colbeck, I. (2019). Bioaerosol biomonitoring: sampling optimization for molecular microbial ecology. *Mol. Ecol. Resour.* 19, 672–690. doi: 10.1111/1755-0998.13002

Filipkowska, Z., Gotkowska-Plachta, A., and Korzeniewska, E. (2008). Moulds, yeasts and yeast-like fungi in the atmospheric air at constructed wetland systems (with aerated and stabilization ponds) and in the surrounding area. *Water Environ. Rural Areas* 8, 69–82.

Filipkowska, Z., Gotkowska-Płachta, A., Korzeniewska, E., and Pawlukiewicz, A. (2007). Micological contamination of the atmospheric air at municipal wastewater treatment plant with activated sludge tanks aerated by CELPOX devices. *Ochr. Środowiska i Zasobów Nat. IOŚ* 32, 240–245.

Filipkowska, Z., Janczukowicz, W., Krzemieniewski, M., and Pesta, J. (2002). Municipal wastewater treatment plant with activated sludge tanks aerated by CELPOX devices as a source of microbiological pollution of the atmosphere. *Polish J. Environ. Stud.* 11, 639–648.

Foladori, P., Cutrupi, F., Segata, N., Manara, S., Pinto, F., Malpei, F., et al. (2020). SARS-CoV-2 from faeces to wastewater treatment: what do we know? A review. *Sci. Total Environ.* 743:140444. doi: 10.1016/j.scitotenv.2020.140444

Fracchia, L., Pietronave, S., Rinaldi, M., and Giovanna Martinotti, M. (2006). Site-related airborne biological hazard and seasonal variations in two wastewater treatment plants. *Water Res.* 40, 1985–1994. doi: 10.1016/j.watres.2006.03.016

Garcia-Alcega, S., Nasir, Z. A., Ferguson, R., Whitby, C., Dumbrell, A. J., Colbeck, I., et al. (2017). Fingerprinting outdoor air environment using microbial

volatile organic compounds (MVOCs)–a review. *TrAC Trends Anal. Chem.* 86, 75–83. doi: 10.1016/j.trac.2016.10.010

Gerardi, M. H., and Zimmerman, M. C. (2005). *Wastewater Pathogens*. Hoboken, NJ: John Wiley & Sons, Inc..

Ghosh, B., Lal, H., and Srivastava, A. (2015). Review of bioaerosols in indoor environment with special reference to sampling, analysis and control mechanisms. *Environ. Int.* 85, 254–272. doi: 10.1016/j.envint.2015.09.018

Griffiths, W. D., and DeCosemo, G. A. L. (1994). The assessment of bioaerosols: a critical review. *J. Aerosol Sci.* 25, 1425–1458. doi: 10.1016/0021-8502(94)90218-6

Haddrell, A. E., and Thomas, R. J. (2017). Aerobiology: experimental considerations, observations, and future tools. *Appl. Environ. Microbiol.* 83, e00809–e00817. doi: 10.1128/AEM.00809-17

Han, Y., Li, L., Wang, Y., Ma, J., Li, P., Han, C., et al. (2020a). Composition, dispersion, and health risks of bioaerosols in wastewater treatment plants: a review. *Front. Environ. Sci. Eng.* 15:38. doi: 10.1007/s11783-020-1330-1

Han, Y., Yang, T., Chen, T., Li, L., and Liu, J. (2019). Characteristics of submicron aerosols produced during aeration in wastewater treatment. *Sci. Total Environ.* 696:134019. doi: 10.1016/j.scitotenv.2019.134019

Han, Y., Yang, T., Han, C., Li, L., and Liu, J. (2020b). Study of the generation and diffusion of bioaerosol under two aeration conditions. *Environ. Pollut.* 267:115571. doi: 10.1016/j.envpol.2020.115571

Healy, D. A., Huffman, J. A., O'Connor, D. J., Pöhlker, C., Pöschl, U., and Sodeau, J. R. (2014). Ambient measurements of biological aerosol particles near Killarney, Ireland: a comparison between real-time fluorescence and microscopy techniques. *Atmos. Chem. Phys.* 14, 8055–8069. doi: 10.5194/acp-14-8055-2014

Heidelberg, J. F., Shahamat, M., Levin, M., Rahman, I., Stelma, G., Grim, C., et al. (1997). Effect of aerosolization on culturability and viability of gram-negative bacteria. *Appl. Environ. Microbiol.* 63, 3585–3588. doi: 10.1128/aem.63.9.3585-3588.1997

Heller, L., Mota, C. R., and Greco, D. B. (2020). COVID-19 faecal-oral transmission: are we asking the right questions? *Sci. Total Environ.* 729:138919. doi: 10.1016/j.scitotenv.2020.138919

Ho, J. (2002). Future of biological aerosol detection. *Anal. Chim. Acta* 457, 125–148. doi: 10.1016/S0003-2670(01)01592-6

Hsiao, T. C., Lin, A. Y. C., Lien, W. C., and Lin, Y. C. (2020). Size distribution, biological characteristics and emerging contaminants of aerosols emitted from an urban wastewater treatment plant. *J. Hazard. Mater.* 388:121809. doi: 10.1016/j.jhazmat.2019.121809

Huffman, J. A., and Santarpia, J. (2017). "Online techniques for quantification and characterization of biological aerosols" in *Microbiology of Aerosols*. eds. A.-M. Delort and P. Amato (John Wiley and Sons, Inc). doi: 10.1002/9781119132318.ch1d

Hung, H. F., Kuo, Y. M., Chien, C. C., and Chen, C. C. (2010). Use of floating balls for reducing bacterial aerosol emissions from aeration in wastewater treatment processes. *J. Hazard. Mater.* 175, 866–871. doi: 10.1016/j.jhazmat.2009.10.090

IWA and OFID (2018). Wastewater Report 2018. The reuse opportunity.

Jones, A. M., and Harrison, R. M. (2004). The effects of meteorological factors on atmospheric bioaerosol concentrations - a review. *Sci. Total Environ.* 326, 151–180. doi: 10.1016/j.scitotenv.2003.11.021

Kabir, E., Azzouz, A., Raza, N., Bhardwaj, S. K., Kim, K. H., Tabatabaei, M., et al. (2020). Recent advances in monitoring, sampling, and sensing techniques for bioaerosols in the atmosphere. *ACS Sensors* 5, 1254–1267. doi: 10.1021/acssensors.9b02585

Karra, S., and Katsivela, E. (2007). Microorganisms in bioaerosol emissions from wastewater treatment plants during summer at a Mediterranean site. *Water Res.* 41, 1355–1365. doi: 10.1016/j.watres.2006.12.014

Kathiriya, T., Gupta, A., and Singh, N. K. (2021). An opinion review on sampling strategies, enumeration techniques, and critical environmental factors for bioaerosols: an emerging sustainability indicator for society and cities. *Environ. Technol. Innov.* 21:101287. doi: 10.1016/j.eti.2020.101287

Katsivela, E., Latos, E., Raisi, L., Aleksandropoulou, V., and Lazaridis, M. (2017). Particle size distribution of cultivable airborne microbes and inhalable particulate matter in a wastewater treatment plant facility. *Aerobiol. Bologna* 33, 297–314. doi: 10.1007/s10453-016-9470-2

Kaye, P. H., Barton, J. E., Hirst, E., and Clark, J. M. (2000). Simultaneous light scattering and intrinsic fluorescence measurement for the classification of airborne particles. *Appl. Opt.* 39, 3738–3745. doi: 10.1364/ao.39.003738

Kaye, P. H., and Hirst, E. (2011). Second generation low-cost particle counter. Patent No: U.S. Patent No. 9116121.

Kim, K. H., Kabir, E., and Jahan, S. A. (2018). Airborne bioaerosols and their impact on human health. *J. Environ. Sci. (China)* 67, 23–35. doi: 10.1016/j.jes.2017.08.027

Kim, K. Y., Kim, Y. S., and Kim, D. (2010). Distribution characteristics of airborne bacteria and fungi in the general hospitals of Korea. *Ind. Health* 48, 236–243. doi: 10.2486/indhealth.48.236

Könemann, T., Savage, N., Klimach, T., Walter, D., Fröhlich-Nowoisky, J., Su, H., et al. (2019). Spectral intensity bioaerosol sensor (SIBS): an instrument for spectrally resolved fluorescence detection of single particles in real time. *Atmos. Meas. Tech.* 12, 1337–1363. doi: 10.5194/amt-12-1337-2019

Korzeniewska, E. (2011). Emission of bacteria and fungi in the air from wastewater treatment plants – a review. *Front. Biosci. Sch.* 3, 393–407. doi: 10.2741/s159

Korzeniewska, E., Filipkowska, Z., and Gotkowska-Płachta, A. (2007). Municipial wastewater treatment plant as a source of *Enterobacteriaceae* bacteria in the air. *Ochr. Środowiska i Zasobów Nat. IOŚ* 32, 178–183.

Korzeniewska, E., Filipkowska, Z., Gotkowska-Płachta, A., Janczukowicz, W., Dixon, B., and Czułowska, M. (2009). Determination of emitted airborne microorganisms from a BIO-PAK wastewater treatment plant. *Water Res.* 43, 2841–2851. doi: 10.1016/j.watres.2009.03.050

Korzeniewska, E., Filipkowska, Z., Gotkowska-Płachta, A., Janczukowicz, W., and Rutkowski, B. (2008). Bacteriological pollution of the atmospheric air at the municipal and dairy wastewater treatment plant area and in its surroundings. *Arch. Environ. Prot.* 34, 13–23.

Kowalski, M., Wolany, J., Pastuszka, J. S., Płaza, G., Wlazło, A., Ulfig, K., et al. (2017). Characteristics of airborne bacteria and fungi in some polish wastewater treatment plants. *Int. J. Environ. Sci. Technol.* 14, 2181–2192. doi: 10.1007/s13762-017-1314-2

Li, Y., Yang, L., Meng, Q., Qiu, X., and Feng, Y. (2013). Emission characteristics of microbial aerosols in a municipal sewage treatment plant in Xi'an. *China. Aerosol Air Qual. Res.* 13, 343–349. doi: 10.4209/aaqr.2012.05.0123

Li, J., Zhou, L., Zhang, X., Xu, C., Dong, L., and Yao, M. (2016). Bioaerosol emissions and detection of airborne antibiotic resistance genes from a wastewater treatment plant. *Atmos. Environ.* 124, 404–412. doi: 10.1016/j.atmosenv.2015.06.030

Liang, L., He, K., and Duan, F. (2012). Progress on quantitative assessment methods of biological aerosols in the atmosphere. *Biogeosci. Discuss.* 9, 1511–1528. doi: 10.5194/bgd-9-1511-2012

Liu, M., Nobu, M. K., Ren, J., Jin, X., Hong, G., and Yao, H. (2020). Bacterial compositions in inhalable particulate matters from indoor and outdoor wastewater treatment processes. *J. Hazard. Mater.* 385:121515. doi: 10.1016/j.jhazmat.2019.121515

Madelin, T. M., and Johnson, H. E. (1992). Fungal and actinomycete spore aerosols measured at different humidities with an aerodynamic particle sizer. *J. Appl. Bacteriol.* 72, 400–409. doi: 10.1111/j.1365-2672.1992.tb01853.x

Mainelis, G. (2020). Bioaerosol sampling: classical approaches, advances, and perspectives. *Aerosol Sci. Technol.* 54, 496–519. doi: 10.1080/02786826.2019.1671950

Masclaux, F. G., Hotz, P., Gashi, D., Savova-Bianchi, D., and Oppliger, A. (2014). Assessment of airborne virus contamination in wastewater treatment plants. *Environ. Res.* 133, 260–265. doi: 10.1016/j.envres.2014.06.002

Mateo-Sagasta, J., Raschid-Sally, L., and Thebo, A. (2015). "Global wastewater and sludge production, treatment and use" in *Wastewater: Economic Asset in an Urbanizing World*. eds P. Drechsel, M. Qadir and D. Wichelns (Dordrecht: Springer), 15–38. doi: 10.1007/978-94-017-9545-6_2

Mbareche, H., Brisebois, E., Veillette, M., and Duchaine, C. (2017). Bioaerosol sampling and detection methods based on molecular approaches: no pain no gain. *Sci. Total Environ.* 599-600, 2095–2104. doi: 10.1016/j.scitotenv.2017.05.076

Mbareche, H., Dion-Dupont, V., Veillette, M., Brisebois, E., Lavoie, J., and Duchaine, C. (2022). Influence of seasons and sites on bioaerosols in indoor wastewater treatment plants and proposal for air quality indicators. *J. Air Waste Manage. Assoc.* 72, 1000–1011. doi: 10.1080/10962247.2022.2066735

Mbareche, H., Veillette, M., Bilodeau, G. J., and Duchaine, C. (2019). Fungal aerosols at dairy farms using molecular and culture techniques. *Sci. Total Environ.* 653, 253–263. doi: 10.1016/j.scitotenv.2018.10.345

Michalkiewicz, M. (2019). Wastewater treatment plants as a source of bioaerosols. *Polish J. Environ. Stud.* 28, 2261–2271. doi: 10.15244/pjoes/90183

Michalkiewicz, M., Pruss, A., Dymaczcwski, Z., Jez-Walkowiak, J., and Kwaśna, S. (2011). Microbiological air monitoring around municipal wastewater treatment plants. *Polish J. Environ. Stud* 20, 1243–1250.

Nasir, Z. A., Hayes, E., Williams, B., Gladding, T., Rolph, C., Khera, S., et al. (2019). Scoping studies to establish the capability and utility of a real-time bioaerosol sensor to characterise emissions from environmental sources. *Sci. Total Environ.* 648, 25–32. doi: 10.1016/j.scitotenv.2018.08.120

Nasir, Z. A., Rolph, C., Collins, S., Stevenson, D., Gladding, T. L., Hayes, E., et al. (2018). A controlled study on the characterisation of bioaerosols emissions from compost. *Atmosphere (Basel).* 9:379. doi: 10.3390/atmos9100379

Negron, A., Deleon-Rodriguez, N., Waters, S. M., Ziemba, L. D., Anderson, B., Bergin, M., et al. (2020). Using flow cytometry and light-induced fluorescence to characterize the variability and characteristics of bioaerosols in springtime in metro Atlanta. *Georgia. Atmos. Chem. Phys.* 20, 1817–1838. doi: 10.5194/acp-20-1817-2020

Niazi, S., Hassanvand, M. S., Mahvi, A. H., Nabizadeh, R., Alimohammadi, M., Nabavi, S., et al. (2015). Assessment of bioaerosol contamination (bacteria and

fungi) in the largest urban wastewater treatment plant in the Middle East. *Environ. Sci. Pollut. Res.* 22, 16014–16021. doi: 10.1007/s11356-015-4793-z

Oconnor, D. J., Healy, D. A., Hellebust, S., Buters, J. T. M., and Sodeau, J. R. (2014). Using the WIBS-4 (waveband integrated bioaerosol sensor) technique for the on-line detection of pollen grains. *Aerosol Sci. Technol.* 48, 341–349. doi: 10.1080/02786826.2013.872768

OECD (2019). *Wastewater Treatment (indicator)*. Paris: OECD.

Orsini, M., Laurenti, P., Boninti, F., Arzani, D., Ianni, A., and Romano-Spica, V. (2002). A molecular typing approach for evaluating bioaerosol exposure in wastewater treatment plant workers. *Water Res.* 36, 1375–1378. doi: 10.1016/S0043-1354(01)00336-0

Pan, Y. (2015). Detection and characterization of biological and other organic-carbon aerosol particles in atmosphere using fluorescence. *J. Quant. Spectrosc. Radiat. Transf.* 150, 12–35. doi: 10.1016/j.jqsrt.2014.06.007

Papke, R. T., and Ward, D. M. (2004). The importance of physical isolation to microbial diversification. *FEMS Microbiol. Ecol.* 48, 293–303. doi: 10.1016/j.femsec.2004.03.013

Pascual, L., Pérez-Luz, S., Yáñez, M. A., Santamaría, A., Gibert, K., Salgot, M., et al. (2003). Bioaerosol emission from wastewater treatment plants. *Aerobiologia (Bologna)* 19, 261–270. doi: 10.1023/B:AERO.0000006598.45757.7f

Patentalakis, N., Pantidou, A., and Kalogerakis, N. (2008). Determination of enterobacteria in air and wastewater samples from a wastewater treatment plant by epi-fluorescence microscopy. Water, air. *Soil Pollut. Focus* 8, 107–115. doi: 10.1007/s11267-007-9135-9

Peccia, J., Milton, D. K., Reponen, T., and Hill, J. (2008). A role for environmental engineering and science in preventing bioaerosol-related disease. *Environ. Sci. Technol.* 42, 4631–4637. doi: 10.1021/es087179e

Perring, A. E., Schwarz, J. P., Baumgardner, D., Hernandez, M. T., Spracklen, D. V., Heald, C. L., et al. (2015). Airborne observations of regional variation in fluorescent aerosol across the United States. *J. Geophys. Res.* 120, 1153–1170. doi: 10.1002/2014JD022495

Pöhlker, C., Huffman, J. A., and Pöschl, U. (2012). Autofluorescence of atmospheric bioaerosols - fluorescent biomolecules and potential interferences. *Atmos. Meas. Tech.* 5, 37–71. doi: 10.5194/amt-5-37-2012

Ranalli, G., Principi, P., and Sorlini, C. (2000). Bacterial aerosol emission from wastewater treatment plants: culture methods and bio-molecular tools. *Aerobiologia (Bologna).* 16, 39–46. doi: 10.1023/A:1007656414770

Rylander, R., Andersson, K., Belin, L., Berglund, G., Bergström, R., Hanson, L. Å., et al. (1976). Sewage Worker'S Syndrome. *Lancet* 308, 478–479. doi: 10.1016/S0140-6736(76)92583-6

Šantl-Temkiv, T., Sikoparija, B., Maki, T., Carotenuto, F., Amato, P., Yao, M., et al. (2020). Bioaerosol field measurements: challenges and perspectives in outdoor studies. *Aerosol Sci. Technol.* 54, 520–546. doi: 10.1080/02786826.2019.1676395

Sauvageat, E., Zeder, Y., Auderset, K., Calpini, B., Clot, B., Crouzy, B., et al. (2020). Real-time pollen monitoring using digital holography. *Atmos. Meas. Tech.* 13, 1539–1550. doi: 10.5194/amt-13-1539-2020

Šikoparija, B. (2020). Desert dust has a notable impact on aerobiological measurements in Europe. *Aeolian Res.* 47:100636. doi: 10.1016/j.aeolia.2020.100636

Singh, N. K., Sanghvi, G., Yadav, M., Padhiyar, H., and Thanki, A. (2021). A state-of-the-art review on WWTP associated bioaerosols: microbial diversity, potential emission stages, dispersion factors, and control strategies. *J. Hazard. Mater.* 410:124686. doi: 10.1016/j.jhazmat.2020.124686

Steele, P. T., Farquar, G. R., Martin, A. N., Coffee, K. R., Riot, V. J., Martin, S. I., et al. (2008). Autonomous, broad-spectrum detection of hazardous aerosols in seconds. *Anal. Chem.* 80, 4583–4589. doi: 10.1021/ac8004428

Stellacci, P., Liberti, L., Notarnicola, M., and Haas, C. N. (2010). Hygienic sustainability of site location of wastewater treatment plants. A case study. II. Estimating airborne biological hazard. *Desalination* 253, 106–111. doi: 10.1016/j.desal.2009.11.024

Thorn, J., Beijer, L., Jonsson, T., and Rylander, R. (2002). Measurement strategies for the determination of airborne bacterial endotoxin in sewage treatment plants. *Ann. Occup. Hyg.* 46, 549–554. doi: 10.1093/annhyg/mef068

Tian, J., Yan, C., Nasir, Z. A., Alcega, S. G., Tyrrel, S., and Coulon, F. (2020). Real time detection and characterisation of bioaerosol emissions from wastewater treatment plants. *Sci. Total Environ.* 721:137629. doi: 10.1016/j.scitotenv.2020.137629

Upadhyay, N., Sun, Q., Allen, J. O., Westerhoff, P., and Herckes, P. (2013). Characterization of aerosol emissions from wastewater aeration basins. *J. Air Waste Manag. Assoc.* 63, 20–26. doi: 10.1080/10962247.2012.726693

Van Hooste, W., Charlier, A. M., Rotsaert, P., Bulterys, S., Moens, G., Van Sprundel, M., et al. (2010). Work-related helicobacter pylori infection among sewage workers in municipal wastewater treatment plants in Belgium. *Occup. Environ. Med.* 67, 91–97. doi: 10.1136/oem.2008.040436

Vestlund, A. T., Al-Ashaab, R., Tyrrel, S. F., Longhurst, P. J., Pollard, S. J. T., and Drew, G. H. (2014). Morphological classification of bioaerosols from composting using scanning electron microscopy. *Waste Manag.* 34, 1101–1108. doi: 10.1016/j.wasman.2014.01.021

Vítězová, M., Vítěz, T., Mlejnková, H., and Lošák, T. (2012). Microbial contamination of the air at the wastewater treatment plant. *Acta Univ. Agric. Silvic. Mendelianae Brun.* 60, 233–240. doi: 10.11118/actaun201260030233

Zhang, M., Klimach, T., Ma, N., Könemann, T., Pöhlker, C., Wang, Z., et al. (2019). Size-resolved single-particle fluorescence spectrometer for real-time analysis of bioaerosols: laboratory evaluation and atmospheric measurements. *Environ. Sci. Technol.* 53, 13257–13264. doi: 10.1021/acs.est.9b01862

frontiers | Frontiers in Microbiology

TYPE Original Research
PUBLISHED 23 February 2023
DOI 10.3389/fmicb.2023.1078469

OPEN ACCESS

EDITED BY
Hongbin Liu,
Hong Kong University of Science and Technology, Hong Kong SAR, China

REVIEWED BY
Jiwen Liu,
Ocean University of China, China
Angelina Lo Giudice,
Department of Earth System Sciences and Technologies for the Environment, Institute of Polar Sciences (CNR), Italy

*CORRESPONDENCE
Viena Puigcorbé
✉ viena.puigcorbe@outlook.com;
✉ vienap@icm.csic.es
Clara Ruiz-González
✉ clararg@icm.csic.es

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

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

RECEIVED 24 October 2022
ACCEPTED 27 January 2023
PUBLISHED 23 February 2023

CITATION
Puigcorbé V, Ruiz-González C, Masqué P and Gasol JM (2023) Impact of particle flux on the vertical distribution and diversity of size-fractionated prokaryotic communities in two East Antarctic polynyas.
Front. Microbiol. 14:1078469.
doi: 10.3389/fmicb.2023.1078469

Impact of particle flux on the vertical distribution and diversity of size-fractionated prokaryotic communities in two East Antarctic polynyas

Viena Puigcorbé[1,2*†], Clara Ruiz-González[1*†], Pere Masqué[2,3] and Josep M. Gasol[1]

[1]Department of Marine Biology and Oceanography, Institut de Ciències del Mar (ICM-CSIC), Barcelona, Catalunya, Spain, [2]Centre for Marine Ecosystems Research, School of Science, Edith Cowan University, Joondalup, WA, Australia, [3]International Atomic Energy Agency, City of Monaco, Monaco

Antarctic polynyas are highly productive open water areas surrounded by ice where extensive phytoplankton blooms occur, but little is known about how these surface blooms influence carbon fluxes and prokaryotic communities from deeper waters. By sequencing the 16S rRNA gene, we explored the vertical connectivity of the prokaryotic assemblages associated with particles of three different sizes in two polynyas with different surface productivity, and we linked it to the magnitude of the particle export fluxes measured using thorium-234 (^{234}Th) as particle tracer. Between the sunlit and the mesopelagic layers (700 m depth), we observed compositional changes in the prokaryotic communities associated with the three size-fractions, which were mostly dominated by *Flavobacteriia*, *Alphaproteobacteria*, and *Gammaproteobacteria*. Interestingly, the vertical differences between bacterial communities attached to the largest particles decreased with increasing ^{234}Th export fluxes, indicating a more intense downward transport of surface prokaryotes in the most productive polynya. This was accompanied by a higher proportion of surface prokaryotic taxa detected in deep particle-attached microbial communities in the station with the highest ^{234}Th export flux. Our results support recent studies evidencing links between surface productivity and deep prokaryotic communities and provide the first evidence of sinking particles acting as vectors of microbial diversity to depth in Antarctic polynyas, highlighting the direct influence of particle export in shaping the prokaryotic communities of mesopelagic waters.

KEYWORDS

prokaryotic communities, marine particles, ^{234}Thorium, particle export, polynyas, Antarctica, particle size fractionation, particle-attached and free-living prokaryotes

1. Introduction

The Southern Ocean (<35°S) is a key player in the global carbon cycle and the regulation of the Earth's climate, accounting for ~40% of the anthropogenic CO_2 oceanic uptake (Gruber et al., 2009 and references therein). Although the Southern Ocean is the largest high-nutrient low-chlorophyll region, highly biologically productive areas are found south of the circumpolar current (del Castillo et al., 2019), particularly on the continental shelf, in coastal zones including polynyas (Sedwick and DiTullio, 1997; Montes-Hugo and Yuan, 2012; Arrigo et al., 2015). Polynyas are ice-free areas created and maintained by katabatic winds and surrounded by consolidated sea ice. In the polynyas, solar

radiation reaches the water column and, when it becomes stratified, light together with nutrient inputs from different sources (e.g., Lannuzel et al., 2010; de Jong et al., 2012; Shadwick et al., 2013) triggers significant phytoplankton blooms (Kang et al., 2001; Arrigo et al., 2015; Jena and Pillai, 2020). This enhanced primary productivity can lead to high local carbon export rates mediated by the sinking of organic particles, thus potentially contributing to the ocean uptake of atmospheric CO_2 (Hoppema and Anderson, 2007; Miller and DiTullio, 2007; Arrigo et al., 2008). However, the timing, extent, and intensity of phytoplankton blooms and primary production can vary significantly even amongst closely located Antarctic polynyas (Arrigo et al., 2015; Moreau et al., 2019), with potential but poorly known implications for the efficiency of the export flux and for the underlying microbial communities that depend on surface-derived carbon inputs.

The structure of surface phytoplankton communities determines the amount, quality, and sinking rates of particles leaving the surface ocean. Besides, physical processes and complex food web interactions, including remineralization processes driven by bacteria, modulate the export and transfer efficiency of the biological carbon pump (Henson et al., 2019; Wiedmann et al., 2020; Nguyen et al., 2022). In turn, sinking particles have been shown to act as microbial diversity vectors between the surface waters and the deep ocean, delivering particle-attached microorganisms that seem able to colonize deeper waters (Mestre et al., 2018), and explaining why meso- and bathypelagic bacterial communities associated with different size-fractions reflect surface gradients of surface phytoplankton productivity (Ruiz-González et al., 2020). In coastal polar ecosystems, pelagic bacteria have been shown to respond quickly to spring and summertime phytoplankton blooms, although the fraction of primary production consumed by heterotrophic bacteria seems highly variable (Ducklow and Yager, 2007; Kirchman et al., 2009; Luria et al., 2016), which could be one of the reasons behind the inverse relationship between primary production and export efficiency often observed in the Southern Ocean (Maiti et al., 2013; le Moigne et al., 2016; Henson et al., 2019). This suggests that particle export fluxes and the interacting prokaryote assemblages are tightly dependent on each other. However, very few studies have coupled measurements of particle export with particle-attached prokaryotic community composition, especially in Antarctic polynyas.

Sinking particle fluxes in the ocean can be directly measured by using sediment traps, which have some advantages (e.g., temporal coverage in the case of moored sediment traps) but have a limited spatial coverage and are time consuming to deploy/recover. Therefore, indirect methods are essential to increase the number of observations across the ocean. One of the most common indirect methods for quantifying the magnitude of sinking particle fluxes in the ocean is the use of parent-daughter radionuclide pairs. Among them, the $^{234}Th/^{238}U$ pair is the most extensively used (Ceballos-Romero et al., 2022). ^{234}Th is continuously produced by the decay of ^{238}U, which has a conservative behavior in oxygenated ocean waters, and is highly particle reactive, so it can be used as a particle tracer. Moreover, the relatively short half-life of ^{234}Th ($T_{1/2}$ = 24.1 d) suits the biologically mediated temporal changes in particle production and export. ^{238}U, on the other hand, has a very long half-life ($T_{1/2}$ = 4.5 10^9 y). The difference in half-lives between ^{234}Th and ^{238}U implies that both radioisotopes should be in secular equilibrium (i.e., $^{234}Th/^{238}U$ activity ratio of 1) in the environment. However, in the presence of marine particles, ^{234}Th sorbs onto them and can be scavenged when the particles sink, thus breaking the secular equilibrium in the water column (i.e., $^{234}Th/^{238}U$ activity ratio <1). The magnitude of the disequilibrium between both radioisotopes allows estimating the magnitude of the particle export flux, which can be converted to fluxes of particulate organic carbon, trace metals, or pollutants (Gustafsson et al., 1997; Black et al., 2019; Puigcorbé et al., 2020; Tesán Onrubia et al., 2020). Thus, combining ^{234}Th fluxes with vertical variations in the composition of the communities attached to sinking particles can provide a direct way to link surface productivity, export fluxes, and impacts on microbial communities.

Here, we explore whether differences in surface productivity among Antarctic polynyas impact differentially the particle export fluxes and the microbial communities from the underlying mesopelagic waters. To do so, prokaryotic communities associated with particles of three different sizes were sampled in three stations located in two East Antarctic polynyas differing largely in their surface productivity and dominant phytoplankton groups. We examined the variability in the prokaryotic communities associated with free-living (0.2–0.8 µm), small (0.8–53 µm), and large (>53 µm) particles collected along the water column and compared them to the magnitude of the particle export fluxes estimated using thorium-234 (^{234}Th) as particle tracer. We hypothesize that mesopelagic prokaryotic communities will be more similar to those in the surface in the stations with the highest surface productivity and higher particle export fluxes.

2. Materials and methods

2.1. Sampling and analyses of physicochemical and biological parameters of the study area

Sampling was performed on board of the RSV Aurora Australis (AU1602, AA-V02 2016/17) between 8 December 2016 and 21 January 2017, in one station belonging to the Dalton polynya (D02) and two stations (M36 and M48) located within the Mertz polynya (East Antarctica; 67.2–66.8 °S and 119.5–145.8°E; Supplementary Figure S1A). These stations were chosen due to their contrasting biological, chemical, and physical characteristics, extensively described and discussed by Moreau et al. (2019) and Ratnarajah et al. (2022). For comparison with Ratnarajah et al. (2022), D02 refers to St.2; M36 is EM03 and M48 is MG08. Briefly, both polynyas had similar sea surface temperatures ranging majorly between 0.0 and 1.5°C, but the Dalton polynya showed higher sea surface salinity than the Mertz polynya (34.0–34.3 g/kg vs. 32.5–33.5 g/kg, respectively), suggesting that the later could have experienced more sea ice melting than the Dalton polynya (Moreau et al., 2019). The Dalton polynya presented deeper euphotic depths (95 ± 56 m) and mixed layer depths (25 ± 12 m, excluding two stations where the mixed layer was down to 100 m and 154 m) than the Mertz polynya (40 ± 9 m and 13 ± 1 m, respectively) (Ratnarajah et al., 2022). In general, chlorophyll-a (Chl-a) concentrations in the surface were higher in the Dalton polynya compared to the Mertz polynya (max of 15 µg L^{-1} vs. 8 µg L^{-1}), but Mertz presented a subsurface Chl-a maximum of ~10 µg L^{-1} located between 20 and 70 m depth that was consistent along the whole polynya, whereas in the Dalton that layer was more variable (Moreau et al., 2019). Both polynyas also presented global differences in nutrient ratios (i.e., Si:N or N:P) suggesting different nutrient sources (i.e., water masses) and different phytoplanktonic communities, something which was also corroborated by microscope analyses (Moreau et al., 2019). During the time of sampling, the dominant phytoplankton groups differed between both polynyas, with *Phaeocystis antarctica* dominating in the Dalton station and diatoms dominating in the much more productive Mertz stations (Moreau et al., 2019).

The sampling and analyses of physicochemical and environmental parameters were performed as described in Moreau et al. (2019) and

Ratnarajah et al. (2022). Temperature and salinity were obtained from the CTD (Rosenberg and Rintoul, 2017) and fluorescence values were obtained with a fluorometer (ECO-AFL/FL 756, Wetlabs, United States) that was installed also on the CTD rosette. Of particular interest for this study is the concentration of Chl-a, particulate organic carbon (POC) and inorganic nutrients. Chl-a concentrations were obtained from ~500 ml of filtered seawater, extracted with acetone and stored at −20°C for 24–48 h in the dark prior to analysis with a Turner Trilogy fluorometer. POC was determined following Knap et al. (1996) and analyzed using a Thermo Finnigan EA 1112 Series Flash Elemental Analyzer. Concentrations of inorganic macronutrients (nitrate, nitrite and silicic acid) were analyzed after the voyage at the CSIRO laboratory (Hobart, Australia) following the methods described in Murphy and Riley (1962), Armstrong et al. (1967), Wood et al. (1967) and Kérouel and Aminot (1997).

2.2. Sampling of free-living and particle-attached prokaryotic communities

Seawater samples were collected using 12 L Niskin bottles attached to a CTD rosette. At each station, 6 depths were sampled covering the entire water column (down to ~50 m above the bottom, see Table 1). After collection, the samples (5–11 L) were filtered onboard at 4°C using a Masterflex peristaltic pump. The filtration was done sequentially through a 53 µm pore-size Nitex screen mesh, followed by a 0.8 µm pore-size polycarbonate membrane filter (Millipore) and finally through a 0.2 µm Sterivex filter, thus obtaining three size-fractions from each sample: >53 µm, 0.8–53 µm and 0.2–0.8 µm (for simplification, hereafter 53 µm, 0.8 µm, and 0.2 µm). After filtration, the filters were stored at −80°C for further analyses at the home laboratory.

2.3. Particle export fluxes derived from the ^{238}U-^{234}Th method

^{234}Th analyses are described in Ratnarajah et al. (2022). Briefly, 4 L seawater samples were collected at 12–14 depths along the water column at each station and processed using the manganese oxide co-precipitation technique (Clevenger et al., 2021), while ^{238}U activity concentrations were derived from salinity data (Owens et al., 2011). Th-234 samples were counted onboard using a gas flow proportional low-level RISO beta counter (counting statistics <5%) and recounted >6 months later to account for background activities. Chemical recoveries were obtained following Puigcorbé et al. (2017a) and measured by inductively coupled plasma mass spectrometry at the Alfred Wegener Institute.

^{234}Th export fluxes (proxies of particle fluxes) were estimated by integrating the ^{234}Th deficit in the upper water column relative to ^{238}U, using a 1D scavenging model assuming steady state conditions and no significant advection nor diffusion transport. The integration depth used here is the depth where ^{234}Th and ^{238}U reached secular equilibrium (i.e., ^{234}Th/^{238}U activity ratio = 1) (see Ratnarajah et al. (2022) for further details).

2.4. Characterization of prokaryotic communities

Prokaryotic community structure was determined by high-throughput Illumina sequencing of the 16S rRNA genes. A total of 54 samples were obtained, which were stored at −80°C until analyses were conducted. At the home laboratory, DNA was extracted using the PowerWater DNA Isolation kit following manufacturer's instructions (MOBIO). The samples were sequenced using Illumina MiSeq 2 × 300 bp

TABLE 1 Characteristics of the sampled sites—location, date, depth, and filtered volume for DNA samples per each size-fraction.

		Depth (m)	Sample volume (L)		
			53µm	0.8µm	0.2µm
Polynya	Dalton	5	6.5*	6.5	5.0
Station	D02	20	9.1	9.1	5.0
Lat (°S)	66.843	100	7.5*	7.5	7.5
Lon (°E)	119.543	300	8.0*	8.0	3.0
Collection date	31/Dec/16	550	8.0	8.0	2.5
		720	9.0	9.0	2.5
Polynya	Mertz	5	7.8	7.8	7.8
Station	M36	25	9.4	9.4	5.0
Lat (°S)	66.908	100	10.6	10.4	5.0
Lon (°E)	145.498	300	9.8	9.8	5.0
Collection date	10/Jan/17	550	10.2	9.8	5.0
		638	10.7	10.5	5.0
Polynya	Mertz	5	5.9	5.9	5.0
Station	M48	50	9.3	5.0	3.8
Lat (°S)	67.219	100	10.2	10.1	3.0
Lon (°E)	145.881	300	10.6	10.6	5.0
Collection date	11/Jan/17	500	9.4	9.4	5.0
		670	4.9	4.9	4.7

*Stations from where sequencing data are not available due to low DNA recoveries.

flow cells at RTL Genomics (Texas, United States) using primers 515F-Y and 926R (Parada et al., 2015) to amplify the V4-V5 region of the 16S rRNA gene. Three of the 53 µm samples from the Dalton polynya did not have enough DNA material and could not be sequenced (see Table 1). The sequences were processed according to Logares (2017). In brief, primers were removed with Cutadapt (Martin, 2011). The paired-end reads were merged with PEAR (Zhang et al., 2014). Quality filtering, chimera checking and operational taxonomic unit (OTU) clustering (99% similarity) were done with the UPARSE pipeline (Edgar, 2013). Singletons and chimeric OTUs were removed, and the remaining OTUs were taxonomically annotated using the SILVA v123 database. OTUs assigned to chloroplasts were removed, resulting in a total of 8,219 OTUs and 402,440 sequences. To enable comparisons between samples, the OTU table was randomly subsampled to ensure an equal number of sequences per sample (5,000 sequences) using *rrarefy* (Vegan package, R Core Team, 2017), retaining 253,545 sequences clustered into 6,546 OTUs. The raw sequence data have been deposited in the Figshare data repository, doi: 10.6084/m9.figshare.21385215.v1.

2.5. Statistical analyses

The spatial differences between prokaryotic communities were visualized using nonmetric multidimensional scaling (NMDS, Vegan *metaMDS* function) based on Bray–Curtis distances. Significant differences in taxonomic composition between depths, stations, or size-fraction were tested using ANOSIM (Vegan *anosim* function). Vertical differences between the prokaryotic communities within each size-fraction were estimated for each individual station as the Bray–Curtis dissimilarity between the surface (5 m) and each of the deeper prokaryotic communities, and the proportion of "surface-derived" OTUs in mesopelagic communities was estimated considering those mesopelagic OTUs that showed presence in surface (<100 m) waters. OTUs unique to a given sampling station (i.e., OTUs present exclusively in one of the three sampled stations) were identified considering all depths together within each station. Statistical analyses and data handling were done in R (R Core Team, 2017).

3. Results

3.1. Overview of the physicochemical and biological conditions

The two polynyas differed largely in their physicochemical conditions and surface productivity, which have been described in more detail in Moreau et al. (2019) and Ratnarajah et al. (2022) and are summarized in section 2.1. In the three sampled stations, we observed a much larger fluorescence peak (proxy of phytoplankton concentrations) at the two Mertz stations (M36 and M48), compared to the Dalton station (D02), coincident with peaks in ammonia (usually produced in the euphotic zone by heterotrophic bacteria and zooplankton grazing; Smith et al., 2022, and references therein) which were also clearly higher in the Mertz stations than in the Dalton one. Surface concentrations of H_4SiO_4, PO_4^{-3}, and NO_3 were much lower in the two Mertz stations than in Dalton surface waters, which was suggested to be due to the higher consumption by phytoplankton, particularly diatoms (Moreau et al., 2019; Ratnarajah et al., 2022) (Supplementary Figure S1).

The Chl-a stocks (in the upper 20 m) were much lower in Dalton (51 mg Chl-a m^{-2}) than in the M36 and M48 Mertz stations (243 mg Chl-a m^{-2} and 300 mg Chl-a m^{-2}, respectively, Figure 1). Accordingly, POC stocks were also lowest at D02 (3.6 mg C m^{-2}) compared to M36 (8.0 mg C m^{-2}) and M48 (10.2 mg C m^{-2}). ^{234}Th export fluxes (indicative of sinking particle fluxes), as expected, followed the same trend of Chl-a and POC stocks and ranged from 167 dpm m^{-2} d^{-1} in D02, to 1,122 dpm m^{-2} d^{-1} and 1977 dpm m^{-2} d^{-1} at M36 and M48, respectively (Figure 1).

3.2. Vertical variations in taxonomic richness and composition of prokaryotic communities

The number of observed OTUs ranged between 130 and 600 across the sampled sites and depths, being lower in communities from the largest particles than in the two other size-fractions (Figure 2). In all studied communities, there was an increase in richness from the most superficial sample to the depth of maximum Chl-a, which was most pronounced in the two smallest size-fractions from M36. In stations D02 and M48, the taxonomic richness decreased from subsurface to deeper waters in the three size-fractions, although this reduction in the number of OTUs was more pronounced in station M48. The richness of the communities from station M36 remained relatively constant throughout the water column below 100 m depth, except for an increase in richness at the deepest site in the two smallest size-fractions (Figure 2). It is also worth noticing that at station M48, which was the station with the highest export flux, the richness of communities from the large particles was much more similar to the richness of the smaller size-fractions, whereas at stations D02 and M36, the large fractions showed markedly lower number of OTUs along the entire water column.

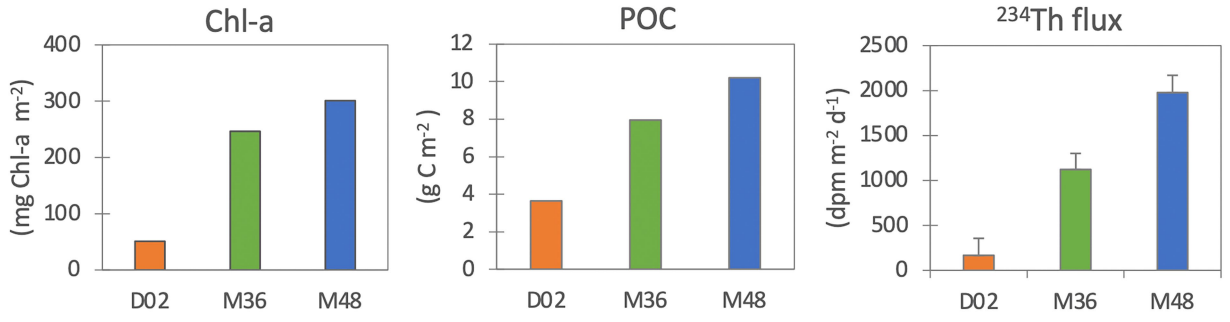


FIGURE 1
Chlorophyll-a (mg Chl-a m^{-2}), particulate organic carbon (mg C m^{-2}) stocks in the upper water column (upper 20 m) and ^{234}Th export fluxes (dpmm^{-2} d^{-1}).

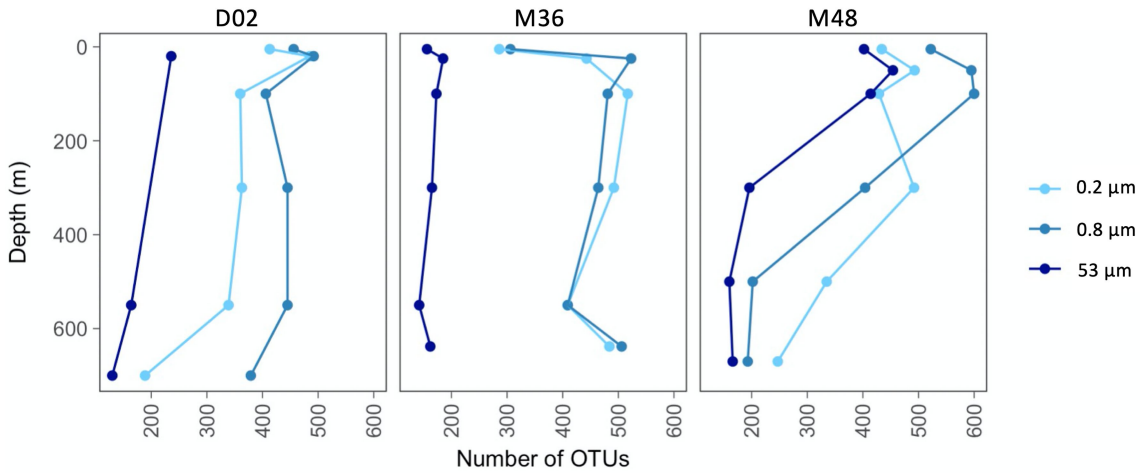


FIGURE 2
Vertical variations in taxonomic richness (number of OTUs per community) across stations and size-fractions.

Overall, the communities were dominated by class *Gammaproteobacteria* (53% of total reads), followed by *Alphaproteobacteria* (26%), *Flavobacteriia* (17%), and the Thaumarchaeota Marine Group I (1.5%). Less than 2% of the sequences were classified as Archaea and < 0.05% as Eukaryotes. Communities from the two smallest size-fractions were generally dominated by *Gammaproteobacteria* and/or *Flavobacteriia* across the three stations, whereas *Alphaproteobacteria* accounted for the majority of the sequences in most assemblages associated with the largest particles (Figure 3).

The relative contribution of these groups also changed vertically. In the free-living size-fraction (0.2 µm–0.8 µm), *Flavobacteriia* decreased their abundances from the surface to mesopelagic waters, where communities comprised mostly *Gammaproteobacteria*. *Alphaproteobacteria* decreased with depth in stations D02 and M48 but increased in M36. Particles of intermediate size were dominated mostly by *Gammaproteobacteria* and *Flavobacteriia*, but their vertical patterns differed across stations. Finally, communities associated with the largest particles showed much higher proportions of *Alphaproteobacteria*, which increased their abundances toward mesopelagic waters in the three stations. The largest vertical variations were observed in D02, where communities changed from a dominance of *Gammaproteobacteria* (95% of the sequences) in the surface toward mesopelagic assemblages comprising mostly *Alphaproteobacteria* (~94% of community sequences, Figure 3). Classes like Deltaproteobacteria, OM190, Planctomycetacia, Actinobacteria, and Sphingobacteriia were also detected locally but at much lower abundances (Figure 3).

At the genera level, the communities were more variable across stations and size-fractions (Supplementary Figure S2). Some groups like *Colwellia*, *Pseudoalteromonas*, *Balneatrix* (*Gammaproteobacteria*), and *Polaribacter* (*Flavobacteriia*) were relatively common and present in many of the studied communities, although at varying proportions with depth and size-fraction. Other groups, such as the *Alphaproteobacteria*, *Brevundimonas* or *Sphingorhabdus*, were present only in specific samples, such as in the mesopelagic large particles of D02, where together they accounted for most of the community sequences. Conversely, *Pseudophaeobacter* (*Alphaproteobacteria*) and *Alcanivorax* (*Gammaproteobacteria*) were mostly found associated with the large particles in the two Mertz stations along most of the sampled depths (Supplementary Figure S2).

All studied communities harbored a relatively large fraction of OTUs that were exclusive from a given sampling station (i.e., unique OTUs, range 14–69% of community OTUs, Figure 4A), although these accounted for a smaller fraction of local community sequences (range 1–13%) except for two communities from M48 where unique OTUs comprised 39 and 76% of local sequences (Figure 4B). In stations D02 and M36, the highest percentage of unique OTUs was found in some of the communities associated with the largest particles, whereas in M48, a similar contribution of unique OTUs was found across size-fractions (Figure 4A). These unique OTUs belonged to different orders within *Gammaproteobacteria* (mostly *Alteromonadales*, *Cellvibrionales*, and *Oceanospirillales*), *Alphaproteobacteria* (*Sphingomonadales* and *Rhodobacterales*) and *Flavobacteriia* across most size-fractions, except for a large contribution of Actinobacteria unique OTUs in the large particles from station M36 (Figure 4C).

A non-metric multidimensional scaling (NMDS) analysis showed that, when pooling all samples together, communities from the largest particles differed from assemblages associated with the other two size-fractions, which were more similar to each other. These differences between size-fractions (Figures 5A,B, $ANOSIM_{bysize}$ $R = 0.46$, $p < 0.001$) were more important than the spatial differences among stations (Figure 5A, $ANOSIM_{bystation}$ $R = 0.14$, $p < 0.001$) and the vertical differences throughout the water column (Figure 5B, $ANOSIM_{bydepth}$ $R = 0.18$, $p < 0.005$).

To determine whether the particle flux explained the vertical differences within each size-fraction, we estimated the Bray–Curtis dissimilarity between the surface (5 m) and each of the deeper prokaryotic communities and explored how these vertical differences varied along environmental gradients related to surface productivity (using Chl-a and POC as proxies) and particle flux (Figure 6). For the two smaller size-fractions, we found that, as expected, the vertical dissimilarity with the surface communities increased with depth, but no clear patterns were found along the productivity gradients (Figure 6). However, we found that the vertical differences between surface and all deeper prokaryotic communities associated with the largest particles decreased along a gradient of Chl-a and POC stocks, and with increasing ^{234}Th export fluxes, and this happened between all depths, suggesting that communities from the large size-fraction were more similar along

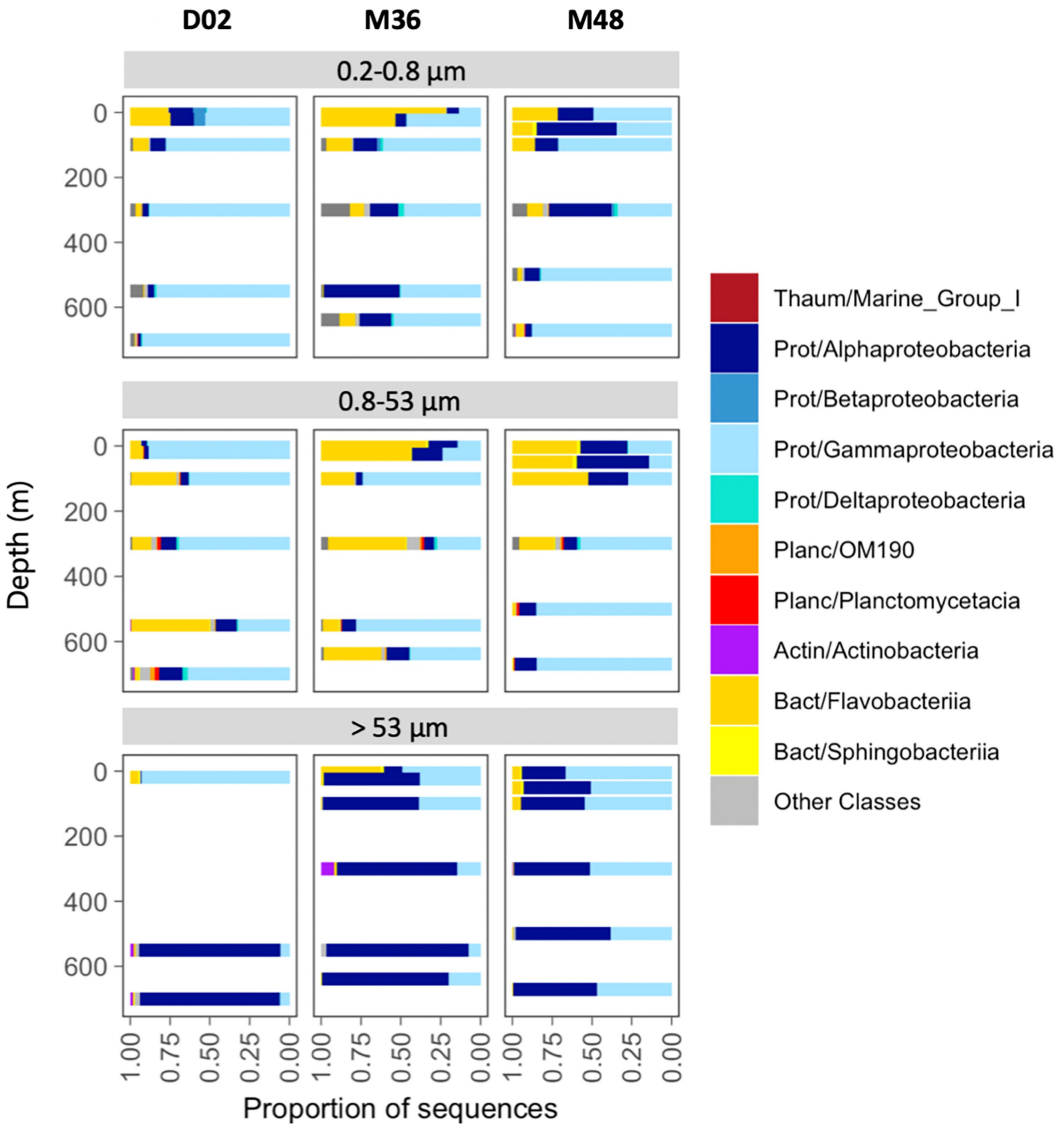


FIGURE 3
Taxonomic composition across size-fractions and depth in the three stations. The classification was performed at the class level, indicating the phylum: Thaum, thaumarchaeota; Actin, actinobacteria; Bact, bacteroidetes; Planc, planctomycetes; Prot, proteobacteria.

the water column in situations of higher surface productivity and ^{234}Th export fluxes.

Finally, we explored whether the contribution of surface-derived OTUs (i.e., OTUs with presence in any of the surface waters studied, i.e., ≤100 m) to the communities of the mesopelagic samples changed along these surface gradients related to productivity and particle fluxes (Figure 7). We found that mesopelagic communities in the >53 µm particles had a larger proportion of surface-derived OTUs along the increasing ^{234}Th flux gradient, ranging from 13 to 63% of the mesopelagic OTUs in this size-fraction (Figure 7A), accounting for 83 to 99% of the local sequences (Figure 7B). In other words, the higher the particle export flux, the higher the proportion of surface-derived taxa present in the large mesopelagic particles. This tendency was not observed for the other two size-fractions, although in general, we found that all mesopelagic communities harbored a large fraction of OTUs (range 28–52%) and of sequences (range 88–97%) which had first been detected in the overlying surface waters (Figure 7).

4. Discussion

The seasonally ice-covered coastal regions of Antarctica host a disproportionately large fraction of its primary production relative to their surface area and support complex food webs with a diversity of upper trophic levels (Arrigo and van Dijken, 2003; Karnovsky et al., 2007), yet these marginal seas and polynyas are heavily understudied. Modeling efforts suggest that about 30% of the Southern Ocean primary production is exported below the euphotic zone (Henson et al., 2012), but high variability has been reported at a smaller spatial scale. For example, in the Amundsen Sea Polynya, the reported most productive coastal Antarctic polynya, bacterial respiration remineralized >95% of the surface-derived particulate organic carbon within the upper 400 m, leading to very low export efficiencies and minimal carbon sequestration (Ducklow et al., 2015; Lee et al., 2017). Particles sinking from the surface ocean undergo remineralization by microbial communities colonizing them, leading to an attenuation of the flux of organic matter toward the deep ocean (Buesseler and Boyd, 2009). This suggests that different

10.3389/fmicb.2023.1078469

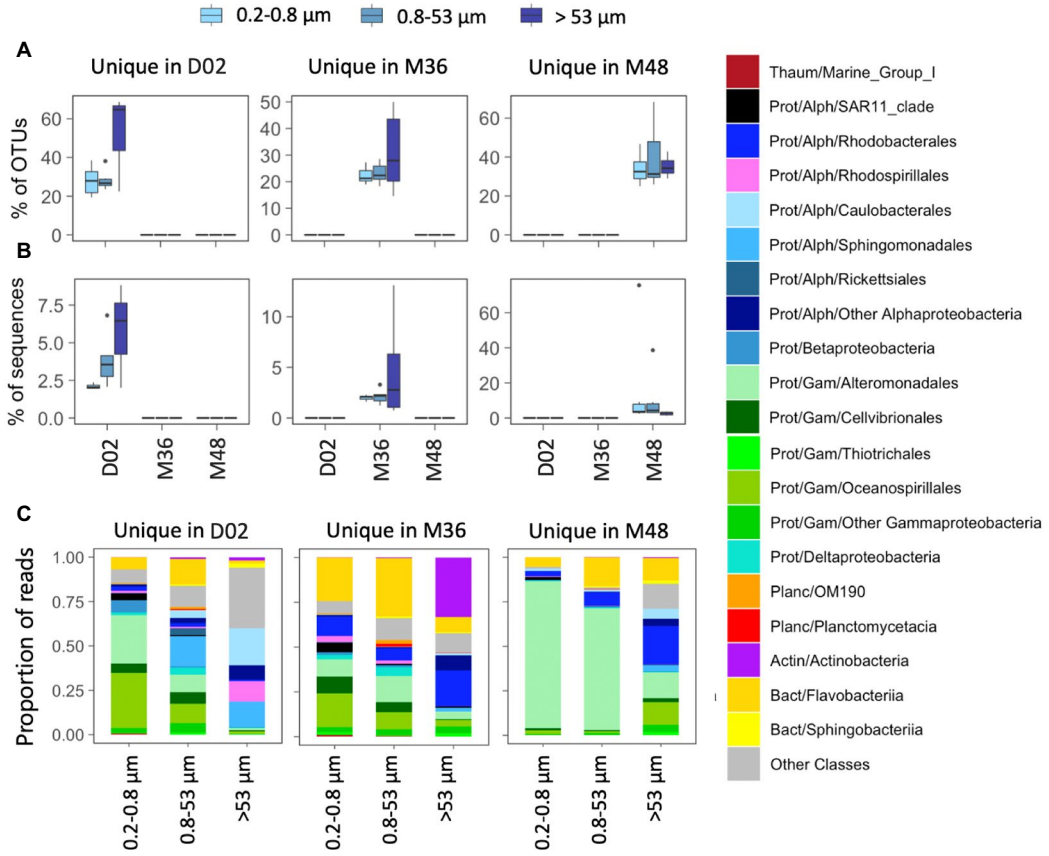


FIGURE 4
(A,B) Contribution in terms of % OTUs **(A)** or sequences **(B)** of those OTUs detected exclusively in one station ("unique" OTUs). **(C)** Taxonomic composition of unique OTUs in the three size-fractions. The classification was performed at the class level although in some cases the main orders are also indicated. The corresponding phyla and classes are indicated in each case: Thaum, thaumarchaeota; Actin, actinobacteria; Bact, bacteroidetes; Planc, planctomycetes; Prot, proteobacteria; Alph-, *Alphaproteobacteria* and Gam-, *Gammaproteobacteria*. Note the different scales of the Y axes.

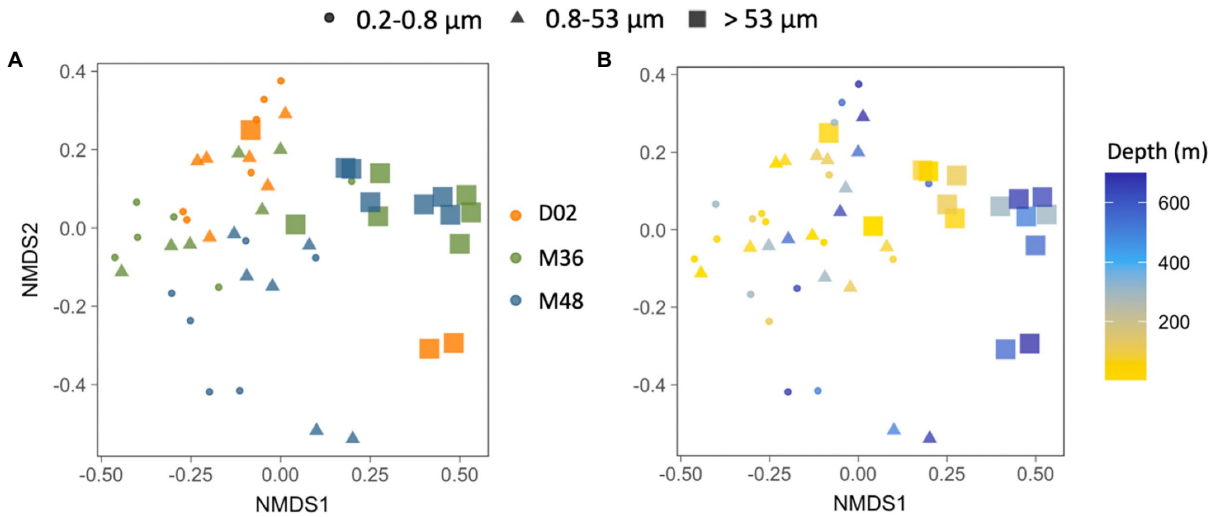


FIGURE 5
(A,B) Non-multidimensional scaling analysis (NMDS) based on the Bray–Curtis dissimilarity between all the studied prokaryotic communities, color-coded by sampling station **(A)** or by sampling depth **(B)** and indicating the different size-fractions. Stress value=0.18.

particle-attached microbial communities may sway the efficiency of the biological carbon pump in these productive systems, but very few studies have characterized prokaryotic communities associated with particles from Antarctic polynyas. Here, by coupling particle export fluxes with vertical variations in prokaryotic communities associated with different particle size-fractions, we show a large influence of surface

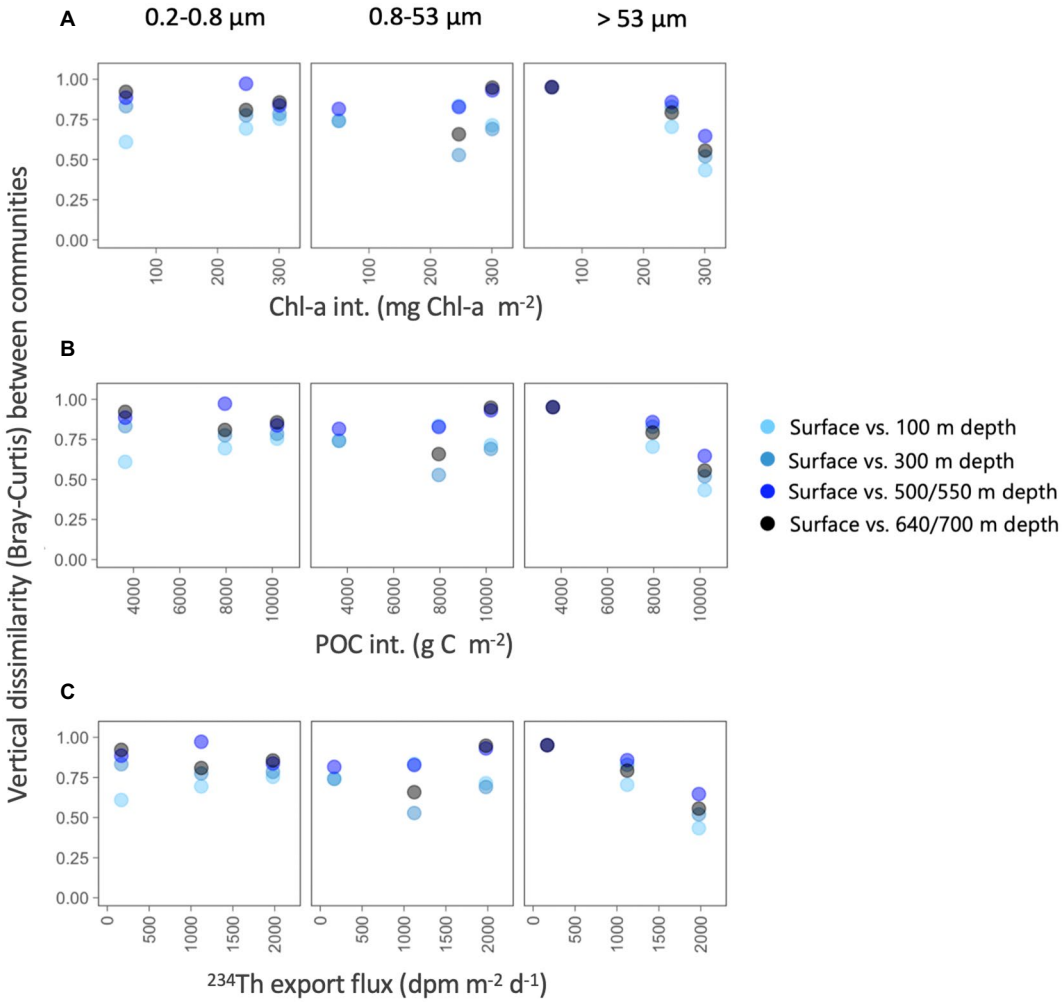


FIGURE 6
Variation in vertical taxonomic differences (Bray–Curtis dissimilarity) between surface and each of the deeper prokaryotic communities (100–700 m depth), along gradients in depth-integrated chlorophyll-a concentration, depth-integrated POC concentration and estimated ^{234}Th export fluxes for each of the three size-fractions.

productivity on the structure of mesopelagic particle-attached communities, suggesting that expected changes in Antarctic productivity due to global change may have large impacts on carbon sequestration and deep-sea microbiota.

Polynyas, despite sharing similarities in size or location, can be biogeochemically very different. These differences result from a myriad of physicochemical processes related to, among others, mixing, ice melting, ocean currents, interaction with the continental shelf, etc., and can lead to local differences in the timing, magnitude, and extension of phytoplankton blooms in these ecosystems even in closely located sites. Actually, despite their relative spatial proximity, the two polynyas sampled for this study were markedly different in terms of physicochemical and biological characteristics. As discussed in Moreau et al. (2019), the lower phytoplankton biomass and net community production in the Dalton Polynya was a consequence of the different water masses present in both polynyas, with warm modified Circumpolar Deep Water (rich in iron, an essential micronutrient that limits the growth of phytoplankton in the Southern Ocean) being widespread down to the ocean floor at Dalton, whereas it was found at shallower depths in the Mertz. Nutrients measured in both polynyas also indicated differences in the phytoplanktonic community between them, as the depletion of silica in surface waters of the Mertz Polynya coincided with the dominance of the diatoms *Fragilariopsis curta* and *F. cylindrus*, whereas small flagellates (mainly *Phaeocystis antarctica*) dominated the community in the Dalton Polynya (Moreau et al., 2019). Within the Mertz Polynya, station M36 presented clearly warmer and saltier surface waters compared to M48 and had a slightly smaller fluorescence peak that was concentrated in the upper 20 m vs. a fluorescence peak that extended down to 100 m at station M48 (Supplementary Figure S1B). M36 also had lower Chl-a and POC concentrations and stocks as well as lower ^{234}Th export flux, thus creating a gradient of surface "productivity" and particle export across the three stations (D02 < M36 < M48).

4.1. Prokaryotic communities in the Dalton and Mertz polynyas

We characterized the prokaryotic assemblages associated with three size-fractions from surface to mesopelagic waters in three stations to explore whether the different surface conditions explained variations in the prokaryotic communities. The largest size-fractions harbored the prokaryotic communities with the lowest number of OTUs at the three

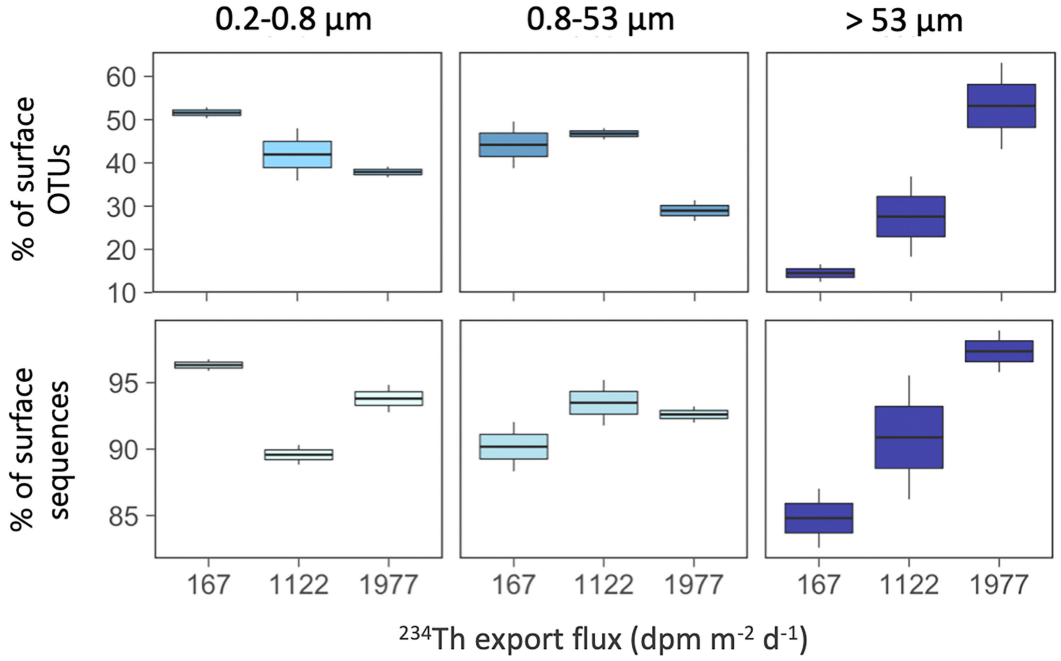


FIGURE 7
Contribution of surface-derived OTUs (top panels) and their sequences (bottom panels) in mesopelagic communities (≥500m) along the gradient of ^{234}Th export fluxes. Surface-derived OTUs are those mesopelagic OTUs that were also present in surface waters (≤100m). Note the different scales of the Y axes.

stations. This pattern is contrary to that found in previous size-fractionation studies in the Mediterranean where prokaryotic richness was shown to increase with particle size and was attributed to higher niche availability (Mestre et al., 2017, 2020) but agrees with the decrease in richness from free-living to particle-attached found during the global Malaspina ocean expedition (Salazar et al., 2015; Mestre et al., 2018).

The communities from the three size-fractions showed pronounced vertical variations in OTU number, but these did not seem related to variations in surface productivity. For example, whereas at stations D02 (the less productive) and M48 (the most productive), richness decreased from the subsurface to mesopelagic waters, at station M36, smaller vertical changes were observed, except for a subsurface (~20–50 m) peak in richness which was observed across the three stations. This contrasts with studies in the Amundsen Sea polynya showing depth-driven increases in richness in free-living prokaryotic communities (Kim et al., 2014; Richert et al., 2019), and highlights a complexity in the nature of the particles and the associated microbial communities depending on the location and likely the surface conditions.

We found a dominance of groups like *Alphaproteobacteria*, *Gammaproteobacteria*, and *Flavobacteriia* which are the most common ones in the polar oceans (Ghiglione et al., 2012), although their abundances varied depending on the size-fraction and depth. Within them, groups like *Colwellia*, *Pseudoalteromonas*, *Alcanivorax* (*Gammaproteobacteria*), *Polaribacter* (*Flavobacteriia*), and *Pseudophaeobacter* (*Alphaproteobacteria*) were the most abundant. Previous reports in the Amundsen Sea polynya have shown a surface dominance of fast-growing copiotrophs including members of *Flavobacteriia*, *Polaribacter*, *Gammaproteobacteria* SAR92, and Oceanospirillaceae and/or of different members within *Flavobacteriia* and Alpha- and *Gammaproteobacteria* in mesopelagic waters (Delmont et al., 2014; Kim et al., 2014; Richert et al., 2015, Choi et al., 2016). Marine Flavobacteria have been described as major components of marine aggregates (e.g., Zhang et al., 2007), with abundances of particle-associated Flavobacteria suggested to be related to enhanced primary production (Abell and Bowman, 2005). However, none of these studies conducted particle size-fractionation analyses, so very little is known about the particle-attached microbiome of Antarctic polynyas. In our study site, we found that their abundance drastically decreased in the largest size-fraction, yet there was a higher abundance of *Flavobacteriia* in the Mertz stations (more productive) compared to the Dalton station. We found a remarkable dominance of *Alphaproteobacteria* in the largest size-fractions across most our samples, which comprised mostly the genera *Brevundimonas* and *Sphingorhabdus* at the Dalton station, and *Pseudophaeobacter* in the case of the two Mertz stations, pointing to large differences in the particle-associated communities between both polynyas. Although *Alphaproteobacteria* are usually found as free-living (Salazar et al., 2016; Mestre et al., 2020), they have been shown to dominate large particles in some areas of the ocean (Mestre et al., 2018) and genera such as *Brevundimonas* have been detected in late stages of particle colonization (Pelve et al., 2017). The differences in the community composition of the large particles from deep waters between the two polynyas could be due to differences in the origin and composition of particles, as well as to differences in the prokaryotic surface inocula (Mestre et al., 2018; Ruiz-González et al., 2020). Actually, we found that all studied communities harbored a relatively high fraction of OTUs that were exclusive from each station (unique OTUs), and this proportion was higher in some of the largest size-fractions from the Dalton D02 and the Mertz M36 stations. This supports that the local physicochemical or biotic conditions established within each polynya may select for specific taxa or that there is dispersal limitation of species across the sampled sites. However, in general, these unique OTUs represented a small fraction of communities in most cases and communities were dominated by taxa that showed presence in both polynyas.

Surface water properties are known to mold marine microbial communities in Antarctic waters (Piquet et al., 2011; Ghiglione et al., 2012; del Negro et al., 2018). For example, Wilkins et al. (2013) conducted a metagenomic survey from Hobart to the Mertz Glacier and found different taxonomic and functional microbial assemblages north and south of the Polar Front. The differences observed concurred with the more oligotrophic characteristics found north of the Polar Front compared to the usually enhanced primary production observed in summer in the Antarctic coastal areas. However, and despite the relevance of Antarctic polynyas in carbon cycling, few studies have characterized their pelagic prokaryotic communities, and most have focused on polynyas from the Amundsen Sea and on the free-living fraction of prokaryotic communities (e.g., Delmont et al., 2014; Kim et al., 2014; Richert et al., 2015, 2019; Choi et al., 2016). These studies have reported vertical differences in the abundances of several prokaryotic groups at different areas within the polynya (Kim et al., 2014) as well as clear linkages between phytoplankton communities and prokaryotic assemblages (Delmont et al., 2014; Kim et al., 2014; Richert et al., 2019). For example, Kim et al. (2014) found that free-living bacterioplankton abundance was strongly correlated with the abundance of *Phaeocystis* spp. and diatoms, and Richert et al. (2019) reported increases in surface bacterioplankton abundance as chlorophyll increased in surface waters during a phytoplankton bloom. Delmont et al. (2014) found that groups such as the SAR92 clade and *Colwellia*, which dominated different size-fractions and depths in our study, were prevalent particle-attached prokaryotic communities at the surface and at 250 m depth, respectively, in the Amundsen Polynya during a *Phaeocystis* bloom, suggesting that they may play important roles at different stages of the bloom. Our results also show highly different communities between surface and deep microbial communities from the two Antarctic polynyas, but the magnitude of these vertical changes differed largely between the studied sites and size-fractions.

4.2. Links between surface and mesopelagic communities

The differences in surface conditions and phytoplankton communities (*Phaeocystis* vs. diatoms) between the two studied polynyas translated into different particle fluxes (estimated through ^{234}Th export fluxes) and carbon export efficiency, which were significantly lower in the Dalton Polynya compared with the Mertz Polynya (5% vs. 15%, respectively; see Ratnarajah et al., 2022). Estimates of particle export using ^{234}Th proxy have not been previously obtained in Antarctic polynyas, but based on basin-wide and global compilations (e.g., Le Moigne et al., 2013; Owens et al., 2015; Puigcorbé et al., 2017b) the ^{234}Th export flux estimated in D02 is very low, characteristic of oligotrophic and low productive ocean areas, whereas the flux at M36 is within the range of values found in temperate areas and the flux observed at M48 resembles fluxes observed under bloom conditions. POC export efficiencies (i.e., fraction of net primary production, NPP, that is exported to certain depth: POC flux/NPP*100) were found to be 3 times lower in the Dalton Polynya (5%) compared to the Mertz Polynya (15%), where transfer efficiencies down to 300 m were >80% (Ratnarajah et al., 2022).

The structure of phytoplankton communities impacts the particles produced in surface waters, consequently affecting carbon export rates (Siegel et al., 2014 and references therein). These sinking particles are known to influence the ecology and assembly of deep-sea microbial communities, not only by delivering surface-derived organic carbon (Arístegui et al., 2009; Herndl and Reinthaler, 2013), but also by directly transporting surface prokaryotes, some of which may colonize deeper waters (Mestre et al., 2018; Wenley et al., 2021). Despite the gradient in surface conditions and particle export, no clear clustering of the deep prokaryotic communities was observed by station (Figure 5), suggesting that the environmental characteristics were not different enough to lead to strong taxonomic variation between the three stations, not even between the highly contrasting Dalton and Mertz polynyas, when all samples were considered together. However, we found that the vertical dissimilarity between surface and deep communities was reduced (i.e., similarity was higher) when there was an increase in the Chl-a and POC inventories or when the flux of particles increased; in other words, the higher the production and export of particles (using Chl-a, POC and ^{234}Th deficit as proxies), the more taxonomically similar were the surface and mesopelagic prokaryotic communities associated with the largest particles. It is worth mentioning that these correlations were observed for the >53 µm size-fraction only. This agrees with the observation that particles of large sizes are more efficient vectors of diversity from the surface to the deep ocean (Mestre et al., 2018), and also with the increases in vertical similarity between surface and meso- or bathypelagic particle-attached communities with increasing surface productivity observed across the global ocean (Ruiz-González et al., 2020).

The estimated variations in particle export fluxes also coincided with a higher contribution of surface-derived OTUs to mesopelagic communities in the largest size-fractions, with the station with the highest particle export fluxes having higher contributions of surface OTUs, supporting the hypothesis of a direct transport of surface bacteria down to mesopelagic waters in highly productive Antarctic polynyas. Similarly, a recent study by Valencia et al. (2022) conducted in the Eastern North Pacific, showed that the taxa found in sinking particles were more similar to the taxa found in the upper water column in highly productive areas than in oligotrophic waters. Also, a global expedition showed that the deep-sea prokaryotic taxa with presence in the overlying surface waters were found to be mainly typical copiotrophs or eukaryote-associated groups (Ruiz-González et al., 2020). All these support that particles reaching deeper waters may comprise larger, fast-sinking material of recent phytoplankton origin that may avoid remineralization processes occurring in shallow layers (Agusti et al., 2015; Grabowski et al., 2019), representing a direct inoculation of surface particle-attached prokaryotic taxa into deep waters, and explaining the increase in the contribution of surface OTUs to large mesopelagic particles in station M48, the one with the highest estimated particle export flux.

Several previous studies have also evidenced tight linkages between the surface and deep-ocean microbial communities. For example, changes in bathypelagic prokaryotic abundance or activity have been related to high carbon fluxes or surface primary production in different oceanic sites (Nagata et al., 2000; Hansell and Ducklow, 2003; Tamburini and Garcin, 2003; Yokokawa et al., 2013) and taxonomic shifts in deep-sea communities have been linked to spatial or temporal variations in surface conditions related to particle formation and sinking (Cram et al., 2015; Parada and Fuhrman, 2017; Santoro et al., 2017; Ruiz-González et al., 2020; Wenley et al., 2021). To our knowledge, however, very few studies have compared the vertical changes in prokaryotic community composition with particle export fluxes, except for the study by Poff et al. (2021), who found that in situations of elevated carbon flux events in the North Pacific Subtropical Gyre, particle-attached bacteria reaching abyssal depths had surface water origins, and Fadeev et al. (2021), who observed that ice-covered areas in the Arctic had higher

carbon export and were also associated with lower dissimilarity between surface and deep sea microbial clades. Our study is in line with these recent findings and provides the first direct attempt to link simultaneous analyses of particle flux with the microbiome of particles in Antarctic polynyas. Our results suggest that changes in surface phytoplankton assemblages and/or productivity may strongly affect the deep-ocean microbial communities associated with the largest sinking particles.

5. Conclusion

The opening timings and size of Antarctic polynyas are likely to increase in coming years due to climate change, which will lead to changes in phytoplankton blooms (Montes-Hugo and Yuan, 2012; Deppeler and Davidson, 2017), as well as changes in the structure and function of microbial communities in these highly productive ecosystems (Piquet et al., 2011; Hörstmann et al., 2022) with yet unknown consequences for ecosystem functioning and carbon export production (Fan et al., 2020). Here, we evidenced that the vertical structuring of particle-attached microbial communities from two contrasting polynyas differed markedly depending on surface conditions. In general terms, there was a segregation of communities between surface and mesopelagic waters, yet differences were less strong at certain stations (M48) and particularly for communities associated with the largest size-fraction.

Alphaproteobacteria, *Gammaproteobacteria*, and *Flavobacteriia* dominated all communities, but different genera were found at different depths and size-fractions. The observed vertical, spatial, and size-fraction dependent variations in prokaryotic community composition support that there are different microbial niches within and across the studied polynyas. Estimates of particle export fluxes based on ^{234}Th coincided with higher chlorophyll-a and particulate organic carbon concentrations and stocks. We found that when the particle flux was higher, mesopelagic prokaryotic communities from the largest size-fraction were more similar to those found in sunlit waters and contained higher proportions of surface-derived taxa, evidencing an intense downward transport of surface bacteria mediated by the largest particles. In consequence, surface conditions will influence differently the deep ocean prokaryotic assemblages depending on the size, origin, and composition of the sinking material. To our knowledge, this is the first direct evidence of compositional differences in particle-attached assemblages linked to the magnitude of the particle flux in Antarctic polynyas. Further research is needed to stablish a mechanistical model that could allow to predict bacterial community structures based on particle export and surface productivity.

Data availability statement

The raw sequence data have been deposited in the Figshare data repository, together with the non-rarefied OTU table, the taxonomy table and the environmental data used in this study, doi: 10.6084/m9.figshare.21385215.v1.

Author contributions

VP, JG, and PM participated in the design of the sampling scheme. VP performed the sampling and sample processing. VP and CR-G compiled the needed data, analyzed the data, and wrote this article. All authors contributed to the article and approved the submitted version.

Funding

The project that gave rise to these results received the support of a fellowship to VP from "la Caixa" Foundation (ID 100010434) and from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement no 847648 (fellowship code LCF/BQ/PI21/11830020). VP also received funding from Edith Cowan University (G1003456) and from the School of Science at Edith Cowan University (G1003362) to support this work. CR-G was supported by the grants RTI2018-101025-B-I00 and a Ramon y Cajal contract (RYC2019-026758-I) and JG by grants CTM2015-70340-R and PID2021-125469NB-C31 of the Spanish Ministry of Science and Innovation and by the Generalitat de Catalunya Consolidated Research Group 2017SGR/1568.

Acknowledgments

We would like to thank the RSV Aurora Australis crew and the Australian Antarctic Division technical support crew for their assistance and their great work during the 2016/2017 AA-v2 voyage (AU1602). Special thanks to the biogeochemistry group, Delphine Lannuzel, Sébastien Moreau, Matthew Corkill, Julie Janssens, Lavenia Ratnarajah, Mar Arroyo, and Cristina Genovese, also like to thank CSIRO hydrochemists for the analyses of nutrients. We warmly thank Montserrat Roca-Martí for her invaluable help with sample collection and preparation onboard of the RSV Aurora Australis, and Vanessa Balagué for DNA extraction and initial data processing. The IAEA is grateful for the support provided to its Environment Laboratories by the Government of the Principality of Monaco. This work acknowledges the "Severo Ochoa Centre of Excellence" accreditation (CEX2019-000928-S).

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/fmicb.2023.1078469/full#supplementary-material

References

Abell, G. C. J., and Bowman, J. P. (2005). Ecological and biogeographic relationships of class Flavobacteria in the Southern Ocean. *FEMS Microbiol. Ecol.* 51, 265–277. doi: 10.1016/j.femsec.2004.09.001

Agusti, S., González-Gordillo, J. I., Vaqué, D., Estrada, M., Cerezo, M. I., Salazar, G., et al. (2015). Ubiquitous healthy diatoms in the deep sea confirm deep carbon injection by the biological pump. *Nat. Commun.* 6:7608. doi: 10.1038/ncomms8608

Arístegui, J., Gasol, J. M., Duarte, C. M., and Herndld, G. J. (2009). Microbial oceanography of the dark ocean's pelagic realm. *Limnol. Oceanogr.* 54, 1501–1529. doi: 10.4319/lo.2009.54.5.1501

Armstrong, F. A. J., Stearns, C. R., and Strickland, J. D. H. (1967). The measurement of upwelling and subsequent biological process by means of the Technicon autoanalyzer® and associated equipment. *Deep-Sea Res. Oceanogr. Abstr.* 14, 381–389. doi: 10.1016/0011-7471(67)90082-4

Arrigo, K. R., and van Dijken, G. L. (2003). Phytoplankton dynamics within 37 Antarctic coastal polynya systems. *J. Geophys. Res. Oceans* 108:108. doi: 10.1029/2002JC001739

Arrigo, K. R., Dijken, G. L.Van, and Bushinsky, S. (2008). Primary production in the Southern Ocean, 1997–2006. *J. Geophys. Res. Oceans* 113:113. doi: 10.1029/2007JC004551

Arrigo, K. R., Dijken, G. L.Van, and Strong, A. L. (2015). Environmental controls of marine productivity hot spots around Antarctica. *J Geophys Res Oceans* 120, 5545–5565. doi: 10.1002/2015JC010888

Black, E. E., Lam, P. J., Lee, J. -M., and Buesseler, K. O. (2019). Insights from the ^{238}U-^{234}Th method into the coupling of biological export and the cycling of cadmium, cobalt, and manganese in the Southeast Pacific Ocean. *Global Biogeochem. Cycles* 33, 15–36. doi: 10.1029/2018GB005985

Buesseler, K. O., and Boyd, P. W. (2009). Shedding light on processes that control particle export and flux attenuation in the twilight zone of the open ocean. *Limnol. Oceanogr.* 54, 1210–1232. doi: 10.4319/lo.2009.54.4.1210

Ceballos-Romero, E., Buesseler, K. O., and Villa-Alfageme, M. (2022). Revisiting five decades of ^{234}Th data: a comprehensive global oceanic compilation. *Earth Syst Sci Data* 14, 2639–2679. doi: 10.5194/essd-14-2639-2022

Choi, S.-B., Kim, J.-G., Jung, M.-Y., Kim, S.-J., Min, U.-G., Si, O.-J., et al. (2016). Cultivation and biochemical characterization of heterotrophic bacteria associated with phytoplankton bloom in the Amundsen Sea polynya, Antarctica. *Deep-Sea Res. II Top. Stud. Oceanogr.* 123, 126–134. doi: 10.1016/j.dsr2.2015.04.027

Clevenger, S. J., Benitez-Nelson, C. R., Drysdale, J., Pike, S., Puigcorbé, V., and Buesseler, K. O. (2021). Review of the analysis of ^{234}Th in small volume (2–4 L) seawater samples: improvements and recommendations. *J. Radioanal. Nucl. Chem.* 329, 1–13. doi: 10.1007/s10967-021-07772-2

Cram, J. A., Chow, C.-E. T., Sachdeva, R., Needham, D. M., Parada, A. E., Steele, J. A., et al. (2015). Seasonal and interannual variability of the marine bacterioplankton community throughout the water column over ten years. *ISME J.* 9, 563–580. doi: 10.1038/ismej.2014.153

de Jong, J., Schoemann, V., Lannuzel, D., Croot, P., de Baar, H., and Tison, J.-L. (2012). Natural iron fertilization of the Atlantic sector of the Southern Ocean by continental shelf sources of the Antarctic peninsula. *J. Geophys. Res.* 117:G01029. doi: 10.1029/2011JG001679

del Castillo, C. E., Signorini, S. R., Karaköylü, E. M., and Rivero-Calle, S. (2019). Is the Southern Ocean getting greener? *Geophys. Res. Lett.* 46, 6034–6040. doi: 10.1029/2019GL083163

del Negro, P., Celussi, M., de Vittor, C., and Fonda Umani, S. (2018). Rapid acclimation of microbes to changing substrate pools in epipelagic waters of an Antarctic polynya during austral summer 2003. *Polar Biol.* 41, 1–10. doi: 10.1007/s00300-017-2165-5

Delmont, T. O., Hammar, K. M., Ducklow, H. W., Yager, P. L., and Post, A. F. (2014). Phaeocystis Antarctica blooms strongly influence bacterial community structures in the Amundsen Sea polynya. *Front. Microbiol.* 5:646. doi: 10.3389/fmicb.2014.00646

Deppeler, S. L., and Davidson, A. T. (2017). Southern Ocean phytoplankton in a changing climate. *Front. Mar. Sci.* 4:40. doi: 10.3389/fmars.2017.00040

Ducklow, H. W., Wilson, S. E., Post, A. F., Stammerjohn, S. E., Erickson, M., Lee, S., et al. (2015). Particle flux on the continental shelf in the Amundsen Sea polynya and Western Antarctic peninsula. *Elementa Sci. Anthropocene* 3:46. doi: 10.12952/journal.elementa.000046

Ducklow, H. W., and Yager, P. L. (2007). "Chapter 10 pelagic bacterial processes in polynyas" in *Polynyas: Windows to the World Elsevier Oceanography Series*. eds. W. O. Smith and D. G. Barber (Elsevier), 323–361.

Edgar, R. C. (2013). UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat. Methods* 10, 996–998. doi: 10.1038/nmeth.2604

Fadeev, E., Rogge, A., Ramondenc, S., Nöthig, E.-M., Wekerle, C., Bienhold, C., et al. (2021). Sea ice presence is linked to higher carbon export and vertical microbial connectivity in the Eurasian Arctic Ocean. *Commun Biol* 4:1255. doi: 10.1038/s42003-021-02776-w

Fan, G., Han, Z., Ma, W., Chen, S., Chai, F., Mazloff, M. R., et al. (2020). Southern Ocean carbon export efficiency in relation to temperature and primary productivity. *Sci. Rep.* 10:13494. doi: 10.1038/s41598-020-70417-z

Ghiglione, J.-F., Galand, P. E., Pommier, T., Pedrós-Alió, C., Maas, E. W., Bakker, K., et al. (2012). Pole-to-pole biogeography of surface and deep marine bacterial communities. *Proc. Natl. Acad. Sci.* 109, 17633–17638. doi: 10.1073/pnas.1208160109

Grabowski, E., Letelier, R. M., Laws, E. A., and Karl, D. M. (2019). Coupling carbon and energy fluxes in the North Pacific subtropical gyre. *Nat. Commun.* 10:1895. doi: 10.1038/s41467-019-09772-z

Gruber, N., Gloor, M., Fletcher, S. E. M., Doney, S. C., Dutkiewicz, S., Follows, M. J., et al. (2009). Oceanic sources, sinks, and transport of atmospheric CO2. *Global Biogeochem. Cycles* 23:GB1005. doi: 10.1029/2008GB003349

Gustafsson, Ö., Gschwend, P. M., and Buesseler, K. O. (1997). Using ^{234}Th disequilibria to estimate the vertical removal rates of polycyclic aromatic hydrocarbons from the surface ocean. *Mar. Chem.* 57, 11–23. doi: 10.1016/S0304-4203(97)00011-X

Hansell, D. A., and Ducklow, H. W. (2003). Bacterioplankton distribution and production in the bathypelagic ocean: directly coupled to particulate organic carbon export? *Limnol. Oceanogr.* 48, 150–156. doi: 10.4319/lo.2003.48.1.0150

Henson, S., le Moigne, F., and Giering, S. (2019). Drivers of carbon export efficiency in the Global Ocean. *Global Biogeochem. Cycles* 33, 891–903. doi: 10.1029/2018GB006158

Henson, S. A., Sanders, R., and Madsen, E. (2012). Global patterns in efficiency of particulate organic carbon export and transfer to the deep ocean. *Global Biogeochem. Cycles* 26:4099. doi: 10.1029/2011GB004099

Herndl, G. J., and Reinthaler, T. (2013). Microbial control of the dark end of the biological pump. *Nat. Geosci.* 6, 718–724. doi: 10.1038/ngeo1921

Hoppema, M., and Anderson, L. G. (2007). "Chapter 6 biogeochemistry of polynyas and their role in sequestration of anthropogenic constituents" in *Polynyas: Windows to the World Elsevier Oceanography Series*. eds. W. O. Smith and D. G. Barber (Elsevier), 193–221.

Hörstmann, C., Buttigieg, P. L., John, U., Raes, E. J., Wolf-Gladrow, D., Bracher, A., et al. (2022). Microbial diversity through an oceanographic lens: refining the concept of ocean provinces through trophic-level analysis and productivity-specific length scales. *Environ. Microbiol.* 24, 404–419. doi: 10.1111/1462-2920.15832

Jena, B., and Pillai, A. N. (2020). Satellite observations of unprecedented phytoplankton blooms in the Maud rise polynya, Southern Ocean. *Cryosphere* 14, 1385–1398. doi: 10.5194/tc-14-1385-2020

Kang, S.-H., Kang, J.-S., Lee, S., Chung, K. H., Kim, D., and Park, M. G. (2001). Antarctic phytoplankton assemblages in the marginal ice zone of the northwestern Weddell Sea. *J. Plankton Res.* 23, 333–352. doi: 10.1093/plankt/23.4.333

Karnovsky, N., Ainley, D. G., and Lee, P. (2007). "Chapter 12 the impact and importance of production in polynyas to top-trophic predators: three case histories" in *Polynyas: Windows to the World Elsevier Oceanography Series*. eds. W. O. Smith and D. G. Barber (Elsevier), 391–410.

Kérouel, R., and Aminot, A. (1997). Fluorometric determination of ammonia in sea and estuarine waters by direct segmented flow analysis. *Mar. Chem.* 57, 265–275. doi: 10.1016/S0304-4203(97)00040-6

Kim, J.-G., Park, S.-J., Quan, Z.-X., Jung, M.-Y., Cha, I.-T., Kim, S.-J., et al. (2014). Unveiling abundance and distribution of planktonic bacteria and archaea in a polynya in Amundsen Sea, Antarctica. *Environ. Microbiol.* 16, 1566–1578. doi: 10.1111/1462-2920.12287

Kirchman, D. L., Morán, X. A. G., and Ducklow, H. (2009). Microbial growth in the polar oceans—role of temperature and potential impact of climate change. *Nat. Rev. Microbiol.* 7, 451–459. doi: 10.1038/nrmicro2115

Knap, A. H., Michaels, A., Close, A. R., Ducklow, H., and Dickson, A. G. (1996). Protocols for the joint global ocean flux study (JGOFS) core measurements. *J. GOFS* 19, 1–210.

Lannuzel, D., Schoemann, V., de Jong, J., Pasquer, B., van der Merwe, P., Masson, F., et al. (2010). Distribution of dissolved iron in Antarctic Sea ice: spatial, seasonal, and inter-annual variability. *J. Geophys. Res. Biogeosci.* 115:1031. doi: 10.1029/2009JG001031

le Moigne, F. A. C., Henson, S. A., Cavan, E., Georges, C., Pabortsava, K., Achterberg, E. P., et al. (2016). What causes the inverse relationship between primary production and export efficiency in the Southern Ocean? *Geophys. Res. Lett.* 43, 4457–4466. doi: 10.1002/2016GL068480

le Moigne, F. A. C., Henson, S. A., Sanders, R. J., and Madsen, E. (2013). Global database of surface ocean particulate organic carbon export fluxes diagnosed from the ^{234}Th technique. *Earth Syst Sci Data* 5, 295–304. doi: 10.5194/essd-5-295-2013

Lee, S., Hwang, J., Ducklow, H. W., Hahm, D., Lee, S. H., Kim, D., et al. (2017). Evidence of minimal carbon sequestration in the productive Amundsen Sea polynya. *Geophys. Res. Lett.* 44, 7892–7899. doi: 10.1002/2017GL074646

Logares, R. (2017). Workflow for Analysing MiSeq Amplicons Based on Uparse v1.5. doi: 10.5281/zenodo.259579

Luria, C. M., Amaral-Zettler, L. A., Ducklow, H. W., and Rich, J. J. (2016). Seasonal succession of free-living bacterial communities in coastal waters of the Western Antarctic peninsula. *Front. Microbiol.* 7:1731. doi: 10.3389/fmicb.2016.01731

Maiti, K., Charette, M. A., Buesseler, K. O., and Kahru, M. (2013). An inverse relationship between production and export efficiency in the Southern Ocean. *Geophys. Res. Lett.* 40, 1557–1561. doi: 10.1002/grl.50219

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

Mestre, M., Borrull, E., Sala, M. M., and Gasol, J. M. (2017). Patterns of bacterial diversity in the marine planktonic particulate matter continuum. *ISME J.* 11, 999–1010. doi: 10.1038/ismej.2016.166

Mestre, M., Höfer, J., Sala, M. M., and Gasol, J. M. (2020). Seasonal variation of bacterial diversity along the marine particulate matter continuum. *Front. Microbiol.* 11:1590. doi: 10.3389/fmicb.2020.01590

Mestre, M., Ruiz-González, C., Logares, R., Duarte, C. M., Gasol, J. M., and Sala, M. M. (2018). Sinking particles promote vertical connectivity in the ocean microbiome. *Proc. Natl. Acad. Sci.* 115, E6799–E6807. doi: 10.1073/pnas.1802470115

Miller, L. A., and DiTullio, G. R. (2007). "Chapter 5 gas fluxes and dynamics in polynyas" in *Polynyas: Windows to the World Elsevier Oceanography Series*. eds. W. O. Smith and D. G. Barber (Elsevier), 163–191.

Montes-Hugo, M. A., and Yuan, X. (2012). Climate patterns and phytoplankton dynamics in Antarctic latent heat polynyas. *J Geophys Res Oceans* 117:6597. doi: 10.1029/2010JC006597

Moreau, S., Lannuzel, D., Janssens, J., Arroyo, M. C., Corkill, M., Cougnon, E., et al. (2019). Sea-ice meltwater and circumpolar deep water drive contrasting productivity in three Antarctic polynyas. *J Geophys Res Oceans* 124, 2943–2968. doi: 10.1029/2019JC015071

Murphy, J., and Riley, J. P. (1962). A modified single solution method for the determination of phosphate in natural waters. *Anal. Chim. Acta* 27, 31–36. doi: 10.1016/S0003-2670(00)88444-5

Nagata, T., Fukuda, H., Fukuda, R., and Koike, I. (2000). Bacterioplankton distribution and production in deep Pacific waters: large–scale geographic variations and possible coupling with sinking particle fluxes. *Limnol. Oceanogr.* 45, 426–435. doi: 10.4319/lo.2000.45.2.0426

Nguyen, T. T. H., Zakem, E. J., Ebrahimi, A., Schwartzman, J., Caglar, T., Amarnath, K., et al. (2022). Microbes contribute to setting the ocean carbon flux by altering the fate of sinking particulates. *Nat. Commun.* 13:1657. doi: 10.1038/s41467-022-29297-2

Owens, S. A., Buesseler, K. O., and Sims, K. W. W. (2011). Re-evaluating the ^{238}U-salinity relationship in seawater: implications for the ^{238}U-^{234}Th disequilibrium method. *Mar. Chem.* 127, 31–39. doi: 10.1016/j.marchem.2011.07.005

Owens, S. A., Pike, S., and Buesseler, K. O. (2015). Thorium-234 as a tracer of particle dynamics and upper ocean export in the Atlantic Ocean. *Deep-Sea Res. II Top. Stud. Oceanogr.* 116, 42–59. doi: 10.1016/j.dsr2.2014.11.010

Parada, A. E., and Fuhrman, J. A. (2017). Marine archaeal dynamics and interactions with the microbial community over 5 years from surface to seafloor. *ISME J.* 11, 2510–2525. doi: 10.1038/ismej.2017.104

Parada, A. E., Needham, D. M., and Fuhrman, J. A. (2015). Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environ. Microbiol.* 18, 1403–1414. doi: 10.1111/1462-2920.13023

Pelve, E. A., Fontanez, K. M., and DeLong, E. F. (2017). Bacterial succession on sinking particles in the Ocean's interior. *Front. Microbiol.* 8:2269. doi: 10.3389/fmicb.2017.02269

Piquet, A. M.-T., Bolhuis, H., Meredith, M. P., and Buma, A. G. J. (2011). Shifts in coastal Antarctic marine microbial communities during and after melt water-related surface stratification. *FEMS Microbiol. Ecol.* 76, 413–427. doi: 10.1111/j.1574-6941.2011.01062.x

Poff, K. E., Leu, A. O., Eppley, J. M., Karl, D. M., and DeLong, E. F. (2021). Microbial dynamics of elevated carbon flux in the open ocean's abyss. *Proc. Natl. Acad. Sci.* 118:e2018269118. doi: 10.1073/pnas.2018269118

Puigcorbé, V., Masqué, P., and le Moigne, F. A. C. (2020). Global database of ratios of particulate organic carbon to thorium-234 in the ocean: improving estimates of the biological carbon pump. *Earth Syst Sci Data* 12, 1267–1285. doi: 10.5194/essd-12-1267-2020

Puigcorbé, V., Roca-Martí, M., Masqué, P., Benitez-Nelson, C. R., Rutgers, V. D., Loeff, M., et al. (2017a). Particulate organic carbon export across the Antarctic circumpolar current at 10°E: differences between north and south of the Antarctic polar front. *Deep-Sea Res. II Top. Stud. Oceanogr.* 138, 86–101. doi: 10.1016/j.dsr2.2016.05.016

Puigcorbé, V., Roca-Martí, M., Masqué, P., Benitez-Nelson, C., Rutgers Van Der Loeff, M., Bracher, A., et al. (2017b). Latitudinal distributions of particulate carbon export across the North Western Atlantic Ocean. *Deep Sea Res 1 Oceanogr Res Pap* 129, 116–130. doi: 10.1016/j.dsr.2017.08.016

Ratnarajah, L., Puigcorbé, V., Moreau, S., Roca-Martí, M., Janssens, J., Corkill, M., et al. (2022). Distribution and export of particulate organic carbon in East Antarctic coastal polynyas. *Deep Sea Research Part I: Oceanographic Research Papers.* 103899. doi: 10.1016/j.dsr.2022.103899

R Core Team (2017). R: A Language and Environment for Statistical Computing. Available at: https://www.R-project.org/

Richert, I., Dinasquet, J., Logares, R., Riemann, L., Yager, P. L., Wendeberg, A., et al. (2015). The influence of light and water mass on bacterial population dynamics in the Amundsen Sea polynya. *Elementa Sci. Anthropocene* 3:44. doi: 10.12952/journal.elementa.000044

Richert, I., Yager, P. L., Dinasquet, J., Logares, R., Riemann, L., Wendeberg, A., et al. (2019). Summer comes to the Southern Ocean: how phytoplankton shape bacterioplankton communities far into the deep dark sea. *Ecosphere* 10:e02641. doi: 10.1002/ecs2.2641

Rosenberg, M., and Rintoul, S. (2017). *Aurora Australis Marine Science Cruise AU1602, Dalton, Mertz and Ninnis CTD's—Oceanographic Field Measurements and Analysis*. Hobart: Australian Antarctic Data Centre.

Ruiz-González, C., Mestre, M., Estrada, M., Sebastián, M., Salazar, G., Agustí, S., et al. (2020). Major imprint of surface plankton on deep ocean prokaryotic structure and activity. *Mol. Ecol.* 29, 1820–1838. doi: 10.1111/mec.15454

Salazar, G., Cornejo-Castillo, F. M., Benitez-Barrios, V., Fraile-Nuez, E., Alvarez-Salgado, X. A., Duarte, C. M., et al. (2016). Global diversity and biogeography of deep-sea pelagic prokaryotes. *ISME J.* 10, 596–608. doi: 10.1038/ismej.2015.137

Salazar, G., Cornejo-Castillo, F. M., Borrull, E., Díez-Vives, C., Lara, E., Vaqué, D., et al. (2015). Particle-association lifestyle is a phylogenetically conserved trait in bathypelagic prokaryotes. *Mol. Ecol.* 24, 5692–5706. doi: 10.1111/mec.13419

Santoro, A. E., Saito, M. A., Goepfert, T. J., Lamborg, C. H., Dupont, C. L., and DiTullio, G. R. (2017). Thaumarchaeal ecotype distributions across the equatorial Pacific Ocean and their potential roles in nitrification and sinking flux attenuation. *Limnol. Oceanogr.* 62, 1984–2003. doi: 10.1002/lno.10547

Sedwick, P. N., and DiTullio, G. R. (1997). Regulation of algal blooms in Antarctic shelf waters by the release of iron from melting sea ice. *Geophys. Res. Lett.* 24, 2515–2518. doi: 10.1029/97GL02596

Shadwick, E. H., Rintoul, S. R., Tilbrook, B., Williams, G. D., Young, N., Fraser, A. D., et al. (2013). Glacier tongue calving reduced dense water formation and enhanced carbon uptake. *Geophys. Res. Lett.* 40, 904–909. doi: 10.1002/grl.50178

Siegel, D. A., Buesseler, K. O., Doney, S. C., Sailley, S. F., Behrenfeld, M. J., and Boyd, P. W. (2014). Global assessment of ocean carbon export by combining satellite observations and food-web models. *Global Biogeochem. Cycles* 28, 181–196. doi: 10.1002/2013GB004743

Smith, S., Altieri, K. E., Mdutyana, M., Walker, D. R., Parrott, R. G., Gallie, S., et al. (2022). Biogeochemical controls on ammonium accumulation in the surface layer of the Southern Ocean. *Biogeosciences* 19, 715–741. doi: 10.5194/bg-19-715-2022

Tamburini, C., and Garcin, J. (2003). Role of deep-sea bacteria in organic matter mineralization and adaptation to hydrostatic pressure conditions in the NW Mediterranean Sea. *Aquat. Microb. Ecol.* 32, 209–218. doi: 10.3354/ame032209

Tesán Onrubia, J. A., Petrova, M. V., Puigcorbé, V., Black, E. E., Valk, O., Dufour, A., et al. (2020). Mercury export flux in the Arctic Ocean estimated from ^{234}Th/^{238}U disequilibria. *ACS Earth Space Chem* 4, 795–801. doi: 10.1021/acsearthspacechem.0c00055

Valencia, B., Stukel, M. R., Allen, A. E., McCrow, J. P., Rabines, A., and Landry, M. R. (2022). Microbial communities associated with sinking particles across an environmental gradient from coastal upwelling to the oligotrophic ocean. *Deep-Sea Res. I Oceanogr. Res. Pap.* 179:103668. doi: 10.1016/j.dsr.2021.103668

Wenley, J., Currie, K., Lockwood, S., Thomson, B., Baltar, F., and Morales, S. E. (2021). Seasonal prokaryotic community linkages between surface and Deep Ocean water. *Front. Mar. Sci.* 8:659641. doi: 10.3389/fmars.2021.659641

Wiedmann, I., Ceballos-Romero, E., Villa-Alfageme, M., Renner, A. H. H., Dybwad, C., van der Jagt, H., et al. (2020). Arctic observations identify phytoplankton community composition as driver of carbon flux attenuation. *Geophys. Res. Lett.* 47:e2020GL087465. doi: 10.1029/2020GL087465

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

Wood, E. D., Armstrong, F. A. J., and Richards, F. A. (1967). Determination of nitrate in sea water by cadmium-copper reduction to nitrite. *J. Mar. Biol. Assoc. U. K.* 47, 23–31. doi: 10.1017/S002531540003352X

Yokokawa, T., Yang, Y., Motegi, C., and Nagata, T. (2013). Large-scale geographical variation in prokaryotic abundance and production in meso- and bathypelagic zones of the Central Pacific and Southern Ocean. *Limnol. Oceanogr.* 58, 61–73. doi: 10.4319/lo.2013.58.1.0061

Zhang, J., Kobert, K., Flouri, T., and Stamatakis, A. (2014). PEAR: a fast and accurate Illumina paired-end reAd mergeR. *Bioinformatics* 30, 614–620. doi: 10.1093/bioinformatics/btt593

Zhang, R., Liu, B., Lau, S. C. K., Ki, J.-S., and Qian, P.-Y. (2007). Particle-attached and free-living bacterial communities in a contrasting marine environment: Victoria Harbor, Hong Kong. *FEMS Microbiol. Ecol.* 61, 496–508. doi: 10.1111/j.1574-6941.2007.00353.x

frontiers | Frontiers in Marine Science

TYPE Original Research
PUBLISHED 15 May 2023
DOI 10.3389/fmars.2023.985514

OPEN ACCESS

EDITED BY
Hongbin Liu,
Hong Kong University of Science and Technology, Hong Kong SAR, China

REVIEWED BY
Heng-Lin Cui,
Jiangsu University, China
Matthias Wietz,
Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research (AWI), Germany

*CORRESPONDENCE
Guan-Jun Chen
guanjun@sdu.edu.cn
Da-Shuai Mu
dashuai.mu@sdu.edu.cn

RECEIVED 04 July 2022
ACCEPTED 02 May 2023
PUBLISHED 15 May 2023

CITATION
Xia H-F, Jia X-Y, Zhou Y-X, Du Z-J, Mu D-S and Chen G-J (2023) Comparative genomics reveal distinct potential of *Tamlana* sp. S12 for algal polysaccharide degradation. *Front. Mar. Sci.* 10:985514. doi: 10.3389/fmars.2023.985514

Comparative genomics reveal distinct potential of *Tamlana* sp. S12 for algal polysaccharide degradation

Hai-Feng Xia[1,2], Xiao-Yu Jia[3], Yan-Xia Zhou[1,2], Zong-Jun Du[1,2], Da-Shuai Mu[1,2]* and Guan-Jun Chen[1,2]*

[1]Marine College, Shandong University, Weihai, China, [2]State key Laboratory of Microbial Technology, Shandong University, Qingdao, China, [3]Department of Pulmonary and Critical Care Medicine, Weihai Municipal Hospital, Cheeloo College of Medicine. Shandong University, Weihai, China

Introduction: Macroalgae contain various polysaccharides that serve as nutrient sources Introduction: Macroalgae contain various polysaccharides that serve as nutrient sources for marine bacteria. Carbohydrate-active enzymes (CAZymes) are the primary feature of marine bacteria that utilize these polysaccharides. In this study, we describe *Tamlana* sp. S12, a novel strain of marine flavobacteria that can degrade alginate and *Laminaria japonica* biomass, isolated from the intestines of the sea cucumber *Apostichopus japonicas* collected at Weihai coast.

Methods: We sequenced the entire genome of strain S12 and constructed a phylogenetic tree using the core genome sequences of related strains. We determined the enzymatic activity of strain S12 using the DNS method and measured its growth curve under different carbon sources using spectrophotometry.

Results: Strain S12 degraded dehydrated *L. japonica* fragments as the sole nutrient source within 48h. Strain S12 harbors a diverse array of CAZymes at multiple polysaccharide utilization loci (PUL). One PUL encoding lyases from PL6, 7, and 17 families may be used for the degradation of alginate. Additionally, strain S12 harbors PULs encoding carrageenan- and agar-targeting CAZymes. Comparative analysis with related flavobacteria from *Algibacter*, *Maribacter*, and *Zobellia* showed shared CAZymes among these strains, potentially derived from a common ancestor and stably maintained within strains. Genomic signatures, algal degradation ability, and CAZyme patterns suggest that strain S12 has the potential to degrade complex algal polysaccharides.

Conclusion: These results expand our knowledge of CAZymes and enrich our understanding of how marine Flavobacteriaceae adapt to marine algal polysaccharide environments. The availability of the genome of *Tamlana* sp. S12 will be beneficial for further analyses of marine Flavobacteriaceae.

KEYWORDS

marine flavobacteria, *Tamlana*, alginate, *Laminaria japonica*, PUL, EPS, *Apostichopus japonicus*, pan-genome

1 Introduction

In marine ecosystems, significant amounts of organic matter are produced through photosynthesis by pelagic phytoplankton and benthic macroalgae (Hehemann et al., 2014; Arnosti et al., 2021). These organisms are the primary sources of polysaccharides in the marine environment (Arnosti et al., 2021), which are important nutrient sources for marine bacteria. For example, brown algae contain carboxylated polysaccharides (alginate), anionic sulfated polysaccharides (fucoidan), and laminarin, none of which are present in land plants (Zhu et al., 2016).

Polysaccharide-degrading bacteria (PDB) specialize in the degradation of algal polysaccharides and colonization of algal surfaces, which are critical events in the global carbon cycle and algal biomass utilization (Ferrer-González et al., 2021). PDB mainly comprises two classes: gammaproteobacteria in the phylum Proteobacteria and flavobacteria in the phylum Bacteroidetes. Marine bacteria have developed specific degradation mechanisms to utilize these polysaccharides, which differ significantly from the degradation of polysaccharides from terrestrial plants (Chernysheva et al., 2021). Marine flavobacteria are vital in the degradation of these complex polysaccharides (Thomas et al., 2017). Polysaccharide degradation characteristics mainly rely on dedicated genomic regions called polysaccharide utilization loci (PUL), which encode proteins that are necessary for the utilization of a specific polysaccharide (Dudek et al., 2020). Flavobacteria carbohydrate-active enzymes (CAZymes) are typically clustered with *susCD* genes in PUL for orchestrated uptake and degradation (Grondin et al., 2017). *Maribacter, Algibacter*, and *Zobellia*, members of the marine flavobacteria, comprise the core community of the cultivable surface microbiota of brown algae, with the ability to degrade algal polysaccharides (Martin et al., 2015). Cha et al. (2021) compared alginate utilization pathways in culturable bacteria isolated from Arctic and Antarctic marine environments, clarifying the function of alginate utilization pathways in Bacteroidetes. Wolter et al. (2021) compared *Maribacter dokdonensis* 62–1 with related strains at the genomic level, revealing *Maribacter* to be a versatile polysaccharide-degrading flavobacterium (Wolter et al., 2021). The study by Chernysheva et al. (2019) demonstrated the specificity of *Zobellia* in the degradation of algal polysaccharides. *Zobellia* members have been extensively studied for their utilization of algal polysaccharides such as agar, porphyrin, mannitol, and alginate, improving our understanding of the molecular mechanisms behind this process (Jam et al., 2005; Thomas et al., 2017; Groisillier et al., 2015; Dudek et al., 2020). Brunet et al. (2022) compared the ability of *Z. galactanivorans* with other *Zobellia* spp. to decompose large fresh algae, evaluating its unique role in the decomposition of large fresh algae. However, research on the ability of marine Flavobacteriaceae to attack algal tissues is still limited.

The genus *Tamlana* is classified under the family Flavobacteriaceae, and was first described by Lee (2007). Currently, there are nine species that have been validly published under the genus *Tamlana. Tamlana* species have been isolated from various sources, including seawater (Yoon et al., 2008; Liu et al., 2015; Jung et al., 2019), algae (Li et al., 2020; Li et al., 2023), marine sediment (Lee, 2007; Romanenko et al., 2014), and the gut of abalone (Cao et al., 2021). Research conducted by Li et al. (2023) has shown that members of *Tamlana* possess a wide range of CAZymes that are associated with the utilization of algal polysaccharides. *Tamlana laminarinivorans* and *Tamlana haliotis* can degrade alginate, *Tamlana sargassicola* can degrade laminarin, and *Tamlana crocina* can degrade utilizing ulvan (Li et al., 2023). Furthhermore, production of κ-carrageenase and alginate lyase have been reported (Sun et al., 2010; Mostafa et al., 2020; Yin et al., 2021). These findings suggest a clear preference among *Tamlana* strains for the utilization of algal polysaccharides.

The sea cucumber, *Apostichopus japonicus*, is an economically important fishery resource in the coastal areas of East Asia (Zhang et al., 2019). These organisms have a significant impact on the cycling of benthic organic matter by consuming various types of organic debris, bottom sediments, microorganisms, algae, and aquatic animal debris (Wang et al., 2018). Flavobacteriaceae, as a component of *A. japonicus* intestine microbiota, enhances the utilization of algal polysaccharides (Zhang et al., 2019), thus benefiting the polysaccharide metabolism of *A. japonicus*.

In this study, the novel marine Flavobacteriaceae strain S12 was isolated from the intestine of *A. japonicus*. Based on whole genome analysis and physiological characteristics, we classified strain S12 as a member of the genus *Tamlana*. Analysis also revealed that strain S12 has the potential to utilize various algal polysaccharides, and effectively degrades *L. japonica* biomass. The degradation of polysaccharides may play an important ecological function in the genus *Tamlana* (Mcbride, 2014; Jung et al., 2019).

Furthermore, we describe the CAZyme content and hydrolytic capacities of strain S12. In the pan-genomic context, we compared the CAZyme characteristics of *Tamlana* (including strain S12) and other related members of marine flavobacteria (*Algibacter*, *Maribacter*, and *Zobellia*) to illustrate the role of CAZyme and PUL-related genes in niche specialization among the four genera. *Algibacter is* most closely related to *Tamlana*, while *Maribacter* and *Zobellia* are widely referenced in research on algal polysaccharides (Wolter et al., 2021; Brunet et al., 2022). The genomic machinery for the degradation of alginate as well as red algal polysaccharides were compared to further understand the utilization ability and degradation mechanism of strain S12 for seaweed polysaccharides. Our findings establish an eco-evolutionary perspective on CAZyme-related niche specialization among marine flavobacteria and its connection to biogeochemical processes.

2 Materials and methods

2.1 Isolation and cultivation

Strain S12 (Yin et al., 2021) was isolated from the intestinal tract of *A. japonicus* collected from a pond in Weihai, Shandong, China (37°30′7″N, 122°07′24″E) in March 2013. The strain was cultured on marine agar 2216 (MA; Difco, Detroit, MI, USA) at 28° for 72 h. After several rounds of sub-culture, the 16S rRNA gene of the isolate was amplified using PCR with universal primers 27F and

1492R (Jung et al., 2019), following the protocol described in previous reports (Leblond-Bourget et al., 1996). The 16S rRNA sequence of strain S12 was extracted and aligned using EzBioCloud to preliminarily determine its taxonomic status (Yoon et al., 2017).

2.2 Phenotypic tests

Commercial systems API ZYM, 20NE, and 50CHB (BioMérieux) were used to evaluate the following enzymatic activities: Amylase (Starch), Esterase (C4) (2-naphthyl butyrate), α- Galactosidase (6-Br-2-naphthyl-αD-galactopyranoside), N-acetyl-β-glucosamindase (1-naphthyl-N-acetyl-βD-glucosaminide), Lipase (C14) (2-naphthyl myristate), α-Fucosidase (2-naphthyl-αL-fucopyranoside), α-Chymotrypsin (N-glutaryl-phenylalanine-2-naphthylamide), Cystine arylamidase (L-cystyl-2-naphthylamide), α- Mannosidase (6-Br-2-naphthyl-αD-mannopyranoside). Bacterial suspensions in a 3% NaCl solution (w/v) were inoculated onto the test plates and cultivated at 28° according to the instructions. To prepare the enzymatic reaction, 200 µL of a 1% solution of beechwood xylan (Megazyme International Ireland, Bray, Ireland) or a 0.2% solution of agar (sigma-aldrich, USA) was mixed with 300 µL of cell-free supernatant obtained from strain S12 fermentation broth after 24 hours of growth in MA medium. The supernatant was obtained by centrifugation at 12000 rpm for 6 minutes. The substrate and enzyme mixture was thoroughly mixed before initiating the enzymatic reaction. The reaction was conducted at 28°C for 30 minutes, and the change in reducing sugar before and after the reaction was determined using the DNS method. As a reference strain, *Tamlana agarivorans* KCTC 22176 was tested under the same conditions as strain S12. Alginate utilization was analyzed in optimized basic salt medium (OBSM; 1 L distilled water, 25 g NaCl, 2 g K_2HPO_4, 2 g $MgSO_4{\cdot}7H_2O$, 0.1 g $FeSO_4{\cdot}7H_2O$, nitrogen source: 5 g $NaNO_3$, pH 7.2 – 7.4). Additionally, 15g of sodium alginate (Cat. no. A2158; Sigma-Aldrich, St. Louis, MO, USA) was supplemented to achieve a concentration of 1.5% in the medium. To assess glucose utilization in strain S12, OBSM was utilized with a final concentration of 1.5% glucose. The seed solution was produced from a single colony grown in marine broth 2216 (MB) for 16 hours at 28° with an inoculation rate of 1% (v/v). *L. japonica* utilization was analyzed in a culture medium prepared from a mixture of 1% (w/v) dehydrated *L. japonica* and sterile seawater. The strains used in this experiment (except strain S12) were purchased from the Korean Collection for Type Cultures, the Japan Collection of Microorganisms, and the German Collection of Microorganisms and Cell Cultures (DSMZ). The information on experimental strains is listed in Table S1. One milliliter of the seed solution was collected, centrifuged (10000 rpm, 2 min), and washed with sterile seawater 3 times. It was then adjusted to an optical density measured at 600 nm (OD_{600}) of 0.1. The medium was inoculated with 1% (v/v) washed bacteria in triplicate, followed by cultivation at 28°C and 100 rpm for 120 h with regular photometric measurements (diluted if OD_{600} >0.4). *L. japonica* residues were well precipitated by low-speed centrifugation at 4000 rpm for 5 seconds was used to precipitate *L. japonica* residues and remove impurities while the bacterial cells remained suspended.

2.3 Biofilm measurement

To determine biofilm formation, 100 µl of 2216E liquid medium (Qingdao Hope Bio-technology Co., Qingdao, China) and OBSM+1.5% alginate solution were added into the wells of a polystyrene microplate. Next, 10 µl of an overnight culture of strain S12 in 2216E liquid medium was inoculated into each well. The plate was then incubated at 28°C for 36 h. After discarding the contents of the wells, 200µl of sterile water was used to wash the wells and remove non-adherent bacteria. The biofilm was fixed with 250 µL of 96% ethanol and air-dried. Then, 100 µl of 1% (v/v) crystal violet (Merck, Germany) was added to each well for staining for 5 minutes. After washing off the excess dye, the plate was air-dried, and then 100 µl of 33% (v/v) glacial acetic acid was added to dissolve the crystal violet. The washed samples were measured for OD_{570}, and the average value was taken as the test value (Gomes et al., 2013). The un-inoculated bacterial medium was used as a negative control, and the negative value was taken twice as the limit value.

2.4 Genome sequencing and taxonomy

Genomic DNA was extracted using a DNA extraction kit (TaKaRa Bio, Shiga, Japan) following the manufacturer's instructions. DNA sequencing was performed using both the Illumina HiSeq and SMRT platforms. The Illumina HiSeq platform was used to construct the Illumina PE library, while the SMRT platform was used for Pacbio library construction. Subsequently, the sequencing data obtained from the Illumina HiSeq platform was utilized to refine the assembly of the SMRT platform-generated data. This optimization procedure resulted in an improved quality of the assembly results. (Mu et al., 2020). The genomic DNA was fragmented using a 26G needle, and fragments > 20 kB were selected by the BluePippin system (Sage Science, Beverly, MA, USA) to prepare the DNA library. Quality control was performed using fastp (quality value threshold: 30, filter length threshold: 50bp) for Illumina and mecat2 (default parameter) for PacBio Sequel (Xiao et al., 2017). After quality control of the samples, *de novo* genome assembly was performed using sequencing data from two platforms with Unicycler (https://github.com/rrwick/Unicycler), followed by error correction using Arrow (Archibald, 2017), and gene prediction using Glimmer3 and Prodigal (Delcher et al., 2007; Hyatt et al., 2010). The complete genome sequence is available at the National Center for Biotechnology Information (NCBI) (https://www.ncbi.nlm.nih.gov/) under NZ_CP068547.

A phylogenetic analysis of 48 selected species (Supplementary Table S1) based on 120 core genes using GTDB-Tk (Chaumeil et al., 2019) was performed, including *Flavobacterium aquatile* ATCC 11947 as an outgroup, with the remaining strains derived from *Tamlana*, *Algibacter* (the genus most closely related to *Tamlana*),

and *Maribacter* and *Zobellia* (which are widely referenced in research on algal polysaccharides). Strain genome assemblies were obtained from the NCBI GenBank facility, and the genome accession numbers of strains were listed in Table S1. RAxML (version 8.2.11) was used to construct a maximum likelihood (ML) phylogenetic tree with 1,000 bootstrap replicates (Tian et al., 2020), using the general time reversible-GTR was used as the model for DNA sequence evolution. An interactive tree of life (iTOL) (https://itol.embl.de) was used for phylogenetic tree display and annotation (Letunic and Bork, 2021).

Phylogenetic analysis of the PL6, PL7, and PL17 was performed using MEGA-X (version 10.1.8). Multiple amino acid sequence alignments were conducted using Clustal W (Kumar S. et al., 2018). Three methods were used to verify the phylogenetic tree, including the Jones–Taylor–Thornton (JTT), p-distance, and subtree-pruning-regrafting (SPR) models, which were used to construct the ML (Chatukuta and Rey, 2020), neighbour-joining (NJ) (Malik et al., 2021), and maximum parsimony (MP) trees (Kumar A. et al., 2018), respectively. If there were no obvious differences, the NJ tree was shown.

2.5 Genome comparison

The bioinformatic software used to compare the genomes of strain S12 and related strains (Supplementary Table S1) included Pyani (https://github.com/widdowquinn/pyani) (Mu et al., 2020) for calculating average nucleotide identities (ANIs), and Genome-to-genome distance calculator 3.0 (GGDC) (https://ggdc.dsmz.de/home.php) (Meier-Kolthoff et al., 2022) for digital DNA-DNA (dDDH) hybridization. The dDDH comparison was performed between the genome of strain S12 and all others from *Tamlana*, with a species boundary threshold of 70% using GGDC (Meier-Kolthoff et al., 2022). Pan-genome analyses were conducted using the Bacterial Pan Genome Analysis (BPGA) tool v1.3 (Chaudhari et al., 2016), with clustering performed using USEARCH with a 50% sequence identity cut-off. Genomic analysis results were obtained by identifying Cluster of Orthologous Groups of proteins (COG) and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases. The BPGA tool was also used to identify genes whose GC % deviated from the average GC % of strain S12 by more than twice the value, and to calculate the GC % of protein-coding genes of strain S12.

CAZymes were identified using blastp and dbCAN2 (Zhang H. et al., 2018), and genomes were annotated by HMMER: dbCAN (E-Value < 1e-15, coverage > 0.35), DIAMOND: CAZy (E-Value < 1e-102), and HMMER: dbCAN-sub (E-Value < 1e-15, coverage > 0.35). The density of alginate lyase genes was determined by calculating the number of alginate lyase genes present in every 1000 open reading frames (ORFs). Gene annotations were done on the CAZy database and UniProtKB-SwissProt database, and PUL boundaries were determined based on the gene annotation results and the analysis of similar PULs in the PULDB (Lombard et al., 2014; Terrapon et al., 2018; UniProt Consortium, 2021). Genes were assigned to the KEGG classes and pathways using the KofamKOALA and KEGG Mapper (Aramaki et al., 2020; Kanehisa and Sato, 2020).

PUL comparison was analyzed using PULDB, and protein function was analyzed using PDB (Burley et al., 2018). Pairwise comparisons of PULs were performed using the BLASTn (BLAST version 2.11.0+) run in EasyFig (version 2.2.5) (Sullivan et al., 2011). SignalP v5.0 was used to predict the signal peptides (Almagro Armenteros et al., 2019).

2.6 Statistical analysis

Barbato et al. (2022) identified polysaccharide lyases (PL) and glycoside hydrolases (GH) as the primary enzymes responsible for degrading algal polysaccharides. Table S2 presents data on PLs + GHs, while Table S1 shows the strain codes. Principal coordinate analysis (PCoA) was conducted at the CAZymes level using the Bray-Curtis distance. The Bray-Curtis matrix distance was built by log-transforming the relative abundance of PLs + GHs. The PCoA plot was generated using ImageGP (Chen et al., 2022). Finally, a grouped percentage stacked bar chart was used to illustrate the relative proportion of different CAZyme components in various strains and genera. This chart displays the percentage contribution of each CAZyme component across different bacterial strains and genera.

3 Results and discussion

3.1 Genomic features and taxonomic analysis of Strain S12

Strain S12 was isolated from the intestine of *A. japonicus*. Colonies on the solid medium were round, smooth, and yellowish. The strain S12 has a whole genome sequence of 3.57 Mb with a GC content of 39%, which includes 224 predicted CAZymes (Supplementary Table S1), corresponding to 7.1% of all protein-encoding genes. This value was higher than the average value for *Tamlana* (6.7%). According to genome annotation, three 16S rRNA gene copies were found in the genome of strain S12, all of which were very similar to the 16S rRNA gene of *Tamlana agarivorans* JW-26T (accession EU221275). The sequence identities, lengths, and coverage were 99.2% (1519 bp, 97% cover), 98.8% (1521 bp, 97% cover), and 98.7% (1521 bp, 97% cover), respectively. The ANI and dDDH between strain S12 and *T. agarivorans* JD-26 were 95.0% and 67.7% (Supplementary Figure S1), respectively. Core genome-based phylogeny revealed clear assignment to *Tamlana* within Flavobacteriaceae (Figure 1). According to the genomic taxonomy parameters, strain S12 might represent a novel species of *Tamlana*.

3.2 Interspecific difference comparison and Phylogenetic reconstruction in the context of CAZymes

To compare the differences in CAZyme composition between strain S12 and other related flavobacteria, 47 publicly available genomes from the four genera were compared using phylogenetic

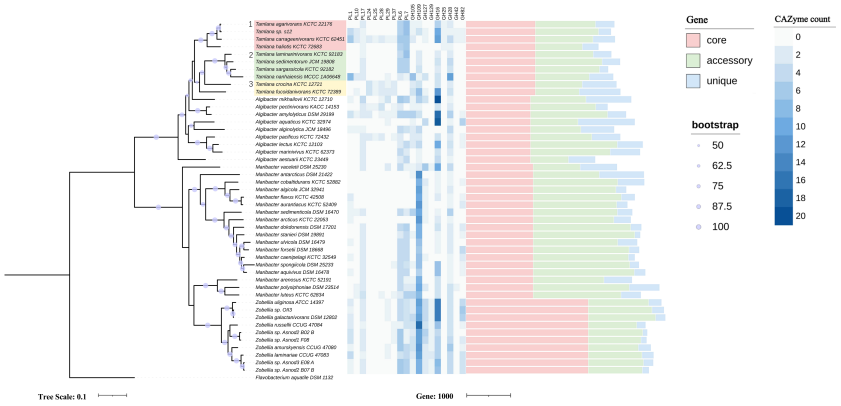


FIGURE 1
Maximum-likelihood phylogenetic tree based on 120 core genes (left panel), numbers of genes from selected carbohydrate-active enzymes (CAZyme) families (middle panel) and pan-genome (right panel) in *Tamlana* sp. S12 and related strains. *Flavobacterium aquatile* DSM 1132 served as outgroup. The three resolved *Tamlana* lineages are numbered and colored. PL, polysaccharide lyase; GH, glycoside hydrolase.

analysis in the context of CAZymes. The four genera include *Tamlana*, to which strain S12 belongs, *Algibacter*, which is most closely related to *Tamlana*, as well as *Maribacter* and *Zobellia*, which are widely referenced in research on algal polysaccharides. In terms of protein-encoding genes, *Maribacter* exhibited the lowest average proportion (5.4%) of CAZyme-encoding genes compared to the other three genera, while *Algibacter* (7.1%) and *Zobellia* (7.5%) had the highest average proportions (Table 1; Supplementary Table S1). We further calculated the standard deviation of the proportion of CAZyme among the four genera members and found that the deviation of *Zobellia* spp. was the smallest (0.4%), and the deviation of the other three genera exceeded 1%. *Zobellia* spp. showed a stable proportion of CAZyme-encoding genes, whereas the other three genera members displayed higher variability in this regard.

The PCoA analysis demonstrated differences in the CAZyme composition between strains (Figure S2). According to PCoA grouping, *Tamlana* and *Algibacter* exhibited similar CAZyme profiles. Conversely, *Zobellia* and *Maribacter* strains formed clusters and displayed distinct CAZyme profiles between the two genera. These findings are consistent with their evolutionary relationships, as supported by the phylogenetic tree (Figure 1).

We analyzed CAZyme patterns in the context of phylogenetic relationships (Figure 1, Supplementary Table S2). Core genome-based phylogeny resolved three lineages of *Tamlana* (Figure 1). Lineages 1 and 2 exhibited different PLs compared to lineage 3, including PL6, 7, and 17, which are involved in alginate degradation. Moreover, PL1 and 10 encoding pectinases were identified in lineages 1 and 2 concurrently, while PL10 was not detected in most genomes outside lineages 1 and 2. GH25 encoding lysozyme was exclusively present in lineage 1. In contrast, PL24, 25, 28, 29, and 37 encoding ulvan lyases were common features of lineage 3 (Figure 1).

PL7 and PL6 accounted for the highest proportion of CAZymes in *Tamlana*, while GH16 and GH2 proportions dominated in *Algibacter*. The proportion of GH109 in the CAZymes of *Zobellia* and *Maribacter* was higher than that in other genera (Figure 2A). GH109 is involved an NAD^+-dependent hydrolysis mechanism that cleaves *N*-acetylgalactosamine residues from various substrates, such as glycolipids, glycopeptides, and glycoproteins (Malinski and Singh, 2019; Gavriilidou et al., 2020). These substrates are abundant in the macroalgal matrix (Kalisch et al., 2016). GH109 in *Zobellia* and *Maribacter* may provide adaptive mechanisms for these substrates.

TABLE 1 CAZyme families encoded by *Tamlana* sp. S12.

CAZyme class	CAZyme family
PL	1, 6, 7, 9, 10, 17
GH	2, 3, 5, 10, 13, 16, 20, 23, 25, 28, 29, 30, 31, 42, 43, 65, 73, 74, 78, 82, 86, 97, 100, 105, 108, 109, 117, 127
CBM	4, 6, 16, 20, 22, 32, 35, 40, 44, 47, 48, 50, 58, 77
CE	1, 2, 4, 8,10, 11, 12, 14
AA	2, 6, 12
GT	2, 4, 5, 8, 9, 12, 19, 27, 28, 30, 51, 57, 81

PL, polysaccharide lyase; GH, glycoside hydrolase; GT, glycosyl transferase; CBM, carbohydrate-binding module; CE, carbohydrate esterase; AA, auxiliary carbohydrate-active oxidoreductase.

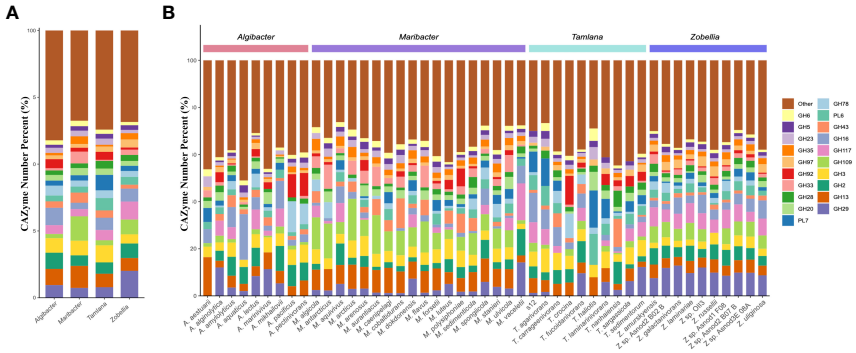


FIGURE 2
Comparison of glycoside hydrolase (GH) and polysaccharide lyase (PL) families composition diversity in the four genera studied. **(A)** The histogram and **(B)** showed the proportion of 20 PL and GH families with the highest abundance at the level of species and genera. Different color represents different CAZyme families.

3.3 Pan-genomic analysis of the four genera

The pan-genome refers to the complete genomic repertoire of a given phylogenetic clade, encoding all possible lifestyles of its organisms (Tian et al., 2016). To compare the genomic information of the four genera, we analyzed of the pan-genomes of each genus using BPGA.

In an attempt to understand the relationships between the pan-genome sizes and number, the pan- genomic numbers combined with the phylogenetic tree are shown in Figure 1, and the data are shown in Supplementary Table S3. The number of new orthologue clusters is directly proportional to the number of genomes, indicating that the four genera have open pan-genomes (Figure S3, Supplementary Table S4). The four genera share only a limited number of genes (833 core genes), with an average percentage of 23.7% for the core genes. Pan-genome analysis revealed that *Zobellia* had the highest number and average proportion of core genes (2614, 65.2%) and the lowest average proportion of unique genes (5.0%) in terms of the pan-genome. A large average proportion of accessory genes was found in the members of *Maribacter* (50.0%), while the average proportion of accessory genes in the *Tamlana* and *Algibacter* was relatively similar (38.8%, 39.8%). Compared to the other two genera, *Tamlana* and *Algibacter* were relatively similar, with a higher average proportion of unique genes (14.8%, 19.4%) (Supplementary Table S3). A significant number of hypothetical genes were identified among the unique genes found in *Tamlana*. This may be due to the fact that not all HGT events may have biological significance (Arnold et al., 2022).

We also utilized BPGA to identify and extract genes whose GC content deviated from the average GC content of strain S12 by more than twice the standard deviation (Chaudhari et al., 2016). These genes are typically acquired through HGT (Liu et al., 2009; Ravenhall et al., 2015; Chaudhari et al., 2016). We then calculated the proportion of genes with atypical GC content in the unique genome, core genome, and accessory genome of the strains. Our findings revealed that the unique genome had the highest proportion of genes with atypical GC content (Supplementary Table S3), which was significantly higher than their proportion in the core genome and accessory genome. However, the deviation between species was substantia (Supplementary Table S3).

Among the genes with atypical GC content, one (WP_068604863.1) showed 95.2% amino acid identity towards a carrageenase GH16 gene (WP_102994428) from *T. carrageenivorans*. To explore the homology of the two sequences, we selected 18 sequences with the highest amino acid sequence identity from NCBI to construct a phylogenetic tree. These two sequences clustered together in the same branch (Figure 3A).

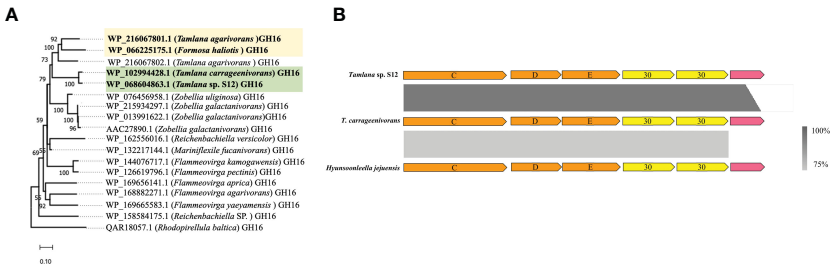


FIGURE 3
Detection of possible horizontal gene transfer (HGT) candidate genes using atypical GC% content and sequence homology analyzis. **(A)** Phylogenetic tree constructed with two GH16 (WP_102994428 and WP_216067801.1) and their closest amino acid sequences. For proteins, locus tags are listed. The organism names are listed in brackets. Bootstrap values lower than 50 are not indicated. Colored boxes mark the HGT candidate and their closest homologous protein sequences. **(B)** Comparison of three homologous polysaccharide utilisation loci (PUL). From top to bottom: a PUL with atypical GC% genes including two GH30 detected in *Tamlana* sp. S12 (top) and homologous clusters in *T. carrageenivorans* (center) and *Hyunsoonleella jejuensis* (bottom).

Although *Formosa haliotis* and *T. agarivorans* are distantly related in Flavobacteriaceae (Tanaka et al., 2015; García-López et al., 2019), GH16 sequences from these species (WP_066225175.1 and WP_216067801.1, respectively) are homologous (70.96% amino acid identity). One PUL-encoded protein from the putative GH30 family with atypical GC content was identified in the genome of strain S12. According to NCBI annotation, two GH30s in PUL were identified as glucosylceramidase, which are speculated to play a role in transglycosylation and synthesizing oligosaccharide and glycoconjugate (Ben Bdira et al., 2018). The homologous gene clusters also existed in *T. carrageenivorans* (NCBI assembly GCA_002893765.1) (>95%) and *Hyunsoonleella jejuensis* (NCBI assembly GCA_000060345.1) (>75%) (Figure 3B).

A cluster for the biosynthesis of exopolysaccharides (EPS) was found in strain S12 (Figure 4A, Supplementary Table S5). The EPS biosynthesis gene polyprenyl glycosylphosphotransferase (JJL45_RS08315) and O-antigen (JJL45_RS08305) were identified as core genes, indicating the existence of EPS in *Tamlana*. Five unique genes in the gene cluster are labelled in Figure 4A. Genes with atypical GC content were identified as members of GT2 (JJL45_RS08185) and GT4 (JJL45_RS08285). Based on the analysis, EPS production may be an adaptive feature among *Tamlana* strains, potentially mediating distinct surface adhesion and host interaction (Decho and Gutierrez, 2017).

3.4 Alginolytic PUL in strain S12

We compared the genome sizes, PUL numbers, and the number and density of alginate lyase genes (Figure 5). Compared to the other 46 strains, strain S12 had a smaller genome size (Figure 5A), and PUL density (Figure 5D), but it displayed a high density of alginate lyases (Figure 5B), indicating specialization toward alginate. The growth curve of strain S12 in various culture media showed growth with 1.5% alginate (w/v) as sole nutrient source (Figure 6). This concentration is much higher than typically used concentrations (0.05 to 0.4%), confirming its ability to utilize alginate (Figure 6). Strain S12 exhibited delayed growth with glucose after being inoculated from the MB seed culture (Figure 6B). This was potentially linked to removal of accumulated mRNA before switching to glucose metabolism (Lin et al., 2011; Ammar et al., 2018; Iosub et al., 2021). This is also supported by the fact that pre-culturing in OBSM+1.5% glucose prior to transfer weakened the growth delay phenomenon of strain S12 caused by glucose. Further research is needed to explore this phenomenon.

Strain S12 is characterized by high proportions of PL7 and PL6 (Figure 2B), which are polysaccharide lyases specialized towards alginate. The genes encoding these alginate lyases are present in one PUL, named ALG PUL (Figure 4B), while seven additional single alginate lyase genes are dispersed throughout the genome. ALG PUL contains five adjacent lyase genes, PL6-6-6-17-7, which co-localize with a *susCD* pair. *T. laminarinivorans*, *T. sargassicola*, and *T. sedimentorum* all contain two similar PULs related to alginate utilization. These PULs contain a similar set of CAZymes as ALG PUL, consisting of PL6, 7, and 17 (Li et al., 2023). However, their gene cluster structure and arrangement differ significantly from ALG PUL, which is consistent with the lineage and phylogenetic tree results (Figure 1).

The ALG PUL encodes a metabolic cascade that involves the complete breakdown of external polymers (PL6 and 7) and hydrolysis of oligosaccharides (PL17) to monomers. The interactions between PLs with catalytic diversity may facilitate the cleavage of alginate. The signal peptide in the 5 PLs (Supplementary Table S6) indicates that one PL6 (JJL45_RS04310) is active intracellularly, while the others are secreted extracellularly. This

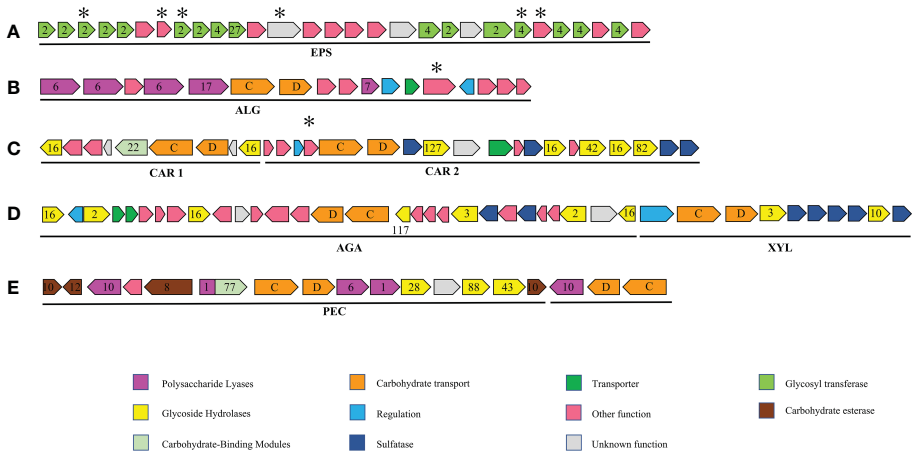


FIGURE 4
Major polysaccharide utilisation loci (PUL) and gene clusters mentioned in this study. numbers designate CAZyme families, letters *susCD* genes. and asterisks, unique genes of strain S12. **(A)** Exopolysaccharide (EPS) gene cluster: putative exopolysaccharide biosynthesis gene cluster only in *Tamlana* sp. S12, *T.* carrageenivorans, *T. agarivorans*; **(B)** ALG PUL: putative PUL with alginate hydrolysis potential, similar gene clusters also exist in *T. haliotis*, *T. agarivorans*; **(C)** CAR PUL1 and CAR PUL2: putative PUL with carrageenan and porphyrin hydrolysis potential, no other similar gene clusters have been found in other genomes of the four genera; **(D)** XLY PUL and AGA PUL: putative PULs with hydrolysis potential of xylan and agar, a similar gene cluster also exists in *T. agarivorans*; **(E)** PEC PUL: putative PUL with pectin hydrolysis potential, no other similar gene clusters have been found in other genomes of the four genera.

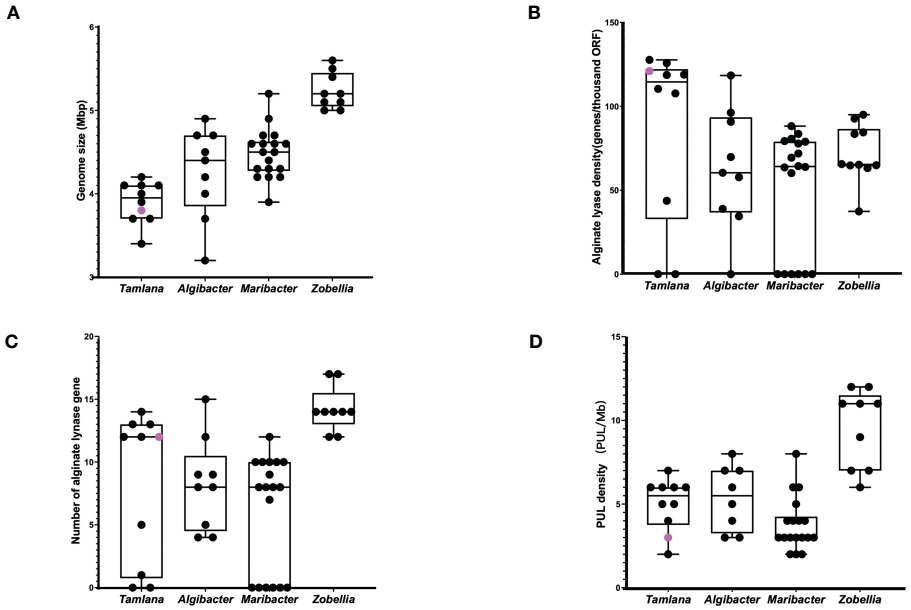


FIGURE 5
(A) Genome size, **(B)** alginate lyase density, **(C)** alginate lynase numbers and **(D)** PUL density of *Tamlana* sp. S12 (pink point) and related strains (black point).

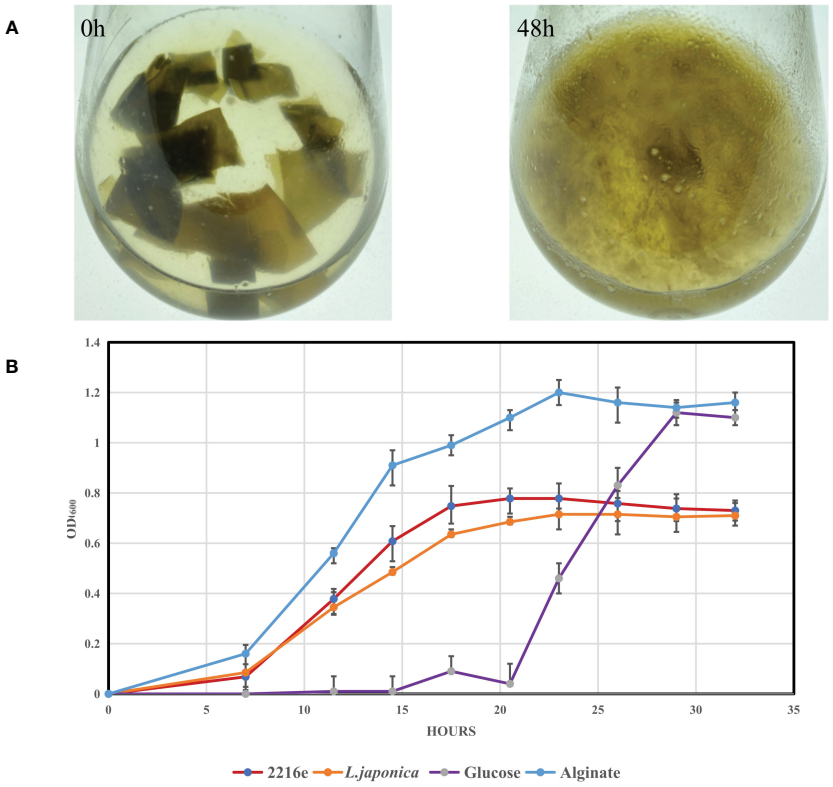


FIGURE 6
Degradation of *L. japonica* by *Tamlana* sp. S12 and growth analysis of *Tamlana* sp. S12 on different mediums. **(A)** Degradation of *L. japonica* by strain S12 in minimal medium with *L. japonica* pieces at 0 h and 48 h post-inoculation, **(B)** Shows the growth curve of strain S12 in four medium.

complementary localization may contribute to the efficient degradation of alginate.

According to genomic analysis, it is hypothesized that strain S12 depolymerizes alginate in a stepwise manner using extracellular lyases PL6 (JJL45_RS04315), PL6 (JJL45_RS04325), and PL7 (JJL45_RS04355). Surface PKD (protein kinase domain) and *susD* proteins facilitate the recruitment of oligosaccharides from the environment and transfer them to the periplasm via *susC*. There, PL17 alginate lyases (JJL45_RS04330) further degrade the oligosaccharides to unsaturated monoaldehydes. The major facility superfamily permease (JJL45_RS04365) (Grondin et al., 2017) transfers alginate oligomers and unsaturated monouronates to the cytoplasm. S12 does not encode a kdgF gene, indicating that unsaturated monouronates are converted into 4-deoxy-l-erythro-5-hexoseulose uronate (DEH) via an alternative (possibly non-enzymatic) mechanism. (Hobbs et al., 2016; Nakata et al., 2019). Subsequently, short-chain dehydrogenase (dehR) and 2-keto-3-deoxy-d-gluconate kinase (KdgK) catalyze their conversion in the cytoplasm. Finally, 2-dehydro-3-deoxyphosphogluconate (KDPG) is assimilated into the central metabolic pathway through the Entner–Doudoroff pathway. All relevant enzymes and pathways have been annotated in strain S12 (Supplementary Table S5).

According to the sequence identity of catalytic domains, the PL7 family is subdivided into seven subfamilies (SF1–SF7) (Lombard et al., 2010; Thomas et al., 2013; Pei et al., 2019; Chernysheva et al., 2021). Similarly, PL17 and PL6 are divided into two (SF1–SF2) and three (SF1–SF3) subfamilies, respectively (Mathieu et al., 2018; Violot et al., 2021). To classify the predicted alginate lyases from the four genera within the ALG PULs, three phylogenetic trees were constructed using characterized PL6, 7, and 17 alginate lyases available in the CAZy database. The representative alginate lyases for each subfamily, along with their target sequences (Figure S4), were included in the analysis.

The phylogenetic tree revealed that seven out of eight *Tamlana* representatives encoded PL7-SF6 lyases, which clustered together as presumptive orthologues. *T. nanhaiensis* PK35_RS15195, *T. laminarinivorans* LG649_RS00180, *T. sargassicola* LG651_RS06140, and *T. sedimentorum* PW52_RS01745 were classified as PL7-SF5 lyases (Figure S4A). Only PL7 from ALG PUL was classified as SF6. To date, only two enzymes in PL7-SF6 have been studied and characterized as endolytic enzymes with a preference for polymannuronate (Inoue et al., 2014; Pei et al., 2019; Chernysheva et al., 2021).

Three PL6 alginate lyases from strain S12 belonged to SF1 and SF2 (Figure S4B), which clustered together with those of *T. agarivorans*, *T. carrageenivorans*, and *T. haliotis* as presumptive orthologues. Twenty-two PL6 alginate lyases from *Zobellia* and *Maribacter* fell strictly into SF1. Mathieu et al. (2016) reported that PL6-SF2 has strict endo-poly-MG lyase activity, whereas PL6-SF1 has either endo-poly-MG/exo-poly-G lyase, endo-poly-MG/endo-poly-G lyase, or strict exo-poly-G lyase activity. The combination of different PL6 subfamilies in the ALG PUL may improve degradation efficiency and substrate adaptability.

PL17 (JJL45_RS04330) alginate lyases from strain S12 showed high amino acid identity (80.8–96.7%) with those of *T. agarivorans*, *T. carrageenivorans*, and *T. haliotis* (A6I66_RS06905, C1A40_RS05220, and F6U93_RS07740) (Figure S4C) and were classified as SF2. However, due to the lack of nearly all the required catalytic and functionally critical residues, PL17-SF2 showed activity towards oligosaccharides containing an unsaturated sugar at the nonreducing end (Jouanneau et al., 2021). Hence, PL17 on ALG PUL was considered to have exolytic oligoalginate lyase activity. PL17 alginate lyases from *T. haliotis* F6U93_RS07745 and *A. marinivivus* GQR97_RS04345 formed a separate branch in addition to PL17-SF1 which may represent a new subfamily. This branch was consistent through the three phylogenetic algorithms (ML, NJ, MP).

Alginate lyases from the PL7 and PL17 families have mainly been isolated from marine organisms (Hehemann et al., 2014; Xu et al., 2020). PL7 from SF6 appears to be conserved only in marine representatives of the Flavobacteriaceae (Chernysheva et al., 2021), and the results also supports this finding.

3.5 Other PULs related to polysaccharide utilization

The genes encoding putative carrageenases and porphyranase were found in two adjacent large gene clusters (one ~19,000 and ~28,000 bp), which are dedicated to carrageenan and porphyrin degradation (Figure 4C). We have named these two gene clusters CAR PUL1 and CAR PUL2, respectively, which encode carrageenolytic and porphyranolytic enzymes (Figure 4C; Supplementary Table S5). The identity level between the putative GHs encoded by CAR PULs and sequences in the Protein Data Bank or Swiss-Prot databases ranged from 24% – 97% (amino acid identity).

Their systems for degrading polysaccharides were represented by GH16.Two putative GH16 (JJL45_RS04765, 04640) showed 62.5% and 45.7% amino acid identity to GH16 κ-carrageenases (ZOBGAL_RS23455) from *Z. galactanivorans*. Additionally, one putative GH82 (JJL45_RS04770) displayed 44.6% identity towards ι-carrageenases from *Z. laminariae* (D9V96_RS20855). GH16 sequences (JJL45_RS04680, 04745) were queried against the sequences with known functions in the Swiss-Prot Database, revealing 25.2% and 24.7% identity, respectively, to β-porphyranases from *Z. galactanivorans* (ZOBGAL_RS04805 and 12330). Blastp analysis also detected these two putative β-porphyranases in *Seonamhaeicola marinus* (NCBI Sequence Number: WP_148539743.1, WP_148539726.1) (77.9% and 74.1% identity, respectively). However, they were not detected in the genomes of the other members of the four genera, suggesting a potential genetic material transfer event between *Seonamhaeicola* and *Tamlana*.

BLASTp analysis revealed that JJL45_RS04760 was conserved in *T. carrageenivorans*, *T. agarivorans*, and *T. fucoidanivorans* (80.4%–97.0% amino acid identity). Based on the PDB database, the protein sequence of JJL45_RS04760 was found to be similar to those of β-galactosidases characterized as GH42, with a high amino acid identity (77.3%) compared to the *S. marinus* sequence (NCBI Sequence Number: WP_187388126.1). A BLASTp search also identified sequences similar to GH42 in several marine bacteria, including *Paraglaciecola hydrolytica* S66^T (NCBI Sequence Number: WP_068375692.1) (60.9% amino acid identity), which

encodes β-galactosidasecan (Schultz-Johansen et al., 2018), suggesting the presence of oligosaccharide enzyme activity of GH42.

Z. galactanivorans was found to harbor three genes encoding GH127 with α-1,3-(3,6)-anhydro-D-galactosidase activity, which play a role in the degradation of carrageenan (Ficko-Blean et al., 2017). JJL45_RS04720 (Supplementary Table S5) showed homology with the three GH127 from *Z. galactanivorans* (ZOBGAL_RS14860, ZOBGAL_RS14875, and ZOBGAL_RS14865), with BLASTp identities ranging from 44.9% to 74.9%. The presence of carrageenan significantly induced the expression of GH127 in *Z. galactanivorans* (Thomas et al., 2017), suggesting that GH127 in CAR PUL2 has 1,3-(3,6)-anhydro-D-galactosidase activity.

CAR PUL1 was found to contain a carbohydrate-binding module family 22 (CBM22) with xylan binding function (Supplementary Table S5) (Liu and Ding, 2016), which is only found in 10 genomes, four of which belong to *Zobellia*. However, CBM22 was not found in *Tamlana* except in strain S12. In red seaweeds, cellulose fibers interact with unusual hemicelluloses embedded in a matrix of sulfated galactans (agars or carrageenans) (Popper et al., 2011). Thomas et al. (2017) showed that one gene coding CBM22 in *Z. galactanivorans* was induced in the presence of agar or porphyrin. Given the presence of two possible porphyranases (JJL45_RS04680, 04745) in CAR PULs, the CBM22 in CAR PUL1 may be induced to act together with porphyranases in the presence of porphyran. However, the role of CBM22 in the utilization of red algal polysaccharides requirs to further study.

Due to the absence of GH127 and/or GH129 for the hydrolysis of anhydrogalactose (AHG) (Supplementary Table S7), it has been suggested that *Maribacter* absorbs hydrolysis products from primary degraders to utilize AHG in red macroalgae (Wolter et al., 2021). CAR PUL 1 and 2 contain one GH127 and a variety of carrageenases and porphyranase, indicating the potential of strain S12 to degrade complex polysaccharides found in red algae. With the assistance of sulfatase, the GHs in CAR PUL1 and 2 are each specific for a particular part of carrageenans and porphyrin, releasing oligosaccharides that can be further desulphated through the action of sulfatases. Further degradation of these oligosaccharides can be performed by GH42 (JJL45_RS04760), resulting in neocarrabiose. Finally, the GH127 α-1,3-(3,6-anhydro)-D-galactosidase homologue (JJL45_RS04720) can release D-AHG from the nonreducing end of neocarrabiose saccharides, leading to depolymerization of carrageenans (Ficko-Blean et al., 2017).

One cluster containing two PULs was found in strain S12 and named AGA PUL and XYL PUL according to the functional annotation. The significant number of GH glycoside hydrolases associated with the utilization of agar polysaccharides in AGA PUL indicates that strain S12 has the potential to degrade these polysaccharides (Figure 4D, Supplementary Table S5). The function of XYL PUL is determined by GH10 encoding for xylanase and GH3 encoding glucosidase. We tested the substrate activity and compared it with *T. agarivorans* KCTC 22176 (Table 2). The results were consistent with the functional prediction. Both strain S12 and *T. agarivorans* can use agar and xylan (Table 2). Additionally, a similar gene cluster was also found in *T. agarivorans*.

AGA PUL contains three GH16 β-agarases (JJL45_RS12010, JJL45_RS12050, JJL45_RS12150) and a GH117 (JJL45_RS12090) that encodes 3,6 anhydro-α-L-galactosidases. GH117 is capable of breaking down neoagarooligosaccharides and generating L-AHG (Jin et al., 2021), playing a role in the degradation of agarose. Additionally, two potential GH2 genes encoding β-galactosidases were found in AGA PUL. These potential GH2s may encode β-galactosidases that remove the non-reducing end of galactose from agarooligosaccharides (Lee et al., 2014). A gene cluster containing GH16, GH117, and GH2 was identified as a putative agarolytic cluster in the human gut bacterium (Pluvinage et al., 2018). The results indicate that the GH16-degraded products of agarose are cyclically degraded into monosaccharides through the collaborative action of GH117 and GH2.

The PL1-PL10 pair exists as accessory genes in a few strains of the four genera (Figure 1). This pair was found in strain S12 and

TABLE 2 Comparison between the novel strain, *Tamlana* sp. S12 and closely related taxon *Tamlana agarivorans* KCTC 22176.

Substrate	Enzyme	S12	*T. agarivorans*
Starch	Amylase (G)	+	-
Agar	Agarase (E)	+	+
xylan	Xylanase (E)	+	+
2-naphthyl butyrate	Esterase (C4) (G)	-	+
6-Br-2-naphthyl-αD-galactopyranoside	α- Galactosidase (G)	-	-
1-naphthyl-N-acetyl-βD-glucosaminide	N-acetyl-β-glucosamindase (G)	-	-
2-naphthyl myristate	Lipase (C14) (G)	-	+
2-naphthyl-αL-fucopyranoside	α- Fucosidase (G)	-	-
N-glutaryl-phenylalanine-2-naphthylamide	α-Chymotrypsin (G)	-	+
L-cystyl-2-naphthylamide	Cystine arylamidase (G)	-	+
6-Br-2-naphthyl-αD-mannopyranoside	α- Mannosidase (G)	-	-

Presence (+) or absence (-) of polymer-degrading abilities; determined via enzymatic (E) or growth (G) assays.

located in two clusters named PEC PUL (Figure 4E), which presumably contributes to pectin utilization, supported by the existence of genes of the families GH28 and 105 that participate in pectinolytic activities (Hobbs et al., 2019). GH28 can encode α-D-galacturonosidase, while GH105 encodes unsaturated uronyl hydrolases, and the catalytic mechanism of GH105 is believed to parallel that of GH88 (Jongkees and Withers, 2011). Pectin is mainly composed of α-(1, 4)-galacturonate and is abundant in terrestrial plants. However, pectin-like polysaccharides also exist in marine microalgae, red and green algae, and seagrasses (Popper et al., 2011; Hehemann et al., 2017). Galacturonate levels can reach 0.3 μM during phytoplankton blooms (Sperling et al., 2017).

In the presence of carbohydrate esterases (CEs) (Supplementary Table S5), specifically CE8 (JJL45_RS15625), 10 (JJL45_RS15605, JJL45_RS15675), and 12 (JJL45_RS15610), the de-esterification of pectin occurs (Hobbs et al., 2019). Additionally, pectinases encoded by PL1 and PL10 may cleave the pectin backbone. This results in the formation of oligouronates with 4,5-unsaturated nonreducing ends. Thees oligouronates can then be further degraded into disaccharides and trisaccharides by exo-poly-α-D-galacturonosidase GH28 (JJL45_RS15655), while GH88 (JJL45_RS15665) results in the direct release of monosaccharides. KDPG is a key intermediate in both pectin and alginate degradation (Koch et al., 2019). The enzyme essential for both pathways is encoded in ALG PUL and may potentially be shared among them (short-chain dehydrogenase, JJL45_RS04370; KdgK, JJL45_RS04385; KDPG aldolase, JJL45_RS04395) (Supplementary Table S5). The PL1-PL10 pair in PEC PUL may also give strain S12 wider substrate adaptability. However, further discussion is needed to fully understand their ecological significance.

3.6 Potential role of S12 in the gut of marine invertebrates

According to Mfilinge and Tsuchiya (2016), sea cucumbers feed on algal fragments (including *L. japonica*) and bacteria. The metabolic activity of Bacteroidetes may enhance the production of carbohydrates, and facilitate the absorption of polysaccharides in the intestine of *A. japonicus*. We sought to investigate the characteristics of strain S12, which enable it to survive in such a marine environment and might contribute to the host's metabolism.

The formation of gut microbiota in marine invertebrates is influenced by bacterial substrate preference and their ability to adhere to the host (Harris, 1993). Utilization of alginate and *L. japonica* biomass as sole nutrient source (Figure 6, Supplementary Figure S5) suggests that strain S12 may contribute to the release of polysaccharides derived from *L. japonica*, facilitating their digestion and absorption within the intestine of *A. japonicus* (Zhang et al., 2019).

Additionally, biofilm assays demonstrated the ability of strain S12 to form biofilms (Figure S6), with poly(β-D-mannuronate) O-acetylase (JJL45_RS08230; Supplementary Table S5) potentially enhancing biofilm formation (Nivens et al., 2001). Biofilm-forming abilities and potential EPS production may facilitates establishment of S12 in the *A. japonicus* gut (Stapper et al., 2004). The formation of a microenvironment around the strain by EPS could potentially reduce the diffusion of secreted CAZymes, which may be advantageous for enhancing degradation efficiency, preserving hydrolysis products (Wolter et al., 2021), and facilitating their absorption within the gut of marine invertebrates. This may provide an added advantage for colonization, reducing competition while preventing excretion along with fecal matter (Harris, 1993; Boyd and Chakrabarty, 1995).

Furthermore, we searched for CAZyme genes that might benefit the invertebrate host, including a GH25 (JJL45_RS06375) encoding a lysozyme. Previous studies have suggested that over 70% of the energy required by herbivorous or omnivorous marine invertebrates is related to bacteria (Zhou et al., 2009). Lysozyme activity has been detected in the gut bacteria of invertebrates. This enzyme has the ability to lyse microbes, and the resulting breakdown products may contribute to extra digestive capacity by providing active enzymes to the host's gut (Harris, 1993). Furthermore, cellulase has been widely detected in the gut microbiota of marine invertebrates (Harris, 1993; Kim et al., 2011), indicating that it may play a role in food preprocessing and improving the host's ability to utilize cellulose (Harris, 1993). Given the promoting effect of cellulase on *L. japonica* degradation (Sun et al., 2020), XYL PUL may facilitate the preprocessing of algal polysaccharides and provide additional enzyme activity to the host's intestine upon *L. japonica* ingestion.

4 Conclusions

The diversity of CAZymes and PUL targeting alginate, carrageenan, and other algal carbohydrates in strain S12 highlights its strong predisposition for polysaccharide degradation. This is evident from its distinct growth in *L. japonica* and alginate as the sole nutrient source, indicating considerable hydrolytic activity in marine systems. In comparison to *Zobellia*, *Algibacter*, and *Maribacter*, *Tamlana* is primarily isolated from marine invertebrates or sediments (Supplementary Table S3). It remains to be determined whether the predisposition of strain S12 towards algal polysaccharides corresponds to the co-occurrence with macroalgae. Growth with *L. japonica* biomass indicated that some members of *Tamlana*, *Zobellia*, *Maribacter*, and *Algibacter* are capable of initiating the release of polysaccharides from algal tissue. Comparative genomic analysis and evidence related to HGT have demonstrated the widespread occurrence of HGT events involving CAZymes among bacteria. The CAZyme diversity in strain S12 adds to the understanding of CAZymes in flavobacteria and the hydrolytic activity of *Tamlana*, with implications for the cycling of algal polysaccharides in the ocean and associated ecological dynamics.

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/bioproject/PRJNA224116.

Author contributions

H-FX conducted genomic analyses and contributed to writing. X-YJ contributed to pan-genome analyses. Y-XZ performed isolation of strain S12 and growth experiments. Z-JD carried out the work of genome data collection and collation. D-SM performed genome sequencing. G-JC designed research and wrote the manuscript. All authors contributed to the final version of the manuscript.

Funding

This work was supported by the National Natural Science Foundation of China (41876166).

Acknowledgments

We would like to thank Editage (www.editage.cn) for English language editing.

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/fmars.2023.985514/full#supplementary-material

References

Almagro Armenteros, J. J., Tsirigos, K. D., Sønderby, C. K., Petersen, T. N., Winther, O., Brunak, S., et al. (2019). SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat. Biotechnol.* 37 (4), 420–423. doi: 10.1038/s41587-019-0036-z

Ammar, E. M., Wang, X., and Rao, C. V. (2018). Regulation of metabolism in *Escherichia coli* during growth on mixtures of the non-glucose sugars: arabinose, lactose, and xylose. *Sci. Rep.* 8, 609. doi: 10.1038/s41598-017-18704-0

Aramaki, T., Blanc-Mathieu, R., Endo, H., Ohkubo, K., Kanehisa, M., Goto, S., et al. (2020). KofamKOALA: KEGG ortholog assignment based on profile HMM and adaptive score threshold. *Bioinformatics* 36 (7), 2251–2252. doi: 10.1093/bioinformatics/btz859

Archibald, A. L. (2017). *Exploiting long read sequencing technologies to establish high quality highly contiguous pig reference genome assemblies* (W137: International Plant and Animal Genome.

Arnold, B. J., Huang, I. T., and Hanage, W. P. (2022). Horizontal gene transfer and adaptive evolution in bacteria. *Nat. Rev. Microbiol.* 20 (4), 206–218. doi: 10.1038/s41579-021-00650-4

Arnosti, C., Wietz, M., Brinkhoff, T., Hehemann, J. H., and Amann, R. (2021). The biogeochemistry of marine polysaccharides: sources, inventories, and bacterial drivers of the carbohydrate cycle. *Ann. Rev. Mar. Sci.* 13, 81–108. doi: 10.1146/annurev-marine-032020-012810

Barbato, M., Vacchini, V., Engelen, A. H., Patania, G., Mapelli, F., Borin, S., et al. (2022). What lies on macroalgal surface: diversity of polysaccharide degraders in culturable epiphytic bacteria. *AMB Express.* 12 (1), 98. doi: 10.1186/s13568-022-01440-8

Ben Bdira,, Artola, M., Overkleeft, H. S., Ubbink, M., and Aerts, J. (2018). Distinguishing the differences in β-glycosylceramidase folds, dynamics, and actions informs therapeutic uses. *J. Lipid Res.* 59 (12), 2262–2276. doi: 10.1194/jlr.R086629

Boyd, A., and Chakrabarty, A. M. (1995). Pseudomonas aeruginosa biofilms: role of the alginate exopolysaccharide. *J. Ind. Microbiol.* 15 (3), 162–168. doi: 10.1007/BF01569821

Brunet, M., Le Duff, N., Barbeyron, T., and Thomas, F. (2022). Consuming fresh macroalgae induces specific catabolic pathways, stress reactions and type IX secretion in marine flavobacterial pioneer degraders. *ISME J.* 16 (8), 2027–2039. doi: 10.1038/s41396-022-01251-6

Burley, S. K., Berman, H. M., Christie, C., Duarte, J. M., Feng, Z., Westbrook, J., et al. (2018). RCSB protein data bank: sustaining a living digital data resource that enables breakthroughs in scientific research and biomedical education. *Protein Sci.* 27 (1), 316–330. doi: 10.1002/pro.3331

Cao, W. R., Liu, B. T., Sun, X. K., Sun, Y. Y., Jiang, M. Y., and Du, Z. J. (2021). *Tamlana haliotis* sp. nov., isolated from the gut of the abalone *Haliotis rubra*. *Arch. Microbiol.* 203 (5), 2357–2364. doi: 10.1007/s00203-021-02216-7

Cha, Q. Q., Wang, X. J., Ren, X. B., Li, D., Wang, P., Li, P. Y., et al. (2021). Comparison of alginate utilization pathways in culturable bacteria isolated from Arctic and Antarctic marine environments. *Front. Microbiol.* 12. doi: 10.3389/fmicb.2021.609393

Chatukuta, P., and Rey, M. E. C. (2020). A cassava protoplast system for screening genes associated with the response to *South African cassava mosaic virus*. *Virol. J.* 17 (1), 184. doi: 10.1186/s12985-020-01453-4

Chaudhari, N. M., Gupta, V. K., and Dutta, C. (2016). BPGA- an ultra-fast pan-genome analysis pipeline. *Sci. Rep.* 6, 24373. doi: 10.1038/srep24373

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 (6), 1925–1927. doi: 10.1093/bioinformatics/btz848

Chen, T., Liu, Y.-X., and Huang, L. (2022). *ImageGP: An easy-to-use data visualization web server for scientific researchers. iMeta 1, e5*. doi: 10.1002/imt2.5

Chernysheva, N., Bystritskaya, E., Likhatskaya, G., Nedashkovskaya, O., and Isaeva, M. (2021). Genome-wide analysis of PL7 alginate lyases in the genus *Zobellia*. *Molecules* 26 (8), 2387. doi: 10.3390/molecules26082387

Chernysheva, N., Bystritskaya, E., Stenkova, A., Golovkin, I., Nedashkovskaya, O., and Isaeva, M. (2019). Comparative genomics and CAZyme genome repertoires of marine *Zobellia amurskyensis* KMM 3526^T and *Zobellia laminariae* KMM 3676^T. *Mar. Drugs* 17 (12), 661. doi: 10.3390/md17120661

Decho, A. W., and Gutierrez, T. (2017). Microbial extracellular polymeric substances (EPSs) in ocean systems. *Front. Microbiol.* 8. doi: 10.3389/fmicb.2017.00922

Delcher, A. L., Bratke, K. A., Powers, E. C., and Salzberg, S. L. (2007). Identifying bacterial genes and endosymbiont DNA with glimmer. *Bioinformatics* 23 (6), 673–679. doi: 10.1093/bioinformatics/btm009

Dudek, M., Dieudonné, A., Jouanneau, D., Rochat, T., Michel, G., Sarels, B., et al. (2020). Regulation of alginate catabolism involves a GntR family repressor in the marine flavobacterium *Zobellia galactanivorans* Dsij^T. *Nucleic Acids Res.* 48 (14), 7786–7800. doi: 10.1093/nar/gkaa533

Ferrer-González, F., Widner, B., Holderman, N. R., Glushka, J., and Moran, M. A. (2021). Resource partitioning of phytoplankton metabolites that support bacterial heterotrophy. *ISME J.* 15 (3), 762–773. doi: 10.1038/s41396-020-00811-y

Ficko-Blean, E., Préchoux, A., Thomas, F., Rochat, T., Larocque, R., Zhu, Y., et al. (2017). Carrageenan catabolism is encoded by a complex regulon in marine heterotrophic bacteria. *Nat. Commun.* 8 (1), 1685. doi: 10.1038/s41467-017-01832-6

García-López, M., Meier-Kolthoff, J. P., Tindall, B. J., Gronow, S., Woyke, T., Kyrpides, N. C., et al. (2019). Analysis of 1,000 type-strain genomes improves taxonomic classification of *Bacteroidetes*. *Front. Microbiol.* 10. doi: 10.3389/fmicb.2019.02083

Gavriilidou, A., Gutleben, J., Versluis, D., Forgiarini, F., and Sipkema, D. (2020). Comparative genomic analysis of *Flavobacteriaceae*: insights into carbohydrate metabolism, gliding motility and secondary metabolite biosynthesis. *BMC Genomics* 21 (1), 569. doi: 10.1186/s12864-020-06971-7

Gomes, L. C., Moreira, J. M., Miranda, J. M., Simões, M., Melo, L. F., and Mergulhão, F. J. (2013). Macroscale versus microscale methods for physiological analysis of biofilms formed in 96-well microtiter plates. *J. Microbiol. Methods* 95 (3), 342–349. doi: 10.1016/j.mimet.2013.10.002

Groisillier, A., Labourel, A., Michel, G., and Tonon, T. (2015). The mannitol utilization system of the marine bacterium zobellia galactanivorans. *Appl. Environ. Microbiol.* 81 (5), 1799–1812. doi: 10.1128/AEM.02808-14

Grondin, J. M., Tamura, K., Déjean, G., Abbott, D. W., and Brumer, H. (2017). Polysaccharide utilization loci: fueling microbial communities. *J. Bacteriol.* 199 (15), e00860–e00816. doi: 10.1128/JB.00860-16

Harris, J. M. (1993). The presence, nature, and role of gut microflora in aquatic invertebrates: a synthesis. *Microb. Ecol.* 25 (3), 195–231. doi: 10.1007/BF00171889

Hehemann, J. H., Boraston, A. B., and Czjzek, M. (2014). A sweet new wave: structures and mechanisms of enzymes that digest polysaccharides from marine algae. *Curr. Opin. Struct. Biol.* 28, 77–86. doi: 10.1016/j.sbi.2014.07.009

Hehemann, J. H., Truong, L. V., Unfried, F., Welsch, N., Kabisch, J., Heiden, S. E., et al. (2017). Aquatic adaptation of a laterally acquired pectin degradation pathway in marine gammaproteobacteria. *Environ. Microbiol.* 19 (6), 2320–2333. doi: 10.1111/1462-2920.13726

Hobbs, J. K., Hettle, A. G., Vickers, C., and Boraston, A. B. (2019). Biochemical reconstruction of a metabolic pathway from a marine bacterium reveals its mechanism of pectin depolymerization. *Appl. Environ. Microbiol.* 85 (1), e02114–e02118. doi: 10.1128/AEM.02114-18

Hobbs, J. K., Lee, S. M., Robb, M., Hof, F., Barr, C., Abe, K. T., et al. (2016). KdgF, the missing link in the microbial metabolism of uronate sugars from pectin and alginate. *PNAS* 113 (22), 6188–6193. doi: 10.1073/pnas.1524214113

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. doi: 10.1186/1471-2105-11-119

Inoue, A., Takadono, K., Nishiyama, R., Tajima, K., Kobayashi, T., and Ojima, T. (2014). Characterization of an alginate lyase, FlAlyA, from flavobacterium sp. strain UMI-01 and its expression in *Escherichia coli*. *Mar. Drugs* 12 (8), 4693–4712. doi: 10.3390/md12084693

Iosub, I. A., Marchioretto, M., van Nues, R. W., McKellar, S., Viero, G., and Granneman, S. (2021). The mRNA derived MalH sRNA contributes to alternative carbon source utilization by tuning maltoporin expression in *E. coli*. *RNA Biol.* 18 (6), 914–931. doi: 10.1080/15476286.2020.1827784

Jam, M., Flament, D., Allouch, J., Potin, P., Thion, L., Kloareg, B., et al. (2005). The endo-beta-agarases AgaA and AgaB from the marine bacterium *Zobellia galactanivorans*: two paralogue enzymes with different molecular organizations and catalytic behaviours. *Biochem. J.* 385 (Pt 3), 703–713. doi: 10.1042/BJ20041044

Jin, Y., Yu, S., Kim, D. H., Yun, E. J., and Kim, K. H. (2021). Characterization of neoagarooligosaccharide hydrolase *Bp* GH117 from a human gut bacterium *Bacteroides plebeius*. *Mar. Drugs* 19 (5), 271. doi: 10.3390/md19050271

Jongkees, S. A. K., and Withers, S. G. (2011). Glycoside cleavage by a new mechanism in unsaturated glucuronyl hydrolases. *J. Am. Chem. Soc.* 133 (48), 19334–19337. doi: 10.1021/ja209067v

Jouanneau, D., Klau, L. J., Larocque, R., Jaffrennou, A., Duval, G., Duff, Le, et al. (2021). Structure-function analysis of a new PL17 oligoalginate lyase from the marine bacterium *Zobellia galactanivorans* Dsij^{T}. *Glycobiology* 31 (10), 1364–1377. doi: 10.1093/glycob/cwab058

Jung, J., Bae, S. S., Chung, D., and Baek, K. (2019). *Tamlana carrageenivorans* sp. nov., a carrageenan-degrading bacterium isolated from seawater. *Int. J. Syst. Evol. Microbiol.* 69 (5), 1355–1360. doi: 10.1099/ijsem.0.003318

Kalisch, B., Dörmann, P., and Hölzl, G. (2016). DGDG and glycolipids in plants and algae. *Subcell Biochem.* 86, 51–83. doi: 10.1007/978-3-319-25979-6_3

Kanehisa, M., and Sato, Y. (2020). KEGG mapper for inferring cellular functions from protein sequences. *Protein Sci.* 29 (1), 28–35. doi: 10.1002/pro.3711

Kim, D., Kim, S. N., Baik, K. S., Park, S. C., Lim, C. H., Kim, J. O., et al. (2011). Screening and characterization of a cellulase gene from the gut microflora of abalone using metagenomic library. *J. Microbiol.* 49 (1), 141–145. doi: 10.1007/s12275-011-0205-3

Koch, H., Dürwald, A., Schweder, T., Noriega-Ortega, B., Vidal-Melgosa, S., Hehemann, J. H., et al. (2019). Biphasic cellular adaptations and ecological implications of alteromonas macleodii degrading a mixture of algal polysaccharides. *Isme J.* 13 (1), 92–103. doi: 10.1038/s41396-018-0252-4

Kumar, A., Jaiswal, V., Kumar, V., Dey, A., and Kumar, A. (2018). Functional redundancy in echinocandin b in-cluster transcription factor *ecd*B of *Emericella rugulosa* NRRL 11440. *Biotechnol. Rep. (Amst).* 19, e00264. doi: 10.1016/j.btre.2018.e00264

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 (6), 1547–1549. doi: 10.1093/molbev/msy096

Leblond-Bourget, N., Philippe, H., Mangin, I., and Decaris, B. (1996). 16S rRNA and 16S to 23S internal transcribed spacer sequence analyses reveal inter- and intraspecific bifidobacterium phylogeny. *Int. J. Syst. Bacteriol.* 46 (1), 102–111. doi: 10.1099/00207713-46-1-102

Lee, S. D. (2007). *Tamlana crocina* gen. nov., sp. nov., a marine bacterium of the family *Flavobacteriaceae*, isolated from beach sediment in Korea. *Int. J. Syst. Evol. Microbiol.* 57 (Pt 4), 764–769. doi: 10.1099/ijs.0.64720-0

Lee, C. H., Kim, H. T., Yun, E. J., Lee, A. R., Kim, S. R., Kim, J. H., et al. (2014). A novel agarolytic β-galactosidase acts on agarooligosaccharides for complete hydrolysis of agarose into monomers. *Appl. Environ. Microbiol.* 80 (19), 5965–5973. doi: 10.1128/AEM.01577-14

Letunic, I., and Bork, P. (2021). Interactive tree of life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* 49 (W1), W293–W296. doi: 10.1093/nar/gkab301

Li, J., Liang, Y., He, Z., An, L., Liu, Y., Zhong, M., et al. (2023). *Tamlana laminarinivorans* sp. nov. and tamlana sargassicola sp. nov., two novel species isolated from *Sargassum*, show genomic and physiological adaptations for a *Sargassum*-associated lifestyle. *Int. J. Syst. Evol. Microbiol.* 73 (3). doi: 10.1099/ijsem.0.005706

Li, J., Xu, Y., Feng, J., Zhong, M., Xie, Q., Peng, T., et al. (2020). *Tamlana fucoidanivorans* sp. nov., isolated from algae collected in China. *Int. J. Syst. Evol. Microbiol.* 70 (3), 1496–1502. doi: 10.1099/ijsem.0.003850

Lin, H. H., Hsu, C. C., Yang, C. D., Ju, Y. W., Chen, Y. P., and Tseng, C. P. (2011). Negative effect of glucose on ompA mRNA stability: a potential role of cyclic AMP in the repression of *hfq* in *Escherichia coli*. *J. Bacteriol.* 193 (20), 5833–5840. doi: 10.1128/JB.05359-11

Liu, S., and Ding, S. (2016). Replacement of carbohydrate binding modules improves acetyl xylan esterase activity and its synergistic hydrolysis of different substrates with xylanase. *BMC Biotechnol.* 16 (1), 73. doi: 10.1186/s12896-016-0305-6

Liu, X., Lai, Q., Du, Y., Li, G., Sun, F., and Shao, Z. (2015). *Tamlana nanhaiensis* sp. nov., isolated from surface seawater collected from the south China Sea. *Antonie Van Leeuwenhoek* 107 (5), 1189–1196. doi: 10.1007/s10482-015-0410-x

Liu, M., Siezen, R. J., and Nauta, A. (2009). In silico prediction of horizontal gene transfer events in *Lactobacillus bulgaricus* and *Streptococcus thermophilus* reveals protocooperation in yogurt manufacturing. *Appl. Environ. Microbiol.* 75 (12), 4120–4129. doi: 10.1128/AEM.02898-08

Lombard, V., Bernard, T., Rancurel, C., Brumer, H., Coutinho, P., and Henrissat, B. (2010). A hierarchical classification of polysaccharide lyases for glycogenomics. *Biochem. J.* 432 (3), 437–444. doi: 10.1042/BJ20101185

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

Malik, S., Rashid, M., Javid, A., Hussain, A., Bukhari, S. M., Suleman, S., et al. (2021). Genetic variations and phylogenetic relationship of genus *Uromastyx* from punjab Pakistan. *Braz. J. Biol.* 84, e254253. doi: 10.1590/1519-6984.254253

Malinski, T. J., and Singh, H. (2019). Enzymatic conversion of RBCs by α -N-Acetylgalactosaminidase from spirosoma linguale. *Enzyme Res.* 2019, 6972835. doi: 10.1155/2019/6972835

Martin, M., Barbeyron, T., Martin, R., Portetelle, D., Michel, G., and Vandenbol, M. (2015). The cultivable surface microbiota of the brown alga *Ascophyllum nodosum* is enriched in macroalgal-Polysaccharide-Degrading bacteria. *Front. Microbiol.* 6. doi: 10.3389/fmicb.2015.01487

Mathieu, S., Henrissat, B., Labre, F., Skjåk-Bræk, G., and Helbert, W. (2016). Functional exploration of the polysaccharide lyase family PL6. *PloS One* 11 (7), e0159415. doi: 10.1371/journal.pone.0159415

Mathieu, S., Touvrey-Loiodice, M., Poulet, L., Drouillard, S., Vincentelli, R., Henrissat, B., et al. (2018). Ancient acquisition of "alginate utilization loci" by human gut microbiota. *Sci. Rep.* 8 (1), 8075. doi: 10.1038/s41598-018-26104-1

Mcbride, M. J. (2014). *The family flavobacteriaceae* (Berlin Heidelberg: Springer), 643–676. doi: 10.1007/978-3-642-38954-2_130

Meier-Kolthoff, J. P., Sardà Carbasse, J., Peinado-Olarte, R. L., and Göker, M. (2022). TYGS and LPSN: a database tandem for fast and reliable genome-based classification and nomenclature of prokaryotes. *Nucleic Acid Res.* 50, D801–D807. doi: 10.1093/nar/gkab902

Mfilinge, P. L., and Tsuchiya, M. (2016). Changes in sediment fatty acid composition during passage through the gut of deposit feeding holothurians: *Holothuria atra* (Jaeger 1883) and *Holothuria leucospilota* (Brandt 1835). *J. Lipids.* 2016, 4579794. doi: 10.1155/2016/4579794

Mostafa, Y. S., Alrumman, S. A., Otaif, K. A., Alamri, S. A., Mostafa, M. S., and Sahlabji, T. (2020). Production and characterization of bioplastic by polyhydroxybutyrate accumulating *Erythrobacter aquimaris* isolated from mangrove rhizosphere. *Molecules* 25 (1), 179. doi: 10.3390/molecules25010179

Mu, D. S., Wang, S., Liang, Q. Y., Du, Z. Z., Tian, R., Ouyang, Y., et al. (2020). Bradymonabacteria, a novel bacterial predator group with versatile survival strategies in saline environments. *Microbiome* 8 (1), 126. doi: 10.1186/s40168-020-00902-0

Nakata, S., Murata, K., Hashimoto, W., and Kawai, S. (2019). Uncovering the reactive nature of 4-deoxy-L-*erythro*-5-hexoseulose uronate for the utilization of alginate, a promising marine biopolymer. *Sci. Rep.* 9 (1), 17147. doi: 10.1038/s41598-019-53597-1

Nivens, D. E., Ohman, D. E., Williams, J., and Franklin, M. J. (2001). Role of alginate and its O -acetylation in formation of pseudomonas aeruginosa microcolonies and biofilms. *J. Bacteriol* 183, 1047–1057. doi: 10.1128/JB.183.3.1047-1057.2001

Pei, X., Chang, Y., and Shen, J. (2019). Cloning, expression and characterization of an endo-acting bifunctional alginate lyase of marine bacterium wenyingzhuangia fucanilytica. *Protein Expr Purif.* 154, 44–51. doi: 10.1016/j.pep.2018.09.010

Pluvinage, B., Grondin, J. M., Amundsen, C., Klassen, L., Moote, P. E., Xiao, Y., et al. (2018). Molecular basis of an agarose metabolic pathway acquired by a human intestinal symbiont. *Nat. Commun.* 9 (1), 1043. doi: 10.1038/s41467-018-03366-x

Popper, Z. A., Michel, G., Hervé, C., Domozych, D. S., Willats, W. G., Tuohy, M. G., et al. (2011). Evolution and diversity of plant cell walls: from algae to flowering plants. *Annu. Rev. Plant Biol.* 62, 567–590. doi: 10.1146/annurev-arplant-042110-103809

Ravenhall, M., Škunca, N., Lassalle, F., and Dessimoz, C. (2015). Inferring horizontal gene transfer. *PloS Comput. Biol.* 11 (5), e1004095. doi: 10.1371/journal.pcbi.1004095

Romanenko, L. A., Tanaka, N., Kurilenko, V. V., and Svetashev, V. I. (2014). *Tamlana sedimentorum* sp. nov., isolated from shallow sand sediments of the Sea of Japan. *Int. J. Syst. Evol. Microbiol.* 64 (Pt 8), 2891–2896. doi: 10.1099/ijs.0.061812-0

Schultz-Johansen, M., Bech, P. K., Hennessy, R. C., Glaring, M. A., Barbeyron, T., Czjzek, M., et al. (2018). A novel enzyme portfolio for red algal polysaccharide degradation in the marine bacterium *Paraglaciecola hydrolytica* S66(T) encoded in a sizeable polysaccharide utilization locus. *Front. Microbiol.* 9. doi: 10.3389/fmicb.2018.00839

Sperling, M., Piontek, J., Engel, A., Wiltshire, K. H., Niggemann, J., Gerdts, G., et al. (2017). Combined carbohydrates support rich communities of particle-associated marine bacterioplankton. *Front. Microbiol.* 8. doi: 10.3389/fmicb.2017.00065

Stapper, A. P., Narasimhan, G., Ohman, D. E., Barakat, J., Hentzer, M., Molin, S., et al. (2004). Alginate production affects *Pseudomonas aeruginosa* biofilm development and architecture, but is not essential for biofilm formation. *J. Med. Microbiol.* 53 (Pt 7), 679–690. doi: 10.1099/jmm.0.45539-0

Sullivan, M. J., Petty, N. K., and Beatson, S. A. (2011). Easyfig: a genome comparison visualizer. *Bioinformatics* 27 (7), 1009–1010. doi: 10.1093/bioinformatics/btr039

Sun, F. X., Ma, Y. X., Wang, Y., and Liu, Q. (2010). Purification and characterization of novel κ-carrageenase from marine tamlana sp. HC4. *Chin. J. Oceanology&Limnology* 28 (6), 7. doi: 10.1007/s00343-010-9012-7

Sun, C., Zhou, J., Duan, G., and Yu, X. (2020). Hydrolyzing *Laminaria japonica* with a combination of microbial alginate lyase and cellulase. *Bioresour Technol.* 311, 123548. doi: 10.1016/j.biortech.2020.123548

Tanaka,, Cleenwerck, I., Mizutani, Y., Iehata, S., Shibata, T., Miyake, H., et al. (2015). Formosa Haliotis sp. nov., a brown-alga-degrading bacterium isolated from the gut of the abalone haliotis gigantea. *Int. J. Syst. Evol. Microbiol.* 65 (12), 4388–4393. doi: 10.1099/ijsem.0.000586

Terrapon, N., Lombard, V., Drula, É., Lapébie, P., Al-Masaudi, S., Gilbert, H. J., et al. (2018). PULDB: the expanded database of polysaccharide utilization loci. *Nucleic Acids Res.* 46 (D1), D677–D683. doi: 10.1093/nar/gkx1022

Thomas, F., Bordron, P., Eveillard, D., and Michel, G. (2017). Gene expression analysis of *Zobellia galactanivorans* during the degradation of algal polysaccharides reveals both substrate-specific and shared transcriptome-wide responses. *Front. Microbiol.* 8. doi: 10.3389/fmicb.2017.01808

Thomas, F., Lundqvist, L., Jam, M., Jeudy, A., and Barbeyron, T. (2013). Comparative characterization of two marine alginate lyases from zobellia galactanivorans reveals distinct modes of action and exquisite adaptation to their natural substrate. *J. Biol. Chem.* 288 (32), 23021–23037. doi: 10.1074/jbc.M113.467217

Tian, R., Ning, D., He, Z., Zhang, P., Spencer, S. J., Gao, S., et al. (2020). Small and mighty: adaptation of superphylum patescibacteria to groundwater environment drives their genome simplicity. *Microbiome* 8 (1), 51. doi: 10.1186/s40168-020-00825-w

Tian, X., Zhang, Z., Yang, T., Chen, M., Li, J., Chen, F., et al. (2016). Comparative genomics analysis of streptomyces species reveals their adaptation to the marine environment and their diversity at the genomic level. *Front. Microbiol.* 7. doi: 10.3389/fmicb.2016.00998

UniProt Consortium (2021). UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res.* 49 (D1), D480–D489. doi: 10.1093/nar/gkaa1100

Violot, S., Galisson, F., Carrique, L., Jugnarain, V., Conchou, L., Robert, X., et al. (2021). Exploring molecular determinants of polysaccharide lyase family 6-1 enzyme activity. *Glycobiology* 31 (11), 1557–1570. doi: 10.1093/glycob/cwab073

Wang, L., Zhao, X., Xu, H., Bao, X., Liu, X., Chang, Y., et al. (2018). Characterization of the bacterial community in different parts of the gut of sea cucumber (*Apostichopus japonicus*) and its variation during gut regeneration. *Aquac Res.* 491987–1996. doi: 10.1111/are.13654

Wolter, L. A., Mitulla, M., Kalem, J., Daniel, R., and Wietz, M. (2021). CAZymes in *Maribacter dokdonensis* 62-1 from the Patagonian shelf: genomics and physiology compared to related flavobacteria and a co-occurring *Alteromonas* strain., (2020). *Front. Microbiol.* 12. doi: 10.3389/fmicb.2021.628055

Xiao, C. L., Chen, Y., Xie, S.-Q., Chen, K.-N., Wang, Y., Han, Y., et al. (2017). MECAT: fast mapping, error correction, and *de novo* assembly for single-molecule sequencing reads. *Nat. Methods* 14, 1072–1074. doi: 10.1038/nmeth.4432

Xu, F., Chen, X. L., Sun, X. H., Dong, F., Li, C. Y., Li, P. Y., et al. (2020). Structural and molecular basis for the substrate positioning mechanism of a new PL7 subfamily alginate lyase from the arctic. *J. Biol. Chem.* 295 (48), 16380–16392. doi: 10.1074/jbc.RA120.015106

Yin, R., Yi, Y. J., Chen, Z., Wang, B. X., Li, X. H., and Zhou, Y. X. (2021). Characterization of a new biofunctional, exolytic alginate lyase from tamlana sp. s12 with high catalytic activity and cold-adapted features. *Mar. Drugs* 19 (4), 191. doi: 10.3390/md19040191

Yoon, S. H., Ha, S. M., Kwon, S., Lim, J., Kim, Y., Seo, H., et al. (2017). Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. *Int. J. Syst. Evol. Microbiol.* 67 (5), 1613–1617. doi: 10.1099/ijsem.0.001755

Yoon, J. H., Kang, S. J., Lee, M. H., and Oh, T. K. (2008). *Tamlana agarivorans* sp. nov., isolated from seawater off jeju island in Korea. *Int. J. Syst. Evol. Microbiol.* 58 (Pt 8), 1892–1895. doi: 10.1099/ijs.0.65704-0

Zhang, H., Wang, Q., Liu, S., Huo, D., Zhao, J., Zhang, L., et al. (2019). Genomic and metagenomic insights into the microbial community in the regenerating intestine of the Sea cucumber *Apostichopus japonicus*. *Front. Microbiol.* 10. doi: 10.3389/fmicb.2019.01165

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 (W1), W95–W101. doi: 10.1093/nar/gky418

Zhu, Y., Chen, P., Bao, Y., Men, Y., Zeng, Y., Yang, J., et al. (2016). Complete genome sequence and transcriptomic analysis of a novel marine strain bacillus weihaiensis reveals the mechanism of brown algae degradation. *Sci. Rep.* 6, 38248. doi: 10.1038/srep38248

Frontiers in Microbiology

Explores the habitable world and the potential of microbial life

The largest and most cited microbiology journal which advances our understanding of the role microbes play in addressing global challenges such as healthcare, food security, and climate change.

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