frontiers | Research Topics

# Methods in aquatic microbiology

**Edited by**
Tony Gutierrez, Frederic Coulon, Terry John McGenity and Kai Ziervogel

**Published in**
Frontiers in Microbiology
Frontiers in Marine Science

ISSN 1664-8714
ISBN 978-2-8325-3472-4
DOI 10.3389/978-2-8325-3472-4

## 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

# Methods in aquatic microbiology

**Topic editors**

Tony Gutierrez — Heriot-Watt University, United Kingdom
Frederic Coulon — Cranfield University, United Kingdom
Terry John McGenity — University of Essex, United Kingdom
Kai Ziervogel — University of New Hampshire, United States

**Citation**

Gutierrez, T., Coulon, F., McGenity, T. J., Ziervogel, K., eds. (2023). *Methods in aquatic microbiology*. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-8325-3472-4

# Table of contents

frontiers | Frontiers in Microbiology

TYPE Editorial
PUBLISHED 04 December 2023
DOI 10.3389/fmicb.2023.1338297



OPEN ACCESS

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

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

RECEIVED 14 November 2023
ACCEPTED 21 November 2023
PUBLISHED 04 December 2023

CITATION
Gutierrez T, Coulon F and Ziervogel K (2023) Editorial: Methods in aquatic microbiology. *Front. Microbiol.* 14:1338297. doi: 10.3389/fmicb.2023.1338297



# Editorial: Methods in aquatic microbiology

Tony Gutierrez[1]*, Frederic Coulon[2] and Kai Ziervogel[3]

[1]School of Engineering and Physical Sciences, Institute of Mechanical Process and Energy Engineering (IMPEE), Heriot-Watt University, Edinburgh, United Kingdom, [2]School of Water, Energy, and Environment, Cranfield University, Cranfield, United Kingdom, [3]Institute for the Study of Earth, Oceans, and Space, University of New Hampshire, Durham, NC, United States

KEYWORDS

**methods, aquatic microbiology, sampling, sequencing, phylogenetics**

Editorial on the Research Topic
Methods in aquatic microbiology

This Research Topic highlights the latest experimental techniques and methods used to investigate fundamental questions in Aquatic Microbiology research, from aquatic microbe sampling and culturing, to sequencing, phylogenetics, and microbial material cycling. The Topic includes new or existing methods/technologies that have been significantly improved or adapted that help advance our understanding of the identification and characterization of microbial species and ecosystems in aquatic environments. We hope that this Research Topic will feed into existing research, stimulate new research and collaboration in the coming years.

It includes a total of 11 articles, with three articles describing new methods to sample and improve prokaryotic or eukaryotic community surveys in aquatic systems. To begin, the article by Ma et al. investigated if bias is introduced into eukaryotic plankton community assessments based on DNA-based community sequencing when using size-fractionation (i.e., different filter pore sizes) when sampling. The authors showed that this can indeed lead to significant differences in community composition, and that this had more influence on unique OTUs than shared OTUs, and on low-abundance OTUs. In the article by Ranchou-Peyruse et al. the authors developed an improved method for sampling of deep aquifers to eliminate contamination. Microbial sampling of deep aquifers is challenging because of the need to minimize, and ideally eliminate, contamination from surface microbial contaminants, as this is key to guaranteeing the best interpretations and understanding of the functioning of the deep biosphere. In this study, the adaptation, preparation, sterilization, and deployment of a commercial downhole sampler (PDSshort, Leutert, Germany) for studying the microbiology of deep aquifers, is presented. The authors showed that this improved sampler allowed the formation water and the autochthonous microbial community to be maintained at *in situ* pressure for laboratory analysis. To improve the automation and time-scale resolution for sampling and data collection to assess prokaryotic community structure during research cruises, the paper by Pernice and Gasol showed the successful coupling of an automated continuous water sampler (OC-300) that samples marine surface waters every 15 minutes and stains the microbial cell community, with a flow-cytometer (Accuri-C6) that then quantifies the community based on size. The method would be an excellent add-on to sequencing surveys as it provides absolute values for community abundances.

To drive the discovery of new biotechnology products from previously unculturable microbes, several methods such as modification of media composition, incubation conditions, single-cell isolation, and *in situ* incubation, have been employed to improve microbial recovery from environmental samples. Here, three papers present new methods to isolate and culture marine bacteria. To improve microbial recovery, the article by Pope et al. examined the effect of microencapsulation followed by *in situ* incubation on the abundance, viability, and diversity of bacteria recovered from marine sediment. The authors showed that encapsulation recovered greater bacterial diversity from the sediment than simple resuspension (41 vs. 31 OTUs, respectively), and overall this study provides another tool that microbiologists can use to access microbial dark matter for environmental, biotechnology bioprospecting. With respect to the latter, to exploit the microbes present in the environment for their beneficial resources, effective selection and isolation of microbes from environmental samples is essential. In the study by Duran et al. the authors used a gel-filled microwell array device, using resin for microbial culture, that was fabricated and validated for isolating and culturing marine bacteria in individual wells. The design essentially comprises a simple, effective and cheap device for the first-step screening of microorganisms from marine environmental samples. In the article by Jung et al. two advanced methods, based on a continuous-flow bioreactor (CF) and *in situ* cultivation (I-tip), were evaluated to isolate previously uncultivated marine sponge-associated bacteria. This approach was found to yield a greater number of novel bacteria compared to conventional direct plate cultivation. To understand the mechanisms operating in each cultivation method, and which could help further improve the methodology, the physiological properties of the isolates from each method were characterized. From this, it was found that some of the bacteria require a "growth initiation factor" that is present in the natural environment, and that it would be beneficial to carry out pre-enrichment cultivation targeting bacteria that are less competitive on conventional cultivation. This work provides a step forward in cultivation methods together with improved understanding of the key factors for cultivating previously uncultivated microbes which can be applied to build new cultivation strategies, and that can be integrated into discovery pipelines for isolating microbes that are sources for valuable secondary metabolites of biotechnological interest.

In some facets of aquatic microbiology, such as in aquaculture, the ability to detect low numbers of microbial cells in water or tissue samples is highly valuable but remains a significant challenge. This Research Topic presents two articles that describe new or improved methods that address this challenge. The article by Luo et al. presents a new method for the detection of *Prymnesium parvum* – a harmful algal bloom (HAB)-forming species that produces a collection of compounds known as prymnesins that can cause harmful effects to fish, shellfish, and molluscs, resulting in significant economic losses. The novel method described in this study utilizes isothermal amplification, known as recombinase polymerase amplification (RPA), in combination with lateral-flow dipstick (LFD). The authors showed the newly designed RPA-LFD method to be highly specific, demonstrating no cross-reaction with distinct control microalgae. It is 100 times more sensitive than the current PCR-based detection test for *P. parvum*, and the test result can be obtained in 20 min. The report by Geraci-Yee et al. describes the development of a detection assay for Quahog Parasite Unknown (QPX) disease in hard clams. This disease has caused mass mortality events in both wild and cultured hard clams since the 1960s, yet progress in our understanding of QPX disease outbreaks has been limited in part by poor understanding of the biology and ecology of the causative agent – an opportunistic thraustochytrid that was first cultivated in the 1990's, but only recently formally described as *Mucochytrium quahogii*. Whilst several methods have been able to detect *M. quahogii* in seawater and sediment, its abundance was typically too low for reliable quantification by those methods. In this study, the authors describe the development of a nqPCR SYBR-based assay for sensitive detection and quantification of *M. quahogii* from marine environmental samples following the Minimum Information for Publication of Quantitative Real-Time PCR Experiments (MIQE) guidelines. They then applied the nqPCR assay to a large set of sediment and seawater samples, confirming that *M. quahogii* is prevalent in the environment, as expected for an opportunistic pathogen, and that the *M. quahogii*-specific nqPCR assay is sensitive enough to quantify *M. quahogii* at its natural environmental abundance. The new assay represents a valuable tool to better understand *M. quahogii* dynamics in the environment and possible relationships with QPX disease outbreaks and provides a model to guide the development of similar assays for other important marine microbes typically present at similarly low abundance.

Traditionally, diatoms have been assessed using microscopy, which in some cases is not reliable or reproducible, and more recently high-throughput sequencing using phylogenetic markers, such as various 18S rDNA regions, are not always reliable in taxonomy studies and diversity surveys because they cannot distinguish between some taxa. In the report by Dermastia et al. the chloroplast-encoded *rbcL* marker was used in a series of monthly samples from the Gulf of Trieste, northern Adriatic Sea, to assess diatom diversity, and it was found to reliably detect more taxa compared to 18S-V9 metabarcoding or microscopy. Whilst the authors showed a very thorough community analysis obtained by *rbcL* and its implications for studies dealing with taxa distribution and population structure, as well as carbon and silica flux models and networks, the authors highlight the incompleteness of reference databases and options for possible improvements in this regard.

The article by Neu and Kuhlicke describes a lectin-based approach to visually capture and analyze the hydrated structure of biofilms and bioaggregates, which has been a significant challenge. The authors present lectin data in form of a barcoding table that can serve as a guide to scientists for the selection of lectins in studies on glycoconjugates of multispecies and environmental biofilm/bioaggregate systems.

Finally, in the study by Mazière et al. a mesocosm approach is proposed on how to expose microbial mats, as a microbial community model, to changes in temperature and pH in order to investigate the impact of climate change on microbial communities representative of coastal areas. Although the method cannot represent the full spectrum of environmental complexity,

it closely mimics *in situ* conditions and could be improved in this respect by, for example, adding rainfall cycles or seasonal variations.

## Author contributions

TG: Conceptualization, Funding acquisition, Validation, Writing – original draft, Writing – review & editing. FC: Conceptualization, Validation, Writing – review & editing. KZ: Conceptualization, Validation, Writing – review & editing.

## Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. Preparation of this manuscript was made possible by a NERC Global Challenges Research Fund (NE/V006088/1) and a Leverhulme International Academic Fellowship (IAF-2019-015) awarded 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

TYPE Methods
PUBLISHED 08 August 2022
DOI 10.3389/fmicb.2022.940280



OPEN ACCESS

EDITED BY
Kai Ziervogel,
University of New Hampshire,
United States

REVIEWED BY
Giovanni Scillitani,
University of Bari Aldo Moro, Italy
Kara B. De Leon,
University of Oklahoma, United States
Katya Georgieva,
IBER-BAS, Bulgaria

*CORRESPONDENCE
Thomas R. Neu
thomas.neu@ufz.de

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

RECEIVED 10 May 2022
ACCEPTED 29 June 2022
PUBLISHED 08 August 2022

CITATION
Neu TR and Kuhlicke U (2022) Matrix glycoconjugate characterization in multispecies biofilms and bioaggregates from the environment by means of fluorescently-labeled lectins. *Front. Microbiol.* 13:940280. doi: 10.3389/fmicb.2022.940280



# Matrix glycoconjugate characterization in multispecies biofilms and bioaggregates from the environment by means of fluorescently-labeled lectins

Thomas R. Neu* and Ute Kuhlicke

Helmholtz Centre for Environmental Research – UFZ, Magdeburg, Germany

Environmental biofilms represent a complex mixture of different microorganisms. Their identity is usually analyzed by means of nucleic acid-based techniques. However, these biofilms are also composed of a highly complex extracellular matrix produced by the microbes within a particular biofilm system. The biochemical identity of this extracellular matrix remains in many cases an intractable part of biofilms and bioaggregates. Consequently, there is a need for an approach that will give access to the fully hydrated structure of the extracellular matrix or at least a major part of it. A crucial compound of the matrix identified as carbohydrate-based polymers represents major structural and functional constituents. These glycoconjugates can be characterized by using fluorescently-labeled lectins in combination with confocal laser scanning microscopy. The lectin approach is defined previously, as fluorescence lectin barcoding (FLBC) and fluorescence lectin-binding analysis (FLBA), where FLBC is equal to the screening of a particular sample with all the commercially available lectins and FLBA is the actual analysis of the matrix throughout an experiment with a selected panel of lectins. As the application of immune-based techniques in environmental biofilm systems is impossible, the lectin approach is currently the only option for probing lectin-specific glycoconjugates in complex biofilms and bioaggregates. From all the commercially available lectins tested, the lectins such as AAL, HAA, WGA, ConA, IAA, HPA, and LEA showed the highest binding efficiency. Furthermore, 20 of the overall lectins tested showed an intermediate signal intensity, nevertheless very useful for the assessment of matrix glycoconjugates. With the data compiled, we shall virtually shed more light on the dark matter of the extracellular matrix and their 3-dimensional distribution in environmental biofilm systems. The results will be helpful in future studies with a focus on the extracellular matrix glycoconjugates present in environmental microbial communities.

KEYWORDS

**biofilm, bioaggregates, biofilm matrix, extracellular polymeric substances, lectin, glycoconjugate, confocal laser scanning microscopy**

# Introduction

The extracellular matrix of microbial communities, such as biofilms and bioaggregates, represents an essential part of their structure and functionality. This self-produced matrix is a complex mixture of different biochemical constituents, such as polysaccharides, proteins, extracellular nucleic acids, amphiphilic compounds, and microbial-derived refractory compounds (Neu and Lawrence, 2017). Usually, a mixture of these constituents summarized as extracellular polymeric substances (EPSs) built the matrix. However, several recent publications used the term extracellular matrix (ECM). The pivotal significance of the matrix in a wider sense was discussed in several overviews with a focus on the following aspects: general matrix facts and issues (Flemming and Wingender, 2010), giving structure to the matrix (Hobley et al., 2015), sensing and signaling of the matrix (Steinberg and Kolodkin-Gal, 2015), the functionality of matrix constituents (Neu and Lawrence, 2017), and matrix function in a social context (Dragoš and Kovács, 2017).

Nevertheless, analyzing the seemingly intractable matrix remains a major challenge. There are several approaches to analyzing the matrix. First, establishing extraction strategies and subsequent biochemical analysis (D'Abzac et al., 2010; Sun et al., 2012; Zhang et al., 2012; Pellicer-Nacher et al., 2013; Loustau et al., 2018); second, new extraction methods (Felz et al., 2016; Pronk et al., 2017; Boleij et al., 2018; Wong et al., 2020); third, taking advantage of genomics and proteomics techniques to examine the extracellular space (Dumas et al., 2008; Paes Leme et al., 2008; Cao et al., 2011; Albertsen et al., 2013; Yu et al., 2016); and fourth, *in-situ* 3-dimensional visualization by confocal laser scanning microscopy (Neu and Lawrence, 2014; Schlafer and Meyer, 2017).

The continuous challenge in analyzing the hydrated matrix *in situ* by means of visual techniques lies in its biochemical heterogeneity and the absence of a general contrasting agent. Ideally, a single probe for the extracellular matrix would be very useful either as a direct stain or as a labeled probe. However, even for one of the major matrix components, e.g., polysaccharides as a prominent example, an overall fluorescence staining approach is not available. Consequently, a compromise had to be established by taking advantage of fluorescently-labeled lectins and their glycoconjugate specificity (Neu and Lawrence, 1999; Neu et al., 2001). The lectin approach is inevitable for environmental samples, as there is no alternative such as the production and application of antibodies. Although immune-based techniques are powerful, they may be only applicable to pure culture studies.

This manuscript represents a follow-up on the lectin screening results already published on several pure culture studies with biofilms and bioaggregates (Neu and Kuhlicke, 2017). At the time, we defined the actual lectin screening as fluorescence lectin barcoding (FLBC) and the subsequent lectin analysis as fluorescence lectin-binding analysis (FLBA). FLBC requires a whole list of commercially available lectins for testing a particular sample type. FLBA stands for the tailor-made employment of selected lectins in a defined experiment. In contrast to the previous publication, this manuscript presents lectin data compiled from several environmental studies. We grouped the various FLBC results into the following biofilm systems: (1) derived from rivers, (2) developed in biofilm reactors, (3) wastewater reactor granules, (4) marine samples, and (5) miscellaneous. The lectin data presented in form of a barcoding table may be useful for similar applications and should serve as a guideline for the selection of lectins in studies on glycoconjugates of multispecies and environmental biofilm systems.

# Materials and methods

## Samples examined

Many samples examined by fluorescence lectin barcoding originate from a range of collaborative projects having very different backgrounds and motivations (see Section Acknowledgments). Publication of major results from these projects included a few selected lectin image data sets. Nevertheless, the presentation of the original and extensive lectin screening data is missing. Details of sample origin and development compiled in form of a table with reference to the original articles are given in Table 1.

## Lectins staining and screening

Commercially available lectins purchased from various suppliers, such as Sigma-Aldrich, EY Laboratories, Vector Laboratories, and Molecular Probes (Supplementary Table 1), had labels of green-emitting fluorochromes. They comprised fluorescein isothiocyanate (FITC), fluorescein, or Alexa Fluor 488. For lectin combinations, orange-/red-/far red-labeled lectins are possible, such as tetramethylrhodamine isothiocyanate (TRITC), Texas Red, Cy5, or various Alexa fluorochromes. Conjugation of unlabeled lectins with Alexa fluorochromes using a kit from Molecular Probes according to their protocol allows the attachment of any fluorochrome matching sample properties. The lectins purchased (1 mg/ml buffer) were divided into aliquots and kept at −20°C. This stock solution was diluted at 1:10 for fluorescence staining of the biofilm matrix. Additional details reported elsewhere give further information (Neu and Lawrence, 1999, 2014).

For staining a particular sample, one biofilm or aggregate needs incubation with one lectin. Thus, a screening using, e.g., 80 different lectins requires 80 subsamples. Lectin staining is straightforward and was described in detail previously

TABLE 1 List of habitats and origin of biofilm samples together with the reference of originally published data.

| Habitat | References |
|---|---|
| **Rivers, creeks** | Zippel and Neu, 2011 |
| Westerhöfer creek, tufa | |
| Deinschwanger creek, tufa | Zippel and Neu, 2011 |
| Chriesbach, column | Derlon et al., 2016 |
| Chriesbach, flow cells | Desmond et al., 2018 |
| Elbe, river snow | Luef et al., 2009 |
| Danube, river snow | Luef et al., 2009 |
| **Flow lanes, reactors** | Zippel et al., 2007 |
| WWTP, flow lane, low light | |
| WWTP, flow lane, high light | Zippel et al., 2007 |
| RAR, Elbe river water | Staudt et al., 2003 |
| RAR, Elbe river water, and glucose | Staudt et al., 2003 |
| RAR, Elbe river water, and methanol | Staudt et al., 2003 |
| Paper mill, white water, reactor | Milferstedt et al., 2012 |
| Cooling tower, industry | http://dottoratobee.uniroma2.it/files/2018/03/Di-Gregorio.pdf |
| **Reactor granules** | Gagliano et al., 2018 |
| Anaerobic granules, low salt | |
| Anaerobic granules, high salt | Gagliano et al., 2018 |
| Anaerobic, anammox granules, EPS glycoprotein | Boleij et al., 2018, 2020 |
| Aerobic reactor, flocs and granules | Weissbrodt et al., 2013 |
| Aerobic granules, WWTP and acetate, acid soluble EPS | Pronk et al., 2017 |
| Aerobic granules, WWTP and seawater, sialic acids | de Graaff et al., 2019 |
| Aerobic granules, hyaluronic acid-like/glycosaminoglycans | Felz et al., 2020 |
| Aerobic/anaerobic granules, nonulosonic acids | Tomás-Martínez et al., 2021 |
| **Marine** | Arp et al., 2012 |
| Hypersaline mat, zone 3 | |
| Hypersaline mat, zone 6 | Arp et al., 2012 |
| Hypersaline mat, zone 12 | Arp et al., 2012 |
| North sea, marine snow | Bennke et al., 2013 |
| **Miscellaneous** | Lu et al., 2013 |
| Iron snow, lignite mining lake | |
| Microbial mat, hot spring | Ward et al., 2006 |
| Cave snotties | Karwautz et al., 2018 |

RAR, rotating annular reactor; WWTP, wastewater treatment plant.

(Neu and Lawrence, 1999). In brief, the hydrated sample is exposed to some droplets of lectin solution and incubated for 20 min in the dark. Washing off the unbound lectins several times results in increased contrast. The washing procedure requires matching liquids, e.g., filter-sterilized water, buffer, or a suitable medium (no complex carbohydrate-containing constituents). There are several options for removing unbound lectins according to sample properties and fragility (Neu and Lawrence, 2014). If paraformaldehyde (PFA) fixed samples are used, replacement of PFA with water or buffer is necessary. The overall procedure of lectin staining was critically examined in an earlier report (Neu et al., 2001).

## Sample mounting and assessment

The mounting of samples was according to their origin, properties, and appearance. Biofilms are usually grown on solid surfaces, e.g., plastic cut into pieces of a few cm². One piece was glued into a small 5 cm petri dish using a silicone sealant. The biofilms were stained with the lectin and washed based on their stability using a variety of options (Neu and Lawrence, 2014). Then, the petri dish was flooded with water or buffer and examined with water-immersible (dipping) lenses. To avoid squeezing the structure of flocs, aggregates, or granules, CoverWell chambers with various spacers proved to be ideal for mounting. After staining and washing bioaggregates, examination of the chamber using water immersion lenses through a No. 1.5H coverslip gave the best results.

In the first step, sample assessment was visually in the epifluorescence mode to identify those with no binding or weak binding patterns (no signal or brownish-green signal). If samples showed a strong or intermediate binding (bright green signal), a reference data set recorded in confocal mode is rather helpful for later comparison of lectin staining patterns.

## Confocal laser scanning microscopy

In the course of our study, two confocal laser scanning microscopes were available: a TCS SP1 and a TCS SP5X both with an upright microscope (Leica Microsystems, Germany). The TCS SP1 system was equipped with traditional laser sources (argon 488 nm, DPSS 561 nm, and HeNe 633 nm), controlled by the LCS software version 2.61. The TCS SP5X system was equipped with a supercontinuum laser light source (470–670 nm) and controlled by the software LAS-AF version 2.4.1. For recording image data sets, usually, excitation at 488 nm and collection of the emission signal from 500 to 550 nm (FITC, fluorescein, and Alexa Fluor 488) became standard for green-emitting fluorochromes. In most cases, the following objective lenses were employed: (1) 25× NA 0.95 VISIR Fluotar and 63× NA 0.9 HCX APO both water immersible (biofilms mounted in Petri dishes without a coverslip) and (2) 25× NA 0.95 VISIR Fluotar and 63× NA 1.2 corr CS HCX PL APO both water immersion lenses (flocs, aggregates, and granules in CoverWell chambers with coverslip).

## Image recording and data presentation

For most of the data sets, the general settings of recording parameters comprised as follows: 8-bit data depth, format or pixel resolution 512 × 512 (TCS SP1) or 1,024 × 1,024 (TCS SP5X), scan speed medium (TCS SP1) or 400 Hz (TCS SP5X), step size 1 μm, no average, and no zoom. As a standardized procedure, all the data sets recorded by using the lookup table "glow over under" (GOU) assured an optimal signal-to-noise ratio. Further, the pixel intensities showed only a few saturated pixels and a background level just above zero. Hence, with GOU, the full dynamic range of the pixel intensities is used.

The results of the screening may be presented as a binary color pattern with only black (binding) and white (no binding) information. More information extracted from a heat map facilitates the differentiation of several binding efficiencies. For this purpose, the voltage settings of the photomultiplier tube (PMT) are available as a convenient measure. The following PMT voltage settings defined 400–600 as a strong signal, 600–800 as an intermediate signal, and 800–1,000 as a weak signal. Compilation of the barcoding patterns presented in form of a color-coded Excel sheet (Microsoft) allows the selection of potential lectins for a more detailed assessment. For the final digital image analysis, the software Imaris (Bitplane) for visualization, Huygens (SVI) for deconvolution, and Photoshop (Adobe) for the presentation were available.

# Results

## Epifluorescence microscopy

For the characterization of the extracellular glycoconjugates in a new unknown biofilm or bioaggregate sample *via* FLBC, a screening with many different lectins was required as a basis for a more detailed examination. The commercially available lectins comprised about 70–80 different types. However, their supply changed depending on the availability of raw materials, biochemical isolation, and legal regulations. Hence, the number of available lectins required the same number of samples subsequently examined visually by epifluorescence microscopy. Samples stained with FITC-, fluorescein-, or Alexa Fluor 488-labeled lectins usually showed different shades of green. Weak binding patterns appeared as a brown-green signal, whereas a bright green signal indicated a strong binding signal. In any case, the signal should be matching a microbiological structure. In other words, one should consider the glycoconjugate structures in relation to microbial features (cells, microcolonies, films, flocs, aggregates, granules, and mats) potentially expected. If a positive lectin signal was identified in the epifluorescence mode visually, a sample data set in confocal mode by using the GOU lookup table was recorded for documentation. Any artificial fluorescence signals originating, e.g., from geometric minerals or detritus particles, were neglected. Examples from various screenings given in Figure 1 show different patterns of lectin-specific glycoconjugate signals from environmental samples.

## Binary barcoding

By means of an FLBC screening, often a Yes/No answer was only needed. For this purpose, a listing of the photomultiplier tube (PMT) setting in form of the voltage (sensitivity) for all the positively tested lectins may be enough. This identified the strongest binding pattern by showing the lowest PMT voltage setting. In addition, the transformation of the results into a simple binary barcoding pattern generated a graphic presentation based on a self-defined cutoff value meaning good binding or weak/no binding. There was, however, one aspect to consider, the nature of the lectin signal, meaning the structure of interest in view of the research focus. For example, the significance of glycoconjugates, e.g., on the cell surface, in microcolonies, throughout the extracellular biofilm matrix, and microbe–microbe or microbe–eukaryote interactions.

## Heat map barcoding

More information revealed by grouping the results into several intensity clusters gave further details on lectin binding characteristics (Figure 2). The basis for this heat map was again the voltage setting of the PMT. From experience, we arbitrarily defined three binding efficiencies meaning three PMT ranges: 400–600, 600–800, and 800–1,000 V. As a result, the low sensitivity of the photomultiplier tube (400–600 V) indicated strong glycoconjugate signals, intermediate sensitivity of the photomultiplier tube (600–800 V) indicated good glycoconjugate signals, and high sensitivity of the photomultiplier tube (800–1,000 V) indicated weak glycoconjugate signals. Of note, the image data recorded using the lookup table GOU may look similar in intensity although recorded at different voltages. Furthermore, CLSM data recorded with a high voltage setting may contain some background noise. Therefore, it was necessary to use background subtraction, filtering, or deconvolution to improve the resolution of the data sets. In a second step, the FLBC data from the heat map provided a frame to perform a more detailed FLBA with selected lectins to follow glycoconjugate patterns throughout, e.g., a time experiment, interaction study, or any other biofilm experiment.

## Lectin information

The lectin applicability as revealed by FLBC should serve as a guideline for similar studies. This would avoid buying all the

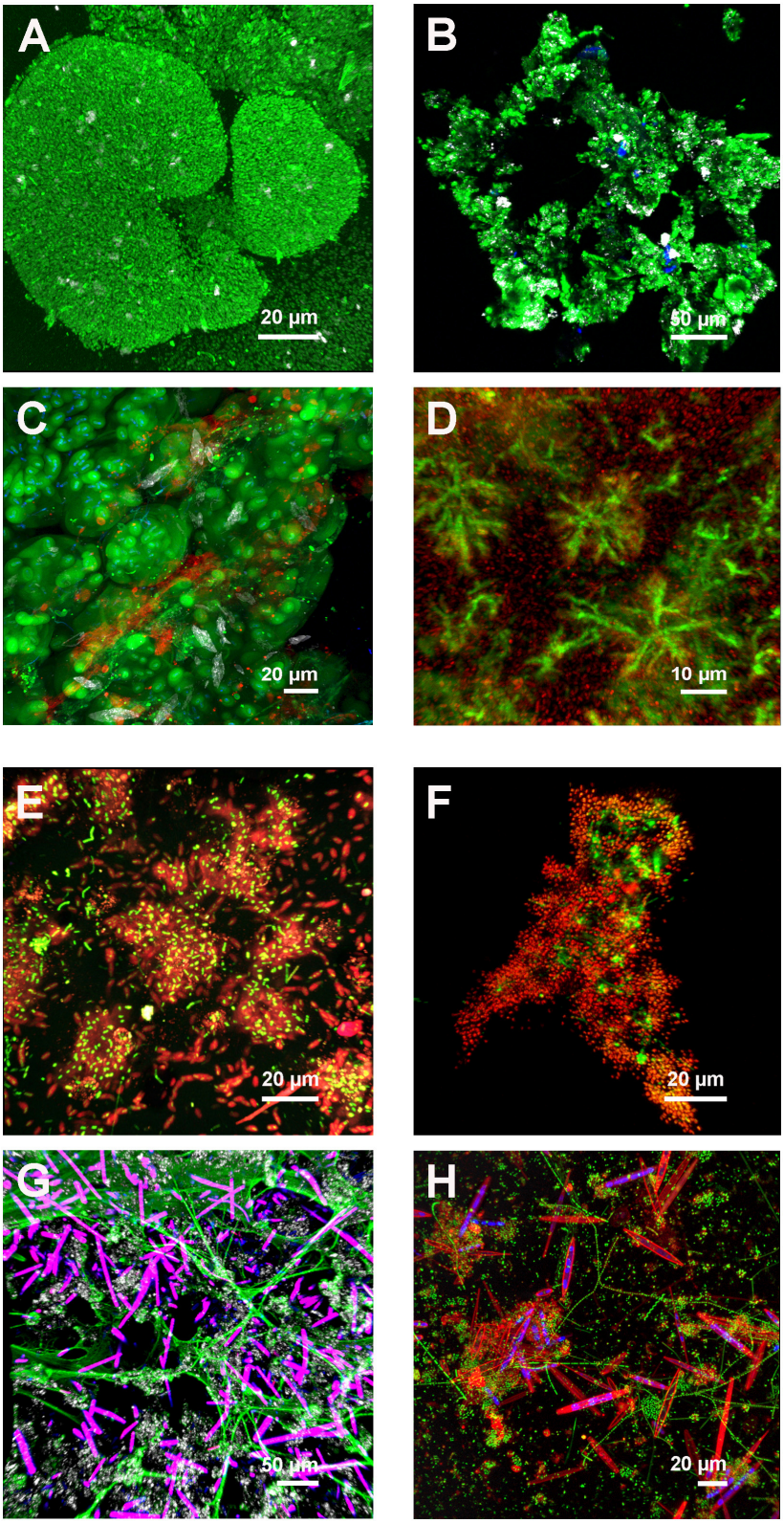



FIGURE 1
Examples of different lectin patterns from environmental biofilm and bioaggregate samples. The multichannel, 3-dimensional data sets recorded by CLSM and projected as maximum intensity projection (MIP) indicate a variety of lectin-specific glycoconjugates. **(A)** Reactor granule
*(Continued)*

FIGURE 1
showing dense bacterial aggregates with cell surface RPA-lectin glycoconjugates and reflection signal. Color allocation: lectin—green, reflection—gray. Axial dimension: 42 μm, 86 optical sections. **(B)** Danube river snow aggregate showing ECA-lectin glycoconjugates. Color allocation: lectin—green, autofluorescence of chlorophyll—blue, and reflection—gray. Axial dimension: 59 μm, 60 optical sections. **(C)** Cave snotty sample with double lectin staining showing globular and filamentous lectin glycoconjugates. Color allocation: AAL-lectin—green, PNA-lectin—red, nucleic acid stain—blue, and reflection—gray. Axial dimension: 74 μm, 75 optical sections. **(D)** Biofilm reactor (white water) with star-like microcolonies linked by VVA-specific lectin glycoconjugates. Color allocation: lectin—green, nucleic acid stain—red. Axial dimension: 87 μm, 88 optical sections. **(E)** Rotating annular reactor biofilm developed from Elbe river water and fed with methanol. The young biofilm shows single cells and microcolonies covered with HPA-lectin glycoconjugates. Color allocation: nucleic acid stain—green, lectin—red. Axial dimension: 29 μm, 30 optical sections. **(F)** Anammox floc from a laboratory reactor showing PHA-E-lectin glycoconjugates. Color allocation: nucleic acid stain—green, lectin—red. Axial dimension: 34 μm, 18 optical sections. **(G)** Freshwater tufa sample showing net-like AAL-lectin glycoconjugates. Color allocation: lectin—green, autofluorescence overlay of phycobilin and chlorophyll—purple, and reflection—gray. Axial dimension: 118 μm, 60 optical sections. **(H)** Flow lane biofilm developed from river Elbe water showing diatoms with cell surface AAL-lectin signal and bacterial colonies embedded in AAL-lectin glycoconjugates. Color allocation: nucleic acid stain—green, lectin—red, and autofluorescence of chlorophyll—blue. Axial dimension: 50 μm, 51 optical sections.

lectins commercially offered. Purchasing the full range of lectins, some of them with different fluorochrome labels would mean an investment in the range of 15–20 K calculated in €.

The signal intensity derived from image data lends itself to counting and shortlisting. In Table 2, those lectins are compiled showing a strong binding efficiency (dark green shading in Figure 2). The lectins listed according to their frequency of binding provided a shortlist of extremely valuable lectins. Table 2 also indicates the inhibitory carbohydrate if known in the literature. Consequently, Table 2 offers a powerful hint of which lectin to select for imaging lectin-specific glycoconjugates within a defined biofilm system. The most frequent lectin identified, AAL, has specificity for α-Fuc. The next best lectins, HAA and WGA, inhibited by β-GlcNAc, have also other specificities (see Table 2). The ConA is probably the most often applied lectin, as it is well characterized, has a bit broader specificity, is readily available, and is also one of the cheapest lectins. The lectin IAA is not well characterized but often showed a rather strong signal. Further down the list, the two lectins, namely HPA and LEA, are specific for α-GalNAc and β-GlcNAc, which were also very successfully applied (see Table 2).

The lectins compiled in Table 3 with intermediate binding efficiency are still very useful for the examination of environmental biofilm systems for their lectin-specific glycoconjugates. The top scorer in this list, GS-I, has specificity for α-Gal and α-GalcNAc. The second lectin, AIA, has again specificity for α-Gal. Interestingly, the next four lectins in the list (LEA, IAA, WGA, and HPA) are also present, as shown in Table 2. This implies that they should be in focus if defining a selection of lectins. The lectins selected either applied as an individual lectin to a single biofilm or as a combination of lectins with different labels applied to one biofilm reveal additional information. The lectins further down the list show a variety of binding specificities and represent valuable probes too. Although some of the lectins might be similar in binding specificity according to the supplier's datasheet, they may show useful signals as these lectins may bind differentially.

The results of Figure 2 can also be evaluated with respect to the habitat and sample origin. Four samples showed a strong binding efficiency with more than 10 lectins, with the "anaerobic reactor–low salt" binding 21 lectins. Four samples showed an intermediate binding efficiency with more than 30 lectins, again with the "anaerobic reactor–low salt" binding of 43 lectins. In general, with all the sample types tested, there were usually enough lectins to select for a more detailed FLBA study. If there are multiple options to choose from, it will be logical either to consider the specificity of the lectins or to look at the lectin signal pattern of interest. By this means, the establishment of a solid base for selecting a panel of useful lectins for a detailed FLBA investigation is feasible.

In view of a possible relationship between lectin binding vs. nonlectin binding and sample type, no clear pattern could be established. The reason might be the different origins of the samples (e.g., rivers and creeks) and the selective enrichment of a specific and differential microbial community (e.g., reactors, flow lanes, and reactor granules).

## Discussion

Visualization of the intact matrix in biofilm systems remains a challenge. Nevertheless, there are attempts to contrast the overall matrix. One approach based on reflection raises some issues and is not established in the field (Swearingen et al., 2016). Another approach, although promoted as a fluorescent matrix stain, the so-called "FilmTracer SYPRO Ruby biofilm matrix stain", remains questionable. In fact, SYPRO Ruby as a general protein-specific fluorochrome will stain all the proteins, including the ones at the cell surface. In the case of polysaccharides as a major matrix polymer, there is no fluorochrome available that binds to all the types of polysaccharides. One fluorochrome sometimes used for matrix visualization is calcofluor white M2R. However, it will detect only one type of polysaccharides having a defined linkage type (β1 → 3, β1 → 4) as, for example, in cellulose (Rasconi et al., 2009). Furthermore, the option of using antibodies, as, for example, in pure culture studies, is not applicable in environmental samples. Consequently, the lectin

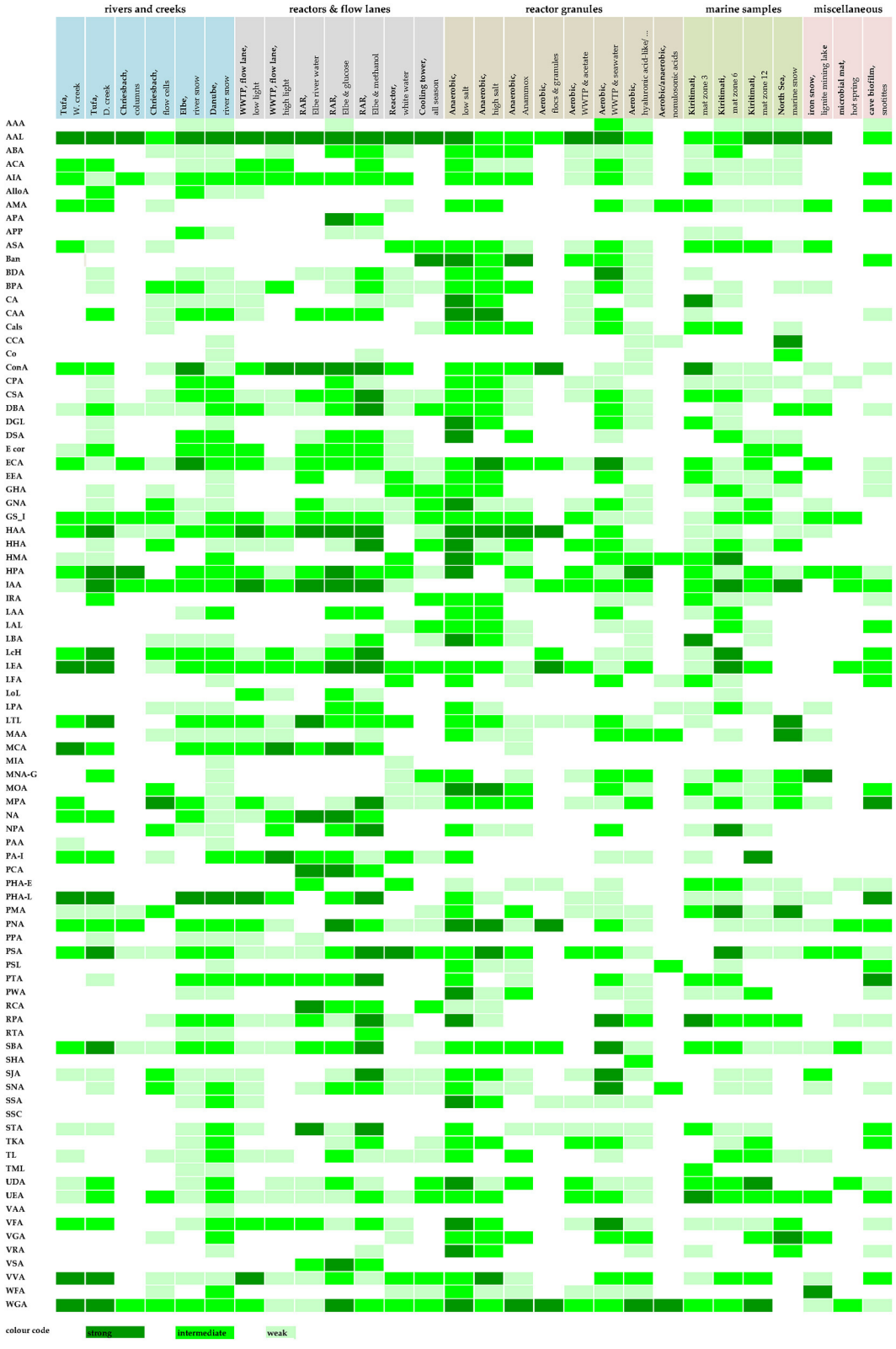



FIGURE 2
Detailed barcoding results of the lectin screening (FLBC) by means of a three color heat map. The signal intensities were color-coded equal to signal intensity: high signal (PMT voltage 400–600), intermediate signal (PMT voltage 600–800), and low signal (PMT voltage 800–1,000). Color coding: strong lectin binding–dark green, intermediate lectin binding–bright green, and low lectin binding–soft green. For abbreviations of lectins, see Supplementary Table 1.

TABLE 2 Shortlist of lectins bound with a strong binding efficiency (see dark green shading in Figure 2).

| Lectin | Specificity |
|---|---|
| AAL | α-Fuc |
| HAA | α-GlcNAc, α-GalNAc |
| WGA | β-GlcNAc, sialic acid |
| ConA | α-Man, α-Glc, α-GlcNAc |
| IAA | n. d. |
| HPA | α-GalNAc |
| LEA | β-GlcNAc |

Lectins are listed according to their binding frequency.
n. d.—lectin specificity not determined.

TABLE 3 Shortlist of lectins bound with an intermediate binding efficiency (see bright green shading in Figure 2).

| Lectin | Specificity |
|---|---|
| GS-I | α-Gal, α-GalcNAc |
| AIA | β-Gal |
| LEA | β-GlcNAc |
| IAA | n.d. |
| WGA | β-GlcNAc, sialic acid |
| HPA | α-GalNAc |
| UEA | α-Fuc, β-GlcNAc |
| ECA | α-Gal, β-Gal, α-GalNAc, β-GalNAc |
| PSA | α-Man, α-Glc, α-GlcNAc |
| ASA | α-Man |
| SBA | α-GalNAc, β-GalNAc |
| PNA | β-Gal |
| VVA | α-Man; α-GalNAc |
| AMA | n. d. |
| CSA | β-Gal |
| MPA | α-Gal, α-GalNAc |
| PA-I | Gal |
| PTA | β-Gal, α-GalNAc, β-GalNAc |
| UDA | β-GlcNAc |
| VFA | α-Man, α-Glc, α-GlcNAc |

Lectins listed according to their binding frequency.
n. d.—lectin specificity not determined.

approach applied to environmental microbiological samples, such as biofilms and bioaggregates, currently represents the only method to detect different matrix glycoconjugates. FLBC and FLBA in combination allow glycoconjugate imaging in hydrated samples, in 3-dimensions and *in situ*. The lectins selected for FLBA can be combined with other fluorochromes as contrasting agents, e.g., for staining cell distribution with nucleic acid-specific fluorochromes. A smart combination of different staining and imaging approaches will allow visualization of a complex 3-dimensional biofilm landscape in multiple channels (Neu and Lawrence, 2015; Lawrence et al., 2016, 2018).

It is important to be aware that if applying lectins, the direction of their binding specificity is not only toward polysaccharides but also to glycoconjugates in general. Glycoconjugates in microbial communities are present in form of glycolipids, glycoproteins, and polysaccharides. Hence, staining biofilm systems with a lectin will reveal different structural features. These may include (1) microbial cell surfaces such as capsules or sheaths (Zippel and Neu, 2011), (2) microbial footprints or holdfasts (Neu and Marshall, 1991; Neu, 1992; Baumgartner et al., 2016), (3) matrix structures within microcolonies (Lawrence et al., 2007, 2016), (4) microbe–microbe or microbe–eukaryote interactions (Kline et al., 2009; Bennke et al., 2013; Ielasi et al., 2016), and (5) overall matrix of the extracellular space (Staudt et al., 2004; Neu et al., 2005). Consequently, one should be aware of the biochemical and structural diversity to which the fluorochrome-labeled lectin may bind.

Lectin application is possible as individual probes and in combination. As indicated, many lectins are available with green-emitting fluorochromes. Other labels comprise orange or red and far-red emitting dyes. If lectins carrying these fluorochromes are not available, there are staining kits, which allow conjugation of any fluorescent color, e.g., with various Alexa dyes. This may be important if there are sample properties occupying a spectral window (e.g., in the far-red by chlorophyll A) or if other fluorochromes are added in combination (e.g., in the green part of the emission spectrum such as SYTO 9). Nevertheless, there is one option of separating multiple green fluorochromes based on the fluorescence lifetime. By means of fluorescence lifetime imaging microscopy (FLIM), separation of fluorochrome emitting in the same region is feasible, if they show a different lifetime. The FLIM technique, however, requires additional hardware and software. Although the FLIM option can be attached to a confocal laser scanning microscopy, it was hardly used in studies of microbial communities. Some examples with a focus on biofilm systems in combination with two-photon laser scanning microscopy are measuring pH in biofilms (Vroom et al., 1999), examination of bacterial activity (Walczysko et al., 2008), and microbial behavior under high pressure (Bourges et al., 2020). To the best of our knowledge, there is no report of FLIM for imaging the complexity of biofilm matrix constituents. In any case, if two or more lectins are applied, one has to make sure that they do not recognize each other forming a precipitate. Pairwise testing of two lectins with different fluorochromes on a microscope slide will easily show potential precipitates (Neu and Kuhlicke, 2017). Thus, a report on the application of a mixture with 20 different lectins has to be looked at critically (Fanesi et al., 2019).

As the lectin information collected by CLSM ends up as a digital data set, it is amenable to quantification. After

thresholding, the pixel (2-dimensional) or voxel (3-dimensional) information may serve as a measure for glycoconjugate production. However, due to the many parameters to be controlled during CLSM, this will be only semiquantitative (Pawley, 2000). Nevertheless, there is a very recent tutorial for quantitative confocal microscopy with guidance to use the CLSM appropriately (Jonkman et al., 2020). Yet another aspect often raised applies to control and inhibition experiments. This issue already discussed reveals differential results, which need evaluation and control with respect to lectin binding characteristics (Zippel and Neu, 2011; Neu and Kuhlicke, 2017).

In the first manuscript on FLBC with pure cultures, the advantages and disadvantages have been elaborated (Neu and Kuhlicke, 2017). A positive aspect is the commercial availability of lectins, whereas a negative aspect may be the limited variation of their specificity. New specificities explored by means of the so-called carbohydrate-binding modules (CBMs) might offer additional probes (Nguyen et al., 2014; Ojima et al., 2015). Other candidates are lectins derived from microbial pili or fimbriae. Clearly, there is a need to have more lectins specific for some of the rare carbohydrates produced by microorganisms and lectins binding to the unique linkages between different microbial monosaccharides. This aspect was addressed in a statistical analysis of the Bacterial Carbohydrate Structure Database (Herget et al., 2008). Furthermore, the potential of lectin microarrays was discussed, as it will allow high-throughput analysis of many samples (Campanero-Rhodes et al., 2020), although this advantage was proven for only pure cultures (Yasuda et al., 2011).

## Conclusion

Fluorescence lectin barcoding constitutes a useful basis for shortlisting many lectins to perform a more detailed fluorescence lectin-binding analysis (FLBA). Thereby, assessment of glycoconjugate distribution and glycoconjugate typing becomes possible (Lawrence et al., 2007, 2016). The lectin approach combined with FISH or CARD-FISH enables the allocation of glycoconjugates to phylogenetic groups of bacteria (Böckelmann et al., 2002; Bennke et al., 2013). In addition to the previous manuscript with a focus on pure culture biofilms (Neu and Kuhlicke, 2017), there is a further important aspect concerning complex environmental biofilm systems. The key point and advantage of the lectin approach, including FLBC and FLBA, lies in its unique feature for detecting diverse glycoconjugates and its direct applicability on hydrated biofilm and bioaggregate samples. This aspect is crucial; as for environmental samples, it is impossible to apply immune techniques, e.g., producing antibodies against the vast diversity of potential matrix glycoconjugates. In terms of the lectin approach, new specificities are needed either by searching for new lectins or looking for similar proteins such as carbohydrate-binding modules as discussed previously (Neu and Kuhlicke, 2017). Finally, there is still a need for a fluorescent carbohydrate stain, which will allow contrasting the overall polysaccharide matrix in biofilm systems. Very likely, multiple strategies are necessary to address the large variety of polysaccharides, carbohydrate linkages, and the partly unique and exotic glycoconjugates present in microbial communities. All of these strategies should ideally match the techniques employed for *in-situ* visualization and analysis of other major matrix compounds, such as extracellular proteins and extracellular nucleic acids, amphiphilic compounds, and microbial-derived refractory constituents.

## Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

## Author contributions

TN and UK worked at the confocal laser scanning microscopy analyzing the samples. UK handled the huge number of data sets recorded and compiled the lectin table/figure of the article. TN wrote the manuscript. All authors contributed to the article and approved the submitted version.

## Funding

TN and UK are thankful for the continuous financial support from the Helmholtz Centre for Environmental Research—UFZ in Leipzig over many years of laser scanning microscopy.

## Acknowledgments

We highly appreciate the cooperation of many people from various research areas who took advantage of the lectin approach in our laboratory. As a result, FLBC, and subsequent FLBA applied to many different environmental biofilm and bioaggregate samples allowed us to characterize the glycoconjugates of the extracellular microbial matrix in complex environmental samples: Christin Bennke (Rudolf Amann), Nicolas Derlon and Peter Desmond (Eberhard Morgenroth), Maria Cristina Gagliano (Caroline M Plugge), Luciana Di Gregorio (Francesca Di Pippo), Clemens Karwautz (Tillmann Lüders), Birgit Luef (Peter Peduzzi), Yuemei Lin (Mark van Loosdrecht), Shipeng Lu (Kirsten Küsel), Kim Milferstedt (Nicolas Bernet), Mari Raulio, Maya Roberts (Uta Passow and Matthias Ulrich), David Weissbrodt (Christof Holliger), and Christian Staudt and Barbara Zippel (both Research Group—Microbiology of Interfaces, UFZ).

## 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.2022.940280/full#supplementary-material

## References


Albertsen, M., Stensballe, A., Nielsen, K. L., and Nielsen, P.-H. (2013). Digging into the extracellular matrix of a complex microbial community using a combined metagenomic and metaproteomic approach. *Water Sci. Technol.* 67, 1650–1656. doi: 10.2166/wst.2013.030

Arp, G., Helms, G., Karlinska, K., Schumann, G., Reimer, A., Reitner, J., et al. (2012). Photosynthesis versus Exopolymer degradation in the formation of microbialites on the atoll of Kiritimati, Republic of Kiribati, Central Pacific. *Geomicrobiol. J.* 29, 29–65. doi: 10.1080/01490451.2010.521436

Baumgartner, M., Neu, T. R., Blom, J. F., and Pernthaler, J. (2016). Protistan predation interferes with bacterial long-term adaptation to substrate restriction by selecting for defence morphotypes. *J. Evol. Biol.* 29, 2297–2310. doi: 10.1111/jeb.12957

Bennke, C. M., Neu, T. R., Fuchs, B. M., and Amann, R. (2013). Mapping glycoconjugate-mediated interactions of marine Bacteroidetes with diatoms. *Syst. Appl. Microbiol.* 36, 417–425. doi: 10.1016/j.syapm.2013.05.002

Böckelmann, U., Manz, W., Neu, T. R., and Szewzyk, U. (2002). A new combined technique of fluorescent *in situ* hybridization and lectin-binding-analysis (FISH-LBA) for the investigation of lotic microbial aggregates. *J. Microbiol. Methods* 49, 75–87. doi: 10.1016/S0167-7012(01)00354-2

Boleij, M., Kleikamp, H., Pabst, M., Neu, T. R., van Loosdrecht, M. C. M., and Lin, Y. (2020). Decorating the anammox house: sialic acids and sulfated glycosaminoglycans in the extracellular polymeric substances of anammox granular sludge. *Environ. Sci. Technol.* 54, 5218–5226. doi: 10.1021/acs.est.9b07207

Boleij, M., Pabst, M., Neu, T. R., van Loosdrecht, M. C. M., and Lin, Y. (2018). Identification of glycoproteins isolated from extracellular polymeric substances of full-scale anammox granular sludge. *Environ. Sci. Technol.* 52, 13127–13135. doi: 10.1021/acs.est.8b03180

Bourges, A. C., Lazarev, A., Declerck, N., Rogers, K. L., and Royer, C. A. (2020). Quantitative high-resolution imaging of live microbial cells at high hydrostatic pressure. *Biophys. J.* 118, 2670–2679. doi: 10.1016/j.bpj.2020.04.017

Campanero-Rhodes, M. A., Palma, A. S., Menéndez, M., and Solís, D. (2020). Microarray strategies for exploring bacterial surface glycans and their interactions with glycan-binding proteins. *Front. Microbiol.* 10, 2909. doi: 10.3389/fmicb.2019.02909

Cao, B., Shi, L., Brown, R. N., Xiong, Y., Fredrickson, J. K., Romine, M. F., et al. (2011). Extracellular polymeric substances from *Shewanella sp.* HRCR-1 biofilms: characterization by infrared spectroscopy and proteomics. *Environ. Microbiol.* 13, 1018–1031. doi: 10.1111/j.1462-2920.2010.02407.x

D'Abzac, P., Bordas, F., van Hullebusch, E., Lens, P. N. L., and Guibaud, G. (2010). Extraction of extracellular polymeric substances (EPS) from anaerobic granular sludge: comparison of chemical and physical extraction protocols. *Appl. Microbiol. Biotechnol.* 85, 1589–1599. doi: 10.1007/s00253-009-2288-x

de Graaff, D. R., Felz, S., Neu, T. R., Pronk, M., van Loosdrecht, M. C. M., and Lin, Y. (2019). Sialic acids in the extracellular polymeric substances of seawater-adapted aerobic granular sludge. *Water Res.* 155, 343–351. doi: 10.1016/j.watres.2019.02.040

Derlon, N., Grütter, A., Brandenberger, F., Sutter, A., Kuhlicke, U., Neu, T. R., et al. (2016). The composition and compression of biofilms developed on ultrafiltration membranes determine hydraulic biofilm resistance. *Water Res.* 102, 63–72. doi: 10.1016/j.watres.2016.06.019

Desmond, P., Best, J. P., Morgenroth, E., and Derlon, N. (2018). Linking composition of extracellular polymeric substances (EPS) to the physical structure and hydraulic resistance of membrane biofilms. *Water Res.* 132, 211–221. doi: 10.1016/j.watres.2017.12.058

Dragoš, A., and Kovács, Á. T. (2017). The peculiar functions of the bacterial extracellular matrix. *Trends Microbiol.* 25, 257–266. doi: 10.1016/j.tim.2016.12.010

Dumas, E., Meunier, B., Berdague, J.-L., Chambon, C., Desvaux, M., and Hebraud, M. (2008). Comparative analysis of extracellular and intracellular proteomes of *Listeria monocytogenes* strains reveals a correlation between protein expression and serovar. *Appl. Environ. Microbiol.* 74, 7399–7409. doi: 10.1128/AEM.00594-08

Fanesi, A., Paule, A., Bernard, O., Briandet, R., and Lopes, F. (2019). The architecture of monospecific microalgae biofilms. *Microorganisms* 7, 352. doi: 10.3390/microorganisms7090352

Felz, S., Al-Zuhairy, S., Aarstad, O. A., van Loosdrecht, M. C. M., and Lin, Y. M. (2016). Extraction of structural extracellular polymeric substances from aerobic granular sludge. *JoVE* 115, e54534. doi: 10.3791/54534

Felz, S., Neu, T. R., van Loosdrecht, M. C. M., and Lin, Y. (2020). Aerobic granular sludge contains Hyaluronic acid-like and sulfated glycosaminoglycans-like polymers. *Water Res.* 169, 115291. doi: 10.1016/j.watres.2019.115291

Flemming, H.-C., and Wingender, J. (2010). The biofilm matrix. *Nat. Rev. Micro.* 8, 623–633. doi: 10.1038/nrmicro2415

Gagliano, M. C., Neu, T. R., Kuhlicke, U., Sudmalis, D., Temmink, H., and Plugge, C. M. (2018). EPS glycoconjugate profiles shift as adaptive response in anaerobic microbial granulation at high salinity. *Front. Microbiol.* 9, 1423. doi: 10.3389/fmicb.2018.01423

Herget, S., Toukach, P. V., Ranzinger, R., Hull, W. E., Knirel, Y. A., and von der Lieth, C.-W. (2008). Statistical analysis of the Bacterial Carbohydrate Structure Data Base (BCSDB): characteristics and diversity of bacterial carbohydrates in comparison with mammalian glycans. *BMC Struct. Biol.* 8, 35. doi: 10.1186/1472-6807-8-35

Hobley, L., Harkins, C., MacPhee, C. E., and Stanley-Wall, N. R. (2015). Giving structure to the biofilm matrix: an overview of individual strategies and emerging common themes. *FEMS Microbiol. Rev.* 39, 649–669. doi: 10.1093/femsre/fuv015

Ielasi, F. S., Alioscha-Perez, M., Donohue, D., Claes, S., Sahli, H., Schols, D., et al. (2016). Lectin-glycan interaction network-based identification of host receptors of microbial pathogenic adhesins. *mBio* 7, e00584. doi: 10.1128/mBio.00584-16

Jonkman, J., Brown, C. M., Wright, G. D., Anderson, K. I., and North, A. J. (2020). Tutorial: guidance for quantitative confocal microscopy. *Nature Protocols* 15, 1585–1611. doi: 10.1038/s41596-020-0313-9

Karwautz, C., Kus, G., Stöckl, M., Neu, T. R., and Lueders, T. (2018). Microbial megacities fueled by methane oxidation in a mineral spring cave. *ISME J.* 12, 87–100. doi: 10.1038/ismej.2017.146

Kline, K. A., Fälker, S., Dahlberg, S., Normark, S., and Henriques-Normark, B. (2009). Bacterial adhesins in host-microbe interactions. *Cell Host Microbe* 5, 580–592. doi: 10.1016/j.chom.2009.05.011

Lawrence, J. R., Swerhone, G. D. W., Kuhlicke, U., and Neu, T. R. (2007). *In situ* evidence for microdomains in the polymer matrix of bacterial microcolonies. *Can. J. Microbiol.* 53, 450–458. doi: 10.1139/W06-146

Lawrence, J. R., Swerhone, G. D. W., Kuhlicke, U., and Neu, T. R. (2016). *In situ* evidence for metabolic and chemical microdomains in the structured polymer matrix of bacterial microcolonies. *FEMS Microbiol. Ecol.* 93, fiw183. doi: 10.1093/femsec/fiw183

Lawrence, J. R., Winkler, M., and Neu, T. R. (2018). Multi-parameter laser imaging reveals complex microscale biofilm matrix in a thick (4,000 μm) aerobic methanol oxidizing community. *Front. Microbiol.* 9, e02186. doi: 10.3389/fmicb.2018.02186

Loustau, E., Rols, J.-L., Leflaive, J., Marcato-Romain, C.-E., and Girbal-Neuhauser, E. (2018). Comparison of extraction methods for the characterization of extracellular polymeric substances from aggregates of three biofilm-forming phototrophic microorganisms. *Can. J. Microbiol.* 64, 887–899. doi: 10.1139/cjm-2018-0182

Lu, S., Chourey, K., Reiche, M., Nietzsche, S., Shah, M. B., Neu, T. R., et al. (2013). Insights into the structure and metabolic function of microbes that shape pelagic iron-rich aggregates (iron snow). *Appl. Environ. Microbiol.* 79, 4272–4281. doi: 10.1128/AEM.00467-13

Luef, B., Neu, T. R., Zweimüller, I., and Peduzzi, P. (2009). Structure and composition of aggregates in two large European rivers, based on confocal laser scanning microscopy and image and statistical analysis. *Appl. Environ. Microbiol.* 75, 5952–5962. doi: 10.1128/AEM.00186-09

Milferstedt, K., Godon, J. J., Escudié, R., Prasse, S., Neyret, C., and Bernet, N. (2012). Heterogeneity and spatial distribution of bacterial background contamination in pulp and process water of a paper mill. *J. Indust. Microbiol. Biotechnol.* 39, 1751–1759. doi: 10.1007/s10295-012-1196-8

Neu, T., and Kuhlicke, U. (2017). Fluorescence lectin bar-coding of glycoconjugates in the extracellular matrix of biofilm and bioaggregate forming microorganisms. *Microorganisms* 5, 5. doi: 10.3390/microorganisms5010005

Neu, T. R. (1992). Microbial "footprints" and the general ability of microorganisms to label interfaces. *Can. J. Microbiol.* 38, 1005–1008. doi: 10.1139/m92-165

Neu, T. R., and Lawrence, J. R. (1999). Lectin-binding-analysis in biofilm systems. *Methods Enzymol.* 310, 145–152. doi: 10.1016/S0076-6879(99)10012-0

Neu, T. R., and Lawrence, J. R. (2014). "Advanced techniques for *in situ* analysis of the biofilm matrix (structure, composition, dynamics) by means of laser scanning microscopy," in *Microbial Biofilms - Methods and Protocols. Methods in Molecular Biology*, eds G. Donelli (New York, NY: Springer), 43–64.

Neu, T. R., and Lawrence, J. R. (2015). Innovative techniques, sensors, and approaches for imaging biofilms at different scales. *Trends Microbiol.* 23, 233–242. doi: 10.1016/j.tim.2014.12.010

Neu, T. R., and Lawrence, J. R. (2017). "The extracellular matrix - an intractable part of biofilm systems," in *The Perfect Slime*, eds H.-C. Flemming, T. R. Neu, and J. Wingender (London: IWA Publishing).

Neu, T. R., and Marshall, K. C. (1991). Microbial "footprints" - a new approach to adhesive polymers. *Biofouling* 3, 101–112. doi: 10.1080/08927019109378166

Neu, T. R., Swerhone, G. D. W., Böckelmann, U., and Lawrence, J. R. (2005). Effect of CNP on composition and structure of lotic biofilms as detected with lectin-specific glycoconjugates. *Aquatic Microb. Ecol.* 38, 283–294. doi: 10.3354/ame038283

Neu, T. R., Swerhone, G. D. W., and Lawrence, J. R. (2001). Assessment of lectin-binding analysis for *in situ* detection of glycoconjugates in biofilm systems. *Microbiology* 147, 299–313. doi: 10.1099/00221287-147-2-299

Nguyen, M. H., Ojima, Y., Sakka, M., Sakka, K., and Taya, M. (2014). Probing of exopolysaccharides with green fluorescence protein-labeled carbohydrate-binding module in Escherichia coli biofilms and flocs induced by bcsB overexpression. *J. Biosci. Bioeng.* 118, 400–405. doi: 10.1016/j.jbiosc.2014.03.005

Ojima, Y., Suparman, A., Nguyen, M. H., Sakka, M., Sakka, K., and Taya, M. (2015). Exopolysaccharide assay in *Escherichia coli* microcolonies using a cleavable fusion protein of GFP-labeled carbohydrate-binding module. *J. Microbiol. Methods* 114, 75–77. doi: 10.1016/j.mimet.2015.05.013

Paes Leme, A. F., Bellato, C. M., Bedi, G., Del Bel Cury, A. A., Koo, H., and Cury, J. A. (2008). Effects of Sucrose on the extracellular matrix of plaque-like biofilm formed *in vivo*, studied by proteomic analysis. *Caries Res.* 42, 435–443. doi: 10.1159/000159607

Pawley, J. B. (2000). The 39 steps: a cautionary tale of quantitative 3-d fluorescence microscopy. *BioTechniques* 28, 884–887. doi: 10.2144/00285bt01

Pellicer-Nacher, C., Domingo-Felez, C., Mutlu, A. G., and Smets, B. F. (2013). Critical assessment of extracellular polymeric substances extraction methods from mixed culture biomass. *Water Res.* 47, 5564–5574. doi: 10.1016/j.watres.2013.06.026

Pronk, M., Neu, T. R., van Loosdrecht, M. C. M., and Lin, Y. M. (2017). The acid soluble extracellular polymeric substance of aerobic granular sludge dominated by Defluviicoccus sp. *Water Res.* 122, 148–158. doi: 10.1016/j.watres.2017.05.068

Rasconi, S., Jobard, M., Jouve, L., and Sime-Ngando, T. (2009). Use of calcofluor white for detection, identification, and quantification of phytoplanktonic fungal parasites. *Appl. Environ. Microbiol.* 75, 2545–2553. doi: 10.1128/AEM.02211-08

Schlafer, S., and Meyer, R. L. (2017). Confocal microscopy imaging of the biofilm matrix. *J. Microbiol. Methods* 138, 50–59. doi: 10.1016/j.mimet.2016.03.002

Staudt, C., Horn, H., Hempel, D. C., and Neu, T. R. (2003). "Screening of lectins for staining lectin-specific glycoconjugates in the EPS of biofilms," in *Biofilms in Medicine, Industry and Environmental Technology*, eds P. Lens, A. P. Moran, T. Mahony, P. Stoodley, and V. O'Flaherty (London, UK: IWA Publishing), 308–327.

Staudt, C., Horn, H., Hempel, D. C., and Neu, T. R. (2004). Volumetric measurements of bacterial cells and extracellular polymeric substance glycoconjugates in biofilms. *Biotechnol. Bioeng.* 88, 585–592. doi: 10.1002/bit.20241

Steinberg, N., and Kolodkin-Gal, I. (2015). The matrix reloaded: how sensing the extracellular matrix synchronizes bacterial communities. *J. Bacteriol.* 197, 2092–2103. doi: 10.1128/JB.02516-14

Sun, M., Li, W. W., Yu, H. Q., and Harada, H. (2012). A novel integrated approach to quantitatively evaluate the efficiency of extracellular polymeric substances (EPS) extraction process. *Appl. Microbiol. Biotechnol.* 96, 1577–1585. doi: 10.1007/s00253-012-4478-1

Swearingen, M. C., Mehta, A., Mehta, A., Nistico, L., Hill, P. J., Falzarano, A. R., et al. (2016). A novel technique using potassium permanganate and reflectance confocal microscopy to image biofilm extracellular polymeric matrix reveals non-eDNA networks in Pseudomonas aeruginosa biofilms. *Pathogens Dis.* 74, ftv104. doi: 10.1093/femspd/ftv104

Tomás-Martínez, S., Kleikamp, H. B. C., Neu, T. R., Pabst, M., Weissbrodt, D. G., van Loosdrecht, M. C. M., et al. (2021). Production of nonulosonic acids in the extracellular polymeric substances of "*Candidatus* Accumulibacter phosphatis". *Appl. Microbiol. Biotechnol.* 105, 3327–3338. doi: 10.1101/2020.11.02.365007

Vroom, J. M., Grauw, d,. C. J., Gerritsen, H. C., Bradshaw, A. M., Marsh, P. D., Watson, G. K., et al. (1999). Depth penetration and detection of pH gradients in biofilms by two-photon excitation microscopy. *Appl. Environ. Microbiol.* 65, 3502–3511. doi: 10.1128/AEM.65.8.3502-3511.1999

Walczysko, P., Kuhlicke, U., Knappe, S., Cordes, C., and Neu, T. R. (2008). *In situ* activity measurement of suspended and immobilized microbial communities by Fluorescence Lifetime Imaging (FLIM). *Appl. Environ. Microbiol.* 74, 294–299. doi: 10.1128/AEM.01806-07

Ward, D. M., Bateson, M. M., Ferris, M. J., Kühl, M., Wieland, A., Koeppel, A., et al. (2006). Cyanobacterial ecotypes in the microbial mat community of Mushroom Spring (Yellowstone National Park, Wyoming) as species-like units linking microbial community composition, structure and function. *Philos. Trans. R. Soc. Biol. Sci.* 361, 1997–2008. doi: 10.1098/rstb.2006.1919

Weissbrodt, D. G., Neu, T. R., Kuhlicke, U., Rappaz, Y., and Holliger, C. (2013). Assessment of bacterial and structural dynamics in aerobic granular biofilms. *Front. Microbiol.* 4, e00175. doi: 10.3389/fmicb.2013.00175

Wong, L. L., Natarajan, G., Boleij, M., Thi, S. S., Winnerdy, F. R., Mugunthan, S., et al. (2020). Extracellular protein isolation from the matrix of anammox biofilm using ionic liquid extraction. *Appl. Microbiol. Biotechnol.* 104, 3643–3654. doi: 10.1007/s00253-020-10465-7

Yasuda, E., Tateno, H., Hirabarashi, J., Iino, T., and Sako, T. (2011). Lectin microarray reveals binding profiles of *Lactobacillus casei* strains in a comprehensive analysis of bacterial cell wall polysaccharides. *Appl. Environ. Microbiol.* 77, 4539–4546. doi: 10.1128/AEM.00240-11

Yu, W., Chen, Z., Shen, L., Wang, Y., Li, Q., Yan, S., et al. (2016). Proteomic profiling of Bacillus licheniformis reveals a stress response mechanism in the synthesis of extracellular polymeric flocculants. *Biotechnol. Bioeng.* 113, 797–806. doi: 10.1002/bit.25838

Zhang, L., Ren, H., and Ding, L. (2012). Comparison of extracellular polymeric substances (EPS) extraction from two different activated sludges. *Water Sci. Technol.* 66, 1558–1564. doi: 10.2166/wst.2012.295

Zippel, B., and Neu, T. R. (2011). Characterization of glycoconjugates of extracellular polymeric substances in tufa-associated biofilms by using fluorescence lectin-binding analysis. *Appl. Environ. Microbiol.* 77, 506–516. doi: 10.1128/AEM.01660-10

Zippel, B., Rijstenbil, J., and Neu, T. R. (2007). A flow-lane incubator for studying freshwater and marine phototrophic biofilms. *J. Microbiol. Methods* 70, 336–345. doi: 10.1016/j.mimet.2007.05.013

frontiers | Frontiers in Microbiology

TYPE Original Research
PUBLISHED 22 August 2022
DOI 10.3389/fmicb.2022.958660



OPEN ACCESS

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

REVIEWED BY
Tan Suet May Amelia,
University of Malaysia Terengganu, Malaysia
Jamseel Moopantakath,
Central University of Kerala, India

*CORRESPONDENCE
Russell G. Kerr
rkerr@upei.ca

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

RECEIVED 31 May 2022
ACCEPTED 05 August 2022
PUBLISHED 22 August 2022

CITATION
Pope E, Cartmell C, Haltli B, Ahmadi A and Kerr RG (2022) Microencapsulation and *in situ* incubation methodology for the cultivation of marine bacteria. *Front. Microbiol.* 13:958660. doi: 10.3389/fmicb.2022.958660



# Microencapsulation and *in situ* incubation methodology for the cultivation of marine bacteria

Emily Pope[1], Christopher Cartmell[2], Bradley Haltli[1,3], Ali Ahmadi[1,4,5] and Russell G. Kerr[1,2,3]*

[1]Department of Biomedical Science, University of Prince Edward Island, Charlottetown, PE, Canada, [2]Department of Chemistry, University of Prince Edward Island, Charlottetown, PE, Canada, [3]Nautilus Biosciences Croda, Charlottetown, PE, Canada, [4]Faculty of Sustainable Design Engineering, University of Prince Edward Island, Charlottetown, PE, Canada, [5]Department of Mechanical Engineering, École de technologie supérieure (ÉTS), Montreal, QC, Canada

Environmental microorganisms are important sources of biotechnology innovations; however, the discovery process is hampered by the inability to culture the overwhelming majority of microbes. To drive the discovery of new biotechnology products from previously unculturable microbes, several methods such as modification of media composition, incubation conditions, single-cell isolation, and *in situ* incubation, have been employed to improve microbial recovery from environmental samples. To improve microbial recovery, we examined the effect of microencapsulation followed by *in situ* incubation on the abundance, viability, and diversity of bacteria recovered from marine sediment. Bacteria from marine sediment samples were resuspended or encapsulated in agarose and half of each sample was directly plated on agar and the other half inserted into modified Slyde-A-Lyzer™ dialysis cassettes. The cassettes were incubated in their natural environment (*in situ*) for a week, after which they were retrieved, and the contents plated. Colony counts indicated that bacterial abundance increased during *in situ* incubation and that cell density was significantly higher in cassettes containing non-encapsulated sediment bacteria. Assessment of viability indicated that a higher proportion of cells in encapsulated samples were viable at the end of the incubation period, suggesting that agarose encapsulation promoted higher cell viability during *in situ* incubation. One hundred and 46 isolates were purified from the study (32–38 from each treatment) to assess the effect of the four treatments on cultivable bacterial diversity. In total, 58 operational taxonomic units (OTUs) were identified using a 99% 16S rRNA gene sequence identity threshold. The results indicated that encapsulation recovered greater bacterial diversity from the sediment than simple resuspension (41 vs. 31 OTUs, respectively). While the cultivable bacterial diversity decreased by 43%–48% after *in situ* incubation, difficult-to-culture (*Verrucomicrobia*) and obligate marine (*Pseudoalteromonas*) taxa were only recovered after *in situ* incubation. These results suggest that agarose encapsulation coupled with *in situ* incubation in commercially available, low-cost, diffusion chambers facilitates the cultivation and improved recovery of bacteria from marine sediments. This study provides another tool that microbiologists can use to access microbial dark matter for environmental, biotechnology bioprospecting.

KEYWORDS

**microencapsulation, marine bacteria, microfluidics, biodiversity, microbiome, natural products**

## Introduction

Environmental microorganisms are the source of a myriad of biotechnological products that have a substantial impact on society. These products range from enzymes that improve the effectiveness of laundry detergents to life-saving antibiotics (Kumar et al., 2008; Bérdy, 2012). The ability to culture a truly representative selection of microbes from environmental samples represents a seminal challenge within the fields of environmental microbiology and natural product discovery. As a result of this elusive microbial dark matter, the true biodiversity of the microbial world remains inaccurately represented by standard bacterial culture to date (Dias et al., 2012). This phenomenon has been named the "Great Plate Count Anomaly" and describes the discrepancy between the diversity of microbes observed in an environmental sample microscopically or detected *via* culture-independent methods, and those that can be cultured in the laboratory. Depending on the habitat, the proportion of microbes that are cultivable using standard culture techniques is estimated to be ~1% (Roberts and Caserio, 1977; Butler, 2004; Kinghorn et al., 2009; Krause and Tobin, 2013; Glatstein et al., 2014; Montinari et al., 2019). Over 100,000 bioactive natural products have been isolated from cultured microorganisms (Bérdy, 2012); thus, unculturable bacteria represent a significant untapped resource for biotechnology bioprospecting (Lewis and Elvin-Lewis, 2003; Cragg and Newman, 2005; Zhang and Demain, 2005). Natural product discovery reached its peak during the twentieth century with the discovery of numerous antibacterial, antifungal, antiviral, immunosuppressant, antiparasitic, anticholesterolemic, and anticancer drugs, however, the discovery rate of new natural product-derived drugs has significantly decreased over the past three decades. Development of new strategies to access the natural product repertoire of uncultured bacteria represents an attractive strategy to reinvigorate natural product discovery (Yazdankhah et al., 2013; Bbosa et al., 2014; Katz and Baltz, 2016; Newman and Cragg, 2016).

To counteract the "Great Plate Count Anomaly" and ultimately assist in the discovery of novel natural products, several strategies have been devised. Careful modifications of cultivation conditions by altering the media composition (e.g., tailoring salinity and pH and addition of habitat-derived extracts) or changing growth conditions (e.g., incubation time, temperature, pressure, salinity, or oxygen levels) have led to the successful cultivation of several new taxa of bacteria (Schut et al., 1993; Vishnivetskaya et al., 2000; Bruns et al., 2002; Sait et al., 2006; Omsland et al., 2009; Pham and Kim, 2012; Lok, 2015; Kawasaki and Kamagata, 2017; Kato et al., 2018); however, limited *a priori* knowledge of the metabolic requirements of diverse taxa limits the use of these strategies. Single-cell isolation strategies have also resulted in the cultivation of unique taxa, however, this methodology does not allow for cell-to-cell communication, which is required for the growth of some bacteria (Perbal, 2003). Cellular communication often involves the transmission of signal molecules, such as auto-inducers or peptides, which may play a role in cell growth (Perbal, 2003). To account for the dependence of some taxa on interactions or signaling from other organisms, co-culturing or community culturing techniques have been developed. These techniques use one or more helper bacteria to promote the growth of dependent bacteria (Pham and Kim, 2012). This approach is complicated by the large number of combinations that may be used, as well as the differing growth requirements and uneven growth rates of bacterial combinations (Nai and Meyer, 2018).

To further improve the recovery of "difficult-to-culture" microorganisms, *in situ* incubation has emerged as an effective technique. This approach is typically combined with single-cell isolation, and then the isolated cells are exposed to the natural chemical and nutritional conditions of their original habitat to promote replication, after which the cells are more amenable to cultivation in the laboratory (Gerwick and Moore, 2012; Jacoby, 2017). This technique has evolved from low throughput, large hollow diffusion chambers constructed of washers with permeable membranes forming the boundary between the chamber and the environments to the Ichip, which miniaturized this concept and encapsulated bacteria in 384 small agar plugs sandwiched between permeable membranes (Aoi et al., 2009; Nichols et al., 2010; Berdy et al., 2017). These devices allow diffusion of nutrients and other signals into the chamber while preventing migration of bacterial cells into, or out of the chamber. The primary advantage of these systems is that they provide the numerous environmental conditions necessary for bacterial growth that are difficult to identify and provide under laboratory settings (Kong et al., 2010; Gerwick and Moore, 2012); however, as the cells are isolated they do not allow for nutrient and/or signal exchange between members of a microbial community, which may limit the ability to culture taxa that require these interactions (Blunt et al., 2018). We recently reported the use of a modified Ichip for use in sponges which led to the isolation of previously uncultured bacteria and the subsequent identification of a new tyrosine derivative (MacIntyre et al., 2019).

In this study, we set out to build upon the diffusion chamber concept. We encapsulated bacteria extracted from marine sediment in agarose-based micro beads and then inserted the beads into a modified dialysis cassette, which was then incubated *in situ* for 1 week. To evaluate the effect of this approach on the recoverable bacterial diversity we compared the bacterial diversity obtained from this method to: (1) bacteria dislodged from the sediment, inserted in the dialysis cassette, and incubated *in situ*, (2) bacteria dislodged from the sediment and immediately plated, and (3) bacteria dislodged from the sediment, encapsulated in agarose, and then immediately plated.

## Materials and methods

### Collection and preparation of environmental bacteria

Intertidal sediment and near-shore seawater were collected from Charlottetown Harbour, PE (46.238 °N, 63.151 °W) in July

(water temperature: 17.8°C; 24.4 ppt salinity; pH 8.11) using sterile Whirl-Pak.® Sediment was separated into 10 ml aliquots and centrifuged at 4,500 *g* for 5 min and the excess seawater was decanted. Ten milliliters of diluted (1:10) Marine Broth; (dMB; BD Difco™, Fisher Scientific, United States) was added to each sediment sample, shaken by hand vigorously for 60 s then centrifuged at 4,500*g* for 5 min. The supernatant was decanted, and this process was repeated twice to remove excess salt water. To dislodge sediment-associated bacteria from sediment particles, 10 ml of dMB was added to the washed sediment sample and vortexed for 60 s. The tube was then shaken horizontally on a shaking platform at 400 rpm for 2 h and then vortexed for an additional 30 s. Large sediment particles were allowed to settle for 60 min at room temperature (RT) and the supernatant was transferred to a new tube. The concentration of bacteria in each sample was determined *via* fluorescent microscopy. Cells were stained using a live/dead bacterial viability kit (PromoCell, Germany) according to the manufacturer's recommendations. Cells were counted using INCYTO C-Chip disposable hemocytometers (VWR, Radnor, PA) observed under a REVOLVE 4 upright inverted brightfield fluorescent microscope (Echo Laboratories, San Diego, CA) according to the manufacturer's recommendations.

The volume necessary to attain ~$8.35 \times 10^6$ cells/ml was centrifuged at 4,500*g* for 10 min to concentrate the cells. Cells obtained from the sediment samples were processed in two ways. Cells were either resuspended in RT dMB by gentle aspiration or gently resuspended in 1% (w/v) low gelling temperature (LGT) agarose (gel strength >200 g/cm$^2$, melting temperature 26°C–30°C, RNase and DNase free; Sigma Canada; product. no. A9414) prepared in 10% dMB at 35°C and then encapsulated into $80 \pm 20$ µm diameter microbeads using a microfluidic chip as described previously (Alkayyali et al., 2021). Our previous research using model marine bacteria indicated that encapsulation using LGT agarose prepared in dMB at 35°C minimized the impact of the encapsulation process on viability (Pope et al., 2022).

## *In situ* incubation

To incubate processed samples in the environment from which the sediment was collected (*in situ* incubation) we utilized Slide-A-Lyzer™ gamma irradiated dialysis cassettes (Thermo Fisher, Canada). A membrane with a 10 kDA cut-off was chosen to retain bacteria but allow nutrients and small molecules to diffuse into the cassette. Prior to using the cassettes, we tested several strategies to aseptically load and unload the cassettes. Cassettes were loaded and unloaded using sterile 1 ml Luer-Lok™ syringes fitted with a sterile 21-gauge hypodermic needle (BD Biosciences, Canada). As the cassettes are sterilized by gamma-irradiation aseptic loading was readily achieved using the standard aseptic technique. After *in situ* incubation, we anticipated that the loading port would be colonized by marine bacteria and that the port would need to be decontaminated to prevent the introduction of external microorganisms into the contents of the cassette during unloading. Two disinfectants were tested for this purpose alone and in combination: 0.5% accelerated hydrogen peroxide (PREempt RTU™, Contec Inc., Canada) and 70% (v/v) isopropanol (Millipore Sigma, Canada). To test decontamination strategies, dialysis cassettes were loaded with sterile seawater and then placed in a 15-gallon aquarium containing a 5 cm bed of sediment and ~60 L of natural seawater and maintained at 22°C with aeration for 24 h. After incubation, the cassettes were unloaded after decontaminating the port and the contents were plated on Marine Agar (MA; BD Difco™, Fisher Scientific, United States). Contamination was assessed after incubation for 7 days at RT (22°C–25°C). Decontamination of the loading port using a sequential three-part process was determined to be the most reliable method and consisted of the following steps: (1) flushed the port with 0.5% (v/v) accelerated hydrogen peroxide followed by a 2 min contact time, (2) flushed the port with 70% (v/v) isopropanol followed by a 10 min contact time (3) flushed the port with sterile water to remove traces of isopropanol before inserting the needle.

To test the integrity of the dialysis cassette during extended incubation in a marine environment, cassettes were loaded with 0.5 ml of sterile water and then placed in the aquarium for 7 days after which the contents were aseptically removed and 100 µl aliquots plated on triplicate MA plates and observed for growth after 7 days at RT. The seal of unmodified dialysis cassettes routinely failed (i.e., microbial growth was observed on MA). Several clamping methods to reinforce the seal of the cassettes were evaluated using the above methodology (6 replicates/trial). The most effective clamping method was to use two pairs of galvanized mending plates (5.00 cm × 1.27 cm; Hillman, Canada) and stainless-steel nuts and bolts and to evenly apply pressure across the surface of the dialysis cassette.

For *in situ* incubation of the resuspended and encapsulated samples, dialysis cassettes were aseptically unpackaged in a biosafety cabinet and soaked in sterile dMB to hydrate the membranes. The cassettes were then loaded with 0.5 ml of either resuspended or encapsulated sediment bacteria using a sterile 1 ml syringe with a 21-gauge needle through one of the four loading ports according to the manufacturer's recommendations. Six dialysis cassettes were loaded for each sediment treatment (i.e., resuspended or encapsulated). It should be noted that the inner diameter of the 21-gauge needle tip was 0.51 mm which was large enough to allow microbead injection without damaging the beads. Each dialysis cassette was then clamped with two pairs of galvanized mending plates and transported back to the collection site (~12 h after initial collection) at RT and gently buried beneath the sediment (~5 cm) in a nylon mesh bag to protect cassettes from damage from marine life. After 7 days the cassettes were gently removed from the sediment and transported back to the laboratory dry and at RT. Samples were processed immediately after collection from the environment. The exterior surface of each dialysis cassette was first washed with sterile dMB to remove excess debris. The removal port was decontaminated as previously

described and a sterile syringe with a 21-gauge needle tip was then used to remove the contents of each dialysis cassette by inserting the needle tip through a port into the center of the dialysis cassette. The contents of each dialysis cassette were placed in a sterile tube for further analysis.

### Cell abundance, viability, and diversity assessment

To determine the effect of *in situ* incubation on both encapsulated and resuspended bacterial samples we assessed bacterial abundance, viability, and diversity before and after *in situ* incubation. To assess abundance, aliquots from encapsulated and resuspended samples, before and after incubation, were serially log-diluted and 100 µl aliquots of each dilution were spread on triplicate MA plates. The plates were incubated at RT for 5 days and the number of colonies observed on each plate was counted to determine colony forming units per mL (CFU/ml) for each sample.

Cell viability was assessed before and after incubation for both encapsulated and resuspended samples *via* a fluorescent Live/Dead bacterial staining kit as described previously (Alkayyali et al., 2021) with the modification that the resuspension buffer was changed from filter-sterilized seawater (FSSW) to dMB. This modification was made to reduce the fluorescent background observed in FSSW caused by marine ultramicrobacteria that are not removed by 0.22 µm filtration (Schut et al., 1993). To assess the effect of *in situ* incubation on the taxonomic composition in the resuspended and encapsulated samples. Assessment of the diversity within each sample was performed *via* random selection of colonies from triplicate plates for each set ($n = 3$). A minimum of 20 colonies were isolated from each sample by serial subculturing and purified isolates were identified by partial 16S rRNA gene sequencing. Template DNA for PCR was prepared by dispersing an isolated colony in 50 µl of dimethyl sulfoxide (DMSO; Millipore Sigma, Canada). The 16S rRNA gene was amplified using the pA (5′-AGAGTTTGATCCTGGCTCAG) and pH (5′-AAGGAGGTGATC CAGCC) primers as described previously (Duncan et al., 2015). Amplicons were directly sequenced at McGill University and Genome Quebec Innovation Centre using the 936R primer (5′-GTGCGGGCCCCCGTCAATT) using an ABI 3730xl DNA Analyzer. Sequences were analyzed in Geneious (Biomatters Inc., v 7.1) and clustered into operational taxonomic units (OTUs) using a 99% sequence identity cutoff value and a representative sequence was used for further analysis (Nichols et al., 2010). The taxonomic affiliation of the OTUs was inferred from a nucleotide search of the 16S rRNA database of GenBank using the Basic Local Alignment Search Tool (BLAST; Zhang et al., 2000). Partial 16S rRNA gene sequences representing OTUs were deposited in GenBank under accession numbers ON385874 to ON385929.

The diversity of each sample was assessed using the Species-richness and Diversity Estimation in R program (SpadeR; Chao and Chiu, 2016). Species richness and diversity were calculated for each sediment treatment. Estimated richness was calculated using the Choa1 estimator (Chao and Chiu, 2016). The Shannon and Simpson diversity indices were calculated within SpadeR and the indices calculated using the maximum likelihood method are reported. The 95% confidence intervals calculated by SpadeR are provided for the Chao1 estimator and diversity indices (Chao, 1984).

### Statistics

All statistical analyses were performed using GraphPad Prism 8 (La Jolla, California, United States). Statistical significance was determined using unpaired, two-tailed *t*-tests and the Holm-Sidak method of correction for *p*-values, with an alpha of 0.05. Each data set was analyzed individually without assuming a consistent standard deviation.

## Results and discussion

### Dialysis cassette modifications

Commercially available Slide-A-Lyzer dialysis cassettes were chosen as vessels for *in situ* incubation experiments conducted in this study because they presented a readymade diffusion chamber that would retain bacteria encapsulated in microbeads while the dialysis membrane would prevent infiltration of external microorganisms into the chamber and allow nutrients to diffuse across the membrane. The 10 kDa dialysis cassettes could be purchased gamma-irradiated providing the sterility required for microbiology studies. Initial testing of the cassettes in a saltwater aquarium over 7 days determined that the seal failed after 2 days, allowing external bacteria into the chamber. This was undesirable as infiltration of external bacteria into the chamber would overwhelm the growth of encapsulated bacteria and confound the study results. Rather than applying adhesives to the cassette, which might present biocompatibility issues, we tested a variety of physical clamping mechanisms to reinforce the cassette seal. The most effective method was the use of small mending plates that were placed across the cassette and held in place with stainless steel bolts and nuts. This method proved to be an effective, simple, and reproducible method of guaranteeing the integrity of the cassette seal during incubation periods of up to 7 days. Another issue encountered when using the dialysis cassettes was microbial fouling of the injection ports during incubation. We tested disinfectants and found that both accelerated hydrogen peroxide (PREempt RTU; 0.5% $H_2O_2$) and 70% IPA were not effective individually, however when we combined the two disinfectants in a sequential process we achieve reproducible sanitization of the injection port, allowing for the aseptic removal of the contents of the dialysis chamber.

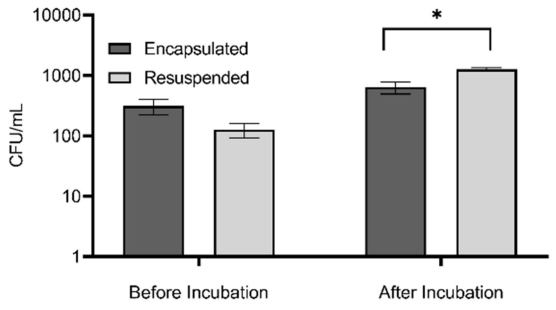



FIGURE 1
Abundance of bacteria in encapsulated and resuspended marine sediment bacteria samples before and after *in situ* incubation based upon colony counts performed in triplicate. Error bars indicate standard error, * denotes significance of $p<0.01$. Is this correct?

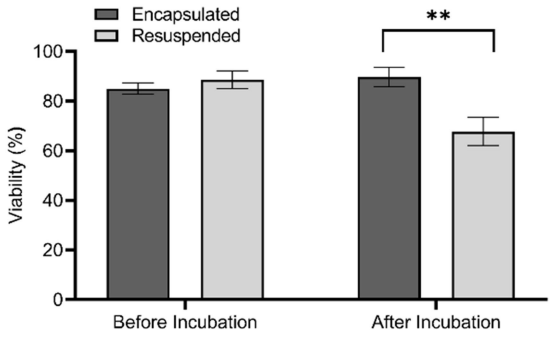



FIGURE 2
Cell viability for encapsulated and resuspended marine sediment bacteria samples before and after *in situ* incubation based upon a Live/Dead assay. Error bars indicate standard error, ** denotes significance of $p\leq0.01$.

## Cell abundance and viability assessment

To evaluate the effect of *in situ* incubation on recoverable bacterial diversity we compared four treatments. First, we prepared microorganisms in two ways. Bacteria dislodged from a marine sediment sample were concentrated by centrifugation and either encapsulated in agarose microbeads or resuspended in dMB. The "encapsulated" and "resuspended" samples were then split, and one portion was plated immediately on agar plates while the other portions were injected into dialysis cassettes and then incubated *in situ* in the environment for 7 days. To assess the effect of these different treatments on bacterial abundance, colony counts were performed. The abundance of culturable bacteria in resuspended and encapsulated bacteria prior to *in situ* incubation was not significantly different ($t(16)=1.93$, $p=0.0728$). This result indicated that the encapsulation process did not adversely affect the viability of cells, confirming our previous results using model marine bacteria (Alkayyali et al., 2021). After *in situ* incubation, the abundance of bacteria obtained from the resuspended samples was significantly greater than that of the encapsulated samples ($t(7)=2.92$, $p=0.0452$; Figure 1). The lower bacterial abundance in the encapsulated sample may be due to a slower growth rate of encapsulated bacteria compared to non-encapsulated bacteria due to the requirement of oxygen and nutrients to diffuse into the agarose bead to be utilized by the bacteria. It may also be because the encapsulated bacteria formed aggregates in the agarose matrix that affected dispersion during plating, resulting in a lower apparent abundance. Future experiments could assess the growth rates of encapsulated bacteria under varying nutrient and oxygen concentrations to determine if diffusion rates limit growth. In addition, hypoxia sensing molecular probes (e.g., Image-iT™ Green Hypoxia Reagent, Fisher Scientific) could be used to assess oxygen limitations in beads. Thorough homogenization of beads prior to plating could also provide a more accurate assessment of abundance of colonies that develop in agar microbeads.

To assess the effect of the different treatment methods on bacterial viability we used a fluorescent Live/Dead assay to microscopically determine the percentage of viable cells present in each sample. This assay stains viable cells green while dead cells appear red. Prior to *in situ* incubation, the proportion of viable cells present in encapsulated and resuspended samples was not significantly different ($t(24)=0.87$, $p=0.399$). Conversely, there was a significantly greater proportion of viable of cells within encapsulated samples compared to resuspended samples after *in situ* incubation ($t(25)=3.23$, $p=0.0035$; Figure 2). The increased viability of encapsulated samples following incubation indicates that although the abundance determined by colony counts on agar plates demonstrates a lower abundance compared to non-encapsulated samples, there are in fact a greater proportion of viable cells within these samples.

The images of encapsulated bacteria before and after incubation provide evidence of the growth of microencapsulated bacteria (Figure 3). Prior to incubation, single colonies or a few individual colonies were observed within the microbeads while after incubation, microbeads contained increased growth of colonies. The bacteria appeared to form microcolonies within the agar beads, which might explain why the bacterial abundance in plate counts of *in situ* incubated encapsulated samples appeared to be lower. Interestingly, the percent viability of encapsulated cells before and after incubation was not significantly different, while the percent viability of resuspended samples was significantly lower after incubation. The reason for this is unknown, but it is possible that the non-encapsulated samples were able to proliferate more quickly than encapsulated bacteria, resulting in a greater number of cells that had entered the death phase. This hypothesis is supported by the data in Figure 1, which shows that the total number of cultivable cells in the resuspended sample following *in situ* incubation was greater than that of the encapsulated sample. Alternatively, it might suggest that agarose encapsulation may protect encapsulated cells from toxic chemicals present in the environment, such as tannins derived from organic matter present in the sediment (Kaczmarek, 2020), or from antimicrobial

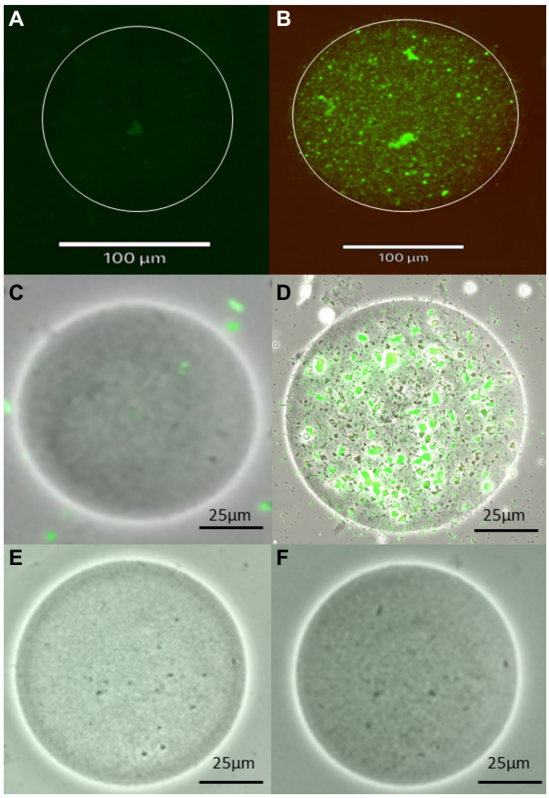



FIGURE 3
Images of agarose microbeads, including a microbead containing a single marine sediment bacterial cell before incubation **(A,C)** and the growth of numerous cells within the agarose microbead following incubation **(B,D)** using fluorescent and brightfield imaging, respectively as well as images of blank beads without any bacteria **(E,F)**. Images were obtained using Live/Dead bacterial stain and Revolve4 microscope with a 20× objective lens.

metabolites produced by microorganisms in the sediment or even within the diffusion chamber (Thomashow et al., 1990).

## Diversity assessment

A primary motivation for this study was to develop a new cultivation method to enable the cultivation of a greater biodiversity of the microbiome present in environmental samples such as marine sediments and soils, as unique bacteria from these habitats have previously been shown to be an excellent source of metabolites with biomedical potential (Gulder and Moore, 2010; Ling et al., 2015). To assess the diversity of bacteria cultured from each of the four treatments, randomly selected colonies from each treatment were purified and 148 isolates (32–38 isolates from each treatment; Supplementary Table 1) were identified. A total of 58 OTUs were obtained using a 99% sequence identity threshold and 8–27 OTUs were obtained from each treatment (Table 1). The number of OTUs was similar between the encapsulated and resuspended samples both before and after incubation, although in both cases fewer OTUs were detected in the resuspended samples (Table 1). The observed richness also decreased following *in situ* incubation in both the encapsulated and resuspended treatments. Similar trends were observed in the Choa1 richness estimator and Shannon and Simpson diversity indices (Table 1), indicating that both species richness and diversity decreased after *in situ* incubation.

In light of this result, we were interested to assess the distribution of OTUs obtained from the four treatments. Pairwise comparisons of OTUs shared between the four samples revealed that there was a low degree of overlap (Figure 4). No OTU was shared between more than two treatments and, as might be expected, the greatest overlap was between the encapsulated and resuspended samples prior to *in situ* incubation (6 OTUs). Only three OTUs were in common between the two encapsulated samples and no OTUs were found in common between the two resuspended samples. This low degree of shared OTUs between samples both before and after *in situ* incubation is somewhat surprising and may be partially due to the limited number of isolates sampled from each treatment. This result also suggests that the different treatments should be selected for the growth of a different subset of the microbial community.

Taxonomic classification of the isolates revealed that *Proteobacteria*, *Bacteroidetes,* and *Firmicutes* were identified in all treatments with the majority of isolates (56.3%–75.7%) recovered from each treatment belonging to the *Proteobacteria* (Figure 5). The dominance of *Proteobacteria* is consistent with previous studies examining microbial diversity of marine sediments and seawater (Hoshino et al., 2020). *Actinobacteria* and *Verrucomicrobia* were only obtained from samples that had been encapsulated and incubated *in situ* (Figure 5). The inability to culture these taxa from samples that were plated immediately may have been because they were so scarce in the environmental samples that they were not represented in the sub-samples that were plated, or because they were incapable of growing on agar media. The process of encapsulation and *in situ* incubation may have enabled the recovery of these taxa by fostering their growth to levels that were detectable when plated after *in situ* incubation. The process may also have acclimatized the bacteria to growing on an agar matrix, allowing them to form colonies when they were plated after *in situ* incubation. In other words, the bacteria were effectively domesticated for cultivation in the laboratory. Interestingly, Nichols and colleagues (Nichols et al., 2010) also only detected *Verrucomicrobia* from in samples that had been embedded in agar and incubated *in situ* in their description of the Ichip. The OTU (OTU24, Supplementary Table 1) belonging to the *Verrucomicrobia* was comprised of two isolates whose 16S rRNA gene sequence exhibited only 95.85% identity to *Luteolibacter algae* (NR_041624), suggesting this OTU represents a novel species of *Luteolibacter. Luteolibacter* has been isolated from a variety of habitats including marine habitats (Zhang and Demain, 2005; Yoon et al., 2008; Dahal et al., 2021), and recently, *Luteolibacter gellanilyticus* was formally described after being isolated from forest soil using an Ichip diffusion

10.3389/fmicb.2022.958660

TABLE 1 Summary of the diversity obtained from two sediment treatments before and after *in-situ* incubation.

| | Before incubation | | After incubation | |
|---|---|---|---|---|
| | Encapsulated | Resuspended | Encapsulated | Resuspended |
| Sample size (*N*) | 38 | 32 | 38 | 38 |
| Richness (OTUs) | 27 | 23 | 14 | 8 |
| Est.Richness (Choa1) | 98.6 | 75.3 | 29.6 | 12.4 |
| [95% CI] | [46.8–285.9] | [37.0–219.0] | [16.9–96.6] | [8.5–47.9] |
| Shannon Diversity | 23.3 | 19.7 | 8.7 | 5.3 |
| [95% CI] | [16.4–30.3] | [13.6–25.8] | [5.2–12.3] | [3.7–6.9] |
| Simpson Diversity | 19.5 | 16.0 | 5.8 | 4.3 |
| [95% CI] | [11.9–27.2] | [9.3–22.7] | [3.4–8.3] | [3.0–5.6] |
| Identity >98.7% (# isolates) | 23 | 18 | 30 | 37 |
| Identity >97–98.7 (# isolates) | 8 | 13 | 4 | 0 |
| Identity >95–<97% (# isolates) | 3 | 1 | 4 | 1 |
| Identity 93–<95% (# isolates) | 4 | 0 | 0 | 0 |

OTUs were defined as sharing >99% sequence identity. Estimated richness (Choa1 index)(38) and Shannon and Simpson diversity indices were calculated using SpadeR.(37) The number of OTUs from each sample exhibiting a range of sequence identities to 16S rRNA gene sequences of type strains contained in the GenBank 16S rRNA database is provided in the bottom four rows.

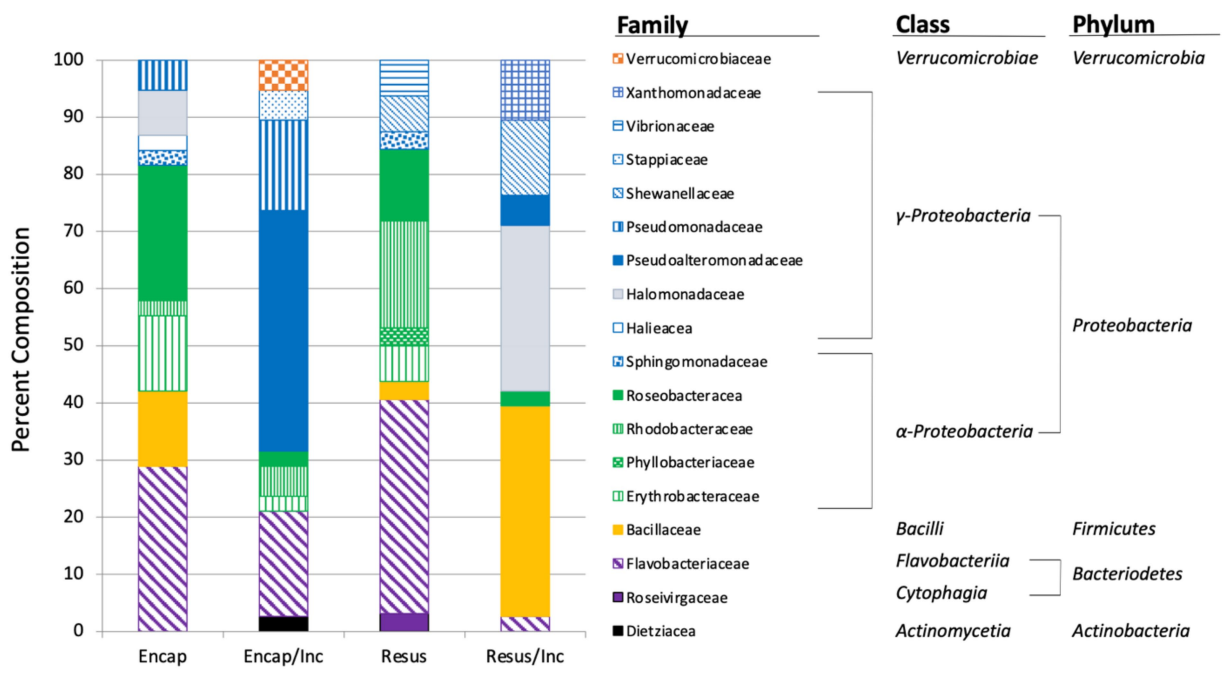



FIGURE 4
Taxonomic composition of bacteria isolated from marine sediment using four treatments: Encap: encapsulated, Encap/Inc. – encapsulated and incubated *in situ*, Resus – resuspended, Resus/Inc. – resuspended and incubated *in situ*. The bar graph shows family-level taxonomic classification.

chamber (Pascual et al., 2017). These results indicate that our methodology provides similar access to difficult-to-culture microbiota that has previously been accessed using the Ichip diffusion chamber design.

At the family level of taxonomic resolution, the differences in microbiota recovered from the four treatments become more apparent (Figure 5). For example, *Flavobacteriaceae* represented 18.5%–37.5% of the communities of three treatments, but only

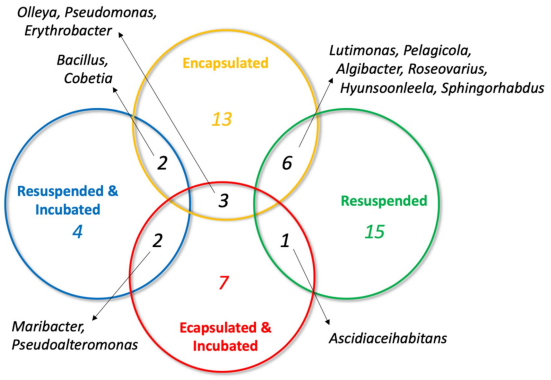



FIGURE 5
Venn diagram illustrating the distribution of shared OTUs between the four sediment treatments. Genera represented by the OTUs are indicated.

2.7% of the community recovered from the resuspended/incubated treatment. Conversely, *Bacillaceae* were a major component of the resuspended/incubated treatment (36.7%) but accounted for a much smaller proportion (0%–13.2%) in the other treatments. One of the most striking differences was the large percentage of isolates belonging to the *Pseudoalteromonadaceae* obtained from the encapsulated/incubated (42.1%) and resuspended/incubated (5.3%) treatments, but completely absent in the other treatments. The *Pseudoalteromonadaceae* component of this community was composed of two OTUs (OTU2 and OTU13) that exhibited close 16S rRNA gene sequence similarity to *Pseudoalteromonas tetraodonis* (NR_114187, 100% identity) and *Pseudoalteromonas shioyasakiensis* (NR_125458, 98.8% identity). Both strains are marine bacteria that require salt for growth (Ivanova et al., 2001; Matsuyama et al., 2014). Although the type strains of these species were not isolated using diffusion chambers, it is tempting to hypothesize that their recovery in the *in situ* incubated treatments was facilitated by the combination of encapsulation and *in situ* incubation in a marine environment. Comparing samples before and after *in situ* incubation, there were also several instances where taxa present prior to incubation were not recovered after incubation (e.g., *Bacillaceae* in the encapsulated samples and *Roseivirgaceae* in the resuspended samples). This may have been due to members of these taxa being reduced in relative abundance due to the more rapid replication of other taxa whose growth was favored by *in situ* incubation. Examining a larger sample size from each treatment may have reduced the incidence of this occurrence by increasing the probably of recovering taxa whose relative abundance decreased during *in situ* incubation.

To assess the ability of the four treatments to recover novel taxa from marine sediment, we compared the number of isolates recovered from each treatment that exhibited different levels of 16S rRNA gene sequence identity to sequences of type strains in GeneBank. Thresholds of 98.7 and 95% identity were used as species and genus-level cut-offs, respectively (Stackebrandt, 2006); however, it should be noted that these thresholds are generalized cutoffs and do not apply to all bacterial taxa (Rossi-Tamisier et al., 2015). To provide further discrimination between the taxonomic novelty of strains obtained from the different treatments we also used two arbitrary sequence identity categories (>97%–98.65%, and >95%–<97%; Table 1). The recovered isolates from the resuspended/incubated sample presented the least taxonomic novelty as 97% of isolates exhibited >98.7% sequence identity with previously described bacterial species. A larger number of taxonomically unique isolates was obtained from the resuspended sample that was directly plated, with 56% of the isolates showing >98.7% identity, 41% showing 97%–98.7% identity, and 3% showing 95%–97% identity to previously described bacteria. In contrast, a much larger proportion of strains isolated from the encapsulated samples exhibited sequence identities <97% to described bacteria (encapsulated – 18%, encapsulated/incubated 11%). Interestingly, fewer potentially novel species were recovered from the encapsulated samples after *in situ* incubation as only 22% of isolates exhibited <98.65% identity to described species compared to 49% from the encapsulated and plated sample. The encapsulated sample prior to *in situ* incubation was also the only treatment to recover novel bacterial genera exhibiting <95% sequence identity to described bacteria.

Overall, the results of this study revealed trends where encapsulated samples exhibited greater richness, diversity, and taxonomic novelty than resuspended samples, while *in situ* incubation reduced the richness, diversity, and taxonomic novelty. Microencapsulation alone may increase the microbial diversity that can be recovered from marine sediments by providing nutrients, a scaffold for growth, and protection from external hazards. Agarose utilization is broadly distributed in marine bacteria (Hehemann et al., 2014; Lombard et al., 2014) and the agarose was supplemented with dMB, thus it is likely that the provision of nutrients contributed to the improved recovery observed from the encapsulated treatments. A similar observation was made by Kaeberlein and colleagues in their seminal report on the use of diffusion chambers (Kaeberlein et al., 2002). Agarose encapsulation has also been used to maintain the viability of bacteria exposed to harsh environmental conditions (Alehosseini et al., 2019), thus the agarose may have protected the viability of sensitive encapsulated bacteria until they reached a cell density where colonies could be established on agar plates. Another explanation for positive effect of encapsulation on recoverable microbial diversity could be due to the exposure of cells to a mild heat shock during the encapsulation process, which could resuscitate viable but non-culturable cells (Li et al., 2014).

Unexpectedly, *in situ* incubation had a negative effect on the diversity and taxonomic novelty of microbes recovered from the sediment sample. This observation is contrary to results obtained in previous studies utilizing combinations of agarose encapsulation and a variety of enrichment or *in situ* incubation techniques (Ji et al., 2012; Ma et al., 2014). Explanations for this difference may be the short duration (7 days) of *in situ* incubation and the use only a single round of *in situ* incubation. *In situ* incubation periods

for diffusion-chamber-based studies have ranged from 2 to 4 weeks and the contents of the chambers have been subjected to multiple iterations of *in situ* incubation (Bollmann et al., 2007; Nichols et al., 2010; Steinert et al., 2014). The convenient format of the dialysis cassettes makes aseptic sample recovery very straightforward, thus future studies could readily assess the effect of multiple *in situ* incubation periods. Our study also utilized a different membrane pore size than previous studies, which typically used membranes with 30–200 nm pores (Bollmann et al., 2007; Nichols et al., 2010; Steinert et al., 2014). The 10 kDa membranes used in this study correspond to a pore size of 3 nm, which is an order of magnitude smaller than the membranes used previously. It is possible that this small pore size restricted the diffusion of macromolecular nutrients (e.g., proteins and polysaccharides) into the diffusion chamber and limited the growth of the bacteria in the chambers.

The objective of many biotechnology-focused bioprospecting efforts is to culture as broad a selection of bacterial taxa as possible so that they can be interrogated for a variety of traits of biotechnological interest. These efforts are understandably focused on culturing novel taxa, which are more likely to possess novel traits than commonly cultured bacteria. Diffusion-chamber-based approaches have already delivered these novel discoveries (Ling et al., 2015; MacIntyre et al., 2019) and will continue to do so into the future as their use becomes more common. In this study, the taxonomic diversity of bacteria recovered from a marine sediment sample was highly dependent on the treatment that the sediment was subjected to. While the small sample size of the study limits the conclusions that can be drawn, the results showed that minimal taxonomic overlap was observed in the bacteria recovered in a random selection from the different treatments (Figure 4). This trend suggests that the use of multiple isolation strategies is key to maximizing the recovery of the viable but uncultured majority of environmental bacteria. The methodology we report here provides an easily implementable approach to improve the diversity of bacteria recovered from environmental samples using microencapsulation and *in situ* incubation of dialysis cassette diffusion chambers.

## Conclusion

Based upon assessments of the abundance, viability, and diversity, microencapsulation of marine sediment bacteria led to the recovery of greater taxonomic diversity of bacteria than the traditional method of simply plating bacteria directly from the environmental sample. As a large proportion of environmental bacteria live attached to surfaces in their natural environment, the agar microbeads likely benefit the growth of difficult-to-culture bacteria providing a surface for growth, nutrients, and protection from stressors (Alain and Querellou, 2009; MacIntyre et al., 2019). Combining microencapsulation with *in situ* incubation further enhanced the recoverable microbial diversity by exposing the bacteria to natural environmental conditions including chemical signaling that is not well understood. The *in situ* method described here employs commercially available dialysis cassettes that can be easily obtained from a myriad of scientific suppliers, making the method readily available to researchers. Further improvements in the *in situ* incubation process can be realized by exploring different membrane pore sizes. The use of microencapsulation technology also makes high throughput downstream processing techniques such as flow cytometry or fluorescent activated cell sorting (FACS) to facilitate the rapid isolation of pure strains after *in situ* incubation possible. Importantly, the use of the new bacterial cultivation method described here has resulted in the isolation of several putatively new species of bacteria and the further use and improvement of this method will provide access to many novel bacteria that can be explored for their biotechnological potential.

## 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

EP, CC, BH, RK, and AA: conceptualization and writing–review and editing. EP, CC, and BH: data curation. EP and BH: formal analysis. RK and AA: funding acquisition, project administration, and resources. EP: investigation, validation, visualization, and writing–original draft. EP and CC: methodology. BH, RK, and AA: supervision. All authors contributed to the article and approved the submitted version.

## Funding

This work was supported by the NSERC Discovery Grant (RGPIN-2017-05272), Canada Research Chair program, New Frontiers in Research Fund (NFRFE-2018-01410), and Canada Foundation for Innovation Grant (Project #37696).

## 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.2022.958660/full#supplementary-material

## References


Alain, K., and Querellou, J. (2009). Cultivating the uncultured: limits, advances and future challenges. *Extremophiles* 13, 583–594. doi: 10.1007/s00792-009-0261-3

Alehosseini, A., Gomez del Pulgar, E.-M., Fabra, M. J., Gómez-Mascaraque, L. G., Benítez-Páez, A., Sarabi-Jamab, M., et al. (2019). Agarose-based freeze-dried capsules prepared by the oil-induced biphasic hydrogel particle formation approach for the protection of sensitive probiotic bacteria. *Food Hydrocoll.* 87, 487–496. doi: 10.1016/j.foodhyd.2018.08.032

Alkayyali, T., Pope, E., Wheatley, S. K., Cartmell, C., Haltli, B., Kerr, R. G., et al. (2021). Development of a microbe domestication pod (MD pod) for in situ cultivation of micro-encapsulated marine bacteria. *Biotechnol. Bioeng.* 118, 1166–1176. doi: 10.1002/bit.27633

Aoi, Y., Kinoshita, T., Hata, T., Ohta, H., Obokata, H., and Tsuneda, S. (2009). Hollow-fiber membrane chamber as a device for *in situ* environmental cultivation. *Appl. Environ. Microbiol.* 75, 3826–3833. doi: 10.1128/AEM.02542-08

Bbosa, G. S., Mwebaza, N., Odda, J., Kyegombe, D. B., and Ntale, M. (2014). Antibiotics/antibacterial drug use, their marketing and promotion during the post-antibiotic golden age and their role in emergence of bacterial resistance. *Health* 6, 410–425. doi: 10.4236/health.2014.65059

Bérdy, J. (2012). Thoughts and facts about antibiotics: where we are now and where we are heading. *J. Antibiot.* 65, 385–395. doi: 10.1038/ja.2012.27

Berdy, B., Spoering, A. L., Ling, L. L., and Epstein, S. S. (2017). *In situ* cultivation of previously uncultivable microorganisms using the ichip. *Nat. Protoc.* 12, 2232–2242. doi: 10.1038/nprot.2017.074

Blunt, J. W., Carroll, A. R., Copp, B. R., Davis, R. A., Keyzers, R. A., and Prinsep, M. R. (2018). Marine natural products. *Nat. Prod. Rep.* 35, 8–53. doi: 10.1039/C7NP00052A

Bollmann, A., Lewis, K., and Epstein, S. S. (2007). Incubation of environmental samples in a diffusion chamber increases the diversity of recovered isolates. *Appl. Environ. Microbiol.* 73, 6386–6390. doi: 10.1128/AEM.01309-07

Bruns, A., Cypionka, H., and Overmann, J. (2002). Cyclic AMP and acyl Homoserine lactones increase the cultivation efficiency of heterotrophic bacteria from the Central Baltic Sea. *Appl. Environ. Microbiol.* 68, 3978–3987. doi: 10.1128/AEM.68.8.3978-3987.2002

Butler, M. S. (2004). The role of natural product chemistry in drug discovery. *J. Nat. Prod.* 67, 2141–2153. doi: 10.1021/np040106y

Chao, A. (1984). Nonparametric estimation of the number of classes in a population. *Scand. J. Stat.* 11, 265–270.

Chao, A., and Chiu, C. (2016). "Nonparametric estimation and comparison of species richness," in *eLS*. eds. N. Balakrishnan, C. Theodore, E. Brian, P. Walter, R. Fabrizio and T. Jef (Hoboken, NJ, USA: John Wiley & Sons, Ltd (Wiley)), 1–11.

Cragg, G. M., and Newman, D. J. (2005). Biodiversity: a continuing source of novel drug leads. *Pure Appl. Chem.* 77, 7–24. doi: 10.1351/pac200577010007

Dahal, R. H., Chaudhary, D. K., Kim, D.-U., and Kim, J. (2021). *Luteolibacter luteus* sp. nov., isolated from stream bank soil. *Arch. Microbiol.* 203, 377–382. doi: 10.1007/s00203-020-02048-x

Dias, D. A., Urban, S., and Roessner, U. (2012). A historical overview of natural products in drug discovery. *Meta* 2, 303–336. doi: 10.3390/metabo2020303

Duncan, K. R., Haltli, B., Gill, K. A., Correa, H., Berrué, F., and Kerr, R. G. (2015). Exploring the diversity and metabolic potential of actinomycetes from temperate marine sediments from Newfoundland, Canada. *J. Ind. Microbiol. Biotechnol.* 42, 57–72. doi: 10.1007/s10295-014-1529-x

Gerwick, W. H., and Moore, B. S. (2012). Lessons from the past and charting the future of marine natural products drug discovery and chemical biology. *Chem. Biol.* 19, 85–98. doi: 10.1016/j.chembiol.2011.12.014

Glatstein, M., Danino, D., Wolyniez, I., and Scolnik, D. (2014). Seizures caused by ingestion of Atropa belladonna in a homeopathic medicine in a previously well infant: case report and review of the literature. *Am. J. Ther.* 21, e196–e198. doi: 10.1097/MJT.0b013e3182785eb7

Gulder, T. A. M., and Moore, B. S. (2010). Salinosporamide natural products: potent 20S proteasome inhibitors as promising cancer chemotherapeutics. *Angew. Chem. Int. Ed. Engl.* 49, 9346–9367. doi: 10.1002/anie.201000728

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

Hoshino, T., Doi, H., Uramoto, G.-I., Wörmer, L., Adhikari, R. R., Xiao, N., et al. (2020). Global diversity of microbial communities in marine sediment. *Proc. Natl. Acad. Sci. U. S. A.* 117, 27587–27597. doi: 10.1073/pnas.1919139117

Ivanova, E. P., Romanenko, L. A., Matté, M. H., Matté, G. R., Lysenko, A. M., Simidu, U., et al. (2001). Retrieval of the species Alteromonas tetraodonis Simidu et al. 1990 as Pseudoalteromonas tetraodonis comb. nov. and emendation of description. *Int. J. Syst. Evol. Microbiol.* 51, 1071–1078. doi: 10.1099/00207713-51-3-1071

Jacoby, G. A. (2017). "History of drug-resistant microbes," in *Antimicrobial Drug Resistance: Mechanisms of Drug Resistance*. eds. D. L. Mayers, J. D. Sobel, M. Ouellette, K. S. Kaye and D. Marchaim, *Vol. 1* (Cham: Springer International Publishing), 3–8.

Ji, S., Zhao, R., Yin, Q., Zhao, Y., Liu, C., Xiao, T., et al. (2012). Gel microbead cultivation with a subenrichment procedure can yield better bacterial cultivability from a seawater sample than standard plating method. *J. Ocean Univ. China* 11, 45–51. doi: 10.1007/s11802-012-1869-y

Kaczmarek, B. (2020). Tannic acid with antiviral and antibacterial activity as A promising component of biomaterials—A Minireview. *Materials (Basel)* 13, 3224. doi: 10.3390/ma13143224

Kaeberlein, T., Lewis, K., and Epstein, S. S. (2002). Isolating "uncultivable" microorganisms in pure culture in a simulated natural environment. *Science* 296, 1127–1129. doi: 10.1126/science.1070633

Kato, S., Yamagishi, A., Daimon, S., Kawasaki, K., Tamaki, H., Kitagawa, W., et al. (2018). Isolation of previously uncultured slow-growing bacteria by using a simple modification in the preparation of agar media. *Appl. Environ. Microbiol.* 84, e00807–e00818. doi: 10.1128/AEM.00807-18

Katz, L., and Baltz, R. H. (2016). Natural product discovery: past, present, and future. *J. Ind. Microbiol. Biotechnol.* 43, 155–176. doi: 10.1007/s10295-015-1723-5

Kawasaki, K., and Kamagata, Y. (2017). Phosphate-catalyzed hydrogen peroxide formation from agar, Gellan, and κ-carrageenan and recovery of microbial cultivability via catalase and pyruvate. *Appl. Environ. Microbiol.* 83, e01366–e01317. doi: 10.1128/AEM.01366-17

Kinghorn, A. D., Chin, Y.-W., and Swanson, S. M. (2009). Discovery of natural product anticancer agents from biodiverse organisms. *Curr. Opin. Drug Discov. Devel.* 12, 189–196. PMID: 19333864

Kong, D.-X., Jiang, Y.-Y., and Zhang, H.-Y. (2010). Marine natural products as sources of novel scaffolds: achievement and concern. *Drug Discov. Today* 15, 884–886. doi: 10.1016/j.drudis.2010.09.002

Krause, J., and Tobin, G. (2013). *Discovery, Development, and Regulation of Natural Products*. IntechOpen.

Kumar, D., Savitri, Thakur, N., Verma, R., and Bhalla, T. C. (2008). Microbial proteases and application as laundry detergent additive. *Res. J. Microbio.* 3, 661–672. doi: 10.3923/jm.2008.661.672

Lewis, W. H., and Elvin-Lewis, M. P. F. (2003). *Medical Botany: Plants Affecting Human Health. 2nd Edn.* New York: John Wiley and Sons. https://www.wiley.com/en-us/Medical+Botany%3A+Plants+Affecting+Human+Health%2C+2nd+Edition-p-9780471628828

Li, L., Mendis, N., Trigui, H., Oliver, J. D., and Faucher, S. P. (2014). The importance of the viable but non-culturable state in human bacterial pathogens. *Front. Microbiol.* 5:258. doi: 10.3389/fmicb.2014.00258

Ling, L. L., Schneider, T., Peoples, A. J., Spoering, A. L., Engels, I., Conlon, B. P., et al. (2015). A new antibiotic kills pathogens without detectable resistance. *Nature* 517, 455–459. doi: 10.1038/nature14098

Lok, C. (2015). Mining the microbial dark matter. *Nature News* 522, 270–273. doi: 10.1038/522270a

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

Ma, L., Kim, J., Hatzenpichler, R., Karymov, M. A., Hubert, N., Hanan, I. M., et al. (2014). Gene-targeted microfluidic cultivation validated by isolation of a gut bacterium listed in human microbiome Project's Most wanted taxa. *Proc. Natl. Acad. Sci. U. S. A.* 111, 9768–9773. doi: 10.1073/pnas.1404753111

MacIntyre, L. W., Charles, M. J., Haltli, B. A., Marchbank, D. H., and Kerr, R. G. (2019). An Ichip-domesticated sponge bacterium produces an N-Acyltyrosine bearing an α-methyl substituent. *Org. Lett.* 21, 7768–7771. doi: 10.1021/acs.orglett.9b02710

Matsuyama, H., Sawazaki, K., Minami, H., Kasahara, H., Horikawa, K., and Yumoto, I. (2014). *Pseudoalteromonas shioyasakiensis* sp. nov., a marine polysaccharide-producing bacterium. *Int. J. Syst. Evol. Microbiol.* 64, 101–106. doi: 10.1099/ijs.0.055558-0

Montinari, M. R., Minelli, S., and De Caterina, R. (2019). The first 3500 years of aspirin history from its roots – a concise summary. *Vasc. Pharmacol.* 113, 1–8. doi: 10.1016/j.vph.2018.10.008

Nai, C., and Meyer, V. (2018). From axenic to mixed cultures: technological advances accelerating a paradigm shift in microbiology. *Trends Microbiol.* 26, 538–554. doi: 10.1016/j.tim.2017.11.004

Newman, D. J., and Cragg, G. M. (2016). Natural products as sources of new drugs from 1981 to 2014. *J. Nat. Prod.* 79, 629–661. doi: 10.1021/acs.jnatprod.5b01055

Nichols, D., Cahoon, N., Trakhtenberg, E. M., Pham, L., Mehta, A., Belanger, A., et al. (2010). Use of Ichip for high-throughput *in situ* cultivation of "uncultivable" microbial species. *Appl. Environ. Microbiol.* 76, 2445–2450. doi: 10.1128/AEM.01754-09

Omsland, A., Cockrell, D. C., Howe, D., Fischer, E. R., Virtaneva, K., Sturdevant, D. E., et al. (2009). Host cell-free growth of the Q fever bacterium *Coxiella burnetii*. *Proc. Natl. Acad. Sci. U. S. A.* 106, 4430–4434. doi: 10.1073/pnas.0812074106

Pascual, J., García-López, M., González, I., and Genilloud, O. (2017). *Luteolibacter gellanilyticus* sp. nov., a gellan-gum-degrading bacterium of the phylum Verrucomicrobia isolated from miniaturized diffusion chambers. *Int. J. Syst. Evol. Microbiol.* 67, 3951–3959. doi: 10.1099/ijsem.0.002227

Perbal, B. (2003). Communication is the key. *Cell Commun. Signaling* 1:3. doi: 10.1186/1478-811X-1-3

Pham, V. H. T., and Kim, J. (2012). Cultivation of unculturable soil bacteria. *Trends Biotechnol.* 30, 475–484. doi: 10.1016/j.tibtech.2012.05.007

Pope, E., Haltli, B., Kerr, R. G., and Ahmadi, A. (2022). Effects of matrix composition and temperature on viability and metabolic activity of microencapsulated marine bacteria. *Microorganisms* 10:996. doi: 10.3390/microorganisms10050996

Roberts, J. D., and Caserio, M. C. (1977). *Basic Principles of Organic Chemistry*, *2nd Edn*. Menlo Park, CA: W. A. Benjamin, Inc.

Rossi-Tamisier, M., Benamar, S., Raoult, D., and Fournier, P.-E. (2015). Cautionary tale of using 16S rRNA gene sequence similarity values in identification of human-associated bacterial species. *Int. J. Syst. Evol. Microbiol.* 65, 1929–1934. doi: 10.1099/ijs.0.000161

Sait, M., Davis, K. E. R., and Janssen, P. H. (2006). Effect of pH on isolation and distribution of members of subdivision 1 of the phylum Acidobacteria occurring in soil. *Appl. Environ. Microbiol.* 72, 1852–1857. doi: 10.1128/AEM.72.3.1852-1857.2006

Schut, F., de Vries, E. J., Gottschal, J. C., Robertson, B. R., Harder, W., Prins, R. A., et al. (1993). Isolation of typical marine bacteria by dilution culture: growth, maintenance, and characteristics of isolates under laboratory conditions. *Appl. Environ. Microbiol.* 59, 2150–2160. doi: 10.1128/aem.59.7.2150-2160.1993

Stackebrandt, E. (2006). Taxonomic parameters revisited: tarnished gold standards. *Microbiol. Today* 33, 152–155.

Steinert, G., Whitfield, S., Taylor, M. W., Thoms, C., and Schupp, P. J. (2014). Application of diffusion growth chambers for the cultivation of marine sponge-associated bacteria. *Mar. Biotechnol.* 16, 594–603. doi: 10.1007/s10126-014-9575-y

Thomashow, L. S., Weller, D. M., Bonsall, R. F., and Pierson, L. S. (1990). Production of the antibiotic phenazine-1-carboxylic acid by fluorescent pseudomonas species in the rhizosphere of wheat. *Appl. Environ. Microbiol.* 56, 908–912. doi: 10.1128/aem.56.4.908-912.1990

Vishnivetskaya, T., Kathariou, S., McGrath, J., Gilichinsky, D., and Tiedje, J. M. (2000). Low-temperature recovery strategies for the isolation of bacteria from ancient permafrost sediments. *Extremophiles* 4, 165–173. doi: 10.1007/s007920070031

Yazdankhah, S., Lassen, J., Midtvedt, T., and Solberg, C. O. (2013). The history of antibiotics. *Tidsskr. Nor. Laegeforen.* 133, 2502–2507. doi: 10.4045/tidsskr.13.0145

Yoon, J., Matsuo, Y., Adachi, K., Nozawa, M., Matsuda, S., Kasai, H., et al. (2008). Description of *Persicirhabdus sediminis* gen. nov., sp. nov., *Roseibacillus ishigakijimensis* gen. nov., sp. nov., *Roseibacillus ponti* sp. nov., *Roseibacillus persicicus* sp. nov., *Luteolibacter pohnpeiensis* gen. nov., sp. nov. and *Luteolibacter algae* sp. nov., six marine members of the phylum "Verrucomicrobia", and emended descriptions of the class Verrucomicrobiae the order Verrucomicrobiales. *Int. J. Syst. Evol. Microbiol.* 58, 998–1007. doi: 10.1099/ijs.0.65520-0

Zhang, L., and Demain, A. (2005). *Natural Products: Drug Discovery and Therapeutic Medicine*. Totowa, NJ, USA: Humana Press Inc.

Zhang, Z., Schwartz, S., Wagner, L., and Miller, W. (2000). A greedy algorithm for aligning DNA sequences. *J. Comput. Biol.* 7, 203–214. doi: 10.1089/10665270050081478

frontiers | Frontiers in Marine Science

TYPE Original Research
PUBLISHED 30 August 2022
DOI 10.3389/fmars.2022.963277



OPEN ACCESS

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

REVIEWED BY
Dagmar Hajkova Leary,
United States Naval Research Laboratory, United States
Cristiane Cassiolato Pires Hardoim,
São Paulo State University, Brazil

*CORRESPONDENCE
Yoshiteru Aoi
yoshiteruaoi@hiroshima-u.ac.jp

†PRESENT ADDRESSES
Dawoon Jung,
Li Dak Sum Yip Yio Chin Kenneth Li Marine Biopharmaceutical Research Center, College of Food and Pharmaceutical Sciences, Ningbo University, Ningbo, China;
Ningbo Institute of Marine Medicine, Peking University, Ningbo, China

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

RECEIVED 07 June 2022
ACCEPTED 12 August 2022
PUBLISHED 30 August 2022

CITATION
Jung D, Machida K, Nakao Y, Owen JS, He S, Kindaichi T, Ohashi A and Aoi Y (2022) Cultivation of previously uncultured sponge-associated bacteria using advanced cultivation techniques: A perspective on possible key mechanisms.
*Front. Mar. Sci.* 9:963277.
doi: 10.3389/fmars.2022.963277



# Cultivation of previously uncultured sponge-associated bacteria using advanced cultivation techniques: A perspective on possible key mechanisms

Dawoon Jung[1†], Koshi Machida[2], Yoichi Nakao[2,3], Jeffrey S. Owen[4], Shan He[5,6], Tomonori Kindaichi[7,8], Akiyoshi Ohashi[7,8] and Yoshiteru Aoi[1,9*]

[1]Department of Molecular Biotechnology, Graduate School of Advanced Sciences of Matter, Hiroshima University, Higashihiroshima, Japan, [2]Research Institute for Science and Engineering, Waseda University, Tokyo, Japan, [3]Department of Chemistry and Biochemistry, Graduate School of Advanced Science and Engineering, Waseda University, Tokyo, Japan, [4]Department of Environmental Science, Hankuk University of Foreign Studies, Yongin, South Korea, [5]Li Dak Sum Yip Yio Chin Kenneth Li Marine Biopharmaceutical Research Center, College of Food and Pharmaceutical Sciences, Ningbo University, Ningbo, China, [6]Ningbo Institute of Marine Medicine, Peking University, Ningbo, China, [7]Department of Civil and Environmental Engineering, Graduate School of Engineering, Hiroshima University, Higashihiroshima, Japan, [8]Graduate School of Advanced Science and Engineering, Hiroshima University, Higashihiroshima, Japan, [9]Unit of Biotechnology, Graduate School of Integrated Sciences for Life, Hiroshima University, Higashihiroshima, Japan

Most of the microbes from natural habitats cannot be cultivated with standard cultivation in laboratory, and sponge-associated microbes are no exception. We used two advanced methods based on a continuous-flow bioreactor (CF) and *in situ* cultivation (I-tip) to isolate previously uncultivated marine sponge-associated bacteria. We also characterized the physiological properties of the isolates from each method and attempted to clarify the mechanisms operating in each cultivation method. A greater number of novel bacteria were isolated using CF and *in situ* cultivation compared to standard direct plating (SDP) cultivation. Most isolates from CF cultivation were poor growers (with lower specific growth rates and saturated cell densities than those of isolates from SDP cultivation), demonstrating that it is effective to carry out pre-enrichment cultivation targeting bacteria that are less competitive on conventional cultivation, especially K-strategists and bacterial types inhibited by their own growth. Isolates from *in situ* cultivation showed a positive influence on cell recovery stimulated by chemical compounds in the extract of sponge tissue, indicating that some of the bacteria require a "growth initiation factor" that is present in the natural environment. Each advanced cultivation method has its

own distinct key mechanisms allowing cultivation of physiologically and phylogenetically different fastidious bacteria for cultivation compared with conventional methods.

KEYWORDS

uncultured microbes, cultivation, continuous-flow bioreactor, slow-growing bacteria, *in situ* cultivation, I-tip cultivation

## Introduction

The vast number of microbial species that remain uncultivated is sometimes referred to as the mysterious dark-matter of the microbial world (Rinke et al., 2013). Although there is argument over the proportion of uncultivated microorganisms in nature (Martiny, 2019; Steen et al., 2019; Martiny, 2020), culture independent surveys have demonstrated high diversity of uncultivated microbial species (Handelsman, 2004; Locey and Lennon, 2016; Hofer, 2018; Lloyd et al., 2018). Cultivation and isolation of microorganisms are essential steps to enable the discovery of natural products with biotechnological potential. Therefore, solving this microbial uncultivability challenge is a prerequisite for bioprospecting of microbial resources.

The mechanisms underlying why most microbes cannot grow using standard cultivation methods are still unclear. A significantly low proportion (generally less than 1%) of plated microbes readily form visible colonies on standard agar plates, thus leading to plate count anomalies (Staley and Konopka, 1985; Amann et al., 1995). However, no studies have led to a full resolution of the phenomenon. Some investigations have focused on modifying the nutrient composition of media and incubation conditions to overcome the limitations of conventional methods. Modifying growth conditions by adding specific organic or inorganic compounds to media (Bruns et al., 2002; Vartoukian et al., 2016), extending incubation times (Davis et al., 2005), lowering nutrient concentrations (Janssen et al., 2002), modifying the preparation of agar media (Tanaka et al., 2014; Kawasaki and Kamagata, 2017; Kato et al., 2018), and the use of alternative gelling agents (Tamaki et al., 2005) has led to increased microbial recovery, but most of the postulated extant microbes in the natural environment remain uncultivated. This suggests that microbial uncultivability cannot simply be explained by the unfitness of specific strains to certain culture conditions such as media composition, gas composition, temperature, or pH (Stewart, 2012). It might be the case that other essential factors that can more critically affect the growth of fastidious microbes are absent from the artificial conditions in these cultivation methods.

In this study, we used two advanced cultivation methods based on a continuous-flow bioreactor (CF) and *in situ* cultivation to successfully cultivate previously uncultivated bacteria associated with marine sponges. The advanced cultivation method using a CF bioreactor has been applied to enrichment culture of specific microbial targets that play a vital ecological role in nitrogen cycling but were previously uncultivated such as nitrite-oxidizing bacteria, ammonium oxidation (anammox) bacteria, and the methanogenic microbial community in sea floor sediments (Isaka et al., 2006; Imachi et al., 2011; Fujitani et al., 2013). The distinctive feature of this system is the use of porous materials such as polyester nonwoven fabric or polyurethane sponge as carrier material in the bioreactor to provide adequate pore space in the medium, providing an enlarged surface area for growth and a longer cell residence time. In addition, the reactor can maintain low substrate concentrations and high flow rate conditions. The use of the continuous-flow bioreactor is expected to facilitate enrichment of slow/poor growing microbial types which do not grow well under the batch-type cultivation methods. However, the effectiveness of this method to cultivate diverse previously uncultivated microbial types (non-targeted approach for novel species) is still unknown. We expected that use of the CF cultivation before standard cultivation on agar plates would lead to the enrichment of diverse slow and/or poor growing sponge-associated bacteria that have been difficult to isolate using the standard direct plating (SDP) cultivation method.

We also employed one *in situ* cultivation method, the I-tip, to isolate previously uncultivated sponge-associated bacteria. Previous studies using *in situ* cultivation methods aiming to better simulate the natural environment in a variety of environments such as activated sludge, sediments, soils, alkaline soda lakes, sponges, and thermal springs, have shown success for isolating novel microbes (Bollmann et al., 2007; Aoi et al., 2009; Jung et al., 2013; Jung et al., 2014; Jung et al., 2016; Jung et al., 2018; Chaudhary et al., 2019). Recently we clarified some of the principles of *in situ* cultivation methods including the presence of growth initiation factors (signaling-like compounds) in natural environments that can stimulate bacterial resuscitation from a nongrowing state (Jung et al., 2021).

Therefore, we expected that *in situ* cultivation methods allow the isolation of sponge-associated bacteria that were considered previously non-growing using standard cultivation methods.

In the present work, we applied those two advanced cultivation methods to a marine sponge, *Theonella swinhoei*, to isolate previously uncultivated bacterial species. Marine sponges have been identified as rich sources of bioactive secondary metabolites of biotechnological interest for their antiviral, antitumoral, antimicrobial and cytotoxic properties (Koopmans et al., 2009; Mehbub et al., 2014; Anjum et al., 2016). Many studies have suggested that symbiotic microbes in marine sponges produce some of these bioactive metabolites (Brantley et al., 1995; Bewley et al., 1996; Wang, 2006). However, these talented producers remain unexplored and unavailable because the microbes associated with the sponges cannot be easily cultivated (Wang, 2006).

The main objectives of this study were to use two different advanced cultivation methods to access the hidden diversity of marine sponge-associated bacteria, isolate previously uncultivated species and identify some of the key factors for successful isolation previously uncultivated bacteria. We hypothesized that each advanced method would incorporate different key factors for successful isolation of fastidious bacteria. To evaluate the hypothesis, we compared patterns of phylogenetic diversity between isolates from two advanced cultivation methods and compared their cultures with the standard agar plating method. In addition, the physiological characteristics were compared among the isolates from each cultivation method to identify some possible mechanisms operating in each cultivation method.

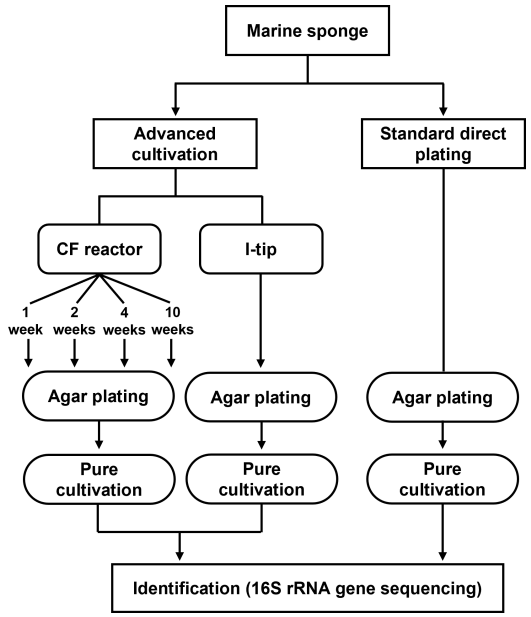



FIGURE 1
Flowchart showing the methods for bacterial cultivation.

# Results

## Identification of isolated bacteria

We applied two advanced cultivation and standard direct plating methods to a marine sponge (Figures 1, 2). The sequencing results of 60 isolates for each cultivation method (in total 120 isolates) were compared. The CF cultivation resulted in isolates representing 31 species (defined as OTUs composed of 16S rRNA gene sequences sharing over 97% identity) from four taxonomic groups (*Actinobacteria*, *Firmicutes*, *Alphaproteobacteria*, and *Gammaproteobacteria*; Figure 3, Table S1). The *in situ* cultivation (I-tip) resulted in isolates representing 25 species from three taxonomic groups (*Bacteroidetes*, *Alphaproteobacteria*, and *Gammaproteobacteria*; Table S2).

We refer to isolates from SDP cultivation performed in the same series of experiments (Jung et al., 2021) using the same sample. SDP cultivation resulted in isolates representing 13 species from three taxonomic groups (*Bacteroidetes*, *Alphaproteobacteria* and *Gammaproteobacteria*; Figure 3, Table S3). Shannon-Weaver diversity indexes (effective number of species) of isolated bacterial species from CF, I-tip and SDP were 21.9, 19.1 and 10.8, respectively. These results showed that advanced cultivation methods yielded higher diversity among the isolates at the species level than among those obtained *via* SDP cultivation. In addition, few common species were found between the three cultivation methods. (Figure 3). There were no species common among the isolates from all three cultivation methods.

The ratios of candidates for novel species, defined as a strain with ≤97% 16S rRNA similarity to the closest known relative among cultivated isolates, were different between the advanced cultivation (I-tip and CF) and SDP methods. Only one SDP isolate (1.6%) was a novel species, while 23% (14 isolates belonging to 5 species) and 45% (27 isolates belonging to 10 species) of I-tip and CF isolates, respectively, were novel species (Figures 3, 4, Tables S1-S3). None of these novel species have more than 95% 16S rRNA similarity to the cultivated isolates obtained from marine sponges in GenBank (www.ncbi.nlm.nih.gov) databases.

## Bacterial community composition determined by the culture-independent method

The microbial community compositions of the marine sponge (*Theonella swinhoei*) samples were analyzed by Illumina-MiSeq

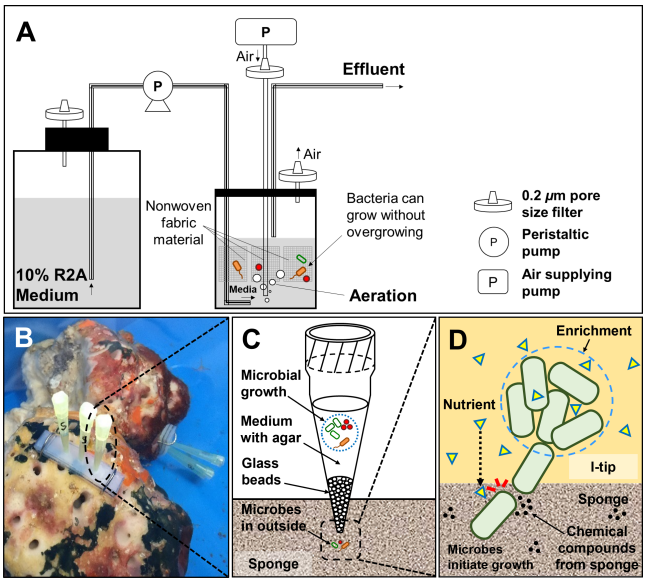



FIGURE 2
Structure and principle of CF (continuous-flow bioreactor) cultivation **(A)**. Schematic showing the *in situ* view **(B)**, structure **(C)** and principle **(D)** of I-tip.

sequencing based on the 16S rRNA gene. A total of 114,624 reads with a median length of 252 base pairs (bp) (V4~533–786 bp) assigned to 1625 OTUs were obtained from the sample. In total 16 taxonomic groups, *Acidobacteria*, *Actinobacteria*, *Bacteroidetes*, *Chlamydiae*, *Chloroflexi*, *Cyanobacteria*, *Firmicutes*, *Nitrospirae*, *Planctomycetes*, *Alpha-*, *Beta-*, *Delta-* *Epsilon-* *Gammaproteobacteria*, *Spirochaetes* and *Verrucomicrobia* were observed in the sponge sample (Figure 5). The isolates from the culture-dependent methods

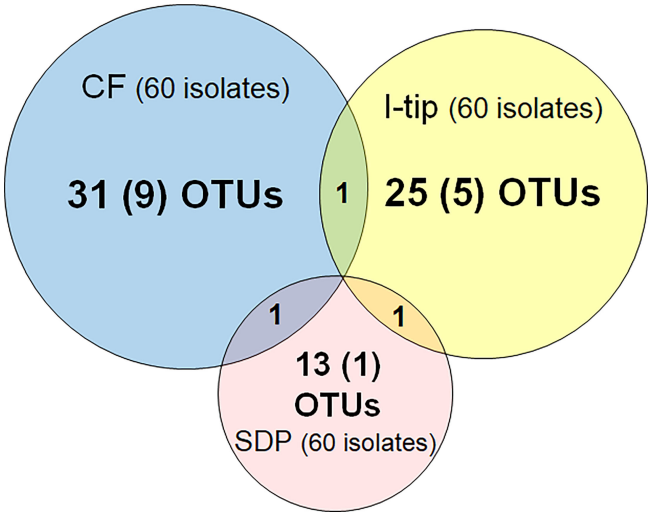



FIGURE 3
Overlap among culture collections obtained using three different cultivation approaches. Values in the center of each circle represent the total number of different isolated species (defined as OTUs composed of 16S rRNA gene sequences sharing over 97% identity) by method; values in parentheses represent the numbers of novel species (defined as a strain with ≤97% 16S rRNA similarity to the closest known relative among cultivated isolates); values in the overlapping areas represent the numbers of co-isolated species. Venn diagrams were made with Venn Diagram Plotter v. 1.5.5228.29250 (Pacific Northwest National Laboratory http://www.pnl.gov/; http://omics.pnl.gov/).

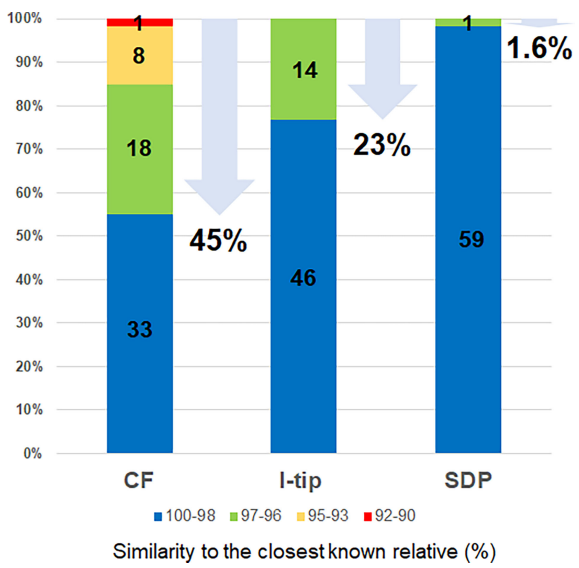



FIGURE 4
Similarity of isolates from each cultivation method based on 16S rRNA gene to the closest known relative in GenBank databases. Numbers in the bar graphs represent OTUs numbers in each similarity level and numbers under arrows represent the ratio of novel species, defined as a strain ≤97% 16S rRNA similarity to the closest known relative.

belonged to 5 groups, *Actinobacteria*, *Bacteroidetes*, *Firmicutes*, *Alpha-*, and *Gammaproteobacteria*.

We also analyzed the 16S rRNA sequences from Illumina-MiSeq sequencing at the genus level and compared the results with those of culture-dependent approaches (Figure 5). For CF cultivation, we used results of isolates from 1, 2, 3 and 4 weeks of CF incubation. A total of 89 genera were identified in the environmental samples. On the other hand, a total of 44 genera appeared in the culture collection for advanced cultivations (33 and 19 in CF and I-tip cultivation, respectively), while the SDP isolates belonged to only 13 genera (Figure 5). Among 44 genera obtained from advanced cultivations, 13 genera appeared in the sponge sample but other 31 genera were obtained only from the culture collection.

Time course of changes in diversity and novelty of isolates from CF

To characterize the phylogenetic variation of isolates from CF over time in the continuous flow bioreactor, 30 isolates were obtained at each sub-cultivation step (at 1, 2, 4 and 10 weeks). Changes in bacterial community composition, diversity and proportion of novel species of isolates from CF cultivation over time were observed (Figure 6). The relative abundance of isolated bacteria at the phylum level dramatically shifted over time. After 1 week of CF cultivation, isolates belonging to *Bacteroidetes*, *Alphaproteobacteria* and *Gammaproteobacteria* were cultivated. The community composition at the phylum level was exactly the same as that from SDP cultivation (Table S3), but dramatically shifted at 2 weeks of CF cultivation.

The diversity at the species level (Shannon-Wiener index) and novelty of isolates (the ratio of novel species among the isolates) also shifted over time (Figure 6). The diversity index (effective number of species) was 4.7 at 1 week increased to 14.9 and 12.0 at 4 and 10 weeks, respectively. While the ratio of novel species among the isolates at 1 week was 6.7% (2/30), the proportion increased to 40% (12/30) and 50% (15/30) at 4 and 10 weeks, respectively.

More than 70% of the organic compounds fed into the reactor were consumed, resulting in a low total organic carbon concentration (TOC) in the reactor (below 30 mg-C/L during most of the incubation) throughout the experiment (Figure S1).

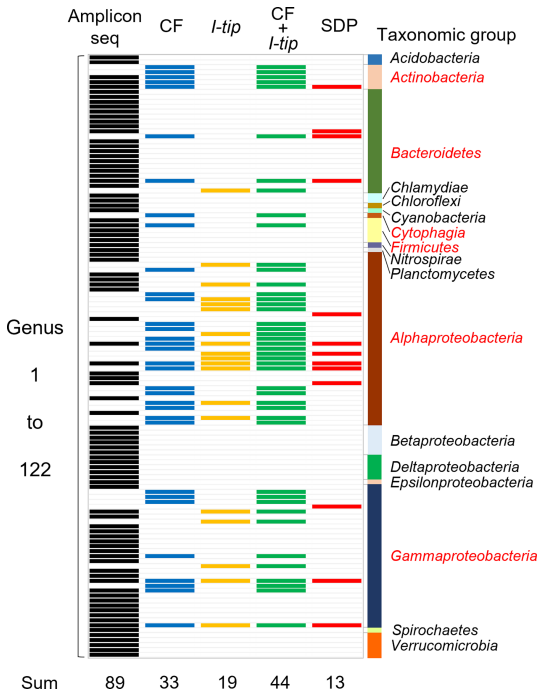



FIGURE 5
Band chart showing the comparison of microbial community compositions at the genera level among the sponge microbial community (amplicon sequencing data), isolates from CF, I-tip, and SDP cultivations. A total of 122 genera were detected from culture-independent and dependent approaches. Numbers at the bottom indicate number of genera detected with each method. Those are 89 from the amplicon sequencing data (black-colored bands), 33 from CF isolates (blue-colored bands), 19 from I-tip isolates (yellow-colored bands), 44 from sum of CF and I-tip isolates (green-colored bands), and 13 from SDP isolates (red-colored bands). Bar graph on the right indicates the taxonomic group to which those genera belong. The red letters for taxonomic group's name represent the group that also isolated by the culture-dependent methods.

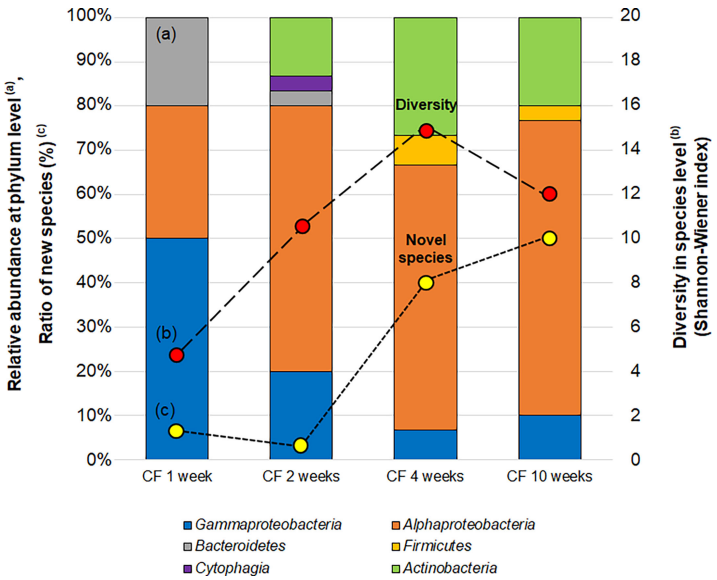



FIGURE 6
Changes in the bacterial community (a; bar graph), diversity (b; red circle) and proportion of novel species (c; yellow circle) of isolates CF (continuous-flow bioreactor) cultivation at 1, 2, 4 and 10 weeks. Bacterial community **(A)** shown as relative abundance at phylum level (class level for *alpha*, *Gamma and Epsilon proteobacteria* group). Diversities of isolates **(B)** are compared by the effective number of species from Shannon-Wiener index, exp $(\sum_{i/1}^{S} p_i \ln p_i)$. Novel species **(C)** were defined as a strain ≤97% 16S rRNA similarity to the closest known relative.

## Growth characteristics of isolated species

For further investigation of the growth characteristics (specific growth rate, saturated cell density and growth curve type), we selected 25, 15 and 11 strains obtained from CF, I-tip and SDP cultivations, respectively, representing identical species of all identified isolates in each cultivation method except some species which had failed to be sub-cultivated during further experiments.

The specific growth rates and saturated cell density of selected strains from CF, I-tip, and SDP cultivation are plotted in Figure 7 (we hereafter define those selected strains as CF, I-tip, SDP strains, respectively). Most of the patterns for the CF strains were distinct from the I-tip, and SDP strains, as the specific growth rates and saturated cell density (carrying capacity) of most CF strains (20/25, 80%) were significantly lower than those of the I-tip, and SDP strains (optical density [OD] at 600 nm less than 0.005 h-1 and 0.115, respectively), except for five strains (belonging to 5 distinct OTUs).

The type of growth curve for CF strains also differed between that of strains from other cultivation methods (Table S4, Figure S2). Most CF strains (19/25, 75%) showed concave down growth curves, whereas all I-tip (15/15) and SDP strains (11/11) had concave up growth curves (logistic growth curve).

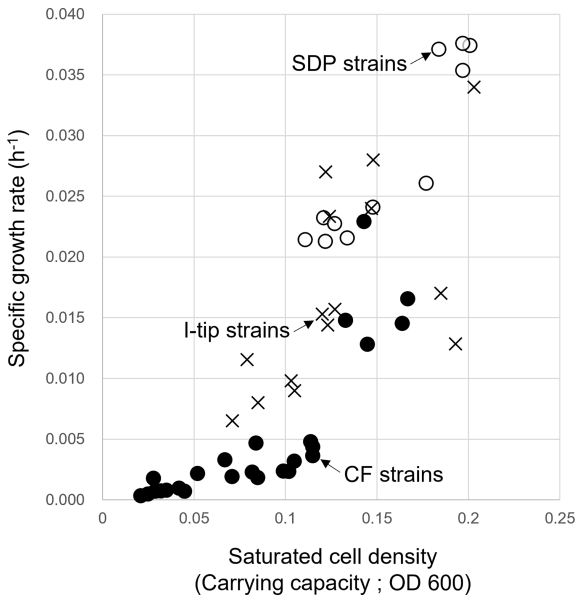



FIGURE 7
Specific growth rate and saturated cell density of CF (continuous-flow bioreactor; filled circle), I-tip (cross), and SDP (standard direct plating; open circle) strains. Data for specific growth rate and saturated cell density of SDP strains have already been reported (Jung et al., 2021).

## Colony size

The average colony size among the cultivation methods was compared using the Kruskal-Wallis procedure followed by *post-hoc* Dunn's tests (Weitzman et al., 2018) (Figure 8). These comparisons indicated that the average colony diameter of CF strains was significantly smaller than that for I-tip and SDP strains (Kruskal-Wallis Rank Test; df =2, P< 0.00001; Figure 8).

## Effect of sponge extract on cell recovery

The colony formation efficiency ratios between the two culture conditions (media with and without the sponge extract) were calculated for selected CF, I-tip and SDP strains to examine the effect of the sponge extract on starvation recovery. The effect of sponge extract on the cell recovery rate on agar media was clearly different between the I-tip and SDP strains (Figure 9). The addition of sponge extract positively affected the recoveries of all I-tip strains. Especially two strains, I14 and I17, had more than 5 times colony numbers on the agar media with the sponge extract compared with the colony numbers on the media with no sponge extract. In contrast, no SDP strains were positively affected or rather negatively affected by the sponge extract addition, except for two strains that showed a slight positive effect (Figure 9). Note that the closest species for one SDP strain (SA7) showing a positive effect was *Pseudovibrio denitrificans*, which was commonly found in I-tip isolates (I9; Table S2).

# Discussion

## How does CF cultivation allow cultivation of fastidious bacterial types?

A variety of cultivation efforts have made use of several alternative techniques to isolate previously uncultivated sponge-associated microbes, and those approaches yielded an increased novelty of sponge isolates and improved cultivability rates up to 14% in some cases (Sipkema et al., 2011). However, most postulated extant microbes in sponges remain uncultivated; thus, further efforts to discover such novel microbes are needed (Wang, 2006).

In this study, CF cultivation enabled obtaining a different culture collection that was larger and more novel than that obtained by SDP cultivation (Figure 4, Table S1). Identified CF isolates shared only one species in common with isolates from I-tip and SDP cultivation, respectively (Figure 3). This indicates that CF cultivation method has its own selectivity, which is different from that of the I-tip and SDP methods. CF cultivation allows enrichment of physiologically different bacterial types (Figure 7) and resulting isolation of phylogenetically different bacterial types. Furthermore, more novel species were obtained from the sample with a longer period of CF cultivation

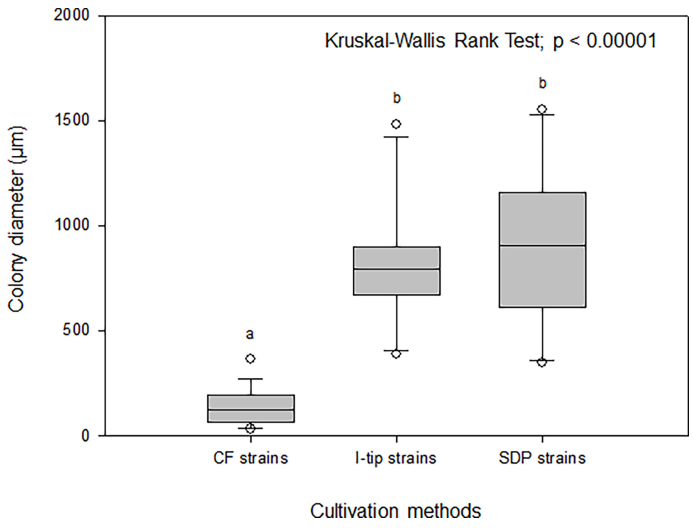



FIGURE 8
Box-and-whisker plots of colony diameter for CF (continuous-flow bioreactor), I-tip and SDP (standard direct plating) strains. The line inside each box indicates the median value. Lines extending from the boxes represent minimum and maximum values excluding outliers. Closed circles indicate the mean. Average colony diameter of each selected strain was measured by randomly selected 20 colonies on agar plate. Average colony diameter was compared among cultivation methods using the Kruskal-Wallis procedure. Letters above boxes indicate significant differences using *post-hoc* Dunn's tests.

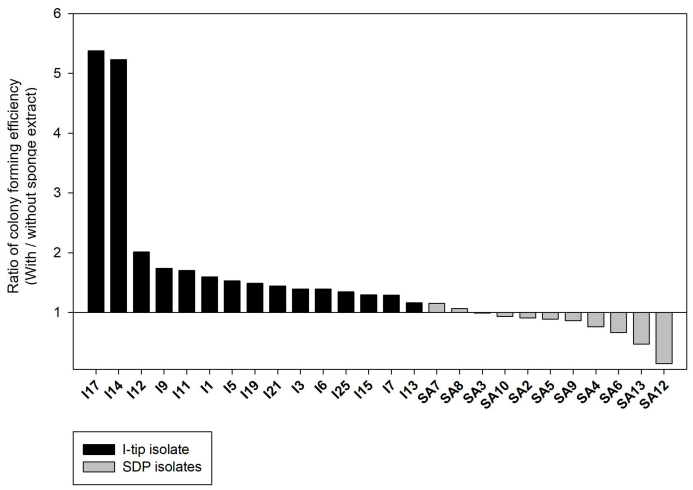



FIGURE 9
The effect of sponge extract on cell recovery rate of I-tip and SDP (standard direct plating) strains. The ratio of colony forming numbers between the two culture conditions (colony numbers on the agar medium with the sponge extract/colony numbers on the agar medium without the sponge extract) was calculated for each selected strain. The ratio was calculated using the average value from triplicate measurements (n=3) of the CFUs on each agar plate. Data for SDP strains have already been reported (Jung et al., 2021).

(Figure 6), suggesting that uncultivated bacterial types unique to CF cultivation require a certain length of time to be enriched in the reactor.

As described earlier, we expected that CF cultivation results in the growth of specific bacterial types that grow poorly on agar plates, which increases the efficiency of isolating previously uncultivated bacterial types. The specific growth rates of CF strains were much lower than those of strains from other methods (Figure 7), indicating that CF cultivation is suitable for the relatively slow-growing bacteria. The nonwoven fabric material used in the CF method was effective in retaining bacterial biomass in the bioreactor under the high flow rate condition, and enriching slow-growing microbes.

Furthermore, medium was continuously provided and completely mixed by aeration in the reactor and continuously discharged to keep the low concentration of the substrate in the CF reactor. Unlike batch or batch-fed type enriched cultivation, the low substrate concentration prevents overgrowth of specific bacterial types, especially r-strategists (higher growth rate under substrate unlimited conditions). This is effective to enrich K-strategists (competitive under substrate limited conditions, and usually slow growers in conventional cultivation methods).

Indeed, previous research successfully demonstrated the enrichment of anaerobic ammonium oxidation (anammox) bacteria, which were hardly cultivated even in enrichment and still not yet purely isolated (Date et al., 2009) and the selective enrichment of slow growing nitrite oxidizing bacteria *Nitrospira* (Fujitani et al., 2013; Fujitani et al., 2014). In this study, we used CF cultivation for the first time to isolate diverse marine sponge-associated bacteria, rather than targeting single species.

However, the favorable conditions for slow-growing bacteria are likely not the only mechanism for isolating a unique bacterial culture collection, because extending the incubation time of SDP cultivation did not affect microbial composition, diversity and novelty of isolates (Figure S3). Other factors in CF cultivation, such as the removal of metabolic waste from bacteria being cultured and difference in medium (solid vs liquid) can explain the result here.

The maximum growth (saturated cell density) of CF strains shown in their growth curves were remarkably lower than those of I-tip and SDP strains (Figure 7). Furthermore, by classifying the type of growth curve for each strain for the different cultivation methods, additional information on the growth characteristics was gained. Most CF strains had the growth curve classified as a concave down type, whereas none of the I-tip and SDP strains showed this type of growth curve but were classified as concave up (logistic growth curve) (Table S4, Figure S2). These results indicate that the growth of CF isolates was suppressed even if their cell concentrations were comparatively low. In other words, it suggests that most strains from CF cultivation are bacterial species that are more sensitive to metabolites or metabolic by-products and, as a result, those strains are less competitive in batch cultivation. This is also suggested by the observation that the average colony diameter of CF strains. was significantly smaller than that of I-tip and SDP strains (Figure 8). The flow-through conditions in the CF reactor minimizes this negative effect and allow enrichment of less competitive bacterial species under the standard cultivation condition. This concept has been previously presumed but not fully explored (Imachi et al., 2011).

## How does *in situ* cultivation (I-tip) allow cultivation of fastidious bacterial types?

The I-tip method produced a richer collection of more novel strains that are fastidious for cultivation with the SDP method (Figure 3, 4). This is likely because the first step of the method, *in situ* cultivation, mimics natural conditions and allows for growth of those species that are only active in the environment (Jung et al., 2014). What mechanisms yield a different culture collection that differs entirely from the culture collection from the standard cultivation method?

We have developed and utilized various types of *in situ* cultivation methods to isolate microbes from a variety of environmental samples (Aoi et al., 2009; Jung et al., 2013; Jung et al., 2014; Jung et al., 2016; Jung et al., 2018). From these previous studies, two key principles of *in situ* cultivation can be identified. The first is that, if microbes are grown *in situ*, e.g., in a diffusion chamber, the cultivation results in higher recovery and increased novelty of isolates (Jung et al., 2013; Jung et al., 2014; Jung et al., 2016; Jung et al., 2018). The second is the observation that some microbes, those which initiate growth and thrive during *in situ* incubation, continue their growth even during the sub-cultivation step *ex situ*. Therefore, we hypothesized that some factors facilitating bacterial resuscitation in natural environments arouse inactive cells and enrich them during *in situ* cultivation, which results in recovery of fastidious bacterial species that cannot be cultivated by using standard cultivation.

This hypothesis was supported by the results of this study, and also those obtained using a diffusion chamber with the same marine sponge samples as in the present study (Jung et al., 2021). The results of the cell recovery experiment clearly support the hypothesis (Figure 9). Most of I-tip strains showed a positive effect on cell recovery using chemical compounds in sponge tissue. Therefore, I-tip cultivation led to isolation of a different bacterial culture collection that could not be cultivated with SDP cultivation. In contrast, most SDP strains showed a negative effect on cell growth with the sponge extract, likely due to some antibacterial metabolites in sponges such as secondary metabolites or antimicrobials from Actinomycetes (Bewley et al., 1996; Schmidt et al., 2000), or high concentration of trace toxic compounds which can accumulate in sponge tissue (Keren et al., 2016). Those bacterial types probably grow rapidly on rich media typical in standard cultivation methods (see growth rate result of SDP strains; Figure 7), and often have negative effects on the growth of more slowly growing oligotrophs (Jung et al., 2013), which results in culture collections with skewed species distributions (Jung et al., 2018).

Adding the sponge extract at higher concentrations than that in the original experimental condition (e.g., 1%, which was 10 times higher than the original experiment condition) strongly inhibited growth even for isolates from I-tip cultivation (data not shown). This might be one reason why SDP cultivation with sponge extract media (including 4% of sponge extract) did not result in isolating diverse and novel microbial species. Although media containing 1-10% of sponge extract have been widely used in several previous studies for isolating sponge-associated microbes (Selvin et al., 2009; Sipkema et al., 2011; Steinert et al., 2014; Esteves et al., 2016), our results suggest that adding lower concentrations of sponge extract (0.1% in this study) is advantageous in isolating a variety of microorganisms.

We performed the same experiment with CF strains, and compared the results with that of I-tip (Figure S4). Unlike for I-tip strains, the addition of the sponge extract did not measurably affect the cell recovery of strains from CF cultivation.

## Bacterial community comparison of the culture-dependent and -independent methods

Comparing the bacterial composition of the 16S rRNA gene from the culture-independent method and that of the culture collections indicates that some bacterial groups that cannot be cultured by standard cultivation can be isolated with advanced cultivation approaches (Figure 5). The advanced cultivation methods, CF and I-tip, led to the isolation of several bacterial groups at the genus level that were not cultivated in the SDP method. On the other hand, there are several genera obtained from culture-dependent approaches (31 and 4 in advanced cultivations and SDP cultivation, respectively) that does not appear in the sponge samples observed by Illumina MiSeq sequencing analysis. These bacterial species are probably present in extremely low numbers in the sponge tissue, and may not be detected in a 16S rRNA gene clone library. This hypothesis is supported by several previous studies about analysis of the diversity in the clone library from sponge and other environmental samples (Sipkema et al., 2009; Sipkema et al., 2011). Some isolates that exclusively exist in culture-dependent methods may correspond to these "hidden" OTUs in the clone library.

## Conclusion and perspectives

The results presented here show the utility of advanced cultivation methods to cultivate diverse previously uncultivated sponge-associated bacteria, including 1) combined use of a CF cultivation system and 2) combined use of *in situ* cultivation (I-tip), prior to agar plating as an isolation step. Furthermore, by demonstrating advanced approaches to microbial cultivation and comparing physiological properties of isolates from different approaches, we clearly identified two key points focused on the efficient cultivation of previously uncultivated microbes. Those are (1) targeting bacterial types that are less competitive (poorly or slow growing) on batch type cultivation, especially K-strategists and bacterial types inhibited

by themselves, and (2) focusing on bacterial types which require a signaling-like factor for restarting to grow (turning on the switch to grow). Achieving this provides crucial insights into microbial uncultivability that can be applied to build new cultivation strategies. In addition, newly discovered microbes are candidates as sources for valuable secondary metabolites.

Nevertheless, there are still microbial species in marine sponges that have been represented in culture-independent methods but were not represented in our culture collection (Figure 5). In addition to the mechanisms explored in this study, there are likely several other important factors operating to cultivate uncultivated microbial types. Continued advances in cultivation methods together with improved understanding of the key factors for cultivating previously uncultivated microbes are needed.

# Materials and methods

## Sample collection

Marine sponges (*Theonella swinhoei*) were collected by scuba diving near Okino Island, (Kochi Prefecture, Japan) in August 2015 at a depth of approximately 15-20 m. Sponges were kept cool in an insulated box with sea water and transported to the aquarium and laboratory for the experiment.

## Media

We used 1/10 diluted Reasoner's 2A (R2A) media (10% of the manufacturer's suggested concentration, Nihon Seiyaku, Japan) for the advanced cultivations, which were adapted from previous study isolating wide variety of sponge-associated bacteria (Jung et al., 2014), and used four different media for sub cultivation in order to cover diverse microbial species: (1) 1/10 diluted R2A medium; (2) marine medium (Difco, Franklin Lakes, NJ, USA); (3) fish extract media (0.2 g fish extract, 0.1 g yeast extract per liter); (4) sponge extract medium (0.01 g peptone, 40 ml aqueous sponge extract per liter). The sponge extract was prepared by mixing homogenized sponge tissue and sterile distilled water including 3.5% of artificial sea salt at a 1:1 (vol/vol) ratio, vortexing for 1 min, centrifugation (10 min, 8,000 rpm), and filter sterilization using a 0.2-μm pore size filter. All media except marine media were supplemented with 3.5% artificial sea salt, SEALIFE (Marine Tech, Tokyo, Japan).

## CF cultivation

The overall experimental design of the cultivation experiments is shown in Figure 1. The CF with porous polyester nonwoven fabric (0.5 cm thickness; Japan Vilene, Tokyo, Japan) was set up to maintain the bacteria in an active state without overgrowing (Figure 2A), according to Fujitani et al. (2013). The bioreactor volume was 500 ml and maintained at 20°C. The reactor was aerated to provide oxygen. Due to the aeration, the liquid phase was completely mixed. A total of 200 ml of homogenized sponge tissue was used as inoculum. To prepare the inoculum, marine sponge tissue (about 150 g) was rinsed three times with sterile artificial sea water and pulverized using a sterile mortar and pestle with the same volume of sterile artificial sea water. The 1/10 diluted R2A including 3.5% of artificial sea salt was supplied from the medium chamber by a peristaltic pump at a velocity of 500 ml per day (hydraulic retention time of 24 h). Therefore, any media overflowing was automatically discharged from the bioreactor in order to reduce the concentration of metabolic by- products and overgrown biomass such as inactive cells, debris, and extracellular polymeric substances. The total organic carbon concentration in the bioreactor was measured during the experiment at intervals of three to seven days with a TOC analyzer (Shimadzu, TOC-500, Kyoto, Japan) to monitor conditions in the bioreactor.

Biomass from the nonwoven fabric material was sub-cultured on agar plate at 1, 2, 4 and 10 weeks of CF cultivation to examine temporal changes in the bacterial community. For sub-culturing, a piece of fabric material was randomly selected and retrieved from the reactor. One gram of fabric was added to 10 ml of sterile water including 3.5% of artificial sea salt and vortexed for 10 min. We then applied the homogenizer by controlling the power at the low level (less than 30% of the maximum power of the machine, Ultrasonic Homogenizer NR-50M, Microtec, Japan) for 10 seconds at a frequency of 20 kHz to dislodge biofilm and the associated bacteria from the surface of fabric material. Liquid with detached bacteria was diluted and inoculated on media described above with 1.5% agar. After one-week incubation at 20°C, 30 colonies were randomly selected for bacterial identification. The isolates identified from 4 and 10 weeks of CF cultivation (total 60) were used as the final isolates to compare among the cultivation methods.

## I-tip *in situ* cultivation

The I-tip method allows both microbes and chemical compounds to enter from the environment, thereby better simulating the natural environment (Figure 2). Because a micropipette tip is used as main element of the I-tip device, the sharp end can be easily installed in invertebrates such as sponges. The I-tip was prepared as described previously (Jung et al., 2014). In brief, 200μL pipette tips were prepared (Eppendorf, No. 022493022, Hamburg, Germany) with the bottom part of the tip filled with acid-washed glass beads (60 to 100 μm in diameter, Sigma-Aldrich, No. G1145, St. Louis, MO, USA) to prevent invasion of larger organisms. Then 70 μl

of sterilized 1/10 diluted R2A medium containing 0.7% (wt/vol) agar was added to the tip for bacterial growth. After autoclaving the assembly, the wide end of the tip was aseptically sealed with a waterproof adhesive (Cemedain, Tokyo, Japan) to prevent contamination from air and water.

For *in situ* cultivation using the I-tip method, sponges were kept in aquaria at the Takehara marine station of Hiroshima University (Takehara, Japan). The aquarium was cylindrical with a holding capacity of ca. 43 L (Ø = 0.35 m, d = 0.45 m), and seawater was continuously supplied. To install I-tip, silicon supporting plates (w = 3.5 cm × 1.2 cm, d = 0.6 cm) were placed on the surface of a sponge, and six I-tips (three per plate) were inserted into the specimen through the silicon plate to a depth of 0.3 cm (Figure 2D). After one week of *in situ* incubation, I-tips were retrieved from the sponges and then transported to the laboratory. The wide end of each I-tip was cut aseptically with a blade, and then the agar with microbes was removed using a sterile loop. The agar material from each tip was added to 1 ml of sterilized water including 3.5% of artificial sea salt, homogenized with a sterile stick, vortexed, diluted with sterile water including 3.5% artificial sea salt, and sub-cultured on media as described above with 1.5% agar for incubation at 20°C. After one week, 60 colonies were selected randomly and pure cultured for further experiments.

## Identification of isolates

Taxonomic identification was carried out by sequencing 643 to 782 long fragments of the 16S rRNA gene. The colony material was used directly as a template for PCR. The 16S rRNA gene was amplified using the universal primer as follows; 27F (5'-AGAGTTTGATCCTGGCTCAG-3') and 1492R (5'-GGTTACCTTGTTACGACTT-3') with a KOD FX Neo system (Toyobo, Osaka, Japan), then purified with fast gene purification kit (Nippon Genetics, Tokyo, Japan). The purified PCR products were sequenced using a commercial sequencer (Takara Bio, Shiga, Japan) by fluorescent dye terminator sequencing. The sequences were compared to those available in GenBank (www.ncbi.nlm.nih.gov) databases using the MEGA program (MEGA software, Tempe, AZ, USA) to determine their closest relatives and their similarity.

## DNA extraction and amplicon sequencing targeting the 16S rRNA gene

To compare the cultivated bacterial diversity with the microbial molecular signatures in the host sponge, 16S amplicon sequencing was performed based on the 16S rRNA genes. After transferring the marine sponges to the laboratory, samples were washed three times with DNA-free water and homogenized. Genomic DNA from the homogenized tissue was extracted using a FastDNA spin kit for soil (MP Biomedicals, Irvine, CA, USA) per the manufacturer's guidelines. The extracted genomic DNA was amplified using the primers, 341F and 805R, with a Hifi hot start ready mix PCR (Kapa Biosystems, Wilmington, MA, USA), then purified using a fast gene purification kit (Nippon Genetics, Tokyo, Japan). DNA was sequenced by Hokkaido System Science (Sapporo, Japan) using an Illumina MiSeq System (Illumina, San Diego, CA, USA). Each read was assigned taxonomically using Quantitative Insights Into Microbial Ecology (QIIME) software (Kuczynski et al., 2011).

## Selection of strains for further experiments

For further investigation, we selected 25, 15 and 11 strains from CF, I-tip and SDP cultivation methods, respectively, except for a few lost isolates. Each strain represents each identified OTU within the group of isolates derived from the same cultivation method.

## Growth characteristics of isolated strains

Growth curves of the selected strains were analyzed to compare bacterial growth among the strains from each cultivation. Diluted R2A medium (5 mL) was inoculated with 5-20 $\mu$L of cell suspension, in triplicate, from pure liquid cultures grown, then incubated at 20°C with shaking at 150 rpm. The optical density (OD) was measured at 600 nm using a spectrophotometer (DR 3900, HACH, USA). Growth curves from OD 600 values were fitted using the nonlinear regression function with logistic model of Sigma plot software (Systat Software Inc). The specific growth rate ($\mu$), representing the maximum rate of change, was calculated by the change in OD 600, and the saturated cell density (maximum growth) was determined by its maximum value on the fitted growth curve.

We also classified the type of growth curve for the strains into two categories. One type was a concave up growth curve (logistic growth curve) that showed exponential growth during the log phase, and the other was a concave down growth curve with the growth increments decreasing and become slower. To determine the type of growth curve, we used the values in fitted growth curves with logistic model of each tested strain above. Then these values were examined on a log scale to determine whether exponential growth occurred during log phase (concave up growth curve) or not (concave down growth curve).

## Colony size of isolated strains

Colony size on agar medium (1/10 R2A) was compared among selected strains for each cultivation method. The average colony diameter of each strain was measured for selected 20 colonies on the agar plate under a digital microscope (KH-8700, Hirox Europe, France). For that, we used colonies of pure cultured strains on agar plate that grew four weeks after inoculation, and

randomly selected 20 colonies. The colony diameter was compared among cultivation methods using the Kruskal-Wallis procedure and *post-hoc* Dunn's tests were carried out on each pair of groups to test for significant differences between the groups (30).

### Effect of sponge extract on cell recovery

The colony formation efficiency ratios between the two culture conditions (media with and without the sponge extract) were calculated for selected strains from CF and I-tip cultivation, and compared with the results for those strains from SDP cultivation to evaluate the effect of chemical compounds on cell recovery. First, one strain per OTU (defined at 97% 16S rRNA gene sequence identity) was pure cultured in 5 ml of 1:10 diluted R2A broth with 3.5% artificial sea salt at 20°C, then the liquid culture was diluted with artificial sea water (1:100) followed by incubation at 5°C for 3 days to inactivate microbes. As test strains after the pre-cultivation were diluted in 100-fold with artificial seawater, this process made microbial cells starved under low temperature, and this condition was enough for the strains from marine sponge to show a different recovery ratio according to two conditions, with and without sponge extract in a previous study (Jung et al., 2021). After 3 days, the liquid culture serial dilutions were inoculated in triplicate on two types of agar media: 1:10-diluted R2A agar medium with 0.1% (vol/vol) sponge extract and the same medium without the sponge extract. The sponge extract was added to the medium before autoclaving as we confirmed that autoclaved sponge extract remains a key function and worked for our purpose in a previous study (Jung et al., 2021). After 5 days of incubation at 20°C, the colony number ratios between the two cultivation methods (with and without the sponge extract) were calculated for each strain.

### Nucleotide sequence accession numbers

Newly determined sequence data have been deposited in GenBank (www.ncbi.nlm.nih.gov) under accession numbers MK697035 to MK697065 for isolates from CF, MK689993 to MK690017 for isolates from I-tip, and MK674856 to MK674892 for isolates from SDP cultivation. The data obtained by Illumina-MiSeq sequencing (microbial community analysis) have been deposited in the DNA Data Bank of Japan (DDBJ) Sequence Read Archive (DRA) under accession number DRA008671.

## 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

DJ, YN and YA conceived and designed the study. DJ and KM performed the laboratory experiments. DJ, JO, SH, TK, AO and YA analyzed the data and compiled the results. DJ and YA wrote the manuscript draft. All authors contributed to the article and approved the submitted version.

## Funding

This work was supported by JSPS KAKENHI Grant Numbers JP17F17098, JP19H02873, JP25630383, 26709063, JP18K19181. DJ was supported by the Japan Society for the Promotion of Science (JSPS) as an International Research Fellow of the JSPS. DJ and SH were supported by the National 111 Project of China (D16013).

## Acknowledgments

The authors would like to acknowledge S. Iwasaki at Takehara Station (Fisheries Research Station), Hiroshima University for his support on operation of the aquarium.

## 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/fmars.2022.963277/full#supplementary-material

# References


Amann, R. I., Ludwig, W., and Schleifer, K. H. (1995). Phylogenetic identification and *in situ* detection of individual microbial cells without cultivation. *Microbiol. Rev.* 59, 143–169. doi: 10.1128/mr.59.1.143-169.1995

Anjum, K., Abbas, S. Q., Shah, S. A. A., Akhter, N., Batool, S., and Hassan, S. S. U. (2016). Marine sponges as a drug treasure. *Biomol. Ther.* 24, 347–362. doi: 10.4062/biomolther.2016.067

Aoi, Y., Kinoshita, T., Hata, T., Ohta, H., Obokata, H., and Tsuneda, S. (2009). Hollow-fiber membrane chamber as a device for *in situ* environmental cultivation. *Appl. Environ. Microbiol.* 75, 3826–3833. doi: 10.1128/AEM.02542-08

Bewley, C. A., Holland, N. D., and Faulkner, D. J. (1996). Two classes of metabolites from *Theonella swinhoei* are localized in distinct populations of bacterial symbionts. *Experientia* 52, 716–722. doi: 10.1007/BF01925581

Bollmann, A., Lewis, K., and Epstein, S. S. (2007). Incubation of environmental samples in a diffusion chamber increases the diversity of recovered isolates. *Appl. Environ. Microbiol.* 73, 6386–6390. doi: 10.1128/AEM.01309-07

Brantley, S. E., Molinski, T. F., Preston, C. M., and DeLong, E. F. (1995). Brominated acetylenic fatty acids from xestospongia sp., a marine sponge bacteria association. *Tetrahedron* 51, 7667–7672. doi: 10.1016/0040-4020(95)00405-W

Bruns, A., Cypionka, H., and Overmann, J. (2002). Cyclic AMP and acyl homoserine lactones increase the cultivation efficiency of heterotrophic bacteria from the central Baltic Sea. *Appl. Environ. Microbiol.* 68, 3978–3987. doi: 10.1128/AEM.68.8.3978-3987.2002

Chaudhary, D. K., Khulan, A., and Kim, J. (2019). Development of a novel cultivation technique for uncultured soil bacteria. *Sci. Rep.* 9, 6666. doi: 10.1038/s41598-019-43182-x

Date, Y., Isaka, K., Ikuta, H., Sumino, T., Kaneko, N., Yoshie, S., et al (2009). Microbial diversity of anammox bacteria enriched from different types of seed sludge in an anaerobic continuous-feeding cultivation reactor. *J. Biosci. Bioeng.* 107, 281–286. doi: 10.1016/j.jbiosc.2008.11.015

Davis, K. E. R., Joseph, S. J., and Janssen, P. H. (2005). Effects of growth medium, inoculum size, and incubation time on culturability and isolation of soil bacteria. *Appl. Environ. Microbiol.* 71, 826–834. doi: 10.1128/AEM.71.2.826-834.2005

Esteves, A. I. S., Amer, N., Nguyen, M., and Thomas, T. (2016). Sample processing impacts the viability and cultivability of the sponge microbiome. *Front. Microbiol.* 7, 499. doi: 10.3389/fmicb.2016.00499

Fujitani, H., Aoi, Y., and Tsuneda, S. (2013). Selective enrichment of two different types of Nitrospira-like nitrite-oxidizing bacteria from a wastewater treatment plant. *Microbes Environ.* 28 (2), 236–243. doi: 10.1264/jsme2.ME12209

Fujitani, H., Ushiki, N., Tsuneda, S., and Aoi, Y. (2014). Isolation of uncultured Nitrospira. *Environ. Microbiol.* 16, 3030–3040. doi: 10.1111/1462-2920.12248

Handelsman, J. (2004). Metagenomics: application of genomics to uncultured microorganisms. *Microbiol. Mol. Biol. Rev.* 68, 669–685. doi: 10.1128/MMBR.68.4.669-685.2004

Hofer, U. (2018). The majority is uncultured. *Nat. Rev. Microbiol.* 16, 716–717. doi: 10.1038/s41579-018-0097-x

Imachi, H., Aoi, K., Tasumi, E., Saito, Y., Yamanaka, Y., Saito, Y., et al. (2011). Cultivation of methanogenic community from subseafloor sediments using a continuous-flow bioreactor. *ISME J.* 5, 1913–1925. doi: 10.1038/ismej.2011.64

Isaka, K., Date, Y., Sumino, T., Yoshie, S., and Tsuneda, S. (2006). Growth characteristic of anaerobic ammonium-oxidizing bacteria in an anaerobic biological filtrated reactor. *Appl. Microbiol. Biotechnol.* 70, 47–52. doi: 10.1007/s00253-005-0046-2

Janssen, P. H., Yates, P. S., Grinton, B. E., Taylor, P. M., and Sait, M. (2002). Improved culturability of soil bacteria and isolation in pure culture of novel members of the divisions Acidobacteria, Actinobacteria, Proteobacteria, and Verrucomicrobia. *Appl. Environ. Microbiol.* 68, 2391–2396. doi: 10.1128/AEM.68.5.2391-2396.2002

Jung, D., Aoi, Y., and Epstein, S. S. (2016). *In situ* cultivation allows for recovery of bacterial types competitive in their natural environment. *Microbes Environ.* 31, 456–459. doi: 10.1264/jsme2.ME16079

Jung, D., Machida, K., Nakao, Y., Kindaichi, T., Ohashi, A., and Aoi, Y. (2021). Triggering growth *via* growth initiation factors in nature: a putative mechanism for *in situ* cultivation of previously uncultivated microorganisms. *Front. Microbiol.* 12, 537194. doi: 10.3389/fmicb.2021.537194

Jung, D., Seo, E., Epstein, S. S., Joung, Y., Han, J., Parfenova, V. V., et al. (2014). Application of a new cultivation technology, I-tip, for studying microbial diversity in freshwater sponges of Lake Baikal, Russia. *FEMS Microbiol. Ecol.* 90, 417–423. doi: 10.1111/1574-6941.12399

Jung, D., Seo, E., Epstein, S. S., Joung, Y., Yim, J. H., Lee, H., et al. (2013). A new method for microbial cultivation and its application to bacterial community analysis in Buus Nuur, Mongolia. *Fundam. Appl. Limnol.* 182, 171–181. doi: 10.1127/1863-9135/2013/0391

Jung, D., Seo, E., Owen, J. S., Aoi, Y., Yong, S., Lavrentyeva, E. V., et al. (2018). Application of the filter plate microbial trap (FPMT), for cultivating thermophilic bacteria from thermal springs in Barguzin area, eastern Baikal, *Biosci. Biotechnol. Biochem.* 82, 1624–1632. doi: 10.1080/09168451.2018.1482194

Kato, S., Yamagishi, A., Daimon, S., Kawasaki, K., Tamaki, H., Kitagawa, W., et al. (2018). Isolation of previously uncultured slow-growing bacteria by using a simple modification in the preparation of agar media. *Appl. Environ. Microbiol.* 84, e00807–e00818. doi: 10.1128/AEM.00807-18

Kawasaki, K., and Kamagata, Y. (2017). Phosphate-catalyzed hydrogen peroxide formation from agar, gellan, and κ-carrageenan and recovery of microbial cultivability *via* catalase and pyruvate. *Appl. Environ. Microbiol.* 83, e01366–e01317. doi: 10.1128/AEM.01366-17

Keren, R., Lavy, A., and Ilan, M. (2016). Increasing the richness of culturable arsenic-tolerant bacteria from Theonella swinhoei by addition of sponge skeleton to the growth medium. *Microb. Ecol.* 71, 873–886. doi: 10.1007/s00248-015-0726-0

Koopmans, M., Martens, D., and Wijffels, H. R. (2009). Towards commercial production of sponge medicines. *Mar. Drugs* 7, 787–802. doi: 10.3390/md7040787

Kuczynski, J., Stombaugh, J., Walters, W. A., González, A. A., Caporaso, J. G, and Knight, R. (2011). Using QIIME to analyze 16S rRNA gene sequences from microbial communities. *Curr. Protoc. Bioinformatics 10* 10, 7. doi: 10.1002/0471250953.bi1007s36

Lloyd, K. G., Steen, A. D., Ladau, J., Yin, J., and Crosby, L. (2018). Phylogenetically novel uncultured microbial cells dominate earth microbiomes. *mSystems* 3, e00055–e00018. doi: 10.1128/mSystems.00055-18

Locey, K. J., and Lennon, J. T. (2016). Scaling laws predict global microbial diversity. *Proc. Natl. Acad. Sci. U.S.A.* 113, 5970–5975. doi: 10.1073/pnas.1521291113

Martiny, A. C. (2019). High proportions of bacteria are culturable across major biomes. *ISME J.* 13, 2125–2128. doi: 10.1038/s41396-019-0410-3

Martiny, A. C. (2020). The '1% culturability paradigm' needs to be carefully defined. *ISME J.* 14, 10–11. doi: 10.1038/s41396-019-0507-8

Mehbub, F. M., Lei, J., Franco, C., and Zhang, W. (2014). Marine sponge derived natural products between 2001 and 2010: Trends and opportunities for discovery of bioactives. *Mar. Drugs* 12, 4539–4577. doi: 10.3390/md12084539

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

Schmidt, E. W., Obraztsova, A. Y., Davidson, S. K., Faulkner, D. J., and Haygood, M. G. (2000). Identification of the antifungal peptide-containing symbiont of the marine sponge Theonella swinhoei as a novel δ-proteobacterium, "Candidatus Entotheonella palauensis". *Mar. Biol.* 136, 969–977. doi: 10.1007/s002270000273

Selvin, J., Gandhimathi, R., Kiran, G. S., Priya, S. S., Ravji, T. R., and Hema, T. A. (2009). Culturable heterotrophic bacteria from the marine sponge *Dendrilla nigra*: isolation and phylogenetic diversity of actinobacteria. *Helgol. Mar. Res.* 63, 239–247. doi: 10.1007/s10152-009-0153-z

Sipkema, D., Holmes, B., Nichols, S. A., and Blanch, H. W. (2009). Biological characterisation of Haliclona (?gellius) sp.: sponge and associated microorganisms. *Microbial. Ecol.* 58, 903–920. doi: 10.1007/s00248-009-9534-8

Sipkema, D., Schippers, K., Maalcke, W. J., Yang, Y., Salim, S., and Blanch, H. W. (2011). Multiple approaches to enhance the cultivability of bacteria associated with the marine sponge Haliclona (gellius) sp. *Appl. Environ. Microbiol.* 77, 2130–2140. doi: 10.1128/AEM.01203-10

Staley, J. T., and Konopka, A. (1985). Measurement of *in situ* activities of nonphotosynthetic microorganisms in aquatic and terrestrial habitats. *Annu. Rev. Microbiol.* 39, 321–346. doi: 10.1146/annurev.mi.39.100185.001541

Steen, A. D., Crits-Christoph, A., Carini, P., DeAngelis, K. M., Fierer, N., Lloyd, K. G., et al. (2019). High proportions of bacteria and archaea across most biomes remain uncultured. *ISME J.* 13, 3126–3130. doi: 10.1038/s41396-019-0484-y

Steinert, G., Whitfield, S., Taylor, M. W., Thoms, C., and Schupp, P. J. (2014). Application of diffusion growth chambers for the cultivation of marine sponge-associated bacteria. *Mar. Biotechnol.* 16, 594–603. doi: 10.1007/s10126-014-9575-y

Stewart, E. J. (2012). Growing unculturable bacteria. *J. Bacteriol.* 194, 4151–4160. doi: 10.1128/JB.00345-12

Tamaki, H., Sekiguchi, Y., Hanada, S., Nakamura, K., Nomura, N., Matsumura, M., et al. (2005). Comparative analysis of bacterial diversity in freshwater sediment of a shallow eutrophic lake by molecular and improved cultivation-based techniques. *Appl. Environ. Microbiol.* 71, 2162–2169. doi: 10.1128/AEM.71.4.2162-2169.2005

Tanaka, T., Kawasaki, K., Daimon, S., Kitagawa, W., Yamamoto, K., Tamaki, H., et al. (2014). A hidden pitfall in the preparation of agar media undermines microorganism cultivability. *Appl. Environ. Microbiol.* 80, 7659–7666. doi: 10.1128/AEM.02741-14

Vartoukian, S. R., Adamowska, A., Lawlor, M., Moazzez, R., Dewhirst, F. E., and Wade, W. G. (2016). *In vitro* cultivation of "Unculturable" oral bacteria, facilitated by community culture and media supplementation with siderophores. *PloS One* 11, e0146926. doi: 10.1371/journal.pone.0146926

Wang, G. (2006). Diversity and biotechnological potential of the sponge-associated microbial consortia. *J. Ind. Microbiol. Biotechnol.* 33, 545–551. doi: 10.1007/s10295-006-0123-2

Weitzman, C. L., Gibb, K., and Christian, K. (2018). Skin bacterial diversity is higher on lizards than sympatric frogs in tropical Australia. *Peer J.* 6, e5960. doi: 10.7717/peerj.5960

frontiers | Frontiers in Microbiology

TYPE Original Research
PUBLISHED 26 September 2022
DOI 10.3389/fmicb.2022.969799



OPEN ACCESS

EDITED BY
Frederic Coulon,
Cranfield University, United Kingdom

REVIEWED BY
Shouliang Huo,
Chinese Research Academy of Environmental Sciences, China
Haihan Zhang,
Xi'an University of Architecture and Technology, China
Lijuan Ren,
Jinan University, China
Qingyun Yan,
Sun Yat-sen University, China
Hera Karayanni,
University of Ioannina, Greece

*CORRESPONDENCE
Yuanyuan Xue
yyxue@iue.ac.cn
Jun Yang
jyang@iue.ac.cn

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

RECEIVED 15 June 2022
ACCEPTED 26 August 2022
PUBLISHED 26 September 2022

CITATION
Ma G, Logares R, Xue Y and Yang J (2022) Does filter pore size introduce bias in DNA sequence-based plankton community studies? *Front. Microbiol.* 13:969799. doi: 10.3389/fmicb.2022.969799



# Does filter pore size introduce bias in DNA sequence-based plankton community studies?

Guolin Ma[1,2], Ramiro Logares[3], Yuanyuan Xue[1,4]* and Jun Yang[1,4]*

[1]Aquatic EcoHealth Group, Fujian Key Laboratory of Watershed Ecology, Ningbo Observation and Research Station, Key Laboratory of Urban Environment and Health, Institute of Urban Environment, Chinese Academy of Sciences, Xiamen, China, [2]College of Life Science, Fujian Agriculture and Forestry University, Fuzhou, China, [3]Institute of Marine Sciences (ICM), Spanish National Research Council (CSIC), Barcelona, Spain, [4]Zhejiang Key Laboratory of Urban Environmental Processes and Pollution Control, CAS Haixi Industrial Technology Innovation Center in Beilun, Ningbo, China

The cell size of microbial eukaryotic plankton normally ranges from 0.2 to 200 μm. During the past decade, high-throughput sequencing of DNA has been revolutionizing their study on an unprecedented scale. Nonetheless, it is currently unclear whether we can accurately, effectively, and quantitatively depict the microbial eukaryotic plankton community using size-fractionated filtration combined with environmental DNA (eDNA) molecular methods. Here we assessed the microbial eukaryotic plankton communities with two filtering strategies from two subtropical reservoirs, that is one-step filtration (0.2–200 μm) and size-fractionated filtration (0.2–3 and 3–200 μm). The difference of 18S rRNA gene copy abundance between the two filtering treatments was less than 50% of the 0.2–200 μm microbial eukaryotic community for 95% of the total samples. Although the microbial eukaryotic plankton communities within the 0.2–200 μm and the 0.2–3 and 3–200 μm size fractions had approximately identical 18S rRNA gene copies, there were significant differences in their community composition. Furthermore, our results demonstrate that the systemic bias introduced by size-fractionation filtration has more influence on unique OTUs than shared OTUs, and the significant differences in abundance between the two eukaryotic plankton communities largely occurred in low-abundance OTUs in specific seasons. This work provides new insights into the use of size-fractionation in molecular studies of microbial eukaryotes populating the plankton.

KEYWORDS

**eukaryotic plankton, environmental DNA, size-fractionated filtering, gene abundance, community composition, subtropical reservoir**

## Introduction

Microbial eukaryotes are key components of aquatic ecosystems, where they can improve the water quality by naturally controlling the nutrient flux and energy transfer in aquatic communities (Cotner and Biddanda, 2002). As useful indicator organisms for assessing the ecosystem status, microbial eukaryotic plankton have been studied in diverse aquatic ecosystems (David et al., 2021; Yang J. et al., 2021). The discovery of many microbial eukaryotes has been aided by advances in instrumentation (e.g., flow cytometer), improved culturing techniques, and the development of molecular assays (Caron and Hu, 2019; Djurhuus et al., 2020). The continuous development and improvement of high-throughput sequencing (HTS) platforms have stimulated interest in the study of complex microbial communities (Jo et al., 2020; Garner et al., 2021). Currently, the most popular approach to studying microbial eukaryotic plankton community composition and dynamics is targeting the small subunit ribosomal RNA (SSU rRNA) gene (e.g., 18S rRNA gene) and sequencing it using HTS methods. The produced sequence data can provide an estimate of the relative abundance of each taxon in each sample. Environmental DNA (eDNA) metabarcoding typically associates HTS sequences with their taxonomy (Liu et al., 2020; Yang J. et al., 2021), allowing the characterization and biomonitoring of complex microbial communities. The eDNA approach is more convenient and comprehensive than conventional microscopy methods for analyzing microbial plankton across spatial and temporal scales, and therefore, for obtaining more comprehensive profiles of freshwater plankton communities (Banerji et al., 2018; Yang and Zhang, 2020). So far, many studies have used eDNA in the investigation of freshwater and marine microbial communities (Simon et al., 2015; Woodhouse et al., 2016; Xue et al., 2018; Lin et al., 2021; Mo et al., 2021).

There are inherent limitations in comparing HTS-generated relative abundances of taxa when analyzing microbiomes due to data compositionality (Gloor et al., 2017; Knight et al., 2018; Morton et al., 2019). The reason is that the increase in one taxon abundance causes an equivalent decrease across the remaining taxa, and it is hard to fully capture how individual microbial taxa differ among samples (Barlow et al., 2020; Yang C. Y. et al., 2021). Thus, in HTS-based analyses, the relative abundance of a taxon is dependent on the abundances of all other taxa, which can lead to high false-positive rates in taxon analyses (Weiss et al., 2017) and negative correlation biases in correlation-based analyses (Tsilimigras and Fodor, 2016). Apart from HTS, the real-time quantitative polymerase chain reaction (qPCR) is one of the most widely used methods of gene quantitation, which is also effective for the detection and quantification of specific eukaryotes from complex natural communities (Zhu et al., 2005). Characterizations of protistan communities based on HTS and qPCR can be affected by the varying rDNA copy numbers (Gong and Marchetti, 2019; Lavrinienko et al., 2020; Sandin et al., 2022), which is also a key trait in protists. All in all, these molecular methods have allowed a better characterization of the diversity of microbial eukaryotic plankton communities and extended our ability to describe them to an unprecedented resolution (Piwosz et al., 2020).

Cell size has been defined as a "master trait," as it is shared by microorganisms across a given taxonomic group (Litchman et al., 2013). Eukaryotic planktonic microbes with different cell sizes might have fundamentally different characteristics and ecological roles in ecosystems (Clarke and Deagle, 2020). Molecular studies of microbial eukaryotic diversity have often used size-fractionated samples to separate unicellular eukaryotes from multicellular microorganisms, as well as to separate microbes with different cell sizes (de Vargas et al., 2015; Liu et al., 2017). Recent studies have revealed the community composition and spatiotemporal dynamic of eukaryotic plankton populating different size fractions (Woodhouse et al., 2016; Giner et al., 2019; Lin et al., 2021).

Lakes and reservoirs are excellent ecosystems for investigating microbial eukaryotic plankton (Zhang et al., 2020). In deep lakes, thermal stratification often facilitates chemical differences along the water column, resulting in the adaptation of microorganisms to different water layers (Boehrer and Schultze, 2008; Yu et al., 2014). In this study, we analyzed 18S rRNA gene abundances in size-fractionated plankton communities by qPCR and HTS. We collected plankton communities from five water layers in two subtropical reservoirs across four seasons. Our research focuses on the following hypotheses: (1) the 18S rRNA gene copies of the 0.2–200 μm size-fractionated eukaryotic plankton community is roughly equal to the sum of 0.2–3 and 3–200 μm fractions; (2) the community composition of the 0.2–200 μm size-fractionated eukaryotic plankton is roughly identical to the sum of the 0.2–3 and 3–200 μm fractions. We analyzed the abundance and composition differences between the two filtering treatments (Treatment 1: 0.2–200 μm and Treatment 2: 0.2–3 and 3–200 μm). Our results can contribute to understanding potential biases introduced by size-fractionation in studies of microbial eukaryotes.

## Materials and methods

### Sampling sites and filtration strategies

Samples were collected from Shidou Reservoir (24°42′N, 118°00′E) and Tingxi Reservoir (24°48′N, 118°08′E), located in Xiamen, southeast China (**Figure 1A**). Both reservoirs are used for water supply, aquaculture, and irrigation. Details of these two reservoirs were described in our previous studies (Yang et al., 2016; Gao et al., 2022). Water samples were taken in January, April, July, and October during 2018. Stratification

appeared based on water temperature and dissolved oxygen concentration in both the Shidou and Tingxi reservoirs during the observation term, while the mixing period occurred in January (Figure 1C). According to the water stratification conditions, five sampling depths were selected for plankton collection, represented by layer A (0.5 m depth), layer B (the average depth between layers A and C), layer C (thermal and hypoxic boundary depth), layer D (the average depth between layers C and E), and layer E (2 m above the bottom sediments), respectively. Water temperature and dissolved oxygen were measured *in situ* with a multi-parameter water quality analyzer (Hydrolab DS5, Hach Company, Loveland, CO, United States). After sampling, water samples were transported to the laboratory quickly. A 200 μm pore size nylon mesh was used to remove macroplankton and other large particles before water filtering (Liu et al., 2017). Next, a part of this water was filtered through 0.22 μm pore size polycarbonate membranes (47 mm diameter, Millipore, Bedford, MA, United States)

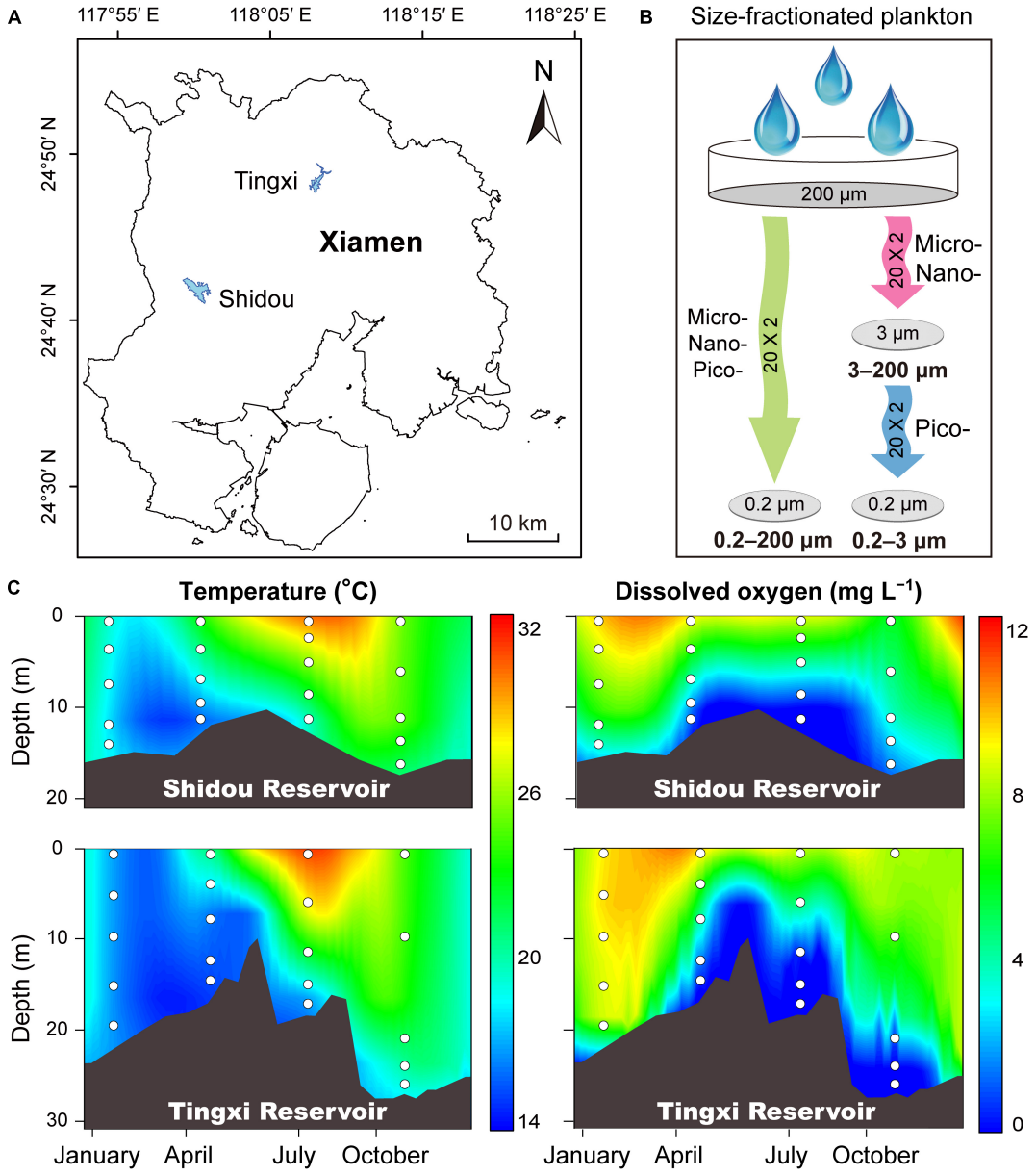



FIGURE 1
Sampling sites of the Shidou and Tingxi reservoirs in Xiamen city **(A)**, and size-fractionated filtering strategies **(B)**. Micro-, nano-, and pico-eukaryotic plankton: 0.2–200 μm; Micro- and nano-eukaryotic plankton: 3–200 μm; Pico-eukaryotic plankton: 0.2–3 μm. Depth-time profiles of water temperature and dissolved oxygen in the Shidou and Tingxi reservoirs, respectively **(C)**. White dots indicate the sampling depth.

to collect pico-, nano-, and micro-eukaryotic plankton cells within the 0.2–200 μm size fraction. The remaining water was filtered firstly through 3 μm pore size polycarbonate membranes (47 mm diameter, Millipore, Bedford, MA, United States) to collect nano- and micro-eukaryotic plankton cells within the 3–200 μm size fraction, and then filtered sequentially through 0.22 μm pore size polycarbonate membranes to obtain pico-eukaryotic plankton cells within the 0.2–3 μm size fraction (**Figure 1B**). The filtration volume ranged from 300 to 900 mL. Specifically, at the beginning of filtering, 300 ml of water sample was added to each filter membrane firstly and recorded the filtering time until completely filtered. If the time exceeded 30 min, we finished the collection of the filter membrane sample; otherwise, added another 50 ml of water sample successively until the filtering time reached 30 min, and the longest filtering time did not exceed 60 min. All these procedures were designed to ensure that each membrane reaches a critical load to collect enough plankton. A total of 120 filtered samples were acquired and the membranes were stored at −80°C until DNA extraction.

## DNA extraction and real-time quantitative PCR

The DNA of microbial eukaryotic plankton communities was extracted directly from the membranes using the FastDNA SPIN Kit (MP Biomedicals, Solon, OH, United States) according to the manufacturer's instructions. The eukaryotic plankton was quantified by targeting the 18S rRNA gene. A pair of universal primers, 1380F (5′-CCCTGCCHTTTGTACACAC-3′) and 1510R (5′-CCTTCYGCAGGTTCACCTAC-3′) were used to amplify the eukaryotic 18S rDNA V9 region (Amaral-Zettler et al., 2009). Each 20 μL reaction mixture contained 10 μL of 2 × *PerfectStart*™ Green qPCR SuperMix, 0.5 μM of forward and reverse primers, 1 μL of DNA template and 8 μL RNase-free water. Negative controls contained the same mixtures with 1 μL of sterile water instead of the DNA template. Each plankton DNA sample was amplified in triplicates for 40 cycles of 30 s at 94°C, 15 s at 60°C, and 10 s at 72°C using an ABI Q6 Real-Time PCR System (Life Technologies, Applied Biosystems, Foster City, CA, United States). The DNA standards were prepared and run during each qPCR reaction to generate standard curves ($r^2 > 0.99$). All qPCR runs yielded amplification efficiencies between 95% and 105%. The cycle number at which the fluorescence signal crosses a certain threshold (threshold cycle [Ct]) was noted. This Ct value is proportional to the logarithm of the target DNA concentration in the assay. From a dilution series of DNA amount corresponding to a known DNA concentration, a standard curve was produced. Combined with qPCR results, we calculated the 18S rRNA gene copies per liter of water and called this the 18S rRNA gene absolute abundance. The 18S rRNA gene absolute abundance of 0.2–3 and 3–200 μm size-fractionated plankton community is expressed as the sum of the 18S rRNA gene copies within 0.2–3 μm and 3–200 μm plankton communities.

## Metabarcoding, sequencing and bioinformatics

The primer pair 1380F and 1510R with barcodes were used to amplify the V9 region of the eukaryotic 18S rRNA gene (Hadziavdic et al., 2014; Xue et al., 2018). Each plankton DNA sample and negative controls were run in triplicates. The PCR reaction mixture contained 10 μL of Phusion High-Fidelity PCR Master Mix (New England Biolabs, Beverly, MA, United States), 0.5 μM of each primer, 1 μL of DNA template and 8 μL RNase-free water. The reactions included an initial denaturation at 95°C for 5 min, followed by 30 cycles of 30 s at 95°C, 30 s at 55°C, and 30 s at 72°C. At the end of the amplification, the amplicons were subjected to a final 5 min extension at 72°C. The triplicate PCR products for each of 120 samples were mixed in equimolar amounts and were confirmed after running in 2% agarose gel, and then isolated and purified using GeneJET Gel Extraction Kit (Thermo Fisher Scientific, Waltham, MA, United States). Sequencing libraries were generated using the NEB Next Ultra DNA Library Prep Kit for Illumina (New England Biolabs, Beverly, MA, United States) according to the manufacturer's instructions. The library quality was evaluated using a Qubit 4.0 Fluorometer (Thermo Fisher Scientific, Waltham, MA, United States). Finally, the library was sequenced on the Illumina X Ten platform (Illumina Inc., San Diego, CA, United States) using a 150 bp paired-end protocol (Mo et al., 2021).

The paired-end V9 region of 18S rRNA gene sequences was processed using VSEARCH v.2.14.1 (Rognes et al., 2016). Chimeras were removed with default settings, and the unoise3 algorithm was used to identify operational taxonomic units (OTUs) at the 97% sequence similarity threshold (Nearing et al., 2018). Representative sequences were assigned taxonomy using the sintax algorithm with a cutoff value of 0.8 against the protist ribosomal reference database (PR2) in VSEARCH (Guillou et al., 2013). To minimize the inclusion of sequencing errors, the OTUs with less than 10 reads were excluded from the dataset before performing downstream analyses (Liu et al., 2017). The final total dataset retained 16,472 OTUs. We used a randomly selected subset of 209,120 sequences from each sample to normalize sequencing effort across samples based on MOTHUR v.1.39.5 (Schloss et al., 2009).

## Statistical analyses

The non-parametric Mann–Whitney *U* test was used to compare the difference in eukaryotic 18S rRNA gene copy abundance between the 0.2–200 μm and 0.2–3 and 3–200 μm

eukaryotic plankton communities using SPSS v22.0 (IBM Corp., Armonk, NY, United States). The analysis of similarities (ANOSIM) was used to evaluate the differences between groups, with the global R representing the separation degree of between-group and within-group mean rank similarities. $R = 0$ indicates no separation, whereas $R = 1$ indicates complete separation (Clarke and Gorley, 2015). A non-metric multidimensional scaling (NMDS) ordination was used to investigate differences in microbial eukaryotic communities within different size fractions in R, using the "vegan" package (R Core Team, 2020). To characterize the beta-diversity of microbial eukaryotic plankton communities, we constructed Bray-Curtis dissimilarity matrices based on the Hellinger transformed read number. To compare the OTUs number among different communities, we constructed a Venn diagram using the "Venn-Diagram" package in R (R Core Team, 2020).

# Results

## Differences in 18S rRNA gene abundance between two filtering treatments

There were 14,739 and 14,540 plankton OTUs retrieved from the Shidou and Tingxi reservoirs, respectively (**Supplementary Figure S1**). The abundance of 18S rRNA gene for 69.58% (11462/16472), 92.73% (13667/14739) and 90.66% (13182/14540) OTUs was not significantly different between the two filtering treatments (0.2–200 μm *vs.* 0.2–3 and 3–200 μm) in both reservoirs, Shidou Reservoir and Tingxi Reservoir, respectively (**Figure 2**). Although the mean abundance in 30.42% OTUs (5010/16472) had a significant difference between

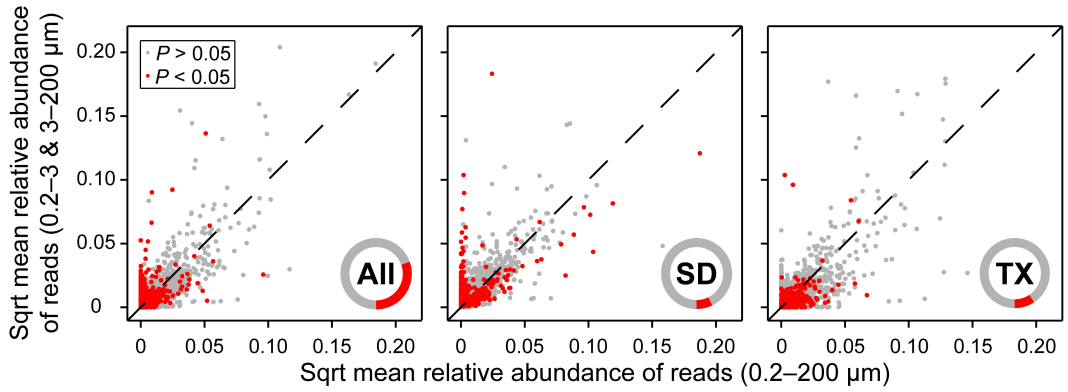



FIGURE 2
Comparison of relative abundance of OTUs between 0.2–200 μm and 0.2–3 and 3–200 μm eukaryotic plankton fractions. More than two-thirds of OTUs (69.58%) showed no significant difference in the relative abundance of reads between the two size-fractionated plankton communities in both reservoirs. The mean relative abundance of reads was square root transformation (Sqrt). Statistical analysis is the non-parametric Mann–Whitney *U* test. Data are presented as the mean of 40 replicates for 16472 OTUs in both reservoirs, 20 replicates for 14739 OTUs in the Shidou Reservoir and 20 replicates for 14540 OTUs in the Tingxi Reservoir, respectively. The operational taxonomic units (OTUs) were defined at 97% sequence similarity threshold. Only 5010 (30.42%), 1072 (7.27%), and 1358 (9.34%) OTUs were significant differences between the two size-fractionated plankton communities in both reservoirs (All), Shidou Reservoir (SD) and Tingxi Reservoir (TX), respectively. The inner red/gray cycles show the composition proportion with/without statistical differences, which more intuitively reflects the proportion with significant differences between the 0.2–200 μm and 0.2–3 and 3–200 μm size-fractionated eukaryotic plankton.

**TABLE 1** The OTUs that are different ($P < 0.05$) and non-significantly different ($P \geq 0.05$) between the 0.2–200 μm and 0.2–3 and 3–200 μm eukaryotic plankton communities in both reservoirs, Shidou Reservoir and Tingxi reservoir, respectively.

| OTUs relative abundance | Shidou and Tingxi reservoirs | | Shidou reservoir | | Tingxi reservoir | |
|---|---|---|---|---|---|---|
| | $P < 0.05$ (5,010 OTUs) | $P \geq 0.05$ (11,462 OTUs) | $P < 0.05$ (1,072 OTUs) | $P \geq 0.05$ (13,667 OTUs) | $P < 0.05$ (1,358 OTUs) | $P \geq 0.05$ (13,182 OTUs) |
| ≥1% | 1 (0.01%) | 14 (0.08%) | 5 (0.03%) | 6 (0.04%) | 0 (0%) | 16 (0.11%) |
| 0.1% ~ 1% | 13 (0.08%) | 108 (0.65%) | 24 (0.16%) | 110 (0.75%) | 7 (0.05%) | 100 (0.69%) |
| 0.01% ~ 0.1% | 123 (0.75%) | 497 (3.02%) | 130 (0.88%) | 424 (2.88%) | 74 (0.51%) | 508 (3.49%) |
| 0.001% ~ 0.01% | 659 (4.00%) | 1,387 (8.42%) | 403 (2.74%) | 1,408 (9.55%) | 353 (2.43%) | 1,463 (10.06%) |
| <0.001% | 4,214 (25.58%) | 9,456 (57.41%) | 510 (3.46%) | 11,719 (79.51%) | 924 (6.35%) | 11,095 (76.31%) |

The operational taxonomic units (OTUs) were defined at 97% sequence similarity threshold. The percentage of OTU number is given in parentheses.

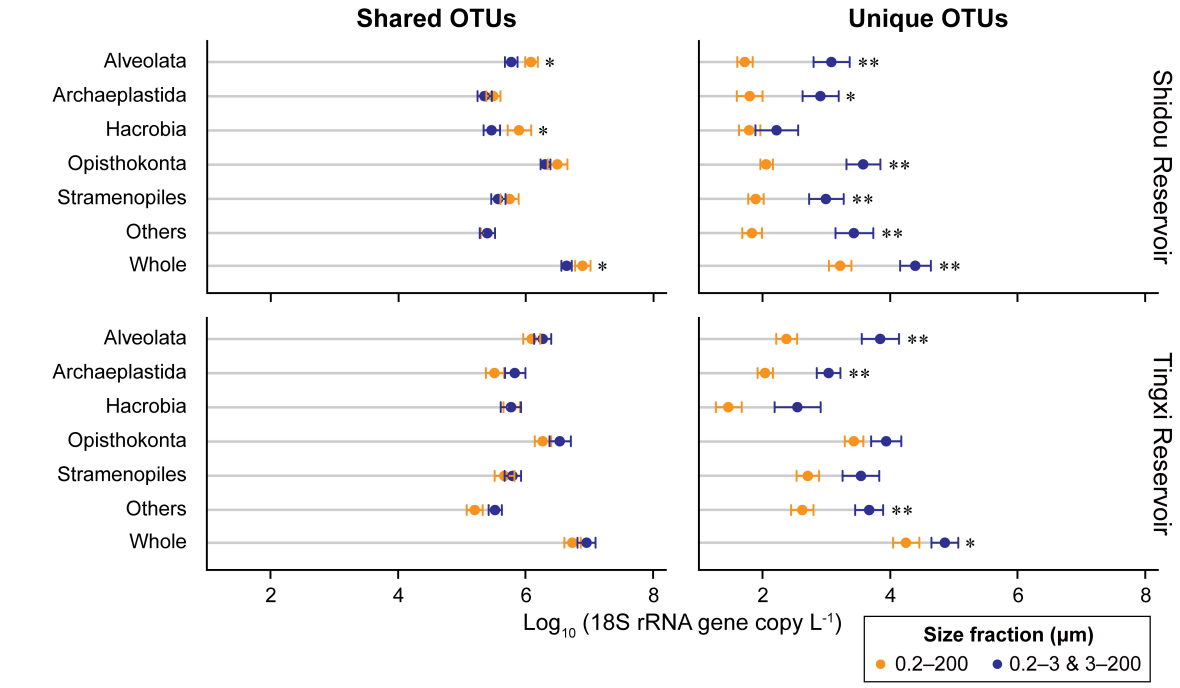



FIGURE 3
Comparison of 18S rRNA gene copy abundance of different supergroups between two size-fractionated plankton communities. The OTUs were defined at 97% sequence similarity threshold. Statistical analysis is non-parametric Mann–Whitney *U* test (*$P < 0.05$, **$P < 0.01$). The error bar represents the standard error of OTU occurrences in all samples. Others includes Amoebozoa, Apusozoa, Excavata, and Rhizaria.

the two size-fractionated eukaryotic plankton communities, all these OTUs only represented < 5% of the total reads (Figure 2), and most of them occurred as low-abundance OTUs (Table 1). More than half of the OTUs were shared (ca. 67.48% of the total OTUs in the Shidou Reservoir and 59.92% of those in the Tingxi Reservoir) (Supplementary Figure S1). The shared OTUs included the OTUs shared between the eukaryotic plankton communities with size fractions of 0.2–3 μm and 0.2–200 μm, 3–200 μm and 0.2–200 μm, also the OTUs shared among 0.2–3 μm, 3–200 μm and 0.2–200 μm size-fractionated eukaryotic communities. However, the systemic bias induced by different filtrations was more significant in the unique OTUs than in shared OTUs (Figure 3). Specifically, for the unique OTUs, the 18S rRNA gene abundance exhibited higher values in 0.2–3 and 3–200 μm size fractions than in the 0.2–200 μm size fraction (Figure 3). In the Shidou Reservoir, there were about 67.48% (9946/14739) shared OTUs for the two size-fractionated eukaryotic plankton communities; while 91.24% (9946/10901) for the 0.2–200 μm size fraction and 72.16% (9946/13784) for the 0.2–3 and 3–200 μm size fraction, respectively (Table 2 and Supplementary Figure S1). Tingxi Reservoir exhibited similar results for shared OTUs. More importantly, these shared OTUs contributed more than 95% of the reads in the two reservoirs for each size-fractionation treatment (Table 2).

The absolute abundance of 18S rRNA gene copies in the three different size-fractionated planktonic communities (0.2–200 μm, 0.2–3 μm, and 3–200 μm) varied from $1.6 \times 10^5$ to $4.4 \times 10^7$ copies $L^{-1}$ in the Shidou Reservoir, and $2.0 \times 10^5$ to $4.9 \times 10^7$ copies $L^{-1}$ in the Tingxi Reservoir, respectively (Supplementary Figure S2). The two filtering treatments (0.2–200 μm *vs.* 0.2–3 and 3–200 μm) had roughly the same number

TABLE 2 The OTUs and sequences that are shared and unique between the 0.2–200 μm ($n = 20$) and 0.2–3 and 3–200 μm ($n = 20$) eukaryotic plankton communities from Shidou and Tingxi reservoirs.

| | 0.2–200 μm (shared) | 0.2–200 μm (unique) | 0.2–3 and 3–200 μm (shared) | 0.2–3 and 3–200 μm (unique) |
|---|---|---|---|---|
| **Shidou** | | | | |
| OTU | 9,946 (91.24%) | 955 (8.76%) | 9,946 (72.16%) | 3,838 (27.84%) |
| Sequence | 262,448,030 (99.95%) | 144,048 (0.05%) | 122,689,638 (95.79%) | 5,397,936 (4.21%) |
| **Tingxi** | | | | |
| OTU | 8,713 (82.03%) | 1,909 (17.97%) | 8,713 (68.98%) | 3,918 (31.02%) |
| Sequence | 216,087,304 (98.38%) | 3,563,669 (1.62%) | 376,846,508 (95.15%) | 19,191,829 (4.85%) |

The operational taxonomic units (OTUs) were defined at 97% sequence similarity threshold. The OTUs with ≤ 10 sequences were removed. The relative contributions (percentages) of shared or unique OTUs to total OTUs in each size fraction are given in parentheses.

of 18S rRNA gene copies in the Shidou Reservoir in January, April and October; and in the Tingxi Reservoir in January and July (Supplementary Figure S2 and Figure 4A). In the Shidou Reservoir, 40% of samples exhibited a higher 18S rRNA gene abundance in the 0.2–200 μm size fraction plankton community, and 55% of that were no significant difference

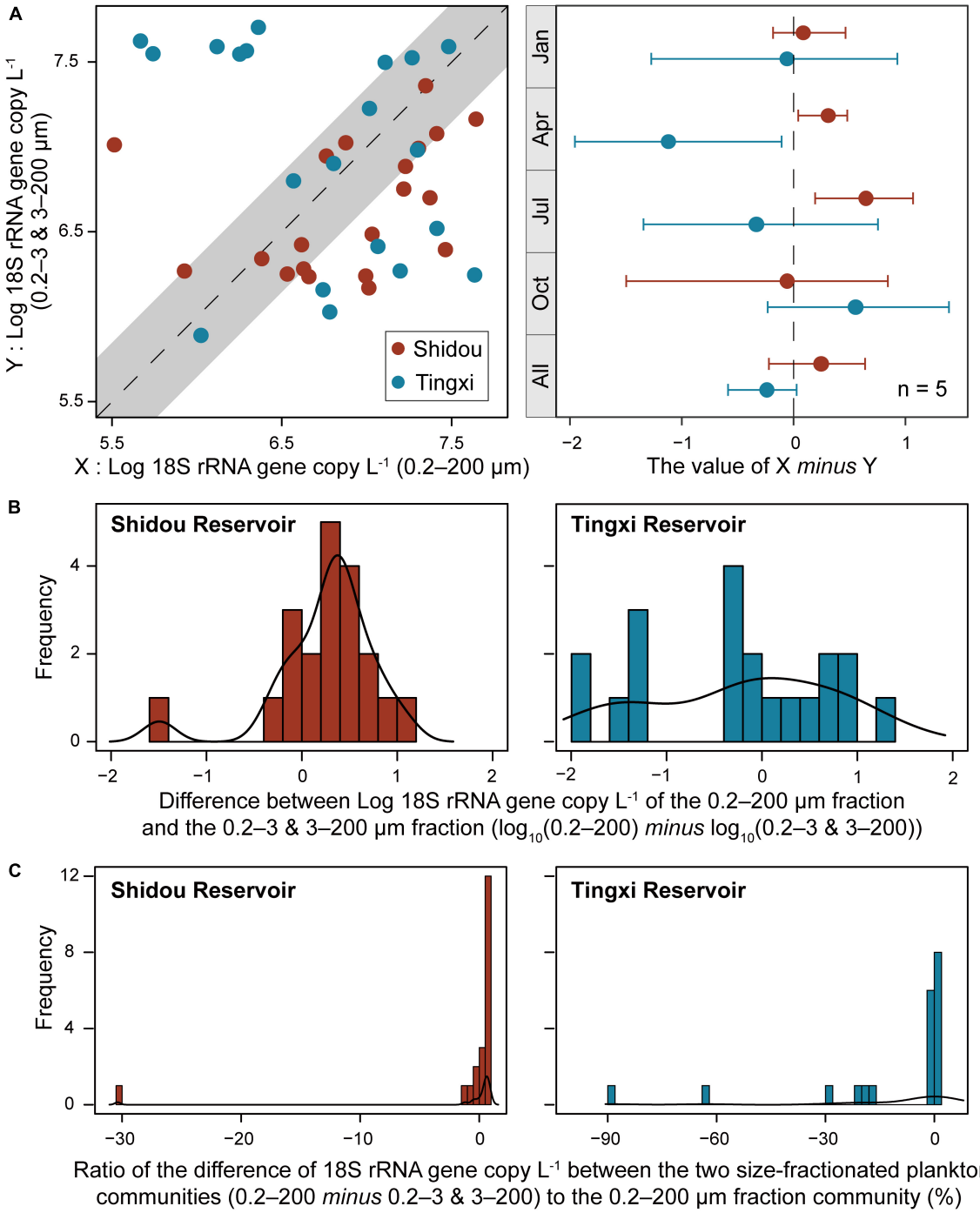



FIGURE 4
Comparison of the 18S rRNA gene abundance between the two size-fractionated plankton communities (with 95% confidence interval) and their seasonal difference (the error bar represents the standard error along with five water layers, *n* = 5) **(A)**. Histogram showing the distribution frequency of log10 18S rRNA gene copy of the 0.2–200 μm minus 0.2–3 and 3–200 μm size fractions **(B)**, and the ratio of the difference to 0.2–200 μm fraction **(C)**. The difference of 18S rRNA gene abundance between the two size-fractionated plankton communities was less than 50% of the 0.2–200 μm microbial eukaryotic community for 95% of the total samples.

between the two filtering treatments (0.2–200 μm *vs.* 0.2–3 and 3–200 μm). In the Tingxi Reservoir, 40% of samples had roughly the same 18S rRNA gene copy abundance, and 30% of samples exhibited a higher abundance in the 0.2–200 μm compared to 0.2–3 and 3–200 μm size-fractionated community (**Figure 4A**). Moreover, between the two size-fractionated plankton communities, the difference of 18S rRNA gene copies for 95% of the total samples was less than 50% of the 0.2–200 μm microbial eukaryotic community (**Figure 4B**), indicating the 18S rRNA gene copies of the two filtering treatments is roughly same across space and time. In addition, the significant difference of 18S rRNA gene abundance among different eukaryotic supergroups between the two filtering treatments only occurred in specific groups in specific seasons (i.e., July in the Shidou Reservoir, and April in the Tingxi Reservoir) (**Figure 5**).

## Differences in plankton community composition between two filtering treatments

The three size-fractionated eukaryotic plankton contributions to community were different at each supergroup level (**Supplementary Figure S3**). Both the Shidou and Tingxi reservoirs are characterized by plenty of OTUs belonging to 0.2–200 μm fraction plankton in almost all the supergroups (except for Apusozoa that with a lower OTUs richness and 18S rRNA gene abundance). Compared with the Tingxi Reservoir, only a small number of OTUs were found enriched in the Shidou Reservoir plankton within the 3–200 μm fraction but a large proportion of the OTUs harbored within the 0.2–3 μm fraction (**Supplementary Figure S3**). This result indicates that the Shidou Reservoir harbored a higher percentage of smaller eukaryotic plankton, while the larger individuals were enriched in the Tingxi Reservoir.

In both Shidou and Tingxi reservoirs, the two filtering treatments (0.2–200 μm *vs.* 0.2–3 and 3–200 μm) exhibited fewer variations in richness (OTUs) than in abundance (18S rRNA gene copies) (**Supplementary Table S1**), and the absolute abundance of the 18S rRNA gene fluctuated greatly among the five water layers and across four seasons (**Supplementary Figures S2, S4**). Additionally, Opisthokonta was the most abundant taxon compared with the other supergroups; it contributed 49.36% and 32.72% of the sequences in the 0.2–200 μm plankton community, while it contributed 43.67 and 46.20% of the sequences in the 0.2–3 and 3–200 μm size fractions in Shidou and Tingxi reservoirs, respectively (**Supplementary Table S1** and **Supplementary Figure S5**). Alveolata was the second most abundant supergroup and represented a mean of 13.99% of the relative abundances in the Shidou Reservoir (0.2–200 μm: 13.51%; 0.2–3 and 3–200 μm: 14.46%) and 21.80% of that in the Tingxi Reservoir (0.2–200 μm: 23.53%; 0.2–3 and 3–200 μm: 20.06%) (**Supplementary Table S1** and **Supplementary Figure S5**).

The non-metric multidimensional scaling (NMDS) analysis showed that the eukaryotic plankton community composition was different between the 0.2–200 μm and the 0.2–3 and 3–200 μm fractions (**Figure 6A**), especially in January and July

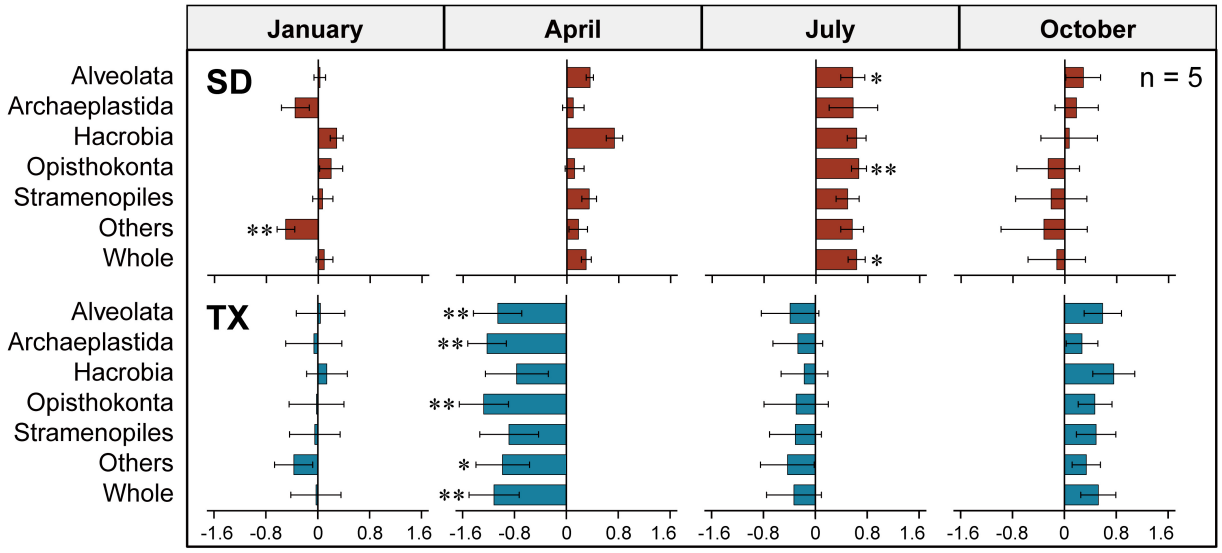



FIGURE 5
Differences of the 18S rRNA gene copy abundance between the two size-fractioned plankton communities at the supergroup level in Shidou (SD) and Tingxi (TX) reservoirs across four seasons, respectively. Size fractions: 0.2–200 μm *vs.* 0.2–3 and 3–200 μm. The data were log10-transformed. The error bar represents the standard error ($n$ = 5). Statistical analysis is non-parametric Mann–Whitney $U$ test (*$P$ < 0.05, **$P$ < 0.01). Others includes Amoebozoa, Apusozoa, Excavata, and Rhizaria.

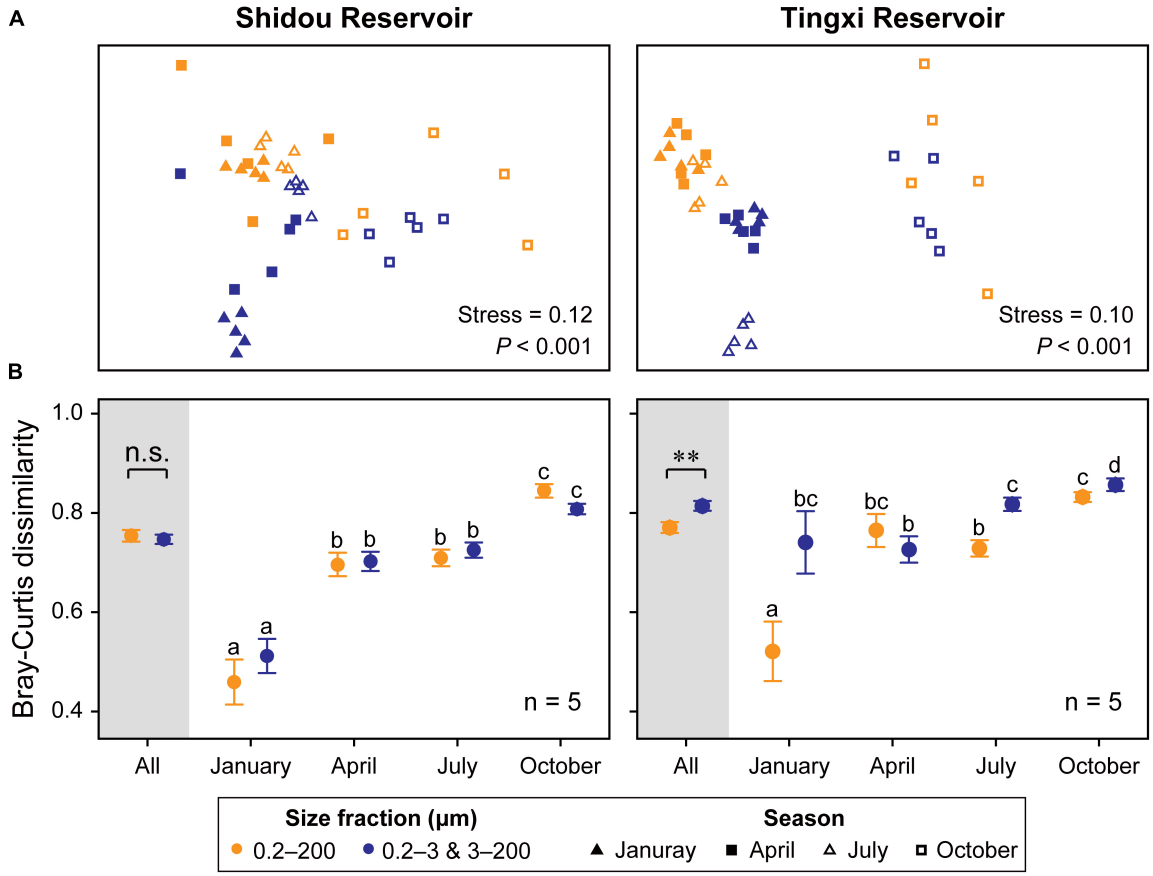



FIGURE 6
Non-metric multidimensional scaling (NMDS) ordination of eukaryotic plankton communities between two size fractions **(Color)** and across four seasons **(Shape)** in both reservoirs (Global R of size fractions is 0.150 and 0.214, while that of seasons is 0.514 and 0.545, respectively, in the Shidou and Tingxi reservoirs) **(A)**. Pairwise Bray–Curtis dissimilarity of different size-fractionated eukaryotic plankton communities. Data are mean ± standard error (SE), and the error bar represents the standard error along with five water layers (n = 5). Significant differences ($P < 0.05$) within-group are indicated with different letters by the non-parametric Mann–Whitney *U* test **(B)**. The small letters are used to distinguish whether there are statistical differences between the two filtering treatments (0.2–200 μm *vs.* 0.2–3 and 3–200 μm) in different seasons (January, April, July, and October). There is no significant difference between groups marked with the same letter, but there is a significant difference between groups marked with different letters.

in the Shidou Reservoir and April in the Tingxi Reservoir (Supplementary Table S2).

## Spatiotemporal dynamics of two size-fractionated eukaryotic plankton communities

Seasonality was evident in both reservoirs for the two filtering treatments (0.2–200 μm *vs.* 0.2–3 and 3–200 μm) (Figure 6A), and the results of ANOSIM further corroborated that the eukaryotic plankton communities could be significantly distinguished across most seasons (Shidou Reservoir: from 0.304 to 0.848; Tingxi Reservoir: from 0.199 to 0.979), with an exception for the comparison between January and April (Supplementary Table S3). The two size-fractionated eukaryotic plankton communities did not show any significant difference in β-diversity in Shidou Reservoir, however, there were significant differences in Tingxi Reservoir, especially in January, July, and October (Figure 6B). Both reservoirs had the smallest Bray–Curtis dissimilarity between all pairs of samples in January, indicating a more similar community composition in winter caused by water mixing (Figures 1C, 6B).

# Discussion

Many studies have used size-fractionation to collect plankton of different size classes and showed community compositional variations among them. Such size-fractionated strategies have expanded our capacity to describe and compare different plankton communities with various resolutions

(de Vargas et al., 2015; Liu et al., 2017; Lin et al., 2021). However, each methodological stage, from sampling to data analysis, can introduce biases; such biases can skew data sets by introducing changes in the observed relative abundance, and result in distorted observations of the true microbial composition within a sample (Costea et al., 2017; Yang C. Y. et al., 2021). Microbiome studies are particularly affected by these biases, especially in DNA sequence-based studies (Nearing et al., 2021). To determine the effects of different size-fractioning strategies on plankton community estimates, we assessed eukaryotic plankton abundance and community composition using DNA metabarcoding. Our results showed that microbial eukaryotic plankton communities between 0.2–200 μm and 0.2–3 and 3–200 μm size fractions had approximately identical 18S rRNA gene abundance in most cases, but there were significant differences in their composition. Besides, we found that the bias introduced by the two size-fractionating filtering strategies had more influence on unique OTUs than on shared OTUs, and the significant differences largely occurred in low-abundance taxonomic groups in specific seasons.

## Differences in 18S rRNA gene abundance between two filtering treatments

Our results showed that the 18S rRNA gene absolute abundance of the 0.2–200 μm eukaryotic plankton was roughly equal to the sum of 0.2–3 μm and 3–200 μm size fractions in most cases (**Figures 2**, **4B,C**). Remarkably, there was a coherence of 18S rRNA gene abundance between the 0.2–200 μm and the 0.2–3 and 3–200 μm size fractions in the Shidou Reservoir, in contrast to the Tingxi Reservoir. This might be a characteristic difference between the two reservoirs. Specifically, Shidou Reservoir had high algal biomass due to the long-term dominance of a bloom-forming and toxic cyanobacterium (*Raphidiopsis raciborskii*) from September 2017 to November 2018 (Gao et al., 2022). The positive correlation between cyanobacteria abundance and other bacterioplankton in the water column has been shown by Woodhouse et al. (2016). Although more than half of the OTUs were shared by the two eukaryotic plankton communities, which may be caused by the eukaryotic plankton cells lysis and release of DNA into the water (Dell'Anno and Danovaro, 2005; Nagler et al., 2018). There still were a few significant differences that occurred in specific seasons.

Collectively, there are two main reasons for the differences in the 18S rRNA gene between the two filtering treatments. First, filtration might cause the loss of some microorganisms. Specifically, in the size-fractionation filtering, the cell morphology of microorganisms changed due to the negative pressure (Cavalier-Smith, 1997; Cavalier-Smith and Chao, 2004; Logares et al., 2014). After that, they would pass through the 3 μm even 0.22 μm pore size membrane, resulting in the loss of cell abundance. Pore size relates to the filter's ability to filter out microbes with certain sizes. For example, a 0.22 μm membrane will filter out particles with a diameter of 0.22 microns or larger from the filtration stream. These biases might further lead to the underestimation of gene abundances in eukaryotic plankton communities. Second, microorganisms might accumulate on the membrane. As water filtering proceeded, a "filter cake layer" appeared on the pore size membranes, which is caused by larger microorganisms (Zheng et al., 2015). Accordingly, small cell plankton may be stuck on the filter membranes (Hasegawa et al., 2003). There were microbial accumulation on the filter membranes both in one-step and two-step filtering. However, microbes were retained more in the two-step filtration than in the one-step due to the longer (double) duration of water filtering. Furthermore, as the sequencing depth increased in the two-step filtration (0.2–3 and 3–200 μm) (**Supplementary Figure S6**), more rare OTUs with lower abundance will be obtained (Liu et al., 2017), which can lead to gene abundance differences between the two size-fractionated plankton communities in unique OTUs.

## Differences in plankton community composition between two filtering treatments

Our second hypothesis was that the community composition of the two size-fractionated eukaryotic plankton was identical. However, the community composition was significantly different between the 0.2–200 μm and the 0.2–3 and 3–200 μm size fractions, although more than half of the OTUs were shared by both fractions in Shidou and Tingxi reservoirs, respectively (**Table 2** and **Supplementary Figure S1**). This is probably because the negative pressure filtering would cause eukaryotic plankton cells to break, and their cell-free DNA would attach to the filter membranes. According to Morrison et al. (2017), water stratification influenced the dominant taxa of the plankton community. Similarly, the differential dominance of eukaryotic microorganisms in different water layers (**Supplementary Figure S4**), further leads to the differences in community composition between the two filtering treatments (0.2–200 μm and 0.2–3 and 3–200 μm). In addition, we infer that the traits of abundant taxa also differentiate the community compositions between two size-fractionated plankton, due to their higher OTU richness and 18S rRNA gene abundance. For example, the flattened inner membrane complex of Apicomplexa (Alveolata), which evolved in association with fusion, is an adaptation for penetrating the host and gliding motility (Cavalier-Smith and Chao, 2004). Also, Opisthokonta is characterized by a single posterior cilium in their unicellular motile stage and by (non-discoid)

flattened plate-like mitochondrial cristae (Cavalier-Smith, 1997). Their deformable morphological structure allows them to change their morphology and freely pass through the filter membranes.

As previous studies have shown, the apparent seasonal variations in microbial plankton communities have been shown in diverse ecosystems, including marine (Djurhuus et al., 2020) and freshwater environments (Simon et al., 2015; Salmaso et al., 2020). Most aquatic organisms have seasonal variation in community composition (de Vargas et al., 2015; Simon et al., 2015; Lin et al., 2021). In our study, the temporal dynamic of the microbial eukaryotes showed a strong seasonality at different taxonomic levels (**Supplementary Figure S4** and **Supplementary Table S3**). Besides, the unique OTUs showed more significant differences between the two filtering treatments (0.2–200 μm *vs.* 0.2–3 and 3–200 μm), yet their relative abundance was much lower than that of shared OTUs. Lynch and Neufeld (2015) have pointed out that rare subcommunities are vital to shaping microbial community structure. In addition, the "rare biosphere" may contribute to the majority of the total richness in microbial communities (Sogin et al., 2006; Nyirabuhoro et al., 2020). Therefore, compared with previous studies (Staley et al., 2013; Logares et al., 2014), our results have illustrated the differences in the composition of different size-fractionated eukaryotic plankton communities largely due to the unique OTUs, although their abundances were very low.

There are still some inherent limitations in our study. For example, the use of a short region of 18S rRNA gene instead of the whole gene tends to exclude some taxa (Maritz et al., 2019). Besides, the choice of universal primers targeting the 18S rRNA gene might result in inaccurate estimates of the eukaryotic plankton community (Engelbrektson et al., 2010; Harder et al., 2016). The rRNA gene copy number varies from one to thousands in eukaryotic genomes and has a greater interspecific and intraspecific variation than in prokaryotic genome (de Vargas et al., 2015; Harder et al., 2016; Lavrinienko et al., 2020), which may also lead to the overestimation of microbes that are present in a sample as well as their proportions (Louca et al., 2018). Moreover, the rRNA gene copy number variability of the prevalent species hinders the accurate translation of metabarcoding data into relative abundance and of qPCR data into their corresponding cell numbers (Lajeunesse and Thornhill, 2011). Fortunately, the rRNA gene copy number per genome is significantly and positively correlated with the plankton size or biovolume (de Vargas et al., 2015; Lavrinienko et al., 2020), therefore we can roughly estimate the plankton biomass based on the rRNA gene copy number. More noticeably, the effects of processing methods in protocols often exceed the biological effects underlining the importance of minimizing these biases (Sinha et al., 2017). Nevertheless, these potential sources of error have not been systematically examined in the development of approaches in microbial eukaryotic plankton studies.

In this study, we have compared the two eukaryotic plankton communities with sizes that overlapping and mutually contained (e.g., 0.2–200 μm *vs.* 0.2–3 and 3–200 μm), trying to illustrate the impact of size-fractionated filtering strategies on the plankton results. Here, we provide evidence that the commonly used size-fractionation particularly affected the low-abundance taxa of the eukaryotic plankton community in specific seasonal eDNA samples.

## Conclusion

We compared the 18S rRNA gene absolute abundance and the community composition of the two filtering treatments (0.2–200 μm *vs.* 0.2–3 and 3–200 μm) from two subtropical reservoirs using qPCR and HTS. Our results reflect the quantitative changes of eukaryotic plankton within two filtering treatments. Both size-fractionated plankton communities had approximately equal 18S rRNA gene abundance in most cases but exhibited dissimilar community composition across four seasons. This suggests that the systemic bias in unique OTUs introduced by different filtration strategies has influenced the consistency between the 0.2–200 μm and 0.2–3 and 3–200 μm size fractions, but the general seasonal patterns were highly similar. In addition, we demonstrated that the difference in temporal scale (seasons) considerably exceeds that in spatial scale (water layers and size fractions). Therefore, we conclude that DNA-based size-fractionated filtering (i.e., 0.2–200 μm and 0.2–3 and 3–200 μm) cannot change the abundance of dominant taxa in most cases, but can alter the results of low-abundance taxa. This work provides new insights into the application of various size-fractionation filtering strategies in the DNA sequencing-based study of microbial eukaryotes.

## Data availability statement

The data presented in this study are deposited in the NCBI database, accession number: PRJNA770434.

## Author contributions

JY led and designed the research. YX performed the sample collection and the PCR. GM performed the qPCR, high-throughput sequencing, and bioinformatics. GM and JY analyzed the data and wrote the first draft of the manuscript. YX and RL reviewed and edited the manuscript. All authors contributed to and approved the final manuscript.

## Funding

This work was supported by the National Natural Science Foundation of China (91851104, 32001152, and 92047204), the Natural Science Foundation of Fujian Province of China (2020J05089), and the "Light of West" Program and the "Fujian STS" Program (2021T3015 and 2022T3015) of the Chinese Academy of Sciences.

## Acknowledgments

We express gratitude to Dr. Lemian Liu, Dr. Lei Jin, Dr. Shuzhen Li, and Dr. Abdullah Al Mamun for their constructive comments on the manuscript. We acknowledge the assistance of Dr. Min Liu for collecting samples.

## 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.2022.969799/full#supplementary-material

## References

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

Banerji, A., Bagley, M., Elk, M., Pilgrim, E., Marinson, J., and Santo Domingo, J. (2018). Spatial and temporal dynamics of a freshwater eukaryotic plankton community revealed via 18S rRNA gene metabarcoding. *Hydrobiologia* 818, 71–86. doi: 10.1007/s10750-018-3593-0

Barlow, J. T., Bogatyrev, S. R., and Ismagilov, R. F. (2020). A quantitative sequencing framework for absolute abundance measurements of mucosal and lumenal microbial communities. *Nat. Commun.* 11:2590. doi: 10.1038/s41467-020-16224-6

Boehrer, B., and Schultze, M. (2008). Stratification of lakes. *Rev. Geophys.* 46:RG2005. doi: 10.1029/2006RG000210

Caron, D. A., and Hu, S. K. (2019). Are we overestimating protistan diversity in nature? *Trends Microbiol.* 27, 197–205. doi: 10.1016/j.tim.2018.10.009

Cavalier-Smith, T. (1997). Amoeboflagellates and mitochondrial cristae in eukaryote evolution: Megasystematics of the new protozoan subkingdoms Eozoa and Neozoa. *Arch. Protistenkd.* 147, 237–258. doi: 10.1016/S0003-9365(97)80051-6

Cavalier-Smith, T., and Chao, E. E. (2004). Protalveolate phylogeny and systematics and the origins of Sporozoa and Dinoflagellates (phylum Myzozoa nom. nov.). *Eur. J. Protistol.* 40, 185–212. doi: 10.1016/j.ejop.2004.01.002

Clarke, K. R., and Gorley, R. N. (2015). *PRIMER v7: User manual/tutorial*. Plymouth: PRIMER-E Ltd.

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

Costea, P. I., Zeller, G., Sunagawa, S., Pelletier, E., Alberti, A., Levenez, F., et al. (2017). Towards standards for human fecal sample processing in metagenomic studies. *Nat. Biotechnol.* 35, 1069–1076. doi: 10.1038/nbt.3960

Cotner, J. B., and Biddanda, B. A. (2002). Small players, large role: Microbial influence on biogeochemical processes in pelagic aquatic ecosystems. *Ecosystems* 5, 105–121. doi: 10.1007/s10021-001-0059-3

David, G. M., Moreira, D., Reboul, G., Annenkova, N. V., Galindo, L. J., Bertolino, P., et al. (2021). Environmental drivers of plankton protist communities along latitudinal and vertical gradients in the oldest and deepest freshwater lake. *Environ. Microbiol.* 23, 1436–1451. doi: 10.1111/1462-2920.15346

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

Dell'Anno, A., and Danovaro, R. (2005). Extracellular DNA plays a key role in deep-sea ecosystem functioning. *Science* 309:2179. doi: 10.1126/science.1117475

Djurhuus, A., Closek, C. J., Kelly, R. P., Pitz, K. J., Michisaki, R. P., Starks, H. A., et al. (2020). Environmental DNA reveals seasonal shifts and potential interactions in a marine community. *Nat. Commun.* 11:254. doi: 10.1038/s41467-019-14105-1

Engelbrektson, A., Kunin, V., Wrighton, K. C., Zvenigorodsky, N., Chen, F., Ochman, H., et al. (2010). Experimental factors affecting PCR-based estimates of microbial species richness and evenness. *ISME J.* 4, 642–647. doi: 10.1038/ismej.2009.153

Gao, X. F., Wang, W. P., Ndayishimiye, J. C., Govaert, L., Chen, H. H., Jeppesen, E., et al. (2022). Invasive and toxic cyanobacteria regulate allochthonous resource use and community niche width of reservoir zooplankton. *Freshw. Biol.* 67, 1344–1356. doi: 10.1111/fwb.13921

Garner, E., Davis, B. C., Milligan, E., Blair, M. F., Keenum, I., Maile-Moskowitz, A., et al. (2021). Next generation sequencing approaches to evaluate water and wastewater quality. *Water Res.* 194:116907. doi: 10.1016/j.watres.2021.116907

Giner, C. R., Balague, V., Krabberod, A. K., Ferrera, I., Rene, A., Garces, E., et al. (2019). Quantifying long-term recurrence in planktonic microbial eukaryotes. *Mol. Ecol.* 28, 923–935. doi: 10.1111/mec.14929

Gloor, G. B., Macklaim, J. M., Pawlowsky-Glahn, V., and Egozcue, J. J. (2017). Microbiome datasets are compositional: And this is not optional. *Front. Microbiol.* 8:2224. doi: 10.3389/fmicb.2017.02224

Gong, W., and Marchetti, A. (2019). Estimation of 18S gene copy number in marine eukaryotic plankton using a next-generation sequencing approach. *Front. Mar. Sci.* 6:219. doi: 10.3389/fmars.2019.00219

Guillou, L., Bachar, D., Audic, S., Bass, D., Berney, C., Bittner, L., et al. (2013). The protist ribosomal reference database (PR2): A catalog of unicellular eukaryote small sub-unit rRNA sequences with curated taxonomy. *Nucleic Acids Res.* 41, D597–D604. doi: 10.1093/nar/gks1160

Hadziavdic, K., Lekang, K., Lanzen, A., Jonassen, I., Thompson, E. M., and Troedsson, C. (2014). Characterization of the 18S rRNA gene for designing universal eukaryote specific primers. *PLoS One* 9:e87624. doi: 10.1371/journal.pone.0087624

Harder, C. B., Ronn, R., Brejnrod, A., Bass, D., Al-Soud, W. A., and Ekelund, F. (2016). Local diversity of heathland Cercozoa explored by in-depth sequencing. *ISME J.* 10, 2488–2497. doi: 10.1038/ismej.2016.31

Hasegawa, H., Naganuma, K., Nakagawa, Y., and Matsuyama, T. (2003). Membrane filter (pore size, 0.22–0.45 μm; thickness, 150 μm) passing-through activity of *Pseudomonas aeruginosa* and other bacterial species with indigenous infiltration ability. *FEMS Microbiol. Lett.* 223, 41–46. doi: 10.1016/s0378-1097(03)00327-6

Jo, J., Oh, J., and Park, C. (2020). Microbial community analysis using high-throughput sequencing technology: A beginner's guide for microbiologists. *J. Microbiol.* 58, 176–192. doi: 10.1007/s12275-020-9525-5

Knight, R., Vrbanac, A., Taylor, B. C., Aksenov, A., Callewaert, C., Debelius, J., et al. (2018). Best practices for analysing microbiomes. *Nat. Rev. Microbiol.* 16, 410–422. doi: 10.1038/s41579-018-0029-9

Lajeunesse, T. C., and Thornhill, D. J. (2011). Improved resolution of reef-coral endosymbiont (*Symbiodinium*) species diversity, ecology, and evolution through *psbA* non-coding region genotyping. *PLoS One* 6:e29013. doi: 10.1371/journal.pone.0029013

Lavrinienko, A., Jernfors, T., Koskimki, J. J., Pirttil, A. M., and Watts, P. C. (2020). Does intraspecific variation in rDNA copy number affect analysis of microbial communities? *Trends Microbiol.* 29, 19–27. doi: 10.1016/j.tim.2020.05.019

Lin, L., Gu, H., Luo, Z., and Wang, N. (2021). Size-dependent spatio-temporal dynamics of eukaryotic plankton community near nuclear power plant in Beibu Gulf, China. *J. Oceanol. Limnol.* 39, 1910–1925. doi: 10.1007/s00343-020-0248-6

Litchman, E., Ohman, M. D., and Kiørboe, T. (2013). Trait-based approaches to zooplankton communities. *J. Plankton Res.* 35, 473–484. doi: 10.1093/plankt/fbt019

Liu, L. M., Liu, M., Wilkinson, D. M., Chen, H. H., Yu, X. Q., 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

Liu, Q., Zhang, Y., Wu, H., Liu, F., Peng, W., Zhang, X., et al. (2020). A review and perspective of eDNA application to eutrophication and HAB control in freshwater and marine ecosystems. *Microorganisms* 8:417. doi: 10.3390/microorganisms8030417

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

Louca, S., Doebeli, M., and Parfrey, L. W. (2018). Correcting for 16S rRNA gene copy numbers in microbiome surveys remains an unsolved problem. *Microbiome* 6:41. doi: 10.1186/s40168-018-0420-9

Lynch, M. D., and Neufeld, J. D. (2015). Ecology and exploration of the rare biosphere. *Nat. Rev. Microbiol.* 13, 217–229. doi: 10.1038/nrmicro3400

Maritz, J. M., Ten Eyck, T. A., Elizabeth Alter, S., and Carlton, J. M. (2019). Patterns of protist diversity associated with raw sewage in New York City. *ISME J.* 13, 2750–2763. doi: 10.1038/s41396-019-0467-z

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

Morrison, J. M., Baker, K. D., Zamor, R. M., Nikolai, S., Elshahed, M. S., and Youssef, N. H. (2017). Spatiotemporal analysis of microbial community dynamics during seasonal stratification events in a freshwater lake (Grand Lake, OK, USA). *PLoS One* 12:e0177488. doi: 10.1371/journal.pone.0177488

Morton, J. T., Marotz, C., Washburne, A., Silverman, J., Zaramela, L. S., Edlund, A., et al. (2019). Establishing microbial composition measurement standards with reference frames. *Nat. Commun.* 10:2719. doi: 10.1038/s41467-019-10656-5

Nagler, M., Insam, H., Pietramellara, G., and Ascher-Jenull, J. (2018). Extracellular DNA in natural environments: Features, relevance and applications. *Appl. Microbiol. Biotechnol.* 102, 6343–6356. doi: 10.1007/s00253-018-9120-4

Nearing, J. T., Comeau, A. M., and Langille, M. G. I. (2021). Identifying biases and their potential solutions in human microbiome studies. *Microbiome* 9:113. doi: 10.1186/s40168-021-01059-0

Nearing, J. T., Douglas, G. M., Comeau, A. M., and Langille, M. G. I. (2018). Denoising the denoisers: An independent evaluation of microbiome sequence error-correction approaches. *PeerJ* 6:e5364. doi: 10.7717/peerj.5364

Nyirabuhoro, P., Liu, M., Xiao, P., Liu, L., Yu, Z., Wang, L., et al. (2020). Seasonal variability of conditionally rare taxa in the water column bacterioplankton community of subtropical reservoirs in China. *Microb. Ecol.* 80, 14–26. doi: 10.1007/s00248-019-01458-9

Piwosz, K., Shabarova, T., Pernthaler, J., Posch, T., Šimek, K., Porcal, P., et al. (2020). Bacterial and eukaryotic small-subunit amplicon data do not provide a quantitative picture of microbial communities, but they are reliable in the context of ecological interpretations. *mSphere* 5:e00052-20. doi: 10.1128/msphere.00052-20

R Core Team (2020). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.

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

Salmaso, N., Boscaini, A., and Pindo, M. (2020). Unraveling the diversity of eukaryotic microplankton in a large and deep perialpine lake using a high throughput sequencing approach. *Front. Microbiol.* 11:789. doi: 10.3389/fmicb.2020.00789

Sandin, M. M., Romac, S., and Not, F. (2022). Intra-genomic rRNA gene variability of Nassellaria and Spumellaria (Rhizaria, Radiolaria) assessed by Sanger, MinION and Illumina sequencing. *Environ. Microbiol.* 24, 2979–2993. doi: 10.1111/1462-2920.16081

Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., et al. (2009). Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol.* 75, 7537–7541. doi: 10.1128/AEM.01541-09

Simon, M., Lopez-Garcia, P., Deschamps, P., Moreira, D., Restoux, G., Bertolino, P., et al. (2015). Marked seasonality and high spatial variability of protist communities in shallow freshwater systems. *ISME J.* 9, 1941–1953. doi: 10.1038/ismej.2015.6

Sinha, R., Abu-Ali, G., Vogtmann, E., Fodor, A. A., Ren, B., Amir, A., et al. (2017). Assessment of variation in microbial community amplicon sequencing by the microbiome quality control (MBQC) project consortium. *Nat. Biotechnol.* 35, 1077–1086. doi: 10.1038/nbt.3981

Sogin, M. L., Morrison, H. G., Huber, J. A., Mark Welch, D., 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

Staley, C., Unno, T., Gould, T. J., Jarvis, B., Phillips, J., Cotner, J. B., et al. (2013). Application of Illumina next-generation sequencing to characterize the bacterial community of the Upper Mississippi River. *J. Appl. Microbiol.* 115, 1147–1158. doi: 10.1111/jam.12323

Tsilimigras, M., and Fodor, A. A. (2016). Compositional data analysis of the microbiome: Fundamentals, tools, and challenges. *Ann. Epidemiol.* 26, 330–335. doi: 10.1016/j.annepidem.2016.03.002

Weiss, S., Xu, Z. Z., Peddada, S., Amir, A., Bittinger, K., Gonzalez, A., et al. (2017). Normalization and microbial differential abundance strategies depend upon data characteristics. *Microbiome* 5:27. doi: 10.1186/s40168-017-0237-y

Woodhouse, J. N., Kinsela, A. S., Collins, R. N., Bowling, L. C., Honeyman, G. L., Holliday, J. K., et al. (2016). Microbial communities reflect temporal changes in cyanobacterial composition in a shallow ephemeral freshwater lake. *ISME J.* 10, 1337–1351. doi: 10.1038/ismej.2015.218

Xue, Y. Y., Chen, H. H., Yang, J. R., Liu, M., Huang, B. Q., 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

Yang, C. Y., Bohmann, K., Wang, X. Y., Cai, W., Wales, N., Ding, Z. L., et al. (2021). Biodiversity soup II: A bulk-sample metabarcoding pipeline emphasizing error reduction. *Methods Ecol. Evol.* 12, 1252–1264. doi: 10.1111/2041-210X.13602

Yang, J., and Zhang, X. (2020). eDNA metabarcoding in zooplankton improves the ecological status assessment of aquatic ecosystems. *Environ. Int.* 134:105230. doi: 10.1016/j.envint.2019.105230

Yang, J., Zhang, X. W., Jin, X. W., Seymour, M., Richter, C., Logares, R., et al. (2021). Recent advances in environmental DNA-based biodiversity assessment and conservation. *Divers. Distrib.* 27, 1876–1879. doi: 10.1111/ddi.13415

Yang, J. R., Lv, H., Yang, J., Liu, L. M., Yu, X. Q., and Chen, H. H. (2016). Decline in water level boosts cyanobacteria dominance in subtropical reservoirs. *Sci. Total Environ.* 557–558, 445–452. doi: 10.1016/j.scitotenv.2016.03.094

Yu, Z., Yang, J., Amalfitano, S., Yu, X. Q., and Liu, L. M. (2014). Effects of water stratification and mixing on microbial community structure in a subtropical deep reservoir. *Sci. Rep.* 4:5821. doi: 10.1038/srep05821

Zhang, H. H., Sekar, R., and Visser, P. M. (2020). Editorial: Microbial ecology in reservoirs and lakes. *Front. Microbiol.* 11:1348. doi: 10.3389/fmicb.2020.01348

Zheng, X., Wang, Q., Chen, L., Wang, J., and Cheng, R. (2015). Photocatalytic membrane reactor (PMR) for virus removal in water: Performance and mechanisms. *Chem. Eng. J.* 277, 124–129. doi: 10.1016/j.cej.2015.04.117

Zhu, F., Massana, R., Not, F., Marie, D., and Vaulot, D. (2005). Mapping of picoeucaryotes in marine ecosystems with quantitative PCR of the 18S rRNA gene. *FEMS Microbiol. Ecol.* 52, 79–92. doi: 10.1016/j.femsec.2004.10.006

frontiers | Frontiers in Marine Science

TYPE Original Research
PUBLISHED 24 October 2022
DOI 10.3389/fmars.2022.1032847



OPEN ACCESS

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

REVIEWED BY
Chuner Cai,
Shanghai Ocean University, China
Muhammed Duman,
Uludağ University,
Turkey

*CORRESPONDENCE
Hailong Huang
huanghailong@nbu.edu.cn
Haibo Jiang
jianghaibo@nbu.edu.cn

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

RECEIVED 07 September 2022
ACCEPTED 10 October 2022
PUBLISHED 24 October 2022

CITATION
Luo N, Huang H and Jiang H (2022) Establishment of methods for rapid detection of *Prymnesium parvum* by recombinase polymerase amplification combined with a lateral flow dipstick. *Front. Mar. Sci.* 9:1032847. doi: 10.3389/fmars.2022.1032847



# Establishment of methods for rapid detection of *Prymnesium parvum* by recombinase polymerase amplification combined with a lateral flow dipstick

Ningjian Luo[1], Hailong Huang[1]* and Haibo Jiang[1,2]*

[1]Key Laboratory of Marine Biotechnology of Zhejiang Province, School of Marine Science, Ningbo University, Ningbo, China, [2]Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), Zhuhai, China

*Prymnesium parvum* is a toxic algal bloom (HAB)-forming species. The toxicity of this alga is a result of a collection of compounds known as prymnesins. Prymnesins exert harmful effects upon fish, shellfish, and mollusks, causing huge economic losses. In the present study, a new method was developed for the detection of *P. parvum*. The novel method utilizes isothermal amplification, known as recombinase polymerase amplification (RPA), in combination with lateral-flow dipstick (LFD). Herein, a set of primers and probes were designed for internal transcribed spacer (ITS) sequences, and a specific and sensitive RPA-LFD rapid detection method was established for *P. parvum*. Meanwhile, we verified its feasibility for the detection of environmental samples. It was demonstrated that the optimal amplification temperature and time for RPA were 39°C and 15 min. RPA/RPA-LFD was experimentally verified to be specific, demonstrating no cross-reaction with distinct control microalgae, and furthermore, the total time required for the RPA-LFD experiment was 20 min. Meanwhile, the detection limit for the genomic DNA of *P. parvum* was $1.5\times10^{-1}$ pg/μL, and the detection limit for plasmids was 2.35 pg/μL. In addition, the results herein revealed that the RPA-LFD assay was 100 times more sensitive than PCR for detection of *P. parvum*. In conclusion, we developed an RPA-LFD that does not require precision instruments, and can be utilized for rapid on-site detection of *P. parvum*. In the future, the RPA-LFD can be considered for practical application for environmental detection of the toxic algal species.

KEYWORDS

***Prymnesium parvum*, rapid detection, recombinase polymerase amplification, lateral flow dipstick, internal transcribed spacer (ITS)**

## Introduction

The haptophyte *Prymnesium parvum* is a freshwater marine species that usually exists in brackish and inland high-mineral waters. It first reported in a Brackish water pond in 1937 (Carter, 1937). As a toxic and harmful alga, the toxin produced by *P. parvum*, prymnesins, exhibits potent neurotoxic, ichthyotoxic, cytotoxic and hemolytic effects (Manning and Claire, 2010b; Seoane et al., 2016). Prymnesins exert a potent toxic effect on fish, regularly resulting in the deaths of a large number of fish and subsequently causing great losses to farmers. Moreover, in humans toxins can be accumulated in by the consumption of seafood leading to the dangerous biomagnification of reactive molecules (Sharifan and Ma, 2017). For phytoplankton, mixed blooms of the killer alga *P. parvum* and known harmful cyanobacteria were one of the major causes of catastrophic changes in Europe's protected wetlands. The alga and cyanobacteria combination caused the phytoplankton to collapse (Maria et al., 2022). Over the previous two decades, the harmful algal blooms of *P. parvum* have been reported in the North America regions, such as Hawaii, Nebraska, etc., causing tens of millions of dollars in economic damage (Southard et al., 2010; Roelke et al., 2016). In China, algal blooms resulting from *P. parvum* occur frequently in Ningxia, Tianjin, Zhejiang, Shandong, Shanxi and other provinces. *P. parvum* has been known to cause the mass mortality of fish in China since 1963 (Guo et al., 1996). In a comparison of the toxicity of three kinds of harmful algae frequently occurring in Sanmen Bay and Xiangshan Bay, *P. parvum* exhibited the most toxic effect, causing acute death within five hours (Yang et al., 2018). Furthermore, *P. parvum* was isolated from the Pearl River Estuary in China (Qin et al., 2020). Therefore, it is necessary to establish a detection method of *P. parvum* that is fast and sensitive enough to detect the algae in advance prior to the outbreak of blooms; the assay must also meet the requirements of field detection, so as to prevent loss as much as possible.

At present, the methods used for the detection of microbiota are mainly divided into two categories; the first are based on the detection of algae morphology, primarily by microscopic examination (Töbe et al., 2006); the second are molecular detection methods based on nucleic acid amplification, including the polymerase chain reaction (PCR) (Manning and Claire, 2010a), quantitative PCR (qPCR) (Manning and Claire, 2013; Beyer et al., 2019), molecular probe techniques (Eckford-Soper and Daugbjerg, 2015) and isothermal amplification techniques (IAT). Upon the basis of these techniques, the detection of *P. parvum* has been broadly established. However, when considering IAT, loop-mediated isothermal amplification (LAMP) is the only method used for *P. parvum* detection (Zhu et al., 2019). Furthermore, all the existing detection methods have some limitations. For example, *P. parvum* is relatively small and fragile and can become distorted during preservation, making its identification and detection difficult (Galluzzi et al., 2008). Examination by microscope is highly uncertain and subjective (Zamor et al., 2012). Only experts skilled in algal identification are qualified for this task. In addition to microscopy methods, PCR, qPCR and other non-isothermal nucleic acid amplification methods possess the advantages of being fast, sensitive, and specific. Although the above molecular methods are useful additions to traditional identification tools, they still need to be performed under laboratory conditions by professional technicians. Moreover, it is difficult to achieve on-site detection of environmental samples. Although a loop-mediated isothermal amplification and lateral flow dipstick (LAMP-LFD) method may possess promise in solving the problems associated with detection in the field, primer design for LAMP technology is complicated (4-6 pairs of primers), and the reaction requires incubation at 65–67°C for 30–60 min. These shortcomings have limited the application of the aforementioned molecular methods for the rapid detection of *P. parvum*. Fortunately, in addition to LAMP, several other isothermal amplification techniques (IAT) have been widely employed for the detection of harmful algal bloom species, including quadruplex priming amplification (QPA) (Lomidze et al., 2018), helicase-dependent amplification (HDA) (Liu et al., 2020), nucleic acid sequence-based amplification (NASBA) (Nai et al., 2022), rolling circle amplification (RCA) (Yang et al., 2022), and recombinase polymerase amplification (RPA) (Cao et al., 2022).

The RPA method was first developed by Piepenburg in 2006 as an isothermal nucleic acid amplification technique (Piepenburg et al., 2006). The RPA reaction process can be completed within 20 minutes at a constant temperature of 37°C without the requirement for high-temperature denaturation. Previous studies have demonstrated that RPA has the same or greater sensitivity and specificity as PCR for the detection of algae, and it only requires a single pair of primers (Fu et al., 2019; Zhai et al., 2021; Toldrà et al., 2018). Finally, RPA can be combined with a variety of endpoint detection techniques for the rapid and simultaneous detection of multiple samples, these include agarose gel electrophoresis (Fu et al., 2020), real-time fluorescence quantitative assays (Wang X.F. et al., 2020), dye assays (Kang et al., 2020), and lateral flow dipsticks (LFD) (Ayfan Abdulrahman et al., 2021). LFDs are simple biosensors that involve building a ternary complex consisting of a biotinylated RPA product hybridized with a fluorescein isothiocyanate (FITC) probe and further complexed with a gold-labeled anti-FITC antibody in non-laboratory environments (Lin et al., 2022). When the control line and the test line simultaneously color, a positive result is indicated; a result than can be detected by the human eye. Indeed, LFD is a convenient assay for immediate visualization after the rapid amplification of RPA (Xia et al., 2022). Compared with traditional gel

electrophoresis detection, LFDs main advantages are that a result can be observed within 1-2 minutes, without the need for precision instruments such as imagers (Wang L. et al., 2020).

At present, due to the technical advantages exhibited by RPA-LFD, the technique has been widely used in various testing fields (Daher et al., 2014; Karakkat et al., 2018; Li et al., 2019; Yang et al., 2020; Jiang T.T. et al., 2022). However, the use of RPA-LFD in detecting harmful algae is still in its infancy. In the present study, we establish a specific and sensitive RPA-LFD assay for the detection of P. *parvum* and investigated its performance in application in the field. The ultimate purpose of this study was to explore novel methods for the rapid, low cost, and convenient field detection of harmful algae blooms species.

# Materials and methods

## Algae culture

A total of 9 algae species were selected for RPA/RPA-LFD specificity analysis in the study. Of these, *Prymnesium parvum* (NMBjih029), *Thalassiosira pseudonana* (NMBguh006), *Chaetoceros curvisetus* (NMBguh003-10), *Skeletonema costatum* (NMBguh0042), *Pseudo-nitzschia multiseries* (NMBguh002-1-1), and *Isochrysis galbana* (NMBjih021-2) were supplied by the Microalgae Collection of Ningbo University. Meanwhile, *Chaetoceros debilis* (MMDL50116), and *Thalassiosira rotula* (MMDL50319) were supplied by Xiamen University, *Trichodesmium erythraeum* (IMS101) was supplied by the University of Southern California and *Phaeodactylum tricornutum* (CCMP2561) was supplied by Westlake University. The culture medium used was f/2 medium. All algal strains were cultured in a 100 ml flask, the light intensity was 15~20 μmol/($m^2$·s), the light-dark period was 12 h:12 h, and the temperature was 16°C, with shaking 1–2 times daily, and stationary culture.

## DNA extraction, PCR, cloning and sequencing of target gene

Genomic DNA was obtained from *P. parvum* cells prepared by harvesting algal culture during their stationary phase, by centrifugation at 12000 × g for 2 min, with the Ezup Column Bacteria Genomic DNA Purification Kit (Sangon Biotech, Shanghai, China). DNA concentrations were measured using a Nano-300 micro spectrophotometer (Allsheng, Hangzhou, China), the concentration was $1.5\times10^4$ pg/μL. The genomic DNA was stored at -20°C.

The ITS of the target sequence was investigated and two PCR primers were designed; these were named Pr-PCR-F (5'-TCTCGCAACGATGAAGAACG-3') and Pr-PCR-R (5'-TCGATGGCACAACGACTTG-3'). The PCR reaction procedure was as follows: an initial denaturation at 95°C for 3 min, 30 cycles of 95°C for 15 s, 56°C for 15 s, and 72°C for 1 min, with a final extension at 72°C for 5 min. The PCR products were purified by Hipure PCR Pure Kits (Magen, Guangzhou, China). The purified PCR products were linked using a pMD 18-T Vector (TaKaRa, Dalian, China) to transform the *competent Escherichia coli* (DH5α). The positive clones identified by colony PCR were sent to Zhejiang Youkang Biotechnology Co., Ltd. for sequencing.

## Design of RPA primers and probes

According to the principles of primer probe design in TwistAmp® DNA Amplification Kits (TwistDX, UK), 5 pairs of primers were designed for ITS sequencing of *P. parvum*. The designed primers underwent further screening experiments to determine the final optimal primers. To obtain specific RPA primers, the Clustalw tool in the Biodeit software package (http://www.mbio.ncsu.edu/BioEdit/bioedit.html) was utilized in addition to BLAST in the NCBI. The sequences of similar species involved in multiple alignment analysis included *Isochrysis galbana* (JX393298.1), *Chrysochromulina ephippum* (KT390082.1) and *Pleurochrysis scherffelii* (AM936922.1). Primer Premier 5 (www.PremierBiosoft.com) was used to design the primers and probes. For the RPA-LFD reaction, biotin was tagged to the 5'-end of the optimum anti-sense primer. The 5'-end of the probe was next labeled with a carboxy fluorescein (FAM), the middle of the probe was labeled with a tetrahydrofuran (THF) site, and the 3'-end was labeled with a C3-spacer blocking group. Synthesis of the primers and probe were performed by Zhejiang Youkang Biotechnology Co., Ltd. The primer design results are shown in Table 1.

## The RPA reaction and system optimization

The RPA reactions were conducted using TwistAmp® Basic kits (TwistDX,UK). RPA reactions were performed in a total volume of 50 μL, containing 29.5 μL of rehydration buffer, 11.2 μL of nuclease-free water, 2.4 μL of sense primer (10 μM), 2.4 μL of anti-sense primer (10 μM), and 2 μL of genomic DNA extracted from *P. parvum*. The reaction mixture was briefly vortexed and centrifuged in a 1.5 ml centrifuge tube. Next, the reaction mixture was transferred to a tube containing lyophilized enzyme pellets that were supplied in the TwistAmp® Basic kits, 2.5 μL of MgOAc (280 mM) was rapidly added to the pellets to resuspend. After vortexing and spinning, the solution

TABLE 1 The design of primers and probe for *Prymnesium parvum*.

| Primer | Sequence (5'-3') | GC (%) | Amplification length (bp) |
|---|---|---|---|
| PR-RPA-1-F | CAGTGAATCATCGAACTTTTGAACGCAACTGG | 43.8 | 242 |
| PR-RPA-1-R | CAACGACTTGGTAGGCGACCTACTAGCACG | 56.7 | |
| PR-RPA-2-F | CAGCGAAATGCGATACGTAATGCGAATTGC | 46.7 | 114 |
| PR-RPA-2-R | GAGGCGCCACTCGAAGAAACATGCTCCCAG | 60.0 | |
| PR-RPA-3-F | CAGTGAATCATCGAACTTTTGAACGCAACTGG | 43.8 | 173 |
| PR-RPA-3-R | CTGCCCTCAGGCGACGCTCGAACCTTGATC | 63.3 | |
| PR-RPA-4-F | GAATCATCGAACTTTTGAACGCAACTGG | 42.9 | 228 |
| PR-RPA-4-R | GTAGGCGACCTACTAGCACGTCGGCACA | 60.7 | |
| PR-RPA-5-F | CATCGAACTTTTGAACGCAACTGGC | 48.0 | 157 |
| PR-RPA-5-R | AGGCGACGCTCGAACCTTGATCTGG | 60.0 | |
| PR-RPA-P | [FAM]GAGCATGTTTCTTCGAGTGGCGCCTCCACCCG[dSpacer] CTGGGCGTGCGCCTCGC [C3-Spacer] | 69.4 | |

Biotin was tagged to the 5'-end of the PR-RPA-4-R. The 5'-end of the PR-RPA-P was labeled with a carboxy fluorescein (FAM), the middle of the probe was labeled with a tetrahydrofuran (THF) site, and the 3'-end was labeled with a C3-spacer blocking group.

was incubated in a metal bath (Allsheng, Hangzhou, China) at 39°C for 20 min. According to the recommended process (Lillis et al., 2016), the reaction tubes were removed after 4 min and the reaction was continued after vortexing. At the end of the RPA reaction, the product was purified using Hipure PCR Pure Kits (Magen, Guangzhou, China) and detected using 2% agarose gel electrophoresis.

System optimization was performed in two parts: reaction temperature optimization and reaction time optimization. Optimization of amplification temperature was performed using the optimum primers, and a total of 3 temperature gradients of 35°C, 37°C, and 39°C. Next, the optimal times were further stratified by setting different amplification times (10 min, 15 min, and 20 min). Subsequent experiments were all the optimum primers and reaction set-up.

## The RPA-LFD reaction

To confirm the RPA results, the LFD (Milenia biotec, Germany) was employed to detect the RPA products. RPA reactions were conducted using the DNA thermostatic rapid amplification kit (Amplification Future, China). Moreover, RPA reactions were performed with a total volume of 50 μL, containing 29.4 μL of buffer A, 11.5 μL of nuclease-free water, 2 μL of sense primer (10 μM), 2 μL of anti-sense primer labeled with biotin (10 μM), 0.6 μL of LFD probe (10 μM) and 2 μL of genomic DNA extracted from *P. parvum*. The subsequent procedure refers to the RPA reaction, but the RPA products were not purified using kits: the 8 μL RPA reaction product was added directly to 92 μL of assay buffer and mixed well in a well of a microtiter plate. The dipsticks were place in sample application assay buffer wells and incubated for 1 minutes in an upright position and the results could be read by eye after the control line was fully colored. The result was positive if both the control line and the detection line were observed. On the contrary, the result was negative if only the control line is observed.

## Specificity and sensitivity analysis

The specificity study was performed using the 10 aforementioned algal species. $ddH_2O$ was selected as the template for the negative control. After extracting the algal genomic DNA using the method mentioned in the "DNA extraction, PCR, cloning and sequencing of target genes" section, RPA, RPA-LFD and PCR amplification were performed. Simultaneously, to ensure the validity of genomic DNA, PCR amplification was first performed with universal primers TW81 (5'-GGGATCCGTTTCCGT AGGTGAACCTGC-3') and AB28 (5'-GGGATCCATATGCT TAAGTTCAGCGGGT-3') (White et al., 1990) before performing the specificity study.

Both the genomic DNA of *P. parvum* and the recombinant plasmids consisting of the ITS sequence of *P. parvum* were used to determine sensitivity. Genomic DNA and a recombinant plasmid containing the *P. parvum* ITS sequence were diluted in a ten-fold serial dilution in $ddH_2O$ for 8 gradients. The result was genomic DNA of concentration ranging from $1.5\times10^{4}$ pg/μL to $1.5\times10^{-3}$ pg/μL, and recombinant plasmid DNA of concentration ranging from $2.35\times10^{5}$ pg/μL to $2.35\times10^{-2}$ pg/μL. The plasmids were extracted with a PurePlasmid Mini Kit (CWBIO, Suzhou, China). Sensitivity of PCR, RPA, and RPA-LFD was detected using the serially diluted genomic DNA, while the plasmid was used as a template with $ddH_2O$ as a blank control. The primers used were the optimal RPA primers identified by earlier screening. The sensitivity of RPA-LFD was

evaluated by comparing the lower limits of detection for PCR, RPA, and RPA-LFD.

## Practicality study for the RPA-LFD assay

The established RPA-LFD method was applied to both the spiked samples and the real sample to verify the feasibility of this study for field application. The natural seawater used for the simulated water sample experiment was obtained from the South China Sea on August – September in 2021, and after microscopic examination it was confirmed that the sample contained no target algal cells. A number of *P. parvum* cells were harvested by centrifugation at 12000 × g for 2 min before the obtained cells were used to prepare the spiked samples using natural seawater as diluents. After a ten-fold gradient dilution of the water sample was generated ($1.12\times10^{5}$ cells/ml-$1.12\times10^{-2}$ cells/ml), genomic DNA was extracted and sensitive detection was performed. Finally, the real environmental water samples from Xiangshan and Sanmen Bay were applied to the practical study. Among them, Xiangshan water samples represented the red tide period. The accuracy of the RPA-LFD results was verified using PCR results.

# Results

## Primer screening and system optimization

As revealed by Figure 1, of the five pairs of primers, only PR-RPA-4-F/R could amplify a clear and bright band. Thus, PR-RPA-4-F/R was selected as the optimum primer for subsequent RPA reactions. For the RPA-LFD reactions, biotin was tagged to the 5'-end of the PR-RPA-4-R, named PR-RPA-4-biotion.

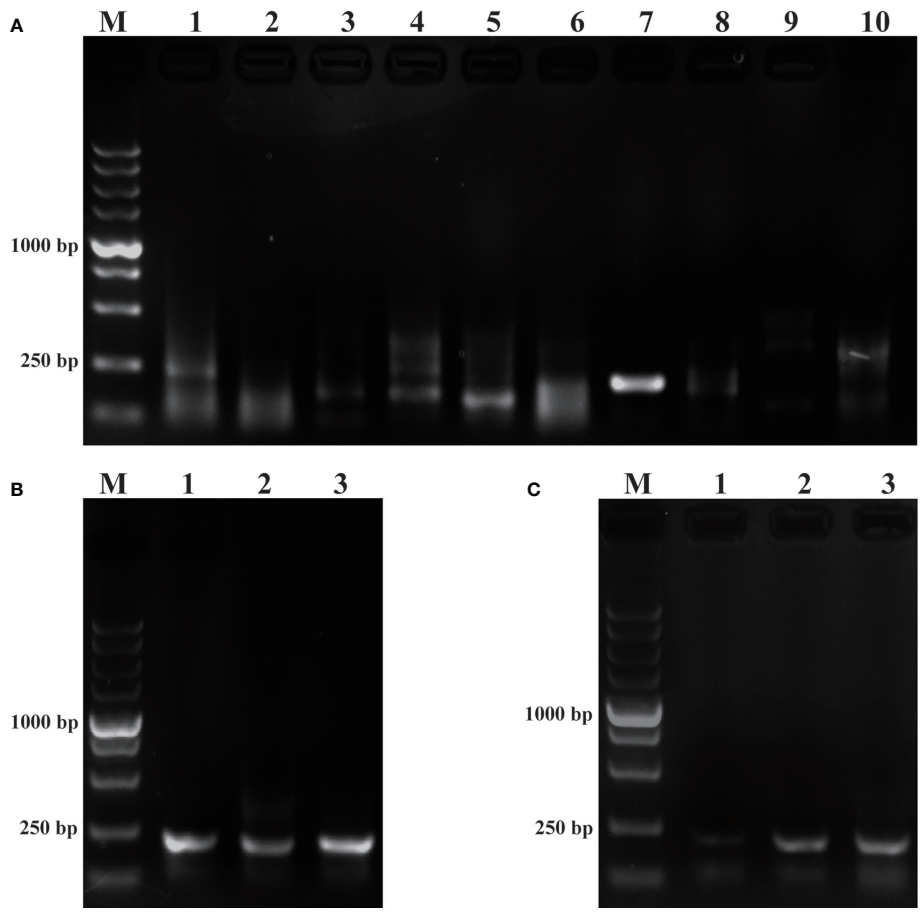



FIGURE 1
Primer screening and system optimization. **(A)**: Primer screening results for *P. parvum*. M: DL5000 DNA marker; 1, 3, 5, 7, 9: positive control of primers 1-5; 2, 4, 6, 8, 10: negative control of primers 1-5. **(B)**: Optimization of amplification temperature. 1: 35°C; 2: 37°C; 3: 39°C. **(C)**: Optimization of amplification time. 1: 10 min; 2: 15 min; 3: 20 min.

Using the screened optimal primers, the RPA amplification temperature was initially optimized at a 20 min time-point. Meanwhile, the temperature was set to a range of 35°C, 37°C, 39°C (Figure 1B). Among the three amplification temperatures, the amplification band representing the 39°C setting was the brightest, so 39°C was selected as the optimal temperature for the RPA amplification system. Time optimization experiments were next performed at the optimal 39°C temperature. The time gradient tested was 10 min, 15 min, 20 min (Figure 1C). When the amplification time was 15 min or 20 min, the LFD bands were brighter than for 10 min, thus 15 min was selected as the optimal amplification time from an economical perspective. Subsequent RPA/RPA-LFD were amplified using the described optimal temperature and time.

## Specificity study

The purpose of this study was to ensure accurate detection of *P. parvum* in complex environments. Thus, the primers and probes designed were required to be highly specific. As shown in Figure 2, the genomic DNA of the 10 species of algae used for specificity testing was successfully amplified with the universal primers (TW28 and AB28) for their ITS sequences; these data rule out possible interference due to degradation of genomic DNA of algae. Meanwhile, through interpretation of the results of PCR and RPA, it was identified that apart from for *P. parvum*, the results of the other algae amplification of all other algae species resulted in negative results in the RPA-LFD are all negative (Figures 3A, B). In addition, during the RPA-LFD detection of *P. parvum*, the control line and the test line were colored at the same time, while analysis of the other algae species resulted in coloring of the control line only (Figure 3C). Therefore, it was concluded that the primers and probes designed in this test are highly specific for *P. parvum* and could be used for further experimentation.

## Sensitivity study

Sensitivity is an important parameter for on-site detection of harmful algae. The methodological sensitivity and applicability determine whether the method could be utilized to predict the occurrence of harmful algal blooms. The detection limit of the RPA-LFD was first quantified using the genomic DNA of *P. parvum* ($1.5\times10^{4}$ pg/μL-$1.5\times10^{-3}$ pg/μL) in comparison to PCR and RPA (Figure 4). The results indicated that the detection limit of RPA-LFD ($1.5\times10^{-1}$ pg/μL) with genomic DNA was 1000 times higher than that of RPA alone ($1.5\times10^{2}$ pg/μL) and 100 times that of PCR ($1.5\times10^{1}$ pg/μL).

Similarly, sensitivity tests were performed using recombinant plasmid ($2.35\times10^{5}$ pg/μL-$2.35\times10^{-2}$ pg/μL). After performing conventional PCR, RPA or RPA-LFD amplification, the results obtained were illustrated in Figure 5. The detection limit for PCR was $2.35\times10^{0}$ pg/μL, while the detection limit for RPA was $2.35\times10^{1}$ pg/μL, and the detection limit for RPA-LFD was $2.35\times10^{0}$ pg/μL. The plasmid detection sensitivity of RPA-LFD was therefore same as that of PCR, with both techniques being 10 times more sensitive than RPA. Compared to the sensitivity for genomic DNA detection, the sensitivity of RPA

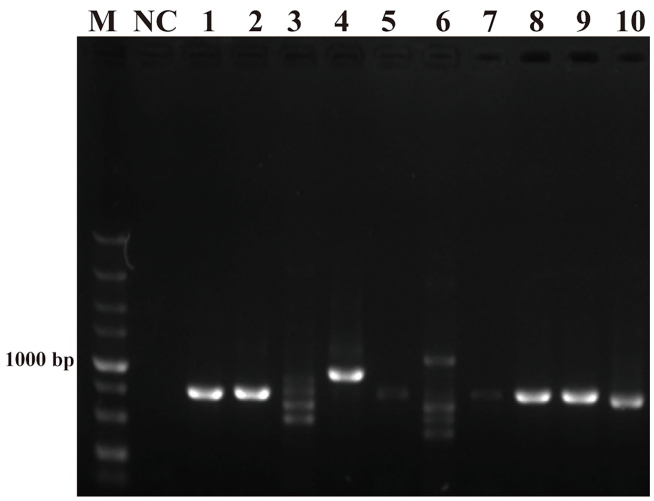



FIGURE 2
Universal primers for amplification of ITS sequences. M: DL5000 DNA marker; NC: negative control; 1: *Prymnesium parvum*; 2: *Thalassiosira pseudonana*; 3: *Chaetoceros debilis*; 4: *Phaeodactylum tricornutum*; 5: *Pseudo-nitzschia multiseries*; 6: *Trichodesmium*; 7: *Skeletonema costatum*; 8: *Thalassiosira rotula*; 9: *Chaetoceros curvisetus*; 10: *Isochrysis galbana*.

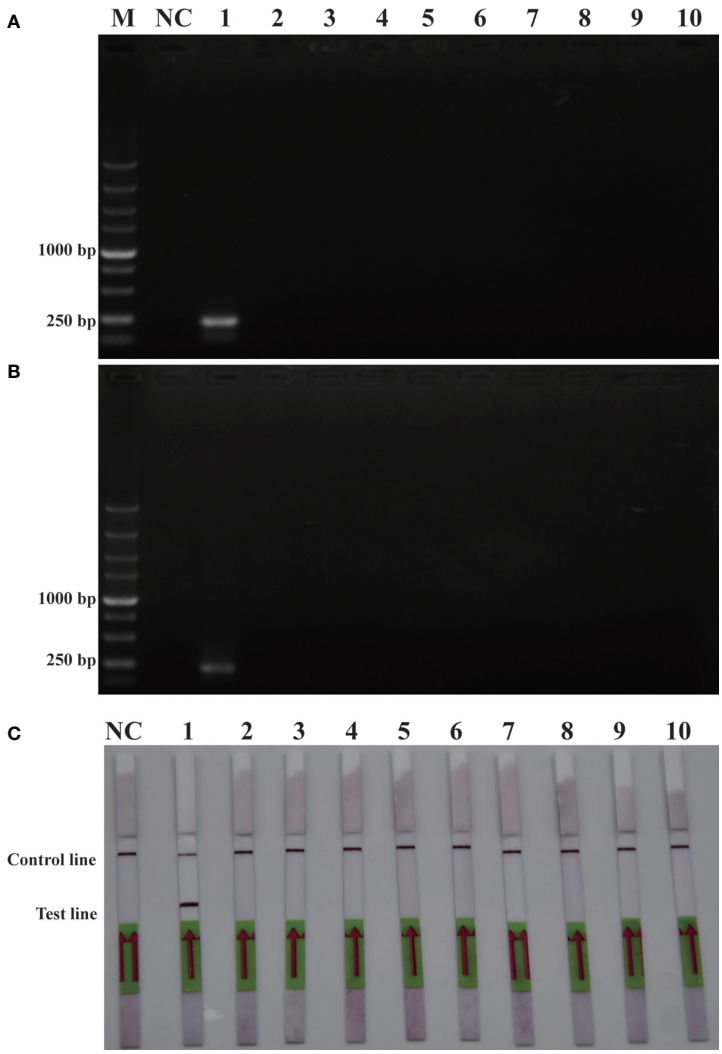



FIGURE 3
Specificity validation. **(A)**: PCR; **(B)**: RPA; **(C)**: RPA-LFD. M: DL5000 DNA marker; NC: negative control; 1: *P. parvum*; 2: *T. pseudonana*; 3: *C. debilis*; 4: *P. tricornutum*; 5: *P. multiseries*; 6: *Trichodesmium*; 7: *S. costatum*; 8: *T. rotula*; 9: *C. curvisetus*; 10: *Isochrysis galbana*.

for plasmid DNA was significantly improved. Overall, the results were suggested that the sensitivity of RPA-LFD is higher than that of conventional PCR.

## Practicality experiments on RPA-LFD

To verify the feasibility of the RPA-LFD assay for on-site environmental detection of *P. parvum*, practical experiments were conducted using spiked samples (Figure 6) and real samples (Figure 7). The results revealed that using the primers and probes optimized in this study, PCR, RPA and RPA-LFD successfully detected *P. parvum* in spiked samples effectively applied during on-site detection. In addition, RPA-LFD was more sensitive than PCR and RPA, with a detection limit of $1.12\times10^{-1}$ cells/ml. However, in the experiment with real environment water samples, both PCR and RPA-LFD test results were negative, which demonstrated that the environmental water sample currently did contain of *P. parvum* cells.

# Discussion

As a toxic haptophyte alga, *Prymnesium parvum* has caused cases of algal bloom outbreaks globally, causing huge economic

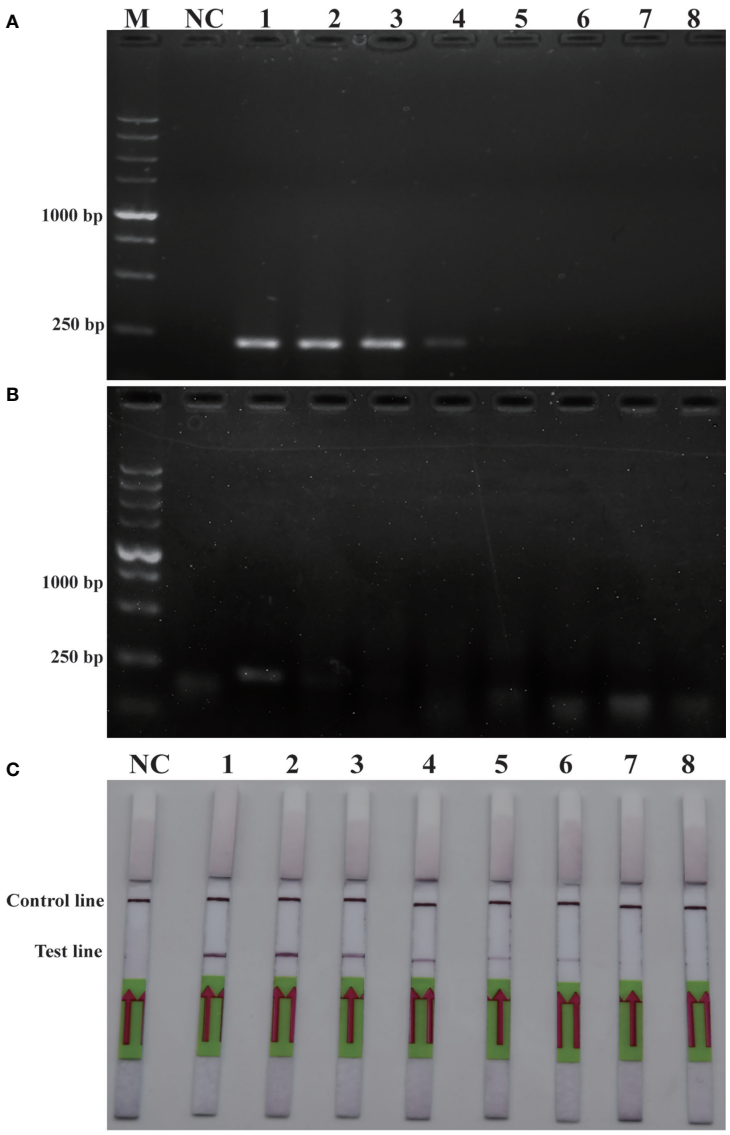



FIGURE 4
Comparison of sensitivities of PCR **(A)**, RPA **(B)** and RPA-LFD **(C)** with gDNA. M: DL5000 DNA marker; NC: negative control; 1-8: gDNA concentration ranges from $1.5\times10^{4}$ pg/μL-$1.5\times10^{-3}$ pg/μL.

losses. Therefore, it has attracted increasing levels of attention from researchers. However, most of the current research on *P. parvum* is focused on toxins and growth factors (Larsen and Bryant, 1998; McCoy Gary et al., 2014; Liu et al., 2022). To date, many tools have been developed for the detection of *P. parvum*; microscopic detection and PCR are still widely utilized, however, these methods are generally time-consuming and laborious. To the best of our knowledge, RPA-LFD-based approaches have not yet been established for monitoring aquatic environments for *P. parvum*.

In the present study, an RPA-LFD detection method targeted to the ITS sequence of *P. parvum* was established. The detection limit of the RPA-LFD for the genomic DNA of the algae was $1.5\times10^{-1}$ pg/μL, and the detection limit for plasmids was $2.35\times10^{0}$ pg/μL, not inferior to conventional PCR. In addition, the RPA-LFD assay took approximately 20 min in its entirety, substantially faster than the 1.5 hours taken for conventional PCR. Moreover, in comparison with other isothermal amplification methods for detecting *P. parvum*, RPA-LFD method possesses the significant advantage

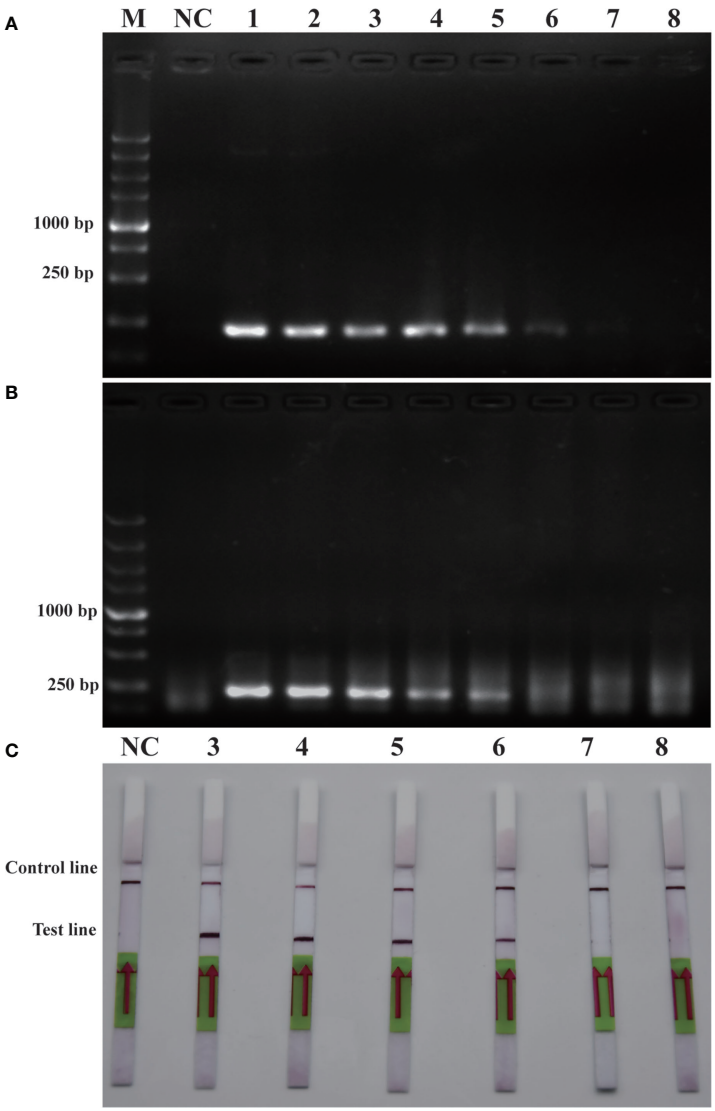



FIGURE 5
Comparison of sensitivities of PCR **(A)**, RPA **(B)** and RPA-LFD **(C)** with plasmid. M: DL5000 DNA marker; NC: negative control; 1-8: plasmid concentration ranges from $2.35\times10^{5}$ pg/μL-$2.35\times10^{-2}$ pg/μL.

of high sensitivity. For example, Zhu et al. (2019) demonstrated that the limit of detection was $3\times10^{1}$ pg/μL when LAMP-LFD was used to detect *P. parvum*. The detection limit of the RPA-LFD utilized in the present study ($1.5\times10^{-1}$ pg/μL) was 200 times higher than that of the that previously recorded for LAMP-LFD. Therefore, the RPA-LFD may be useful in rapid and sensitive detection of *P. parvum* in the laboratory or in the field while maintaining high specificity. As a caveat, in our study, the results of sensitivity of RPA-AGE were much lower than that of RPA-LFD and PCR, which was inconsistent with findings from other studies. The reason for this phenomenon may be related to the method of product purification used in RPA. In this study, we used a PCR product purification kit, while others have used phenol-chloroform mixtures for purification; the latter may have resulted in a loss of RPA product. This result in fact reflected the advantages of RPA-LFD, because the technique does not require purification of RPA products.

Unfortunately, RPA reactions produce more non-specific amplifications relative to PCR, a difficult phenomenon to avoid (Kim et al., 2017). Therefore, it is necessary to screen primers to minimize the interference of non-specific amplification for subsequent experiments. RPA is an emerging method, and

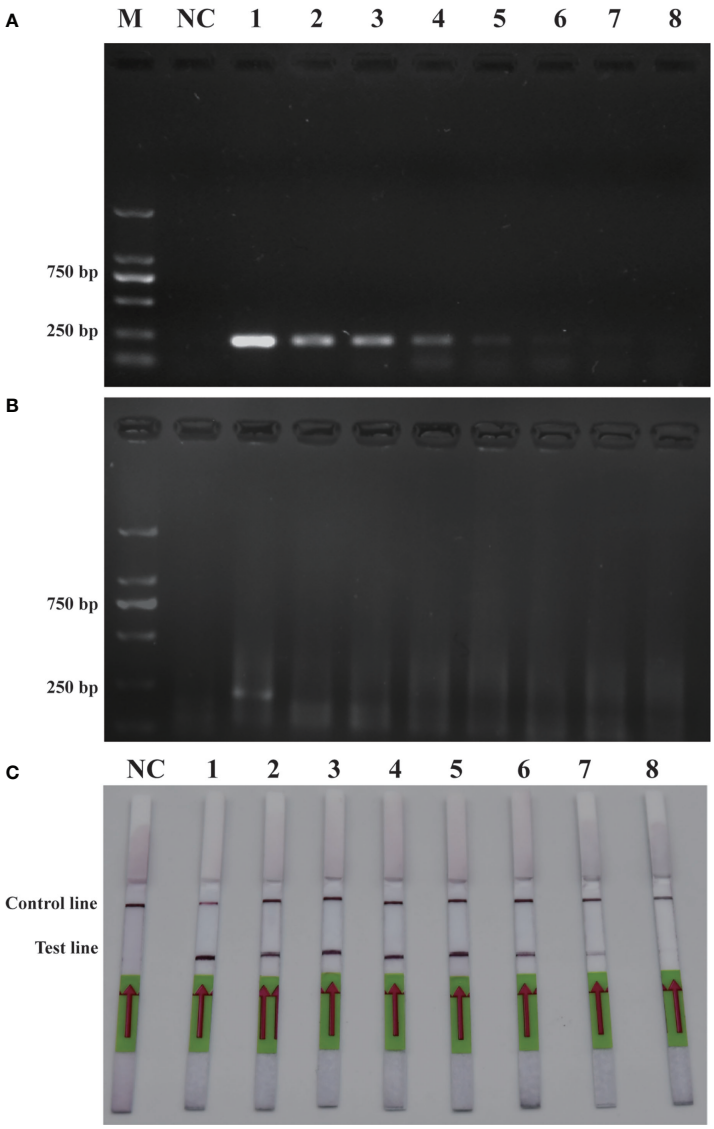



FIGURE 6
Comparison of sensitivities of PCR **(A)**, RPA **(B)** and RPA-LFD **(C)** with gDNA of spiked sample. M: DL2000 DNA marker; NC: negative control; 1-8: P. parvum cell concentration ranges from $1.12\times10^{5}$ cells/ml-$1.12\times10^{-2}$ cells/ml.

many researchers have conducted studies regarding the optimization of this method, including buffered conditions, specificity enhancement, stirred conditions, and additional methods of detection such as RPA combined with LFD (Moody et al., 2016; Luo et al., 2019; Tomar et al., 2022). It is worth noting that combination of RPA with the latest CRISPR technology has been applied to various bioassay fields (Patchsung et al., 2020; Meng et al., 2021; Jiang H.J. et al., 2022; Zhao et al., 2022). For example, an RPA-CRISPR/Cas13a detection system named Sherlock has been developed (Gootenberg et al., 2017). The combination of CRISPR and RPA resulted in high specificity and sensitivity of detection of RPA assays. At present, however, there are still few applications employed for the detection of harmful algal blooms species, and there is a huge potential for development (Durán-Vinet et al., 2021). In the future, it would be of great benefit to design RPA-CRISPR-LFD method for the rapid on-site detection of harmful algal bloom species.

To verify that the application of the RPA-LFD is reliable and stable for the analysis of field samples, mock environment

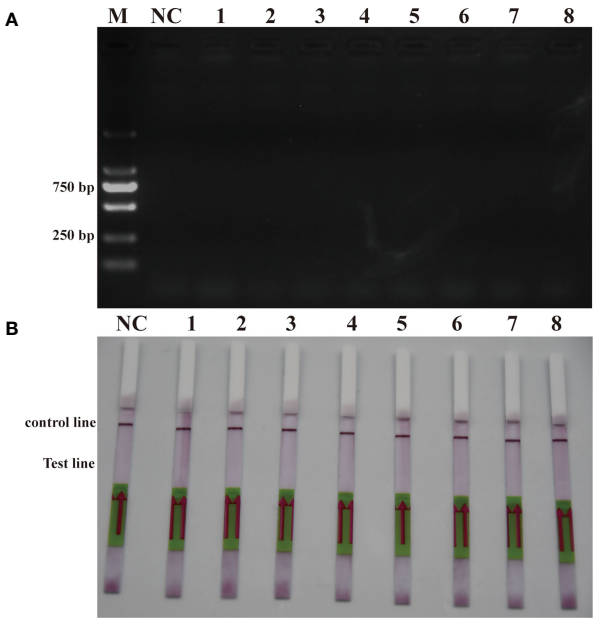



FIGURE 7
Real sample detection **(A)**: PCR; **(B)**: RPA-LFD. M: DL2000 DNA marker; NC: negative; 1: Xiangshan 1; 2: Xiangshan 2; 3: Sanmen Bay 1; 4: Sanmen Bay 2; 5: Sanmen Bay 3; 6: Sanmen Bay 4; 7: Sanmen Bay 5; 8: Sanmen Bay 6.

sample groups were established by spiking *P. parvum* into natural seawater in the absence of target algal cells. In the present study, we initially spiked samples to verify the feasibility of the method in the detection of field samples, and then tested using a real sample. However, the final experimental results demonstrated no presence of *P. parvum* in the real samples used in this study. The most likely reason for this result was that the environmental water sample contained no *P. parvum*. Therefore, additional field samples isolated during different seasons should be used for testing, perhaps during periods of algal bloom.

## Conclusion

This study established a rapid, highly specific and sensitive RPA-LFD method for detection of *Prymnesium parvum*. The method exhibits advantageous simple operation, short detection time and visual detection results. RPA-LFD requires just 20 min for the experimental process, substantially less than for PCR. Furthermore, this method can be employed in field for detection of *P. parvum*, and compared with traditional PCR, its sensitivity is substantially improved. These advantages mean this novel method can be used for early detection of *P. parvum* and prevention of algal bloom. Meanwhile, the RPA-LFD established in this study may also be utilized in largescale field screening for *P. parvum*.

## Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding authors.

## Author contributions

NL: Conceptualization, Methodology, Investigation, Data Curation, Writing-Original Draft; HH: Conceptualization, Methodology, Writing-Review and Editing, Supervision, Funding acquisition; HJ: Conceptualization, Methodology, Writing-Review and Editing, Supervision, Funding acquisition. All authors contributed to the article and approved the submitted version.

## Funding

This study was funded by the National Natural Science Foundation of China (Grant No. 91951111, No. 32170108, and No. 42188102) and the Independent Research Projects of Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai, Grant No. SML2021SP204).

## Acknowledgments

We are grateful to Prof. H.J Zhang from Ningbo University for help in field sampling.

## 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.2022.1032847/full#supplementary-material

## References

Ayfan Abdulrahman, K., Macdonald, J., Harris, P. N. A., Heney, C., Paterson, D. L., Trembizki, E., et al. (2021). Rapid detection of NDM and VIM carbapenemase encoding genes by recombinase polymerase amplification and lateral flow-based detection. *Eur. J. Clin. Microbiol. Infect. Dis.: Off. Publ. Eur. Soc. Clin. Microbiol.* 40 (11), 2447–2453. doi: 10.1007/s10096-021-04267-6

Beyer, J. E., Zamor, R. M., and Hambright, K. D. (2019). Detection limits affect the predictability of the presence of an invasive harmful alga across geographic space. *Biol. Invasions* 21 (7), 2301–2311. doi: 10.1007/s10530-019-01977-z

Cao, G. H., Huo, D. Q., Chen, X. L., Wang, X. F., Zhou, S. Y., Zhao, S. X, et al. (2022). Automated, portable, and high-throughput fluorescence analyzer (APHF-analyzer) and lateral flow strip based on CRISPR/Cas13a for sensitive and visual detection of SARS-CoV-2. *Talanta* 248, 123594–123594. doi: 10.1016/j.talanta.2022.123594

Carter, N. (1937). New or interesting algae from brackish water. *Arch. Protistenkd.* 90, 1–68.

Daher, R. K., et al, Stewart, G., Boissinot, M., Bergeron, M. G. (2014). Isothermal recombinase polymerase amplification assay applied to the detection of group b streptococci in vaginal/anal samples. *Clin. Chem.* 60 (4), 660–666. doi: 10.1373/clinchem.2013.213504

Durán-Vinet, B., Araya-Castro, K., Chao, T. C., Wood, S. A., Gallardo, V., Godoy, K, et al. (2021). Potential applications of CRISPR/Cas for next-generation biomonitoring of harmful algae blooms: A review. *Harmful Algae* 103, 102027–102027. doi: 10.1016/j.hal.2021.102027

Eckford-Soper, L. K., and Daugbjerg, N. (2015). Examination of six commonly used laboratory fixatives in HAB monitoring programs for their use in quantitative PCR based on taqman probe technology. *Harmful Algae* 42, 52–59. doi: 10.1016/j.hal.2014.12.007

Fu, M. Q., Chen, G. F., Zhang, C. Y., Wang, Y. Y., Sun, R., and Zhou, J. (2019). Rapid and sensitive detection method for *Karlodinium veneficum* by recombinase polymerase amplification coupled with lateral flow dipstick. *Harmful Algae* 84, 1–9. doi: 10.1016/j.hal.2019.01.011

Fu, M. Q., Yang, Y. Y., Zhang, C. Y., Chen, G. F., Wang, Y. Y., et al. (2020). Recombinase polymerase amplification combined with lateral-flow dipstick for rapid detection of *Prorocentrum minimum*. *J. Appl. Phycol.* 32, 1–14. doi: 10.1007/s10811-020-02079-3

Galluzzi, L., Bertozzini, E., Penna, A., Perini, F., Pigalarga, A., Graneli, E, et al. (2008). Detection and quantification of *Prymnesium parvum* (Haptophyceae) by real-time PCR. *Lett. Appl. Microbiol.* 46 (2), 261–266. doi: 10.1111/j.1472-765X.2007.02294.x

Gootenberg, J. S., Abudayyeh, O. O., Lee, J. W., Essletzbichler, P., Dy, A. J., Joung, J., et al. (2017). Nucleic acid detection with CRISPR-Cas13a/C2c2. *Science* 356 (6336), 438–442. doi: 10.1126/science.aam9321

Guo, M., Harrison, P. J., and Taylor, F. J. R. (1996). Fish kills related to *Prymnesium parvum* n. carter (Haptophyta) in the people's republic of China. *J. Appl. Phycol.* 8 (2), 111–117. doi: 10.1007/BF02186313

Jiang, T. T., Wang, Y. C., Jiao, W. W., Song, Y. Q., Zhao, Q., Wang, T. Y., et al. (2022). Recombinase polymerase amplification combined with real-time fluorescent probe for *Mycoplasma pneumoniae* detection. *J. Clin. Med.* 11 (7), 1780–1780. doi: 10.3390/jcm11071780

Jiang, H. J., Tan, R., Jin, M., Yin, J., Gao, Z. X., Li, H. B, et al. (2022). Visual detection of *Vibrio parahaemolyticus* using combined CRISPR/Cas12a and recombinase polymerase amplification. *Biomed. Environ. Sci.* 35 (6), 518–527. doi: 10.3967/bes2022.069

Kang, J. Y., Jang, H., Yeom, G., and Kim, M. G. (2020). Ultrasensitive detection platform of disease biomarkers based on recombinase polymerase amplification with h-sandwich aptamers. *Analytical Chem.* 93 (2), 992–1000. doi: 10.1021/acs.analchem.0c03822

Karakkat, B., Hockemeyer, K., Franchett, M., Olson, M., Mullenberg, C., Koch, P., et al. (2018). Data for designing two isothermal amplification assays for the detection of root-infecting fungi on cool-season turfgrasses. *Data Brief* 20, 471–479. doi: 10.1016/j.dib.2018.08.021

Kim, J. M., and truongatai Yoon, B. (2017). Comparison between specific DNA-amplifications using recombinase polymerase amplification (RPA) and using polymerase chain reaction (PCR). *J. Apicult.* 32 (1), 41–50. doi: 10.17519/apiculture.2017.04.32.1.41

Larsen, A., and Bryant, S. (1998). Growth rate and toxicity of *Prymnesium parvum* and *Prymnesium patelliferum*(haptophyta) in response to changes in salinity, light and temperature. *Sarsia* 83 (5), 409–418. doi: 10.1080/00364827.1998.10413700

Li, J. J., Wang, C. M., Yu, X., Lin, H. R., Hui, C., Shuai, L, et al. (2019). Rapid detection of *Cyanobacteria* by recombinase polymerase amplification combined with lateral flow strips. *Water Supply* 19 (4), 1181–1186. doi: 10.2166/ws.2018.174

Lillis, L., Siverson, J., Lee, A., Cantera, J., Parker, M., Piepenburg, O., et al. (2016). Factors influencing recombinase polymerase amplification (RPA) assay outcomes at point of care. *Mol. Cell. Probes* 30, 74–78. doi: 10.1016/j.mcp.2016.01.009

Lin, H., Zhao, S ., Liu, Y. H., Shao, L., Ye, Y. Y., Jiang, N. Z., et al. (2022). Rapid visual detection of *Plasmodium* using recombinase-aided amplification with lateral flow dipstick assay. *Front. Cell. Infect. Microbiol.* 12. doi: 10.3389/fcimb.2022.922146

Liu, M., Li, C. C., Luo, X. L., Ma, F., and Zhang, C. Y. (2020). 5-hydroxymethylcytosine glucosylation-triggered helicase-dependent amplification-based fluorescent biosensor for sensitive detection of β-glucosyltransferase with zero background signal. *Analytical Chem.* 92 (24), 16307–16313. doi: 10.1021/acs.analchem.0c04382

Liu, S. Y., Zhao, R. Z., Qiu, X. C., and Guo, Q. (2022). Optimization analysis to evaluate the relationships between different ion concentrations and *Prymnesium parvum* growth rate. *Water* 14 (6), 928–928. doi: 10.3390/w14060928

Lomidze, L., Williford, T. H., Musier-Forsyth, K., and Kankia, B. (2018). Isothermal amplification of long DNA segments by quadruplex priming amplification. *Analytical Methods: Advancing Methods Appl.* 10 (25), 2972–2979. doi: 10.1039/c8ay00843d

Luo, G. C., Yi, T. T., Jiang, B., Guo, X. L., and Zhang, G. Y. (2019). Betaine-assisted recombinase polymerase assay with enhanced specificity. *Analytical Biochem.* 575, 36–39. doi: 10.1016/j.ab.2019.03.018

Manning, S. R., and Claire, J. W. (2010a). Multiplex PCR methods for the species-specific detection and quantification of *Prymnesium parvum* (Haptophyta). *J. Appl. Phycol.* 22 (5), 587–597. doi: 10.1007/s10811-009-9498-6

Manning, S. R., and La Claire, J. W. (2010b). Prymnesins: Toxic metabolites of the golden alga, *Prymnesium parvum* carter (Haptophyta). *Mar. Drugs* 8 (3), 678–704. doi: 10.3390/md8030678

Manning, S. R., and La Claire, J. W.I. I. (2013). Isolation of polyketides from *Prymnesium parvum* (Haptophyta) and their detection by liquid chromatography/mass spectrometry metabolic fingerprint analysis. *Analytical Biochem.* 442 (2), 189–195. doi: 10.1016/j.ab.2013.07.034

Maria, D., Genitsaris, S., Mazaris, A. D., Kyparissis, A., Voutsa, D., Kozari, A., et al. (2022). A catastrophic change in a european protected wetland: From harmful phytoplankton blooms to fish and bird kill. *Environ. Pollution* 312, 120038. doi: 10.1016/j.envpol.2022.120038

McCoy Gary, R., Kegel, J. U., Touzet, N., Fleming, G. T. A., Medlin, L K., Raine, R., et al. (2014). An assessment of RNA content in *Prymnesium parvum*, prymnesium polylepis, cf. chattonella sp. and *Karlodinium veneficum* under varying environmental conditions for calibrating an RNA microarray for species detection. *FEMS Microbiol. Ecol.* 88 (1), 140–159. doi: 10.1111/1574-6941.12277

Meng, Q. Z., Yang, H. M., Zhang, G. Q., Sun, W. J., Ma, P. X., Liu, X. Y., et al. (2021). CRISPR/Cas12a-assisted rapid identification of key beer spoilage bacteria. *Innovative Food Sci. Emerg. Technol.* 74, 1–9. doi: 10.1016/j.ifset.2021.102854

Moody, C., Newell, H., Viljoen, H., et al (2016). A mathematical model of recombinase polymerase amplification under continuously stirred conditions. *Biochem. Eng. J.* 112, 193–201. doi: 10.1016/j.bej.2016.04.017

Nai, Y. H., Doeven, E. H., and Guijt, R. M. (2022). An improved nucleic acid sequence-based amplification method mediated by T4 gene 32 protein. *PloS One* 17 (3), e0265391–e0265391. doi: 10.1371/journal.pone.0265391

Patchsung, M., Jantarug, K., Pattama, A., Aphicho, K ., Suraritdechachai, S., Meesawat, P., et al. (2020). Clinical validation of a Cas13-based assay for the detection of SARS-CoV-2 RNA. *Nat. Biomed. Eng.* 4 (12), 1140–1149. doi: 10.1038/s41551-020-00603-x

Piepenburg, O., Williams, C. H., Stemple, D. L., and Armes, N. A. (2006). DNA Detection using recombination proteins. *PloS Biol.* 4 (7), e204. doi: 10.1371/journal.pbio.0040204

Qin, J. L., Hu, Z. X., Zhang, Q., Xu, N., and Yang, Y. F. (2020). Toxic effects and mechanisms of prymnesium parvum (Haptophyta) isolated from the pearl river estuary, China. *Harmful Algae* 96 (C), 101844. doi: 10.1016/j.hal.2020.101844

Roelke, D. L., Barkoh, A., Brooks, B. W., Grover, J. P., Hambright, K. D., Laclaire, J. W. II, et al. (2016). A chronicle of a killer alga in the west: ecology, assessment, and management of *Prymnesium parvum* blooms. *Hydrobiologia* 764 (1), 29–50. doi: 10.1007/s10750-015-2273-6

Seoane, S., Riobó, P., and Franco, J. (2016). Haemolytic activity in different species of the genus prymnesium (Haptophyta). *J. Mar. Biol. Assoc. United Kingdom* 97 (3), 491–496. doi: 10.1017/S0025315416001077

Sharifan, H., and Ma, X. M. (2017). Potential photochemical interactions of UV filter molecules with multichlorinated structure of prymnesins in harmful algal bloom events. *Mini-Rev. Org. Chem.* 14 (5), 391–399. doi: 10.2174/1570193X14666170518124658

Southard, G. M., Fries, L. T., and Barkoh, A. (2010). *Prymnesium parvum*: the Texas experience. *J. Am. Water Resour. Assoc.* 46 (1), 14–23. doi: 10.1111/j.1752-1688.2009.00387.x

Töbe, K., Eller, G., and Medlin, L. K. (2006). Automated detection and enumeration for toxic algae by solid-phase cytometry and the introduction of a new probe for *Prymnesium parvum* (Haptophyta: Prymnesiophyceae). *J. Plankton Res.* 28 (7), 643–657. doi: 10.1093/plankt/fbi147

Toldrà, A., Jauset-Rubio, M., Andree, K. B., Fernandez-Tejedor, M., Diogene, J., Katakis, I., et al. (2018). Detection and quantification of the toxic marine microalgae *Karlodinium veneficum* and *Karlodinium armiger* using recombinase polymerase amplification and enzyme-lined oligonucleotide assay. *Analytica Chim. Acta* 1039, 140–148. doi: 10.1016/j.aca.2018.07.057

Tomar, S., Lavickova, B., and Guiducci, C. (2022). Recombinase polymerase amplification in minimally buffered conditions. *Biosensors Bioelectronics* 198, 113802–113802. doi: 10.1016/j.bios.2021.113802

Wang, L., Chen, G. F., Zhang, C. Y., Wang, Y. Y., Sun, R., et al. (2020). Application of loop-mediated isothermal amplification combined with lateral flow dipstick to rapid and sensitive detection of *Alexandrium catenella*. *Environ. Sci. pollut. Res. Int.* 27 (4), 4246–4257. doi: 10.1007/s11356-019-06889-y

Wang, X. F., Xie, S. L., Chen, X. Y., Peng, C., Xu, X. L., Wei, W., et al. (2020). A rapid and convenient method for on-site detection of MON863 maize through real-time fluorescence recombinase polymerase amplification. *Food Chem.* 324, 126821. doi: 10.1016/j.foodchem.2020.126821

White, T. J., Bruns, T., Lee, S., and Taylor, J. (1990). "Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics," in *PCR protocols: A guide to methods and applications*. Eds. M. Innis, J. Gel-fand and J. Sainsky (San Diego: Academic Press), 315–322.

Xia, W. L., Chen, K., Liu, W. S., Yin, Y., Yao, Q., Ban, Y., et al. (2022). Rapid and visual detection of *Mycoplasma synoviae* by recombinase-aided amplification assay combined with a lateral flow dipstick. *Poultry Sci.* 101 (7), 101860–101860. doi: 10.1016/j.psj.2022.101860

Yang, X. Q., Wen, X., Zhou, C. X., Zhu, X. J., Meng, R., Luo, Q.J, et al. (2018). Comparative study of brine shrimp bioassay-based toxic activities of three harmful microalgal species that frequently blooming in aquaculture ponds. *J. Oceanol. Limnol.* 36 (5), 1697–1706. doi: 10.1007/s00343-018-7140-7

Yang, X. H., Zhang, X., Wang, Y., Shen, H., Jiang, G., Dong, J. Q, et al. (2020). A real-time recombinase polymerase amplification method for rapid detection of *Vibrio vulnificus* in seafood. *Front. Microbiol.* 11. doi: 10.3389/fmicb.2020.586981

Yang, Q., Yang, H. Y., Yuan, N., Zuo, S. A., Zhang, Y. Z., and Zhang, W. (2022). Closed-tube saltatory rolling circle amplification with hydroxynaphthol blue for visual on-site detection of peanut as an allergenic food. *Food Chem.* 393, 133408–133408. doi: 10.1016/j.foodchem.2022.133408

Zamor, R. M., Glenn, K. L., and Hambright, K.D. (2012). Incorporating molecular tools into routine HAB monitoring programs: Using qPCR to track invasive prymnesium. *Harmful Algae* 15, 1–7. doi: 10.1016/j.hal.2011.10.028

Zhai, Y. D., Li, R. Q., Liu, F. G., Zhang, C. Y., Wang, Y. Y., and Chen, G. F. (2021). Recombinase polymerase amplification combined with lateral flow dipstick for the rapid detection of *Prorocentrum donghaiense*. *Mar. Biol. Res.* 17, 7–8. doi: 10.1080/17451000.2021.2010279

Zhao, G., Wang, J., Yao, C. Y., Xie, P. C., Li, X. M., Xu, Z. L, et al. (2022). Alkaline lysis-recombinase polymerase amplification combined with CRISPR/Cas12a assay for the ultrafast visual identification of pork in meat products. *Food Chem.* 383, 132318–132318. doi: 10.1016/j.foodchem.2022.132318

Zhu, P., Huang, H. L., Zhou, C. X., Xu, J. L., Qiao, L. L., Dang, C. Y, et al. (2019). Sensitive and rapid detection of *Prymnesium parvum* (Haptophyceae) by loop-mediated isothermal amplification combined with a lateral flow dipstick. *Aquaculture* 505, 199–205. doi: 10.1016/j.aquaculture.2019.02.059

frontiers | Frontiers in Marine Science

TYPE Methods
PUBLISHED 27 October 2022
DOI 10.3389/fmars.2022.988918



OPEN ACCESS

EDITED BY
Frederic Coulon,
Cranfield University, United Kingdom

REVIEWED BY
Inbakandan Dhinakarasamy,
Sathyabama Institute of Science and Technology, India
Conor McManus,
Rhode Island Department of Environmental Management, United States

*CORRESPONDENCE
Jackie L. Collier
jackie.collier@stonybrook.edu

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

RECEIVED 07 July 2022
ACCEPTED 14 October 2022
PUBLISHED 27 October 2022

CITATION
Geraci-Yee S, Allam B and Collier JL (2022) A nested quantitative PCR assay for detection of the hard clam pathogen *Mucochytrium quahogii* (=QPX) in environmental samples. *Front. Mar. Sci.* 9:988918. doi: 10.3389/fmars.2022.988918



# A nested quantitative PCR assay for detection of the hard clam pathogen *Mucochytrium quahogii* (=QPX) in environmental samples

Sabrina Geraci-Yee, Bassem Allam and Jackie L. Collier*

School of Marine and Atmospheric Sciences, Stony Brook University, Stony Brook, NY, United States

Progress in understanding and managing QPX disease outbreaks in hard clams (*Mercenaria mercenaria*) has been limited by lack of insight into basic aspects of the biology and ecology of the opportunistic pathogen *Mucochytrium quahogii* (formerly QPX or Quahog Parasite Unknown). One barrier is that while several methods have been able to detect *M. quahogii* in seawater and sediment, its abundance was typically too low for reliable quantification by those methods. Here we describe the development and validation of a sensitive, *M. quahogii*-specific, nested quantitative PCR (nqPCR) assay following the Minimum Information for Publication of Quantitative Real-Time PCR Experiments (MIQE) guidelines. The assay reaches the theoretical limit of detection (LOD) of a PCR assay at 3 copies per reaction with excellent efficiency, linearity, and minimal sample PCR inhibition. The functionality of the assay was evaluated by quantifying *M. quahogii* in sediment and seawater samples, which revealed that *M. quahogii* was broadly distributed throughout the marine environment, detected in 75% of samples, with mean estimated abundance of 0.21 cells per mg sediment, 0.55 cells per ml bottom seawater, and 0.02 cells per ml surface seawater. *M. quahogii* was most prevalent and most abundant in sediment and bottom seawater samples, suggesting that the flocculent layer at the sediment-water interface is an important environmental reservoir where *M. quahogii* may interact with hard clams. This assay will serve as a valuable tool to better understand QPX disease dynamics and offers a model to guide development of similar assays for other important marine microbes typically present at similarly low abundance.

KEYWORDS

qPCR, *Mucochytrium quahogii*, labyrinthulomycetes, *Mercenaria mercenaria*, QPX disease, low-abundance template, microbial detection methods, MIQE

# 1 Introduction

Quahog Parasite Unknown (QPX) disease in hard clams, *Mercenaria mercenaria*, has caused mass mortality events in both wild and cultured hard clams since the 1960s. Progress in our understanding of QPX disease outbreaks has been limited in part by poor understanding of the biology and ecology of the opportunistic thraustochytrid pathogen, first cultivated in the 1990's but only recently formally described as *Mucochytrium quahogii* (Geraci-Yee et al., 2021). The distribution and abundance of *M. quahogii* in the marine environment is poorly described and its preferred habitats are unknown, as is how QPX disease is transmitted to the hard clam host (Geraci-Yee et al., submitted; Geraci-Yee, 2021).

*M. quahogii* has been detected in marine aggregates by *in situ* hybridization and by PCR-based assays in sediment, seawater, and other environmental samples (Lyons et al., 2005; Gast et al., 2006; Lyons et al., 2006; Gast et al., 2008; Liu et al., 2009). For example, the *M. quahogii*-specific real-time quantitative PCR (qPCR) assay, used routinely for monitoring QPX disease in hard clams (Liu et al., 2009; Geraci-Yee et al., 2022), detected *M. quahogii* in 9% of sediment samples but in no seawater samples (Liu et al., 2009). Similar challenges have been encountered with efforts to quantify the natural abundance of other labyrinthulomycete genera using qPCR (Nakai et al., 2013). Greater sensitivity was achieved by the addition of extra PCR steps to create a nested PCR assay, which was able to detect *M. quahogii* in 40% of environmental samples from Massachusetts (Gast et al., 2008). This assay was sensitive but non-quantitative, being evaluated by denaturing gradient gel electrophoresis (DGGE) (Gast et al., 2006). Adding an additional PCR step to a qPCR assay (i.e., nested qPCR assay or nqPCR) can also increase the sensitivity and specificity of DNA amplification compared to a conventional single-step qPCR assay (Takahashi and Nakayama, 2006), and offers a path for quantifying the distribution of not only *M. quahogii* and other specific species of labyrinthulomycetes, but also other important but low-abundance marine microbes.

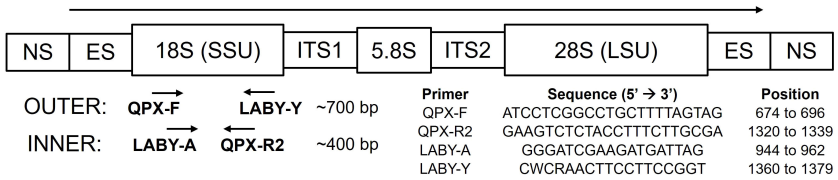
In a nested PCR assay, there are two (or more) sets of primers used and two (or more) rounds of thermal cycling. Typically, a set of primers spaced relatively far apart is used to amplify the target DNA during the first or outer step. Then another pair of primers is used in the second or inner step to amplify a fragment nested within the 'outer' fragment. A nested PCR can be used in a qPCR assay as a nested qPCR (nqPCR) by performing regular PCR reactions before using the product as template in a real-time qPCR assay (Takahashi and Nakayama, 2006; Banno et al., 2011). In order for the assay to be quantitative, the initial phase of amplification (first, outer step) must take place entirely in exponential phase, meaning the PCR reaction must not go to completion (Haff, 1994).

Here we describe the development of a nqPCR SYBR-based assay for sensitive detection and quantification of *M. quahogii* from marine environmental samples following the Minimum Information for Publication of Quantitative Real-Time PCR Experiments (MIQE) guidelines (Bustin et al., 2009; Bustin et al., 2010). We apply the nqPCR assay to a large set of sediment and seawater samples, confirming that *M. quahogii* is prevalent in the environment, as expected for an opportunistic pathogen, and that the *M. quahogii*-specific nqPCR assay is sensitive enough to quantify *M. quahogii* at its natural environmental abundance.

# 2 Methods

## 2.1 Assay development

Primers designed by Stokes et al. (2002) were used to develop the *M. quahogii*-specific nqPCR assay, illustrated in Figure 1 (Stokes et al., 2002; Qian et al., 2007). The first, outer PCR reaction used QPX-F and LABY-Y, which produce a product of ~700 bp, followed by LABY-A and QPX-R2 in the second, inner qPCR reaction, which produce a product of ~400 bp. Both PCR reactions use one *M. quahogii*-specific primer so the final product should be *M. quahogii*-specific. All quality control




| Primer | Sequence (5' → 3') | Position |
|---|---|---|
| QPX-F | ATCCTCGGCCTGCTTTTAGTAG | 674 to 696 |
| QPX-R2 | GAAGTCTCTACCTTTCTTGCGA | 1320 to 1339 |
| LABY-A | GGGATCGAAGATGATTAG | 944 to 962 |
| LABY-Y | CWCRAACTTCCTTCCGGT | 1360 to 1379 |

FIGURE 1
Schematic of the ribosomal RNA (rRNA) genes showing primers, respective locations and directions in which they prime, modified from Qian et al. (2007) with primers from Stokes et al. (2002). Primer positions are based on *M. quahogii* (QPX) sequence AY052644 in GenBank. NS, nontranscribed region; ES, external transcribed spacer; SSU, small subunit or 18S rRNA gene; ITS1, internal transcribed spacer 1; 5.8S, 5.8S rRNA gene; ITS2, internal transcribed spacer 2; LSU, large subunit or 28S rRNA gene.

measures described in Geraci-Yee et al. (2022) were followed, such as positive and negative controls, and at least triplicate replicate reactions. In addition, PCR setup and contamination procedures were augmented by including a new, separate workspace for setup of the outer reaction.

Under standard PCR, the outer and inner PCR reactions were tested using *M. quahogii* genomic DNA (gDNA) extracted from a geographically representative mixture of 4 isolates: two from New York (8BC7 (ATCC TSD-50) and 20AC1), one from Massachusetts (MA; ATCC 50749), and one from Virginia (VA), using the NucleoSpin Genomic DNA Tissue kit (Macherey-Nagel, Inc., Bethlehem, PA) and diluted to 0.5 ng/µl. In addition, linear plasmid DNA containing the *M. quahogii* rRNA region amplified by QPX-F and 28S46Rev (5'-ATATGCTTA ARTTCAGCGGGT-3'), developed by Liu et al. (2009) and described in Geraci-Yee et al. (2022), was also tested in the nested PCR design over a $10^5$ range of concentrations (10, $10^3$, and $10^6$ copies/µl). Primers were synthesized by Integrated DNA Technologies, Inc. (Coralville, IA). Each 20 µl PCR reaction contained 3.3 µl of nuclease-free water, 2 µl of 25 mM $MgCl_2$, 2 µl of 2 mM dNTPs, 2 µl of 2 µM forward and reverse primers (QPX-F and LABY-Y, final concentration 200 nM), 3.2 µl of 50% glycerol (8% final concentration), 4 µl of 5X GoTaq Flexi buffer (colorless), 0.5 µl GoTaq G2 Flexi DNA Polymerase (Promega, Madison, WI), and 1 µl of template DNA. The PCR program used to test the reactions was 35 cycles of 95°C for 30 seconds, 55°C for 1 minute, 72°C for 2 minutes (Liu et al., 2009). The same PCR parameters were used for both outer and inner PCR reactions. The inner PCR reaction contained 1 µl of PCR product from the outer reaction with the same PCR reagents as the outer reaction except the primers, LABY-A and QPX-R2. PCR products from the inner reaction were examined by agarose gel electrophoresis, which confirmed the expected amplicon size of 400 bp.

The PCR conditions for the outer reaction were optimized by standard approaches, according to Haff (1994). The outer reaction was performed as a qPCR reaction to determine the cycle limit for the exponential phase of amplification. Serial dilutions of plasmid DNA from 10 to $10^6$ copies were used as template and prepared in the same way as described in Geraci-Yee et al. (2022). The qPCR thermocycler used was the Mastercycler realplex$^4$ ep gradient S (Eppendorf, Hamburg, Germany) with realplex software (version 2.2) and twin.tec PCR 96 well, semi-skirted, colorless plates (Eppendorf, Hamburg, Germany), sealed with TempPlate RT Optical Film (USA Scientific, Ocala, FL). Each 12.5 µl reaction contained 6.25 µl of Takyon No Rox SYBR MasterMix dTTP Blue (Eurogentec, Fremont, CA), 0.75 µl of nuclease-free water, 1.25 µl of 2 µM forward and reverse primers (QPX-F and LABY-Y, final concentration 200 nM), 2 µl of 50% glycerol, and 1 µl of template DNA. For the first trial, the outer reaction qPCR used a PCR program of 10 minutes at 95°C, followed by 40 cycles of denaturation at 95°C for 30 seconds, anneal gradient from 50-60°C for 1 minute, and extension at 72°C for 2 minutes. Fluorescence was measured for individual wells at the end of each cycle with the detection threshold determined automatically by the qPCR software using the noiseband threshold and automatic baseline setting to determine quantification cycle or $C_q$ values (also known at $C_T$ values). Melt curve analysis as well as agarose gel electrophoresis were used to confirm the expected amplicon. The anneal gradient revealed a thermal optimum between 53 and 56°C with PCR efficiencies around 95%, with lower and higher anneal temperatures producing efficiencies below or above the 90-110% accepted range. Since the initial testing of the outer PCR reaction was done at an anneal of 55°C, we decided to use that anneal temperature for the remaining trials, using the PCR program: 10 minutes at 95°C, followed by 40 cycles of 95°C for 30 seconds, 55°C for 1 minute, and 72°C for 2 minutes. We performed primer titration to determine the lowest concentration of primers for the outer reaction. Primer concentrations tested were (forward/reverse, nM): 200/200, 100/100, 75/75, and 50/50. The only combination that had acceptable PCR efficiency was 200/200 nM of forward/reverse primer, with suppression of PCR efficiency observed with the lower concentrations. We also tested the addition of 1% DMSO to the master mix, which did not enhance the PCR efficiency or linearity. The outer reaction using 200 nM of primers was exponential for all concentrations of the standard curve until 25 cycles with a PCR efficiency of 95% and linearity of 0.951.

After testing the outer reaction as a qPCR, we went back to standard PCR to test the number of cycles (5, 10, 15, 20, and 25 cycles) using serial dilutions of linear plasmid DNA from 1 to $10^5$ copies and 3 µl of template to construct the standard curve, as 3 copies represents theoretical limit of detection (LOD) of a PCR reaction (Bustin et al., 2009). In addition, *M. quahogii* gDNA was used as a positive control. Each 20 µl PCR reaction contained 1.3 µl of nuclease-free water, 2 µl of 25 mM $MgCl_2$, 2 µl of 2 mM dNTPs, 2 µl of 2 µM forward and reverse primers (QPX-F and LABY-Y, final concentration 200 nM), 3.2 µl of 50% glycerol, 4 µl of 5X GoTaq Flexi buffer (colorless), 0.5 µl GoTaq G2 Flexi DNA Polymerase (Promega, Madison, WI), and 3 µl of template DNA (Supplementary File 1 Table S1). The PCR cycle was 95°C for 30 seconds, 55°C for 1 minute, 72°C for 2 minutes (Supplementary File 1 Table S2). After the outer reaction was performed, 1 µl of PCR product was used in the inner qPCR assay, where each reaction contained 6.25 µl of Takyon, 0.625 µl of 2 µM forward and reverse primers (LABY-A and QPX-R2, final concentration 100 nM), 2 µl of 50% glycerol, and 0.125 µl of DMSO (Supplementary File 1 Table S3). The real-time PCR program used was our standard program of 10 minutes at 95°C, followed by 40 cycles of 95°C for 30 seconds, 55°C for 1 minute, 72°C for 1 minute (Supplementary File 1 Figure S1). Melt curve analysis and agarose gel electrophoresis revealed a single peak at 81°C that coincided with the expected 400 bp amplicon. PCR efficiency and linearity were within acceptable limits and

therefore did not require additional optimization, compliant with the MIQE guidelines. The best PCR efficiency and linearity were observed after 10 and 15 cycles of the outer PCR reaction, so additional testing was performed at 10, 12, and 15 cycles. From these, the best PCR efficiency and linearity were observed at 12 cycles with an average efficiency of 96.5% and linearity of 0.985. We retested the reduction of primer concentration in the outer reaction to reduce potential carryover into the inner reaction; however, as before, this reduced PCR efficiency. The nqPCR conditions described here and in detail in Supplementary File 1 (detailed methods protocol: Tables S1-3, Figure S1) were the conditions used to construct the standard curve for quantification, determine analytical specificity and sensitivity, and evaluate functionality using environmental samples (seawater and sediment).

### 2.1.1 Analytical specificity and sensitivity

The nqPCR assay was tested for specificity against species belonging to each of the four groups comprising the cultivated labyrinthulomycetes: labyrinthulids (*Labyrinthula* sp. isolate KIE13), aplanochytrids (*Aplanochytrium stocchinoi* isolate GSB06), oblongichytrids (*Oblongichytrium* sp. isolate 606), and several thraustochytrids (*Aurantiochytrium limacinum* ATCC MYA-1381, *Schizochytrium aggregatum* ATCC28209, *Thraustochytrium aureum* ATCC34304, and *Japanochytrium marinum* ATCC 28207). gDNA was extracted using the NucleoSpin Genomic DNA Tissue kit (Macherey-Nagel, Inc., Bethlehem, PA), following the manufacturer's protocol for cultured cells, and diluted to 0.5 ng/µl for use as template DNA in the assay. Both melt curve and agarose gel electrophoresis analyses were used to evaluate the assay's specificity. The sensitivity or limit of detection (LOD) of the assay was tested by determining the number of failed reactions at the lowest concentration of the standard curve (3 copies). The standard curve was performed to validate the LOD of 3 copies using at least 6 replicates per run, totaling 42 reactions over 5 independent determinations.

### 2.1.2 Inter-run calibrator

Due to the large number of samples analyzed with this assay, it was impossible for all samples to be run on the same plate. Therefore, an inter-run calibrator (IRC), also known as an inter-plate calibrator (IPC), was used to correct run-to-run differences (Bustin et al., 2010) and a single, highly precise standard curve was used for quantification (Svec et al., 2015) following the general strategy described in Geraci-Yee et al. (2022). All $C_q$ values, including the standard curve, were corrected by subtracting the IRC from the individual run and adding the average IRC value for all plates (Equation 1) (Kubista, 2010). Varying concentrations (152, 15.2, and 1.52 pg/µl) of *M. quahogii* gDNA (from mixed isolates) run in at least triplicate were tested as the IRC to determine the most stable and robust calibrator. The IRC used was 152 pg/µl of *M. quahogii* gDNA.

$$C_q corrected = C_q measured - C_q calibrator + Average\ C_q calibrators \quad \text{(Eq. 1)}$$

## 2.2 Assay evaluation

Environmental samples (sediment and seawater) were collected from a variety of sampling sites across Long Island, New York, during 2014 and 2015 (Figure 2 and Supplementary File 2 Table S5), including a QPX-enzootic site in Barnstable, Massachusetts, as part of a large-scale hard clam field survey. Sediment was collected using a ponar sediment grab (0.04 $m^2$) and surface sediment (top 2 cm) was homogenized and stored in a sterile Whirl-Pak®. Surface (SSW) and bottom (BSW) seawater were collected using a 2 L Niskin bottle (General Oceanics, Miami, FL). All samples were put on ice for transport back to the laboratory and processed immediately within several hours. SSW and BSW samples were inverted gently to mix and up to 350 ml was filtered under low vacuum pressure "< 100 in Hg" on a 0.4 µm (47 mm) polycarbonate filter (GE Osmonics Inc., Minnetonka, MN). Seawater filters were carefully folded using sterile forceps and stored in cryovials. Seawater filters and sediment were stored at -80°C until DNA extraction using the MO BIO PowerSoil DNA isolation kit (Qiagen, Hilden, Germany) according to the manufacturer's protocol. For sediment samples, an average of 314 ± 61 (SD) mg sediment was used for DNA extraction. For BSW samples an average of 333 ± 41 (SD) ml and for SSW samples an average of 331 ± 50 (SD) ml was used for DNA extraction. Extracted DNA was quantified using Quant-iT PicoGreen (Molecular Probes, Eugene, OR). Most samples had low DNA concentrations, so we did not assess DNA purity with Nanodrop (Thermo Fisher Scientific, Wilmington, DE), as reliability of purity ratios below 20 ng/µl is often compromised. Sample DNA was stored at -20°C until being used in the nqPCR assay. Each sample was run in triplicate using 3 µl of template DNA in the outer reaction and 1 µl of the outer PCR product in the inner qPCR assay. $C_q$ values were accepted if the standard deviation was less than 0.5 for at least two of the three replicates; samples that did not meet this criterion were rerun until replicates agreed. Melt curve and agarose gel electrophoresis analysis were performed to ensure specificity ($T_m$ at 81°C and 400 bp amplicon). Furthermore, a representative set of nqPCR products were purified and Sanger sequenced using the inner reaction forward primer (LABY-A) to confirm assay specificity.

### 2.2.1 Inhibition testing

Due to the large number of samples analyzed, it was impractical to test each sample for inhibition. Therefore,

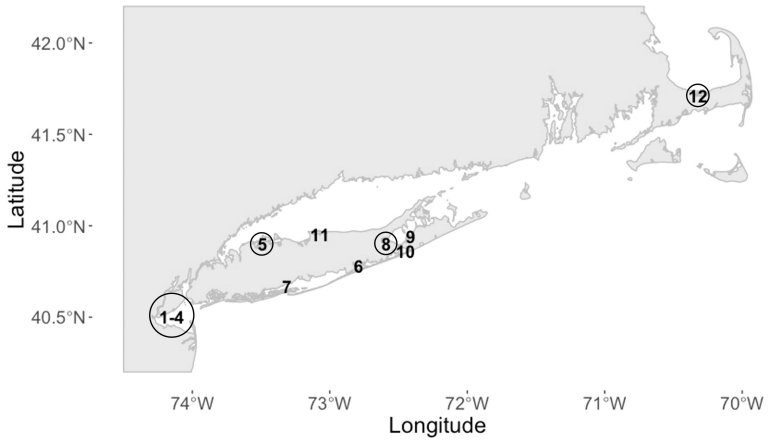



FIGURE 2
Location of sampling sites in Long Island, New York, and Cape Cod, Massachusetts (MA). 1-4 = Raritan Bay (RB), 5 = Oyster Bay (OB), 6 = Moriches Bay (MB), 7 = Babylon Bay (BB), 8 = Birch Creek (BC), 9 = Peconic Estuary (PE), 10 = Shinnecock Bay (SB), 11 = Port Jefferson Harbor (PJ), and 12 = Massachusetts (MA). Sites with a previous history of QPX disease are circled (RB, OB, BC, and MA; circle size is arbitrary). Location coordinates are provided in Supplementary File 2; Table S5.

inhibition testing was carried out on a representative set of samples. After nqPCR, at least 6 samples of each type (i.e., sediment, BSW, and SSW) were assessed, including negative and a range of positive samples from low to high copy number, representing samples from each sampling bay, season, and year. Since there were no samples with enough *M. quahogii* to perform the dilution series spanning several orders of magnitude needed to assess PCR amplification efficiency, 17 µl of sample template DNA from these samples were spiked with 2 µl of $10^7$ *M. quahogii* linear plasmid copies/µl (final concentration of $10^6$ copies/µl). The spiked sample template DNA was used to create a dilution series, in the same manner as the standard curve. *M. quahogii*-spiked sample PCR efficiency and linearity were determined, and the efficiency was considered acceptable if it was within 10% of the efficiency of the standard curve (Irwin et al., 2012) with a linearity of at least 0.98 (Johnson et al., 2013).

# 3 Results

## 3.1 Assay performance

We assessed assay performance as efficiency and linearity of the standard curve over 5 independent trials each with at least 6 technical replicates (Figure 3). The average PCR efficiency was 99.21% with a range of 96 to 101% and linearity > 0.988, in compliance with the MIQE guidelines (Bustin et al., 2009; Bustin et al., 2010). There were 2 failed reactions of 42 total reactions (4.76%) at the LOD of 3 copies, which is within the failed reaction limit (5%) for the LOD according to the MIQE guidelines. The $C_q$ standard deviation (SD) of replicates of the individual points on the standard curve from the 5 independent trials ranged from 0.36 to 0.89. These $C_q$ SDs represent the raw $C_q$ values not corrected with the IRC; therefore, these $C_q$ SDs are impacted by run-to-run variation. No amplification was observed in the qPCR or nqPCR negative control (no template added) after implementation of the quality control protocols.

The assay was specific for *M. quahogii* as revealed by specificity testing (Table 1). While there were late amplification of products (high $C_q$ values) for most of the labyrinthulomycetes tested, melt curve analysis and agarose gel electrophoresis revealed different melt temperatures ($T_m$) and product sizes. $C_q$ values for these samples ranged from 31.6 to 37.2 with large SDs (> 0.5) between technical replicates, suggesting that the PCR was heavily influenced by stochastic processes. Most of the labyrinthulomycetes tested had a product that was ~150 bp with a broad $T_m$ peak between 78-79°C, likely representing a PCR artifact, such as a large primer concatemer. Additionally, *Aurantiochytrium* and *Aplanochytrium* had a larger product (~450 bp) associated with a $T_m$ peak around 80°C. We think this is an rRNA amplification product produced in the absence of the *M. quahogii* target, when the leftover general labyrinthulomycete LABY-Y primer interacts with the nonspecific LABY-A primer in the inner qPCR reaction. These non-specific products were very faint, and we were unable to isolate and verify them by sequencing.

The mean $C_q$ value of the IRC across 15 independent runs (1 run for the standard curve and 14 runs with the environmental samples) was 14.3 with a SD of 0.34, indicating excellent inter-

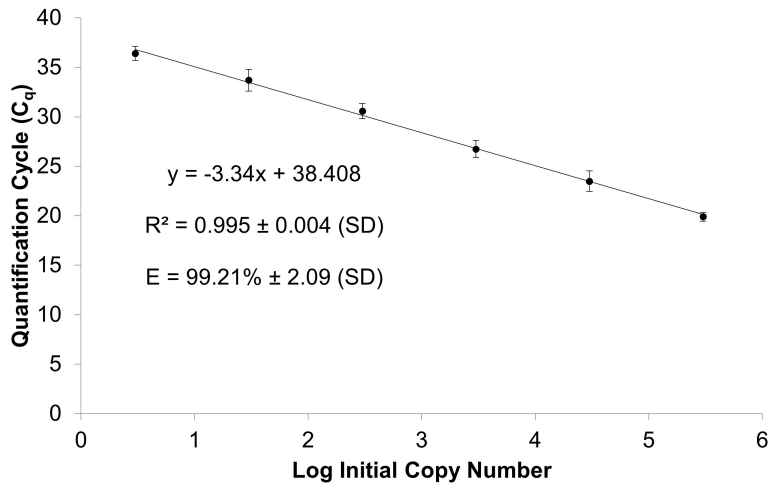



FIGURE 3
Standard curve of linearized plasmid serial dilutions containing 3 to 3 x $10^5$ *M. quahogii* copies averaged over 5 independent trials using at least 6 replicates per run with 95% confidence intervals (CI). The efficiency (E) and linearity ($R^2$) are expressed as the average and standard deviation (SD).

TABLE 1 Results from melt curve and gel electrophoresis analyses from the nqPCR assay's analytical specificity testing.

| Species | Group | # of Products | Melt Curve Tm °C | Product Size (bp) |
|---|---|---|---|---|
| *Mucochytrium quahogii* (QPX)* | Thraustochytrid | 1 | 81 | 400 |
| *Aurantiochytrium limacinum* (ATCC MYA-1381) | Thraustochytrid | 2 | 79, 80 | ~150, faint ~450 |
| *Schizochytrium aggregatum* (ATCC 28209) | Thraustochytrid | 1 | 79 | 150 |
| *Thraustochytrium aureum* (ATCC 34304) | Thraustochytrid | 1 | 79 | 150 |
| *Japanochytrium marinum* (ATCC 28207) | Thraustochytrid | 1 | 79 | 150 |
| *Oblongichytrium* sp. (isolate 606) | Oblongichytrid | 1 | 78 | 150 |
| *Aplanochytrium stocchinoi* (isolate GSB06) | Aplanochytrid | 2 | 79, 80 | ~150, faint ~450 |
| *Labyrinthula* sp. (isolate KIE13) | Labyrinthulid | 0 | none | none |

*mixed gDNA from several *M. quahogii* isolates.

run variation (reproducibility). A highly precise, robust standard curve consisting of 8 replicates was run and $C_q$ values were corrected using the IRC and Equation 1 (Supplementary File 1 Table S4 and Figure S4). The replicates of this standard curve had excellent intra-run variation (repeatability) with SDs less than 0.5 for each concentration. The standard curve gave an efficiency of 99.37% and linearity of 0.996. The $C_q$ corrected standard curve was used to convert IRC-corrected $C_q$ values into *M. quahogii* copy number using Equation 2. Using the IRC as a measure of assay reproducibility, the IRC raw $C_q$ values were translated into copy number using Equation 2 to determine the coefficient of variation (CV) between independent runs, which was 20%.

$$Copy\ \ number = 10^{\wedge}\left[\frac{C_q\ \ corrected - 37.881}{-3.3371}\right] \qquad \text{(Eq. 2)}$$

## 3.2 Assay evaluation

A total of 206 samples collected from 12 sampling sites from May to November during 2014 and 2015, comprising 71 sediment, 64 BSW, and 71 SSW samples, were analyzed with the nqPCR assay. The average DNA concentration (ng/µl) for sediment samples was 15.74 ± 6.74 (SD) with a range of 3.99 to 36.98, for BSW samples was 13.33 ± 8.49 (SD) with a range of 1.76 to 37, and for SSW samples was 13.13 ± 9.67 (SD) with a range of 1.18 to 58.48. Of the 206 samples, 154 (75%) were positive for *M. quahogii* by nqPCR, with 89% of sediment samples, 83% of BSW samples, and 56% of SSW samples positive (Figure 4). The average $C_q$ SD of replicate positive samples was 0.31 ± 0.19 (SD) indicating excellent intra-run variation (repeatability). Melt curve analysis of nqPCR products revealed a single peak at 81°C (Figure 5 and Supplementary File 1

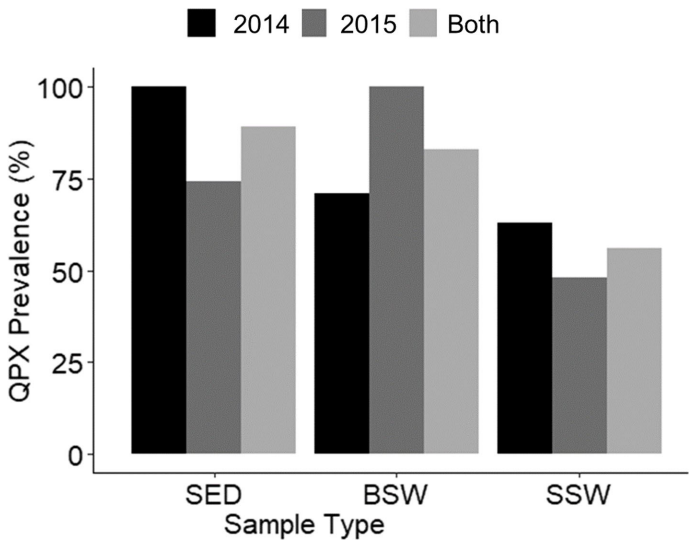



FIGURE 4
*M. quahogii* (QPX) prevalence (% positive samples) by sample type and year. SED represents sediment samples, BSW represents bottom seawater samples, and SSW represents surface seawater samples.

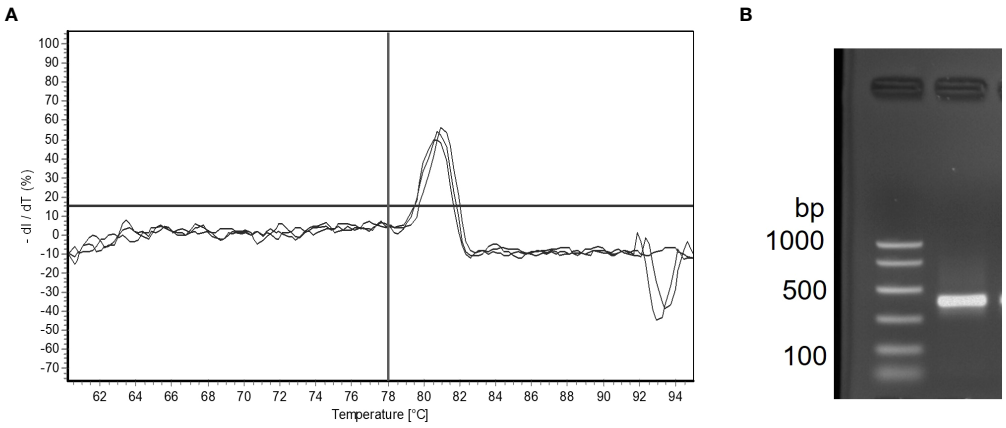



FIGURE 5
Melt curve analysis showing three technical replicates with one product melting at 81°C in a *M. quahogii* positive bottom seawater (BSW) sample **(A)**. The horizontal line is oriented at 0% dI/dT and the vertical line is oriented at 78°C. Gel electrophoresis analysis for the same sample showing one product at ~400 bp **(B)**. The nqPCR product was sequenced and was > 99% identical to *M. quahogii* (QPX) sequences in GenBank.

Figure S2), which was the same observed $T_m$ of both *M. quahogii* gDNA and plasmid standards. From the positive samples, 6 samples from each sample type (n=18) were randomly selected and visualized by agarose gel electrophoresis, from which the target amplicon was gel-purified and Sanger sequenced. All 18 samples had > 99% sequence identity to *M. quahogii* (QPX) sequences in NCBI GenBank. Average sample PCR efficiency and linearity were 91.8% ± 3.3 (SD) and 0.99 ± 0.007 (SD) (Table 2), suggesting that inhibition was minimal with PCR efficiencies within 10% of the efficiency of the standard curve used for quantification.

Close examination of amplification and melt curves revealed that some samples had linear instead of exponential quantification curves with a broad melt curve peak ($T_m$ ~72° C). The linear amplification curves were easily identified as they also crossed the threshold very early (low $C_q$ value;

TABLE 2 Mean ± standard deviation (SD) of PCR efficiency (E) and linearity ($R^2$) of spiked sample serial dilutions for inhibition testing by sample type from a mix of representative samples.

| Sample Type | N | Efficiency (E) | Linearity ($R^2$) |
|---|---|---|---|
| Sediment | 6 | 92.33 ± 1.97 | 0.992 ± 0.004 |
| Bottom Seawater | 6 | 90.54 ± 2.14 | 0.992 ± 0.007 |
| Surface Seawater | 6 | 92.45 ± 5.22 | 0.985 ± 0.008 |
| All | 18 | 91.80 ± 3.3 | 0.990 ± 0.007 |

Supplementary File 1 Figure S3). Agarose gel electrophoresis revealed these samples had a smear, sometimes with the target *M. quahogii* band. These samples were rerun with either 1 µl of template DNA in the outer PCR reaction (as opposed to 3 µl) or 1:10 and/or 1:100 dilutions of the outer PCR product in the inner qPCR. Either reduction of the template DNA in the outer PCR reaction or dilution of the outer reaction PCR product was effective in producing exponential amplification, and removed the smear when visualized in a gel, leaving the target *M. quahogii* amplicon if the sample was positive; in the case of negative samples, the smear was effectively removed with no amplification observed. The estimated *M. quahogii* concentrations of the 1:10 and 1:100 dilutions for individual samples were in agreement. Of the 206 samples tested, 165 (80%) did not require adjustment to the protocol, while 5 samples (2%) required 1 µl of template DNA used in the outer PCR reaction, plus 24 (12%) samples required a 1:10 dilution and 12 (6%) samples required a 1:100 dilution of the outer PCR product for the inner qPCR reaction. These samples either contained too much DNA or PCR product for the inner qPCR assay to function properly or promoted an artifactual interaction between the general labyrinthulomycete primers. Average sample DNA concentration was slightly greater for the sediment and BSW samples that required adjustment to the protocol (e.g., less template or dilution of outer product) in comparison to the samples that did not require adjustment to the protocol (sediment mean DNA concentration 17.53 ng/µl, n=12 vs. 15.49 ng/µl, n=59 and BSW 18.68 ng/µl, n=23 vs. 10.24 ng/µl, n=41; for SSW, DNA concentration was similar between samples that required adjustment or did not: 13.6 ng/µl, n=6 v. 13.09 ng/µl, n=65). Occasionally, some *M. quahogii*-negative samples had late amplification (high $C_q$ values past the LOD) with broad melt curves and a $T_m$ between 77-79°C, sometimes accompanied by the $T_m$ of the smear (~72°C). When visualized on a gel, the amplicon was smaller (100-150 bp) than the *M. quahogii* amplicon (~400 bp) and very faint, similar to the non-specific products found during specificity testing (Table 1).

From the mean corrected $C_q$ value of each sample (Equation 1), the *M. quahogii* (QPX) copy number was determined using Equation 2, and the number of copies per mg sediment or ml seawater was determined using Equation 3, where 100 µl represents the DNA elution volume.

$$\mathrm{M.\ } quahogii \frac{copies}{mg\ \ or\ \ ml} = \frac{Copy\ \ number \times 100\ \mu l \times Dilution\ \ Factor}{Sample\ \ weight\ (mg)\ \text{or}\ Volume\ (ml) \times DNA\ \ template\ \ (\mu l)} \quad \text{(Eq. 3)}$$

The concentration of *M. quahogii* ranged from below the detection limit of the assay (or negative) to 703 copies/mg sediment, 2,980 copies/ml BSW, and 176 copies/ml SSW with mean concentrations of 93 copies/mg sediment, 241 copies/ml BSW, and 7 copies/ml SSW (Figure 6 and Table 3). We converted *M. quahogii* gene copies to cellular abundance, using 440 copies/mononucleate cell as described in Geraci-Yee et al. (2022), assuming that the rDNA copy number of *M. quahogii* cells in culture is similar to that of *M. quahogii* in the environment. Aspects of the *M. quahogii* nqPCR assay for MIQE compliance are summarized in our modified version of the MIQE checklist (Table 4) (Bustin et al., 2009; Bustin et al., 2010; Geraci-Yee et al., 2022).

## 3.3 M. quahogii in New York waters

*M. quahogii* was detected at every sampling time point in at least one of the environmental sample types, except at BC on July 14, 2015 (sediment and SSW were negative and there was not a BSW sample since the site can be shallow depending on the tides). *M. quahogii* was detected in 100% of sediment samples in 2014 and in 100% of BSW samples in 2015 (Figure 4), and although *M. quahogii* appeared more prevalent (% positive samples) in sediment and BSW than SSW samples, the difference was not significant (Kruskal-Wallis rank sum test). The abundance of *M. quahogii* varied by orders of magnitude among samples of each type, suggesting that while *M. quahogii* is widespread, it has very patchy spatiotemporal distribution (Figure 6). Coinciding with the difference in sediment and BSW prevalence between years was significantly greater *M. quahogii* abundance in sediment in 2014 than in 2015 (p=6.25E-11) and significantly greater *M. quahogii* abundance in BSW in 2015 than in 2014 (p=6.24E-05) (Wilcoxon rank sum test; Supplementary File 2 Table S6 and Figure S5). Environmental metadata revealed than 2014 was a "wetter"

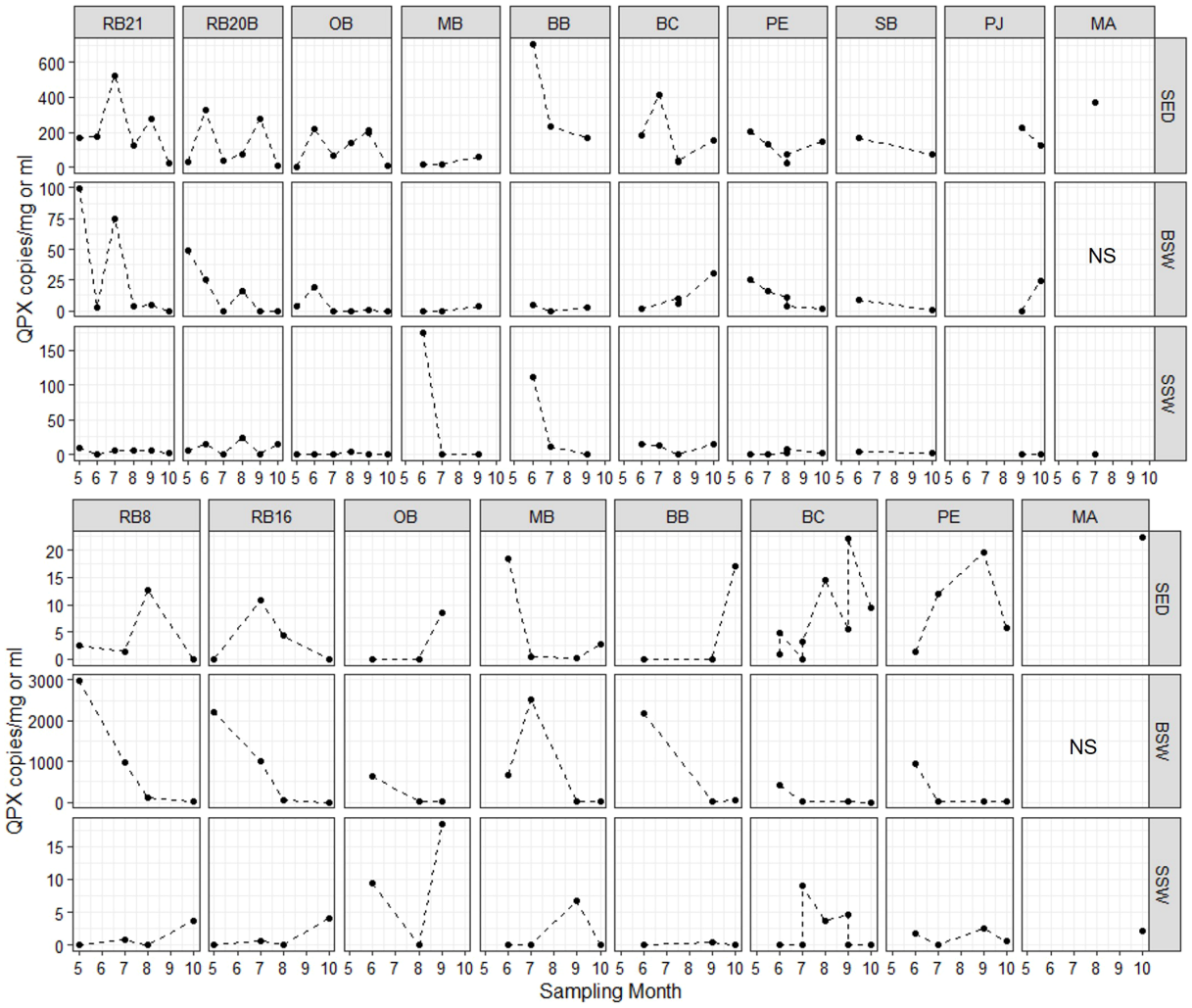



FIGURE 6
*M. quahogii* (QPX) abundance in copies per mg sediment or ml seawater as determined by the nqPCR assay. Note the varying scales used in the y-axis per each row. SED, sediment; BSW, bottom seawater; and SSW, surface seawater; NS, no sample (intertidal site).

TABLE 3 Summary statistics of *M. quahogii* (QPX) concentration in 18S gene copies per mg sediment or ml seawater, as well as theoretical conversion to cellular concentration using 440 copies per mononucleate cell.

| Sample Type | N | % | QPX copies/mg or ml | | | QPX cells/mg or ml | | |
|---|---|---|---|---|---|---|---|---|
| | | Positive | Mean | SD | Range | Mean | SD | Range |
| Sediment | 71 | 89 | 93 | 135 | 0.27 – 702.61 | 0.211 | 0.31 | 0.0006 – 1.6 |
| BSW | 64 | 83 | 241 | 630 | 0.61 – 2,980 | 0.548 | 1.43 | 0.0013 – 6.77 |
| SSW | 71 | 56 | 7.4 | 24.6 | 0.43 – 176.37 | 0.0168 | 0.06 | 0.001 – 0.4 |

(more freshwater, lower salinity) year compared to 2015 (less freshwater, higher salinity) (Supplementary File 2; Table S7), which may account for the observed differences, as changes in precipitation can affect many other environmental factors, including salinity, delivery of terrestrial organic matter, and production of marine organic matter, particularly for thraustochytrids (Ueda et al., 2015; Geraci-Yee, 2021). Otherwise, there were no clear patterns in distribution of *M.*

TABLE 4 Modified MIQE checklist for the nqPCR assay.

| ASSAY CHECKLIST | *Mucochytrium quahogii* (QPX)-specific nqPCR Assay |
|---|---|
| **Sample/Template** | |
| Source | Sediment and seawater processed within 24 hours after collection |
| Method of preservation | Frozen at -80°C immediately after processing |
| Storage time (if appropriate) | DNA extracted within 1 year after freezing |
| Extraction method | MO BIO PowerSoil DNA isolation kit (Qiagen, Hilden, Germany) |
| DNA storage and time | DNA stored at -20°C until assayed within 1 year |
| Concentration/Purity | Sediment 15.74 ng/µl ± 6.74 (SD); Bottom seawater (BSW) 13.33 ng/µl ± 8.49 (SD); Surface seawater (SSW) 13.13 ng/µl ± 9.67 by PicoGreen; Purity by Nanodrop was not determined due to low DNA concentration |
| Inhibition assessment | Acceptable within 10% of the standard curve PCR efficiency; Dilution series on sample template DNA matrix spiked with QPX plasmid performed on a representative set of samples (18 samples, 6 of each type) with E = 91.80% ± 3.3 (SD) and $R^2$ = 0.99 ± 0.007 (SD); Table 2 |
| **Assay Optimization/Validation** | |
| Sequence accession number | AY052644, QPX isolate from Massachusetts (Stokes et al., 2002); Figure 1 |
| Amplicon details | 18S region 400 bp; Figure 1 |
| Primer sequence | Outer = QPX-F and LABY-Y; Inner = LABY-A and QPX-R2 (Stokes et al., 2002); Figure 1 |
| *In silico* | See Stokes et al., 2002 |
| Empirical | See Methods 'Assay Development' |
| Priming conditions | Summarized in Supplementary File 1 Table S2 and Figure S1 |
| Specificity | Single Tm at 81°C Figure 4; Tested against 7 other labyrinthulomycetes Table 1 |
| PCR efficiency | PCR efficiency = 99.21% ± 2.09 (SD); Figure 3 |
| Linear dynamic range | 3 to 3 x $10^5$ copies, $R^2$ = 0.995 ± 0.004 (SD); Figure 3 |
| Limits of detection (LOD) | 3 copies per reaction |
| Intra-assay variation (repeatability) | Average Cq SD of triplicate replicates of samples = 0.31 ± 0.19 (SD) |
| Inter-assay variation (reproducibility) | IRC Cq SD = 0.34, copy number CV = 20% across 15 independent runs |
| **PCR** | |
| Protocols | See Methods and Supplementary File 1 detailed methods protocol |
| Thermocycler | Mastercycler realplex$^4$ p gradient S (Eppendorf) |
| Reagents | Takyon No Rox SYBR MasterMix dTTP Blue (Eurogentec); detailed in Supplementary File 1 Tables S1 and S3 |
| Negative control | No amplification of qPCR NTC and nqPCR NTC (no Cq or Tm) |
| Positive control | 0.152 ng of QPX gDNA (used as IRC) |
| Replicates | Triplicate |
| **Data Analysis** | |
| Software | Eppendorf realplex software (version 2.2), noiseband threshold and automatic baseline setting |

*quahogii*. There were no significant differences in *M. quahogii* abundance by QPX disease history (sites with and without a history of QPX disease in hard clams), sampling site, or sampling month (Supplementary File 2; Table S6). Moreover, *M. quahogii* abundance was not significantly correlated to measured environmental metadata: seawater temperature, salinity, or dissolved oxygen (Spearman's rank order method).

# 4 Discussion

The MIQE-compliant *M. quahogii*-specific nqPCR developed in this study is sensitive enough to detect and quantify *M. quahogii* in sediment and seawater samples against the background of other labyrinthulomycetes. Although the final amplicon size is larger than traditionally used in qPCR assays, adjustment to the extension time and use of a SYBR-based assay demonstrates that the assay functions properly within the MIQE guidelines and accepted qPCR practices (Bustin et al., 2009; Bustin et al., 2010). The nested design enabled the detection of *M. quahogii* in 75% of the environmental samples tested, representing a significant increase from the non-nested qPCR assay used by Liu et al. (2009), which was only able to detect *M. quahogii* in 4 of 43 sediment samples (9%) and none of 40 seawater samples, even though we cannot rule out the remote possibility that abundance of *M. quahogii* may have changed in the environment over the decade that separates both studies. *M. quahogii* prevalence in our 2014-2015 environmental samples from New York (89% of sediment, 83% of BSW samples, and 56% of SSW) was similar to the 55% of seawater and 46% of sediment samples from QPX disease enzootic sites in

Massachusetts that were positive using a non-quantitative nested PCR assay (Gast et al., 2006; Gast et al., 2008). Although the *M. quahogii* nqPCR assay did occasionally produce non-specific products, this only happened in the absence of the *M. quahogii* target and could be easily identified by melt curve analysis (lower $T_m$ compared to the *M. quahogii* $T_m$) and/or agarose gel electrophoresis. Given the many rounds of PCR (52 total cycles) and use of four primers, including two general labyrinthulomycete primers, the non-specific products could arise from PCR artifacts such as large primer dimers or from real amplification of non-target rRNA genes. The smaller artifact bands (100-150 bp) in the environmental nqPCR were similar to those seen in the specificity testing using other labyrinthulomycetes (Table 1).

For the few *M. quahogii*-positive sediment samples, Liu et al. (2009) calculated concentrations ranging from 35 to 215 *M. quahogii* cells/mg sediment with a mean of 117.5 cells/mg ± 88.4 (SD), which are much higher than the estimates obtained from this nqPCR assay (Table 3). As discussed by Geraci-Yee et al. (2022), the Liu et al. (2009) values likely represent overestimation derived from the use of circular plasmid for the standard curve, lower copy number conversion to cellular concentration (181 vs. 440 copies/cell), and correction factors applied for PCR inhibition and DNA recovery rate of the extraction method. Due to a combination of improvements in our methods, from DNA extraction to following the MIQE-recommended dilution series approach to determine PCR efficiency (revealing negligible PCR inhibition, Table 2) rather than an alien spike (Bustin et al., 2009; Bustin et al., 2010), we can now avoid such potentially inflationary correction factors. With their greater PCR inhibition, Liu et al. (2009) more often needed to dilute the template DNA, which may sometimes have reduced the *M. quahogii* target DNA to below the detection limit of the assay, generating false negatives.

Measured by nqPCR, *M. quahogii* abundance was highly variable but typically less than 1 cell/ml or mg in seawater and sediment (Table 3, Figure 6), with maximums of 0.4 cells/ml SSW, 6.77 cells/ml BSW, and 1.6 cells/mg sediment, suggesting that *M. quahogii* does not exist at high densities in environmental samples. Supporting these cellular concentration estimates are the low abundances found in Gast et al. (2008) by *in situ* hybridization in macrophytes and scraping of detritus with less than 5 *M. quahogii* cells per slide for each sample examined. Taken qualitatively, this suggests that *M. quahogii* is not particularly abundant even when associated directly with potential substrates like macrophytes and detritus.

To our knowledge only one study has attempted to quantify individual genera of labyrinthulomycetes. Nakai et al. (2013) were able to quantify two genera, *Aurantiochytrium* and *Oblongichytrium*, from seawater in coastal Japan. Mean (± SD) concentration of *Aurantiochytrium* and *Oblongichytrium* was 12.15 ± 1.55 cells/ml and 16.75 ± 2 cells/ml seawater, respectively. Although these estimates are greater than our *M. quahogii* abundance estimates for seawater samples (Table 3), the difference could reflect that both *Aurantiochytrium* and *Oblongichytrium* comprise many species. The other 5 labyrinthulomycete genera that Nakai et al. (2013) designed qPCR assays for were not successfully quantified from any of their seawater samples. Moreover, labyrinthulomycetes were able to be quantified by qPCR in only 8 of 212 (3.8%) seawater samples, despite the fact that labyrinthulomycetes were present in 104 additional samples when quantified by acriflavine direct detection method (AfDD), a common method used to enumerate labyrinthulomycetes (Raghukumar and Schaumann, 1993). Low and sporadic abundance pose a challenge for investigating the ecology not only of specific taxa for labyrinthulomycetes, but for other microorganisms as well. Our study demonstrates that nqPCR offers a potential solution for quantification of low-abundance microbes, which has been used widely in diagnostic applications for human pathogens, such as malaria (Tran et al., 2014) and tuberculosis (Takahashi and Nakayama, 2006), as well as important plant pathogens (Banno et al., 2011; Coy et al., 2014). We believe that this is the first time that a nqPCR assay has been used for the quantification of a marine pathogen, representing a novel approach that can be applied to ecological investigations of other marine microorganisms where conventional qPCR assays, such as those used in Nakai et al. (2013), have limited success. The only pitfall that may occur when using the nqPCR approach is the increased risk of contamination due to the additional manipulation of DNA template and extremely high sensitivity of the nested reactions. However, there are well-established protocols for minimizing PCR contamination, such as the use of filter tips and separate rooms or areas for each step (Zehr and Turner, 2001; Takahashi and Nakayama, 2006), which were utilized here and further described in Geraci-Yee et al. (2022) and Supplementary File 1 (detailed methods protocol).

We have described the development of a *M. quahogii* (QPX)-specific nqPCR assay that is MIQE-compliant and sensitive enough to quantify *M. quahogii* from natural samples, amid a background of related organisms, as demonstrated by assaying 206 environmental samples, of which 75% were positive. Such information is needed to identify potential reservoirs in which *M. quahogii* may be able to survive and grow outside its hard clam host, as a pathogen with an environmental reservoir has the potential to cause widespread disease because it is not limited by its host's density (Harvell et al., 2004; Burge et al., 2013). Results from this assay represent new insights into both prevalence and abundance of *M. quahogii* in the environment, suggesting that *M. quahogii* may be associated with the flocculent layer at the sediment-water interface, making interaction with hard clams likely as they live in the same habitat buried in the sediment. Besides the interannual differences found in the distribution of *M. quahogii* in sediment and bottom seawater, there were no clear patterns in the distribution or abundance of *M. quahogii* in the environment, as well as no significant correlations with environmental metadata. Other environmental conditions not measured in this study may more strongly determine *M. quahogii* distribution and abundance, offering avenues for future research. Nevertheless, this assay

represents a valuable tool to better understand *M. quahogii* dynamics in the environment and possible relationships with QPX disease outbreaks, and provides a model to guide the development of similar assays for other important marine microbes typically present at similarly low abundance.

## Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding author.

## Author contributions

JC and BA secured funding to develop the methodology and perform the field survey. SGY developed and validated the method, and wrote the manuscript with guidance from JC and BA. All authors contributed to the article and approved the submitted version.

## Funding

This research was supported by projects R/FBM-36 and R/XG-32 funded by the National Sea Grant College Program of NOAA to the Research Foundation of State University of New York on behalf of New York Sea Grant.

## Acknowledgments

We are grateful to the members of the Marine Animal Disease Laboratory (MADL) and Collier Lab for Microbial Ecology at Stony Brook University for assistance with sample processing, including but not limited to Dr. Ewelina Rubin, Dr. Mariana Rius, and Rachel Hartman. We also thank Raymond Czaja for assistance with creating a map of the sampling locations. Lastly, we thank the New York State Department of Environmental Conservation (NYSDEC), specifically Captain Todd Smith, Dr. Soren Dahl, Wade Carden, and Debra Barnes, for field and administrative support. We are also thankful to Joshua Reitsma (Cape Cod Cooperative Extension) for providing samples from Massachusetts.

## 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.2022.988918/full#supplementary-material

## References

Banno, S., Saito, H., Sakai, H., Urushibara, T., Ikeda, K., Kabe, T., et al. (2011). Quantitative nested real-time PCR detection of verticilliumlongisporum and v. dahliae in the soil of cabbage fields. *J. Gen. Plant Pathol.* 77 (5), 282–291. doi: 10.1007/s10327-011-0335-9

Burge, C. A., Kim, C. J., Lyles, J. M., and Harvell, C. D. (2013). Special issue oceans and humans health: the ecology of marine opportunists. *Microb. Ecol.* 65 (4), 869–879. doi: 10.1007/s00248-013-0190-7

Bustin, S. A., Beaulieu, J. F., Huggett, J., Jaggi, R., Kibenge, F. S., Olsvik, P. A., et al. (2010). MIQE précis: Practical implementation of minimum standard guidelines for fluorescence-based quantitative real-time PCR experiments. *BMC Mol. Biol.* 11, 74. doi: 10.1186/1471-2199-11-74

Bustin, S. A., Benes, V., Garson, J. A., Hellemans, J., Huggett, J., Kubista, M., et al. (2009). The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin. Chem.* 55 (4), 611–622. doi: 10.1373/clinchem.2008.112797

Coy, M., Hoffmann, M., Gibbard, H. K., Kuhns, E., Pelz-Stelinski, K., and Stelinski, L. (2014). Nested-quantitative PCR approach with improved sensitivity for the detection of low titer levels of candidatus liberibacter asiaticus in the Asian citrus psyllid, diaphorina citri kuwayama. *J. microbiological Methods* 102, 15–22. doi: 10.1016/j.mimet.2014.04.007

Gast, R. J., Cushman, E., Moran, D. M., Uhlinger, K. R., Leavitt, D., and Smolowitz, R. (2006). DGGE-based detection method for quahog parasite unknown (QPX). *Dis. Aquat Organ* 70 (1-2), 115–122. doi: 10.3354/dao070115

Gast, R. J., Moran, D. M., Audemard, C., Lyons, M. M., DeFavari, J., Reece, K. S., et al. (2008). Environmental distribution and persistence of quahog parasite unknown (QPX). *Dis. Aquat Organ* 81 (3), 219–229. doi: 10.3354/dao01948

Geraci-Yee, S. (2021). Taking the "X" out of QPX disease: Distribution and dynamics of the hard clam pathogen, *Mucochytrium quahogii* (formerly QPX = quahog parasite unknown). *[Ph.D.] Stony Brook NY: Stony Brook University*.

Geraci-Yee, S., Allam, B., and Collier, J. L. (2022). Keeping up with advances in qPCR pathogen detection: an example for QPX disease in hard clams. *Dis. Aquat Organ* 148, 127–144. doi: 10.3354/dao03648

Geraci-Yee, S., Brianik, C. J., Rubin, E., Collier, J. L., and Allam, B. (2021). Erection of a new genus and species for the pathogen of hard clams 'Quahog

parasite unknown' (QPX): *Mucochytrium quahogii* gen. nov., sp. nov. *Protist* 172 (1), 125793. doi: 10.1016/j.protis.2021.125793

Haff, L. A. (1994). Improved quantitative PCR using nested primers. *Genome Res.* 3, 332–337. doi: 10.1101/gr.3.6.332

Harvell, D., Aronson, R., Baron, N., Connell, J., Dobson, A., Ellner, S., et al. (2004). The rising tide of ocean diseases: unsolved problems and research priorities. *Front. Ecol. Environ.* 2 (7), 375–382. doi: 10.1890/1540-9295(2004)002[0375:TRTOOD]2.0.CO;2

Irwin, P. L., Nguyen, L. H., Chen, C. Y., Uhlich, G. A., and Paoli, G. C. (2012). A method for correcting standard-based real-time PCR DNA quantitation when the standard's polymerase reaction efficiency is significantly different from that of the unknown's. *Anal. Bioanal Chem.* 402 (9), 2713–2725. doi: 10.1007/s00216-012-5737-9

Johnson, G., Nolan, T., and Bustin, S. A. (2013). ""Real-time quantitative PCR, pathogen detection and MIQE," in *PCR detection of microbial pathogens: Second editions, methods in molecular biology*. Ed. M. Wilks (Totowa, NJ: Springer).

Kubista, M. (2010). TATAA interplate calibrator SYBR protocol. *User Manual tataabiocenter*.

Liu, Q., Allam, B., and Collier, J. L. (2009). Quantitative real-time PCR assay for QPX (Thraustochytriidae), a parasite of the hard clam (*Mercenaria mercenaria*). *Appl. Environ. Microbiol.* 75 (14), 4913–4918. doi: 10.1128/AEM.00246-09

Lyons, M. M., Smolowitz, R., Dungan, C. F., and Roberts, S. (2006). Development of a real time quantitative PCR assay for the hard clam pathogen quahog parasite unknown (QPX). *Dis. Aquat. Organisms* 72, 45–52. doi: 10.3354/dao072045

Lyons, M. M., Smolowitz, R., Uhlinger, K. R., Gast, R. J., and Ward, J. E. (2005). Lethal marine snow: Pathogen of bivalve mollusc concealed in marine aggregates. *Limnology Oceanography* 50, 1983–1988. doi: 10.4319/lo.2005.50.6.1983

Nakai, R., Nakamura, K., Jadoon, W. A., Kashihara, K., and Naganuma, T. (2013). Genus-specific quantitative PCR of thraustochytrid protists. *Mar. Ecol. Prog. Ser.* 486, 1–12. doi: 10.3354/meps10412

Qian, H., Liu, Q., Allam, B., and Collier, J. L. (2007). Molecular genetic variation within and among isolates of QPX (Thraustochytridae), a parasite of the hard clam *Mercenaria mercenaria*. *Dis. Aquat. Organisms* 77, 159–168. doi: 10.3354/dao01848

Raghukumar, S., and Schaumann, K. (1993). An epifluorescence microscopy method for direct detection and enumeration of the fungi like marine protist, the thraustochytrids. *Limnology Oceanography* 38, 182–187. doi: 10.4319/lo.1993.38.1.0182

Stokes, N. A., Calvo, L. M. R., Reece, K. S., and Burreson, E. M. (2002). Molecular diagnostics, field validation, and phylogenetic analysis of quahog parasite unknown (QPX), a pathogen of the hard clam *Mercenaria mercenaria*. *Dis. Aquat. Organisms* 52, 233–247. doi: 10.3354/dao052233

Svec, D., Tichopad, A., Novosadova, V., Pfaffl, M. W., and Kubista, M. (2015). How good is a PCR efficiency estimate: Recommendations for precis and robust qPCR efficiency assessments. *Biomolecular Detection Quantification* 3, 9–16. doi: 10.1016/j.bdq.2015.01.005

Takahashi, T., and Nakayama, T. (2006). Novel technique of quantitative nested real-time PCR assay for *Mycobacterium tuberculosis* DNA. *J. Clin. Microbiol.* 44 (3), 1029–1039. doi: 10.1128/JCM.44.3.1029-1039.2006

Tran, T. M., Aghili, A., Li, S., Ongoiba, A., Kayentao, K., Doumbo, S., et al. (2014). A nested real-time PCR assay for the quantification of plasmodium falciparum DNA extracted from dried blood spots. *Malaria J.* 13 (1), 1–8. doi: 10.1186/1475-2875-13-393

Ueda, M., Nomura, Y., Doi, K., Nakajima, M., and Honda, D. (2015). Seasonal dynamics of culturable thraustochytrids (Labyrinthulomycetes, stramenopiles) in estuarine and coastal waters. *Aquat. Microbial Ecol.* 74, 187–204. doi: 10.3354/ame01736

Zehr, J. P., and Turner, P. J. (2001). ""Nitrogen fixation: nitrogenase genes and gene expression.","," in *Methods in microbiology*. Ed. J. H. Paul (St. Petersburg, FL: San Diego: Academic Press), 271–286.

frontiers | Frontiers in Microbiology

TYPE Original Research
PUBLISHED 16 December 2022
DOI 10.3389/fmicb.2022.1031439



OPEN ACCESS

EDITED BY
Frederic Coulon,
Cranfield University,
United Kingdom

REVIEWED BY
Surekha K. Satpute,
Savitribai Phule Pune University, India
Simona Bartkova,
Tallinn University of Technology, Estonia

*CORRESPONDENCE
Robert Duran
robert.duran@univ-pau.fr
Fumihiro Sassa
sassa@ed.kyushu-u.ac.jp

†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 30 August 2022
ACCEPTED 28 November 2022
PUBLISHED 16 December 2022

CITATION
Duran C, Zhang S, Yang C, Falco ML, Cravo-Laureau C, Suzuki-Minakuchi C, Nojiri H, Duran R and Sassa F (2022) Low-cost gel-filled microwell array device for screening marine microbial consortium. *Front. Microbiol.* 13:1031439. doi: 10.3389/fmicb.2022.1031439



# Low-cost gel-filled microwell array device for screening marine microbial consortium

Clelia Duran[1†], Shiyi Zhang[2†], Chongyang Yang[3], Maria Lorena Falco[1†], Cristiana Cravo-Laureau[1], Chiho Suzuki-Minakuchi[3,4], Hideaki Nojiri[3,4], Robert Duran[1*] and Fumihiro Sassa[2*]

[1]Universite de Pau et des Pays de l'Adour, E2S UPPA, CNRS, IPREM, Pau, France, [2]Graduate School of Information Science and Electrical Engineering, Kyushu University, Fukuoka, Japan, [3]Agro-Biotechnology Research Center, Graduate School of Agricultural and Life Sciences, The University of Tokyo, Tokyo, Japan, [4]Collaborative Research Institute for Innovative Microbiology, The University of Tokyo, Tokyo, Japan

In order to exploit the microbes present in the environment for their beneficial resources, effective selection and isolation of microbes from environmental samples is essential. In this study, we fabricated a gel-filled microwell array device using resin for microbial culture. The device has an integrated sealing mechanism that enables high-density isolation based on the culture of microorganisms; the device is easily manageable, facilitating observation using bright-field microscopy. This low-cost device made from polymethyl methacrylate (PMMA)/polyethylene terephthalate (PET) has 900 microwells (600μm×600μm×700μm) filled with a microbial culture gel medium in glass slide-sized plates. It also has grooves for maintaining the moisture content in the micro-gel. The partition wall between the wells has a highly hydrophobic coating to inhibit microbial migration to neighboring wells and to prevent exchange of liquid substances. After being hermetically sealed, the device can maintain moisture in the agarose gels for 7days. In the bacterial culture experiment using this device, environmental bacteria were isolated and cultured in individual wells after 3days. Moreover, the isolated bacteria were then picked up from wells and re-cultured. This device is effective for the first screening of microorganisms from marine environmental samples.

KEYWORDS

**gel-filled microwell array, bacterial screening, high-throughput, microbial consortium, microbial isolation method**

## Introduction

To exploit the numerous beneficial microbes present in our environment, effective and rapid selection of beneficial microorganisms with useful functions from the vast number of microbes inhabiting our planet is essential; more than $5 \times 10^{30}$ prokaryotic cells (Wooley et al., 2010) and more than 1 trillion ($10^{12}$) microbial species (Locey and Lennon, 2016) are estimated to exist on the earth. The utilization of microorganisms as a natural resource has gained

attention in many fields, including human health, agro-industry, environmental remediation processes, and chemical production using biotechnology. Moreover, microbes have been successfully utilized for the development of various products and industrial processes, such as the development of new anti-HIV (Banyal et al., 2021) and anti-cancer drugs, bioethanol production using highly functional yeast (Mohd Azhar et al., 2017), and establishment of wastewater treatment plants using microbial consortia (Cydzik-Kwiatkowska and Zielinska, 2016). Thus, microbial resource-based processes have become the key technologies in this century; they are particularly important in the fields of clean energy, biotechnology, drug discovery, and environmental technology.

The conventional method for screening microorganisms involves a microbial suspension extracted from an environmental sample (soil or any other natural sample), which is subjected to selection pressure on an agar plate containing adequate medium and screening components according to the criteria of microbial selection under controlled conditions, including constant temperature and oxygen levels. Then, pure microbial strains, collected from the colonies formed on the solid medium plate, are cultured for variable periods (from overnight to several days) in order to evaluate their metabolic functions. However, this approach has some disadvantages because it is time-consuming and requires several days, and the number of cultures that can be processed by hand is limited to several dozens. In addition, due to the competition among the different microorganisms present in an environmental sample, it is difficult to isolate microorganisms with low growth rate because of competitive inhibition by microorganisms with fast growth rate (Jiang et al., 2016).

In recent years, microfabrication techniques, such as Bio-Micro Electro Mechanical System (BioMEMS) and microfluidics have been used to develop microbial research tools (Weibel et al., 2007). Various devices have been developed for microbial culturing using microfluidic technology, such as a combination of optical tweezers (Enger et al., 2004), a tool for the evaluation of microbial motility by chemical gradient (Walker et al., 2005), and disease model devices by co-culture with human cells (Shah et al., 2016). Some devices can handle single-cell bacteria with sub-micrometer scale microfluidic structures (Balaban et al., 2004; Boedicker et al., 2009; Mannik et al., 2009; Moffitt et al., 2012; Taheri-Araghi et al., 2015). In particular, an advanced micro-liquid manipulation system using a poly dimethylpolysiloxane (PDMS) micro-pneumatic valve (Unger et al., 2000; Thorsen et al., 2002) has shown high versatility. It can be used for individual bacterial cell isolation and selective manipulation of solutions containing medium (nanoliter scale), as well as digital PCR (Ottesen et al., 2006) and microchemostat culture (Balagadde et al., 2005). However, it requires a special and expensive setup, including a large external drive pneumatic pump array system to drive each system.

Abbreviations: BioMEMS: Bio-micro electro mechanical system; LB: Lysogeny broth; MSM: Minimal salt medium; OD: Optical density.

On the other hand, high-throughput and low-cost microbial isolation and culture systems using droplets (liquid plugs) have been proposed (Boedicker et al., 2008; Walter et al., 2011; Guo et al., 2012; Jiang et al., 2016; Kaminski et al., 2016). However, it is difficult to track the growth of cells in individual droplets because the position of the droplet is indefinite. Therefore, statistical batch processing using a flow cytometer is widely used for analysis and bacteria isolation (Li et al., 2021). In addition, the aqueous phase (the droplet) and the oil phase surrounding the droplet are both liquids that allow the exchange of soluble substances between the droplets since it is difficult to completely block such exchange *via* the carrier fluid (oil phase), even if with the addition of appropriate surfactant to reduce the it (Courtois et al., 2009; Chowdhury et al., 2019).

A number of devices for culture evaluation in microenvironments using microfabrication techniques and microbial isolation culture array techniques, such as microbial cell patterns utilizing soft lithography, have been proposed (Cerf et al., 2008; Schirhagl et al., 2012; Arnfinnsdottir et al., 2015). In particular, devices that combine the color-coded droplet technique with microwell arrays enabled high throughput combinatorial co-culture analysis (Kehe et al., 2019). Moreover, large-scale bacterial culture techniques involving small gel spots of microbial suspensions formed on large gel plates (Srinivasan et al., 2013) and separation devices using micro through-hole array device placed on gel plates (Ingham et al., 2007) have been proposed. Since these devices can handle a large number of isolated microorganisms in independent compartments, individual growth can be tracked sequentially through microscopic observation, and all wells can be directly accessed for isolation. In recent years, agarose-made gel microwell arrays formed by soft lithography techniques that can be used for microorganism isolation (Zhang et al., 2019) and a technique for photo addressable recovery of isolates from wells based on photodegradable polyethylene glycol hydrogel membranes have been developed (Barua et al., 2020).

However, these devices are directly connected to a common gel plate at the bottom to prevent the drying of the gel throughout the duration of the culture, and it is difficult to prevent the exchange and diffusion of liquid substances *via* the plate as well as the droplet system. In a previous study, we developed a gel-filled microwell array device using photolithography that can be used for culture for more than 18 h without exchange or diffusion of liquid substances (Nomura et al., 2014). It contains a completely independent micro-gel medium with 3,900 wells per $cm^2$. We used spacers and sealing technique for maintaining the small volume of the culture vessel, thus moisture could be prevented from entering the picoliter-scale gel. However, it was difficult to culture the cells for longer than one night with this device due to gel drying, which was a trade-off for the small size of the gel-filled wells. Moreover, the device was not cost-effective for one-time use due to the materials used for its fabrication, such as Pyrex glass, and photolithography processes. Furthermore, the sealing of the device required trained individuals.

In this study, we have developed a low cost, easy to use gel-filled microwell array device for long term microbial culturing. The device was fabricated using inexpensive plastic materials and fabrication methods.

# Materials and methods

## Bacterial strains and culture media

In order to determine whether bacteria can grow on the gel-filled microwell array device, we used *Pseudomonas putida* KT2440 (Bagdasarian et al., 1981) harboring the plasmid pBBR1MCS-5::mCherry, which possesses a red fluorescent protein (mCherry) encoding gene at the BamHI site of pBBR1MCS-5 (Kovach et al., 1995). We also used microorganisms from surface sediment marine samples (around the first 5 cm) collected at the Marina Port of Anglet (France) under a low tide. The sediments contained 31 ± 5 mg/kg of total petroleum hydrocarbon. An enrichment culture with acetate (4 g/l) as carbon source obtained from the hydrocarbon-contaminated marine sediments was used for the experiment.

The composition of the minimal salt medium (MSM) used in this study is as follows (g/L): KCl, 0.428; $CaCl_2.2H_2O$, 0.84; $NH_4Cl$, 0.015; $MgSO_4.7H_2O$, 3.794; $MgCl_2.6H_2O$, 3.017; NaCl, 15.14; and $Na_2CO_3$, 0.15. After sterilization, the medium was supplemented with 100 µl of 50 mM phosphate buffer solution; 1 ml of trace element stock solution with the following composition (g/L): $Fe(NH_4)_2(SO_4)_2.6H_2O$, 0.2; $Na_2SO_3$, 0.2; $CoCl_2.6H_2O$, 0.1; $MnSO_4.2H_2O$, 0.1; $Na_2MoO_4.2H_2O$, 0.1; $ZnSO_4.7H_2O$, 0.1; $AlCl_3.6H_2O$, 0.04; $NiCl_2.6H_2O$, 0.025; $H_3BO_3$, 0.01; $CuSO_4.5H_2O$, 0.01; and 1 ml of V7 vitamin stock solution with the following composition (mg per 200 ml): vitamin B12, 10.0; p-aminobenzoic acid, 10.0; D-biotin, 2.0; acid nicotinic, 20.0; calcium pantothenate, 5.0; pyridoxine. HCl, 50.0; thiamine. $HCl.2H_2O$, 10.0. Initial pH of the medium was adjusted to 7.0–7.2 with sterile 10 M NaOH. Filter sterilized trace element and V7 solutions were added to the sterile MSM as described in previous studies (Kovach et al., 1995; Alvarez-Barragan et al., 2021). The composition of the lysogeny broth (LB) is as follows (g/L): tryptone, 10; yeast extract, 5; NaCl, 10 (Sambrook, 2001). Gentamicin (Gm; 15 µg/ml) was added to the LB when needed.

## Structure of gel-filled microwell array device

The schematics of the developed device are shown in Figure 1A. The device consists of a well-array plate, an airtight sealing lid made of polymethyl methacrylate (PMMA) for observation, and a sealing frame made of polyethylene terephthalate (PET) that serves as a spacer (Figure 1B). The footprint shape and size of the device are similar to that of a standard glass slide (76 × 26 mm); it can be mounted on a standard glass slide holder of a microscope. The thickness of the device is 3.7 mm after the three parts are stacked. The well-array plate is made of PMMA (thickness: 1.5 mm), and one plate has a total of 900 gel-filling wells, which are divided into four subareas with 15 × 15 wells each (size of a well: 600 µm × 600 µm; depth: 700 µm). Scales for visual observation are present around each well area. The top surface on the partition wall between the wells is coated with a highly hydrophobic material (contact angle of 110° or higher) to avoid cross-contamination caused by the transfer of bacteria between the wells.

In addition, rectangular gel-filling grooves (size: 65 mm × 1.5 mm; depth: 0.5 mm) are present on both sides of the four well areas. These grooves are filled with the same gel as that used for the medium to keep the micro-gel in the wells from drying on both sides of the device seal during incubation. The outer edge of the plate has a groove (1.8 mm) for sealing frame (0.5 mm depth), which keeps the spacing between the observation lid and the culture gel surface constant; it also maintains humidity inside the device. The sealing frame made from PET is rectangular in shape (width: 1.5 mm; height: 1.7 mm). This frame is designed to fit into the groove of the well-array plate described above. The lid for observation and the airtight seal are made of 1.5 mm thick PMMA and have a groove (0.5 mm) for fitting and assembling the sealing frame as well as the well-array plate. They can be easily stacked and disassembled. If necessary, the outer edge of the assembled device can be sealed with a removable adhesive (Paper Bond Ta-100, KOKUYO, Japan) or polyimide tape to increase airtightness.

## Fabrication of microwell array device

The device was fabricated using a $CO_2$ laser engraving machine (LaserPRO C180II, GCC, Taiwan). The surfaces of PMMA (1.5 mm) and PET (1.7 mm) plates were cleaned with clean room grade wipes (BEMCOT PS-2, OZU CORPRATION, Japan). Then, they were set in a laser cutting machine, cut, and processed. Next, the device was brushed under running water to remove the dust formed during processing and then dried. The well-array plate was coated with a highly hydrophobic coating agent (FS-1060TH-2.0, Fluorotechnology, Japan) only on the top surface of the device to make the partitions between the wells hydrophobic; then the device was dried and thermally treated at 100°C for curing. Finally, to remove water-soluble impurities, the well-array plate was immersed in pure water overnight, and then dried. After assembly, the three completed parts were wrapped in aluminum foil and stored in a cool and dark place.

## Filling gel to microwell array device

After filling the gel medium into the wells, the fabricated device was used for culturing microorganisms. In the experiments described in this paper, the gel-filling process was performed before the inoculation step. Initially, the gel to be filled into the wells was

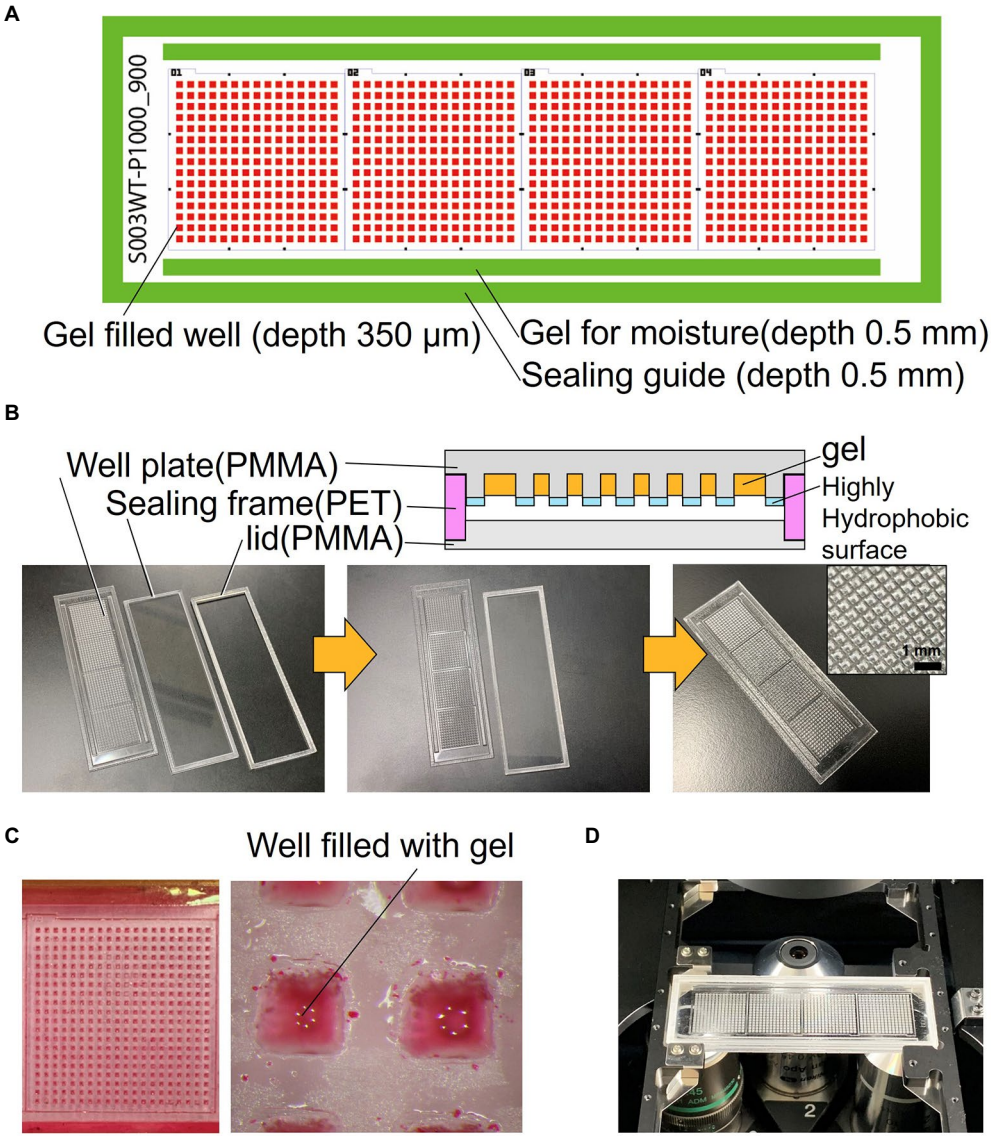



FIGURE 1
Structure of the resin-made gel-filled microwell array. **(A)** 2D arrangement of the device components. **(B)** Cross sectional view and photograph of the device. **(C)** Micrograph of gel-filled wells. **(D)** Photograph of the device mounted on microscope.

melted in an autoclave. Next, the gel was poured into a glass Petri dish and kept warm at 60°C on a hot plate, and the stored device was disassembled and the well-forming plate was removed. The plate was immersed in pure water in a beaker at room temperature, the beaker was placed in a vacuum desiccator, and the pressure was reduced to approximately 1 kPa using a vacuum pump; air bubbles in the wells expanded due to the reduced pressure and were released from the wells. Consequently, the pressure returned to normal. Next, the wells of the plate were filled with pure water and immersed in the molten medium gel. To replace pure water with molten gel, the setup was not disturbed for 3 min. The Petri dish with the plate was placed in the vacuum desiccator again, and the pressure was reduced to remove the air bubbles present inside the well. The plate was then removed from the molten gel and cooled down for solidification in air. The remaining layer of the gel formed on the top surface of the plate was removed using polypropylene film and a PDMS spatula so that the water-repellent coating is exposed at the partitions between the wells. The plate was immersed in liquid medium, which had the same composition as that of the gel medium except for the gel component, at room temperature to prevent drying; the plate can be stored for several hours until the next step. Almost all of these operations were carried out on a clean bench.

## Inoculation and cultivation of bacteria

The schematics of the inoculation and cultivation of bacteria are shown in Figure 2A. To cultivate *P. putida* KT2440

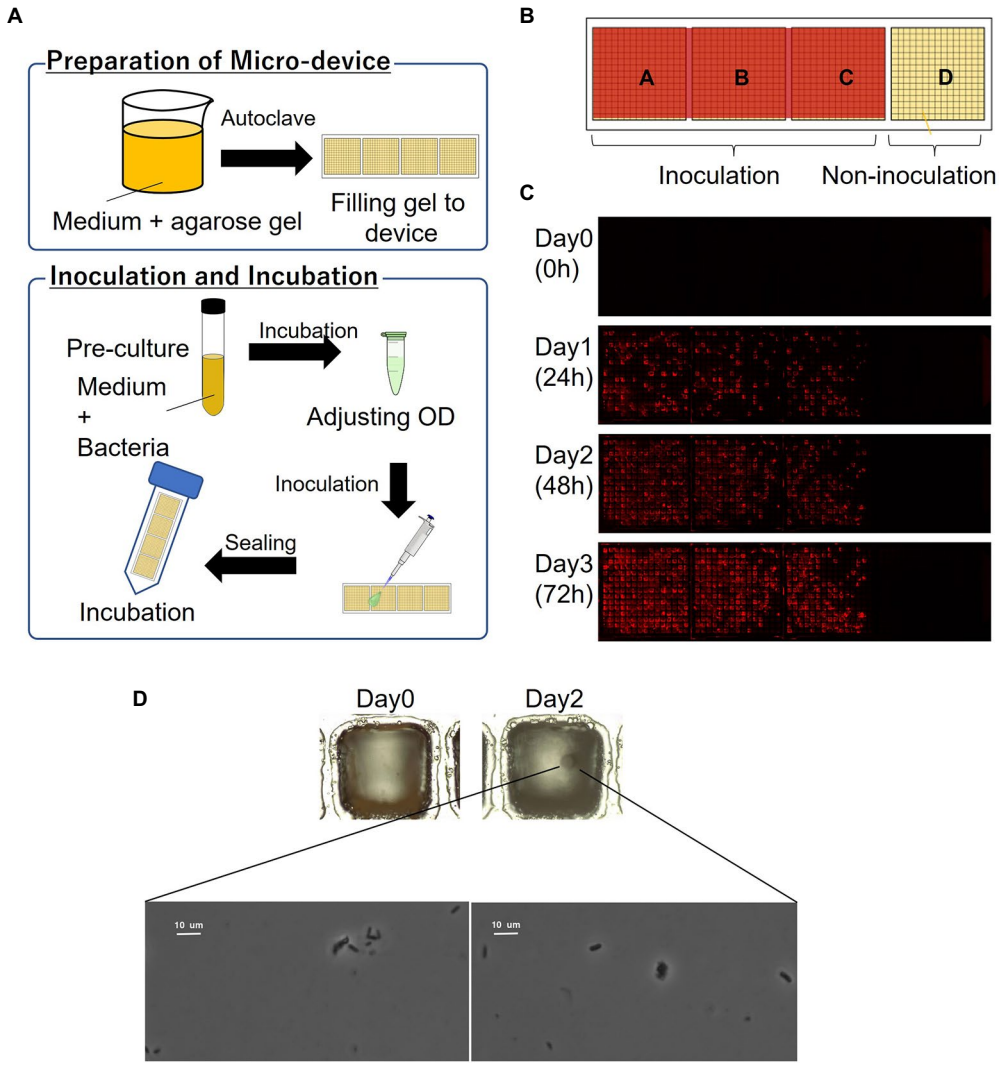



FIGURE 2
Culturing microorganisms in the gel-filled microwell array device. **(A)** Schematics of the culturing procedure using the device. **(B)** Arrangement of inoculation area on the device. **(C)** Micrographs of a gel-filled well just after inoculation to after 72h of incubation. **(D)** Bacterial colony and bacterial cells from marine hydrocarbon-contaminated sediments obtained using gel-filled microwell array device. Microscopic observation of bacterial colonies (top; magnification ×100) and bacterial cells (down; magnification ×1,000).

(pBBR1MCS-5::mCherry), LB containing Gm with 1% agarose was used to fill the wells of the device. The strain was pre-cultivated in LB supplemented with Gm at 30°C for 14–16 h. The optical density (OD) at 600 nm of the bacterial suspension was adjusted to 0.1 and diluted 10,000 times before inoculation. The resultant bacterial suspension was dispensed onto the gel-filled device by pipettes, and then removed. The partitions between the wells were hydrophobic so that the liquid remains as droplets only on the hydrophilic gel surface in each well. To remove the suspension liquid protruding from the few wells, the wells were air-blown and then sealed. The device was placed in a 50-mL conical tube with a wet paper towel and incubated at 30°C. mCherry-expressing cells were visualized by fluorescence microscopy.

To transfer the cells from the device into the liquid media, the wells on the device where mCherry-expressing cells existed were stabbed by a needle. The cells were inoculated into 96-deep well plates containing 1 ml of LB supplemented with Gm in each well. The plates were incubated at 30°C with shaking for 40 h. The growth in each well was evaluated by OD at 600 nm.

For the microorganisms from the marine samples, MSM (Dias et al., 2008; Alvarez-Barragan et al., 2021) containing 4 g/l sodium acetate as carbon source with 1% agarose was used to fill the wells of the device. The microbial suspension was adjusted to OD 0.1 at 560 nm and diluted 10,000 times before inoculation. The inoculation was performed as described above, and the device was incubated at 30°C.

## Results and discussion

Figure 1C shows the device filled with gel using the method described in Section 2. To enable easy observation, the gel (1% agarose) used for the experiments was mixed with red food dye. Figure 1D shows the device sealed with a lid and mounted on a microscope. Most of the wells were filled with gel, and no gel was left in the partitions between the wells. Microscopic images of a typical gel-filled well from the lid side and the bottom side are shown in Figures 3A,B, respectively. The height of the gel was slightly lower than the height of the well. On this surface, microbial colonies can be optically observed even by bright-field microscopy.

To evaluate the moisture retaining property of this device, the device filled with gel was stored at 30°C. Figure 3C shows the changes in the gel-filled device after 7 days. After 7 days of storage, the volume of the gel decreased by 74% ($n = 6$), as observed by measuring the volume of the gel by changes in the

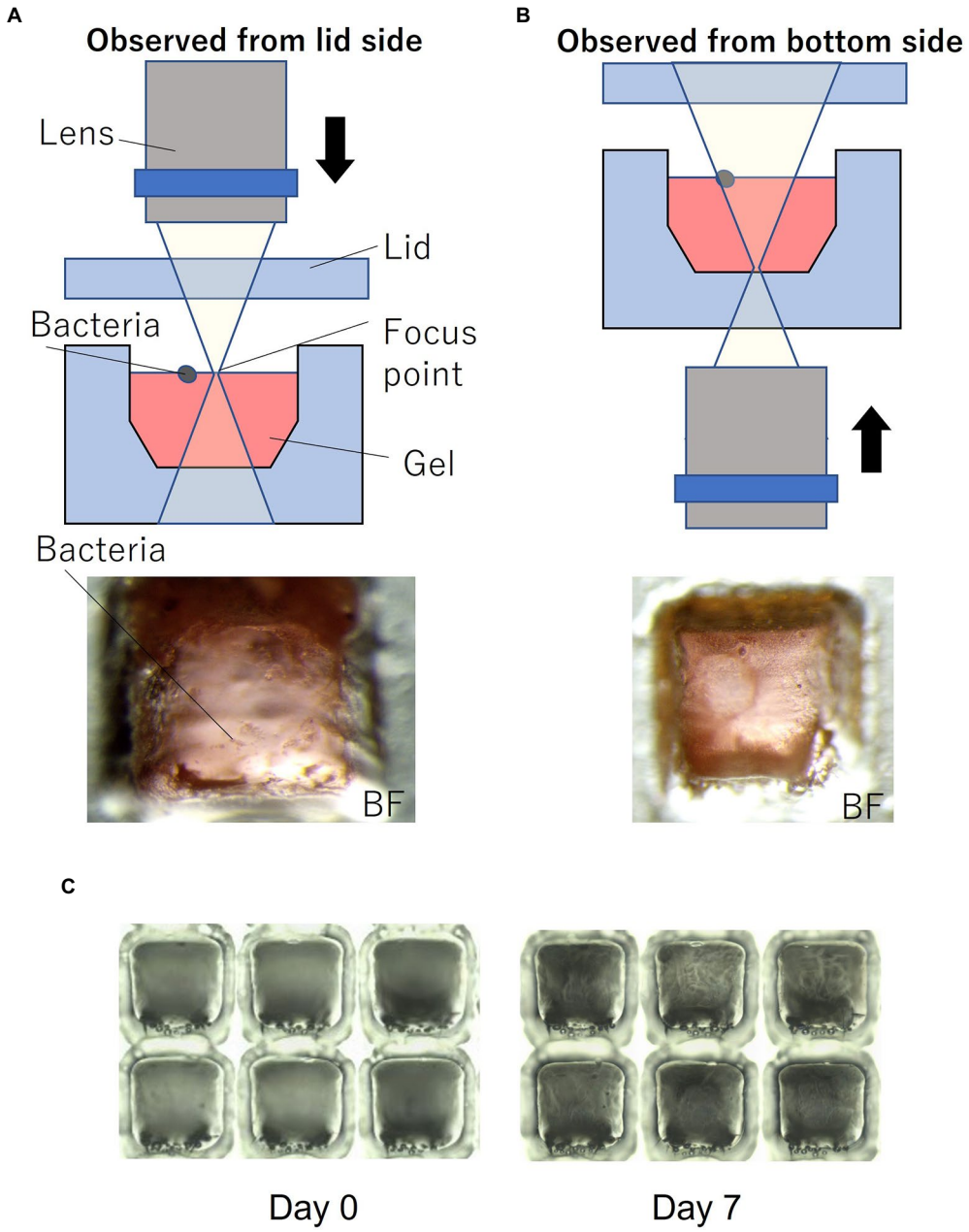



FIGURE 3
Microorganisms on the gel-filled microwell array. **(A)** Observation set up and micrograph from lid side. **(B)** Observation set up and micrograph from bottom side. **(C)** Micrograph of changes in the gel-filled device after 7days at 30°C.

distance from the focus point on the gel surface to the lid surface.

In order to evaluate the basic characteristics of the fabricated gel-filled device, a microbial culture experiment was conducted using *P. putida* KT2440 (pBBR1MCS-5::mCherry). The schematics of the procedure are shown in Figure 2A. By adjusting the concentration of the microbial suspension appropriately, each well was randomly inoculated with microbial cells with zero to several, often one. When a concentration of particle suspension is randomly inoculated into an array of a constant-size well, the number of bacteria entering each well is theoretically expected to follow a Poisson distribution (Hansen et al., 2016). In this experiment, the inoculated concentration was determined by a preliminary experiment in which the concentration of the suspension sample was adjusted. As shown in Figure 2B, a compartment of the device was not inoculated and acted as a control. A micrograph of gel-filled wells just after inoculation to after 72 h of incubation is shown in Figure 2C. The microbial colonies expressing mCherry were observed on the gel surface of the wells after incubating the device for 24 h. In addition, several wells without microbial growth (isolated wells) were observed, indicating that the movement of the microorganisms between the wells was inhibited. From the device incubated for 72 h, 46 wells were stabbed by a needle and the cells were transferred into a 96-deep well plate containing liquid medium. Bacterial growth was observed in 98% of the wells (Table 1), confirming that it is possible to pick up and re-cultivate the bacteria grown on this device.

Then we applied this method to cultivate microorganisms from marine samples. An enrichment culture obtained from the hydrocarbon-contaminated marine sediments was used as the microbial suspension for inoculation. MSM (Dias et al., 2008; Alvarez-Barragan et al., 2021) containing acetate as carbon source with 1% agarose was filled on the device. A single bacterial colony was observed on the gel surface of a well after incubating the device for 2 days (Figure 2D). Then, the colony was picked and the bacterial cells were observed under a microscope (Figure 2D). Thus, the proposed culturing method using the new device allowed us to obtain isolated colony formed by viable bacterial cells, showing that our approach can be applied for high-throughput screening of cultivable microbial cells from stock marine microbial communities.

Enrichment culture is often used to select beneficial microorganisms from environmental samples. Many microorganisms with valuable functions, such as the degradation of xenobiotic compounds, have been obtained by enrichment culture. This method selects fast-growing microorganisms, which can be dominant in the specific competitive conditions used for the enrichment process. These selected microorganisms, however, are not always dominant when applied in the actual environment. To obtain various microbes, it is useful to avoid competition in the consortia. One possible method is spreading the microbial suspension before enrichment on the selective solidified media using Petri dishes, though it is time- and space-consuming. The gel-filled microwell array device developed herein showed a possibility of solving this problem. The 900 gel-filled wells on the glass slide-sized

TABLE 1 Growth of *P. putida* KT2440 (pBBR1MCS-5::mCherry) transferred from the gel-filled well-array device into a 96-deep well plate.

| No. | OD at 600 nm |
|---|---|
| 1 | 1.263 |
| 2 | 1.013 |
| 3 | 1.036 |
| 4 | 1.177 |
| 5 | 1.107 |
| 6 | 1.073 |
| 7 | 1.278 |
| 8 | 1.073 |
| 9 | 1.114 |
| 10 | 1.150 |
| 11 | 1.173 |
| 12 | 1.114 |
| 13 | 1.258 |
| 14 | 1.057 |
| 15 | 1.001 |
| 16 | 1.029 |
| 17 | 1.142 |
| 18 | 1.073 |
| 19 | 1.098 |
| 20 | 1.150 |
| 21 | 1.144 |
| 22 | 1.067 |
| 23 | 1.154 |
| 24 | 1.110 |
| 25 | 1.075 |
| 26 | 1.084 |
| 27 | 1.127 |
| 28 | 1.067 |
| 29 | 1.065 |
| 30 | 1.129 |
| 31 | 1.131 |
| 32 | 1.246 |
| 33 | 1.127 |
| 34 | 1.070 |
| 35 | 1.238 |
| 36 | 1.115 |
| 37 | 1.358 |
| 38 | 1.079 |
| 39 | 1.020 |
| 40 | 1.118 |
| 41 | 0.986 |
| 42 | 0.996 |
| 43 | 1.103 |
| 44 | 1.095 |
| 45 | 0 |
| 46 | 1.147 |

plate work as independent solidified media and can be easily used for the first screening of target microbes from environmental samples. We should note that this method is applicable only to microorganisms able to grow on gel-solidified media. Moreover, the compounds added to the media should be soluble in water; it is challenging to use highly hydrophobic compounds, such as polycyclic aromatic hydrocarbons, on this device as for any aqueous medium. Recently, Nakai and colleagues succeeded in the isolation of a phylogenetically novel *Rhodospirillales* bacterium using the prototype of the device used in this study (Nakai et al., 2021). This indicates that the device reported here has the potential to isolate a microorganism that has been overlooked by conventional screening methods.

## Conclusion

In this study, we fabricated a glass slide-sized gel-filled microwell array device from resin for microbial culture screening. The device was fabricated by laser processing of PMMA and PET, and it is a stand-alone container with a lid that can be hermetically sealed. The device has a total of 900 independent gel-filled wells, and the partitions between the wells are coated with a highly hydrophobic barrier to prevent cross-contamination between the wells. The device consists of a small number of parts and is designed to be sealed in a few steps so that it can be easily handled by non-specialists. The gel-filled wells of the device showed the presence of gel even after storage for 7 days.

The microbial culture experiments using this device confirmed the formation of microbial colonies in the wells isolated to a single or a small number of bacteria. In addition, the colonies picked up from these wells could be regrown in liquid medium. This device was successfully applied for the cultivation of environmental bacteria able to grow on acetate as a carbon source, demonstrating that it can be easily used to perform high-throughput isolation and re-cultivation of microorganisms. This device has been fabricated using inexpensive materials and is low-cost. In addition, the well size and depth can be easily changed by design according to the microbial sample and the culture period of each experiment. The device is appropriate as a first screening tool in the search for useful microorganisms under various culture conditions. It represents a general microbial research tool for the screening of microorganisms from environmental samples as well as the selection of clones from metagenomic libraries.

## Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## Author contributions

RD and FS designed the experiments. CD, MF, and RD provided the new bacterial sample. CD, MF, CC-L, and CY performed bacterial experiments with the micro device. SZ and FS designed and fabricated the micro device. CD, MF, CS-M, CC-L, HN, RD, and FS analyzed data. CD, CS-M, HN, RD, and FS wrote the main text. All authors contributed to the article and approved the submitted version.

## Funding

This work was the financially supported by the JSPS KAKENHI (grant numbers: 19H05679, 19H05680, 19H05686, 21K14768, and 22H01502), Adaptable and Seamless Technology Transfer Program through Target-driven R&D (A-STEP) from Japan Science and Technology Agency (JST) (grant number: JPMJTM20GW), Initiative for Realizing Diversity in the Research Environment, MEXT Initiative for Realizing Diversity in the Research Environment, MEXT Initiative for Realizing Diversity in the Research Environment, and Japan-France Integrated action Program (SAKURA) (grant number: JPJSBP120213205, Japanese-French HYBAM project through the bilateral program PHC SAKURA, project no. 46981TD) and EU-Marie Skłodowska-Curie grant No. 892764.

## 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

Alvarez-Barragan, J., Cravo-Laureau, C., Wick, L. Y., and Duran, R. (2021). Fungi in PAH-contaminated marine sediments: cultivable diversity and tolerance capacity towards PAH. *Mar. Pollut. Bull.* 164:112082. doi: 10.1016/j.marpolbul.2021.112082

Arnfinnsdottir, N. B., Ottesen, V., Lale, R., and Sletmoen, M. (2015). The Design of Simple Bacterial Microarrays: development towards immobilizing single living bacteria on predefined micro-sized spots on patterned surfaces. *PLoS One* 10:e0128162. doi: 10.1371/journal.pone.0128162

Bagdasarian, M., Lurz, R., Rückert, B., Franklin, F. C. H., Bagdasarian, M. M., Frey, J., et al. (1981). Specific-purpose plasmid cloning vectors II. Broad host range, high copy number, RSF 1010-derived vectors, and a host-vector system for gene cloning in pseudomonas. *Gene* 16, 237–247. doi: 10.1016/0378-1119(81)90080-9

Balaban, N. Q., Merrin, J., Chait, R., Kowalik, L., and Leibler, S. (2004). Bacterial persistence as a phenotypic switch. *Science* 305, 1622–1625. doi: 10.1126/science.1099390

Balagadde, F. K., You, L., Hansen, C. L., Arnold, F. H., and Quake, S. R. (2005). Long-term monitoring of bacteria undergoing programmed population control in a microchemostat. *Science* 309, 137–140. doi: 10.1126/science.1109173

Banyal, A., Thakur, V., Thakur, R., and Kumar, P. (2021). Endophytic microbial diversity: a new Hope for the production of novel anti-tumor and anti-HIV agents as future therapeutics. *Curr. Microbiol.* 78, 1699–1717. doi: 10.1007/s00284-021-02359-2

Barua, N., Herken, A. M., Stern, K. R., Reese, S., Powers, R. L., Morrell-Falvey, J. L., et al. (2020). Simultaneous discovery of positive and negative interactions among rhizosphere bacteria using microwell recovery arrays. *Front. Microbiol.* 11:601788. doi: 10.3389/fmicb.2020.601788

Boedicker, J. Q., Li, L., Kline, T. R., and Ismagilov, R. F. (2008). Detecting bacteria and determining their susceptibility to antibiotics by stochastic confinement in nanoliter droplets using plug-based microfluidics. *Lab Chip* 8, 1265–1272. doi: 10.1039/b804911d

Boedicker, J. Q., Vincent, M. E., and Ismagilov, R. F. (2009). Microfluidic confinement of single cells of bacteria in small volumes initiates high-density behavior of quorum sensing and growth and reveals its variability. *Angew. Chem. Int. Ed. Engl.* 48, 5908–5911. doi: 10.1002/anie.200901550

Cerf, A., Cau, J. C., and Vieu, C. (2008). Controlled assembly of bacteria on chemical patterns using soft lithography. *Colloids Surf. B Biointerfaces* 65, 285–291. doi: 10.1016/j.colsurfb.2008.04.016

Chowdhury, M. S., Zheng, W., Kumari, S., Heyman, J., Zhang, X., Dey, P., et al. (2019). Dendronized fluorosurfactant for highly stable water-in-fluorinated oil emulsions with minimal inter-droplet transfer of small molecules. *Nat. Commun.* 10:4546. doi: 10.1038/s41467-019-12462-5

Courtois, F., Olguin, L. F., Whyte, G., Theberge, A. B., Huck, W. T. S., Hollfelder, F., et al. (2009). Controlling the retention of small molecules in emulsion microdroplets for use in cell-based assays. *Anal. Chem.* 81, 3008–3016. doi: 10.1021/ac802658n

Cydzik-Kwiatkowska, A., and Zielinska, M. (2016). Bacterial communities in full-scale wastewater treatment systems. *World J. Microbiol. Biotechnol.* 32:66. doi: 10.1007/s11274-016-2012-9

Dias, M., Salvado, J. C., Monperrus, M., Caumette, P., Amouroux, D., Duran, R., et al. (2008). Characterization of *Desulfomicrobium salsuginis* sp. nov. and *Desulfomicrobium aestuarii* sp. nov., two new sulfate-reducing bacteria isolated from the Adour estuary (French Atlantic coast) with specific mercury methylation potentials. *Syst. Appl. Microbiol.* 31, 30–37. doi: 10.1016/j.syapm.2007.09.002

Enger, J., Goksor, M., Ramser, K., Hagberg, P., and Hanstorp, D. (2004). Optical tweezers applied to a microfluidic system. *Lab Chip* 4, 196–200. doi: 10.1039/b307960k

Guo, M. T., Rotem, A., Heyman, J. A., and Weitz, D. A. (2012). Droplet microfluidics for high-throughput biological assays. *Lab Chip* 12, 2146–2155. doi: 10.1039/c2lc21147e

Hansen, R. H., Timm, A. C., Timm, C. M., Bible, A. N., Morrell-Falvey, J. L., Pelletier, D. A., et al. (2016). Stochastic assembly of bacteria in microwell arrays reveals the importance of confinement in community development. *PLoS One* 11:e0155080. doi: 10.1371/journal.pone.0155080

Ingham, C. J., Sprenkels, A., Bomer, J., Molenaar, D., van den Berg, A., van Hylckama Vlieg, J. E., et al. (2007). The micro-petri dish, a million-well growth chip for the culture and high-throughput screening of microorganisms. *Proc. Natl. Acad. Sci. U. S. A.* 104, 18217–18222. doi: 10.1073/pnas.0701693104

Jiang, C. Y., Dong, L., Zhao, J. K., Hu, X., Shen, C., Qiao, Y., et al. (2016). High-throughput single-cell cultivation on microfluidic streak plates. *Appl. Environ. Microbiol.* 82, 2210–2218. doi: 10.1128/AEM.03588-15

Kaminski, T. S., Scheler, O., and Garstecki, P. (2016). Droplet microfluidics for microbiology: techniques, applications and challenges. *Lab Chip* 16, 2168–2187. doi: 10.1039/c6lc00367b

Kehe, J., Kulesa, A., Ortiz, A., Ackerman, C. M., Thakku, S. G., Sellers, D., et al. (2019). Massively parallel screening of synthetic microbial communities. *Proc. Natl. Acad. Sci. U. S. A.* 116, 12804–12809. doi: 10.1073/pnas.1900102116

Kovach, M. E., Elzer, P. H., Hill, D. S., Robertson, G. T., Farris, M. A., Roop, R. M., et al. (1995). Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. *Gene* 166, 175–176. doi: 10.1016/0378-1119(95)00584-1

Li, M., Liu, H., Zhuang, S., and Goda, K. (2021). Droplet flow cytometry for single-cell analysis. *RSC Adv.* 11, 20944–20960. doi: 10.1039/d1ra02636d

Locey, K. J., and Lennon, J. T. (2016). Scaling laws predict global microbial diversity. *Proc. Natl. Acad. Sci. U. S. A.* 113, 5970–5975. doi: 10.1073/pnas.1521291113

Mannik, J., Driessen, R., Galajda, P., Keymer, J. E., and Dekker, C. (2009). Bacterial growth and motility in sub-micron constrictions. *Proc. Natl. Acad. Sci. U. S. A.* 106, 14861–14866. doi: 10.1073/pnas.0907542106

Moffitt, J. R., Lee, J. B., and Cluzel, P. (2012). The single-cell chemostat: an agarose-based, microfluidic device for high-throughput, single-cell studies of bacteria and bacterial communities. *Lab Chip* 12, 1487–1494. doi: 10.1039/c2lc00009a

Mohd Azhar, S. H., Abdulla, R., Jambo, S. A., Marbawi, H., Gansau, J. A., Mohd Faik, A. A., et al. (2017). Yeasts in sustainable bioethanol production: a review. *Biochem Biophys Rep* 10, 52–61. doi: 10.1016/j.bbrep.2017.03.003

Nakai, R., Kusada, H., Sassa, F., Morigasaki, S., Hayashi, H., Takaya, N., et al. (2021). Draft genome sequence of novel filterable *Rhodospirillales* bacterium strain TMPK1, Isolated from Soil. *Microbiol Resour Announc* 10:e0039321. doi: 10.1128/MRA.00393-21

Nomura, N., Suzuki, H., Sassa, F., Kiyokawa, T., Yokogawa, M., Toyofuku, M., et al. (2014). *Cell culture device*. Japan Patent Application No JP2014009737A.

Ottesen, E. A., Hong, J. W., Quake, S. R., and Leadbetter, J. R. (2006). Microfluidic digital PCR enables multigene analysis of individual environmental bacteria. *Science* 314, 1464–1467. doi: 10.1126/science.1131370

Sambrook, J.R., (2001). *Molecular Cloning: A Laboratory Manual*, *3rd Edn.*. NY, USA: Cold Spring Harbor Laboratory Press.

Schirhagl, R., Hall, E. W., Fuereder, I., and Zare, R. N. (2012). Separation of bacteria with imprinted polymeric films. *Analyst* 137, 1495–1499. doi: 10.1039/c2an15927a

Shah, P., Fritz, J. V., Glaab, E., Desai, M. S., Greenhalgh, K., Frachet, A., et al. (2016). A microfluidics-based in vitro model of the gastrointestinal human-microbe interface. *Nat. Commun.* 7:11535. doi: 10.1038/ncomms11535

Srinivasan, A., Leung, K. P., Lopez-Ribot, J. L., and Ramasubramanian, A. K. (2013). High-throughput nano-biofilm microarray for antifungal drug discovery. *mBio* 4:e00331-13. doi: 10.1128/mBio.00331-13

Taheri-Araghi, S., Bradde, S., Sauls, J. T., Hill, N. S., Levin, P. A., Paulsson, J., et al. (2015). Cell-size control and homeostasis in bacteria. *Curr. Biol.* 25, 385–391. doi: 10.1016/j.cub.2014.12.009

Thorsen, T., Maerkl, S. J., and Quake, S. R. (2002). Microfluidic large-scale integration. *Science* 298, 580–584. doi: 10.1126/science.1076996

Unger, M. A., Chou, H. P., Thorsen, T., Scherer, A., and Quake, S. R. (2000). Monolithic microfabricated valves and pumps by multilayer soft lithography. *Science* 288, 113–116. doi: 10.1126/science.288.5463.113

Walker, G. M., Sai, J., Richmond, A., Stremler, M., Chung, C. Y., and Wikswo, J. P. (2005). Effects of flow and diffusion on chemotaxis studies in a microfabricated gradient generator. *Lab Chip* 5, 611–618. doi: 10.1039/b417245k

Walter, A., Marz, A., Schumacher, W., Rosch, P., and Popp, J. (2011). Towards a fast, high specific and reliable discrimination of bacteria on strain level by means of SERS in a microfluidic device. *Lab Chip* 11, 1013–1021. doi: 10.1039/c0lc00536c

Weibel, D. B., Diluzio, W. R., and Whitesides, G. M. (2007). Microfabrication meets microbiology. *Nat. Rev. Microbiol.* 5, 209–218. doi: 10.1038/nrmicro1616

Wooley, J. C., Godzik, A., and Friedberg, I. (2010). A primer on metagenomics. *PLoS Comput. Biol.* 6:e1000667. doi: 10.1371/journal.pcbi.1000667

Zhang, L., Chen, P., Zhou, Z., Hu, Y., Sha, Q., Zhang, H., et al. (2019). Agarose-based microwell array chip for high-throughput screening of functional microorganisms. *Talanta* 191, 342–349. doi: 10.1016/j.talanta.2018.08.090

frontiers | Frontiers in Microbiology

TYPE Original Research
PUBLISHED 04 January 2023
DOI 10.3389/fmicb.2022.1012400



OPEN ACCESS

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

REVIEWED BY
Daniel Lipus,
GFZ German Research Centre for Geosciences, Germany
Simon Poirier,
IFP Energies nouvelles,
France

*CORRESPONDENCE
Anthony Ranchou-Peyruse
✉ anthony.ranchou-peyruse@univ-pau.fr

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

RECEIVED 05 August 2022
ACCEPTED 06 December 2022
PUBLISHED 04 January 2023

CITATION
Ranchou-Peyruse M, Guignard M, Haddad PG, Robin S, Boesch F, Lanot M, Carrier H, Dequidt D, Chiquet P, Caumette G, Cézac P and Ranchou-Peyruse A (2023) A deep continental aquifer downhole sampler for microbiological studies. *Front. Microbiol.* 13:1012400. doi: 10.3389/fmicb.2022.1012400



# A deep continental aquifer downhole sampler for microbiological studies

Magali Ranchou-Peyruse[1,2,3], Marion Guignard[1], Perla G. Haddad[2], Sylvain Robin[4], Fabrice Boesch[4], Maud Lanot[4], Hervé Carrier[3,5], David Dequidt[6], Pierre Chiquet[3,7], Guilhem Caumette[3,7], Pierre Cézac[2,3] and Anthony Ranchou-Peyruse[1,3]*

[1]E2S-UPPA, CNRS, IPREM, Universite de Pau & Pays Adour, Pau, France, [2]E2S-UPPA, LaTEP, Universite de Pau & Pays Adour, Pau, France, [3]Joint Laboratory SEnGA, E2S-UPPA-Teréga, Pau, France, [4]Modis, Pau, France, [5]E2S-UPPA, CNRS, TOTAL, LFCR, Universite de Pau & Pays Adour, Pau, France, [6]STORENGY – Geosciences Department, Bois-Colombes, France, [7]Teréga, Pau, France

To be effective, microbiological studies of deep aquifers must be free from surface microbial contaminants and from infrastructures allowing access to formation water (wellheads, well completions). Many microbiological studies are based on water samples obtained after rinsing a well without guaranteeing the absence of contaminants from the biofilm development in the pipes. The protocol described in this paper presents the adaptation, preparation, sterilization and deployment of a commercial downhole sampler (PDSshort, Leutert, Germany) for the microbiological studying of deep aquifers. The ATEX sampler (i.e., explosive atmospheres) can be deployed for geological gas storage (methane, hydrogen). To validate our procedure and confirm the need to use such a device, cell counting and bacterial taxonomic diversity based on high-throughput sequencing for different water samples taken at the wellhead or at depth using the downhole sampler were compared and discussed. The results show that even after extensive rinsing (7 bore volumes), the water collected at the wellhead was not free of microbial contaminants, as shown by beta-diversity analysis. The downhole sampler procedure was the only way to ensure the purity of the formation water samples from the microbiological point of view. In addition, the downhole sampler allowed the formation water and the autochthonous microbial community to be maintained at *in situ* pressure for laboratory analysis. The prevention of the contamination of the sample and the preservation of its representativeness are key to guaranteeing the best interpretations and understanding of the functioning of the deep biosphere.

KEYWORDS

**deep aquifer, microorganisms, downhole sampler, UGS, ATEX, geological storage**

## Introduction

The latest estimates of the volume represented by all the world's aquifers included in the first two kilometers of the crust reach 22.6 million km$^3$, while the total volume of fresh surface water is estimated to be only 100,000 km$^3$ (Gleeson et al., 2015). In 2018, Magnabosco and her collaborators published a review compiling microbial concentration and diversity data from 3,800 continental subsurface studies, a third of which were groundwater studies. The authors estimated that this deep continental biomass could represent 23 to 31 petagrams of carbon (C), hundreds of times more than that comprised by the total of humanity. Approximately 98% of the world's freshwater reserves are located in aquifers (Margat and van der Gun, 2013). Shallow, deep, freshwater or saline aquifer resources must be managed, as they play strategic roles in our societies in terms of both water resources (drinking water, irrigation) and participating in energy transitions (energy storage through underground gas storage (UGS) and geothermal energy) and carbon sequestration in greenhouse gas reduction approaches (De Silva and Ranjith, 2012; Sainz-Garcia et al., 2017; Limberger et al., 2018).

The microorganisms present in aquifers, whether indigenous or nonindigenous, can be involved in natural or stimulated bioattenuation processes (with hydrocarbons, pesticides, chlorinated solvents, etc.) and can be used to reduce the concentrations of contaminants, such as nitrate (Calderer et al., 2010; Matteucci et al., 2015; Lueders, 2017; Aldas-Vargas et al., 2021). The Deep Carbon Observatory program[1] has clearly shown that a detailed and exhaustive understanding of the carbon cycle necessitates these environments being taken into account. It is obvious that studies on the impacts of climate change on these ecosystems, including deep aquifers, will proliferate in the coming years. The topics of interest are diverse and may include studies on decreases in water reserves and their quality, bioattenuation and facilitated biodegradation, and the impacts of artificial groundwater recharge activities (Russo and Lall, 2017; Abdelmohsen et al., 2019; Jasrotia et al., 2019; Ranchou-Peyruse et al., 2021). Sampling is the starting point for all microbiological studies and therefore is a key step that is a difficult challenge in deep environments (Lehman, 2007; Mangelsdorf and Kallmeyer, 2010; Wilkins et al., 2014; Cario et al., 2019; Soares et al., 2019).

Unlike surface ecosystems, the study of continental aquifers in general, especially those several hundreds of meters deep, is complicated because of access-related difficulties, safety and sparse sampling sites, which often involve water sampling *via* drilled wells. These studies involve the withdrawal of formation water *via* surface installations, wellheads, or control or operating wells. For the deepest aquifers, these steel pipes can exceed 1 km in depth and represent important sites that can be characterized as "windows" into deep environments (Sorensen et al., 2013; Kadnikov et al., 2018). Conditions in these pipes are different from those in aquifers: these are open ecosystems rather than microporous ecosystems, with the possible presence of oxygen, metal alloys, pipe maintenance grease, and wellhead contamination by surface ecosystems (Basso et al., 2005; Lehman, 2007; Korbel et al., 2017; Hershey et al., 2018; Mullin et al., 2020). To limit the microbial contamination inherent in a well, most standard sampling protocols involve at least a purge of 1 (Garvis and Stuermer, 1980; Stevens et al., 1993), 2 (Pionke and Urban, 1987), or even 3 to 5 bore volumes (Sundaram et al., 2009; Hose and Lategan, 2012; Smith et al., 2012, 2015; Gründger et al., 2015; Hershey et al., 2018) to remove stagnant water. The use of a pump is necessary when the well is not eruptive (López-Archilla et al., 2007; Kirs et al., 2020). Some researchers have introduced a polyamide tube factory cleaned with a back-pressure valve at the lower end of the tubes to prevent water from flowing out during recovery (Nurmi and Kukkonen, 1986; Itävaara et al., 2011). The tube was made up of several sections sized 100 meters each to avoid contamination from the well. This method of sampling could be associated with inflatable packers to isolate specific fracture zones (Purkamo et al., 2013) but cannot be used to collect water from a monitoring well in the context of a gas storage aquifer for safety reasons. In 2005, Basso and her coauthors published a study on the procedure for cleaning an 800-meter-deep well before sampling the formation water to study microbial diversity. This work has shown that purging alone, regardless of its duration, cannot guarantee that the samples are uncontaminated by the microorganisms developing in biofilms colonizing the steel surfaces inside the well, valve seals and grease used for wellhead maintenance. The proposed procedure first involved a very large purge throughout the various stages with more than 25 times the volume of the tubing. Mechanical cleaning was carried out to remove the biofilms. Finally, three volumes of chlorine (4 liters of 9.6% active chlorine solution each) were injected into the bottom of the well before being eliminated by purge water, thus preventing injection into the aquifer. This procedure was applied at 11 different sites and gave rise to several published research works (Basso et al., 2005, 2009; Klouche et al., 2009; Berlendis et al., 2010; Aüllo et al., 2016; Ranchou-Peyruse et al., 2017; Godin et al., 2020; Ranchou-Peyruse et al., 2021). In addition to the quality of the samples, this last technique has the advantage of having no restrictions on the volume of water to be withdrawn after the well cleaning protocol is completed. Although this procedure is very effective since it drastically limits the risk of contamination, it is very complex and time-consuming and must be established at the site (Wireline Combi Unit, evacuation of purged water, etc.). As a result, it requires very close collaboration with the field operator and a substantial but necessary financial investment. It is therefore necessary to develop a sampling method that is just as effective but less restrictive. Downhole-sampling approaches appear to be the most relevant for ensuring the noncontamination of water samples and their associated microbial communities. The use of a downhole sampler is common in the field of geosciences in the broad sense for collecting samples of fluids such as water, dissolved or undissolved gases and oils to analyze their physicochemical or

1 https://deepcarbon.net/

isotopic compositions (Rivard et al., 2018; Struchkov and Rogachev, 2018; Banks et al., 2019; Osselin et al., 2019). In a UGS context, the study of microbial communities from geological reservoirs represents a real challenge because of the hundreds of meters of pipe, pressure, potential toxicities of gases (methane, sulfide, carbon dioxide, etc.), and explosive atmosphere of the study environment.

Here, a new sampling procedure with a commercial downhole sampler (Positive Displacement Sampler, PDSshort, Leutert, Germany) is presented. This sampler was previously used to sample deep wells as far as 4,240 mbs (meters below the surface) to maintain fluid samples at an *in situ* pressure to study the isotopic compositions of saline formation waters and dissolved gases (Regenspurg et al., 2010; Kietäväinen et al., 2012; Kampman et al., 2013; Feldbusch et al., 2018). There are a few examples of sampling deep continental aquifers with downhole samplers, but the procedures are not detailed and are difficult to replicate. An older sampler model (Sampler Model 60', Leutert) was employed to study the microbial community evolving in the formation water of an aquifer used for the storage of town gas in Lobodice (Czech Republic) in the early 1990s (Šmigáň et al., 1990). From the end of the 1990s, a similar pressurized groundwater sampling instrument called "PAVE" developed by Anttila et al. (1999) was used to study the microbiology of deep igneous rock aquifers in Finland (Haveman et al., 1999; Haveman and Pedersen, 2002; Öhberg, 2006; Pedersen et al., 2008). This wire-line PAVE downhole equipment used in deep biosphere studies in Fennoscandian bedrock had a small sampling volume (150 to 250 ml, Haveman et al., 1999; Haveman and Pedersen, 2002; Öhberg, 2006; Pedersen et al., 2008). However, deep aquifers are most often oligotrophic environments with microbial concentrations that are often very low and on the order of $10^1$ to $10^5$ cells.ml$^{-1}$ (Lerm et al., 2013; Bomberg et al., 2015, 2016; Magnabosco et al., 2018). These low concentrations necessitate an ability to work with the highest possible volumes. In addition, experiments in a pressurized reactor aimed at simulating these environments over periods of several months require volumes of more than 1 L of water, i.e., two samplers in the context of the study of Haddad and collaborators (Haddad et al., 2022a, b). Moreover, the electric current presents a danger during PAVE system deployment in an explosive atmosphere and prohibits its use on aquifers used to store natural gas (Ruotsalainen and Snellman, 1996). Later, to study the impact of $CO_2$ injection on microbial communities, a downhole sampler (Doppelkugelbüchse, DKB) was deployed in a saline aquifer at 675 mbs (Morozova et al., 2013).

In this article, we present the procedure for preparing and deploying a downhole sampler to sample water from a deep aquifer while respecting microbiological conditions such as anoxia, sterility and pressure maintenance. We recently used this sampler in two sampling campaigns related to geological gas storage and their interactions with autochthonous microbial communities (Haddad et al., 2022a, b). Here, the samples were taken from a monitoring well of a deep aquifer (−582 mbs; 69 bar; 36°C) used for the storage of natural gas. The formation water was maintained at the *in situ* pressure from the aquifer to the laboratory before undergoing controlled depressurization. To validate our procedure and prove its importance, basic but fundamental microbiological analyses were carried out: (i) the physiological state of the microbial cells was assessed by epifluorescence microscopy approaches and (ii) the bacterial taxonomic diversity was monitored at the wellhead throughout the sampling operation, as well as in the water collected using the two downhole samplers sent successively to the bottom of the well.

# Materials and methods

## Sampling site

Water was sampled from the AB_L_1 control well of an aquifer used for the storage of natural gas in the Aquitaine geological basin (southwest France) in January 2019. This well has already been used to study microbial diversity in formation water (Ranchou-Peyruse et al., 2019; Haddad et al., 2022a). The tubing of the monitoring well descends to a depth of 582 mbs and has an estimated volume of 7.6 m$^3$. No water from the upper formation can penetrate inside the tubing. The formation water evolves at a geological level dating from the Eocene-Lutetian that consists of inframolassic sands with sandy detrital facies (sandstone to numulites) at 36°C. The pressure at the bottom was estimated to be approximately 69 bar and fluctuated depending on the storage gas in place, reaching 3 GNm$^3$. At the time of sampling, the well was eruptive, and the stored gas was located approximately 250 meters from the well in the inframollassic sands. Although the sampler can be deployed in nonartesian drillholes, a purge water flow rate of 10 m$^3$.h$^{-1}$ was maintained to guarantee that the samples would not be contaminated by pieces of biofilm torn from the surface of the pipe during the descent of the sampler. During sampling, two downhole samplers were deployed consecutively to obtain formation water with representative autochthonous microbial communities. Simultaneously, several water samples were collected at the wellhead.

## Description of the downhole sampler

A LEUTERT One Phase Sampler OPS sampler (Adendorf, Germany) was used during a one-day sampling campaign (Figure 1A) to avoid contamination by biofilms developing on the surfaces of the well, preserve anoxic conditions and maintain the *in situ* pressure during the process. With a length of 4.63 m and a diameter of 43 mm, the sampler is made of stainless steel according to NACE MR-01-75 and a bronze alloy. It operates up to 1,035 bar and 180°C (supplier information). Fortunately, this sampler has a steam-sterilizable sampling chamber that opens and closes only at the depth to be sampled, thus making it possible to obtain a representative sample of the target microbial environment. The

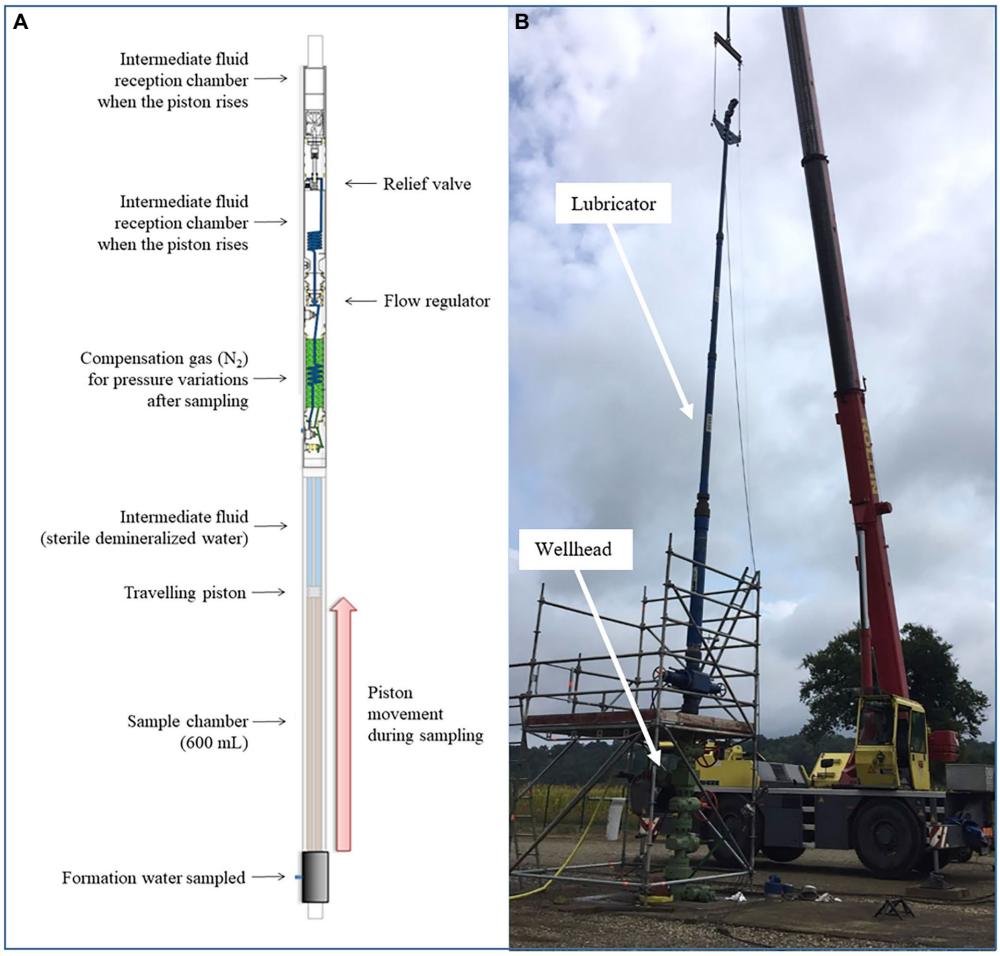



FIGURE 1
Use of the LEUTERT OPS downhole sampler to collect microorganisms from a deep aquifer. **(A)** Descriptive diagram of the downhole sampler; **(B)** Deployment of the sampler on site: the sampler is inserted into a chamber. Then, the chamber is connected to the wellhead before it is opened, and the sampler descends into the well.

sampler is ATEX-certified and can be used in explosive atmospheres (ATEX). Indeed, it is completely mechanical and lacks any source that can generate a spark or energy that can cause an explosion. Practically, the sampler is composed of three compartments: (i) The upper compartment is filled with a biologically inert gas such as nitrogen (Linde) during this sampling, and it is mechanically isolated from the other compartments of the tool until the piston, at the end of its stroke (i.e., at the end of sampling), activates a set of linkages allowing the following actions: closure of the sample chamber, sealing of the connection between the intermediate fluid reception chamber and the sample compartment, and connection of the nitrogen compartment with the sample compartment *via* the traveling piston. This latter connection allows the fluid in the nitrogen compartment to exert a pressure force on the downstream face of the piston once the sample has been taken. This force ensures that the sample is maintained at a minimum pressure close to the pressure imposed on the nitrogen compartment at the surface (i.e., 69 bar). (ii) The intermediate fluid keeps the piston in its low position until the opening of the channel connecting the sample compartment and the sterile demineralized water compartment. This connection is made mechanically at the end of the countdown programmed on the clock/mechanical actuator. (iii) The lower compartment (i.e., sampling chamber), 600 $cm^3$, was used to collect the water sample. A flow regulator adds a calibrated pressure drop to control the transfer rate and therefore the speed of sample collection. When the sampled fluid occupies the entire dedicated compartment, i.e., 600 $cm^3$, the compartment containing nitrogen and the intermediate fluid are mechanically isolated from the other components of the tool. During ascent and once at the surface, the bottom pressure is maintained inside the sampler with a slight modification linked to the change in temperature.

## Preparation of the sampler for microbiological studies

In the laboratory, all of the instruments that would come in contact with the collected sample were cleaned, degreased and

disinfected with 70% ethanol. Each operation that could lead to contamination from the external environment was subject to cleaning with alcohol and precautions to avoid microbial contamination until the equipment was deployed in the well. This was followed by a step of rinsing all the parts with sterile water. The water collection chamber, filled with 50 ml of distilled water, was then sterilized for 2 h at 125°C using an enveloping jacket with heating resistance (Supplementary Figure S1). Finally, the timer released a mechanical system allowing the piston to go up. The latter triggered the opening of the pressurized gas responsible for maintaining the pressure of the collected sample and at the same time the closing of the inlet orifices of the sampling chamber. The operation settings allowed us to work with time steps from 15 min to 5 h, but the equipment could be reprogrammed to go up to 24 h.

## Field deployment

The downhole sampler was operated by Modis (Pau, France). To allow the descent of the sampler with a crane along the monitoring well, the downhole sampler was first inserted into a lubricator (Figure 1B), and centralizers were positioned to prevent the sampler from coming into contact with the internal surface of the lubricator and the well at all times during descent. Once the sampler was lowered to the desired depth, the fluid to be sampled exerted a pressure force on the upstream surface of the piston. During this sampling campaign, two downhole samplers were used approximately 2 h apart. Back in the laboratory, the water samples could be slowly depressurized to be analyzed, as was done in this experiment.

## Depressurization of the sample

All connections, tubes and flasks that were in contact with the water collection chamber and with the collected water were sterilized beforehand *via* steam sterilization, high temperature overnight or cleaning with 70% ethanol. Because deep aquifers are anoxic environments, the air inside the sterilized flasks used to collect water containing microorganisms was replaced by nitrogen. To obtain as representative of a sample as possible, the rate of depressurization had to be controlled to avoid lysing the microbial cells. To this end, a pressure gauge was connected to the filling opening of the nitrogen chamber, and a pressure compensation system consisting of a sterilized high-pressure manual pump equipped with a pressure gauge and one two-way valve was used (Supplementary Figure S2). This setup ensured an equal pressure when the sampler was connected to the transfer tubing, and the pressure drop was controlled while the sample was depressurized. The valve of the outlet of the sampling chamber was first slowly opened to balance the two upstream/downstream pressure taps without creating a depressurization shock that could lead to the death of the microorganisms in the sample. The gradual purging of nitrogen was carried out from the transfer cylinder pressure compensation chamber. Once the sample pressure was reduced to approximately 5 bar, it was very gradually brought to atmospheric pressure during its transfer into sterile glass bottles equipped with overpressure exhaust filters.

## Biomass filtration and nucleic acid extraction

Some of the samples collected with the both downhole samplers during this campaign (2 × 500 ml) were kept separately in order to simulate the conditions of the deep aquifer inside a high-pressure reactor (Haddad et al., 2022a). Here, 100 ml of the formation waters sampled with each of the two downhole samplers (DS1 and DS2) and four 1 L flasks of the wellhead waters (WHS1 to WHS4) were filtered with 0.22 µm porosity filters (cellulose nitrate filter, Sartorius Stedim) for microbial community taxonomic diversity analyses. Subsequently, the eluate was filtered a second time with 0.1 µm porosity filters (polyethersulfone filter, Sartorius Stedim) to collect cells smaller than 0.22 µm. Each filter was stored at-80°C for future use to conserve nucleic acids. The filters were then ground in a mortar with liquid nitrogen. All of the nucleic acids were extracted with the Fast RNA Pro-Soil (MP Bio) kit following the manufacturer's instructions until the nucleic acids were eluted in 50 µl of DEPC water included in the kit. Then, the extracted DNA and RNA were separated with the All Prep DNA/RNA kit (QIAGEN). For this study, we were only interested in DNA analysis. The DNA concentrations were quantified with a Qubit fluorometer (Invitrogen by Life Technologies, Carlsbad, CA). DNA concentrations were not detectable for the 100 ml samples from the downhole samplers.

## Nested PCR and high-throughput sequencing

A conventional PCR approach did not allow sufficient amplification of all samples. For each sample, the V3-V4 region of the 16S rRNA gene was amplified by nested PCR (Yu et al., 2015) with the PCR CORE kit (Roche). The primers used to target this region were 8F/1489R (Weisburg et al., 1991), 344F_5'–ACGGRAGGCAGCAG-3' and 801R_5'-CGGCGTGGACTTCCAGGGTATC-3' (Simmon et al., 2006; Wichels et al., 2006). The first amplification was carried out as follows: a strand separation at 94°C for 2 min; 15 cycles of 94°C for 40 s; annealing at 55°C for 40 s; and strand extension at 72°C for 45 s. A final stage of 7 min at 72°C for further strand extension. The second amplification differed in the 30 amplification cycles: 94°C for 30 s; 65°C for 30 s; 72°C for 40 s. The 344F/801R primers contained the adapters 5'-CTTTCCCTACACGACGCTCTTCCGATCT-3' and 3'-GGAGTTCAGACGTGTGCTCTTCCGATCT-5' to achieve high-throughput sequencing. High-throughput sequencing was performed with a GenoToul genomic platform (Toulouse, France)

that used MiSeq Illumina 2 × 250 bp technology, according to the manufacturer's protocol. The raw sequencing data were deposited in the NCBI SRA under bioproject ID PRJNA769063. The sequencing data were then checked for their quality and processed *via* the FROGS analysis pipeline developed by the GenoToul genomic platform in the Galaxy interface (Escudié et al., 2018) using Flash to merge the paired-end reads (Magoc and Salzberg, 2011), Swarm for sequence clustering based on the Sellers' evolutionary distance (Sellers, 1974; Mahé et al., 2014), and VSEARCH with the *de novo* UCHIME method to eliminate chimeras (Edgar et al., 2011; Rognes et al., 2016). From the initial 518,289 reads, the pre-processing and filtration steps led to 409,781 reads. Rarefaction curves obtained for each sample (data not shown) indicated that sequencing was deep enough to estimate microbial composition and there did not seem to be any effect of filtration of different volumes of water between WHS and DS. After normalization, there were 20,682 reads per sample. The processed dataset was analyzed using the R "phyloseq" package (McMurdie and Holmes, 2013). The graphs were constructed using the R "ggplot2" package (Wickham, 2009). Alpha-diversity indices were calculated and comparison of microbial compositions between samples was obtained by performing a hierarchical clustering using the Ward D2 method and based on Jaccard's analysis of beta-diversity. A heatmap representation was also generated. The taxonomic classification was based on the Silva database (version 138.1).

## Cell counts and microscopy

From all the water collected (samplers and well-head), the proportions of living and dead cells were determined by epifluorescence microscopy with the LIVE/DEAD BacLight Bacterial Viability Kit (Thermo Fisher Scientific), as described by the supplier. Briefly, 1 ml of water was spiked with 1.5 µl of SYTO9 and 1.5 µl of propidium iodide. After incubation in the dark for 15 min, the water was filtered through 0.2 µm pore-size black polycarbonate (Millipore) under vacuum as described in Ranchou-Peyruse et al. (2021). In each of the measurements, 20 randomly selected fields were observed, and 85 to 3,300 microbial cells were counted. To our knowledge, there are no black polycarbonate filters with a porosity of 0.1 µm, which explains the absence of this analysis. A Zeiss Observer.Z1c epifluorescence microscope equipped with a mercury light source was used.

# Results and discussion

## A tool adapted for deep aquifer sampling in the UGS context

Before being deployed on site, several laboratory tests were carried out to verify the maintenance of the sterility of the sampler and its proper functioning under pressures ranging from 50 to 100 bar. This pressure range was selected because it corresponds to French gas storage in deep aquifers (–500 mbs to –1,200 mbs). This tool was designed to be able to operate at 1035 bar in a petroleum environment. We successfully tested its operation at lower pressures without the use of grease (which is a source of contamination and exogenous molecules) for the moving parts in contact with the sampling chamber. Here, the downhole sampler was lowered in tubing with a minimum internal diameter of 69 mm, but it can reasonably be used down to 55 mm. The first tests were carried out without the use of centralizers and led to the scraping of the pipe during the descent and the contamination of the inlet to the sampling chamber (data not shown). We deduce that the use of such downhole samplers in inclined drillholes could compromise the quality of the microbiological sampling. In this type of well, the sampler, the slickline train and the cable continuously rub against the tubing.

## Control of the depressurization rate

Throughout our operation, the monitoring well was in eruptive conditions with a water flow rate maintained at 10 $m^3.h^{-1}$, allowing the evacuation of any piece of biofilm or any particle possibly torn off during the descent of the sampler. Taking into account the flow rate as well as the volume of the well, raising a microbial cell out of the deep aquifer from 582 mbs represents a depressurization of 60 bar at atmospheric pressure in 46 min, or 1.3 $bar.min^{-1}$. It is still difficult to assess the impact of this depressurization on the physiological state of microorganisms. This depressurization could explain the high percentages of dead cells, ranging from 28 to 54% of the community in samples WHS1, 2, 3 and 4 (WHS, wellhead sampling; Figure 2). However, in a study on MEOR (Microbial Enhanced Oil Recovery), the authors considered that with a depressurization of 1 $bar.min^{-1}$, the deleterious effect on the cells was not "pronounced" (Krüger et al., 2016). Nevertheless, it should be noted the depressurization seems to have been made by stages of several hours, even days. Sampling with these downhole samplers enabled us to maintain the pressure and/or regulate the pressure decrease. In this work, the chosen depressurization speeds were lower than those experienced during wellhead sampling, with one relatively close (DS, downhole sampling; DS1: 1 $bar.min^{-1}$) and the other significantly lower (DS2: ≈0.2 $bar.min^{-1}$). The higher cell mortality in the case of the DS1 sampler supports this first hypothesis, and a comparison with the results obtained in the case of DS2 validates a depressurization rate of approximately 0.2 $bar.min^{-1}$. Likewise, the lowest apparent cellular concentrations for water taken directly from the wellhead (between $1.6{\cdot}10^4 \pm 4.4{\cdot}10^3$ $cell.mL^{-1}$ and $4.4{\cdot}10^4 \pm 2.0{\cdot}10^4$ $cell.mL^{-1}$) compared to those taken with the sampler (DS1: $2.3{\cdot}10^5 \pm 7.2{\cdot}10^4$ $cell.mL^{-1}$ and DS2: $6.3.10^5 \pm 2.0{\cdot}10^5$ $cell.mL^{-1}$) support the idea that the lysis of the cell membrane is more sensitive to depressurization, which could have led to the dispersion of genomic material outside the cell, making it impossible to count these cells by fluorescent nucleic acid stains (SYTO9 and propidium iodide).

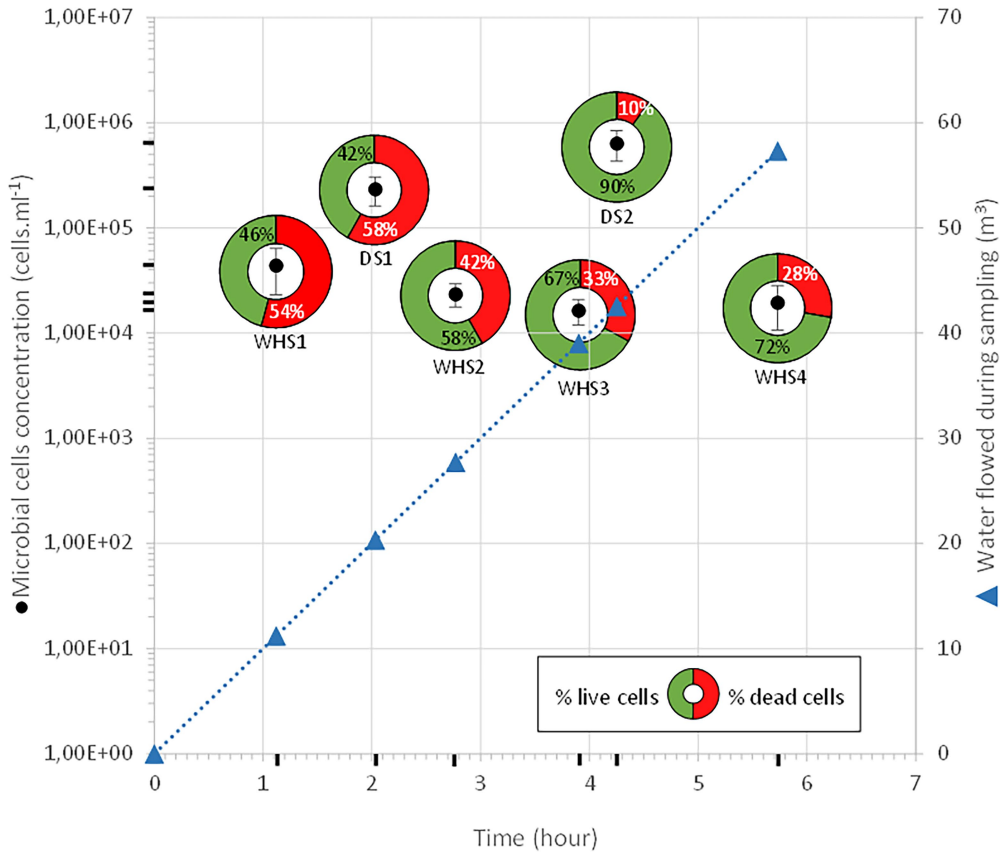



FIGURE 2
Enumeration and cell survival during the sampling campaign. The black circles represent total cell counts. For each of these counts, the cells with intact membranes, qualified as living cells, are represented in green, while the cells with lysed membranes, qualified as dead cells, are shown in red. As the well is artesian throughout the procedure, the volume of water purged is represented with blue triangles. DS: downhole sampling; WHS: wellhead sampling.

## Samples free from microbial contamination

Regarding the water sampled at the wellhead, the water flow before sampling corresponded to more than 1 ($1 \times 7.6\,m^3$ of well volume $< 11.3\,m^3$ of water discharged), 3 ($22.8\,m^3 < 27.7\,m^3$), 5 ($38\,m^3 < 39\,m^3$) and 7 ($53.6\,m^3 < 57.3\,m^3$) bore volumes for WHS1, WHS2, WHS3 and WHS4, respectively (Figure 2). The alpha-diversity of the microbial communities from the water sampled at the well-head decreased as the bore volumes increased, with an exception for WHS3 0.22 µm (Table 1). The differences in various metrics (number of observed OTU, Chao1 index, Shannon index, InvSimpson index) between the biomass filtered at 0.22 µm and 0.1 µm suggested the presence of a sub-community of smaller cell size (< 0.1 µm; 26 OTU in Figure 3). Despite the volumes of water discharged throughout the sampling day, the beta-diversity analysis clearly shows a difference between the microbial communities sampled at the wellhead and those sampled using the downhole sampler (Figure 4), even after 7 bore volumes were discharged, as for the WHS4 sample. These results clearly show that the simple water purges described in many studies of deep aquifers may not be sufficient to prevent contamination and therefore do not guarantee the purity of the microbial communities studied (Garvis and Stuermer, 1980; Pionke and Urban, 1987; Stevens et al., 1993; Sundaram et al., 2009; Hose and Lategan, 2012; Smith et al., 2012, 2015; Gründger et al., 2015; Hershey et al., 2018). In all cases, an evolution of the microbial taxonomic diversity was revealed during the withdrawal of the formation water (Table 1; Figure 4). We interpret this by the heterogeneity of the aquifer close to the well, discussed in more detail below. The low microbial concentrations often encountered in oligotrophic deep aquifer water make it necessary to filter the biomass to concentrate it for molecular or even culturing approaches (Basso et al., 2005; Trias et al., 2017; Guðmundsdóttir et al., 2019). Living in a deep environment promotes cell shrinkage to reduce the metabolic needs of microorganisms in an environmental context of extreme oligotrophy (Miyoshi et al., 2005; Luef et al., 2015; Wu et al., 2016). The use of filters with a porosity of 0.1 µm is recommended but is not necessarily easy when filtering water that may be clogged with mineral particles, in particular iron sulfides, which can be present in abundance in this kind of anoxic environment rich in ferrous iron and sulfide leading to precipitates such as pyrite. By continuing to rely on betadiversity,

TABLE 1 Alpha-diversity values of microbial community in the formation water collected at the wellhead (WHS) and with the downhole samplers (DS) after a biomass concentration with filters with a porosity of 0.22μm and 0.1μm from the 0.22μm-eluate.

| | Observed | Chao | Shannon | InvSimpson |
|---|---|---|---|---|
| DS1_0.1 μm | 32 | 32.25 | 2.2026627057447 | 6.68692475963143 |
| DS1_0.22 μm | 16 | 16 | 1.6071602891354 | 4.06101605909724 |
| DS2_0.1 μm | 35 | 36.5 | 1.69487857291484 | 4.00962888083069 |
| DS2_0.22 μm | 16 | 16 | 1.05375703144571 | 2.1445596371457 |
| WHS1_0.1 μm | 54 | 55.4285714285714 | 0.919883527901311 | 1.52772507061242 |
| WHS2_0.22 μm | 37 | 39 | 0.506904400990633 | 1.22670219064637 |
| WHS2_0.1 μm | 43 | 46 | 1.35404072016678 | 2.12060845612808 |
| WHS2_0.22 μm | 34 | 34.25 | 0.751805115394659 | 1.4350021942113 |
| WHS3_0.1 μm | 39 | 41 | 1.83916566155245 | 3.587061073388 |
| WHS3_0.22 μm | 56 | 58.3333333333333 | 1.55171796086355 | 2.97583926302589 |
| WHS4_0.1 μm | 24 | 25 | 0.739682176064375 | 1.3862512153454 |
| WHS4_0.22 μm | 25 | 25 | 2.05107884523552 | 4.46583265503757 |

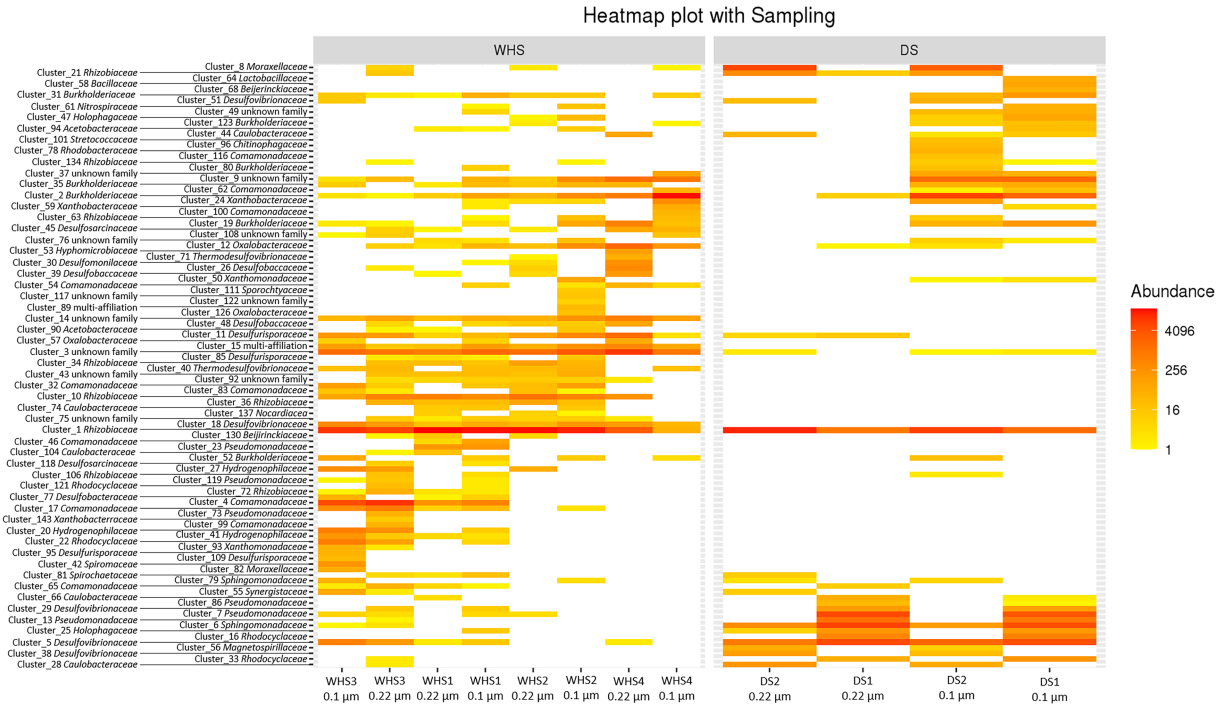



FIGURE 3
Heatmap showing the distributions of 108 bacterial OTUs in the formation water collected at the wellhead (WHS) and with the downhole samplers (DS) after a biomass concentration with filters with a porosity of 0.22μm and 0.1μm from the 0.22μm-eluate. The taxa associated with the OTUs are presented to the left of the figure and in Supplementary Table S1.

there was no difference found between the microbial communities according to the filtration porosity from the water collected at the wellhead (Figure 4). On the other hand, there seems to be an effect of the filter porosity in the case of the bacterial communities from the water sampled *via* the downhole samplers (DS), suggesting that the microbial contamination from the pipe surface (WHS) masks a possible cell size effect. Regarding the microbial communities, an analysis of the different taxonomic diversities based on the 16S rRNA gene by heatmap representation (Figure 3) and a comparison of the 20 dominant bacterial families (Figure 4) show variability between the different waters sampled at the wellhead (WHS) and in the samplers (DS). These differences could be explained by the expected heterogeneity of the microbial communities throughout the sampling (Goldscheider et al., 2006). The studied aquifer, which is

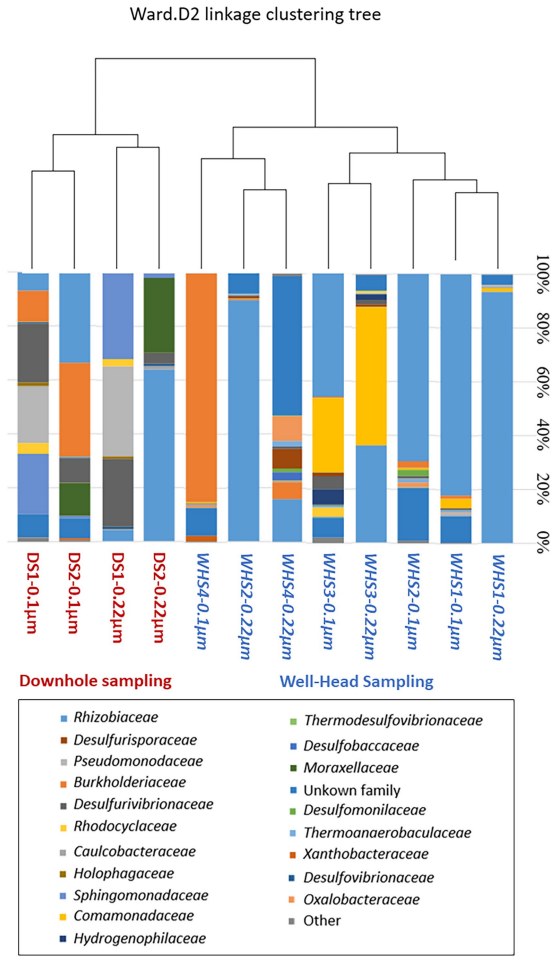



FIGURE 4
Jaccard's analysis of the beta diversity and comparison representation of the dominant bacterial families based on the 16S rRNA gene (v3-v4) in different formation water samples from the downhole sampler (DS, red) and the wellhead sampler (WHS, blue). The deep aquifer bacterial community was obtained by 0.22µm filtration or 0.1µm filtration from the 0.22µm eluate.

homogeneous in its overall constitution and characterized by inframolassic sands, presents heterogeneity on a more local scale, with the presence of calcite, clay and iron sulfides (Haddad et al., 2022a). Obviously, the longer the sampling and the greater the volume discharged, the further the sampled bacterial populations are from the well. With a flow rate of 10 $m^3.h^{-1}$, the formation water circulating in the sands of the aquifer, with a porosity between 25 and 35% (Ranchou-Peyruse et al., 2019), can come from several meters around the well, approximately 3 to 4 m, which can be represented by imagining a sphere of influence centered on the strainers at the bottom of the well. With 108 operational taxonomic units (OTUs) present in at least one analyzed sample, it appears that there are 54 OTUs that are only found in the water sampled at the wellhead and not in the samplers (Figures 3, 4). These OTUs were mainly distributed among 6 bacterial families: *Desulfobaccaceae*, *Desulfomonilaceae*, *Hydrogenophilaceae*, *Thermoanaerobaculaceae*, *Thermodesulfovibrionaceae*, and *Xanthobacteraceae*. Three of these families are sulfate reducers (*Desulfobaccaceae*, *Desulfomonilaceae,* and *Thermodesulfovibrionaceae*), and one family includes bacteria exhibiting fermentation metabolism (*Thermonaerobaculaceae*). Mesothermal conditions raise questions about the presence of bacteria classified as thermophilic microorganisms, but these results suggest that this physiological criterion is perhaps not decisive in the characterization of both bacterial families. The presence of members of these 6 families in anoxic or micro-oxic conditions is consistent with the lifestyle of these microorganisms at the interface with the steel of the pipeline, the corrosion of which can release $H_2$ as energy and electron sources, for example (Rajala et al., 2017). In the water samples taken with the downhole samplers, 54 OTUs were found, including 11 not highlighted from the WHS samples. Seven OTUs had a higher abundance of between 89 and 95.6% of all the OTUs detected in these samples. These main OTUs are affiliated with *Rhizobiaceae* (Cluster_1), *Burkholderiaceae* (Cluster_2), *Desulfurivibrionaceae* (Cluster_5), *Sphingomonadaceae* (Cluster_6), *Pseudomonadaceae* (Cluster_7 and Cluster_13), *Moraxellaceae* (Cluster_8), and an unknown family (Cluster_9). These families have already been found in other deep continental anoxic environments and have been supposed to be involved in sulfur and nitrogen cycles (Mu et al., 2014; Chiriac et al., 2018; Nuppunen-Puputti et al., 2018; Soares et al., 2019; Bell et al., 2020). Because some of these families include bacteria described to be aerobic or nitrate-reducing, being certain of the quality of the sampling is essential. Since deep aquifers are free of $O_2$ and $NO_3^-$, these families are often considered as surface and/or soil contaminants during the sampling from deep aquifers (Pedersen et al., 1997; Kadnikov et al., 2020). These microorganisms can adapt and survive in anoxic oligotrophic conditions without terminal electron acceptors such as $O_2$ and $NO_3^-$, probably by fermenting the organic molecules trapped in clays and released during the displacement of formation water during sampling. Our understanding of microbial environments at great depths is still limited. Samplings must be irreproachable so as not to add doubt to results that can sometimes be surprising. Thus, cyanobacteria and photosynthetic microorganisms were discovered in a 613-m-deep borehole from aseptic subsamples of cores (Puente-Sánchez et al., 2018). Using metagenomic approaches, the authors explained the ability of these organisms to colonize the deep continental subsurface, a light-deprived environment, by possible lithoautotrophic growth based on hydrogen.

## Conclusion

Similar to surface ecosystems, deep aquifers are shaped by the living organisms that develop there. These ecosystems are by nature oligotrophic and therefore poor in energy. Their use in the storage of energy or $CO_2$ modifies their physicochemical conditions and can be the source of new nutrients and/or energy, increasing the activity of anaerobic heterotrophs, fermenters, and even hydrogenotrophs. The quality of the samples guarantees that the expressed results represent the microbial diversity of the deep subsurface, and the noncontamination of the well from the surface

is essential in this regard. Doubt with respect to contamination can result in the classification of certain microorganisms, such as *Pseudomonadaceae*, as contaminants and prevent microbiologists from looking for new possible metabolic pathways that allow them to be maintained in these environments. In this sense, the use of a downhole sampler, common in the study of deep oceanic microbial environments, must be popularized in the study of deep aquifers when wells are used for access. In this paper, the LEUTERT One Phase Sampler OPS sampler commonly used for geoscience studies was successfully prepared to meet the specifics of a microbiological study to collect representative formation waters and was perfectly suited for deployment in areas presenting an explosive risk, such as natural gas storage areas. As we have shown here, the use of this type of sampler allows the elimination of microorganisms present in the well but absent from the original formation water. Finally, the use of a pressurized sampler makes it possible to control depressurization in an optimal manner for the survival of the microorganisms. Depending on the needs of future studies, this protocol can be easily adapted to directly transfer the collected pressurized water from the downhole sampler to a high-pressure reactor or storage in a high-pressure cell.

## 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

MR-P, SR, FB, ML, HC, PCh, PCé, and AR-P: co-conceived the study. MR-P, MG, PH, and AR-P: carried out sampling and microbiological studies. MR-P, MG, FB, ML, HC, and PCh participated in the preparation of the downhole sampler and the sampling. All authors contributed to interpretation of results and paper writing.

## Funding

Storengy and Teréga are acknowledged for funding this research project. MR-P salary was supported by E2S-UPPA.

## Acknowledgments

The authors thank Yannick Bouet for its comments during the writing of the article.

## Conflict of interest

DD is employed by STORENGY – Geosciences Department. PCh and GC are employed by Teréga. SR, FB, and ML are employed by Modis.

The remaining 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.2022.1012400/full#supplementary-material

SUPPLEMENTARY TABLE S1
Taxa of the OTUs presented in Figure 3 (heatmap).

SUPPLEMENTARY FIGURE S1
Sterilization of the sampler using a heating cell. On the left, a photograph of the heating cell. On the right, thermal imaging during the sterilization step at 125 °C for two hours in the presence of distilled water.

SUPPLEMENTARY FIGURE S2
Pressure compensation system for precisely controlling the depressurization of the downhole sampler. On the left, a schematic representation of the sampler; on the right, a sterilized high-pressure manual pump equipped with a pressure gauge and one two-way valve.

## References

Abdelmohsen, K., Sultan, M., Ahmed, M., Save, H., Elkaliouby, B., Emil, M., et al. (2019). Response of deep aquifers to climate variability. *Sci. Total Environ.* 677, 530–544. doi: 10.1016/j.scitotenv.2019.04.316

Aldas-Vargas, A., van der Vooren, T., Rijnaarts, H. H. M., and Sutton, N. B. (2021). Biostimulation is a valuable tool to assess pesticide biodegradation capacity of groundwater microorganisms. *Chemosphere* 280:130793. doi: 10.1016/j.chemosphere.2021.130793

Anttila, P., Ahokas, H., Front, K., Hinkkanen, H., Johansson, E., Paulamäki, S., et al. (1999). Final disposal of spent nuclear fuel in Finnish bedrock – Olkiluoto site report. Posiva Technical Reprot 99-10, Posiva: Rauma, Finland.

Aüllo, T., Berlendis, S., Lascourrèges, J. F., Dessort, D., Duclerc, D., Saint-Laurent, S., et al. (2016). New bio-indicators for long term natural attenuation of monoaromatic compounds in deep terrestrial aquifers. *Front. Microbiol.* 7:122. doi: 10.3389/fmicb.2016.00122

Banks, E. W., Hatch, M., Smith, S., Underschultz, J., Lamontagne, S., Suckow, A., et al. (2019). Multi-tracer and hydrogeophysical investigation of the hydraulic connectivity between coal seam gas formations, shallow groundwater and stream network in a faulted sedimentary basin. *J. Hydrol.* 578:124132. doi: 10.1016/j.jhydrol.2019.124132

Basso, O., Lascourrèges, J. F., Jarry, M., and Magot, M. (2005). The effect of cleaning and disinfecting the sampling well on the microbial communities of deep

subsurface water samples. *Environ. Microbiol.* 7, 13–21. doi: 10.1111/j.1462-2920.2004.00660.x

Basso, O., Lascourreges, J. F., Le Borgne, F., Le Goff, C., and Magot, M. (2009). Characterization by culture and molecular analysis of the microbial diversity of a deep subsurface gas storage aquifer. *Res. Microbiol.* 160, 107–116. doi: 10.1016/j.resmic.2008.10.010

Bell, E., Lamminmäki, T., Alneberg, J., Andersson, A. F., Qian, C., Xiong, W., et al. (2020). Active sulfur cycling in the terrestrial deep subsurface. *ISME J.* 14, 1260–1272. doi: 10.1038/s41396-020-0602-x

Berlendis, S., Lascourreges, J. F., Schraauwers, B., Sivadon, P., and Magot, M. (2010). Anaerobic biodegradation of BTEX by original bacterial communities from an underground gas storage aquifer. *Environ. Sci. Technol.* 44, 3621–3628. doi: 10.1021/es100123b

Bomberg, M., Lamminmäki, T., and Itävaara, M. (2016). Microbial communities and their predicted metabolic characteristics in deep fracture groundwaters of the crystalline bedrock at Olkiluoto. *Finland. Biogeosci. Discuss.* 13, 6031–6047. doi: 10.5194/bg-13-6031-2016

Bomberg, M., Nyyssönen, M., Pitkänen, P., Lehtinen, A., and Itävaara, M. (2015). Active microbial communities inhabit sulphate-methane interphase in deep bedrock fracture fluids in Olkiluoto. *Finland. BioMed Res. Int.* 2015:979530. doi: 10.1155/2015/979530

Calderer, M., Gibert, O., Martí, V., Rovira, M., de Pablo, J., Jordana, S., et al. (2010). Denitrification in presence of acetate and glucose for bioremediation of nitrate-contaminated groundwater. *Environ. Technol.* 31, 799–814. doi: 10.1080/09593331003667741

Cario, A., Oliver, G. C., and Rogers, K. L. (2019). Exploring the deep marine biosphere: challenges, innovations, and opportunities. *Front. Earth Sci.* 7. doi: 10.3389/feart.2019.00225

Chiriac, C. M., Baricz, A., Szekeres, E., Rudi, K., Dragoș, N., and Coman, C. (2018). Microbial composition and diversity patterns in deep hyperthermal aquifers from the western plain of Romania. *Microb. Ecol.* 75, 38–51. doi: 10.1007/s00248-017-1031-x

De Silva, P. N. K., and Ranjith, P. G. (2012). A study of methodologies for $CO_2$ storage capacity estimation of saline aquifers. *Fuel* 2012, 13–27. doi: 10.1016/j.fuel.2011.07.004

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

Escudié, F., Auer, L., Bernard, M., Mariadassou, M., Cauquil, L., Vidal, K., et al. (2018). FROGS: find, rapidly, OTUs with galaxy solution. *Bioinformatics* 34, 1287–1294. doi: 10.1093/bioinformatics/btx791

Feldbusch, E., Wiersberg, T., Zimmer, M., and Regenspurg, S. (2018). Origin of gases from the geothermal reservoir Groß Schönebeck (north German Basin). *Geothermics* 71, 357–368. doi: 10.1016/j.geothermics.2017.09.007

Garvis, D. G., and Stuermer, D. H. (1980). A well-head instrument package for multi-parameter measurements during well water sampling. *Water Res.* 14, 1525–1527. doi: 10.1016/0043-1354(80)90019-6

Gleeson, T., Befus, K. M., Jasechko, S., Luijendijk, E., and Bayani Cardenas, M. (2015). The global volume and distribution of modern groundwater. *Nat. Geosci.* 9, 161–167. doi: 10.1038/ngeo2590

Godin, S., Kubica, P., Ranchou-Peyruse, A., Le Hécho, I., Patriarche, D., Caumette, G., et al. (2020). An LC-MS/MS method for a comprehensive determination of metabolites of BTEX anaerobic degradation in bacterial cultures and groundwater. *Water* 12:1869. doi: 10.3390/w12071869

Goldscheider, N., Hunkeler, D., and Rossi, P. (2006). Review: microbial biocenoses in pristine aquifers and an assessment of investigative methods. *Hydrogeol. J.* 14, 926–941. doi: 10.1007/s10040-005-0009-9

Gründger, F., Jiménez, N., Thielemann, T., Straaten, N., Lüders, T., Richnow, H. H., et al. (2015). Microbial methane formation in deep aquifers of a coal-bearing sedimentary basin. *Germany. Front. Microbiol.* 6:200. doi: 10.3389/fmicb.2015.00200

Guðmundsdóttir, R., Kreiling, A. K., Kristjánsson, B. K., Marteinsson, V. Þ., and Pálsson, S. (2019). Bacterial diversity in Icelandic cold spring sources and in relation to the groundwater amphipod Crangonyx islandicus. *PLoS One* 14:e0222527. doi: 10.1371/journal.pone.0222527

Haddad, P. G., Mura, J., Casteran, F., Guignard, M., Ranchou-Peyruse, M., Sénéchal, P., et al. (2022a). Biological, geological and chemical effects of oxygen injection in underground gas storage aquifers in the setting of biomethane deployment. *Sci. Total Environ.* 806:150690. doi: 10.1016/j.scitotenv.2021.150690

Haddad, P. G., Ranchou-Peyruse, M., Guignard, M., Mura, J., Casteran, F., Ronjon-Magand, L., et al. (2022b). Geological storage of hydrogen in deep aquifers – an experimental multidisciplinary study. *Energy Environ. Sci.* 15, 3400–3415. doi: 10.1039/D2EE00765G

Haveman, S. A., and Pedersen, K. (2002). Distribution of culturable microorganisms in Fennoscandian shield groundwater. *FEMS Microbiol. Ecol.* 39, 129–137. doi: 10.1111/j.1574-6941.2002.tb00914.x

Haveman, S. A., Pedersen, K., and Ruotsalainen, P. (1999). Distribution and metabolic diversity of microorganisms in deep igneous rock aquifers of Finland. *Geomicrobiol J.* 16, 277–294. doi: 10.1080/014904599270541

Hershey, O. S., Kallmeyer, J., Wallace, A., Barton, M. D., and Barton, H. A. (2018). High microbial diversity despite extremely low biomass in a deep karst aquifer. *Front. Microbiol.* 9:2823. doi: 10.3389/fmicb.2018.02823

Hose, G. C., and Lategan, M. J. (2012). Sampling strategies for biological assessment of groundwater ecosystems. CRC CARE technical report no 21 (CRC for Contamination Assessment and Rehabilitaion of the Environment, Adelaide, Australia, 2012).

Itävaara, M., Nyyssönen, M., Kapanen, A., Nousiainen, A., Ahonen, L., and Kukkonen, I. (2011). Characterization of bacterial diversity to a depth of 1500 m in the Outokumpu deep borehole. *Fennoscandian Shield. FEMS Microbiol. Ecol.* 77, 295–309. doi: 10.1111/j.1574-6941.2011.01111.x

Jasrotia, A. S., Kumar, A., Taloor, A. K., and Saraf, A. K. (2019). Artificial recharge to groundwater using geospatial and groundwater modelling techniques in North Western Himalaya. *India. Arab. J. Geosci* 12:774. doi: 10.1007/s12517-019-4855-5

Kadnikov, V. V., Mardanov, A. V., Beletsky, A. V., Banks, D., Pimenov, N. V., Frank, Y. A., et al. (2018). A metagenomic window into the 2-km-deep terrestrial subsurface aquifer revealed multiple pathways of organic matter decomposition. *FEMS Microbiol. Ecol.* 94. doi: 10.1093/femsec/fiy152

Kadnikov, V. V., Mardanov, A. V., Beletsky, A. V., Karnachuk, O. V., and Ravin, N. V. (2020). Microbial life in the deep subsurface aquifer illuminated by metagenomics. *Front. Microbiol.* 11:572252. doi: 10.3389/fmicb.2020.572252

Kampman, N., Maskell, A., Bickle, M. J., Evans, J. P., Schaller, M., Purser, G., et al. (2013). Scientific drilling and downhole fluid sampling of a natural $CO_2$ reservoir, Green River. *Utah. Sci. Drill.* 1, 1–11. doi: 10.5194/sd-16-33-2013

Kietäväinen, R., Ahonen, L., Kukkonen, I. T., Hendriksson, N., Nyyssönen, M., and Itävaara, M. (2012). Characterisation and isotopic evolution of saline waters of the Outokumpu deep drill hole, Finland – implications for water origin and deep terrestrial biosphere. *Appl. Geochem.* 32, 37–51. doi: 10.1016/j.apgeochem.2012.10.013

Kirs, M., Kisand, V., Nelson, C. E., Dudoit, T., and Moravcik, P. S. (2020). Distinct bacterial communities in tropical island aquifers. *PLoS One* 15:e0232265. doi: 10.1371/journal.pone.0232265

Klouche, N., Basso, O., Lascourrèges, J.-F., Cayol, J. L., Thomas, P., Fauque, G., et al. (2009). *Desulfocurvus vexinensis* gen. Nov., sp. nov., a sulfate-reducing bacterium isolated from a deep subsurface aquifer. *Int. J. Sysyt. Evol. Microbiol.* 59, 3100–3104. doi: 10.1099/ijs.0.010363-0

Korbel, K., Chariton, A., Stephenson, S., Greenfield, P., and Hose, G. C. (2017). Wells provide a distorted view of life in the aquifer: implications for sampling, monitoring and assessment of groundwater ecosystems. *Sci. Rep.* 7:40702. doi: 10.1038/srep40702

Krüger, M., Dopffel, N., Sitte, J., Mahler, E., Mukherjee, S., Herold, A., et al. (2016). Sampling for MEOR: comparison of surface and subsurface sampling and its impact on field applications. *J. Pet. Sci. Eng.* 146, 1192–1201. doi: 10.1016/j.petrol.2016.08.020

Lehman, R. M. (2007). Understanding of aquifer microbiology is tightly linked to sampling approaches. *Geomicrobiol J.* 24, 331–341. doi: 10.1080/01490450701456941

Lerm, S., Westphal, A., Miethling-Graff, R., Alawi, M., Seibt, A., Wolfgramm, M., et al. (2013). Thermal effects on microbial composition and microbiologically induced corrosion and mineral precipitation affecting operation of a geothermal plant in a deep saline aquifer. *Extremophiles* 17, 311–327. doi: 10.1007/s00792-013-0518-8

Lueders, T. (2017). The ecology of anaerobic degraders of BTEX hydrocarbons in aquifers. *FEMS Microbiol. Ecol.* 93:fiw220. doi: 10.1093/femsec/fiw220

Limberger, J., Boxem, T., Pluymaekers, M. P. D., Bruhn, D. F., Manzella, A., Calcagno, P., et al. (2018). Geothermal energy in deep aquifers: a global assessment of the resource base for direct heat utilization. *Renew. Sust. Energ. Rev.* 82, 961–975. doi: 10.1016/j.rser.2017.09.084

López-Archilla, A. I., Moreira, D., Velasco, S., and Lopez-Garcia, P. (2007). Archaeal and bacterial community composition of a pristine coastal aquifer in Doñana National Park. *Spain. Aquat. Microb. Ecol.* 47, 123–139. doi: 10.3354/ame047123

Luef, B., Frischkorn, K. R., Wrighton, K. C., Holman, H. Y., Birarda, G., Thomas, B. C., et al. (2015). Diverse uncultivated ultra-small bacterial cells in groundwater. *Nat. Commun.* 6:6372. doi: 10.1038/ncomms7372

Magnabosco, C., Lin, L.-H., Dong, H., Bomberg, M., Ghiorse, W., Stan-Lotter, H., et al. (2018). The biomass and biodiversity of the continental subsurface. *Nat. Geosci.* 11, 707–717. doi: 10.1038/s41561-018-0221-6

Magoc, T., and Salzberg, S. L. (2011). FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* 27, 2957–2963. doi: 10.1093/bioinformatics/btr507

Mahé, F., Rognes, T., Quince, C., de Vargas, C., and Dunthorn, M. (2014). Swarm: robust and fast clustering method for amplicon-based studies. *Peer J.* 2:e593. doi: 10.7717/peerj.593

Mangelsdorf, K., and Kallmeyer, J. (2010). Integration of deep biosphere research into the international continental scientific drilling program. *Sci. Dril.* 10, 46–55. doi: 10.5194/sd-10-46-2010

Margat, J., and van der Gun, J. (2013). *Groundwater Around the World: A Geographic Synopsis*, London: CRC Press.

Matteucci, F., Ercole, C., and Del Gallo, M. (2015). A study of chlorinated solvent contamination of the aquifers of an industrial area in Central Italy: a possibility of bioremediation. *Front. Microbiol.* 6:924. doi: 10.3389/fmicb.2015.00924

McMurdie, P. J., and Holmes, S. (2013). Phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* 8:e61217. doi: 10.1371/journal.pone.0061217

Miyoshi, T., Iwatsuki, T., and Naganuma, T. (2005). Phylogenetic characterization of 16S rRNA gene clones from deep-groundwater microorganisms that pass through 0.2-micrometer-pore-size filters. *Appl. Environ. Microbiol.* 71, 1084–1088. doi: 10.1128/AEM.71.2.1084-1088.2005

Morozova, D., Let, D., and Würdemann, H. (2013). Analysis of the microbial community from a saline aquifer prior to CO2 injection in Ketzin using improved fluorescence *in situ* hybridisation method. *Energy Procedia* 40, 276–284. doi: 10.1016/j.egypro.2013.08.032

Mu, A., Boreham, C., Leong, H. X., Haese, R. R., and Moreau, J. W. (2014). Changes in the deep subsurface microbial biosphere resulting from a field-scale $CO_2$ geosequestration experiment. *Front. Microbiol.* 209. doi: 10.3389/fmicb.2014.00209

Mullin, S. W., Wanger, G., Kruger, B. R., Sackett, J. D., Hamilton-Brehm, S. D., Bhartia, R., et al. (2020). Patterns of *in situ* mineral colonization by microorganisms in a ~60°C deep continental subsurface aquifer. *Front. Microbiol.* 11:536535. doi: 10.3389/fmicb.2020.536535

Nuppunen-Puputti, M., Purkamo, L., Kietäväinen, R., Nyyssönen, M., Itävaara, M., Ahonen, L., et al. (2018). Rare biosphere archaea assimilate acetate in Precambrian terrestrial subsurface at 2.2 km depth. *Geosciences* 8:418. doi: 10.3390/geosciences8110418

Nurmi, P. A., and Kukkonen, I. T. (1986). A new technique for sampling water and gas from deep drill holes. *Can. J. Earth Sci.* 23, 1450–1454. doi: 10.1139/e86-138

Öhberg, A. (2006). Investigation equipment and methods used by Posiva (no. POSIVA-WR--06-81) Posiva Oy.

Osselin, F., Saad, S., Nightingale, M., Hearn, G., Desaulty, A. M., Gaucher, E. C., et al. (2019). Geochemical and sulfate isotopic evolution of flowback and produced waters reveals water-rock interactions following hydraulic fracturing of a tight hydrocarbon reservoir. *Sci. Total Environ.* 687, 1389–1400. doi: 10.1016/j.scitotenv.2019.07.066

Pedersen, K., Arlinger, J., Eriksson, S., Hallbeck, A., Hallbeck, L., and Johansson, J. (2008). Numbers, biomass and cultivable diversity of microbial populations relate to depth and borehole-specific conditions in groundwater from depths of 4–450 m in Olkiluoto. *Finland. ISME J.* 2, 760–775. doi: 10.1038/ismej.2008.43

Pedersen, K., Hallbeck, L., Arlinger, J., Erlandson, A. C., and Jahromi, N. (1997). Investigation of the potential for microbial contamination of deep granitic aquifers during drilling using 16S rRNA gene sequencing and culturing methods. *J. Microbiol. Methods* 30, 179–192. doi: 10.1016/S0167-7012(97)00066-3

Pionke, H. B., and Urban, J. B. (1987). Sampling the chemistry of shallow aquifer systems – a case study. *Ground Water Monit. Remidiat.* 7, 79–88. doi: 10.1111/j.1745-6592.1987.tb01046.x

Puente-Sánchez, F., Arce-Rodríguez, A., Oggerin, M., García-Villadangos, M., Moreno-Paz, M., Blanco, Y., et al. (2018). Viable cyanobacteria in the deep continental subsurface. *Proc. Natl. Acad. Sci. U. S. A.* 115, 10702–10707. doi: 10.1073/pnas.1808176115

Purkamo, L., Bomberg, M., Nyyssönen, M., Kukkonen, I., Ahonen, L., Kietäväinen, R., et al. (2013). Dissecting the deep biosphere: retrieving authentic microbial communities from packer-isolated deep crystalline bedrock fracture zones. *FEMS Microbiol. Ecol.* 85, 324–337. doi: 10.1111/1574-6941.12126

Rajala, P., Bomberg, M., Vepsäläinen, M., and Carpén, L. (2017). Microbial fouling and corrosion of carbon steel in deep anoxic alkaline groundwater. *Biofouling* 33, 195–209. doi: 10.1080/08927014.2017.1285914

Ranchou-Peyruse, M., Auguet, J.-C., Mazière, C., Restrepo-Ortiz, C. X., Guignard, M., Dequidt, D., et al. (2019). Geological gas-storage shapes deep life. *Environ. Microbiol.* 21, 3953–3964. doi: 10.1111/1462-2920.14745

Ranchou-Peyruse, M., Gasc, C., Guignard, M., Aüllo, T., Dequidt, D., Peyret, P., et al. (2017). The sequence capture by hybridization: a new approach for revealing the potential of mono-aromatic hydrocarbons bioattenuation in a deep oligotrophic aquifer. *Microb. Biotechnol.* 10, 469–479. doi: 10.1111/1751-7915.12426

Ranchou-Peyruse, M., Guignard, M., Casteran, F., Abadie, M., Defois, C., Peyret, P., et al. (2021). Microbial diversity under the influence of natural gas storage in a deep aquifer. *Front. Microbiol.* 12:688929. doi: 10.3389/fmicb.2021.688929

Regenspurg, S., Wiersberg, T., Brandt, W., Huenges, E., Saadat, A., Schmidt, K., et al. (2010). Geochemical properties of saline geothermal fluids from the *in-situ* geothermal laboratory Groß Schönebeck (Germany). *Chemie der Erde – Geochem.* 3, 3–12. doi: 10.1016/j.chemer.2010.05.002

Rivard, C., Bordeleau, G., Lavoie, D., Lefebvre, R., and Malet, X. (2018). Temporal variations of methane concentration and isotopic composition in groundwater of the St. Lawrence lowlands, eastern Canada. *Hydrogeol. J.* 26, 533–551. doi: 10.1007/s10040-017-1677-y

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

Ruotsalainen, P., and Snellman, M. (1996). *Hydrogeochemical baseline characterization at Romuvaara, Kivetty and Olkiluoto, Finland*. Work Report PATU-96-91e Posiva Oy, Helsinki, Finland.

Russo, T. A., and Lall, U. (2017). Depletion and response of deep groundwater to climate-induced pumping variability. *Nat. Geosci.* 10, 105–108. doi: 10.1038/ngeo2883

Sainz-Garcia, A., Abarca, E., Rubi, V., and Grandia, F. (2017). Assessment of feasible strategies for seasonal underground hydrogen storage in a saline aquifer. *Int. J. Hydrog. Energy* 42, 16657–16666. doi: 10.1016/j.ijhydene.2017.05.076

Sellers, P. (1974). On the theory and computation of evolutionary distances. *SIAM J. Appl. Math.* 26, 787–793. doi: 10.1137/0126070

Simmon, K. E., Croft, A. C., and Petti, C. A. (2006). Application of SmartGene IDNS software to partial 16S rRNA gene sequences for a diverse group of bacteria in a clinical laboratory. *J. Clin. Microbiol.* 44, 4400–4406. doi: 10.1128/JCM.01364-06

Šmigáň, P., Greksak, M., Kozankova, J., Buzek, F., Onderka, V., and Wolf, I. (1990). Methanogenic bacteria as a key factor involved in changes of town gas in an underground reservoir. *FEMS Microbiol. Ecol.* 73, 221–224. doi: 10.1111/j.1574-6968.1990.tb03944.x

Smith, R. J., Jeffries, T. C., Roudnew, B., Fitch, A. J., Seymour, J. R., Delpin, M. W., et al. (2012). Metagenomic comparison of microbial communities inhabiting confined and unconfined aquifer ecosystems. *Environ. Microbiol.* 14, 240–253. doi: 10.1111/j.1462-2920.2011.02614.x

Smith, R. J., Paterson, J. S., Sibley, C. A., Hutson, J. L., and Mitchell, J. G. (2015). Putative effect of aquifer recharge on the abundance and taxonomic composition of endemic microbial communities. *PLoS One* 10:e0129004. doi: 10.1371/journal.pone.0129004

Soares, A., Edwards, A., An, D., Bagnoud, A., Bomberg, M., Budwill, K., et al. (2019). A global perspective on microbial diversity in the terrestrial deep subsurface. *bioRxiv* 602672. doi: 10.1101/602672

Sorensen, J. P., Maurice, L., Edwards, F. K., Lapworth, D. J., Read, D. S., Allen, D., et al. (2013). Using boreholes as windows into groundwater ecosystems. *PLoS One* 8:e70264. doi: 10.1371/journal.pone.0070264 PMID: 23936176.

Stevens, T. O., McKinley, J. P., and Fredrickson, J. K. (1993). Bacteria associated with deep, alkaline, anaerobic groundwaters in Southeast Washington. *Microb. Ecol.* 25, 35–50. doi: 10.1007/BF00182128

Struchkov, I. A., and Rogachev, M. K. (2018). The challenges of waxy oil production in a Russian oil field and laboratory investigations. *J. Pet. Sci. Eng.* 163, 91–99. doi: 10.1016/j.petrol.2017.12.082

Sundaram, B., Feitz, A., de Caritat, P., Plazinska, A., Brodie, R., Coram, J., et al. (2009). *Groundwater sampling and analysis – A field guide, geoscience Australia, record 2009/27*.

Trias, R., Ménez, B., le Campion, P., Zivanovic, Y., Lecourt, L., Lecoeuvre, A., et al. (2017). High reactivity of deep biota under anthropogenic $CO_2$ injection into basalt. *Nat. Commun.* 8:1063. doi: 10.1038/s41467-017-01288-8

Weisburg, W. G., Barns, S. M., Pelletier, D. A., and Lane, D. J. (1991). 16S ribosomal DNA amplification for phylogenetic study. *J. Bacteriol.* 173, 697–703. doi: 10.1128/jb.173.2.697-703.1991

Wichels, A., Würtz, S., Döpke, H., Schütt, C., and Gerdts, G. (2006). Bacterial diversity in the breadcrumb sponge *Halichondria panicea* (Pallas). *FEMS Microbiol. Ecol.* 56, 102–118. doi: 10.1111/j.1574-6941.2006.00067.x

Wickham, H. (2009). *ggplot2, 2nd Editio. ggplot2*. New York, NY: Springer New York.

Wilkins, M. J., Daly, R. A., Mouser, P. J., Trexler, R., Sharma, S., Cole, D. R., et al. (2014). Trends and future challenges in sampling the deep terrestrial biosphere. *Front. Microbiol.* 5:481. doi: 10.3389/fmicb.2014.00481

Wu, X., Holmfeldt, K., Hubalek, V., Lundin, D., Åström, M., Bertilsson, S., et al. (2016). Microbial metagenomes from three aquifers in the Fennoscandian shield terrestrial deep biosphere reveal metabolic partitioning among populations. *ISME J.* 10, 1192–1203. doi: 10.1038/ismej.2015.185

Yu, G., Fadrosh, D., Goedert, J. J., Ravel, J., and Goldstein, A. M. (2015). Nested PCR biases in interpreting microbial community structure in 16S rRNA gene sequence datasets. *PLoS One* 10:e0132253. doi: 10.1371/journal.pone.0132253

frontiers | Frontiers in Microbiology

TYPE Methods
PUBLISHED 06 January 2023
DOI 10.3389/fmicb.2022.1064112



OPEN ACCESS

EDITED BY
Frederic Coulon,
Cranfield University,
United Kingdom

REVIEWED BY
Cristiana Cravo-Laureau,
Université de Pau et des Pays de l'Adour,
France
Susann Müller,
Helmholtz Association of German Research Centres (HZ), Germany

*CORRESPONDENCE
Massimo C. Pernice
✉ pernice@icm.csic.es

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

RECEIVED 07 October 2022
ACCEPTED 12 December 2022
PUBLISHED 06 January 2023

CITATION
Pernice MC and Gasol JM (2023) Automated flow cytometry as a tool to obtain a fine-grain picture of marine prokaryote community structure along an entire oceanographic cruise. *Front. Microbiol.* 13:1064112. doi: 10.3389/fmicb.2022.1064112



# Automated flow cytometry as a tool to obtain a fine-grain picture of marine prokaryote community structure along an entire oceanographic cruise

Massimo C. Pernice* and Josep M. Gasol

Departament de Biologia Marina i Oceanografia, Institut de Ciències del Mar-CSIC, Barcelona, Spain

On a standard oceanographic cruise, flow cytometry data are usually collected sparsely through a bottle-based sampling and with stations separated by kilometers leading to a fragmented view of the ecosystem; to improve the resolution of the datasets produced by this technique here it is proposed the application of an automatic method of sampling and staining. The system used consists of a flow-cytometer (Accuri-C6) connected to an automated continuous sampler (OC-300) that collects samples of marine surface waters every 15 min. We tested this system for five days during a brief Mediterranean cruise with the aim of estimating the abundance, relative size and phenotypic diversity of prokaryotes. Seawater was taken by a faucet linked to an inlet pump (*ca.* 5 m depth). Once the sample was taken, the Oncyt-300 stained it and sent it to the flow cytometer. A total of 366 samples were collected, effectively achieving a fine-grained scale view of microbial community composition both through space and time. A significative positive relationship was found comparing data obtained with the automatic method and 10 samples collected from the faucet but processed with the standard protocol. Abundance values retrieved varied from $3.56{\cdot}10^5$ cell $mL^{-1}$ in the coastal area till $6.87\ 10^5$ cell $mL^{-1}$ in open waters, exceptional values were reached in the harbor area where abundances peaked to $1.28\ 10^6$ cell $mL^{-1}$. The measured features (abundance and size) were associated with metadata (temperature, salinity, conductivity) also taken in continuous, of which conductivity was the one that better explained the variability of abundance. A full 24 h measurement cycle was performed resulting in slightly higher median bacterial abundances values during daylight hours compared to night. Alpha diversity, calculated using computational cytometry techniques, showed a higher value in the coastal area above 41° of latitude and had a strong inverse relationship with both salinity and conductivity. This is the first time to our knowledge that the OC-300 is directly applied to the marine environment during an oceanographic cruise; due to its high-resolution, this set-up shows great potential both to cover large sampling areas, and to monitor day-night cycles *in situ*.

KEYWORDS

**flow cytometry, marine prokaryotes, DNA stain, automatization, online cytometry**

## Introduction

Flow cytometry is a well-established technique for the measurement of the abundance of prokaryotes, pico- and nanophytoplankton and even viruses; through this method microbial populations are discriminated based on the optical properties of the cells and/or in response to stains targeting specific cellular structures or molecules, e.g., nucleic acids (Gasol and Morán, 2015). Nevertheless, it is still a field in continuous evolution. The standard protocol for prokaryotic abundance estimation (Box 1) is simple and fast but, during an oceanographic cruise, the number of samples taken per day is limited both by the navigation time between stations and by the complexity and slowness of the CTD rosette casts; in fact, this sampling method involves the collection of seawater by deploying a set of bottles that are lowered till the desired depth and then closed and recovered on board, an operation that can last between 30 min and 1 h. The combination of these two factors gives, as a result, a low-resolution picture of the distribution of microbial communities along a surface transect. A way to improve the number of samples without increasing sampling effort is the automatization both of the sampling process and of the flow cytometry determinations.

In microbial ecology, the first attempts with automated flow cytometry were developed for phytoplankton with the use of submersible flow cytometers like FlowCytobot, CytoSub-CytoBuoy, SeaFlow (Dubelaar et al., 1999; Olson and Sosik, 2007; Thyssen et al., 2008, 2011; Swalwell et al., 2011) and the automation of the process of sampling and fixation was used for prokaryotes and picoplankton to obtain samples every 12 min which were analyzed 4 h later (Martin et al., 2005, 2008, 2010). More recently, the usefulness of automatization for the study of prokaryotes was highlighted by two papers by Besmer et al., (2014, 2016) dealing with the monitoring of microbial communities in drinking waters and freshwater environments. In this case, the entire process was automated including sampling, staining and reading of the sample signal by a flow cytometer. Based on these studies, Besmer and his team developed and now commercialize an automatic sampler (OC-300, OnCyt microbiology AG, Zurich, Switzerland) that works in tandem with a simple bench-top flow cytometer (Accuri C6, BD Accuri, San Jose CA, United States), this coupled system is the one used in the present work.

Even though the OC-300 is in the market since 2017, to our knowledge there are only three publications based on its use, two of them about monitoring microbial abundance in biofuel cultures (Haberkorn et al., 2021), and bioreactor performance (Hess et al., 2021), and the third one focus on the identification of outlying observations with machine learning (Russo et al., 2021). It seems that the potential of the OC-300 for environmental monitoring, particularly in seawater, is still unexploited.

Here we present the first application of this methodology of automatic flow cytometry directly on an oceanographic vessel for the semicontinuous assessment of abundance, relative size and biomass of prokaryotes. We evaluate its feasibility, advantages and limitations and discuss the reliability of the obtained results. In addition, the huge dataset produced by this automated system (which includes the optical properties of every single cell measured) is also perfectly tailored for computational cytometry, a set of bioinformatic tools that can define populations based on common optical properties and calculate prokaryotic community phenotypic diversity. This methodology was tested during a week-long Mediterranean cruise in early June 2021. We report here on the setup and arrangements used, and present the type of data acquired. Moreover, we highlight how this method is suited for testing hypotheses about the spatial abundance, diversity and community structure of marine prokaryotes in the surface ocean.

**BOX 1**

Sampling
Seawater is collected from a Niskin bottle and prefiltered through a mesh of 20 µm.

Fixation
1800 µL of seawater +200 µL of a mix of paraformaldehyde 1% and glutaraldehyde 0.05%.

Stain
1 µL of SYBR Green I for each 100 µL of fixed sample, incubation of 10 min in the dark at RT.

Flow cytometry
The sample it is excited by a blue laser and the light emission collected by a bandpass 530/30nm filter, it runs for 2 minutes at low speed (14 µL $min^{-1}$).

## Material and equipment

The system tested in this work was constituted by the OC-300 automatic sampler connected with a BD Accuri™ C6 Plus Personal Flow Cytometer. The Accuri C6 has two lasers, blue (488 nm) and red (640 nm), four filters (FICT-533 nm, PE-585 nm, PerCP-670 nm and APC-675 nm) and two scatters, Forward (FSC, 0° ± 13°) and Side (SSC, 90° ± 13°); laser power and detector voltage are fixed in this type of machine and cannot be manipulated. The flux of the machine was calibrated before the cruise and the correct functioning of lasers and filters was checked through a quality control with the appropriate beads (CS & RUO Beads, BD). This operation was also repeated after the cruise to ensure that the flow cytometer worked correctly throughout the cruise.

The reagents to reproduce this method are (i) Ultrapure or distilled water, (ii) Sodium hypochlorite solution (approx. 1% active chlorine), (iii) Sodium thiosulfate solution (Sigma-Aldrich, 72,049-250G, purum p.a. ≥98%), (iv) SYBR Green I stain (10,000X, Sigma-Aldrich, S9430), (v) TRIS buffer (Sigma-Aldrich, T6791-1 Kg, purity ≥99.9%). Other required materials are: three Pyrex bottles of 1 liter, an amber bottle for the SGI solution, a

plastic or glass container for liquid waste, and 0.2 syringe filters (Thermo Fisher).

## Preparation of solutions

The total volume of the solutions is calculated based on the duration of the entire cruise, or the part of the cruise where measurements will be taken. The OC-300 needs *ca.* 300 mL per day for cleaning operation and, as a general rule, one liter of each of the cleaning solutions (sodium thiosulfate, sodium hypochlorite and ultrapure water) has to be prepared for a 3 days cruise. It is worth to mention that these reagents dissolve quickly in ultrapure water and, in case it is not feasible to transport ready-made solutions to the boat, it is suggested to weight the exact amount of sodium thiosulfate and TRIS at the home laboratory and prepare the final solutions directly on the ship. For the solution of TRIS buffer and sodium thiosulfate, which is the dilution media for the SGI stain, the volume is 125 mL since it has to be changed every 7–10 days to ensure the correct working of the stain.

### Sodium hypochlorite solution

Dilute 300 mL of commercial bleach, which contains around 3% of active chlorine, with 600 mL of ultrapure water to obtain a final solution with *ca.* 1% of active chlorine. Filter the solution through a 0.2 µm filter. If it has to be prepared on the ship, it will be useful to bring 60 mL plastic syringes with sterile syringe filters. Store at Room Temperature (RT).

### Sodium thiosulfate (100mM)

Mix 15.8 g of sodium thiosulfate powder in 1 L of ultrapure water in a glass bottle, shake the mix until all powder is dissolved and autoclave it. For a long cruise, bring 1 L bottle for each 3 days, plus an extra one. This solution can be stored at RT.

### TRIS buffer (10mM, pH 8) with sodium thiosulfate (50mM)

Add 1.2 g of TRIS base and 7.9 g of sodium thiosulfate to 900 mL of ultrapure water. Measure and adjust pH to 8.0 by adding HCl. Autoclave and store it in the fridge.

### SYBR green I solution

Filter 125 mL of solution number 3 through a 0.2 Swinnex filter and deposit it into an amber bottle, then add 25 µL of SYBR Green I (SGI) stock solution, mix well (Final concentration 2X). Prepare this solution right before use. This solution lasts for 7–10 days at RT.

# Materials and methods

The objective of this procedure is to improve the resolution of standard flow-cytometry sampling for surface heterotrophic prokaryotes during an oceanographic cruise. In order to validate this procedure, we tested the automatic sampler directly on a ship to (1) compare with values of abundance obtained with the standard protocol, (2) identify possible issues connected with the ship's environment.

## A brief description of the OC-300 functioning

The OC-300 is an automatic sampler which was first developed to work in conjunction with a flow cytometer of the Becton Dickinson Accuri type, although the developers informed us recently about the possibility to connect the machine to a flow cytometer belonging to other brands (e.g., Beckman Coulter Cytoflex). Once the sampler is connected to the flow cytometer (Figure 1), all operations are mediated by a software (cyON) that takes control over the Accuri's regular software and works based on python scripts. In a normal routine, and after the initial cleaning, the system uses a glass syringe to take both the sample and the stain, mix them (1:1) and incubate them in a chamber for 10 min at 37°C. After that, the syringe mediated again the movement of the sample from the incubation chamber to the flow cytometer in order to be analyzed. Sample run for 1 min at fast speed (66 µL $min^{-1}$), and then a cleanup round is imposed. Other procedures involving several staining steps can be applied (e.g., NADS staining, which involve the use of Propidium Iodide at the same time of SGI). Specific characteristics, such as the pacing of the sampling, the duration and temperature of the incubation, the number of sampling tubes involved and the number of stains used, can be set by manipulating the python scripts. A minimum time of 15 min between samples is mandatory due to the cleaning operation; the machine cleans the tubing system, the syringe and the mix chamber with sodium thiosulfate (which is a quencher of SGI), sodium hypochlorite, and ultrapure water. The machine allows a maximum of 11 different samplings tubes (i.e., 11 samples can be taken at each run) but for environmental measures one tube is enough; nevertheless, considering the duration of the cruise, it was preferred to sample seawater with 3 tubes that took a sample in succession every 15 min, since in the event that one or two tubes get clogged, without operator notice, at least one measurement would be obtained every 45 min. There is a distance limitation between the automatic sampler and the source of water based on the maximum length of the tubing (120 cm is the max distance tested by the developers).

## Setting of the system

Samples are obtained from a faucet connected with the ship's intake pump which samples water at 5 m. Submerged-pump systems are generally present on a normal research vessel. The water was let run for around 30 min before starting. The continuous flow of seawater fell directly from the faucet into a 1 l bucket which was tilted to let the water pour over. Three sampling tubes were taped to the bucket in order to allow the collection of the circulating water near the surface and avoid to take the

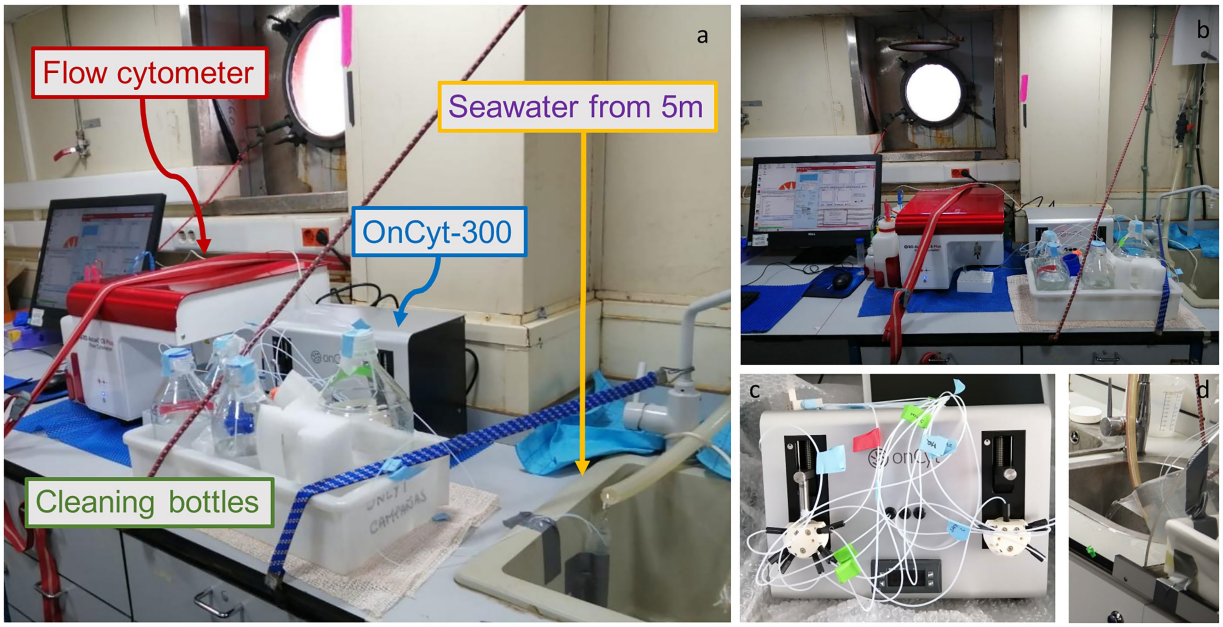



FIGURE 1
Photos of the system set-up, **(A)** Lateral view of the system with a falcon as a water collecting system from the pump. The little white tubes bring water directly to the OC-300 (grey box behind the cleaning bottles) that then sends the sample to the flow cytometer (red lid machine) **(B)** Frontal view of the system **(C)** OC-300 tubes preparation **(D)** Bucket as alternative water collector.

relatively still water toward the bottom of the bucket. It is important to fix the bucket so that the water keeps being continuously renewed. An alternative, as shown in Figure 1 is to direct the flow to a falcon tube and sample water from the top of it, in this case it is fundamental to build a system resembling a half funnel to compensate for the change of the position of the water flow due to ship movement (Supplementary Figure S1). Once the sampling tube is taped to the bucket/falcon, it is essential to measure the length of the sampling tube and incorporate that value into the script for a correct functioning of the sampler, since the total volume incorporated depends on the length of the tube. With a longer tube, more pressure will be needed to deliver the right amount of liquid into the syringe.

Considering the daily need to refill with cleaning reagents, it is useful to close the bottles with parafilm, fixing the tubes to the bottles with tape. For the SGI a hole was drilled in the middle of the amber bottle lid. All the bottles were placed together in a box, well subjected and protected from movement and potential breaking. Special care was taken with the amber glass bottle containing SGI which is a known mutagenic substance (Bourzac et al., 2003). The entire system occupies a minimum surface of around 1 $m^2$ placing the computer on top of the Accuri-C6.

## Procedures

Table 1 lists a series of steps and tips important for the functioning of automatic flow cytometer on a ship before, during, and after the cruise with the goal of facilitating system operations, from sampling to data management. When the cyOn software takes over the Accuri, a predefined template is automatically opened to visualize the measuring of events for each sample and check the effectiveness of the cleaning steps through the absence of events in the cytograms when ultrapure water was running. The template establishes the threshold of data collection (in this case 800 in green fluorescence, FL1), the speed of reading (fast) and a gate for the population of interest in a cytogram of green versus red fluorescence (see Gasol and Morán, 2015), which separates cyanobacteria from heterotrophic bacteria. Histograms of green fluorescence versus counts, and green fluorescence over time are also displayed in the template. Since only one stain was used, a normal practice for the abundance quantification protocol with marine microbial samples, no quenching depletion was expected and, for the same reason, no compensation was applied to correct for spillover. SGI does have spillover over the red channel (FL3), which leads to a visualization of heterotrophic bacteria as a diagonal population in a cytogram of green versus red fluorescence, but this normally does not interfere with the gating of the autotrophic population (Supplementary Figure S2). Once the cruise was completed, the total prokaryote community was manually gated for all cytograms of side scatter (SSC) versus green fluorescence (FITC) using the second software associated with the OC-300 named cyPlot. The adequacy of the gate was visually checked for the entire dataset, and the number of events per sample extracted as an excel table. Abundance then was calculated from gated events $\mu L^{-1}$ times 2 (because of the dilution with SGI is 1:1), and then times 1,000 to obtain cells $mL^{-1}$. The raw cytograms of this cruise ($n=366$) are deposited in the SEANOE repository (seanoe.org) and available through this link https://doi.org/10.17882/90671. Abiotic parameters (Temperature, Salinity and Conductivity) of the surface water were collected in

TABLE 1 Key steps of the procedure.

| Before the cruise |
|---|
| Prepare the solutions as described in the protocol |
| Prepare scissors, ruler and tape for make new tubes on boat |
| Set parameters (e.g., fluorescence threshold) in the Accuri C6 template |
| Check the functioning of the OC-300 and of the Flow-cytometer |
| Check the presence of ultrapure water machine on boat |
| Check the access to a continuous seawater pump |
| Check the space near the seawater pump (at least 1 square meter) |
| **During the cruise** |
| Check the correct functioning of Flow Cytometer |
| Set up the system (Accuri C6 + Computer + OC-300 + Bottles) |
| Fill the cleaning bottles with their respective solutions |
| Be sure that the sampling tubes are less than 120 cm |
| Change the volume in the script according to the sample tube length |
| Start the initializing procedure |
| Start to measure |
| Refill everyday till 1 L the cleaning solutions |
| At the end clean the system |
| **After the cruise** |
| **For abundance** |
| Open the cytogram with cyPlot |
| Gate the desired population |
| Collect events, average and mean of the fluorescent channels |
| **For computational cytometry application** |
| Extract .fcs files |
| Rename the files and group it in one folder |
| Use FlowCore to: Visualize, transform, normalize, gates and subset your samples |
| Use PhenoFlow to calculate phenotypic diversity |
| Use FlowSOM to identify different groups |

continuous mode along the entire cruise by a SBE 21 SeaCAT Thermosalinograph (Sea-Bird Scientific, United States).

## Statistical analysis and computational analysis

Regression analysis was used to compare between regular flow cytometry protocol and the automatic method whereas the correlation between the prokaryotic abundance and the abiotic parameters was tested with a Pearson correlation coefficient (both performed with Rstudio). The computational analysis required the cytograms of the entire sampling period to be pooled together, in this regard it is important to highlight that, although the cyON software generated folders named with a time signature (e.g., 2021-06-03_11-01-38), the files contained in each folder are always named with the same alphanumeric code (e.g., A04.fcs, A08.fcs and so on). For the present analysis each file was renamed as foldername_filename.fcs (e.g., 2021-06-03_11-01-38_A04.fcs). The python script used for this task is presented as Supplementary material.

Cytograms were visualized with the flowCore R package (Hahne et al., 2009), first the fluorescence values were transformed (using the arcsin function) for better visualization, in a second step the events of interest were manually gated in a cytogram of forward scatter versus green fluorescence to delete noise, and finally all cytograms were normalized against the maximum green fluorescence being the normalization a fundamental step for successive fingerprint analysis (Props et al., 2016). In order to allow diversity comparisons and also to reduce computational time, the number of cells was subsampled to the minimum number of cells retrieved among our samples (in this case, 10,303).

For the definition of clusters, groups of cells with similar values of fluorescence and scatter were grouped in 100 clusters and were identified with R package FlowSOM (Van Gassen et al., 2015). These clusters were grouped again to simplify the analyzes in 10 entities named metacluster.

# Results

## Comparison of the OnCyt-300 method versus flow cytometry standard protocol

A general stain used for acid nucleic in bacteria is SGI which is usually diluted in DMSO and unfrozen the same day of the measurement. SGI is chemically very stable and allows for several cycles of freezing/unfreezing without evident changes in its efficiency. With the automatic sampler, although the stain used is the same, the conditions are quite different. In this method SGI is dissolved in a solution of TRIS plus sodium thiosulfate and kept at room temperature for 7–10 days. The efficiency of this SGI solution was tested by the developer company (Besmer et al., 2014) and will be discussed in the following paragraphs.

The developers found a very good relationship ($R^2$ of 0.99) when comparing unfixed samples enumerated with automatic and not automatic methods (personal communication). Since the automatic sampling does not include a fixation step, and knowing from the literature (Bullock, 1984) that fixed samples tend to stain better due to cell wall permeabilization caused by fixation, allowing an easier penetration of the stain inside the microbes, it was decided to test the similarities between our usual protocol (Box 1, fixation with a mix of paraformaldehyde and glutaraldehyde, incubation at room temperature, Gasol and Morán, 2015) against the OC-300 protocol (no fixation, incubation at 37°C). For the usual protocol, water samples were taken from the bucket with a pipette near the area where the OC-300 tubes were taped and, as much as possible, at the same time of the automatic sampling. A total of 10 samples were collected. A significant positive relationship was observed between the two methods ($n = 10$, $R^2 = 0.30$, $p = 0.0472$, slope = 1.01), with the OC-300 protocol giving slightly higher values (average 20%

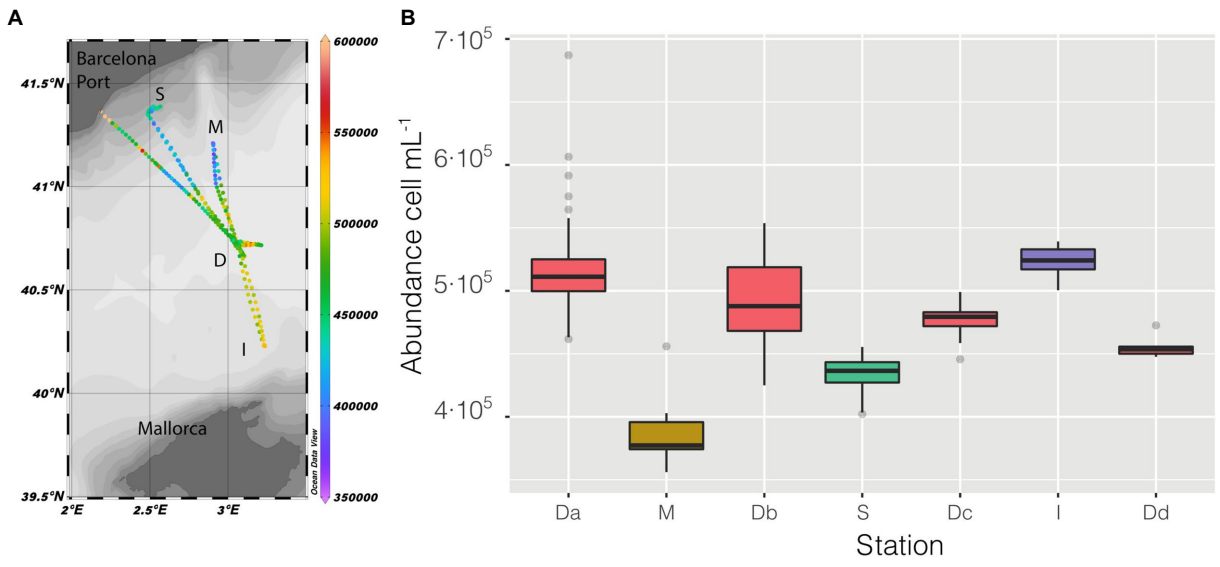



FIGURE 2
**(A)** Map of the cruise, with prokaryotic abundance (cell $mL^{-1}$) along the entire cruise presented in a color scale, the depth of the stations is also presented as shaded. **(B)** distribution of prokaryotic abundance, presented as a boxplot for each station, the time of permanence in each station varies (Da=21h, M=2h, Db=6h, S=9h, Dc=7h, I=3h, Dd=1h).

more). This difference can possibly be explained by likely errors in the manual sampling, which has to be done at the same time as the automatic machine for this comparison, and by the incubation at 37°C, in addition to the fixation itself.

## Temporal and spatial variations of prokaryotic abundance

Sampling started at station D (Figure 2A), and followed a kind of flower shape going out from the center and back to it, twice toward the coast (stations M and S) and once between the islands of Mallorca and Minorca (Station I), the vessel stopped at the center four times (named in Figure 2B as Da, Db, Dc, Dd) and the measurements lasted until the moment it made landfall in the Barcelona Harbor.

Abundance varied through space and time from a minimum of $3.56 \cdot 10^5$ cell $mL^{-1}$ in the coastal station M till a maximum of $6.87 \cdot 10^5$ cell $mL^{-1}$ in the clear waters of station D, this excepting the last 3 points when entering the Barcelona port, where abundances peaked to $1.28 \cdot 10^6$ cell $mL^{-1}$. Of the abiotic variables measured, conductivity was the one that better explained the variability of abundance ($n=366$, $R=0.36$, $p<0.001$) but, when not considering the harbor area (the 3 final values), the $R$ increased to 0.60. The correlation of abundance with temperature and salinity were $R=0.51$ and $R=0.38$, respectively, (for both $n=363$, $p<0.001$). Excluding the extreme values found in the harbor area, the values of prokaryotic abundances decreased going from station D to more coastal areas (M and S, Figure 2B), while they increased when sailing South-West toward the islands. Considering only station D (Figure 2 boxplot, red squares), the median value of prokaryotic abundance decreased over time from day 1 to day 4. This trend


FIGURE 3
Prokaryotic abundance along the transect between station D and S **(A)** and salinity (psu) of the same transits **(B)**.

could easily have been overlooked with Niskin-based sampling since we would have had a single value instead of a median of multiple values for the same point. Interestingly, comparing the transits back and forth from station D to station S (Figure 3A), the

abundance patterns showed a strong similarity, this is particularly important considering that the ship spent 9 h at station S. A similar situation was found comparing the transects back and forth from station D to M (data not shown). It appeared that different water masses had a stable bacterial abundance and that there was a marked change at *ca.* 41° of latitude, the change of water conditions is evidenced by the change in salinity (Figure 3B).

The observed variations in abundance along the cruise could be due by a progressive weakening of the stain or could be a true response of the prokaryotic community to changes in the environment. It is true that considering only station D, the abundance was better correlated with time ($n = 149$, $R = -0.59$, $p < 0.001$) than with conductivity ($n = 149$, $R = 0.47$, $p < 0.001$) but, considering the entire cruise ($n = 366$), time did not have a significant correlation with abundance ($p = 0.4332$) while conductivity (an environmental factor) became significant ($R = 0.36$, $p < 0.001$). If we do not consider the extreme high values of abundance found in the harbor, the correlation with conductivity ($n = 363$, $R = 0.59$, $p < 0.001$) is even stronger than the one with time ($n = 363$, $R = -0.36$, $p < 0.001$). This led us to think that differences in abundance are determined by the environmental difference and not by the weakening of the stain.

We also measured a full 24 h cycle (85 samples) in station D. The median of bacterial abundances during the daylight hours (Median = $5.20{\cdot}10^5$ cells $mL^{-1}$) was higher than the median of the samples taken during the night (Median = $5.01{\cdot}10^5$ cells $mL^{-1}$, with a difference of $0.18{\cdot}10^5$, that is 3.6%, Figure 4), a difference statistically supported (*t*-student $p < 0.0001$). This day-night effect disappeared when considering the entire cruise, indicating that the daily variability was lower than the spatial one. The net growth rate, calculated as the slope of the natural logarithm of abundance versus time was $0.04 \pm 0.01$ which would indicate a duplication time of 18 days, quite reasonable (see compilation in Kirchman, 2016).

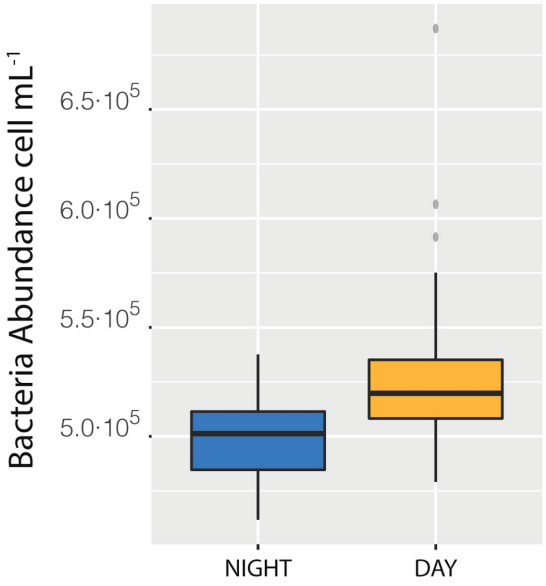



FIGURE 4
Difference between the median of prokaryotic abundance of the samples collected during the day as compared with sample collected during the night in station Da.

## Storm and movement of the ship

In the early hours of June 5th, a medium-intensity storm occurred during the transit from station D to S. As it can be observed in the variations of prokaryotic abundance over time (Supplementary Figure S3), the storm did not seem to affect the measurements: the abundance values were in the same range as in the hours before the storm and, as shown before in Figure 3, they closely resembled the values of the day after the storm in a similar geographic position. The lack of effect of (reasonable) ship movement on machine performance was also observed and confirmed numerically on a successive cruise where ship movement data (pitch and roll) were available (Supplementary Figure S4). Variation of prokaryotic abundance seems independent from the movement of roll. When the ship was still, it was found from the lowest to the highest value, whereas when the roll was higher a narrow range of values was observed and this movement was not correlated with abundance. The progressive narrowing of the range of abundances can be easily explained by a greater homogenization of the waters as movement increases. Actually, the abundance values corresponding to periods of high roll motion were in the range of most other data points, in other words, the abundance measurements appeared to be independent of ship movement.

## Relative size and biomass

Forward scatter (FSC) after illumination with, e.g., a blue laser can be taken as a relative measure of size (Allman et al., 1990; Gasol and Del Giorgio, 2000). An advantage of the automatic sampler is that its gating software (cyPlot) allows easy extraction of the mean and the median intensity of all variables, including FCS, for each sample point of the entire cruise. In the case study, higher average cell sizes were observed near the coast compared to stations D and I (Supplementary Figure S5), and were generally inversely proportional to abundance: excluding the 3 points in the harbor, where size and abundance changed abruptly, the Pearson correlation between mean relative size per sample and abundance was significant with a moderate inverse relationship ($n = 363$, $R = -0.57$, $p < 0.001$). Relative size was also negatively correlated with several environmental variables including temperature ($n = 363$, $R = -0.52$, $p < 0.001$), salinity ($n = 363$, $R = -0.55$, $p < 0.001$) and conductivity ($n = 363$, $R = -0.67$, $p < 0.001$). An approximation

of the actual bacterial size could be made by calibrating with the FCS of beads of known size (0.5 µm Fluoresbrita Multifluorescent Microsphere by Polysciences Inc.), and the diameter of the retrieved cells ranged from 0.1 and 0.4 µm. Taking this diameter as a starting point and considering each cell in the shape of a sphere, it was possible to calculate the bacterial biovolume, and then, using the following equation biomass pgC $cell^{-1} = 0.2*V^{0.72}$ (Moran et al., 2015), the average biovolume was converted into biomass; average cell biomass times cell abundance resulted in total biomasses ranging between $1.58 \cdot 10^2$ and $4.77 \cdot 10^3$ pgC $mL^{-1}$ being higher in coastal stations S, M and in the harbor area.

## Computational cytometry

As stated above, the automated process produces a dataset useful for computational cytometry. Each gated cell is described by a set of measurements consisting of the height and area of the emission peak for each collector filter and representing a unique optical signature for each particle (Rubbens and Props, 2021). Based on this information it is possible to (i) identify groups of particles with similar characteristics, (ii) calculate alpha diversity, and (iii) compare communities (beta diversity).

Among the groups identified by the algorithm as metaclusters, three were easily recognizable as known classical populations: High NA content prokaryotes, Low NA content Bacteria, and cyanobacteria (Figure 5A). Cyanobacteria (Supplementary Figure S6) were more abundant in coastal stations compared with station D and decreased in abundance near the harbor. High NA content bacteria could be caused by the presence of more RNA copies and/or dividing cells, and it has been suggested to be a proxy of more active cells (Li et al., 1995). The ratio High/Low has often been used as a possible indicator of the degree of community activity (a more active community has a higher ratio). These higher activity bacteria were more abundant at coastal stations compared with station D and I. The ratio also increased in the samples closer to the harbor (Figure 5B).

Alpha diversity was calculated with the R package PhenoFlow as Hill diversity (Props et al., 2016). Hill diversity could be interpreted as 'effective number of species' and, in this case, similarly to the Simpson index, the calculated value represents the number of equally abundant species required to generate an identical diversity as that of the observed microbial community (Jost, 2006). Here "species" should be intended as phenotypes, i.e., each one of the bins collecting cells with similar fluorescence and scatter properties. Briefly, each cytogram was divided into a grid of $128 \times 128$ bins and the particles density in each bin was the basis for diversity calculations. Hill diversity index D2 (Figure 6A) ranged from 5,742 to 7,731, and excluding the harbor, had a strong inverse relationship with abundance ($n = 356$, $R = -0.65$ and $p < 0.001$) and, actually, had a higher value in the coastal area above 41° of latitude (Figure 6B), except for the harbor where went back down. Alpha diversity had a strong inverse relationship both with salinity ($n = 356$, $R = -0.67$, $p < 0.001$) and conductivity ($n = 356$, $R = -0.63$, $p < 0.001$) and less so with temperature ($n = 363$, $R = -0.42$, $p < 0.001$).

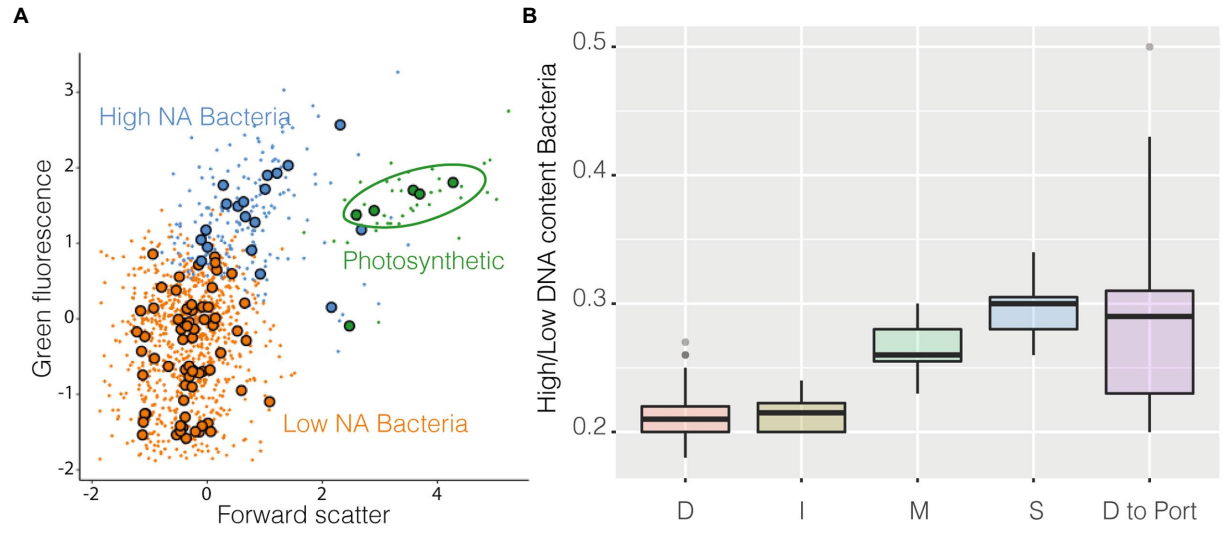



FIGURE 5
FlowSOM results, panel **(A)** is a resume cytogram that shows three of the groups identified by the algorithm (metaclusters), photosynthetic prokaryotes in green, high NA content prokaryotes in blue, and low NA content bacteria in orange, large dots represent the geometrical center of each of the 100 clusters on which metaclusters are built; on panel **(B)** is shown as boxplot the ratio High/Low of all the samples belonging to the different stations and the final transect between station D and the harbor.

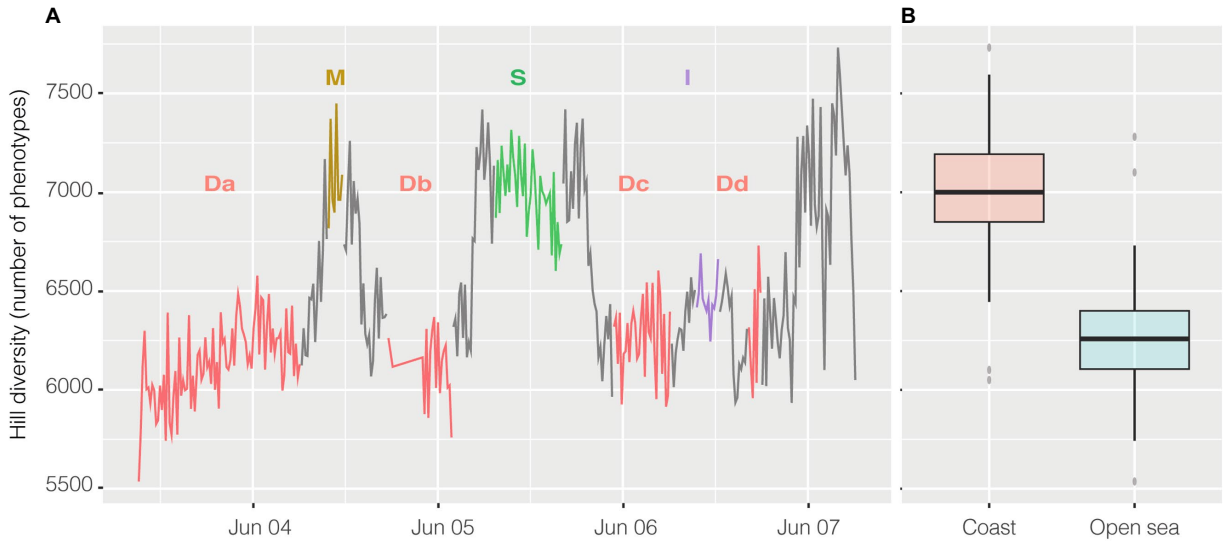



FIGURE 6
Hill Diversity, panel **(A)** shows the variation of alpha diversity along the entire cruise. On the x axis time is plotted whereas on the y axis the values of Hill diversity are expressed as "number of phenotypes" i.e. the number of bins collecting cells with shared characteristics. The same data are visualized in panel **(B)** divided in coastal samples (from 41° of latitude to higher latitudes) and open ocean samples (from 41° to lower latitudes). Transits between stations reported on the graph are in grey.

## Possible problems, pitfalls, artifacts and suggested solutions

Our results highlight that the method presented here is robust and very useful, however, some problems/risks should be taken in account. They are listed below from least to most important.

Use of glass bottles: Considering the use of corrosive, toxic and teratogen chemicals, it is worth to stress the importance of a careful setting of the system. Special attention should be paid to the protection of the SGI amber bottle. If the cleaning solution is closed with only parafilm, remember to fix it as much as possible and away from the computer and other electrical components. To prevent incidents, it is possible to use special lids.

Storm and spills: Although we were able to demonstrate that the method works well during a storm, there are 3 issues to be aware of: (1) the liquid inside the incubation chamber could spill out with strong movements (2) the collection tubes of the cleaning bottles could be above of the liquid level at the time they operate, (3) the pipe of the inlet seawater pump could move and not pour water into the collector container where the sample tubes are placed. There are no possible solutions for an internal spill of the incubation chamber and it is important to double check the data to ensure that they are biologically sound. On the other hand, there is no risk for the operator since the spill would occur inside the chamber. About possible problems with the bottles, it is important to periodically check during a storm that the tubes are sitting near the bottom of the bottle and tape them in place. To keep collecting water in stormy conditions, try using a larger collection container (if there is a strong flow that ensures the continuous renewal of water) or, if a falcon is used as the collector, a "leaky" half funnel might be useful for intercept the moving stream of water (Supplementary Figure S1).

Unfiltered water and clogging of the tubes: Water collected from a Niskin is normally prefiltered through a 20 µm mesh, this step is absent with automatic sampling. A bloom of large phytoplankton or, in general, high turbidity levels may cause clogging in both the flow cytometer and the OC-300. A possible solution for this situation is to add a mesh to the faucet that provides water, cleaning it every day. This could be important especially in coastal areas but not in clear water oceanic areas. The mesh should be rinsed daily.

Clogging of the valve: The valve where the syringe is mounted could get clogged during the cruise. If this happens, less liquid is incorporated by the system and irregular bubbles appear. The fastest way to solve this problem is to disassemble the valve, clean it (with compressed air through each hole, one night in ultrapure water, and again a round of compressed air in each hole) and put it back in place. If the machine is not new, it is probably a good idea to do it anyway before a cruise.

# Discussion

Automated flow cytometry has the obvious advantage of improving sampling resolution; for example, in the present study, the regular sampling plan for the MIAU oceanographic cruise involved 4 surface samples, one for each station, while with the OC-300 we sampled surface waters 366 times, thus increasing the number of samples by more than 90 times. In terms of economics,

automating the process allows for better use of sailing time, even during a storm, and the effort to set up the system on the first day of cruising is fully balanced by the small amount of work required during the rest of the days, and by the number of samples taken. Considering that for each sample point we can extract information about abundance, cellular size, biomass, diversity and (phenotypic) community composition, the dataset produced is extremely useful.

Data obtained with the automatic sampler were compared with samples analyzed with the usual protocol (Box 1) which was applied to samples belonging to water from the inlet pump ($n = 10$), this comparison shown that OC-300 tend to give slightly higher values showing a moderate direct relationship with data from the standard protocol.

Overall, the results here presented are sound, they appear to have an ecological meaning and they are in agreement with the results found with the normal method. There is a clear geographical structure for all the studied parameters, probably determined by a change in the main water mass around latitude 41° (Supplementary Figure S7), which is associated with a change in salinity. Compared to stations D and I, the coastal stations M and S presented lower abundance (perhaps counterintuitively), higher diversity, and an apparently more active prokaryotic community. Interestingly, the area near the Barcelona harbor has a very different pattern compared with the coastal stations. When we entered the port, in the last hour of sailing, prokaryotic abundance peaked reaching the highest values of the entire cruise, whereas diversity and average cell size went down abruptly. The ratio High/Low NA content prokaryotes was also highest in the port area whereas photosynthetic microbes decreased remarkably. The effect of the harbor structure over nearby coastal communities is seldom inspected, yet thanks to the automatic sampling we were able to obtain hints of the microbial communities inside the man-made structure.

The growth rate observed during the diel cycles, of $0.04 \pm 0.01$, is similar to that observed at the same station in 1995, which was of $0.07 \pm 0.03$ (Pedrós-Alió et al., 1999). This low value could be explained by biotic interactions (grazing and/or viral lysis) or by the lack of sufficient nutrients for the growth of the population. We have not measured either of these variables in continuum.

Most aquatic microbial studies associate flow cytometry with abundance values only, but the huge amount of data produced (for each of the 366 sampling points there were 10,000 cells subsampled and for each cell there were 6 difference characteristics, scatter or fluorescences, collected) highlight the power of flow cytometry as a high-throughput technique. Handling of this type of dataset could have been difficult and tedious in the past but luckily, paralleling what has happened with DNA sequencing, several bioinformatic tools have been developed recently under the name of "computational cytometry." Applying these techniques to the dataset allowed to identify the classical groups (High and Low NA content prokaryotes, photosynthetic prokaryotes) in a standardized way (i.e., not depending on the operator decisions), which results on a decrease of the variability between operators and between studies when a subpopulation is defined. The ratio High/Low suggests that more active cells were present in the coastal area (Station M and S), although it is worth to mention that the number of chromosome copies, as well as their size, are strain specific, and hence, it could be that an inactive or not dividing cell with a high number of chromosomes is stained with the same intensity, and so occupies the same position in a cytogram, as a small very active cell which has more mRNA copies or is actively synthesizing DNA (Cichocki et al., 2020).

Moreover, it was also obtained an estimation of diversity (i.e., phenotypic diversity) based on flow cytometry data, which was found to correlate well with the 16S rRNA gene diversity in several environments including the ocean (García et al., 2015). Here the estimated diversity values were not random, they were higher in the coastal area than in the open sea, as there were probably more niches and resources to exploit compared to the open ocean, but they dropped when approaching the harbor likely because of the presence of a more polluted environment. It is important to stress the relevance and novelty of determining microbial diversity at this unprecedent scale. The diversity of a similar transect was analyzed by Pommier and colleagues in 2010 (Pommier et al., 2010), based on pyrosequencing of the 16S rRNA gene. These authors found that station D and M had lower values of diversity compared with more coastal samples. At that time, they observed, as we did, a lower value of diversity in station D.

Although we presented here only data for prokaryotes, and given that FC appears well suited to analyze a whole set of microbial populations (Gasol and Morán, 2015), this method could easily be applied to phytoplankton, and, with some adjustments, to heterotrophic flagellates and viruses, opening the possibility of reaching a comprehensive picture of the environment at a fine scale, collecting information about interactions such as grazing, infection and carbon flux in general.

In addition, the OC-300 system allows monitoring of the prokaryote abundance in real time. Considering the detection of the abrupt change in abundance around 41° of latitude it could be used in the future, in association with physical parameter measurements, as an easy way to detect changes in water masses defined for their different microbial communities. The potential of this methodology to detect rapid environmental changes and microenvironments along the track is huge and yet unexplored.

Automation of flow cytometry presents some limitations, in the specific the main mechanical issue we encountered was the clogging of the valve and its relative cleaning which, although fixable, involves the loss of one day of sampling, as we experienced in a posterior cruise. It is also important to stress that the automated sampling is limited to the ocean surface only since, at the moment, there is no technology available to pump waters continuously from other depths. Despite these limitations, overall

we found the automatic sampler to be very useful and a clear improvement to the toolbox of oceanographic cruises. In particular, this methodology is especially adequate to detect temporal trends, e.g., between day and night, or spatial ones between crossed water masses, e.g., the change in abundances along the cruise in station D could have easily been missed with standard sampling.

## Data availability statement

The datasets presented in this study can be found in online repositories. The name of the repository and accession link can be found at: SEANOE; https://doi.org/10.17882/90671.

## Author contributions

MP performed all the tasks during the cruise and the analyzes of the dataset. MP and JG wrote the manuscript. All authors contributed to the article and approved the submitted version.

## Funding

The OC-300 and the cruise were funded by project MIAU (RTI2018-101025-B-I00) from the Spanish Ministry of Science and Innovation. We acknowledge project PID2021-125469NB-C31 form the same Ministry, and the generic support by the 'Severo Ochoa Centre of Excellence' accreditation (CEX2019-000928-S).

## Acknowledgments

We thank the crew of the Garcia del Cid for the help with the setting of the system. We would like to thank Xabier López for the help with the python scripts. We specially thank Konstanze Schiessl and the OnCyt team for all the support and tips on how to use the machine.

## 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.2022.1064112/full#supplementary-material

## References

Allman, R., Hann, A. C., Phillips, A. P., Martin, K. L., and Lloyd, D. (1990). Growth of *Azotobacter vinelandii* with correlation of coulter cell size, flow cytometric parameters, and ultrastructure. *Cytometry* 11, 822–831. doi: 10.1002/cyto.990110708

Besmer, M. D., Epting, J., Page, R. M., Sigrist, J. A., Huggenberger, P., and Hammes, F. (2016). Online flow cytometry reveals microbial dynamics influenced by concurrent natural and operational events in groundwater used for drinking water treatment. *Sci. Rep.* 6, 1–10. doi: 10.1038/srep38462

Besmer, M. D., Weissbrodt, D. G., Kratochvil, B. E., Sigrist, J. A., Weyland, M. S., and Hammes, F. (2014). The feasibility of automated online flow cytometry for in-situ monitoring of microbial dynamics in aquatic ecosystems. *Front. Microbiol.* 5, 1–12. doi: 10.3389/fmicb.2014.00265

Bourzac, K. M., LaVine, L. J., and Rice, M. S. (2003). Analysis of DAPI and SYBR green I as alternatives to ethidium bromide for nucleic acid staining in agarose gel electrophoresis. *J. Chem. Educ.* 80, 1292–1296. doi: 10.1021/ed080p1292

Bullock, G. R. (1984). The current status of fixation for electron microscopy: a review. *J. Microsc.* 133, 1–15. doi: 10.1111/j.1365-2818.1984.tb00458.x

Cichocki, N., Hübschmann, T., Schattenberg, F., Kerckhof, F. M., Overmann, J., and Müller, S. (2020). Bacterial mock communities as standards for reproducible cytometric microbiome analysis. *Nat. Protoc.* 15, 2788–2812. doi: 10.1038/s41596-020-0362-0

Dubelaar, G. B. J., Gerritzen, P. L., Beeker, A. E. R., Jonker, R. R., and Tangen, K. (1999). Design and first results of CytoBuoy: a wireless flow cytometer for in situ analysis of marine and fresh waters. *Cytometry* 37, 247–254. doi: 10.1002/(SICI)1097-0320(19991201)37:4<247::AID-CYTO1>3.0.CO;2-9

García, F. C., Alonso-Sáez, L., Morán, X. A. G., and López-Urrutia, Á. (2015). Seasonality in molecular and cytometric diversity of marine bacterioplankton: the re-shuffling of bacterial taxa by vertical mixing. *Environ. Microbiol.* 17, 4133–4142. doi: 10.1111/1462-2920.12984

Gasol, J. M., and Del Giorgio, P. A. (2000). Using flow cytometry for counting natural planktonic bacteria and understanding the structure of planktonic bacterial communities. *Sci. Mar.* 64, 197–224. doi: 10.3989/scimar.2000.64n2197

Gasol, J. M., and Morán, X. A. G. (2015). *Flow Cytometric Determination of Microbial Abundances and its Use to Obtain Indices of Community Structure and Relative Activity*. Berlin: Springer, 159–187.

Haberkorn, I., Off, C. L., Besmer, M. D., Buchmann, L., and Mathys, A. (2021). Automated online flow cytometry advances microalgal ecosystem management as in situ, high-temporal resolution monitoring tool. *Front. Bioeng. Biotechnol.* 9, 1–13. doi: 10.3389/fbioe.2021.642671

Hahne, F., Lemeur, N., Brinkman, R. R., Ellis, B., Haaland, P., Sarkar, D., et al. (2009). FlowCore: a Bioconductor package for high throughput flow cytometry. *BMC Bioinformatics* 10:106. doi: 10.1186/1471-2105-10-106

Hess, A., Baum, C., Schiessl, K., Besmer, M. D., Hammes, F., and Morgenroth, E. (2021). Stagnation leads to short-term fluctuations in the effluent water quality of biofilters: a problem for greywater reuse? *Water Res. X* 13:100120. doi: 10.1016/j.wroa.2021.100120

Jost, L. (2006). Entropy and diversity. *Oikos* 113, 363–375. doi: 10.1111/j.2006.0030-1299.14714.x

Kirchman, D. L. (2016). Growth rates of microbes in the oceans. *Annu. Rev. Mar. Sci.* 8, 285–309. doi: 10.1146/annurev-marine-122414-033938

Li, W. K. W., Jellett, J. F., and Dickie, P. M. (1995). DNA distributions in planktonic bacteria stained with TOTO or TO-PRO. *Limnol. Oceanogr.* 40, 1485–1495. doi: 10.4319/lo.1995.40.8.1485

Martin, A. P., Zubkov, M. V., Burkill, P. H., and Holland, R. J. (2005). Extreme spatial variability in marine picoplankton and its consequences for interpreting Eulerian time-series. *Biol. Lett.* 1, 366–369. doi: 10.1098/rsbl.2005.0316

Martin, A. P., Zubkov, M. V., Fasham, M. J., Burkill, P. H., and Holland, R. J. (2008). Microbial spatial variability: an example from the Celtic Sea. *Prog. Oceanogr.* 76, 443–465. doi: 10.1016/j.pocean.2008.01.004

Martin, A. P., Zubkov, M. V., Holland, R. J., Tarran, G., and Burkill, P. (2010). Variability in ultraplankton at the porcupine abyssal plain study site. Deep. *Res. II Top. Stud. Oceanogr.* 57, 1336–1345. doi: 10.1016/j.dsr2.2010.01.010

Moran, A. G., Alonso-sa, L., Nogueira, E., Ducklow, H. W., Gonza, N., Calvo-dı, A., et al. (2015). *More, Smaller Bacteria in Response to Ocean's Warming?* London: Royal Society Publishing.

Olson, R. J., and Sosik, H. M. (2007). A submersible imaging-in-flow instrument to analyze nano-and microplankton: imaging FlowCytobot. *Limnol. Oceanogr. Methods* 5, 195–203. doi: 10.4319/lom.2007.5.195

Pedrós-Alió, C., Calderón-Paz, J. I., Guixa-Boixereu, N., Estrada, M., and Gasol, J. M. (1999). Bacterioplankton and phytoplankton biomass and production during summer stratification in the northwestern Mediterranean Sea. *Deep Res. I Oceanogr. Res. Pap.* 46, 985–1019. doi: 10.1016/S0967-0637(98)00106-X

Pommier, T., Neal, P. R., Gasol, J. M., Coll, M., Acinas, S. G., and Pedrós-Alió, C. (2010). Spatial patterns of bacterial richness and evenness in the NW Mediterranean Sea explored by pyrosequencing of the 16S rRNA. *Aquat. Microb. Ecol.* 61, 221–233. doi: 10.3354/ame01484

Props, R., Monsieurs, P., Mysara, M., Clement, L., and Boon, N. (2016). Measuring the biodiversity of microbial communities by flow cytometry. *Methods Ecol. Evol.* 7, 1376–1385. doi: 10.1111/2041-210X.12607

Rubbens, P., and Props, R. (2021). Computational analysis of microbial flow cytometry data. *mSystems* 6:20. doi: 10.1128/msystems.00895-20

Russo, S., Besmer, M. D., Blumensaat, F., Bouffard, D., Disch, A., Hammes, F., et al. (2021). The value of human data annotation for machine learning based anomaly detection in environmental systems. *Water Res.* 206:117695. doi: 10.1016/j.watres.2021.117695

Swalwell, J. E., Ribalet, F., and Armbrust, E. V. (2011). Seaflow: a novel underway flow-cytometer for continuous observations of phytoplankton in the ocean. *Limnol. Oceanogr. Methods* 9, 466–477. doi: 10.4319/lom.2011.9.466

Thyssen, M., Beker, B., Ediger, D., Yilmaz, D., Garcia, N., and Denis, M. (2011). Phytoplankton distribution during two contrasted summers in a Mediterranean harbour: combining automated submersible flow cytometry with conventional techniques. *Environ. Monit. Assess.* 173, 1–16. doi: 10.1007/s10661-010-1365-z

Thyssen, M., Tarran, G. A., Zubkov, M. V., Holland, R. J., Grégori, G., Burkill, P. H., et al. (2008). The emergence of automated high-frequency flow cytometry: revealing temporal and spatial phytoplankton variability. *J. Plankton Res.* 30, 333–343. doi: 10.1093/plankt/fbn005

Van Gassen, S., Callebaut, B., Van Helden, M. J., Lambrecht, B. N., Demeester, P., Dhaene, T., et al. (2015). FlowSOM: using self-organizing maps for visualization and interpretation of cytometry data. *Cytom. A* 87, 636–645. doi: 10.1002/cyto.a.22625

frontiers | Frontiers in Microbiology

TYPE Original Research
PUBLISHED 02 March 2023
DOI 10.3389/fmicb.2023.1071379



OPEN ACCESS

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

REVIEWED BY
Jean-François Mangot,
Spanish National Research Council (CSIC),
Spain
Owen S. Wangensteen,
University of Barcelona,
Spain

*CORRESPONDENCE
Timotej Turk Dermastia
✉ timotej.turkdermastia@nib.si

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

RECEIVED 16 October 2022
ACCEPTED 10 February 2023
PUBLISHED 02 March 2023

CITATION
Turk Dermastia T, Vascotto I, Francé J, Stanković D and Mozetič P (2023) Evaluation of the *rbcL* marker for metabarcoding of marine diatoms and inference of population structure of selected genera.
*Front. Microbiol.* 14:1071379.
doi: 10.3389/fmicb.2023.1071379



# Evaluation of the *rbcL* marker for metabarcoding of marine diatoms and inference of population structure of selected genera

Timotej Turk Dermastia[1,2]*, Ivano Vascotto[1,2], Janja Francé[1], David Stanković[3] and Patricija Mozetič[1]

[1]Marine Biology Station Piran, National Institute of Biology, Piran, Slovenia, [2]Jožef Stefan International Postgraduate School, Ljubljana, Slovenia, [3]Department of Organisms and Ecosystems Research, National Institute of Biology, Ljubljana, Slovenia

Diatoms are one of the most important phytoplankton groups in the world's oceans. There are responsible for up to 40% of the photosynthetic activity in the Ocean, and they play an important role in the silicon and carbon cycles by decoupling carbon from atmospheric interactions through sinking and export. These processes are strongly influenced by the taxonomic composition of diatom assemblages. Traditionally, these have been assessed using microscopy, which in some cases is not reliable or reproducible. Next-generation sequencing enabled us to study diversity in a high-throughput manner and uncover new distribution patterns and diversity. However, phylogenetic markers used for this purpose, such as various 18S rDNA regions, are often insufficient because they cannot distinguish between some taxa. In this work, we demonstrate the performance of the chloroplast-encoded *rbcL* marker for metabarcoding marine diatoms compared to microscopy and 18S-V9 metabarcoding using a series of monthly samples from the Gulf of Trieste (GoT), northern Adriatic Sea. We demonstrate that *rbcL* is able to detect more taxa compared to 18S-V9 metabarcoding or microscopy, while the overall structure of the diatom assemblage was comparable to the other two methods with some variations, that were taxon dependent. In total, 6 new genera and 22 new diatom species for the study region were identified. We were able to spot misidentification of genera obtained with microscopy such as *Pseudo-nitzschia galaxiae*, which was mistaken for *Cylindrotheca closterium*, as well as genera that were completely overlooked, such as *Minidiscus* and several genera from the Cymatosiraceae family. Furthermore, on the example of two well-studied genera in the region, namely *Chaetoceros* and particularly *Pseudo-nitzschia*, we show how the *rbcL* method can be used to infer even deeper phylogenetic and ecologically significant differences at the species population level. Despite a very thorough community analysis obtained by *rbcL* the incompleteness of reference databases was still evident, and we shed light on possible improvements. Our work has further implications for studies dealing with taxa distribution and population structure, as well as carbon and silica flux models and networks.

KEYWORDS

***rbcL*, metabarcoding, monitoring, diatoms, population genetics, *Pseudo-nitzschia*, Adriatic**

## 1. Introduction

Diatoms belong to the phylum Ochrophyta and are an obligate autotrophic group. According to some estimates, they are responsible for up to 40% of the primary productivity of the ocean (Nelson et al., 1995). They are a highly ecologically successful and diverse group that also plays a key role in the biogenic silica cycle by forming silicate frustules that protect their cells. They also act as ballast, contributing significantly to carbon export and the biological carbon pump, especially in the oligotrophic stratified ocean (Nelson et al., 1995). This is also directly influenced by their diversity, as diatoms vary in the size, shape, and thickness of their frustules (Tréguer et al., 2018 and references therein). These frustules are also the key morphological features for species identification. The occupation of different ecological niches, from oceanic, coastal, benthic and epiphytic environments, has led to a wide distribution of this group (Round et al., 1990). However, the actual number of species is very difficult to estimate because cultivation and detailed morphological or genetic analysis are often required to identify species. New groups and species are constantly being discovered through the elucidation of phylogenetic relationships among different groups, while high-throughput genetic techniques and advanced imaging techniques that help decipher differences among species and reveal diversity have only recently begun to be applied on a large scale. Phytoplankton monitoring, which helps to detect changes in phytoplankton communities and establish long-term ecological changes and processes, utilizes so-called long-term ecological research (LTER) sites (Edwards et al., 2010) where a wide range of data, including phytoplankton community structure, is collected over long periods of time. The role of such sites has been shown to be immensely important, although they are often underfunded and/or neglected as "routine monitoring" (Zingone et al., 2019). Phytoplankton monitoring has traditionally relied on phytoplankton counts, largely based on the Utermöhl method (Utermöhl, 1958). However, in recent years, many LTER sites, including those in the Mediterranean, have incorporated different molecular methods into their programs, the most powerful of which is environmental DNA (eDNA) analysis using metabarcoding (Piredda et al., 2017; Stern et al., 2018; Armeli Minicante et al., 2020). Metabarcoding has gained appeal as it offers several advantages for natural community analysis compared to traditional methods (De Vargas et al., 2015; Malviya et al., 2016; Penna et al., 2017; Piredda et al., 2018; Cristescu, 2019). These include high throughput; scalability, i.e., large-scale sampling campaigns and analyzes are possible and comparable to small-scale campaigns; interoperability, i.e., datasets can be compiled based on the same bioinformatics pipelines and are thus comparable, while data obtained through counts or observations are always observer-dependent; detection of rare and cryptic taxa; detection of non-indigenous species. On the other hand, eDNA analysis brings its own problems. First, it depends on reference databases, which are usually not complete, since most marine organisms have not yet been cultured (Weigand et al., 2019). However, even those that have been cultured and barcoded may have considerable genetic diversity, leading to conceptual and technical problems in assigning species or even higher taxonomic levels. Most studies focus on the V4 or V9 regions of the universal 18S eukaryotic domain because it allows comparison across different taxonomic levels and eukaryotic groups (De Vargas et al., 2015; Piredda et al., 2017). However, 18S is not a very informative marker and species-level resolution in case of diatoms is often difficult to achieve even when the entire gene is considered (Moniz and Kaczmarska, 2009; Guo et al., 2015). In addition, the definition of molecular species is problematic since usually no ecological and morphological reference can be defined, especially with environmental samples. To overcome this problem, the operational taxonomic unit (OTU) concept was introduced (Caron et al., 2009), which enables clustering of similar sequences to retrieve relative species-like resolution, while accounting for sequencing errors. However, this method inevitably leads to loss of information obtained by sequencing, while it does not account for sequences that are potentially shared among different species that get clustered within the same OTU. To prevent information loss, error corrected sequences–using algorithms such as *dada2* (Callahan et al., 2016)–can be clustered at 100% identity to produce amplicon sequence variants (ASVs). Conversely, in order to increase the resolution for species detection, other more variable markers can be used, but so far we have seen limited application in the field of marine phytoplankton monitoring. For example, the *rbcL* marker has previously been used in metabarcoding of freshwater diatoms (Vasselon et al., 2017b) and in reconstructing clone libraries of chromophytic phytoplankton in mangrove forests (Samanta and Bhadury, 2014). The resolution of the 312 bp barcoding region (Rimet et al., 2016) has been demonstrated on the example of the marine diatom *Pseudo-nitzschia* (Turk Dermastia et al., 2020), while different regions of the marker have been evaluated several times for barcoding suitability of diatoms (Daugbjerg and Andersen, 1997; Watson and Tabita, 2006; Guo et al., 2015). They all showed that the marker is more suitable for barcoding compared to 18S, especially for smaller fragments. On the other hand, it may lack resolution compared to the more diverse COI and ITS-2 markers, although these two often exhibit difficulties in amplification and alignment (Moniz and Kaczmarska, 2009). Based on these findings, *rbcL* appears to be an ideal candidate for marine diatom metabarcoding.

In the Gulf of Trieste, the northernmost part of the Adriatic Sea, phytoplankton have been consistently monitored using light microscopy for almost 40 years (Cabrini et al., 2012; Mozetič et al., 2012; Cerino et al., 2019). This long period is also characterized by a significant decline in phytoplankton biomass observed throughout the northern Adriatic basin over the last two decades (Mozetič et al., 2012; Brush et al., 2021), largely due to phosphorus limitation exacerbated during the drought in major rivers (Mozetič et al., 2010; Brush et al., 2021). The observed regime shift is clearly reflected in the changes in the main phytoplankton groups and in species diversity, including diatoms, which account for the largest proportion of the total biomass (Vascotto et al., 2021). Thus, the typical pattern of diatom assemblage in recent times is characterized by two seasonal peaks in summer and fall (Brush et al., 2021). The typical spring bloom (February–March), which was predominant in earlier decades (Cabrini et al., 2012), decreased and developed in later months (May–June), but with lower abundances and different species. For example, the diatom *Skeletonema marinoi* bloomed in late winter and early spring until 2000 and was in later years replaced by blooms of smaller diatoms of the genus *Chaetoceros*, which may be more efficient in an oligotrophic environment (Cabrini et al., 2012). In addition, diversity can also be disrupted by the introduction of non-indigenous species (NIS), which in some cases can even be harmful. One such NIS, *Pseudo-nitzschia multistriata*, was recently found in Adriatic ports (Mozetič et al., 2019). To follow such changes in more detail, a deeper,

taxonomist-independent, high-throughput analysis of the phytoplankton community at this stage is highly desired and necessary.

In this work, we evaluated the *rbcL* plastid marker to describe diatom assemblages in a half year-long study and compared it to 18S-V9 metabarcoding and light microscopy. The study is the first of its kind in the area and demonstrates the power of molecular monitoring strategies as previously unknown community members were discovered. We have also evaluated the *rbcL* marker as a tool for analyzing population structure by analyzing haplotype composition in diatoms of the genera *Pseudo-nitzschia* and *Chaetoceros* and comparing the resulting composition to known haplotype structure (Turk Dermastia et al., 2020).

## 2. Methods

### 2.1. Sampling campaign

Samples were collected from September 2019 to February 2020 and an additional one in October 2020 (October-20) at the 00BF station (LTER-SI; Figure 1). A total of 36 samples were collected. Seawater was collected at 0 m and 5 m using 6-liter Niskin bottles. For metabarcoding, 1 l of seawater from both depths was filtered in triplicate on 0.8 µm polycarbonate filters without prefiltration. The filters were frozen and stored at −80°C. Simultaneous phytoplankton counts using the Utermöhl technique (Utermöhl, 1958) were performed as part of ongoing monitoring activities.

Sampling in October-20 was conducted using a phytoplankton net by towing the net horizontally five times from the depth of 5 m. The collected sample was subsequently filtered the same way as described above.

Alongside sampling, several environmental parameters were measured. CTD temperature and salinity profiles were obtained with a SBE 19plus SEACAT multiparametric probe. Discrete seawater samples were collected with 6-liter Niskin bottles at two depths (0.5, 5). Dissolved inorganic nutrient concentrations were determined colorimetrically on filtered samples with a QuAAtro (Seal Analytical), according to Hansen and Koroleff (1999).

### 2.2. Extraction of DNA for metabarcoding analysis

DNA from environmental samples destined for metabarcoding was extracted in two ways, as samples were filtered in triplicate. After

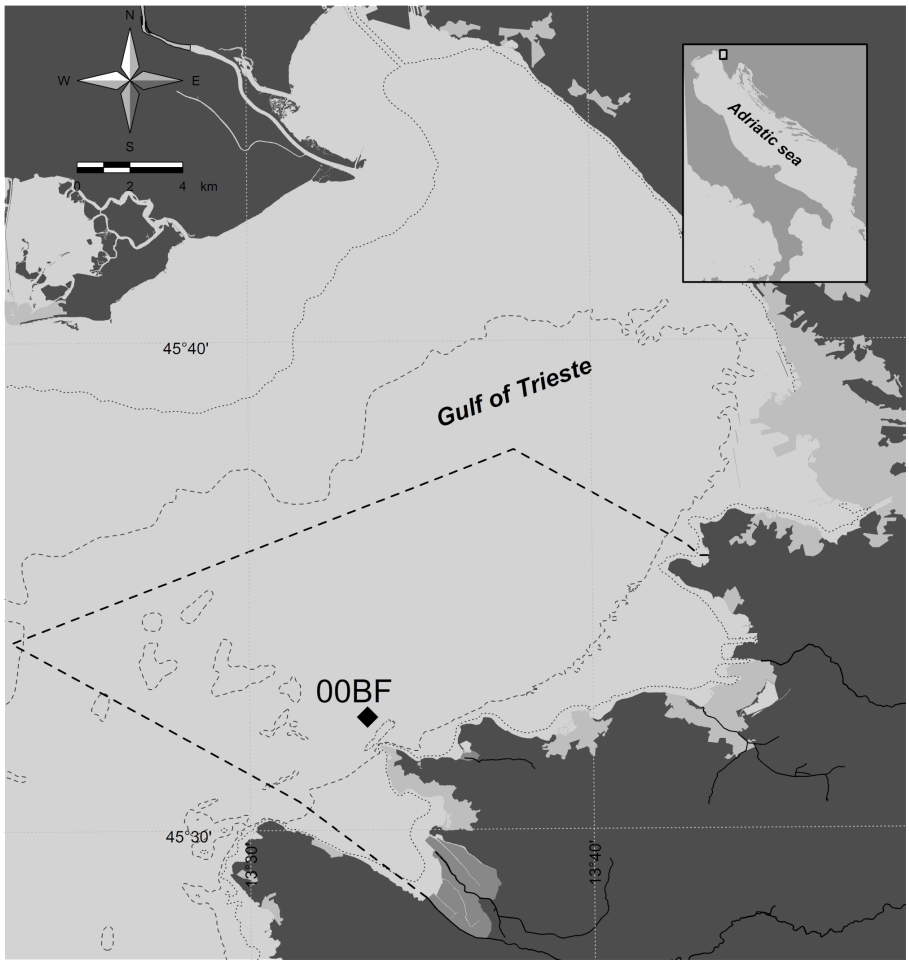



FIGURE 1
Sampling location. Dashed line represents the border of the Slovenian territorial sea.

a filter crushing step, in which the tubes containing the filters were immersed in liquid nitrogen and crushed into small pieces using a sterile metal spatula, one replicate was extracted using the E.Z.N.A Mollusc DNA Kit according to the manufacturer's guidelines. The other two were extracted using the phenol-chloroform extraction procedure described in the protocol of Angel (2012) and modified for extraction from filters. These were immersed in liquid nitrogen and crushed with a sterile metal spatula. 1 ml of 120 mM phosphate buffer (pH 8) and 125 µl of TNS buffer (500 mM Tris base, 100 mM NaCl, 10% SDS) were added to the crushed filters. Extraction then proceeded as in the original procedure. DNA concentration and quality of both phenol and E.Z.N.A replicates were measured by Nanodrop spectrophotometer and on an agarose gel (1%) by electrophoresis (60 V, 1 h). The reason for different extraction approaches was to achieve maximal DNA recovery and amplification success, and not for comparative purposes.

## 2.3. Amplification and Illumina MiSeq sequencing

Two markers were chosen for metabarcoding, namely the 150 base pair (bp) 18S-V9 region and a ~ 312 bp barcode within the *rbcL* chloroplast gene. The primers used to amplify the 18S gene were the universal eukaryotic 18S-V9F (TTGTACACACCGCCCGTCGC) and 18S-V9R (CCTTCYGCAGGTTCACCTAC; Piredda et al., 2017). Libraries for the *rbcL* barcode were built using the diatom-specific primers 708F-DEG (CCTTCYGCAGGTTCACCTAC) and R3-DEG (CCTTCTAATTTACCWACWACWG), both modified from Vasselon et al. (2017a). All primers were modified according to the Illumina protocol by adding universal Illumina tails. The initial amplification step using primers with Illumina adapters was performed by the authors. Samples that failed to amplify were not further considered, while those with visible amplification products were sent to BMR Genomics srl (Padua, Italy) where further library preparation and sequencing were performed. Sequencing libraries for the entire series of samples were generated with 18S, whereas we were unable to obtain sequences from October with *rbcL*. The list of samples and associated metadata can be found in Supplementary Table S1. 18S-V9 amplicons were sequenced with 2 x 150 bp MiSeq reagent kits, while *rbcL* amplicons were sequenced using 2 x 300 bp kits at 300 and 600 cycles, respectively. Both runs produced roughly 1.5 M reads, but the 18S run was sequenced together with samples not reported in this study which resulted in a lower average number of reads per sample.

## 2.4. Bioinformatics analysis

The sequences provided by BMR Genomics had already been demultiplexed with primers removed. For inference of diatom assemblage data, sequence denoising and taxonomy assignment were performed using the *dada2* software package (Callahan et al., 2016) embedded in the R software framework (R Core Team, 2019). Reads were filtered and pruned based on quality scores. The quality of reads was good, except at the beginning of the reads. Error rates were assumed using the implemented *dada2* algorithm and applied to the sample inference step where ASVs and their relative abundance are derived. For both markers, 13 bases of the 5′ ends were trimmed while the length was eventually truncated to 150 bp with the first run and to 300 bp with the second run for *rbcL*. The maxEE score was derived with Figaro software (Weinstein et al., 2019) and set to 2 for both reverse (RR) and forward reads (FR). Reads with ambiguous nucleotides were excluded. FR and RR were then merged. The sequence table obtained was cleaned of chimeric sequences and singletons. To further decrease noise we applied the LULU curation algorithm (Frøslev et al., 2017) for inferring erroneous ASVs based on co-occurrence and similarity. The minimum threshold of sequence similarity for considering any ASV as an error of another was set to 97% but the result was the same even with default settings (84%). All other parameters were left as default. Since *rbcL* is a coding gene, ASVs were manually analyzed for the presence of stop codons in translation. Sequences containing stop codons were removed. Codon entropy ratios between the entropy of position 2 and position 3 were calculated according to Turon et al. (2020) to determine the influence of the curation process. Sequences from two different depths were for the purpose of this study pooled by averaging to represent the monthly surface diatom assemblages, except in certain cases where fine-scale resolution was of interest.

For taxonomy assignment, two approaches were followed, both dealing with ASVs. The first was the naïve Bayesian classifier (NBC) with 50% bootstrap thresholds (NBC50) for classifying any given taxonomy rank implemented in the *dada2* package. 80% thresholds were also tested (NBC80) but are not reported. For the 18S marker, we used the PR2 database, version 4.12 (Guillou et al., 2013), while for *rbcL* we used the Rsyst::diatom database, version 9 (Rimet et al., 2016), which we modified by adding local taxa sequences obtained during this and previous works (Turk Dermastia et al., 2020, 2022), as well as other underrepresented marine diatom taxa, increasing the number of diatom taxa in the database to 1,454. The database used can be accessed through Zenodo (https://doi.org/10.5281/zenodo.7064747). An alternative taxonomy was assigned using local BLAST implemented in the software MALT (Herbig et al., 2016) with a cutoff value of 97% and a maximum E value of -40E10 and by saving the top 10 hits, using the top hit for visualization. Reference databases for 18S and *rbcL* BLAST were obtained from GenBank using the search terms available in the supplemental data (Code Piece 1). The resulting assignments were visualized in MEGAN (Huson et al., 2007, RRID:SCR_011942). 18S data was first cleared of metazoan sequences by applying filtration methods implemented in the R package *phyloseq* (McMurdie and Holmes, 2013). This data was used to visualize the protist assemblages, before proceeding with the analysis of diatoms. Diatom species accumulation curves were obtained with the *vegan* package (Oksanen et al., 2019) with the *rarecurve* function on species-agglomerated data, since ASVs are not directly comparable between different phylogenetic markers. Visualization, filtration and agglomeration of abundance data was performed using *phyloseq*, together with *ggplot2* (Wickham, 2016).

To infer the haplotype network of *Pseudo-nitzschia* we selected ASVs assigned to the *Pseudo-nitzschia* genus with BLAST. Haplotype networks were constructed using the *pegas* package in R (Paradis, 2018), and phylogenetic trees and heat maps were constructed and drawn using the *ape* (Paradis and Schliep, 2019, RRID:SCR_017343) and *ggtree* (Yu et al., 2016, RRID:SCR_018560) packages.

## 2.5. Comparative analysis of diversity

Since the 18S marker recovered the entire eukaryote community, we have for the purpose of this study filtered out all metazoan taxa at the first stage. These data from 18S served as an overview of the structure of the phytoplankton community and as a comparison to microscopy-derived structure. Two data normalization approaches were conducted. The first was the $\chi^2$ transformation implemented in the *decostand* function in *vegan*. The second, followed the normalization procedures described in Gloor et al. (2017) for treating high-throughput sequencing data as compositional based on the centered log-ratio (CLR) transformation (Aitchison, 1982). Here zeros were replaced by pseudocounts prior to the transformation but were then back-traced and replaced by zeros again.

The progression of diatom abundance was inspected based on the CLR values of the complete ASV datasets. α-diversity estimates including confidence intervals were calculated following the Aitchison log-ratio model for compositional data (Aitchison, 1982) using the *divnet* (Willis and Martin, 2022) and *breakaway* packages (Willis and Bunge, 2015). Richness, Shannon Diversity and Simpson's Diversity were estimated for a series of methods and taxonomical approach pairs. The differences between different methods and approaches were tested using the TukeyHSD test. Based on these and previous results we selected methods and taxonomical approaches that were the most robust for each barcoding marker. These were then used in a β-diversity analysis and similarity assessment. For this process we used the R package *CoDaSeq* (Gloor et al., 2016) to perform the CLR transformation, followed by the calculation of expected CLR values with the *ALDEx2* package (Fernandes et al., 2014). A principal component analysis (PCA) was conducted on these data, where environmental variables were also fitted to the ordination using the *envfit* function implemented in *vegan* (Oksanen et al., 2019). Significance of differences was tested using the *anosim* function in the *vegan* package. Although the CLR transformation is stable when subsetting abundance data, we report the sequence of data filtering prior to transformation. First, data were agglomerated by depth, then diatom ASVs were selected and agglomerated to the genus level, prior to CLR transformation. Unassigned diatom ASVs were kept in the data. In addition, a correspondence analysis based on the weighted $\chi^2$ distances was also performed (Legendre and Legendre, 2012). Because the resulting ordination demonstrated a specific triangular shape, we used the *decorana* function to perform a detrended correspondence analysis as suggested by Legendre and Legendre (2012). Statistical significance of this relationship was tested using coinertia analysis implemented in the *ade4* package (Dray and Dufour, 2007) and Mantel test for matrix correlation. The comparative analysis was for the most part conducted using genus-agglomerated data, since many species could not be resolved using 18S-V9 and microscopy. The qualitative comparison of the relative abundance of genera was performed on non-normalized data.

# 3. Results

## 3.1. Diversity estimates from 18S and *rbcL* metabarcoding

The average number of reads per sample was about 50,000 ± 9,000 with 18S and 142,000 ± 50,000 with *rbcL* (Supplementary Table S1). 18S reads presented in this study amounted to roughly 600 k, while there were 1.5 M *rbcL* reads. Quality filtering, merging, denoising and chimera removal resulted in greatly reduced dataset of *rbcL* (20–30% retained) whereas with 18S the majority of reads were retained (80–90%). Diatom representation following NBC taxonomy implemented in *dada2* was low in 18S, as more than half of the ASVs belonged to metazoan organisms, while others belonged to other unicellular eukaryotes. Thus in some samples the number of diatom ASVs was less than 1%. With *rbcL*, for which primers targeting diatoms were used, the representation was much higher from 60 to 100%. With 18S irrespective of the taxonomy assignment method, 148 unique diatom ASVs were recovered. With *rbcL* this number depended on the assignment strategy. NBC50 assigned 1,113 ASVs while BLAST assigned 1,021. LULU curation was applied to both taxonomy datasets. Interestingly, the curation did not change the 18S data, whereas the *rbcL* data was greatly reduced following LULU, from the 1,113 ASVs assigned with NBC50 to 731 diatom ASVs. The result of LULU curation was further evaluated by calculating codon position entropies. The original dataset had a position2:position3 ratio of 1.09 suggesting a high error rate, since the second position is expected to be less variable. After LULU, this ratio dropped to 0.25. The species accumulation curves (Supplementary Figure S1) with 18S (NBC50) and *rbcL* LULU curated BLAST data show that with *rbcL* the sequencing depth was sufficient and saturation is achieved with a relatively low library size. On the other hand, with 18S the depth was at least for some samples not high enough. This was particularly evident with September samples. The number of species recovered was higher with *rbcL*.

Although our time series was not long, we can look at the changes in CLR-normalized diatom abundances and diatom classified amplicons by our methods (Figure 2). The inferred progression was similar between 18S and *rbcL* but diverged from that of microscopy. The absolute CLR values were the highest with *rbcL*. Both metabarcoding markers showed an increase in diatom classified amplicons from September to November, followed by a decline in December with *rbcL* and a continued increase with 18S. The momentary decline in December was registered also with microscopy. All three methods concurred in the decline of abundance in February. The higher abundance in October-20 was recorded with all three methods, as well but with the metabarcoding data the CLR value was the highest in the entire series. The peak of abundance with microscopy in October was not as profound with 18S, while we lack this data for *rbcL*.

*rbcL* recovered the highest richness of ASVs, genera and species among all the methods (Figure 3A). Particularly on the ASV level, there were also significant differences between LULU curated and non-currated data and also between taxonomical approaches (Supplementary Figure S2). The differences between *rbcL* data and the other two methods were significant for all tested combinations. On the other hand, microscopy and 18S inferred richness were comparable and non-significantly different. The estimated Simpson's Index (Figure 3B) was similar particularly between metabarcoding data, irrespective of the method, while the estimated index for microscopy data was different. Wherever the index for metabarcoding decreased it seemed to increase for microscopy. Therefore, with metabarcoding the least diverse samples were September and October-20, whereas with microscopy these were the winter months December and January. Nevertheless, these differences were not significant with the TukeyHSD test (Supplementary Figure S2), although the sample

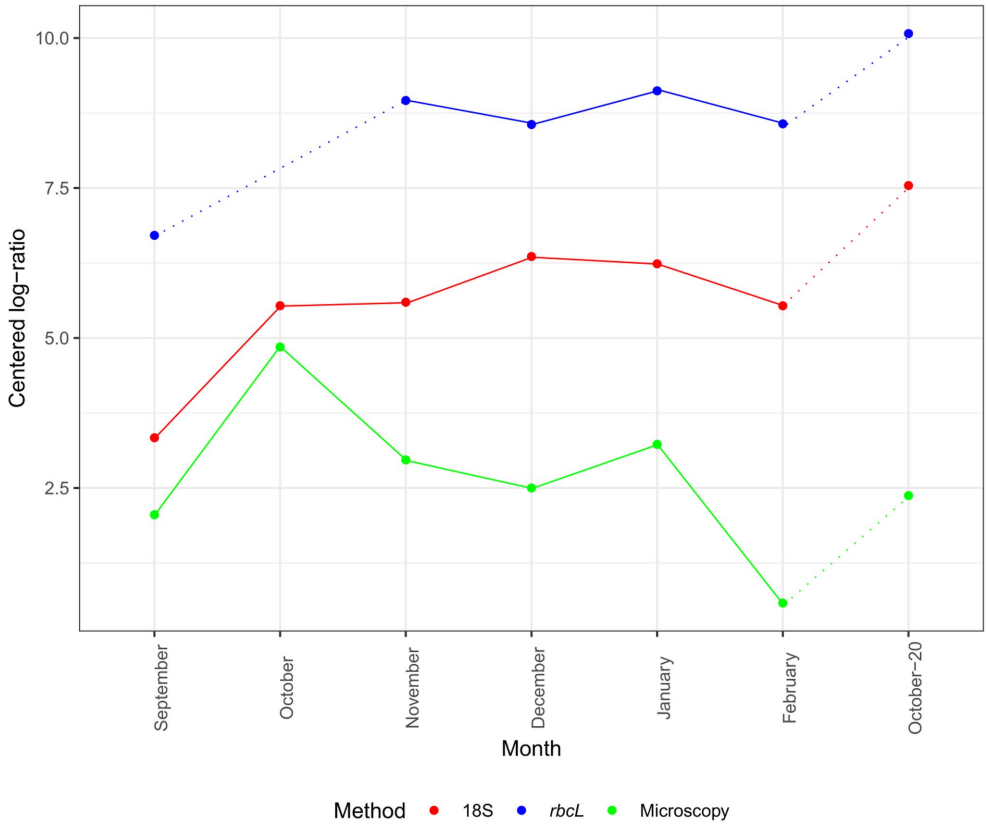



FIGURE 2
Monthly succession of diatom ASV number and cell abundance in the epipelagic layer (pooled depths) at certain months. Data are normalized with the CLR transformation. 18S-V9 data obtained with the NBC50 classifier, *rbcL* data curated with LULU and classified with BLAST. Dotted lines represent time point that are separated by more than 1 month.

number was low. A similar case was observed for the Shannon Index (Figure 3C), although on the ASV level, the differences were significant between 18S and *rbcL* and also between LULU-currated and non-currated *rbcL* data. There were no significant differences on the genus level, while on the species level *rbcL* NBC data was different from 18S NBC. From these results, we established that LULU curated *rbcL* data were more robust and comparable to the data obtained by the other methods and were considered also in the β-diversity analyzes, while the others were not.

## 3.2. Diatom assemblage composition and comparison of methods

The relative abundance of genera differed on the qualitative scale quite substantially between the applied methods (Figure 4). The most obvious differences are between microscopy and each of the metabarcoding markers. The taxonomy of *rbcL* data did not differ largely between BLAST and NBC50 and interestingly, despite the clear change in diversity estimates not with LULU curated data. 18S data with BLAST taxonomy differed from 18S data with NBC50 taxonomy, particularly for a few genera. For example, ASVs classified as *Pseudo-nitzschia* with NBC50 were consistently classified as *Fragilariopsis* with BLAST. When we inspected these ASVs, we realized that they were also classified as *Pseudo-nitzschia* with identical BLAST scores. Given that *rbcL* data also recovered similar percentages of *Pseudo-nitzschia* in these months, we believe that NBC50 classification in the case of 18S was more appropriate. Similarly, BLAST did not identify *Cylindrotheca* as the first hit, but close examination showed that sequences classified as *Bacillaria* had the same BLAST scores and 100% identity with both *Cylindrotheca* and *Bacillaria*. A total of 18 and 26 genera were found using 18S NBC50 and 18S BLAST, respectively; 48, 43 and 42 with *rbcL* NBC50, BLAST and LULU, respectively; and 16 with microscopy. From this, it follows that 18S and microscopy were more comparable, but bear in mind, that sequencing was much deeper and targeting diatoms with *rbcL*.

The proportion of diatom ASVs without genus assignments (Figure 4A) was largely homogenous between the methods. This proportion was especially high in certain months such as September, where it reached almost 80%. Since the various methods, including microscopy, showed similar proportions of unassigned genera and because these proportions seemed unrelated to sequencing depth (e.g., the month of September), we hypothesize this is not a result of incomplete reference databases, but rather represent undescribed diatom diversity. The relative abundances of the most represented genera were comparable at least between *rbcL* and 18S (NBC50). These include *Chaetoceros*, *Minidiscus*, *Pseudo-nitzschia* and *Thalassiosira*. *Minidiscus* represented important relative abundances and is one of the genera that was never found in the studied region before. These are one of the smallest diatoms and are difficult to

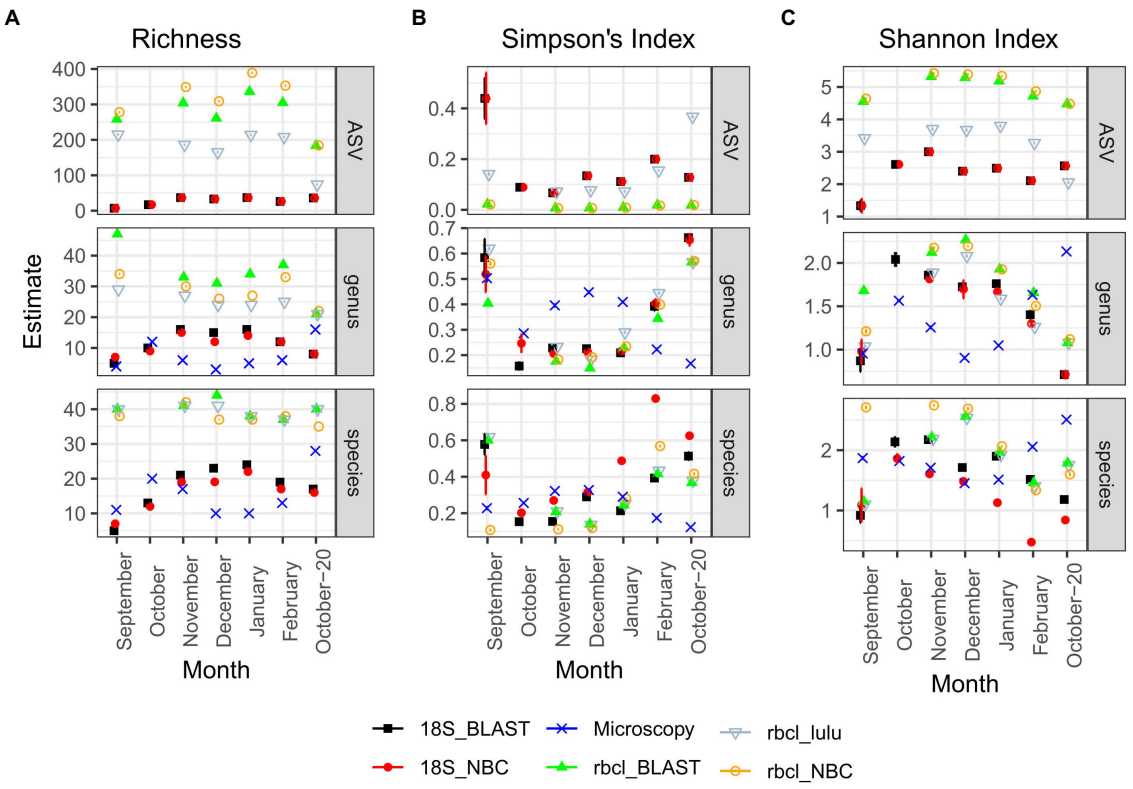



FIGURE 3
Diversity estimates obtained after CLR transformation with *divnet* and *breakaway* for different methods and taxonomy classification approaches. **(A)** Richness; **(B)** Simpson's Index; **(C)** Shannon Index. Error bars are included but are visible only in a few cases. "Genus" and "species" panels represent data agglomerated to the genus or species levels, respectively. Microscopy data is by nature not included in the ASV column.

identify with light microscopy. Particularly high relative abundances of this genus were recorded in December and January. One clear difference between the methods is the large proportion of *Skeletonema* assigned reads with 18S data, but a complete lack of this genus in *rbcL* as well as in microscopy data. Another distinctive feature between the microscopic data and metabarcoding data is the seeming substitution of *Pseudo-nitzschia* in metabarcodes with *Cylindrotheca* in microscopy. This is perhaps not surprising as the majority of *Pseudo-nitzschia* ASVs belonged to *P. galaxie*, which could have been easily misidentified for *C. closterium* in preserved microscopy samples.

September showed a very similar assemblage recovered by microscopy and *rbcL*, where relative abundances of certain genera were also in accordance (*Cyclotella*, *Guinardia*, *Cylindrotheca*, *Nitzschia*, *Rhizosolenia*). Here, 18S underperformed but this was likely also to low diatom read recovery in general. The October samples where we lacked *rbcL* data, showed some similarity between microscopy and 18S but also some notable differences, for example the relatively large proportion of *Guinardia* ASVs recovered with 18S-BLAST. Fall samples showed high diversity with a more even distribution of abundance. In November and December, there was an increase in smaller taxa such as *Thalassiosira* and *Minidiscus* compared to the month of October. *Pseudo-nitzschia* also occupied a larger proportion in the fall, gradually increasing during the winter months. In February, the relative abundance of this genus exceeded 50% of the diatom assemblage, although this still represented only about 10% of the total phytoplankton community (Supplementary Figure S3). *Skeletonema* also occupied significant relative abundances in winter months, but only with 18S data. The most uniform month considering metabarcoding was October-20, for which the samples were obtained differently than in the other months. Here there was a prevalence of *Chaetoceros* and *Pleurosigma* ASVs. *rbcL* also recovered ASVs that belonged to genera also detected by microscopy, namely *Thalassiosira*, *Nitzschia*, *Pseudo-nitzschia* and *Leptocylindrus*. *Bacteriastrum* was detected also by *rbcL* and 18S, and *Thalassionema* with *rbcL*, albeit with very low frequencies therefore they are not visible on the figure.

When we consider the number of times a certain genus was recovered across samples we start to see how related the methods actually are (Figure 5). The most common genera such as *Chaetoceros*, *Pseudo-nitzschia*, *Thalassiosira*, *Cyclotella*, *Minidiscus* had similar representation in the metabarcoding data. *Cyclotella* was found in all samples with microscopy, but it is likely that some cells classified as *Cyclotella* actually belonged to other similar genera such as *Actinoptychus* that was also quite common in the *rbcL* data. On the other hand, the occurrence of *Nitzchia* and *Cylindrotheca* was more similar between microscopy and *rbcL*, since the former was not even found by 18S, while the latter was present only in one sample.

For the multivariate comparison of methods, the data were normalized. We show here results of two different normalization procedures, one based on Aitchison distances (Figure 6A) and the other based on $\chi^2$-distances (Figure 6B). Despite the fact, that some samples appeared to be very similar between methods when the data were not normalized (Figure 4), the differences were larger and

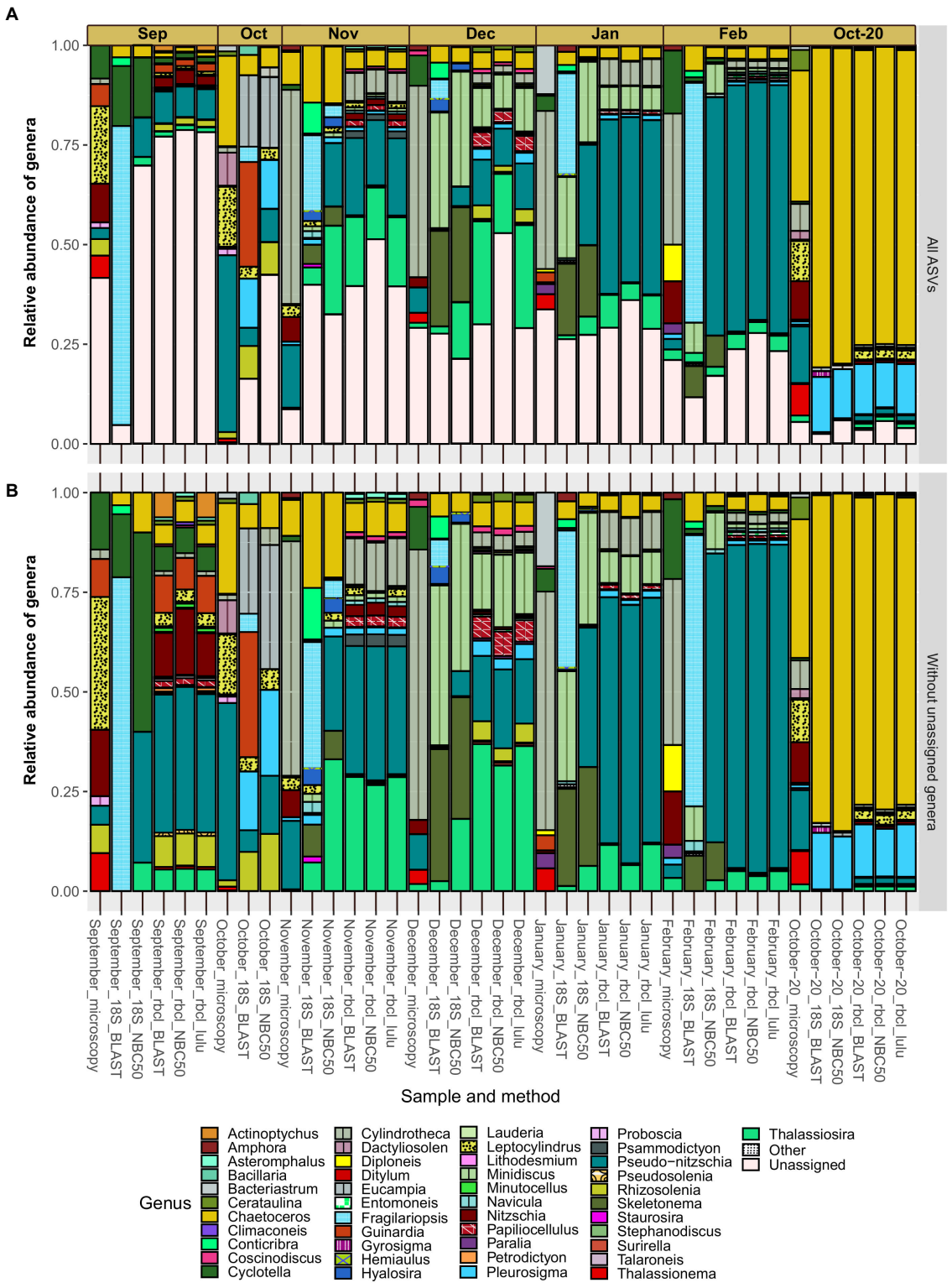



FIGURE 4
Monthly distribution and relative abundances of diatom genera recovered with microscopy, 18S-V9 and *rbcL* using different taxonomy classification approaches. Assemblages with included diatom ASVs with unassigned genera **(A)** and without unassigned genera **(B)**. Genera classified as "Other" had less than 50 reads in any sample and include: *Cocconeis*, *Cyclophora*, *Cymatosira*, *Extubocellulus*, *Gedaniella*, *Grammonema*, *Halamphora*, *Licmophora*, *Meuneria*, *Plagiotropis* and *Tryblionella*.

statistically significant (anosim $R = 0.51$, $p = 0.001$) when the data were normalized with CLR. Even samples that had very similar relative compositions (e.g., October-20) were separated after the transformation. Significant covariates were temperature and nitrite, clearly distinguishing the fall and winter samples. It appears that the winter samples of 18S and *rbcL* were mainly separated by PC1. The

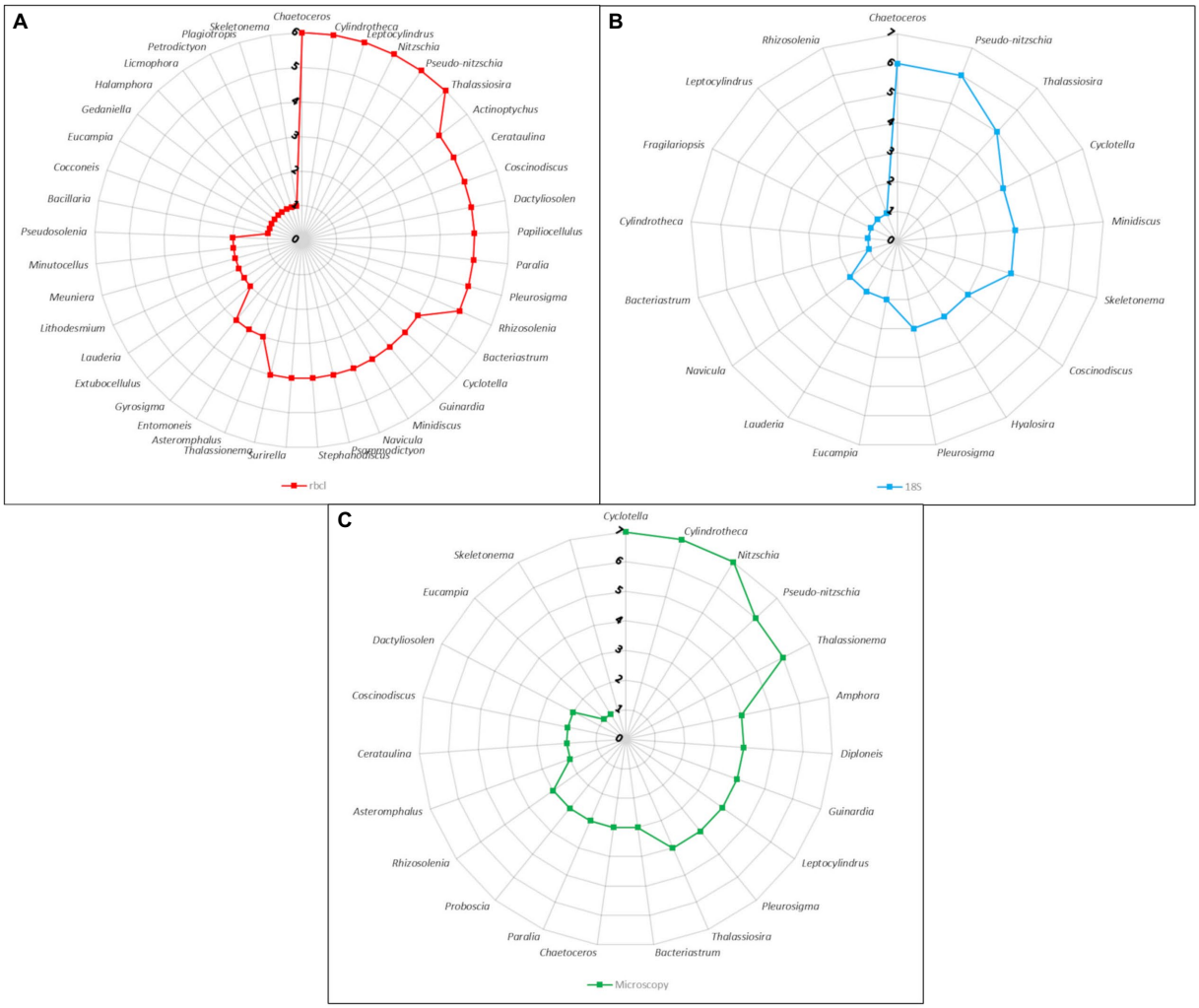



FIGURE 5
Radar charts of the number of samples harboring genera as identified by different methods. **(A)** *rbcL* data with LULU curation and BLAST assignment. **(B)** 18S-V9 data with NBC50 taxonomy assignment. **(C)** Microscopy count data.

highest loadings for this principal component were for *Skeletonema*, *Hyalosira*, *Cylindrotheca*, *Papiliocellulus* and *Cerataulina* (data not shown). Of these genera, only *Skeletonema* and *Cylindrotheca* represented high non-normalized relative abundances, whereas the other three genera were less abundant. *Cerataulina*, *Hyalosira* and *Papiliocellulus* were genera that were uniquely represented by one of the markers. *Hyalosira* only with 18S, and the other two only with *rbcL*. This may point to the fact that the CLR transformation gives more weight to low read abundance taxa that are nevertheless represented in a sample, rather than completely missing.

The ordination stemming from $\chi^2$ normalization resulted in more close clustering of 18S and *rbcL* samples (Figure 6B). Here we present the results of the detrended correspondence analysis (DCA) because the correspondence analysis resulted in a triangular shape of sample scores, which points to high correlation. The structure of the DCA resembled that of Figure 6A with a distribution significantly correlated with temperature and in this case with salinity. Interestingly, samples that appeared to be quite different with the non-normalized data such as September, clustered closely together here. This structure is a consequence of the $\chi^2$ transformation that takes into account not only the change within samples but also individual taxa. It is thus not surprising that genera with the highest loadings in this case were those, that appeared only in one sample in the whole dataset such as *Plagiotropis*, *Halamphora*, *Petrodictyon* and *Extubocellulus*, which were all bunched under "Others" in Figure 4. *Skeletonema* however still had the highest loading with DCA2, similar to the CLR-transformed data. A pairwise comparison of the $\chi^2$ transformed metabarcoding abundance data showed a significant correlation between 18S and *rbcL* data (Mantel statistic based on Pearson's product–moment correlation: $R = 0.64$, $p = 0.05$), different to the analysis of the CLR transformed data. As you may have noticed, the microscopy data was not included in these ordinations. The microscopy cluster formed a completely separate group with the $\chi^2$ transformed data, except for both October samples, while only the September sample clustered with the metabarcoding data in the CLR transformed dataset (Supplementary Figure S4). The most influential taxa separating metabarcoding and microscopy samples with the CLR data in this analysis were *Minidiscus*, *Thalassionema*, *Nitzschia*, *Cylindrotheca*, and *Skeletonema*, therefore those that had very different relative abundances or were missing from the microscopy data completely.

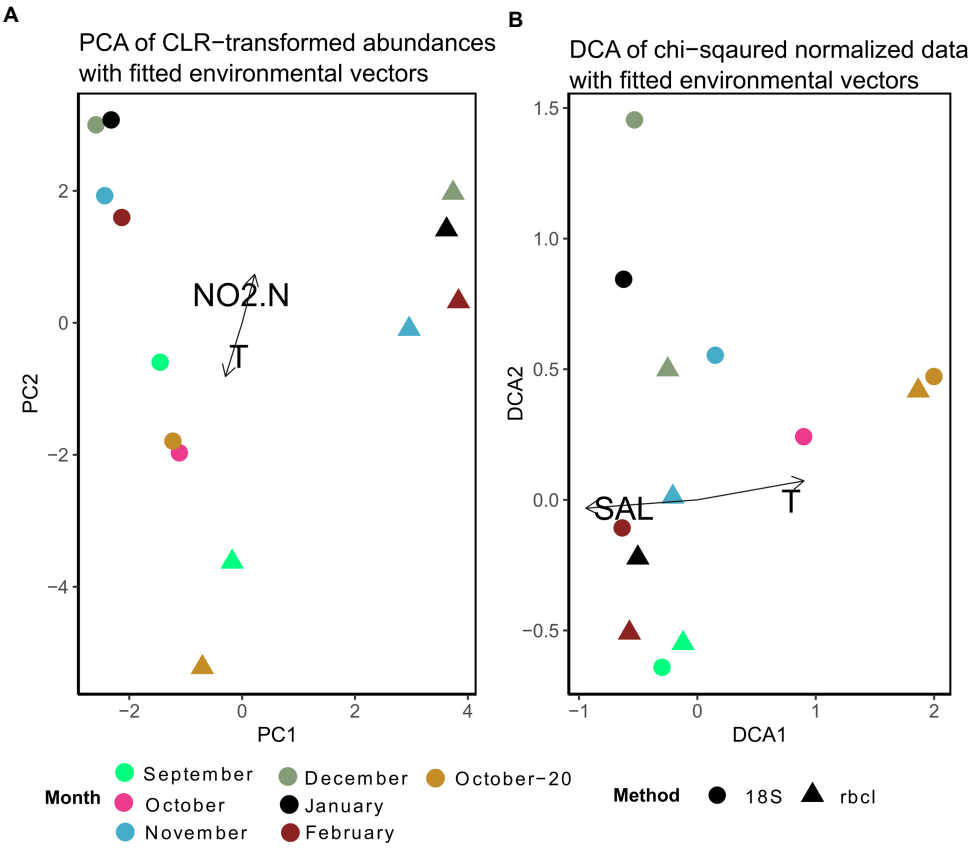



FIGURE 6
**(A)** Principal component analysis of expected CLR data agglomerated to the genus level for 18S and *rbcL*. Fitted are environmental variables with significant correlations ($p<0.05$) as determined by the *envfit* function implemented in *vegan*. **(B)** Detrended correspondence analysis based on $\chi^2$–distances of non-transformed data. Ordinations including microscopy samples are shown in Supplementary Figure S4.

With the $\chi^2$ transformed data, the most influential data in the DCA analysis were different and included *Skeletonema*, *Diploneis*, *Hyalosira*, *Gedaniella*, and *Cocconeis* (data not shown). In addition, no environmental variables were significantly correlated with these ordinations. A pairwise comparison of the $\chi^2$ data did not reveal a significant correlation with the Mantel test (*rbcL*-counts: $R = 0.54$, $p = 0.13$; 18S-counts: $R = 0.4$, $p = 0.09$). This suggested a slightly stronger relationship between 18S and microscopic counts, confirmed by a subsequent coinertia analysis. The inertia was borderline significant when comparing 18S with counts after a Monte Carlo test with 999 replications ($p = 0.056$), whereas the effect was not significant when comparing the *rbcL* and count data ($p = 0.23$).

Apart from *Chaetoceros* and *Pseudo-nitzschia*, genera such as *Bacteriastrum* (2), *Papiliocellulus* (2), *Guinardia* (2), *Rhizosolenia* (2), and *Thalassiosira* (7) were represented by more than one species (Supplementary Table S2) both with 18S and *rbcL*. Many of these taxa, especially from the family Cymatosiraceae, were found for the first time in the region, highlighting the power of *rbcL*. A total of 22 new species were identified for the area with *rbcL*_BLAST, including several new genera (Supplementary Table S2). Not surprisingly, given the ambiguity of some genus assignments, identification of diatom species with 18S-V9 appeared to be very unreliable and many species remained unidentified. Most genera were represented by one or two different species. The ability to discriminate species with 18S-V9 was strongly taxon dependent. For example, with *Pseudo-nitzschia*, a common taxon that forms harmful algal blooms, only one species, *P. delicatissima*, was identified with 18S-NBC, and even here the maximum bootstrap support was 77%. This was also evidenced by BLAST, which resulted in several different species assignments with very similar or identical E-scores for *Pseudo-nitzschia* ASVs. *rbcL* on the other hand, identified nine. Contrarily, all reads assigned to *Cyclotella*, for example, were unambiguously assigned to *Cyclotella choctawhatcheeana* by both markers. Where both markers were comparable in terms of species identification was *Chaetoceros* with multiple species identified. This is also why we chose this genus to compare the species composition with that recovered by *rbcL* (Figure 7). Following the results presented above and the fact that NBC50 was based on the curated PR2 database, i.e., more reliable taxonomy, we chose these results for the comparison, even though of 148 diatom ASVs, BLAST identified 49 as *Chaetoceros*, while NBC50 identified only 35. Of these, 31 were identical between the two methods. In total, 10 species were identified with 18S-V9, while *rbcL* identified 16. Microscopic identification of *Chaetoceros* species was qualitative so we can provide only species lists for comparison (Supplementary Table S3). The composition was somewhat comparable between 18S and *rbcL* especially in the winter months and in October-20, similar to the genus data presented above. In September, 18S recovered only *C. socialis debilis*, which was not found by *rbcL* at all. Microscopy also identified *C. cf. vixvisibilis* and *C. decipiens* that were

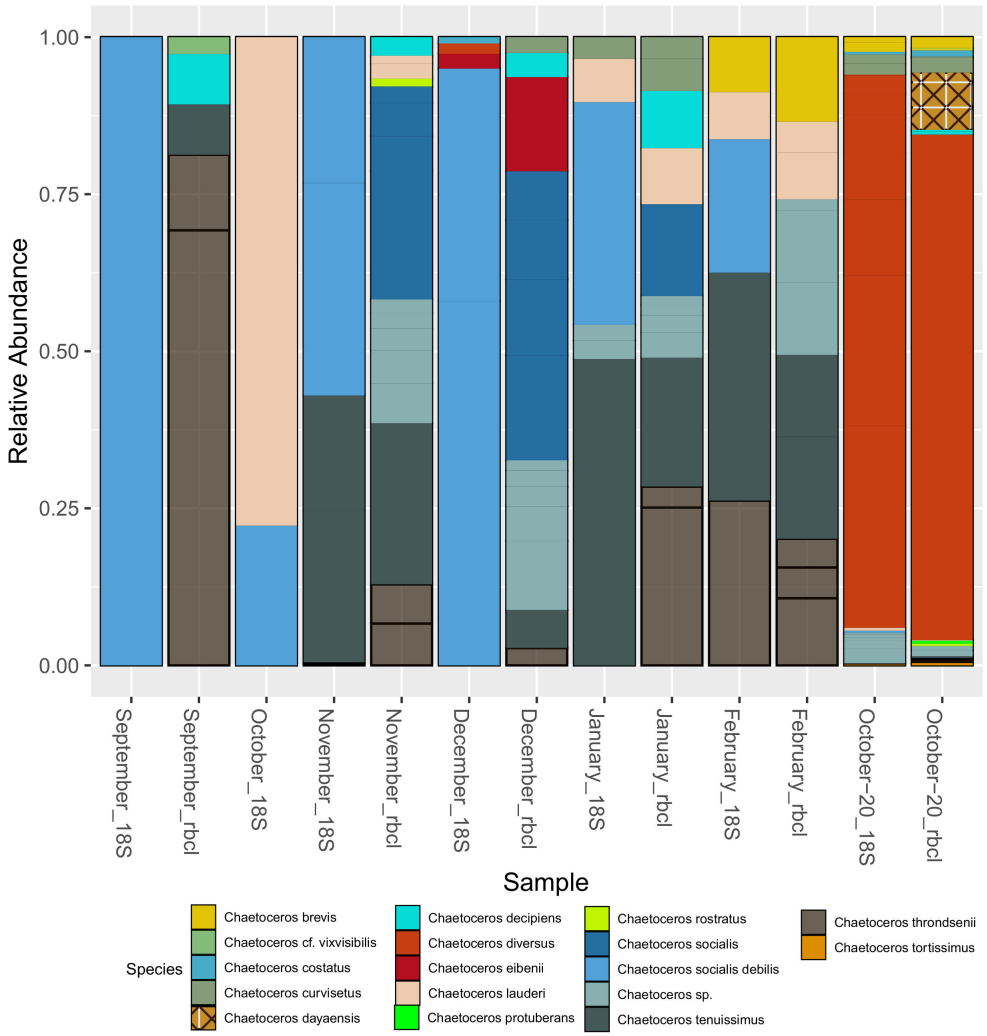



FIGURE 7
Seasonal distribution of *Chaetoceros* species recovered with 18S-V9 (NBC50) and *rbcL* (LULU+BLAST). The unassigned species in January and February with 18S were assigned to *C. cf. wighmani* with BLAST.

identified by *rbcL*, along with several other species. The two species that were found by 18S in October (*C. lauderi* and *C. socialis*) were also found by microscopy, but the list of October species was far greater for microscopy. The November assemblage was more comparable between 18S and *rbcL* with both *C. tenuissimus* and *C. socialis* recovered by both markers, while *rbcL* recovered some additional species. Interestingly, microscopy did not find any *Chaetoceros* in this month. A similar case was observed in December, where only *C. socialis* was found by microscopy, while the metabarcoding markers recovered more species. In January and February, the metabarcoding assemblages were highly similar, while they differed from the assemblage recovered by microscopy. October-20 showed highly similar assemblages, which were similar to those identified by microscopy. Common species included *C. decipiens*, *C. curvisetus*, *C. lauderi*, *C. rostratus*, *C. socialis*, and *C. tortissimus*. *rbcL* additionally identified *C. dayaensis*, a novel species for the Gulf of Trieste. The obvious difference between metabarcoding and microscopy was the high prevalence of *C. diversus* in metabarcoding data, which was not found at all by microscopy. Another novel species for the Slovenian side of the GoT, *Chaetoceros tenuissimus* was present in almost all samples with significant relative abundance. This species could have been identified as *C. simplex* with microscopy and may be missed in counts due to its inconspicuous morphology. It is worth noting, that particularly *rbcL* identified several additional *Chaetoceros* without formal description, including undescribed but commonly referred to strains (e.g., *Chaetoceros* sp. Na28A1 (also with 18S), *Chaetoceros* sp. Na11C3) which were all grouped to *Chaetoceros* sp. in Figure 7. This shows that for *rbcL* the reference database is underpopulated with several species that have not been barcoded yet, but also some that have not been taxonomically resolved.

## 3.3. *Pseudo-nitzschia rbcL* population structure

With *Pseudo-nitzschia*, a heavily studied genus in the area, the NBC50 approach identified 7 species, namely *P. calliantha*, *P. delicatissima*, *P. fraudulenta*, *P. galaxiae*, *P. linea*, *P. mannii*, and

*P. subfraudulenta*, while several species-level reads remained unclassified. Three different morphotypes of *P. galaxiae* were also found, which have been shown to be genetically distinct in previous work (Turk Dermastia et al., 2020). Additional species were identified using BLAST. These included *P. delicatissima* and *P. cf. delicatissima*, which represents a different clade within the *P. delicatissima* complex, while *P. linea* was detected only at the 95% identity threshold. The distribution of *Pseudo-nitzschia* species was month and for some taxa also depth dependent (Figure 8). Here LULU-curated data are shown in order to ensure robust representation of the population structure. The most abundant ASVs of strains that were previously isolated and barcoded in the area were identified by BLAST as almost identical to these strains. This was the case for *P. delicatissima*, *P. fraudulenta*, *P. subfraudulenta*, *P. mannii*, all three morphotypes of *P. galaxiae* and *P. calliantha*. *P. galaxiae* was the most dominant species. In particular, the small morphotype was present throughout the sampling period except in October-20, but with a marked increase in February and January. ASVs belonging to the small morphotype were the most abundant *P. galaxiae* ASVs. ASVs of the large morphotype of *P. galaxiae* had increased

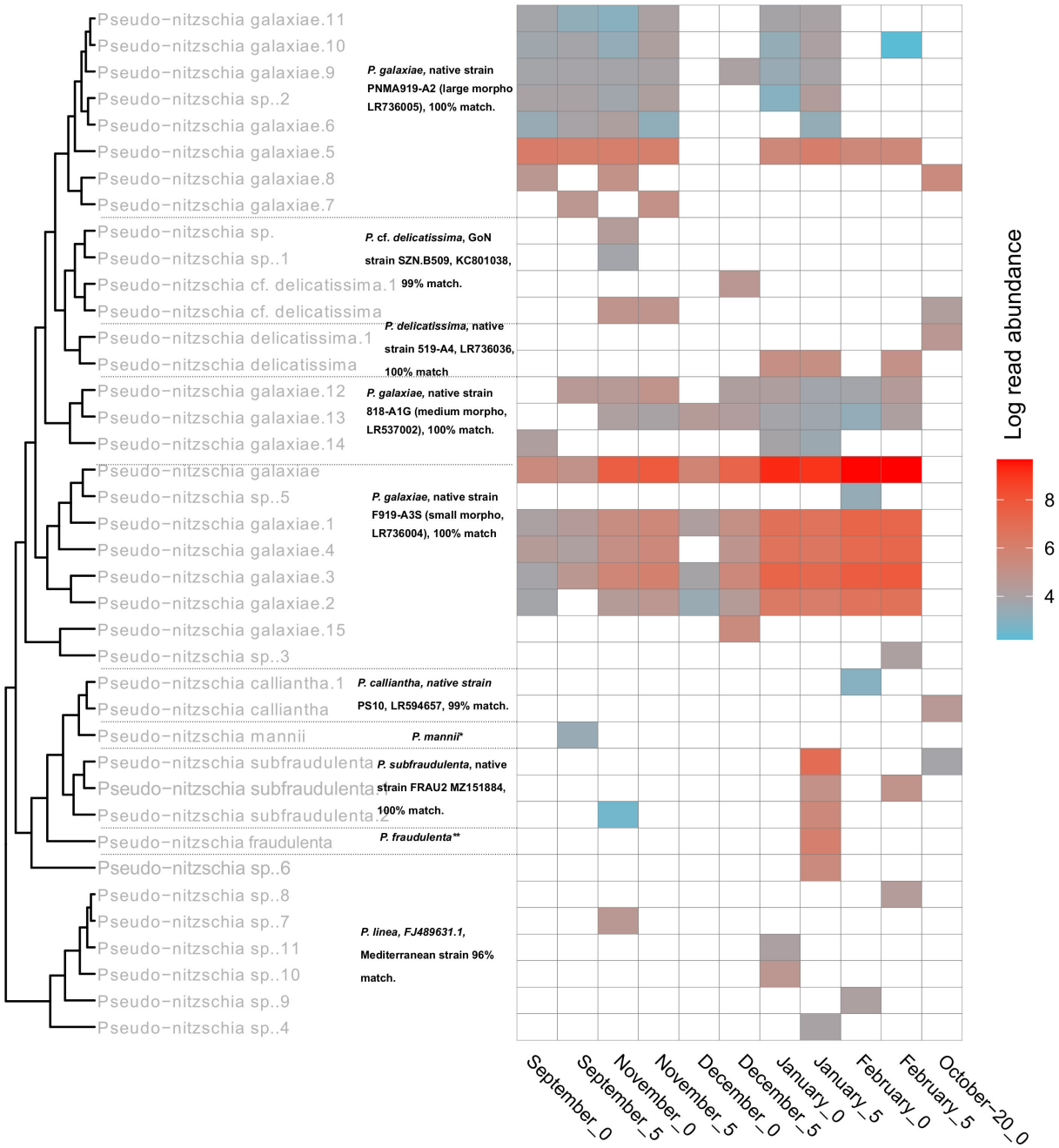



FIGURE 8
Seasonal succession and phylogeny of different *Pseudo-nitzschia* haplotypes inferred from *rbcL*. Tree labels in gray represent classifications following the NBC50 assignment, labels in black represent top hits of BLAST. The asterixes were used to report top hits in the caption for a clearer representation. **P. mannii*, native strain PS9, LR594653. GoN=Gulf of Naples. ***P. fraudulenta*, native strain 217-A2, LR537014.1.

abundance in September and November compared to the winter months. The medium morphotype had the lowest number of ASVs of the three and a peak amplicon abundance in November. Some species, such as *P. fraudulenta*, *P. subfraudulenta*, and *P. mannii*, showed distinct occurrence patterns, with the former two peaking in January at a distinct depth (5 m), while *P. mannii* occurred only in September. *P. delicatissima* ASVs were present in January, February and October-20, while *P. cf. delicatissima* was present in November and December, thus separated from the *P. delicatissima* ASVs. The reference sequence most similar to these reads was strain SZN-B509, which was classified as *P. cf. delicatissima*, a sister group to *P. arenysensis* but ultrastructurally similar to *P. delicatissima* (Lamari et al., 2013). *P. calliantha* appeared in February in low abundance at 0 m and with a higher abundance in October-20. Lastly, several ASVs classified as *P. linea* with 96% identity occurred sporadically throughout the sampling period.

Most assigned species had multiple haplotypes, the number of which correlated with the number of reads for that species (Pearson $r = 0.71$, $p < 0.001$). Most taxa were represented by more than one haplotype, and dominant haplotypes were also recorded, accounting for the majority of reads, especially in January and February 2020, where very dominant haplotypes of *P. galaxiae* and *P. subfraudulenta* occurred. *P. mannii* and *P. fraudulenta* were represented by only one haplotype.

To further clarify the suitability of the *rbcL* marker for population genetic studies, we separately analyzed the haplotypes classified as *P. galaxiae* using haplotype network analysis. We see three distinct haplogroups, each corresponding to different morphotypes of *P. galaxiae* (Figure 9). The haplogroups were seasonally well distributed even though the time series did not even encompass the entire year. The large morphotype was most abundant in late summer and fall, the medium represented a transition to winter, and the small morphotype, which was also represented by most of the haplotypes was most abundant in winter. Nevertheless, most haplotypes co-occurred. Haplogroups were likely reproductively isolated as they either represent several sympatric populations that cannot interbreed due to size differences and are already in the process of speciation, or they may already represent different species. Note that Figure 9 shows the maximum parsimony network in which alternative connections are not shown. This may indicate that haplotypes within the same cluster are more distant than those between clusters, which is not the case as the links between clusters show additional mutational steps from the basal state.

# 4. Discussion

## 4.1. Methodological overview

This work presents a comprehensive analysis of the *rbcL* marker for metabarcoding marine diatoms, not only for diversity assessment,

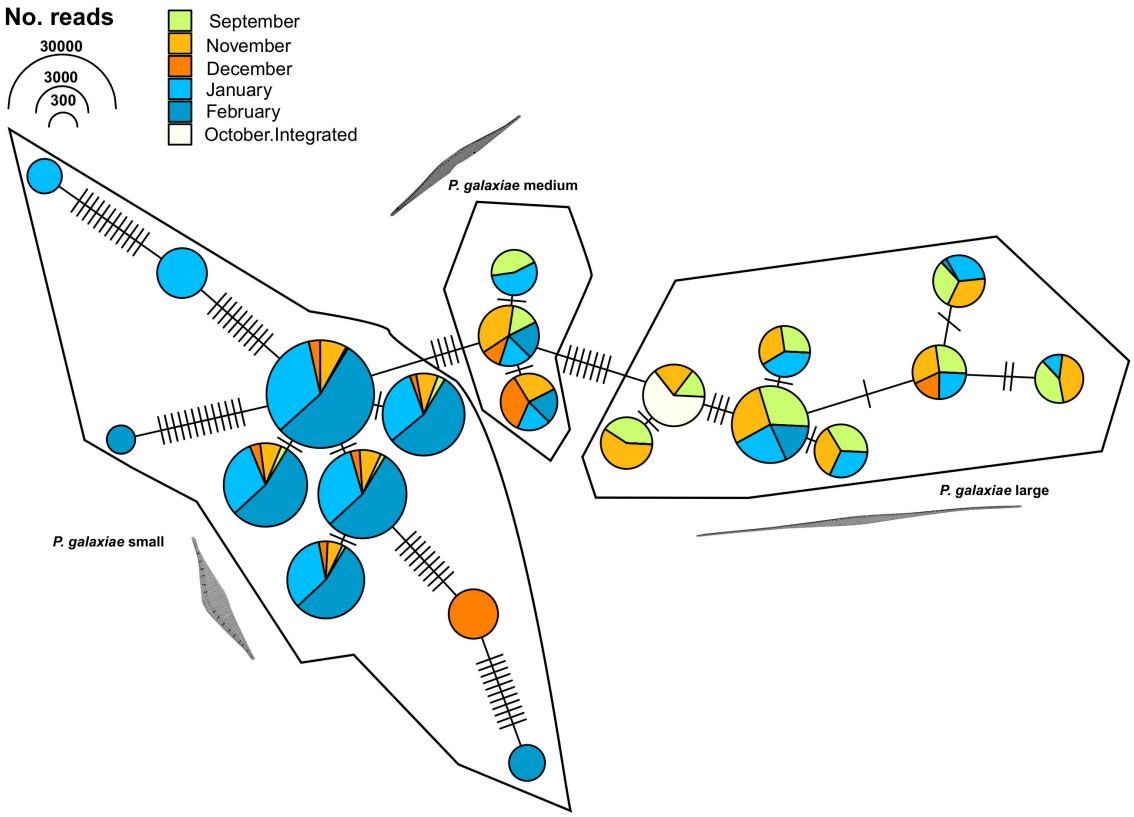



FIGURE 9
TCS maximum parsimony haplotype network for ASVs classified as *Pseudo-nitzschia galaxiae* by *rbcL* (LULU+BLAST). Three different clusters, corresponding to morphological variants of *P. galaxiae* were recognized. Note that alternative network links are not shown, therefore adjacent haplotypes of different clusters appear to be more related than members of the same cluster, while they are in fact additionally separated for the mutational steps depicted between haplogroups.

but also for population genetics and microevolutionary studies. It is also the first metabarcoding study of planktonic protists in the Gulf of Trieste and also one of the first in the Adriatic Sea, apart from the metabarcoding study in the Venice Lagoon (Armeli Minicante et al., 2020). The study, conducted with a half-year series of samples collected monthly, was short but provided insight into the hidden diversity and community patterns unknown in more than 30 years of continuous microscopic monitoring.

The choice of markers in our study, the 18S-V9 region and *rbcL*, was based on previous practice of metabarcoding marine plankton and novelty. 18S-V9 is a well-established eukaryotic marker that has been used in many metabarcoding experiments and campaigns, including the global Tara Oceans cruise and the annual Ocean Sampling Day (De Vargas et al., 2015; Bradley et al., 2016; Malviya et al., 2016; Piredda et al., 2017; Stefanni et al., 2018; Tragin et al., 2018). In this sense, the 18S-V9 data are useful for comparison with other published metabarcoding surveys and the present *rbcL* assessment, while they lack the perhaps desired species-specific robustness. This was also evident from our results, with the 18S marker generally underperforming compared to *rbcL* in both diatom ASVs and recovery of classified diatom taxa. This is likely due to two factors, the first being greater sequencing depth and the use of diatom targeted primers with *rbcL*, exemplified by species accumulation curves (Supplementary Figure S1). The second factor is the fact that the *rbcL* region used in this study is in fact more variable than 18S-V9. However, this observation was taxon dependent. For example, in the genus *Chaetoceros*, species recoveries were similar between the two methods in certain months, although with different relative abundances. On the other hand, the number of *Pseudo-nitzschia* sequences recovered with 18S was very low, probably also due to the incompleteness of the reference database (Piredda et al., 2017). This was illustrated by the two different taxonomic classification approaches, with NBC50 correctly recognizing winter-blooming *Pseudo-nitzschia galaxiae* and classifying reads as *Pseudo-nitzschia* spp. while BLAST assigned these sequences to *Fragilariopsis kerguelensis*. The lack of a reference for *P. galaxiae* 18S-V9 has already been pointed out by Piredda et al. (2017). In this sense, 18S-V4 might be a more appropriate marker, and we acknowledge that the assessment performed in this work would benefit from 18S-V4 data. On the other hand, barcoding effort was far greater with 18S than with *rbcL* and many diatom taxa lack their reference *rbcL* barcodes (Guo et al., 2015). Our study also shows that 18S data resulted in less unassigned ASVs than *rbcL* although the proportions were still quite similar. Furthermore, taxonomic assignment of *rbcL* ASVs with BLAST resulted in less unassigned amplicons compared to the NBC50 approach that used a curated database, with the caveat of the assignments being less reliable.

Traditionally, the phytoplankton community is determined by counting phytoplankton cells under a light microscope. This method is still widely used and the counts have been the gold standard in environmental assessments by government agencies and research institutes for decades. However, the method has its drawbacks, namely dependence on a trained observer, i.e., the taxonomist who counts the cells, inability to distinguish cryptic or inconspicuous taxa, insensitivity to rare taxa, and low sample throughput. The latter may not be important when the number of examined samples is small, but becomes critical when sampling is intensified in time and space. In addition, phytoplankton count data can be valuable in their own right, but they are rarely accompanied by other important cellular characteristics such as biomass that could be more ecologically informative (Godhe et al., 2008; Santi et al., 2021). Thus, it has been shown in recent years that identification and counting of cells under the microscope is not sufficient. Metabarcoding solves all the above problems, but also brings some problems of its own. First, it is PCR-dependent. PCR may preferentially amplify some targets while not amplifying others or amplifying them at a lower rate (Kelly et al., 2019). Second, taxonomic assignment relies on reference databases that do not cover all genetic diversity because most organisms have never been cultured and sequenced (Weigand et al., 2019). Therefore, assignment may be ambiguous or, in some cases, not possible at all, as it is very likely that metabarcoding in a new environment will yield a large number of new sequences. A conceptual problem also arises in translating sequence data into taxonomic richness at the species level. What is a genetic species and up to what point are dissimilar sequences considered a species?

Finally, there are no simple quantification strategies to convert abundance from sequence data into ecologically meaningful abundance, especially given the bias mentioned earlier. Sequencing machines have pre-defined upper thresholds of obtained reads and thus more abundant (or more amplified sequences) can take up a large proportion of the designated sequencing space on chips, artificially skewing the signal (Gloor et al., 2017). Here we approach this by treating the data as compositional, as proposed by Gloor et al. (2017) using the Aitchison CLR transformations. The crucial advantage of the CLR normalization is that it takes into account the exponential differences between more abundant and less abundant taxa that occur after PCR (signal amplification) but also because it does not remove data. This is important for the comparison of two different markers, each of which had a different sequencing depth and priming properties. Furthermore, the sampling was performed in a heterogeneous and fast-changing environment for which the existing diversity is well known from years of biological observation, mainly through microscopy. Therefore, we did not want to omit genera or species recovered with metabarcoding through rarefaction, which is the more standard yet widely disputed (McMurdie and Holmes, 2014) method of normalization. Authors such as Cameron et al. (2021) suggest the limitations of classic rarefaction could be tackled by repetitive rarefaction, however this still preserves the amplification bias. We believe such an approach could be beneficial in less known and understudied systems such as microbiomes, but for comparative purposes with traditional methodology of higher organisms, the benefits of such an approach are less clear. On the other hand, we stress that no normalization procedures is perfect and each has its associated biases. With CLR this is clearly the means of treating zeros, either with pseudocounts such as in our case, or with different imputation methods that estimate the actual values from prior distributions (e.g., Bayesian-multiplicative replacement).

Still, the resulting transformed data is hard to interpret and compare with microscopy, particularly in terms of abundance of different genera. A better way to compare the results of morphological and genetic methods might therefore be data on the biomass or biovolume of different phytoplankton groups (Olenina, 2006; Godhe et al., 2008; Leonilde et al., 2017), although these data are rarely obtained. There is a clear relationship between cell volume and rDNA content that has been demonstrated in dinoflagellates (Prokopowich et al., 2003). They are known to be over-represented

in metabarcoding datasets, with a high proportion of unclassified reads below the phylum level (Piredda et al., 2017; Santi et al., 2021). Dinoflagellate rDNA from the 18S-V9 metabarcoding had previously been analyzed for agreement with true abundances of organisms in mock communities, and the percentages found were quite variable (Guo et al., 2016). The authors found that the actively translated actin gene is a better representative of the community. In our study, the *rbcL* gene was sequenced to enable a different type of identification and quantification strategy, although the relationship between biovolume and *rbcL* copy number is believed to be similar to the relationship between biovolume and 18S copy number– it varies significantly from species to species (Godhe et al., 2008; Vasselon et al., 2018). Vasselon et al. (2018) proposed correction factors based on biovolume that significantly improved abundance estimates of freshwater diatoms. In the case of natural populations this is difficult to apply, since even members of the same species (e.g., *Pseudo-nitzschia galaxiae*) can vary significantly in this characteristic. We thus opted not to apply these correction factors to the *rbcL* data, even though they could improve the correlation with microscopic data. This fine-tuning was not the purpose of this study, but we acknowledge it can be applied if accurate abundance estimates from metabarcoding are sought.

The use of diatom-specific *rbcL* primers has limited us to diatoms, although many "phytoplankton taxa" such as different dinoflagellates do not actually have chloroplasts and therefore a different type of bias would be introduced if universal *rbcL* primers were used. Another problem arising from next-generation sequencing platforms and subsequent bioinformatics analysis is the inflation of operational taxonomic units (OTUs) due to errors in sequencing data generation. OTUs are consensus sequences from multiple reads that differ up to an arbitrary threshold set at 97% in most studies. They are used to infer the taxonomic identity of sequences and to measure the diversity of the communities under study. Alternatively, amplicon sequence variants (ASVs) can be used, such as is in this study so that maximum genetic information was retained. ASVs are sequences that are 100% different and have been obtained from metabarcoding, but have been informatically corrected for errors introduced by the sequencing platform (denoising). Algorithms such as *dada2* are capable of such correction, but the efficiency is still questionable (Nearing et al., 2018). For the most part, sequencing errors do not pose a problem for diversity estimation, as erroneous sequences do not usually occur in large numbers (Frøslev et al., 2017; Turon et al., 2020). However, when intraspecific or haplotypic diversity is desired, these errors can artificially inflate the number of haplotypes and lead to erroneous results. In this study, we have applied the LULU curation algorithm (Frøslev et al., 2017) in order to minimize noise in the data. The noise was drastically decreased as was evidenced by a substantial decrease in ASV number as was in the calculations of the entropy ratios, which could be obtained as *rbcL* is a coding gene. The obtained haplotypes used in the analysis of *Pseudo-nitzschia* population structure are thus robust as they are highly similar to haplotypes of strains previously isolated in the region (Turk Dermastia et al., 2020, 2022). To our knowledge, this is the first study to examine the performance of the *rbcL* marker for metabarcoding of marine diatoms. *rbcL* has already been used for metabarcoding of freshwater diatoms (Vasselon et al., 2017a). Our study yielded comparable results to 18S-V9, but with higher resolution and detail. We acknowledge that using diatom-specific 18S primers or greatly increased sequencing depth with 18S would benefit the final verdict on whether *rbcL* is superior to either region of 18S for diatom metabarcoding.

## 4.2. Diatom assemblages revealed by metabarcoding

The composition of the diatom assemblage determined by 18S-V9 and *rbcL* metabarcoding agreed quite well with long-term patterns in this region described from phytoplankton microscopy counts, while providing new insights into the structure of the assemblage. Diatom blooms in autumn are expected in the northern Adriatic (Mozetič et al., 1998; Cabrini et al., 2012; Mozetič et al., 2012; Cerino et al., 2019). Autumn and early winter, at least in the Italian part of the Gulf of Trieste, are characterized by low diatom abundance and greater diversity (Cabrini et al., 2012; Cerino et al., 2019), while in the Slovenian part autumn blooms are more abundant (Mozetič et al., 2012; Vascotto et al., 2021). In this study, microscopy followed the long-term trends and showed a peak in diatom cells in both October samples. With the normalized metabarcoding data a peak relative to other samples was observed only in October-20 (Figure 2). It is worth noting here that the progression of 18S and *rbcL* was similar, even though *rbcL* detected predominantely diatom reads, whereas 18S detected the entire eukaryotic community. This suggest that the normalization procedure with CLR does a good job at approximating the true composition.

Amplicon sequencing is a suitable approach for comparing diversity estimates with other methods. While *rbcL* recovered much more ASVs, which was also reflected in a greater number of species and genera, the Simpson's diversity and Shannon's diversity indexes were comparable between methods. Even though estimates for individual samples were different, the TukeyHSD test did not show significant differences between the methods and taxonomic approaches. This result is encouraging since it suggests that *rbcL* can be reliably used as an estimator of diatom diversity, while providing a much greater resolution when used as it was in this study.

At first glance, the qualitative comparison of genus and species recovery between metabarcoding and microscopy showed large differences, while the two metabarcoding markers were much more comparable. This was confirmed also with the statistical tests we applied, which also indicated that 18S and microscopy were more correlated. A reason for this could be that due to deeper sequencing and greater variability, *rbcL* recovered more "rare" diversity which is difficult to find also by microscopy, producing differences in datasets. The other reason is hinted in samples where the abundance of diatoms determined with microscopy was low such as in the winter samples. Here the discrepancies were high because *rbcL* amplified the already rare diatom signal which likely skewed the composition. The non-normalized relative abundance of taxa found only by *rbcL* was generally low, but following normalization these differences started to influence the changes between the compositions obtained by *rbcL* and 18S leading to the observed differences in sample clustering (Figure 6). As already hinted in the results, the most influential taxa where in fact those that had either large differences in relative abundances or were entirely absent from one or the other dataset. On the other hand, in certain months such as September, where the number of 18S diatom ASVs was unusually low, microscopy and *rbcL* were more similar. The

only month where all methods really concurred was October-20, especially when we consider the correspondence analysis. We must be cautious with these samples however, since they were obtained with a phytoplankton net of 20 µm mesh size, thus favoring larger diatoms. These samples thus inherently miss the smaller diatom fraction, which could hardly be identified by microscopy, leading to higher similarity. Interestingly, the Simpson's Diversity index obtained for this month was actually the most different from 18S and *rbcL*. We used different normalization procedures (CLR and $\chi^2$) to perform the β-analysis. The CLR method normalized the read numbers sample-wise while the $\chi^2$ method normalized both taxon and sample wise. Technically, the sample wise logarithmic normalization is more appropriate for high-throughput sequencing data, since amplification bias is expressed exponentially and is related to the individual sample (Gloor et al., 2017). Some similar patterns between the two approaches emerged, but the correspondence analysis using $\chi^2$ transformed data resulted in more common structure between 18S and *rbcL*. In both cases, the seasonal clustering related to temperature was evident, while the most influential taxa, except for *Skeletonema* differed. The CLR approach clearly gives a lot of weight to rare but present taxa, whereas the $\chi^2$ is less sensitive to the absence of a particular genus in a certain sample. This is because the expected abundance of a rare taxon from the $\chi^2$ distribution will still be low, whereas in the CLR approach the normalization could increase the importance of rare taxa, accounting for PCR and sequencing bias (Gloor et al., 2016).

Several genera and species were found for the first time in the area, and several of them had significant relative abundances. Examples include *Minidiscus* and certain *Thalassiosira* species. In particular, *Minidiscus* is globally overlooked in phytoplankton counts, while its biomass and abundance potential has been captured in other metabarcoding studies (Leblanc et al., 2018; Arsenieff et al., 2020). Our work adds to these findings, as *Minidiscus* was never recorded by microscopy, although it accounted for up to a quarter of the diatom assemblage in December. Still, this represented small relative numbers within the entire phytoplankton community, as diatom numbers in December were the lowest. In a recent study of Adriatic ports, *Minidiscus* was tentatively identified only in the port of Trieste among 12 studied ports across the entire Adriatic basin (Mozetič et al., 2019). We confirm the presence of this genus here with the identification of *Minidiscus trioculatus*. 18S recovered a relatively small number of genera, comparable to the number of microscopy. Some common genera such as *Thalassionema* and *Nitzschia*, identified under the microscope and with *rbcL*, were absent among the taxa found. However, we must keep in mind that 18S sequencing was not tailored to diatoms. Therefore, a large proportion of the ASVs in the 18S dataset belonged to taxa that were not diatoms, whereas the majority of *rbcL* were diatoms. Among identified species, this discrepancy between markers was even greater, as we have shown using *Pseudo-nitzschia* and *Chaetoceros* as examples. The *rbcL* marker provided much higher resolution when considering species or even genera, with 22 new species in the study area, including 6 new genera. Since metabarcoding is an eDNA-based method, the source of amplicons could also be dead cells, cell fragments, resting stages and solubilized nucleic acids (Corinaldesi et al., 2008). This could be especially true for organelle-encoded genes, as organelle membranes provide additional protection against degradation. We speculate that at least some of the assigned taxa, such as *Pseudosolenia calcar-avis*, may be derived from such sources, as it is an extremely large diatom that is unlikely to be overlooked during microscopy, but was not found in the analyzed microscopy samples. Both data sets (18S and *rbcL*) contained a very high proportion of putative diatom sequences without genus or even species assignment. This is a common phenomenon in metabarcoding datasets obtained in novel regions, but such patterns are also common in areas with extensive sampling (Malviya et al., 2016; Piredda et al., 2018). Malviya et al. (2016) have shown that in global datasets 30–80% of all reads are unassigned at the genus level, with the percentage strongly dependent on the size fraction of diatoms, with smaller diatoms having larger percentages of unidentified amplicons. This is exemplified with our October-20 samples, obtained with a 20 µm phytoplankton net, and where the proportion of unassigned genera was in fact the smallest. The precentages of unassigned reads reported by Malviya et al. (2016) are comparable to both our 18S and *rbcL* data, with *rbcL* generally showing slightly higher unassigned rates. For *rbcL*, there is certainly also a lack of reference sequences, as the sequencing effort for cultured phytoplankton was much lower compared to 18S (Guo et al., 2015). Therefore, the larger proportion of unassigned reads is not surprising. On the other hand, it was surprising that the proportion of unassigned taxa in microscopy was very similar to that in metabarcoding, suggesting that taxa that have not yet been cultured can be detected by microscopy and selectively cultured to increase our knowledge of marine diatoms and improve reference databases.

Results of the comparative study are similar to Malviya et al. (2016), who showed that correlations between metabarcoding and microscopy exist but are not initially obvious. The discriminatory power of light microscopy is known to be much lower than that of metabarcoding. Metabarcoding recovered nearly twice as many taxa than microscopy in freshwater diatoms, for example (Zimmermann et al., 2015). Indeed, the main cause of the differences between the morphological and metabarcoding datasets were small diatom taxa such as *Minidiscus trioculatus* and *Thalassiosira* (several species), but also *Pseudo-nitzschia galaxiae*, which were largely overlooked by microscopy. Another obvious source of divergence between phytoplankton counts and metabarcoding data was the high number of *Cylindrotheca* counts in winter, which were replaced by *Pseudo-nitzschia* reads in the metabarcoding data. Since most of these reads belonged to *P. galaxiae*, we hypothesize that these differences are actually due to misclassification of *P. galaxiae* as *Cylindrotheca closterium* because the cells look very similar (Turk Dermastia et al., 2020). Many taxa that appeared in the *rbcL* dataset were benthic (e.g., *Psammodictyon*, *Cocconeis*), which is not surprising since most known diatom species are benthic and their presence in planktonic datasets with low reads has also been demonstrated in other diatom metabarcoding studies (Piredda et al., 2018). In addition, *Amphora*, another benthic genus has also been detected with microscopy. Interestingly, some of these genera including *Minidiscus*, *Cylindrotheca*, *Cocconeis*, *Amphora*, were among the most influential for the clustering of microscopy and mentabarcoding samples in the performed multivariate analyzes.

As for species data, we focused on *Chaetoceros* and *Pseudo-nitzschia*, as these genera have already been well studied (Turk Dermastia et al., 2020; Janja Francé, personal communication), with the former being even the most diverse diatom genus of the nearshore waters of the Adriatic (Mozetič et al., 2019). The number of

*Chaetoceros* species was similar for both methods, although the *rbcL* marker provided higher resolution and more details. Most of the species found have been previously recorded in the northern Adriatic (Bosak et al., 2009; Godrijan et al., 2013), with the exception of *C. dayaensis*, which was identified by *rbcL*. This species was described in Chinese tropical waters (Li et al., 2015). Some of the ASVs classified as *C. dayaensis* had 100% identical sequences to published ones, while others showed only 97% sequence similarity, suggesting there may be other related species present that have not been barcoded yet. The bloom of *C. socialis* found with 18S in September was also in accordance with previous studies investigating this genus (Bosak et al., 2009; Godrijan et al., 2013). In the records of the Slovenian harmful algal bloom (HAB) monitoring program, *C. socialis* was also found in this period (Janja Francé, personal communication, Supplementary Table S3). With *rbcL*, maximum abundance occurred in November and December, similar to one of the published studies (Godrijan et al., 2013). *C. throndsenii* and *C. tenuissimus* were present throughout the study period and were recovered by both markers. They also represented the majority of *Chaetoceros* amplicons. The other recovered *Chaetoceros* species had restricted occurrence windows that were very similar to the published data. One group of reads in January and February with *rbcL* could not be identified to species level. This may be *C. cf. wighamii*, which was recognized by 18S but only with BLAST, which was not shown in Figure 6. *rbcL* reference sequences for this species were not available at the time of analysis. This species was previously found in the region by microscopy but it is a relatively new recognized taxon. The microscopic phytoplankton monitoring program in the Slovenian part of the Gulf of Trieste does not distinguish between many *Chaetoceros* species. On the other hand, many species are recognized but not quantified as part of the HAB monitoring program. This data was used on our case for comparison. Many more *Chaetoceros* species were identified by microscopy in certain months compared particularly to 18S metabarcodes. The assemblage compositions were comparable but not the same. This can be explained by the unsystematic way in which the presence of species was detected under the microscope, but possibly also due to misclassification. For example, *C. tenuissimus*, a very common representative in the metabarcoding data is not found in microscopy but may instead be classified as *C. simplex*. With *rbcL* the lack of reference sequences was really evident, since there was a large share of ASVs that were classified as *Chaetoceros* sp., while with 18S there were only a handful. This is thus perhaps the largest burden of *rbcL* metabarcoding at present, especially for genera with complex taxonomies and hundreds of species such as *Chaetoceros*.

Diatoms of the genus *Pseudo-nitzschia* were detected with both the 18S and *rbcL* markers. In terms of number of species identified, *rbcL* performed much better, but at the genus level the markers were comparable. 18S was only able to discriminate *P. delicatissima*, although given the discriminatory power of 18S-V9 and the fact that *P. delicatissima* was represented by only a handful of *rbcL* reads, we assume that these sequences actually belonged to another species, possibly *P. galaxiae*. *rbcL* detected 8 different species (*P. calliantha*, *P. delicatissima*, *P. cf. delicatissima*, *P. galaxiae*, *P. linea*, *P. mannii*, *P. fraudulenta*, *P. subfraudulenta*). Most were previously found in the GoT with very similar seasonal distributions determined by a combination of environmental and cultural techniques (Turk Dermastia et al., 2020). Of these a new confirmation for the area was *P. linea*, although it is thought to be overlooked in microscopic monitoring due to its small size and epiphytic nature, usually associated with *Chaetoceros* (Lundholm et al., 2002, 2012; Ruggiero et al., 2015). ASVs classified as this species had lower similarity to published sequences than the other *Pseudo-nitzschia* ASVs. This is probably also due to the lack of reliable reference sequences. The ASVs belonging to this species occurred sporadically, with more reads present in winter. *P. linea* represents a new species for the area. *P. linea* is known to occur in the Mediterranean and has been found mainly in the winter months (Quijano-Scheggia et al., 2010; Ruggiero et al., 2015). Among the species found, there are also several clusters of sequences identified as *P. galaxiae*, representing different morphological types (Figure 9). The phylogeny of *P. galaxiae* is complex and it has been demonstrated several times that it is probably a species complex in which the morphological types are genetically distinct, which could also be related to their toxicity (Ruggiero et al., 2015; Turk Dermastia et al., 2020, 2022). The strains of *P. galaxiae* show great diversity, which was also confirmed by our correlation analysis between the number of haplotypes and ASV frequency. Another novel species in the study area we identified is *P. cf. delicatissima*, which was also time-separated from its sister species *P. delicatissima*. This taxon was previously described from the Gulf of Naples, where it showed a very similar pattern of seasonal occurrence, appearing from September to December as in our data (Lamari et al., 2013; Ruggiero et al., 2015). We could not detect *P. pungens* and *P. multistriata* in our metabarcodes, although these species are known to occur here (Turk Dermastia et al., 2020). However, in recent years, the former species was no longer found in phytoplankton monitoring samples, while the other is very rare (Janja Francé, personal communication). Our study shows, that with complete reference databases *rbcL* metabarcoding can provide near strain-level resolution of the assemblage. This is important for monitoring HAB species. *Pseudo-nitzschia galaxiae* is known to be toxic in this area (Turk Dermastia et al., 2022), with the toxicity being strain dependent. Our study provided insight into the occurrence and frequency of different haplotypes (strains), although this strategy is still too time-consuming because actions to manage HABs must be taken in near real time. It is however, a valuable tool to identify novel strains, species or even genera in a particular study site, that are missed by microscopy or are even new to science.

Two of the most surprising results are the dominance of *P. galaxiae* ASVs in this study and the rarity of *P. calliantha*. The high number of *P. galaxiae* ASVs in January and February was surprising mainly because *Pseudo-nitzschia* is rarely detected in high numbers by cell counting in these months (Turk Dermastia et al., 2020; Vascotto et al., 2021), including cell counts in the year 2020 when samples were collected for metabarcoding. The tiny size of the small *P. galaxiae* morphotype could make the counts very unreliable, first because of the low detection rate and second because of misclassification. One cause of this misclassification is already apparent in our data, because there was a disproportionately large difference in the relative abundance of *Cylindrotheca* and *Pseudo-nitzschia* obtained by microscopy and metabarcoding in these months (Figure 4). *Pseudo-nitzschia galaxiae* is easily misinterpreted as *Cylindrotheca closterium*, especially in fixed samples. This is therefore a clear example of how biased long-term microscopic data sets can be and why it is critical to introduce barcoding techniques

into protist monitoring programs (Stern et al., 2018). As noted earlier, another cause of this discrepancy could again be signal amplification by metabarcoding, since the abundance of diatoms in winter months was low. This should at least partly be addressed by data normalization, which in fact diminished the influence of *Pseudo-nitzschia* but the influence of *Cylindrotheca* remained high. The rarity of *P. calliantha* was surprising, as it was one of the most abundant species in previous studies (Turk Dermastia et al., 2020) and is usually present throughout the year. It is easily confused with *P. mannii* under light microscopy, although previous studies also confirmed its presence using genetic methods as well.

### 4.3. *Rbcl* metabarcoding as a population-genetics tool

Lastly, we discuss the potential of *rbcL* metabarcoding to draw meaningful conclusions about population structure and microevolution. We focused on two widely studied groups, and in the case of *Pseudo-nitzschia galaxiae*, one species for which speciation hypotheses exist (Ruggiero et al., 2015, 2022; Turk Dermastia et al., 2022). When using metabarcoding as a population genetics tool, caution must be exercised with sequencing technology and ASV calling algorithms, which are not error-prone and therefore can artificially inflate haplotype composition (Nearing et al., 2018; Turon et al., 2020). However, when function and speciation are of interest, higher resolution can lead to discriminating taxa based on functionally irrelevant variation and decreasing predictive power for observing environmental and biological covariation (Needham et al., 2017). Thus, a balance must be struck. Several pipelines have been developed to address this problem (Frøslev et al., 2017; Turon et al., 2020; Kleine Bardenhorst et al., 2022). The simplest way to detect errors is to use coding genes and translate the obtained sequences to identify stop codons. In our case, *rbcL* performs this function well, and our haplotypes were all screened for translation errors. Next, we can proceed with OTU picking, although this inevitably reduces the number of true haplotypes. Post-clustering methods can also be used (Frøslev et al., 2017; Turon et al., 2020), which were used in this study to avoid OTU-picking.

In our case, *P. galaxiae* haplotypes clearly showed time-separation that was also related to the inferred morphologies of these haplotypes. The morphological assignments were based on the minimal E-scores in BLAST, which matched the cultured specimens from our culture collection with analyzed morphologies (Turk Dermastia et al., 2020). The most common haplotypes showed identical or near-identical sequences to these strains. This is to be expected as it is statistically very likely that the most common haplotype will be isolated if the sampling effort is not very large, which was not the case in our case as we isolated only 6 strains (Turk Dermastia et al., 2020). The dominant haplotype was usually accompanied by less frequent but still abundant auxiliary haplotypes. Marker genes say nothing about the abundance of evolutionary processes taking place at the genome level, but they can be informative enough (Kashtan et al., 2014). There are many factors that drive these microevolutionary processes. The simplest are mitotic division errors, followed by clonal expansion and bottlenecks that follow blooming periods (Ruggiero et al., 2018). As in classical evolutionary theory, some of these mutations provide an advantage to the new haplotype. For example, it has been shown that virus infection can drive haplotypic differentiation of phytoplankton and that host ASV variation is relevant for protection against the virus (Needham et al., 2017). The existence of *Pseudo-nitzschia* viruses has been demonstrated (Carlson et al., 2016) and also recently isolated from the GoT (unpublished data), while strain specificity for viral infections has also been demonstrated and is common in diatom viruses (Nagasaki et al., 2004). Another commonly observed pattern in metabarcoding data is that more common taxa tend to be more microdiverse (Needham et al., 2017; Ruggiero et al., 2022). In our case, this was true for both *Pseudo-nitzschia* and *Chaetoceros*, with the number of haplotypes positively correlated with haplotype reads. Recently, a study on the haplotypic composition of *Pseudo-nitzschia* 18S-V4 in the Gulf of Naples was conducted (Ruggiero et al., 2022). In this study, no curation of ASV tables was performed, so the haplotypic data sets are likely inflated, but we can see that very similar patterns emerged compared to our *rbcL* analysis. They also identified three clusters of *P. galaxiae* with different seasonal patterns and a very large number of haplotypes. On the other hand, *P. mannii* and *P. fraudulenta* were the least diverse species, similar to our study. The second most abundant species in our study, *P. subfraudulenta*, also had a larger number of haplotypes, while this species was not found in the Gulf of Naples.

In conclusion, we have shown that *rbcL* is a suitable marker for both assessment of diatom assemblages and eDNA-based population studies. This study will serve as a basis for future efforts to bring eDNA-based monitoring into the mainstream as it demonstrates that improved as well as complementary data to traditional monitoring can be obtained. We acknowledge that this study is burdened by a short time-series, while it could benefit from additional sequencing efforts using 18S diatom specific markers, and other regions of the 18S gene. However, due to remaining caveats, it is still better to use a combination of approaches to assess diatom assemblages. This way accurate and reproducible datasets can be obtained, which can serve as a basis for further studies that rely on accurate assessment of diversity. This is important because diatoms contribute immensely to the ocean's primary production, carbon and silicon budgets. These processes are dependent on the diversity and composition of diatom assemblages (Tréguer et al., 2018), therefore accurate diversity assessments are crucial.

## 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

TT has conducted the sampling, analysis and has written the draft of the manuscript. IV has contributed to the diversity analysis and provided comments to the manuscript. JF has conducted the analysis of microscopy samples and has provided comments and insight on the manuscript. DS has contributed to the bioinformatic analysis. PM has supervised the study and provided comments and insight on the manuscript. All authors contributed to the article and approved the submitted version.

## Funding

This work was conducted with the help of RI-SI-2 LifeWatch (Operational Program for the Implementation of the EU Cohesion Policy in the period 2014–2020, Development of Research Infrastructure for International Competition of Slovenian Development of Research infrastructure area–RI-SI, European Regional Development Fund, Republic of Slovenia Ministry of Education, Science, and Sport). The authors acknowledge the financial support from the Slovenian Research Agency (ARRS) research core funding No. P1-0237. DS was additionally supported by core funding No. P1-0255 and research project No. J1-3015. TT was supported by ARRS program for young researchers, in accordance with the agreement on (co) financing research activities.

## Acknowledgments

We acknowledge the effort of Tihomir Makovec, Tristan Bartole, and Leon Lojze Zamuda for collecting the samples, Petra Slavinec for the occasional help with filtration and Milijan Šiško for sampling and taxonomic identification of microscopy samples.

## 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.2023.1071379/full#supplementary-material

## References

Aitchison, J. (1982). The statistical analysis of compositional data. *J. R. Stat. Soc. B Methodol.* 44, 139–160. doi: 10.1111/j.2517-6161.1982.tb01195.x

Angel, R. (2012). Total nucleic acid extraction from soil. *Protoc. Exch.* 2012:46. doi: 10.1038/protex.2012.046

Armeli Minicante, S., Piredda, R., Finotto, S., Bernardi Aubry, F., Acri, F., Pugnetti, A., et al. (2020). Spatial diversity of planktonic protists in the lagoon of Venice (LTER-Italy) based on 18s rDNA. *Adv. Oceanogr. Limnol.* 11, 35–44. doi: 10.4081/aiol.2020.8961

Arsenieff, L., le Gall, F., Rigaut-Jalabert, F., Mahé, F., Sarno, D., Gouhier, L., et al. (2020). Diversity and dynamics of relevant nanoplanktonic diatoms in the Western English Channel. *ISME J.* 14, 1966–1981. doi: 10.1038/s41396-020-0659-6

Bosak, S., Burić, Z., Djakovac, T., and Viličić, D. (2009). Seasonal distribution of plankton diatoms in Lim Bay, northeastern Adriatic Sea. *Acta Bot. Croat.* 68, 351–365.

Bradley, I. M., Pinto, A. J., and Guest, J. S. (2016). Design and evaluation of Illumina MiSeq-compatible, 18S rRNA gene-specific primers for improved characterization of mixed phototrophic communities. *Appl. Environ. Microbiol.* 82, 5878–5891. doi: 10.1128/AEM.01630-16

Brush, M. J., Mozetič, P., Francé, J., Aubry, F. B., Djakovac, T., Faganeli, J., et al. (2021). "Phytoplankton dynamics in a changing environment" in *Coastal Ecosystems in Transition: A Comparative Analysis of the Northern Adriatic and Chesapeake Bay*. eds. T. C. Malone, A. Malej and J. Faganeli, vol. 9 (Hoboken: American Geophysical Union), 49–74.

Cabrini, M., Fornasaro, D., Cossarini, G., Lipizer, M., and Virgilio, D. (2012). Phytoplankton temporal changes in a coastal Northern Adriatic site during the last 25 years. *Estuar. Coast. Shelf Sci.* 115, 113–124. doi: 10.1016/j.ecss.2012.07.007

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

Cameron, E. S., Schmidt, P. J., Tremblay, B. J., Emelko, M. B., and Muller, K. M. (2021). Enhancing diversity analysis by repeatedly rarefying next generation sequencing data describing microbial communities. *Sci. Rep.* 11:22302. doi: 10.1038/s41598-021-01636-1

Carlson, M. C., McCary, N. D., Leach, T. S., and Rocap, G. (2016). *Pseudo-nitzschia* challenged with co-occurring viral communities display diverse infection phenotypes. *Front. Microbiol.* 7:527. doi: 10.3389/fmicb.2016.00527

Caron, D. A., Countway, P. D., Savai, P., Gast, R. J., Schnetzer, A., Moorthi, S. D., et al. (2009). Defining DNA-based operational taxonomic units for microbial-eukaryote ecology. *Appl. Environ. Microbiol.* 75, 5797–5808. doi: 10.1128/AEM.00298-09

Cerino, F., Fornasaro, D., Kralj, M., Giani, M., and Cabrini, M. (2019). Phytoplankton temporal dynamics in the coastal waters of the North-Eastern Adriatic Sea (Mediterranean Sea) from 2010 to 2017. *Nat. Conserv.* 34, 343–372. doi: 10.3897/natureconservation.34.30720

Corinaldesi, C., Beolchini, F., and Dell'Anno, A. (2008). Damage and degradation rates of extracellular DNA in marine sediments: implications for the preservation of gene sequences. *Mol. Ecol.* 17, 3939–3951. doi: 10.1111/j.1365-294X.2008.03880.x

Cristescu, M. E. (2019). Can environmental RNA revolutionize biodiversity science? *Trends Ecol. Evol.* 34, 694–697. doi: 10.1016/j.tree.2019.05.003

Daugbjerg, N., and Andersen, R. A. (1997). A molecular phylogeny of the Heterokont algae based on analyses of chloroplast-encoded rbcl sequence data. *J. Phycol.* 33, 1031–1041. doi: 10.1111/j.0022-3646.1997.01031.x

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., and Dufour, A.-B. (2007). The ade4 package: implementing the duality diagram for ecologists. *J. Stat. Softw.* 22, 1–20. doi: 10.18637/jss.v022.i04

Edwards, M., Beaugrand, G., Hays, G. C., Koslow, J. A., and Richardson, A. J. (2010). Multi-decadal oceanic ecological datasets and their application in marine policy and management. *Trends Ecol. Evol.* 25, 602–610. doi: 10.1016/j.tree.2010.07.007

Fernandes, A. D., Reid, J. N., Macklaim, J. M., McMurrough, T. A., Edgell, D. R., and Gloor, G. B. (2014). Unifying the analysis of high-throughput sequencing datasets: characterizing RNA-seq, 16S rRNA gene sequencing and selective growth experiments by compositional data analysis. *Microbiome* 2, 1–13. doi: 10.1186/2049-2618-2-15

Frøslev, T. G., Kjøller, R., Bruun, H. H., Ejrnæs, R., Brunbjerg, A. K., Pietroni, C., et al. (2017). Algorithm for post-clustering curation of DNA amplicon data yields reliable biodiversity estimates. Nature. *Communications* 8:1188. doi: 10.1038/s41467-017-01312-x

Gloor, G. B., Macklaim, J. M., Pawlowsky-Glahn, V., and Egozcue, J. J. (2017). Microbiome datasets are compositional: and this is not optional. *Front. Microbiol.* 8:2224. doi: 10.3389/fmicb.2017.02224

Gloor, G. B., Wu, J. R., Pawlowsky-Glahn, V., and Egozcue, J. J. (2016). It's all relative: analyzing microbiome data as compositions. *Ann. Epidemiol.* 26, 322–329. doi: 10.1016/j.annepidem.2016.03.003

Godhe, A., Asplund, M. E., Harnstrom, K., Saravanan, V., Tyagi, A., and Karunasagar, I. (2008). Quantification of diatom and dinoflagellate biomasses in coastal marine seawater samples by real-time PCR. *Appl. Environ. Microbiol.* 74, 7174–7182. doi: 10.1128/AEM.01298-08

Godrijan, J., Marić, D., Tomažić, I., Precali, R., and Pfannkuchen, M. (2013). Seasonal phytoplankton dynamics in the coastal waters of the North-Eastern Adriatic Sea. *J. Sea Res.* 77, 32–44. doi: 10.1016/j.seares.2012.09.009

Guillou, L., Bachar, D., Audic, S., Bass, D., Berney, C., Bittner, L., et al. (2013). The Protist ribosomal reference database (PR2): a catalog of unicellular eukaryote small sub-unit rRNA sequences with curated taxonomy. *Nucleic Acids Res.* 41, D597–D604. doi: 10.1093/nar/gks1160

Guo, L., Sui, Z., and Liu, Y. (2016). Quantitative analysis of dinoflagellates and diatoms community via Miseq sequencing of actin gene and v9 region of 18S rDNA. *Sci. Rep.* 6, 1–9. doi: 10.1038/srep34709

Guo, L., Sui, Z., Zhang, S., Ren, Y., and Liu, Y. (2015). Comparison of potential diatom 'barcode' genes (the 18S rRNA gene and ITS, COI, rbcL) and their effectiveness in

discriminating and determining species taxonomy in the Bacillariophyta. *Int. J. Syst. Evol. Microbiol.* 65, 1369–1380. doi: 10.1099/ijs.0.000076

Hansen, H. P., and Koroleff, F. (1999). "Determination of nutrients" in *Methods of Seawater Analysis*. eds. K. Grasshoff, K. Kremling and M. Ehrhardt. *3rd* ed (Weinheim, Germany: Wiley-VCH Verlag GmbH), 159–228.

Herbig, A., Maixner, F., Bos, K. I., Zink, A., Krause, J., and Huson, D. H. (2016). MALT: fast alignment and analysis of metagenomic DNA sequence data applied to the Tyrolean iceman. *BioRxiv* 2016:050559. doi: 10.1101/050559

Huson, D. H., Auch, A. F., Qi, J., and Schuster, S. C. (2007). MEGAN analysis of metagenomic data. *Genome Res.* 17, 377–386. doi: 10.1101/gr.5969107

Kashtan, N., Roggensack, S. E., Rodrigue, S., Thompson, J. W., Biller, S. J., Coe, A., et al. (2014). Single-cell genomics reveals hundreds of coexisting subpopulations in wild *Prochlorococcus*. *Science* 344, 416–420. doi: 10.1126/science.1248575

Kelly, R. P., Shelton, A. O., and Gallego, R. (2019). Understanding PCR processes to draw meaningful conclusions from environmental DNA studies. *Sci. Rep.* 9:12133. doi: 10.1038/s41598-019-48546-x

Kleine Bardenhorst, S., Vital, M., Karch, A., and Rubsamen, N. (2022). Richness estimation in microbiome data obtained from denoising pipelines. *Comput. Struct. Biotechnol. J.* 20, 508–520. doi: 10.1016/j.csbj.2021.12.036

Lamari, N., Ruggiero, M. V., D'Ippolito, G., Kooistra, W. H. C. F., Fontana, A., and Montresor, M. (2013). Specificity of Lipoxygenase pathways supports species delineation in the marine diatom genus *Pseudo-nitzschia*. *PLoS One* 8, –e73281. doi: 10.1371/journal.pone.0073281

Leblanc, K., Quéguiner, B., Diaz, F., Cornet, V., Michel-Rodriguez, M., Durrieu De Madron, X., et al. (2018). Nanoplanktonic diatoms are globally overlooked but play a role in spring blooms and carbon export. Nature. *Communications* 9:953. doi: 10.1038/s41467-018-03376-9

Legendre, P., and Legendre, L. (2012). *Numerical Ecology*. Amsterdam, Netherlands: Elsevier.

Leonilde, R., Elena, L., Elena, S., Francesco, C., and Alberto, B. (2017). Individual trait variation in phytoplankton communities across multiple spatial scales. *J. Plankton Res.* 39, 577–588. doi: 10.1093/plankt/fbx001

Li, Y., Zhu, S., Lundholm, N., and Lu, S. (2015). Morphology and molecular phylogeny of *Chaetoceros dayaensis* sp. Nov. (Bacillariophyceae), characterized by two 90 degrees rotations of the resting spore during maturation. *J. Phycol.* 51, 469–479. doi: 10.1111/jpy.12290

Lundholm, N., Bates, S. S., Baugh, K. A., Bill, B. D., Connell, L. B., Léger, C., et al. (2012). Cryptic and pseudo-cryptic diversity in diatoms-with descriptions of *Pseudo-nitzschia hasleana* sp. nov. and P. fryxelliana sp. nov. *J. Phycol.* 48, 436–454. doi: 10.1111/j.1529-8817.2012.01132.x

Lundholm, N., Hasle, G. R., Fryxell, G. A., and Hargraves, P. E. (2002). Morphology, phylogeny and taxonomy of species within the *Pseudo-nitzschia americana* complex (Bacillariophyceae) with descriptions of two new species, *Pseudo-nitzschia brasiliana* and *Pseudo-nitzschia* linea. *Phycologia* 41, 480–497. doi: 10.2216/i0031-8884-41-5-480.1

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

McMurdie, P. J., and Holmes, S. (2013). Phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* 8, –e61217. doi: 10.1371/journal.pone.0061217

McMurdie, P. J., and Holmes, S. (2014). Waste not, want not: why rarefying microbiome data is inadmissible. *PLoS Comput. Biol.* 10:e1003531. doi: 10.1371/journal.pcbi.1003531

Moniz, M. B., and Kaczmarska, I. (2009). Barcoding diatoms: is there a good marker? *Mol. Ecol. Resour.* 9, 65–74. doi: 10.1111/j.1755-0998.2009.02633.x

Mozetič, P., Cangini, M., France, J., Bastianini, M., Bernardi Aubry, F., Buzancic, M., et al. (2019). Phytoplankton diversity in Adriatic ports: lessons from the port baseline survey for the management of harmful algal species. *Mar. Pollut. Bull.* 147, 117–132. doi: 10.1016/j.marpolbul.2017.12.029

Mozetič, P., Fonda Umani, S., Cataletto, B., and Malej, A. (1998). Seasonal and inter-annual plankton variability in the Gulf of Trieste (Northern Adriatic). *ICES J. Mar. Sci.* 55, 711–722. doi: 10.1006/jmsc.1998.0396

Mozetič, P., Francé, J., Kogovšek, T., Talaber, I., and Malej, A. (2012). Plankton trends and community changes in a coastal sea (Northern Adriatic): bottom-up vs. top-down control in relation to environmental drivers. *Estuar. Coast. Shelf Sci.* 115, 138–148. doi: 10.1016/j.ecss.2012.02.009

Mozetič, P., Solidoro, C., Cossarini, G., Socal, G., Precali, R., Francé, J., et al. (2010). Recent trends towards oligotrophication of the Northern Adriatic: evidence from chlorophyll a time series. *Estuar. Coasts* 33, 362–375. doi: 10.1007/s12237-009-9191-7

Nagasaki, K., Tomaru, Y., Katanozaka, N., Shirai, Y., Nishida, K., Itakura, S., et al. (2004). Isolation and characterization of a novel single-stranded RNA virus infecting the bloom-forming diatom *Rhizosolenia setigera*. *Appl. Environ. Microbiol.* 70, 704–711. doi: 10.1128/AEM.70.2.704-711.2004

Nearing, J. T., Douglas, G. M., Comeau, A. M., and Langille, M. G. I. (2018). Denoising the Denoisers: an independent evaluation of microbiome sequence error–correction approaches. *PeerJ* 6, –e5364. doi: 10.7717/peerj.5364

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

Nelson, D. M., Tréguer, P., Brzezinski, M. A., Leynaert, A., and Quéguiner, B. (1995). Production and dissolution of biogenic silica in the ocean: revised global estimates, comparison with regional data and relationship to biogenic sedimentation. *Glob. Biogeochem. Cycles* 9, 359–372. doi: 10.1029/95GB01070

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., et al. (2019). *Vegan: Community Ecology Package*. R Package Version 2.2-0.

Olenina, I. (2006). *Biovolumes and Size-classes of Phytoplankton in the Baltic Sea*. Helsinki: HELCOM Baltic Sea Environment Proceedings No. 106, pp. 144.

Paradis, E. (2018). Pegas: an R package for population genetics with an integrated-modular approach. *Bioinformatics* 26, 419–420. doi: 10.1093/bioinformatics/btp696

Paradis, E., and Schliep, K. (2019). Ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* 35, 526–528. doi: 10.1093/bioinformatics/bty633

Penna, A., Casabianca, S., Guerra, A. F., Vernesi, C., and Scardi, M. (2017). Analysis of phytoplankton assemblage structure in the Mediterranean Sea based on high-throughput sequencing of partial 18S rRNA sequences. *Mar. Genomics* 36, 49–55. doi: 10.1016/j.margen.2017.06.001

Piredda, R., Claverie, J. M., Decelle, J., de Vargas, C., Dunthorn, M., Edvardsen, B., et al. (2018). Diatom diversity through HTS-metabarcoding in coastal European seas. *Sci. Rep.* 8, 18059–18012. doi: 10.1038/s41598-018-36345-9

Piredda, R., Tomasino, M. P., D'Erchia, A. M., Manzari, C., Pesole, G., Montresor, M., et al. (2017). Diversity and temporal patterns of planktonic protist assemblages at a Mediterranean long term ecological research site. *FEMS Microbiol. Ecol.* 93:fiw200. doi: 10.1093/femsec/fiw200

Prokopowich, C. D., Gregory, T. R., and Crease, T. J. (2003). The correlation between rDNA copy number and genome size in eukaryotes. *Genome* 46, 48–50. doi: 10.1139/g02-103

Quijano-Scheggia, S., Garcés, E., Andree, K. B., de la Iglesia, P., Diogène, J., Fortuño, J. M., et al. (2010). Especies *Pseudo-nitzschia* en la costa Catalana: Caracterización y contribución al conocimiento actual de la distribución del género en el mar Mediterráneo. *Sci. Mar.* 74, 395–410. doi: 10.3989/scimar.2010.74n2395

R Core Team. (2019). *R: A Language and Environment for Statistical Computing*. Vienna, Austria.

Rimet, F., Chaumeil, P., Keck, F., Kermarrec, L., Vasselon, V., Kahlert, M., et al. (2016). R-Syst::diatom: an open-access and curated barcode database for diatoms and freshwater monitoring. *Database* 2016:baw016. doi: 10.1093/database/baw016

Round, F. E., Crawford, R. M., and Mann, D. G.. (1990). *The Diatoms: Biology and Morphology of the Genera*. Cambridge: Cambridge University Press.

Ruggiero, M. V., D'Alelio, D., Ferrante, M. I., Santoro, M., Vitale, L., Procaccini, G., et al. (2018). Clonal expansion behind a marine diatom bloom. *ISME J.* 12, 463–472. doi: 10.1038/ismej.2017.181

Ruggiero, M. V., Kooistra, W. H. C. F., Piredda, R., Sarno, D., Zampicinini, G., Zingone, A., et al. (2022). Temporal changes of genetic structure and diversity in a marine diatom genus discovered via metabarcoding. *Environ. DNA* 4, 763–775. doi: 10.1002/edn3.288

Ruggiero, M. V., Sarno, D., Barra, L., Kooistra, W. H. C. F., Montresor, M., and Zingone, A. (2015). Diversity and temporal pattern of *Pseudo-nitzschia* species (Bacillariophyceae) through the molecular lens. *Harmful Algae* 42, 15–24. doi: 10.1016/j.hal.2014.12.001

Samanta, B., and Bhadury, P. (2014). Analysis of diversity of chromophytic phytoplankton in a mangrove ecosystem using rbcL gene sequencing. *J. Phycol.* 50, 328–340. doi: 10.1111/jpy.12163

Santi, I., Kasapidis, P., Karakassis, I., and Pitta, P. (2021). A comparison of DNA metabarcoding and microscopy methodologies for the study of aquatic microbial eukaryotes. *Diversity* 13:180. doi: 10.3390/d13050180

Stefanni, S., Stankovic, D., Borme, D., de Olazabal, A., Juretic, T., Pallavicini, A., et al. (2018). Multi-marker metabarcoding approach to study mesozooplankton at basin scale. *Sci. Rep.* 8:12085. doi: 10.1038/s41598-018-30157-7

Stern, R., Kraberg, A., Bresnan, E., Kooistra, W. H. C., Lovejoy, C., Montresor, M., et al. (2018). Molecular analyses of protists in long-term observation programmes–current status and future perspectives. *J. Plankton Res.* 40, 519–536. doi: 10.1093/plankt/fby035

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

Tréguer, P., Bowler, C., Moriceau, B., Dutkiewicz, S., Gehlen, M., Aumont, O., et al. (2018). Influence of diatom diversity on the ocean biological carbon pump. *Nat. Geosci.* 11, 27–37. doi: 10.1038/s41561-017-0028-x

Turk Dermastia, T., Cerino, F., Stanković, D., Francé, J., Ramšak, A., Žnidarič Tušek, M., et al. (2020). Ecological time series and integrative taxonomy unveil seasonality and diversity of the toxic diatom *Pseudo-nitzschia* H. Peragallo in the northern Adriatic Sea. *Harmful Algae* 93:101773. doi: 10.1016/j.hal.2020.101773

Turk Dermastia, T., Dallara, S., Dolenc, J., and Mozetič, P. (2022). Toxicity of the diatom genus *Pseudo-nitzschia* (Bacillariophyceae): insights from toxicity tests and genetic screening in the northern Adriatic Sea. *Toxins* 14:60. doi: 10.3390/toxins14010060

Turon, X., Antich, A., Palacín, C., Præbel, K., and Wangensteen, O. S. (2020). From metabarcoding to metaphylogeography: separating the wheat from the chaff. *Ecol. Appl.* 30, e02036–e02019. doi: 10.1002/eap.2036

Utermöhl, H. (1958). Zur Vervollkommnung der quantitativen phytoplankton-Methodik. *SIL Commun. 1953–1996* 9, 1–38. doi: 10.1080/05384680.1958.11904091

Vascotto, I., Mozetič, P., and Francé, J. (2021). Phytoplankton time-series in a LTER site of the Adriatic Sea: methodological approach to decipher community structure and indicative taxa. *Water* 13:2045. doi: 10.3390/w13152045

Vasselon, V., Bouchez, A., Rimet, F., Jacquet, S., Trobajo, R., Corniquel, M., et al. (2018). Avoiding quantification bias in metabarcoding: application of a cell biovolume correction factor in diatom molecular biomonitoring. *Methods Ecol. Evol.* 9, 1060–1069. doi: 10.1111/2041-210X.12960

Vasselon, V., Domaizon, I., Rimet, F., Kahlert, M., and Bouchez, A. (2017a). Application of high-throughput sequencing (HTS) metabarcoding to diatom biomonitoring: do DNA extraction methods matter? *Freshw. Sci.* 36, 162–177. doi: 10.1086/690649

Vasselon, V., Rimet, F., Tapolczai, K., and Bouchez, A. (2017b). Assessing ecological status with diatoms DNA metabarcoding: scaling-up on a WFD monitoring network (Mayotte Island, France). *Ecol. Indic.* 82, 1–12. doi: 10.1016/j.ecolind.2017.06.024

Watson, G. M. F., and Tabita, F. R. (2006). Microbial ribulose 1,5-bisphosphate carboxylase/oxygenase: a molecule for phylogenetic and enzymological investigation. *FEMS Microbiol. Lett.* 146, 13–22. doi: 10.1111/j.1574-6968.1997.tb10165.x

Weigand, H., Beermann, A. J., Čiampor, F., Costa, F. O., Csabai, Z., Duarte, S., et al. (2019). DNA barcode reference libraries for the monitoring of aquatic biota in Europe: gap-analysis and recommendations for future work. *Sci. Total Environ.* 678, 499–524. doi: 10.1016/j.scitotenv.2019.04.247

Weinstein, M. M., Prem, A., Jin, M., Tang, S., and Bhasin, J. M. (2019). FIGARO: an efficient and objective tool for optimizing microbiome rRNA gene trimming parameters. *bioRxiv* 2019:610394. doi: 10.1101/610394

Wickham, H.. (2016). *ggplot2: elegant graphics for data analysis*. Berlin, Germany: Springer Science and Business Media.

Willis, A., and Bunge, J. (2015). Estimating diversity via frequency ratios. *Biometrics* 71, 1042–1049. doi: 10.1111/biom.12332

Willis, A. D., and Martin, B. D. (2022). Estimating diversity in networked ecological communities. *Biostatistics* 23, 207–222. doi: 10.1093/biostatistics/kxaa015

Yu, G., Smith, D. K., Zhu, H., Guan, Y., Tsan, T., and Lam, Y. (2016). 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

Zimmermann, J., Glockner, G., Jahn, R., Enke, N., and Gemeinholzer, B. (2015). Metabarcoding vs. morphological identification to assess diatom diversity in environmental studies. *Mol. Ecol. Resour.* 15, 526–542. doi: 10.1111/1755-0998.12336

Zingone, A., D'Alelio, D., Mazzocchi, M. G., Montresor, M., and Sarno, D.LTER-MC Team (2019). Time series and beyond: multifaceted plankton research at a marine Mediterranean LTER site. *Nat. Conserv.* 34, 273–310. doi: 10.3897/natureconservation.34.30789

frontiers | Frontiers in Microbiology

TYPE Methods
PUBLISHED 15 June 2023
DOI 10.3389/fmicb.2023.1039658



OPEN ACCESS

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

REVIEWED BY
Andrew Decker Steen,
The University of Tennessee, Knoxville, United States
Panagiotis D. Dimitriou,
University of Crete, Greece

*CORRESPONDENCE
C. Mazière
✉ camille.maziere@univ-pau.fr
C. Cravo-Laureau
✉ cristiana.cravo-laureau@univ-pau.fr

RECEIVED 08 September 2022
ACCEPTED 30 May 2023
PUBLISHED 15 June 2023

CITATION
Mazière C, Duran R, Dupuy C and Cravo-Laureau C (2023) Microbial mats as model to decipher climate change effect on microbial communities through a mesocosm study.
*Front. Microbiol.* 14:1039658.
doi: 10.3389/fmicb.2023.1039658



# Microbial mats as model to decipher climate change effect on microbial communities through a mesocosm study

C. Mazière[1,2]*, R. Duran[1], C. Dupuy[2] and C. Cravo-Laureau[1]*

[1]Université de Pau et des Pays de l'Adour, E2S UPPA, CNRS, IPREM UMR 525—Bât. IBEAS, BP1155, Pau, France, [2]La Rochelle Université, CNRS, UMR 7266 LIENSs (Littoral Environnement et Sociétés)—2, rue Olympe de Gouges, Bât. ILE, La Rochelle, France

Marine environments are expected to be one of the most affected ecosystems by climate change, notably with increasing ocean temperature and ocean acidification. In marine environments, microbial communities provide important ecosystem services ensuring biogeochemical cycles. They are threatened by the modification of environmental parameters induced by climate change that, in turn, affect their activities. Microbial mats, ensuring important ecosystem services in coastal areas, are well-organized communities of diverse microorganisms representing accurate microbial models. It is hypothesized that their microbial diversity and metabolic versatility will reveal various adaptation strategies in response to climate change. Thus, understanding how climate change affects microbial mats will provide valuable information on microbial behaviour and functioning in changed environment. Experimental ecology, based on mesocosm approaches, provides the opportunity to control physical-chemical parameters, as close as possible to those observed in the environment. The exposure of microbial mats to physical-chemical conditions mimicking the climate change predictions will help to decipher the modification of the microbial community structure and function in response to it. Here, we present how to expose microbial mats, following a mesocosm approach, to study the impact of climate change on microbial community.

KEYWORDS

**global change, ocean warming, ocean acidification, microbial mat, experimental ecology, mesocosms**

## 1. Introduction

In the marine environment, climate change is predicted to have great and long-term impacts (Pachauri, 2014; IPCC, 2021). The increase in air temperature due to the enhanced greenhouse effect will impact first the surface water temperature and then the deep ocean temperature (Orr et al., 2005; Pachauri, 2014; Hutchins and Fu, 2017; IPCC, 2021). The global surface temperature of the ocean is projected to exceed 2°C (relative to 1850–1900

temperatures) for the 22nd century according to RCP8.5 (Pachauri, 2014). Moreover, global warming is predicted to release faster $CO_2$ into the atmosphere, increasing its concentration from 390 ppm to 700 ppm by 2100 (Pachauri, 2014). This rise will lead to a decrease in pH between 0.3 and 0.4 units. Ocean acidification is associated with increasing concentrations of $HCO_3^-$, $CO_2$, and $H^+$ and decreasing concentrations of $CO_3^{2-}$. Thus, the global carbon cycle is directly affected by ocean acidification. Ocean models also predicted a decline in the dissolved oxygen inventory of 1–7% by 2100 (Schmidtko et al., 2017). These physical-chemical changes will have a strong impact on marine life. Microorganisms are abundant and diverse in the oceans, with $4 \times 10^{29}$ cells in the deep oceanic subsurface, $5 \times 10^{28}$ cells in the upper oceanic sediment and $1 \times 10^{29}$ cells in the oceans (Flemming and Wuertz, 2019). They dominate the metabolic activity, for example, ensuring approximately half of the global primary production of the oceans (O'Brien et al., 2016). They play a central role in the biogeochemical cycles, as well as in the exchange of trace gases that have direct impacts on local climate (O'Brien et al., 2016). They are also crucial for the good functioning of the aquatic ecosystems being at the basis of food webs. Thus, marine microbial communities are expected to play a central role in the response of the ecosystem faced with environmental change because of their key functions in the oceans (Azam and Malfatti, 2007).

Microbial mats are laminated microbial structures (van Gemerden, 1993) observed at the water-sediment interface in several environments (Fourçans et al., 2004; Bolhuis et al., 2014; Mazière et al., 2021). They show a vertical stratification according to different physical-chemical gradients such as oxygen, sulphide, and pH. In such organization different interacting metabolisms coexist allowing microbial mats to be self-sustaining structure ensuring important ecosystem services in coastal areas (Pajares et al., 2015). In response to environmental changes, microbial mats can show various strategies due to their versatility and large diversity, as well as to their functional resilience (Bordenave et al., 2007). They represent ideal microbial community model to study the effect of global changes.

Current climate change effects on microbial communities can be investigated but it is difficult to simulate the conditions induced by climate change *in situ*, especially the acidification that remains difficult to control in open natural environments. Different approaches have been developed to study climate change effects on organisms inhabiting marine environments. Some studies focused on specific natural environments defined as models where one or several environmental parameters are similar to those predicted by IPCC, such as hot springs to simulate a high water temperature (Kitidis et al., 2011; Danovaro et al., 2017). However, the microorganisms inhabiting these environments are already adapted to the physical-chemical changes occurring on them and no temporal changes comparable to those due to climate change are observed. Moreover, the environmental complexity does not allow a definitive conclusion on the cause(s) of the observed changes. Thus, the precise impact on the environmental functioning and the dynamic of the microbial community in response to climate change cannot be studied.

Global changes are rapid simultaneous changes of environmental factors due to anthropogenic activities which threaten marine microorganisms. So far, little knowledge has been acquired to predict the effect of combined alteration of several environmental factors on marine microbial communities. Understanding the microbial communities shifts in response to the modification of environmental factors would allow to predict the effect of global changes on ecosystem functioning (Pajares et al., 2015). In order to gain such knowledge, it is thus necessary to develop accurate and controlled strategies to expose model microbial community to controlled climate change conditions. The mesocosms are a good approach because they allow the implementation of artificial scenarios by controlling physical-chemical parameters mimicking as close as possible the natural environment or the environmental conditions to be simulated (Cravo-Laureau and Duran, 2014). It is also possible to distinguish each factor independently allowing thus the characterization of their precise impact. Other advantages are that replicates and analysis can be multiplied, as well as manipulations can be performed without the constraints encountered *in situ* (Petersen et al., 2009). Data collection and analysis are thus facilitated. Mesocosms are enclosed ecosystem experiments that have gained in popularity as research tools in ecological science, particularly in the study of coastal aquatic environments (Petersen et al., 2009). Mesocosms have been developed for many applications, including the study of environmental stress (Pajares et al., 2012), hydrocarbon degradation (Cravo-Laureau and Duran, 2014), simulated oil spill (Chronopoulou et al., 2013), climate change simulation (Stewart et al., 2013; Mazière et al., 2022). Mesocosms have been used to analyse the behaviour of specific microbial communities (Chronopoulou et al., 2013; Stauffert et al., 2013), and of complex microbial structure such as biofilm (Agogué et al., 2014) and microbial mats (Mazière et al., 2022). They allow the control of different physical-chemical parameters and the association of them that can act synergistically (Baragi et al., 2015; Baragi and Anil, 2016; Li et al., 2016; Mazière et al., 2022), which thus allows to disentangle the impact of the studied factors from other environmental factors (Cravo-Laureau and Duran, 2014). So far, most studies have focused on the impact on one specific population or microorganisms (Widdicombe et al., 2009; Beman et al., 2011; Gingold et al., 2013; Guanyong et al., 2017; Lee et al., 2017; Tan et al., 2019) in a specific environmental compartment, such as water column and deep-sea benthos (Danovaro et al., 2017). However, the investigation of complex microbial assemblages is necessary to obtain data representative of the reality (Dimitriou et al., 2017). For that, a good microbial community model, representing a complex network of microorganisms, diversified at both the taxonomic and functional levels, is required. Microbial mats have been shown to represent good microbial community models in experimental ecology, particularly revealing the effect of oil spills on microbial community structure (Bordenave et al., 2004b, 2007) and functioning (Bordenave et al., 2004a, 2008). Moreover, microorganisms inhabiting microbial mats are well-adapted to physical-chemical fluctuations (Fourçans et al., 2006) making them adequate models to study climate change in microbial ecosystems. To the best of our knowledge, only a few studies have reported the impact of climate change on microbial mats (Verleyen et al., 2010; Mazière et al., 2022). In this paper, an experimental methodology is proposed where microbial mats were used as a microbial community model in a mesocosm experiment, representative of coastal areas, which showed to be a good way

for deciphering climate change effect on complex microbial community.

## 2. Materials and methods

Climate change was simulated by ocean acidification and warming according to the most pessimistic scenario (RCP8.5) predicted by IPCC (Pachauri, 2014).

### 2.1. Mesocosms design

Microbial mats for mesocosm experiments came from a non-exploited salt marsh located in Ars-en-Ré (46°13′29.9″N 1°31′07.5″W, Ré Island, France). These microbial mats were selected as microbial model to study the climate change impact because they presented layers more developed than those from exploited ponds (Mazière et al., 2021). A constant water height of 3 cm was maintained above them by the salt marsh owner to prevent microbial mats from being exposed to the air. Microbial mats were sampled with a core collector (2 cm mat and sediment depth), and transferred while maintaining their integrity, into mesocosms boxes (48 × 33 cm). Their transport to the laboratory was done without water, in the dark and at room temperature to avoid the destruction of the vertical structure.

Four sets of six mesocosms were built, each representing a treatment (**Figure 1**). The water was distributed independently from the seawater reservoir on each microbial mats *via* stainless steel taps. This seawater was filtered at 80 μm and passed under UV light to prevent the presence of other marine organisms in the mesocosms system. The salinity was adjusted to that measured *in situ* [60 practical salinity unit (psu)] by adding salt coming from the salt marshes of the Ré Island.

A stable water level was maintained at 3 cm above the mat with a water outlet in front of its inlet. The microbial mats in salt marshes were under a continuous flow-through water depending on the evaporation rate and thus, of the daily temperature making the determination of the water flux difficult. The microbial mats in the mesocosms were under a continuous slow (0.5 L/h; dilution rate estimated around $10^{-5}$/h; water turnover in 1h20) flow-through water diffusion simulating a slow water inlet occurring in salt marshes, which ensures water renewal supplying natural nutrients for the development of microbial mats (**Figure 1**). The microbial mats were illuminated by LED (TOP-24H company by SYLED, France) 12 h a day following the day/night cycle observed during the spring in France (**Figure 1**). They provided white cold light colour, limiting the evaporation rate, and intensity of 12 ± 1 μmol.photons.$m^{-2}$.$s^{-1}$ (corresponding to 874 ± 70 lux) (HOBO Pendant® Temperature/Light Data Logger, Onset Computer Corporation, USA) on the surface of the microbial mats. The LED approximates a solar spectrum according to the manufacturer's instructions. It is certified like a natural light, used for the cultivation of aquatic plants.

For each treatment, the desired temperature water above the microbial mats was maintained by placing the six containers with the sediment samples in a thermostated water bath with a pump (EHEIM universal 600, Germany) connected to a thermoregulating device (Teco®) (**Figure 2**).

### 2.2. Simulated scenarios

In order to acclimate the microbial mats to their new environment, a stabilisation period was applied (Stauffert et al., 2013; Gette-Bouvarot et al., 2015) maintaining the temperature (20°C) and salinity (60 psu) observed *in situ*. In this experiment, this period lasted 5 weeks ending when the physical-chemical parameters (salinity, temperature, pH, and dissolved oxygen concentration) in the water above microbial mats were stabilised for 1 week [see figure 2 in Mazière et al. (2022) and figure A in its Supplementary Materials].

A control (C) in which the parameters were not changed, and three treatments were then applied to the microbial mats for seven additional weeks [see figure A in the Supplementary Materials of Mazière et al. (2022)]. The first treatment was the warming treatment (W) in which the temperature was increased by 0.5°C every 2 days for 2 weeks until reaching 24°C. The acidification treatment (A) represented the second condition. Water acidification was performed in the water reservoir with the addition of $CO_2$ (IKS aquastar, iks ComputerSysteme GmbH, Germany), allowing a drop of initial water pH of 0.1 unit every 4 days for 2 weeks until reaching a decrease of the initial pH of the water reservoir equal to 0.4 unit, i.e., 7.6. The water pH in the water reservoir was precisely monitored with a pH probe (826 pH mobile, Ω Metrohm, Swiss) to adjust the pH on IKS device by taking into account the alkalinity using the seacarb package on RStudio software (version 4.2.1© RStudio, R Core Team, 2020). This probe was calibrated in standard solutions (pH 4, 7, and 10, HANNA (instruments, France). The third treatment combined warming and acidification treatments (WA). Acidification and water warming were performed over 2 weeks to mimic a continuous temporal variation and not a sudden change in these parameters like the predicted IPCC scenarios (Pachauri, 2014). The mesocosms were then maintained with the stable conditions for five further weeks.

Many parameters were monitored to allow the most accurate experimental ecology approach possible, combining physical-chemical, diversity, functional, and biochemical analyses (**Table 1**). The sampling strategy was described by Mazière et al. (2022). Briefly, measurement of all these parameters was done before the change of conditions (t9) and regularly during their variation and stabilization (t16, t23, t30, t37, t44, t51, and t58). Among the six replicates, one was specifically dedicated for the evaluation of the community's maximum potential for photosynthetic activity by measuring the chlorophyll *a* fluorescence with pulse amplitude modulation (PAM) as described in Mazière et al. (2022). The specific replicate was dedicated to PAM analysis because sampling for the analyses listed in **Table 1** were performed immediately beforehand, leading to a suspension of biological material, which can disturb PAM analysis.

### 2.3. DNA extraction and sequencing

Microbial mats subsamples of 0.25 g were done and used for DNA extraction using the DNeasy PowerSoil kit (Qiagen) according to the manufacturer's instructions. The bacterial V3-V4 region of the 16S rRNA gene was amplified using the

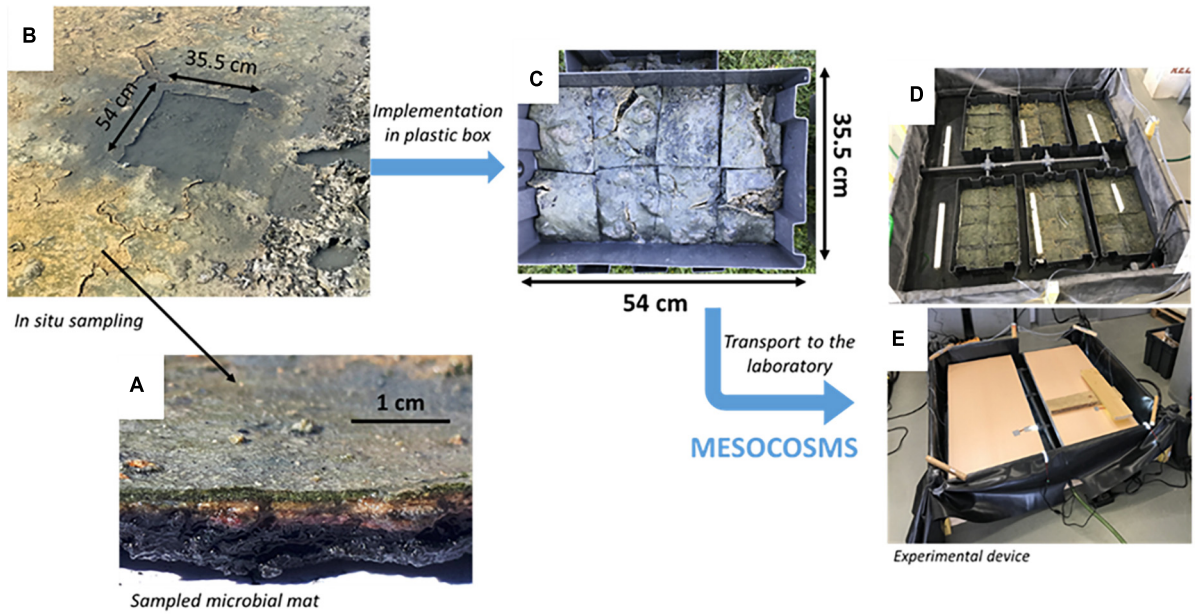



FIGURE 1
*In situ* sampling of the microbial mats for mesocosms experiment. Microbial mats were placed in containers with the sediment samples and transported to the laboratory for the mesocosms experiment. **(A)** Sampled microbial mat and its laminated structure. **(B)** *In situ* sampling of the microbial mat with a mold. **(C)** Implementation of the sampled microbial mat cut in eight parts to transfer it by avoiding damaging it as much as possible in a representing one mesocosm. **(D)** Six mesocosms circled by a pool representing one treatment. **(E)** The same than panel **(D)** with the light system above.

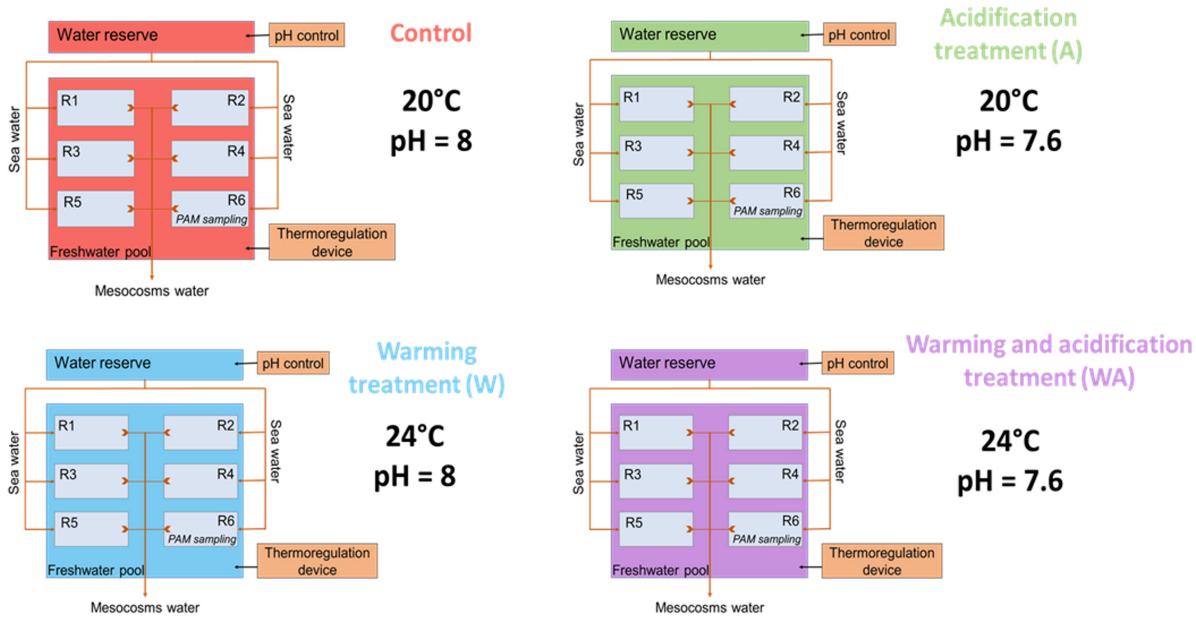



FIGURE 2
Schematic representation of the experimental design with the control (C), acidification (A), warming (W), and mixed (WA) treatments. The letter R represents the replicates. The first five replicates were used for sampling to perform the analysis listed in **Table 1** and the sixth was dedicated to evaluate the photosynthetic activity potential measuring the chlorophyll *a* fluorescence by pulse amplitude modulation (PAM) and physical-chemical gradients. This figure was modified from Mazière (2021) and Mazière et al. (2022).

primers 344F (5′-ACGGRAGGCAGCAG-3′) and 801R_m (5′-ACCAGGGTATCTAATCCT-3′) (Liu et al., 2007). PCR mix consisted in 12.5 μL of AmpliTaq Gold 360 master mix (Applied Biosystems), 1 μL of each primer (10 μM) and 1 μL of genomic DNA, in a final volume of 25 μL (adjusted with distilled water). All amplifications were performed on a Veriti 96 Well Thermal Cycler (Applied Biosystem) using the following PCR program: 10 min at 95°C, 30 cycles of 30 s at 95°C, 30 s at 63°C, and 40 s at 72°C,

TABLE 1 List of samplings performed from the mesocosms for the different analyses.

| Sampled material | Analysis | | References |
|---|---|---|---|
| Microbial mat | Diversity | | Mazière et al., 2021 |
| Microbial mat | Cell number | | Lavergne et al., 2014 |
| Microbial mat | EPS | | Mazière et al., 2022 |
| Surface water | EPS | | Mazière et al., 2022 |
| Microbial mat | Pigment composition | | Mazière et al., 2022 |
| Microbial mat | Meiofauna diversity and abundance | | Mazière et al., 2021 |
| Microbial mat | Bacterial production | | Garet and Moriarty (1996) modified by Pascal et al. (2009) |
| Surface water | Bacterial production | | |
| Surface water | Nutrients | Nitrate, nitrite, ammonium, phosphate | Aminot and Kérouel, 2007 |
| | | Silicon | |

A total of eight sampling points were done after the stabilisation period of 5 weeks. EPS, extracellular polymeric substances.

and finally, 10 min at 72°C. The sequencing was performed by the Genomic platform of Roscoff (France), using Illumina MiSeq technology.

## 2.4. Statistical analyses

The bioinformatic analyses on the DNA raw data obtained after sequencing were performed using the SAMBA (Standardized and Automated MetaBarcoding Analyses workflow) (v3.0.0) workflow written by SeBiMER, the IFREMER's Bioinformatics Core Facility (Noël et al., 2021). Two replicates of the control condition at t44 and t58 were removed because they didn't pass the quality control test during the data integrity step of the workflow. The taxonomic affiliation was performed against the Silva database v138 (Quast et al., 2012; Yilmaz et al., 2013) with 99% of similarities. All calculations and statistical analyses were performed on RStudio software (version 4.2.1 (RStudio, R Core Team, 2020). The statistical analyses used to interpret the physical-chemical data, the pigmentary composition, and the extracellular polymeric substances (EPS) concentration were described in Mazière et al. (2022). Non-metric multi-dimensional scaling (NMDS) analysis was based on the Bray–Curtis distance matrices calculated from the bacterial Amplicon Sequence Variants (ASVs) table (**Supplementary Table 1**). In order to define whether the treatment at a sampling time explained the variance among the Bray–Curtis distance matrices, a permutational multivariate analysis of variance (PERMANOVA) was performed with the *adonis* function of the vegan package. Linear discriminant analysis effect size (LEfSe) (Segata et al., 2011) on the 1,000 more abundant ASVs was performed on Galaxy web application to determine bacterial genera biomarkers for each treatment. The non-parametric Kruskal–Wallis sum-rank test ($\alpha = 0.05$) was performed to detect taxa with significant differential abundance. The biological consistency was investigated by performing a pairwise Wilcoxon test ($\alpha = 0.05$). A linear discriminant analysis (LDA) threshold score of 2.0 was applied.

# 3. Results and discussion

Although the decrease of 0.4 unit of pH (upH) in the seawater reservoir of the acidification treatments was effective, the water pH was not significantly different between the treatments [see figure 2 in Mazière et al. (2022) and figures B, D in its Supplementary Materials]. Such phenomenon has been previously described; it can be explained by abiotic factors controlling the $CO_2$ chemistry in seawater (Zeebe and Wolf-Gladrow, 2001) and by biotic factors linked to oxygenic photosynthesis (Giordano et al., 2005; Crawfurd et al., 2011; Ma et al., 2019). Thus, it can be hypothesized that the acidification might result in $CaCO_3$ dissolution, which in turn consume the protons in excess counteracting the acidification (Middelburg et al., 2020), explaining that pH was not decreased in our experiment. Regarding the biotic factors, it is known that many microbial species possess mechanisms allowing the storage of high $CO_2$ concentrations named carbon concentration mechanisms (CCMs). The CCMs have been described in oxygenic phototrophs like algae (Giordano et al., 2005) and cyanobacteria (Ma et al., 2019). The storage of the $CO_2$ added for acidification in CCMs might also explain why the decrease of pH was not observed in our experiment. However, when the maximum $CO_2$ concentration capacity of the CCMs will be reached in the cells, it could be expected to observe a pH decrease in the water column due to the added $CO_2$ (Black et al., 2019). Moreover, the dissolved oxygen concentration was observed to be more increased than in the other treatments, likely due to an increase in photosynthesis because of the carbon input (Black et al., 2019; Mazière et al., 2022). As expected, the temperature increased by 4°C from the initial temperature in the warming treatments [see figure 2 in Mazière et al. (2022)]. An increase in salinity was also observed in the warming treatments probably because of water evaporation [see figure 2 in Mazière et al. (2022)].

Thus, the mesocosms experiment allowed to change and maintain the conditions as expected (acidification of the water input and water warming on the mesocosms) (Petersen et al., 2009; Cravo-Laureau and Duran, 2014; Mazière et al., 2022). Such changes in environmental conditions are difficult to produce *in situ* because of the environmental complexity (day/night alternation, weather hazards, human impact. . .). By changing the conditions one by one, the changes in the other measured parameters could be attributed to the treatment. In our study, the warming treatment impacted the salinity whereas the acidification treatment affected the dissolved oxygen concentration. Thus, the mesocosms were effective to control the physical-chemical parameters for simulating climate change.

After changing the conditions, the bacterial community composition differed in the acidification treatment 3 weeks after (t44) in comparison to the other treatments (**Figure 3**) (PERMANOVA, $p < 0.05$). At the end of the experiment (t58), no more difference was observed between the acidification treatment and the control and warming treatments (**Figure 3**), suggesting that the bacterial community was resilient. However, the bacterial community composition differed, certainly due to the WA treatment (PERMANOVA, $p < 0.05$).

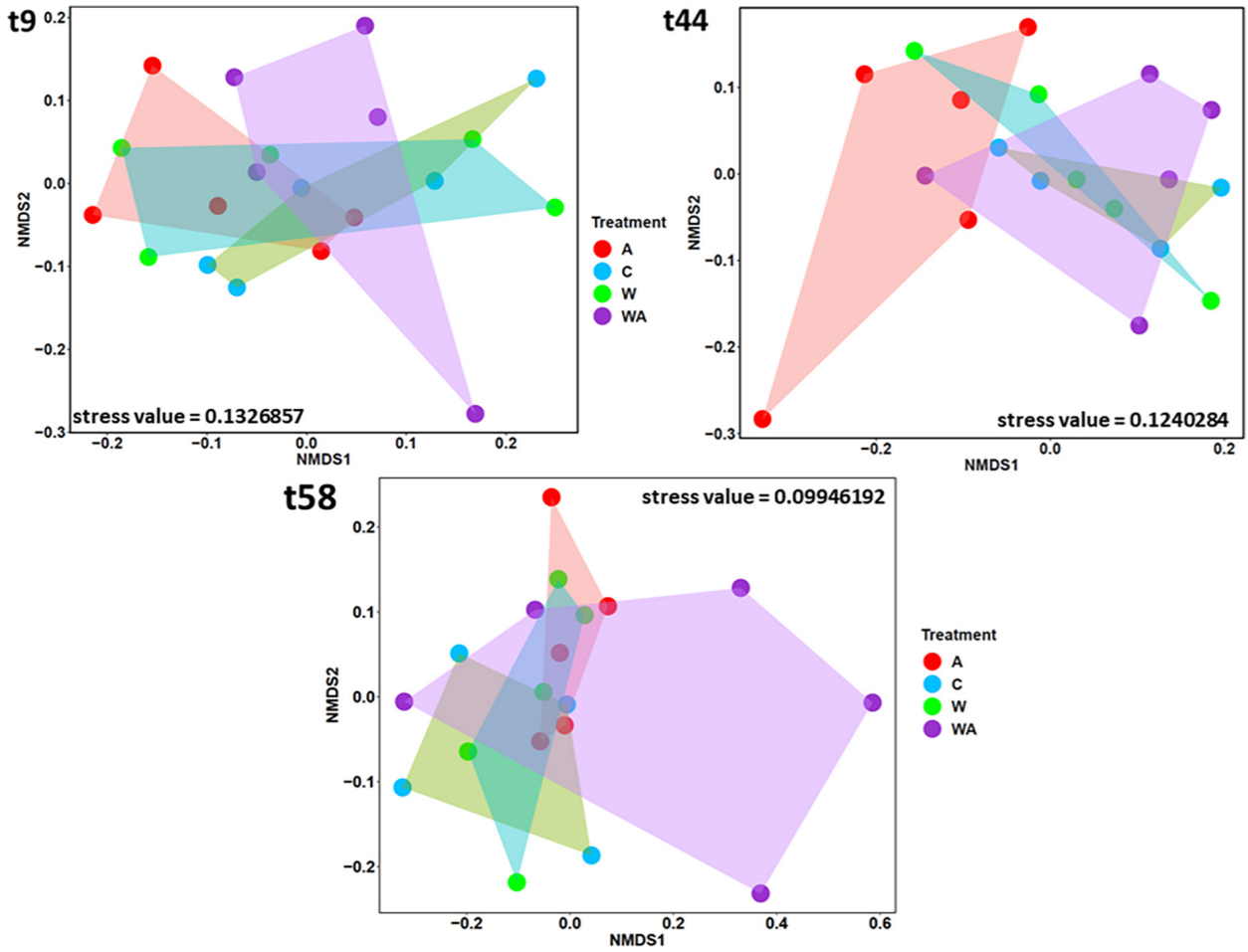



FIGURE 3
Comparison of bacterial communities according to the treatments. The Non-metric Multi-dimensional Scaling (NMDS), based on the bacterial 16S rRNA gene sequences, was performed to compare the bacterial communities according to the treatments [control (C), acidification (A), warming (W), and warming and acidification together (WA)] at the beginning (t9), the middle (t44), and the end (t58) of the experiment. The shapes correspond to the replicates ($n$ = 5, except for control $n$ = 4 at t44 and t58).

Indeed, LEfSe analyses on the 1,000 more abundant ASVs revealed bacterial taxa differentially abundant between the four treatments at t58. LEfSe revealed that the acidification was characterized by bacterial biomarkers, genera significantly more abundant in the acidification treatment, belonging to genera affiliated to Rhodobacteraceae and *Sulfurimonas* (**Figure 4**). The *Sulfurimonas* species are widespread in the environment and present large flexibility to colonize various habitats due to their versatile energy metabolisms and their adaptive abilities (Han and Perner, 2015). They play key roles in sulphur, hydrogen, nitrogen, oxygen, and carbon cycles (Han and Perner, 2015). The *Sulfurimonas* species are known to grow over a large range of pH (Inagaki, 2003; Takai et al., 2006) with the capacity to tolerate oxygen (Sievert et al., 2008), features allowing *Sulfurimonas* to outcompete under the conditions observed after acidification in our experiment. The Rhodobacteraceae play a key role in biogeochemical cycles and are often eukaryote mutualists (Simon et al., 2017). The abundance of Rhodobacteraceae has ever been shown to be affected under a high $pCO_2$ (Hu et al., 2021). *Desulfuromonas* genus, sulphate reducer bacteria (SRB) obligate anaerobe and obligate sulphur reducer (Margulis and Chapman, 2009), was revealed as a bacterial biomarker in the warming treatment (**Figure 4**). SRB are known to exhibit versatile metabolism, an asset to adapt to the modification of environmental conditions in various ecosystems (Giloteaux et al., 2013), particularly in microbial mats (Fourçans et al., 2008). The WA treatment presented *Guyparkeria*, *Ilumatobacter*, and MAT-CR-H6-H10 as bacterial biomarkers (**Figure 4**). *Guyparkeria* genus includes autotrophic sulphur oxidizers (Nosalova et al., 2022). *Ilumatobacter* species are generally found in contaminated places by hydrocarbons (Peng et al., 2015; Papale et al., 2020). MAT-CR-H6-H10 was described in a hypersaline microbial mat (Kirk Harris et al., 2013). The results suggest that the different treatments simulating climate change impact bacteria playing a role in the sulfur cycle. The acidification treatments (A and WA) affected sulfur oxidizers while the warming treatment impacted sulphate reducers.

Concomitantly, the increasing of the concentration of bound carbohydrate EPS in the acidification treatment was observed (Mazière et al., 2022) indicating that EPS are involved in the adaptation of communities to their environment as previously

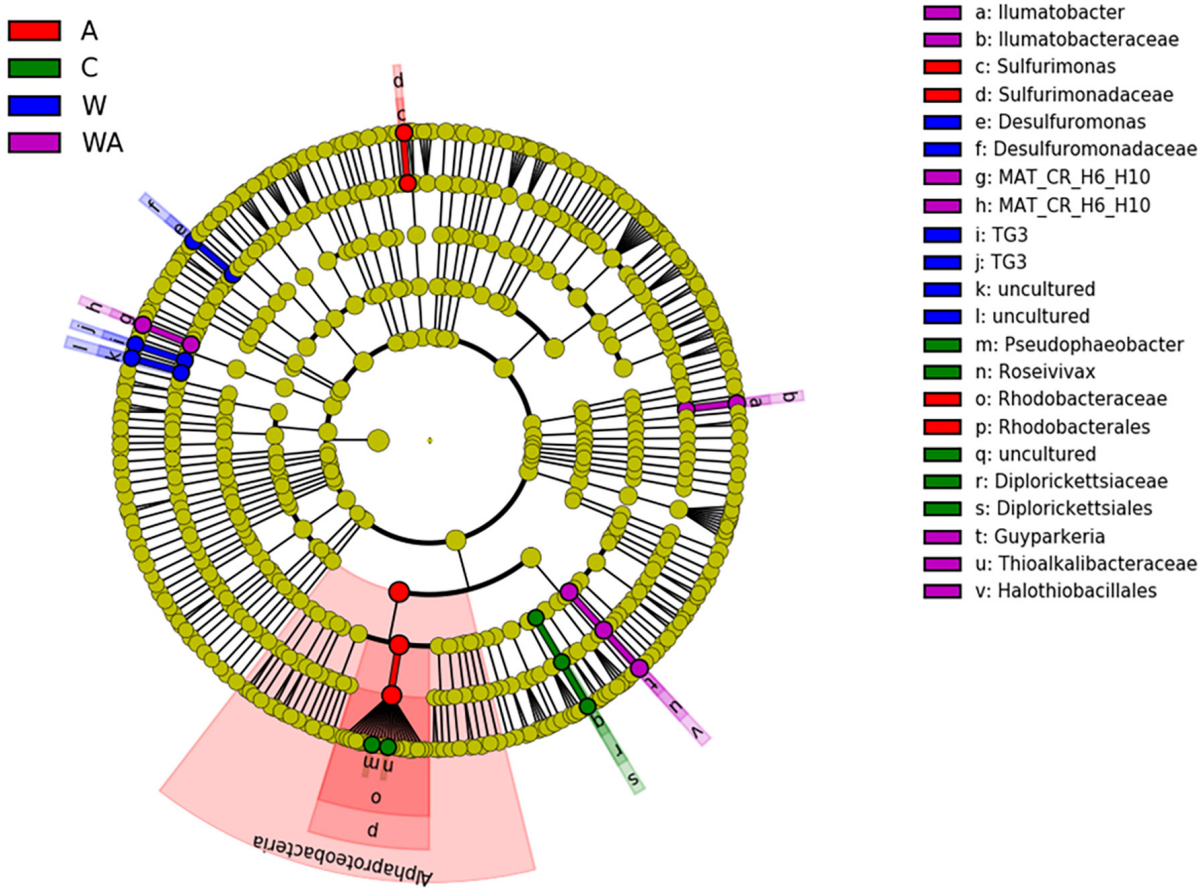



FIGURE 4
Linear discriminant analysis effect (LefSe). Comparison of bacterial communities at t58 in the different treatments [control (C), acidification (A), warming (W), and warming and acidification together (WA)] by linear discriminant analysis effect size (LEfSe). The analysis was performed with the 1,000 more abundant bacterial ASVs of each treatment.

proposed (Dupraz and Visscher, 2005; Hubas, 2018; Prieto-Barajas et al., 2018). The most probable explanation is that the studied microbial mats are already naturally exposed to fluctuating environmental conditions (pH, light, temperature, salinity, etc.) with a great amplitude in salt marshes (Fourçans et al., 2004, 2006, 2008; Abed et al., 2007; Boujelben et al., 2012; Baumann et al., 2015; Mazière et al., 2021). Therefore, they are already confronted to temperature such those predicted by the IPCC for 2100. Changes in the bacterial metabolism probably enabled bacteria to maintain a high activity potential when exposed to the conditions of the simulated scenarios. Noteworthy, a change in the pigment composition of the microbial mats submitted to acidification was observed (Mazière et al., 2022), that suggested a shift in the photosynthetic capacities corresponding to a modification in the photosynthetic communities or their metabolism (Mazière et al., 2022). The microbial mats maintained in mesocosms allowed us to observe the effect of the acidification alone on photosynthetic microorganisms, which was not observable in the treatment combining both acidification and warming.

To summarize, the acidification treatment affected the photosynthetic communities by changing the pigment proportion of the microbial mat, which probably resulted from a change in either their composition or their metabolisms (**Figure 5**). Concomitantly, an increase of dissolved oxygen concentration was observed in the water, likely due to an increase of the photosynthesis (**Figure 5**). The modifications of the conditions generated by the acidification impacted also the sulfur-oxidizers belonging to the *Sulfurimonas* genus and the Rhodobacteraceae family, which were significantly more abundant in this condition (**Figure 4**). In contrast, the warming of the water, characterized by an upsurge of salinity due to evaporation (**Figure 5**), increased the relative abundance of the sulphate-reducers *Desulfurimonas* (**Figure 4**). Finally, the warming and acidified treatment resulted in an increase of the water salinity but the dissolved oxygen concentration was not impacted (**Figure 5**), affecting the sulfur-oxidizers *Guyparkeria* (**Figure 4**).

Maintaining microbial mats in mesocosms following an experimental ecology approach allowed to apply conditions to simulate climate change, which is difficult to perform *in situ*. It also made possible to increase the number of replicates and analysis strengthening the statistical analyses. We were thus able to unveil the relationships between the microorganisms inhabiting

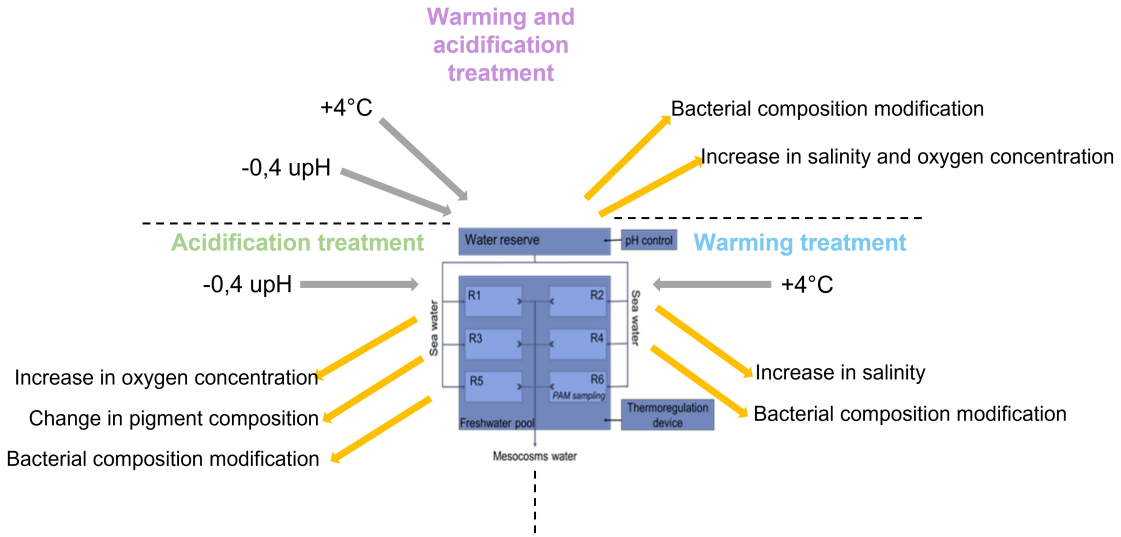



FIGURE 5
Summary of the impact of climate change on microbial mats.

the microbial mat and their environment, helpful information to refine the hypotheses of the impact of climate change at the ecosystem level. More generally, the gained information provides new insights to better understand the importance of microbial mats in their environment.

## 4. Conclusion and perspectives

Climate change is already happening, but it is difficult to determine its effect on complex microbial assemblages, particularly because *in situ* simulation experiments are difficult to perform. An alternative to overcome this difficulty is to follow an experimental ecology approach maintaining complex microbial community in mesocosm applying conditions simulating predicted climate change conditions. Microbial mats used allowed to unveil microbial structural modification in response to climate change. The mesocosm approach mimics *in situ* conditions excluding environmental complexity, which was useful to characterize the impact of each changed parameter independently. The total cost of the mesocosms set up was less than 3.000$. The described method in this paper allowed to simulate the change of two parameters, the water pH and temperature, alone and associated. The regular monitoring of several abiotic and biotic parameters and the presence of five replicates led to the complete description of the ecosystem. After applying climate change scenario, several physical-chemical and biological modifications were observed, which provide new insights on microbial mats behaviour. The gained information is useful to understand the impact of climate change at the ecosystem level. Microbial mats, easy to handle and to maintain in mesocosm, are thus precious complex microbial community model to investigate ecological issues of concern for coastal marine ecosystems. Mesocosms experiment can mimic as close as possible the environment but it can't reproduce all *in situ* variations and stresses. The method presented could therefore be improved by adding rainfall cycles or seasonal variations. Additional analyses could be performed to specify the functional role of microbial mats in its environment, such as vertical microprofiling of oxygen and sulphur, or a day/night monitoring. This method could be adapted with additional climate change parameters as for example, deoxygenation or the alternance of drought and extreme rainfall. Other issues could also be investigated with this system, such as the impact of oil pollution, metals contamination, organic enrichment, etc. The presented method could also be extended on sediments coming from aquatic environments, as for example estuaries, beach, deep ocean, but also rivers, lakes, by adjusting the sampling method and the parameters applied.

## Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## Author contributions

CM: formal analysis, methodology, and validation—original draft. RD, CD, and CC-L: funding acquisition, project administration, resources, and supervision. All authors: conceptualization, writing—review and editing, contributed to the article, and approved the submitted version.

## Funding

CM was supported by a Ph.D. grant from E2S-UPPA program and the Région Nouvelle-Aquitaine. We thank the funding support from the European programme ERANETMED AQUASALT

(NMED-0003-01) and from the ACI politique d'établissement Université de La Rochelle.

## Acknowledgments

The authors are grateful to the salterns owner Michel Jauffrais and all the people who participated in this work. The content of this manuscript is a part of the CM's thesis.

## 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.2023.1039658/full#supplementary-material

## References

Abed, R. M. M., Kohls, K., and de Beer, D. (2007). Effect of salinity changes on the bacterial diversity, photosynthesis and oxygen consumption of cyanobacterial mats from an intertidal flat of the Arabian Gulf. *Environ. Microbiol.* 9, 1384–1392. doi: 10.1111/j.1462-2920.2007.01254.x

Agogué, H., Mallet, C., Orvain, F., De Crignis, M., Mornet, F., and Dupuy, C. (2014). Bacterial dynamics in a microphytobenthic biofilm: a tidal mesocosm approach. *J. Sea Res.* 92, 36–45.

Aminot, A., and Kérouel, R. (2007). *Dosage automatique des nutriments dans les eaux marines: Méthodes d'analyse en milieu marin. Méthodes en flux continu.* Paris: QUAE.

Azam, F., and Malfatti, F. (2007). Microbial structuring of marine ecosystems. *Nat. Rev. Microbiol.* 5, 782–791.

Baragi, L. V., and Anil, A. C. (2016). Synergistic effect of elevated temperature, $pCO_2$ and nutrients on marine biofilm. *Mar. Pollut. Bull.* 105, 102–109. doi: 10.1016/j.marpolbul.2016.02.049

Baragi, L. V., Khandeparker, L., and Anil, A. C. (2015). Influence of elevated temperature and $pCO_2$ on the marine periphytic diatom Navicula distans and its associated organisms in culture. *Hydrobiologia* 762, 127–142.

Baumann, H., Wallace, R. B., Tagliaferri, T., and Gobler, C. J. (2015). Large natural pH, $CO_2$ and $O_2$ fluctuations in a temperate tidal salt marsh on diel, seasonal, and interannual time scales. *Estuaries Coasts* 38, 220–231.

Beman, J. M., Chow, C.-E., King, A. L., Feng, Y., Fuhrman, J. A., Andersson, A., et al. (2011). Global declines in oceanic nitrification rates as a consequence of ocean acidification. *Proc. Natl. Acad. Sci.* 108, 208–213. doi: 10.1073/pnas.1011053108

Black, J. G., Stark, J. S., Johnstone, G. J., McMinn, A., Boyd, P., McKinlay, J., et al. (2019). In-situ behavioural and physiological responses of Antarctic microphytobenthos to ocean acidification. *Sci. Rep.* 9:1890. doi: 10.1038/s41598-018-36233-2

Bolhuis, H., Silvia, M., and Stal, L. J. (2014). Molecular ecology of microbial mats. *FEMS Microbiol. Ecol.* 90, 335–350.

Bordenave, S., Fourçans, A., Blanchard, S., Goñi-Urriza, M. S., Caumette, P., and Duran, R. (2004a). Structure and functional analyses of bacterial communities changes in microbial mats following petroleum exposure. *Ophelia* 58, 195–203.

Bordenave, S., Goñi-Urriza, M. S., Caumette, P., and Duran, R. (2007). Effects of heavy fuel oil on the bacterial community structure of a pristine microbial mat. *Appl. Environ. Microbiol.* 73, 6089–6097. doi: 10.1128/AEM.01352-07

Bordenave, S., Goñi-Urriza, M. S., Vilette, C., Blanchard, S., Caumette, P., and Duran, R. (2008). Diversity of ring-hydroxylating dioxygenases in pristine and oil contaminated microbial mats at genomic and transcriptomic levels. *Environ. Microbiol.* 10, 3201–3211. doi: 10.1111/j.1462-2920.2008.01707.x

Bordenave, S., Jézéquel, R., Fourçans, A., Budzinski, H., Merlin, F. X., Fourel, T., et al. (2004b). Degradation of the "Erika" oil. *Aquat. Living Resour.* 17, 261–267.

Boujelben, I., Gomariz, M., Martínez-García, M., Santos, F., Peña, A., López, C., et al. (2012). Spatial and seasonal prokaryotic community dynamics in ponds of increasing salinity of Sfax solar saltern in Tunisia. *Antonie Van Leeuwenhoek* 101, 845–857. doi: 10.1007/s10482-012-9701-7

Chronopoulou, P.-M., Fahy, A., Coulon, F., Païssé, S., Goñi-Urriza, M. S., Peperzak, L., et al. (2013). Impact of a simulated oil spill on benthic phototrophs and nitrogen-fixing bacteria in mudflat mesocosms. *Environ. Microbiol.* 15, 242–252. doi: 10.1111/j.1462-2920.2012.02864.x

Cravo-Laureau, C., and Duran, R. (2014). Marine coastal sediments microbial hydrocarbon degradation processes: contribution of experimental ecology in the omics' era. *Front. Microbiol.* 5:39. doi: 10.3389/fmicb.2014.00039

Crawfurd, K. J., Raven, J. A., Wheeler, G. L., Baxter, E. J., and Joint, I. (2011). The response of thalassiosira pseudonana to long-term exposure to increased $CO_2$ and decreased pH. *PLoS One* 6:e26695. doi: 10.1371/journal.pone.0026695

Danovaro, R., Corinaldesi, C., Dell'Anno, A., and Rastelli, E. (2017). Potential impact of global climate change on benthic deep-sea microbes. *FEMS Microbiol. Lett.* 364:fnx214.

Dimitriou, P. D., Papageorgiou, N., Geropoulos, A., Kalogeropoulou, V., Moraitis, M., Santi, I., et al. (2017). A novel mesocosm setup for benthic-pelagic coupling experiments. *Limnol. Oceanogr. Methods* 15, 349–362.

Dupraz, C., and Visscher, P. T. (2005). Microbial lithification in marine stromatolites and hypersaline mats. *Trends Microbiol.* 13, 429–438. doi: 10.1016/j.tim.2005.07.008

Flemming, H.-C., and Wuertz, S. (2019). Bacteria and archaea on earth and their abundance in biofilms. *Nat. Rev. Microbiol.* 17, 247–260.

Fourçans, A., Oteyza, T. G., Wieland, A., Solé, A., Diestra, E., Bleijswijk, J., et al. (2004). Characterization of functional bacterial groups in a hypersaline microbial mat community (Salins-de-Giraud, Camargue, France). *FEMS Microbiol. Ecol.* 51, 55–70. doi: 10.1016/j.femsec.2004.07.012

Fourçans, A., Ranchou-Peyruse, A., Caumette, P., and Duran, R. (2008). Molecular analysis of the spatio-temporal distribution of sulfate-reducing bacteria (SRB) in camargue (France) hypersaline microbial mat. *Microb. Ecol.* 56, 90–100. doi: 10.1007/s00248-007-9327-x

Fourçans, A., Solé, A., Diestra, E., Ranchou-Peyruse, A., Esteve, I., Caumette, P., et al. (2006). Vertical migration of phototrophic bacterial populations in a hypersaline microbial mat from Salins-de-Giraud (Camargue, France). *FEMS Microbiol. Ecol.* 57, 367–377. doi: 10.1111/j.1574-6941.2006.00124.x

Garet, M. J., and Moriarty, D. J. W. (1996). Acid extraction of tritium label from bacterial DNA in clay sediment. *J. Microbiol. Methods* 25, 1–4. doi: 10.1007/BF02097401

Gette-Bouvarot, M., Mermillod-Blondin, F., Lemoine, D., Delolme, C., Danjean, M., Etienne, L., et al. (2015). The potential control of benthic biofilm growth by macrophytes—a mesocosm approach. *Ecol. Eng.* 75, 178–186.

Giloteaux, L., Duran, R., Casiot, C., Bruneel, O., Elbaz-Poulichet, F., Goñi-Urriza, M. S., et al. (2013). Three-year survey of sulfate-reducing bacteria community structure in Carnoulès acid mine drainage (France), highly contaminated by arsenic. *FEMS Microbiol. Ecol.* 83, 724–737. doi: 10.1111/1574-6941.12028

Gingold, R., Moens, T., and Rocha-Olivares, A. (2013). Assessing the response of nematode communities to climate change-driven warming: a microcosm experiment. *PLoS One* 8:e66653. doi: 10.1371/journal.pone.0066653

Giordano, M., Beardall, J., and Raven, J. A. (2005). $CO_2$ concentrating mechanisms in algae: mechanisms, environmental modulation, and evolution. *Annu. Rev. Plant Biol.* 56, 99–131. doi: 10.1002/jez.2367

Guanyong, O., Wang, H., Ranran, S., and Guan, W. (2017). The dinoflagellate Akashiwo sanguinea will benefit from future climate change: the interactive effects of ocean acidification, warming and high irradiance on photophysiology and hemolytic activity. *Harmful Algae* 68, 118–127. doi: 10.1016/j.hal.2017.08.003

Han, Y., and Perner, M. (2015). The globally widespread genus Sulfurimonas: versatile energy metabolisms and adaptations to redox clines. *Front. Microbiol.* 6:989. doi: 10.3389/fmicb.2015.00989

Hu, C., Li, X., He, M., Jiang, P., Long, A., and Xu, J. (2021). Effect of ocean acidification on bacterial metabolic activity and community composition in oligotrophic oceans, inferred from short-term bioassays. *Front. Microbiol.* 12:583982. doi: 10.3389/fmicb.2021.583982

Hubas, C. (2018). Biofilms, Tapis et Agrégats Microbiens: vers Une Vision Unificatrice (HDR (Habilitation à Diriger les Recherches)). Muséum National D'Histoire Naturelle. Available online at: https://doi.org/10.5281/zenodo.3784703 (accessed September 8, 2022).

Hutchins, D. A., and Fu, F. (2017). Microorganisms and ocean global change. *Nat. Microbiol.* 2:17058.

Inagaki, F. (2003). Sulfurimonas autotrophica gen. nov., sp. nov., a novel sulfur-oxidizing -proteobacterium isolated from hydrothermal sediments in the mid-okinawa trough. *Int. J. Syst. Evol. Microbiol.* 53, 1801–1805. doi: 10.1099/ijs.0.02682-0

IPCC (2021). "Climate change 2021: the physical science basis. contribution of working group I to the sixth assessment report of the intergovernmental panel on climate change. summary for policymakers," in *Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, eds V. Masson-Delmotte, P. Zhai, and A. Pirani (Geneva: IPCC).

Kirk Harris, J., Gregory Caporaso, J., Walker, J. J., Spear, J. R., Gold, N. J., Robertson, C. E., et al. (2013). Phylogenetic stratigraphy in the Guerrero Negro hypersaline microbial mat. *ISME J.* 7, 50–60. doi: 10.1038/ismej.2012.79

Kitidis, V., Laverock, B., McNeill, L. C., Beesley, A., Cummings, D., Tait, K., et al. (2011). Impact of ocean acidification on benthic and water column ammonia oxidation. *Geophys. Res. Lett.* 38:L21603.

Lavergne, C., Beaugeard, L., Dupuy, C., Courties, C., and Agogué, H. (2014). An efficient and rapid method for the enumeration of heterotrophic prokaryotes in coastal sediments by flow cytometry. *J. Microbiol. Methods* 105, 31–38. doi: 10.1016/j.mimet.2014.07.002

Lee, M. R., Torres, R., and Manríquez, P. H. (2017). The combined effects of ocean warming and acidification on shallow-water meiofaunal assemblages. *Mar. Environ. Res.* 131, 1–9. doi: 10.1016/j.marenvres.2017.09.002

Li, W., Xu, X., Fujibayashi, M., Niu, Q., Tanaka, N., and Nishimura, O. (2016). Response of microalgae to elevated $CO_2$ and temperature: impact of climate change on freshwater ecosystems. *Environ. Sci. Pollut. Res.* 23, 19847–19860. doi: 10.1007/s11356-016-7180-5

Liu, Z., Lozupone, C., Hamady, M., Bushman, F. D., and Knight, R. (2007). Short pyrosequencing reads suffice for accurate microbial community analysis. *Nucleic Acids Res.* 35:e120. doi: 10.1093/nar/gkm541

Ma, J., Wang, P., Wang, X., Xu, Y., and Paerl, H. W. (2019). Cyanobacteria in eutrophic waters benefit from rising atmospheric $CO_2$ concentrations. *Sci. Total Environ.* 691, 1144–1154.

Margulis, L., and Chapman, M. J. (2009). "Chapter one - kingdom prokaryotae (Bacteria, Monera, Prokarya)," in *Kingdoms and Domains*, 4th Edn, eds L. Margulis and M. J. Chapman (London: Academic Press).

Mazière, C. (2021). *Exploration and Study of the Impact of Climate Change on Microbial Mats in the Nouvelle-Aquitaine Region*. France: Université de Pau et des Pays de l'Adour.

Mazière, C., Agogué, H., Cravo-Laureau, C., Cagnon, C., Lanneluc, I., Sablé, S., et al. (2021). New insights in bacterial and eukaryotic diversity of microbial mats inhabiting exploited and abandoned salterns at the Ré Island (France). *Microbiol. Res.* 252:126854. doi: 10.1016/j.micres.2021.126854

Mazière, C., Bodo, M., Perdrau, M. A., Cravo-Laureau, C., Duran, R., Dupuy, C., et al. (2022). Climate change influences chlorophylls and bacteriochlorophylls metabolism in hypersaline microbial mat. *Sci. Total Environ.* 802:149787. doi: 10.1016/j.scitotenv.2021.149787

Middelburg, J. J., Soetaert, K., and Hagens, M. (2020). Ocean alkalinity, buffering and biogeochemical processes. *Rev. Geophys.* 58:e2019RG000681.

Noël, C., Quintric, L., Cormier, A., Leroi, L., and Durand, P. (2021). *SAMBA: Standardized and Automated MetaBarcoding Analyses workflow*. Available online at: https://doi.org/10.48546/WORKFLOWHUB.WORKFLOW.156.1 (accessed September 8, 2022).

Nosalova, L., Piknova, M., Bonova, K., and Pristas, P. (2022). Deep subsurface hypersaline environment as a source of novel species of halophilic sulfur-oxidizing bacteria. *Microorganisms* 10:995. doi: 10.3390/microorganisms10050995

O'Brien, P. A., Morrow, K. M., Willis, B. L., and Bourne, D. G. (2016). Implications of ocean acidification for marine microorganisms from the free-living to the host-associated. *Front. Mar. Sci.* 3:14. doi: 10.3389/fmars.2016.00047

Orr, J. C., Fabry, V. J., Aumont, O., Bopp, L., Doney, S. C., and Feely, R. A. (2005). Anthropogenic ocean acidification over the twenty-first century and its impact on calcifying organisms. *Nature* 437, 681–686. doi: 10.1038/nature04095

Pachauri, R. K. (2014). "Climate change 2014: synthesis report. contribution of working groups I, II and III to the fifth assessment report of the intergovernmental panel on climate change," in *Climate Change 2014: Synthesis Report*, eds The Core Writing Team, R. K. Pachauri, and L. Meyer (Geneva: IPCC).

Pajares, S., Bonilla-Rosso, G., Travisano, M., Eguiarte, L. E., and Souza, V. (2012). Mesocosms of aquatic bacterial communities from the cuatro cienegas basin (Mexico): a tool to test bacterial community response to environmental stress. *Microb. Ecol.* 64, 346–358. doi: 10.1007/s00248-012-0045-7

Pajares, S., Souza, V., and Eguiarte, L. E. (2015). Multivariate and phylogenetic analyses assessing the response of bacterial mat communities from an ancient oligotrophic aquatic ecosystem to different scenarios of long-term environmental disturbance. *PLoS One* 10:e0119741. doi: 10.1371/journal.pone.0119741

Papale, M., Rappazzo, A., Mikkonen, A., Rizzo, C., Moscheo, F., Conte, A., et al. (2020). Bacterial diversity in a dynamic and extreme sub-arctic watercourse (pasvik river, norwegian arctic). *Water* 12:3098.

Pascal, P.-Y., Dupuy, C., Richard, P., Mallet, C., telet, E. A., du, C., et al. (2009). Seasonal variation in consumption of benthic bacteria by meio- and macrofauna in an intertidal mudflat. *Limnol. Oceanogr.* 54, 1048–1059.

Peng, M., Zi, X., and Wang, Q. (2015). Bacterial community diversity of oil-contaminated soils assessed by high throughput sequencing of 16S rRNA genes. *Int. J. Environ. Res. Public. Health* 12, 12002–12015. doi: 10.3390/ijerph121012002

Petersen, J., Kennedy, V. S., Dennison, W., and Kemp, W. M. (2009). *Enclosed Experimental Ecosystems and Scale: Tools For Understanding and Managing Coastal Ecosystems*. Berlin: Springer.

Prieto-Barajas, C. M., Valencia-Cantero, E., and Santoyo, G. (2018). Microbial mat ecosystems: structure types, functional diversity, and biotechnological application. *Electron. J. Biotechnol.* 31, 48–56.

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

R Core Team (2020). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. Available online at: https://www.R-project.org/

Schmidtko, S., Stramma, L., and Visbeck, M. (2017). Decline in global oceanic oxygen content during the past five decades. *Nature* 542, 335–339. doi: 10.1038/nature21399

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

Sievert, S. M., Scott, K. M., Klotz, M. G., Chain, P. S. G., Hauser, L. J., and Hemp, J. (2008). Genome of the epsilonproteobacterial chemolithoautotroph sulfurimonas denitrificans. *Appl. Environ. Microbiol.* 74, 1145–1156. doi: 10.1128/AEM.01844-07

Simon, M., Scheuner, C., Meier-Kolthoff, J. P., Brinkhoff, T., Wagner-Döbler, I., Ulbrich, M., et al. (2017). Phylogenomics of Rhodobacteraceae reveals evolutionary adaptation to marine and non-marine habitats. *ISME J.* 11, 1483–1499. doi: 10.1038/ismej.2016.198

Stauffert, M., Cravo-Laureau, C., Jezequel, R., Barantal, S., Cuny, P., Gilbert, F., et al. (2013). Impact of oil on bacterial community structure in bioturbated sediments. *PLoS One* 8:e65347.

Stewart, R. I. A., Dossena, M., Bohan, D. A., Jeppesen, E., Kordas, R. L., and Ledger, M. E. (2013). "Chapter two - mesocosm experiments as a tool for ecological climate-change research," in *Advances in Ecological Research, Global Change in Multispecies Systems: Part 3*, eds G. Woodward and E. J. O'Gorman (Cambridge, MA: Academic Press).

Takai, K., Suzuki, M., Nakagawa, S., Miyazaki, M., Suzuki, Y., Inagaki, F., et al. (2006). Sulfurimonas paralvinellae sp. nov., a novel mesophilic, hydrogen- and sulfur-oxidizing chemolithoautotroph within the Epsilonproteobacteria isolated from a deep-sea hydrothermal vent polychaete nest, reclassification of Thiomicrospira denitrificans as Sulfurimonas denitrificans comb. nov. and emended description of the genus Sulfurimonas. *Int. J. Syst. Evol. Microbiol.* 56, 1725–1733. doi: 10.1099/ijs.0.64255-0

Tan, Y.-H., Lim, P.-E., Beardall, J., Poong, S.-W., and Phang, S.-M. (2019). A metabolomic approach to investigate effects of ocean acidification on a polar microalga Chlorella sp. *Aquat. Toxicol.* 217:105349. doi: 10.1016/j.aquatox.2019.105349

van Gemerden, H. (1993). Microbial mats: a joint venture. *Mar. Geol.* 113, 3–25.

Verleyen, E., Sabbe, K., Hodgson, D. A., Grubisic, S., Taton, A., Cousin, S., et al. (2010). Structuring effects of climate-related environmental factors on Antarctic microbial mat communities. *Aquat. Microb. Ecol.* 59, 11–24.

Widdicombe, S., Dashfield, S. L., McNeill, C. L., Needham, H. R., Beesley, A., McEvoy, A., et al. (2009). Effects of $CO_2$ induced seawater acidification on infaunal diversity and sediment nutrient fluxes. *Mar. Ecol. Prog. Ser.* 379, 59–75.

Yilmaz, P., Wegener Parfrey, L., Yarza, P., Gerken, J., Pruesse, E., Quast, C., et al. (2013). The SILVA and "all-species living tree project (LTP)" taxonomic frameworks. *Nucleic Acids Res.* 42, D643–D648. doi: 10.1093/nar/gkt 1209

Zeebe, R., and Wolf-Gladrow, D. (2001). *$CO_2$ in Seawater: Equilibrium, Kinetics, Isotopes, Elsevier Oceanography Book Series*. Amsterdam: Elsevier.

# 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