frontiers | Research Topics

Diversity, dynamics and function of host-associated microbiomes in marine organisms

Edited by
Yunyun Zhuang, Senjie Lin, Caiwen Li and Guangyi Fan

Published in
Frontiers in Marine Science
Frontiers in Microbiology

ISSN 1664-8714
ISBN 978-2-8325-4372-6
DOI 10.3389/978-2-8325-4372-6

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

Diversity, dynamics and function of host-associated microbiomes in marine organisms

Topic editors

Yunyun Zhuang — Ocean University of China, China
Senjie Lin — University of Connecticut, United States
Caiwen Li — Institute of Oceanology, Chinese Academy of Sciences (CAS), China
Guangyi Fan — Beijing Genomics Institute (BGI), China

Citation

Zhuang, Y., Lin, S., Li, C., Fan, G., eds. (2024). *Diversity, dynamics and function of host-associated microbiomes in marine organisms*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-4372-6

Table of contents

frontiers | Frontiers in Marine Science

TYPE Original Research
PUBLISHED 01 November 2022
DOI 10.3389/fmars.2022.1024117

OPEN ACCESS

EDITED BY
Yunyun Zhuang,
Ocean University of China, China

REVIEWED BY
Susan Smith,
Mystic Aquarium, United States
Jianwei Chen,
Beijing Genomics Institute (BGI), China

*CORRESPONDENCE
Jinsong Zheng
zhengjinsong@ihb.ac.cn

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

RECEIVED 21 August 2022
ACCEPTED 17 October 2022
PUBLISHED 01 November 2022

CITATION
Wan X, Li J, Tian R, McLaughlin RW, Hao Y, Wu J, Wang Z, Fan F, Wang D and Zheng J (2022) The effects of human care on the blowhole and gut microbiotas of two cohabiting dolphin species based on a year-round surveillance.
Front. Mar. Sci. 9:1024117.
doi: 10.3389/fmars.2022.1024117

The effects of human care on the blowhole and gut microbiotas of two cohabiting dolphin species based on a year-round surveillance

Xiaoling Wan[1], Jia Li[1], Renmao Tian[2], Richard William McLaughlin[3], Yujiang Hao[1,4], Junyi Wu[5], Zhengyang Wang[5], Fei Fan[1,4], Ding Wang[1,4] and Jinsong Zheng[1,4]*

[1]The Key Laboratory of Aquatic Biodiversity and Conservation of the Chinese Academy of Sciences, Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan, China, [2]Institute for Food Safety and Health, Illinois Institute of Technology, Bedford Park, IL, United States, [3]General Studies, Gateway Technical College, Kenosha, WI, United States, [4]National Aquatic Biological Resource Center, Wuhan, China, [5]Dolphinarium of Nanning Zoo, Nanning, China

Understanding the effects of human care on the dynamics of the host-associated microbiota is critical for the health management of dolphins living in an aquarium. Yet this aspect remains relatively unexplored. Here, by utilizing 16S rRNA gene sequencing, we profiled the blowhole and gut bacterial communities of two bottlenose dolphins (*Tursiops truncatus*) and a Chinese white dolphin (*Sousa chinensis*) reared in the same indoor pool, based on year-round surveillance. In addition, we compared these dolphin microbiotas with those previously published datasets from wild dolphins. Our results showed that both the blowhole and the gut of the two dolphin species under human care shared a more similar microbiome than members of the same dolphin species across different habitats (human care vs wild). However, the effects of human care on the dolphin microbiome from the two body sites varied. In the aquarium, bacterial alpha diversities differed significantly between the two body sites, and the seasonal stability of the bacterial community was more evident in the gut than in the blowhole. Additionally, the blowhole bacterial composition and the predicted functional capacity from the two dolphin species showed differences and were less convergent than their gut microbiota over a decade-long cohabitation. Further analyses showed that heterogeneous and homogeneous selections (i.e., deterministic processes) contributed more to the blowhole than to the gut bacterial communities, while a dispersal limitation (i.e., a stochastic process) was more important for the gut microbiota. The present study provides the first comparative evidence that the gut microbiota may be more plastic in response to the human care environment than the blowhole microbiota. This improves our understanding of dolphin health

management under human care and helps to predict their microbial responses to environmental changes.

KEYWORDS

dolphin, gut microbiota, blowhole microbiota, convergence, human care

Introduction

As top predators in marine ecosystems, cetaceans are critical sentinel and indicator species for ecosystem health and integrity, which has caught global attention (Parsons et al., 2015). However, due to the pervasiveness of anthropogenic pressures, at least a quarter of the world's cetaceans have been confirmed as endangered or even extinct (Turvey et al., 2007; Simmonds and Eliott, 2009). Human care has been put forward as one of the complementary approaches to *in situ* and natural *ex situ* conservation practices for some endangered small cetacean species (Wang et al., 2005; Curry et al., 2013; Muka and Zarpentine, 2021). Although the care of Dolphinarium animal husbandry specialists can facilitate the diagnosis and treatment of injured or sick individuals and preserve genetic diversity (Muka and Zarpentine, 2021), it also changes the animals' diet and restricts their behaviors, which may lead to physiological changes, disease susceptibility (e.g., gastrointestinal and respiratory diseases) and early mortality (Lott and Williamson, 2017; Nabi et al., 2020; Segawa et al., 2020). As a result, there is an urgent need to develop effective assessment tools that can improve cetaceans' long-term survival and welfare under human care.

In recent decades there has been a growing appreciation of the ability to alter the microbiome to positively affect the health of the animal which could aid in the conservation of threatened species (Jin et al., 2019; West et al., 2019). Microbes, especially those which colonize the gut and respiratory tract, are extremely important in maintaining the health of the host by affecting the immune defenses, digestion and nutrient absorption, behavior, reproduction, etc., (Vuong et al., 2017; Yadav et al., 2018; Invernizzi et al., 2020; Mallott et al., 2020; Bosco and Noti, 2021). There is strong support for systemic microbial communication between the microbiota found in the gastrointestinal and respiratory tracts (Enaud et al., 2020). Multiple lines of evidence show that the host-associated microbiota co-diversifies with their host lineages and can be further shaped by the external environment (e.g., dietary and habitat conditions) (Bik et al., 2016; Bai et al., 2021; Jin et al., 2021; Mallott and Amato, 2021). Compared to wild animals, animals under human care may undergo a series of changes that influence their microbiome, e.g., dietary restrictions, antibiotic interventions, and increased exposure to human-associated microbes (Hyde et al., 2016). The disturbance of a healthy microbiome may disrupt immune homeostasis and promote opportunistic bacterial pathogens (Toor et al., 2019; Kaul et al., 2020). Therefore, as recently proposed, integrating human care along with host microbial research is becoming more important in going forward for conservation programs (Dallas and Warne, 2022). Despite recent advancements, studies on the microbiome of cetaceans under human care remain limited.

Several recent studies have reported the influences of the human care environment on the cetacean microbiota. For example, Suzuki et al. (2019) found that the gut microbial alpha diversity and the dominant phyla of common bottlenose dolphins (*Tursiops truncatus*) differed in three aquaria. Suzuki et al. (2021) found that Indo-Pacific bottlenose dolphins (*T. aduncus*) under human care and wild environments harbored habitat-specific microbial groups and that several pathogenic bacterial genera were only detected in the dolphins under human care, which highlighted the importance of monitoring pathogenic bacteria in cetaceans under human care. Another study provided evidence that the blowhole microbiota of common bottlenose dolphins was individual-specific and stable over eight months (Vendl et al., 2021). Regardless of the potential importance of human care for the health of captive cetaceans, the seasonal stability, and the degree to which human care humanizes the microbiota of different cetacean taxa have not been explored.

For these reasons, by using 16S rRNA gene-based high-throughput sequencing, we characterized the taxonomic and potential functional profiles of the blowhole and fecal microbiota of two common bottlenose dolphins and a Chinese white dolphin (*Sousa chinensis*) reared in the same indoor pool, based on year-round surveillance. Firstly, we uncovered the seasonal dynamics of the blowhole and gut microbiota of both dolphin species. Then we examined whether and how the host microbiota (blowhole and gut) converged between these two cohabiting dolphin species. Lastly, we explored the effects of human care on the dolphins' microbiota through a comparison with their wild counterparts from the previously published datasets. In this study, we aim to enhance our knowledge about how human care influences the dolphin microbiota to facilitate their health monitoring in the future.

Materials and methods

Specimen collection

Two bottlenose dolphins (ID: BD1 and BD2) and one Chinese white dolphin (ID: CWD) housed in the Dolphinarium of Nanning Zoo, China, were included in this study (Table S1). These three dolphins were rescued from the wild and have lived in the Dolphinarium since 2007 under the same housing conditions (i.e., the same diet and aquarium parameters). There were no relations between these animals. They were able to breathe out and defecate promptly on command, allowing for a biological sample (exhaled blow or feces) to be gathered in no more than 10 min in a non-invasive way. Sample collection was generally performed at the end of the month around 4 pm during the whole course of the study. To eliminate potential contamination by pool water and skin, the skin regions near the blowhole and the anus were wiped gently using alcohol pads before sampling. Specifically, feces were collected using a sterilized plastic spoon immediately after defecating and were then transferred to a 5-mL sterile Eppendorf tube. At least 5 g of feces was collected for each dolphin at each sampling time point. For blow samples, a 50-mL sterile Eppendorf tube was placed 2-3 cm above the blowhole, and 2-4 exhalations from the same individual were collected per tube. In total two tubes were used to collect the blow from one individual. After fecal and blowhole sampling each month, a total of the 5-L pool water was gathered and filtered through a mixed cellulose esters filter paper (47-um pore diameter and 0.45-um pore size). The filters were stored with 95% ethanol until DNA extraction. Feces, blow, and water samples were collected monthly over a period of 12 months from Dec. 2015 to Nov. 2016, except for the Mar. 2016 sample. Sample details are shown in Table S2. Monthly mean water temperatures were recorded by the dolphin trainers (Figure S1). All samples were preserved at -80 °C until DNA extraction.

One week before each sampling event, the blood biochemical and blood routine tests from three dolphins were evaluated by veterinarians. The test results were within the normal range and all dolphins were considered healthy during all sampling events. No antibiotics were given to dolphins two weeks before each sampling event. The sampling procedures used in this study were approved by the Regulations of the People's Republic of China for the Implementation of Wild Aquatic Animal Protection (promulgated in 1993), adhering to all ethical guidelines and legal requirements in China.

DNA extraction, high-throughput sequencing, and sequence processing

The ZR Faecal DNA Kit (Zymo Research Incorporated, CA, USA) was applied to extract the metagenomic DNA from frozen samples. Universal primers 341F-806R were then used to amplify the V3-V4 regions of bacterial 16S rRNA genes for library constructions (Roggenbuck et al., 2014). PCR products of different samples were pooled and run on the Illumina HiSeq 2500 platform (2 × 250 bp paired ends).

The raw sequencing reads were analyzed *via* an internal pipeline Amplicon Sequencing Analysis Pipeline (Wan et al., 2018). In brief, after a quality check of raw sequences using FastQC, sequences with high quality were then trimmed to remove primer sequences (Brown et al., 2017). Clean reads were then processed to generate operational taxonomic units (OTUs) by UPARSE at the 97% cutoff (Edgar, 2013). The representative sequence of each OTU was classified into taxa with a confidence threshold of 0.8 on the RDP classifier using the SILVA database (Cole et al., 2014). FastTree was applied to generate phylogenetic trees from all representative sequences (Price et al., 2009). To enable comparisons with previously-published datasets of the bottlenose dolphin and the Chinese white dolphin living in other habitats, we downloaded the respiratory (PRJNA279427) and rectal sequences (PRJNA174530) of bottlenose dolphins inhabiting San Diego Bay and Sarasota Bay, USA (Bik et al., 2016) and intestinal sequences (PRJNA610950) of the wild Chinese white dolphin stranding in Guangdong Province, China (Wan et al., 2021). The above steps were repeated for sequencing processing for all downloaded 16S rRNA gene sequences.

Functional predictions

PICRUST version 2.0 (Douglas et al., 2020) was applied to predict functions based on the taxonomy affiliations from the 16S rRNA gene sequences. To obtain the potential pathogens from 16S rRNA gene sequences, the Functional Annotation of PROkaryotic TAXa (FAPROTAX) database was further used to sort out potentially pathogenic groups, including 'animal parasites or symbionts', 'intracellular parasites', 'invertebrate parasites', 'human pathogens all', 'human pathogens diarrhea', 'human pathogens gastroenteritis', 'human pathogens meningitis', 'human pathogens nosocomia', 'human pathogens pneumonia', 'human pathogens septicemia', 'plant pathogen' and 'fish parasites', which has been experimentally verified to threaten the welfare of humans, aquatic animals or plants (Louca et al., 2016; Zhou et al., 2020).

Data analysis

To eliminate potential contamination originating from the water samples, we identified contaminated OTUs using the R package {decontam} following the same parameters outlined in the study by Vendl et al., 2021. In total, we discarded 339 OTUs

from the dolphin microbiota which likely originated from the pool water (Figure S2). We then removed seven blowhole samples and two fecal samples from the dataset, as the number of reads was extremely low (< 6,000) (Table S2). The resultant OTU abundance tables from these remaining samples were rarefied to the same number (at the minimum sequence number of 7,551) of sequences per sample to ensure equal sampling depth using the *rrarefy* function (R package {vegan}) (Weiss et al., 2017).

Microbial α-diversity indices, including taxonomic (richness) and phylogenetic (Faith's PD) diversities, were computed (R package {Picante}) (Kembel et al., 2010). To evaluate whether both α-diversity indices were correlated with the sampling month for each dolphin, we fitted linear mixed effects models (Lindstrom and Bates, 1990) with a fixed effect of 'sampling month' and random effects of 'dolphin ID'. The α-diversity indices between groups were further tested by applying the Wilcoxon rank sum test. To assess the difference in the microbial communities between the groups, non-parametric permutation tests, including the multiple response permutation procedure (MRPP), analysis of similarity (ANOSIM), and permutational multivariate analysis of variance (ADONIS), were computed, based on Bray-Curtis and weighted UniFrac distances. Non-metric multidimensional scaling (NMDS) analyses were further calculated to visualize patterns of community dissimilarity across samples. The LEfSe (Linear discriminant analysis Effect Size) was applied to identify biomarker functional potentials between groups ($P < 0.05$, LDA score > 3.0) *via* the website http://huttenhower.sph.harvard.edu/galaxy/.

The microbial community assembly processes were initially quantified using the null model approach (Stegen et al., 2013). Bray-Curtis-based Raup-Crick (RCBray) and β-mean nearest taxon distance (β-NTI) values of each pairwise comparison between two samples were calculated using the iCAMP (Ning et al., 2020). A β-NTI value < -2 or > 2 indicates that homogeneous selection or variable selection (deterministic processes) is predominant in shaping the community composition, whereas $|\beta\text{-NTI}| < 2$ indicates that neutral assembly processes account for the communities. Further, RCBray was used to partition the stochastic ecological processes into dispersal limitation, undominated and homogenizing dispersal at the cutoff of ± 0.95 (Stegen et al., 2013).

Results

Sequencing data overview

For all 80 samples, a total of 2,831,917 sequences were obtained, with an average of 35,398 (ranging from 19,415 to 44,181) sequences per sample, resulting in 2,714 OTUs. After removing 339 contaminated OTUs that were likely from the pool water (Figure S2), as well as nine dolphin samples with low reads number of less than 6,000 (Table S2), 2,539 OTUs were obtained from 71 samples in total. The rarefaction curves of the observed OTUs suggested that the diversity of blow, fecal, and pool water microbiota sampled in this study were well represented after resampling at the minimum sequence number of 7,551 (Figure S3).

Dolphin microbiota between the two body sites differed and yet were distinct from the pool water

Based on the taxonomic assignment, we found that the blow, feces, and pool water samples exhibited unique microbial profiles, respectively. Proteobacteria (~51.17%) and Bacteroidetes (~32.20%) constitute the main component (~83.37%) of the dolphin blow microbiota, whereas the dolphin fecal microbiota was dominated by Firmicutes (~73.99%) and a lesser degree Proteobacteria (~16.57%) (Figure 1A). Pool water samples were overwhelmingly dominated by Proteobacteria (~96.21%) (Figure 1A). The blow microbiota was mainly composed of *Tenacibaculum* (~6.35%) and *Arcobacter* (~4.35%) (Figure 1B). The fecal bacteria were dominated by *Romboutsia* (~19.28%), *Clostridium* XI (~16.22%), *Clostridium sensu stricto* (~12.82%), and unclassified Peptostreptococcaceae (~10.03%) (Figure 1B). In the pool water, unclassified Alphaproteobacteria (~50.05%), *Halomonas* (~23.94%), and unclassified Hyphomonadaceae (~18.32%) were the most prevalent (Figure 1B). Non-metric multidimensional scaling (NMDS) analyses further demonstrated that, for each dolphin species, the blow and feces contained disparate bacterial communities (Adonis: $P < 0.001$), which were well separated from the pool water community (Adonis: $P < 0.001$) (Figures 1C, D). As for α-diversity parameters (richness and PD), even though there was no significant inter-individual differentiation (Figures S4A, B), both α-diversity indices of the blow bacteria were significantly higher than either fecal or pool water bacteria (Figures S4C, D, $P < 0.05$). Nevertheless, fecal and pool water bacterial diversities were comparable (Figures S4C, D, $P > 0.05$).

Temporal fluctuation in the blow and fecal microbiota

Neither blow nor fecal α-diversity indices (richness and PD) differed significantly over time (Table S3, $P > 0.05$). In general, for each dolphin, the microbial diversity indices of fecal samples over one year were more variable than the blow samples, as revealed by CV coefficient of variation (CV) values (Table 1). Only the blow PD of dolphin BD2 showed a higher CV value

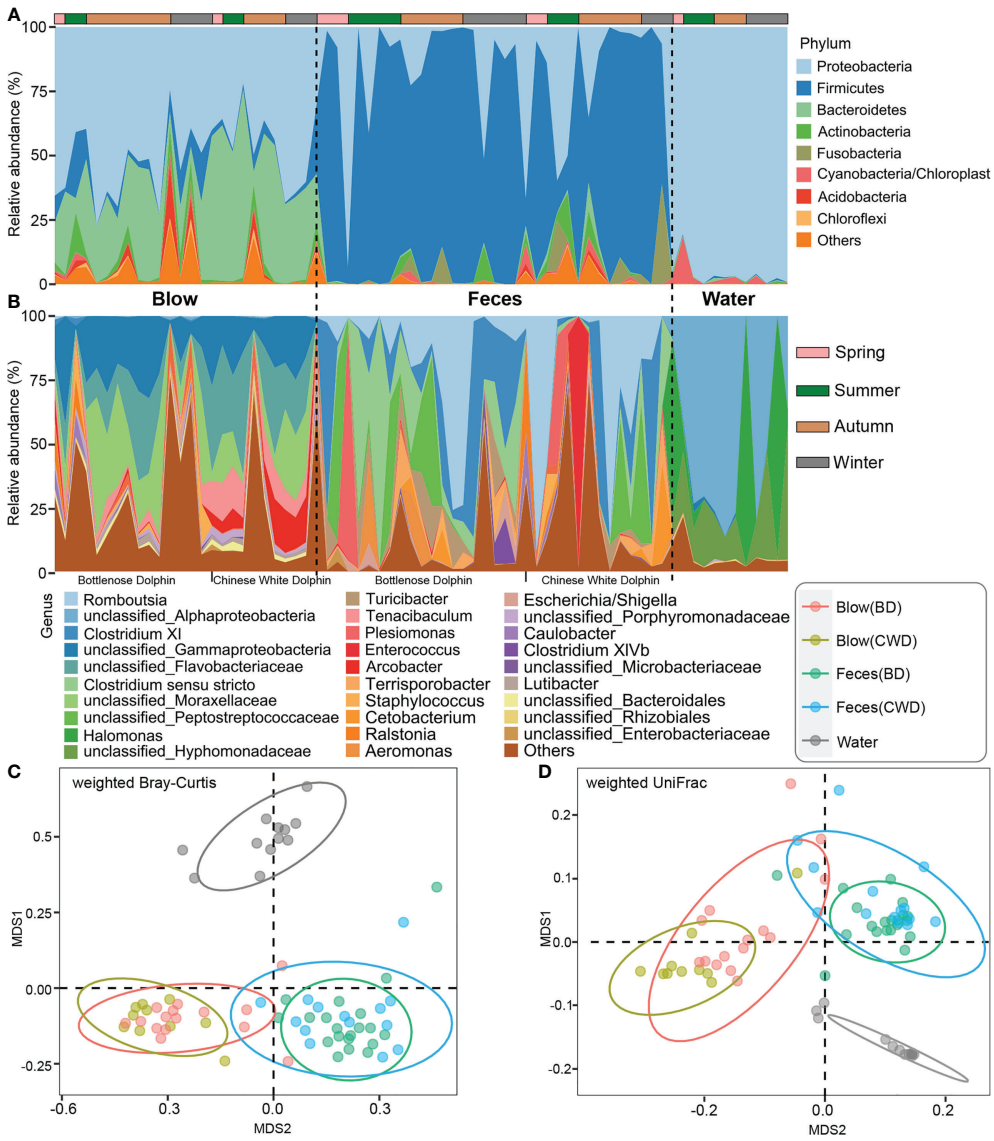


FIGURE 1
Composition and comparison of microbiota from the dolphin blow, feces, and pool water. **(A)**, the relative abundance of each phylum. A total of 28 phyla were detected. Only the top 8 most abundant phyla are shown here; the remaining 20 phyla as well as unclassifiable sequences were grouped as 'Others'. **(B)**, the relative abundance of each genus. A total of 571 genera, including unclassified genera, were detected. Only the top 29 most abundant genera are shown here; the remaining genera including unclassifiable sequences were grouped as 'Others'. The horizontal lines at the top indicate the collecting season of each specimen. **(C)**, the non-metric multidimensional scaling (NMDS) analysis of different sampling groups at the OTU level based on weighted Bray-Curtis distances. **(D)**, the NMDS analysis of different sampling groups at the OTU level based on weighted UniFrac distances. BD indicates the bottlenose dolphin, and CWD indicates the Chinese white dolphin.

than that in the feces (Table 1). No significant seasonal variation in the blow or fecal microbial composition at the OTU level was detected among the four seasonal groups (winter, spring, summer, and autumn) (Adonis, $P > 0.05$; Table S4) or two seasonal groups (winter-spring and summer-autumn) (Adonis, $P > 0.05$; Table S5). Interestingly, when samples were divided into two temperature groups (high-temperature and low-temperature groups) based on the monthly mean pool water temperature, feces rather than blow samples, exhibited significant microbial composition shifts (Adonis, $P < 0.05$; Table S6), signifying that fecal microbial composition is more sensitive to water temperature compared to the blow microbiota.

We further named the OTUs that were detected in each dolphin across 80% of all sampling events as the "intra-core" (i.e., temporal core). A minimal part (0.07-13.17%) of the total OTUs were detected to constitute a dolphin's temporal core, which

TABLE 1 The coefficient of variation of microbial diversity indices over one year in each dolphin.

Dolphin ID	Richness		PD	
	Blow	Feces	Blow	Feces
BD1	0.853	0.915	0.556	0.745
BD2	0.797	0.815	0.749	0.684
CWD	0.323	0.769	0.315	0.844

Richness: the number of observed OTUs; PD, phylogenetic diversity. BD indicates the bottlenose dolphin, and CWD indicates the Chinese white dolphin.

accounted for a relatively high proportion (54.35-86.55%) of the total reads of each dolphin, except for the blow dataset of the bottlenose dolphin BD1 (0.84%) (Table S7). In total, 70 intra-core OTUs were shared by both blow and fecal samples, accounting for on average 59.10% of all the dolphin microbial reads. The blow-associated intra-core of the Chinese white dolphin harbored a higher proportion of Bacteroidetes (43.91%) compared to bottlenose dolphins (21.17%) (Figure 2A). As for the fecal intra-cores, Firmicutes (63.02%) was most prevalent, followed by Proteobacteria (6.46%) (Figure 2A). A higher proportion of intra-core bacteria in the blow (59.79%) compared to the feces (10.97%) were novel and could therefore not be assigned a genus (Figure 2B). Among the classifiable genera, *Tenacibaculum* (7.05%), *Arcobacter* (5.15%), and *Staphylococcus* (1.36%) were abundant in the blow samples (Figure 2B). Whereas, *Romboutsia* (20.20%), *Clostridium* XI (15.72%), *Clostridium sensu stricto* (5.48%), *Terrisporobacter* (2.76%), *Plesiomonas* (4.63%), *Turicibacter* (4.29%) and *Escherichia/Shigella* (1.23%) were prevalent in the feces (Figure 2B).

Human care-driven gut microbiota compositional convergence in both dolphin species

The hierarchical clustering analysis showed that the fecal microbiota from both dolphin species clustered closer to each other, whereas the blow microbiota differed substantially between species (Figure 3). This pattern is further supported by three different non-parametric multivariate statistical tests (Table 2). Then we identified OTUs of blow microbiota which was able to discriminate dolphin species classification, with an overall accuracy of 96% (random forest models). Interestingly, the top 30 most discriminatory OTUs (1.72% of the overall blow OTUs) accounted for an average of 30.06% of the relative read abundance of the blow microbiota (Figure S5 and Table S8). Among them, nine OTUs were assigned at the genus level, with *Tenacibaculum* (4.35%), *Staphylococcus* (1.29%), and *Ralstonia* (1.24%) at a higher proportion (Table S8).

By quantifying deviations in the microbial phylogenetic turnover, we found that both blow and fecal bacterial communities were primarily governed by neutral processes (dispersal limitation and drift) (Figure 4). However, feces (-0.49) exhibited significantly lower β-NTI than the blow (0.056) (Figure 4A). Specifically, deterministic assembly (heterogeneous and homogeneous selections, 2.54% and 22.87%, respectively) showed greater importance to the blow (0.46% and 8.36%, respectively), while the stochastic process was higher in the fecal bacterial community (dispersal limitation: 38.70%) than the blow bacterial community (dispersal limitation: 30.55%) (Figure 4B).

Human care-driven gut microbiota functional convergence in both dolphin species

Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis showed that a total of 297 pathways were detected across all dolphin samples. Based on functional profiles, the blow microbiota was distinct from the feces (Adonis: $P < 0.001$), having a significantly higher relative abundance of *Vibrio cholerae* infection, biofilm formation-*Vibrio cholerae*, epithelial cell signaling in *Helicobacter pylori* infection, *Salmonella* infection,

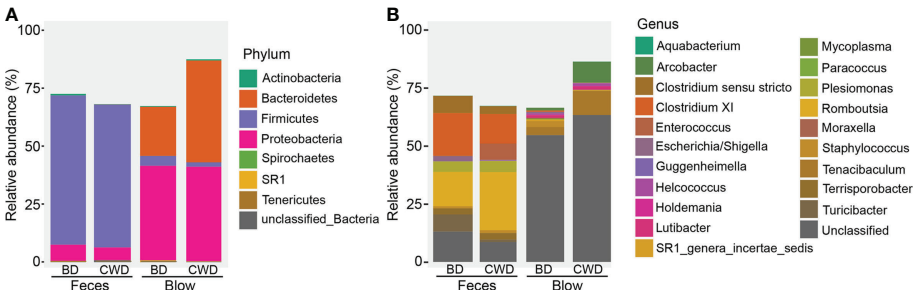


FIGURE 2
Taxonomic composition of 70 intra-cores detected from each dolphin. **(A)**, Phylum level. **(B)**, Genus level. BD indicates the bottlenose dolphin, and CWD indicates the Chinese white dolphin.

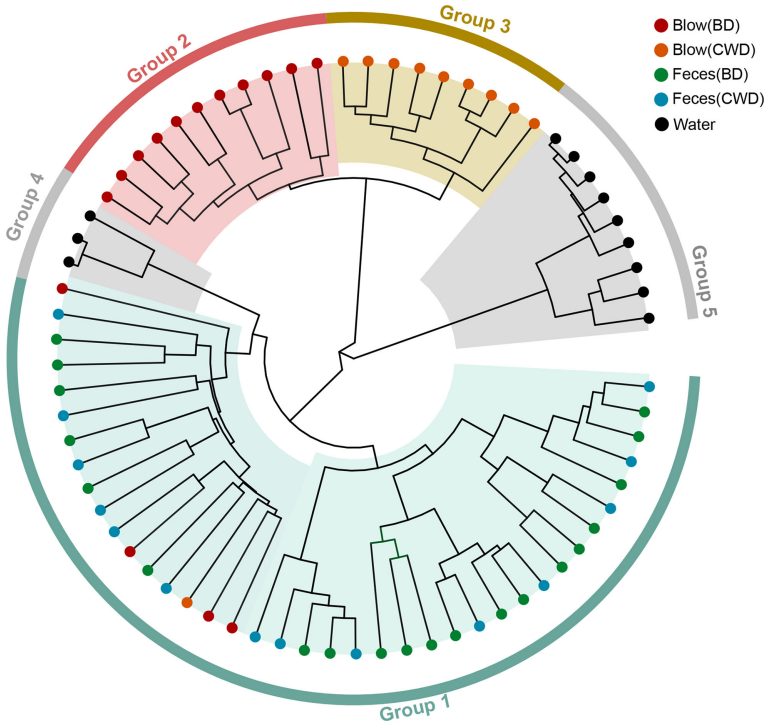


FIGURE 3
The hierarchical clustering analysis of all samples. Different colors of the outside ring represent that all samples were grouped into five clusters (group 1, group 2, group 3, group 4 and group 5) using unweighted pair group method with arithmetic mean (UPGMA) based on the microbial taxonomic composition at the OTU level. The terminal node of each branch corresponded to the individual sample. Colors of the nodes reflected different sampling groups in the legend. BD indicates the bottlenose dolphin, and CWD indicates the Chinese white dolphin.

Influenza A, lipid metabolism, etc. (Figures 5A, B). Whereas, pathways such as bacterial invasion of epithelial cells, *Staphylococcus aureus* infection, amino acid metabolism, and energy metabolism were enriched in the feces (Figure 5A).

Consistent with the taxonomic profiles, the predicted functional composition of fecal microbiomes from both dolphin species was similar, while the predicted functional composition of the blow microbiomes from the two dolphin species was discriminated against each other significantly (Figures 6A, B; Table S9, $P < 0.05$). A LEfSe analysis was used to further characterize the significant pathways in the blow samples between the two dolphin species. The results showed that quorum sensing, secretion system, two-component system, and ABC transporters were abundant in the bottlenose dolphin, while ribosome was enriched in the Chinese white dolphin (Figure 6C).

Comparing the microbiota of dolphins under human care to those of wild animals from published datasets

To further evaluate the effects of the human care environment on the dolphin microbiota, blow and fecal microbiota data derived

TABLE 2 Significance tests on the effects of species on the respiratory and intestinal microbial community composition.

Sample type	Distance	MRPP		ANOSIM		ADONIS	
		δ	*P*	R	*p*	F	*p*
Blow	Bray-Curtis	0.612	**0.001**	0.178	**0.019**	3.563	**0.001**
	Weighted Unifrac	0.224	**0.007**	0.140	**0.048**	3.688	**0.014**
Feces	Bray-Curtis	0.821	0.618	0.123	**0.021**	1.621	0.064
	Weighted Unifrac	0.222	0.170	0.088	0.065	1.435	0.209

Three different permutation tests were performed, including the multiple response permutation procedure (MRPP), analysis of similarity (ANOSIM), and permutational multivariate analysis of variance (ADONIS), calculated with Bray-Curtis and weighted Unifrac distances. Bold values indicate $p < 0.05$.

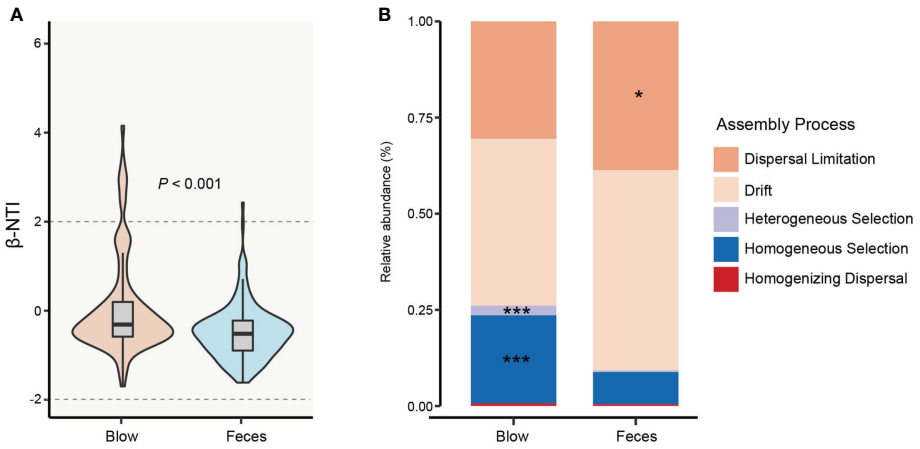


FIGURE 4
Bacterial community assembly processes for blow and feces in both dolphin species. **(A)**, the distribution of the bacterial weighted beta nearest taxon index (β-NTI) in the blow and feces. The *P* value is from Wilcox rank sum test. **(B)**, the relative abundance of different assembly processes in the blow and feces. Asterisks represent statistically greater proportions at the following significance levels: * $P < 0.05$, *** $P < 0.001$.

from this present study were compared with published datasets found in the studies (Bik et al., 2016; Wan et al., 2021). Our analyses revealed that the microbial composition of dolphin species under human care tended to be separate from their wild counterparts as determined by NMDS clustering analysis (Figure S6). Further statistical tests corroborated the NMDS results. Specifically, for both the blow and feces, the community dissimilarity, as revealed by four different distance matrices, between both captive dolphin species was significantly lower than that between dolphins under human care and their wild conspecifics, except the blow samples as calculated by the weighted UniFrac distance matrix (Figure 7).

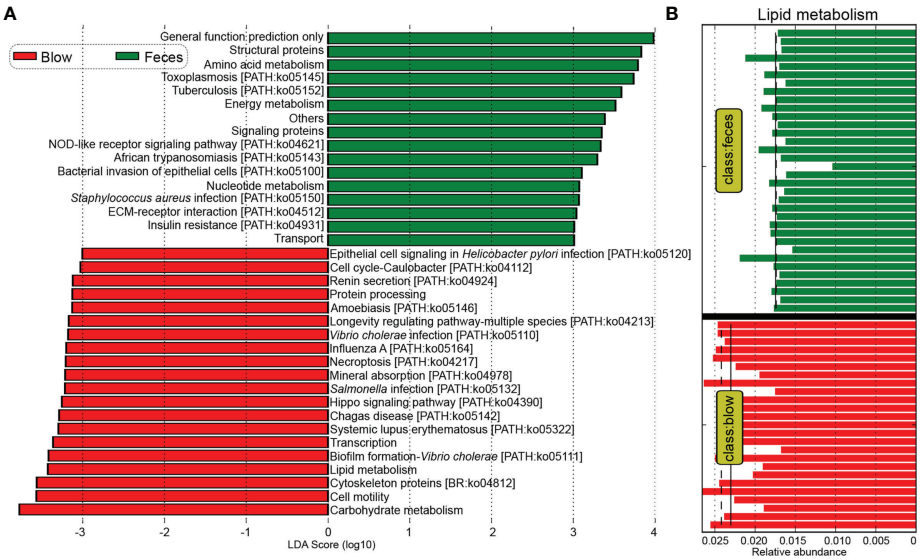


FIGURE 5
Comparative analyses of potential functions between the blow and fecal microbiomes. **(A)**, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways enriched in the blowhole and fecal microbiomes by the Linear discriminant analysis Effect Size (LEfSe) analysis. **(B)**, the lipid metabolism pathway was significantly higher in the blow than in the fecal microbiome. The mean and median relative abundance are indicated with solid and dashed lines, respectively.

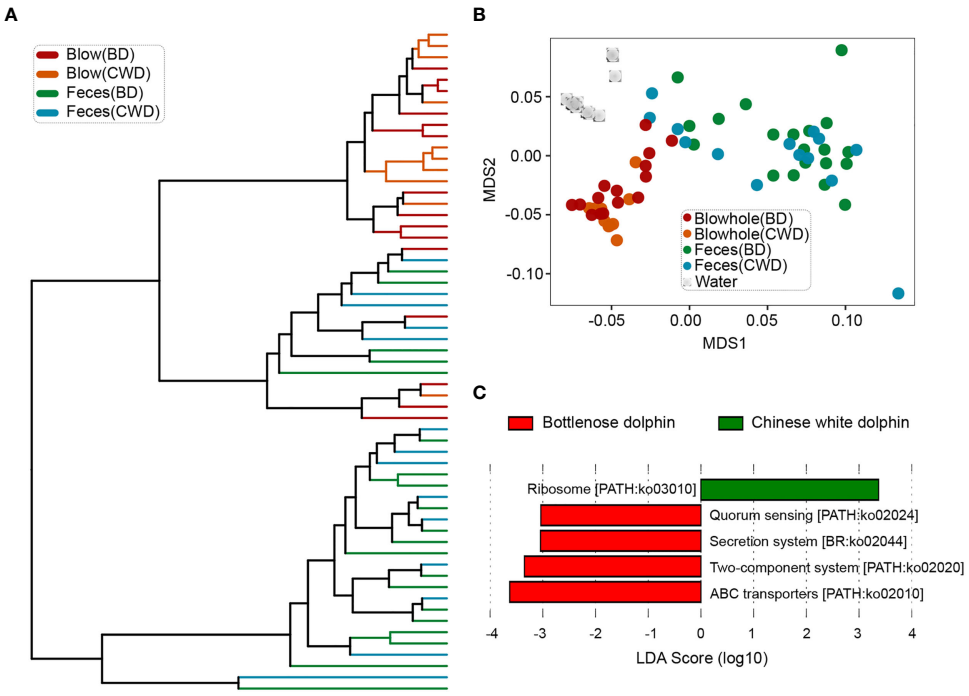


FIGURE 6
Convergence of the predicted functional composition of the dolphin microbiomes. **(A)**, the hierarchy clustering analysis of the dolphin microbial functional composition. **(B)**, the non-metric dimensional scaling (NMDS) analysis of microbial functional compositions based on the Bray-Curtis distances. **(C)**, pathways that were over-represented for both dolphin species identified by the Linear discriminant analysis Effect Size (LEfSe) analysis. BD indicates the bottlenose dolphin, and CWD indicates the Chinese white dolphin.

Potential bacterial pathogens detected in the dolphins under human care

Among all the OTUs obtained from our current study, 57 OTUs (2,579 sequences) were assigned to three potential pathogenic groups (animal parasites or symbionts, intracellular parasites, and human pathogens all) based on the FAPROTAX database (Figure 8). This only accounted for an average of 0.48% of the total abundance. These potential pathogenic OTUs included species from *Prevotella*, *Chlamydiales*, *Roseburia*, *Roseomonas*, *Campylobacter*, *Stenotrophomonas*, *Moraxella*, *Legionella*, *Edwardsiella*, *Proteus*, *Aquicella*, Coxiellaceae and Pasteurellaceae (Figure 8). *Prevotella* and *Stenotrophomonas* were widespread in the blow, feces, and pool water samples, with a higher relative abundance in the blow. *Moraxella* was only detected in the blow, and *Coxiella* was only detected in the feces (Figure 8).

Discussion

Although keeping cetaceans in human care remains controversial, it cannot be denied that populations under human care are ideal study subjects that can be used to address some biological questions which will ultimately contribute to species conservation efforts. Recent research has revealed the influences of the human care environment on the mammalian gut microbiome (McKenzie et al., 2017; Frankel et al., 2019). However, how the human care environment impacts the microbial communities of different dolphin species is yet to be assessed. In this study, utilizing a one-year time series of 16S rRNA gene sequencing data, we compared the blow and fecal microbial composition and functional potentials of two cohabiting dolphin species (two bottlenose dolphins and one Chinese white dolphin) in an indoor pool. We found that although their blow and fecal bacterial communities stayed stable over one year, the blow diversity was less variable and its microbial composition was less sensitive to pool water temperature compared to the feces. Moreover, the blow microbial composition and functional potentials from two dolphin species under human care were species-specific and less convergent than their fecal microbiota over a decade-long cohabitation. This result indicates that bacteria from the interior respiratory tract are likely to be less alterable by human care, whereas bacteria from the intestinal tract are prone to be more plastic in response to the human care environment.

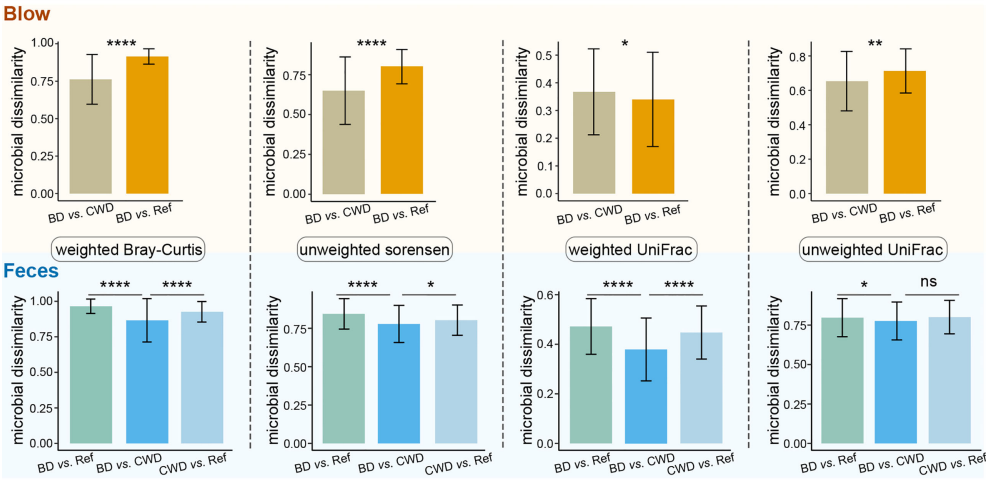


FIGURE 7
Information on within-group community dissimilarities for blow and fecal microbiotas. In these panels, higher microbial dissimilarity values indicate that these microbial samples are more dissimilar in composition. BD vs. CWD, the dissimilarity between bottlenose dolphins (BD) and the Chinese white dolphin (CWD). BD vs. Ref, the dissimilarity between bottlenose dolphins (BD) and the referred published datasets of their conspecifics. CWD vs. Ref, the dissimilarity between the Chinese white dolphin (CWD)and the referred published datasets of their conspecifics. Statistical significance is based on Wilcoxon rank sum test; **** $P < 0.0001$, ** $P < 0.01$, * $P < 0.05$. ns, not significant.

The bacterial community composition in the blow and fecal samples were distinct (Figure 1). For example, *Romboutsia*, which mainly has an intestinal origin (Gerritsen et al., 2019), was detected abundantly (~19.28%) in the feces, rather than in the blow, from all sampled dolphins (Figure 1B). This abundant genus may play essential roles in dolphins' intestinal nutrient absorptions as it covers a variety of metabolic capabilities regarding carbohydrate utilization, fermentation of single amino acids, anaerobic respiration, and metabolic end products (Gerritsen et al., 2019). Similar to the previous findings of the blow from free-ranging bottlenose dolphins (Robles-Malagamba et al., 2020), *Tenacibaculum* (~6.35%) and *Arcobacter* (~4.35%) (Figure 1B) also dominated in the blow specimens in the present study. Thus, both genera may be commensal residents of the respiratory tract of dolphins. However, these genera also contained several pathogenic species that widely infect both fish and invertebrates (Habib et al., 2014; Ramees et al., 2017). The specific roles that these species play in the respiratory tract need further experimental validation.

The respiratory tract and gut represent two major protective barriers against the external environment in mammals and could

FIGURE 8
Taxonomic and functional information of potential pathogenic taxa. 'Microbial taxa' indicates the classification details of each potential pathogenic taxa. UC denotes 'unclassified'. The heatmap in the middle indicates the log-transformed and column-scaled abundance data of potential pathogenic taxa present in five different sampling groups. BD, bottlenose dolphins; CWD, the Chinese white dolphin. Decimals indicate the relative abundance (%) of each pathogenic taxa, and the empty value indicates zero. 'No. of OTUs' indicates the number of OTUs detected from each potential pathogenic taxon. 'Pathogen group' indicates potentially pathogenic group of the pathogenic taxa predicted based on the FAPROTAX database. AnPS, animal parasites or symbionts; IntP, intracellular parasites; HuPA, human pathogens all.

get infected easily. Despite the speculated significance of dysbacteriosis in respiratory diseases, there are relatively few studies on the cetacean respiratory microbiome and it has been less explored than their gut microbiome. The respiratory tract, historically considered sterile in healthy humans, harbors diverse communities of microbes (Dickson et al., 2016). As cetaceans surface from the water through a blowhole, fresh DNA, lipids, proteins, bacteria, and viruses are forcefully exhaled out from their lungs with their secretions (Foot et al., 2006). Therefore, unlike gut microbes which are constantly receiving exogenous dietary ingredients, microbiota associated with the interior respiratory tract are prone to receive less disturbance from the external environment and may remain relatively unchanged once colonized. This may at least partly explain why the blow microbiotas of the three dolphins under human care showed relatively higher seasonal constancy, during the one year, and less convergence between dolphin species than the fecal microbiome in this study. However, when compared with their wild conspecifics, both the blow and fecal microbiome of the two dolphin species shared a more similar microbiome than members of the same dolphin species across different habitats (Figures 7 and S6), signifying that the human care environment may exert strong effects on structuring dolphin microbiotas distinct from the wild.

Shared bacterial communities between individuals often occur from vertical and/or horizontal transmission (Inoue and Ushida, 2003). In the present study, horizontal transmission resulting from shared environments and socialization between individuals may account for some of the shared bacteria observed between the two dolphin species under human care. To be specific, they were fed with the same thawed fish that had been washed with saline and they swam near one another. It's likely that in the process they transferred microbes. The prevalent influence of the human care environment suggests that environmental microorganisms or potential pathogens might colonize a dolphin easily. In this study, a considerable number (4/7) of potentially pathogenic members detected in the pool water also were found in the feces (Figure 8). It's likely that the potentially fecal pathogens derived from the adjacent pool water, or that the dolphins defecated potential pathogens into the pool water. Therefore, ensuring water quality and reducing or eliminating potential pathogens in the aquarium should aid in maintaining the gut health of dolphins in the human care environment. The dolphins in this study lived in a relatively constant indoor condition, where animals are not exposed to a diurnal cycle in water temperature, nor is there the same level of habitat complexity that is found in the natural environment. For example, the annual water temperature fluctuation amplitude in the aquarium was artificially regulated to approximately 6.50°C, with the minimum monthly mean water temperature at 21.79°C in February and the maximum at 28.29°C in July (Figure S1). The pool water was periodically disinfected with chlorine. Therefore, as revealed in our study, the seasonal dynamics of the fecal bacterial composition from dolphins under human care were likely driven by pool water temperature (Table S6). It is well known that temperature is a crucial abiotic factor structuring the composition and function of animal gut microbiomes (Sepulveda and Moeller, 2020). Our results are likely to confirm previous findings that gut microbiota may assist hosts in coping with seasonal thermoregulatory demands (Baniel et al., 2021). Since gut microbiota is critical in regulating many hosts' physiological processes, such as metabolic and immune system functions (Kau et al., 2011; Frame et al., 2020), as well as their adaptive evolution (Moeller et al., 2019), gut microbial dysbiosis may provide a mechanism for temperature variation influences on animal fitness in endangered populations (Sepulveda and Moeller, 2020). Since fecal bacteria are sensitive to low-level water temperature changes, even in controlled indoor conditions, it is speculated that dolphin gut microbiome could be disrupted rapidly by acute temperature changes. On the other hand, hosts can reshape their gut microbial communities to compensate for the environmental effects (e.g., temperature), resulting in gut microbial resilience to environmental change (Kokou et al., 2018). For example, in COVID-19 patients, the gut microbiota restored more rapidly with increased bacterial diversity and enrichment of beneficial commensal than the respiratory microbiome, and the former then promoted the restoration of the latter *via* bacterial cross-talk and resulted in a synchronous restoration of both organs (Xu et al., 2021). Thus, maintaining intestinal flora homeostasis is highly significant for both intestinal and respiratory homeostatic balance and health.

Respiratory illnesses, especially those caused by bacteria, are considered a threat to cetaceans in both human care and wild populations (Acevedo-Whitehouse et al., 2010; Waltzek et al., 2012). If non-invasive blow sampling can aid in the respiratory health monitoring of cetaceans has been one of the major concerns for conservation biologists. In agreement with previous observations (Nelson et al., 2019; Vendl et al., 2021), our findings supported that blow sampling of captive dolphins can be used for their health monitoring in conservation practices. Firstly, the dolphin blow displayed more highly diverse bacterial communities compared to the feces and pool water samples, even if the potentially contaminated reads originating from the pool water were removed (Figure S1). Furthermore, the blow bacterial composition and functional potentials were distinct from the feces and pool water (Figure 1). Therefore, dolphin blow samples may mirror the pulmonary microbial communities. However, a handful of studies showed potential water-borne microbes contaminated in the blow (Nelson et al., 2019). This may be overcome by first wiping the blowhole, using alcohol pads, before sampling the dolphins living under human care or collecting biological replicates. Secondly, the blow bacterial community stayed relatively stable over the sampling period of one year. Species-specific bacteria were detected in each dolphin species even with an over decade-long cohabitation in which potential horizontal

transmission could have occurred. In addition, the dolphin blows maintained highly abundant intra-cores (Table S7). This longitudinal observation provides additional evidence that the cetacean blow microbiota may be persistent and relatively independent of environmental influences (i.e., seasonal fluctuations and host phylogeny) in the human care environment. Lastly, we observed a small portion (0.94%) of potential bacterial pathogens in the blow, including members of the genera *Prevotella*, *Roseburia*, *Roseomonas*, *Stenotrophomonas*, *Moraxella*, *Legionella*, and *Proteus* (Figure 8). All of these are known potential pathogens of both animals and humans (Louca et al., 2016). Since we find no obvious signs of illness during the sampling, the presence of potential pathogens in the blow may be commensals to the dolphin lungs. However, this result showed that blow samples could be a method used to detect putative pathogenic bacteria in the respiratory tract of dolphins.

This study casts new light on explaining the differences between the blow and fecal bacterial communities by calculating their community assembly processes using the null model framework, which is widely acknowledged in microbial ecology (Stegen et al., 2013; Xie et al., 2022). The importance of homogeneous selection in structuring blow bacterial communities over one year is consistent with the strong relationships between bacterial communities and host species and individuality (e.g., host species-specificity and individual-specificity) (Lima et al., 2012; Vendl et al., 2021). In contrast, we observed that stochastic processes contributed more to the fecal community structure, primarily due to the drift and dispersal limitation processes (Figure 4), which may indicate that fecal bacteria exchanged between communities and fluctuated due to chance events in the dolphins under human care. These differences in community assembly processes likely reflect different transmission routes of the blow and fecal bacteria. For example, as observed in the present study, compared to the blow bacteria, the fecal bacteria tended to be more horizontally transferred between animals in the human care facility. It's important to notice that the differences between reared dolphins and their wild counterparts from previous studies may be due to different experimental methods (including 16S amplicon region, PCR process, and sequencing platform), which may not be as accurate as sampling and sequencing both reared and wild dolphins under the same experiment condition.

In conclusion, this study demonstrated that human care exerted different impacts on the blow and gut bacterial communities of two dolphin species over a decade-long cohabitation. Typically, compared to the blow bacteria, the fecal bacteria between the two-dolphin species were more variable over one year, more convergent between dolphin species, and more sensitive to the pool water temperature. These results improve our understanding of how the human care environment alters the dolphin microbiome from different body sites, which provides an empirical baseline for diagnosing abnormal shifts in symbiont communities of dolphins under human care. In the future blow and fecal monitoring should be considered as an important part of cetacean health management.

Data availability statement

The data presented in the study are deposited in the NCBI repository, accession number PRJNA865874.

Ethics statement

The animal study was reviewed and approved by People's Republic of China for the Implementation of Wild Aquatic Animal Protection.

Author contributions

XW and JZ designed this research. XW and JL collected the samples and performed the lab experiments. XW and RT completed the data analysis. XW and JZ wrote the manuscript. RM, YH, JW, ZW, FF, and DW assisted with experiments and advice on the manuscript. All authors contributed to discussions, and revisions, and approved the final manuscript.

Funding

This study was partly supported by grants from the National Natural Science Foundation of China (No. 31870372), China Postdoctoral Science Foundation (No. 2020M682530), and the Bureau of Science & Technology for Development, Chinese Academy of Sciences (No. ZSSD-004).

Acknowledgments

We thank Dr. Xianyuan Zeng, Junhua Ji, and the dolphin trainers in Nanning Zoo for their assistance in collecting specimens. We also appreciate Dr. Youchuan Wei in Guangxi University for helping with sample preprocessing.

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.2022.1024117/full#supplementary-material

References

Acevedo-Whitehouse, K., Rocha-Gosselin, A., and Gendron, D. (2010). A novel non-invasive tool for disease surveillance of free-ranging whales and its relevance to conservation programs. *Anim. Conserv.* 13 (2), 217–225. doi: 10.1111/j.1469-1795.2009.00326.x

Bai, S., Zhang, P., Zhang, C., Du, J., Du, X., Zhu, C., et al. (2021). Comparative study of the gut microbiota among four different marine mammals in an aquarium. *Front. Microbiol.* 12. doi: 10.3389/fmicb.2021.769012

Baniel, A., Amato, K. R., Beehner, J. C., Bergman, T. J., Mercer, A., Perlman, R. F., et al. (2021). Seasonal shifts in the gut microbiome indicate plastic responses to diet in wild geladas. *Microbiome* 9 (1), 1–20. doi: 10.1186/s40168-020-00977-9

Bik, E. M., Costello, E. K., Switzer, A. D., Callahan, B. J., Holmes, S. P., Wells, R. S., et al. (2016). Marine mammals harbor unique microbiotas shaped by and yet distinct from the sea. *Nat. Commun.* 7 (1), 1–13. doi: 10.1038/ncomms10516

Bosco, N., and Noti, M. (2021). The aging gut microbiome and its impact on host immunity. *Genes Immun.* 22 (5), 289–303. doi: 10.1038/s41435-021-00126-8

Brown, J., Pirrung, M., and McCue, L. A. (2017). FQC dashboard: integrates FastQC results into a web-based, interactive, and extensible FASTQ quality control tool. *Bioinformatics* 33 (19), 3137–3139. doi: 10.1093/bioinformatics/btx373

Cole, J. R., Wang, Q., Fish, J. A., Chai, B., McGarrell, D. M., Sun, Y., et al. (2014). Ribosomal database project: data and tools for high throughput rRNA analysis. *Nucleic Acids Res.* 42 (D1), D633–D642. doi: 10.1093/nar/gkt1244

Curry, B. E., Ralls, K., and Brownell, R. L.Jr. (2013). Prospects for captive breeding of poorly known small cetacean species. *Endanger. Species Res.* 19 (3), 223–243. doi: 10.3354/esr00461

Dallas, J. W., and Warne, R. W. (2022). Captivity and animal microbiomes: potential roles of microbiota for influencing animal conservation. *Microb. Ecol.*, 1–19. doi: 10.1007/s00248-022-01991-0

Dickson, R. P., Erb-Downward, J. R., Martinez, F. J., and Huffnagle, G. B. (2016). The microbiome and the respiratory tract. *Annu. Rev. Physiol.* 78, 481. doi: 10.1146/annurev-physiol-021115-105238

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

Edgar, R. C. (2013). UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat. Methods* 10 (10), 996–998. doi: 10.1038/nmeth.2604

Enaud, R., Prevel, R., Ciarlo, E., Beaufils, F., Wieërs, G., Guery, B., et al. (2020). The gut-lung axis in health and respiratory diseases: A place for inter-organ and inter-kingdom crosstalks. *Front. Cell. Infect. Microbiol.* 10. doi: 10.3389/fcimb.2020.00009

Foot, N. J., Orgeig, S., and Daniels, C. B. (2006). The evolution of a physiological system: the pulmonary surfactant system in diving mammals. *Resp. Physiol. Neurobi.* 154 (1-2), 118–138. doi: 10.1016/j.resp.2006.04.012

Frame, L. A., Costa, E., and Jackson, S. A. (2020). Current explorations of nutrition and the gut microbiome: a comprehensive evaluation of the review literature. *Nutr. Rev.* 78 (10), 798–812. doi: 10.1093/nutrit/nuz106

Frankel, J. S., Mallott, E. K., Hopper, L. M., Ross, S. R., and Amato, K. R. (2019). The effect of captivity on the primate gut microbiome varies with host dietary niche. *Am. J. Primatol.* 81 (12), e23061. doi: 10.1002/ajp.23061

Gerritsen, J., Hornung, B., Ritari, J., Paulin, L., Rijkers, G. T., Schaap, P. J., et al. (2019). A comparative and functional genomics analysis of the genus *Romboutsia* provides insight into adaptation to an intestinal lifestyle. *BioRxiv* 845511. doi: 10.1101/845511

Habib, C., Houel, A., Lunazzi, A., Bernardet, J.-F., Olsen, A. B., Nilsen, H., et al. (2014). Multilocus sequence analysis of the marine bacterial genus *Tenacibaculum* suggests parallel evolution of fish pathogenicity and endemic colonization of aquaculture systems. *Appl. Environ. Microbiol.* 80, 5503–5514. doi: 10.1128/AEM.01177-14

Hyde, E. R., Navas-Molina, J. A., Song, S. J., Kueneman, J. G., Ackermann, G., Cardona, C., et al. (2016). The oral and skin microbiomes of captive komodo dragons are significantly shared with their habitat. *mSystems* 1 (4), e00046–e00016. doi: 10.1128/mSystems.00046-16

Inoue, R., and Ushida, K. (2003). Vertical and horizontal transmission of intestinal commensal bacteria in the rat model. *FEMS Microbiol. Ecol.* 46 (2), 213–219. doi: 10.1016/S0168-6496(03)00215-0

Invernizzi, R., Lloyd, C. M., and Molyneaux, P. L. (2020). Respiratory microbiome and epithelial interactions shape immunity in the lungs. *Immunology* 160 (2), 171–182. doi: 10.1111/imm.13195

Jin, L., Huang, Y., Yang, S., Wu, D., Li, C., Deng, W., et al. (2021). Diet, habitat environment and lifestyle conversion affect the gut microbiomes of giant pandas. *Sci. Total Environ.* 770, 145316. doi: 10.1016/j.scitotenv.2021.145316

Jin Song, S., Woodhams, D. C., Martino, C., Allaband, C., Mu, A., Javorschi-Miller-Montgomery, S., et al. (2019). Engineering the microbiome for animal health and conservation. *Exp. Biol. Med.* 244 (6), 494–504. doi: 10.1177/1535370219830075

Kau, A. L., Ahern, P. P., Griffin, N. W., Goodman, A. L., and Gordon, J. I. (2011). Human nutrition, the gut microbiome and the immune system. *Nature* 474 (7351), 327–336. doi: 10.1038/nature10213

Kaul, D., Rathnasinghe, R., Ferres, M., Tan, G. S., Barrera, A., Pickett, B. E., et al. (2020). Microbiome disturbance and resilience dynamics of the upper respiratory tract during influenza a virus infection. *Nat. Commun.* 11 (1), 1–12. doi: 10.1038/s41467-020-16429-9

Kembel, S. W., Cowan, P. D., Helmus, M. R., Cornwell, W. K., Morlon, H., Ackerly, D. D., et al. (2010). Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* 26 (11), 1463–1464. doi: 10.1093/bioinformatics/btq166

Kokou, F., Sasson, G., Nitzan, T., Doron-Faigenboim, A., Harpaz, S., Cnaani, A., et al. (2018). Host genetic selection for cold tolerance shapes microbiome composition and modulates its response to temperature. *eLlife* 7, e36398. doi: 10.7554/eLife.36398

Lima, N., Rogers, T., Acevedo-Whitehouse, K., and Brown, M. V. (2012). Temporal stability and species specificity in bacteria associated with the bottlenose dolphins respiratory system. *Env. Microbiol. Rep.* 4 (1), 89–96. doi: 10.1111/j.1758-2229.2011.00306.x

Lindstrom, M. J., and Bates, D. M. (1990). Nonlinear mixed effects models for repeated measures data. *Biometrics* 46, 673–687. doi: 10.2307/2532087

Lott, R., and Williamson, C. (2017). "Cetaceans in captivity," in *Marine mammal welfare. animal welfare*, vol. 17 . Ed. A. Butterworth (Cham: Springer). doi: 10.1007/978-3-319-46994-2_11

Louca, S., Parfrey, L. W., and Doebeli, M. (2016). Decoupling function and taxonomy in the global ocean microbiome. *Science* 353 (6305), 1272–1277. doi: 10.1126/science.aaf4507

Mallott, E. K., and Amato, K. R. (2021). Host specificity of the gut microbiome. *Nat. Rev. Microbiol.* 19 (10), 639–653. doi: 10.1038/s41579-021-00562-3

Mallott, E. K., Borries, C., Koenig, A., Amato, K. R., and Lu, A. (2020). Reproductive hormones mediate changes in the gut microbiome during pregnancy and lactation in phayre's leaf monkeys. *Sci. Rep.* 10 (1), 1–9. doi: 10.1038/s41598-020-66865-2

McKenzie, V. J., Song, S. J., Delsuc, F., Prest, T. L., Oliverio, A. M., Korpita, T. M., et al. (2017). The effects of captivity on the mammalian gut microbiome. *Integr. Comp. Biol.* 57 (4), 690–704. doi: 10.1093/icb/icx090

Moeller, A. H., Gomes-Neto, J. C., Mantz, S., Kittana, H., Segura Munoz, R. R., Schmaltz, R. J., et al. (2019). Experimental evidence for adaptation to species-specific gut microbiota in house mice. *mSphere* 4 (4), e00387–e00319. doi: 10.1128/mSphere.00387-19

Muka, S., and Zarpentine, C. (2021). Cetacean conservation and the ethics of captivity. *Biol. Conserv.* 262, 109303. doi: 10.1016/j.biocon.2021.109303

Nabi, G., McLaughlin, R. W., Khan, S., Hao, Y., and Chang, M. (2020). Pneumonia in endangered aquatic mammals and the need for developing low-coverage vaccination for their management and conservation. *Anim. Health Res. Rev.* 21 (2), 122–130. doi: 10.1017/S1466252320000158

Nelson, T. M., Wallen, M. M., Bunce, M., Oskam, C. L., Lima, N., Clayton, L., et al. (2019). Detecting respiratory bacterial communities of wild dolphins:

implications for animal health. *Mar. Ecol. Prog. Ser.* 622, 203–217. doi: 10.3354/meps13055

Ning, D., Yuan, M., Wu, L., Zhang, Y., Guo, X., Zhou, X., et al. (2020). A quantitative framework reveals ecological drivers of grassland microbial community assembly in response to warming. *Nat. Commun.* 11 (1), 1–12. doi: 10.1038/s41467-020-18560-z

Parsons, E. C. M., Baulch, S., Bechshoft, T., Bellazzi, G., Bouchet, P., Cosentino, A. M., et al. (2015). Key research questions of global importance for cetacean conservation. *Endanger. Species Res.* 27 (2), 113–118. doi: 10.3354/esr00655

Price, M. N., Dehal, P. S., and Arkin, A. P. (2009). FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Mol. Biol. Evol.* 26 (7), 1641–1650. doi: 10.1093/molbev/msp077

Ramees, T. P., Dhama, K., Karthik, K., Rathore, R. S., Kumar, A., Saminathan, M., et al. (2017). *Arcobacter*: an emerging food-borne zoonotic pathogen, its public health concerns and advances in diagnosis and control-a comprehensive review. *Vet. Quart.* 37 (1), 136–161. doi: 10.1080/01652176.2017.1323355

Robles-Malagamba, M. J., Walsh, M. T., Ahasan, M. S., Thompson, P., Wells, R. S., Jobin, C., et al. (2020). Characterization of the bacterial microbiome among free-ranging bottlenose dolphins (*Tursiops truncatus*). *Heliyon* 6 (6), e03944. doi: 10.1016/j.heliyon.2020.e03944

Roggenbuck, M., Bærholm Schnell, I., Blom, N., Bælum, J., Bertelsen, M. F., Sicheritz-Pontén, T., et al. (2014). The microbiome of new world vultures. *Nat. Commun.* 5, 5498. doi: 10.1038/ncomms6498

Segawa, T., Ohno, Y., Tsuchida, S., Ushida, K., and Yoshioka, M. (2020). *Helicobacter delphinicola* sp. nov., isolated from common bottlenose dolphins *Tursiops truncatus* with gastric diseases. *Dis. Aquat. Organ.* 141, 157–169. doi: 10.3354/dao03511

Sepulveda, J., and Moeller, A. H. (2020). The effects of temperature on animal gut microbiomes. *Front. Microbiol.* 11. doi: 10.3389/fmicb.2020.00384

Simmonds, M. P., and Eliott, W. J. (2009). Climate change and cetaceans: concerns and recent developments. *J. Mar. Biol. Assoc.* 89 (1), 203–210. doi: 10.1017/S0025315408003196

Stegen, J. C., Lin, X., Fredrickson, J. K., Chen, X., Kennedy, D. W., Murray, C. J., et al. (2013). Quantifying community assembly processes and identifying features that impose them. *ISME J.* 7 (11), 2069–2079. doi: 10.1038/ismej.2013.93

Suzuki, A., Akuzawa, K., Kogi, K., Ueda, K., and Suzuki, M. (2021). Captive environment influences the composition and diversity of fecal microbiota in indo-pacific bottlenose dolphins, *Tursiops aduncus*. *Mar. Mammal Sci.* 37 (1), 207–219. doi: 10.1111/mms.12736

Suzuki, A., Segawa, T., Sawa, S., Nishitani, C., Ueda, K., Itou, T., et al. (2019). Comparison of the gut microbiota of captive common bottlenose dolphins *Tursiops truncatus* in three aquaria. *J. Appl. Microbiol.* 126 (1), 31–39. doi: 10.1111/jam.14109

Toor, D., Wasson, M. K., Kumar, P., Karthikeyan, G., Kaushik, N. K., Goel, C., et al. (2019). Dysbiosis disrupts gut immune homeostasis and promotes gastric diseases. *Int. J. Mol. Sci.* 20 (10), 2432. doi: 10.3390/ijms20102432

Turvey, S. T., Pitman, R. L., Taylor, B. L., Barlow, J., Akamatsu, T., Barrett, L. A., et al. (2007). First human-caused extinction of a cetacean species? *Biol. Lett.* 3 (5), 537–540. doi: 10.1098/rsbl.2007.0292

Vendl, C., Nelson, T., Ferrari, B., Thomas, T., and Rogers, T. (2021). Highly abundant core taxa in the blow within and across captive bottlenose dolphins provide evidence for a temporally stable airway microbiota. *BMC Microbiol.* 21 (1), 20. doi: 10.1186/s12866-020-02076-z

Vuong, H. E., Yano, J. M., Fung, T. C., and Hsiao, E. Y. (2017). The microbiome and host behavior. *Annu. Rev. Neurosci.* 40 (1), 21–49. doi: 10.1146/annurev-neuro-072116-031347

Waltzek, T. B., Cortés-Hinojosa, G., Wellehan, J. F. X.Jr., and Gray, G. C. (2012). Marine mammal zoonoses: a review of disease manifestations. *Zoonoses Public Helth.* 59 (8), 521–535. doi: 10.1111/j.1863-2378.2012.01492.x

Wang, D., Hao, Y., Wang, K., Zhao, Q., Chen, D., Wei, Z., et al. (2005). Aquatic resource conservation. the first Yangtze finless porpoise successfully born in captivity. *Environ. Sci. Pollut. R.* 12 (5), 247–250. doi: 10.1065/espr2005.08.284

Wan, X., Li, J., Cheng, Z., Ao, M., Tian, R., Mclaughlin, R. W., et al. (2021). The intestinal microbiome of an indo-pacific humpback dolphin (*Sousa chinensis*) stranded near the pearl river estuary, China. *Integr. Zool.* 16 (3), 287–299. doi: 10.1111/1749-4877.12477

Wan, X., McLaughlin, R. W., Zheng, J., Hao, Y., Fan, F., Tian, R., et al. (2018). Microbial communities in different regions of the gastrointestinal tract in East Asian finless porpoises (*Neophocaena asiaeorientalis sunameri*). *Sci. Rep.* 8 (1), 14142. doi: 10.1038/s41598-018-32512-0

Weiss, S., Xu, Z., Peddada, S., Amir, A., Bittinger, K., Gonzalez, A., et al. (2017). Normalization and microbial differential abundance strategies depend upon data characteristics. *Microbiome* 5 (1), 1–18. doi: 10.1186/s40168-017-0237-y

West, A. G., Waite, D. W., Deines, P., Bourne, D. G., Digby, A., McKenzie, V. J., et al. (2019). The microbiome in threatened species conservation. *Biol. Conserv.* 229, 85–98. doi: 10.1016/j.biocon.2018.11.016

Xie, Y., Ouyang, Y., Han, S., Se, J., Tang, S., Yang, Y., et al. (2022). Crop rotation stage has a greater effect than fertilisation on soil microbiome assembly and enzymatic stoichiometry. *Sci. Total Environ.* 815, 152956. doi: 10.1016/j.scitotenv.2022.152956

Xu, R., Lu, R., Zhang, T., Wu, Q., Cai, W., Han, X., et al. (2021). Temporal dynamics of human respiratory and gut microbiomes during the course of COVID-19 in adults. *Commun. Biol.* 4 (1), 1-11. doi: 10.1101/2020.07.21.20158758

Yadav, M., Verma, M. K., and Chauhan, N. S. (2018). A review of metabolic potential of human gut microbiome in human nutrition. *Arch. Microbiol.* 200 (2), 203–217. doi: 10.1007/s00203-017-1459-x

Zhou, L., Liu, L., Chen, W., Sun, J., Hou, S., Kuang, T., et al. (2020). Stochastic determination of the spatial variation of potentially pathogenic bacteria communities in a large subtropical river. *Environ. Pollut.* 264, 114683. doi: 10.1016/j.envpol.2020.114683

frontiers | Frontiers in Marine Science

TYPE Original Research
PUBLISHED 05 December 2022
DOI 10.3389/fmars.2022.1077595

OPEN ACCESS

EDITED BY
Yunyun Zhuang,
Ocean University of China, China

REVIEWED BY
Lingxin Chen,
Yantai Institute of Coastal Zone Research (CAS), China
Tingting Yin,
Nanyang Technological University, Singapore
Yi Xin,
Hainan University, China

*CORRESPONDENCE
Xin Zhang
xzhang@qdio.ac.cn
Chaolun Li
lcl@qdio.ac.cn

†These authors have contributed equally to this work and share first authorship

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

RECEIVED 23 October 2022
ACCEPTED 22 November 2022
PUBLISHED 05 December 2022

CITATION
He W, Wang M, Li M, Zhong Z, Chen H, Xi S, Luan Z, Li C and Zhang X (2022) Confocal Raman microscopy for assessing effects of preservation methods on symbiotic deep-sea mussel gills.
Front. Mar. Sci. 9:1077595.
doi: 10.3389/fmars.2022.1077595

Confocal Raman microscopy for assessing effects of preservation methods on symbiotic deep-sea mussel gills

Wanying He[1,2,3†], Minxiao Wang[3,4,5†], Mengna Li[3,4,5], Zhaoshan Zhong[4,5], Hao Chen[4,5], Shichuan Xi[1,2], Zhendong Luan[1,2,3], Chaolun Li[6*] and Xin Zhang[1,2,3*]

[1]Chinese Academy of Sciences Key Laboratory of Marine Geology and Environment and Center of Deep Sea Research, Institute of Oceanology, Chinese Academy of Sciences (CAS), Qingdao, China, [2]Laboratory for Marine Geology, Pilot Laboratory for Marine Science and Technology, Qingdao, China, [3]University of Chinese Academy of Sciences, Beijing, China, [4]Chinese Academy of Sciences (CAS) Key Laboratory of Marine Ecology and Environmental Sciences and Center of Deep Sea Research, Institute of Oceanology, Chinese Academy of Sciences (CAS), Qingdao, China, [5]Laboratory for Marine Ecology and Environmental Science, Qingdao National Laboratory for Marine Science and Technology, Qingdao, China, [6]South China Sea Institute of Oceanology, Chinese Academy of Sciences (CAS), Guangzhou, China

Confocal Raman microscopy (CRM) is a powerful tool for biological research, which can provide information regarding the composition and distribution of biomolecules in an in situ, label-free, non-destructive manner and with high spatial resolution. Sample preservation is often an unavoidable step, especially for symbiotic deep-sea samples. Moreover, protocols for the preservation of samples for CRM have not been established and specific effects of different preservation methods on biomolecules have not been studied for relevant samples. In this study, we used deep-sea mussel *Gigantidas platifrons*, an ideal model in the study of deep-sea symbiosis and investigated the effect of four common preservation methods on the results of CRM imaging and signals. The methods included snap-freeze (SF), SF followed by rapid fixation in methanol (SF-MeOH), 2.5% glutaraldehyde and 2% paraformaldehyde fixation (SF-GP), and 4% paraformaldehyde and alcohol fixation (PS-PA). The results of this study indicate that SF was the most effective method for the comprehensive analysis of the biomolecular composition although the sectioning success rate was relatively low. Moreover, SF-MeOH was found to be effective when SF is not sufficient in obtaining good morphology in sections, or when the effect of chemical bonding on the composition of biomolecules upon SF-MeOH can be neglected. Finally, SF-GP and PS-PA were found to be the most effective methods considering the overall morphological observation. However, they were less suitable for metabolic studies. We believe our results can provide guidance for further studies of Raman on symbiotic deep-sea biological samples. It is of great importance for the wide application of Raman technique.

KEYWORDS

confocal Raman microscopy, symbiotic deep-sea mussel, preservation methods, Raman imaging, Raman signals

Introduction

Raman techniques reflect chemical bond vibrations using inelastic light scattering (Gomes Da Costa et al., 2019). It is a promising analytical technique in biology (Clemens et al., 2014; Kallepitis et al., 2017), chemistry (Wetzel and LeVine, 1999), physics (Moerner and Orrit, 1999), materials (Chen et al., 2012), and geology (Dodd et al., 2017). In biological research, it has been used for the identification and detection of microorganisms (Clemens et al., 2014; Pahlow et al., 2015; Stöcke et al., 2016; Lorenz et al., 2017), metabolites within living cells (Du et al., 2020), and different types of cancer (Clemens et al., 2014). It is particularly important for biological research because it allows rapid, *in situ*, label-free, and non-destructive acquisition of information related to the composition and distribution of biomolecules. In addition, Raman analysis usually does not require extensive sample preparation or is not hindered by interference from water molecules. Among Raman techniques, confocal Raman microscopy (CRM) has a higher spatial resolution; thus, it can be combined with imaging and quantitative analysis to determine the composition of biomolecules. Moreover, it can be used to study the complex metabolic processes in cells in a spatiotemporal manner (Clemens et al., 2014; Butler et al., 2016; Kallepitis et al., 2017; Lazarevic et al., 2019; Høgset et al., 2020). These advantages have led to a growing interest in exploring CRM's potential for application in biologically relevant fields (Clemens et al., 2014; Gomes Da Costa et al., 2019; Lazarevic et al., 2019).

Sample preservation is critical for spectroscopy-based analysis. For example, without chemical fixation, exposure of bone tissues to room temperature affects the carbonate-to-phosphate ratio owing to natural degradation (Fiedler et al., 2018). Although Raman spectroscopy can be conducted to directly analyze fresh tissues or cells (Fiedler et al., 2018), preservation is unavoidable in most cases. Some samples need to be transported over long distances and preserved for long periods of time. In order to accurately analyze the structural and biomolecular composition of such samples (e.g., tissues or cells) and to mimic their original state *in vivo*, sample preservation is critical (Mariani et al., 2009; Butler et al., 2016). Sample preservation includes non-chemical preservation (e.g., direct snap-freezing) and chemical fixation methods (Mariani et al., 2009). In unfixed frozen sections, tissue or cell components can be analyzed closest to their *in situ* state; however, the sections are fragile and may undergo degradation (Butler et al., 2016). On the other hand, chemical preservation methods make it easier to obtain sections from tissues. However, although various chemical fixation methods are available, contamination from the fixative may lead to the degradation or loss of components of certain cells, preventing the accurate analysis through Raman techniques (Lyng et al., 2011; Galli et al., 2014). Therefore, the selection of the appropriate sample preservation method based on experimental needs is critical to the interpretation of spectroscopic and imaging results (Mariani et al., 2009; Meade et al., 2010; Kuzmin et al., 2014).

The effect of the sample preservation method on the Raman signal has been explored, and it was found that each method has its advantages and disadvantages because of different fixation mechanisms (Mariani et al., 2009; Meade et al., 2010; Kuzmin et al., 2014). At present, some conventional organisms have been extensively researched, because of their easy access and simple composition; also, the samples can be preserved and processed in a more effective way, allowing their direct analysis through CRM (Høgset et al., 2020). However, in several symbiotic deep-sea organisms, wherein the intracellular mechanisms remain unclear, several problems are encountered, such as difficulty in sampling and preservation, presence special structures in tissues or cells, high complexity of tissues, and low light transmission (Wang et al., 2021). Since different preservation methods have different effects on cell morphology and biomolecular content (i.e., nucleic acids, lipids, and proteins (Kuzmin et al., 2014)), the resulting compositional and morphological changes are likely to vary depending on the organism sample or cell type (Mariani et al., 2009; Meade et al., 2010). A systematic comparison of sample processing methods for symbiotic deep-sea organism samples is lacking. To promote research, further exploration of sample preservation methods is necessary before proceeding with CRM to resolve metabolic mechanisms.

To this end, we selected deep-sea mussel *Gigantidas platifrons* (*G. platifrons*), the dominant species functioning as secondary producers with a large number of symbionts dwelling inside their gills, as the object of our study (Dubilier et al., 2008; Xu et al., 2019). Since the sharing of the metabolites among the symbionts and hosts extend the ecological niches of both sides greatly (Sun et al., 2017; Wang et al., 2021), it is of great value to study the distribution pattern of the key metabolites. The findings will undoubtedly increase our understanding how deep-sea life thrives in these extreme environments and the cycling of elements such as carbon and sulfur (Childress et al., 1986; Petersen et al., 2011; Laming et al., 2018). However, the gill is fragile with variable morphology, complex metabolic composition, and complex intra- and intercellular interactions that hinder further studies of its cellular function (Wang et al., 2021). This organism requires preservation and sectioning before Raman spectroscopy owing to the soft texture of gills and impossible on-site analysis. Different preservation methods may have different effects on the composition and morphology of the tissues, which in turn may mislead us to make wrong judgments (Clemens et al., 2014; Fiedler et al., 2018). Therefore, we need to evaluate different preservation methods before embarking on the relevant research regarding its composition and the distribution of its biomolecules (Geier et al., 2020), in order to select the appropriate method for the subsequent focus of attention.

Herein, we explored four commonly used preservation methods for CRM: snap freeze, snap freeze followed by rapid fixation in methanol, 2.5% glutaraldehyde and 2% paraformaldehyde fixation, and 4% paraformaldehyde and alcohol fixation. We then analyzed the effect of the Raman assay results on different regions of the gill preserved using different methods after sectioning, to show whether CRM is sensitive enough to detect these changes. The results of this study can facilitate the improvement in the performance of Raman techniques used for relevant samples and the pretreatment of other biological methods.

Materials and methods

Sample collection

The deep-sea mussels (*G. platifrons*) were collected during the "Kexue" cruise in 2021 from the cold seeps of the South China Sea (119°17′08.297″E; 22°06′55.435″N, temperature: 3.57°C, salinity: 34.65 psu). The samples were brought to the deck using a thermal-preserve sampler carried by a remotely operated vehicle (ROV), and rapidly dissected for further processing.

Preservation treatment of samples

We compared samples obtained by four common preservation methods. For the samples obtained by first method, SF, paper was used to absorb the liquid around the gill block to prevent the formation of ice crystals. A metal beaker was then filled with isopentane and placed in liquid nitrogen. Next, the tissue block was placed in the metal beaker by using forceps to completely submerge the tissue in isopentane until it was completely frozen. This step increases the rate of freezing and restricts ice crystal formation. The snap-frozen gill was then placed in a −80°C freezer for long-term storage. For slicing, the gill was broken into small pieces and placed directly into a pre-chilled optimal cutting temperature compound (OCT) embedding agent for frozen sectioning, with a section thickness of 10 µm (Leica CM1950). The second samples obtained by method, SF-MeOH, follows the same steps as those involved in SF till the sectioning process. For sectioning in SF-MeOH, the sections were fixed in methanol for 10 min, washed three times with 0.01 M phosphate-buffered saline (1 × PBS), and then placed in a pre-chilled OCT embedding agent for frozen sectioning (thickness: 10 µm). For the third samples obtained by method (SF-GP), the dissected gills were directly transferred to a pre-cooled paraformaldehyde-glutaraldehyde (2%/2.5%) fixative, fixed at 4°C for 24 h, and placed in a fresh fixative. The slicing operation was the same as that conducted in SF. For the fourth samples obtained by method (PS-PA), the dissected gills were stored in a 4% paraformaldehyde fixative (PFA) for 24 h, rinsed twice with 1 × PBS, and stored in 75% ethanol at 4°C. The samples were dehydrated in different gradients of ethanol (70%, 80%, 95%, and 100%) for 1 h, and placed in xylene for 2 h. After transparency, the samples were embedded in paraffin wax at 60°C for 3 h. The embedded samples were sectioned using a Leica microtome (Leica RM 2016) to achieve a thickness of 7 µm. The samples were dewaxed with xylene before use and then rehydrated with different gradients of ethanol (100%, 95%, 80%, and 70%) before use.

Spectral acquisition

Parts of the gill were selected for Raman spectral acquisition using a confocal Raman microspectrometer (alpha 300R, WITec, Ulm, Germany) equipped with a laser operating at 532 nm and a 600 grooves/mm grating (UHTS 300). The sample was placed on a calcium fluoride substrate, and the excitation laser was focused onto the sample for spectral acquisition using a ZEISS EC Epiplan (Carl ZEISS, Jena, Germany) with 50×/0.75 objectives. Before use, systematic calibration was performed using the characteristic peak of the silicon wafer at 520 cm^{-1}. The step size, total acquisition time, and laser power of the data set acquired by each method were optimized to obtain the best results. The laser power was 15–20 mW for one point in the X- and Y-axis directions of 310–670 nm.

Processing of Raman spectra

For preprocessing the raw data, baseline correction and cosmic ray removal were performed using WITec Project plus, followed by principal component analysis (PCA) to improve the signal-to-noise ratio in order to retain the necessary signal. The tissue was then partitioned for imaging using clustering analysis to further obtain characteristic spectrum of the tissue. The spectra were fitted with a Gaussian using the GRAMS/AI software, and the model peak positions were chosen based on the values of the relevant material peak positions reported by different authors. A range of positions for each Raman peak was restricted for the fit, and the peak positions were fine-tuned on the basis of the fit. The result with the smallest sum of squared deviations between the curve fitting model and the experimental data was considered to obtain the relevant information such as peak position, peak intensity, peak width, and peak area.

Results and discussion

The pipeline for CRM analysis in gills

We have established a pipeline to show the metabolic distribution feature of key metabolites of the gills of the

symbiotic deep-sea mussels. A well-preserved SF sample was selected for the metabolic profile constructions. It has been demonstrated that SF is closest to the *in situ* state in terms of component composition because it has not been disturbed by immobilizers (Lyng et al., 2011; Butler et al., 2016). We performed classification imaging of the Raman dataset by multivariate analysis and found that the resulting image could be classified into four major categories: the nucleic acid enrichment region, protein enrichment region, phospholipid enrichment region, and the lysosome (Figure 1A). By analyzing the characteristic spectra of each region, the peak for the red nucleic acid region was found at 1,063 cm^{-1} which can be attributed to DNA (Okotrub et al., 2015), suggesting the existence of the nucleus. The peak at 1,261 cm^{-1} was the cyan protein enrichment region and is assigned to =C-H bending (protein) (Chan et al., 2006), indicating the presence of the cytoplasm containing symbionts. Moreover, a peak 1,270 cm^{-1} indicated the blue phospholipid enrichment region, and can be attributed to the presence of typical phospholipids (Malini et al., 2006; Kochan et al., 2013), representing the basal membrane (Figure 1B). Lastly, the characteristic peak of 780 cm^{-1} for the yellow region is likely ascribed to the presence of phosphatidylinositol (Xu et al., 2020), which can be used as a lipid signaling molecule for the lysosome (Hohman et al., 1982). This imaging distribution is consistent with conventional methods observed in the published literature (Ponnudurai et al., 2017; Sun et al., 2017; Wang et al., 2021). Therefore, we chose SF samples as standards to analyze the effects of different preservation methods on subsequent Raman studies and biological studies.

Effects of different methods on the morphology and biomolecular distribution of gills

Using the same spectral processing method, we found that the spectral datasets obtained by all four methods could allow the visualization of the three major classes of the nucleus, cytoplasm containing symbionts, and basal membrane after imaging analysis (Figures 1A and 2). However, cells at SF during sectioning are very prone to extensive fragmentation, leaving only the basal membrane, or adhering together indistinguishably, requiring many sections for local selection detection (Supplementary Figure S2). The basal membrane detected upon SF was found to be significantly thicker than that through other methods *via* CRM imaging (Figure 1A); this thickness is probably due to being unfixed, resulting in scraping during sectioning, which may have an impact on the biomolecular distribution of the biomolecules. The overall morphology acquired upon SF-MeOH is more prominent than that upon SF, allowing the clearer distinction of the boundary of the gill cells and facilitating significantly less fragmentation (Supplementary Figure S2). Slight internal breakage was found after imaging of the selected areas (Figure 2), probably due to the rapid fixation of methanol, which maintained the overall morphology of the gill cells but did not provide good maintenance of the cell interior. SF-GP and PS-PA allowed the efficient maintenance of the overall morphology of the gill cells after sectioning due to prolonged fixation, and the degree of fragmentation was very low (Figure 2). However, prolonged immobilization upon SF-GP and PS-PA changed the

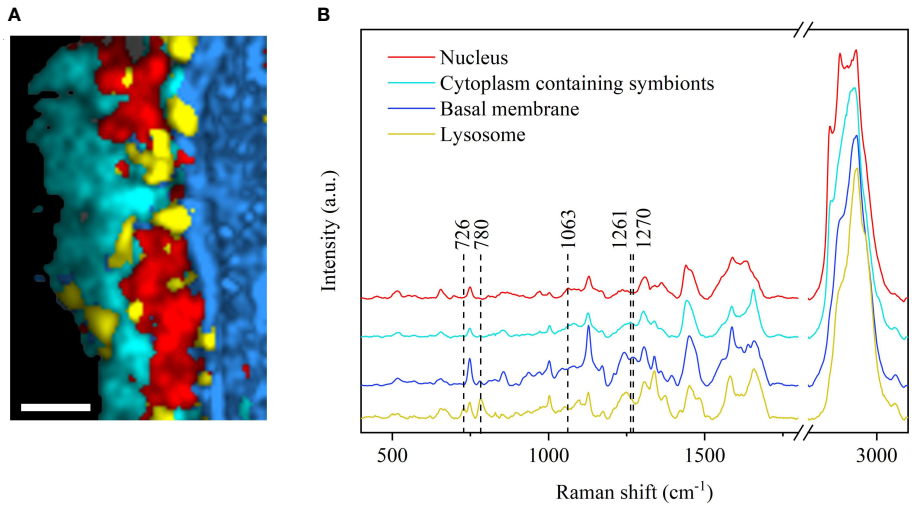


FIGURE 1
CRM imaging analysis of *G. platifrons* gills following the CS preservation method. **(A)** Raman imaging of gills showing the distribution of cellular components, including the nucleus (red), cytoplasm containing symbionts (cyan), basal membrane (blue), and lysosome (yellow). Scale bars: 8 μm. The bright field corresponding to Raman imaging images were shown in Supplementary Figure S1. **(B)** Average Raman spectrum of the nucleus, cytoplasm containing symbionts, basal membrane, and lysosome.

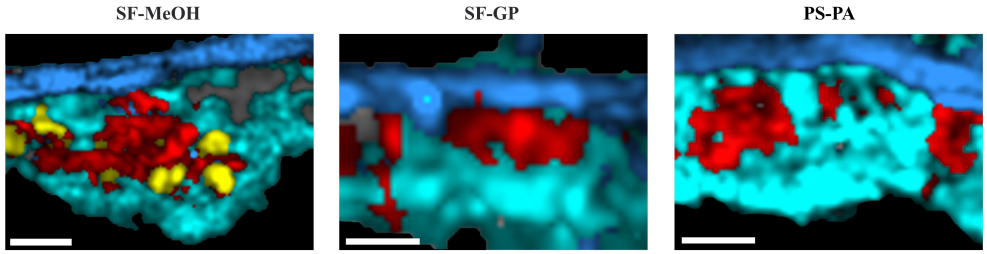


FIGURE 2
Comparative analysis of Raman imaging maps of the gills preserved using different methods. Raman imaging of the gills showing the distribution of cellular components, including the nucleus (red), cytoplasm containing symbionts (cyan), basal membrane (blue) and lysosome (yellow). The gray part is composed of feature spectrum of the slides, which is caused by the slight breakage of the tissue. Scale bars: 8 μm. The bright field corresponding to Raman imaging images were shown in Supplementary Figure S1.

intensity of the Raman signals of some substances, such as that of phosphatidylinositol (780 cm^{-1}), restricting the Raman imaging of some important organelles such as lysosomes (Figure 2). Moreover, prolonged immobilization resulted in poor preservation of the organelle structures.

Lysosomes are important for the study of symbiotic regulation and have been demonstrated in symbiotic systems of multiple taxa (Hohman et al., 1982; Nishikori et al., 2009; Castillo et al., 2015). Lysosomes are involved in regulating the homeostasis of symbiotic bacteria and determining the fate of the symbiont (Trench, 1971). In symbiotic deep-sea mussels, lysosomes play an important role in the intracellular digestion of the pasture. Here, SF-GP and PS-PA resulted in failure of imaging of the lysosomes; the spectra of the as-preserved samples did not show the characteristic peak for the lysosomes at 726 cm^{-1}, which is typically attributed to DNA (Talari et al., 2015) (Figure 1B)—an important evidence of the host's consumption of symbionts through lysosomes. Although the presence of lysosomes can also be verified *via* electron microscopy, the Raman signal of phosphatidylinositol is disrupted, leading to an inability to discriminate by Raman imaging. The absence of the peak corresponding to phosphatidylinositol, a membrane-forming molecule, indicates the loss of membrane integrity, which can lead to drastic changes in the distribution of some substances inside and outside the lysosome and affect the further study of substance distribution.

In summary, SF-GP and PS-PA are more effective toward the maintenance of the overall morphology; however, SF and SF-MeOH have a weaker effect on the distribution of substances due to the weak influence upon fixation, which is more favorable for some detailed imaging studies.

Effects of the different methods on the biomolecular composition

First, we took the total average spectra of the tissues after area scans of the gill sections obtained from the four different preservation methods (Figure 3). Significant variability in the peak positions and intensities in the regions within the spectra were observed; however, most of the molecular contributions could be identified (Table 1). The spectra were normalized to improve clarity (Faoláin, 2005). Considering that the signal of the Raman peaks may be influenced by the laser focusing, we discussed the differences in peak intensities together with the spectra before and after normalization for comparative analysis.

Following dewaxing, the peak in the spectra of the PS-PA samples were clearly different from those prepared by other methods at 2,881 cm^{-1}. Further analysis revealed that this peak was consistent with the Raman peak of paraffin wax. Further comparison with other paraffin wax main peaks indicated that paraffin wax likely has a stronger signal in this section. Other characteristic peaks of paraffin wax, e.g., those near 1,063 cm^{-1} (C-C stretch), 1,131 cm^{-1} (C-C stretch), 1,296 cm^{-1} (CH_2 deformation), and 1,418 cm^{-1} (CH_3 deformation), could be observed in the spectra (Faoláin, 2005; Tfayli et al., 2009). The combination of the average spectra obtained by the other treatments revealed that the peaks of the samples overlapped significantly with those of the biological tissues (Tfayli et al., 2009). It was difficult to distinguish their contribution and to determine whether the peak was due to the distortion of the biological components caused by extensive treatment. For example, in sections other than those preserved *via* PS-PA, the peak at 1,063 cm^{-1} can be judged as the characteristic peak for DNA; however, in the PS-PA-preserved samples, it may also be the characteristic peak of paraffin. Although it does not affect the H&E staining observation, incomplete dewaxing can interfere with the acquisition of Raman data and analysis. However, it is possible to study spectral regions of interest that do not overlap with the paraffin signal; also, for the study of overlapping regions, further optimization of the proposed method is needed to verify that paraffin is removed completely for more effective Raman analysis.

The Raman peak at 517 cm^{-1} corresponding to the S-S stretching (Van Wart et al., 1973; Maquelin et al., 2002) and at

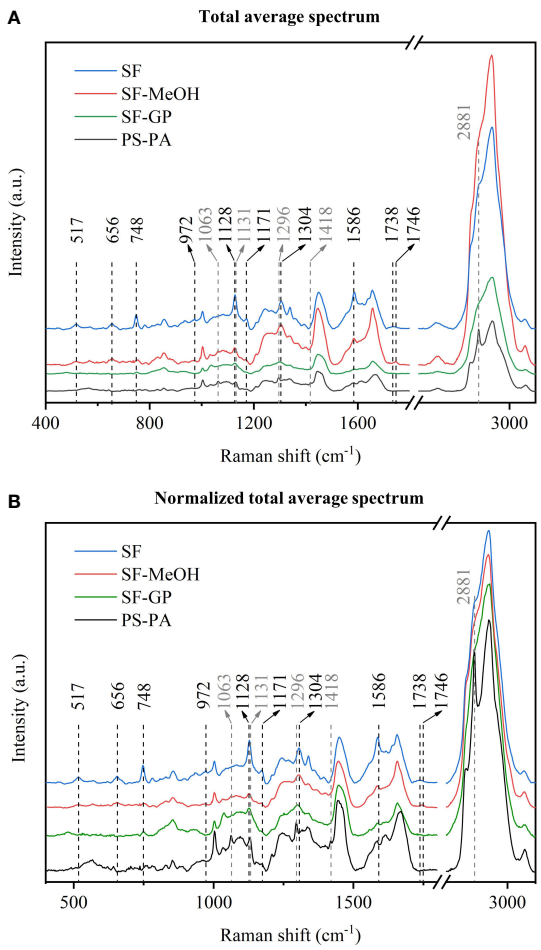


FIGURE 3
Comparative analysis of the spectra by the four different methods. **(A)** Average spectra. **(B)** Normalized average spectra.

656 cm^{-1} ascribed to the C-S stretching were present in the SF and SF-MeOH samples, indicating the ability of these methods to preserve some spatial structures and chemical bonds that are not preserved *via* SF-GP and PS-PA. The Raman peak at 748 cm^{-1} is derived from lanosterol (Supplementary Figure S3) and that at 1,304 cm^{-1} represents the amide II band (Fujioka et al., 2003; Pijanka et al., 2010), which are absent in the PS-PA-preserved samples, indicating that the PS-PA method alters the amino acid composition and disrupts the spatial structure of proteins. The 972 cm^{-1} peak is specific to the SF and SF-MeOH samples and can be assigned to lipids (Czamara et al., 2015). This indicates that fixatives can destroy lipids and should be used with caution during the interpretation of lipid changes. The samples preserved *via* SF have a unique Raman peak at 1,738 cm^{-1}, and while those *via* SF-MeOH have a unique Raman peak at 1,746 cm^{-1}—both of which can be attributed to the C=O stretching (lipids) (Lakshmi et al., 2002; Silveira et al., 2002; Huang et al., 2003; Stone et al., 2004; Krafft et al., 2005; Dukor, 2006; Huang et al., 2011; Czamara et al., 2015; Gieroba et al., 2020); these results indicate that SF and SF-MeOH are effective for the preservation of triacylglycerols (Czamara et al., 2015). The disappearance of the C=O band in the spectra of the samples preserved by the other two methods may be due to the fixation of aldehydes and the xylene treatment, resulting in the massive removal of cellular lipids from the tissues (Lyng et al., 2011).

In addition, we found that the Raman spectra for the SF samples were significantly stronger (with several peaks at around 517, 656, 748, 972, 1,128, 1,171, and 1,586 cm^{-1}, which may represent the S-S stretching (Van Wart et al., 1973; Maquelin et al., 2002), C-S stretching (Sugeta, 1975), lanosterol (Supplementary Figure S3), lipid (Czamara et al., 2015), carbohydrate (Wiercigroch et al., 2017), lanosterol, and amide II bands (Chen et al., 1995a; Chen et al., 1995b), respectively) than those for the other samples, suggesting that the concentration of these substances is higher when they are not contaminated with chemical fixatives. Therefore, the method associated with fixation will reduce the concentration of these substances.

Differences in the overall composition of the Raman spectra are expected; unfortunately, the spectra may cause some weaker peaks to be missed in the averaging process. Therefore, to further investigate the effects of different preservation methods on different biomolecules, we compared the enriched regions of nucleic acids, proteins, and lipids separately (Figure 4). In the nucleic acid enriched region (nucleus), the peaks in the Raman spectra of the SF and SF-MeOH samples at 1,095 cm^{-1} can be attributed to DNA (Malini et al., 2006), indicating that SF-MeOH is a good fixative for preserving nucleic acids. The Raman peak at 1,102 cm^{-1} is present in the SF-GP and PS-PA samples and belongs to amide III (Lakshmi et al., 2002). This indicates that SF and SF-MeOH may disrupt the secondary structures of proteins and that different methods have different effects on the spatial structure of proteins. In the region of protein enrichment (cytoplasm containing symbionts), the SF-GP sample possessed unique Raman peaks near 505 cm^{-1} representing S-S stretching (Sugeta, 1975; Van Wart and Scheraga, 1986), indicating that aldehyde fixation may well preserve the tertiary structure of proteins, while other methods would likely lead to tertiary denaturation. Alcohols denature proteins by disrupting their tertiary structures (Shao et al., 2012). The SF method may have led to conformational changes in the proteins (Lyng et al., 2011). In the phospholipid enrichment region (basal membrane), the spectra of the SF sample contained a peak at 896 cm^{-1} corresponding to the CH_3 rocking band, which can be attributed to the presence of the fatty acid chain (Maquelin et al., 2002). In addition, the spectra of the SF sample contained a peak at 1,309 cm^{-1}, which is ascribed to the CH_3/CH_2 twisting or bending mode of lipid/collagen (Cheng et al., 2005), and one at 1,360 cm^{-1}, corresponding to the tryptophan bands (Huang et al., 2003; Cheng et al., 2005). The band at 450 cm^{-1} in the spectra of the SF and SF-MeOH samples is due to the skeletal

TABLE 1 Raman band assignment observed in the spectra of the gills.

Raman shift (cm^{-1})	Assignment	Reference
450	skeletal mode, β(CCC) (carbohydrates)	(Wiercigroch et al., 2017)
505	S-S stretch	(Sugeta, 1975; Van Wart and Scheraga, 1986)
517	S-S stretch	(Van Wart et al., 1973; Maquelin et al., 2002)
656	C-S stretch	(Sugeta, 1975)
726	DNA	(Talari et al., 2015)
748	Lanosterol	Supplementary Figure S3
780	Phosphatidylinositol	(Krafft et al., 2005)
896	CH_3 rocking (fatty acid chain)	(Maquelin et al., 2002)
972	β(CH) (lipids)	(Czamara et al., 2015)
1063	C–C stretch (paraffin wax) DNA	(Faoláin, 2005; Tfayli et al., 2009; Okotrub et al., 2015)
1095	DNA	(Malini et al., 2006)
1102	Amide III and other groups (proteins)	(Lakshmi et al., 2002)
1128	Carbohydrates	(Wiercigroch et al., 2017)
1131	C–C stretch (paraffin wax)	(Faoláin, 2005)
1171	Tyrosine	(Lakshmi et al., 2002; Huang et al., 2003; Cheng et al., 2005)
1261	C-H bend (protein)	(Chan et al., 2006)
1270	Typical phospholipids	(Malini et al., 2006; Kochan et al., 2013)
1296	deformation (paraffin wax)	(Faoláin, 2005; Tfayli et al., 2009)
1304	Amide II	(Fujioka et al., 2003; Pijanka et al., 2010)
1309	CH_3/CH_2 twisting or bending mode (lipid/collagen)	(Cheng et al., 2005)
1360	Tryptophan	(Cheng et al., 2005)
1395	Uracil ring stretches	(Clemens et al., 2014)
1418	CH_3 deformation (paraffin wax)	(Tfayli et al., 2009)
1586	Amide II	(Chen et al., 1995a; Chen et al., 1995b)
1738-1746	C=O (Triacylglycerols)	(Huang et al., 2003; Gieroba et al., 2020)

mode of carbohydrates (Wiercigroch et al., 2017). Only the SF-GP samples did not undergo the uracil ring stretching, whose peak is located at 1,395 cm^{-1} (Clemens et al., 2014). These results suggest that SF-MeOH, SF-GP, and PS-PA disrupt the fatty acid chains, SF and SF-MeOH are more effective for the observation of the Raman peaks ascribed to carbohydrates, and SF-GP leads to protein destruction.

In addition to the direct analysis of the different peak positions, we can further identify the overlapping biomolecules by fitting the profiles *via* deconvolution analysis in order to obtain more information. We can observe the spectra of the samples obtained by the different methods to show significantly different profiles in terms of peak position, intensity, and width in the range of 1,500–1,700 cm^{-1}. The deconvolution analysis results (Figure 5 and Table 2) confirmed the existence of vibrational modes such as β-sheet, α-helix, unordered, turn, and aromatic amino acid ring modes, and side chains of the amide band (Goormaghtigh et al., 1994; Chen et al., 1995a; Chen et al., 1995b; Pelton and McLean, 2000; Barth and Zscherp, 2002; Faoláin et al., 2005; Chan et al., 2006; Lefèvre et al., 2007; Kochan et al., 2013; Yang et al., 2015; Rivas-Arancibia et al., 2017; Gieroba et al., 2020; Kowalska et al., 2020; Talaikis et al., 2020). Among them, a peak corresponding to a mixture of lipid and protein may exist at around 1,664 cm^{-1} (Kochan et al., 2013). The peak width range of 1,570-1,579 cm^{-1} was fitted in the total average spectra of the nucleus and cytoplasm containing symbionts for both the SF and SF-MeOH samples; these peaks can be attributed to DNA (Chan et al., 2006; Gieroba et al., 2020) (Supplementary Figure S4 and Table 2). This result reconfirmed that methanol can better preserve nucleic acids in cell nuclei, and has a stronger preservation effect for DNA in the cytoplasm containing symbionts. In terms of the protein structure, the SF-GP and PS-PA samples possessed unique turn structures, and the α-helix structure was completely disrupted in PS-PA.

To further investigate specific changes induced by the different preservation methods, we performed semi-quantitative analyses. Because the area of each peak corresponds to their conformational contribution (Kozicki et al., 2015), the area of each fraction was divided by the sum of the areas of all amide I band fractions to determine their relative contents. Table 3 summarizes the total average spectra of the different treatments and specific information on the subpeaks obtained after fitting the different sites for each method. From the results of the total

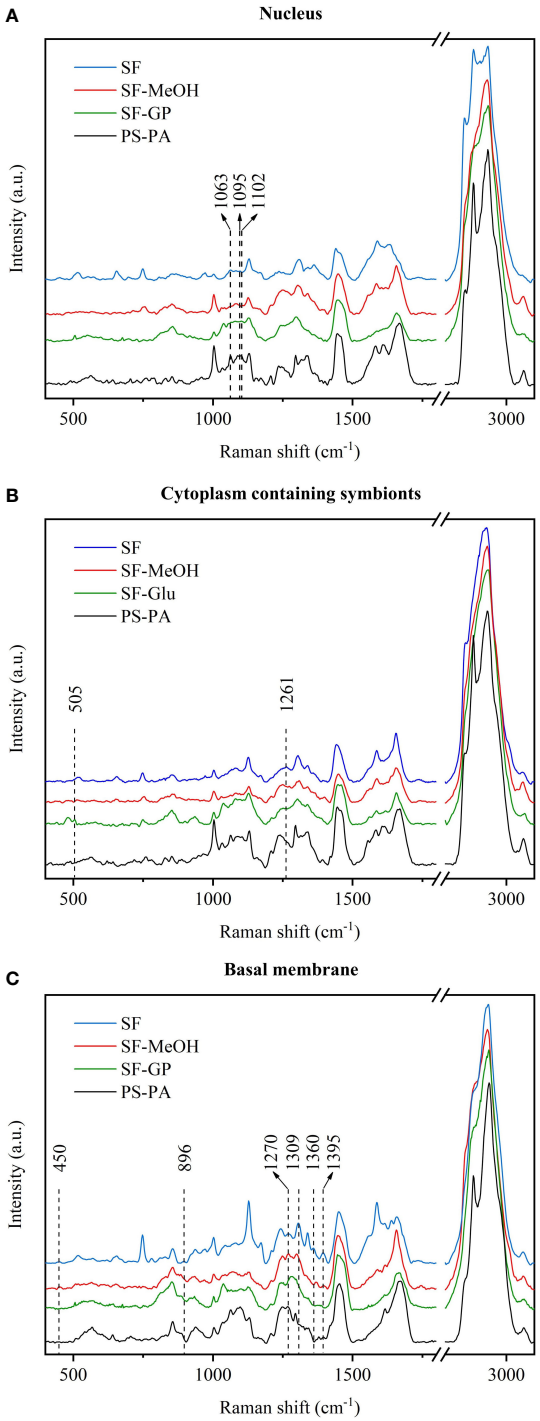


FIGURE 4
Comparative analysis of the characteristic spectra of different regions of the samples subjected to the four different preservation methods. **(A)** Nucleus. **(B)** Cytoplasm containing symbionts. **(C)** Basal membrane.

average spectra fitting, the spectra of the SF and SF-MeOH samples contain the same vibrational modes, with differences in content; however, both were dominated by unordered structures, probably due to the depolymerization of proteins in the frozen sections during direct snap freezing (Lyng et al., 2011).

The turn structure accounts for the greater proportion of the cytoplasm containing symbionts region in the SF-GP samples, and the nucleus region in the PS-PA samples (Supplementary Figure S4 and Supplementary Table S1). The β-sheet structures of the SF-GP and PS-PA samples were redundant with SF and SF-MeOH. The differences here indicate a significant change in the protein structure. The increased β-sheet amount indicates increased stiffness, easier tissue sectioning, and less fragmentation of the cells (Herrero et al., 2014). The SF-GP and PS-PA samples had a higher content of aromatic amino acids ring mode than the others, and the amino acids side chain structure accounted for the least. The lowest percentage of aromatic amino acids ring mode was found in the SF-MeOH sample. It may be because aldehyde fixation can cause cross-linking between functional groups in the side chains of amino acids, including cross-linking between adjacent amino acid chains and different positions of the same amino acid chain (Hobro and Smith, 2017). This process causes the protein to become insoluble. Concurrently, chemical changes occur, although it is usually more mechanically stable and suitable for subsequent processing, such as paraffin and resin embedding for sectioning.

In the basal membrane region, the content of β-sheet structures were high for all four methods in the three regions. Compared with SF, the immobilized samples showed lower contents of aromatic amino acid ring modes and amino acid side chain structures. The content of amino acid side chain structures in SF-GP was below the detection limit (Supplementary Figure S4 and Supplementary Table S1). It has been suggested that the basal membrane is the most morphologically stable region in the cell and can maintain its integrity without fixation during sectioning. It is speculated that the fixation of this region is less effective for the preservation of amino acids. This result is in good agreement with the conclusion that we found a strong link between the ease of obtaining integrity sections and fixation during our experiments.

Conclusion

Choosing an appropriate preservation method that causes minimal changes in the composition, distribution, and morphological integrity of biomolecules is important for the study of scientific problems (Mariani et al., 2009; Lyng et al., 2011). To the best of our knowledge, this is the first study that

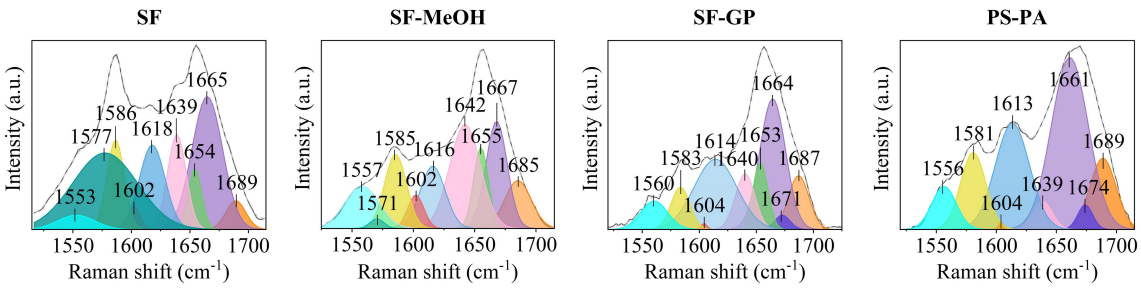


FIGURE 5
Fitting of the Raman spectra with peaks at 1,550-1,700 cm^{-1} obtained for different regions in the samples subjected to the four different preservation methods.

TABLE 2 Location of the Raman bands in the range of 1,550-1,700 cm^{-1}, and their assignments for the gills.

Raman shift (cm^{-1})	Assignment	Reference
1681-1691	β-sheet	(Goormaghtigh et al., 1994; Chen et al., 1995a; Chen et al., 1995b; Pelton and McLean, 2000; Barth and Zscherp, 2002; Faoláin et al., 2005; Chan et al., 2006; Lefèvre et al., 2007; Kochan et al., 2013; Yang et al., 2015; Rivas-Arancibia et al., 2017; Gieroba et al., 2020; Kowalska et al., 2020; Talaikis et al., 2020)
1671-1675	Turn	
1661-1668	C=C lipids stretching or Amide I (C=O stretching mode of protein, α-helix/random coil; stretching)	
1652-1655	α-helix	
1634-1644	Unordered	
1613-1622	Aromatic amino acids ring mode	
1601-1609	Amino acids side chains	
1580-1589	Amide II	
1570-1579	DNA	
1550-1568	Amide II	

TABLE 3 Quantitative estimation related to protein spatial structure in different preservation methods located at 1,600-1,700 cm^{-1}.

Assignment	SF	SF-MeOH	SF-GP	PS-PA
β-sheet	9.84 ± 0.16	14.61 ± 0.15	20.13 ± 0.12	29.89 ± 0.12
Turn	–	–	5.38 ± 0.06	10.39 ± 0.07
α-helix	20.79 ± 0.23	24.51 ± 0.19	25.02 ± 0.13	-
Unordered	32.61 ± 0.3	31.82 ± 0.22	20.87 ± 0.12	11.89 ± 0.08
Aromatic amino acids ring mode	28.97 ± 0.3	18.60 ± 0.17	26.24 ± 0.13	44.53 ± 0.15
Amino acids side chains	7.80 ± 0.1	10.46 ± 0.13	2.37 ± 0.04	3.31 ± 0.05

analyzing the effect of different preservation methods on the Raman signal of symbiotic deep-sea biological tissues. We analyzed the effects of four common preservation methods on Raman analysis by using the deep-sea mussel as a model (Table 4). To summarize, there are differences in the effect of preservation of different biomolecules in different areas of the tissue, probably due to differences in biomolecular composition. Our results highlight the strength of CRM in complex biomolecular composition recovery at the cellular scale and provide a guideline for other symbiotic deep-sea species with fragile tissues and complex metabolite compositions. The CRM can be extended to other species after the sample preservation and preparation refinement.

For this study model, the best preservation method needs to be selected on the basis of further experimental requirements. For spectroscopic analysis and other biological methods, SF is the best method to comprehensively analyze the composition of biological macromolecules in biological tissues if the focus is on metabolites. However, the sections can be damaged and sample degradation must be controlled. Fixatives can improve

TABLE 4 Comprehensive evaluation summary of different preservation methods.

	Raman imaging		Biomolecules	
	Success rate in obtaining integrity slices	Compartmentalization visualization	Composition	Content
SF	Low	The result is better.	Proteins may be denatured.	
SF-MeOH	Medium	The result is better.	Closest to SF, especially in the nucleus and basal membrane.	The concentration of S-S stretch, C-S stretch, tryptophan, lipid, carbohydrate, tyrosine, and amide II were reduced.
SF-GP	High	There is a significant impact.	Affects the composition of biomolecules, the spatial structure of proteins is changed, but some tertiary structures of proteins are preserved.	
PS-PA	High	There is a significant impact.	The presence of paraffin interference. There is a significant effect on the composition, and the spatial structure of the protein is the most altered.	

sectioning success but affect intracellular biomolecules to varying degrees and have different sensitivities for tracking different biochemical components. The SF-MeOH method is an acceptable choice when SF is not sufficient in obtaining good morphology in sections, and if the user is not concerned with the fact that SF-MeOH affects the composition of biomolecules. If the user is more concerned with overall morphological observations, SF-GP and PS-PA are the best choices. However, they are not well suited for metabolic studies.

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

WH conceived and designed the experiments, performed research, analyzed data, and wrote the manuscript. MW and XZ conceived and designed the experiments, analyzed data, wrote the manuscript, project administration, and funding acquisition. ML, ZZ, HC, and SX performed research. CL and ZL, project administration, and funding acquisition. All authors contributed to the article and approved the submitted version.

Funding

This research was supported by the following grants: the Strategic Priority Research Program of Chinese Academy of Sciences (XDA22050102, XDA19060402), Key project of Ocean Research Center, Chinese Academy of Sciences (COMS2020J03), the National Natural Science Foundation of China (41822604, 42076091), the Young Taishan Scholars Program (tsqn201909158).

Acknowledgments

We thank all the crews onboard the R/V Kexue for their assistance in sample collection and all the laboratory staff for continuous technical advice and helpful discussions.

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.2022.1077595/full#supplementary-material

References

Barth, A., and Zscherp, C. (2002). What vibrations tell about proteins. *Quart. Rev. Biophys.* 35 (4), 369–430. doi: 10.1017/S0033583502003815

Butler, H. J., Ashton, L., Bird, B., Cinque, G., Curtis, K., Dorney, J., et al. (2016). Using raman spectroscopy to characterize biological materials. *Nat. Protoc.* 11 (4), 664–687. doi: 10.1038/nprot.2016.036

Castillo, M. G., Salazar, K. A., and Joffe, N. R. (2015). The immune response of cephalopods from head to foot. *Fish Shellfish Immunol.* 46 (1), 145–160. doi: 10.1016/j.fsi.2015.05.029

Chan, J. W., Taylor, D. S., Zwerdling, T., Lane, S. M., Ihara, K., and Huser, T. (2006). Micro-raman spectroscopy detects individual neoplastic and normal hematopoietic cells. *Biophys. J.* 90 (2), 648–656. doi: 10.1529/biophysj.105.066761

Chen, X. G., Asher, S. A., Schweitzer-Stenner, R., Mirkin, N. G., Krimm, S., and de Mul, F. F. M. (1995a). UV Raman determination of the $\pi\pi^*$ excited state geometry of *N*-methylacetamide: Vibrational enhancement pattern. *J. Am. Chem. Soc* 117, 2884–2895. doi: 10.1021/ja00115a021

Cheng, W., Liu, M., Liu, H., and Lin, S. (2005). Micro-raman spectroscopy used to identify and grade human skin pilomatrixoma. *Microsc. Res. Tech.* 68 (2), 75–79. doi: 10.1002/jemt.20229

Chen, X. G., Schweitzer-Stenner, R., Asher, S. A., Mirkin, N. G., and Krimm, S. (1995b). Vibrational assignments of trans-*N*-methylacetamide and some of its deuterated isotopomers from band decomposition of IR, visible, and resonance raman spectra. *J. Phys. Chem.* 99, 3074–3083. doi: 10.1021/j100010a017

Chen, S., Wu, Q., Mishra, C., Kang, J., Zhang, H., Cho, K., et al. (2012). Thermal conductivity of isotopically modified graphene. *Nat. Mater.* 11 (3), 203–207. doi: 10.1038/NMAT3207

Childress, J. J., Fisher, C. R., Brooks, J. M., Kennicutt, M. N., Bidigare, R., and Anderson, A. E. (1986). A methanotrophic marine molluscan (bivalvia, mytilidae) symbiosis: Mussels fueled by gas. *Science* 233 (4770), 1306–1308. doi: 10.1126/science.233.4770.1306

Clemens, G., Hands, J. R., Dorling, K. M., and Baker, M. J. (2014). Vibrational spectroscopic methods for cytology and cellular research. *Analyst* 139 (18), 4411–4444. doi: 10.1039/c4an00636d

Czamara, K., Majzner, K., Pacia, M. Z., Kochan, K., Kaczor, A., and Baranska, M. (2015). Raman spectroscopy of lipids: A review. *J. Raman Spectrosc.* 46 (1), 4–20. doi: 10.1002/jrs.4607

Dodd, M. S., Papineau, D., Grenne, T., Slack, J. F., Rittner, M., Pirajno, F., et al. (2017). Evidence for early life in earth's oldest hydrothermal vent precipitates. *Nature* 543 (7643), 60. doi: 10.1038/nature21377

Dubilier, N., Bergin, C., and Lott, C. (2008). Symbiotic diversity in marine animals: the art of harnessing chemosynthesis. *Nat. Rev. Microbiol.* 6 (10), 725–740. doi: 10.1038/nrmicro1992

Dukor, R. K. (2006). Vibrational spectroscopy in the detection of cancer. *Handb. Vibrational Spectrosc.* 5, 3335–3361. doi: 10.1002/9780470027325.s8107

Du, J., Su, Y., Qian, C., Yuan, D., Miao, K., Lee, D., et al. (2020). Raman-guided subcellular pharmaco-metabolomics for metastatic melanoma cells. *Nat. Commun.* 11 (1), 4830. doi: 10.1038/s41467-020-18376-x

Faoláin, E. O. (2005). Raman spectroscopic evaluation of efficacy of current paraffin wax section dewaxing agents. *J. Histochem. Cytochem* 53 (1), 121–129. doi: 10.1369/jhc.4A6536.2005

Faoláin, E. O., Hunter, M. B., Byrne, J. M., Kelehan, P., McNamara, M., Byrne, H. J., et al. (2005). A study examining the effects of tissue processing on human tissue sections using vibrational spectroscopy. *Vibrational Spectrosc.* 38 (1-2), 121–127. doi: 10.1016/j.vibspec.2005.02.013

Fiedler, I. A. K., Casanova, M., Keplinger, T., and Busse, B. (2018). Effect of short-term formaldehyde fixation on raman spectral parameters of bone quality. *J. Biomed. Opt.* 23 (11), 1. doi: 10.1117/1.JBO.23.11.116504

Fujioka, N., Morimoto, Y., Takeuchi, K., Yoshioka, M., and Kikuchi, M. (2003). Difference in infrared spectra from cultured cells dependent on cell harvesting method. *Appl. Spectrosc.* 57 (2), 241–243. doi: 10.1366/000370203321535187

Galli, R., Uckermann, O., Koch, E., Schackert, G., Kirsch, M., and Steiner, G. (2014). Effects of tissue fixation on coherent anti-stokes raman scattering images of brain. *J. Biomed. Opt.* 19 (7), 071402. doi: 10.1117/1.JBO.19.7.071402

Geier, B., Sogin, E. M., Michellod, D., Janda, M., Kompauer, M., Spengler, B., et al. (2020). Spatial metabolomics of *in situ* host–microbe interactions at the micrometre scale. *Nat. Microbiol.* 5 (3), 498–510. doi: 10.1038/s41564-019-0664-6

Gieroba, B., Krysa, M., Wojtowicz, K., Wiater, A., Pleszczyńska, M., Tomczyk, M., et al. (2020). The FT-IR and raman spectroscopies as tools for biofilm characterization created by cariogenic streptococci. *Int. J. Mol. Sci.* 21 (11), 3811. doi: 10.3390/ijms21113811

Gomes Da Costa, S., Richter, A., Schmidt, U., Breuninger, S., and Hollricher, O. (2019). Confocal raman microscopy in life sciences. *Morphologie* 103 (341), 11–16. doi: 10.1016/j.morpho.2018.12.003

Goormaghtigh, E., Cabiaux, V., and Ruysschaert, J. (1994). Determination of soluble and membrane protein structure by fourier transform infrared spectroscopy. *Subcell Biochem.* 23, 329–362. doi: 10.1007/978-1-4615-1863-1_9

Høgset, H., Horgan, C. C., Armstrong, J. P. K., Bergholt, M. S., Torraca, V., Chen, Q., et al. (2020). *In vivo* biomolecular imaging of zebrafish embryos using confocal raman spectroscopy. *Nat. Commun.* 11 (1), 6172. doi: 10.1038/s41467-020-19827-1

Herrero, A. M., Ruiz-Capillas, C., Jiménez-Colmenero, F., and Carmona, P. (2014). Raman spectroscopic study of structural changes upon chilling storage of frankfurters containing olive oil bulking agents as fat replacers. *J. Agric. Food Chem.* 62 (25), 5963–5971. doi: 10.1021/jf501231k

Hobro, A. J., and Smith, N. I. (2017). An evaluation of fixation methods: Spatial and compositional cellular changes observed by raman imaging. *Vibrational Spectrosc.* 91, 31–45. doi: 10.1016/j.vibspec.2016.10.012

Hohman, T. C., McNeil, P. L., and Muscatine, L. (1982). Phagosome-lysosome fusion inhibited by algal symbionts of hydra viridis. *J. Cell Biol.* 94 (1), 56–63. doi: 10.1083/jcb.94.1.56

Huang, Z., McWilliams, A., Lui, H., McLean, D. I., Lam, S., and Zeng, H. (2003). Near-infrared raman spectroscopy for optical diagnosis of lung cancer. *Int. J. Cancer* 107 (6), 1047–1052. doi: 10.1002/ijc.11500

Huang, N., Short, M., Zhao, J., Wang, H., Lui, H., Korbelik, M., et al. (2011). Full range characterization of the raman spectra of organs in a murine model. *Opt. Express* 19 (23), 22892–22909. doi: 10.1364/OE.19.022892

Kallepitis, C., Bergholt, M. S., Mazo, M. M., Leonardo, V., Skaalure, S. C., Maynard, S. A., et al. (2017). Quantitative volumetric raman imaging of three dimensional cell cultures. *Nat. Commun.* 8 (1), 14843. doi: 10.1038/ncomms14843

Kochan, K., Marzec, K. M., Chruszcz-Lipska, K., Jasztal, A., Maslak, E., Musiolik, H., et al. (2013). Pathological changes in the biochemical profile of the liver in atherosclerosis and diabetes assessed by raman spectroscopy. *Analyst* 138 (14), 3885–3890. doi: 10.1039/c3an00216k

Kowalska, A. A., Berus, S., Szleszkowski, L., Kaminska, A., Kmiecik, A., Ratajczak-Wielgomas, K., et al. (2020). Brain tumour homogenates analysed by surface-enhanced raman spectroscopy: Discrimination among healthy and cancer cells. *Spectrochim Acta A Mol. Biomol. Spectrosc.* 231, 117769. doi: 10.1016/j.saa.2019.117769

Kozicki, M., Czepiel, J., Biesiada, G., Nowak, P., Garlicki, A., and Weselucha-Birczynska, A. (2015). The ring-stage of plasmodium falciparum observed in RBCs of hospitalized malaria patients. *Analyst* 140 (23), 8007–8016. doi: 10.1039/c5an01598g

Krafft, C., Neudert, L., Simat, T., and Salzer, R. (2005). Near infrared raman spectra of human brain lipids. *Spectrochim Acta A Mol. Biomol. Spectrosc.* 61 (7), 1529–1535. doi: 10.1016/j.saa.2004.11.017

Kuzmin, A. N., Pliss, A., and Prasad, P. N. (2014). Changes in biomolecular profile in a single nucleolus during cell fixation. *Anal. Chem.* 86 (21), 10909–10916. doi: 10.1021/ac503172b

Lakshmi, R. J., Kartha, V. B., Krishna, C. M., J., G. R. S., Ullas, G., and Devi, P. U. (2002). Tissue raman spectroscopy for the study of radiation damage: Brain irradiation of mice. *Radiat. Res.* 157 (2), 175–182. doi: 10.1667/0033-7587(2002)157[0175:TRSFTS]2.0.CO;2

Laming, S. R., Gaudron, S. M., and Duperron, S. (2018). Lifecycle ecology of deep-Sea chemosymbiotic mussels: A review. *Front. Mar. Sci.* 5 (282). doi: 10.3389/fmars.2018.00282

Lazarevic, J. J., Ralevic, U., Kukolj, T., Bugarski, D., Lazarevic, N., Bugarski, B., et al. (2019). Influence of chemical fixation process on primary mesenchymal stem cells evidenced by raman spectroscopy. *Spectrochim Acta A Mol. Biomol. Spectrosc.* 216, 173–178. doi: 10.1016/j.saa.2019.03.012

Lefèvre, T., Rousseau, M., and Pézolet, M. (2007). Protein secondary structure and orientation in silk as revealed by raman spectromicroscopy. *Biophys. J.* 92 (8), 2885–2895. doi: 10.1529/biophysj.106.100339

Lorenz, B., Wichmann, C., Stöckel, S., Rösch, P., and Popp, J. (2017). Cultivation-free raman spectroscopic investigations of bacteria. *Trends Microbiol.* 25 (5), 413–424. doi: 10.1016/j.tim.2017.01.002

Lyng, F., Gazi, E., and Gardner, P. (2011). "Preparation of tissues and cells for infrared and raman spectroscopy and imaging," in *RSC Analytical spectroscopy monographs*. Ed. D. Moss (Cambs, England: Royal Society of Chemistry), 147–185.

Malini, R., Venkatakrishna, K., Kurien, J., M. Pai, K., Rao, L., Kartha, V. B., et al. (2006). Discrimination of normal, inflammatory, premalignant, and malignant oral

tissue: A raman spectroscopy study. *Biopolymers* 81 (3), 179–193. doi: 10.1002/bip.20398

Maquelin, K., Kirschner, C., Choo-Smith, L. P., van den Braak, N., Endtz, H. P., Naumann, D., et al. (2002). Identification of medically relevant microorganisms by vibrational spectroscopy. *J. Microbiol. Methods* 51 (3), 255–271. doi: 10.1016/s0167-7012(02)00127-6

Mariani, M. M., Lampen, P., Popp, J., Wood, B. R., and Deckert, V. (2009). Impact of fixation on *in vitro* cell culture lines monitored with raman spectroscopy. *Analyst* 134 (6), 1154. doi: 10.1039/b822408k

Meade, A. D., Clarke, C., Draux, F., Sockalingum, G. D., Manfait, M., Lyng, F. M., et al. (2010). Studies of chemical fixation effects in human cell lines using raman microspectroscopy. *Anal. Bioanal. Chem.* 396 (5), 1781–1791. doi: 10.1007/s00216-009-3411-7

Moerner, W. E., and Orrit, M. (1999). Illuminating single molecules in condensed matter. *Science* 283 (5408), 1670. doi: 10.1126/science.283.5408.1670

Nishikori, K., Morioka, K., Kubo, T., and Morioka, M. (2009). Age- and morph-dependent activation of the lysosomal system and *Buchnera* degradation in aphid endosymbiosis. *J. Insect Physiol.* 55 (4), 351–357. doi: 10.1016/j.jinsphys.2009.01.001

Okotrub, K. A., Surovtsev, N. V., Semeshin, V. F., and Omelyanchuk, L. V. (2015). Raman spectroscopy for DNA quantification in cell nucleus. *Cytometry A.* 87 (1), 68–73. doi: 10.1002/cyto.a.22585

Pahlow, S., Meisel, S., Cialla-May, D., Weber, K., Rösch, P., and Popp, J. (2015). Isolation and identification of bacteria by means of raman spectroscopy. *Adv. Drug Deliv. Rev.* 89, 105–120. doi: 10.1016/j.addr.2015.04.006

Pelton, J. T., and McLean, L. R. (2000). Spectroscopic methods for analysis of protein secondary structure. *Anal. Biochem.* 277 (2), 167–176. doi: 10.1006/abio.1999.4320

Petersen, J. M., Zielinski, F. U., Pape, T., Seifert, R., Moraru, C., Amann, R., et al. (2011). Hydrogen is an energy source for hydrothermal vent symbioses. *Nature* 476 (7359), 176–180. doi: 10.1038/nature10325

Pijanka, J., Sockalingum, G. D., Kohler, A., Yang, Y., Draux, F., Parkes, G., et al. (2010). Synchrotron-based FTIR spectra of stained single cells. Towards a clinical application in pathology. *Lab. Invest.* 90 (5), 797–807. doi: 10.1038/labinvest.2010.8

Ponnudurai, R., Kleiner, M., Sayavedra, L., Petersen, J. M., Moche, M., Otto, A., et al. (2017). Metabolic and physiological interdependencies in the *Bathymodiolus azoricus* symbiosis. *ISME J.* 11 (2), 463–477. doi: 10.1038/ismej.2016.124

Rivas-Arancibia, S., Rodríguez-Martínez, E., Badillo-Ramírez, I., López-González, U., and Saniger, J. M. (2017). Structural changes of amyloid beta in hippocampus of rats exposed to ozone: A raman spectroscopy study. *Front. Mol. Neurosci.* 10. doi: 10.3389/fnmol.2017.00137

Shao, Q., Fan, Y., Yang, L., and Qin Gao, Y. (2012). From protein denaturant to protectant: Comparative molecular dynamics study of alcohol/protein interactions. *J. Chem. Phys.* 136 (11), 115101. doi: 10.1063/1.3692801

Silveira, L., Sathaiah, S., Zângaro, R. A., Pacheco, M. T. T., Chavantes, M. C., and Pasqualucci, C. A. G. (2002). Correlation between near-infrared raman spectroscopy and the histopathological analysis of atherosclerosis in human coronary arteries. *Lasers Surg. Med.* 30 (4), 290–297. doi: 10.1002/lsm.10053

Stöcke, S., Kirchhoff, J., Neugebauer, U., Rösch, P., and Popp, J. (2016). The application of raman spectroscopy for the detection and identification of microorganisms. *J. Raman Spectrosc.* 47, 89–109. doi: 10.1002/jrs.4844

Stone, N., Kendall, C., Smith, J., Crow, P., and Barr, H. (2004). Raman spectroscopy for identification of epithelial cancers. *Faraday Disc.* 126, 141. doi: 10.1039/b304992b

Sugeta, H. (1975). Normal vibrations and molecular conformations of dialkyl disulfides. *Spectrochimica Acta Part A: Mol. Spectrosc.* 31 (11), 1729–1737. doi: 10.1016/0584-8539(75)80116-4

Sun, Y., Wang, M., Li, L., Zhou, L., Wang, X., Zheng, P., et al. (2017). Molecular identification of methane monooxygenase and quantitative analysis of methanotrophic endosymbionts under laboratory maintenance in *Bathymodiolus platifrons* from the south China Sea. *PeerJ* 5, e3565. doi: 10.7717/peerj.3565

Sun, J., Zhang, Y., Xu, T., Zhang, Y., Mu, H., Zhang, Y., et al. (2017). Adaptation to deep-sea chemosynthetic environments as revealed by mussel genomes. *Nat. Ecol. Evol.* 1 (5), 121. doi: 10.1038/s41559-017-0121

Talaikis, M., Strazdaitė, S., Žiaunys, M., and Niaura, G. (2020). Far-off resonance: Multiwavelength raman spectroscopy probing amide bands of amyloid-β-(37–42) peptide. *Molecules* 25 (15), 3556. doi: 10.3390/molecules25153556

Talari, A. C. S., Movasaghi, Z., Rehman, S., and Rehman, I. U. (2015). Raman spectroscopy of biological tissues. *Appl. Spectrosc. Rev.* 50 (1), 46–111. doi: 10.1080/05704928.2014.923902

Tfayli, A., Gobinet, C., Vrabie, V., Huez, R., Manfait, M., and Piot, O. (2009). Digital dewaxing of raman signals: discrimination between nevi and melanoma spectra obtained from paraffin-embedded skin biopsies. *Appl. Spectrosc.* 63 (5), 564–570. doi: 10.1366/000370209788347048

Trench, R. K. (1971). The physiology and biochemistry of zooxanthellae symbiotic with marine coelenterates. i. the assimilation of photosynthetic products of zooxanthellae by two marine coelenterates. *Proc. R. Soc B* 177 (1047), 225–235. doi: 10.1098/rspb.1971.0024

Van Wart, H. E., Lewis, A., Scheraga, H. A., and Saeva, F. D. (1973). Disulfide bond dihedral angles from raman spectroscopy. *Proc. Natl. Acad. Sci.* 70 (9), 2619–2623. doi: 10.1073/pnas.70.9.2619

Van Wart, H. E., and Scheraga, H. A. (1986). Agreement with the disulfide stretching frequency-conformation correlation of sugeta, go, and miyazawa. *Proc. Natl. Acad. Sci.* 83 (10), 3064–3067. doi: 10.1073/pnas.83.10.3064

Wang, H., Zhang, H., Zhong, Z., Sun, Y., Wang, M., Chen, H., et al. (2021). Molecular analyses of the gill symbiosis of the bathymodiolin mussel *Gigantidas platifrons*. *iScience* 24 (1), 101894. doi: 10.1016/j.isci.2020.101894

Wetzel, D. L., and LeVine, S. M. (1999). Microspectroscopy - imaging molecular chemistry with infrared microscopy. *Science* 285 (5431), 1224–1225. doi: 10.1126/science.285.5431.1224

Wiercigroch, E., Szafraniec, E., Czamara, K., Pacia, M. Z., Majzner, K., Kochan, K., et al. (2017). Raman and infrared spectroscopy of carbohydrates: A review. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* 185, 317–335. doi: 10.1016/j.saa.2017.05.045

Xu, T., Feng, D., Tao, J., and Qiu, J. (2019). A new species of deep-sea mussel (Bivalvia: Mytilidae: *Gigantidas*) from the south China Sea: Morphology, phylogenetic position, and gill-associated microbes. *Deep-Sea Res. Part I Oceanogr. Res. Pap.* 146, 79–90. doi: 10.1016/j.dsr.2019.03.001

Xu, T., Gan, Q., Wu, B., Yin, M., Xu, J., Shu, X., et al. (2020). Molecular basis for PI(3,5)P2 recognition by SNX11, a protein involved in lysosomal degradation and endosome homeostasis regulation. *J. Mol. Biol.* 432 (16), 4750–4761. doi: 10.1016/j.jmb.2020.06.010

Yang, H., Yang, S., Kong, J., Dong, A., and Yu, S. (2015). Obtaining information about protein secondary structures in aqueous solution using Fourier transform IR spectroscopy. *Nat. Protoc.* 10 (3), 382–396. doi: 10.1038/nprot.2015.024

frontiers | Frontiers in Marine Science

TYPE Original Research
PUBLISHED 09 January 2023
DOI 10.3389/fmars.2022.1073264

OPEN ACCESS

EDITED BY
Senjie Lin,
University of Connecticut,
United States

REVIEWED BY
Mikhail Solovyev,
Institute of Systematics and Ecology of Animals (RAS), Russia
Guoxing Nie,
Henan Normal University, China

*CORRESPONDENCE
Yanjie Zhang
yanjiezhang@hainanu.edu.cn
Zhiqiang Guo
guozq@hainanu.edu.cn

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

RECEIVED 18 October 2022
ACCEPTED 19 December 2022
PUBLISHED 09 January 2023

CITATION
Zou Y, Wu D, Wei L, Xiao J, Zhang P, Huang H, Zhang Y and Guo Z (2023) Mucus-associated microbiotas among different body sites of wild tuna from the South China Sea. *Front. Mar. Sci.* 9:1073264. doi: 10.3389/fmars.2022.1073264

Mucus-associated microbiotas among different body sites of wild tuna from the South China Sea

Ying Zou[1], Di Wu[1], Lu Wei[1], Juan Xiao[1], Pengfei Zhang[1], Hai Huang[2], Yanjie Zhang[1]* and Zhiqiang Guo[1]*

[1]School of Life Sciences, State Key Laboratory of Marine Resource Utilization in South China Sea, Hainan University, Haikou, China, [2]Key Laboratory of Utilization and Conservation for Tropical Marine Bioresources, Hainan Key Laboratory for Conservation and Utilization of Tropical Marine Fishery Resources, College of Fisheries and Life Science, Hainan Tropical Ocean University, Sanya, China

The mucus-associated symbionts have profound impacts on the pathogen defense, metabolism, and development of aquatic animals. To understand the microbial structure of regional endothermic fish, a total of 52 samples from the skin, oral, gill, and hindgut of wild tuna *Thunnus albacares* and *T. obesus* were determined by 16S amplicon sequencing. The results showed the diversity and composition of microbial communities varied in the four different body sites of tunas, with a greater heterogeneity between the external surface and the gut. Phyla Proteobacteria, Firmicutes, Actinobacteria and genus *Acinetobacter* were found in high relative abundance in all body sites. The other abundant taxa were enriched in different body sites, such as *Lactobacillus* and *Kocuria* in the skin and *Geobacillus* in the gut. The core taxa interacted with each other to different degrees in the four body sites, which may be related to species' co-evolution and microbial community stability. Finally, the correlation between biomarkers and COG functions highlighted the importance of microbial biomarkers to the host. This work firstly characterized the microbial feature in different body sites of wild tunas, providing a foundational dataset to understand the microbial role in endothermic fish and to find key microbial components beneficial to farmed tunas.

KEYWORDS

***Thunnus*, mucosal microbiome, core taxa, biomarkers, 16S rRNA**

Introduction

The host and its diverse microbial communities are so closely linked that they are often described as a single entity: the holobiont, which is considered a unit of selection in host-microbiota co-evolution (Rosenberg and Zilber-Rosenberg, 2018). In that pattern, host genetics is considered a dominant driver in shaping host microbiotas, and

microbiotas colonized in different body mucus play a paramount role in host phenotypes and traits (Lynch and Hsiao, 2019). Gut microbiotas are well-known contributors of host symbionts to a broad set of functions related to host immunity, metabolism, and development (Banerjee and Ray, 2017; Wang et al., 2018). The outer mucosal microbiotas serve as a preliminary dynamic interface between the fish and the environment and have essential roles in resisting pathogen invasion (Ross et al., 2019), and take surface mucus as an intermediate niche between the water and digestive tract (Carda-Dieguez et al., 2017).

The microbial stability in the mucus layer is pivotal for host health promotion and mutualistic microbiotas configuration (Wang et al., 2018; Fassarella et al., 2021). It depends on the high microbial diversity and abundance, which makes sure that microorganisms with similar functions can act as substitutes when probiotics are reduced (Fassarella et al., 2021). Further, the complex interactions within microbial communities from mutualism to competition, and the symbiotic relationship between microbes and their host, are essential for host homeostasis (Foster et al., 2017; Fassarella et al., 2021). For instance, gut microbiotas members, especially Firmicutes, were essential to lipid droplet formation and fatty acid uptake in the intestinal epithelium of zebrafish (Semova et al., 2012). The commensal skin-microbiotas are key factors in skin wounds healing, mediated by triggering IFN-dependent innate repair responses (Di Domizio et al., 2020). Additionally, core microbial community, including bacteria, Archaea, microeukaryotes and viruses, are more relatively conserved in composition and function, which is expected to stabilize the ecosystem (Shetty et al., 2017).

Fish can acquire microbiotas through surrounding water from the early developmental phases (Spor et al., 2011). Its larval microbiotas depend greatly on water quality, salinity, nutrients, and oxygen content (Dehler et al., 2017; Wang et al., 2018; Sylvain et al., 2020). Once the microbiome dynamic balance is affected by environmental changes, such as water contamination, living space limitation, or antibiotics exposure, the marked compositional change and diversity decrease will be found in microbial communities (Ross et al., 2019), resulting in host chronic diseases and recurrent infection over time. For example, common pathogens *Vibrio*, *Flavobacterium*, *Arcobacter* and *Allorhizobium* are persistently dominant in the ulcer mucus of unhealthy fish (Karlsen et al., 2017; Sultana et al., 2022). Anthropogenic antibiotics intervention is generally used to resist pathogen expansion but may lead to the accumulation of antibiotic resistance genes in the microbiome (Willmann et al., 2019). In a microbiota-mediated way, pathogen overgrowth can be prevented by the microbial colonization resistance, which is performed by means of niche and nutrient competitions, conjugation-dependent killing, and antagonistic molecules (Buffie and Pamer, 2013). A previous study has confirmed that *Bacteroides* spp. can produce short-chain fatty acid (SCFA) propionate and adjust intracellular pH homeostasis to directly inhibit *Salmonella Typhimurium* growth (Jacobson et al., 2018). Moreover, Stressmann et al. (2020) demonstrated that conventional zebrafish with 10 culturable bacterial species showed sufficiently reduced infection susceptibility than germ-free individuals. Therefore, the specific microbial components are critical for the resilience and stability of the microbial community when encountering perturbance.

Tuna is one of the most commercially valuable marine fish with high nutrients (FAO, 2020). Open-net pens aquaculture industry for tuna is in high demand because of limited wild resources. However, cultured tunas with high density are more susceptible to opportunistic pathogenic bacteria and infectious parasites (Nowak et al., 2021), since their living environment was changed and the structure of host-associated microbial communities is altered (Minich et al., 2020a). However, our understanding of mucosal microbial symbionts in wild tuna is limited, especially from multiple body sites.

Here, yellowfin tuna (*Thunnus albacares*) and bigeye tuna (*Thunnus obesus*) are chosen as wild hosts to study mucus-associated symbionts. They have similar habitats, shapes, and diets, and are common in the South China Sea (Varela et al., 2017; Ohshimo et al., 2018). The primary objectives of this study were (1) to compare the diversity and structure of the mucus-associated microbial community in three dimensionalities (intraspecies, interspecies, and interindividual), and (2) to determine the composition and feature of microbiotas of healthy tunas among different body sites, and (3) to detect potential microbial biomarkers of tuna's health status. This study will let us better understand the relationship between the symbiotic microbes and host health, and further contribute to the aquaculture industry of tuna.

Materials and methods

Sample collection

Wild tunas (11 yellowfin tunas and 4 bigeye tunas) were captured by line lures from the South China Sea (17°24′N, 110°36′E) in August 2021. The two species were classified preliminarily on the spot and the final identification was determined by comparing cytochrome C oxidase subunit I (COI) gene sequences of muscle tissues to NCBI (Supplementary Table 1). The detailed process of mucus bacteria sampling is shown below. When tuna was hooked, external (skin, oral, and gill) mucus bacteria were wiped from alive fish immediately by sterile cotton swabs, and each site was wiped with at least two swabs to make sure enough mucus to extract DNA. After dissection, the hindgut mucus-associated bacteria were collected from hindgut contents squeezed out by sterilized scissors and tweezers. Sampling sites among individuals were the same. According to fork length, body weight, and species, the tuna individuals were divided into three groups and a total of 60 biological samples were collected (Table 1). We also collected environmental microorganisms from 40 m depth by filtering 250 mL of seawater through a

TABLE 1 Samples information collected from tunas.

Tunas	Individuals	Fork length (cm)	Weight (kg)	Skin	Oral	Gill	Hindgut
Juvenile *Thunnus albacares*	A1	38.88	0.95	★	☆	★	★
	A2	36.52	0.97	★	★	★	★
	A3	39.54	1.12	★	★	★	★
	A4	40.78	1.18	☆	★	★	★
	A5	51.88	3.56	★	★	★	★
	A6	54.76	3.96	★	★	★	★
Adult *Thunnus albacares*	B1	62.22	4.62	★	★	★	★
	B2	72.28	6.54	★	★	★	☆
	B3	74.85	8.12	★	★	★	★
	B4	75.10	8.78	★	★	☆	☆
	B5	86.54	11.18	★	★	★	★
Adult *Thunnus obesus*	C1	82.77	12.15	★	★	★	★
	C2	85.78	13.55	★	★	★	☆
	C3	94.12	16.52	★	★	★	★
	C4	98.52	18.56	★	☆	☆	★

"★" stands for sequencing is successful; "☆" stands for sequencing is failed.

0.22-μm-pore-size polycarbonate membrane, and 6 seawater samples were preserved. All the biological and environmental samples were quickly frozen in liquid nitrogen and transferred to a -80°C refrigerator until the next procedure.

16S amplicon sequencing

The DNA extraction was processed using the FastDNA® Spin Kit for Soil (MP Biomedicals, Norcross, GA, U.S.) according to the manufacturer's protocols. To obtain amplification of V3-V4 hypervariable regions of the 16S rRNA gene, triplicate PCR reactions of each sample were started at 95°C for denaturation and followed by 27 cycles at 95°C for 30 s, annealed at 55°C for 30 s, elongated at 72°C for 45 s, and finalized with an extension at 72°C for 10 min. Each 20 μL PCR mixture contained 4 μL of 5×FastPfu Buffer, 2 μL of 2.5 mM dNTPs, 0.8 μL of 5 μM primers (338F: 5′-ACTCCT ACGGGAGGCAGCAG-3′ and 806R: 5′-GGACTACHVG GGTWTCTAAT-3′; Liu et al., 2016), 0.4 μL of FastPfu Polymerase (TransGen, Beijing, China), and 10 ng of template DNA. The amplified fragments were sent to Majorbio Bio-Pharm Technology Co. Ltd. (Shanghai, China) for paired-end sequencing (2×300 bp) using the Illumina MiSeq platform (Illumina, San Diego, USA). All raw reads were deposited in the NCBI Sequence Read Archive (SRA) database (BioProject ID: PRJNA884520 and PRJNA902642).

Bioinformatics processing

Quality control was done by fastp version 0.20.0 (Chen et al., 2018). The reads were truncated when their average quality score was <20 and filtrated when the read length was <50 bp after quality-controlling. PE reads were merged according to their overlap (>10 bp, allowing 2 bp mismatching) by FLASH version 1.2.7 (Magoč and Salzberg, 2011). Sequences of each sample were screened according to barcodes (exactly matching) and primers (allowing 2 bp mismatching).

Operational taxonomic units (OTUs) were clustered with a 97% similarity cut-off by Uparse version 7.0.1090 (Edgar, 2013). The taxonomy of each 16S rRNA gene sequence was analyzed by the RDP Classifier Bayesian algorithm against the Silva v1.3.8 16S rRNA database (default confidence threshold of 0.7) (Wang et al., 2007). After that, the OTUs were normalized to the smallest library to eliminate sample heterogeneity, by "subsample" function in Mothur version 1.30.2 (Schloss et al., 2009) following the method in Minniti et al. (2017).

Data analysis

The alpha diversity of the microbiome was estimated using Mothur. Significant differences of alpha diversity indices were tested by Welch's t-test at the OTU level. Alpha diversity indices were visualized by Graphpad Prism (version 9.0.0). Beta

diversity analysis was based on Bray-Curtis distance and visualized by principal coordinates analysis (PCoA) to conduct clustering at the OTU level of the sample community. Pairwise comparisons of beta diversity distances between groups were calculated by both Bray-Curtis metrics and weighted Unifrac metrics (Rosado et al., 2019). Briefly, Bray-Curtis distance was calculated by the vegan package in R version 4.0.3 based on the OTUs table. Weighted UniFrac distance was calculated by GUniFrac package, based on the phylogenetic tree constructed by the OTUs table, using phyloseq package. Significant differential species among four body parts were analyzed by Kruskal-Wallis H Test at the genus level by SPSS version 26.0 (Zhang et al., 2019). To analyze the difference between host-associated microbiotas and seawater microbiotas, Spearman correlation heatmap was analyzed in R, using the "corr.test" function and pheatmap package. To find the main species contributing microbial difference between body sites and seawater, similarity percentage (SIMPER) was utilized by PRIMER version 5.2.8 (Gardner et al., 2019).

Linear discriminant analysis Effect Size (LEfSe) was conducted to estimate the biomarkers from phylum to genus with the threshold of linear discriminant analysis (LDA) score of 3.5 (default setting), under the premise of the Kruskal-Wallis H Test (Segata et al., 2011).

The network analysis was based on core OTUs that occurred in at least 70% of all samples in each group. The network was constructed at the genus level by the NetworkX version 1.11 (Hagberg et al., 2008). The top 50 abundant genera in each body site were connected by Spearman's correlation coefficient ($r>0.75$, $p<0.05$).

The function of the OTUs was predicted by the Phylogenetic Investigation of Communities by Reconstruction of Unobserved States (PICRUSt version 2.0; Douglas et al., 2020). The COG (Cluster of Orthologous Group) annotation was obtained by mapping to the EggNOG library version 5.0. The relative abundance difference of COG function classifications was determined by the Kruskal-Wallis H Test.

Results

In this study, 86.67% (52/60) of biological samples from four body sites of tunas were sequenced successfully, including 14 from skin, 13 from oral, 13 from gill, and 12 from hindgut. A total of 1,520,324 sequences were subsampled and 10,208 OTUs were obtained, representing 61 phyla, 182 classes, 445 orders, 786 families, and 1,924 genera. The number of sequences per sample ranged from 30,000 to 60,000, with an average length of 417 bp.

Variation of the microbial community does not correlate with host body size or species

The taxonomy composition of OTUs was used to test whether the body size or species of the host is related to the variation of the microbial community in different body sites. The results showed that no significant difference was found in mucus-associated microbial diversity with body size or species (Supplementary Table 2. Shannon index, $p>0.05$; PERMANOVA, $p>0.05$). Therefore, the following analysis will focus on the correlation between the microbial community and four body sites, i.e. group S (skin), group O (oral), group G (gill), and group H (hindgut).

Microbial diversity in different body sites

The results of six alpha diversity indices showed significantly distinct richness and diversity across body sites (Figure 1). In detail, the Sobs index detected four groups that had significant differences in community richness (Figure 1A. Welch's t-test, S vs O: $p=0.0018$; S vs G: $p=0.0008$; O vs H: $p=0.0018$; G vs H: $p=0.0009$). And similar significances were observed in Chao index (Figure 1B). The Shannon index revealed a significantly high diversity in gill compared to skin or hindgut (Figure 1C, Welch's t-test, G vs S: $p=0.0017$; G vs H: $p=0.0284$). The Shannoneven index indicated skin had a significant difference with gill ($p=0.0143$) and hindgut ($p=0.0378$) (Figure 1D). Significant differences of community coverage were detected by the Good's coverage index (Figure 1E. Welch's t-test, S vs O: $p=0.0031$; O vs H: $p=0.0002$; G vs H: $p=0.0198$). The Pd index showed a significant difference in phylogenetic diversity for all pairwise comparisons except oral vs gill (Figure 1F). No significant difference in microbial diversity index was found between the oral and gill (Figures 1A–F). In addition, the difference in diversity of microbiotas between the host surface mucus and the environment was compared. The results showed environmental microbial indices were significantly less than that in the skin (Welch's t-test, Sobs: $p=0.0002$; Shannon: $p=0.0001$; Shannoneven: $p=0.0007$; Pd: $p=0.0006$) (Supplementary Table 3).

The results of PCoA analysis showed that the factor of body sites explained approximately 20% of difference of microbial structure: 20.35% for four body sites and 19.2% without hindgut (Figures 2A, B). According to Figure 2A, the cluster of hindgut samples was completely separated from the external samples. From Figure 2B, the overlap of confidence ellipse within skin and gill or oral was much less than that of gill and oral, and correlation heatmap showed highest Spearman correlation coefficient between gill and oral (Supplementary Figure 1),

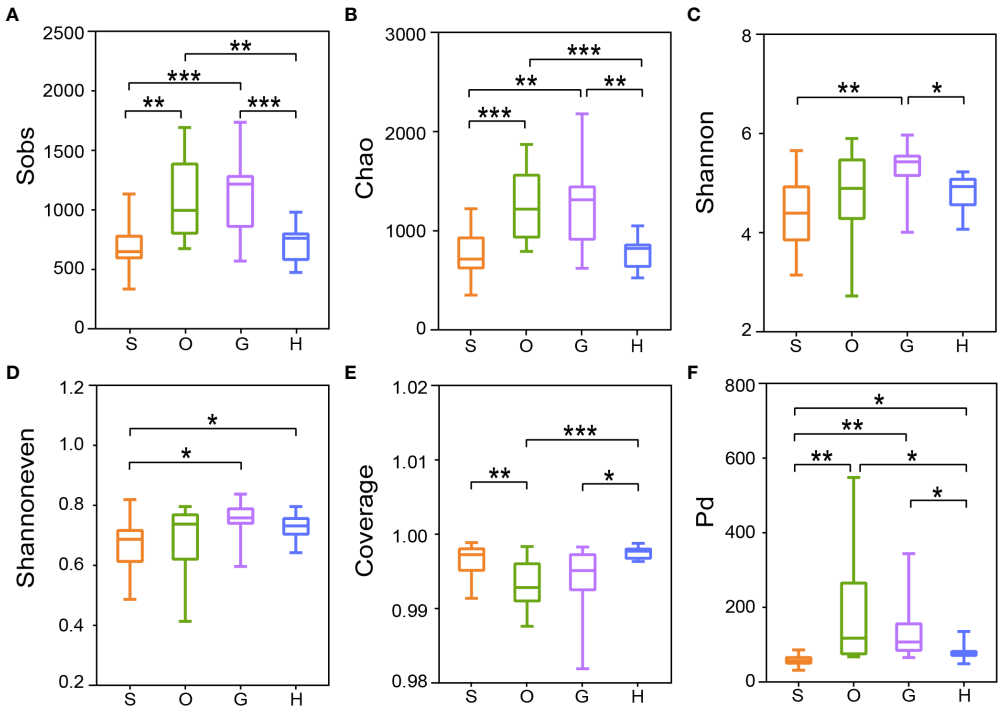


FIGURE 1
Alpha diversity of bacterial communities in four body sites of tunas. The statistical significance of six alpha diversity indices of **(A)** Sobs index, **(B)** Chao index, **(C)** Shannon index, **(D)** Shannoneven index, **(E)** Coverage index, and **(F)** Pd index were calculated by the Welch's t-test (*$p \leq 0.05$; **$p \leq 0.01$; ***$p \leq 0.001$). S, skin; O, oral; G, gill; H: hindgut.

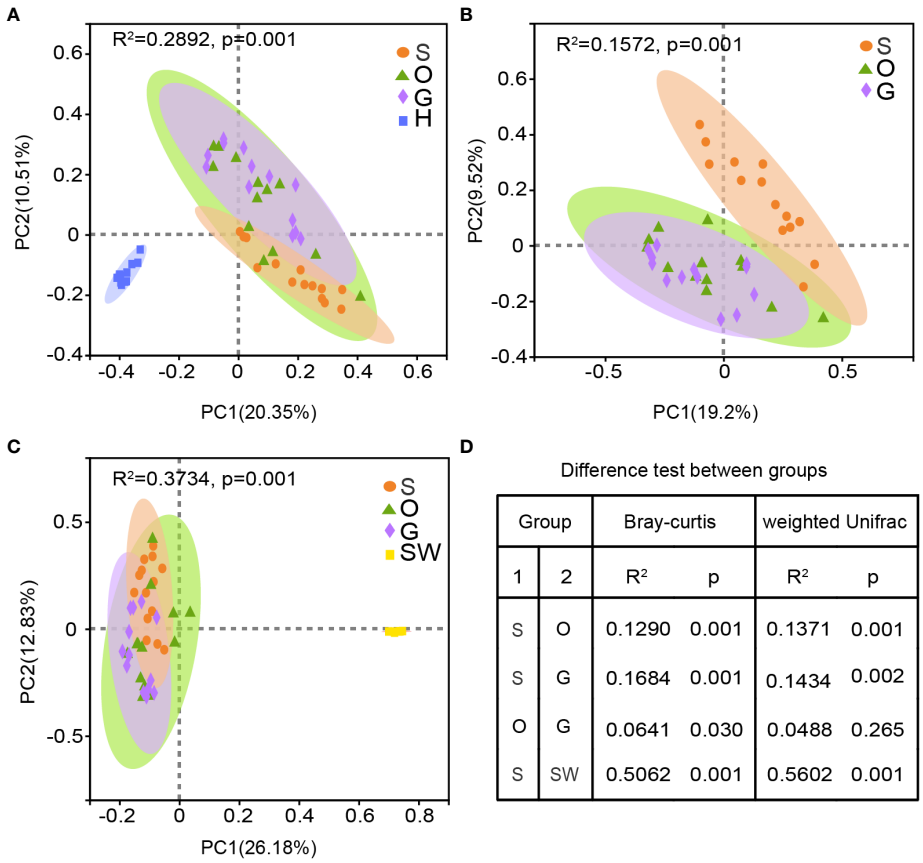
indicating the microbiome communities of the oral and gill were similar. The statistically significant differences were shown among all pairwise comparisons based on Bray-Curtis metrics and weighted Unifrac metrics, except oral vs gill for weighted Unifrac metrics (Figure 2D).

Microbial diversity in external body sites showed highly taxonomic differentiation with seawater (Figure 2C), and host-associated microbiome showed weak correlation with environment microbiome (Supplementary Figure 1). According to similarity percentage analysis (SIMPER), hindgut had a highest average dissimilarity with seawater at the genus level (Supplementary Table 4). There were 46 species cumulatively contributed 60% dissimilarity in all groups, with *Prochlorococcus* MIT9313 contributed 21.19%~22.63% dissimilarity in each group.

Microbial relative abundance in different body sites

Bacteria belonging to 61 phyla were detected from all mucus-associated samples, including 44 from skin, 53 from oral, 50 from gill, and 48 from hindgut, respectively (Supplementary Figure 2A). There were 62% (38/61) of phyla shared in all body sites. Bacteria belonging to 5 phyla (i.e. Proteobacteria, Firmicutes, Actinobacteria, Bacteroidota, and Cyanobacteria) had high relative abundance and they account for about 90% of microbial composition in four groups (Figure 3A). The abundance of Proteobacteria in the oral was observed significantly higher than skin (Figure 3C. Kruskal-Wallis H Test, O *vs* S: p=0.0499), whereas Firmicutes in the oral was significantly lower (Kruskal-Wallis H Test, O *vs* S: p=0.0397; O *vs* H: p=0.0028). For Actinobacteria, no significant difference was found across all groups. Much fewer phyla (23) were identified from seawater compared with that from the host. Bacteria of Cyanobacteria and Proteobacteria had the highest relative abundance in seawater, making up more than 80% of the microbiome community (Figure 3A).

The number of genera ranged from 1053 to 1455 in each group, and the four groups shared 644 genera (Supplementary Figure 2B). Among the external group, three genera (i.e. *Acinetobacter*, *Kocuria* and *Geobacillus*) constituted 20.04% ~28.56% of the surface mucus microbial community (Figure 3B). In the hindgut, the top 3 abundant genera were *Geobacillus*, *Acinetobacter* and g_norank_Propionibacteriaceae,



Group		Bray-curtis		weighted Unifrac	
1	2	R^2	p	R^2	p
S	O	0.1290	0.001	0.1371	0.001
S	G	0.1684	0.001	0.1434	0.002
O	G	0.0641	0.030	0.0488	0.265
S	SW	0.5062	0.001	0.5602	0.001

FIGURE 2
Beta diversity of bacterial communities in four body sites of tunas and seawater. **(A)** PCoA analysis of all microbiome across the external and hindgut. **(B)** PCoA analysis of external mucus microbiome. **(C)** PCoA analysis of external mucus microbiome and seawater. **(D)** Pairwise comparisons of beta diversity distances between groups are calculated by the Bray-Curtis metrics and weighted Unifrac metrics. S, skin; O, oral; G, gill; H, hindgut; SW, seawater.

accounting for 28.44% of the microbial community (Figure 3B). In addition, the most dominant genus *Acinetobacter* accounted for 9.44%~16.54% for each external group and was highest in oral. It showed no marked difference between the four groups. Kruskal-Wallis H test screened out 15 abundant genera showing significant difference in four groups (i.e. *Kocuria*, *Geobacillus*, *Lactobacillus*, *Pseudomonas*, *Paracoccus*, g_norank_f_Propionibacteriaceae, *Deinococcus*, *Chroococcidiopsis*_PCC_7203, *Macrococcus*, *Sphingomonas*, *Brevundimonas*, *Knoellia*, Supplementary Figure 3). Among them, *Kocuria* had the highest relative abundance in the skin (*Kocuria*: 11.32%) compared to that in the hindgut (*Kocuria*: 0.09%) (Figure 3D). *Lactobacillus* was one of the most abundant genera in the skin and accounted for 8.29% of the community, showing a significant difference from other groups (Figure 3D). The relative abundance of *Geobacillus* was significantly highest in the hindgut (H: 12.65%; S: 2.51%; O: 3.08%; G: 4.32%).

Core microbes and co-occurrence networks in each group

A total of 324 core OTUs (S: 115; O: 132; G: 176; H: 140) were presented in at least 70% replicates of mucus-associated samples. The relative abundance in each group was shown in Supplementary Figure 4. In each group, they composed only 2.46%~4.13% of OTUs but covered 55.20%~71.58% of the sequence reads. Proteobacteria (117 OTUs), Firmicutes (91 OTUs), Actinobacteria (66 OTUs), and Bacteroidota (21 OTUs) were the four most abundant phyla within the core taxa. Among the 324 core OTUs, 13 OTUs were present in all replicates. They represented only 0.13% of all OTUs but covered 18.66% of all the sequence reads. They were *Acinetobacter* (OTU6807, OTU6806, OTU7089, OTU6848), *Pseudomonas* (OTU7873, OTU1419), *Geobacillus* (OTU6817), *Knoellia* (OTU7006), *Microbacterium* (OTU7122), *Escherichia-Shigella* (OTU6820), *Brevundimonas*

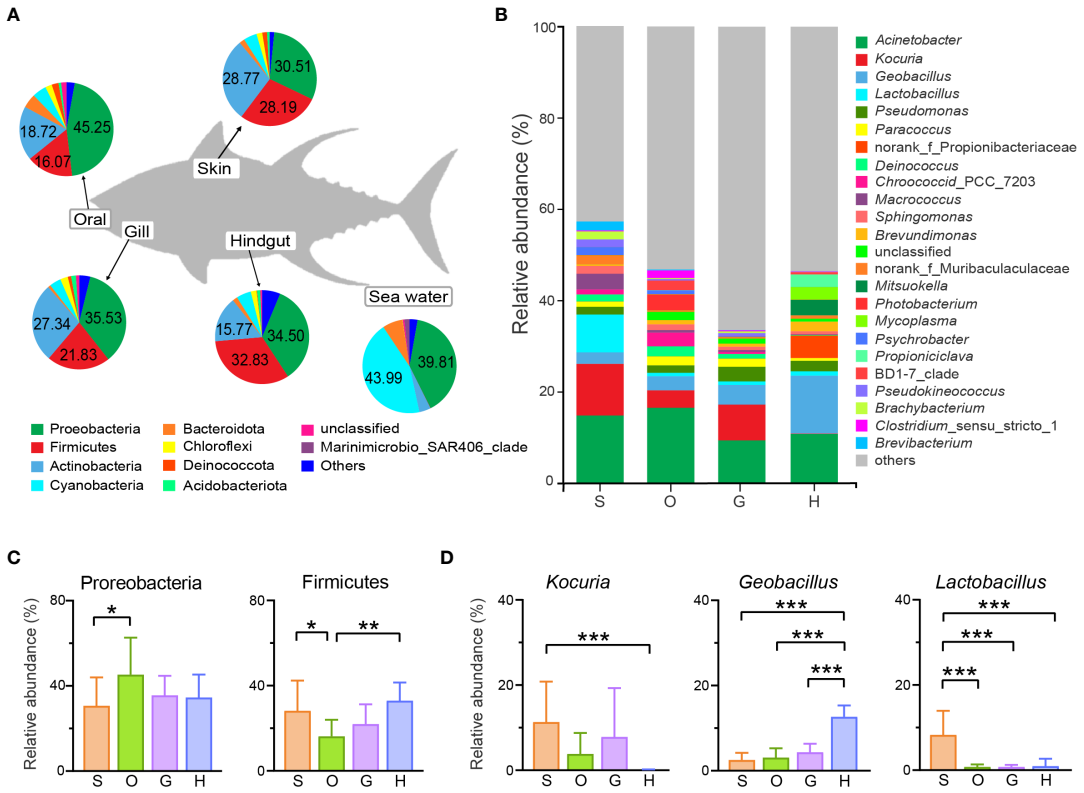


FIGURE 3
Relative abundance of the bacterial community. The microbial abundance of tunas and seawater at the **(A)** phylum level and **(B)** genus level. OTUs with low relative abundance (<1.5%) were clustered in "others". The significant differences analysis of **(C)** abundant phyla and **(D)** abundant genera. Statistical test was carried out by the Kruskal-Wallis H Test followed by Bonferroni correction (*$p \leq 0.05$; **$p \leq 0.01$; ***$p \leq 0.001$). S, skin; O, oral; G, gill; H, hindgut.

(OTU7071), *Staphylococcus* (OTU6907), and unclassified OTU6955 (Supplementary Table 5).

The co-occurrence network of core OTUs in each group was analyzed at the genus level (Figure 4). Different nodes and edges numbers presented significant differences in four groups: 37 and 56 in skin, 40 and 71 in oral, 38 and 61 in gill, 33 and 31 in hindgut. Generally, the positive edges of the network were much more than the negative edges in each group. The ratio of negative correlations was the lowest in oral (5.63%) and highest in hindgut (25.81%), indicating more competition relationships within the hindgut microbial community. For the most abundant core OTUs at the genus level, in the skin, *Lactobacillus* had a positive relationship with *Streptococcus* and *Pediococcus*, and *Kocuria* was in positive correlation with *Microbacterium* (Figure 4A). In the oral, *Kocuria* was in a negative correlation with *Acinetobacter*, whereas *Geobacillus* was in a positive relationship with others, including *Lactobacillus* (Figure 4B). *Kocuria* in the gill was positively related to many Actinobacteria species (such as *Knoellia*, *Pseudokineococcus*, *Kytococcus*, etc.), and they were negatively related to other phylum species (Figure 4C). Moreover, *Acinetobacter* had a positive relationship with *Geobacillus*. In the hindgut, *Acinetobacter* showed a negative correlation with *Mitsuokella*, which was also in high abundance (3.42%), and *Geobacillus* had a positive correlation with *Massilia* (Figure 4D).

Biomarkers in each group

Biomarkers were discovered in four different groups (Figure 5). The phylogenetic distribution of microbial communities in different groups was shown in Figure 5A, and 7 phyla clades contained at least one biomarker, i.e. Proteobacteria, Firmicutes, Bacteroidota, Cyanobacteria, Actinobacteria, and Synergistota.

From Figure 5B and Supplementary Table 6, there were 10 biomarkers were identified in the skin, mainly belonging to the phyla of Firmicutes and Actinobacteria. Among them, *Kocuria* and *Lactobacillus* were abundant biomarkers. The abundance of other biomarkers was low, for example, *Psychrobacter* (1.78%), *Brachybacterium* (1.75%), *Pseudokineococcus* (1.65%) (Supplementary Table 6). In the oral, there were only 4 biomarkers from different phyla. Biomarkers in the gill belong to Proteobacteria, Firmicutes, and Actinobacteria, and they had

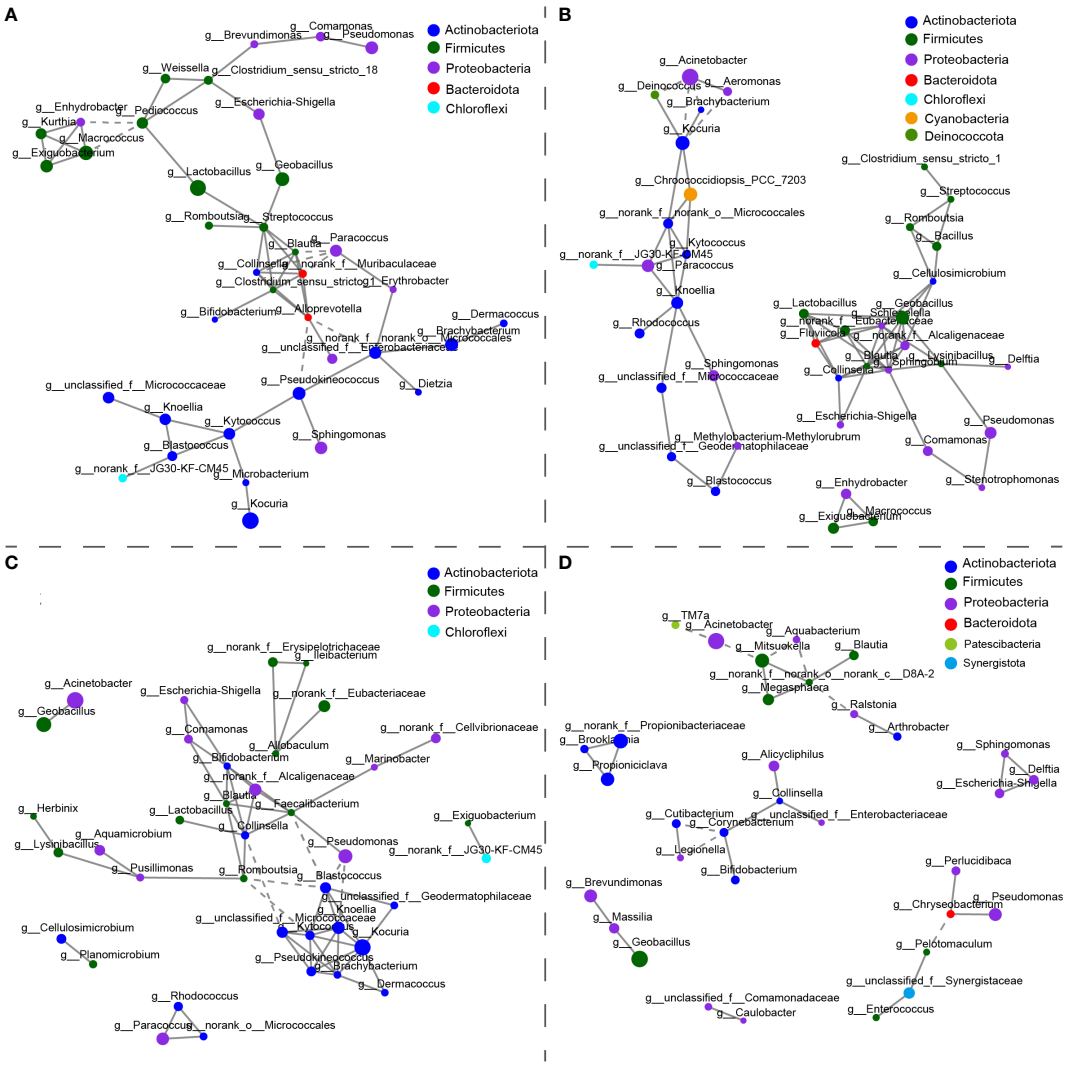


FIGURE 4
The co-occurrence networks of core microbiotas in the **(A)** skin, **(B)** oral, **(C)** gill, and **(D)** hindgut. The size of the node shows the relative abundance at the genus level. Node color represents taxonomic classification at the phylum level. Positive interactions are depicted in the solid line and negative interactions in the dashed line. The patterns of networks were measured by Spearman's correlation coefficient ($r>0.75$, $p<0.05$).

low relative abundance, such as *Knoellia* (1.49%), *Bacillus* (1.12%), and *Aquamicrobium* (0.84%). *Geobacillus* was specifically enriched in the hindgut, as well as the *Mitsuokella*, *Propioniciclava*, and *Brevundimonas*. They had a relative abundance of more than 2%.

Microbial function prediction

According to function analysis, the top 5 most abundant classifications among 24 COGs were "Amino acid transport and metabolism", "Translation, ribosomal structure and biogenesis", "Energy production and conversion", "Inorganic ion transport and metabolism", and "Transcription" (Supplementary Figure 5). Thirteen COGs showed significant differences in the four groups (Figure 6A). COGs of L (Replication, recombination and repair) and F (Nucleotide transport and metabolism) had a significantly highest relative abundance in the skin, and T (Signal transduction and metabolism) and N (Cell motility) were significantly lowest. In the hindgut, COGs of M (Cell wall/membrane/envelope biogenesis) and U (Intracellular trafficking, secretion, and vesicular transport) had higher relative abundance than that in gill and skin, and lower in K (Transcription) compared to gill and skin. COG of O (Posttranslational modification, protein turnover, chaperones) had a higher relative abundance in the oral than skin or gill. COG of E (Amino acid transport and metabolism) had a higher relative abundance in gill than skin.

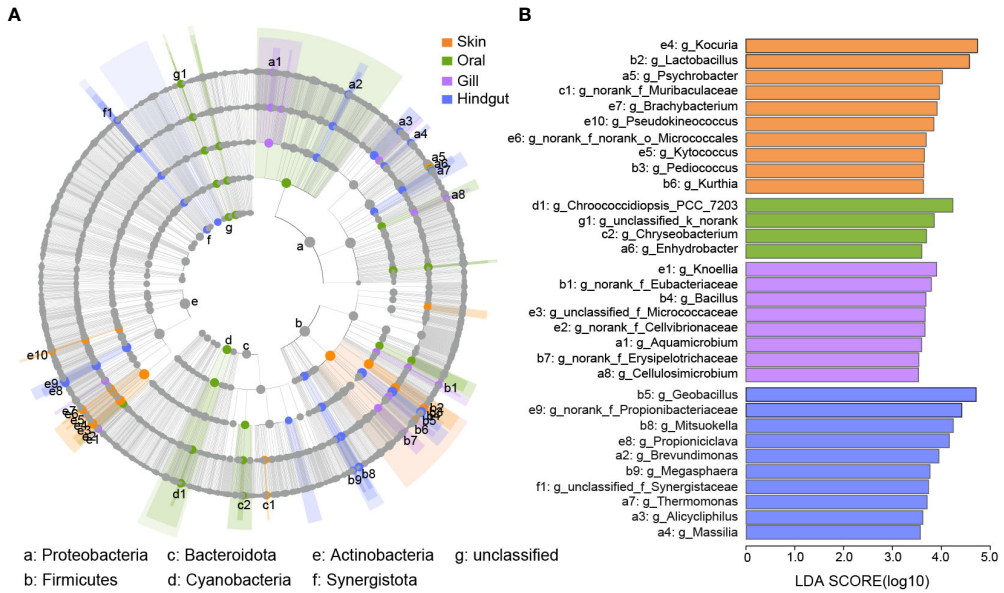


FIGURE 5
Biomarkers in four body sites. **(A)** Cladogram of the phylogenetic distribution of the bacterial lineages from phylum to the genus in the skin, oral, gill, and hindgut. The diameter of each dot repents the relative abundance. The species with a significant difference were represented by different colors among the four groups. The grey nodes indicate that there was no significant difference. **(B)** Bar plots showed the biomarkers from **(A)** with an LDA ≥ 3.5.

The correlation heatmap showed COG functional enrichment was related to most microbial biomarkers in each group (Figures 6B, C). In the skin, *Kocuria* was significantly related to the COGs of E, K, and F. Biomarkers identified in Figure 5B, such as *Brachybacterium*, *Pseudokineococcus*, g_norank_f_norank_o_Micrococcales, and *Kytococcus* in the skin were proved to have a positive correlation in COGs of E, G and T (Figure 6B). In the hindgut, the biomarker *Brevundimonas* was positively related to COGs of M and U (Figure 6C). COG of Amino acid transport and metabolism was significantly positively correlated with *Brevundimonas* and *Massilia*. Carbohydrate transport and metabolism was significantly positively correlated with *Propioniciclava*, but negatively correlated with *Thermomonas*. *Mitsuokella* was negatively correlated with most COGs except for Carbohydrate transport and metabolism. In the oral, *Chroococcidiopsis*_PCC_7203 was positively related to COGs G, and an unclassified genus was negatively related to most COGs (Supplementary Figure 6A). *Bacillus* in the gill was positively related to COGs of N, M and U (Supplementary Figure 6B).

Discussion

Microbial community structure varied in different body sites

Mucus-associated microbial communities were subjected to host-associated selection factors (Pratte et al., 2018). Body sites was proved as the dominant factor of microbial community diversity in our study, but no significant difference was found in the microbial structure from different body size or species. The high convergence of microbiome in yellowfin tuna and bigeye tuna may be caused by the process of "phylosymbiosis", since the two tuna species are in a close phylogenetical relationship and host species with close phylogenetic relationships have more similar microbiotas (Brooks et al., 2016). It is worth noting that sample size of bigeye tuna was limited in this study. It could be a larger variation between microbiomes from yellowfin and bigeye tunas when the sample size is larger. In addition, an overlap in the ecological niche of the two tunas might be the other primary cause of their similarity in bacterial composition.

The marked microbial heterogeneity was obviously reflected in external and internal sites. It suggested different bacterial assemblages and microbial niches differentiation at the organ scale, further providing insight into how microbiome adapt to the host (Chiarello et al., 2015; Zhang et al., 2019; Sylvain et al., 2020). The different microenvironments among body sites were proposed as the leading cause of microbial heterogeneity (Chiarello et al., 2015). For surficial microorganisms, they are more susceptible to environmental influence or disturbance (Chiarello et al., 2015). While the gut microecosystem has relatively stable pH and temperature to keep homeostasis, promoting the specialized and modular populations to colonize (Ross et al., 2019; Sylvain et al., 2020). Furthermore, unlike most poikilothermic fish, tuna represents a regional warm-blood fish with a stable celiac temperature at 25~28°C and approximately 10°C difference from the external surface (Block et al., 2001). The relatively stable temperature in the celiac area may allow tuna to selectively

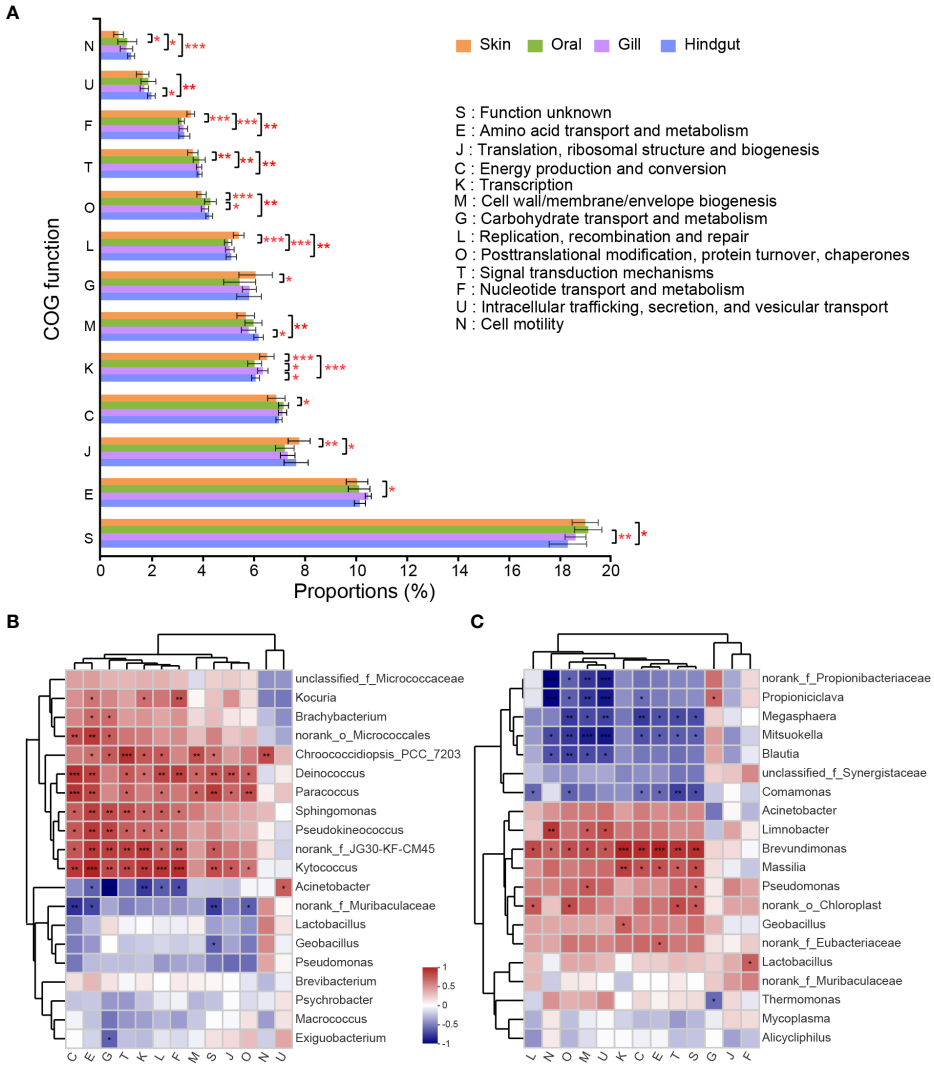


FIGURE 6
COG function classifications with a significant difference in the four groups and its correlation with high abundant species. **(A)** Statistical test of the COGs relative abundance was carried out by the Kruskal-Wallis H Test followed by Bonferroni correction. Heatmap for correlation between COG function classifications and top 20 high abundant species in the **(B)** skin and **(C)** hindgut. Correlation coefficients were calculated by Spearman. (*$p \leq 0.05$; **$p \leq 0.01$; ***$p \leq 0.001$).

recruit and assemble organ-specific symbionts for long-term co-evolution. It is necessary to point out that the microbiome of hindgut in this paper was from the content of the hindgut instead of its surface mucus and there might be difference of biodiversity and abundance of microbiota between gut mucus layer and the content (Kashinskaya et al., 2017). The aim in this study is to compare the mucus-associated microbiota from different body sites in tuna. We believe that the content of the hindgut can reflect the mucus-associated microbiota from hindgut, since sampling of feces and intestinal contents has been widely used in the study on gut mucosal microbiome, including the tuna (Minich et al., 2020a; Minich et al., 2020b). On the other hand, to make the description more clearly and let the comparison more reasonable, it would be better to pay attention to the difference of microbiome from gut mucus and content, and process the related results carefully in the future work.

Among external surface groups, the alpha diversity in the gill was significantly higher than that of skin, which was consistent with the cultured southern bluefin tunas (Minich et al., 2020a). However, it was the skin that possessed a higher Shannon index than gill in other studies on Pacific chub mackerel (Minich et al., 2020b), seabass and seabream (Rosado et al., 2019), grass carp and southern catfish (Zhang et al., 2019). The inconsistent results between tunas and other fish may be related to host habits and physiological characteristics. As we all know, tuna is remarkable in swimming performance and high speed. As the skin of tuna lacks scales, it would be heavily scoured by the water

current to the disadvantage of bacterial colonization. Whereas, the gill is in a semi-enclosed space and the gill filaments are interlamellar fusional to relieve the impact of water flow (Evans et al., 2005).

The oral, primarily responsible for energy intake, was similar to the gill in community composition and diversity, probably because they are spatially connected. The environment of the host oral is complex, and the microbes that live here are closely related to diet and environmental parameters (Abdelhafiz et al., 2021). However, the related studies on fish oral microbiotas are limited. More work needs to be done to investigate the relationship between the diet and the microbiome community in the oral.

Microbial discrepancy between the external surface and the environment

The external surface of fish touched with seawater directly, but the bacterial diversity and components of surrounding seawater were significantly lower than that of the external surface. The same cases had been reported previously (Minniti et al., 2017; Sylvain et al., 2020; Steiner et al., 2021; McMurtrie et al., 2022). The high discrepancy between host-associated and environmental microbiotas suggests that fish surface microbial communities are not simple reflections of the microbial assemblages in their habitat. Heterotrophic bacteria were scarce in the water than on the external surface of fish, due to the oligotrophic nutrients in the ocean (Larsen et al., 2013). By contrast, fish body is taken as a eutrophic "island", as there are multitudinous components secreted by mucosal layer cells, such as mucins, gel-forming glycoproteins and glycosaminoglycans (Chiarello et al., 2015; Ross et al., 2019). The higher bacterial diversity on the external surface reflects the host selection effect from the surrounding "bacterial pool" (Chiarello et al., 2015). Therefore, the quality of the aquaculture water system is crucial to the individual health of farmed fish.

Core species have a vital status and interaction in microbial community

Core microbiotas play a critical role in the formation of symbiotic communities, and many studies aimed to reduce the complexity of host-associated and determine the correlation between the core microbiome and host health (Dong et al., 2021). The four groups shared a part of core taxa including highly abundant phyla of Proteobacteria, Firmicutes, and Actinobacteria, which were also detected in farmed tuna (Minich et al., 2020a) and most teleost species (Wilson et al., 2008; Chiarello et al., 2015). At the genus level, *Acinetobacter* spp. were universally dominant in all body sites of wild tuna in this paper, as well as farmed southern bluefin tuna (Minich et al., 2020a), indicating that mucus of tuna is one of their natural niches. A previous study confirmed that the probiotic *Acinetobacter* strain (*Acinetobacter* KU011TH) isolated from the skin mucus of bighead catfish can significantly improve growth performance (Bunnoy et al., 2019).

The core taxa can affect the community structure by their high abundance or strong biological interaction with other species (Agler et al., 2016; Dong et al., 2021). Moreover, the dominant species may influence selection pressure on other resident microbial strains (Ferreiro et al., 2018). In the gut, 25.81% of relationships were defined as negative, such as *Acinetobacter* and *Mitsuokella*, which demonstrated potential competition. As described in the "Red Queen hypothesis", competitions could accelerate microbial evolution in the microbial ecosystem, resulting in apparent stability of microecosystem (Ferreiro et al., 2018). Additionally, the Firmicutes species *Lactobacillus* and *Streptococcus* are both lactic acid bacteria (LAB), showing commensal interaction in the skin co-occurrence network. They are helpful in resisting aquatic pathogen colonization by producing inhibitory compounds and competing for nutrients (Pérez-Sánchez et al., 2011). Whereas, *Pseudoalteromonas*, *Psychrobacter*, and *Vibrio* were prominent in the skin and gill of the diseased tuna but lack of LAB (Minich et al., 2020a). *Geobacillus* spp. are important probiotics beneficial to the host's digestion and absorption of nutrients (Miao et al., 2018). They were found significantly enriched in gut of wild tuna in this paper, and in a positive relationship with the colonization of *Massilia*, which was proposed to contribute to fish development (Califano et al., 2017; Fujimoto et al., 2020). This commensal relationship implied the co-evolution of their ecological niche and performed the parallel function for tunas' health.

From the core microbial networks, some taxa were not dominant in abundance, but have a close relationship with others, such as *Alloprevotella* and *Streptococcus* in the skin and g_norank_c_DBA-2, *Corynebacterium* in the gut. They might be instrumental in stabilizing and regulating the community, as similar conditions were reported in a previous paper (Jousset et al., 2017; Dong et al., 2021). For the dominant core taxa, take *Acinetobacter*, *Kocuria* and *Geobacillus* for example, they had more simple interactions with others. These taxa were stable in population dynamics and played a vital role in the structure of host-associated microbial networks (Dong et al., 2021). Therefore, all the core microbiomes, no matter with high abundance or low, have an essential role in maintaining the relative stability of the community and more studies need to be done to further determine their functions.

Microbial biomarkers provide insights into the microecological function

The microbes on the external surface of fish are high environment-dependent. Some of them commonly act as indicator species when the host homeostasis or the surrounding environment is perturbed (Sylvain et al., 2020). Based on our result, the abundant *Lactobacillus* in the skin could be considered the best biomarker candidate for tuna to monitor the host health status and environmental quality. It has been also proposed as a biomarker for toxicogenic exposure (Spilsbury et al., 2022).

Host microbiotas are in highly complex interactions, and function prediction can help us to better understand their roles (Hicks et al., 2021). In the gut, several microbial biomarkers, such as *Propioniciclava*, *Mitsuokella*, *Massilia*, and *Brevundimonas*, were positively linked with functions related to the metabolism of nutrients and nucleotides. They were also reported to have similar functions or the other metabolism: *Propioniciclava* is propionate-producing bacteria, and was remarkably increased in presence in the seabream gut after intaking a high level of protein (Solé-Jiménez et al., 2021); *Mitsuokella* was related to carbohydrate metabolism (Tsukahara et al., 2002) and amino acid metabolism (Dai et al., 2010), and *Mitsuokella multiacida* can use lactate and acetate to form butyrate, which is an important SCFA in the intestine (Tsukahara et al., 2002); *Massilia* could be promoted by a commercial probiotic *Pediococcus acidilactici*, resulting in synergistic roles (Rasmussen et al., 2022). Although *Lactobacillus* is not abundant in the gut, it was positively related to energy metabolism. It was helpful for probiotics to increase the host metabolic level and obtain enough nutrients to meet the energy requirement of high-performance fish (Rasmussen et al., 2022). *Brevundimonas* was detected as a biomarker in the gut and is a potential pathogen in most situations (Minich et al., 2020a), indicating a complex co-evolution between commensal symbionts, pathogens, and host. It is essential to further work on the co-evolution of symbiotic and pathogenic bacteria.

Conclusion

The diversity and composition of microbial communities in the four different body sites (skin, oral, gill, and gut) of wild tunas were reported in this paper. They varied among body sites instead of species or body size. Proteobacteria, Actinobacteria, and Firmicutes were the dominant phyla in all body sites. Some abundant species *Acinetobacter*, *Lactobacillus*, and *Geobacillus* were in tight interaction with other probiotic species identified from core taxa, and may contribute to host nutrition and immunity. Tightly connected core microbiotas may promote the stability of the microbial community. Biomarkers are predicted in the four body sites. The skin biomarker *Lactobacillus* could be used in environmental disturbance monitoring. Gut biomarkers were shown to be closely related to the metabolism functions. The results of this study highlight the importance of symbiotic microorganisms for host health, especially for cultured tuna in the process of aquaculture.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.ncbi.nlm.nih.gov/, BioProject ID: PRJNA884520; https://www.ncbi.nlm.nih.gov/, BioProject ID: PRJNA902642.

Ethics statement

Ethical review and approval was not required for the animal study because all tunas were harvested by private company and as part of commerce.

Author contributions

YZ: sampling, data processing and analysis, drafted the manuscript. DW: sampling and data analysis. LW and JX: writing-review and editing, pre-processing of the samples. PZ and HH: sampling and pre-processing of the samples. YJZ and ZG: research design, manuscript editing. All authors discussed and approved the final version of the manuscript.

Funding

This work was supported by the Opening Foundation of Hainan Tropical Ocean University & Yazhouwan Innovation Research Institute (2022RHDKFKT03) and the Modern Agricultural Industry Technology System of Grouper in Hainan Province (HNARS2022-3-G4 & HNARS2022-3).

Acknowledgments

The authors would like to thank Shuai Pan and Anfu Sun at Hainan University for their generous help in the sample collection.

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.2022.1073264/full#supplementary-material

References

FAO (2020). "The state of world fisheries and aquaculture 2020," in *Sustainability in action* Rome: Food and Agriculture Organization of the United Nations. doi: 10.4060/ca9229en

Abdelhafiz, Y., Fernandes, J. M. O., Stefani, E., Albanese, D., Donati, C., and Kiron, V. (2021). Power play of commensal bacteria in the buccal cavity of female Nile tilapia. *Front. Microbiol.* 12. doi: 10.3389/fmicb.2021.773351

Agler, M. T., Ruhe, J., Kroll, S., Morhenn, C., Kim, S. T., Weigel, D., et al. (2016). Microbial hub taxa link host and abiotic factors to plant microbiome variation. *PloS Biol.* 14, e1002352. doi: 10.1371/journal.pbio.1002352

Banerjee, G., and Ray, A. K. (2017). Bacterial symbiosis in the fish gut and its role in health and metabolism. *Symbiosis* 72 (1), 1–11. doi: 10.1007/s13199-016-0441-8

Block, B. A., Dewar, H., Blackwell, S. B., Williams, T. D., Prince, E. D., Farwell, C. J., et al. (2001). Migratory movements, depth preferences, and thermal biology of Atlantic bluefin tuna. *Science* 293 (5533), 1310–1314. doi: 10.1126/science.1061197

Brooks, A. W., Kohl, K. D., Brucker, R. M., van Opstal, E. J., and Bordenstein, S. R. (2016). Phylosymbiosis: relationships and functional effects of microbial communities across host evolutionary history. *Plos. Biol.* 14 (11), e2000225. doi: 10.1371/journal.pbio.2000225

Buffie, C. G., and Pamer, E. G. (2013). Microbiota-mediated colonization resistance against intestinal pathogens. *Nat. Rev. Immunol.* 13, 790–801. doi: 10.1038/nri3535

Bunnoy, A., Na-Nakorn, U., and Srisapoome, P. (2019). Probiotic effects of a novel strain, *Acinetobacter* KU011TH, on the growth performance, immune responses, and resistance against *Aeromonas hydrophila* of bighead catfish (*Clarias macrocephalus* Gunther 1864). *Microorganisms* 7 (12), 613. doi: 10.3390/microorganisms7120613

Califano, G., Castanho, S., Soares, F., Ribeiro, L., Cox, C. J., Mata, L., et al. (2017). Molecular taxonomic profiling of bacterial communities in a gilthead seabream (*Sparus aurata*) hatchery. *Front. Microbiol.* 8. doi: 10.3389/fmicb.2017.00204

Carda-Dieguez, M., Ghai, R., Rodriguez-Valera, F., and Amaro, C. (2017). Wild eel microbiome reveals that skin mucus of fish could be a natural niche for aquatic mucosal pathogen evolution. *Microbiome* 5, 162. doi: 10.1186/s40168-017-0376-1

Chen, S. F., Zhou, Y. Q., Chen, Y. R., and Gu, J. (2018). Fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatic* 34 (17), 884–890. doi: 10.1093/bioinformatics/bty560

Chiarello, M., Villéger, S., Bouvier, C., Bettarel, Y., and Bouvier, T. (2015). High diversity of skin-associated bacterial communities of marine fishes is promoted by their high variability among body parts, individuals and species. *FEMS Microbiol. Ecol.* 91 (7), fiv061. doi: 10.1093/femsec/fiv061

Dai, Z. L., Zhang, J., Wu, G. Y., and Zhu, W. Y. (2010). Utilization of amino acids by bacteria from the pig small intestine. *Amino Acids* 39 (5), 1201–1215. doi: 10.1007/s00726-010-0556-9

Dehler, C. E., Secombes, C. J., and Martin, S. A. M. (2017). Environmental and physiological factors shape the gut microbiota of Atlantic salmon parr (*Salmo salar l.*). *Aquaculture* 467, 149–157. doi: 10.1016/j.aquaculture.2016.07.017

Di Domizio, J., Belkhodja, C., Chenuet, P., Fries, A., Murray, T., Mondejar, P. M., et al. (2020). The commensal skin microbiota triggers type I IFN-dependent innate repair responses in injured skin. *Nat. Immunol.* 21 (9), 1034–1045. doi: 10.1038/s41590-020-0721-6

Dong, C. B., Shao, Q. Y., Zhang, Q. Q., Yao, T., Huang, J. Z., Liang, Z. Q., et al. (2021). Preferences for core microbiome composition and function by different definition methods: Evidence for the core microbiome of *Eucommia ulmoides* bark. *Sci. Total. Environ.* 790, 148091. doi: 10.1016/j.scitotenv.2021.148091

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

Edgar, R. C. (2013). UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat. Methods* 10 (10), 996–998. doi: 10.1038/nmeth.2604

Evans, D. H., Piermarini, P. M., and Choe, K. P. (2005). The multifunctional fish gill: dominant site of gas exchange, osmoregulation, acid-base regulation, and excretion of nitrogenous waste. *Physiol. Rev.* 85 (1), 97–177. doi: 10.1152/physrev.00050.2003

Fassarella, M., Blaak, E. E., Penders, J., Nauta, A., Smidt, H., and Zoetendal, E. G. (2021). Gut microbiome stability and resilience: elucidating the response to perturbations in order to modulate gut health. *Gut* 70 (3), 595–605. doi: 10.1136/gutjnl-2020-321747

Ferreiro, A., Crook, N., Gasparrini, A. J., and Dantas, G. (2018). Multiscale evolutionary dynamics of host-associated microbiomes. *Cell* 172 (6), 1216–1227. doi: 10.1016/j.cell.2018.02.015

Foster, K. R., Schluter, J., Coyte, K. Z., and Rakoff-Nahoum, S. (2017). The evolution of the host microbiome as an ecosystem on a leash. *Nature* 548 (7665), 43–51. doi: 10.1038/nature23292

Fujimoto, M., Marsh, T. L., and Scribner, K. T. (2020). Effects of water filtration and temperature on microbial colonization and survival of lake sturgeon eggs. *N. Am. J. Aquacult.* 83 (1), 26–37. doi: 10.1002/naaq.10169

Gardner, S. G., Camp, E. F., Smith, D. J., Kahlke, T., Osman, E. O., Gendron, G., et al. (2019). Coral microbiome diversity reflects mass coral bleaching susceptibility during the 2016 El niño heat wave. *Ecol. Evol.* 9, 938–956. doi: 10.1002/ece3.4662

Hagberg, A. A., Schult, D. A., and Swart, P. J. (2008). "Exploring network structure, dynamics, and function using networkx," in *Proceedings of the 7th Python in science conference. SciPy 2008: 7th Python in Science Conference.* (Pasadena, California, USA), 11–15. Available at: https://www.osti.gov/biblio/960616.

Hicks, L. C., Frey, B., Kjoller, R., Lukac, M., Moora, M., Weedon, J. T., et al. (2021). Toward a function-first framework to make soil microbial ecology predictive. *Ecology* 103 (2), e03594. doi: 10.1002/ecy.3594

Jacobson, A., Lam, L., Rajendram, M., Tamburini, F., Honeycutt, J., Pham, T., et al. (2018). A gut commensal-produced metabolite mediates colonization resistance to *Salmonella* infection. *Cell Host Microbe* 24 (2), 296–307. doi: 10.1016/j.chom.2018.07.002

Jousset, A., Bienhold, C., Chatzinotas, A., Gallien, L., Gobet, A., Kurm, V., et al. (2017). Where less may be more: how the rare biosphere pulls ecosystems strings. *ISME J.* 11 (4), 853–862. doi: 10.1038/ismej.2016.174

Karlsen, C., Ottem, K. F., Brevik, O. J., Davey, M., Sorum, H., and Winther-Larsen, H. C. (2017). The environmental and host-associated bacterial microbiota of Arctic seawater-farmed Atlantic salmon with ulcerative disorders. *J. Fish Dis.* 40 (11), 1645–1663. doi: 10.1111/jfd.12632

Kashinskaya, E. N., Andree, K. B., Simonov, E. P., and Solovyev, M. M. (2017). DNA extraction protocols may influence biodiversity detected in the intestinal microbiome: a case study from wild Prussian carp, *Carassius gibelio*. 2016. *FEMS Microbiol. Ecol* 93 (2). doi: 10.1093/femsec/fiw240

Larsen, A., Tao, Z., Bullard, S. A., and Arias, C. R. (2013). Diversity of the skin microbiota of fishes: evidence for host species specificity. *FEMS Microbiol. Ecol.* 85 (3), 483–494. doi: 10.1111/1574-6941.12136

Liu, J. H., Zhang, M. L., Zhang, R. Y., Zhu, W. Y., and Mao, S. Y. (2016). Comparative studies of the composition of bacterial microbiota associated with the ruminal content, ruminal epithelium and in the faeces of lactating dairy cows. *Microb. Biotechnol.* 9 (2), 257–268. doi: 10.1111/1751-7915.12345

Lynch, J. B., and Hsiao, E. Y. (2019). Microbiomes as sources of emergent host phenotypes. *Science* 365 (6460), 1405–1408. doi: 10.1126/science.aay0240

Magoč, T., and Salzberg, S. L. (2011). FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* 27 (21), 2957–2963. doi: 10.1093/bioinformatics/btr507

McMurtrie, J., Alathari, S., Chaput, D. L., Bass, D., Ghambi, C., Nagoli, J., et al. (2022). Relationships between pond water and tilapia skin microbiomes in aquaculture ponds in Malawi. *Aquaculture* 558, 738367. doi: 10.1016/j.aquaculture.2022.738367

Miao, S. Y., Zhu, J. Y., Zhao, C. Z., Sun, L. S., Dong, X. J., and Chen, G. H. (2018). Effect of dietary soybean meal associated with feeding time on the growth performance and intestinal microbiota composition of northern snakehead. *Aquac. Res.* 50 (10), 2751–2759. doi: 10.1111/are.14035

Minich, J. J., Petrus, S., Michael, J. D., Michael, T. P., Knight, R., and Allen, E. E. (2020b). Temporal, environmental, and biological drivers of the mucosal microbiome in a wild marine fish, *Scomber japonicus*. *mSphere* 5 (3), e00401–e00420. doi: 10.1128/mSphere.00401-20

Minich, J. J., Power, C., Melanson, M., Knight, R., Webber, C., Rough, K., et al. (2020a). The southern bluefin tuna mucosal microbiome is influenced by husbandry method, net pen location, and anti-parasite treatment. *Front. Microbiol.* 11. doi: 10.3389/fmicb.2020.02015

Minniti, G., Hagen, L. H., Porcellato, D., Jorgensen, S. M., Pope, P. B., and Vaaje-Kolstad, G. (2017). The skin-mucus microbial community of farmed Atlantic salmon (*Salmo salar*). *Front. Microbiol.* 8. doi: 10.3389/fmicb.2017.02043

Nowak, B. F., Dang, M., Webber, C., Neumann, L., Bridle, A., Bermudez, R., et al. (2021). Changes in the splenic melanomacrophage centre surface area in southern bluefin tuna (*Thunnus maccoyii*) are associated with blood fluke infections. *Pathogens* 10 (1), 9. doi: 10.3390/pathogens10010079

Ohshimo, S., Hiraoka, Y., Sato, T., and Nakatsuka, S. (2018). Feeding habits of bigeye tuna (*Thunnus obesus*) in the north pacific from 2011 to 2013. *Mar. Freshw. Res.* 69 (4), 585–606. doi: 10.1071/MF17058

Pérez-Sánchez, T., Balcázar, J. L., Garcia, Y., Halaihel, N., Vendrell, D., de Blas, I., et al. (2011). Identification and characterization of lactic acid bacteria isolated from rainbow trout, *Oncorhynchus mykiss* (Walbaum), with inhibitory activity against *Lactococcus garvieae*. *J. Fish. Dis.* 34 (7), 499–507. doi: 10.1111/j.1365-2761.2011.01260.x

Pratte, Z. A., Besson, M., Hollman, R. D., and Stewart, F. J. (2018). The gills of reef fish support a distinct microbiome influenced by host-specific factors. *Appl. Environ. Microbiol.* 84 (9), e00063–e00018. doi: 10.1128/AEM.00063-18

Rasmussen, J. A., Villumsen, K. R., Ernst, M., Hansen, M., Forberg, T., Gopalakrishnan, S., et al. (2022). A multi-omics approach unravels metagenomic and metabolic alterations of a probiotic and synbiotic additive in rainbow trout (*Oncorhynchus mykiss*). *Microbiome* 10 (1), 21. doi: 10.1186/s40168-021-01221-8

Rosado, D., Pérez-Losada, M., Severino, R., Cable, J., and Xavier, R. (2019). Characterization of the skin and gill microbiomes of the farmed seabass (*Dicentrarchus labrax*) and seabream (*Sparus aurata*). *Aquaculture* 500, 57–64. doi: 10.1016/j.aquaculture.2018.09.063

Rosenberg, E., and Zilber-Rosenberg, I. (2018). The hologenome concept of evolution after 10 years. *Microbiome* 6, 78. doi: 10.1186/s40168-018-0457-9

Ross, A. A., Hoffmann, A. R., and Neufeld, J. D. (2019). The skin microbiome of vertebrates. *Microbiome* 7 (1), 79. doi: 10.1186/s40168-019-0694-6

Schloss, P. D., Westcott, S. L., Ryabin, T., Hartmann, M., Hollister, E. B., Lesniewski, R. A., 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

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

Semova, I., Carten, J., Stombaugh, J., Mackey, L. C., Knight, R., Farber, S. A., et al. (2012). Microbiota regulate intestinal absorption and metabolism of fatty acids in the zebrafish. *Cell Host Microbe* 12 (3), 277–288. doi: 10.1016/j.chom.2012.08.003

Shetty, S. A., Hugenholtz, F., Lahti, L., Smidt, H., and de Vos, W. M. (2017). Intestinal microbiome landscaping: insight in community assemblage and implications for microbial modulation strategies. *FEMS Microbiol. Rev.* 41 (2), 182–199. doi: 10.1093/femsre/fuw045

Solé-Jiménez, P., Naya-Català, F., Piazzon, M. C., Estensoro, I., Calduch-Giner, J.À., Sitjà-Bobadilla, A., et al. (2021). Reshaping of gut microbiota in gilthead sea bream fed microbial and processed animal proteins as the main dietary protein source. *Front. Mar. Sci.* 8. doi: 10.3389/fmars.2021.705041

Spilsbury, F., Foysal, M. J., Tay, A., and Gagnon, M. M. (2022). Gut microbiome as a potential biomarker in fish: dietary exposure to petroleum hydrocarbons and metals, metabolic functions and cytokine expression in juvenile *Lates calcarifer*. *Front. Microbiol.* 13. doi: 10.3389/fmicb.2022.827371

Spor, A., Koren, O., and Ley, R. (2011). Unravelling the effects of the environment and host genotype on the gut microbiome. *Nat. Rev. Microbiol.* 9 (4), 279. doi: 10.1038/nrmicro2540

Steiner, K., Heasman, K., Laroche, O., Pochon, X., Preece, M., Bowman, J. P., et al. (2021). The microbiome of Chinook salmon (*Oncorhynchus tshawytscha*) in a recirculation aquaculture system. *Aquaculture* 534, 736227. doi: 10.1016/j.aquaculture.2020.736227

Stressmann, F. A., Bernal-Bayard, J., Perez-Pascual, D., Audrain, B., Rendueles, O., Briolat, V., et al. (2020). Mining zebrafish microbiota reveals key community-level resistance against fish pathogen infection. *ISME J.* 15 (3), 702–719. doi: 10.1038/s41396-020-00807-8

Sultana, S., Khan, M. N., Hossain, M. S., Dai, J., Rahman, M. S., and Salimullah, M. (2022). Community structure and functional annotations of the skin microbiome in healthy and diseased catfish, *Heteropneustes fossilis*. *Front. Microbiol.* 13. doi: 10.3389/fmicb.2022.856014

Sylvain, F.-E., Holland, A., Bouslama, S., Audet-Gilbert, E., Lavoie, C., Val, A. L., et al. (2020). Fish skin and gut microbiomes show contrasting signatures of host species and habitat. *Appl. Environ. Microbiol.* 86 (16), e00789–e00720. doi: 10.1128/AEM.00789-20

Tsukahara, T., Koyama, H., Okada, M., and Ushida, K. (2002). Stimulation of butyrate production by gluconic acid in batch culture of pig cecal digesta and identification of butyrate-producing bacteria. *J. Nutr.* 132 (8), 2229–2234. doi: 10.1093/jn/132.8.2229

Varela, J. L., Intriago, K. M., Flores, J. C., and Lucas-Pilozo, C. R. (2017). Feeding habits of juvenile yellowfin tuna (*Thunnus albacares*) in Ecuadorian waters assessed from stomach content and stable isotope analysis. *Fish Res.* 194, 89–98. doi: 10.1016/j.fishres.2017.05.017

Wang, Q., Garrity, G. M., Tiedje, J. M., and Cole, J. R. (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* 73 (16), 5261–5267. doi: 10.1128/AEM.00062-07

Wang, A. R., Ran, C., Ringo, E., and Zhou, Z. G. (2018). Progress in fish gastrointestinal microbiota research. *Rev. Aquacult.* 10 (3), 626–640. doi: 10.1111/raq.12191

Willmann, M., Vehreschild, M. J. G. T., Biehl, L. M., Vogel, W., Dorfel, D., Hamprecht, A., et al. (2019). Distinct impact of antibiotics on the gut microbiome and resistome: a longitudinal multicenter cohort study. *BMC Biol.* 17 (1), 76. doi: 10.1186/s12915-019-0692-y

Wilson, B., Danilowicz, B. S., and Meijer, W. G. (2008). The diversity of bacterial communities associated with Atlantic cod *Gadus morhua*. *Microb. Ecol.* 55 (3), 425–434. doi: 10.1007/s00248-007-9288-0

Zhang, Z. M., Li, D. P., Xu, W. T., Tang, R., and Li, L. (2019). Microbiome of co-cultured fish exhibits host selection and niche differentiation at the organ scale. *Front. Microbiol.* 10. doi: 10.3389/fmicb.2019.02576

frontiers | Frontiers in Marine Science

TYPE Original Research
PUBLISHED 19 January 2023
DOI 10.3389/fmars.2023.1055848

OPEN ACCESS

EDITED BY
Senjie Lin,
University of Connecticut, United States

REVIEWED BY
Jiayuan Liang,
Guangxi University, China
Kefu Yu,
Guangxi University, China

*CORRESPONDENCE
Zhiyong Li
zyli@sjtu.edu.cn

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

RECEIVED 28 September 2022
ACCEPTED 04 January 2023
PUBLISHED 19 January 2023

CITATION
Xu J, Chai G, Xiao Y, Xie Z, Yang X, Liao B, Xiao B and Li Z (2023) The community profiles of symbiotic bacteria at the different life stages of coral *Dipsastraea favus*. *Front. Mar. Sci.* 10:1055848. doi: 10.3389/fmars.2023.1055848

The community profiles of symbiotic bacteria at the different life stages of coral *Dipsastraea favus*

Jianjun Xu[1], Guangjun Chai[1], Yilin Xiao[1], Ziqiang Xie[2], Xiaodong Yang[2], Baolin Liao[2], Baohua Xiao[2] and Zhiyong Li[1]*

[1]State Key Laboratory of Microbial Metabolism, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai, China, [2]Shenzhen Institute of Guangdong Ocean University, Shenzhen, China

Corals live in a symbiotic relationship with various bacteria that are fundamental to host fitness, health, and survival. Though the diversity of symbiotic bacteria has been revealed in the early life stages of some corals, the dynamic bacterial community profiles of one coral are still poorly characterized, particularly the stage-specific bacteria. In this study, the bacterial communities in the patent, eggs, and 4-day-old planula larvae of a hermaphrodite coral *Dipsastraea favus* were investigated by high-throughput sequencing of 16S ribosomal RNA gene. As a result, dynamic profiles of bacterial community in the parent, eggs and larvae of *D. favus* were suggested. The bacterial diversity in the planula larvae was a bit higher than that in the *D. favus* parent, and distinct stage-specific symbiotic bacteria were detected, e.g., Oceanospirillaceae, *Kordia*, and Legionellaceae in *D. favus* larvae, and Kiloniellales and *Prosthecochloris* in adult *D. favus*. The dynamic change of bacterial community in coral adults and larvae may expand our understanding of the complex relationship between coral host and its symbiotic microbiota.

KEYWORDS

***Dipsastraea favus*, life stage, bacterial community, high-throughput sequencing, horizontal transmission**

Introduction

Coral holobionts harbor a complex consortium of microorganisms, including zooxanthellae, bacteria, archaea, fungi, viruses, and protists (Knowlton and Rohwer, 2003). Coral microbial flora has been attracting attention for their potentials to protect corals from environmental stresses and their relevance with coral hosts, such as defense against pathogens, promotion of coral health, development, nutrition-based symbioses, and environmental adaption (Krediet et al., 2013; Soffer et al., 2015; Lema et al., 2016; Yang et al., 2016; Zhou et al., 2017; Bernasconi et al., 2019; Babbin et al., 2021; Gavish et al., 2021; Kitamura et al., 2021). For example, the confirmation of anaerobic nitrogen metabolism in coral microorganisms sheds new light on coral and reef productivity (Babbin et al., 2021). Pollock et al. (2018) found evidence of coral–

microbe phylosymbiosis, in which coral microbiome composition and richness reflect coral phylogeny.

The majority of coral microbiology studies focus on coral adults (Rohwer et al., 2002; Ceh et al., 2011). To date, the coral–bacteria symbiosis particularly the relationship between the coral host and its intracellular bacteria is just beginning to be acknowledged (Kitamura et al., 2021; Maire et al., 2021). Bernasconi et al. (2019) indicated that the uptake and structure of bacterial communities were developmental, supporting the idea that microbial communities are likely to play specific roles within the distinct life history stages of the coral host. The establishment of coral microbial community in the early developmental stages of corals is fundamental to coral fitness, health, and survival (Apprill et al., 2009; Lema et al., 2014; Lema et al., 2016). However, though the diversity of symbiotic bacteria has been revealed in early life stages of some corals (Sharp et al., 2010; Sharp et al., 2012; Lema et al., 2014; Zhou et al., 2017; Epstein et al., 2019), we still rarely know about the dynamic profiles of bacteria at the different development stages of one coral because it is difficult to get enough male and female gametes and larvae at the early development stage of corals. Thus, investigation on the dynamic community profiles of symbiotic bacteria at the different life stages of one coral is with great importance for our understanding of the possible roles of bacterial symbionts, particularly stage-specific bacteria, in the host's development and coral–bacteria symbioses (Sharp et al., 2010; Lema et al., 2014).

Dipsastraea favus is a hermaphrodite, broadcast-spawning species coral belonging to the family Merulinidae (Shlesinger and Woesik, 2021), which is able to release gamete bundles with both sperms and eggs during its spawning period. In this study, with the seawater as control, the bacterial diversity in the coral *D. favus* parent, eggs and larvae was investigated, respectively, using high-throughput sequencing of 16S ribosomal RNA gene, with the aim to reveal the coral's dynamic profiles of symbiotic bacteria and consequently find stage-specific bacteria for the future understanding of *D. favus*–microbe symbioses and utilization of *D. favus* symbiotic microbes.

Materials and methods

Coral spawning and larval cultivation

Eight colonies of hermaphrodite *D. favus* (ca. 30 × 40 × 5 cm^3), 10 m away from each other, were collected at the Egong bay (114°29'E, 22°29'N), at a depth of 5 m, Shenzhen, China, approximately 1 month before spawning, and cultivated in pools with circulating natural seawater, and the temperature was kept at 26°C. The spawning occurred on 9 June 2017. The gamete bundles from each of the eight coral parents were captured and transferred to separate experimental tanks. After floating to the surface, the gamete bundles ruptured and released eggs and sperm. The eggs were collected and fixed immediately after being washed by sterile seawater. The seawater in tanks was not changed after spawning to avoid loss of sperm and kept at 26°C. Approximately 4 days after fertilization, most larvae tended to swim close to the substrate and searched for settlement locations, indicating that they were ready to settle and metamorphose soon; thus, we collected the 4-day-old planula larvae before they became settled juveniles. Because of the number limitation of eggs, no juveniles were collected in this study. The coral eggs, 4-day-old larvae, and parent fragments (Figure 1) were washed by sterile seawater several times to remove the surface microbes and fixed in RNALater. In order to provide enough biomass for successful PCR amplification, at least 50 eggs and 20 individuals of 4-day-old larvae were collected as one replicate for each coral parent. Approximately 2,000 ml of seawater in tanks was filtered by a 0.22-µm filter as one replicate. All samples with three duplicates were frozen in −80°C for further processing.

DNA extraction and high-throughput sequencing

For coral adult DNA extraction, fragments of approximately 1 cm^2 including mucus, tissue, and skeleton were ground in liquid nitrogen by mortar and pestle. Proteinase (20 ml) was added to egg using Qiagen DNeasy Plant Mini Kit (cat. no. 69104), following the manufacturer's instructions. The extracted DNA was preserved at −20°C. DNA samples, including eggs, larvae, adult clones, and tank seawater (3 replications for each), were PCR-amplified using the following primers: forward primer (CCTACGGRRBGCASCAGKVRVGAAT) and reverse primer (GGACTACNVGGGTWTCTAATCC), which were designed by GENEWIZ (Suzhou, China), targeting the hypervariable regions V3–V4 of 16S ribosomal RNA gene for bacteria. The PCR solution contained 2.5 µl of TranStart Buffer, 2 µl of dNTPs, 0.5 µl of TransStart Taq DNA polymerase, 20 ng of total DNA, 1 µl of each primer, and ddH_2O in a final volume of 25 µl. The PCR amplification

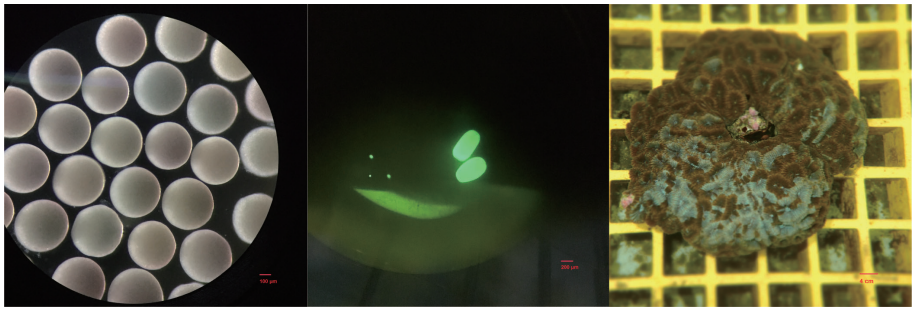


FIGURE 1
The eggs (left), larvae (middle), and adult (right) of *D. favus*.

reaction was performed with the following procedure: initial denaturing at 94°C for 3 min; 35 cycles at 94°C for 5 s, 57°C for 1.5 min, and 72°C for 10 s; final extension at 72°C for 5 min. The library quality of PCR products was validated using Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA, USA). The library concentration was quantified by Qubit2.0 Fluorometer (Invitrogen, Carlsbad, CA). All the amplicon products were sequenced on an Illumina Miseq platform according to the manufacturer's instructions, using 2X300 bp mode at GENEWIZ (Suzhou, China).

Sequence analysis

Analysis of the 16S rDNA sequences was conducted using the bioinformatics package QIIME 1.6 (Caporaso et al., 2010). A total of 74,9631 raw reads were generated from the nine samples (three replications for coral parent, eggs, and larvae, respectively). The sequences containing the adapter were discarded by Trimmomatic (Bolger et al., 2014). The paired-end sequence was merged using Pandaseq (Masella et al., 2012) with the parameter (minimum sequence length of 430 bp and maximum length of 470 bp). The chimera sequences detected by Usearch 61 were abandoned. The clean reads were clustered using UCLUST. Operational taxonomic units (OTUs) were selected at 97% identity by pick_open_reference_otus.py with the greengenes database (DeSantis et al., 2006). The OTUs annotated as chloroplast, mitochondria, and unassigned were discarded. After quality control, an average of 27,839 sequences were selected from each sample, ranging from 11,299 to 46,463. For a fair comparative analysis, samples were subsampled to an equal depth of 11,229 and were conducted by core_diversity_analyses.py. Summarized taxonomic abundance at the phylum, class, order, family, and genus levels was analyzed by QIIME.

Statistical analysis

Alpha diversity statistics including OTU, Chao1, and Shannon index were calculated in QIIME by alpha_diversity.py. The differences among groups were analyzed by one-way ANOVA with Tukey's HSD test. OTU count data were square-root-transformed. Bray–Curtis distance matrix was built to examine the patterns of community structure and was visualized using non-metric multidimensional scaling (NMDS) analysis. Pearson correlation vectors were overlaid to demonstrate which taxon has strong correlation with group difference. Multidimensional statistical analysis was performed in PRIMER v7 with the PERMANOVA+ add-on (PRIMER-e, Quest Research Limited, Auckland, NZ).

Results and discussion

A total of 250,555 high-quality 16S rRNA gene sequences (clean reads in Table 1) were obtained from all of the nine larva and adult samples of *D. favus* and seawater sample (three replications each). The rarefaction curves showed that the diversity reached a plateau (Figure S1). It was a pity that no 16S rRNA gene sequences were retrieved from the egg samples by egg DNA sequencing. The number of 16S rRNA gene reads, number of OTUs, and alpha diversity (Shannon and Chao) estimate were calculated with OTUs at 3% dissimilarity (Table 1). The alpha diversity of the bacterial community of *D. favus* larvae is a bit higher than that of adult *D. favus* based on the Shannon index, and the bacterial diversity shows statistical difference among the coral parent, larvae, and seawater samples (Table 1). The main shared bacterial OTUs among the three samples belong to the class of Rhodobacteraceae and Alteromonadaceae (Table S1). The shared OTUs in the coral parent and its larvae mainly belong to Bacteriodaceae and Flavobacteriales, despite the fact that these bacteria were not detected in the seawater (Table S2). In the case of seawater control, its bacterial community is composed of Alphaproteobacteria, Gammaproteobacteria, and Bacteroidetes (Figure 2).

The dominant 16S rRNA sequences from 4-day-old coral larvae are Gammaproteobacteria and Bacteroidetes (Figure 2). In contrast, Alphaproteobacteria dominates the 16S rRNA gene sequences recovered from adult *F. favus*, representing ~71% of the total sequences (Figure 2). The most abundant OTUs in the larvae are unidentified Oceanospirillaceae (OTU2) and *Kordia* sp. (OTU5), while the unidentified Kiloniellales (OTU1) and *Prosthecochloris* sp. (OTU10) are abundant in the adults (Figure 3). Based on non-metric multi-dimensional scaling analysis and permutational multivariate analysis of variance (PERMANOVA), we found obvious variations in the bacterial community composition among the *D. favus* adults, larvae, and seawater (PERMANOVA: $p = 0.03$) (Figure 4), e.g., Oceanospirillaceae, *Kordia*, and Legionellaceae in *D. favus* larvae, Kiloniellales and *Prosthecochloris* in adult *D. favus*, while Oceanospirillaceae, Rhodobacteraceae, and Sediminicole in seawater. Clearly, Figure 4 shows that the *D. favus* adults and larvae host different predominant bacteria. Similarly, significant stage-specific changes were detected in bacterial profiles during the

TABLE 1 Alpha diversity of bacteria in coral and the environmental seawater samples.

Sample	Raw reads	Clean reads	Observed OTUs	Chao1	Shannon
Coral egg	\	\	\	\	\
Coral larva	99,108 ± 16,883.19	29,069 ± 11,217.87	557.76 ± 32.69(a)	717.72 ± 31.632(a)	5.79 ± 0.14(b)
Coral adult	100,181 ± 5,852.11	12,723 ± 1,281.38	613.8 ± 26.94(a)	752.14 ± 20.46(a)	5.42 ± 0.07(a)
Seawater	50,588 ± 35,24.03	41,725 ± 5,064.39	917.2 ± 12.55(b)	1,729.39 ± 28.83(b)	6.14 ± 0.03(c)

Data are the average of three replicates. Letters indicate the significant differences based on the analysis of one-way ANOVA and Tukey's HSD test: same letter indicates that there is no significant difference.

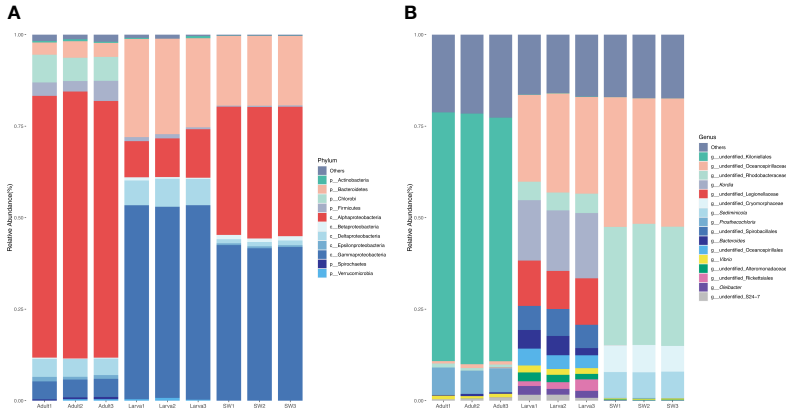


FIGURE 2
Percentage of the microbes in different samples at phylum **(A)** and genus **(B)** levels. A sequence number of less than 0.5% of the total sequence in the sample is classified as other.

development of scleractinian coral *Acropora gemmifera* (Zhou et al., 2017). Lema et al. (2014) suggested dynamic overall bacterial communities associated with the coral *Acropora millepora* at early life stages (larva, 2-week-old juveniles, and 3-, 6-, 12-month-old juveniles). Damjanovic et al. (2020) also suggested that *Pocillopora acuta* recruits harbored dynamic and diverse bacterial assemblages.

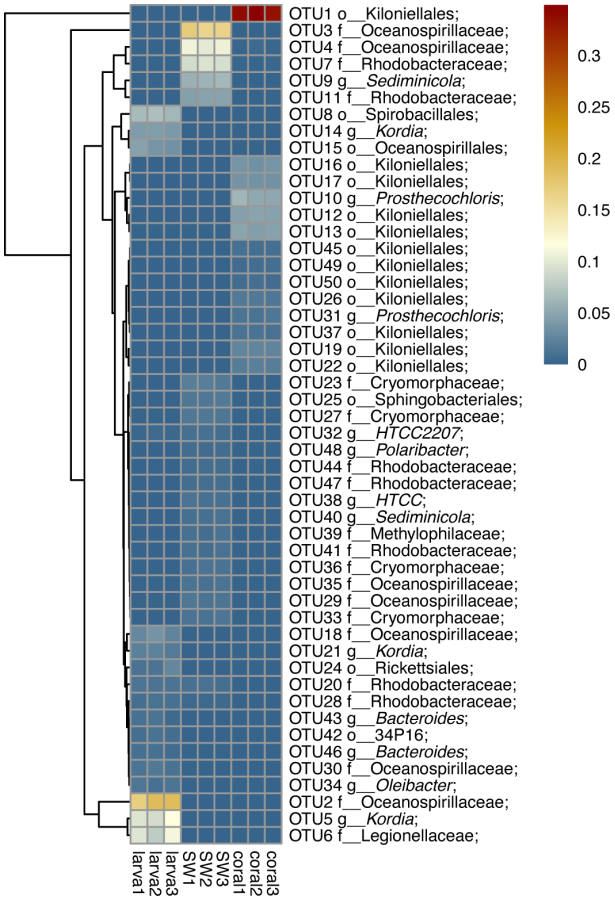


FIGURE 3
Heat map reflecting the relative abundance and distributions of the top 20 OTUs. The scales 0.01 and 0.1 show the abundance at 5% and 10%, respectively.

We detected abundant Oceanospirillaceae in the larvae of *D. favus* rather than in adult *D. favus* (Figures 3, 4). Oceanospirillaceae has been identified from some stony corals (Morrow et al., 2012; Rodriguez-Lanetty et al., 2013). Oceanospirillaceae was only detected in 1-week old juveniles of *A. millepora* and disappeared in the later development juveniles (Lema et al., 2014). The same pattern occurred in the Caribbean coral *Porites astreoides*, which harbors high abundant Oceanospirillaceae in 4-day-old larvae while less Oceanospirillaceae is presented in 11-day-old larvae (Sharp et al., 2012). Previous studies demonstrated that the Oceanospirillales, belonging to Oceanospirillaceae, are able to metabolize DMSP in the coral *A. millepora*, which indicates that Oceanospirillaceae may play an important role in the coral–microbiome interactions because DMSP is a carbon and sulfur source for bacteria and produced by symbiotic dinoflagellates (Bourne et al., 2013). Moreover, the genus *Kordia*, which was abundant in the *D. favus* larvae (Figures 3, 4), was previously detected in the sea urchin and the sea anemone (Becker et al., 2009; Herrera et al., 2017). The genus *Kordia* is one of many genera affiliated with the family Flavobacteriaceae of the phylum Bacteroidetes, well known for its degradation of high-molecular-weight organic matters. The genome of Kordia antarctica IMCC3317T encodes many carbohydrate-active enzymes, some of which are clustered into approximately seven polysaccharide utilization loci, thereby demonstrating the potential for polysaccharide utilization (Lim et al., 2020).

In contrast to *D. favus* larvae, we detected abundant bacteria Kiloniellales (Figures 3, 4) in the adult *D. favus*, which were previously detected in sponges and coral larvae (Sharp et al., 2012; Cleary et al., 2013). In the coral *Montastraea annularis*, Kiloniellales is significantly more abundant in healthy corals than the diseased corals and able to utilize nitrates through denitrification, indicating its positive function on coral's health (Soffer et al., 2015). The future research on the metabolic function of Kiloniellales associated with corals may explain its high abundance in *D. favus*. Another abundant bacterium in adult *D. favus* is *Prosthecochloris* (Figures 3, 4). *Prosthecochloris* sp. is a green sulfur bacterium, frequently detected in corals *Platygyra carnosa* and *Porites lutea* (Li et al., 2014; Cai et al., 2017).

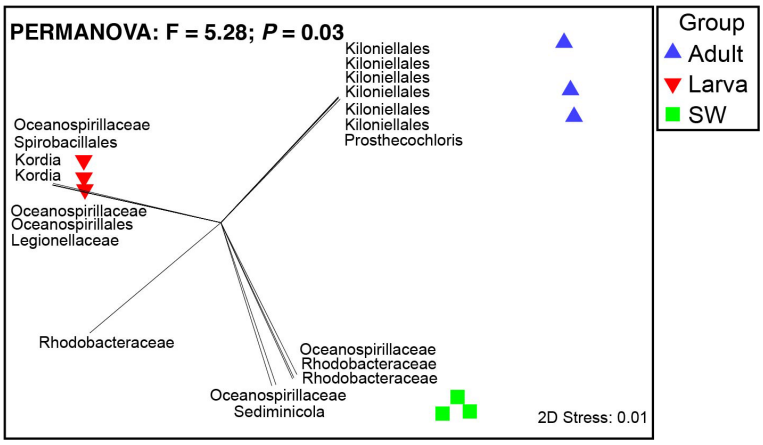


FIGURE 4
Non-metric multi-dimensional scaling analysis of microbial communities in seawater, coral larvae, and adults. Pearson correlation vectors represent the most resolved taxonomy for bacterial OTUs.

Prosthecochloris sp. may provide organic and nitrogenous nutrients to coral host (Yang et al., 2016).

In this study, the most abundant bacteria shared by the larvae and adults are the unidentified *Rhodobacteraceae* (Table S1), which has been detected at the early development stages of corals (Sharp et al., 2012; Zhou et al., 2017). Despite the fact that *Rhodobacteraceae* is generally associated with diseased corals (Roder et al., 2014), the OTUs belonging to *Rhodobacteraceae* are also observed in healthy corals, such as the *Oculina patagonica* and *Isopora palifera* (Chen et al., 2011; Rubio-Portillo et al., 2016). The genus *Roseobacter* belonging to the family *Rhodobacteraceae* is reported to have an important function, that is, supplying fixed organic nitrogen to Symbiodiniaceae or defending against pathogenic bacteria (Nissimov et al., 2009). The *Alteromans* (OTU836759) and HTCC2207 (OTU805719) belonging to the *Alteromonadaceae* were also detected in both *D. favus* larvae and adults in this study (Table S1). A similar bacterial community was detected in the larvae of *D. favus* like other corals, indicating that these bacteria may be important at the early stages of corals.

Based on this study, some core bacteria were shared among the coral larvae, adults, and seawater despite their small proportion, suggesting that the coral *D. favus* is able to obtain bacteria from environmental seawater. The fact that abundant bacterial population in the coral larvae, adults, and the seawater is different (Figure 4) indicates that the selection of bacterial partners by the coral host is not random. Similarly, some studies also suggest the possible horizontal transmission of symbiotic bacteria from the environmental seawater to corals during their development because of the dynamic microbial community change. For instance, Lema et al. (2014); Lema et al. (2016) demonstrated horizontal (environmental) uptake of coral bacteria, e.g., nitrogen-fixing bacteria. Apprill et al. (2009) revealed that the composition of bacteria differed between life stages, i.e., prespawned oocyte bundles, spawned eggs, and week-old planulae, and concluded that the onset of association between microorganisms and the coral *P. meandrina* appeared to occur through horizontal uptake by planulae older than 79 h. Epstein et al. (2019) found that only a small number of bacterial strains were shared between offspring and their respective parents, revealing bacterial establishment as largely environmentally driven at the early life stages and suggested that recruits harbored more variable microbiomes compared to their parents, indicating that winnowing occurs as corals mature (Epstein et al., 2019).

On the contrary, Sharp et al. (2010) suggested a bacterial vertical transmission strategy based on the finding of bacteria in the ectoderm layer in newly released planula larvae of *P. astreoides*, but it was not known whether there were bacteria in the gametes of *P. astreoides*. Leite et al. (2017) found bacteria from the gamete bundles to the planula larvae and parental colony of *M. hispida* based on sequencing, and suggested that spawning coral *M. hispida* could vertically transfer microbes to the offspring based on the similar bacterial community in different stages of *M. hispida*, but the presence of bacterial cells in association with the (female and male) gametes of *M. hispida* through microscopy was not verified. Zhou et al. (2017) found that the microbial communities in the eggs of *A. gemmifera* were highly similar to those in parental colonies, indicating a transmission from coral parents to eggs. However, the possibility that some bacteria are eventually transmitted inside the gametes or larvae from the environments cannot be excluded. Most likely, the coral offspring can be partly colonized postspawning through the uptake of microbial associates released by the parental colonies into the surrounding seawater. In the case of coral *D. favus* in this study, no bacteria were detected by the 16S rRNA gene amplification from the egg samples, but we could not exclude the presence of bacteria in *D. favus* eggs because the failure in detecting the bacterial sequences in eggs may result from the very low abundance of bacteria or the primers and PCR procedure used. Meanwhile, in this study, some of the bacteria shared between the coral parents and larvae, and stage-specific bacteria were not detected in the seawater, suggesting the possible vertically transfer. Whether there is a vertical transmission of bacteria by coral *D. favus* needs future in-depth study, e.g., the confirmation of bacteria in the eggs or gametes of *D. favus* using electron microscopy and fluorescence *in situ* hybridization (FISH). Meanwhile, the characteristics of the symbiotic community profiles of *D. favus* in different regions need to be wholly investigated using abundant

samples from all life stages, e.g., gametes, larvae, and the early juvenile.

According to the results from this study, the larvae of *D. favus* harbor a bit higher number of diverse bacteria in comparison to adult *D. favus*. This may be explained by the different selective ability of corals at different life stages. It has been proposed that morphology change at different life stages may provide different complex niches, which may lead to a significant shift in the microbial community structure (Zhou et al., 2017). An initial higher diversity of bacterial community at the early life stage of corals, followed by low species diversity associated with the adults, has been observed (Franzenburg et al., 2013; Lema et al., 2014). The same situation occurs in the coral *Acropora gemmifera*, which has higher bacterial diversity at the larvae and 6-day-old juvenile stages than in adults (Zhou et al., 2017). The higher diversity of bacteria at the early life stage (e.g., larvae period) of corals provides a higher possibility for nutrient cycling due to the presence of special bacteria, such as *Alteromonas* and *Roseobacter*, which have been suggested to help coral larvae acquire additional nutrients in the oligotrophic environment and consequently increase the survival rate and fitness of corals at the early life stage (Lema et al., 2016). Moreover, the higher diversity of bacteria in the larvae may increase the possibility for more production of antibiotics, which can protect coral larvae from pathogens (Sharp et al., 2012; Lema et al., 2014).

In summary, we investigated the dynamic profiles of bacteria in the coral *D. favus* adult and its eggs and larvae by the high-throughput sequencing method, with the seawater as control. The *D. favus* parent and its larvae host variant bacterial community with larvae have a higher diversity. The shared bacteria among *D. favus* adult and its larvae, and seawater suggest that *D. favus* is able to acquire specific bacteria from the environment. The observed stage-specific bacteria, e.g., Oceanospirillaceae, *Kordia*, and Legionellaceae in *D. favus* larvae, and Kiloniellales and *Prosthecochloris* in adult *D. favus*, could be an important clue to explain the biological or ecological functions of bacteria in *D. favus* holobionts.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: NCBI with SRR number SRR12687352-12687360 and SRR12687344-12687349.

Author contributions

JX, ZX, and XY finished the experiment. JX, GC, and YX analyzed the data. BL and BX participated in the experiment design. JX and ZL designed the study plan and wrote the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This work was funded by the National Natural Science Foundation of China (42176146 and 41776138), the Major National Scientific Research Project, China (2013CB956100), and the Shenzhen Science and Technology R&D Fund (KJYY20180213182720347).

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.2023.1055848/full#supplementary-material

References

Apprill, A., Marlow, H. Q., Martindale, M. Q., and Rappe, M. S. (2009). The onset of microbial associations in the coral *Pocillopora meandrina*. *ISME J.* 3, 685–699. doi: 10.1038/ismej.2009.3

Babbin, A. R., Tamasi, T., Dumit, D., Weber, L., Rodríguez, M. V. I., Schwartz, S. L., et al. (2021). Discovery and quantification of anaerobic nitrogen metabolisms among oxygenated tropical Cuban stony corals. *ISME J.* 15, 1222–1235. doi: 10.1038/s41396-020-00845-2

Becker, P. T., Gillan, D. C., and Eeckhaut, I. (2009). Characterization of the bacterial community associated with body wall lesions of *Tripneustes gratilla* (Echinoidea) using culture-independent methods. *J. Invertebr. Pathol.* 100, 127–130. doi: 10.1016/j.jip.2008.11.002

Bernasconi, R., Stat, M., Koenders, A., Paparini, A., Bunce, M., and Huggett, M. J. (2019). Establishment of coral-bacteria symbioses reveal changes in the core bacterial community with host ontogeny. *Front. Microbiol.* 10. doi: 10.3389/fmicb.2019.01529

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170

Bourne, D. G., Dennis, P. G., Uthicke, S., Soo, R. M., Tyson, G. W., and Webster, N. (2013). Coral reef invertebrate microbiomes correlate with the presence of photosymbionts. *ISME J.* 7, 1452–1458. doi: 10.1038/ismej.2012.172

Cai, L., Zhou, G., Tian, R. M., Tong, H., Zhang, W., Sun, J., et al. (2017). Metagenomic analysis reveals a green sulfur bacterium as a potential coral symbiont. *Sci. Rep.* 7, 9320. doi: 10.1038/s41598-017-09032-4

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

Ceh, J., Van Keulen, M., and Bourne, D. G. (2011). Coral-associated bacterial communities on ningaloo reef, Western Australia. *FEMS Microbiol. Ecol.* 75, 134–144. doi: 10.1111/j.1574-6941.2010.00986.x

Chen, C. P., Tseng, C. H., Chen, C. A., and Tang, S. L. (2011). The dynamics of microbial partnerships in the coral *Isopora palifera*. *ISME J.* 5, 728–740. doi: 10.1038/ismej.2010.151

Cleary, D. F., Becking, L. E., de Voogd, N. J., Pires, A. C., Polonia, A. R., Egas, C., et al. (2013). Habitat- and host-related variation in sponge bacterial symbiont communities in Indonesian waters. *FEMS Microbiol. Ecol.* 85, 465–482. doi: 10.1111/1574-6941.12135

Damjanovic, K., Blackall, L. L., Menéndez, P., and van Oppen, M. J. H. (2020). Bacterial and algal symbiont dynamics in early recruits exposed to two adult coral species. *Coral Reefs* 39, 189–202. doi: 10.1007/s00338-019-01871-z

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

Epstein, H. E., Torda, G., and Munday, P. L. (2019). Parental and early life stage environments drive establishment of bacterial and dinoflagellate communities in a common coral. *ISME J.* 13, 1635–1638. doi: 10.1038/s41396-019-0358-3

Franzenburg, S., Fraune, S., Altrock, P. M., Kunzel, S., Baines, J. F., Traulsen, A., et al. (2013). Bacterial colonization of hydra hatchlings follows a robust temporal pattern. *ISME J.* 7, 781–790. doi: 10.1038/ismej.2012.156

Gavish, A. R., Shapiro, O. H., Kramarsky-Winter, E., and Vardi, A. (2021). Microscale tracking of coral-vibrio interactions. *ISME Commun.* 1, 1–18. doi: 10.1038/s43705-021-00016-0

Herrera, M., Ziegler, M., Voolstra, C. R., and Aranda, M. (2017). Laboratory-cultured strains of the sea anemone *Exaiptasia* reveal distinct bacterial communities. *Front. Mar. Sci.* 4. doi: 10.3389/fmars.2017.00115

Kitamura, R., Miura, N., Ito, M., Takagi, T., Yamashiro, H., and Nishikawa, Y., et al. (2021). Specific detection of coral-associated Ruegeria, a potential probiotic bacterium, in corals and subtropical seawater. *Mar Biotechnol.* 23, 576–589. doi: 10.1007/s10126-021-10047-2

Knowlton, N., and Rohwer, F. (2003). Multispecies microbial mutualisms on coral reefs: The host as a habitat. *Am. Nat.* 162, S51–S62. doi: 10.1086/378684

Krediet, C. J., Ritchie, K. B., Paul, V. J., and Teplitski, M. (2013). Coral-associated micro-organisms and their roles in promoting coral health and thwarting diseases. *Proc. Biol. Sci.* 280, 20122328. doi: 10.1098/rspb.2012.2328

Leite, D. C., Leao, P., Garrido, A. G., Lins, U., Santos, H. F., Pires, D. O., et al. (2017). Broadcast spawning coral *Mussismilia hispida* can vertically transfer its associated bacterial core. *Front. Microbiol.* 8. doi: 10.3389/fmicb.2017.00176

Lema, K. A., Bourne, D. G., and Willis, B. L. (2014). Onset and establishment of diazotrophs and other bacterial associates in the early life history stages of the coral *Acropora millepora*. *Mol. Ecol.* 23, 4682–4695. doi: 10.1111/mec.12899

Lema, K. A., Clode, P. L., Kilburn, M. R., Thornton, R., Willis, B. L., and Bourne, D. G. (2016). Imaging the uptake of nitrogen-fixing bacteria into larvae of the coral *Acropora millepora*. *ISME J.* 10, 1804–1808. doi: 10.1038/ismej.2015.229

Li, J., Chen, Q., Long, L. J., Dong, J. D., Yang, J., and Zhang, S. (2014). Bacterial dynamics within the mucus, tissue and skeleton of the coral *Porites lutea* during different seasons. *Sci. Rep.* 4, 7320. doi: 10.1038/srep07320

Lim, Y., Kang, I., and Cho, J.-C. (2020). Genome characteristics of *Kordia antarctica* IMCC3317T and comparative genome analysis of the genus kordia. *Sci. Rep.* 10, 14715. doi: 10.1038/s41598-020-71328-9

Maire, J., Girvan, S. K., Barkla, S. E., Perez-Gonzalez, A., Suggett, D. J., and Blackall, L. L., et al. (2021) Intracellular bacteria are common and taxonomically diverse in cultured and in hospite algal endosymbionts of coral reefs. *ISME J.* 15, 2028–2042. doi: 10.1038/s41396-021-00902-4

Masella, A. P., Bartram, A. K., Truszkowski, J. M., Brown, D. G., and Neufeld, J. D. (2012). PANDAseq: paired-end assembler for illumina sequences. *BMC Bioinf.* 13, 31. doi: 10.1186/1471-2105-13-31

Morrow, K. M., Moss, A. G., Chadwick, N. E., and Liles, M. R. (2012). Bacterial associates of two Caribbean coral species reveal species-specific distribution and geographic variability. *Appl. Environ. Microbiol.* 78, 6438–6449. doi: 10.1128/AEM.01162-12

Nissimov, J., Rosenberg, E., and Munn, C. B. (2009). Antimicrobial properties of resident coral mucus bacteria of *Oculina patagonica*. *FEMS Microbiol. Lett.* 292, 210–215. doi: 10.1111/j.1574-6968.2009.01490.x

Pollock, F. J., McMinds, R., Smith, S., Bourne, D. G., Willis, B. L., and Medina, M., et al. (2018). Coral-associated bacteria demonstrate phylosymbiosis and cophylogeny. *Nat Commun.* 9, 4921. doi: 10.1038/s41467-018-07275-x

Roder, C., Arif, C., Bayer, T., Aranda, M., Daniels, C., Shibl, A., et al. (2014). Bacterial profiling of white plague disease in a comparative coral species framework. *ISME J.* 8, 31–39. doi: 10.1038/ismej.2013.127

Rodriguez-Lanetty, M., Granados-Cifuentes, C., Barberan, A., Bellantuono, A. J., and Bastidas, C. (2013). Ecological inferences from a deep screening of the complex bacterial consortia associated with the coral, *Porites astreoides*. *Mol. Ecol.* 22, 4349–4362. doi: 10.1111/mec.12392

Rohwer, F., Seguritan, V., Azam, F., and Knowlton, N. (2002). Diversity and distribution of coral-associated bacteria. *Mar. Ecol. Prog. Ser.* 243, 1–10. doi: 10.3354/meps243001

Rubio-Portillo, E., Santos, F., Martinez-Garcia, M., de Los Rios, A., Ascaso, C., Souza-Egipsy, V., et al. (2016). Structure and temporal dynamics of the bacterial communities associated to microhabitats of the coral *Oculina patagonica*. *Environ. Microbiol.* 18, 4564–4578. doi: 10.1111/1462-2920.13548

Sharp, K. H., Distel, D., and Paul, V. J. (2012). Diversity and dynamics of bacterial communities in early life stages of the Caribbean coral *Porites astreoides*. *ISME J.* 6, 790–801. doi: 10.1038/ismej.2011.144

Sharp, K. H., Ritchie, K. B., Schupp, P. J., Ritson-Williams, R., and Paul, V. J. (2010). Bacterial acquisition in juveniles of several broadcast spawning coral species. *PloS One* 5, e10898. doi: 10.1371/journal.pone.0010898

Shlesinger, T., and Woesik, R. (2021). Different population trajectories of two reef-building corals with similar life-history traits. *J. Anim. Ecol.* 90, 1379–1389. doi: 10.1111/1365-2656.13463

Soffer, N., Zaneveld, J., and Vega Thurber, R. (2015). Phage-bacteria network analysis and its implication for the understanding of coral disease. *Environ. Microbiol.* 17, 1203–1218. doi: 10.1111/1462-2920.12553

Yang, S. H., Lee, S. T. M., Huang, C. R., Tseng, C. H., Chiang, P. W., Chen, C. P., et al. (2016). Prevalence of potential nitrogen-fixing, green sulfur bacteria in the skeleton of reef-building coral *Isopora palifera*. *Limnol. Oceanogr.* 61, 1078–1086. doi: 10.1002/lno.10277

Zhou, G., Cai, L., Yuan, T., Tian, R., Tong, H., Zhang, W., et al. (2017). Microbiome dynamics in early life stages of the scleractinian coral *Acropora gemmifera* in response to elevated pCO_2. *Environ. Microbiol.* 19, 3342–3352. doi: 10.1111/1462-2920.13840

frontiers | Frontiers in Marine Science

TYPE Original Research
PUBLISHED 26 January 2023
DOI 10.3389/fmars.2023.1089349

OPEN ACCESS

EDITED BY
Guangyi Fan,
Beijing Genomics Institute (BGI), China

REVIEWED BY
Shijie Bai,
Institute of Deep-Sea Science and Engineering, Chinese Academy of Sciences (CAS), China
Du Hong,
Shantou University, China
Meng Yan,
City University of Hong Kong, Hong Kong, Hong Kong SAR, China

*CORRESPONDENCE
Jie Li
lijie@ysfri.ac.cn

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

RECEIVED 04 November 2022
ACCEPTED 05 January 2023
PUBLISHED 26 January 2023

CITATION
Yan Y, Wang S, Liu K, Mo Z, Yang H, Rong X and Li J (2023) Divergence of epibacterial community assemblage correlates with malformation disease severity in *Saccharina japonica* seedlings. *Front. Mar. Sci.* 10:1089349. doi: 10.3389/fmars.2023.1089349

Divergence of epibacterial community assemblage correlates with malformation disease severity in *Saccharina japonica* seedlings

Yongwei Yan[1], Shanshan Wang[2], Kuimei Liu[2], Zhaolan Mo[1,3], Huichao Yang[1], Xiaojun Rong[1] and Jie Li[1]*

[1]Laboratory for Marine Fisheries Science and Food Production Processes, Pilot National Laboratory for Marine Science and Technology (Qingdao), Key Laboratory of Maricultural Organism Disease Control, Ministry of Agriculture and Rural Affairs, Yellow Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Qingdao, China, [2]Harbin University of Science and Technology, Weihai, China, [3]Key Laboratory of Tropical Aquatic Germplasm of Hainan Province, Sanya Oceanology Institute, Ocean University of China, Sanya, China

Malformation disease (MD) is one of the major constraints in the mariculture of the commercially important seaweed *Saccharina japonica*, which can cause severe losses of the kelp seedlings in the production process. However, the pathogenesis of MD remains largely undetermined. In this study, using cultivation experiments under laboratory conditions, MD severity was estimated for the young sporophytes of *S. japonica*, which were separately cultivated under different treatment conditions, including light intensity, duration of light exposure, and maturity level of the parent kelp. The associations between MD severity and epibacterial community divergence and assembly patterns were characterized. Higher MD severity significantly correlated with longer light exposure, and with both unmatured and overmatured parent kelp. The bacterial classes γ-Proteobacteria and α-Proteobacteria dominated all samples, but different MD severities were associated with different epibacterial communities. Comparisons of predicted functions for epibacterial communities revealed that longer light exposure led to a depletion in development and regeneration, while overmatured parent kelp resulted in a depletion in glycan biosynthesis and metabolism. Functional comparisons of the epibacterial communities between normal and unmatured parent kelp-generated seedlings revealed significantly different categories related to metabolism, environmental information processing, cellular processes, drug resistance, and bacterial infection. The significantly different pathways between groups, notably those related to bacterial infection and cellular processes, were partly consistent with the differences in bacterial compositions. The assembly of epiphytic bacterial communities was predominately governed by deterministic processes, and less impact was determined when there was significantly higher MD severity except when using overmatured parent kelp. Co-occurrence networks of the epibacterial communities associated with higher MD severity contained fewer nodes and exhibited lower modularity, but had higher graph density and degrees compared with those of seedlings with lower MD severity, indicating more complicated interactions. *Nesterenkonia*, *Glycocaulis*, *Halomonas*, *Pseudoalteromonas*,

Pseudomonas, *Loktanella*, and *Cobetia* were frequently determined keystone taxa in communities associated with higher MD severity. The present study enhances our understanding of the factors significantly associated with MD severity and the potential roles of epimicrobiome in determining the disease severity, which will be useful for disease management in the future.

KEYWORDS

disease severity, epiphytic bacterial community, light exposure, malformation disease, maturity level, *Saccharina japonica*

Introduction

Seaweeds are a phylogenetically diverse group of macroscopic marine algae in marine ecosystems, which include brown, red, and green algae. These multicellular photosynthetic eukaryotes are well-equipped to function as ecosystem engineers by contributing to primary productivity, and by providing food and shelter for aquatic life (Watt and Scrosati, 2014; Hasselström et al., 2018). The cultivation of commercially important seaweed species has become a significant and growing industry owing to their significant value in food, medicines, and cosmetics (Lomartire et al., 2021). However, like other living organisms, seaweeds are susceptible to diseases, leading to significant losses in the natural population and economic losses in commercial species. For example, the annual losses caused by oomycete pathogens can reach approximately 15–20% and 25–30% for the cultivated laver in Korea and China respectively, with up to a 64% loss rate reported in Japan (Gachon et al., 2010).

The fitness and ecological traits of seaweeds are influenced by mutual interactions with the associated microbial community, which can be referred to as the "holobiont" (Wahl et al., 2012; Dittami et al., 2021). The microbial community associated with seaweeds consists of potentially beneficial species along with pathogenic species; thus, bacterial-host interactions have the potential to influence seaweed health and development in two different ways. On the one hand, these interactions can be beneficial by supplying nutrients or assisting with nutrient acquisition (Singh and Reddy, 2015; Dogs et al., 2017), providing chemicals for morphogenesis and development (Goecke et al., 2010; Alsufyani et al., 2020), and conferring resistance to stressful environments (Dittami et al., 2016; Morrissey et al., 2021) and pathogen infections (Saha and Weinberger, 2019; Li et al., 2022). On the other hand, these interactions can be detrimental if opportunistic pathogens colonize and proliferate, which notably occur under unfavorable environments when the host's immunity is undermined (Egan and Gardiner, 2016; Ward et al., 2021; Behera et al., 2022). It is now recognized that seaweed diseases are tightly associated with interactions that occur among macroalgae, bacteria, and the environment. Emerging evidence demonstrates that disruption of the host's natural microbiota under suboptimal conditions is correlated with disease development (Zozaya-Valdés et al., 2015; Ghaderiardakani et al., 2020; Yan et al., 2022). For an infection disease, it is typically characterized by the increase in potentially pathogenic bacteria and/or virulence-related functional genes. For example, under bleaching conditions, the red alga *Delisea pulchra* is significantly enriched in certain members of the bacterial families Rhodobacteraceae and Flavobacteriaceae, in addition to functional genes related to chemotaxis, motility, oxidative stress response, vitamin biosynthesis, and nutrient acquisition (Zozaya-Valdés et al., 2015; Zozaya-Valdés et al., 2017). Collectively, the investigation on seaweed microbiome requires a broader perspective by combining their interactions with host and environments.

Currently, it is difficult to disentangle the causal relationship between the changes in epibacterial community and the occurrences of seaweed disease due to the existence of multiple disease stages and the complexity of potential opportunistic pathogens (Zozaya-Valdés et al., 2015). Unraveling the mechanisms that govern community assembly is crucial to predicting the incidence of disease, and can also guide the establishment of new strategies against community disruption and/or pathogenic infections (De Schryver et al., 2012; Pearson et al., 2018; Li et al., 2022). Recently, two principles have been proposed. The "lottery hypothesis," combining both selective and neutral aspects, suggests that the community assembly depends more on functional level of genes (Burke et al., 2011a). Another hypothesis argues that communities are structured by both untargeted recruitment and targeted deterrence processes from the host (Saha and Weinberger, 2019). These principles combined with other studies suggest that the assembly of microorganisms on seaweeds is controlled by both host- and non-host-related factors, such as the already present colonizers (Longford et al., 2019), host health status (Yan et al., 2022), growth stage of the host (Lemay et al., 2018), and environmental conditions (Florez et al., 2019), which can be categorized into two different ecological processes: deterministic (niche-based) and stochastic (neutral) processes. A deterministic process is expected to drive microbial niche structures that favor one species over the other through host selection and species-to-species interactions (Zhou and Ning, 2017). Alternatively, stochastic processes assume species to be ecologically equivalent, emphasizing the importance of chance colonization, random extinction, and ecological drift (Chase and Myers, 2011; Zhou and Ning, 2017). However, scarce data are available to quantify the relative importance of deterministic and stochastic processes in bacterial assembly on the surfaces of seaweeds, notably when they are plagued by disease.

Increasing evidence demonstrates that responses of a community to disease depend on species-to-species interactions (Zhu et al., 2016; Dai et al., 2020). The alterations can be indicative by the presence of certain taxa within the microbial communities that have greater biotic connectivity to the resident members, irrespective of their abundance;

these can be referred to as "keystone taxa" (Herren and McMahon, 2018). Previous studies have identified keystone taxa (also denoted as "gatekeepers") from the gut microbiota of shrimp, demonstrating their disproportional roles in sustaining stability, dynamics, and function. Such keystone taxa are identified as candidate pathogens in diseased shrimps, further validated by their enrichment in terms of relative abundance and encoded virulence genes (Dai et al., 2018). Moreover, keystone taxa can function as potential probiotics under certain contexts (Xiong, 2018; Dai et al., 2019). Therefore, identifying keystone taxa can provide clues for the identification of causal agents of a disease, and thus offer strategies for the management of maintaining a "healthy" microbiota in disease prevention. Network analysis has emerged as a powerful tool for the integrated understanding of microbe-to-microbe interactions, and the network scores can statistically identify the potential importance of specific microorganisms to the network, such as keystone taxa (Berry and Widder, 2014; Banerjee et al., 2018). For example, taxa with high degree, high closeness centrality, and low betweenness centrality scores in the network are considered as the keystone taxa in the community (Berry and Widder, 2014). Thus, network analysis has been widely applied in the analyses of interactions occurring in multiple ecosystems, including aquaculture systems (Dai et al., 2020; Lin et al., 2021).

The brown seaweed *Saccharina japonica* (Areschoug) C. E. Lane, C. Mayes, Druehl & G. W. Saunders, 2006 has been cultivated in China since the 1920s as a commercial marine algal species (Tseng, 2001). *S. japonica* has an alternate life history comprising generations of micro-gametophyte and macro-sporophyte, and its culture process consists of indoor cultivation of seedlings (i.e., sporelings or young sporophytes) and outdoor cultivation of mature sporophytes (Duan et al., 2015). The artificial seedling-rearing techniques developed in the 1950s substantially contributed to the rapid development of the kelp aquaculture industry in China (Tseng, 2001). Recently, *S. japonica* has been increasingly threatened by various diseases at both the seedling-rearing and open-sea cultivation stages, notably diseases that target young sporophytes, leading to complete failure of seedling production (Peng and Li, 2013; Wang et al., 2014; Li et al., 2020). Malformation disease (MD) is one of the disastrous and frequent diseases in seaweed mariculture, which is becoming an increasingly serious bottleneck in kelp cultivation by causing massive losses of seedlings (Qian et al., 2016; Sheng et al., 2016). According to our field investigations during 2018–2021, the malformation rates in certain tanks were > 50% in kelp genotypes collected in northern China. MD frequently occurs in the early stage of young sporophytes, notably when reaching the stage of 4–8 columns of cells (Wu et al., 1979; Qian et al., 2016). MD was observed to cause abnormal division in some or all of the cells, characterized by expanded cell size and/or irregular division. Eventually, these malformed seedlings disintegrate, leaving no seedlings remaining on the cultivation curtains. The pathogenesis for MD remains largely unknown, although associated bacteria and/or potential factors (i.e. duration of light exposure, maturity level of the parent kelp) have been proposed to play a role (Wu et al., 1979; Qian et al., 2016; Sheng et al., 2016).

In the present study, the MD severity of young sporophytes of *S. japonica* was estimated under laboratory conditions. Sporophytes were separately cultivated under different treatment conditions, including variation in light intensity, duration of light exposure, and maturity level of the parent kelp, respectively. With these experiments, we aimed to reveal the following: (1) factors significantly correlated with MD severity, (2) divergence of the epiphytic bacterial community associated with different MD severities, and (3) alterations in the bacterial assembly patterns associated with different MD severities. These findings are of great significance for the production of kelp seedlings in the field of disease management.

Material and methods

Experimental design and sample collection

The parent kelp used in the present study was collected from Ailun Bay (122°33'32" E, 37°9'14") in July, 2021, and a cultivar that has long been cultivated in this area was used for zoospore collection. Approximately 6 individuals (2 replicate individuals × 3 maturity levels) of the cultivar were collected and then washed, incubated, and stimulated to release zoospores in the laboratory according to standard methods (DB37/T 1190—2009). After reaching a density of approximately 3–5 zoospores per microscopic field (100×) with completed attachment on microscopic slides, the zoospores were incubated under various experimental conditions for seedling incubation (Table 1). Factors previously found to be associated with MD in our investigations and other published reports (Qian et al., 2016; Sheng et al., 2016) were selected to further investigate their associations with MD severity, including light intensity, light duration, and maturity level of the parent kelp (estimated based on standard methods SC/T 2024-2006). For each factor, three treatment groups were established with varying levels of the given factor (referred to as a "set"; Table 1). For the set of light intensity, seedlings generated by normally-matured parent kelp were cultivated at 8°C under different light intensities (500 lx, 1500 lx, and 3000 lx) with a photoperiod of 10h: 14h (light: dark). These seedlings were separately grouped and denoted as L500, L1500, and L3000, respectively (Table 1). For the set of light duration, seedlings generated by normally-matured parent kelp were cultivated at 8°C under the light intensity 1500 lx with different photoperiods (5h: 19h, 10h: 14h, and 15h: 9h), and they were grouped and denoted as T5, T10, and T15, respectively (Table 1). For the set of maturity level, seedlings were generated by different maturity levels of the parent kelp, including unmatured, normally-matured, and overmatured, which were grouped and denoted as UM, NM, and OM, respectively (Table 1). These seedlings were cultivated under

TABLE 1 Experimental sets and treatment groups.

Sets	Treatment groups
Light intensity	500 lx (L500); 1500 lx (L1500) *; 3000 lx (L3000)
Light duration	5h (T5); 10h (T10) *; 15h (T15)
Maturity level	Unmatured (UM), Normal (NM) *, Overmatured (OM)

*These indicate groups derived from normally-matured parent kelp and cultivated under standard conditions including temperature 8°C, light intensity 1500lx, and light exposure 10h/day (DB37/T 1190—2009).

standard environmental conditions, including 8°C, 1500 lx, 10-h light exposure (DB37/T 1190—2009). Within these groups, the seedlings of the groups L1500, T10, and NM represented the standard cultivation conditions (normally-matured parent kelp and standard environmental conditions), which can function as the control group for each set.

Assays for each set were performed in separate laboratories. Nitrogen and phosphorus nutrients, which are essential for seedling growth, were supplied in sterile natural seawater within the culture Petri dishes according to the standard protocols (DB37/T 1190—2009), and the seawater was exchanged every 3 days. Twelve replicate slides were established as biological replicates for each treatment group and one slide per Petri dish was incubated. After approximately 20 days of incubation when sporophytes entered the stage of 4–8 columns of cells, the malformation severity was estimated based on microscopic observation. For each slide, 10 microscopic fields were selected to count the total number of malformed seedlings. MD severity was then calculated as (total observed malformed seedlings/total observed individuals) × 100%. It has been reported that certain gametophytes cannot convert to sporophytes at a normal pace due to slow growth or death under inappropriate conditions (e.g., lower or higher light intensity, shorter or longer light exposure) (Duan et al., 2015). Therefore, an unconverted proportion was also calculated for each replicate as (total observed individuals not converted to sporophytes/total observed individuals) ×100%. This is useful to evaluate seedling growth and development status so that we can select groups at approximately the same status for the comparisons of MD severity and epibacterial community. Finally, six randomly selected replicates were used to scrape seedlings on cotton swabs for analysis of the epiphytic bacterial communities.

DNA extraction, PCR amplification, and illumina sequencing

For detection and quantification of epiphytic bacteria, DNA was isolated from each replicate using the EZNA Soil DNA Kit (OMEGA Bio-Tek, Norcross, GA, USA). To avoid any loss of the firmly attached bacteria, both cotton swab and the collected seedlings on it were used for DNA extraction. The extracted DNA was assessed by gel electrophoresis, and the qualified DNA was then stored at −20°C until used for amplification. To preclude amplification of plastid sequences, the primers 799F and 1193R were used to amplify the V5–V7 hypervariable regions of the 16S rRNA gene which has been used in the study of plants and algae microbiota (Zhang et al., 2019; Aires et al., 2021). The amplicon libraries were constructed using TruSeq Nano DNA LT Library Prep Kit (Illumina, San Diego, CA, USA), following the manufacturer's recommendations, and index codes were added. Subsequently, the amplicon libraries were sequenced on a NovaSeq 6000 sequencer (Illumina), and 250-bp paired-end reads were generated. The raw sequence data have been deposited in the Genome Sequence Archive (Chen et al., 2021) in National Genomics Data Center (Xue et al., 2022), China National Center for Bioinformation/Beijing Institute of Genomics, Chinese Academy of Sciences (GSA: CRA008545) that are publicly accessible at https://ngdc.cncb.ac.cn/gsa.

Sequence processing

The paired-end reads were joined using the command "fastq_mergepairs" in VSEARCH (version 2.15.1) (Rognes et al., 2016), and then truncated (fastq_stripleft 19 and fastq_stripright 18) and quality-filtered (fastq_maxee 1.0) using the command "fastx_filter" in VSEARCH. High-quality reads were clustered into zero-radius operational taxonomic units (ZOTUs) using the unoise3 algorithm in USEARCH (version 10.0.240) (Edgar, 2010). Taxonomy assignment for representative ZOTU sequences was also performed in USEARCH using the sintax algorithm against the RDP 16S training set 18 database (cutoff = 0.8). Sequences assigned as plastids and mitochondria, as well as those not belonging to the domain Bacteria, were discarded. Taxonomic information of ZOTUs in the following analyses, e.g., community comparisons and keystone taxa, were also checked using Basic Local Alignment Search Tool (BLAST). A rarefied ZOTU table at the same depth was used to calculate α- and β-diversities in R using the "vegan" package (version 2.5-6) (Oksanen et al., 2019). Relative codes for taxa filtering, taxa circular visualization plots, and diversity calculation were obtained from a direct application documented in Zhang et al. (2019).

Functional prediction of the epibacterial community

The ZOTU table was normalized by dividing the abundance of each ZOTU by its predicted 16S rRNA gene copy number to produce the Kyoto Encyclopedia of Genes and Genomes (KEGG) orthology (KO) functional categories using PICRUSt2 (Douglas et al., 2020). The script "summarizeAbundance.py" was used to calculate pathway count abundance at different levels, which is documented in EasyMicrobiome (version 1.16)[1].

Null model analyses of epiphytic bacterial communities

To disentangle the relative contributions of stochastic and deterministic processes to microbial community assembly, the NST value was used to quantify ecological stochasticity in communities (Ning et al., 2019). This index measures the relative importance of stochasticity versus determinism considering both the situations where deterministic factors drive the communities to be more similar or dissimilar than random patterns (Ning et al., 2019); NST < 0.5 indicates a more deterministic community and NST > 0.5 indicates a more stochastic community assembly. NST was calculated based on Jaccard and Bray–Curtis similarity metrics using the null model algorithm PF (with fixed data richness and proportional taxa occurrence frequency). NST analysis was performed using the "NST" R package (Ning et al., 2019).

Construction of the co-occurrence network

The networks were individually constructed to infer associations between ZOTUs for each treatment group. To construct the co-

occurrence network, compositionality-robust correlations from the median of 20 iterations and 100 bootstrap samples were used to infer pseudo *p* values (Kurtz et al., 2015). The inferred correlations were restricted to those having correlation coefficients > 0.8 or < −0.8 ($p < 0.05$, two-sided). Visualization was performed with the Gephi 0.9.3 platform, using a Fruchterman Reingold layout (Bastian et al., 2009). To describe the topology of the networks, the following metrics were calculated: average path length, network diameter, average degree, clustering coefficient, and graph density. Average path length refers to the average network distance between all pairs of nodes, network diameter refers to the greatest distance between the nodes that exist in the network, average degree refers to the average connections of each node with another unique node in the network, clustering coefficient represents the degree to which the nodes tend to cluster together, and graph density refers to the intensity of connections among nodes (Newman, 2002; Newman, 2006). A higher average degree, clustering coefficient, and graph density suggest a more connected network. In addition, lower average path lengths and diameters indicate closer associations in the network (Barberán et al., 2012; Ma et al., 2016). The node-level topological features, including degree, betweenness, and closeness centrality, were also calculated for each node. The ZOTUs with the highest degree, highest closeness centrality, and lowest betweenness centrality scores were considered as the keystone taxa (Berry and Widder, 2014).

Statistical analysis

One-way analysis of variance (ANOVA) was used to test the significance in differences of MD severity, α-diversity indices, and normalized stochasticity ratio (NST) values among groups within each set (Chambers et al., 1992). Constrained principal co-ordinates analysis (CPCoA) was performed based on the Bray–Curtis distance matrix and visualized using the web server ImageGP (Chen et al., 2022). A random forest (RF) algorithm was used to screen significantly differentiated bacteria ZOTUs among treatment groups using the "randomForest" R package (Breiman, 2001). The significance of differentially abundant ZOTUs was also tested in STAMP (version 2.1.3) based on ANOVA with a Tukey *post-hoc* test ($p < 0.05$) (Parks et al., 2014). STAMP (version 2.1.3) was also used to determine bacterial ZOTUs that are significantly different between groups with significantly different MD severities, and Welch's *t*-test was used to test the pair-wise significance ($p < 0.05$). All the aforementioned analyses performed in STAMP were also applied to determine significantly different functions based on their relative abundance, and categories related to humans were excluded from the statistical analysis.

Results

Malformation severity for different treatments

The presence and severity of MD were observed when the experiments were terminated. As shown in Figure 1, several types of individuals were observed in the present study, including normal, unconverted, and malformed individuals. Given unsuccessful cultivation for zoospores within certain replicates, the MD severity was estimated using the remaining replicates for each treatment

FIGURE 1
Typical characteristics of different types of seedlings observed in the present study. **(A)** Normal seedlings. Image from T15 group; **(B)** Individual not converted to sporophyte as indicated by a black arrow. Image from T10 group; **(C, D)** Malformed seedlings indicated by red arrows. Image from T15 group.

condition (Table 2). For the treatments with different light intensities, the MD severity was not significantly different among groups. Lower light intensity also prevented more gametophytes from transforming into sporophytes, resulting in a significantly higher unconversion rate (Table 2). Longer light exposure generated a significantly higher MD severity than found in the seedlings cultivated with standard light conditions (1500 lx, L1500). Although shorter light exposure (5 h, T5) generated significantly lower MD severity, more gametophytes could not properly convert into sporophytes under this condition, with an unconversion rate reaching over 66% (Table 2). In particular, the normally-matured parent kelp (NM) generated significantly lower MD severity of the seedlings than that derived from both unmatured (UM) and overmatured (OM) parent kelp (Table 2). These results demonstrated that groups of T10 and T15, NM and UM, and NM and OM warrant further comparisons concerning epibacterial communities due to their significantly different MD severity and similar growth status (i.e., unconverted proportion).

Epiphytic bacterial community composition and diversity

All samples were dominated by the bacterial classes γ-Proteobacteria and α-Proteobacteria, which accounted for 92.85 ± 6.59% (mean ± standard deviation) of the total sequences (Figure 2). The relative abundance of γ-Proteobacteria declined as the light intensity increased (Figure 2A). In contrast, the relative abundance of α-Proteobacteria increased as the light intensity increased (Figure 2A). Similar patterns were also observed for the groups cultivated under different light durations (Figure 2B). The relative abundance of γ-Proteobacteria was the highest in the group cultivated with normally matured parent kelp (NM), but was the lowest in the UM group, and did not tend to decline progressively with maturity level (Figure 2C). Conversely, the changes in the relative abundance of α-Proteobacteria showed the opposite trend with respect to the maturity level of parent kelp (Figure 2C).

Observed species and Shannon index was used to estimate the α-diversity for all treatment groups. The light intensity did not have a significant effect on α-diversity, although the diversity of the L1500 group (representing the standard culture condition) was higher than that of the other two groups (Figure 3A). The T5 group showed a significantly lower Shannon index compared with that of the groups exposed to light for 10 h (T10) and 15 h (T15), respectively (Figure 3A). Although the highest α-diversity was found for the T10 group, there was no significant difference from that of the T10 group (Figure 3A). With respect to the maturity of the parent kelp, NM showed a higher species richness value and a relatively lower Shannon index than the other two groups, although no significant difference was observed (Figure 3A). CPCoA plots revealed that the bacterial communities were separately clustered according to the different treatments (Figure 3B). Further, perMANOVA confirmed significant associations between treatments and the community differences ($p < 0.05$), with light intensity, light duration, and maturity level explaining 16.6%, 14.8%, and 16.9% of the community variations, respectively (Figure 3B).

TABLE 2 Malformation severity for different treatment groups.

	L500 (n = 9)	L1500 (n = 10)	L3000 (n = 10)
MD severity*	34.67 ± 17.21% **a**	30.75 ± 6.61% **a**	39.52 ± 13.38% **a**
Unconverted proportion**	39.55 ± 19.37 **a**	19.44 ± 10.60 **b**	17.64 ± 10.04 **b**
	T5 (n = 12)	**T10 (n = 12)**	**T15 (n = 9)**
MD severity	10.90 ± 11.15 **a**	35.06 ± 15.52 **b**	50.55 ± 9.96 **c**
Unconverted proportion	66.59 ± 21.91 **a**	4.58 ± 3.04 **b**	0.69 ± 1.37 **b**
	UM (n = 11)	**NM (n = 6)**	**OM (n = 10)**
MD severity	44.34 ± 9.27% **a**	31.66 ± 6.86% **b**	45.04 ± 11.69% **a**
Unconverted proportion	10.69 ± 5.42 **a**	10.82 ± 8.95 **a**	14.89 ± 11.43 **a**

*This severity indicates the malformation disease (MD) severity, which is calculated as follows for each replica: the total observed malformed seedlings/total observed individuals × 100%.
**This value indicates the proportion of individuals not converted to sporophytes, which is calculated as follows for each replica: total observed individuals not converted to sporophytes/total observed individuals × 100%.
Severity percentage values are indicated as mean ± standard deviation, and the letter "n" indicates the number of biological replicates. Different bolded letters behind the severity percentages indicate significant differences (ANOVA, $p < 0.05$) among treatment groups.

Divergent bacterial communities among different treatments

The associated bacterial communities were then compared at the ZOTU level to reveal the community divergence according to treatment condition. Based on the RF algorithm and ANOVA, approximately 6, 16, and 3 ZOTUs were determined to differentiate treatment groups for the sets of light intensity, light duration, and maturity level, respectively (Figure 4A). This indicated that these bacterial ZOTUs, though distributed in various genera (Table S1), can potentially serve as indicators of host status. Bacterial communities associated with seedlings with different MD severities were further compared based on Welch's *t*-test. Approximately 12, 17, and 18 ZOTUs were determined to be significantly different between groups in the comparisons of light duration (T10 vs. T15), NM vs. UM, and NM vs. OM, respectively (Figure 4B). Based on BLAST and phylogenetic analysis, these ZOTUs were phylogenetically diverse notably those affiliated with Rhodobacteraceae (Figure S1 and Table S2). Specifically, seven of the 12 significantly different ZOTUs between T10 and T15 were enriched in T10, which were distributed in seven genera. By contrast, the relative abundances of five ZOTUs affiliated with *Sulfitobacter*, *Fluviibacterium* (two ZOTUs), *Psychroserpens*, and *Pacificibacter* were higher in T15 than in T10 (Figure 4B). Ten of the 17 significantly different ZOTUs between NM and UM were more significantly abundant in NM, including those affiliated with *Colwellia* (two ZOTUs), *Pseudoalteromonas* (four ZOTUs), *Pseudomonas* (two ZOTUs), and *Cobetia* (two ZOTUs) (Figure 4B). Compared with those in OM, seven ZOTUs were significantly enriched in NM, distributed in *Colwellia* (two ZOTUs), *Cobetia* (two ZOTUs), *Pseudoalteromonas*, *Yoonia*, and

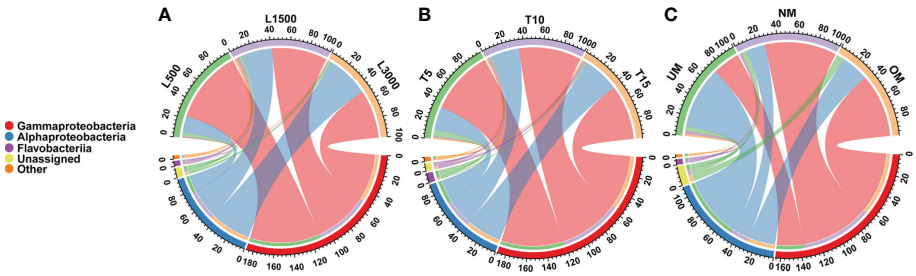


FIGURE 2
Distribution of major bacterial taxa in different treatment groups at the class level. **(A)** Groups treated with different light intensities; **(B)** Groups treated with different durations of light exposure; **(C)** Groups derived from different maturity levels of parent kelps.

Pseudomonas, respectively. Whereas, eleven ZOTUs were significantly depleted in NM, including those assigned to *Pseudomonas* (four ZOTUs), *Parasedimentitalea* (two ZOTUs), *Halopseudomonas* (two ZOTUs), *Paracoccus* (two ZOTUs), and *Lentibacter* (Figure 4B).

Divergent potential functional groups among different treatments

From the comparisons of PICRUSt2 predictions, no significant functional difference was determined among the groups treated with different light durations. However, significant differences among groups treated with different light intensities were detected, which were mainly derived from pathways related to amino acid metabolism, carbohydrate metabolism, and metabolism-related protein families (Figure 5A). For groups derived from parent kelp with different maturity levels, significant differences were mainly distributed in pathways related to metabolism, including energy metabolism, metabolism of other amino acids, and metabolism of terpenoids and polyketides, in addition to the aforementioned three categories that were divergent according to light intensity (Figure 5A). Moreover, pathways related to environmental information processing (signal transduction) and cellular processes (cell motility) were also significantly differentially distributed among groups (Figure 5A).

To further reveal functional differences between groups, functional profiles between T10 and T15, NM and UM, and NM and OM were respectively compared. T10 was significantly enriched

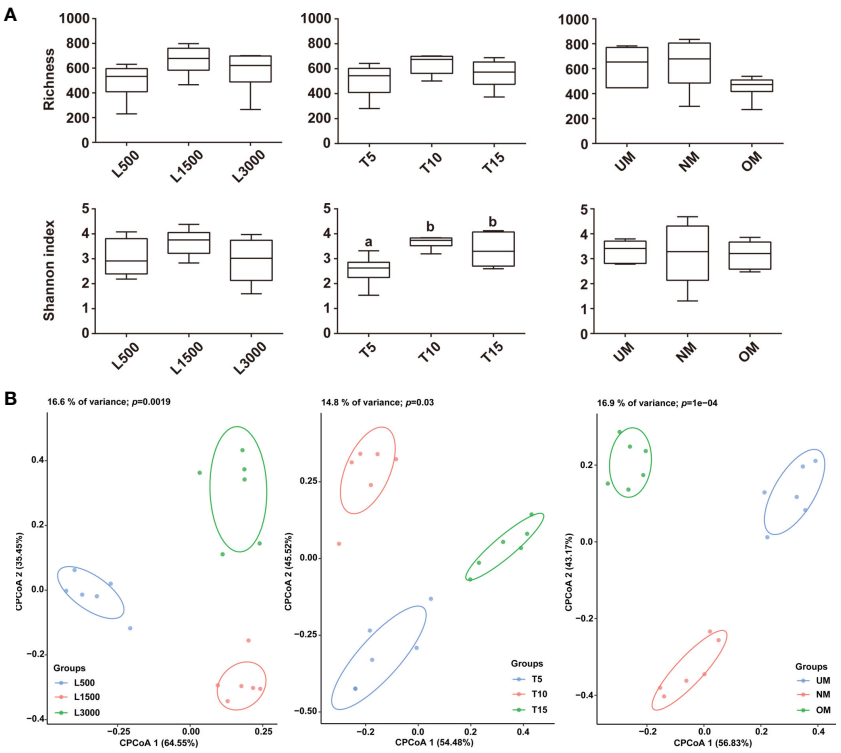


FIGURE 3
Comparisons of α **(A)** and β **(B)** diversities among the epiphytic bacterial communities associated with different treatment groups. The α-diversity was estimated using Observed species and Shannon index, and letters on the boxes indicate significant difference ($p < 0.05$, ANOVA, Tukey *post-hoc* test) among groups. Plots for the β-diversity were performed based on the Bray–Curtis dissimilarity.

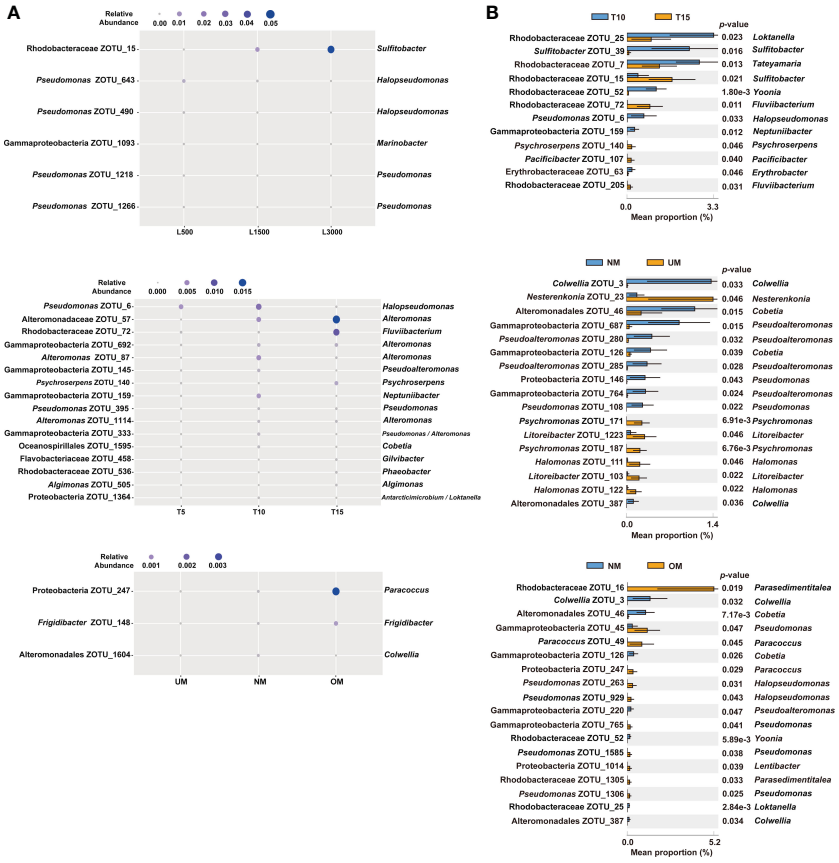


FIGURE 4
Significantly different epiphytic bacterial ZOTUs among treatment groups **(A)** and between groups with different MD severity **(B)**. For **(A)**, significant differences were determined using RF and ANOVA ($p < 0.05$, Tukey *post-hoc* test); for **(B)**, significant differences were determined using Welch's *t*-test (effect size >0.1, $p < 0.05$). Taxonomic information on the right of all the plots is determined using BLAST.

in categories related to development and regeneration compared with T15 (Figure 5B). Compared with UM, approximately six functional categories were significantly enriched in NM, including certain protein families (metabolism, genetic information processing, and signaling and cellular processes), environmental information processing (signal transduction), cellular process (cell motility), and drug resistance (Figure 5B). In addition, 10 pathways were found to be significantly depleted in NM compared with those of UM, which mainly included metabolism-related pathways (e.g., amino acid, carbohydrate, energy, lipid, and biosynthesis of other secondary metabolites), environmental information processing (membrane transport), bacterial infection, and cellular process (cell growth and death) (Figure 5B). For the comparison between the NM and OM groups, glycan biosynthesis and metabolism was the only category that showed a significant difference, which was more enriched in the NM group (Figure 5B).

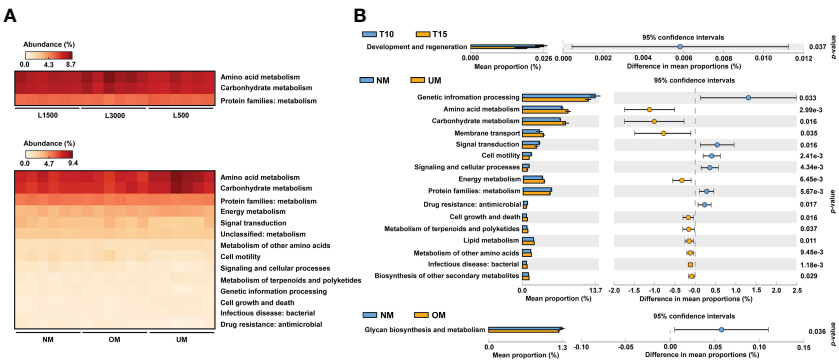


FIGURE 5
Comparisons of predicted functional profiles among treatment groups **(A)** and between groups with significantly different MD severity **(B)**. Significant differences were determined based on relative abundance using ANOVA ($p < 0.05$, Tukey *post-hoc* test) and Welch's *t*-test ($p < 0.05$), respectively.

Assembly process of epiphytic bacterial community

NST values based on both the Bray–Curtis distance and Jaccard dissimilarity revealed consistent trends among different treatment groups (Figure 6). Specifically, the NST values for all groups ranged from 16.79% to 35.06% and from 19.23% to 42.99% based on the Bray–Curtis distance (Figure 6A) and Jaccard dissimilarly (Figure 6B), respectively, indicating that the assembly of epiphytic bacterial communities on the seedlings was predominately governed by deterministic processes. For seedlings treated with different light intensities, the epiphytic bacterial community on seedlings cultivated under standard conditions (L1500) had a lower NST value (16.79% and 19.35% for the Bray–Curtis distance and Jaccard similarity, respectively) compared with those of the other two groups (L500: 35.06% and 38.30%; L3000: 31.41% and 42.99%, respectively), although no statistically significant difference was determined. Similar patterns were also observed for the groups treated with different light durations; notably, the T10 group had a significantly lower value (17.06%) than the T15 group (29.60%) based on the Bray–Curtis distance. For seedlings cultivated using different maturity level of parent kelp, NST values showed a slightly decreasing trend in accordance with the maturity level.

Correlation networks

Networks constructed for the groups under stressful conditions, notably those with significant MD severity, tended to harbor fewer nodes but a relatively higher average degree and graph density compared with those constructed under the standard cultivation conditions in each set (Table 3 and Figure S2). For example, the T15 network had 99 nodes, which was less than the number of nodes in the T10 network. The average degree and density of the T15 network were 2.222 and 0.023, respectively, which were also both higher than those of the T10 network. Similar trends were also determined for the comparisons between the NM and UM networks, and between the NM and OM networks, respectively (Table 3). The modularity indices for all networks were greater than 0.4, indicating that all networks harbored a modularized structure (Newman, 2006). Networks for groups with significantly higher MD severities had modularity indices lower than those of the groups cultivated under standard conditions (Table 3). Specifically, the modularity of the T15 network was 0.791, which was lower than the network of the T10 group (0.856) that also had a lower MD severity than the T15 group. Modularity indices of the UM (0.641) and OM (0.806) networks were also both lower than that of the NM group (0.818).

For all networks, ZOTUs with degree ≥ 5, closeness centrality > 0.15, and betweenness centrality < 200 were selected as the keystone taxa. Consequently, different profiles of keystone taxa were determined for all groups. The taxonomic information of these keystone taxa and their relative abundances are listed in Table S3. All keystone taxa belonged to the predominant groups identified in all networks: γ-Proteobacteria, α-Proteobacteria, and Actinobacteria. The keystone taxa of the T15 groups were determined to be *Halomonas* (ZOTU_81) and *Pseudoalteromonas* (ZOTU_1 and ZOTU_220), whereas those of the T10 group were identified as *Phaeobacter* (ZOTU_245), *Sulfitobacter* (ZOTU_462), and *Pseudomonas* (ZOTU_2). The two groups also shared three keystone taxa: *Nesterenkonia* (ZOTU_23), *Glycocaulis* (ZOTU_34), and *Aliihoeflea* (ZOTU_38). Both the UM and OM groups were associated with more complicated profiles of keystone taxa since only two genera (i.e., *Pseudomonas* and *Halopseudomonas*) were determined as keystone taxa in the NM group. ZOTUs belonging to *Nesterenkonia*, *Glycocaulis*, *Halomonas*, *Pseudoalteromonas*, *Pseudomonas*, *Loktanella*, and *Cobetia* were frequently determined keystone taxa in groups with significantly higher MD severity.

Discussion

In the present study, different MD severities of kelp seedlings were determined under various laboratory cultivation conditions. The epiphytic bacterial communities of the groups varying in MD severity were then analyzed and compared with respect to their

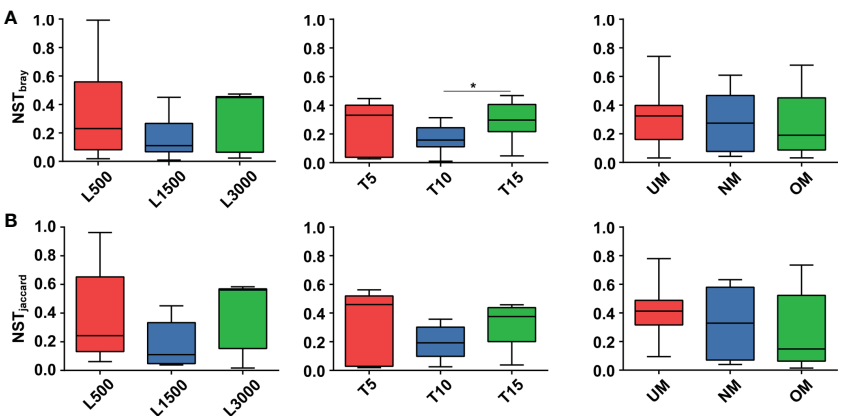


FIGURE 6
Comparisons of NST values among different treatment groups based on Bray-Curtis **(A)** and Jaccard **(B)** dissimilarity. Asterisk on the boxes indicate significant difference (*$p < 0.05$ Tukey *post-hoc* test) among groups.

TABLE 3 Comparison of the topological parameters for all networks.

	L500	L1500	L3000	T5	T10	T15	UM	NM	OM
Node	108	124	116	107	128	99	99	136	120
Edge	178	154	147	140	137	110	137	164	178
Positive (%)	63.48	76.62	69.39	73.57	74.45	75.45	66.42	80.49	79.21
Negative (%)	35.52	23.38	30.61	26.43	25.55	24.55	33.58	19.51	20.79
Average degree	3.296	2.484	2.534	2.617	2.141	2.222	2.768	2.412	2.967
Diameter	16	16	15	13	14	14	11	16	18
Density	0.031	0.02	0.022	0.025	0.017	0.023	0.028	0.018	0.025
Modularity	0.693	0.743	0.765	0.791	0.856	0.791	0.641	0.818	0.806
CC[a]	0.282	0.198	0.209	0.227	0.223	0.157	0.303	0.198	0.286
APL[b]	5.556	4.835	5.556	5.268	4.827	5.420	3.999	6.641	6.165

[a]Clustering coefficient; [b]Average path length.

structural and predicted functional profiles, in addition to the bacterial assemblage patterns.

Overall, the results demonstrated that the maturity level of parent kelp has a significant association with MD severity. Longer light exposure also showed a significant correlation with MD severity. These findings are consistent with previous reports proposing that either unfavorable light exposure or the maturity level of parent kelp is highly associated with the occurrence of MD (Qian et al., 2016; Sheng et al., 2016). However, scarce field investigations or experimental data have ever been provided to confirm such speculation. Exceeding the limits of environmental adaptation and tolerance is known to have negative impacts on the host state and defenses (Harvell et al., 2002; Moenne et al., 2016), while parent kelp of poor quality can also produce gametophytes and young sporophytes of poor quality (Qian et al., 2016; Sheng et al., 2016). The overall undermined physiology and performance of a host makes it potentially more susceptible to disruption of the associated microbiota, which can further facilitate pathogen infection, thereby providing good opportunities for disease occurrence. Multiple lines of evidence have demonstrated that bacteria are significant participants in macroalgal diseases, including rotting (Wang et al., 2008; Yan et al., 2022), bleaching (Zozaya-Valdés et al., 2017; Ward et al., 2021), gall formation (Ashen and Goff, 1996), and others. However, in many cases, a specific causal bacterial pathogen of the disease has not yet been conclusively identified, because diverse bacterial genera may be involved in the same disease or even induce similar symptoms as found in field observations. In such a context, it is necessary to investigate the pathogenesis at the level of the bacterial community. In the present study, significant community divergence was determined among different treatment groups, as confirmed by the CPCoA plots and perMANOVA for these groups. This indicates that seedlings of different health status had different bacterial communities.

To further reveal the nature of the observed divergence, the bacterial communities were compared between seedling groups with different MD severities. The predicted pathway changes (i.e., those involved in development and regeneration) enriched in T10 might indicate that the microbiota disruption caused by longer light exposure in T15 led to a lack of bacteria responsible for the normal development of seedlings. It has been demonstrated that disruption of associated microbiomes under stressful environments could cause abnormal phenotypes in macroalgae, e.g., malformations (Wichard, 2015; Ghaderiardakani et al., 2020). This is also a plausible mechanism related to MD given the characteristic of the disease in displaying abnormal morphology in seedling development. The enrichment of ZOTUs affiliated with the genus *Loktanella* (ZOTU_25) and *Sulfitobacter* (ZOTU_39) on the T10 seedlings is consistent with this observation, because both taxa have been reported to include essential members that stimulate the growth and/or morphology of macroalgal species (Spoerner et al., 2012; Dogs et al., 2017). As such, the differences detected between NM and UM further suggest the significant role played by bacteria in MD severity, including enriched signal transduction/signaling and cellular processes in NM. Signal-mediated quorum sensing allows bacteria to assemble as biofilms on seaweeds, and some bacteria (e.g., *Pseudoalteromonas*) are known to regulate zoospore settlement through cross-kingdom signaling (Marshall et al., 2006; Joint et al., 2007). Some bacteria are also well-equipped to have a positive impact on the morphology and growth of algae (Marshall et al., 2006), which is somewhat in line with our structural comparison at the ZOTU level. We found that ZOTUs belonging to *Pseudoalteromonas* (ZOTU_687, ZOTU_280, ZOTU_285, and ZOTU_764) and *Pseudomonas* (ZOTU_146 and ZOTU_108) were significantly enriched in NM. A significant proportion of the *Pseudomonas* members have been reported to induce morphological changes in green algal species (Nakanishi et al., 1996).

In the comparison between NM and OM, it is noted that the growth and/or morphology-related ZOTUs were significantly enriched in OM, such as ZOTU_45, ZOTU_765, ZOTU_1585, and ZOTU_1306 assigned to *Pseudomonas*, although only one predicted pathway (i.e., glycan biosynthesis and metabolism) was significantly enriched in NM. In addition, the OM-enriched taxon *Paracoccus* (ZOTU_49 and ZOTU_247) has also been reported to promote growth and morphogenesis in *Ulva* species (Ghaderiardakani et al., 2017). Collectively, these results suggest that the overgrowth of parent kelp might also be associated with an outcome of serious MD, potentially by influencing bacteria that impact host growth and

morphology. Based on these observations, we cannot preclude the possibility that certain bacterial members are capable of infecting seedlings as opportunistic pathogens, hence manipulating host cellular processes to cause malformation. The production of molecular mimics of algal hormones in marine bacteria have been proposed to facilitate the bypass of host defenses, thus facilitating invasion of the macroalgal tissue to interfere with host cell growth (Egan et al., 2014). We also found that certain morphology and/or growth-related ZOTUs were significantly enriched in seedlings with significantly higher MD severity, such as ZOTU_15 assigned as *Sulfitobacter* (Dogs et al., 2017) in T15, ZOTU_111 and ZOTU_122 assigned as *Halomonas* (Spoerner et al., 2012) in UM, and the aforementioned four ZOTUs assigned as *Pseudomonas* (Nakanishi et al., 1996) in OM. All of these taxa have been determined to constitute a significant proportion of the epiphytic bacterial community on diseased *S. japonica* (Wang et al., 2008; Zhang et al., 2020). Their capabilities of degrading algal polysaccharides (e.g., structural elements for cell walls and storage materials) and manipulating cellular processes by inducing development and growth have been well documented (Wang et al., 2008; Spoerner et al., 2012; Tapia et al., 2016; Dogs et al., 2017), which are conceived as significant virulence traits to be potential opportunistic pathogens for seaweeds (Egan et al., 2014). Additionally, the pathways related to bacterial infection and cell growth and death were significantly enriched in UM, in contrast to downregulation in the pathways of signaling and cellular processes. This indicates a functional transition from mild to more severe MD by disrupting the host's cell signaling and development. Taken together, our results suggest that disruption of the epiphytic microbiota caused by suboptimal growth and development conditions might provide a good opportunity for the invasion of opportunistic pathogens to interfere with the host's cellular processes, hence leading to more severe MD in the seedlings. However, more efforts are required to investigate how the epiphytic microbiome may interact with the host across stages of disease progression, which can provide further insight into MD pathogenesis.

In the present study, the assembly of all epiphytic bacterial communities was predominantly governed by deterministic processes. There is evidence that algal-associated bacterial communities differ from planktonic communities, and that seaweeds harbor species-specific and temporally adapted epiphytic bacterial communities on their surfaces (Lachnit et al., 2009; Lachnit et al., 2011; Burke et al., 2011b). Such differences between seaweed-associated communities imply selective mechanisms of assembly of the bacteria, further supporting a significant role of deterministic processes. We found variations in the assembly of bacterial communities on seedlings with significantly higher MD severity, particularly a decrease in determinism for the T15 and UM groups. This indicates that the relative contribution of the ecological processes that govern the establishment of epiphytic bacterial communities is linked to MD severity. Similar cases have also been reported for assembly of the gut microbiota in aquaculture animals during disease occurrence (Zhu et al., 2016; Xiong et al., 2017; Yao et al., 2018). Specifically, previous studies have suggested that the emergence of host disease reduces the relative importance of deterministic processes. One possible explanation is that the host's selective capability on planktonic bacteria is reduced under environmental stresses (e.g., longer light exposure in T15). Conversely, it is reported that both environmental stress and poor-quality parent kelp can lead to a decline in the health state of seedlings (Sheng et al., 2016), which may in turn lead to disruptions of the proper functions of the epiphytic bacterial community. Therefore, it is also possible that changes in the epiphytic bacterial assembly provide an open niche (a consequent lack of competition) for ambient bacterial species. Health-undermined seedlings are subject to stochastic colonization by dispersed microorganisms from the ambient water, including bacteria capable of causing infection, which is partly supported by the significant enrichment of the bacterial infection pathway in the UM group (Figure 5). Hence, this might support the possibility of opportunistic pathogen infection from the angle of community assembly. From a practical point, it is necessary to reduce such process variations to resist disease occurrence. Notably, the deterministic process became more important in OM compared with NM. This demonstrates that the ability of selection on microbiota increased as the hosts matured (Burns et al., 2016), further indicating an increase in the relative importance of deterministic processes. However, more detailed investigations are required to determine if overmatured parent kelp will have a direct influence on the assembly of the epiphytic bacterial community in their seedlings.

Bacterial interactions were also determined and their co-occurrence patterns were then compared between treatment groups with different MD severity. Groups with significantly higher MD severity (T15, UM, and OM) tended to contain fewer nodes but had higher graph density and degree, indicating more connected and complicated networks. Lower modularity was also determined for these networks, indicating less modularized structures. A module in the network is generally recognized as a group of taxa that are highly connected, which may represent similar niche preferences for these taxa (Ma et al., 2020). In such a context, lower modularization for the severe MD networks may indicate a relatively wide range of niche preferences for the colonized bacteria, thus reflecting disruptions of the epiphytic bacterial communities and increased complexity from the perspective of species–species interactions. Similar cases have also been reported for both shrimp and kelp diseases (Dai et al., 2020; Ling et al., 2022). In particular, the diseased shrimp also harbored a more highly connected and more complex bacterial network of gut microbiota (Dai et al., 2020). In addition, more diverse keystone taxa were determined for diseased shrimp compared with those identified in the healthy cohorts (Dai et al., 2020). This previous finding is to a large extent consistent with the more complicated profiles of keystone taxa detected in the present study for seedlings with significantly higher MD severity. Because keystone taxa play indispensable roles in the microbiome, and their removal could cause devastating changes in microbial composition and function (Banerjee et al., 2016; Banerjee et al., 2018; McMahon, 2018). This suggests that more diverse keystone taxa sustain and drive the changes in the epiphytic microbiota of seedlings with more severe MD. In this regard, multiple keystone taxa identified in groups with more severe MD, including *Halomonas*, *Pseudoalteromonas*, *Pseudomonas*, *Loktanella*, and *Cobetia*, have been reported to be capable of causing rotting, spotting, or degrading polysaccharides in macroalgae (Wang et al., 2008; Martin et al., 2015; Han, 2015), in addition to the aforementioned capabilities of inducing morphology and/or growth and development. Therefore, these taxa could be

significant contributors to disease severity. Additionally, ZOTUs related to these taxa were also determined to be significantly different in relative abundance between mild and more severe groups. A recent study indicated that the functions of a certain bacterial taxon, i.e., beneficial or harmful to the algal host, depend on the species or specific strains to a large extent (Li et al., 2022). Therefore, the significant roles of these keystone taxa in severe MD samples, notably specific bacterial strains, warrants deeper investigation in future studies to gain further insights into the mechanisms of disease occurrence.

Conclusions

MD severity was estimated for young *S. japonica* sporophytes under laboratory conditions. The findings illustrated that light exposure and the maturity level of parent kelp are significantly associated with MD severity. The divergence of the epiphytic bacterial communities and assembly patterns were also determined to be closely linked to MD severity. The present study enhances our understanding of the factors that significantly correlate with MD severity, and provides insights into the potential roles that epiphytic bacteria may play in the mechanism driving this association. Thus, it is proposed that external factors such as longer light exposure and parent kelp may be sufficient to cause a poor health status for young sporophytes, which then induces negative effects on their subsequent growth and development. Under such a stressed context, the structures and functions of the epimicrobiome were significantly disrupted compared with those of seedlings cultivated under standard conditions, which in turn significantly increased the MD severity through modifying host–bacteria interactions, such as lacking sufficient functions to fulfill host growth and development and/or inducing potential pathogens to manipulate cellular processes. However, more data (e.g., metagenomic and metatranscriptomic data) are still required to explore how the influencing factors, host cellular processes, and epiphytic microbiota interact with disease severity to determine the specific underlying causes of the disease. Besides, the determined keystone taxa, e.g., *Halomonas*, *Pseudoalteromonas*, *Pseudomonas*, *Loktanella*, and *Cobetia*, also require these fine-scale data in addition to their co-incubation with seedlings to provide deeper insights into their significant roles for the disease.

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

YY, ZM, and JL designed the experiments. SW, KL, and HY conducted experiments. YY performed the microbiome analysis. YY and JL prepared the manuscript. XR was involved in designing the experiments. All authors contributed to the article and approved the submitted version.

Funding

This work was financially supported by the China Agriculture Research System (CARS-50).

Acknowledgments

We would like to thank Weihai Changqing Ocean Science and Technology Co., Ltd. for their many help in the investigation of malformation disease, including providing samples for surveys and laboratories for the incubation in the present study.

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmars.2023.1089349/full#supplementary-material

References

Aires, T., Stuij, T. M., Muyzer, G., Serrão, E. A., and Engelen, A. H. (2021). Characterization and comparison of bacterial communities of an invasive and two native Caribbean seagrass species sheds light on the possible influence of the microbiome on invasive mechanisms. *Front. Microbiol.* 12, 653998. doi: 10.3389/fmicb.2021.653998

Alsufyani, T., Califano, G., Deicke, M., Grueneberg, J., Weiss, A., Engelen, A. H., et al. (2020). Macroalgal-bacterial interactions: identification and role of thallusin in morphogenesis of the seaweed *Ulva* (Chlorophyta). *J. Exp. Bot.* 71, 3340–3349. doi: 10.1093/jxb/eraa066

Ashen, J. B., and Goff, L. J. (1996). Molecular identification of a bacterium associated with gall formation in the marine red alga *Prionitis lanceolata*. *J. Phycol.* 32, 286–297. doi: 10.1111/j.0022-3646.1996.00286.x

Banerjee, S., Kirkby, C. A., Schmutter, D., Bissett, A., Kirkegaard, J. A., and Richardson, A. E. (2016). Network analysis reveals functional redundancy and keystone taxa amongst

bacterial and fungal communities during organic matter decomposition in an arable soil. *Soil Biol. Biochem.* 97, 188–198. doi: 10.1016/j.soilbio.2016.03.017

Banerjee, S., Schlaeppi, K., and van der Heijden, M. G. A. (2018). Keystone taxa as drivers of microbiome structure and functioning. *Nat. Rev. Microbiol.* 16, 567–576. doi: 10.1038/s41579-018-0024-1

Barberán, A., Bates, S. T., Casamayor, E. O., and Fierer, N. (2012). Using network analysis to explore co-occurrence patterns in soil microbial communities. *ISME J.* 6, 343–351. doi: 10.1038/ismej.2011.119

Bastian, M., Heymann, S., and Jacomy, M. (2009) *Gephi: an open source software for exploring and manipulating networks*. Available at: http://gephi.org.

Behera, D. P., Ingle, K. N., Mathew, D. E., Dhimmar, A., Sahastrabudhe, H., Sahu, S. K., et al. (2022). Epiphytism, diseases and grazing in seaweed aquaculture: a comprehensive review. *Rev. Aquacult.* 14, 1345–1370. doi: 10.1111/raq.12653

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

Breiman, L. (2001). Random forests. *Mach. Learn.* 45, 5–32. doi: 10.1023/A:1010933404324

Burke, C., Steinberg, P., Rusch, D., Kjelleberg, S., and Thomas, T. (2011a). Bacterial community assembly based on functional genes rather than species. *Proc. Natl. Acad. Sci. U.S.A.* 108, 14288–14293. doi: 10.1073/pnas.1101591108

Burke, C., Thomas, T., Lewis, M., Steinberg, P., and Kjelleberg, S. (2011b). Composition, uniqueness and variability of the epiphytic bacterial community of the green alga *Ulva australis*. *ISME J.* 5, 590–600. doi: 10.1038/ismej.2010.164

Burns, A. R., Stephens, W. Z., Stagaman, K., Wong, S., Rawls, J. F., Guillemin, K., et al. (2016). Contribution of neutral processes to the assembly of gut microbial communities in the zebrafish over host development. *ISME J.* 10, 655–664. doi: 10.1038/ismej.2015.142

Chambers, J. M., Freeny, A., and Heiberger, R. M. (1992). "Analysis of variance: designed experiments," in *Statistical models in s*. Eds. J. M. Chambers and T. J. Hastie (Pacific Grove, CA: Wadsworth & Brooks/Cole).

Chase, J. M., and Myers, J. A. (2011). Disentangling the importance of ecological niches from stochastic processes across scales. *Philos. Trans. R. Soc B Biol. Sci.* 366, 2351–2363. doi: 10.1098/rstb.2011.0063

Chen, T., Chen, X., Zhang, S., Zhu, J., Tang, B., Wang, A., et al. (2021). The genome sequence archive family: Toward explosive data growth and diverse data types. *Genom. Proteom. Bioinf.* 19, 578–583. doi: 10.1016/j.gpb.2021.08.001

Chen, T., Liu, Y., and Huang, L. (2022). ImageGP: an easy-to-use data visualization web server for scientific researchers. *iMeta* 1, e5. doi: 10.1002/imt2.5

Dai, W., Chen, J., and Xiong, J. (2019). Concept of microbial gatekeepers: positive guys? *Appl. Microbiol. Biotechnol.* 103, 633–641. doi: 10.1007/s00253-018-9522-3

Dai, W., Sheng, Z., Chen, J., and Xiong, J. (2020). Shrimp disease progression increases the gut bacterial network complexity and abundances of keystone taxa. *Aquaculture* 517, 734802. doi: 10.1016/j.aquaculture.2019.734802

Dai, W., Yu, W., Xuan, L., Tao, Z., and Xiong, J. (2018). Integrating molecular and ecological approaches to identify potential polymicrobial pathogens over a shrimp disease progression. *Appl. Microbiol. Biotechnol.* 102, 3755–3764. doi: 10.1007/s00253-018-8891-y

De Schryver, P., Defoirdt, T., Boon, N., Verstraete, W., and Bossier, P. (2012). "Managing the microbiota in aquaculture systems for disease prevention and control," in *Infectious disease in aquaculture: Prevention and control*. Ed. B. Austin (Sawston: Woodhead Publishing), 394–418. doi: 10.1533/9780857095732.3.394

Dittami, S. M., Arboleda, E., Auguet, J.-C., Bigalke, A., Briand, E., Cárdenas, P., et al. (2021). A community perspective on the concept of marine holobionts: current status, challenges, and future directions. *PeerJ* 9, e10911. doi: 10.7717/peerj.10911

Dittami, S. M., Duboscq-Bidot, L., Perennou, M., Gobet, A., Corre, E., Boyen, C., et al. (2016). Host-microbe interactions as a driver of acclimation to salinity gradients in brown algal cultures. *ISME J.* 10, 51–63. doi: 10.1038/ismej.2015.104

Dogs, M., Wemheuer, B., Wolter, L., Bergen, N., Daniel, R., Simon, M., et al. (2017). Rhodobacteraceae on the marine brown alga *Fucus spiralis* are abundant and show physiological adaptation to an epiphytic lifestyle. *Syst. Appl. Microbiol.* 40, 370–382. doi: 10.1016/j.syapm.2017.05.006

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

Duan, D., Miao, G., and Wang, X. (2015). *Aquacultural biology of saccharina japonica* (Beijing: Science Press).

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

Egan, S., Fernandes, N. D., Kumar, V., Gardiner, M., and Thomas, T. (2014). Bacterial pathogens, virulence mechanism and host defence in marine macroalgae. *Environ. Microbiol.* 16, 925–938. doi: 10.1111/1462-2920.12288

Egan, S., and Gardiner, M. (2016). Microbial dysbiosis: rethinking disease in marine ecosystems. *Front. Microbiol.* 7, 991. doi: 10.3389/fmicb.2016.00991

Florez, J. Z., Camus, C., Hengst, M. B., Marchant, F., and Buschmann, A. H. (2019). Structure of the epiphytic bacterial communities of *Macrocystis pyrifera* in localities with contrasting nitrogen concentrations and temperature. *Algal Res.* 44, 101706. doi: 10.1016/j.algal.2019.101706

Gachon, C. M. M., Sime-Ngando, T., Strittmatter, M., Chambouvet, A., and Kim, G. H. (2010). Algal diseases: spotlight on a black box. *Trends Plant Sci.* 15, 633–640. doi: 10.1016/j.tplants.2010.08.005

Ghaderiardakani, F., Coates, J. C., and Wichard, T. (2017). Bacteria-induced morphogenesis of *Ulva intestinalis* and *Ulva mutabilis* (Chlorophyta): a contribution to the lottery theory. *FEMS Microbiol. Ecol.* 93, fix094. doi: 10.1093/femsec/fix094

Ghaderiardakani, F., Quartino, M. L., and Wichard, T. (2020). Microbiome-dependent adaptation of seaweeds under environmental stresses: A perspective. *Front. Mar. Sci.* 7, 575228. doi: 10.3389/fmars.2020.575228

Goecke, F., Labes, A., Wiese, J., and Imhoff, J. F. (2010). Chemical interactions between marine macroalgae and bacteria. *Mar. Ecol. Prog. Ser.* 409, 267–299. doi: 10.3354/meps08607

Han, X. (2015). Identification and complete genome sequencing of a cobetia sp. causing green rot disease in porphyra haitanensis (Qingdao: Ocean University of China).

Harvell, C. D., Mitchell, C. E., Ward, J. R., Altizer, S., Dobson, A. P., Ostfeld, R. S., et al. (2002). Climate warming and disease risks for terrestrial and marine biota. *Science* 296, 2158–2162. doi: 10.1126/science.1063699

Hasselström, L., Visch, W., Gröndahl, F., Nylund, G. M., and Pavia, H. (2018). The impact of seaweed cultivation on ecosystem services - a case study from the west coast of Sweden. *Mar. pollut. Bull.* 133, 53–64. doi: 10.1016/j.marpolbul.2018.05.005

Herren, C. M., and McMahon, K. D. (2018). Keystone taxa predict compositional change in microbial communities. *Environ. Microbiol.* 20, 2207–2217. doi: 10.1111/1462-2920.14257

Joint, I., Tait, K., and Wheeler, G. (2007). Cross-kingdom signalling: exploitation of bacterial quorum sensing molecules by the green seaweed *Ulva*. *Philos. Trans. R. Soc B Biol. Sci.* 362, 1223–1233. doi: 10.1098/rstb.2007.2047

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

Lachnit, T., Blümel, M., Imhoff, J. F., and Wahl, M. (2009). Specific epibacterial communities on macroalgae: phylogeny matters more than habitat. *Aquat. Biol.* 5, 181–186. doi: 10.3354/ab00149

Lachnit, T., Meske, D., Wahl, M., Harder, T., and Schmitz, R. (2011). Epibacterial community patterns on marine macroalgae are host-specific but temporally variable. *Environ. Microbiol.* 13, 655–665. doi: 10.1111/j.1462-2920.2010.02371.x

Lemay, M. A., Martone, P. T., Hind, K. R., Lindstrom, S. C., and Wegener Parfrey, L. (2018). Alternate life history phases of a common seaweed have distinct microbial surface communities. *Mol. Ecol.* 27, 3555–3568. doi: 10.1111/mec.14815

Li, J., Majzoub, M. E., Marzinelli, E. M., Dai, Z., Thomas, T., and Egan, S. (2022). Bacterial controlled mitigation of dysbiosis in a seaweed disease. *ISME J.* 16, 378–387. doi: 10.1038/s41396-021-01070-1

Ling, F., Egan, S., Zhuang, Y., Chang, L., Xiao, L., Yang, Q., et al. (2022). Epimicrobiome shifts with bleaching disease progression in the brown seaweed *Saccharina japonica*. *Front. Mar. Sci.* 9, 865224. doi: 10.3389/fmars.2022.865224

Lin, W., Lu, J., Yao, H., Lu, Z., He, Y., Mu, C., et al. (2021). Elevated pCO_2 alters the interaction patterns and functional potentials of rearing seawater microbiota. *Environ. pollut.* 287, 117615. doi: 10.1016/j.envpol.2021.117615

Li, J., Pang, S., Shan, T., and Su, L. (2020). Changes of microbial community structures associated with seedlings of *Saccharina japonica* at early stage of outbreak of green rotten disease. *J. Appl. Phycol.* 62, 191. doi: 10.1007/s10811-019-01975-7

Lomartire, S., Marques, J. C., and Gonçalves, A. M. M. (2021). An overview to the health benefits of seaweeds consumption. *Mar. Drugs* 19, 341. doi: 10.3390/md19060341

Longford, S. R., Campbell, A. H., Nielsen, S., Case, R. J., Kjelleberg, S., and Steinberg, P. D. (2019). Interactions within the microbiome alter microbial interactions with host chemical defences and affect disease in a marine holobiont. *Sci. Rep.* 9, 1363. doi: 10.1038/s41598-018-37062-z

Marshall, K., Joint, I., Callow, M. E., and Callow, J. A. (2006). Effect of marine bacterial isolates on the growth and morphology of axenic plantlets of the green alga *Ulva linza*. *Microb. Ecol.* 52, 302–310. doi: 10.1007/s00248-006-9060-x

Martin, M., Barbeyron, T., Martin, R., Portetelle, D., Michel, G., and Vandenbol, M. (2015). The cultivable surface microbiota of the brown alga *Ascophyllum nodosum* is enriched in macroalgal-polysaccharide-degrading bacteria. *Front. Microbiol.* 6, 1487. doi: 10.3389/fmicb.2015.01487

Ma, B., Wang, H., Dsouza, M., Lou, J., He, Y., Dai, Z., et al. (2016). Geographic patterns of co-occurrence network topological features for soil microbiota at continental scale in eastern China. *ISME J.* 10, 1891–1901. doi: 10.1038/ismej.2015.261

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

Moenne, A., González, A., and Sáez, C. A. (2016). Mechanisms of metal tolerance in marine macroalgae, with emphasis on copper tolerance in chlorophyta and rhodophyta. *Aquat. Toxicol.* 176, 30–37. doi: 10.1016/j.aquatox.2016.04.015

Morrissey, K. L., Iveša, L., Delva, S., D'Hondt, S., Willems, A., and De Clerck, O. (2021). Impacts of environmental stress on resistance and resilience of algal-associated bacterial communities. *Ecol. Evol.* 11, 15004–15019. doi: 10.1002/ece3.8184

Nakanishi, K., Nishijima, M., Nishimura, M., Kuwano, K., and Saga, N. (1996). Bacteria that induce morphogenesis in *Ulva pertusa* (Chlorophyta) grown under axenic conditions. *J. Phycol.* 32, 479–482. doi: 10.1111/j.0022-3646.1996.00479.x

Newman, M. E. J. (2002). The structure and function of networks. *Comput. Phys. Commun.* 147, 40–45. doi: 10.1016/S0010-4655(02)00201-1

Newman, M. E. J. (2006). Modularity and community structure in networks. *Proc. Natl. Acad. Sci. U.S.A.* 103, 8577–8582. doi: 10.1073/pnas.0601602103

Ning, D., Deng, Y., Tiedje, J. M., and Zhou, J. (2019). A general framework for quantitatively assessing ecological stochasticity. *Proc. Natl. Acad. Sci. U.S.A.* 116, 16892–16898. doi: 10.1073/pnas.1904623116

Oksanen, J., Simpson, G., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P., et al. (2019) *Vegan: community ecology package. package version 2.5-6*. Available at: https://cran.r-project.org/package=vegan.

Parks, D. H., Tyson, G. W., Hugenholtz, P., and Beiko, R. G. (2014). STAMP: statistical analysis of taxonomic and functional profiles. *Bioinformatics* 30, 3123–3124. doi: 10.1093/bioinformatics/btu494

Pearson, D. E., Ortega, Y. K., Eren, Ö., and Hierro, J. L. (2018). Community assembly theory as a framework for biological invasions. *Trends Ecol. Evol.* 33, 313–325. doi: 10.1016/j.tree.2018.03.002

Peng, Y., and Li, W. (2013). A bacterial pathogen infecting gametophytes of saccharina japonica (Laminariales, phaeophyceae). *Chin. J. Oceanol. Limnol.* 31, 366–373. doi: 10.1007/s00343-013-2136-9

Qian, R., Zhang, Z., Li, X., Pan, J., Sheng, B., and Jiang, Y. (2016). Disease occurrence and controlling methods in nursery-cultivated *Saccharina japonica*. *China Fish.* 6, 96–99.

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

Saha, M., and Weinberger, F. (2019). Microbial "gardening" by a seaweed holobiont: surface metabolites attract protective and deter pathogenic epibacterial settlement. *J. Ecol.* 107, 2255–2265. doi: 10.1111/1365-2745.13193

Sheng, B., Wang, Q., Zhang, Z., Pan, J., and Qu, S. (2016). Prevention and treatment of sporeling malformation disease in the production of *Saccharina japonica*. *Hebei Fish.* 10, 45–46. doi: 10.3969/j.issn.1004–6755.2016.10.014

Singh, R. P., and Reddy, C. R. K. (2015). Unraveling the functions of the macroalgal microbiome. *Front. Microbiol.* 6, 1488. doi: 10.3389/fmicb.2015.01488

Spoerner, M., Wichard, T., Bachhuber, T., Stratmann, J., and Oertel, W. (2012). Growth and thallus morphogenesis of *Ulva mutabilis* (Chlorophyta) depends on a combination of two bacterial species excreting regulatory factors. *J. Phycol.* 48, 1433–1447. doi: 10.1111/j.1529-8817.2012.01231.x

Tapia, J. E., González, B., Goulitquer, S., Potin, P., and Correa, J. A. (2016). Microbiota influences morphology and reproduction of the brown alga *Ectocarpus* sp. *Front. Microbiol.* 7, 197. doi: 10.3389/fmicb.2016.00197

Tseng, C. K. (2001). Algal biotechnology industries and research activities in China. *J. Appl. Phycol.* 13, 375–380. doi: 10.1023/A:1017972812576

Wahl, M., Goecke, F., Labes, A., Dobretsov, S., and Weinberger, F. (2012). The second skin: ecological role of epibiotic biofilms on marine organisms. *Front. Microbiol.* 3, 292. doi: 10.3389/fmicb.2012.00292

Wang, G., Lu, B., Shuai, L., Li, D., and Zhang, R. (2014). Microbial diseases of nursery and field-cultivated *Saccharina japonica* (Phaeophyta) in China. *Algol. Stud.* 145, 39–51. doi: 10.1127/1864-1318/2014/0167

Wang, G., Shuai, L., Li, Y., Lin, W., Zhao, X., and Duan, D. (2008). Phylogenetic analysis of epiphytic marine bacteria on hole-rotten diseased sporophytes of *Laminaria japonica*. *J. Appl. Phycol.* 20, 403–409. doi: 10.1007/s10811-007-9274-4

Ward, G. M., Kambey, C. S. B., Faisan, J. P., Tan, P., Daumich, C. C., Matoju, I., et al. (2021). Ice-ice disease: an environmentally and microbiologically driven syndrome in tropical seaweed aquaculture. *Rev. Aquacult.* 47, 361. doi: 10.1111/raq.12606

Watt, C. A., and Scrosati, R. A. (2014). Seaweeds as ecosystem engineers. *Bull. Ecol. Soc Am.* 95, 250–251. doi: 10.1890/0012-9623-95.3.250

Wichard, T. (2015). Exploring bacteria-induced growth and morphogenesis in the green macroalga order ulvales (Chlorophyta). *Front. Plant Sci.* 6, 86. doi: 10.3389/fpls.2015.00086

Wu, C., Gao, N., Che, D., Zhou, B., Cai, P., Dong, S., et al. (1979). On the malformation disease of *Laminaria* sporelings. *Oceanol. Et Limnol. Sin.* 10, 238–247.

Xiong, J. (2018). Progress in the gut microbiota in exploring shrimp disease pathogenesis and incidence. *Appl. Microbiol. Biotechnol.* 102, 7343–7350. doi: 10.1007/s00253-018-9199-7

Xiong, J., Zhu, J., Dai, W., Dong, C., Qiu, Q., and Li, C. (2017). Integrating gut microbiota immaturity and disease-discriminatory taxa to diagnose the initiation and severity of shrimp disease. *Environ. Microbiol.* 19, 1490–1501. doi: 10.1111/1462-2920.13701

Xue, Y., Bao, Y., Zhang, Z., Zhao, W., Xiao, J., He, S., et al. (2022). Database resources of the national genomics data center, China national center for bioinformation in 2022. *Nucleic Acids Res.* 50, D27–D38. doi: 10.1093/nar/gkab951

Yan, Y., Wang, S., Li, J., Liu, F., and Mo, Z. (2022). Alterations in epiphytic bacterial communities during the occurrence of green rot disease in *Saccharina japonica* seedlings. *J. Mar. Sci. Eng.* 10, 730. doi: 10.3390/jmse10060730

Yao, Z., Yang, K., Huang, L., Huang, X., Qiuqian, L., Wang, K., et al. (2018). Disease outbreak accompanies the dispersive structure of shrimp gut bacterial community with a simple core microbiota. *AMB Express* 8, 120. doi: 10.1186/s13568-018-0644-x

Zhang, R., Chang, L., Xiao, L., Zhang, X., Han, Q., Li, N., et al. (2020). Diversity of the epiphytic bacterial communities associated with commercially cultivated healthy and diseased *Saccharina japonica* during the harvest season. *J. Appl. Phycol.* 32, 2071–2080. doi: 10.1007/s10811-019-02025-y

Zhang, J., Liu, Y.-X., Zhang, N., Hu, B., Jin, T., Xu, H., et al. (2019). NRT1.1B is associated with root microbiota composition and nitrogen use in field-grown rice. *Nat. Biotechnol.* 37, 676–684. doi: 10.1038/s41587-019-0104-4

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

Zhu, J., Dai, W., Qiu, Q., Dong, C., Zhang, J., and Xiong, J. (2016). Contrasting ecological processes and functional compositions between intestinal bacterial community in healthy and diseased shrimp. *Microbiol. Ecol.* 72, 975–985. doi: 10.1007/s00248-016-0831-8

Zozaya-Valdés, E., Egan, S., and Thomas, T. (2015). A comprehensive analysis of the microbial communities of healthy and diseased marine macroalgae and the detection of known and potential bacterial pathogens. *Front. Microbiol.* 6, 146. doi: 10.3389/fmicb.2015.00146

Zozaya-Valdés, E., Roth Schulze, A. J., Egan, S., and Thomas, T. (2017). Microbial community function in the bleaching disease of the marine macroalgae *Delisea pulchra*. *Environ. Microbiol.* 19, 3012–3024. doi: 10.1111/1462-2920.13758

frontiers | Frontiers in Marine Science

TYPE Methods
PUBLISHED 30 January 2023
DOI 10.3389/fmars.2023.1087447

OPEN ACCESS

EDITED BY
Neil Ross McEwan,
Robert Gordon University, United Kingdom

REVIEWED BY
Matthew Hegarty,
Aberystwyth University, United Kingdom
Shuyuan Wang,
Harbin Medical University, China

*CORRESPONDENCE
Guangyi Fan
fanguangyi@genomics.cn
Xin Liu
liuxin@genomics.cn

†These authors have contributed equally to this work.

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

RECEIVED 02 November 2022
ACCEPTED 16 January 2023
PUBLISHED 30 January 2023

CITATION
Xu M, Guo L, Qi Y, Shi C, Liu X, Chen J, Han J, Deng L, Liu X and Fan G (2023) Symbiont-screener: A reference-free tool to separate host sequences from symbionts for error-prone long reads. *Front. Mar. Sci.* 10:1087447. doi: 10.3389/fmars.2023.1087447

Symbiont-screener: A reference-free tool to separate host sequences from symbionts for error-prone long reads

Mengyang Xu[1,2†], Lidong Guo[1,3†], Yanwei Qi[1†], Chengcheng Shi[1], Xiaochuan Liu[1], Jianwei Chen[1], Jinglin Han[1], Li Deng[1], Xin Liu[1,2,4*] and Guangyi Fan[1,2,4*]

[1]BGI-Qingdao, BGI-Shenzhen, Qingdao, China, [2]BGI-Shenzhen, Shenzhen, China, [3]College of Life Sciences, University of Chinese Academy of Sciences, Beijing, China, [4]State Key Laboratory of Agricultural Genomics, BGI-Shenzhen, Shenzhen, China

Metagenomic sequencing facilitates large-scale constitutional analysis and functional characterization of complex microbial communities without cultivation. Recent advances in long-read sequencing techniques utilize long-range information to simplify repeat-aware metagenomic assembly puzzles and complex genome binning tasks. However, it remains methodologically challenging to remove host-derived DNA sequences from the microbial community at the read resolution due to high sequencing error rates and the absence of reference genomes. We here present Symbiont-Screener (https://github.com/BGI-Qingdao/Symbiont-Screener), a reference-free approach to identifying high-confidence host's long reads from symbionts and contaminants and overcoming the low sequencing accuracy according to a trio-based screening model. The remaining host's sequences are then automatically grouped by unsupervised clustering. When applied to both simulated and real long-read datasets, it maintains higher precision and recall rates of identifying the host's raw reads compared to other tools and hence promises the high-quality reconstruction of the host genome and associated metagenomes. Furthermore, we leveraged both PacBio HiFi and nanopore long reads to separate the host's sequences on a real host-microbe system, an algal-bacterial sample, and retrieved an obvious improvement of host assembly in terms of assembly contiguity, completeness, and purity. More importantly, the residual symbiotic microbiomes illustrate improved genomic profiling and assemblies after the screening, which elucidates a solid basis of data for downstream bioinformatic analyses, thus providing a novel perspective on symbiotic research.

KEYWORDS

symbiosis, decontamination, metagenomic sequencing, long reads, bioinformatics, alignment-free, reference-free, *de novo* assembly

1 Introduction

Powered by advanced biotechnologies and big data analytics, genome sequences become the significant biological basis and genomic resources of modern life science. Simultaneous genome sequencing of the host-symbiont ecosystem reveals the bioinformatic information of both the host species and associated microbial communities, thus prompting the symbiotic studies to move from gene-centric to genome-centric fields (Xie et al., 2016; Xie et al., 2020).

However, valuable insights into the construction and dynamics of the symbiotic ecosystem require successful separation of the host, symbiont, and contaminant data (Cornet and Baurain, 2022). Complicated and laborious experimental approaches have been developed to isolate host sequences from prokaryotic contamination (Arimoto et al., 2019; Cheng et al., 2019; Wang et al., 2020). Current bioinformatic methods rely on the species differentiation in statistical features (Woyke et al., 2006; Alneberg et al., 2014), or nucleotide and protein similarity of pre-assemblies to known genomes or public databases (Coghlan et al., 2019). Unfortunately, most methods designed for next-generation sequencing (NGS) short reads provide incomplete or inaccurate information and are facing challenges such as strong dependence on public data libraries or pre-assembly quality. It has been demonstrated that public genomic data may contain foreign sequences, leading to erroneous genetic characteristics (Neimark, 2015; Steinegger and Salzberg, 2020; Douvlataniotis et al., 2020). Moreover, those sequences that cannot be aligned to the reference genomes usually provide more critical findings, for instance, the identification of novel COVID-19 variants (Ricker et al., 2012; Kim et al., 2020; Cheng et al., 2022).

Short-read assemblies cannot resolve the highly nonuniform coverage of the composing species and the presence of long intra-genomic and inter-genomic repeats (Kolmogorov et al., 2020), resulting in uncompleted draft genomes with gaps (Fraser et al., 2002; Xu et al., 2020). Third-generation sequencing (TGS) long reads, including Pacific Biosciences (PacBio) and Oxford Nanopore Technologies (ONT), provide unique long-range information, which straightforwardly simplifies complicated biomedical problems (Nagarajan and Pop, 2013; Rhoads and Au, 2015; Qi et al., 2022). The reconstruction of the complete genome sequence of hosts and microbes enables the analysis of the symbiotic relationship and microbial diversity, including the detection of horizontal gene transfer of mobile elements, large-scale structural rearrangements, and search for biosynthetic gene clusters (Chin et al., 2013; Bertrand et al., 2019). Long-read metagenomic classifiers such as Centrifuge (Kim et al., 2016), Kraken2 (Wood et al., 2019), and MetaMaps (Dilthey et al., 2019) can build indexed databases according to acknowledged reference genomes and NCBI taxonomy, and assign corresponding long reads to the host. Meanwhile, MetaProb (Girotto et al., 2016), BusyBee (Laczny et al., 2017) and MetaBCC-LR (Wickramarachchi et al., 2020) do not require references to classify long reads based on the unsupervised clustering results of *k*-mer coverage or oligonucleotide composition, but cannot indicate which cluster belongs to the host. But the relatively high sequencing error rates of TGS data might be greater than the genetic difference between organisms, resulting in a low capture ratio of the host's data and large computational consumption (Bharti and Grimm, 2019). It becomes almost impossible to classify highly similar sequences shared by both the host and symbionts, for instance, symbiotic algae in a floating island of seaweeds (Thiel and Gutow, 2005; Rothäusler et al., 2012). Besides, it is even more complex for the *de novo* projects without sufficient priori knowledge, that is, lack of reference genomes or libraries.

The combination of TGS's unprecedented read lengths and the trio's global inherited information has been demonstrated to improve the genome assembly and further reconstruct haplotypes (Koren et al., 2018; Ebert et al., 2021; Xu et al., 2021). This idea also motivates us to introduce a reference-free and alignment-free way to solve the screening problem prior to assembly. Laboratory cultivation of sexually reproducing diploid host with associated microorganisms enables us to gather the parent-offspring pedigree information without loss of symbiotic microbial information. In this work, we established a novel screening model of TGS raw reads according to the transmissibility of heterozygous variants in the trio of host species, the stability of symbiotic relations, and the randomness of contaminant sources. Based on this model, Symbiont-Screener selects high-confidence host's reads. Then it captures more host sequences by an unsupervised clustering algorithm. The final data, in which most of the foreign genomes have been screened out, can recover a high-quality host genome. On the other hand, the residual microbial long reads can enhance the variation profiling, metagenomic assemblies, and taxonomic binning.

2 Methods

2.1 Trio-based screening model

The design of Symbiont-Screener focuses on the stepwise purification of the host's sequences with sufficient precision and recall rates to reconstruct the genome by combining the advantage of long read lengths with trio-binning markers and minimizing the effect of high error rates (Figure 1). The mixed sample possibly comprises steady symbionts and random DNA contaminants induced by laboratory pollution or artificial experimental errors other than the host genome. Among them, symbionts sharing highly similar sequences with the host are the most difficult to be isolated. According to the species sources and the relation of parent-offspring trios, the possible foreign genomes in the offspring's data can be categorized into four types: the offspring-only (OC), perhaps random contaminants; shared by father and offspring (POC); shared by mother and offspring (MOC); and shared by all three (SC), perhaps steady symbionts. Theoretically, the trio-specific markers inherited from contaminated parents allow the identification and reconstruction of the host chromosomes, and meanwhile, SC can be filtered out since their markers will no longer occur in the parent-specific marker

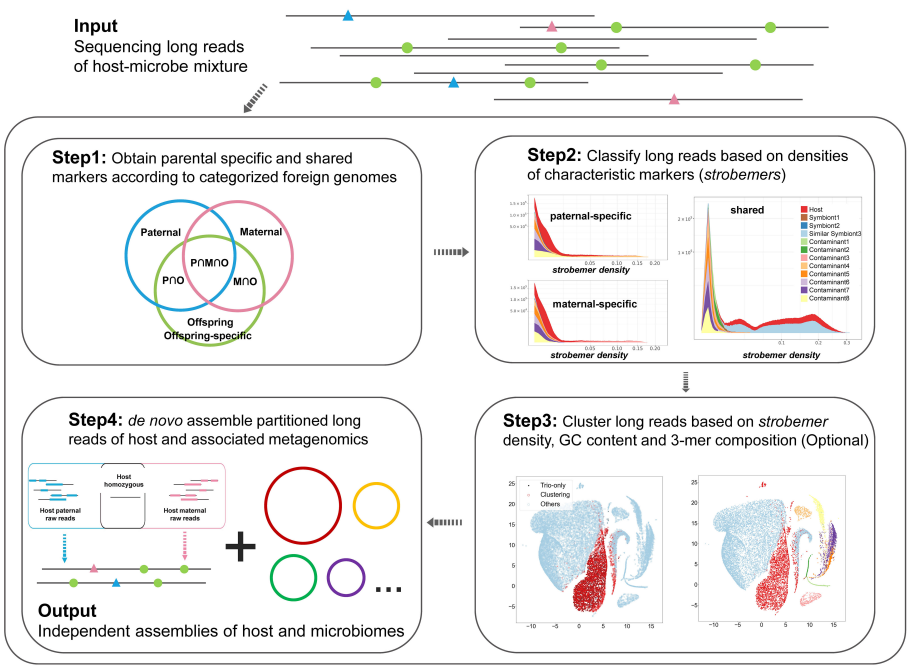


FIGURE 1
Workflow of Symbiont-Screener. The input sequencing long reads of the host-microbe mixture can be categorized based on whether they contain characteristic markers or not. The tool starts by calculating and matching the parent-specific and shared *strobemers*. The second step is to classify long reads according to the species differentiation of characteristic marker densities, and detect the high-confidence host's data, which satisfies the relatively more accurate PacBio reads (single-base accuracy ≥95%). Then, for ONT long reads with higher error rates (single-base accuracy<95%), all the long reads including those without any matched markers are clustered by the features of characteristic *strobemer* density, GC content, and trinucleotide composition. Next, the clusters which belong to the host are detected by high-confidence long reads preselected in Step2, and other remaining clusters are labeled as metagenomic long reads. The final output includes a high-quality haplotype-collapsed or two haplotype-resolved assemblies for the diploid host as well as complete metagenome-assembled assemblies for the host-associated microbiomes.

libraries after set operations. The set difference, however, cannot remove parent-specific foreign genomes, POC or MOC. On the other hand, the intersection of marker libraries for all individuals ascertains the host autosomes alongside with SC. Nevertheless, the foreign species shared only by one parent (POC or MOC) will be discarded. At last, none of the markers will be enriched in the sequences of OC. Parental samples of the trio are required to be collected and sequenced to provide paternal- and maternal-specific markers. In principle, the set operations of characteristic marker libraries can eliminate most of the foreign species if the host heterozygosity, read length, and read sing-base accuracy meet certain thresholds.

The read length of current PacBio or ONT data meets the requirement of this model, although the high sequencing error rate remains the greatest obstacle. Theoretically, the expected number of characteristic *k*-mers (discriminability) in a single read is proportional to the read length, heterozygosity ratio, and *k*th power of single-base sequencing accuracy. In addition to the use of error-tolerate *strobemers* (Sahlin, 2021), we identify raw long-read data based on the genomic feature, which are further clustered with the unsupervised Bayesian Gaussian Mixture model (BGMM) method after principal component analysis (PCA) dimensionality reduction. In practice, the parent-specific and shared *strobemers*, the species differentiation in GC content and oligonucleotide frequencies consist of the main features in the algorithm. The procedure of genome binning is equivalent to the read clustering based on the Mahalanobis distances to centers in the 36-dimensional feature space. Moreover, the characteristic markers can be used to identify which read cluster belongs to the host.

2.2 Generation of characteristic markers

According to the above model, characteristic markers existing in the paternal group other than the maternal are defined as paternal-only makers, while those only existing in the maternal group are defined as maternal-only. Meanwhile, markers shared by both parents are defined as shared. We removed markers in low- and high-frequency regions and then ran set operations to calculate characteristic markers (Supplementary Figure 1). Supplementary Figure 2 shows the marker category of reference assemblies for the host, symbionts, and contaminants, which reflects the feasibility of this screening model in identifying different types of foreign species.

Plots of normalized densities of parent-specific and shared markers demonstrate that only host long reads synchronously own abundant parent-specific and shared markers (Supplementary Figure 3). Thus, the high-confidence host reads can be determined by the following formula

$$(x_{1,\,i} - C_1 > 0 \;\|\; x_{2,i} - C_1 > 0) \&\& \; x_{3,i} - C_2 > 0$$

where $x_{1,i}$and $x_{2,i}$ refer to the parent-only maker densities in the ith long read, $x_{3,i}$ refers to the shared maker density, while C_1 and C_2 refer to different thresholds.

```
Input Long read LR, marker set POK, MOK, and SK
Output the source type of read LR
flag_pok = KmerLookup ( LR , POK )
flag_mok = KmerLookup ( LR , MOK )
flag_sk = KmerLookup ( LR , SK )
if ( ( flag_pok || flag_mok ) && flag_sk )
return Host
else if ( flag_pok && ! ( flag_mok && flag_sk ) )
return POC
else if ( flag_mok && ! ( flag_pok && flag_sk ) )
return MOC
else if ( ! ( flag_pok && flag_mok && flag_sk ) )
return OC
else if ( ! ( flag_pok || flag_mok ) && flag_sk ) )
return SC
end
```

ALGORITHM 1. READSOURCETYPE

2.3 strobemer vs. k-mer

Characteristic markers can be *k*-mers as used in conventional trio-binning and genome-binning approaches (Koren et al., 2018; Ebert et al., 2021; Xu et al., 2021), or error-tolerant *strobemers* (Sahlin, 2021). The utilization of *k*-mers can statistically capture parent-specific markers in raw data as long as the reads are sufficiently long and the host homologous chromosomes have enough heterozygous sites (Koren et al., 2018). However, the captured ratio is still severely limited by the sequencing errors for such a complex application of screening. Therefore, we chose *strobemer* implementation for the error-tolerant indexing and matching to provide more evenly distributed matches with higher genome coverage and less sensitive to sequencing errors, especially for insertions and deletions (Sahlin, 2021). In *k*-mer mode, we applied meryl (Rhie et al., 2020) to generating and counting 21-mers. In *strobemer* mode, we first used Jellyfish (Marcais and Kingsford, 2011) to build large canonical *k*-mers (*k*=40) of two contaminated parental datasets, and then transformed them to *strobemers* [*randstrobes* (20,10,10,30)] by custom C++ scripts.

To benchmark the effect of *k*-mers and *strobemers*, we generated random sequences with a fixed length of 100 kbp and varied the mutation (error) rates. The mutation spots were randomly selected, in which the error probabilities of substitutions, insertions, and deletions were equal. Benchmarking of matched markers under different sampling protocols with mutation rates of 1%, 5%, and 10% was listed in Supplementary Table 1. Each test was independently run 100 times. Either *minstrobes* or *randstrobes* (two types of *strobemers*) match more mutation spots than *k*-mers, especially for higher error rates. This benchmark is available at https://github.com/BGI-Qingdao/strobemer_cpptest. Supplementary Figure 4 shows the precision-recall curve with trio-binning *strobemers* and *k*-mers for three simulated datasets. The implementation of *strobemers* obtains relatively better performance on average.

2.4 Unsupervised clustering of remaining reads

The following procedure of screening relies on the raw read clustering analogous to the genomic binning. The high-dimensional feature space for the unsupervised clustering of long reads consists of characteristic marker densities as Feature 1-3, the GC content and canonical 3-mer frequencies as Feature 4-36 (Supplementary Table 2). The preprocessing of PCA decomposes those features into new K independent variables of a 36-dimensional matrix. We assume that the K principal components belong to the same parametric family of Gaussian distribution but with various parameters (mean, variance). The preprocessing of whitening has been employed to reduce information redundancy. The BGMM algorithm is selected according to the applicable geometry and running speed from the comparison of clustering algorithms in scikit-learn (Pedregosa et al., 2011). Briefly, the Bayesian framework infers the posterior distribution of the parameters $\widetilde{\phi}_i$, $\widetilde{\mu}_i, \widetilde{\sigma}_i$, and the expectation–maximization algorithm is utilized to update these parameters.

$$p(q \mid x) = \sum_{i=1}^{K} \widetilde{\phi}_i \cdot \mathrm{N}(\widetilde{\mu}_i, \widetilde{\sigma}_i)$$

where $\widetilde{\phi}_i$, $\widetilde{\mu}_i$, $\widetilde{\sigma}_i$ are the weight, mean and variance of the ith component. Application Programming Interface from scikit-learn (Pedregosa et al., 2011) was used to implement these steps. This unsupervised machine learning algorithm is irrelevant to training. Therefore, no reference genomes or high-quality public databases are needed.

We illustrated the differences in screening results before and after clustering in the *k*-mer mode in Figure 2A and Supplementary Figure 5. Overall, the clustering increased classification F1-scores by improving the recall rates but sacrificing fractional precision. Note that we only used long reads longer than 5,000 bp for clustering since the inaccurate sequence statistics of shorter reads might mislead the results. The error-tolerant *strobemer* mode usually did not need further clustering, as the recall rate was already satisfactory. We did not cluster the *k*-mer-based result of the real ONT dataset, because the performance was adequate and extra clustering could not improve it further.

Besides, the randomness of BGMM clustering may affect the final classification of host reads. Thus, we independently ran the clustering procedure multiple times (default 10) and achieved consensus results. The best host read cluster was annotated according to the criterion that whether this cluster contained the most preselected high-confidence host reads with the smallest variance. The final host

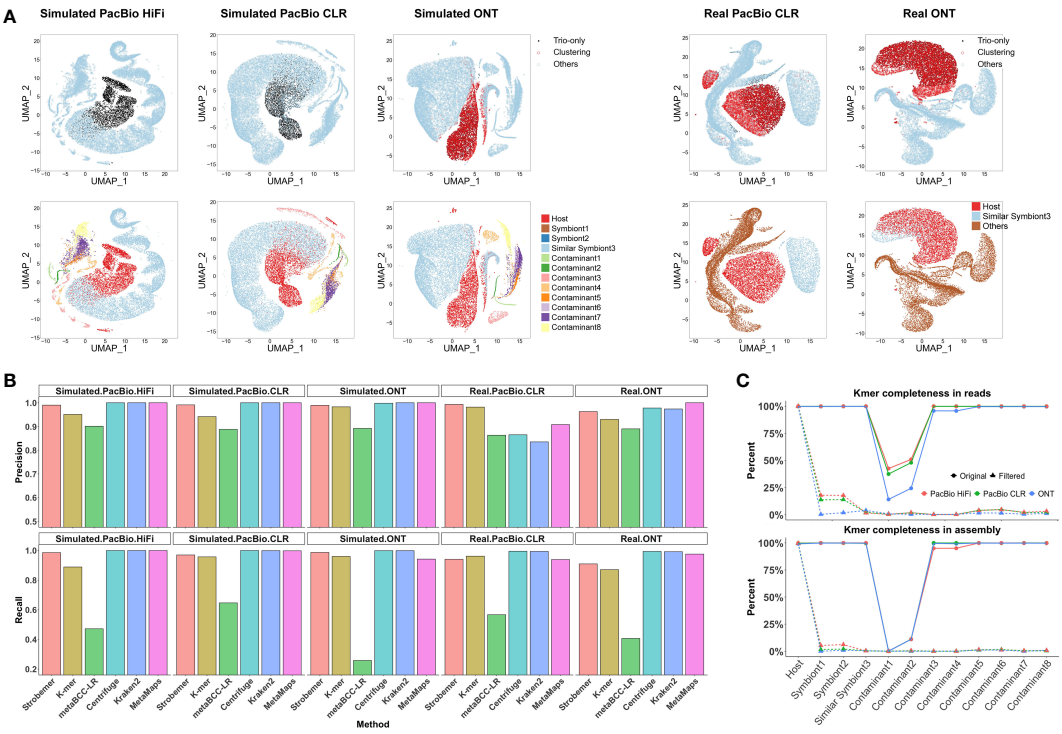


FIGURE 2
Performance of screening over different benchmarking datasets. **(A)** Visualization of identified host reads by trio-binning markers and unsupervised clustering after dimension reduction. We arbitrarily selected 20,000 long reads to cancel image blurring. Each point refers to one long read. For each dataset, top figure is colored by screening results *via* Symbiont-Screener, while bottom is colored by the original species. Note that the simulated PacBio HiFi and CLR reads with higher single-base accuracy are classified by trio-binning markers only. The other three datasets are further identified by clustering. **(B)** Comparisons of screening precision and recall rates over five simulated and real PacBio and ONT long-read datasets. Note that for the *de novo* tool, MetaBCC-LR, we benchmarked the result based on the extracted clusters with more than 5,000 reads and precision >0.5 as the host clusters cannot be identified without the reference genome. **(C)** Distinct *k*-mer completeness for each species in long-read data (top) and final assemblies (bottom) after screening. The low *k*-mer completeness of Contaminant 1 and 2 before filtering is because of the 1x input. On average, distinct *k*-mers for other foreign species including Symbiont3 sharing highly similar sequences with the host are reduced from 99.6% to 0.5% in raw long-read data and from 99.4% to 0.2% in assemblies, while >98.2% of host's *k*-mers are retained in both raw data and following assemblies.

group was automatically produced by merging raw reads, which repeatedly occurred in the best and second-best clusters. Supplementary Figure 6 illustrates the frequencies of occurrence of different species in the best and second-best clusters for 10 runs of BGMM clustering. In the dumbbell-shaped overall profile, the foreign genomic data aggregate in the low-frequency region while the host's stay in the high-frequency region. Users can determine the frequency threshold based on the position of the second peak in the high-frequency region to extract host data for *de novo* projects.

2.5 Reconstruction of the host and metagenomic genomes

TGS assemblers, for instance, Canu, Flye, and metaFlye (Koren et al., 2017; Kolmogorov et al., 2019; Kolmogorov et al., 2020) can reconstruct the host chromosomes and symbiotic microbial genomes. We employed default parameters to assemble PacBio HiFi, PacBio CLR, and ONT data types. Note that although the trio-binning markers can partition paternal and maternal reads, we did not show the result of haplotype-resolved assemblies because the host sequencing coverage depth in our datasets is insufficient.

2.6 Benchmarking datasets

We used the following five long-read datasets for evaluation.

Dataset1: Simulated PacBio HiFi dataset: human chromosome 19 of HG002 as the host, two bacteria from the Unified Human Gastrointestinal Genome (UHGG) collection (Almeida et al., 2021) as inter-phyla symbionts, additional eight UHGG bacteria as random contaminants, and mouse chromosome 7 as a symbiont sharing similar sequences with the host genome. PacBio HiFi long reads were simulated by PBSIM2 (Ono et al., 2021) based on the reference genomes with an average read length of 10 kbp and an average error rate of 1% (Supplementary Table 3).

Dataset2: Simulated PacBio CLR dataset: same composition as *Dataset1*. The average read length is 10 kbp and the average error rate is 5%.

Dataset3: Simulated ONT dataset: same composition as *Dataset1*. The average read length is 30 kbp and the average error rate is 15%.

Dataset4: Real PacBio RSII CLR dataset: human chromosome 19 of HG002 as the host, two bacteria of mock microbial datasets from the ZymoBIOMICS Microbial Community Standards as inter-phyla symbionts, other eight bacteria and yeasts as random contaminants. In this case, we challenged the chimpanzee chromosome 21 as a

symbiont owning highly similar sequences (Supplementary Table 4). The human HG002 raw reads were downloaded from NCBI GIAB, and those mapped to chromosome 19 of the HG002 reference assembly (GCA_011064465.1) were extracted (Shumate et al., 2020). The chimpanzee raw reads were downloaded from (Logsdon et al., 2021), and those mapped to the chromosome 21 reference (GCA_000001515.5) were extracted. The mock raw reads were downloaded from (McIntyre et al., 2019).

Dataset5: Real ONT PromethION dataset: same composition as *Dataset4*. The human and chimpanzee raw reads were downloaded and extracted in the same way. The mock data were downloaded from (Nicholls et al., 2019).

We used default error profiles (substitution: insertion: deletion=6: 50: 54 for PacBio and 23:31:46 for ONT) provided by PBSIM2 for all simulations. 50× host (human chromosome 19) data were simulated using the HG002 reference (GCA_011064465.1), while 50× mouse chromosome 7 data were simulated based on the GRCm39 reference (GCA_000001635.9). Other bacteria were simulated based on UHGG references with various coverages. For the real PacBio and ONT data, we also gained 50× human and 50× chimpanzee raw reads, and maintained the same proportions of 10 bacteria and yeasts in the mock sequencing data.

All the long reads were annotated by their species names before binning to benchmark the precision and recall rates of screening. The recall rate is defined as the ratio of the number of correctly identified host reads to the total number of host reads in the mixed input. The precision rate is defined as the ratio of the number of correctly identified host reads to the total number of final extracted reads.

2.7 Evaluation methods

To assess the effect of screening on the host genome assembly for three simulated datasets, we generated species-specific non-repeating canonical 21-mers (distinct *k*-mers) according to the reference genomes. We calculated the completeness and contamination rates in the mixed and purified long-read raw data and assemblies, respectively (Supplementary Table 5). The completeness is defined as the ratio of distinct *k*-mer numbers belonging to each species in reads or assemblies to the total number in the reference. The contamination rate is defined as the ratio of the distinct *k*-mer number belonging to other species to the total distinct *k*-mer number of the host reference assembly.

The host reference assembly was used to validate the assemblies before and after screening. The assembly statistics were reported by QUAST (Gurevich et al., 2013) with default parameters except -m 1000.

The metagenomic analysis of host-associated microbial community is supposed to be improved after screening if we ignore the random contamination. However, the evaluation of this effect does not apply to our benchmarking datasets.

2.8 Parameters of other tools

We used recommended or default parameters for MetaProb, BusyBee, MetaBCC-LR, Centrifuge, Kraken2, and MetaMaps. However, MetaProb and BusyBee cannot support the large data size. We performed parameter sweeps for MetaBCC-LR to obtain its best performance (–sample-count 1%, –bin-size 10). Note that the host clusters generated by MetaBCC-LR cannot be detected as the host without references. In addition, the host data are usually split into several clusters. Thus, we extracted all clusters with more than 5,000 reads and precision >0.5 for evaluation. Centrifuge, Kraken2, and MetaMaps are reference-based classifiers. We first built indexed databases according to host and foreign reference assemblies, as well as NCBI taxonomy. Next, we classified datasets and extracted all reads assigned to the host species for benchmarking.

3 Results

3.1 Screening of simulated and real datasets

We applied Symbiont-Screener to three simulated and two real long-read datasets with varying read lengths and error rates, covering PacBio HiFi, PacBio CLR, and ONT types. Each dataset consists of a host species, inter-phyla symbionts, a symbiont with similar sequences, and several random microbial contaminants.

Figure 2A shows the host's data identified by characteristic markers and further clustered by other genomic features. Compared with the corresponding ground truth, the parent-specific and shared *strobemers* precisely detect host long reads for three simulated datasets. Limited by the sequencing errors, the simulated ONT dataset requires species differentiation in GC content and trinucleotide frequencies to obtain more host data as supplementary features. Long reads sufficiently close to the preselected reads in the space after dimension reduction are also marked as host's and extracted for the following assembly. Real long-read datasets for symbiotic samples are rare. Thus, we chose chimpanzee to imitate an indistinguishable symbiont, sharing approximately 98% of the genome with the host, human, which is a conundrum of alignment-based screening. Although a few foreign reads are misidentified by trio-binning markers in the real PacBio dataset, they are further corrected by genomic signatures. The relatively higher error rate accompanied by the high sequence similarity leads to the lower accuracy of screening for real ONT dataset. Besides, we evaluated the contribution of each feature to the final clustering results (Supplementary Figure 7). For three simulated datasets, the relative importance score for characteristic *strobemers* is 24.4% on average, among which the parent-specific markers are more important. None of their contributions are negligible in the clustering, which proves those features are highly complementary.

We have also tested several state-of-the-art reference-free or reference-based tools for screening using the same datasets. Figure 2B shows the classification precision and recall of the *strobemer* mode, *k*-mer mode, MetaBCC-LR, Centrifuge, Kraken2, and MetaMaps. Hampered by the computing performance, neither MetaProb nor BusyBee can support clustering the entire data. Overall, Symbiont-Screener outperforms MetaBCC-LR for all five datasets. The implementation of trio-binning *strobemers* and clustering allows Symbiont-Screener to surmount the obstacle of high error rates, thus extracting more host long reads. By contrast, Centrifuge, Kraken2,

and MetaMaps benefit from accurate and complete reference genomes as well as taxonomic relationships, and obtain relatively higher precision and recall rates, especially for the simulated datasets. Nevertheless, for more complex relations in the real PacBio dataset where the similar symbiont, chimpanzee shares approximately 98% of the genome with the host, the highest F1-scores are found in both modes of Symbiont-Screener. It indicates that the trio-binning information is a qualified substitute for reference genomes or public databases if they are not available.

3.2 Effect on host assembly

The effect of screening on the final assembly was first assessed by the species-specific genomic non-repeating canonical 21-mers. The distinct *k*-mer completeness ratio indicated the ability to reconstruct the whole host genome. Meanwhile, the distinct *k*-mer contamination ratio represented the assembly accuracy.

For three simulated datasets, the filtered long reads obtained up to 99.4% of the host's 42,988,682 distinct *k*-mers regardless of repeats or non-ACGT bases on average (Figure 2C), thus ensuring the high quality of genome assembly (~99.2% completeness). On the contrary, *k*-mer completeness was significantly reduced to 0.3% in raw reads and 0.1% in final assembly for all ten bacteria after screening. For the challenging symbiont with highly similar sequences, the most difficult component to be cleaned, only 2.4% of *k*-mers were retained in the results. They could not support the foreign genome reconstruction in the following host assembly.

The QUAST-based evaluations also reflected the advantage after a nearly perfect screening. Supplementary Table 7 showed the Canu assembly statistics for three simulated and two real datasets. The assembled total length and unaligned length were significantly reduced as the foreign genomes were removed after the screening, while the genome fraction and misassemblies remained almost the same. The comparison was nearly consistent with that of Flye assemblies as shown in Supplementary Table 8.

3.3 Application to a red seaweed with symbionts

We applied Symbiont-Screener to an economically important red seaweed, *Neoporphyra haitanensis*, to demonstrate the success of screening results in the natural world. Previous studies have shown a complex relationship between the host algae and the associated metagenomes, involving the microbial components, functional microbial lineages, and the exchange of diverse chemical currencies, which mainly rely on the sequence alignments of short-read reads or genome assemblies (Brawley et al., 2017; Wang et al., 2022). Here, we sequenced 51.6 Gb ONT long reads (read N50 = 25.2 kb) and 23.2 Gb PacBio HiFi reads (read N50 = 18.3 kb) and separate the host seaweed sequences from the symbionts. Symbiont-Screener employed the characteristic *strobemers* to automatically identify host raw reads without reference genomes. The 108,837,094 characteristic *strobemers* were generated by the trio pedigree relations of the lab-cultured parents and offspring. The clustering procedure further gathered host long reads to overcome the limit of high sequencing error rate. Finally, the whole identified host raw reads were assembled by Flye, while the remaining associated bacteria were assembled by metaFlye, respectively. We regarded all foreign genomes as symbiotic bacteria. We also applied metaFlye to the assembly of the mixed data for comparison.

The total length of the reconstructed seaweed genome was 45,172,822 bp, consisting of 59 contigs (Supplementary Table 9). The assembly contiguity reached a chromosomal level with a contig N50 of 7,218,067 bp, compared to the previously published closely-related species (Chen et al., 2022). The 81.454% genome fraction against the 53.3 Mbp closely-related genome implied the assembly completeness. Only 2,084,431 bp were unaligned. The significant difference of the chromatin contacts from pair-wise Hi-C reads confirmed the thorough isolation of the host, which further constructed 5 complete chromosomes (Figure 3A). On the other hand, the metagenome-assembled genomes involved 608 contigs with a contig N50 of 4,376,601 bp. The total genome size was 260,496,743 bp, of which only 3.829% could be aligned to the closely-related reference genome. GC-depth plot is an alternative method to benchmark the screening result. Multiple peaks in the preliminary assembly of the whole mixed data indicated different species with various GC contents and covered read depths (Figure 3B). In contrast, the purified host assembly after screening presented a more concentrated peak with a more convergent distribution of GC ratio.

Additionally, the metagenomes were binned and taxonomically annotated by Kraken2 using the indexed NCBI RefSeq database k2_pluspfp_20200919, which illustrated that the sequences annotated as eukaryote were dramatically eliminated after screening (Figure 3C). The profiling results also disclosed a bias of different sequencing platforms, possibly due to the more misalignments induced by the relatively higher error rates for ONT. Bandage (Wick et al., 2015) were used to visualize the *de novo* assembly graphs with sequence connections. There were totally 14 complete, closed and circularized metagenomes were assembled (Figure 3D). Although the 322,277,090 bp-long mixed assembly reconstructed an additional circularized genome, it was annotated as eukaryote. We investigated the mixed assembly and found 84.4 Mbp contigs could be aligned to the closely-related reference of the host seaweed, implying the host assembly errors. The high-quality genomes of *Neoporphyra haitanensis* and associated bacteria might provide a comprehensive approach for elucidating genome coevolution and the influence of symbiotic metagenomes to the adaptation of *Pyropia* to intertidal zone habitats.

3.4 Performance

We benchmarked the performance of Symbiont-Screener on a Linux system with Intel Core Processor (Broadwell, IBRS), 15 CPU cores and 30 threads. We individually recorded the CPU and memory usage for each assembly and calculated the percentage of saved consumption after screening. Supplementary Figure 8 recorded the computational consumption for three simulated PacBio HiFi, PacBio CLR and ONT datasets, representing that the screening result saves considerable CPU and memory usage.

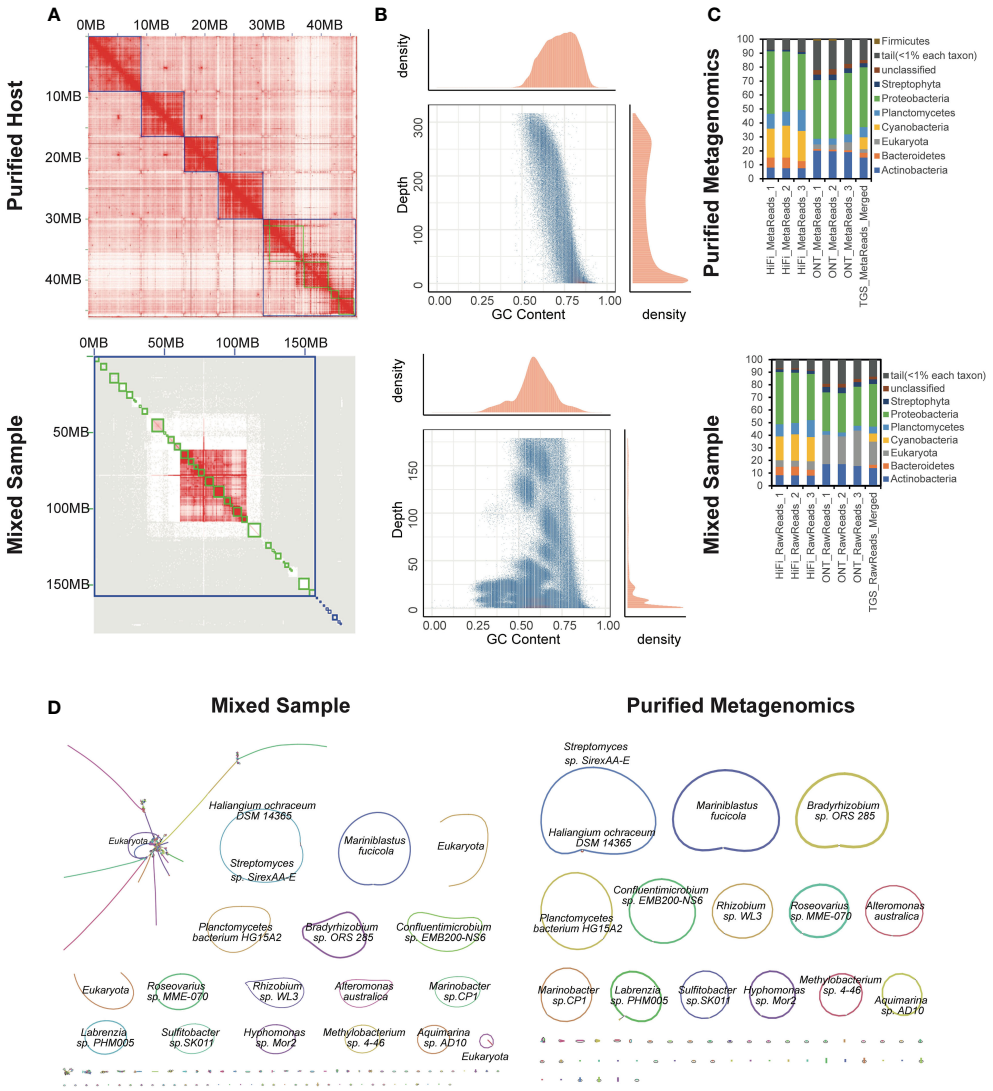


FIGURE 3
Application to an algal-bacterial ecosystem. **(A)** Hi-C interaction heatmaps of purified (top) and mixed (bottom) assemblies. **(B)** GC-Coverage plots for simulated PacBio HiFi assemblies of purified (top) and mixed (bottom) data. **(C)** Differences in the taxonomic profiling of associated microbial community structures at the phylum level based on the indexed RefSeq database between purified (top) and mixed (bottom) data. **(D)** Bandage visualization of assembly graphs with connections for the mixed and purified data. Note that the primary assembled sequences are annotated by Kraken2.

4 Discussion

We introduce a novel but accurate model for screening that classifies reliable host raw long reads from the mixed sample according to the trio-binning information, which is computationally efficient without the requirement of reference genomes or sequence alignments. Symbiont-Screener further utilizes other supplementary features to directly cluster error-prone long reads. The multi-dimensional clustering system is open-ended and accepts additional features such as remaining genomic markers after sterilization to avoid overreliance on the presence or accuracy of one specific feature. Moreover, the trio-binning markers support the haplotype-resolved partitioning and genome assembly of extracted host's long reads. We did not show the haplotype-resolved assemblies due to the insufficient sequencing coverage depth of host's data.

The application of this algorithm requires parental sequencing data with or without symbionts and contamination for sexually reproducing diploid or allotetraploid species. Therefore, the patient's parents need to provide their clean or contaminated DNA samples for the microbial pathogen identification in clinical applications. For samples of animals or plants collected from the wild, sexual reproduction in the field or laboratory culture is required to eliminate symbionts and random contaminants. If parental data are unobtainable, then reference assemblies of closely-related species if available, or parental lines for cross-bred crops can be used to mark long reads corresponding to the conserved genomic regions instead.

Data availability statement

The simulated PacBio HiFi, PacBio CLR and ONT datasets have been deposited in the CNGB Sequence Archive (CNSA, https://db.cngb.org/cnsa) under the accession number CNP0001829. We downloaded real PacBio CLR and ONT ultra-long data of HG002/NA24385 as host

from GIAB (https://ftp://ftp-trace.ncbi.nlm.nih.gov/giab/ftp/data/AshkenazimTrio/HG002_NA24385_son/PacBio_MtSinai_NIST and https://ftp://ftp-trace.ncbi.nlm.nih.gov/giab/ftp/data/AshkenazimTrio/HG002_NA24385_son/Ultralong_OxfordNanopore). The real PacBio and ONT datasets for chimpanzee are available in the NCBI under the accession number PRJNA659034. PacBio data for the mock microbial community from ZymoBIOMICS Microbial Community Standards are extracted, which are publicly available from (McIntyre et al., 2019). The ONT data for the same mock standard are obtained from (Nicholls et al., 2019). The algal-bacterial data have been deposited in the CNSA under the accession number CNP0003571. But restrictions apply to the availability of these algal-bacterial data, which are not publicly available. Data are however available from the corresponding author upon request. The RefSeq-based database used for the Kraken2 analysis can be downloaded at https://genome-idx.s3.amazonaws.com/kraken/k2_pluspfp_20200919.tar.gz. The source code used in this manuscript is available at https://github.com/BGI-Qingdao/Symbiont-Screener.

Author contributions

MX and LG performed software design and implementation. MX, LG, YQ, CS, XCL, JC, and JH contributed to data modeling, data curation, and benchmarking. MX wrote the draft manuscript, and LG, LD, and GF contributed to manuscript editing. XL and GF supervised the project. MX and GF secured funding. All authors read and approved the final manuscript. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the National Natural Science Foundation of China (Grant No. 32100514); and the National Key Research and Development Program of China (Grant No. 2018YFD0900301-05).

Acknowledgments

We thank Kristoffer Sahlin for help with implementation of *strobemers*. We thank Xiaobin Liu for fruitful discussions during the development and performance test. The data that support the findings of this study have been deposited into CNGB Nucleotide Sequence Archive (CNSA) (Guo et al., 2020) of China National GeneBank DataBase (CNGBdb) (Chen et al., 2020) with accession numbers CNP0001829.

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.2023.1087447/full#supplementary-material

References

Almeida, A., Nayfach, S., Boland, M., Strozzi, F., Beracochea, M., Shi, Z. J., et al. (2021). A unified catalog of 204,938 reference genomes from the human gut microbiome. *Nat. Biotechnol.* 39, 105–114. doi: 10.1038/s41587-020-0603-3

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

Arimoto, A., Hikosaka-Katayama, T., Hikosaka, A., Tagawa, K., Inoue, T., Ueki, T., et al. (2019). A draft nuclear-genome assembly of the acoel flatworm praesagittifera naikaiensis. *Gigascience* 8. doi: 10.1093/gigascience/giz023

Bertrand, D., Shaw, J., Kalathiyappan, M., Ng, A. H. Q., Kumar, M. S., Li, C., et al. (2019). Hybrid metagenomic assembly enables high-resolution analysis of resistance determinants and mobile elements in human microbiomes. *Nat. Biotechnol.* 37, 937–944. doi: 10.1038/s41587-019-0191-2

Bharti, R., and Grimm, D. G. (2019). Current challenges and best-practice protocols for microbiome analysis. *Briefings Bioinf.* 22, 178–193. doi: 10.1093/bib/bbz155

Brawley, S. H., Blouin, N. A., Ficko-Blean, E., Wheeler, G. L., Lohr, M., Goodson, H. V., et al. (2017). Insights into the red algae and eukaryotic evolution from the genome of porphyra umbilicalis (Bangiophyceae, rhodophyta). *Proc. Natl. Acad. Sci. U.S.A.* 114, E6361–e6370. doi: 10.1073/pnas.1703088114

Chen, H., Chu, J. S., Chen, J., Luo, Q., Wang, H., Lu, R., et al. (2022). Insights into the ancient adaptation to intertidal environments by red algae based on a genomic and multiomics investigation of neoporphyra haitanensis. *Mol. Biol. Evol.* 39. doi: 10.1093/molbev/msab315

Cheng, X.-W., Li, J., Zhang, L., Hu, W.-J., Zong, L., Xu, X., et al. (2022). Identification of SARS-CoV-2 variants and their clinical significance in hefei, China. *Front. Med.* 8. doi: 10.3389/fmed.2021.784632

Cheng, S., Xian, W., Fu, Y., Marin, B., Keller, J., Wu, T., et al. (2019). Genomes of subaerial zygnematophyceae provide insights into land plant evolution. *Cell* 179, 1057–1067.e14. doi: 10.1016/j.cell.2019.10.019

Chen, F. Z., You, L., Yang, F., Wang, L. N., Guo, X. Q., Gao, F., et al. (2020). CNGBdb: China national GeneBank DataBase. *Hereditas* 42, 799–809. doi: 10.16288/j.yczz.20-080

Chin, C. S., Alexander, D. H., Marks, P., Klammer, A. A., Drake, J., Heiner, C., et al. (2013). Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat. Methods* 10, 563–569. doi: 10.1038/nmeth.2474

Coghlan, A., Tyagi, R., Cotton, J. A., Holroyd, N., Rosa, B. A., Tsai, I. J., et al. (2019). Comparative genomics of the major parasitic worms. *Nat. Genet.* 51, 163–174. doi: 10.1038/s41588-018-0262-1

Cornet, L., and Baurain, D. (2022). Contamination detection in genomic data: More is not enough. *Genome Biol.* 23, 60. doi: 10.1186/s13059-022-02619-9

Dilthey, A. T., Jain, C., Koren, S., and Phillippy, A. M. (2019). Strain-level metagenomic assignment and compositional estimation for long reads with MetaMaps. *Nat. Commun.* 10, 3066. doi: 10.1038/s41467-019-10934-2

Douvlataniotis, K., Bensberg, M., Lentini, A., Gylemo, B., and Nestor, C. E. (2020). No evidence for DNA N^6-methyladenine in mammals. *Sci. Adv.* 6, eaay3335. doi: 10.1126/sciadv.aay3335

Ebert, P., Audano, P. A., Zhu, Q., Rodriguez-Martin, B., Porubsky, D., Bonder, M. J., et al. (2021). Haplotype-resolved diverse human genomes and integrated analysis of structural variation. *Science.* 372 doi: 10.1126/science.abf7117

Fraser, C. M., Eisen, J. A., Nelson, K. E., Paulsen, I. T., and Salzberg, S. L. (2002). The value of complete microbial genome sequencing (you get what you pay for). *J. Bacteriol* 184, 6403–5; discusion 6405. doi: 10.1128/JB.184.23.6403-6405.2002

Girotto, S., Pizzi, C., and Comin, M. (2016). MetaProb: accurate metagenomic reads binning based on probabilistic sequence signatures. *Bioinformatics* 32, i567–i575. doi: 10.1093/bioinformatics/btw466

Guo, X., Chen, F., Gao, F., Li, L., Liu, K., You, L., et al. (2020). CNSA: a data repository for archiving omics data. *Database* 2020. doi: 10.1093/database/baaa055

Gurevich, A., Saveliev, V., Vyahhi, N., and Tesler, G. (2013). QUAST: Quality assessment tool for genome assemblies. *Bioinformatics* 29, 1072–1075. doi: 10.1093/bioinformatics/btt086

Kim, D., Lee, J. Y., Yang, J. S., Kim, J. W., Kim, V. N., and Chang, H. (2020). The architecture of SARS-CoV-2 transcriptome. *Cell* 181, 914–921.e10. doi: 10.1016/j.cell.2020.04.011

Kim, D., Song, L., Breitwieser, F. P., and Salzberg, S. L. (2016). Centrifuge: Rapid and sensitive classification of metagenomic sequences. *Genome Res.* 26, 1721–1729. doi: 10.1101/gr.210641.116

Kolmogorov, M., Bickhart, D. M., Behsaz, B., Gurevich, A., Rayko, M., Shin, S. B., et al. (2020). metaFlye: Scalable long-read metagenome assembly using repeat graphs. *Nat. Methods* 17, 1103–1110. doi: 10.1038/s41592-020-00971-x

Kolmogorov, M., Yuan, J., Lin, Y., and Pevzner, P. A. (2019). Assembly of long, error-prone reads using repeat graphs. *Nat. Biotechnol.* 37, 540–546. doi: 10.1038/s41587-019-0072-8

Koren, S., Rhie, A., Walenz, B. P., Dilthey, A. T., Bickhart, D. M., Kingan, S. B., et al. (2018). *De novo* assembly of haplotype-resolved genomes with trio binning. *Nat. Biotechnol.* 36, 1174–1182. doi: 10.1038/nbt.4277

Koren, S., Walenz, B. P., Berlin, K., Miller, J. R., Bergman, N. H., and Phillippy, A. M. (2017). Canu: Scalable and accurate long-read assembly *via* adaptive k-mer weighting and repeat separation. *Genome Res.* 27, 722–736. doi: 10.1101/gr.215087.116

Laczny, C. C., Kiefer, C., Galata, V., Fehlmann, T., Backes, C., and Keller, A. (2017). BusyBee web: Metagenomic data analysis by bootstrapped supervised binning and annotation. *Nucleic Acids Res.* 45, W171–w179. doi: 10.1093/nar/gkx348

Logsdon, G. A., Vollger, M. R., Hsieh, P., Mao, Y., Liskovykh, M. A., Koren, S., et al. (2021). The structure, function and evolution of a complete human chromosome 8. *Nature.* 593, 101–107 doi: 10.1038/s41586-021-03420-7

Marcais, G., and Kingsford, C. (2011). A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. *Bioinformatics* 27, 764–770. doi: 10.1093/bioinformatics/btr011

McIntyre, A. B. R., Alexander, N., Grigorev, K., Bezdan, D., Sichtig, H., Chiu, C. Y., et al. (2019). Single-molecule sequencing detection of N6-methyladenine in microbial reference materials. *Nat. Commun.* 10, 579. doi: 10.1038/s41467-019-08289-9

Nagarajan, N., and Pop, M. (2013). Sequence assembly demystified. *Nat. Rev. Genet.* 14, 157–167. doi: 10.1038/nrg3367

Neimark, J. (2015). Line of attack. *Science* 347, 938–940. doi: 10.1126/science.347.6225.938

Nicholls, S. M., Quick, J. C., Tang, S., and Loman, N. J. (2019). Ultra-deep, long-read nanopore sequencing of mock microbial community standards. *GigaScience* 8. doi: 10.1093/gigascience/giz043

Ono, Y., Asai, K., and Hamada, M. (2021). PBSIM2: a simulator for long-read sequencers with a novel generative model of quality scores. *Bioinformatics* 37, 589–595. doi: 10.1093/bioinformatics/btaa835

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* 12, 2825–2830. doi: 10.48550/arXiv.1201.0490

Qi, Y., Gu, S., Zhang, Y., Guo, L., Xu, M., Cheng, X., et al. (2022). MetaTrass: A high-quality metagenome assembler of the human gut microbiome by cobarcoding sequencing reads. *iMeta*, 1, e46. doi: 10.1002/imt2.46

Rhie, A., Walenz, B. P., Koren, S., and Phillippy, A. M. (2020). Merqury: reference-free quality, completeness, and phasing assessment for genome assemblies. *Genome Biol.* 21, 245. doi: 10.1186/s13059-020-02134-9

Rhoads, A., and Au, K. F. (2015). PacBio sequencing and its applications. *Genomics Proteomics Bioinf.* 13, 278–289. doi: 10.1016/j.gpb.2015.08.002

Ricker, N., Qian, H., and Fulthorpe, R. R. (2012). The limitations of draft assemblies for understanding prokaryotic adaptation and evolution. *Genomics* 100, 167–175. doi: 10.1016/j.ygeno.2012.06.009

Rothäusler, E., Gutow, L., and Thiel, M. (2012). Floating Seaweeds and Their Communities. In: Wiencke, C., Bischof, K. (eds) Seaweed Biology. Ecological Studies. (Berlin, Heidelberg: Springer) 219. doi: 10.1007/978-3-642-28451-9_17

Sahlin, K. (2021). Effective sequence similarity detection with strobemers. *Genome Res.* 31, 2080–2094. doi: 10.1101/gr.275648.121

Shumate, A., Zimin, A. V., Sherman, R. M., Puiu, D., Wagner, J. M., Olson, N. D., et al. (2020). Assembly and annotation of an ashkenazi human reference genome. *Genome Biol.* 21, 129. doi: 10.1186/s13059-020-02047-7

Steinegger, M., and Salzberg, S. L. (2020). Terminating contamination: large-scale search identifies more than 2,000,000 contaminated entries in GenBank. *Genome Biol.* 21, 115. doi: 10.1186/s13059-020-02023-1

Thiel, M., and Gutow, L. (2005). The ecology of rafting in the marine environment. II. the rafting organisms and community. *Oceanography Mar. Biol.* 43, 279–418. doi: 10.1201/9781420037449.ch7

Wang, J., Tang, X., Mo, Z., and Mao, Y. (2022). Metagenome-assembled genomes from pyropia haitanensis microbiome provide insights into the potential metabolic functions to the seaweed. *Front. Microbiol.* 13. doi: 10.3389/fmicb.2022.857901

Wang, D., Yu, X., Xu, K., Bi, G., Cao, M., Zelzion, E., et al. (2020). Pyropia yezoensis genome reveals diverse mechanisms of carbon acquisition in the intertidal environment. *Nat. Commun.* 11, 4028. doi: 10.1038/s41467-020-17689-1

Wickramarachchi, A., Mallawaarachchi, V., Rajan, V., and Lin, Y. (2020). MetaBCC-LR: metagenomics binning by coverage and composition for long reads. *Bioinformatics* 36, i3–i11. doi: 10.1093/bioinformatics/btaa441

Wick, R. R., Schultz, M. B., Zobel, J., and Holt, K. E. (2015). Bandage: interactive visualization of *de novo* genome assemblies. *Bioinformatics* 31, 3350–3352. doi: 10.1093/bioinformatics/btv383

Wood, D. E., Lu, J., and Langmead, B. (2019). Improved metagenomic analysis with kraken 2. *Genome Biol.* 20, 257. doi: 10.1186/s13059-019-1891-0

Woyke, T., Teeling, H., Ivanova, N. N., Huntemann, M., Richter, M., Gloeckner, F. O., et al. (2006). Symbiosis insights through metagenomic analysis of a microbial consortium. *Nature* 443, 950–955. doi: 10.1038/nature05192

Xie, M., Ren, M., Yang, C., Yi, H., Li, Z., Li, T., et al. (2016). Metagenomic analysis reveals symbiotic relationship among bacteria in microcystis-dominated community. *Front. Microbiol.* 7. doi: 10.3389/fmicb.2016.00056

Xie, H., Yang, C., Sun, Y., Igarashi, Y., Jin, T., and Luo, F. (2020). PacBio long reads improve metagenomic assemblies, gene catalogs, and genome binning. *Front. Genet.* 11. doi: 10.3389/fgene.2020.516269

Xu, M., Guo, L., Du, X., Li, L., Peters, B. A., Deng, L., et al. (2021). Accurate haplotype-resolved assembly reveals the origin of structural variants for human trios. *Bioinformatics.* 37. 2095–2102 doi: 10.1093/bioinformatics/btab068

Xu, M., Guo, L., Gu, S., Wang, O., Zhang, R., Peters, B. A., et al. (2020). TGS-GapCloser: A fast and accurate gap closer for large genomes with low coverage of error-prone long reads. *GigaScience* 9. doi: 10.1093/gigascience/giaa094

frontiers | Frontiers in Microbiology

TYPE Original Research
PUBLISHED 08 February 2023
DOI 10.3389/fmicb.2023.1078171

OPEN ACCESS

EDITED BY
Manuel Maldonado,
Center for Advanced Studies of Blanes (CSIC), Spain

REVIEWED BY
Lucia Pita,
Institute of Marine Sciences (CSIC), Spain
Rodrigo Costa,
University of Lisbon, Portugal
Laura Steindler,
University of Haifa, Israel

*CORRESPONDENCE
Zhao-Ming Gao
✉ gaozm@idsse.ac.cn
Yong Wang
✉ wangyong@sz.tsinghua.edu.cn

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

RECEIVED 24 October 2022
ACCEPTED 12 January 2023
PUBLISHED 08 February 2023

CITATION
Wei T-S, Gao Z-M, Gong L, Li Q-M, Zhou Y-L, Chen H-G, He L-S and Wang Y (2023) Genome-centric view of the microbiome in a new deep-sea glass sponge species *Bathydorus* sp. *Front. Microbiol.* 14:1078171. doi: 10.3389/fmicb.2023.1078171

Genome-centric view of the microbiome in a new deep-sea glass sponge species *Bathydorus* sp.

Tao-Shu Wei[1,2], Zhao-Ming Gao[1]*, Lin Gong[3], Qing-Mei Li[1], Ying-Li Zhou[1], Hua-Guan Chen[1,2], Li-Sheng He[1] and Yong Wang[1,4]*

[1]Institute of Deep-Sea Science and Engineering, Chinese Academy of Sciences, Sanya, Hainan, China, [2]University of Chinese Academy of Sciences, Beijing, China, [3]Institute of Oceanology, Chinese Academy of Sciences, Qingdao, Shandong, China, [4]Institute for Ocean Engineering, Shenzhen International Graduate School, Tsinghua University, Shenzhen, China

Sponges are widely distributed in the global ocean and harbor diverse symbiotic microbes with mutualistic relationships. However, sponge symbionts in the deep sea remain poorly studied at the genome level. Here, we report a new glass sponge species of the genus *Bathydorus* and provide a genome-centric view of its microbiome. We obtained 14 high-quality prokaryotic metagenome-assembled genomes (MAGs) affiliated with the phyla Nitrososphaerota, Pseudomonadota, Nitrospirota, Bdellovibrionota, SAR324, Bacteroidota, and Patescibacteria. In total, 13 of these MAGs probably represent new species, suggesting the high novelty of the deep-sea glass sponge microbiome. An ammonia-oxidizing Nitrososphaerota MAG B01, which accounted for up to 70% of the metagenome reads, dominated the sponge microbiomes. The B01 genome had a highly complex CRISPR array, which likely represents an advantageous evolution toward a symbiotic lifestyle and forceful ability to defend against phages. A sulfur-oxidizing Gammaproteobacteria species was the second most dominant symbiont, and a nitrite-oxidizing Nitrospirota species could also be detected, but with lower relative abundance. *Bdellovibrio* species represented by two MAGs, B11 and B12, were first reported as potential predatory symbionts in deep-sea glass sponges and have undergone dramatic genome reduction. Comprehensive functional analysis indicated that most of the sponge symbionts encoded CRISPR–Cas systems and eukaryotic-like proteins for symbiotic interactions with the host. Metabolic reconstruction further illustrated their essential roles in carbon, nitrogen, and sulfur cycles. In addition, diverse putative phages were identified from the sponge metagenomes. Our study expands the knowledge of microbial diversity, evolutionary adaption, and metabolic complementarity in deep-sea glass sponges.

KEYWORDS

ammonia-oxidizing archaea, symbiont *Bdellovibrio*, the South China Sea, CRISPR, phage

Introduction

Marine sponges in the phylum Porifera are important members of marine benthic communities and emerged on Earth at ~600 mya (Yin et al., 2015). Sponges are extensively found throughout the global oceans, from shallow water to deep sea, from temperate to arctic regions, along shelves, on ridges, and on seamounts (Howell et al., 2016; Maldonado et al., 2017). Sponges usually harbor dense and diverse prokaryotic communities, which can account for up to 35% of the sponge biomass (Webster and Thomas, 2016; Pita et al., 2018). In terms of taxonomic diversity, up to 63 prokaryotic phyla have been

recovered from marine sponges (Schmitt et al., 2012; Moitinho-Silva et al., 2017b). Considering the diversity and potential functional importance of the symbiotic microbiome, sponges are frequently referred to "holobionts", a complex and interdependent consortium that comprises the sponge host and the entire microbiome (Pita et al., 2018). Because of their ancient origin and intense association with microbes, sponges and their microbiomes from shallow waters have been extensively studied (Thomas et al., 2016; Zhang et al., 2019a).

Deep-sea sponges are of importance in ecology functions such as microhabitat provision, substrate modification, benthic–pelagic coupling, and nutrient cycling (Bell, 2008; Maldonado et al., 2012; de Goeij et al., 2013; Kutti et al., 2013; Hawkes et al., 2019). The diversity and novelty of deep-sea sponges have drawn increasing attention, as well as the associated microbiomes (Steinert et al., 2020; Busch et al., 2022). Nutrient conversions by the ammonia-oxidizing archaea (AOA), the nitrite-oxidizing bacterium (NOB), and the sulfur-oxidizing bacterium (SOB) inhabiting the deep-sea glass sponge *Lophophysema eversa* were revealed to carry out relatively complete carbon, sulfur, and nitrogen cycles (Tian et al., 2016). The major microbial lineages in the deep-sea glass sponge *Vazella pourtalesii* underlined probably benefit from their small genome sizes and low GC contents likely due to adaptation to the unique seawater environment (Bayer et al., 2020). Especially, a comprehensive analysis of prokaryotic communities associated with 13 phylogenetically diverse deep-sea sponge species in the South Pacific Ocean revealed that archaeal 16S rRNA gene numbers were up to three orders of magnitude higher than those in shallow-water sponges, which highlighted the importance of the archaea for deep-sea sponges (Steinert et al., 2020). A large-scale analysis of microbial diversity further revealed the biodiversity, environmental drivers, and sustainability of the global deep-sea sponge microbiome, and underscored the uniqueness of each deep-sea sponge ground (Busch et al., 2022). However, knowledge about the diversity and novelty of deep-sea sponge microbial consortia remains limited. The metabolic potential and mutualistic strategies of deep-sea sponge symbionts are also poorly studied.

Following the deepening research of sponge symbionts, viruses in sponges and their functions on the sponge holobionts have also drawn increasing attention recently. Viruses are the most abundant organisms in marine environments, infecting nearly all organisms and having a direct impact on energy flux in marine food webs by regulating prokaryotic and eukaryotic populations (Suttle, 2005; Roux et al., 2015a). Because of the lifestyle of the host for water filtration, sponge symbionts are also exposed to the microenvironment with high-flux viruses (Pascelli et al., 2020). In 1978, transmission electron micrographs revealed the presence of viral-like particles (VLPs) in sponges (Vacelet and Gallissian, 1978). Viral ecology analysis in nine sponge species from the Great Barrier Reef and seven from the Red Sea sponges has provided a comprehensive insight into sponge-associated phage communities (Pascelli et al., 2020). Imaging and bioinformatics analyses indicated the importance of animal–phage–bacterium tripartite interplay in a sponge holobiont (Jahn et al., 2021). Briefly, sponge-associated viromes are becoming new research hotspots, particularly for deep-sea sponge-associated viromes.

In this study, we report an undescribed deep-sea glass sponge species from the South China Sea and provide a genome-centric view of its microbiome. We successfully retrieved 14 microbial metagenome-assembled genomes (MAGs). Among them, a novel Nitrososphaerota AOA species was found to dominate the prokaryotic consortium and exhibit typical symbiotic characteristics. We also identified two *Bdellovibrio* species that have probably undergone symbiotic evolution for adaptation to a sponge-associated lifestyle. Finally, we propose a new conceptual framework based on all possible interactions in the sponge holobiont, especially with the potential involvement of phages as we identified many phage-like contigs in the sponge metagenome.

Materials and methods

Sample collection and DNA extraction

The *R/V Tansuoyihao* TS-7 cruise was carried out in April 2018 in the South China Sea (SCS). A single sponge specimen (hereafter referred to as SQW35) was collected by dive No. 35 of the manned submersible *Deep-Sea Warrior* from a north SCS site (18° 41′ N, 113° 22′ E) at ~983 m depth. Upon arrival at the main deck of the *R/V*, the specimen was rinsed several times with 0.22-μm membrane-filtered seawater to remove loosely attached microbes and debris. The cleaned samples were placed into separate sterile plastic bags, transported to the laboratory, and stored at −80°C. Two pieces of tissue (~0.5 cm^3 per piece) from an inner part of the sponge body as technological replicates (SQW35-1 and SQW35-2) were separately put into 1 ml of DNA extraction buffer (50 mM Tris–HCl, 40 mM EDTA, 500 mM NaCl, 0.75 M sucrose, pH = 8) and fully cut into tiny pieces with sterile scissors. Total suspensions of two replicates were separately subjected to DNA extraction with the PowerSoil® DNA Isolation Kit (MoBio Laboratories, Carlsbad, CA, USA) following the default experimental procedure. The extracted DNA was quantified by the Qubit® dsDNA HS Assay Kit with Qubit 2.0 fluorometer (Invitrogen, Carlsbad, CA, USA) and stored at −80°C for further processing.

Metagenome DNA sequencing and prokaryotic community analyses

One hundred nanograms of genomic DNA for each replicate were randomly fragmented to ~350 bp by Covaris M220 Focused-ultrasonicator (Covaris, Massachusetts, USA). Two metagenomic libraries were constructed using the TruSeq® Nano DNA LT Kit (Illumina, San Diego, CA, USA) and sequenced on a HiSeq2500 platform (Illumina, San Diego, CA, USA) to produce 2 × 150 bp paired-end reads. Raw reads were trimmed and filtered by Fastp v0.20.0 (Chen et al., 2018) with default settings. The data were further processed with FastUniq v1.1 (Xu et al., 2012) to remove duplicated reads. 16S rRNA gene sequences were predicted from clean reads using the rRNA_HMM program for analysis (subsequently referred to as 16S miTags) (Huang et al., 2009) and subjected to QIIME1 pipelines for microbial community analyses (Caporaso et al., 2010). Briefly, the 16S rRNA reads with a shared similarity of 97% were clustered to operational taxonomic units (OTUs) using UCLUST. The longest read of each OTU was selected as the representative for further taxonomic classification with the SILVA 138 database as a reference (Quast et al., 2013). OTUs that were annotated as chloroplasts, mitochondria, and eukaryotes were excluded. False positive OTUs that were not assigned to any taxa were also removed. Finally, the taxonomic relative abundance of

the microbial communities in two replicates was calculated at the phylum level.

Metagenome assembly and genome binning

Clean data of two technical replicates, SQW35-1 and SQW35-2, were separately assembled using both SPAdes v3.1.1 with the recommended settings for metagenomes (Bankevich et al., 2012) and MEGAHIT v1.2.6 with default settings (Li et al., 2015). Four assembled datasets (SQW35-1-spades, SQW35-2-spades, SQW35-1-megahit, and SQW35-2-megahit) were subjected to MetaWRAP v1.2.1 (Uritskiy et al., 2018) for genome binning with default settings, which generated 38 MAGs. These retrieved MAGs were then de-replicated together using dRep v.2.2.3 (Olm et al., 2017), which resulted in 14 high-quality MAGs. CheckM v1.0.11 was used to evaluate the completeness and contamination of the MAGs (Parks et al., 2015). Those with completeness > 50% and contamination < 10% were retained. Taxonomic annotation and relative evolutionary distance (RED) calculation of the MAGs were carried out by GTDB-tk v0.2.2 (Chaumeil et al., 2019). The nomenclature of the MAG was based on the List of Prokaryotic names with Standing in Nomenclature (LPSN) (https://lpsn.dsmz.de/).

Relative abundance and global distribution of the dominant sponge-associated microbes

The relative abundance of a MAG was estimated as the percentage of mapped metagenomic reads with the removal of eukaryotic reads. First, metagenomic contigs with a length of >1,000 bp were subjected to EukRep v.0.6.7 to identify eukaryotic contigs (West et al., 2018). Then, the metagenomic reads were mapped to the eukaryotic contigs by Bowtie2 (Langdon, 2015), and the mapped reads were removed from the metagenomes. Finally, the remaining reads were subjected to coverM v0.2.0 (https://github.com/wwood/CoverM) to calculate the relative abundance of sponge-associated MAGs. To investigate the global distribution of a sponge symbiont, the V4 region of the 16S rRNA gene extracted from the MAG was searched by BLASTN (e-value cutoff: 1e-05) against the datasets of 16S rRNA gene sequences in the Sponge Microbiome Project (SMP) (Moitinho-Silva et al., 2017b) and the Deep-sea Sponge Microbiome Project database (D-SMP) (Busch et al., 2022). Target sequences with >96% identity were retained to calculate their relative abundance in respective sponges.

Genome annotation

Close relatives of symbiotic MAGs were obtained by the phylogenomic analysis of the GTDB-tk program, and their genome sequences were downloaded from the NCBI and Integrated Microbial Genomes (IMG) databases. Open reading frames (ORFs) of our MAGs and reference genomes were predicted by Prodigal v2.6.2 (Hyatt et al., 2012). The predicted genes were annotated by KofamScan v1.1.0 (Aramaki et al., 2020) against KEGG databases. Eukaryotic-like protein (ELP) domains were annotated by PfamScan script against the PFAM database (El-Gebali et al., 2019). ORFs were searched by BLASTP v2.5.0 against the COG_2019_v11.0 database to find genes encoding a transposase. Clustered regularly interspaced short palindromic repeats (CRISPRs) and Cas proteins were predicted using the online CRISPRminer2, a toolkit comprising CRISPRCasFinder, CRT, and PILER-CR programs (Zhang et al., 2018). Pairwise ANI values of MAGs were calculated by PyAni v0.2.10 (Pritchard et al., 2016) with the ANIb model. AAI values of MAGs were calculated by CompareM v0.0.23 (https://github.com/dparks1134/CompareM).

Phylogenetic analysis

The *coxI* (cytochrome c oxidase subunit I) gene of the deep-sea sponge sample was annotated from metagenome contigs by Prokka v1.13.7 (Seemann, 2014). 16S rRNA genes of symbiotic MAGs were predicted using the RNA_HMM program (Huang et al., 2009). For phylogenetic analysis of both *coxI* and 16S RNA genes, the targeted gene sequences were searched by the BLASTN program against the NCBI GenBank database to identify close relatives. The collected gene sequences were aligned using MAFFT v7.427 (Katoh and Toh, 2010) and trimmed using trimAl v1.4 (Capella-Gutiérrez et al., 2009). Phylogenetic trees were built using IQ-TREE v1.6.10 (Nguyen et al., 2015) with the "TIM2+F+I+G4" model. For phylogenomic analysis, the alignment of 43 concatenated conserved proteins deduced from our MAGs and reference genomes was produced using the CheckM program with default settings and further treated with trimAl v1.4 (Capella-Gutiérrez et al., 2009) to remove poorly aligned regions. The maximum likelihood (ML) tree was built using IQ-TREE v1.6.10 (Nguyen et al., 2015) with the "LG+F+R7" models. Bootstrap values of the trees were calculated based on 1,000 replicates.

Identification of phage-like assembled contigs

Sponge metagenomic contigs in size of >5,000 bp were imported into DeepVirFinder (Ren et al., 2020), VIBRANT (Kieft et al., 2020), VirFinder (Ren et al., 2017), VirSorter (Roux et al., 2015b), and VirSorter2 (Guo et al., 2021) to identify viral genomes. CheckV v0.8.1 (Nayfach et al., 2021) was used to assess the viral genome quality, to identify and remove potential host contamination in integrated proviruses, and to match closely relative viral genomes from public databases. Viral genome completeness was estimated by searching against a database that comprises 76,262 complete viral genomes from publicly available metagenomes, metatranscriptomes, and metaviromes by IMG/M (Chen et al., 2019), MGnify (Mitchell et al., 2020), and the study of the human microbiome (Nayfach et al., 2019) and ocean virome (Gregory et al., 2019). Taxonomic assignment of positive viral contigs was performed using vConTACT2 (Bin Jang et al., 2019), which was designed to cluster the protein sequences with a RefSeq database based upon shared protein clusters. The relative abundance of each viral contig in a deep-sea sponge sample was calculated by mapping clean reads to assembled contigs using BWA-MEN with default settings (Li and Durbin, 2010).

Results

Sponge morphology and taxonomic identification

A thick-walled tubular deep-sea sponge was collected from the South China Sea at a depth of 983 m. The top of the sponge body is a moderately sized, trumpet-shaped osculum. The body color is pale beige. Prostalia protrudes over the body by several centimeters (Figure 1A). Phylogenetic inference based on the *coxI* gene (1,575 bp in size) suggested that the sponge is closely related to *Bathydorus laniger* from the coast of California (Kahn et al., 2013) and *Bathydorus spinosus* in the Weddell Sea (Dohrmann et al., 2011), two known species of the genus *Bathydorus* in the glass sponge class Hexactinellida (Figure 1B). The *coxI* gene of our sponge sample shared an identity of 94.83 and 93.27% with its two *Bathydorus* relatives. The *coxI* genes had respective matching thresholds at order, family, and genus levels (45, 73, and 91%, respectively) based on a sponge identification protocol proposed by a previous study based on 37 sponge species belonging to 10 orders from South Australia (Yang et al., 2017). Thus, our sample should represent a novel sponge species in the genus *Bathydorus* and is preliminarily named *Bathydorus* sp. SQW35.

Prokaryotic composition in the sponge *Bathydorus* sp. SQW35

Taxonomic classification of 16S miTags demonstrated highly consistent prokaryotic communities in two technical replicates of the sponge *Bathydorus* sp. SQW35. The bacterial symbionts were mainly affiliated with the phyla Pseudomonadota (synonyms Proteobacteria) (Oren and Garrity, 2021), Planctomycetota, Nitrospirota, and Bacteroidota, and archaeal symbionts were dominated by the phylum Nitrososphaerota (synonyms Thermoproteota) (Supplementary Figure 1; Supplementary Table 1). Among them, the highest proportion belonged to Nitrososphaerota, which accounted for 81.81 ± 2.09% of the prokaryotic communities. Further taxonomic analysis revealed the dominance of the Nitrosopumilaceae family in Nitrososphaerota, known as AOA, which can obtain energy from ammonia oxidation and use CO_2 as a carbon source for chemoautotrophy (Wang et al., 2019). Gammaproteobacteria that mainly belonged to SOBs were the second dominant group in the sponge *Bathydorus* sp. SQW35 and accounted for 10.60 ± 0.85% of the prokaryotic communities. Alternatively, Nitrospirota that is involved in nitrite oxidation carbon fixation (Lücker et al., 2013) occupied 0.85 ± 0.11% of the prokaryotic communities.

Recovery of novel prokaryotic genomes

In total, 14 metagenome-assembled genomes (MAGs) with >50% estimated completeness and <10% contamination were obtained from the sponge metagenomes of the same sponge individual, and five of them were more than 90% complete (Table 1). The MAGs have a genome size ranging from 0.54 to 3.63 Mbp and GC content of 31.93–52.87%. Classification by the GTDB-tk program revealed that these MAGs were affiliated with the phyla Nitrososphaerota, Pseudomonadota, Nitrospirota, Bdellovibrionota, SAR324, Bacteroidota, and Patescibacteria (Supplementary Table 2). Relative evolutionary divergence (RED) that infers a phylogenetic distance between the last common ancestor (set to RED = 0) and all extant taxa (RED = 1) can be used to establish taxonomic ranks (Parks et al., 2018). RED values indicated that only MAG B05 fell into known species and the others likely represent novel species. Among these novel species, five MAGs have RED values of 0.71–0.89, indicative of novel genera. Three MAGs with RED values of 0.64–0.69 are suggested to present novel families (Parks et al., 2018).

Phylogenomic inference also revealed that the 14 symbiotic MAGs fell into seven phyla, and the far distance of these MAGs with known species further indicated their taxonomic novelty (Figure 2A). Five MAGs (B01–B05) fell into the clade composed of AOAs in the phylum Nitrososphaerota, whereas only B03 was closely related to sponge-associated AOAs. ANIs between B03 and its close relatives *Ca.* Nitrosopumilus sp. ESC from the sponge *Hymedesmia (Stylopus) methanophila* (Haber et al., 2020) and *Ca. Nitrosopumilus* sp. LS from the glass sponge *Lophophysema eversa* (Tian et al., 2016) were 90.40 and 90.86%, respectively (Supplementary Table 3). The affiliation of MAG B14 with sulfur-oxidizing Gammaproteobacteria Gsub from the cold seep sponge *Suberites* sp. (Tian et al., 2017) in the phylogenomic tree was in line with their ANI of 70.80%. Nitrite-oxidizing bacterium (NOB) MAG B08 in the phylum Nitrospirota was affiliated with the genus SPGG5 and was placed together with Nitrospinae bacterium UBA9942 (Parks et al., 2018), which shared an ANI of 67.37%. MAGs B12 and B11 were classified as the phylum Bdellovibrionota and shared the ANIs of 68.07 and 67.91%, respectively, with its relatives Bdellovibrionales RBG_16_40_8 (Kauffman et al., 2018) and Bdellovibrionales bacterium SXSP01 (Xing et al., 2020).

The relative abundances of the sponge symbionts in metagenomes of the sponge *Bathydorus* sp. SQW35 are summarized in Figure 2B. The AOA MAG B01 recruited 70.31 ± 1.15% of the prokaryotic metagenome reads and represented the sole dominant symbiont (Figure 2B; Supplementary Table 4). The SOB MAG 14 as the second most abundant symbiont was mapped by 2.83 ± 0.40% of the prokaryotic metagenomic reads. The NOB MAG B08 could also be mapped with less abundant metagenomic reads. Relative abundances of the representative MAGs including AOA B01, SOB B14, and NOB B08 were consistent with the result of the 16S miTags analysis. In contrast, the Bdellovibrionota MAGs B11 and B12 that accounted for 0.29 ± 0.29% and 0.07 ± 0.07% metagenomic reads were not detected in 16S miTags analysis.

A specialist sponge AOA symbiont lineage

Our detailed phylogenomic inference revealed that the dominant symbiont AOA B01 was closely related to the *Cenarchaeum* clade that is now composed of four lineages including *Cenarchaeum symbiosum* A from the demosponge *Axinella mexicana*, *Cenarchaeum* sp. bin74s, *Cenarchaeum* sp. bin90o, and *Cenarchaeum* sp. bin143o (phylum: Nitrososphaerota) from *Vazella pourtalesii*, Nitrososphaerota archaeon TS (phylum: Nitrososphaerota) from *Theonella swinhoei*, and *Candidatus* Cenporiarchaeum stylissum S13–S15 (phylum: Nitrososphaerota) from *Stylissa flabelliformis* (Figure 3). ANIs between B01 and these reference genomes range from 68.96 to

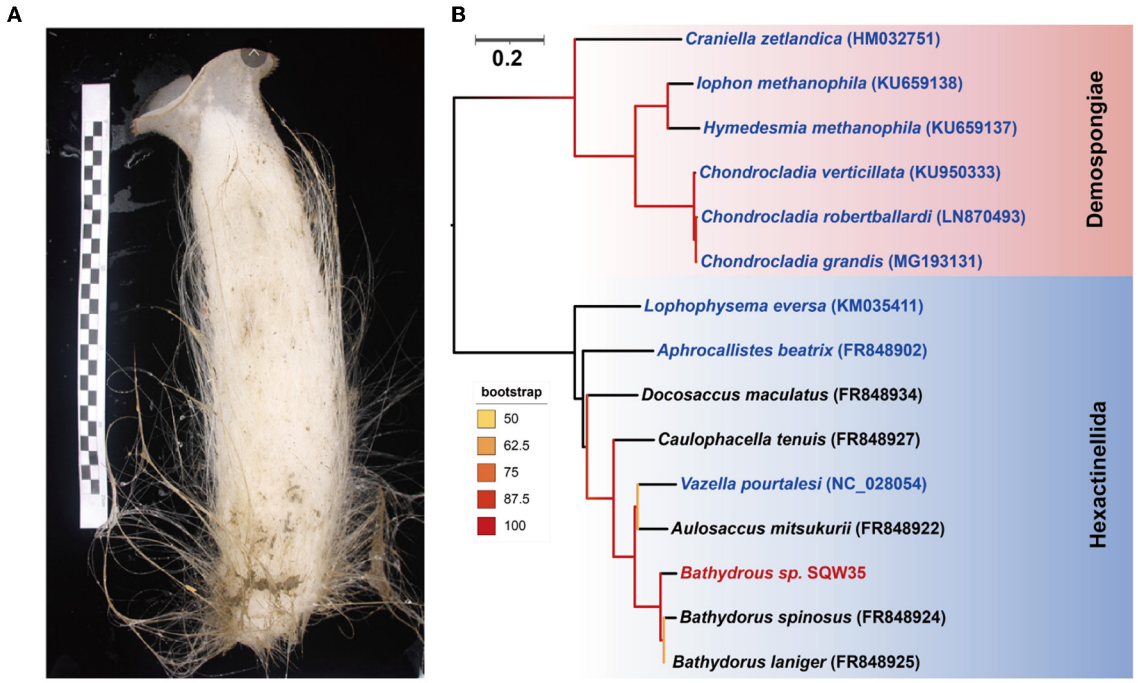


FIGURE 1
New glass sponge from the South China Sea. **(A)** A photograph of the sponge *Bathydorus* sp. SQW35. The left scale bar unit is 10 mm, and the body length is about 350 mm. **(B)** The *coxI*-based maximum-likelihood phylogenetic tree was constructed using the "TIM2+F+I+G4" model. The scale bar represents 0.2 substitutions per nucleotide position.

TABLE 1 General genomic features of the 14 sponge-associated prokaryotic MAGs.

MAG	Phylum[a]	Genome size (Mbp)	N_{50} (kb)	No. of contigs	No. of CDSs	% GC	Compl. (%)[b]	Contam. (%)[b]	RED[c]
B01	Nitrososphaerota	1.45	0.23	100	1,692	33.92%	98.54	0	0.87
B02	Nitrososphaerota	1.61	0.21	124	2,027	33.00%	99.03	0.97	0.98
B03	Nitrososphaerota	1.00	0.47	37	1,206	31.93%	75.73	0.07	0.99
B04	Nitrososphaerota	1.65	0.28	97	2,119	32.16%	95.63	2.91	0.99
B05	Nitrososphaerota	0.54	0.04	145	758	35.34%	56.15	0.97	N/A
B06	SAR324	1.49	2.84	10	1,321	40.31%	87.96	0	0.51
B07	Pseudomonadota	1.09	1.27	16	1,054	44.39%	84.62	0	0.64
B08	Nitrospirota	3.63	0.14	358	4,177	49.76%	97.67	1.82	0.96
B09	Bacteroidota	3.61	0.04	1,017	3,515	35.26%	62.46	1.76	0.69
B10	Patescibacteria	0.44	0.09	61	478	32.72%	69.49	0	0.74
B11	Bdellovibrionota	1.24	0.96	17	1,144	32.44%	90.12	0	0.80
B12	Bdellovibrionota	1.67	0.05	343	1,741	35.41%	86.14	2.68	0.67
B13	Pseudomonadota	1.88	0.13	165	1,908	52.87%	79.49	4.69	0.71
B14	Pseudomonadota	1.62	0.38	73	1,576	35.34%	89.94	0.61	0.89

[a]The nomenclature of the MAG was based on the List of Prokaryotic names with Standing in Nomenclature (LPSN) (https://lpsn.dsmz.de/).
[b]Genome completeness (Compl.) and contamination (Contam.) were estimated by CheckM (Parks et al., 2015).
[c]RED, relative evolutionary divergence indicates taxonomic novelty (Chaumeil et al., 2019).

69.97% (Supplementary Table 3). With its RED value of 0.87 (Table 1), MAG B01 should represent a new genus. Global distribution analysis against the SMP database (Moitinho-Silva et al., 2017b) revealed that the V4 region of the 16S rRNA gene in MAG B01 showed 100% identity with amplicons from three glass sponge individuals in the family Dactylocalycidae (Figure 4), with relative abundances of 23.25, 8.78, and 8.60% in the respective communities (Supplementary Table 5). B01 also shared 97–99%

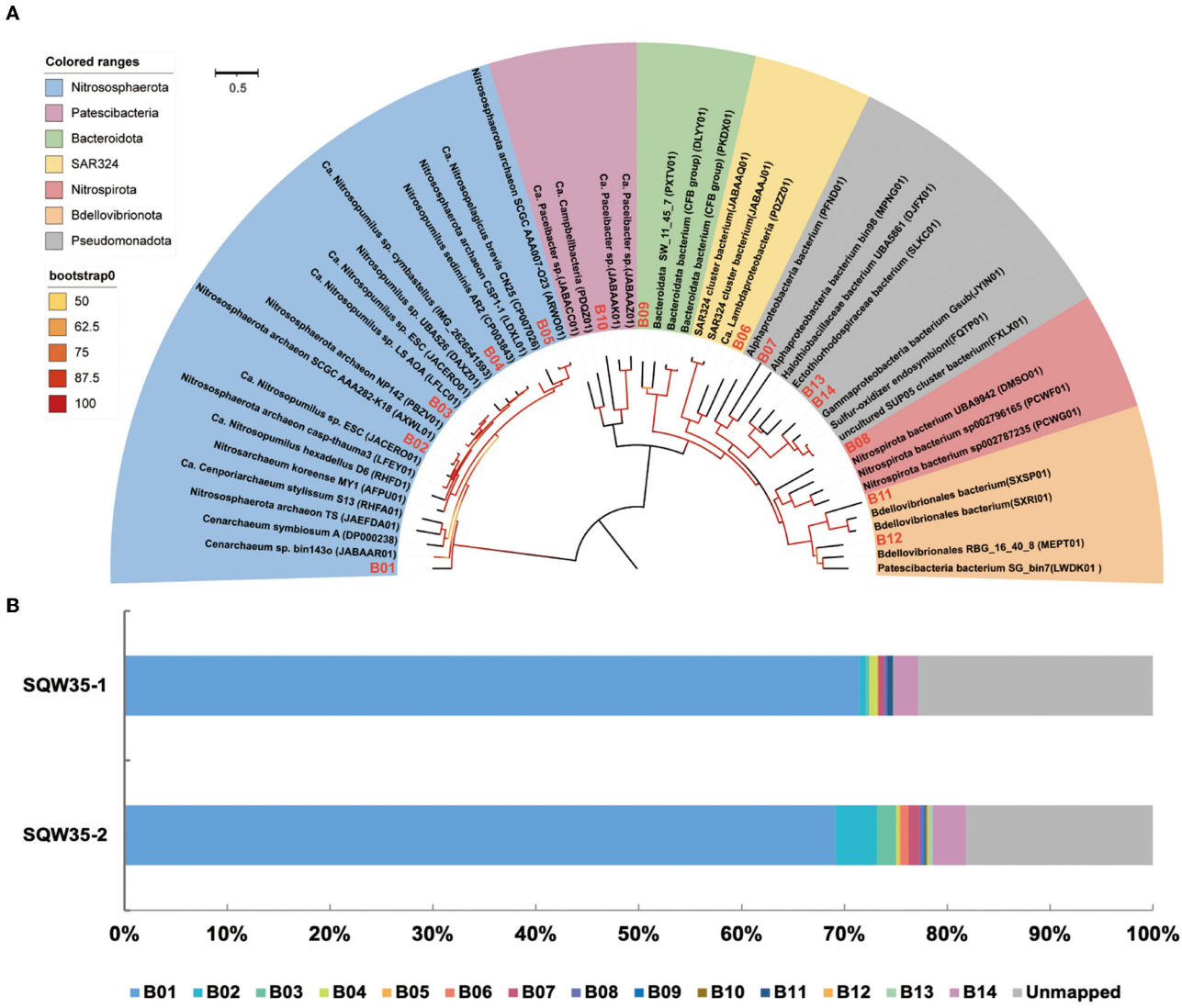


FIGURE 2
Phylogenomic tree of sponge-associated MAGs and their relative abundance. **(A)** Phylogenomic maximum-likelihood tree of the 14 sponge-associated MAGs was constructed using the "LG+F+R7" model. Sponge-associated MAGs of *Bathydorus* sp. SQW35 are marked with red in the tree. All the reference genomes are listed in Supplementary Table 12. **(B)** Relative abundance of sponge-associated MAGs in the metagenomes of *Bathydorus* sp. SQW35. The relative abundance of MAGs was calculated with coverM v0.2.0 using metagenome reads after eliminating reads assigned to eukaryotic contigs.

identities with sponge symbionts in three other sponge species, *Aphrocallistes beatrix*, *Axinella* sp., and an unnamed species, but was only 96% similar to its closest relatives from seawater samples (Supplementary Table 5). By a query against the D-SMP database (Busch et al., 2022), B01 shared more than 99% identity with symbionts in *Pheronema carpenteri*, *Bathydorus* sp., *Aphrocallistes beatrix*, *Saccocalyx tetractinus*, *Rossellinae indet*, and *Lophocalyx* sp. The relative abundances of B01 in these sponges span between 0.007 and 0.657% (Supplementary Table 5). Because the D-SMP database employed bacterial primers for amplification, the relative abundance of archaeal relatives of MAG B01 in this database is likely underestimated. Sponge symbionts can be classified into generalists (found in a wide range of sponge species) and specialists (living in a small number of sponge species) according to their distribution patterns (Haber et al., 2020). In brief, the distribution patterns against both the SMP and D-SMP databases indicate that AOA B01 is a specialist.

Bdellovibrionota predators as sponge symbionts

Phylogenomic inference showed that both MAGs B11 and B12 fell in the Bdello-group2 in the phylum Bdellovibrionota but form two separate deep branches (Figure 5). Bdello-group2 comprised of the representative *Bdellovibrio* predators including *Bdellovibrio exovorus* JSS, *Bdellovibrio bacteriovorus* HD100, *Bdellovibrio bacteriovorus* Tiberius, and *Bdellovibrio bacteriovorus* 109J (Li et al., 2021), suggesting that B11 and B12 may also represent

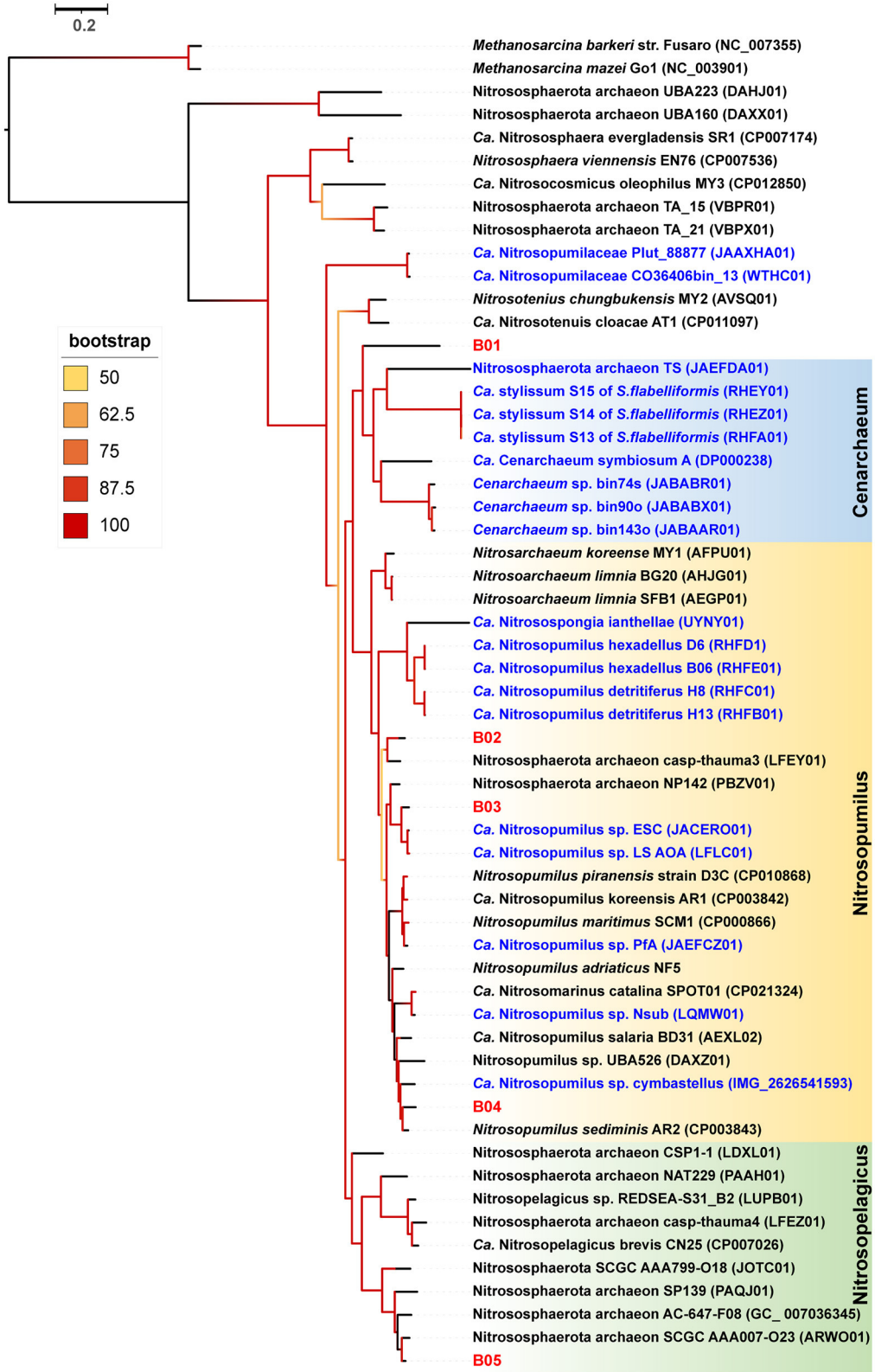


FIGURE 3
Phylogenomic tree of the sponge-associated AOA MAGs. The phylogenomic maximum-likelihood tree of sponge-associated AOA MAGs was constructed using the "LG+F+R7" model. *Bathydorus* sp. SQW35 sponge-associated MAGs are marked with red. Sponge-associated reference MAGs from other environments are marked with blue. *Methanosarcina barkeri* and *Methanosarcina mazei* Go1 (p_Euryarchaeota) are used as outgroups. All the reference genomes are listed in Supplementary Table 12.

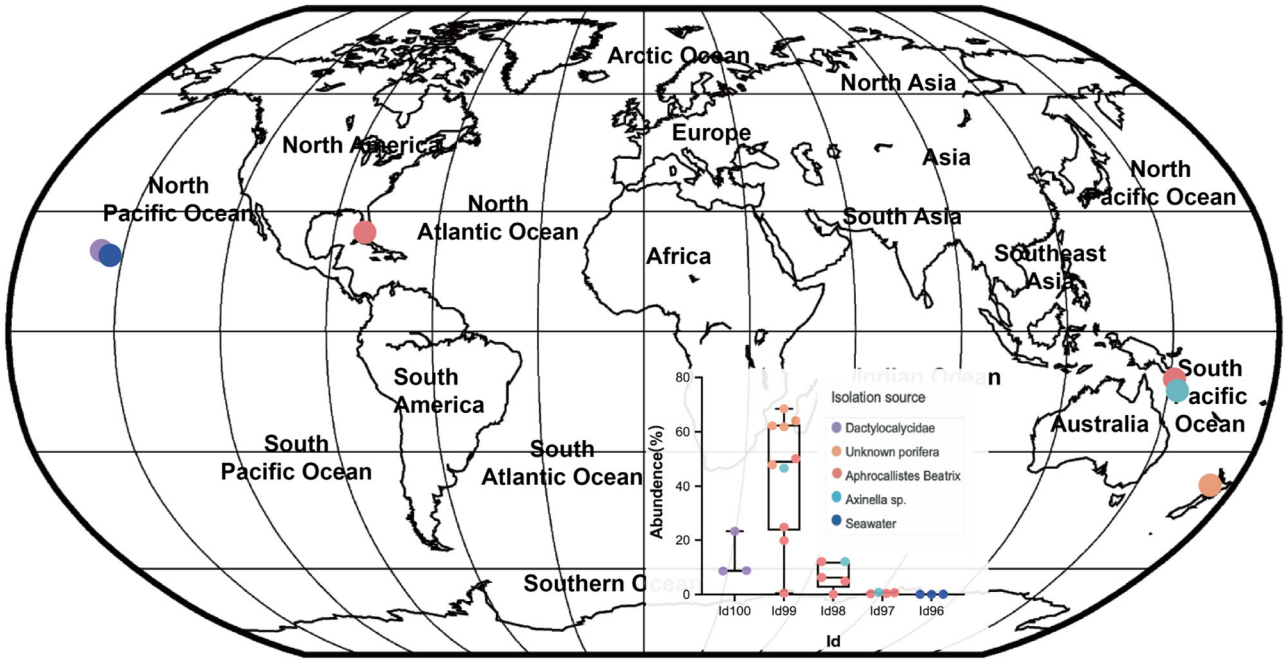


FIGURE 4
Global distribution of the sponge symbiont Nitrososphaerota B01. To investigate the distribution of B01 in sponges and surrounding environmental samples, its 16S rRNA gene sequence as a query was searched against the Sponge Microbiome Project (SMP) using BLASTN, and the target sequences with identity values of >96% were selected and assigned to Id100 (identity = 100%), Id99 (identity ≥ 99% but < 100%), Id98, Id97, and Id96 groups. The relative abundance of the B01 relatives in the SMP samples was calculated. Target sponge species and environmental samples were labeled on the world map.

bacterial predators. The Bdellovibrionota MAGs B11 and B12 have remarkably smaller genome sizes and lower GC contents than those free-living Bdellovibrionota species (Supplementary Figure 2). The comparative genomic analysis further revealed that the free-living Bdellovibrionota bacteria harbored genes encoding chemotaxis proteins and mobility systems (flagellum and type IV pili), yet MAGs B11 and B12 lacked these genes (Supplementary Table 6). Predatory *Bdellovibrio* bacteria invade the periplasm of bacterial prey cells by penetrating the peptidoglycan layer to form transient structures there. Penicillin-binding proteins (PBPs) are important for Bdellovibrionota to lyze the cell wall of a prey bacterium (Lerner et al., 2012). Annotation analysis showed that symbiotic Bdellovibrionota MAGs lack the PBP1C coding gene compared to free-living relatives but retain PBP1A, PBP1B, PBP2, PBP3, and PBP4 coding genes (Supplementary Table 6), which suggests that MAGs B11 and B12 may still retain the predatory potential.

Carbon, nitrogen, and sulfur metabolisms of the sponge symbionts

We searched the three main prokaryotic autotrophic carbon fixation pathways: 3-hydroxypropionate/4-hydroxybutyrate (3HP/4HB) cycle, reductive citric acid (rTCA) cycle, and the Calvin–Benson–Bassham (CBB) cycle. The 3HP/4HB cycle was identified in sponge-associated AOA (family *Nitrosopumilaceae*) with the presence of key genes encoding 3-hydroxypropionyl-CoA synthetase (K18594) and 4-hydroxybutyrate-CoA ligase (K18593) (Supplementary Table 7). Almost all genes of the rTCA cycle were identified in Nitrospirota B08 (family UBA8639, genus SPGG5), including those coding for citrate synthase (K01647), fumarate hydratase (K01676), and succinate dehydrogenase (K00240), which were considered as evidence for the rTCA cycle. Furthermore, more than 90% of the genes for carrying out the CBB cycle including those encoding subunits of ribulose-bisphosphate carboxylase (K01602 and K01601) were found in the SOB MAGs (class Gammaproteobacteria). These findings indicate the existence of various carbon fixation pathways in the microbial consortium of the sponge *Bathydorus* sp. SQW35.

The *amoABC* genes (K10944, K10945, K10946) encoding ammonia monooxygenase subunits were identified in three Nitrososphaerota MAGs (B01, B04, and B05) but not in any bacterial MAGs (Supplementary Table 7). This suggests that AOA is the only symbiotic group in charge of ammonia oxidation in the sponge *Bathydorus* sp. SQW35. Nitrospirota B08 had the gene encoding nitrate/nitrite transporter (K02575) but lacked genes encoding nitrite oxidoreductase. In addition, the B08 has a gene set involving urea decomposition that produces ammonia for nitrification and biomass production, where urea would be imported from the extracellular milieu by the urea ABC transporter urtABCDE (K11959, K11960, K11961, K11962, and K11963) and be hydrolyzed by urease ureABC (K01428, K01429, and K01430) (Supplementary Table 7). These results further imply the potential role of NOB B08 in urea utilization.

Microbial sulfur oxidation is frequently reported in sponges to remove toxic sulfide produced by other bacterial symbionts that use sulfate as an electron acceptor (Tian et al., 2014). Gammaproteobacteria SOB B13 and B14 encode adenylyl-sulfate reductase A (K00394), adenylyl-sulfate reductase B (K00395), and sulfate adenylyltransferase (K00958) to oxidize sulfite to sulfate

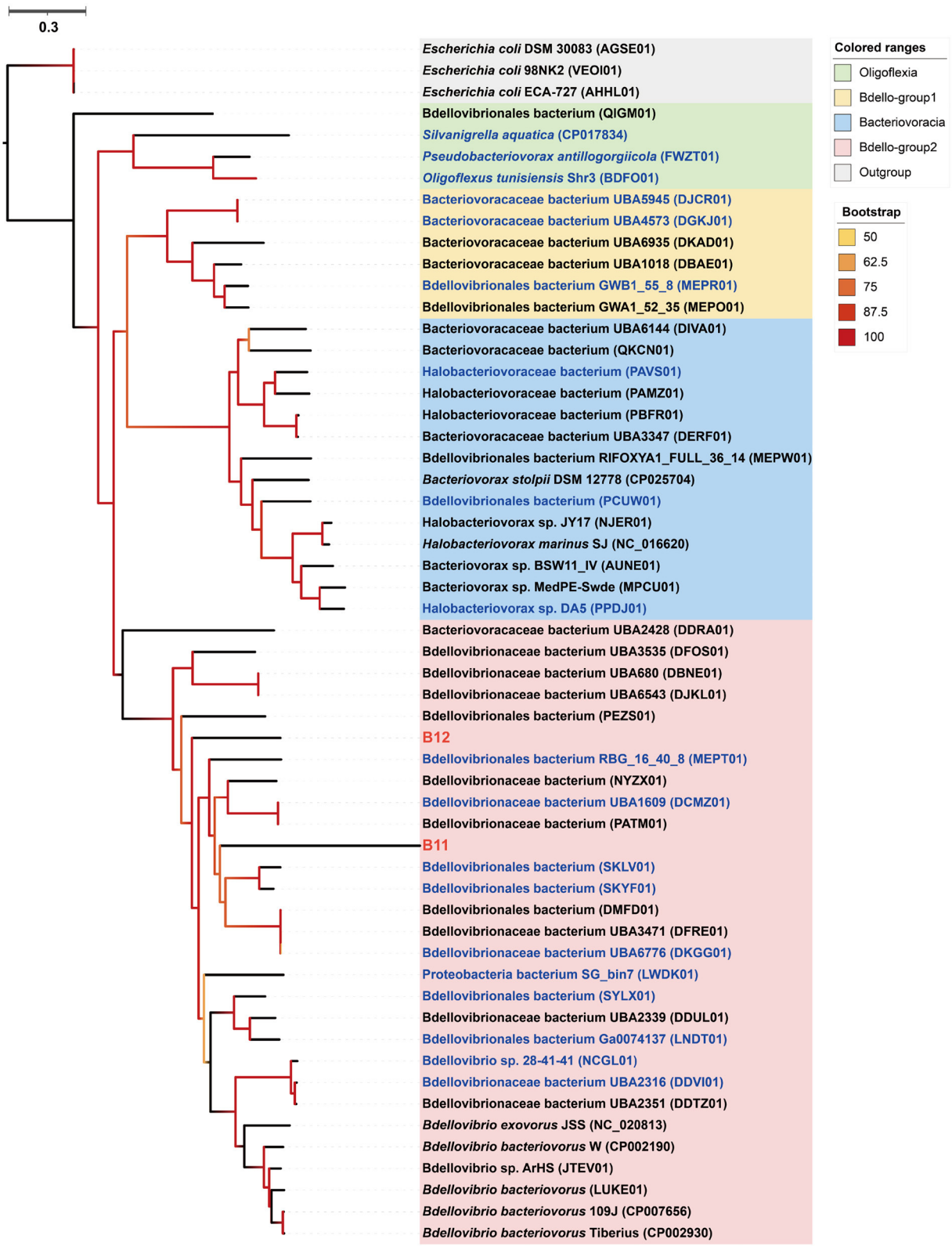


FIGURE 5
Phylogenomic tree of the sponge-associated Bdellovibrionota MAGs. The maximum-likelihood tree was constructed using the "LG+F+R7" model. *Escherichia coli* DSM 30083, *Escherichia coli* 98NK2, and *Escherichia coli* ECA-727 were used as outgroups. Sponge-associated Bdellovibrionota MAGs B11 and B12 were colored in red. Free-living Bdellovibrionota MAGs with completeness > 90% were colored in blue.

(Supplementary Table 7). SOB B14 also contains the *soxAXYZ* gene clusters (K17222, K17223, K17226, and K17227) that code for the SOX complex, confirming potential capacity for sulfur oxidation (Supplementary Table 7). Previous studies showed that taurine is a natural product of sponges, and sponge symbionts can import and utilize taurine as suggested by the presence of ABC transporter

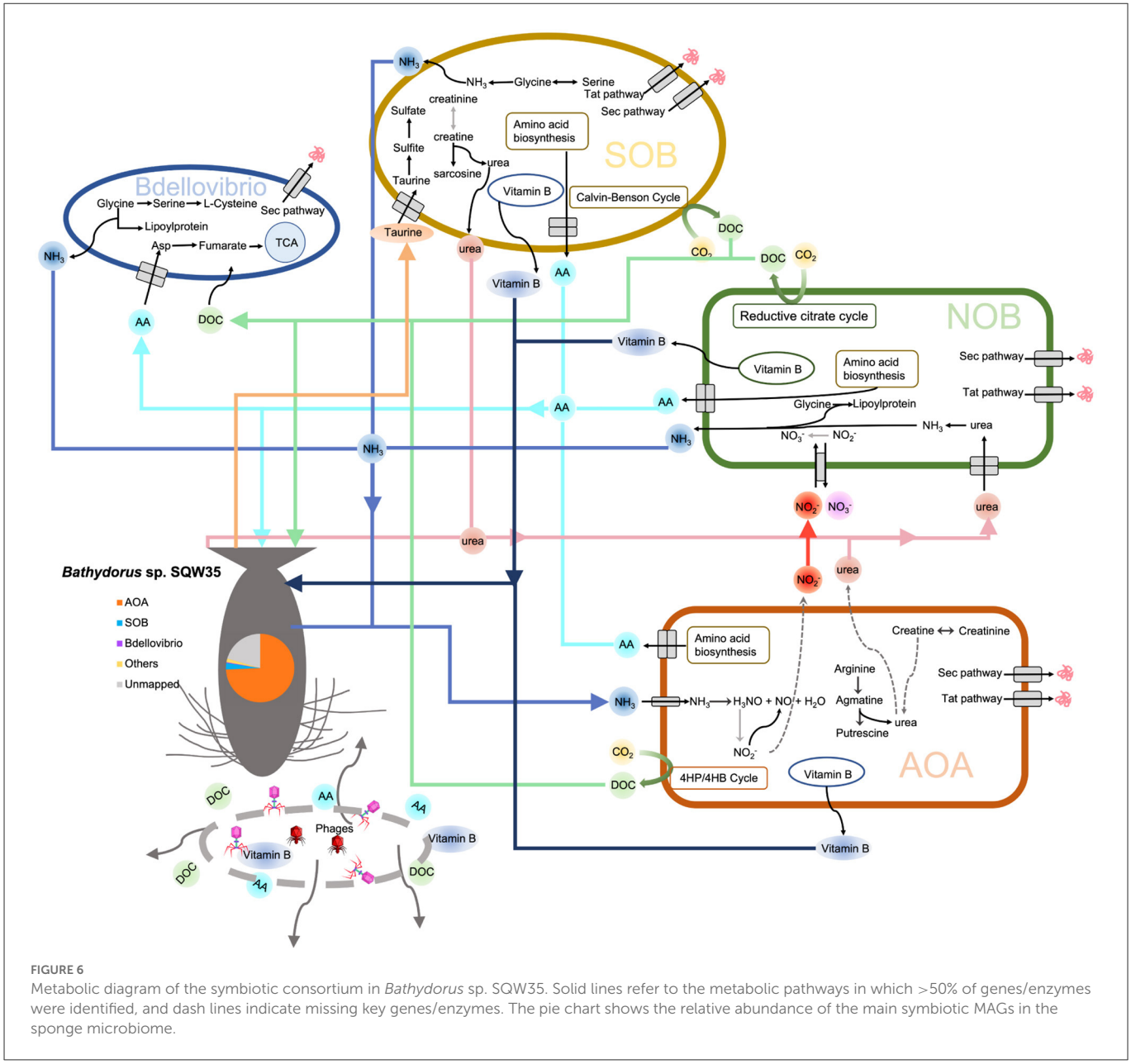


FIGURE 6
Metabolic diagram of the symbiotic consortium in *Bathydorus* sp. SQW35. Solid lines refer to the metabolic pathways in which >50% of genes/enzymes were identified, and dash lines indicate missing key genes/enzymes. The pie chart shows the relative abundance of the main symbiotic MAGs in the sponge microbiome.

genes (*tauABC*) (Emura et al., 2006; Karimi et al., 2018). The SOB MAGs B13 and B14 also have genes encoding taurine ABC transporter (*tauAC*; K15551 and K15552), taurine-pyruvate aminotransferase (K03851), and sulfoacetaldehyde acetyltransferase (K03852), which can catalyze taurine to sulfite for further oxidation by the SOX complex.

Eukaryotic-like proteins encoded by the sponge symbionts

We analyzed the distribution of ELPs including ankyrin repeats (Ank), tetratricopeptide repeats (TPRs), NCL-1, HT2A and Lin-41 repeats (NHL), fibronectin type III (Fn3) and cadherin (CAD) CUB, bacterial Ig-like domain (Big), WD40, and pyrroloquinoline quinone repeat (PQQ) in the deep-sea glass sponge microbiome (Supplementary Table 8). We found that five of the nine bacterial MAGs encoded Ank domains, including SOB B14 and NOB B08. Notably, Bdellovibrionota B11 has 12 Ank domains, which was the highest number among these bacterial MAGs. However, none of the AOA MAGs had Anks. The bacterial and archaeal MAGs, except B10, had TPR domains, especially in B08 and B09. As the most prevalent symbiont, AOA B01 encoded the maximum number of NHL domains. NOB B08 also showed the richness of NHL domains ($n = 13$).

CRISPR–Cas systems in the sponge symbionts

One major strategy of prokaryotic defense against phages is dependence on the Clustered Regularly Interspaced Short

Palindromic Repeats (CRISPR)–Cas system (Makarova et al., 2013). Here, all the MAGs were analyzed for their CRISPR–Cas systems. Almost all the symbiotic MAGs in the sponge *Bathydorus* sp. SQW35 has genes encoding Cas proteins (Supplementary Figure 3A). CRISPRminer2 analysis further indicated that CRISPR arrays could be found in six MAGs (B01, B08, B09, B11, B13, and B14), which represented almost all the dominant microbial groups (Supplementary Figure 3B; Supplementary Table 9). Interestingly, the dominant symbiont AOA B01 has a CRISPR array with the highest number of spacers ($n = 85$). The second abundant symbiont SOB B14 also has a CRISPR array with 81 spacers, followed by a CRISPR array in Bdellovibrionota B11 consisting of 51 spacers.

Phage diversity of the sponge microbiome

To understand the high prevalence of CRISPR–Cas systems in the dominant symbiotic inhabitants, we analyzed the phages in the deep-sea glass sponge *Bathydorus* sp. SQW35. In total, 125 contigs in SQW35-1 and 148 contigs in SQW35-2 were predicted as potential phage genomes (>5,000 bp) (Supplementary Table 10). The potential viral contigs of SQW35-1 and SQW35-2 were predicted to harbor 237 and 236 viral genes by CheckV (Nayfach et al., 2021), respectively. There are 232 positive virus sequences after redundancy removal of the contigs, which indicated a large phage community in the sponge *Bathydorus* sp. SQW35. Among them, the most abundant phage fragment accounted for 1.48 and 2.19% of metagenome reads, respectively, in the two replicate samples (Supplementary Table 11).

Discussion

While shallow-water sponge holobionts have been extensively studied in the last two decades, our knowledge about deep-sea sponges and their associated microbes remains rare. The present study collected a new deep-sea sponge species affiliated with the genus *Bathydorus* (*Bathydorus* sp. SQW35) from the South China Sea. Both 16S miTags-based prokaryotic community analysis and relative abundance analysis of MAGs demonstrated an AOA-dominant microbiome. AOAs together with SOBs and NOBs comprise the key prokaryotic players involved in carbon, sulfur, and nitrogen metabolism in the sponge holobiont, which is consistent with previous findings of prokaryotic communities in a large number of deep-sea sponge species (Busch et al., 2022). In total, 14 prokaryotic MAGs were successfully retrieved from the sponge microbiome, and 13 MAGs belong to potential new species, indicating the novelty of prokaryotic genome resources in deep-sea sponges (Wang et al., 2022).

Ammonia oxidation plays a significant role in the nitrogen cycle in marine. AOAs in the phylum Nitrososphaerota have key roles in the global nitrogen cycle and are also widely found in sponge holobionts (Moitinho-Silva et al., 2017a; Zhang et al., 2019b; Steinert et al., 2020). *Cenarchaeum symbiosum* A is the first reported sponge AOA symbiont in the family *Nitrosopumilaceae* (Hallam et al., 2006). Recent genomic work has expanded the *Cenarchaeum* clade to four lineages (Zhang et al., 2019b; Bayer et al., 2020; Haber et al., 2020). Till now, this clade has only been reported in sponge holobionts and likely represents a sponge-specific cluster (Preston et al., 1996; Schleper et al., 1998; Hallam et al., 2006; Bayer et al., 2020). The dominant AOA MAG B01 of *Bathydorus* sp. SQW35 was closely related to the *Cenarchaeum* clade but probably represents a novel species in a new genus, which thus discloses the 50 lineages of the *Cenarchaeum* clade and expands the diversity of sponge AOA symbionts. This finding of novel Nitrososphaerota symbionts helps to uncover unique adaptation mechanisms of AOAs to sponge hosts. Sponges can excrete ammonia as a metabolic waste, which produces an ammonia-rich microenvironment in their bodies, and thus benefits the nitrification process (Moitinho-Silva et al., 2017a). AOA B01 is the sole dominant symbiont of *Bathydorus* sp. SQW35 and its relative abundance are much higher than those AOAs associated with shallow-water and even deep-sea sponges. Thus, AOA B01 likely plays a much more important role in ammonia oxidization in this deep-sea sponge species. Meanwhile, the dependency of this deep-sea sponge on ammonia oxidization for ammonia elimination and nutrition supply is likely much higher than shallow-water sponges.

Bdellovibrio (in the phylum Bdellovibrionota) and *Bdellovibrio*-like bacteria (BALOs) are gram-negative bacterial predators living in various environments (Jurkevitch et al., 2000; Sockett, 2009) and mainly consists of four groups: Bacteriovoracia, Oligoflexia, Bdello-group1, and Bdello-group2 (Li et al., 2021). Till now, sponge-associated BALO had only been reported in the shallow-water sponge *Cymbastela concentrica*, and this predator was proposed to live with cyanobacteria (Tian et al., 2017). The MAGs B11 and B12 of the sponge *Bathydorus* sp. SQW35 are affiliated with the Bdello-group2 in the Bdellovibrionota, providing additional examples of BALOs with a possible role of sponge symbionts. Genomic reduction is common for microbial symbionts when some genes are no longer required during adaptive evolution (McCutcheon and Moran, 2011). Low GC content correlates with symbiotic evolution because of mutational bias in symbionts (Wernegreen and Funk, 2004; Bohlin et al., 2020). The smaller genome size and lower GC content of MAGs B11 and B12 compared to their free-living relatives also support their symbiotic lifestyle indicative of reductive genome evolution. Motility machines are needed for the free-living predatory Bdellovibrionota to capture the bacterial prey (Lambert et al., 2009). The loss of genes encoding motility machines in MAGs B11 and B12 indicates that the sponge Bdellovibrionota symbionts do not need motility after their symbiotic evolution. Although the presence of PBP genes involved in predator roles suggests the predatory potential of both MAGs B11 and B12, their predator roles and prey target in the sponge *Bathydorus* sp. SQW35 remains unclear and needs further exploration.

Eukaryotic-like protein coding genes in prokaryotic genomes are markers of symbionts with a long history of coevolution with their hosts (Cazalet et al., 2004). ELPs such as Anks and TPRs are commonly enriched in sponge-associated metagenomes and are perceived to be important for the interaction between symbionts and their hosts and for helping the symbiont to avoid decomposition by the host (Díez-Vives et al., 2017). Heterologous expression of Anks from sponge symbionts has even been proven to be capable of avoiding phagocytosis of amoeba (Nguyen et al., 2014). Most of the bacterial MAGs including Bdellovibrionota B11 and SOB B14 encoded Anks, speculating their symbiotic role in *Bathydorus* sp. SQW35. Consistent with previous studies (Zhang et al., 2019b; Haber et al., 2020), none of our AOA MAGs encoded Anks; thus, their symbiotic interaction mechanism with the host would be different. TPRs mediate bacterium–eukaryote interactions and have been revealed to allow entry of *Legionella pneumophila*

into epithelial cells and to regulate exopolysaccharide biosynthesis in *Pseudomonas aeruginosa* (Mittl and Schneider-Brachert, 2007). NHL was previously found to be enriched in sponge metagenomes (Fan et al., 2012), which were found in serine/threonine protein kinases and were indicated to affect phagosome processing in sponge symbionts (Reynolds and Thomas, 2016). The distribution patterns of these ELPs are discrepant among symbiotic MAGs of *Bathydorus* sp. SQW35, suggesting that symbionts employed different strategies to interact with their hosts.

Phages act as important prokaryotic killers in the oceans. They dominate seawater viral communities and lyse 20–50% of marine surface bacteria per day (Fuhrman, 1999). Sponges are likely exposed to high-concentration phages because of their filtration activity (Pascelli et al., 2020). Recent studies have illustrated diverse phage sponge-associated phage communities and tripartite sponge–phage–bacterium interplays (Pascelli et al., 2020; Jahn et al., 2021; Carrier et al., 2022). Our study also showed the involvement of diverse phages in the microbiome of the deep-sea sponge *Bathydorus* sp. SQW35. According to the virome analysis of coral and sponge-associated viruses, the virome communities were dominated by double-stranded DNA (dsDNA) bacteriophage of the order Caudovirales and a diverse community of single-stranded DNA (ssDNA) viruses of the family Microviridae (Laffy et al., 2018). However, our deep-sea sponge-associated phages could not be identified as any known viral taxa, and all predicted phage proteins were unknown in function. We suppose that deep-sea sponges contained viruses and phages of unknown taxonomy, which prompts us to investigate with further efforts. Because no reference seawater samples were analyzed in parallel, it must be declared that those phages may come from the ambient seawater.

For potential prokaryotic defense strategy against phages, genomic analyses highlight the distribution of CRISPR–Cas systems in 14 symbiotic MAGs. The dominant symbiont AOA B01 has the highest spacers ($n = 85$), and the SOB B14 comes ranks second ($n = 81$). A previous study has shown that containing more spacers in CRISPR–Cas systems could maximize the prokaryotic cell survival rate (Martynov et al., 2017). The highly complex CRISPR array might help the sponge symbionts gain a competitive advantage because of their efficient defense against phages. These findings suggest that symbiont defense mechanisms have evolved in the context of the sponge holobiont to maintain their dominant symbiont status.

Based on the earlier findings, we propose a metabolic network of the microbial consortium in the sponge *Bathydorus* sp. SQW35 (Figure 6). Sponges could obtain nutrients by feeding on both exogenetic carbon sources (Wilkinson, 1983) and autotrophically fixed carbon by their microbial symbionts such as AOAs, NOBs, and SOBs. By autotrophic CO_2 fixation, symbionts of the deep-sea sponge provide the necessary primary metabolites and continuously stable carbon sources that enabled the sponge host to adapt to the diverse deep-sea environment. Compared with the AOAs, NOBs, and SOBs that are involved in the elimination of toxic ammonia and sulfide waste produced by the sponge, *Bdellovibrio* predators acquire organic carbons from the autotrophies and/or sponge host for a parasitic lifestyle. In addition, phages can probably infect all the sponge symbiotic inhabitants and break down the cells to release organic carbon to nourish the sponge host. This might also regulate the population size of the microbial inhabitants in sponges as suggested previously (Jahn et al., 2021).

Conclusion

In this study, we report a novel deep-sea glass sponge species *Bathydorus* sp. SQW35 inhabiting the South China Sea and describe an AOA-dominant sponge microbiome. We uncovered a previously undescribed AOA species B01 dominating the microbiome and highlighted the Bdellovibrionota predators B11 and B12 as sponge symbionts undergoing reductive genome reduction. We further revealed the roles of CRISPR–Cas systems and ELPs in adaptive evolution to the deep-sea sponge holobiont and the involvement of phages in the symbiotic network. In summary, our results explored the symbiotic diversity, evolutionary adaptation, and symbiotic network in a deep-sea sponge holobiont. This cumulative knowledge base is a reference for subsequent research on the origin of deep-sea life. However, because of sampling limitations for deep-sea inhabitants and environment samples, we do not have multiple sponge individuals and ambient seawater and sediment samples. Whether all the analyzed microbes and phages are inherent to the sponge *Bathydorus* sp. SQW35 requires further investigations with regard to the distribution and generalization of both virome and prokaryotic microbiome.

Data availability statement

Metagenome data and MAGs obtained for the deep-sea sponge *Bathydorus* sp. SQW35 have been submitted to the NCBI database under BioProject PRJNA871057.

Author contributions

Z-MG and YW conceived the project. Z-MG and T-SW designed the experiments. L-SH collected the samples and took photos. LG described and identified the sponge samples. T-SW and H-GC extracted the DNA and performed the analyses. Q-ML and Y-LZ analyzed the data. T-SW drafted the manuscript. All authors edited the manuscript and approved its submission.

Funding

This study was supported by the National Natural Science Foundation of China (42076136), the National Key R&D Program of China (2022YFC2805505), and the Key Science and Technology Program of Hainan Province (ZDKJ2019011).

Acknowledgments

The authors are grateful to the cruise members of TS-7 and the research vessel (*R/V*) *Tansuoyihao* submersible pilots for their efforts in sampling. We thank the Supercomputing Center of the University of Sanya for providing the computation assistance.

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.2023.1078171/full#supplementary-material

References

Aramaki, T., Blanc-Mathieu, R., Endo, H., Ohkubo, K., Kanehisa, M., Goto, S., et al. (2020). KofamKOALA: KEGG Ortholog assignment based on profile HMM and adaptive score threshold. *Bioinformatics* 36, 2251–2252. doi: 10.1093/bioinformatics/btz859

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

Bayer, K., Busch, K., Kenchington, E., Beazley, L., Franzenburg, S., Michels, J., et al. (2020). Microbial strategies for survival in the glass sponge *Vazella pourtalesii*. *mSystems* 5, e00473–e00420. doi: 10.1128/mSystems.00473-20

Bell, J. J. (2008). The functional roles of marine sponges. *Estuar. Coast. Shelf Sci.* 79, 341–353. doi: 10.1016/j.ecss.2008.05.002

Bin Jang, H., Bolduc, B., Zablocki, O., Kuhn, J. H., Roux, S., Adriaenssens, E. M., et al. (2019). Taxonomic assignment of uncultivated prokaryotic virus genomes is enabled by gene-sharing networks. *Nat. Biotechnol.* 37, 632–639. doi: 10.1038/s41587-019-0100-8

Bohlin, J., Rose, B., Brynildsrud, O., and De Blasio, B. F. (2020). A simple stochastic model describing genomic evolution over time of GC content in microbial symbionts. *J. Theor. Biol.* 503, 110389. doi: 10.1016/j.jtbi.2020.110389

Busch, K., Slaby, B. M., Bach, W., Boetius, A., Clefsen, I., Colaco, A., et al. (2022). Biodiversity, environmental drivers, and sustainability of the global deep-sea sponge microbiome. *Nat. Commun.* 13, 5160. doi: 10.1038/s41467-022-32684-4

Capella-Gutiérrez, S., Silla-Martínez, J. M., and Gabaldón, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25, 1972–1973. doi: 10.1093/bioinformatics/btp348

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

Carrier, T. J., Maldonado, M., Schmittmann, L., Pita, L., Bosch, T. C. G., and Hentschel, U. (2022). Symbiont transmission in marine sponges: reproduction, development, and metamorphosis. *BMC Biol.* 20, 1–19. doi: 10.1186/s12915-022-01291-6

Cazalet, C., Rusniok, C., Bruggemann, H., Zidane, N., Magnier, A., Ma, L., et al. (2004). Evidence in the *Legionella pneumophila* genome for exploitation of host cell functions and high genome plasticity. *Nat. Genet.* 36, 1165–1173. doi: 10.1038/ng1447

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

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

Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34, i884–i890. doi: 10.1093/bioinformatics/bty560

de Goeij, J. M., van Oevelen, D., Vermeij, M. J. A., Osinga, R., Middelburg, J. J., de Goeij, A. F. P. M., et al. (2013). Surviving in a marine desert: the sponge loop retains resources within coral reefs. *Science* 342, 108–110. doi: 10.1126/science.1241981

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

Dohrmann, M., Haen, K. M., Lavrov, D. V., and Wörheide, G. (2011). Molecular phylogeny of glass sponges (Porifera, Hexactinellida): increased taxon sampling and inclusion of the mitochondrial protein-coding gene, cytochrome oxidase subunit I. *Hydrobiologia* 687, 11–20. doi: 10.1007/s10750-011-0727-z

El-Gebali, S., Mistry, J., Bateman, A., Eddy, S. R., Luciani, A., Potter, S. C., et al. (2019). The Pfam protein families database in 2019. *Nucl. Acids Res.* 47, D427–D432. doi: 10.1093/nar/gky995

Emura, C., Higuchi, R., and Miyamoto, T. (2006). Irciniasulfonic acid B, a novel taurine conjugated fatty acid derivative from a Japanese marine sponge, Ircinia sp. *Tetrahedron* 62, 5682–5685. doi: 10.1016/j.tet.2006.03.087

Fan, L., Reynolds, D., Liu, M., Stark, M., Kjelleberg, S., Webster, N. S., et al. (2012). Functional equivalence and evolutionary convergence in complex communities of microbial sponge symbionts. *Proc. Natl. Acad. Sci. U. S. A.* 109, E1878–1887. doi: 10.1073/pnas.1203287109

Fuhrman, J. A. (1999). Marine viruses and their biogeochemical and ecological effects. *Nature* 399, 541. doi: 10.1038/21119

Gregory, A. C., Zayed, A. A., Conceicao-Neto, N., Temperton, B., Bolduc, B., Alberti, A., et al. (2019). Marine DNA viral macro- and microdiversity from pole to pole. *Cell* 177, 1109–1123 e1114. doi: 10.1016/j.cell.2019.03.040

Guo, J., Bolduc, B., Zayed, A. A., Varsani, A., Dominguez-Huerta, G., Delmont, T. O., et al. (2021). VirSorter2: a multi-classifier, expert-guided approach to detect diverse DNA and RNA viruses. *Microbiome* 9, 37. doi: 10.1186/s40168-020-00990-y

Haber, M., Burgsdorf, I., Handley, K. M., Rubin-Blum, M., and Steindler, L. (2020). Genomic insights into the lifestyles of Thaumarchaeota inside sponges. *Front. Microbiol.* 11, 622824. doi: 10.3389/fmicb.2020.622824

Hallam, S. J., Konstantinidis, K. T., Putnam, N., Schleper, C., Watanabe, Y., Sugahara, J., et al. (2006). Genomic analysis of the uncultivated marine crenarchaeote *Cenarchaeum symbiosum*. *Proc. Natl. Acad. Sci. U. S. A.* 103, 18296–18301. doi: 10.1073/pnas.0608549103

Hawkes, N., Korabik, M., Beazley, L., Rapp, H. T., Xavier, J. R., and Kenchington, E. (2019). Glass sponge grounds on the Scotian Shelf and their associated biodiversity. *Mar. Ecol. Prog. Ser.* 614, 91–109. doi: 10.3354/meps12903

Howell, K. L., Pi Ec Haud, N., Downie, A. L., and Kenny, A. (2016). The distribution of deep-sea sponge aggregations in the North Atlantic and implications for their effective spatial management. *Deep Sea Res.* 115, 309–320. doi: 10.1016/j.dsr.2016.07.005

Huang, Y., Gilna, P., and Li, W. (2009). Identification of ribosomal RNA genes in metagenomic fragments. *Bioinformatics* 25, 1338–1340. doi: 10.1093/bioinformatics/btp161

Hyatt, D., LoCascio, P. F., Hauser, L. J., and Uberbacher, E. C. (2012). Gene and translation initiation site prediction in metagenomic sequences. *Bioinformatics* 28, 2223–2230. doi: 10.1093/bioinformatics/bts429

Jahn, M. T., Lachnit, T., Markert, S. M., Stigloher, C., Pita, L., Ribes, M., et al. (2021). Lifestyle of sponge symbiont phages by host prediction and correlative microscopy. *ISME J.* 15, 2001–2011. doi: 10.1038/s41396-021-00900-6

Jurkevitch, E., Minz, D., Ramati, B., and Barel, G. (2000). Prey range characterization, ribotyping, and diversity of soil and rhizosphere *Bdellovibrio* spp. isolated on phytopathogenic bacteria. *Appl. Environ. Microbiol.* 66, 2365–2371. doi: 10.1128/AEM.66.6.2365-2371.2000

Kahn, A. S., Geller, J. B., Reiswig, H. M., and Smith, K. L. (2013). *Bathydorus laniger* and *Docosaccus maculatus* (Lyssacinosida; Hexactinellida): two new species of glass sponge from the abyssal eastern North Pacific Ocean. *Zootaxa* 3646, 386–400. doi: 10.11646/zootaxa.3646.4.4

Karimi, E., Slaby, B. M., Soares, A. R., Blom, J., Hentschel, U., and Costa, R. (2018). Metagenomic binning reveals versatile nutrient cycling and distinct adaptive features in alphaproteobacterial symbionts of marine sponges. *FEMS Microbiol. Ecol.* 94, fiy074. doi: 10.1093/femsec/fiy074

Katoh, K., and Toh, H. (2010). Parallelization of the MAFFT multiple sequence alignment program. *Bioinformatics* 26, 1899–1900. doi: 10.1093/bioinformatics/btq224

Kauffman, K. M., Hussain, F. A., Yang, J., Arevalo, P., Brown, J. M., Chang, W. K., et al. (2018). A major lineage of non-tailed dsDNA viruses as unrecognized killers of marine bacteria. *Nature* 554, 118–122. doi: 10.1038/nature25474

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

Kutti, T., Bannister, R. J., and Fossså, J. H. (2013). Community structure and ecological function of deep-water sponge grounds in the Traenadypet MPA—Northern Norwegian continental shelf. *Cont. Shelf Res.* 69, 21–30. doi: 10.1016/j.csr.2013.09.011

Laffy, P. W., Wood-Charlson, E. M., Turaev, D., Jutz, S., Pascelli, C., Botte, E. S., et al. (2018). Reef invertebrate viromics: diversity, host specificity and functional capacity. *Environ. Microbiol.* 20, 2125–2141. doi: 10.1111/1462-2920.14110

Lambert, C., Hobley, L., Chang, C. Y., Fenton, A., Capeness, M., and Sockett, L. (2009). A predatory patchwork: membrane and surface structures of *Bdellovibrio bacteriovorus*. *Adv. Microb. Physiol.* 54, 313–361. doi: 10.1016/S0065-2911(08)00005-2

Langdon, W. B. (2015). Performance of genetic programming optimised Bowtie2 on genome comparison and analytic testing (GCAT) benchmarks. *BioData Min.* 8, 1. doi: 10.1186/s13040-014-0034-0

Lerner, T. R., Lovering, A. L., Bui, N. K., Uchida, K., Aizawa, S., Vollmer, W., et al. (2012). Specialized peptidoglycan hydrolases sculpt the intra-bacterial niche of predatory *Bdellovibrio* and increase population fitness. *PLoS Pathog.* 8, e1002524. doi: 10.1371/journal.ppat.1002524

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

Li, H., and Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 26, 589–595. doi: 10.1093/bioinformatics/btp698

Li, Q.-M., Zhou, Y.-L., Wei, Z.-F., and Wang, Y. (2021). Phylogenomic insights into distribution and adaptation of Bdellovibrionota in marine waters. *Microorganisms* 9, 757. doi: 10.3390/microorganisms9040757

Lücker, S., Nowka, B., Rattei, T., Spieck, E., and Daims, H. (2013). The genome of *Nitrospina gracilis* illuminates the metabolism and evolution of the major marine nitrite oxidizer. *Front. Microbiol.* 4, 27. doi: 10.3389/fmicb.2013.00027

Makarova, K. S., Wolf, Y. I., and Koonin, E. V. (2013). Comparative genomics of defense systems in archaea and bacteria. *Nucl. Acids Res.* 41, 4360–4377. doi: 10.1093/nar/gkt157

Maldonado, M., Aguilar, R., Bannister, R. J., Bell, J. J., Conway, K. W., Dayton, P. K., et al. (2017). "Sponge grounds as key marine habitats: a synthetic review of types, structure, functional roles, and conservation concerns," in *Marine Animal Forests: The Ecology of Benthic Biodiversity Hotspots*, eds Rossi, S., Bramanti, L., Gori, A., and Orejas, C. (Cham: Springer International Publishing), 145–183.

Maldonado, M., Ribes, M., and van Duyl, F. C. (2012). Nutrient fluxes through sponges. Biology, budgets, and ecological implications. *Adv. Mar. Biol.* 62, 113–182. doi: 10.1016/B978-0-12-394283-8.00003-5

Martynov, A., Severinov, K., and Ispolatov, I. (2017). Optimal number of spacers in CRISPR arrays. *PLoS Comput. Biol.* 13:e1005891. doi: 10.1371/journal.pcbi.1005891

McCutcheon, J. P., and Moran, N. A. (2011). Extreme genome reduction in symbiotic bacteria. *Nat. Rev. Microbiol.* 10, 13–26. doi: 10.1038/nrmicro2670

Mitchell, A. L., Almeida, A., Beracochea, M., Boland, M., Burgin, J., Cochrane, G., et al. (2020). MGnify: the microbiome analysis resource in 2020. *Nucleic Acids Res.* 48, D570–D578. doi: 10.1093/nar/gkz1035

Mittl, P. R. E., and Schneider-Brachert, W. (2007). Sel1-like repeat proteins in signal transduction. *Cell Signal.* 19, 20–31. doi: 10.1016/j.cellsig.2006.05.034

Moitinho-Silva, L., Diez-Vives, C., Batani, G., Esteves, A. I., Jahn, M. T., and Thomas, T. (2017a). Integrated metabolism in sponge-microbe symbiosis revealed by genome-centered metatranscriptomics. *ISME J.* 11, 1651–1666. doi: 10.1038/ismej.2017.25

Moitinho-Silva, L., Nielsen, S., Amir, A., Gonzalez, A., Ackermann, G. L., Cerrano, C., et al. (2017b). The sponge microbiome project. *Gigascience* 6, 1–7. doi: 10.1093/gigascience/giy145

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

Nayfach, S., Shi, Z. J., Seshadri, R., Pollard, K. S., and Kyrpides, N. C. (2019). New insights from uncultivated genomes of the global human gut microbiome. *Nature* 568, 505–510. doi: 10.1038/s41586-019-1058-x

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

Nguyen, M. T. H. D., Liu, M., and Thomas, T. (2014). Ankyrin-repeat proteins from sponge symbionts modulate amoebal phagocytosis. *Mol. Ecol.* 23, 1635–1645. doi: 10.1111/mec.12384

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

Oren, A., and Garrity, G. M. (2021). Valid publication of the names of forty-two phyla of prokaryotes. *Int. J. Syst. Evol. Microbiol.* 71:005056. doi: 10.1099/ijsem.0.005056

Parks, D. H., Chuvochina, M., Waite, D. W., Rinke, C., Skarshewski, A., Chaumeil, P. A., et al. (2018). A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat. Biotechnol.* 36, 996–1004. doi: 10.1038/nbt.4229

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

Pascelli, C., Laffy, P. W., Botte, E., Kupresanin, M., Rattei, T., Lurgi, M., et al. (2020). Viral ecogenomics across the Porifera. *Microbiome* 8, 144. doi: 10.1186/s40168-020-00919-5

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

Preston, C. M., Wu, K. Y., Molinski, T. F., and DeLong, E. F. (1996). A psychrophilic crenarchaeon inhabits a marine sponge: *Cenarchaeum symbiosum* gen. nov., sp. nov. *Proc. Natl. Acad. Sci. U. S. A.* 93, 6241–6246. doi: 10.1073/pnas.93.13.6241

Pritchard, L., Glover, R.H., Humphris, S., Elphinstone, J.G., and Toth, I.K. (2016). Genomics and taxonomy in diagnostics for food security: soft-rotting enterobacterial plant pathogens. *Anal. Methods* 8, 12–24. doi: 10.1039/C5AY02550H

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

Ren, J., Ahlgren, N. A., Lu, Y. Y., Fuhrman, J. A., and Sun, F. (2017). VirFinder: a novel k-mer based tool for identifying viral sequences from assembled metagenomic data. *Microbiome* 5, 69. doi: 10.1186/s40168-017-0283-5

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

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

Roux, S., Enault, F., Hurwitz, B. L., and Sullivan, M. B. (2015b). VirSorter: mining viral signal from microbial genomic data. *PeerJ* 3, e985. doi: 10.7717/peerj.985

Roux, S., Hallam, S. J., Woyke, T., and Sullivan, M. B. (2015a). Viral dark matter and virus–host interactions resolved from publicly available microbial genomes. *Elife* 4, e08490. doi: 10.7554/eLife.08490

Schleper, C., DeLong, E. F., Preston, C. M., Feldman, R. A., Wu, K. Y., and Swanson, R. V. (1998). Genomic analysis reveals chromosomal variation in natural populations of the uncultured psychrophilic archaeon *Cenarchaeum symbiosum*. *J. Bacteriol.* 180, 5003–5009. doi: 10.1128/JB.180.19.5003-5009.1998

Schmitt, S., Tsai, P., Bell, J., Fromont, J., Ilan, M., Lindquist, N., et al. (2012). Assessing the complex sponge microbiota: core, variable and species-specific bacterial communities in marine sponges. *ISME J.* 6, 564–576. doi: 10.1038/ismej.2011.116

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

Sockett, R. E. (2009). Predatory lifestyle of *Bdellovibrio bacteriovorus*. *Annu. Rev. Microbiol.* 63, 523–539. doi: 10.1146/annurev.micro.091208.073346

Steinert, G., Busch, K., Bayer, K., Kodami, S., Arbizu, P. M., Kelly, M., et al. (2020). Compositional and quantitative insights into bacterial and archaeal communities of south pacific deep-sea sponges (Demospongiae and Hexactinellida). *Front. Microbiol.* 11, 716. doi: 10.3389/fmicb.2020.00716

Suttle, C. A. (2005). Viruses in the sea. *Nature* 437, 356–361. doi: 10.1038/nature04160

Thomas, T., Moitinho-Silva, L., Lurgi, M., Bjork, J. R., Easson, C., Astudillo-Garcia, C., et al. (2016). Diversity, structure and convergent evolution of the global sponge microbiome. *Nat. Commun.* 7, 11870. doi: 10.1038/ncomms11870

Tian, R.-M., Wang, Y., Bougouffa, S., Gao, Z.-M., Cai, L., Bajic, V., et al. (2014). Genomic analysis reveals versatile heterotrophic capacity of a potentially symbiotic sulfur-oxidizing bacterium in sponge. *Environ. Microbiol.* 16, 3548–3561. doi: 10.1111/1462-2920.12586

Tian, R. M., Sun, J., Cai, L., Zhang, W. P., Zhou, G. W., Qiu, J. W., et al. (2016). The deep-sea glass sponge *Lophophysema eversa* harbours potential symbionts responsible for the nutrient conversions of carbon, nitrogen and sulfur. *Environ. Microbiol.* 18, 2481–2494. doi: 10.1111/1462-2920.13161

Tian, R. M., Zhang, W., Cai, L., Wong, Y. H., Ding, W., and Qian, P. Y. (2017). Genome reduction and microbe-host interactions drive adaptation of a sulfur-oxidizing bacterium associated with a cold seep sponge. *mSystems* 2, e00184–e00116. doi: 10.1128/mSystems.00184-16

Uritskiy, G. V., DiRuggiero, J., and Taylor, J. (2018). MetaWRAP-a flexible pipeline for genome-resolved metagenomic data analysis. *Microbiome* 6, 158. doi: 10.1186/s40168-018-0541-1

Vacelet, J., and Gallissian, M.-F. (1978). Virus-like particles in cells of the sponge *Verongia cavernicola* (Demospongiae, Dictyoceratida) and accompanying tissues changes. *J. Invertebr. Pathol.* 31, 246–254. doi: 10.1016/0022-2011(78)90014-9

Wang, P., Li, M., Dong, L., Zhang, C., and Xie, W. (2022). Comparative genomics of Thaumarchaeota from deep-sea sponges reveal their niche adaptation. *Front. Microbiol.* 13, 869834. doi: 10.3389/fmicb.2022.869834

Wang, Y., Huang, J. M., Cui, G. J., Nunoura, T., Takaki, Y., Li, W. L., et al. (2019). Genomics insights into ecotype formation of ammonia-oxidizing archaea in the deep ocean. *Environ. Microbiol.* 21, 716–729. doi: 10.1111/1462-2920.14518

Webster, N. S., and Thomas, T. (2016). The sponge hologenome. *MBio* 7, e00135–e00116. doi: 10.1128/mBio.00135-16

Wernegreen, J. J., and Funk, D. J. (2004). Mutation exposed: a neutral explanation for extreme base composition of an endosymbiont genome. *J. Mol. Evol.* 59, 849–858. doi: 10.1007/s00239-003-0192-z

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

Wilkinson, C. R. (1983). Net primary productivity in coral reef sponges. *Science* 219, 410–412. doi: 10.1126/science.219.4583.410

Xing, P., Tao, Y., Luo, J., Wang, L., Li, B., Li, H., et al. (2020). Stratification of microbiomes during the holomictic period of Lake Fuxian, an alpine monomictic lake. *Limnol. Oceanogr.* 65, S134–S148. doi: 10.1002/lno.11346

Xu, H., Luo, X., Qian, J., Pang, X., Song, J., Qian, G., et al. (2012). FastUniq: a fast de novo duplicates removal tool for paired short reads. *PLoS ONE* 7, e52249. doi: 10.1371/journal.pone.0052249

Yang, Q., Franco, C. M. M., Sorokin, S. J., and Zhang, W. (2017). Development of a multilocus-based approach for sponge (phylum Porifera) identification: refinement and limitations. *Sci. Rep.* 7, 1–14. doi: 10.1038/srep41422

Yin, Z., Zhu, M., Davidson, E. H., Bottjer, D. J., Zhao, F., and Tafforeau, P. (2015). Sponge grade body fossil with cellular resolution dating 60 Myr before the Cambrian. *Proc. Natl. Acad. Sci. U. S. A.* 112, E1453–E1460. doi: 10.1073/pnas.1414577112

Zhang, F., Jonas, L., Lin, H., and Hill, R. T. (2019a). Microbially mediated nutrient cycles in marine sponges. *FEMS Microbiol. Ecol.* 95. doi: 10.1093/femsec/fiz155

Zhang, F., Zhao, S., Ren, C., Zhu, Y., Zhou, H., Lai, Y., et al. (2018). CRISPRminer is a knowledge base for exploring CRISPR-Cas systems in microbe and phage interactions. *Commun. Biol.* 1, 180. doi: 10.1038/s42003-018-0184-6

Zhang, S., Song, W., Wemheuer, B., Reveillaud, J., Webster, N., and Thomas, T. (2019b). Comparative genomics reveals ecological and evolutionary insights into sponge-associated Thaumarchaeota. *mSystems* 4, e00288–e00219. doi: 10.1128/mSystems.00288-19

frontiers | Frontiers in Marine Science

TYPE Original Research
PUBLISHED 29 June 2023
DOI 10.3389/fmars.2023.1151119

OPEN ACCESS

EDITED BY
Senjie Lin,
University of Connecticut, United States

REVIEWED BY
Santosh Thapa,
Baylor College of Medicine, United States
Jinsong Zheng,
Chinese Academy of Sciences (CAS), China

*CORRESPONDENCE
Francisco Marco-Jiménez
fmarco@dca.upv.es

RECEIVED 25 January 2023
ACCEPTED 13 June 2023
PUBLISHED 29 June 2023

CITATION
Muñoz-Baquero M, Lorenzo-Rebenaque L, García-Vázquez FA, García-Párraga D, Martínez-Priego L, De Marco-Romero G, Galán-Vendrell I, D'Auria G and Marco-Jiménez F (2023) Unveiling microbiome signature in inner body fluids: comparison between wild and aquarium small-spotted catshark (*Scyliorhinus canicula*). *Front. Mar. Sci.* 10:1151119. doi: 10.3389/fmars.2023.1151119

Unveiling microbiome signature in inner body fluids: comparison between wild and aquarium small-spotted catshark (*Scyliorhinus canicula*)

Marta Muñoz-Baquero[1,2], Laura Lorenzo-Rebenaque[1], Francisco Alberto García-Vázquez[3], Daniel García-Párraga[2,4], Llucia Martínez-Priego[5], Griselda De Marco-Romero[5], Inmaculada Galán-Vendrell[5], Giuseppe D'Auria[5] and Francisco Marco-Jiménez[6]*

[1]Departamento de Producción y Sanidad Animal, Salud Pública Veterinaria y Ciencia y Tecnología de los Alimentos, Instituto de Ciencias Biomédicas, Facultad de Veterinaria, Universidad Cardenal Herrera-CEU, CEU Universities, Calle Santiago Ramón y Cajal, Alfara del Patriarca, Spain, [2]Research Department. Fundación Oceanogràfic de la Comunitat Valenciana, Ciudad de las Artes y las Ciencias, C/Eduardo Primo Yúfera, Valencia, Spain, [3]Departamento de Fisiología, Facultad de Veterinaria, Universidad de Murcia, Murcia, Spain, [4]Veterinary Services, Avanqua-Oceanogràfic S.L, Ciudad de las Artes y las Ciencias, C/Eduardo Primo Yúfera, Valencia, Spain, [5]Servicio de Secuenciación y Bioinformática, Fundación para el Fomento de la Investigación Sanitaria y Biomédica de la Comunidad Valenciana (FISABIO-Salud Pública), Valencia, Spain, [6]Instituto de Ciencia y Tecnología Animal, Universitat Politècnica de València, Valencia, Spain

The microbiome is a crucial influencer in animal development, immune function and health, and it has complex and dynamic interactions with the environment, but little is known about the microbial signatures of inner body fluids. Recent evidence suggests that inner body fluids could be an indicator of the environmental interactions that fish experience. In the present study, we provide a comparative analysis of the microbial profile found in small-spotted catsharks' blood plasma and seminal plasma and how microbial signatures vary between aquarium and wild animals. In the blood plasma microbiome, the habitat did not affect the α- and β-diversity, while in seminal plasma, both α- and β-diversity differed between both habitats. Proteobacteria are the main bacteria dominated independently the inner body fluid and habitats. No core microbiome was identified at the genus level, with only *Pseudomonas* and *Cloacibacterium* present in both inner body fluids and habitats. Of the 14 genera identified in blood plasma, only four were shared between habitats (making up 45.17% and 51.03% of relative abundance for wild and aquarium, respectively). Similarly, of the 100 genera identified in seminal plasma, only 41 were shared between habitats (84.79% and 64.83%, respectively). Moreover, in the seminal plasma, using ANCOM approaches, *Serratia*, *Salinisphaera* and *Cutibacterium* were found significantly enriched in aquarium animals. None potentially pathogenic bacteria were identified in the blood samples, while *Coxiella*, *Prevotella*, *Coprococcus*, *Haemophilus* and *Phocoenobacter* were potentially pathogenic bacteria identified in the seminal plasma samples. In

summary, this study provides evidence of a circulating blood and seminal plasma microbiome in healthy small-spotted catsharks. Furthermore, dynamic changes were observed in the microbiome of these inner body fluids, which differed between the aquarium and wild habitats.

KEYWORDS

microbial community, sharks, aquaria, aquatic animals, chondrichthyan, 16S rRNA

1 Introduction

As with other vertebrates, the fish microbiome is fundamental to the health of its host and has complex and dynamic interactions with the environment (Luna et al., 2022). In this context, it is well known that environmental stressors can compromise fish health and fitness (Uren-Webster et al., 2020). Certainly, animals under managed care offer an ideal framework for identifying microbial components that may have significant associations with host metabolism and overall health. It is also important during the identification of potential drivers influencing microbiome assembly, such as environmental changes (Clavere-Graciette et al., 2022). Certainly, large aquaria, designed using recirculation systems to mimic natural marine environments as closely as possible, including host resident microbes, can be particularly useful (Patin et al., 2018; Clavere-Graciette et al., 2022). It is now widely accepted that all multicellular organisms associated with microbes (bacteria, protists, fungi, viruses and protist) that contribute to the biology of their respective host have variable functional impacts, ranging from beneficial to unimpactful to detrimental (Bosch and McFall-Ngai, 2011; Human Microbiome Project Consortium, 2012; Moran, 2015). Indeed, host-associated microbiomes are considered one functional form (biological unit) (Knowlton and Rohwer, 2003; Bosch and McFall-Ngai, 2011; Bang et al., 2018; Jaspers et al., 2019), but there is little literature focused on the microbe communities in large aquaria and their effects on animal physiology (Rowe et al., 2020; Perry et al., 2021; Clavere-Graciette et al., 2022). Likewise, microbial structure in wild animals has been poorly studied, although natural fish populations are increasingly subject to multiple anthropogenic stressors that threaten their conservation status (Perry et al., 2020; Rowe et al., 2020; Uren-Webster et al., 2020).

In recent decades, numerous studies have explored oral, gill, skin, cloacal, and gut microbial microbiomes (Perry et al., 2020; Sehnal et al., 2021; Clavere-Graciette et al., 2022). In contrast, little is known about the existence of a microbiome signature in inner body fluids. For instance, recent data support the existence of a microbial profile in the blood of various domesticated mammals and birds (Sze et al., 2014; Mandal et al., 2016; Vientós-Plotts et al., 2017; Scarsella et al., 2020); however, this statement is not enterily clear in relation to healthy human individuals due to opposite results found in the literature (Whittle et al., 2018; Castillo et al., 2019; D'Aquila et al., 2021; Tan et al., 2023) and, although due to opposite results recently published (Tan et al., 2023). In fish, very little is known about the microbial profile found in blood (Grimes et al., 1985; Buck, 1990; Mylniczenko et al., 2007; Tao et al., 2014; Tarnecki et al., 2018). The bacteria that can access the circulatory compartment could potentially be used as an indicator of the environmentally-derived taxa (Castillo et al., 2019). Likewise, a reproductive microbial profile could significantly affect the reproductive function and fitness of males and females, although these are rarely considered from an ecological or evolutionary perspective (Weinberg, 1974; Karin et al., 2006; Culligan and Sleator, 2016; Taylor et al., 2018; Theodosopoulos et al., 2019; Rowe et al., 2020). Understanding microbial signatures of reproductive fluids is an emerging and critical component of wildlife conservation (Comizzoli et al., 2021). Actually, metagenomic study of the reproductive microbiome is being considered a pivotal tool for species conservation, as it will allow the study of evolutionary and adaptive mechanisms that could be involved in the lack of reproductive efficiency and potentially, its intervention could collaborate in the preservation of endangered species (Contreras et al., 2023).

However, the key elements connecting microbial communities in reproductive fluids and reproductive success are not defined (Koh et al., 2019; Rowe et al., 2020).

In the present study, we compare the microbial profile found in the peripheral blood and seminal plasma of small-spotted male catsharks (*Scyliorhinus canicula*). In addition, we explored how microbial signatures of these inner body fluids vary between wild and aquarium (Oceanogràfic, Ciudad de las Artes y las Ciencias, Valencia, Spain) animals.

2 Materials and methods

2.1 Sample collection

Small-spotted catsharks were sampled in January 2020. Blood and seminal samples were collected from 17 wild small-spotted catsharks from the Mediterranean Sea and 7 aquarium-housed small-spotted catsharks in collaboration with the Oceanogràfic. Ubication (wild *vs* aquarium), date of sampling, body weight, length, width and clasper length for all catsharks are provided in Supplementary Table 1. Wild animals were donated by local fisheries in the Valencian Community and were part of accidental

captures destined for commercial fisheries from the ports of Valencia (39°26′45″N 0°19′12″W), Jávea (38°47′21″N 0°09′47″E) and Cullera (39°09′58″N 0°15′10″W). Mediterranean Sea water parameters, measured by Valencia buoy (39°52′N0°20′E) during the sampling period were 14.6-19°C temperature and 34-37 g/l salinity (http://www.puertos.es/es-es/oceanografia/Paginas/portus.aspx). Animals were maintained in a closed and recirculation system under controlled conditions, monitored water quality (17–21°C, 5.1 mg/l oxygen, 36 g/l salinity and 7.6–8.2 pH), fixed photoperiod (12:12 h) and disinfection using UV light and ozone. Animals were kept isolated from females for at least 1 year before the experiment started. Aquarium animals' diet was based on frozen-thawed whole fish: herring (*Clupea harengus*), mackerel (*Scomber scombrus*) and squid (*Loligo* sp). All animals from the study were classified as adults, displaying calcified claspers ensuring reproductive maturity (Kousteni et al., 2010). All wild animals were determined healthy based on physical examination, body conditions and confirmed *in vitro* sperm quality (Muñoz-Baquero et al., 2021). Aquarium animals were also determined healthy based on physical exams and clinical history. Each sampling day and biometrics were recorded for all the small-spotted catsharks, the weight ranged from 165 to 370 g (mean ± SD was 281.2 ± 48.6 g). The maturity of the sharks was determined by their size according to Ebert and Dando, 2020, together with the gonad development and clasper calcification length ranged from 39 to 50 cm (mean ± SD was 44.9 ± 2.7 cm), and clasper length ranged from 2.9 to 5 cm (mean ± SD was 3.7 ± 0.5 cm), considering also the total width.

Peripheral blood and semen samples from wild and aquarium animals were collected on the same day across 2 sessions. Therefore, during the collection of samples from the aquarium sharks instructed by the veterinarian, small-spotted catsharks were also collected from local fisheries on the same day. Peripheral blood was collected from the ventral coccygeal vein by caudal venipuncture, using 23-gauge needles and transferred to lithium-heparin tubes (Figure 1). Semen samples were collected by applying external pressure on the ampulla of the vas deferens, using a 5 ml syringe for semen collection (Figure 1). In aquarium animals, to rule out possible contamination with skin microbiome during the sampling, an sterile zone was created, both at the venipuncture and cloacal sites, by washing the area with sterile elasmobranch Ringer's solution (22 g/l urea and 9 g/l NaCl in distilled water) and sterile gauzes (Mylniczenko and Clauss, 2017), following the Elasmobranch Husbandry Guidelines. Moreover, semen samples were checked for possible urine and fecal contamination under the microscope. After the collection of the samples from aquarium animals, the small-spotted catsharks were released back into the quarantine tank, and their recovery was carefully monitored by the aquarium veterinarian. Samples were immediately transferred to the laboratory at 4°C in a dark container after collection. For wild animals, firstly, disinfection of area with an ethanol-soaked (96°) cloth was carried out, followed by washing as was done for the

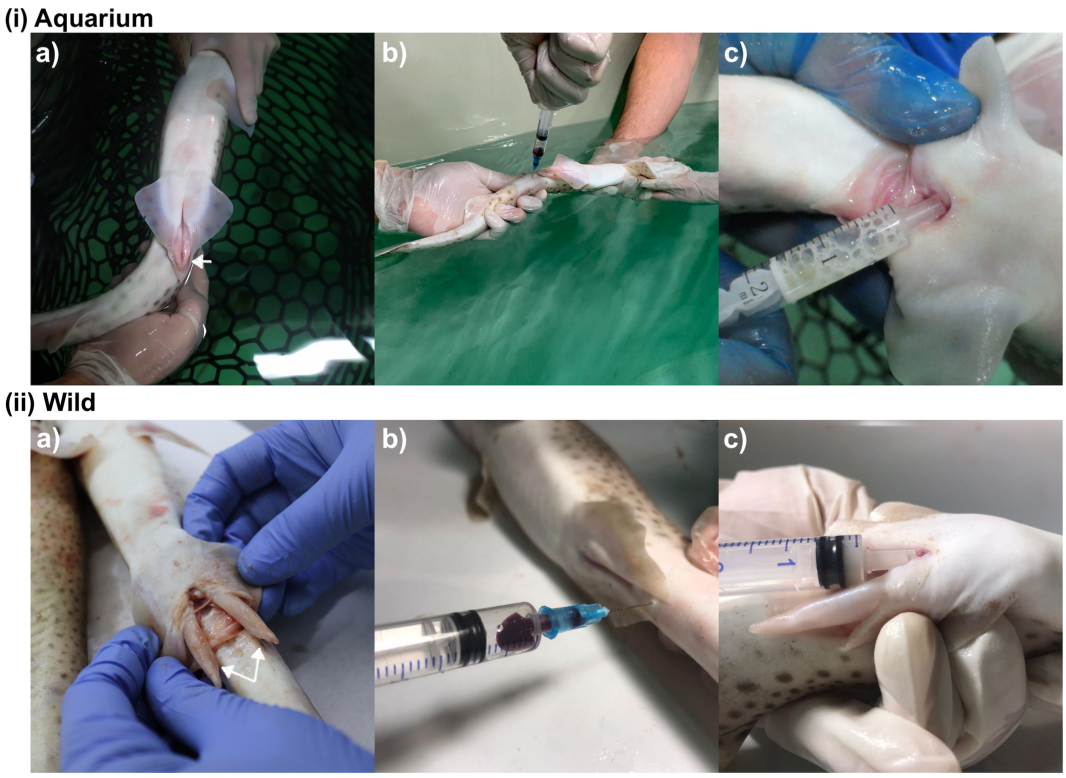


FIGURE 1
Pictures demonstrating blood and semen samples collection from aquarium (i) and wild (ii) small-spotted catshark (*Scyliorhinus canicula*): **(A)** (i) an aquarium small-spotted catshark in holding tank and (ii) a wild small-spotted catshark in lab with clasper detail (white arrows), **(B)** (i, ii) blood sampling, **(C)** (i, ii) semen sampling, Semen samples were collected by abdominal massage.

aquarium animals. Samples were transferred to the laboratory under the same conditions, within a maximum of 4 h post-fishing.

Only animals from which both semen (100 µl) and blood (100 µl) were obtained were included in the study. A total of 7 pooled samples (4 for wild and 3 for aquarium animals) were established, mixing 100 µl of each individual sample. The 4 pooled samples from the wild animals included 5, 3, 3, and 6 animals for Valencia, Cullera (2 pooled samples) and Jávea ports, respectively. Regarding the samples from the aquarium animals, due to the limited number of animals (n=7), 3 pooled samples were obtained. The 2 first pooled samples included 3 and 3 animals. The third pooled sample includes the remaining animals, 1 from the first sampling session and 1 from the second pool with an excess in semen production (more than 200 µl). The sample pooling strategy is described in Supplementary Table 1. Pooling samples can serve as an effective method for screening bacterial nucleic acids (Federer, 1994; Furstenau et al., 2020). When both, blood and seminal plasma were available from each individual sample, a random selection of 3 samples (for the aquarium group) or 4 samples (for the wild group) was made to establish a single pooled sample per group. This was achieved by combining 100 µl of each individual sample.

Both blood and semen pooled samples were centrifuged at 3000 g and 7400 g, respectively, for 10 min at 4°C. After centrifugation, the seminal plasma fraction was verified under a microscope ensuring no spermatozoa. All plasma samples were directly flash-frozen in liquid nitrogen and stored at −80°C until further processing.

2.2 DNA extraction, library preparation and sequencing

Total DNA from generated pool samples was extracted using the NucliSens easyMag automated system (bioMérieux, Marcy l'Etoile, France) according to the manufacturer's instructions. The amplification and sequencing protocol started with the microbial genomic DNA (5 ng/µl in 10 mM Tris pH 8.5) as the starting point. The V3–V4 region of the bacterial 16S rRNA gene sequences was amplified using the primer pair S-D-Bact-0341-b-S-17 -S-D-Bact-0785-a-A-21 (Klindworth et al., 2013) containing the gene-specific and Illumina adapter overhang nucleotide sequences. Primer sequences were forward primer: (5′- TCGTCGGCAGCGTC AGATGTGTATAAGAGACAG-CCTACGGGNGGCWGCAG-3'; reverse primer: 5'-GTCTCGTGGGCTCGGAGATGTG TATAAGAGACAG-GACTACHVGGGTATCTAATCC-3'). After 16S rRNA gene amplification, the multiplexing step was performed using the Nextera XT Index Kit (FC-131-1096). One-µl of the PCR product on a Bioanalyzer DNA 1000 chip (Agilent Technologies, USA, Santa Clara) was run to verify the size, with ~550 bp as the expected size on the Bioanalyzer trace. Thus, the libraries were sequenced using a 2 × 300 bp paired-end run (MiSeq Reagent kit v3, MS-102-3001) on a MiSeq Sequencer, following the manufacturer's instructions (Illumina, Inc., San Diego, CA, USA). In addition, a negative control was employed during library preparation to ensure accuracy and identify any potential background or contamination issues.

DNA-free water (UltraClean PCR Water, MO BIO Laboratories, Carlsbad, CA, USA) was employed as the input for the post-extraction steps of the protocol, starting with library generation, to assess contamination in downstream processes.

2.3 Bioinformatics analysis

Raw sequencing data were processed by QIIME2 v2021.4. The DADA2 pipeline included into QIIME2 was used to carry out denoising, filtering, and chimera removal of the sequences and reads were derreplicated in Amplicon Sequence Variants (ASVs). Each ASV was taxonomical assigned using the SILVA v138 database (Quast et al., 2013; Campos et al., 2022). Sequences classified as Chloroplast or Mitochondria were filtered out from the analysis. Identification of contaminant ASVs was performed using the decontam (v1.8.0) R package (Davis et al., 2018). The negative control exhibited the presence of 4 bacterial genera (Supplementary Table 2). R Decontam package did not detected rRNA contaminant gene reads among the datasets (Supplementary Figure 1). These low numbers of contaminant reads did not have any effect on the statistical analyses.

The α, β and taxonomic diversity analyses were mainly based on the QIIME2 pipeline. For the α-diversity, the richness and the richness with their relative abundance (evenness) were measured by Chao1 and Shannon indexes, respectively. The significance of differences among wild and aquarium sharks for seminal plasma and serum plasma was evaluated by Kruskal–Wallis test. Box-and-whisker plots for α-diversity were generated using Graphad Prism 8. A Venn diagram was drawn up to show the shared, and unique features among groups, based on the occurrence of features in a sample group regardless of their relative abundance and an interactive Venn software was used for Venn diagram construction (Heberle et al., 2015). Analysis of the composition of microbiomes (ANCOM) was run to determine if samples changed significantly between habitats by measuring the W value corresponding to the number of times an ASV abundance is significantly different for a group of samples (Mandal et al., 2015). The Bray-curtis distance, unweighted and weighted Unifrac values were calculated to estimate β-diversity between samples. Principal Coordinate Analysis (PCoA) plots were obtained for each sample's habitats from Bray-Curtis distances using ClustVist software (Metsalu and Vilo, 2015). Microbial taxa that are shared by all groups was considered as core microbiome.

The processed 16S rRNA gene sequencing data were analyzed using PICRUSt (Phylogenetic Investigation of Communities by Reconstruction of Unobserved States) (Douglas et al., 2020). This allowed for the prediction of the microbial communities functional capabilities on the taxonomic composition inferred from the 16S rRNA gene data. To identify biomarker functional potentials between groups, the LEfSe (Linear discriminant analysis Effect Size) method was employed through the website http://huttenhower.sph.harvard.edu/galaxy/. Additionally, the Functional Annotation of PROkaryotic TAXa (FAPROTAX) database (Louca et al., 2016; Zhou et al., 2020) was implemented to identify potential pathogens based on the 16S rRNA gene sequences.

3 Results

A total of 37 and 410 unique sequences were used for taxonomic assignment for peripheral blood and seminal plasma samples, respectively (Table 1). For peripheral blood plasma, there were no differences in α-diversity indexes between the wild and aquarium animals (Figure 2A), while for the seminal plasma, lower richness (Chao1 index) for the wild animals was observed (Figure 2B). To observe differences regarding the β-diversity between samples, a PCoA ordination plot was performed using Bray-Curtis distances (Figure 3), where the two first principal components (PCs) explain the 29.8% and 25.5% for blood plasma and the 19.6% and 16.7% for seminal plasma of the observed variance, respectively. Moreover, results of PERMANOVA (Permutational multivariate analysis of variance) on unweighted and weighted UniFrac distances were used to compare microbiome taxonomic composition between body fluids and between the aquarium and wild animals. In pairwise comparisons of inner body fluids, no significant differences were found between the wild and aquarium small-spotted catsharks in blood plasma. However, there was a significant difference in seminal plasma for unweighted UniFrac distance (blood plasma: $p = 0.474$ and $p = 0.774$ and, for wild and aquarium animals respectively; and seminal plasma: $p = 0.102$ and $p = 0.036$, for wild and aquarium animals, respectively).

At the phyla level, peripheral blood plasma microbiome in wild animals was dominated by *Proteobacteria, Firmicutes, Actinobacteria* and *Bacteroidetes*, while in aquarium small-spotted catsharks, *Firmicutes* were not present (Figures 4A, B). In the seminal plasma, the microbial composition did not differ between wild and aquarium animals, being dominated in both habitats by *Proteobacteria, Firmicutes, Actinobacteria* and *Bacteroidetes*

TABLE 1 General features of 16S rRNA amplicon sequencing of peripheral blood and seminal plasma microbiota.

	Peripheral Blood Plasma	Seminal Fluid
Total raw reads	375,009	334,893
Average sequences length (bp)	376.7 ± 64.35	412.8 ± 65.73
Average number of sequences per sample	46,876.12	41,861.65
Minimum number of sequences per sample	32,709	11,195
Maximum number of sequences per sample	60,524	107,421
Unique sequences (a)	38,456	155,169
ASVs generated	3,485	989
Sequences removed	37,972	37,463
Mapped Sequences (b)	484	117,706
ASVs generated for taxonomic assignment	37	410

(a) Sequences obtained after a series of data processing steps such as data filtering, denoising, merging and chimera removing.
(b) Against the reference dataset SILVA.

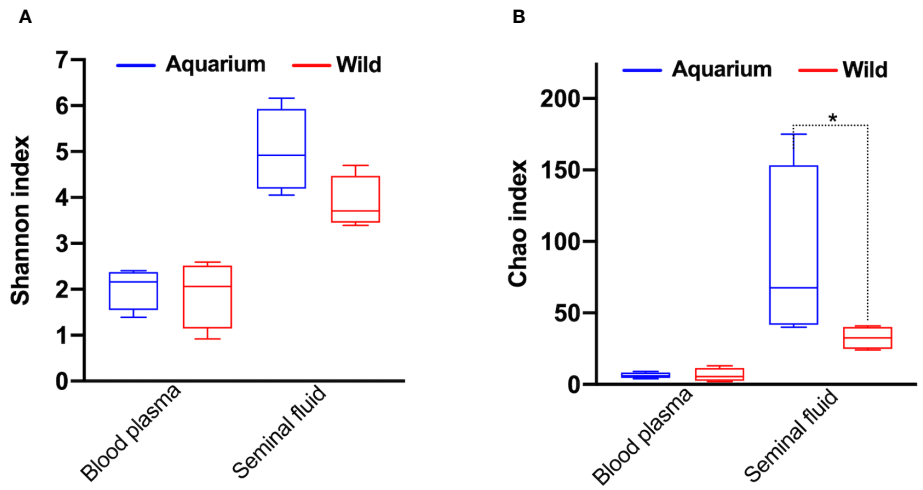


FIGURE 2
α-Diversity metrics for **(A)** Shannon and **(B)** Chao diversity indices across inner body fluids (peripheral blood plasma and seminal plasma) of aquarium and wild small-spotted catshark (*Scyliorhinus canicula*). * indicates significant differences ($p < 0.05$) between aquarium and wild animals in the same inner body fluid analysed.

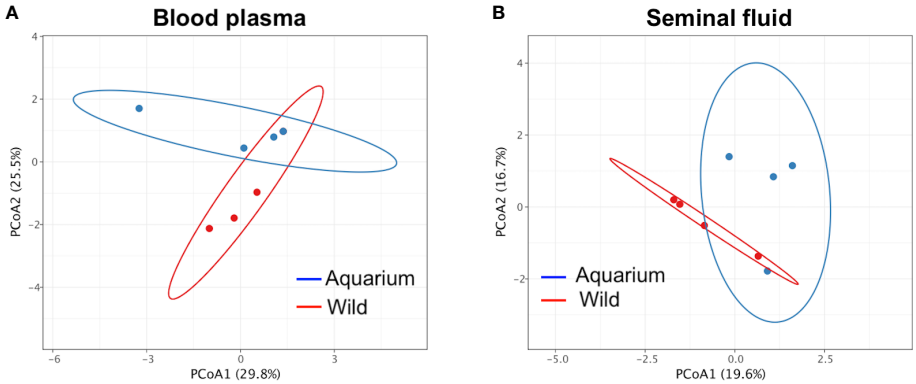


FIGURE 3
Principal coordinate analysis (PCoA) of β-diversity comparison using Bray–Curtis distances across inner body fluids of wild and aquarium small-spotted catshark (*Scyliorhinus canicula*). For **(A)** peripheral blood plasma and **(B)** seminal plasma samples.

(Figures 4C, D). Across both groups of animals (aquarium and wild), *Proteobacteria* was the predominant phylum independently of the inner body fluid. To examine the existence of an identifiable common core microbiome, defined as the group of members shared among the microbial community, a Venn diagram was represented by overlapping areas in the circles at the genus level of the bacterial community. Thus, no core microbiome could be identified in both inner body fluids (Figure 5A), with only *Pseudomonas* and *Cloacibacterium* being present blood plasma and seminal fluid in both wild and aquarium animals (Figure 5B). In peripheral blood plasma, 4 shared genera were identified between both experimental groups (wild and aquarium) (Figure 5C), while in seminal plasma, a total of 41 shared genera were identified between wild and aquarium animals (Figure 5D). At finer taxonomic levels, such as genus, the blood microbiome was distinct between wild and aquarium individuals (Table 2). Differences were observed in several genera present only in wild animals and other genera present only in aquarium catsharks (Table 2). Likewise, the

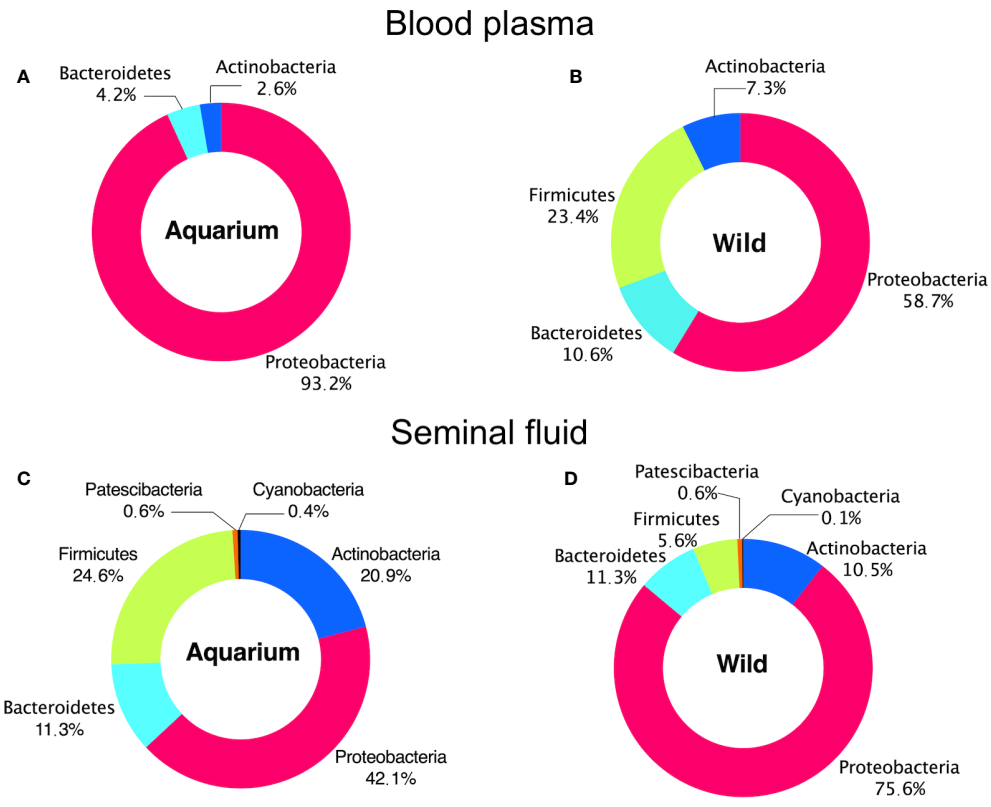


FIGURE 4
Phylum distribution in different inner body fluids from aquarium and wild small-spotted catshark (*Scyliorhinus canicula*). For **(A, B)** peripheral blood plasma and **(C, D)** seminal plasma samples.

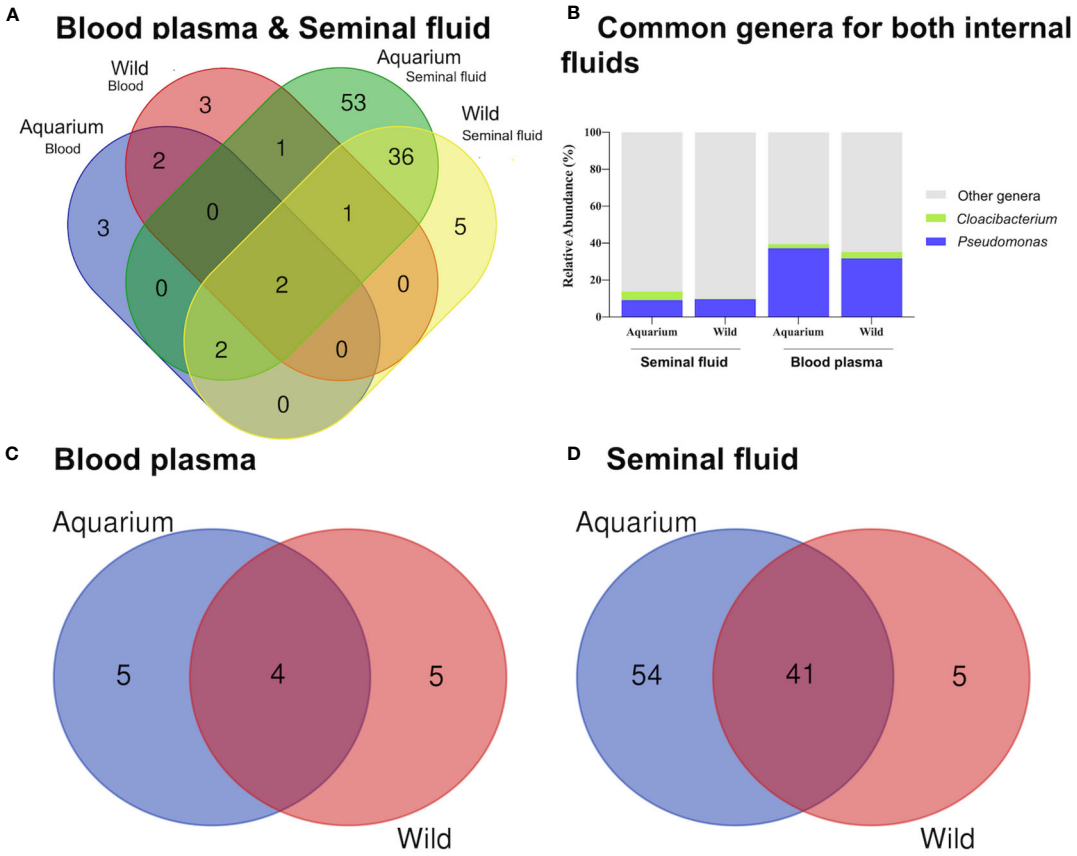


FIGURE 5
Venn diagram representation of shared and unique genera in different inner body fluids from aquarium and wild small-spotted catshark (*Scyliorhinus canicula*) using Venny (https://bioinfogp.cnb.csic.es/tools/venny_old/venny.php). For **(A)** both inner body fluids, **(B)** Shared genera **(C)** peripheral blood plasma and **(D)** seminal plasma samples.

seminal plasma fluid microbiome differs between wild and aquarium organisms (Table 3— seminal plasma microbiome at an average relative abundance of more than 1% in at least one sample group). ANCOM approaches demonstrated that *Salinisphaera, Serratia*, and *Cutibacterium* were significantly enriched in the aquarium small-spotted catshark seminal plasma (W=32, 11, and 11, respectively). In addition, these findings were observed for several genera that were exclusively present in wild animals and other genera that were exclusively present in aquarium animals (Table 3).

Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis revealed the detection of a total of 194 pathways across all small-spotted catshark samples. The PICRUSt results elucidated slight differences in the composition of metabolic functions in the blood plasma samples (Supplementary Table 4). The significant metabolic functions identified were glucose and glucose-1-phosphate degradation (GLUCOSE1PMETAB-PWY), L-methionine biosynthesis III (HSERMETANA.PWY), urate biosynthesis/inosine 5-phosphate degradation (PWY.5695), and pyruvate fermentation to isobutano (PWY.7111). It is worth noting that pyruvate fermentation to isobutano is an engineered pathway, not occurring naturally in any known organism, but constructed through metabolic engineering in a living cell. Comparing the seminal plasma samples, no significant differences in the composition of metabolic functions were observed between aquarium and wild animals (Supplementary Table 4).

Furthermore, FAPROTAX analysis revealed that the relative abundance of dominant functions exhibited varied compositions among different samples. A total of 16 and 27 categories of microbial functions linked to the bacterial communities were identified in the blood and seminal plasma samples, respectively (Supplementary Table 5). The predominant microbial functions assigned were chemoheterotrophy and aerobic chemoheterotrophy for both inner fluids. Additionally, nitrate reduction and fermentation were observed in the seminal plasma samples. In order to assess the risks of potential pathogens for aquatic wildlife (e.g., aquatic animals and plants), we focused on groups relevant to pathogens in this study. No pathogenic groups were identified in the blood samples, whether from aquarium or wild animals (Figure 6A). In the seminal plasma, pathogenic groups were detected, including animal parasites or symbionts (in 2 pooled samples from aquarium animals and 1 pooled sample from wild animals, representing 2.36%) and intracellular parasites (in 1 pooled sample from aquarium animals, representing 0.44%) (Figure 6B). The potentially pathogenic bacteria detected were *Coxiella*, *Prevotella*, *Coprococcus*, *Haemophilus* and *Phocoenobacter*.

TABLE 2 Relative abundance of bacteria at the genus level in peripheral blood plasma from aquarium and wild small-spotted catshark (*Scyliorhinus canicula*), based on the ANCOM test.

Phylum	Family	Genera	Habitat	
			Aquariumm (%)	Wild (%)
Proteobacteria	*Pseudomonadaceae*	*Pseudomonas*	37.28	31.72
	Beijerinckiaceae	*Methylobacterium-Methylorubrum*	29.82	–
	Xanthobacteraceae	*Rhodopseudomonas*	7.46	–
		Bradyrhizobium	3.07	3.45
	Rhodobacteraceae	*Paracoccus*	–	10.34
	Burkholderiaceae	*Pandoraea*	6.14	–
	Sphingomonadaceae	*Novosphingobium*	7.02	–
	Yersiniaceae	*Serratia*	4.39	–
Firmicutes	*Bacillaceae*	*Anoxybacillus*	–	13.10
	Monoglobaceae	*Monoglobus*	–	9.66
	Thermicanaceae	*Thermicanus*	–	8.28
Bacteroidota	*Dysgonomonadaceae*	*Dysgonomonas*	2.63	12.41
	Weeksellaceae	*Cloacibacterium*	2.19	3.45
Actinobacteriota	*Nocardiaceae*	*Gordonia*	–	7.59

(–) indicates undetectable levels.

TABLE 3 Relative abundance of more than 1% in at least one sample group of bacteria at the genus level in seminal plasma from aquarium and wild small-spotted catshark (*Scyliorhinus canicula*).

Phylum	Family	Genera	Habitat	
			Aquarium (%)	Wild (%)
Proteobacteria	*Pseudomonadaceae*	*Pseudomonas*	9.15	9.69
	Saccharospirillaceae	*Oleispira*	6.07	1.40
	Yersiniaceae	*Serratia**	2.47	2.07
	Moraxellaceae	*Enhydrobacter*	2.36	0.77
		Acinetobacter	2.21	2.47
	Salinisphaeraceae	*Salinisphaera**	2.12	–
	Marinobacteraceae	*Marinobacter*	2.07	0.28
	Pseudoalteromonadaceae	*Pseudoalteromonas*	2.06	29.68
	Comamonadaceae	*Comamonas*	1.91	0.91
	Sphingomonadaceae	*Novosphingobium*	1.32	0.09
		U.m. family Sphingomonadaceae	0.52	2.00
	Halomonadaceae	*Halomonas*	1.20	0.23
	Rhizobiaceae	*Allorhizobium-Neorhizobium-Pararhizobium*	0.37	1.29
		U.m. family Rhizobiaceae	0.23	1.31
	Vibrionaceae	*Vibrio*	0.01	14.84
		Photobacterium	–	6.58
Firmicutes	*Staphylococcaceae*	*Staphylococcus*	4.75	0.05

(Continued)

TABLE 3 Continued

Phylum	Family	Genera	Habitat	
			Aquarium (%)	Wild (%)
	Streptococcaceae	*Streptococcus*	1.84	0.39
	Enterococcaceae	*Enterococcus*	1.74	–
	Ruminococcaceae	*Subdoligranulum*	1.24	–
		Faecalibacterium	1.56	–
	Aerococcaceae	*Aerococcus*	1.55	–
		Globicatella	1.71	–
	Erysipelotrichaceae	*Holdemanella*	1.34	–
	Lachnospiraceae	*Blautia*	1.48	–
		[Eubacterium]_hallii_group	1.30	–
		Lachnoclostridium	1.25	–
	Bacillaceae	*Geobacillus*	0.69	1.24
	Mycoplasmataceae	*Mycoplasma*	–	3.18
Bacteroidota	*Weeksellaceae*	*Cloacibacterium*	4.64	0.23
		U.m. family Weeksellaceae	1.52	–
		Chryseobacterium	0.30	1.97
	Rikenellaceae	*Alistipes*	1.30	–
	Flavobacteriaceae	*U.m. family Flavobacteriaceae*	–	2.24
Actinobacteriota	*Propionibacteriaceae*	*Cutibacterium**	7.36	4.44
	Corynebacteriaceae	*Corynebacterium*	4.35	1.12
		Lawsonella	4.13	0.05
	Microbacteriaceae	*U.m. family Microbacteriaceae*	1.53	–
	Rubrobacteriaceae	*Rubrobacter*	0.07	1.45
	Promicromonosporaceae	*Promicromonospora*	–	1.72

*indicates significant differences in abundance levels between groups based on the ANCOM test. (–) indicates undetectable levels.

4 Discussion

Our results are consistent with previous studies that have reported bacterial colonization in the blood of healthy marine life, terrestrial animals, and humans (Vientós-Plotts et al., 2017; Tarnecki et al., 2018; Castillo et al., 2019; Anhê et al., 2020). This supports the existence of a blood plasma microbiota in healthy small-spotted catsharks. In this study, no statistical associations were observed between the pathogen community and blood samples, regardless of the sample's origin. However, it is important to note that functional annotations based on the FAPROTAX database provide inferred functions relying on 16S rRNA fragments, which may not be as precise as a comprehensive shotgun metagenomic study. This consideration should be taken into account when evaluating the accuracy of pathogen assignment in this study. These findings are consistent with the observations made during medical evaluations of sharks from both Oceanogràfic and the wild. In fact, the sharks at Oceanogràfic had been housed in the aquarium facility for at least two years and were determined to be healthy. The sharks from the wild were also examined post-mortem, with accidental capture considered as the cause of death. These results support the hypothesis that the microbial profile found in blood plasma is a credible biological phenomenon with no pathological implications. Previous studies on healthy Chondrichthyes have shown the occurrence of positive bacterial cultures in their hosts without any clinical signs, considering non-sterile blood as their baseline condition (Grimes et al., 1985; Buck, 1990; Mylniczenko et al., 2007; Tao et al., 2014; Tarnecki et al., 2018).

Admittedly, we found a similar α- and β-diversity between wild and aquarium animals in blood plasma. Consequently, we can rule out the possibility of external contamination (skin bacteria) in the wild animals where an aseptic field was created by washing with sterile gauze without previous cleaning using alcohol. At the phylum level, blood plasma microbial composition differs between both experimental groups, being dominated in aquarium animals only by *Proteobacteria*, while in addition, *Proteobacteria, Firmicutes, Actinobacteria* and *Bacteroidota* were also present in

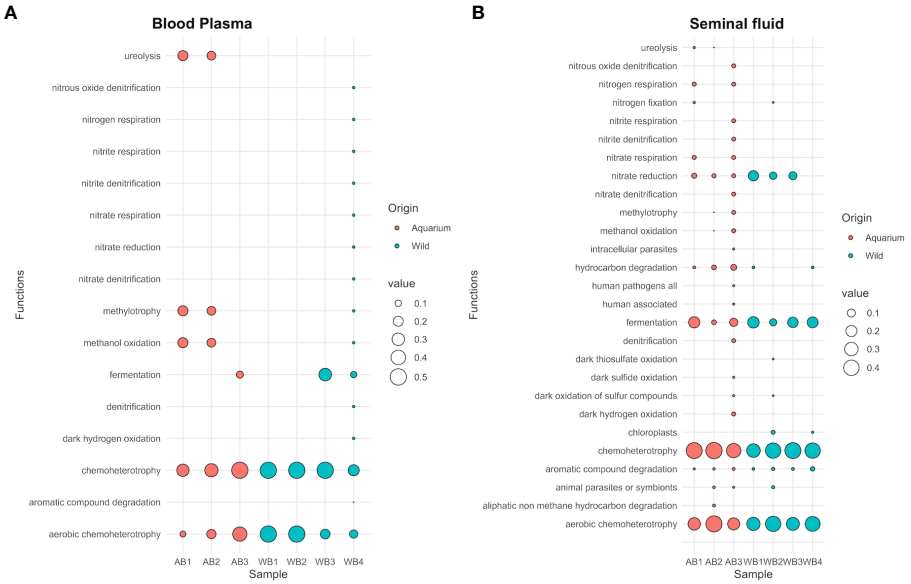


FIGURE 6
Bubble plot showing the relative abundance of functional annotations within the microbial communities for both aquarium and wild small-spotted catshark (*Scyliorhinus canicula*) samples. The plot includes data from **(A)** peripheral blood plasma and **(B)** seminal plasma samples.

wild animals. In the white-spotted eagle ray, *Proteobacteria, Firmicutes, Actinobacteria* and *Bacteroidota* also dominate the microbiota of the cloaca, gills and skin (Clavere-Graciette et al., 2022). Likewise, a substantial environmental influence on microbiome structuring has been documented in marine animals, including the common thresher shark (*Alopias vulpinus*, Doane et al., 2017), the black-tip reef shark (*Carcharhinus melanopterus*, Pogoreutz et al., 2019), killer whales (*Orcinus orca*, Hooper et al., 2019), humpback whales (*Megaptera novaeangliae*, Apprill et al., 2014). Although the environment was not explicitly studied, our findings suggest that the environment has an evident influence on the blood microbiome, confirming the recent results of Clavere-Graciette et al. (2022) in several body compartments in white-spotted eagle rays. *Proteobacteria* has been repeatedly described as the most abundant phylum in marine organisms, environments and water (Pogoreutz et al., 2019; Ruiz-Rodríguez et al., 2020; Serra et al., 2021; Clavere-Graciette et al., 2022). In fact, the shark skin microbiome is predominantly composed of this genus (Pogoreutz et al., 2019; Clavere-Graciette et al., 2022). Likewise, *Proteobacteria* were also the dominant phylum in the gut microbiome of fish and sharks (Givens et al., 2015; Huang et al., 2020) and has been identified as the predominant phylum in teleost fish and water in the Mediterranean Sea (Crespo et al., 2013; Ruiz-Rodríguez et al., 2020). However, other phyla identified in wild teleosts from the Mediterranean Sea were *Fusobacteria* and *Tenericutes*, phyla not identified in this study (Ruiz-Rodríguez et al., 2020). These results would support previous studies claiming that the autochthonous microbes are not a passive reflection of their habitat communities, i.e., each species has its own microbiota (Sullam et al., 2015; Ruiz-Rodríguez et al., 2020; Clavere-Graciette et al., 2022). Of the 14 genera detected in blood microbiome in wild or aquarium, we only found 4 shared genera (making up 45.17% and 51.03% of relative abundance for wild and aquarium sharks, respectively), with *Pseudomonas, Bradyrhizobium, Dysgonomonas* and *Cloacibacterium*. The *Pseudomonas* genus contains well-known pathogenic species, such as *P. aeruginosa*, but also contains species with probiotics potential, widely used in the aquaculture industry (Sehnal et al., 2021). Moreover, *Pseudomonas* and *Cloacibacterium* have been identified in the blood of healthy captive Red Drum (Tarnecki et al., 2018). In addition, *Pseudomonas* has been associated with protein biosynthesis and degradation, involved in nitrate and nitrite ammonification and denitrification, leading to ammonia assimilation and urea decomposition (Doane et al., 2022), as a common characteristic from chondrichthyan metabolism in wild and aquarium environments. Additionally, wild and aquarium animals showed significant differences in blood plasma. For example, *Paracoccus, Anoxybacillus, Monoglobus, Thermicanus*, and *Gordonia* genera were only present in wild animals. At the same time, *Methylobacterium-Methylorubrum, Rhodopseudomonas, Pandoraea, Novosphingobium*, and *Serratia* genera were only present in aquarium catsharks and these differences have been previously associated with the environment and diet influences in microbiome compositional shifts (Clavere-Graciette et al., 2022). *Firmicutes* was the only phylum present in the wild blood plasma microbiome, not present in aquarium sharks. The presence of *Firmicutes* is normally associated with the gastrointestinal bacteria present in the blood of wild animals (Clavere-Graciette et al., 2022). Regarding this fact, we cannot ignore possible bacterial contamination related to the capture of wild animals, as the

substrate contact originated during the trawling, in contact with a highly perfused tissue such as gills, or to the possible breach of the gastrointestinal barrier for bacterial blood contamination.

Here we provide the first description of the taxonomic composition of the seminal plasma present in fish. We found a significantly lower α-diversity in seminal plasma microbiomes in wild organisms compared to aquarium counterparts, but with a similar β-diversity (microbiome composition). As with blood samples, we can rule out the potential external contamination (ampulla bacteria) during sampling in the aquarium animals. Previous studies also described a higher richness and diversity (fecal and skin samples) in wild counterparts compared to under managed care (Uren-Webster et al., 2018: 2020). Nevertheless, contradictory results have recently been found for the white-spotted eagle ray (Clavere-Graciette et al., 2022). These discrepancies could be explained in part by the differences in microbial community composition across body sites (Minich et al., 2020; Clavere-Graciette et al., 2022). In the seminal fluid, the diverse microbiota was predominantly composed of 4 main phyla, which include *Proteobacteria, Firmicutes, Bacteroidetes* and *Actinobacteria* (Koh et al., 2019), also found in a study based on denaturing gradient gel electrophoresis.

The taxonomic composition of each sampling habitat carried a unique signature. At the genus level, collectively, of the 100 genera detected, only 41 were shared between wild and aquarium catsharks. It should be noted that different species reared in the same aquatic environment vary in their microbiome (Li et al., 2015; Sullam et al., 2015; Reverter et al., 2017). Notably, a total of 54 genera (making up 35.2% of the relative abundance) were present only in aquarium animals (Supplementary Table 3), while 5 genera (*Photobacterium, Mycoplasma, Lactobacillus, U.m. family Flavobacteriaceae* and *Promicromonospora*, making up 14.6% of the relative abundance) were only present in the wild small-spotted catsharks (Supplementary Table 3). Such compositional differences may be environmentally driven, in agreement with the results of a previous study by Clavere-Graciette et al. (2022) on white-spotted eagle rays (*Aetobatus narinari*).

Major taxonomic composition shifts were observed (ANCOM approaches) between wild and aquarium small-spotted catsharks in *Serratia*, *Salinisphaera* and *Cutibacterium* genera, being higher in managed care sharks, a factor associated with dysbiosis in other species (Garcia-Segura et al., 2022). *Salinisphaera* was ubiquitous in the Malaysian Mahseer sperm microbiota (Koh et al., 2019). Correspondingly, *Cutibacterium* has been identified in rays and skates (Gonçalves et al., 2020; Mika et al., 2021; Garcia-Segura et al., 2022). Even if this observation has not been put forward in fish studies, significant differences between fertile and infertile men were found in the relative presence of the *Cutibacterium* genus (Garcia-Segura et al., 2022). Likewise, *Serratia*has not been put forward in fish studies, although ovine and porcine seminal plasma samples have tested positive for *Serratia* genera (Althouse and Lu, 2005; Tvrdá et al., 2022). There is some congruence in the identity of bacterial taxa we found in the seminal plasma with other marine fish belonging to different species. Briefly, as part of the seminal plasma microbiome shared between wild and aquarium small-spotted catsharks, an ASV identified as *Mycoplasma* and *Vibrio* are found in other marine fish species (Ciric et al., 2019; Ruiz-Rodríguez et al., 2020). Finally, within the shared seminal plasma microbiome of wild and aquarium sharks, an ASV identified as *Staphylococcus* has been identified in the skin and sting stripes of both wild and aquarium animals (Gonçalves et al., 2020). Notably, *Staphylococcus* is predominantly composed of the sperm microbiota in the Malayan mahseer; although their function is still unclear, but appears to play an essential role in spermatogenesis in fish testis (Kousteni et al., 2010). The higher diversity (in both α- and β-) found in the seminal plasma from both habitats, may be due to the anatomy of the reproductive tract in males, which is directly linked to the exterior through the genital pore, connecting the seminal ampullae with the cloaca (García-Salinas et al., 2021). It is well established that body sites have unique microbial signatures and are differentially influenced by the environment (Minich et al., 2020; Clavere-Graciette et al., 2022). Nevertheless, there is little information about the microbiome of internal fluids in fish; therefore, a meaningful discussion about the effects of specific genera is not yet possible. For instance, further research is needed to understand the role of *Pseudomonas* and *Cloacibacterium*, as both genera were present in the inner fluids of both wild and aquarium spotted sharks.

PICRUSt and FAPROTAX have been recently used to compare the microbiota of dolphins under human care to those of wild animals (Wan et al., 2022). Our results show that, based on the PICRUSt, the blood microbiota in aquarium animals is rich in the glucose-1-phosphate degradation pathway and L-methionine III biosynthesis. This enrichment can be mainly explained the daily food accessibility of the aquarium animals that increased availability of substrates necessary for L-methionine synthesis and glucose metabolism, resulting in heightened activity of genes and enzymes involved in these processes within the bacteria. A recent study by Pinchaud et al. (2022) observed that the urate biosynthesis/5-phosphate-inosine degradation pathways were more prevalent in mice gut microbiota when the animals were fed a lipid-rich diet. Based on the FAPROTAX, our results show that the predominant microbial functions assigned were chemoheterotrophy and aerobic chemoheterotrophy for both inner fluids, while nitrate reduction and fermentation were also predominant in the seminal plasma samples. Further experiments such as metabolomics are needed to verify the functions and roles that these enriched pathways play in aquarium small-spotted catsharks. The potential pathogens mentioned here mainly focus on aquatic wildlife. However, the database for pathogens of marine species is still lacking (Peng et al., 2021). Phylogenetic analysis identified five genera of potentially pathogenic bacteria in the seminal plasma of small-spotted catsharks-associated bacterial communities: *Coxiella*, *Prevotella*, *Coprococcus*, *Haemophilus*, and Phocoenobacter. All of these are recognized potential pathogens for both animals and humans (Louca et al., 2016). Given the absence of evident signs of illness during sampling and the fact that all males exhibit normal *in vitro* sperm quality (Muñoz-Baquero et al., 2021), the presence of potential pathogens in the seminal plasma may indicate a commensal relationship with the ejaculates of small-spotted

catsharks. The findings demonstrated the potential significance of the male genital tract microbiota and its potential implications for the fertility and pathophysiology of small-spotted catsharks.

There are certain limitations to consider in the study. The environmental microbiome analysis for the wild group was not conducted to the unavailability of accurate water samples. Wild small-spotted catsharks inhabit a wide geographical range, spanning from shallow waters to 550 meters. The lack of this information in the analysis prevented us from confirming whether the bacterial composition is influenced by the environment. Additionally, our study only utilized pooled samples, which limited our ability to explore individual microbiota variations. Regarding the wild individual samples, the four-hour refrigerator storage may have led to some modifications in their composition. These changes could have affected our ability to detect similarities in ASV levels between samples. In the case of aquarium animals, both blood and seminal plasma samples may have been minimally contaminated by skin bacteria, even though the area was thoroughly disinfected with Elasmobranch's Ringer solution, as the use of alcohol was not recommended.

In summary, this study provides evidence of a circulating blood and seminal plasma microbiome in healthy small-spotted catsharks. Furthermore, dynamic changes were observed in the microbiome of these inner body fluids, which differed between the wild and aquarium habitats. Additional studies are necessary to identify the physiological traits that contribute to bacterial colonization in different ecosystems among small-spotted catsharks. Moreover, further investigation into the potential association between bacteria found in seminal plasma and male fertility is warranted.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.ncbi.nlm.nih.gov/, PRJNA830653, https://www.ncbi.nlm.nih.gov/, PRJNA974610.

Ethics statement

All the procedures involving aquarium animals in this study were approved by the Animal Care and Welfare Committee of the Oceanogràfic Valencia (Reference number: OCE-18–19) following the aquarium animal care protocol and policies. In the case of wild animals, they were fresh accidental captures with commercial value, donated by local fisheries.

Author contributions

Conceptualization: MM-B and FM-J. Methodology: MM-B, FM-J, LM-P, GM-R, IG-V, LL-R and GD'A. Investigation and data analysis: LL-R, GD'A and FM-J. Writing—original draft: MM-B and FM-J. Writing—review and editing: MM-B, LL-R, FG-V, DG-P, GD'A and FM-J. Supervision: FM-J. All authors contributed to the article and approved the submitted version.

Funding

LL-R was supported by a research grant from the Generalitat Valenciana-Fondo Social Europeo (ACIF/2020/376).

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.2023.1151119/full#supplementary-material

SUPPLEMENTARY FIGURE 1
Negative control using 16S rRNA gene amplicon sequencing (V4 region).

References

Althouse, G. C., and Lu, K. G. (2005). Bacteriospermia in extended porcine semen. *Theriogenology* 63, 573–584. doi: 10.1016/j.theriogenology.2004.09.031

Anhê, F. F., Jensen, B. A. H., Varin, T. V., Servant, F., Van Blerk, S., Richard, D., et al. (2020). Type 2 diabetes influences bacterial tissue compartmentalisation in human obesity. *Nat. Metab.* 2, 233–242. doi: 10.1038/s42255-020-0178-9

Apprill, A., Robbins, J., Eren, A. M., Pack, A. A., Reveillaud, J., Mattila, D., et al. (2014). Humpback whale populations share a core skin bacterial community: towards a health index for marine mammals? *PloS One* 9, e90785. doi: 10.1371/journal.pone.0090785

Bang, C., Dagan, T., Deines, P., Dubilier, N., Duschl, W. J., Fraune, S., et al. (2018). Metaorganisms in extreme environments: do microbes play a role in organismal adaptation? *Zool. (Jena)* 127, 1–19. doi: 10.1016/j.zool.2018.02.004

Bosch, T. C. G., and McFall-Ngai, M. J. (2011). Metaorganisms as the new frontier. *Zool. (Jena)* 114, 185–190. doi: 10.1016/j.zool.2011.04.001

Buck, J. D. (1990). Potentially pathogenic marine vibrio species in seawater and marine animals in the Sarasota, Florida, area. *J. Coast. Res.* 6, 943–948.

Campos, P. M., Darwish, N., Shao, J., and Proszkowiec-Weglarz, M. (2022). Research note: choice of microbiota database affects data analysis and interpretation in chicken cecal microbiota. *Poultry Sci.* 101, 101971. doi: 10.1016/j.psj.2022.101971

Castillo, D. J., Rifkin, R. F., Cowan, D. A., and Potgieter, M. (2019). The healthy human blood microbiome: fact or fiction? *Front. Cell. infect. Microbiol.* 9, 148. doi: 10.3389/fcimb.2019.00148

Ciric, M., Waite, D., Draper, J., and Jones, J. B. (2019). Characterisation of gut microbiota of farmed Chinook salmon using metabarcoding. *Arch. Biol. Sci.* 71 (4), 577–587. doi: 10.2298/ABS190402040C

Clavere-Graciette, A. G., McWhirt, M. E., Hoopes, L. A., Bassos-Hull, K., Wilkinson, K. A., Stewart, F. J., et al. (2022). Microbiome differences between wild and aquarium whitespotted eagle rays (Aetobatus narinari). *Anim. Microbiome* 4, 34. doi: 10.1186/s42523-022-00187-8

Comizzoli, P., Power, M. L., Bornbusch, S. L., and Muletz-Wolz, C. R. (2021). Interactions between reproductive biology and microbiomes in wild animal species. *Anim. Microbiome* 3, 87. doi: 10.1186/s42523-021-00156-7

Contreras, M.J., Núñez-Montero, K., Bruna, P., Zárate, A., Pezo, F., García, M., Leal, K., and Barrientos, L. (2023). Mammals' sperm microbiome: current knowledge, challenges, and perspectives on metagenomics of seminal samples. *Front. Microbiol.* 14, 1167763. doi: 10.3389/fmicb.2023.1167763

Crespo, B. G., Pommier, T., Fernández-Gómez, B., and Pedrós-Alió, C. (2013). Taxonomic composition of the particle-attached and free-living bacterial assemblages in the Northwest Mediterranean Sea analyzed by pyrosequencing of the 16S rRNA. *MicrobiologyOpen* 2, 541–552. doi: 10.1002/mbo3.92

Culligan, E. P., and Sleator, R. D. (2016). Advances in the microbiome: applications to clostridium difficile infection. *J. Clin. Med.* 5 (9), 83. doi: 10.3390/jcm5090083

D'Aquila, P., Giacconi, R., Malavolta, M., Piacenza, F., Bürkle, A., Villanueva, M. M., et al. (2021). Microbiome in blood samples from the general population recruited in the MARK-AGE project: a pilot study. *Front. Microbiol.* 12, 707515. doi: 10.3389/fmicb.2021.707515

Davis, N. M., Proctor, D. M., Holmes, S. P., Relman, D. A., and Callahan, B. J. (2018). Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome* 6, 226. doi: 10.1186/s40168-018-0605-2

Doane, M. P., Haggerty, J. M., Kacev, D., Papudeshi, B., and Dinsdale, E. A. (2017). The skin microbiome of the common thresher shark (Alopias vulpinus) has low taxonomic and gene function β-diversity: skin microbiome of the common thresher shark. *Environ. Microbiol. Rep.* 9, 357–373. doi: 10.1111/1758-2229.12537

Doane, M. P., Johnson, C. J., Johri, S., Kerr, E. N., Morris, M. M., Desantiago, R., et al. (2022). The epidermal microbiome within an aggregation of leopard sharks (Triakis semifasciata) has taxonomic flexibility with gene functional stability across three time-points. *Microbial Ecol.* 7, 1–18. doi: 10.1007/s00248-022-01969-y

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

Ebert, D. A., and Dando, M. (2020). *Field guide to sharks, rays & chimaeras of Europe and the Mediterranean* (US: Princeton University Press).

Federer, W. T. (1994). Pooling and other designs for analysing laboratory samples more efficiently. *J. R. Stat. Society: Ser. D. (The Statistician)* 43 (3), 413–422. doi: 10.2307/2348577

Furstenau, T. N., Cocking, J. H., Hepp, C. M., and Fofanov, V. Y. (2020). Sample pooling methods for efficient pathogen screening: practical implications. *PloS One* 15 (11), e0236849. doi: 10.1371/journal.pone.0236849

García-Salinas, P., Gallego, V., and Asturiano, J. F. (2021). Reproductive anatomy of chondrichthyans: notes on specimen handling and sperm extraction. II. sharks and chimaeras. *Animals* 11, 2191. doi: 10.3390/ani11082191

Garcia-Segura, S., Del Rey, J., Closa, L., Garcia-Martínez, I., Hobeich, C., Castel, A. B., et al. (2022). Seminal microbiota of idiopathic infertile patients and its relationship with sperm DNA integrity. *Front. Cell Dev. Biol.* 10, 937157. doi: 10.3389/fcell.2022.937157

Givens, C., Ransom, B., Bano, N., and Hollibaugh, J. (2015). Comparison of the gut microbiomes of 12 bony fish and 3 shark species. *Mar. Ecol. Prog. Ser.* 518, 209–223. doi: 10.3354/meps11034

Gonçalves, E. S. F., Dos Santos, H. F., de Assis, L. D. C., Lutfi, D. S., Vianna, M., and Rosado, A. S. (2020). Skin and stinger bacterial communities in two critically endangered rays from the south Atlantic in natural and aquarium settings. *MicrobiologyOpen* 9, e1141. doi: 10.1002/mbo3.1141

Grimes, D. J., Brayton, P., Colwell, R. R., and Gruber, S. H. (1985). Vibrios as autochthonous flora of neritic sharks. *Syst. Appl. Microbiol.* 6, 221–226. doi: 10.1016/S0723-2020(85)80056-4

Heberle, H., Meirelles, G. V., da Silva, F. R., Telles, G. P., and Minghim, R. (2015). InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams. *BMC Bioinf.* 16, 169. doi: 10.1186/s12859-015-0611-3

Hooper, R., Brealey, J. C., van der Valk, T., Alberdi, A., Durban, J. W., Fearnbach, H., et al. (2019). Host-derived population genomics data provides insights into bacterial and diatom composition of the killer whale skin. *Mol. Ecol.* 28, 484–502. doi: 10.1111/mec.14860

Huang, Q., Sham, R. C., Deng, Y., Mao, Y., Wang, C., Zhang, T., et al. (2020). Diversity of gut microbiomes in marine fishes is shaped by host-related factors. *Mol. Ecol.* 29, 5019–5034. doi: 10.1111/mec.15699

Human Microbiome Project Consortium (2012). Structure, function and diversity of the healthy human microbiome. *Nature* 486, 207–214. doi: 10.1038/nature11234

Jaspers, C., Fraune, S., Arnold, A. E., Miller, D. J., Bosch, T. C. G., and Voolstra, C. R. (2019). Resolving structure and function of metaorganisms through a holistic framework combining reductionist and integrative approaches. *Zool. Jena Ger.* 133, 81–87. doi: 10.1016/j.zool.2019.02.007

Karin, M., Lawrence, T., and Nizet, V. (2006). Innate immunity gone awry: linking microbial infections to chronic inflammation and cancer. *Cell* 124, 823–835. doi: 10.1016/j.cell.2006.02.016

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

Knowlton, N., and Rohwer, F. (2003). Multispecies microbial mutualisms on coral reefs: the host as a habitat. *Am. Nat.* 162, S51–S62. doi: 10.1086/378684

Koh, I. C. C., Badrul Nizam, B. H., Muhammad Abduh, Y., Abol Munafi, A. B., and Iehata, S. (2019). Molecular characterization of microbiota associated with sperm of Malaysian mahseer tor tambroides. *Evolutionary Bioinf. Online* 15, 117693431985082. doi: 10.1177/1176934319850821

Kousteni, V., Kontopoulou, M., and Megalofonou, P. (2010). Sexual maturity and fecundity of scyliorhinus canicula (Linnaeus 1758) in the Aegean Sea. *Mar. Biol. Res.* 6, 390–398. doi: 10.1080/17451000903233771

Li, T., Long, M., Gatesoupe, F. J., Zhang, Q., Li, A., and Gong, X. (2015). Comparative analysis of the intestinal bacterial communities in different species of carp by pyrosequencing. *Microbial Ecol.* 69, 25–36. doi: 10.1007/s00248-014-0480-8

Louca, S., Parfrey, L. W., and Doebeli, M. (2016). Decoupling function and taxonomy in the global ocean microbiome. *Science* 353, 1272–1277. doi: 10.1126/science.aaf4507

Luna, G. M., Quero, G. M., Kokou, F., and Kormas, K. (2022). Time to integrate biotechnological approaches into fish gut microbiome research. *Curr. Opin. Biotechnol.* 73, 121–127. doi: 10.1016/j.copbio.2021.07.018

Mandal, R. K., Jiang, T., Al-Rubaye, A. A., Rhoads, D. D., Wideman, R. F., Zhao, J., et al. (2016). An investigation into blood microbiota and its potential association with bacterial chondronecrosis with osteomyelitis (BCO) in broilers. *Sci. Rep.* 6, 25882. doi: 10.1038/srep25882

Mandal, S., Van Treuren, W., White, R. A., Eggesbø, M., Knight, R., and Peddada, S. D. (2015). Analysis of composition of microbiomes: a novel method for studying microbial composition. *Microbial Ecol. Health Dis.* 26, 27663. doi: 10.3402/mehd.v26.27663

Metsalu, T., and Vilo, J. (2015). ClustVis: a web tool for visualizing clustering of multivariate data using principal component analysis and heatmap. *Nucleic Acids Res.* 43, W566–W570. doi: 10.1093/nar/gkv468

Mika, K., Okamoto, A. S., Shubin, N. H., and Mark Welch, D. B. (2021). Bacterial community dynamics during embryonic development of the little skate (Leucoraja erinacea). *Anim. Microbiome* 3, 72. doi: 10.1186/s42523-021-00136-x

Minich, J. J., Petrus, S., Michael, J. D., Michael, T. P., Knight, R., and Allen, E. E. (2020). Temporal, environmental, and biological drivers of the mucosal microbiome in a wild marine fish, scomber japonicus. *mSphere* 5, e00401-20. doi: 10.1128/mSphere.00401-20

Moran, M. A. (2015). The global ocean microbiome. *Science* 350, 8455. doi: 10.1126/science.aac8455

Muñoz-Baquero, M., Marco-Jiménez, F., García-Domínguez, X., Ros-Santaella, J. L., Pintus, E., Jiménez-Movilla, M., et al. (2021). Comparative study of semen parameters and hormone profile in small-spotted catshark (Scyliorhinus canicula): aquarium-housed vs. wild-captured. *Animals* 11, 2884. doi: 10.3390/ani11102884

Mylniczenko, N., and Clauss, T. (2017). "Pharmacology of elasmobranchs: updates and techniques," in *The elasmobranch husbandry manual II: recent advances in the care of sharks, rays and their relatives*. Eds. M. D. Smith, D. Warmolts, D. Thoney, R. Hueter, M. Murray and J. Ezcurra (Columbus, OH: Ohio Biological Survey), 289–302.

Mylniczenko, N. D., Harris, B., Wilborn, R. E., and Young., F. A. (2007). Blood culture results from healthy captive and free-ranging elasmobranchs. *J. Aquat. Anim. Health* 19, 159–167. doi: 10.1577/H06-039.1

Patin, N. V., Pratte, Z. A., Regensburger, M., Hall, E., Gilde, K., Dove, A. D. M., et al. (2018). Microbiome dynamics in a Large artificial seawater aquarium. *Appl. Environ. Microbiol.* 84, e00179-18. doi: 10.1128/AEM.00179-18

Peng, S., Hao, W., Li, Y., Wang, L., Sun, T., Zhao, J., et al. (2021). Bacterial communities associated with four blooming scyphozoan jellyfish: potential species-specific consequences for marine organisms and humans health. *Front. Microbiol.* 12, 647089. doi: 10.3389/fmicb.2021.647089

Perry, W. B., Lindsay, E., Payne, C. J., Brodie, C., and Kazlauskaite, R. (2020). The role of the gut microbiome in sustainable teleost aquaculture. *Proc. R. Soc. London Ser. B.* 287, 20200184. doi: 10.1098/rspb.2020.0184

Perry, C. T., Pratte, Z. A., Clavere-Graciette, A., Ritchie, K. B., Hueter, R. E., Newton, A. L., et al. (2021). Elasmobranch microbiomes: emerging patterns and implications for host health and ecology. *Anim. Microbiome* 3, 61. doi: 10.1186/s42523-021-00121-4

Pinchaud, K., Hafeez, Z., Auger, S., Chatel, J. M., Chadi, S., Langella, P., et al. (2022). Impact of dietary arachidonic acid on gut microbiota composition and gut-brain axis in Male BALB/C mice. *Nutrients* 14, 5338. doi: 10.3390/nu14245338

Pogoreutz, C., Gore, M. A., Perna, G., Millar, C., Nestler, R., Ormond, R. F., et al. (2019). Similar bacterial communities on healthy and injured skin of black tip reef sharks. *Anim. Microbiome* 1, 9. doi: 10.1186/s42523-019-0011-5

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

Reverter, M., Sasal, P., Tapissier-Bontemps, N., Lecchini, D., and Suzuki, M. (2017). Characterisation of the gill mucosal bacterial communities of four butterflyfish species: a reservoir of bacterial diversity in coral reef ecosystems. *FEMS Microbiol. Ecol.* 93, 1–10. doi: 10.1093/femsec/fix051

Rowe, M., Veerus, L., Trosvik, P., Buckling, A., and Pizzari, T. (2020). The reproductive microbiome: an emerging driver of sexual selection, sexual conflict, mating systems, and reproductive isolation. *Trends Ecol. Evol.* 35, 220–234. doi: 10.1016/j.tree.2019.11.004

Ruiz-Rodríguez, M., Scheifler, M., Sanchez-Brosseau, S., Magnanou, E., West, N., Suzuki, M., et al. (2020). Host species and body site explain the variation in the microbiota associated to wild sympatric Mediterranean teleost fishes. *Microbial Ecol.* 80, 212–222. doi: 10.1007/s00248-020-01484-y

Scarsella, E., Sandri, M., Monego, S. D., Licastro, D., and Stefanon, B. (2020). Blood microbiome: a new marker of gut microbial population in dogs? *Vet. Sci.* 7, 198. doi: 10.3390/vetsci7040198

Sehnal, L., Brammer-Robbins, E., Wormington, A. M., Blaha, L., Bisesi, J., Larkin, I., et al. (2021). Microbiome composition and function in aquatic vertebrates: small organisms making big impacts on aquatic animal health. *Front. Microbiol.* 12, 567408. doi: 10.3389/fmicb.2021.567408

Serra, C. R., Oliva-Teles, A., Enes, P., and Tavares, F. (2021). Gut microbiota dynamics in carnivorous European seabass (Dicentrarchus labrax) fed plant-based diets. *Sci. Rep.* 11, 447. doi: 10.1038/s41598-020-80138-y

Sullam, K. E., Rubin, B. E. R., Dalton, C. M., Kilham, S. S., Flecker, A. S., and Russell, J. A. (2015). Divergence across diet, time and populations rules out parallel evolution in the gut microbiomes of trinidadian guppies. *ISME J.* 9, 1508–1522. doi: 10.1038/ismej.2014.231

Sze, M. A., Tsuruta, M., Yang, S. W., Oh, Y., Man, S. F., Hogg, J. C., et al. (2014). Changes in the bacterial microbiota in gut, blood, and lungs following acute LPS instillation into mice lungs. *PloS One* 9, e111228. doi: 10.1371/journal.pone.0111228

Tan, C. C., Minghao, C., Ko, K. K., Chen, H., Liu, J., Loh, M., et al. (2023). No evidence for a common blood microbiome based on a population study of 9,770 healthy humans. *Nat. Microbiol.* 8, 973–985. doi: 10.1038/s41564-023-01350-w

Tao, Z., Bullard, S. A., and Arias, C. R. (2014). Diversity of bacteria cultured from the blood of lesser electric rays caught in the northern gulf of Mexico. *J. Aquat. Anim. Health* 26, 225–232. doi: 10.1080/08997659.2014.922513

Tarnecki, A. M., Rhody, N. R., and Walsh, C. J. (2018). Health characteristics and blood bacterial assemblages of healthy captive red drum: implications for aquaculture and fish health management. *J. Aquat. Anim. Health* 30, 339–353. doi: 10.1002/aah.10047

Taylor, M. J., Bordenstein, S. R., and Slatko, B. (2018). Microbe profile: wolbachia: a sex selector, a viral protector and a target to treat filarial nematodes. *Microbiol. (Reading England)* 164, 1345–1347. doi: 10.1099/mic.0.000724

Theodosopoulos, A. N., Hund, A. K., and Taylor, S. A. (2019). Parasites and host species barriers in animal hybrid zones. *Trends Ecol. Evol.* 34, 19–30. doi: 10.1016/j.tree.2018.09.011

Tvrdá, E., Lovíšek, D., Gálová, E., Schwarzová, M., Kováčiková, E., Kunová, S., et al. (2022). Possible implications of bacteriospermia on the sperm quality, oxidative characteristics, and seminal cytokine network in normozoospermic men. *Int. J. Mol. Sci.* 23, 8678. doi: 10.3390/ijms23158678

Uren-Webster, T. M., Consuegra, S., Hitchings, M., and Garcia de Leaniz, C. (2018). Interpopulation variation in the Atlantic salmon microbiome reflects environmental and genetic diversity. *Appl. Environ. Microbiol.* 18, e00691. doi: 10.1128/AEM.00691-18

Uren-Webster, T. M., Rodriguez-Barreto, D., Consuegra, S., and Garcia de Leaniz, C. (2020). Cortisol-related signatures of stress in the fish microbiome. *Front. Microbiol.* 11, 1621. doi: 10.3389/fmicb.2020.01621

Vientós-Plotts, A. I., Ericsson, A. C., Rindt, H., Grobman, M. E., Graham, A., Bishop, K., et al. (2017). Dynamic changes of the respiratory microbiota and its relationship to fecal and blood microbiota in healthy young cats. *PloS One* 12, e0173818. doi: 10.1371/journal.pone.0173818

Wan, X., Li, J., Tian, R., McLaughlin, R.W., Hao, Y., Wu, J., Wang, Z, et al. (2022). The effects of human care on the blowhole and gut microbiotas of two cohabiting dolphin species based on a year-round surveillance. *Front. Mar. Sci.* 9, 1024117. doi: 10.3389/fmars.2022.1024117

Weinberg, E. D. (1974). Iron and susceptibility to infectious disease. *Science* 184, 952–956. doi: 10.1126/science.184.4140.952

Whittle, E., Leonard, M. O., Harrison, R., Gant, T. W., and Tonge, D. P. (2018). Multi-method characterization of the human circulating microbiome. *Front. Microbiol.* 9, 3266. doi: 10.3389/fmicb.2018.03266

Zhou, L., Liu, L., Chen, W., Sun, J., Hou, S., Kuang, T. X., et al. (2020). Stochastic determination of the spatial variation of potentially pathogenic bacteria communities in a large subtropical river. *Environ. Pollut.* 264, 114683. doi: 10.1016/j.envpol.2020.114683

frontiers | Frontiers in Microbiology

TYPE Original Research
PUBLISHED 13 July 2023
DOI 10.3389/fmicb.2023.1087750

OPEN ACCESS

EDITED BY
Senjie Lin,
University of Connecticut, United States

REVIEWED BY
Zheng Zhang,
Shandong University, China
Jinwei Zhang,
University of Exeter, United Kingdom

*CORRESPONDENCE
Chang-Yun Wang
✉ changyun@ouc.edu.cn
Lu Liu
✉ liulu@qibebt.ac.cn

RECEIVED 02 November 2022
ACCEPTED 28 June 2023
PUBLISHED 13 July 2023

CITATION
Pei P-T, Liu L, Jing X-L, Liu X-L, Sun L-Y, Gao C, Cui X-H, Wang J, Ma Z-L, Song S-Y, Sun Z-H and Wang C-Y (2023) Meta-analysis reveals variations in microbial communities from diverse stony coral taxa at different geographical distances.
Front. Microbiol. 14:1087750.
doi: 10.3389/fmicb.2023.1087750

Meta-analysis reveals variations in microbial communities from diverse stony coral taxa at different geographical distances

Peng-Tao Pei[1,2,3,4], Lu Liu[2,3,5]*, Xiao-Li Jing[6], Xiao-Lu Liu[3,5], Lu-Yang Sun[3,5], Chen Gao[3], Xiao-Han Cui[3,5], Jing Wang[3,7], Zhong-Lian Ma[1,4], Shu-Yue Song[1,4], Zhi-Hua Sun[7] and Chang-Yun Wang[1,4]*

[1]Key Laboratory of Marine Drugs, The Ministry of Education of China, Institute of Evolution and Marine Biodiversity, School of Medicine and Pharmacy, Ocean University of China, Qingdao, China, [2]School of Pharmacy, Fujian Health College, Fuzhou, China, [3]Single-Cell Center, Chinese Academy of Science Key Laboratory of Biofuels, Shandong Key Laboratory of Energy Genetics, Shandong Energy Institute, Qingdao New Energy Shandong Laboratory, Qingdao Institute of Bioenergy and Bioprocess Technology, Chinese Academy of Sciences, Qingdao, China, [4]Laboratory for Marine Drugs and Bioproducts, National Laboratory for Marine Science and Technology (Qingdao), Qingdao, China, [5]University of Chinese Academy of Sciences, Beijing, China, [6]High Performance Computing and System Simulation Platform, National Laboratory for Marine Science and Technology (Qingdao), Qingdao, China, [7]Department of Mathematics, Ocean University of China, Qingdao, China

Coral-associated microbial communities play a vital role in underpinning the health and resilience of reef ecosystems. Previous studies have demonstrated that the microbial communities of corals are affected by multiple factors, mainly focusing on host species and geolocation. However, up-to-date, insight into how the coral microbiota is structured by vast geographic distance with rich taxa is deficient. In the present study, the coral microbiota in six stony coral species collected from the coastal area of three countries, including United States of America (USA), Australia and Fiji, was used for analysis. It was found that the geographic influence on the coral microbiota was stronger than the coral host influence, even though both were significant. Interestingly, the contribution of the deterministic process to bacterial community composition increased as geographical distance grew. A total of 65 differentially abundant features of functions in coral microbial communities were identified to be associated with three geolocations. While in the same coastal area of USA, the similar relationship of coral microbiota was consistent with the phylogenetic relationship of coral hosts. In contrast to the phylum Proteobacteria, which was most abundant in other coral species in USA, Cyanobacteria was the most abundant phylum in *Orbicella faveolata*. The above findings may help to better understand the multiple natural driving forces shaping the coral microbial community to contribute to defining the healthy baseline of the coral microbiome.

KEYWORDS

microbiome, stony coral, large geographical distance, diversity analysis, 16S rRNA gene

Introduction

As the most species-rich and abundant biomass marine ecosystem, coral reefs are vital for maintaining ecological balance, developing marine drugs and protecting coastlines (Sharp and Ritchie, 2012; Bourne et al., 2016; Rinkevich, 2021). The symbiont, particularly special microbial communities, was recognized as crucially important in both organism function and ecosystem health (Hernandez-Agreda et al., 2017). Microbial communities participate in the nutrient cycling of corals by producing glucose and taking part in the metabolism of nitrogenous and sulfurous substances and provide important protective services for corals by generating antibiotics and competing with pathogens (Pantos et al., 2015; Ziegler et al., 2017). Coral core microbiome is a stable system across hosts, space and time in the communities (Ainsworth and Gates, 2016; Hernandez-Agreda et al., 2016, 2017). While the overall coral microbiome are influenced by various factors, including coral host and geolocation.

According to previous studies, the coral microbial community composition has been demonstrated to change with coral hosts. Earlier studies showed that the coral host was the strongest driving force for the microbial community structures of six Caribbean corals over place and time (Chu and Vollmer, 2016). Later, host identity was proven to play a dominant role in structuring the microbiome of Caribbean corals, while microbial community dissimilarity increased with geographical distance (Dunphy et al., 2019). Recently, the coral host was also reported as the main driver for structural differences in microbial communities in reef corals of the Malay Peninsula (Kanisan et al., 2022). In addition, a study on *Pocillopora verrucosa* and *Turbinaria peltata*, endemic corals in the South China Sea, showed that both host and geographical factors affect coral microbial communities (Chen et al., 2021).

However, other previous studies suggested that the different viewpoints of coral microbiota vary through geolocation. It was found that the microbial communities of five *Acropora loripes* in the four sites of Davies reef have significant differences (Damjanovic et al., 2020). Even at a small distance, the coral microbiota in *Pocillopora acuta* collected from the offshore islands of Singapore were significantly different (Wainwright et al., 2019). These studies revealed that it is necessary to accumulate comprehensive and in-depth knowledge about the variation pattern of the coral microbiome structured by complex driving forces, especially the global range incorporating diverse coral taxa. There is a high possibility of missing the information on driving influence with broad background based on the findings from individual studies using limited locations and taxa.

To date, research on multiple driving forces in shaping coral microbial communities is still insufficient, especially at large geographical distances. In this study, a meta-analysis of six stony coral species from the Pacific Ocean and the Atlantic Ocean with a distance of more than 2,700 kilometers was used to study the coral microbiota variation associated with coral hosts and a large geographic distance. The similarity relationship of microbial communities in the same sea coastal area was constructed and compared with the phylogenetic relationship of coral hosts. The functional traits of coral microbial communities related to geolocation were also discussed.

Results

Sequence information

After a series of processing, a total of 21,997,761 high-quality sequences were obtained from 307 samples from three countries (USA: 24–26°N, 80–82°W, Australia: 23–24°S, 151–152°E, Fiji: 18.13°S, 177.42°E; Table S1 for detailed information): 62 samples of *Montastraea cavernosa*, 25 samples of *Orbicella faveolata*, 128 samples of *Pocillopora damicornis*, 36 samples of *Porites astreoides*, 34 samples of *Porites* and 22 samples of *Pseudodiploria strigosa*. A total of 71,182 amplicon sequence variants (ASV) were obtained by clustering the high-quality sequences using DADA2.

Structural variation pattern of stony coral microbiota in multiple coral species across a large geographical distance

To understand the effects on variations in microbial communities from diverse coral taxa across a large geographical distance, the alpha- and beta-diversity of the microbiome were calculated based on the resampled ASV tables. Jaccard distance-based principal coordinate analysis (PCoA) showed that geolocation has a strong and significant impact on the microbial community composition of stony corals (PERMANOVA, permutations = 999, Pseudo-F = 21.46, R^2 = 0.1238, p = 0.001; Figure 1A). Significant differences in microbial community structures were also observed at the coral family and species levels (family level: PERMANOVA, permutations = 999, Pseudo-F = 5.701, R^2 = 0.0951, p = 0.001; species level: PERMANOVA, permutations = 999, Pseudo-F = 6.784, R^2 = 0.0828, p = 0.001). It should be noted that the large geographical distance had a stronger impact on the coral microbiota than the host taxon. The family- and species-level coral hosts had a similar impact degree. Moreover, the hierarchical clustering analysis revealed that all samples were classified into three prominent groups corresponding to the geolocations of Australia, Fiji and USA (Figure 1B). The above results showed that geolocation was the main driving force for the microbial communities of the six coral species across a large geographical distance. Additionally, the β-nearest taxon index (βNTI) analysis revealed that deterministic processes became increasingly influential in shaping the structure of the bacterial community with distance growth (Figure 2).

Furthermore, the Shannon index and observed features index in alpha diversity were calculated, as shown in Figure 3. Kruskal–Wallis pairwise analysis results showed that both Shannon and observed features index (Shannon, mean = 5.1494 ± 2.5049 s.d.; observed features, mean = 448.8101 ± 392.9550 s.d.) of the coral microbiome in the northern subtropical group (USA) was significantly higher than that in the tropical group (Fiji) (Shannon, mean = 3.6470 ± 0.3671 s.d.; observed features, mean = 47.0952 ± 21.1652 s.d.) and southern subtropical group (Australia) (Shannon, mean = 1.2665 ± 0.3655 s.d.; observed features, mean = 40.2045 ± 14.9675 s.d.) (p < 0.001). These results suggested that the species diversity, evenness and ASV numbers of the coral microbial communities in the northern subtropical group were significantly higher than those in the tropical and southern subtropical groups, while those in the tropical group were significantly higher than those in the southern subtropical group.

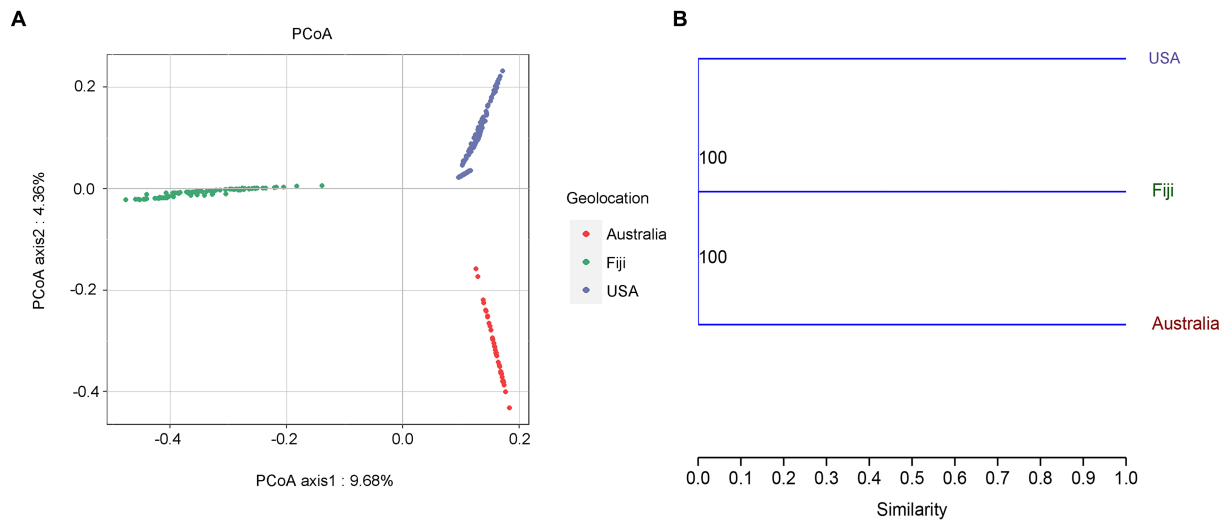


FIGURE 1
Principal coordinates analysis (PCoA) based on jaccard distance **(A)**. A hierarchical clustering tree of the coral microbial community based on geolocations was constructed by the unweighted pair group method with arithmetic mean (UPGMA). Bootstrap values based on percentage agreement with resampled ASV tables **(B)**.

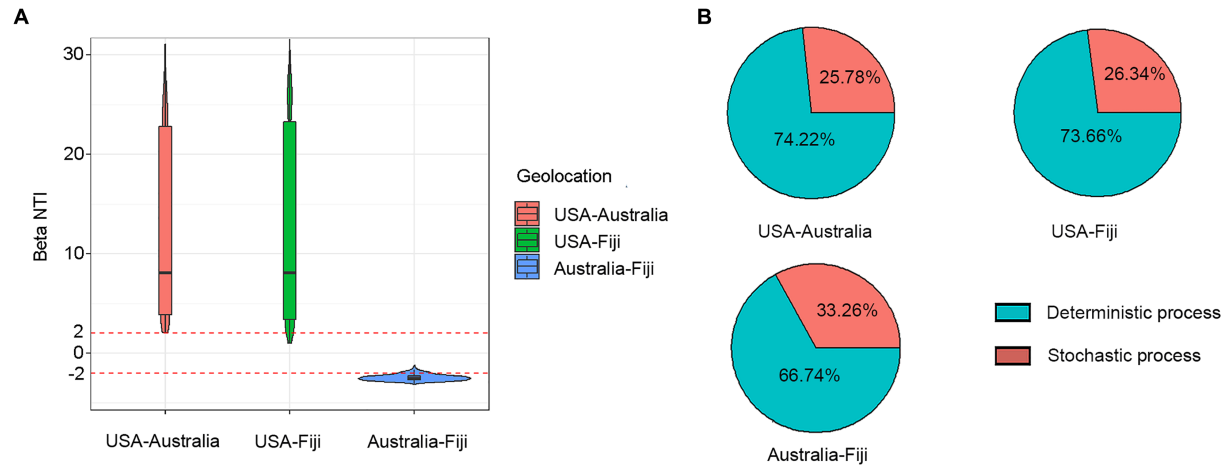


FIGURE 2
Variation in βNTI median at different geolocations **(A)**. Relative contribution of deterministic and stochastic processes to coral microbial community composition between USA, Australia, and Fiji **(B)**.

Structural variation pattern of stony coral microbiota in multiple coral species in the same geographical region

To research the variations in coral microbial communities in the same geographical region, microbiome samples collected from five coral species in the same sea area of USA with a distance of less than 230 kilometers were analyzed. The hierarchical clustering analysis suggested that the similarity of coral microbial community composition could reflect the phylogeny of coral hosts. Except for *P. porites* corals excluded in the phylogenetic tree of a previous report (Quek and Huang, 2019), the similarity relationships of coral microbiota from four coral species were completely consistent with the phylogenetic relationships of coral hosts (Figure 4). The above results revealed that corals with closer phylogenetic relationships could exhibit more similar microbiota compositions in the same geographical region.

Proteobacteria and Cyanobacteria were the dominant phyla in all five coral species, accounting for 52.74 and 14.21% of the relative abundance of total prokaryotes, respectively (Figure 5A). The three most abundant phyla in Porites coral were Proteobacteria, Cyanobacteria and unidentified phyla. The order Rickettsiales was most abundant in *M. cavernosa*, *P. porites* and *P. astreoides*, corresponding to relative abundances of 49.22, 19.08 and 6.17%, respectively, while it was no more than 2.2% in other coral species. Despite permeability to NAD, Rickettsiales was also suggested to be a pathogenic factor that could cause coral diseases in previous studies (Casas et al., 2004; Garcia et al., 2013; Angly et al., 2016; Driscoll et al.,

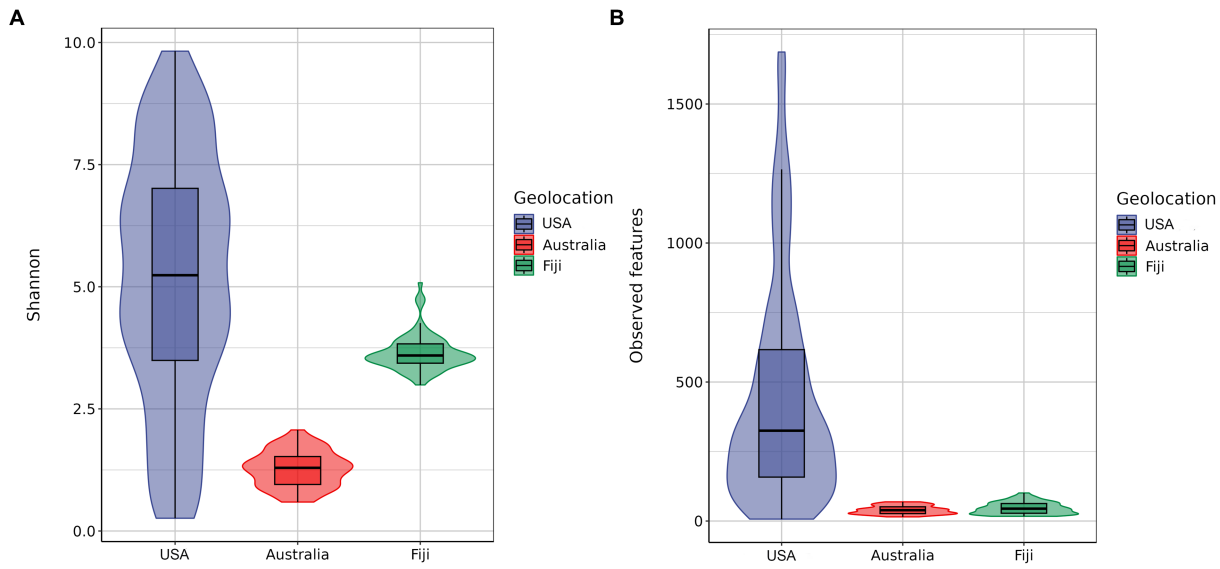


FIGURE 3
Alpha diversity based on the Shannon diversity index **(A)**. Alpha diversity based on the observed feature diversity index **(B)**.

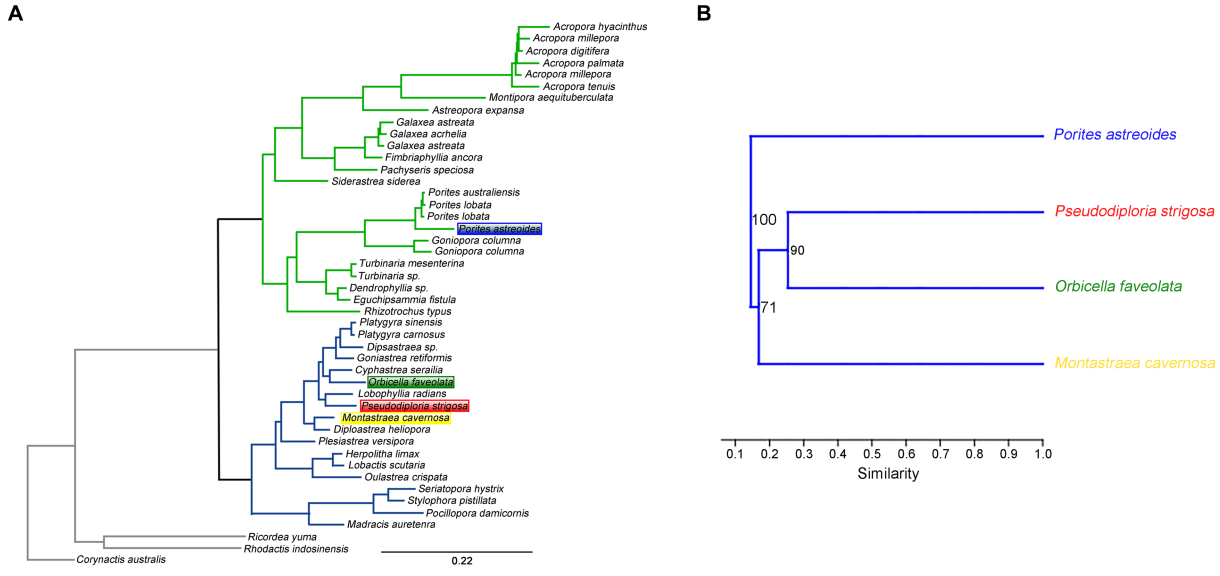


FIGURE 4
The phylogenetic relationship of stony coral species **(A)**. The similarity of coral microbial community structures constructed by hierarchical clustering **(B)**.

2017; Klinges et al., 2019). Notably, the unidentified phylum was found to account for 12.98% of the total prokaryotic abundance in coral Porites.

The three most abundant phyla in corals *O. faveolata* and *P. strigosa* were Proteobacteria, Bacteroidetes and Cyanobacteria, respectively. It was interesting that the most abundant phylum in coral *O. faveolata* was Cyanobacteria (35.96%), followed by Proteobacteria (32.54%), which was obviously different from the other four coral species. Cyanobacteria were proven to be responsible for nitrogen fixation for corals (Wiebe et al., 1975; Wang et al., 2000; Lesser et al., 2004).

The three most abundant phyla in coral *M. cavernosa* were Proteobacteria, Cyanobacteria and Bacteroidetes. According to the linear discriminant analysis effect size (LEfSe) analysis (Supplementary Table S1), *M. cavernosa* harbored three characteristic bacteria, Proteobacteria, Crenarchaeota and Nitrospirae, at the phylum level compared with the other coral species (Figure 5B). The genus *Nitrosopumilus* belonging to Crenarchaeota was observed to account for 6.68% in coral *M. cavernosa*, while it is almost nonexistent in other corals. *Nitrosopumilus* is an aerobic bacterium, which can efficiently convert ammonia into nitrate (Könneke et al., 2005; Siboni et al., 2008; Jensen et al., 2012; Chekidhenkuzhiyil et al., 2021; Polónia

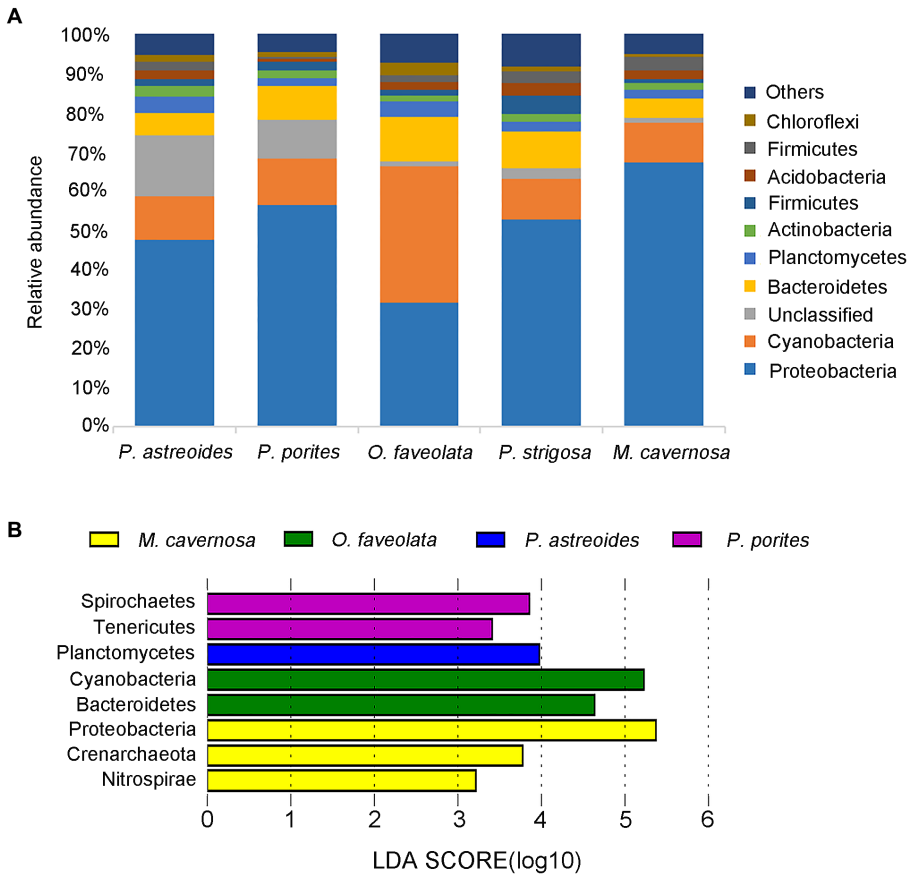


FIGURE 5
Relative phylum-level abundance of microbial communities for 5 corals **(A)**. LDA value distribution map of different microorganisms. The length of the histogram represents the contribution of different microorganisms (LDA score). The figure shows the microorganisms with significant differences in abundance between different groups when the LDA score was greater than the set value (set to 3). The coral species are highlighted with different colors **(B)**.

et al., 2021). Nitrospirae is an aerobic bacterium that efficiently converts ammonia into nitrate and is reported to be a common aerobic symbiont in sponges (Hoffmann et al., 2009; Han et al., 2013; Karlińska-Batres and Wörheide, 2013).

Functional traits of stony coral microbiota in multiple coral species across a large geographical distance

To further understand the distance-based shifts in microbial communities, the microbial functional traits from stony corals in three countries were elucidated by LEfSe analysis (Supplementary Table S1; Supplementary Figure S1). A total of 65 differentially abundant features of functions in coral microbial communities were identified to be associated with three geolocations. In the southern subtropical group, 37 functional traits, especially several degradation characteristic functions, e.g., degradation of organic acids, nucleosides, acetylene and sucrose, were enriched. In the northern subtropical group, 21 functional traits involved in metabolic pathways, e.g., pentose phosphate pathway and glycolysis, were revealed to be more abundant. In the tropical group, 7 functional traits, such as biotin and organic acid synthesis, were more abundant. In addition, according to the Venn diagram analysis, 168 functions presented in all three groups could be suggested as common functions, accounting for more than 70.29% of all predicted functions (Figure 6). Six functional traits were shared by the northern and southern subtropical groups, including biosynthetic pathways. A total of 38 functional traits were common between the southern subtropical group and tropical group, such as superpathway for organic synthesis. Interestingly, there were no common functional traits between the northern subtropical group and the tropical group. Only 23, 3 and 1 functions were unique to the southern subtropical, northern subtropical and tropical groups, respectively.

Discussion

Diverse microbial communities play a crucial role in maintaining and shifting the coral holobiont equilibrium. The small conserved coral core microbiome is consistently stable in the communities (Ainsworth and Gates, 2016; Hernandez-Agreda et al., 2016, 2017). However, the overall coral microbiome is dynamic. The detrimental shift in microbial community composition could promote the invasion of opportunistic pathogens and cause coral mortality. Thus, it is necessary to comprehensively and deeply understand the normal

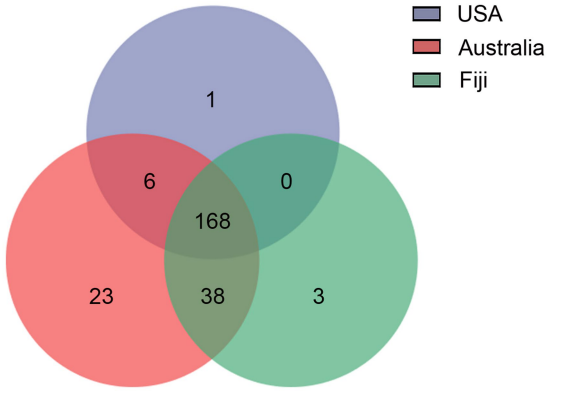


FIGURE 6
Venn diagram of functional traits of coral microbial communities associated with a large geographical distance.

variation pattern of the coral microbiome driven by multiple factors to accurately distinguish abnormal changes. Until now, studies researching the composition variation of coral microbial communities have focused on specific geographic regions and coral taxa while ignoring variation patterns by multiple driving forces with various coral taxa from a macro perspective. Previous studies have suggested that both host taxa and geolocation could become the driving forces of coral microbial community structure. However, it presented different conclusions according to the individual studies, even contradictorily (Damjanovic et al., 2020; Kanisan et al., 2022). Recently, it was revealed that microbial community dissimilarity increased with geographical distances of less than 30 kilometers, but the coral host taxon still played a leading role (Dunphy et al., 2019).

In our study, a meta-analysis was used to research the microbiome variation patterns influenced by different geographic distances and diverse host taxa. The results showed that the influence of geolocation factors exceeding the host taxon became the dominant driving force on the coral microbial community at a sufficiently large geographic distance. However, the host taxon was the main influence on the variations in coral microbial community composition in the same coastal area. Moreover, the deterministic process was found to dominate microbial community composition at a larger geographic distance rather than stochastic process, and even increase with distance growth, which was consistent with the previous research (Shi et al., 2018). The result suggested that most ASVs could not interchange between two distinctively environmental condition, leading to predictable differences of their community composition. In contrast, most ASVs could be surviving in the same coastal area where the environmental condition was sufficiently similar. Exploring the variation pattern of a large geographic region with diverse coral taxa would help to provide comprehensive basic knowledge to identify the various healthy states of the coral microbiome.

In addition, the phylum Proteobacteria, as a dominant taxon reported by other studies, was also found to be the most abundant in the coral microbiome of five coral species from USA, Australia and Fiji. Only the phylum Cyanobacteria was the most abundant phylum in the coral *O. faveolata* of USA rather than Proteobacteria (Rubio-Portillo et al., 2018; Wainwright et al., 2019). Cyanobacteria have built coral reefs on Earth for more than three million years. In modern coral reef ecosystems, Cyanobacteria are still ubiquitous, forming a major component and enriching the ecosystem by providing nitrogen through nitrogen fixation. However, Cyanobacteria could also associate with other microbes to play a detrimental role in causing coral tissue lysis and death (Charpy et al., 2012). Coral *O. faveolata*, as a special case, may provide a pathway to promote understanding of the functional evolution mechanism of Cyanobacteria.

It should be pointed out that the structures of coral microbial communities could be affected by diverse factors, including coral hosts, geolocations, coral conditions (health or disease), seasons, sampling methods, and environmental changes (Li et al., 2014; Sharp et al., 2017; Dunphy et al., 2019; Gardner et al., 2019; Vega et al., 2020; Williams et al., 2022). According to the investigation of previous studies for coral microbiome, it was found that the top three most studies are related to coral hosts, geolocations, and coral conditions. In the literature about coral conditions, more than 90% of studies showed that health/disease conditions had a significant effect on the structures of coral microbial communities. Furthermore, the previous research also suggested that 16S rRNA gene amplicon data from different sequencing platforms, and amplified regions could affect resolution and microbial diversity (Schmalenberger et al., 2001; Engelbrektson et al., 2010; Kellogg, 2019). Considering that the metadata about coral hosts, geolocations, coral conditions, sequencing platforms, and amplified regions was provided in most previous studies, yet the metadata about seasons and environmental changes was rarely provided. Therefore, in this study, the filtering conditions of data involved the same sequencing platform, amplified region, and coral condition. After data filtering, the number of coral samples in each geolocation was not exactly the same. In addition, the natural distribution of coral species varied on a large geographical distance. In view of this, the method of PERMANOVA, which is non-parametric, was used for analysis in our study. PERMANOVA is used to compare differences among groups in multivariate datasets without making specific assumptions about the data distribution. Furthermore, we used data from coral adults and larvae of one study (Beatty et al., 2018). Fortunately, it was found that the PERMANOVA statistical results for microbiome of only coral adults were consistent with those of both coral adults and larvae (Supplementary Table S1; Supplementary Figure S2). While our data may not be flawless, the analysis could still provide insights into prevailing trends. In the future, it is necessary to continue efforts to provide detailed information and expand the database of coral microbial communities to identify various patterns of the healthy state. These findings contribute to defining the healthy baseline of the coral microbiome and open the vast possibilities for coral microbiome application in diagnosis and intervention.

Materials and methods

Sample data

16S rRNA gene amplicon samples involving six stony coral species collected from three previous publications were merged and reanalyzed (Apprill et al., 2016; Beatty et al., 2018; van Oppen et al., 2018). These original sequence data can be obtained from the NCBI

Sequence Read Archive under BioProjects PRJNA324417, PRJNA380169, PRJNA382809, and PRJNA872539 of NCBI and the publication of original papers. The stony coral samples across these four studies to sequence microbial amplicons were taken from coral tissue, mucus and holobiont. The V4 hypervariable region of the 16S rRNA gene was amplified with primers, and then the samples were sequenced on the Illumina platform (Supplementary Table S1).

Bioinformatics analysis

For the sequences converted to fastq format, Trim Galore v.0.6.6[1] software was used to filter the sequences with low scores and short lengths maintaining a Phred score threshold of 20 and a length threshold of 100 bp. The paired sequences after quality control were combined using FLASH v.1.2.11 (Magoč and Salzberg, 2011) with a merge rate of more than 80%, then the abiotic sequences were removed with VSEARCH v.2.7.0 (Rognes et al., 2016). FastQC v.0.11.9[2] software was used to evaluate the quality of data, and unqualified samples were removed. Later, samples with sequencing depth less than 5,000 were filtered. Downstream bioinformatics analysis was performed using QIIME2 v.2020.11.1 (Bolyen et al., 2019). High-quality sequences were clustered into ASV, and Greengenes release 13_8 was used to annotate the sequence as the reference database (DeSantis et al., 2006; McDonald et al., 2012).

Statistical analysis

Prior to calculating the diversity index, samples were randomly rarefied to the minimum library size (5,800 sequences) with QIIME2 v.2020.11.1 due to large differences in library sizes between different samples. Alpha diversity metrics, including the Shannon index and observed features index, were calculated using QIIME2 v.2020.11.1. The mean value and standard deviation of each metric group were calculated by SPSS v.25 software. Differences in beta diversity were computed using Jaccard dissimilarity matrices and tested via permutational multivariate analysis of variance (PERMANOVA). PERMANOVA analysis was performed in R v.4.2.1 using the vegan package. Variation in community composition among samples was visualized with PCoA. Hierarchical clustering analysis was obtained by using the unweighted pair group method with arithmetic mean (UPGMA) clustering in PAST v.4.0 software, and bootstrap values were based on percentage agreement with the resampled ASV table. All βNTI values were conducted using Picante in R v.4.2.1. If βNTI >2 or βNTI < − 2, deterministic processes were important in shaping the community composition, whereas if values were between −2 < βNTI <2, stochastic processes would play an important role. The zero model was used to quantify the deviation from absolute phylogenetic distance to stochastic phylogenetic distance in this method. The greater the degree of deviation, the more the community was affected by deterministic factors, and the smaller the degree of deviation, the more the community was affected by random factors (Kembel et al., 2010). The phylogenetic tree in Figure 4A was constructed with RAxML v.8.2.11 by Quek in previous study (Stamatakis, 2014; Quek and Huang, 2019).

Functional profiling based on the microbial community

To better understand the potential functional profiles of the specific microbial taxa in reef habitats, Phylogenetic Investigation of Communities by Reconstruction of Unobserved States (PICRUSt) was applied (Douglas et al., 2020). PICRUSt2 can predict metagenomic functional composition from 16S rRNA gene using marker gene data and a reference genome database. The ASV table of microbial communities with large geographical distances was imported into the "Picrust2_pipelines.py" script in QIIME2 format. The enzyme commission (EC) metagenomes, KEGG orthology (KO) metagenomes and MetaCyc pathway abundances were predicted. The weighted nearest sequenced taxon index (NSTI) score for each sample was calculated (mean NSTI = 0.29 ± 0.16 s.d.). The LEfSe method was used to further analyze the metagenomic prediction count table to identify significantly different functional traits of microbial communities across a large geographical distance [Linear discriminant analysis (LDA) > 3.0 for individual KOs].[3] A Venn diagram was drawn to highlight shared functions among the different geolocations.

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

C-YW and LL conceived the study. P-TP, X-LL, CG, Z-LM, S-YS, X-LJ, and Z-HS collected the samples. P-TP, LL, X-LL, X-HC, and JW analyzed the data and prepared the figures and tables. C-YW, LL, P-TP, and L-YS wrote the manuscript. All authors commented and contributed to improving the manuscript and preparing the manuscript for submission, read and approved the final manuscript.

Funding

This work was supported by the Key Program of National Natural Science Foundation of China (No. 41830535), the Shandong Natural Science Foundation, China (ZR2019BC087), the Major Project of Qingdao Marine Science and Technology Center (2022QNLM030003-1), the Program of Open Studio for Druggability Research of Marine Natural Products, National Laboratory for Marine Science and Technology (Qingdao, China) Directed by Kai-Xian Chen and Yue-Wei Guo, and the Taishan Scholars Program, China.

1 https://github.com/FelixKrueger/TrimGalore

2 https://www.bioinformatics.babraham.ac.uk/projects/fastqc

3 https://cloud.oebiotech.cn/task/detail/LEfSe-oehw/

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.2023.1087750/full#supplementary-material

References

Ainsworth, T. D., and Gates, R. D. (2016). Ocean biology. Corals' microbial sentinels. *Science* 352, 1518–1519. doi: 10.1126/science.aad9957

Angly, F. E., Heath, C., Morgan, T. C., Tonin, H., Rich, V., Schaffelke, B., et al. (2016). Marine microbial communities of the great barrier reef lagoon are influenced by riverine floodwaters and seasonal weather events. *PeerJ.* 4:e1511. doi: 10.7717/peerj.1511

Apprill, A., Weber, L. G., and Santoro, A. E. (2016). Distinguishing between microbial habitats unravels ecological complexity in coral microbiomes. *mSystems.*:1. doi: 10.1128/mSystems.00143-16

Beatty, D. S., Clements, C. S., Stewart, F. J., and Hay, M. E. (2018). Intergenerational effects of macroalgae on a reef coral: major declines in larval survival but subtle changes in microbiomes. *Mar. Ecol. Prog. Ser.* 589, 97–114. doi: 10.3354/meps12465

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., et al. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* 37, 852–857. doi: 10.1038/s41587-019-0209-9

Bourne, D. G., Morrow, K. M., and Webster, N. S. (2016). Insights into the coral microbiome: underpinning the health and resilience of reef ecosystems. *Annu. Rev. Microbiol.* 70, 317–340. doi: 10.1146/annurev-micro-102215-095440

Casas, V., Kline, D. I., Wegley, L., Yu, Y., Breitbart, M., and Rohwer, F. (2004). Widespread association of a Rickettsiales-like bacterium with reef-building corals. *Environ. Microbiol.* 6, 1137–1148. doi: 10.1111/j.1462-2920.2004.00647.x

Charpy, L., Casareto, B. E., Langlade, M. J., and Suzuki, Y. (2012). Cyanobacteria in coral reef ecosystems: a review. *J. Mar. Biol.* 2012, 1–9. doi: 10.1155/2012/259571

Chekidhenkuzhiyil, J., Anas, A., Thomas, P. A., Tharakan, B., and Nair, S. (2021). Characterization of archaeal symbionts of sponges from the coral reef ecosystems of the Gulf of Mannar, southeast coast of India. *Saudi J Biol Sci.* 28, 3783–3788. doi: 10.1016/j.sjbs.2021.03.054

Chen, B., Yu, K., Liao, Z., Yu, X., Qin, Z., Liang, J., et al. (2021). Microbiome community and complexity indicate environmental gradient acclimatisation and potential microbial interaction of endemic coral holobionts in the South China Sea. *Sci. Total Environ.* 765:142690. doi: 10.1016/j.scitotenv.2020.142690

Chu, N. D., and Vollmer, S. V. (2016). Caribbean corals house shared and host-specific microbial symbionts over time and space. *Environ. Microbiol. Rep.* 8, 493–500. doi: 10.1111/1758-2229.12412

Damjanovic, K., Blackall, L. L., Peplow, L. M., and van Oppen, M. J. H. (2020). Assessment of bacterial community composition within and among *Acropora loripes* colonies in the wild and in captivity. *Coral Reefs* 39, 1245–1255. doi: 10.1007/s00338-020-01958-y

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

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

Driscoll, T. P., Verhoeve, V. I., Guillotte, M. L., Lehman, S. S., Rennoll, S. A., Beier-Sexton, M., et al. (2017). Wholly Rickettsia! Reconstructed metabolic profile of the quintessential bacterial parasite of eukaryotic cells. *MBio* 8:e00859-17. doi: 10.1128/mBio.00859-17

Dunphy, C. M., Gouhier, T. C., Chu, N. D., and Vollmer, S. V. (2019). Structure and stability of the coral microbiome in space and time. *Sci. Rep.* 9:6785. doi: 10.1038/s41598-019-43268-6

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

Garcia, G. D., Gregoracci, G. B., Santos, E. O., Meirelles, P. M., Silva, G. G., Edwards, R., et al. (2013). Metagenomic analysis of healthy and white plague-affected *Mussismilia braziliensis* corals. *Microb. Ecol.* 65, 1076–1086. doi: 10.1007/s00248-012-0161-4

Gardner, S. G., Camp, E. F., Smith, D. J., Kahlke, T., Osman, E. O., Gendron, G., et al. (2019). Coral microbiome diversity reflects mass coral bleaching susceptibility during the 2016 El Nino heat wave. *Ecol. Evol.* 9, 938–956. doi: 10.1002/ece3.4662

Han, M., Li, Z., and Zhang, F. (2013). The ammonia oxidizing and denitrifying prokaryotes associated with sponges from different sea areas. *Microb. Ecol.* 66, 427–436. doi: 10.1007/s00248-013-0197-0

Hernandez-Agreda, A., Gates, R. D., and Ainsworth, T. D. (2017). Defining the Core microbiome in Corals' microbial soup. *Trends Microbiol.* 25, 125–140. doi: 10.1016/j.tim.2016.11.003

Hernandez-Agreda, A., Leggat, W., Bongaerts, P., and Ainsworth, T. D. (2016). The microbial signature provides insight into the mechanistic basis of coral success across reef habitats. *MBio* 7:7. doi: 10.1128/mBio.00560-16

Hoffmann, F., Radax, R., Woebken, D., Holtappels, M., Lavik, G., Rapp, H. T., et al. (2009). Complex nitrogen cycling in the sponge *Geodia barretti*. *Environ. Microbiol.* 11, 2228–2243. doi: 10.1111/j.1462-2920.2009.01944.x

Jensen, S., Bourne, D. G., Hovland, M., and Murrell, J. C. (2012). High diversity of microplankton surrounds deep-water coral reef in the Norwegian Sea. *FEMS Microbiol. Ecol.* 82, 75–89. doi: 10.1111/j.1574-6941.2012.01408.x

Kanisan, D. P., Quek, Z. B. R., Oh, R. M., Afiq-Rosli, L., Lee, J. N., Huang, D., et al. (2022). Diversity and distribution of microbial communities associated with reef corals of the Malay peninsula. *Microb. Ecol.* 85, 37–48. doi: 10.1007/s00248-022-01958-1

Karlińska-Batres, K., and Wörheide, G. (2013). Microbial diversity in the coralline sponge Vaceletia crypta. *Antonie Van Leeuwenhoek* 103, 1041–1056. doi: 10.1007/s10482-013-9884-6

Kellogg, C. A. (2019). Microbiomes of stony and soft deep-sea corals share rare core bacteria. *Microbiome* 7:90. doi: 10.1186/s40168-019-0697-3

Kembel, S.W., Cowan, P. D., Helmus, M. R., Cornwell, W. K., Morlon, H., Ackerly, D. D., et al. (2010). Picante: R tools for integrating phylogenies and ecology. *Bioinformatics.* 26, 1463–4. doi: 10.1093/bioinformatics/btq166

Klinges, J. G., Rosales, S. M., McMinds, R., Shaver, E. C., Shantz, A. A., Peters, E. C., et al. (2019). Phylogenetic, genomic, and biogeographic characterization of a novel and ubiquitous marine invertebrate-associated Rickettsiales parasite, Candidatus *Aquarickettsia rohweri*, gen. Nov., sp. nov. *ISME J.* 13, 2938–2953. doi: 10.1038/s41396-019-0482-0

Könneke, M., Bernhard, A. E., de la Torre, J. R., Walker, C. B., Waterbury, J. B., and Stahl, D. A. (2005). Isolation of an autotrophic ammonia-oxidizing marine archaeon. *Nature* 437, 543–546. doi: 10.1038/nature03911

Lesser, M. P., Mazel, C. H., Gorbunov, M. Y., and Falkowski, P. G. (2004). Discovery of symbiotic nitrogen-fixing cyanobacteria in corals. *Science* 305, 997–1000. doi: 10.1126/science.1099128

Li, J., Chen, Q., Long, L. J., Dong, J. D., Yang, J., and Zhang, S. (2014). Bacterial dynamics within the mucus, tissue and skeleton of the coral *Porites lutea* during different seasons. *Sci. Rep.* 4:7320. doi: 10.1038/srep07320

Magoč, 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

McDonald, D., Price, M. N., Goodrich, J., Nawrocki, E. P., DeSantis, T. Z., Probst, A., et al. (2012). An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *ISME J.* 6, 610–618. doi: 10.1038/ismej.2011.139

Pantos, O., Bongaerts, P., Dennis, P. G., Tyson, G. W., and Hoegh-Guldberg, O. (2015). Habitat-specific environmental conditions primarily control the microbiomes of the coral *Seriatopora hystrix*. *ISME J.* 9, 1916–1927. doi: 10.1038/ismej.2015.3

Polónia, A. R. M., Cleary, D. F. R., Gauvin-Bialecki, A., and de Voogd, N. J. (2021). Archaeal communities of low and high microbial abundance sponges inhabiting the remote western Indian Ocean island of Mayotte. *Antonie Van Leeuwenhoek* 114, 95–112. doi: 10.1007/s10482-020-01503-5

Quek, Z. B. R., and Huang, D. (2019). Effects of missing data and data type on phylotranscriptomic analysis of stony corals (Cnidaria: Anthozoa: Scleractinia). *Mol. Phylogenet. Evol.* 134, 12–23. doi: 10.1016/j.ympev.2019.01.012

Rinkevich, B. (2021). Augmenting coral adaptation to climate change via coral gardening (the nursery phase). *J. Environ. Manag.* 291:112727. doi: 10.1016/j.jenvman.2021.112727

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

Rubio-Portillo, E., Kersting, D. K., Linares, C., Ramos-Esplá, A. A., and Antón, J. (2018). Biogeographic differences in the microbiome and Pathobiome of the coral *Cladocora caespitosa* in the Western Mediterranean Sea. *Front. Microbiol.* 9:22. doi: 10.3389/fmicb.2018.00022

Schmalenberger, A., Schwieger, F., and Tebbe, C. C. (2001). Effect of primers hybridizing to different evolutionarily conserved regions of the small-subunit rRNA gene in PCR-based microbial community analyses and genetic profiling. *Appl. Environ. Microbiol.* 67, 3557–3563. doi: 10.1128/AEM.67.8.3557-3563.2001

Sharp, K. H., Pratte, Z. A., Kerwin, A. H., Rotjan, R. D., and Stewart, F. J. (2017). Season, but not symbiont state, drives microbiome structure in the temperate coral *Astrangia poculata*. *Microbiome.* 5:120. doi: 10.1186/s40168-017-0329-8

Sharp, K. H., and Ritchie, K. B. (2012). Multi-partner interactions in corals in the face of climate change. *Biol. Bull.* 223, 66–77. doi: 10.1086/BBLv223n1p66

Shi, Y., Li, Y., Xiang, X., Sun, R., Yang, T., He, D., et al. (2018). Spatial scale affects the relative role of stochasticity versus determinism in soil bacterial communities in wheat fields across the North China plain. *Microbiome.* 6:27. doi: 10.1186/s40168-018-0409-4

Siboni, N., Ben-Dov, E., Sivan, A., and Kushmaro, A. (2008). Global distribution and diversity of coral-associated Archaea and their possible role in the coral holobiont nitrogen cycle. *Environ. Microbiol.* 10, 2979–2990. doi: 10.1111/j.1462-2920.2008.01718.x

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313. doi: 10.1093/bioinformatics/btu033

van Oppen, M. J. H., Bongaerts, P., Frade, P., Peplow, L. M., Boyd, S. E., Nim, H. T., et al. (2018). Adaptation to reef habitats through selection on the coral animal and its associated microbiome. *Mol. Ecol.* 27, 2956–2971. doi: 10.1111/mec.14763

Vega, T. R., Mydlarz, L. D., Brandt, M., Harvell, D., Weil, E., Raymundo, L., et al. (2020). Deciphering coral disease dynamics: integrating host, microbiome, and the changing environment. *Front. Ecol. Evol.* 8:575927. doi: 10.3389/fevo.2020.575927

Wainwright, B. J., Afiq-Rosli, L., Zahn, G. L., and Huang, D. (2019). Characterisation of coral-associated bacterial communities in an urbanised marine environment shows strong divergence over small geographic scales. *Coral Reefs* 38, 1097–1106. doi: 10.1007/s00338-019-01837-1

Wang, Q., Li, H., and Post, A. F. (2000). Nitrate assimilation genes of the marine diazotrophic, filamentous cyanobacterium *Trichodesmium* sp. strain WH9601. *J. Bacteriol.* 182, 1764–1767. doi: 10.1128/jb.182.6.1764-1767.2000

Wiebe, W. J., Johannes, R. E., and Webb, K. L. (1975). Nitrogen fixation in a coral reef community. *Science* 188, 257–259. doi: 10.1126/science.188.4185.257

Williams, S. D., Klinges, J. G., Zinman, S., Clark, A. S., Bartels, E., Villoch Diaz Maurino, M., et al. (2022). Geographically driven differences in microbiomes of *Acropora cervicornis* originating from different regions of Florida's coral reef. *PeerJ.* 10:e13574. doi: 10.7717/peerj.13574

Ziegler, M., Seneca, F. O., Yum, L. K., Palumbi, S. R., and Voolstra, C. R. (2017). Bacterial community dynamics are linked to patterns of coral heat tolerance. *Nat. Commun.* 8:14213. doi: 10.1038/ncomms14213

frontiers | Frontiers in Marine Science

TYPE Original Research
PUBLISHED 18 July 2023
DOI 10.3389/fmars.2023.1201434

OPEN ACCESS

EDITED BY
Zhiyong Li,
Shanghai Jiao Tong University, China

REVIEWED BY
Wang Minxiao,
Chinese Academy of Sciences (CAS), China
Yongyu Zhang,
Chinese Academy of Sciences (CAS), China

*CORRESPONDENCE
Yantao Liang
liangyantao@ouc.edu.cn
Andrew McMinn
andrew.mcminn@utas.edu.au
Min Wang
mingwang@ouc.edu.cn

†These authors have contributed equally to this work

RECEIVED 06 April 2023
ACCEPTED 03 July 2023
PUBLISHED 18 July 2023

CITATION
Sun J, Zhang X, Liang Y, Zheng K, Kennedy F, Han M, Liu G, Liu Y, Wang Z, Chen X, Sung YY, Mok WJ, Wong LL, McMinn A and Wang M (2023) Abundance and ecological footprint of Pseudoalteromonas phage vB_PhoS_XC in the *Ulva prolifera* green tide. *Front. Mar. Sci.* 10:1201434. doi: 10.3389/fmars.2023.1201434

Abundance and ecological footprint of Pseudoalteromonas phage vB_PhoS_XC in the *Ulva prolifera* green tide

Jianhua Sun[1†], Xinran Zhang[1†], Yantao Liang[1,2*], Kaiyang Zheng[1], Fraser Kennedy[3], Meiaoxue Han[1], Gang Liu[1], Yundan Liu[1], Ziyue Wang[1], Xuechao Chen[1], Yeong Yik Sung[2,4], Wen Jye Mok[2,4], Li Lian Wong[2,4], Andrew McMinn[1,3*] and Min Wang[1,2,5,6*]

[1]College of Marine Life Sciences, Institute of Evolution and Marine Biodiversity, Frontiers Science Center for Deep Ocean Multispheres and Earth System, Key Lab of Polar Oceanography and Global Ocean Change, and Center for Ocean Carbon Neutrality, Ocean University of China, Qingdao, China, [2]Universiti Malaysia Terengganu (UMT) - Ocean University of China (OUC) Joint Centre for Marine Studies, Qingdao, China, [3]Institute for Marine and Antarctic Studies, University of Tasmania, Hobart, TAS, Australia, [4]Institute of Marine Biotechnology, Universiti Malaysia Terengganu (UMT), Kuala Nerus, Malaysia, [5]Haide College, Ocean University of China, Qingdao, China, [6]The Affiliated Hospital of Qingdao University, Qingdao, China

Pseudoalteromonas is a ubiquitous and abundant genus of marine bacteria commonly associated with algae. In this study, a novel siphoviral-morphological bacteriophage, vB_PhoS_XC, was isolated from the coastal seawaters of Qingdao (China) during a bloom of the *Ulva prolifera* (*U. prolifera*) green tide. The morphology of this phage (icosahedron head 51 ± 1 nm in diameter; a tail length of 86 ± 1 nm) was characterized through transmission electron microscope. The biological properties of this virus showed a short latent period (45 minutes), a large burst size (241 virions per cell) and a relatively wide range of temperatures/pH level tolerance (-20°C to 45°C and pH 4 to pH 10, respectively). The vB_PhoS_XC has a 46,490-bp double-stranded DNA genome with a G+C content of 40.0%, and encodes 72 open reading frames (ORFs). Thirty-five of these ORFs were assigned into known functions based on BLAST-based algorithm against NR database of GenBank. In addition, eco-genomic analysis provides the evidence of vB_PhoS_XC accompanied by bloom of *U. prolifera*, and confirmed the high expression of two phosphatase-metabolism-related auxiliary metabolic genes (AMGs). This study provides new insights into the functional and ecological roles of the Pseudoalteromonas phage vB_PhoS_XC, shedding light on the virological study approach combined with traditional isolation and meta-omics data.

KEYWORDS

bacteriophage, Pseudoalteromonas, Ulva prolifera blooms, genomic and meta-omics analysis, ecological footprint

Introduction

In coastal marine ecosystems, algae are a naturally abundant food source for marine organisms and contribute to the production of polysaccharide-rich detritus (Jain et al., 2020; Akter et al., 2022). However, macroalgae blooms create great pressure on these ecosystems, resulting in hypoxia and acidification (Zhang et al., 2019). The largest causal agent of green tides is the macroalgae *Ulva prolifera* (*U. prolifera*), which has been affecting Qingdao since 2007. Following a bloom, huge amounts of dissolved organic matter (DOM) are produced due to the degradation of *U. prolifera* (Zhang and Wang, 2017). This leads to serious negative economic impacts and complex ecological consequences, most notably on the structure of the microbial community (Xiao et al., 2021; Zhang et al., 2021). Nevertheless, the majority of studies on *U. prolifera* control and microenvironment typically concentrated on the relationships between *U. prolifera* and bacteria, for instance, changes in bacterial communities and responses to *U. prolifera* without taking into consideration the potential contribution of viruses (Qu et al., 2020; Liang et al., 2021; Qu et al., 2021; Zhao et al., 2022).

Viruses are the most abundant and diverse biological entities in the ocean (Culley et al., 2006), playing an important role in global biogeochemical cycles (Fuhrman, 1999). They are also the most abundant and genetically diverse "life forms" on Earth (Fuhrman, 1999). Viruses mediate horizontal gene transfer by infecting host cells, and they also play a key role in the formation of bacterial community structure. Moreover, they can reduce the abundance of the major host by infecting and inducing the community structure into homeostasis (Zhang et al., 2007). The largest community of viruses in marine ecosystems are bacteriophages, which are bacterial specific. It is estimated that abundance of bacteriophages is around 10-15 times higher than bacteria (Culley et al., 2006), and the expression of phage-encoded auxiliary metabolic genes (AMGs) can influence bacterial host with reprogramming metabolic progress and improves their adaptability to environmental changes (Gregory et al., 2019; Zimmerman et al., 2020).

Pseudoalteromonas is the most abundant bacteria genus found in coastal seawater and is dominant during *U. prolifera* blooms (Wang et al., 2020; Qu et al., 2021). Metabolically, this species exhibits high productivity and can effectively use a variety of carbon sources, providing a competitive advantage over other microorganisms. Further, it has been explored to secrete a range of extracellular substances, including enzymes, toxins, antibiotics and polysaccharides (Ivanova et al., 2003), and found to make significant contributions to dissolved algae through the action of alginate lyases. This organism's strong capacity to produce extracellular degrading enzymes, as well as a variety of bioactive substances, leads to its important role in marine biogeochemical cycles (Bowman, 2007). Although *Pseudoalteromonas* is a significant alginate-degrading group in the ocean, with its presence being particularly prominent in coastal seawater (Rong et al., 2018; Ren et al., 2022), few Pseudoalteromonas phages have been isolated. Thus, we sought to explore the possible regulatory role of viruses in the *U. prolifera* blooms in the coastal marine ecosystems of Qingdao in this study. To do this, the Pseudoalteromonas phage vB_PhoS_XC was isolated from these waters, its biological characteristics were investigated, and its genome was analyzed. Additionally, the metagenome and metatranscriptome outcomes of the phage-host interaction were used as models to suggest that the virus could influence the dynamic population of *U. prolifera*. This work thus adds to our understanding of *U. prolifera* blooms and provides a framework for further studies on the complex interactions between *U. prolifera*, *Pseudoalteromonas*, and their phages.

Materials and methods

Sampling and virion enrichment

Seven samples were collected from the coastal seawaters of Qingdao between July to December in 2020 year. Eighty-liters of surface seawater were sampled from the coast of Qingdao (36.06° N, 120.31° E), and then pre-filtered on nylon sieves with a mesh of 200 μm and 20 μm to remove large zooplankton. Seawater samples were pre-filtered 3 μm filters and then followed by filtration through 0.22 μm Polycarbonate Membranes filters (Millipore, MA, USA) using a peristaltic pump (Stewart et al., 2012; Pesant et al., 2015). Filters were transferred immediately to DNase/RNase-Free tubes with RNAlater and stored at -80°C for RNA extraction. To further concentrate the filtered seawater containing viruses, tangential flow filtration (laboratory scale, 50 kDa; Millipore) was used to reduce samples to 100 mL for DNA extraction. The samples and the original seawater were stored at 4°C in the dark until experimentation (Wang et al., 2015).

Preparation of bacterial strain

Host bacterial strain was isolated from the original seawater at the coast seawaters of Qingdao during the *U. prolifera* blooms in 2020 year, and then incubated in liquid Zobell medium at 28°C for next experiments (Liu et al., 2018). The 16S rRNA gene of host was amplified by PCR. The phylogenetic tree was calculated on the bias of the 16S rRNA sequences using IQ-TREE with MFP model as the best-fitted substitution model undergoing 1,000 bootstraps (Supplementary Figure S1).

Isolation, purification of bacteriophage

The phage vB_PhoS_XC was prepared using the standard virus enrichment approach and doubled-layer agar methods (Jamalludeen et al., 2007), followed by using the soft-agar overlay method for plaque analysis (Hyman and Abedon, 2010). Briefly, 1 mL host cell culture was incubated with 0.2 mL seawater, and then was filtered through 0.22 μm pore-size membranes (Millipore) for 20min. The mixture was placed into 5 mL of soft and warm agar (0.6%), and poured onto petri dishes to form plaques. Phages were purified by picking a single plaque, suspending it in SM buffer (100 mM NaCl, 81.2 mM $MgSO_4$, 50 mM Tris HCl [pH 7.5], 0.01%

gelatin) for 1 h at 37°C. The suspension was filtered through 0.22 μm pore-size membranes (Millipore), and was used to obtain the plaque through double-layer plate method. The purification step was repeated for three times. Then the purified phages were amplified and stored at 4°C.

Morphological study by transmission electron microscopy

20 μL viral suspension was negatively stained with 2% uranyl acetate, and was examined by energy-filtering transmission electron microscopy (TEM, JEOL Model JEM-1200EX) at 100 kV at NICEM to obtain the viral morphological data (Li Y et al., 2016). The average size of virion was determined by measuring three phage particles within the same field of vision.

One-step growth assay

The one-step growth curve assay was used to identify the burst size of the phage vB_PhoS_XC following the previous method with some modifications (Middelboe et al., 2010; Ma et al., 2021). The burst size was calculated by dividing the final number of virions by the initial number of infected host cells at the beginning of assay (Wang et al., 2017). A bacterial culture in the exponential growth phase (2 x 10^8 CFU/mL) was mixed with 1 mL of the phage suspension (MOI=0.01) at 25°C for 20 min. After centrifugation at 12,800 g for 30 seconds to remove unabsorbed phage, samples were re-suspended to 50 mL liquid Zobell medium. The re-suspended virions were used to infect bacteria with the 5-min/10-min/30-min interval in first/second/third hour, respectively (Gong et al., 2017). This assay was re-performed by three times (Li et al., 2016). The virion/bacteria co-culture was plated on a double-layer plate to incubate overnight. The number of phage plaques were counted to determine the titers of the phage at different time points to assess its growth states.

pH/thermal stability

We tested pH/thermal stability to gain the information on the biological properties of the phage vB_PhoS_XC. Briefly, to test capsid tolerance to pH, 100 μL phage suspension mixed with 900 μL SM buffer with a range of pH from 3 to 12 at 30 °C for 2 h. To test the capsid tolerance to temperature, the phage suspension was incubated at -20°C, 4°C, 25°C, 35°C, 45°C, 55°C, 65°C, and 75°C for 2 h, then incubated with host at 25°C for 20min. The double-layer plate method was used to calculate the loss of virus titer of each sample. Each sample was performed in triplicate.

Genome sequencing and function analysis

Bacteriophage DNA was extracted by Virus DNA Kit (OMEGA) and sequenced using the Illumina HiSeq PE150 paired-end sequence method by Novogene (Tianjin, China) company (Ma et al., 2021). The gaps between the remaining contigs were closed using GapCloser and GapFiller, with purified genomic DNA as the template (Nadalin et al., 2012; Xu et al., 2020). Open reading frames (ORFs) were predicted by GeneMarkS (http://topaz.gatech.edu/GeneMark/) and RAST (http://rast.nmpdr.org/). The protein function and homology of ORFs were predicted through the BLASTP-algorithm and Hidden-Markov-Module (HMM) based algorithm against NR (2023/02) and Pfam (v.35) database with 1e-5 as the E-value threshold (Jacob et al., 2008; Katoh and Standley, 2013; Mistry et al., 2013; Mistry et al., 2021). The genomic organization diagram was visualized by the CLC Main Workbench 20. The tRNAscan-SE (http://lowelab.ucsc.edu/tRNAscan-SE/) was used to detect tRNA sequences is in the viral genome (Chan et al., 2021). The complete viral genome sequence of vB_PhoS_XC is now available to access under the accession number MT002874.2 in GenBank. The auxiliary metabolic genes (AMGs) were blast through the Integrated Microbial Genomes/Virus 4 database with the E-value 1e-5, high search quality genes were chosen for homologous genes, and details were given in Supplementary Table 1.

Metagenome, metatranscriptome sequencing and analysis

Total viral DNA was extracted using QIAamp DNA MiniKit (QIAGEN) according to the manufacturer's instructions. Library construction and high-throughput sequencing were carried out by Novogene (Tianjin, China) using Illumina NovaSeq 6000 (pair-end sequencing, 2 × 150 bp). Raw reads were quality-controlled by removing the adapters using Cutadapt (Kechin et al., 2017). The high-quality paired-end reads were filtered by Perl scripts following these criteria: 1) without N; 2) no more than 20% bases with a quality score less than 20 (Q20); 3) no more than 30% bases with a quality score less than 30 (Q30) (Yang et al., 2019). To calculate the relative abundances of the vB_PhoS_XC in all meta-omics samples, clean reads were mapped to the vB_PhoS_XC contig using the contig section of CoverM v0.6.1 (https://github.com/wwood/CoverM) with default parameter. The relative abundance of the vB_PhoS_XC was quantified as reads per kilobase million mapped reads (RPKM) (Zhang et al., 2023). We also calculated the relative abundance of other representative phages abundant in the ocean including those infecting *Pelagibacter* (Zhao et al., 2013; Chen et al., 2019), *Cyanobacteria* (Labrie et al., 2013), and all *Pseudoalteromonas* which belong to the same subfamily as vB_PhoS_XC (Table 1).

Total RNA extraction, library construction and next-generation sequencing of viral RNA was carried out by Novogene (Tianjin, China) using Illumina NovaSeq 6000 (pair-end sequencing, 2 × 150 bp). Clean reads were picked from the raw reads by Novagene. Clean reads were then quality trimmed using Trimmomatic v0.39 (Bolger et al., 2014), and then were dropped below SLIDINGWINDOW:4:30, MINLEN:70. The potential rRNA-associated reads were removed using the SortMeRNA v4.3.2 against all databases (Kopylova et al., 2012). Then we got the

TABLE 1 The genome information of *Pseudoalteromonas* phages belonging to the same subfamily as *Pseudoalteromonas* phage vB_PhoS_XC.

Name	Accession number	Length (bp)	G+C content (%)	Number of ORF	Number of tRNA	Proposed subfamily/ genus	Viral family
XC	MT002874	46609	40.01	76	1	SG15	Siphoviridae
Station137_DCM_ALL_assembly_NODE_68	Station137_DCM_ALL_assembly_NODE_68	45654	39.53	91	1	SG15	Siphoviridae
DTR_889654	DTR_889654	45599	39.52	91	1	SG15	Siphoviridae
vB_PspS-H40/1	KU747973	45306	40.15	73	0	SG15	Siphoviridae

high-quality reads. The relative abundances of the genes encoded by vB_PhoS_XC were calculated on the basis of all meta-omics samples through CoverM v0.6.1 with default parameter (Salazar et al., 2019). The relative abundances of these genes were then determined as reads per kilobase million mapped reads (RPKM).

Correlation analysis of the phage vB_PhoS_XC with environmental parameters

Temperature, salinity, Dissolved Oxygen (DO) and pH were measured at the sampling site using a YSI proplus multi-parameter water quality analyzer (YSI, Yellow Springs, OH, USA). Seawater filtered through membranes, The concentrations of nitrate (NO_3^-), nitrite (NO_2^-), ammonium (NH_4^+) and phosphate ($PO4^{3-}$) were measured with a QuAAtro nutrient auto analyzer (Seal Analytical Ltd., King's Lynn, UK). Picoplankton (Pico), Synechococcus (Syn), Bacterioplankton (Bac) counting using flow cytometry (Beckman Coulter Inc, USA). All environmental parameters described in Supplementary Table S2. Pearson correlation coefficient (R-value) was used to describe the correlation between vB_PhoS_XC and environmental parameters, including biological parameters (such as Pico, Syn, Bac) and non-biological parameters (such as NH_4: Ammonium (NH_3-N), Nitrite (NO_2-N), Phoshate (PO_4^{3-}), pH value, Nitrate (NO_3-N), Salinity and Dissolved Oxygen (DO)). The analysis was performed by psych and vegan which are both powered by R (Dixon, 2003).

Result and discussion

Morphological and biological properties of phage vB_PhoS_XC

A novel marine phage, vB_PhoS_XC, was isolated from the coastal waters of Qingdao during an *U. prolifera* bloom (Figure 1A). By the 16srRNA analysis, we determined that the host of the phage vB_PhoS_XC is *Pseudoalteromonas hodoensis* strain H7 (Supplementary Figure S1). Characterization of vB_PhoS_XC revealed an icosahedral head (~51 ± 1 nm in diameter) with a long non-contractile tail (~86 ± 1 nm) (Figures 1B, C), combined with previous research, indicating that it could be classified into the Sipovirus subfamily (Zheng et al., 2023).

The phage vB_PhoS_XC demonstrated a stable titer across a range of pH levels, from 4 to 10. While it could withstand environments of slightly lower (pH3) and higher (pH10-12) acidity, damage did occur in both cases (Figure 2A). Moreover, the phage displayed a more resistance to acidic environments over alkaline. In addition, thermal stability was observed from -20°C to 45°C, although a rapid decrease in phage titer occurred at temperatures of 55°C or higher, with a maximum tolerance of 65° C (Figure 2B). Overall, the phage vB_PhoS_XC displayed a wide range of tolerance to pH and temperature.

The one-step growth curve of phage vB_Phos_XC demonstrated a latent period of approximately 45 minutes,

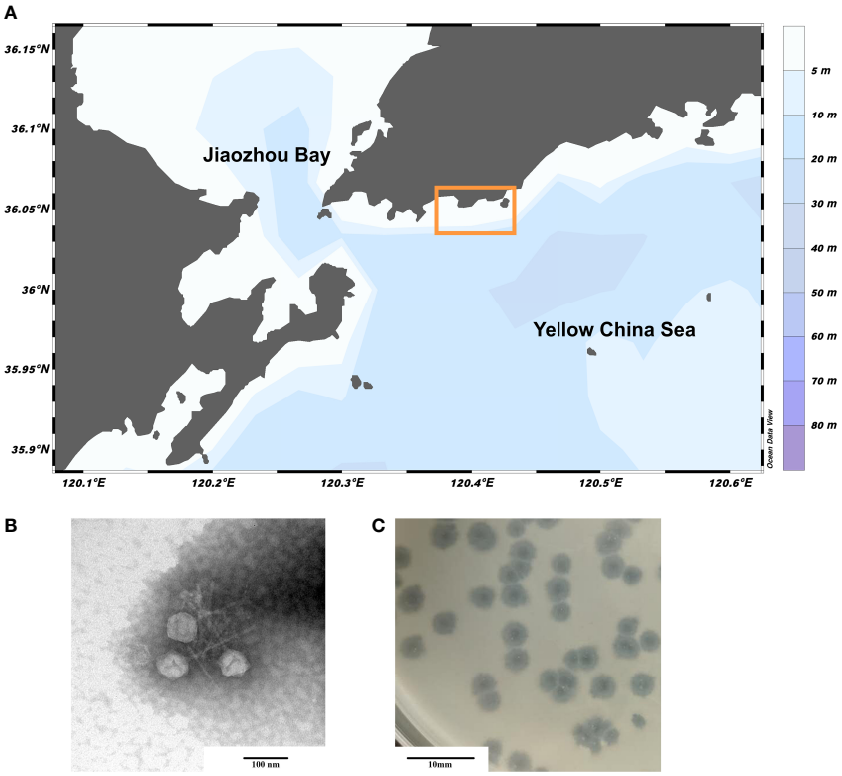


FIGURE 1
Sampling area of Qingdao and Morphology of phage vB_PhoS_XC. **(A)** Sampling region bounded by the yellow box. **(B)** Transmission electron micrograph (TEM) of *Pseudoalteromonas* phage vB_PhoS_XC. The scale bar is 100 nm. **(C)** The phage plaques of *Pseudoalteromonas* phage vB_PhoS_XC. The scale bar is 10 mm.

reaching a plateau stage after 120 minutes (Figure 2C). The final count of released phage particles was 5.0×10^{12} PFU/mL, indicating a burst size of approximately 241 virions per cell. This suggests that VB_PHOS_XC has a lytic activity and a relatively short replication cycle (Yuan et al., 2022).

Overall genome features of the phage vB_PhoS_XC

The genome of the phage vB_PhoS_XC consisted a double-stranded DNA (dsDNA) molecule with a GC content of 40.0% and a total length of 46,490 bp (Figure 3). The genome contained 72 open reading frames (ORFs) and one tRNA. Of the ORFs, 37 had no homologous proteins in current public databases, and 35 had known functions, which were grouped into three functional modules: DNA replication and nucleotide metabolism, phage packing and lysis, and phage structure (Table 2).In these 35 known genes, two auxiliary metabolic genes (AMGs) were also predicted to be associated with its host (ORF 11 and ORF 51).

Phage structure and packing genes of the phage vB_PhoS_XC

Most of the ORFs associated with the phage structure in the VB_PHOS_XC genome. In addition to common structure proteins of Pseudoalteromonas phage, the genome contained two unique ORFs associated with Siphovirus structure proteins, namely Tail Tape Measure Proteins (TMPs; ORF 65 and ORF 66). This family of proteins is characterized by short repeats and its function is related to the assembly of the phage, as it participates in determining tail length (Pedersen et al., 2000). TMPs are also known to assist in phage genome delivery into the host cell cytoplasm by interacting with a cognate receptor on the cells surface (Mahony et al., 2016). Studies on the mechanisms of host infection by the Myoviridae phage T4 and Siphoviridae phage T5 have revealed the role of TMP in forming a transmembrane channel that allows the passage of phage DNA into the cell cytoplasm (Boulanger et al., 2008; Hu et al., 2015). Moreover, it has been found that the TMP is a multifunction protein responsible for both tail length determination and viral genome delivery. Consequently, the bacteriophage vB_PhoS_XC containing a TMPs may be used as a model to further investigate the mechanism of *Pseudoalternomonas* infection. Genes related to the packing of the genome in the vB_PhoS_XC were identified. The terminase, composed of two unique subunits (TerS and TerL), was found to be responsible for this process (Esterman et al., 2021). TerS (ORF 35) was found to be responsible for recognizing and binding to the packaging initiation site, as well as regulating the ATPase activity of the TerL (Dixit et al., 2019; Lokareddy et al., 2022b). Additionally, TerL (ORF 36) acted as an ATP-driven molecular motor that transferred viral DNA into the capsid regions, and as an endonuclease that cut the viral genome from the concatemer

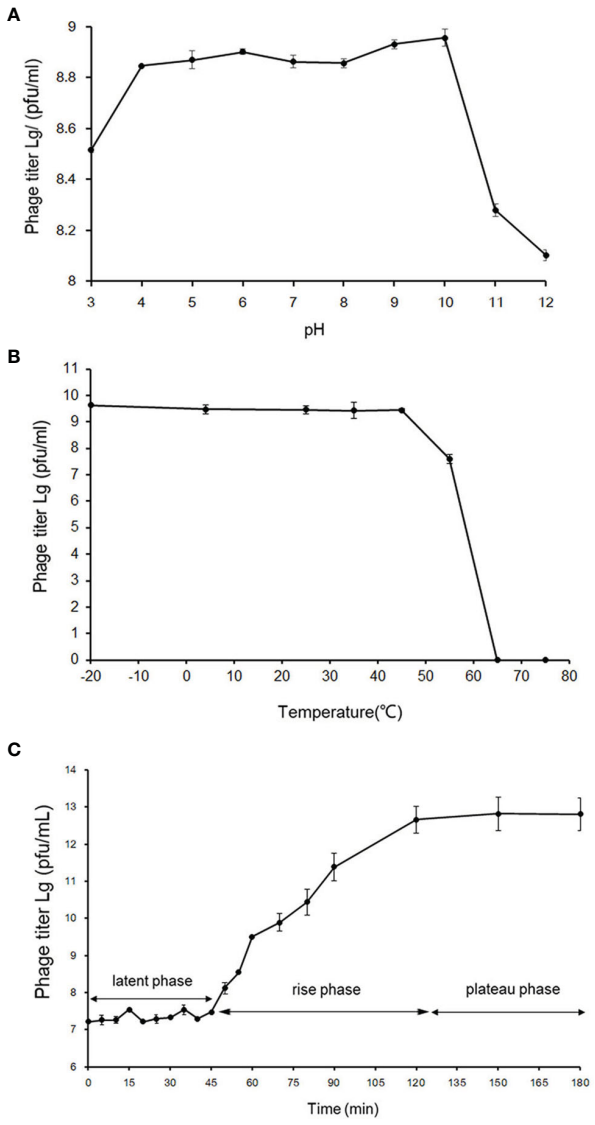


FIGURE 2
Biological properties of phage vB_PhoS_XC. **(A)** pH stability of *Pseudoalteromonas* phage vB_PhoS_XC. Y-axis shows the log of plaque-forming units per milliliter (PFU/mL). **(B)** Temperature stability of *Pseudoalteromonas* phage vB_PhoS_XC. Y-axis shows the log of plaque-forming units per milliliter (PFU/mL). **(C)** One-step growth curve of *Pseudoalteromonas* phage vB_PhoS_XC. Y-axis shows the log of plaque-forming units per milliliter (PFU/mL).

(Esterman et al., 2021; Lokareddy et al., 2022a). Finally, these two genes were docked on the portal protein (ORF 8) to facilitate the generation of circular phage heads (Al-Zahrani et al., 2009).

Phage metabolic genes of the phage vB_PhoS_XC

Analysis of the vB_PhoS_XC genome revealed the presence of genes related to some important metabolic processes. ORF 2 was predicted to encode a member of the M15 family metallopeptidase, which is involved in bacterial cell wall biosynthesis and metabolism (Rawlings and Barrett, 2004). Additionally, ORF 15 was observed to encode a phosphoadenosine phosphosulfate (PAPS) reductase enzyme, a member of sulfonucleotide reductases (SRs) that catalyzes the reduction of adenylated sulfate to sulfite and an essential part of the cysteine biosynthesis pathway (Chartron et al., 2007). PAPS reductase was an important enzyme found in many organisms and played a crucial role in bacteria as it is involved in the metabolism of sulfur-containing compounds such as cysteine and methionine (Kredich and Tomkins, 1966, Chartron et al., 2007). PAPS reductase regulates these compounds by reducing PAPS to 3'-phosphoadenosine-5'-phosphate (PAP). This process involves NADPH (nicotinamide adenine dinucleotide phosphate) as a cofactor and helps maintain proper levels of sulfur-containing compounds within bacterial cells (Murray-Rust et al., 2001).

ORF 17 encodes a DNA-binding transcriptional regulator, which regulates gene transcription by recognizing promoters and enhancers near the transcription initiation site (Wang et al., 2013). ORF 18 is a protein with the catalytic structural domain of GIY-YIG, used to catalyze phosphodiester bond breaking, DNA repair, recombination and genetic element transfer (Truglio et al., 2005; Dunin-Horkawicz et al., 2006). ORF 19 is a DEAD/DEAH box helicase belonging to the family of SvB_ValS_NF2, containing sequence motifs similar to those of DNA helicase proteins; involved in various aspects of transcriptional regulation, recombination and DNA repair (Xu et al., 2021). Additionally, Exodeoxyribonuclease VIII (ORF 25) acts to break dsDNA on both 5' ends, allowing kinked and abnormal portions of the genome to be straightened and repaired through homologous recombination. Interestingly, Exodeoxyribonuclease VIII facilitates genomic repair even in low energy environments and remains stable (Sharma et al., 2019; Kumar et al., 2021).

ORF 28 encodes a bifunctional DNA primase-polymerase, which is widespread in diverse bacteria, archaea, and viruses (Lipps et al., 2003; Geibel et al., 2009; Halgasova et al., 2012). This enzyme has the capacity to act as a primase and a polymerase, and can synthesize long DNA from dNTPs without a preexisting primer bound to the DNA template (Zhu et al., 2017). ORF 57 is predicted to encode a DNA methylase that modifies the DNA of the phage, ultimately avoiding degradation of enzymes from the host cell, as well as protecting phage marked with methylation from its own nucleases which degrade host DNA during the early period of infection (Aravind et al., 2013). Therefore, it may play an important role in the lytic life cycle of the phage (Aravind et al., 2013). The phage vB_PhoS_XC has a relatively well-developed functional composition of the genome. In addition to structural, packing and metabolic genes, some other genes also contribute to its independent survival strategy, such as DNA polymerase (ORF 21), Nucleic acid-binding (ORF22), transcriptional regulator (ORF 27) and ATPase (ORF 40) and most nucleotide metabolism genes.

Two phosphatase genome-biosynthetic-related AMGs

Two host-derived auxiliary metabolic genes (AMGs) were identified in the genome of phage vB_PhoS_XC: a nucleotide

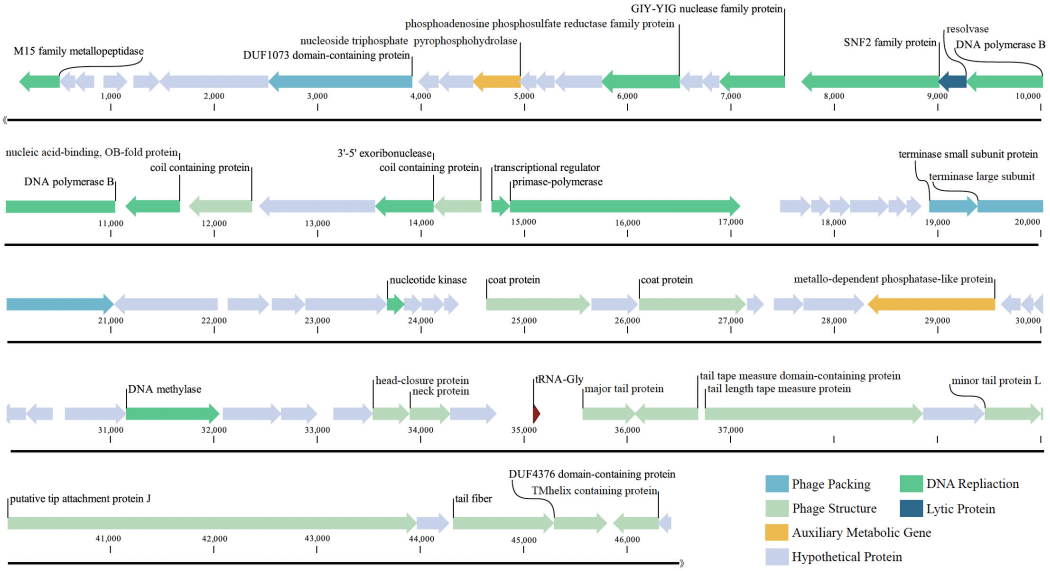


FIGURE 3
Genome map of *Pseudoalteromonas* phage vB_PhoS_XC. Putative functional categories were defined according to annotation and are represented by different colors. The length of each arrow represents the length of each gene.

TABLE 2 The conserved domains information of *Pseudoalteromonas* phage vB_PhoS_XC genome.

ORF	Start	Stop	Strand	Function	Match phage	Accession	E value
1	35084	35154	+				
2	504	112	–	M15 family metallopeptidase	Pseudoalteromonas sp.	MCP4060181.1	2e-37
7	2541	1468	–	putative minor head protein	Pseudoalteromonas phage vB_PspS-H40/1	YP_010655240.1	0
8	3916	2522	–	portal protein	Pseudoalteromonas phage vB_PspS-H40/1	YP_010655241.1	0
11	4966	4499	–	nucleotide pyrophosphohydrolase	Pseudoalteromonas phage TW1	YP_010655192.1	1e-36
15	6502	5750	–	phosphoadenosine phosphosulfate reductase family protein	Vibrio phage 1.047.O._10N.286.55.F2	AUR84066.1	2e-138
17	6886	6716	–	DNA-binding transcriptional regulator	Oceanospirillum phage vB_OliS_GJ44	QSM00634.1	6.00e-10
18	7523	6888	–	GIY-YIG nuclease family protein	Pseudomonas paracarnis	WP_220381685.1	1e-32
19	9018	7681	–	DEAD/DEAH box helicase	Pseudoalteromonas phage AL	QIG62466.1	0
20	9280	9005	–	resolvase	Pseudoalteromonas phage PHS21	AQN32316.2	1e-57
21	11048	9273	–	putative DNA polymerase	Pseudoalteromonas phage AL	QIG62464.1	0
22	11667	11140	–	nucleic acid-binding, OB-fold protein	Vibrio phage 1.017.O._10N.286.55.C11	AUR82036.1	2e-68
23	12365	11754	–	coil containing protein	Vibrio phage 1.133.O._10N.222.51.E4	AUR89805.1	3e-31
25	14123	13557	–	exodeoxyribonuclease VIII	Pseudoalteromonas phage PHS21	AQN32313.1	3e-106
26	14584	14126	–	coil containing protein	Vibrio phage LP.1	AZU97906.1	2e-07
27	14683	14862	+	transcriptional regulator	Pseudoalteromonas phage AL	QIG62460.1	9e-22
28	14862	17090	+	bifunctional DNA primase/polymerase	Vibrio phage LP.1	AZU97904.1	0
35	18918	19385	+	terminase small subunit protein	Cupriavidus lacunae	WP_208647650.1	3e-39

(Continued)

TABLE 2 Continued

ORF	Start	Stop	Strand	Function	Match phage	Accession	E value
36	19385	21028	+	phage terminase large subunit	Candidatus Enterovibrio escacola	WP_150137941.1	0
40	22880	23671	+	ATPase	Pseudoalteromonas phage vB_PspS-H40/1	YP_010655271.1	7e-35
45	24631	25638	+	coat protein	Pseudoalteromonas phage H103	YP_009200228.1	1e-55
46	25650	26099	+	head decoration	Pseudoalteromonas phage H103	YP_009200227.1	4.00e-85
47	26110	27144	+	virion structural protein	Pseudoalteromonas phage H103	YP_009200226.1	0
51	29555	28323	–	metallo-dependent phosphatase-like protein	Vibrio phage 1.106.O._10N.286.51.F7	AUR87948.1	4e-141
57	31145	32050	+	DNA methylase	Vibrio phage 1.160.O._10N.261.48.B11	AUR91427.1	6e-173
58	32083	32649	+	minor head protein	Pseudoalteromonas phage vB_PspS-H40/1	YP_010655287.1	1.00e-68
61	33532	33888	+	head-closure protein	Vibrio phage 1.122.A._10N.286.46.F8	AUR89308.1	4e-05
62	33890	34285	+	neck protein	Vibrio phage 1.076.O._10N.286.51.B7	AUR85547.1	3e-16
64	35564	36079	+	major tail protein	Pseudoalteromonas phage vB_PspS-H40/1	YP_010655294.1	9e-110
65	36683	36066	–	tail tape measure domain-containing protein	Pseudoalteromonas phage XCL1123	QGJ84348.1	3e-32
66	36747	38858	+	tail length tape measure protein	Pseudoalteromonas phage vB_PspS-H40/1	YP_010655297.1	0
68	39455	40000	+	minor tail protein L	Pelagibacter phage HTVC112P	QGZ18222.1	3e-21
69	40000	43962	+	putative tip attachment protein J	Prokaryotic dsDNA virus sp.	QDP59295.1	0
71	44309	45286	+	tail fiber protein	Alteromonas phage vB_AmeM_PT11-V22	YP_009855710.1	3e-08
72	45286	45801	+	tail fiber protein	Pseudoalteromonas phage vB_PspS-H40/1	YP_010655232.1	2e-15
73	46303	45860	–	TMhelix containing protein	Vibrio phage LP.2	AZU97851.1	1e-26

pyrophosphohydrolase gene (ORF 11) and a metallo-dependent phosphatase-like protein gene (ORF 51). These genes were found to contribute to enhanced viral replication through regulation of the host metabolism (Hurwitz and U'Ren, 2016).

ORF 11 encodes nucleotide pyrophosphohydrolase (NTPPHase), a protein found in many marine bacteriophages, suggesting it has an important role in phage proliferation (Bryan et al., 2008; Duhaime et al., 2011). Studies on bacteria have shown NTPPHase is involved in stress responses and removal of noncanonical nucleotides (Lu et al., 2010; Gonçalves et al., 2011), potentially improving the survival of its host in nutrient-depleted environments (Duhaime et al., 2011; Kang et al., 2013; Wang et al., 2021). In addition, ORF 51 encodes a metallo-dependent phosphatase-like protein (PPM). This protein was highly conserved in its sequence and structure with homologous genes identified from prokaryotes, animals, and plants (Kamada et al., 2020). It also regulates the activity of other proteins involved in glycolysis and amino acid biosynthesis (Kamada et al., 2020). Evidence also indicates that PPM has an important role in signaling processes, such as proliferation, apoptosis and metabolism (Tamura et al., 2006; Lammers and Lavi, 2007, Lu and Wang, 2008). As the viral AMG, it is particularly crucial for bacterial survival and growth under certain conditions (Aravind and Koonin, 1998). Furthermore, its homologs were also found to facilitate viral replication by manipulating host cell metabolism (Wang et al., 2021).

Abundance and ecological function of phage vB_PhoS_XC

The temporal abundance of the phage vB_PhoS_XC was characterized in several viral metagenome data from the coastal seawaters of Qingdao in 2020 year. After analysis of viral community structure in each viral metagenome and together with other common ocean bacteriophages. The results revealed that the relative abundance of Pelagibacter phages was consistently high because it was the most abundance phages in the ocean (Zhao et al., 2019), while the relative abundance of Pseudoalternomonas vB_PhoS_XC phage increased in July (Figure 4). This finding

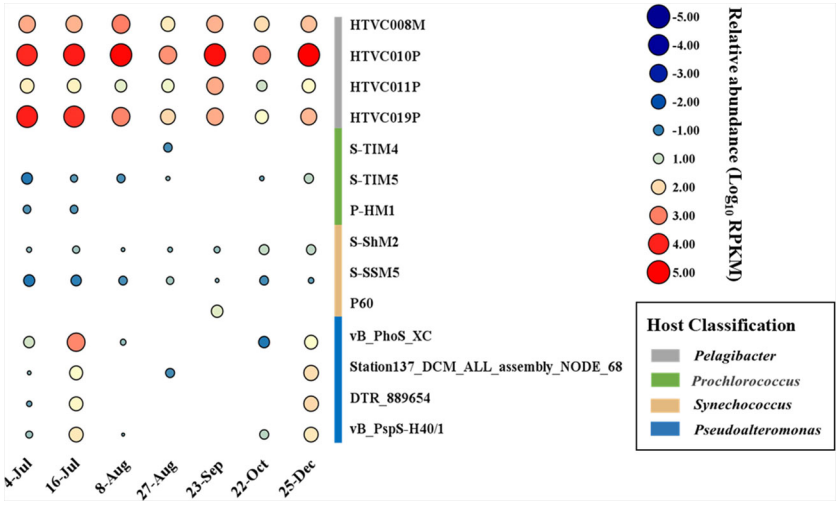


FIGURE 4
Temporal changes of *Pseudoalteromonas* phage vB_PhoS_XC. Bubble size shows relative abundance on *Pseudoalteromonas* phage vB_PhoS_XC of different time series, expressed by RPKM (reads per kilobase per million mapped reads) values.

correlated with an observed outbreak and extinction process of *U. prolifera* green tide in the seawater region of Qingdao in the same time period (Figure 5A). This suggests that the increase in availability of *Pseudoalternomonas* species likely resulted in the corresponding increase in abundance of its specific phage vB_PhoS_XC. These results demonstrate the dynamic relationship between prokaryotic and eukaryotic components of the coastal ecological system. The abundance of this phage significantly increased might be with the accumulation of high amounts of algal-derived DOM after the bloom arrival (Han et al., 2023). By the end of the bloom termination stage, most of the labile DOM had been utilized by bacteria, it might result the decrease of viral abundance like phage vB_PhoS_XC. Although there was no obvious correlation between the phage vB_PhoS_XC and environment factors (Supplementary Figure S2), further implied that there might be a complex interaction between the phage vB_PhoS_XC and its host.

Thus, we also analyzed the expression of AMGs and other functional genes in the corresponding metatranscriptome data to explore the correlation between the phage vB_PhoS_XC and its host. Figure 5B demonstrates that ORF 11 and ORF 51 and ORF 57 (AMGs or accessory metabolic genes) had a high relative abundance in the phage on July 4th and December 25th (Figure 5B). ORF 11 is largely involved in proliferation, stress response, and increasing the survival rate of the host. This may assist the host cells in preventing programmed cell death resulting from rapid changes in abiotic factors caused by *U. prolifera* blooms. ORF 51 regulates the activity of proteins associated with a variety of metabolic pathways. Every year from June to July, a large-scale *U. prolifera* bloom typically occurs in Qingdao. This bloom rapidly causes a decline in inorganic carbon, increasing the pH of the seawater (Han et al., 2023). When the bloom terminates, pH, DOM, DOC, and nutrients are significantly altered to an extent that disrupts the dynamic balance of bacterial community structure and function (Li H et al., 2016; Zhang and Wang, 2017; Chen et al., 2020; Liang et al., 2021). Environmental changes result in the increased expression of these AMGs, which helps the host to survive and maximizes their own replication.

Interestingly, the relative abundance of some genes had a higher expression in December (ORF 22, ORF 36, ORF 40, ORF57 and ORF 8) (Figure 5B). At this time, concentration of nutrients is poor, the water temperature is low and the abundance of bacterioplankton are significantly reduced (Uyà et al., 2017; Guo et al., 2020; Zhao et al., 2022). We speculated that during this period the phage vB_PhoS_XC is manipulating the host through AMGs to improve metabolic efficiency and survival. It may also improve its replication efficiency by repacking genes and increasing nucleotide metabolism, helping create a "virus factory" in the host cells to maximize the output when favorable conditions return.

Conclusion

Isolating and culturing viruses that infect bacteria provides a new and valuable insight into novel viral sequences from metagenomic and metatranscriptomic data. Few Pseudoalteromonas phages have been isolated from the bacterial hosts and research on them is limited. This study aimed to isolate the phage vB_PhoS_XC from coastal seawaters of Qingdao during a U. prolifera bloom. Results indicate that the phage has a short latent period and a large burst size, as well as being active in a wide range of environmental conditions. Additionally, abundance of the phage was observed to be higher

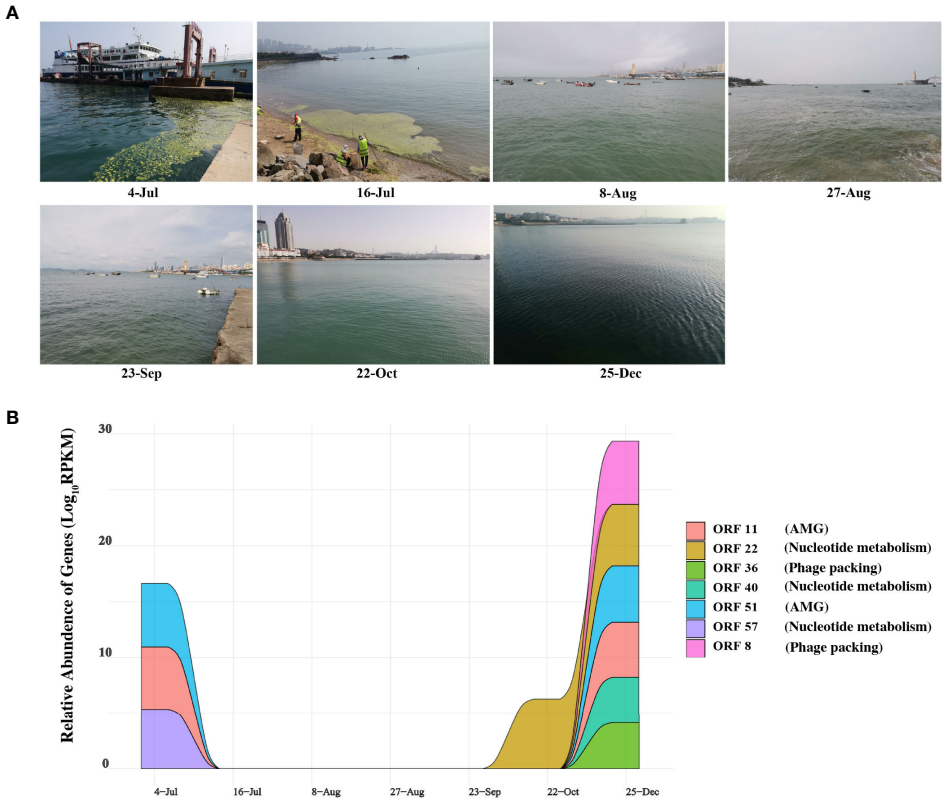


FIGURE 5
Expression of genes and sampling scene pictures. **(A)** The real scene pictures of different sample taken time. **(B)** Auxiliary metabolic genes (AMGs) and nucleotide metabolic Genes with Temporal changes. Colored area shows the relative abundance on the ORFs encoding AMGs and nucleotide metabolic of different time series, expressed by RPKM (reads per kilobase per million mapped reads) values.

during *U. prolifera* blooms and winter, two AMGs and nucleotide metabolism system were also expressed a higher abundance in these periods. This suggests that the phage may be able to improve the host's survival rate and enhance its metabolic efficiency to improve its own replication. Consequently, further research will enable a better understanding of the role of phages in marine ecosystems, as well as their interactions with their hosts.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.cncb.ac.cn, GSA: CRA010505.

Author contributions

YTL, AM and MW: conceptualization, revision, project administration, supervision and funding acquisition. JS, XZ, KZ, MH, GL, YDL and FK: methodology, formal analysis, writing, and original draft preparation. JS, XZ and GL: sample collection and expedition organization. YS, WM, LW and FK: review and editing. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the Laoshan Laboratory (No. LSKJ202203201), National Key Research and Development Program of China (2022YFC2807500), Natural Science Foundation of China (No. 42120104006, 41976117 and 42176111), and the Fundamental Research Funds for the Central Universities (202172002, 201812002, 202072001 and AM).

Acknowledgments

We sincerely appreciate the Center for High Performance Computing and System Simulation, Pilot National Laboratory for Marine Science and Technology (Qingdao). We appreciate the computing resources provided by IEMB-1, a high-performance computing cluster operated by the Institute of Evolution and Marine Biodiversity, and Marine Big Data Center of Institute for the Advanced Ocean Study of the Ocean University of China.

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.2023.1201434/full#supplementary-material

SUPPLEMENTARY FIGURE 1
Phylogenetic tree of the *Pseudoalteromonas* genus.

SUPPLEMENTARY FIGURE 2
Pearson correlation coefficients among the phage vB_PhoS_XC, Picoplankton (Pico), Synechococcus (Syn), Bacterioplankton (Bac), NH4: Ammonium (NH_3-N), Nitrite (NO_2-N), Phoshate (PO_4^{3-}), pH value, Nitrate (NO_3-N), Salinity and Dissolved Oxygen (DO). *Correlation is significant at the 0.05 level, ** Correlation is significant at the 0.01 level, *** Correlation is significant at the 0.001 level.

SUPPLEMENTARY TABLE 1
Homologous genes of auxiliary metabolic genes (AMGs).

SUPPLEMENTARY TABLE 2
The sampling environmental parameters.

FILTER_NO_SINGLE.PL.ZIP
The perl script can be used for removing N from the data.

FILTER_NO_SINGLE2030.PL.ZIP
The perl script can be used for removing less than 20% and 30% bases with a quality score.

References

Akter, L., Ullah, M. A., Hossain, M. B., Karmaker, A. R., Hossain, M. S., Albeshr, M. F., et al. (2022). Diversity and assemblage of harmful algae in homestead fish ponds in a tropical coastal area. *Biol. (Basel)* 11 (9), 1335. doi: 10.3390/biology11091335

Al-Zahrani, A. S., Kondabagil, K., Gao, S., Kelly, N., Ghosh-Kumar, M., and Rao, V. B. (2009). The small terminase, gp16, of bacteriophage T4 is a regulator of the DNA packaging motor. *J. Biol. Chem.* 284 (36), 24490–24500. doi: 10.1074/jbc.M109.025007

Aravind, L., Anand, S., and Iyer, L. M. (2013). Novel autoproteolytic and DNA-damage sensing components in the bacterial SOS response and oxidized methylcytosine-induced eukaryotic DNA demethylation systems. *Biol. Direct* 8, 20. doi: 10.1186/1745-6150-8-20

Aravind, L., and Koonin, E. V. (1998). The HD domain defines a new superfamily of metal-dependent phosphohydrolases. *Trends Biochem. Sci.* 23 (12), 469–472. doi: 10.1016/S0968-0004(98)01293-6

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for illumina sequence data. *Bioinformatics* 30 (15), 2114–2120. doi: 10.1093/bioinformatics/btu170

Boulanger, P., Jacquot, P., Plançon, L., Chami, M., Engel, A., Parquet, C., et al. (2008). Phage T5 straight tail fiber is a multifunctional protein acting as a tape measure and carrying fusogenic and muralytic activities. *J. Biol. Chem.* 283 (20), 13556–13564. doi: 10.1074/jbc.M800052200

Bowman, J. P. (2007). Bioactive compound synthetic capacity and ecological significance of marine bacterial genus pseudoalteromonas. *Mar. Drugs* 5 (4), 220–241. doi: 10.3390/md504220

Bryan, M. J., Burroughs, N. J., Spence, E. M., Clokie, M. R., Mann, N. H., and Bryan, S. J. (2008). Evidence for the intense exchange of MazG in marine cyanophages by horizontal gene transfer. *PloS One* 3 (4), e2048. doi: 10.1371/journal.pone.0002048

Chan, P. P., Lin, B. Y., Mak, A. J., and Lowe, T. M. (2021). tRNAscan-SE 2.0: improved detection and functional classification of transfer RNA genes. *Nucleic Acids Res.* 49 (16), 9077–9096.

Chartron, J., Shiau, C., Stout, C. D., and Carroll, K. S. (2007). 3'-Phosphoadenosine-5'-phosphosulfate reductase in complex with thioredoxin: a structural snapshot in the catalytic cycle. *Biochemistry* 46 (13), 3942–3951. doi: 10.1021/bi700130e

Chen, J., Li, H., Zhang, Z., He, C., Shi, Q., Jiao, N., et al. (2020). DOC dynamics and bacterial community succession during long-term degradation of ulva prolifera and their implications for the legacy effect of green tides on refractory DOC pool in seawater. *Water Res.* 185, 116268. doi: 10.1016/j.watres.2020.116268

Chen, L. X., Zhao, Y., McMahon, K. D., Mori, J. F., Jessen, G. L., Nelson, T. C., et al. (2019). Wide distribution of phage that infect freshwater SAR11 bacteria. *mSystems* 4 (5), e00410-19. doi: 10.1128/mSystems.00410-19

Culley, A. I., Lang, A. S., and Suttle, C. A. (2006). Metagenomic analysis of coastal RNA virus communities. *Science* 312 (5781), 1795–1798.

Dixit, A. B., Ray, K., and Black, L. W. (2019). A viral small terminase subunit (TerS) twin ring pac synapsis DNA packaging model is supported by fluorescent fusion proteins. *Virology* 536, 39–48. doi: 10.1016/j.virol.2019.07.021

Dixon, P. (2003). VEGAN, a package of r functions for community ecology. *J. Vegetation Sci.* 14 (6), 927–930. doi: 10.1111/j.1654-1103.2003.tb02228.x

Duhaime, M. B., Wichels, A., Waldmann, J., Teeling, H., and Glöckner, F. O. (2011). Ecogenomics and genome landscapes of marine pseudoalteromonas phage H105/1. *Isme J.* 5 (1), 107–121. doi: 10.1038/ismej.2010.94

Dunin-Horkawicz, S., Feder, M., and Bujnicki, J. M. (2006). Phylogenomic analysis of the GIY-YIG nuclease superfamily. *BMC Genomics* 7, 98. doi: 10.1186/1471-2164-7-98

Esterman, E. S., Wolf, Y. I., Kogay, R., Koonin, E. V., and Zhaxybayeva, O. (2021). Evolution of DNA packaging in gene transfer agents. *Virus Evol.* 7 (1), veab015. doi: 10.1093/ve/veab015

Fuhrman, J. A. (1999). Marine viruses and their biogeochemical and ecological effects. *Nature* 399 (6736), 541–548.

Geibel, S., Banchenko, S., Engel, M., Lanka, E., and Saenger, W. (2009). Structure and function of primase RepB′ encoded by broad-host-range plasmid RSF1010 that replicates exclusively in leading-strand mode. *Proc. Natl. Acad. Sci.* 106 (19), 7810–7815. doi: 10.1073/pnas.0902910106

Gonçalves, A. M. D., de Sanctis, D., and McSweeney, S. M. (2011). Structural and functional insights into DR2231 protein, the MazG-like nucleoside triphosphate pyrophosphohydrolase from deinococcus radiodurans. *J. Biol. Chem.* 286 (35), 30691–30705. doi: 10.1074/jbc.M111.247999

Gong, Z., Wang, M., Yang, Q., Li, Z., Xia, J., Gao, Y., et al. (2017). Isolation and complete genome sequence of a novel pseudoalteromonas phage PH357 from the Yangtze river estuary. *Curr. Microbiol.* 74 (7), 832–839. doi: 10.1007/s00284-017-1244-8

Gregory, A. C., Zayed, A. A., Conceição-Neto, N., Temperton, B., Bolduc, B., Alberti, A., et al. (2019). Marine DNA viral macro- and microdiversity from pole to pole. *Cell* 177 (5), 1109–1123.e1114. doi: 10.1016/j.cell.2019.03.040

Guo, J., Yuan, H., Song, J., Li, X., and Duan, L. (2020). Hypoxia, acidification and nutrient accumulation in the yellow Sea cold water of the south yellow Sea. *Sci. Total Environ.* 745, 141050. doi: 10.1016/j.scitotenv.2020.141050

Halgasova, N., Mesarosova, I., and Bukovska, G. (2012). Identification of a bifunctional primase-polymerase domain of corynephage BFK20 replication protein gp43. *Virus Res.* 163 (2), 454–460. doi: 10.1016/j.virusres.2011.11.005

Han, M., Sun, J., Yang, Q., Liang, Y., Jiang, Y., Gao, C., et al. (2023). Spatiotemporal dynamics of coastal viral community structure and potential biogeochemical roles affected by an ulva prolifera green tide. *mSystems* 8 (2), e0121122. doi: 10.1128/msystems.01211-22

Hu, B., Margolin, W., Molineux, I. J., and Liu, J. (2015). Structural remodeling of bacteriophage T4 and host membranes during infection initiation. *Proc. Natl. Acad. Sci. U.S.A.* 112 (35), E4919–E4928. doi: 10.1073/pnas.1501064112

Hurwitz, B. L., and U'Ren, J. M. (2016). Viral metabolic reprogramming in marine ecosystems. *Curr. Opin. Microbiol.* 31, 161–168. doi: 10.1016/j.mib.2016.04.002

Hyman, P., and Abedon, S. T. (2010). Bacteriophage host range and bacterial resistance. *Adv. Appl. Microbiol.* 70, 217–248. doi: 10.1016/S0065-2164(10)70007-1

Ivanova, E. P., Bakunina, I. Y., Nedashkovskaya, O. I., Gorshkova, N. M., Alexeeva, Y. V., Zelepuga, E. A., et al. (2003). Ecophysiological variabilities in ectohydrolytic enzyme activities of some pseudoalteromonas species, p. citrea, p. issachenkonii and p. nigrifaciens. *Curr. Microbiol.* 46 (1), 6–10.

Jacob, A., Lancaster, J., Buhler, J., Harris, B., and Chamberlain, R. D. (2008). Mercury BLASTP: accelerating protein sequence alignment. *ACM Trans. Reconfigurable Technol. Syst.* 1 (2), 9. doi: 10.1145/1371579.1371581

Jain, A., Krishnan, K. P., Begum, N., Singh, A., Thomas, F. A., and Gopinath, A. (2020). Response of bacterial communities from kongsfjorden (Svalbard, Arctic ocean) to macroalgal polysaccharide amendments. *Mar. Environ. Res.* 155, 104874. doi: 10.1016/j.marenvres.2020.104874

Jamalludeen, N., Johnson, R. P., Friendship, R., Kropinski, A. M., Lingohr, E. J., and Gyles, C. L. (2007). Isolation and characterization of nine bacteriophages that lyse O149 enterotoxigenic escherichia coli. *Vet. Microbiol.* 124 (1-2), 47–57. doi: 10.1016/j.vetmic.2007.03.028

Kamada, R., Kudoh, F., Ito, S., Tani, I., Janairo, J. I. B., Omichinski, J. G., et al. (2020). Metal-dependent Ser/Thr protein phosphatase PPM family: evolution, structures, diseases and inhibitors. *Pharmacol. Ther.* 215, 107622. doi: 10.1016/j.pharmthera.2020.107622

Kang, I., Oh, H. M., Kang, D., and Cho, J. C. (2013). Genome of a SAR116 bacteriophage shows the prevalence of this phage type in the oceans. *Proc. Natl. Acad. Sci. U.S.A.* 110 (30), 12343–12348. doi: 10.1073/pnas.1219930110

Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30 (4), 772–780. doi: 10.1093/molbev/mst010

Kechin, A., Boyarskikh, U., Kel, A., and Filipenko, M. (2017). cutPrimers: a new tool for accurate cutting of primers from reads of targeted next generation sequencing. *J. Comput. Biol.* 24 (11), 1138–1143. doi: 10.1089/cmb.2017.0096

Kopylova, E., Noé, L., and Touzet, H. (2012). SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics* 28 (24), 3211–3217. doi: 10.1093/bioinformatics/bts611

Kredich, N. M., and Tomkins, G. M. (1966). The enzymic synthesis of l-cysteine in escherichia coli and salmonella typhimurium. *J. Biol. Chem.* 241 (21), 4955–4965. doi: 10.1016/S0021-9258(18)99657-2

Kumar, P., Meghvansi, M. K., and Kamboj, D. V. (2021). "Phenotypic characterization and whole-genome analysis of a novel bacteriophage HCF1 infecting citrobacter amalonaticus and c. *freundii.*" *Front. Microbiol.* 12, 644013. doi: 10.3389/fmicb.2021.644013

Labrie, S. J., Frois-Moniz, K., Osburne, M. S., Kelly, L., Roggensack, S. E., Sullivan, M. B., et al. (2013). Genomes of marine cyanopodoviruses reveal multiple origins of diversity. *Environ. Microbiol.* 15 (5), 1356–1376. doi: 10.1111/1462-2920.12053

Lammers, T., and Lavi, S. (2007). Role of type 2C protein phosphatases in growth regulation and in cellular stress signaling. *Crit. Rev. Biochem. Mol. Biol.* 42 (6), 437–461. doi: 10.1080/10409230701693342

Li, Y., Wang, M., Liu, Q., Song, X., Wang, D., Ma, Y., et al. (2016). Complete genomic sequence of bacteriophage H188: a novel vibrio kanaloae phage isolated from yellow Sea. *Curr. Microbiol.* 72 (5), 628–633. doi: 10.1007/s00284-015-0984-6

Li, H., Zhang, Y., Han, X., Shi, X., Rivkin, R. B., and Legendre, L. (2016). Growth responses of ulva prolifera to inorganic and organic nutrients: implications for macroalgal blooms in the southern yellow Sea, China. *Sci. Rep.* 6, 26498. doi: 10.1038/srep26498

Liang, J., Liu, J., Zhan, Y., Zhou, S., Xue, C. X., Sun, C., et al. (2021). Succession of marine bacteria in response to ulva prolifera-derived dissolved organic matter. *Environ. Int.* 155, 106687. doi: 10.1016/j.envint.2021.106687

Lipps, G., Röther, S., Hart, C., and Krauss, G. (2003). A novel type of replicative enzyme harbouring ATPase, primase and DNA polymerase activity. *EMBO J.* 22 (10), 2516–2525. doi: 10.1093/emboj/cdg246

Liu, Z., Li, H., Wang, M., Jiang, Y., Yang, Q., Zhou, X., et al. (2018). Isolation, characterization and genome sequencing of the novel phage SL25 from the yellow Sea, China. *Mar. Genomics* 37, 31–34. doi: 10.1016/j.margen.2017.09.008

Lokareddy, R. K., Hou, C. D., Doll, S. G., Li, F., Gillilan, R. E., Forti, F., et al. (2022a). Terminase subunits from the pseudomonas-phage E217. *J. Mol. Biol.* 434 (20), 167799. doi: 10.1016/j.jmb.2022.167799

Lokareddy, R. K., Hou, C. D., Li, F., Yang, R., and Cingolani, G. (2022b). Viral small terminase: a divergent structural framework for a conserved biological function. *Viruses* 14 (10), 2215. doi: 10.3390/v14102215

Lu, L. D., Sun, Q., Fan, X. Y., Zhong, Y., Yao, Y. F., and Zhao, G. P. (2010). Mycobacterial MazG is a novel NTP pyrophosphohydrolase involved in oxidative stress response. *J. Biol. Chem.* 285 (36), 28076–28085. doi: 10.1074/jbc.M109.088872

Lu, G., and Wang, Y. (2008). Functional diversity of mammalian type 2C protein phosphatase isoforms: new tales from an old family. *Clin. Exp. Pharmacol. Physiol.* 35 (2), 107–112. doi: 10.1111/j.1440-1681.2007.04843.x

Ma, R., Lai, J., Chen, X., Wang, L., Yang, Y., Wei, S., et al. (2021). A novel phage infecting alteromonas represents a distinct group of siphophages infecting diverse aquatic copiotrophs. *mSphere* 6 (3), e0045421. doi: 10.1128/mSphere.00454-21

Mahony, J., Alqarni, M., Stockdale, S., Spinelli, S., Feyereisen, M., Cambillau, C., et al. (2016). Functional and structural dissection of the tape measure protein of lactococcal phage TP901-1. *Sci. Rep.* 6, 36667. doi: 10.1038/srep36667

Middelboe, M., Chan, A., and Bertelsen, S. (2010). Isolation and life cycle characterization of lytic viruses infecting heterotrophic bacteria and cyanobacteria. In S. W. Wilhelm, M. G. Weinbauer, & C. A. Suttle (Eds.), Manual of aquatic viral ecology (pp. 118–133). doi: 10.4319/mave.2010.978-0-9845591-0-7.118

Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G. A., Sonnhammer, E. L. L., et al. (2021). Pfam: the protein families database in 2021. *Nucleic Acids Res.* 49 (D1), D412–d419. doi: 10.1093/nar/gkaa913

Mistry, J., Finn, R. D., Eddy, S. R., Bateman, A., and Punta, M. (2013). Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions. *Nucleic Acids Res.* 41 (12), e121. doi: 10.1093/nar/gkt263

Murray-Rust, J., Leiper, J., McAlister, M., Phelan, J., Tilley, S., Santa Maria, J., et al. (2001). Structural insights into the hydrolysis of cellular nitric oxide synthase inhibitors by dimethylarginine dimethylaminohydrolase. *Nat. Struct. Biol.* 8 (8), 679–683. doi: 10.1038/90387

Nadalin, F., Vezzi, F., and Policriti, A. (2012). GapFiller: a *de novo* assembly approach to fill the gap within paired reads. *BMC Bioinf.* 13 Suppl 14 (Suppl 14), S8. doi: 10.1186/1471-2105-13-S14-S8

Pedersen, M., Ostergaard, S., Bresciani, J., and Vogensen, F. K. (2000). Mutational analysis of two structural genes of the temperate lactococcal bacteriophage TP901-1 involved in tail length determination and baseplate assembly. *Virology* 276 (2), 315–328. doi: 10.1006/viro.2000.0497

Pesant, S., Not, F., Picheral, M., Kandels-Lewis, S., Le Bescot, N., Gorsky, G., et al. (2015). Open science resources for the discovery and analysis of Tara oceans data. *Sci. Data* 2, 150023.

Qu, T., Hou, C., Zhao, X., Zhong, Y., Guan, C., Lin, Z., et al. (2021). Bacteria associated with ulva prolifera: a vital role in green tide formation and migration. *Harmful Algae* 108, 102104. doi: 10.1016/j.hal.2021.102104

Qu, T., Zhao, X., Hao, Y., Zhong, Y., Guan, C., Hou, C., et al. (2020). Ecological effects of ulva prolifera green tide on bacterial community structure in qingdao offshore environment. *Chemosphere* 244, 125477. doi: 10.1016/j.chemosphere.2019.125477

Rawlings, N. D., and Barrett, A. J. (2004). "67 - introduction: metallopeptidases and their clans," in *Handbook of proteolytic enzymes (Second edition)*. Eds. A. J. Barrett, N. D. Rawlings and J. F. Woessner (London: Academic Press), 231–267.

Ren, L., Lu, Z., Xia, X., Peng, Y., Gong, S., Song, X., et al. (2022). Metagenomics reveals bacterioplankton community adaptation to long-term thermal pollution through the strategy of functional regulation in a subtropical bay. *Water Res.* 216, 118298. doi: 10.1016/j.watres.2022.118298

Rong, J. C., Cha, Q. Q., Ren, X. B., Xie, B. B., and Song, X. Y. (2018). Complete genome sequence of pseudoalteromonas espejiana DSM9414(T), an amino-acid-requiring strain from seawater. *Mar. Genomics* 38, 21–23. doi: 10.1016/j.margen.2017.08.006

Salazar, G., Paoli, L., Alberti, A., Huerta-Cepas, J., Ruscheweyh, H. J., Cuenca, M., et al. (2019). Gene expression changes and community turnover differentially shape the global ocean metatranscriptome. *Cell* 179 (5), 1068–1083.e1021. doi: 10.1016/j.cell.2019.10.014

Sharma, R., Pielstick, B. A., Bell, K. A., Nieman, T. B., Stubbs, O. A., Yeates, E. L., et al. (2019). A novel, highly related jumbo family of bacteriophages that were isolated against erwinia. *Front. Microbiol.* 10, 1533. doi: 10.3389/fmicb.2019.01533

Stewart, F. J., Ulloa, O., and DeLong, E. F. (2012). Microbial metatranscriptomics in a permanent marine oxygen minimum zone. *Environ. Microbiol.* 14 (1), 23–40. doi: 10.1111/j.1462-2920.2010.02400.x

Tamura, S., Toriumi, S., Saito, J., Awano, K., Kudo, T. A., and Kobayashi, T. (2006). PP2C family members play key roles in regulation of cell survival and apoptosis. *Cancer Sci.* 97 (7), 563–567. doi: 10.1111/j.1349-7006.2006.00219.x

Truglio, J. J., Rhau, B., Croteau, D. L., Wang, L., Skorvaga, M., Karakas, E., et al. (2005). Structural insights into the first incision reaction during nucleotide excision repair. *EMBO J.* 24 (5), 885–894. doi: 10.1038/sj.emboj.7600568

Uyà, M., Maggi, E., Mori, G., Nuccio, C., Gribben, P. E., and Bulleri, F. (2017). Carry over effects of nutrient addition on the recovery of an invasive seaweed from the winter die-back. *Mar. Environ. Res.* 126, 37–44. doi: 10.1016/j.marenvres.2017.02.006

Wang, C., Chen, Q., Zhang, C., Yang, J., Lu, Z., Lu, F., et al. (2017). Characterization of a broad host-spectrum virulent salmonella bacteriophage fmb-p1 and its application on duck meat. *Virus Res.* 236, 14–23. doi: 10.1016/j.virusres.2017.05.001

Wang, Q., Liu, Y., Liu, Q., Liu, X., Yang, F., Wang, M., et al. (2020). Isolation and complete genome of the marine pseudoalteromonas phage C7 from coastal seawater of yellow Sea, China. *Curr. Microbiol.* 77 (2), 279–285. doi: 10.1007/s00284-019-01815-4

Wang, F., Marshall, C. B., and Ikura, M. (2013). Transcriptional/epigenetic regulator CBP/p300 in tumorigenesis: structural and functional versatility in target recognition. *Cell Mol. Life Sci.* 70 (21), 3989–4008. doi: 10.1007/s00018-012-1254-4

Wang, D. B., Sun, M. Q., Shao, H. B., Li, Y., Meng, X., Liu, Z. Y., et al. (2015). Characterization and genome sequencing of a novel bacteriophage PH101 infecting pseudoalteromonas marina BH101 from the yellow Sea of China. *Curr. Microbiol.* 71 (5), 594–600. doi: 10.1007/s00284-015-0896-5

Wang, Z., Zhang, F., Liang, Y., Zheng, K., Gu, C., Zhang, W., et al. (2021). Genome and ecology of a novel alteromonas podovirus, ZP6, representing a new viral genus, mareflavirus. *Microbiol. Spectr.* 9 (2), e0046321.

Xiao, J., Wang, Z., Liu, D., Fu, M., Yuan, C., and Yan, T. (2021). Harmful macroalgal blooms (HMBs) in china's coastal water: green and golden tides. *Harmful Algae* 107, 102061. doi: 10.1016/j.hal.2021.102061

Xu, J., Cai, Y., Ma, Z., Jiang, B., Liu, W., Cheng, J., et al. (2021). DEAD/DEAH-box helicase 5 is hijacked by an avian oncogenic herpesvirus to inhibit interferon beta

production and promote viral replication. *Dev. Comp. Immunol.* 119, 104048. doi: 10.1016/j.dci.2021.104048

Xu, M., Guo, L., Gu, S., Wang, O., Zhang, R., Peters, B. A., et al. (2020). TGS-GapCloser: a fast and accurate gap closer for large genomes with low coverage of error-prone long reads. *Gigascience* 9 (9), giaa094. doi: 10.1093/gigascience/giaa094

Yang, Q., Gao, C., Jiang, Y., Wang, M., Zhou, X., Shao, H., et al. (2019). Metagenomic characterization of the viral community of the south Scotia ridge. *Viruses* 11 (2), 95. doi: 10.3390/v11020095

Yuan, Y., Li, Q., Zhang, S., Gu, J., Huang, G., Qi, Q., et al. (2022). Enhancing thermal stability and lytic activity of phage lysin PlyAB1 from acinetobacter baumannii. *Biotechnol. Bioengineering* 119 (10), 2731–2742. doi: 10.1002/bit.28187

Zhang, Y., He, P., Li, H., Li, G., Liu, J., Jiao, F., et al. (2019). Ulva prolifera green-tide outbreaks and their environmental impact in the yellow Sea, China. *Natl. Sci. Rev.* 6 (4), 825–838. doi: 10.1093/nsr/nwz026

Zhang, X., Liang, Y., Zheng, K., Wang, Z., Dong, Y., Liu, Y., et al. (2023). Characterization and genomic analysis of phage vB_ValR_NF, representing a new viral family prevalent in the ulva prolifera blooms. *Front. Microbiol.* 14, 1161265. doi: 10.3389/fmicb.2023.1161265

Zhang, T., and Wang, X. (2017). Release and microbial degradation of dissolved organic matter (DOM) from the macroalgae ulva prolifera. *Mar. pollut. Bull.* 125 (1-2), 192–198. doi: 10.1016/j.marpolbul.2017.08.029

Zhang, R., Weinbauer, M. G., and Qian, P. Y. (2007). Viruses and flagellates sustain apparent richness and reduce biomass accumulation of bacterioplankton in coastal marine waters. *Environ. Microbiol.* 9 (12), 3008–3018. doi: 10.1111/j.1462-2920.2007.01410.x

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

Zhao, G., He, H., Wang, H., Liang, Y., Guo, C., Shao, H., et al. (2022). Variations in marine bacterial and archaeal communities during an ulva prolifera green tide in coastal qingdao areas. *Microorganisms* 10 (6), 1204. doi: 10.3390/microorganisms10061204

Zhao, Y., Qin, F., Zhang, R., Giovannoni, S. J., Zhang, Z., Sun, J., et al. (2019). Pelagiphages in the podoviridae family integrate into host genomes. *Environ. Microbiol.* 21 (6), 1989–2001. doi: 10.1111/1462-2920.14487

Zhao, Y., Temperton, B., Thrash, J. C., Schwalbach, M. S., Vergin, K. L., Landry, Z. C., et al. (2013). Abundant SAR11 viruses in the ocean. *Nature* 494 (7437), 357–360.

Zheng, K., Dong, Y., Liang, Y., Liu, Y., Zhang, X., Zhang, W., et al. (2023). Genomic diversity and ecological distribution of marine Pseudoalteromonas phages. *Mar. Life Sci. Technol* 5 (2), 271–285. doi: 10.1007/s42995-022-00160-z

Zhu, B., Wang, L., Mitsunobu, H., Lu, X., Hernandez, A. J., Yoshida-Takashima, Y., et al. (2017). Deep-sea vent phage DNA polymerase specifically initiates DNA synthesis in the absence of primers. *Proc. Natl. Acad. Sci.* 114 (12), E2310–E2318. doi: 10.1073/pnas.1700280114

Zimmerman, A. E., Howard-Varona, C., Needham, D. M., John, S. G., Worden, A. Z., Sullivan, M. B., et al. (2020). Metabolic and biogeochemical consequences of viral infection in aquatic ecosystems. *Nat. Rev. Microbiol.* 18 (1), 21–34. doi: 10.1038/s41579-019-0270-x

frontiers | Frontiers in Marine Science

TYPE Original Research
PUBLISHED 14 December 2023
DOI 10.3389/fmars.2023.1168623

OPEN ACCESS

EDITED BY
Yunyun Zhuang,
Ocean University of China, China

REVIEWED BY
Wang Minxiao,
Chinese Academy of Sciences (CAS), China
Jinsong Zheng,
Chinese Academy of Sciences (CAS), China
Jianwei Chen,
Beijing Genomics Institute (BGI), China

*CORRESPONDENCE
Susan A. Smith
susansmith@mysticaquarium.org;
susan.smith@uconn.edu

RECEIVED 17 February 2023
ACCEPTED 23 October 2023
PUBLISHED 14 December 2023

CITATION
Smith SA, Ropati DV, Santoferrara LF, Romano TA and McManus GB (2023) The respiratory microbiota of three cohabiting beluga whales (*Delphinapterus leucas*) under human care.
Front. Mar. Sci. 10:1168623.
doi: 10.3389/fmars.2023.1168623

The respiratory microbiota of three cohabiting beluga whales (*Delphinapterus leucas*) under human care

Susan A. Smith[1,2*], Destiny V. Ropati[1,2], Luciana F. Santoferrara[3], Tracy A. Romano[1,2] and George B. McManus[2]

[1]Research Department, Sea Research Foundation, Inc. d/b/a Mystic Aquarium, Mystic, CT, United States, [2]Department of Marine Sciences, University of Connecticut, Groton, CT, United States, [3]Department of Biology, Hofstra University, Hempstead, NY, United States

We sampled the respiratory mucus from voluntary blowhole exhalations ("blow") of three healthy beluga whales (*Delphinapterus leucas*) under professional human care. Blow samples were collected from three resident belugas, one adult male (M1) and two adult females (F1, F2), with voluntary behaviors via non-invasive methods over three days in July 2021 (four days for M1). Samples were weighed and examined microscopically for the enumeration of eukaryotic and prokaryotic microbes, and then were used to evaluate carbon substrate use and taxonomic diversity of prokaryotic communities in the host respiratory sytem. Microscopical observations and 18S rRNA gene sequencing indicated the presence of eukaryotic microbiota, the ciliate genera *Planilamina* and *Kyaroikeus* in all three individuals. Exposure of samples to different metabolic carbon substrates indicated significant differences in the number of carbon sources usable by the prokaryotic communities of different whales (range: 11-25 sources), as well as a signficantly decreased diversity of carbon sources used by the community in the habitat water (5 sources). Sequencing of the hypervariable V4 region of the 16S rRNA gene revealed 19 amplicon sequence variants (ASVs) that were present in all whale samples. The oldest female *D. leucas* (F2) had the lowest overall diversity, and was significantly different from M1 and F1 in taxon composition, including an anomalously low ratio of Baccillota: Bacteroidota (0.01) compared to the other whales. In comparisons of microbial community composition, M1 had a significantly higher diversity than F1 and F2. These results suggest that attention should be given to regular microbiome sampling, and indicate a need for the pairing of microbiome and clinical data for animals in aquaria. Overall, these data contribute to the growing database on the core respiratory microbiota in cohabiting cetaceans under professional human care, indicate the utility of non-invasive sampling, and help characterize a baseline for healthy *D. leucas.*

KEYWORDS

microbiome, *Delphinapterus leucas*, microbiota, ciliate, blowhole, blowhole microbiota, marine mammal

Introduction

The respiratory microbiome in humans and model organisms is widely-recognized as a vital physiological component in the protection against pathogens and the maintainence of a dynamic immune system (Lombardo, 2008; Fraune and Bosch, 2010). Disruptions in the natural microbiome community (dysbiosis) can indicate symptomatic response to immune system dysregulation, disease manifestation, nutritional deficiencies, medicinal interactions, as well as psychological and physical stress from external or environmental factors (O'Dwyer et al., 2016; Shukla et al., 2017). However, limited analyses exist for marine mammals. In cetacea, respiratory illness is a primary cause of mortality (Venn-Watson et al., 2012), and contributes to the vulnerability of wild and endangered populations (Waltzek et al., 2012). However, blowhole exhalate (blow) sampling can provide a unique window into the health of the respiratory system, including identification of pathogens, illness, and dysbiosis caused by external stressors (Lombardo, 2008; Fraune and Bosch, 2010; Lima et al., 2012; Nelson et al., 2015) using non-invasive methods. Regular monitoring of the blow microbiome can serve as a warning sign in the detection of early dysbiosis, indicating the presence of an active or ongoing health issue (e.g. infection, external-stress). For cetaceans under professional human care, the non-invasive sampling of blow microbiota also allows for the continuous collection of biological data, which can inform and supplement clinical health monitoring, including the identification of dysbiosis, the effect of antibiotic treatments, and testing for infectious pathogens. Further, as opposed to more canonical health assessments accomplished via blood draw, the respiratory blow of cetaceans can be collected using entirely non-invasive methods, which has application for monitoring health in wild populations. The limited microbiome data available for toothed whales has provided a foundational context, including the identification of respiratory pathogens and the presence of microbial taxa which suggests a core microbial community (Apprill, 2017). However, more data is needed to examine the host-microbe relationship between cohabitating animals.

In this study, we used a non-invasive sampling procedure to collect blow from the beluga whale, *Delphinapterus leucas*. Access to three resident *D. leucas* individuals, in combination with specific learned voluntary behaviors for blow sample collection, allowed the opportunity to collect and analyze the metabolic and taxonomic composition of the associated microbial community. To characterize the microbiome, we used three methods. Clone libraries of 16S and 18S ribosomal gene sequences provide a comprehensive phylogenetic assessment of which taxonomic groups are present, including relative abundances. Traditional fluorescence microscopic observation and enumeration of prokaryotic and eukaryotic microbes present in the microbiome provided information on the magnitude of microbial populations, but minimal information on diversity. Both genetic and microscopic analyses are "culture independent" methods that avoid post-sampling changes in the microbial assemblage. The third method we used required incubation of blow samples with single carbon source substrates (Garland and Mills, 1991). This method allowed us to evaluate metabolic diversity of the microbiome community.

These data serve as a valuable examination in the viability of respiratory microbiome testing on *D. leucas* using non-invasive means. Use of these methods would be ideal to track changes in the health of wild populations, including the detection of pathogens and correlations with existing clinical health assessments (Acevedo-Whitehouse et al., 2010; Thompson et al., 2022). Further, the tested application of these non-invasive sampling methods is particularly critical in the study of endangered populations of *D. leucas*, which are in immediate need of health monitoring and less-invasive sampling methods.

Materials and methods

Sampling

We collected the exhaled breath condensate (blow) from the voluntary exhalations of three healthy beluga whales under professional human care at the Mystic Aquarium in Mystic, Connecticut. Individuals included M1, a 19 year old adult male, and two adult females, F1 and F2, 38 and 39 years old, respectively. Samples from three dates (7/1/21, 7/14/21, and 7/21/21) were collected from all three whales, with an additional sample collected from M1 on 8/2/21, and used to extract and sequence DNA from the blow microbiome. The whales have been cohabiting in a circulating saltwater system that is disinfected using ozone. Regular clinical health assessments by on-site veterinarians (e.g. blood cell counts, ultrasound, daily visual inspection, etc.) confirmed that all three whales remained stable and healthy throughout the sampling period.

Blow was collected on an inverted, pre-weighed, sterile plastic petri dish via established voluntary positive reinforcement behaviors. Briefly, each animal responded to learned visual cues developed over long-term training from an animal husbandry specialist that signaled the whales to position their head above the water surface, and then exhale on cue. After an initial exhalation to clear seawater, three consecutive exhalations were collected on one petri dish, and two dishes were used at each sampling event. Two mL of filter-sterilized seawater was added to prevent desiccation. The dish was then covered and sealed with parafilm, and transported to the laboratory in an insulated cooler. The amount of mucus per sample was estimated by reweighing the dishes. In addition to the blow samples, 100 mL of water from the whale habitat ("Habitat") was filtered onto a 0.2 µm glass fiber filter (GFF), cut into pieces using sterilized forceps, and then preserved in a 2.0 mL tube containing lysis buffer (100 mM NaCl, 0.5% SDS, Tris–EDTA at pH 8) and stored at −20°C until DNA extraction.

Bacterial counting

Samples for bacterial counts were preserved with formaldehyde (0.5% final concentration v/v). One mL of preserved sample was stained with the fluorochrome DAPI, collected on a black-stained polycarbonate membrane filter (0.2 µm pore-size), and counted with a fluorescence microscope under ultraviolet illumination. A minimum of 200 bacterial cells were counted per sample. Bacterial abundance is reported as cells per mL of the diluted sample (blow sample plus the added 2 mL filtered seawater).

Prokaryotic carbon metabolism

To assess the carbon substrates used by prokaryotes living in the blowhole and in the habitat water, we used EcoPlates from Biolog, Inc. (Hayward, CA, USA; catalog number 1506). These 96-well plates contain 31 different carbon sources plus a blank, in triplicate. After sample incubation, the presence of species capable of using a given substrate as sole carbon source is indicated by color development of a redox-sensitive tetrazolium dye. Samples of blow or aquarium water were added to each well (150 µl) and diluted with 300 µl of phosphate-buffered saline to avoid calcium precipitation (Pierce et al., 2014). We standardized experimental incubations to three days at 20˚C and measured color development at t=0 and t=3d using a BioTek ELx808 microplate reader, with absorbance at 590 nm. Each well was blanked by its own t=0 value. Wells were considered positive for a given carbon source if they had absorbance values greater than 0.25. The whales and aquarium water were assayed in triplicate (i.e. one 96-well plate for each sample). For the purpose of comparison, we focused on the number of different carbon substrates that could be metabolized in each sample.

DNA extraction, sequencing, and bioinformatics

Frozen habitat water samples collected on GFF were thawed, 15uL of proteinase K was added, and then tubes were incubated in a 56°C water bath overnight. Filters were then processed using the Quick-DNA^{TM} Fecal/Soil Microbe Microprep Kit (Zymo Research; model D6012) using the manufacturer's protocol for fecal samples. For blow samples, the Qiagen Dneasy Blood and Tissue Kit was used for DNA extraction, following manufacturer's instructions for tissue samples, except for an additional overnight incubation period in lysis buffer to improve the efficiency of DNA extraction. Briefly, one mL of lysis buffer (for 40 mL: 2mL 10% SDS, 8mL EDTA 0.5 M pH 8, 0.4 mL Tris HCL 1M pH8 + 29.6 mL ddH2O), 1mL of AL buffer, and 5 µL of proteinase K were added to 1 mL of blow sample and digested overnight as above. Extracted DNA concentration was quantified using Qubit high sensitivity DNA assays (ThermoFisher; model Q32851) and through gel visualization. PCR amplification with V4 16S rRNA primers for Bacteria and Archaea (Walters et al., 2016) was completed using DreamTaq polymerase, with PCR amplification conditions as follows: 95°C for 3 min; 35 cycles of 94°C for 45 s, 50°C for 60 s, and 72°C for 90 s; 72°C for 10 min, and then a 10°C hold. For the examination of ciliate presence, PCR amplification of the 18S rRNA gene was achieved using the "Anti-metazoan" primer set (i.e. biased against the amplification of metazoan 18S rRNA; del Campo et al., 2019) 574F (forward: 5'-CGGTAAYTCCAGCTCYV-3'; Hugerth et al., 2014) and UNonMet_DB (reverse: 5'-CTTTAARTTTCASYCTTGCG-3'; Bass and del Campo, 2020) and the following conditions: 95°C for 3 min; 35 cycles of 98°C for 10 s, 51.1°C for 30 s, and 72°C for 1min; with a final 72°C for 10 min. Products were then sent for sequencing with the Illumina MiSeq instrument at the Microbial Analysis, Resources and Services (MARS) facility at the University of Connecticut. Sequences were demultiplexed and quality-filtered in BaseSpace (Illumina).

Subsequent processing was done in QIIME 2 (Bolyen et al., 2019), including primer removal and additional quality filtering. Denoising, dereplication and chimera removal were completed with DADA2 (Callahan et al., 2016). For taxonomic classification, a Naïve Bayes classifier (Pedregosa et al., 2011) pre-trained on a V4-trimmed version of the 16S SILVA database (v.132) (Quast et al., 2012) was used for prokaryote sequences, and a Naïve Bayes classifier pre-trained on pr2 (v. 5.0.1) (Guillou et al., 2013) was used in the classification of sequences from Anti-metazoan primer sequencing. Statistical analyses regarding alpha-diversity indices, multivariate analyses, ANOSIM, and taxonomic visualizations were completed post rarefaction at 8,000 reads using QIIME 2. Eukaryotic sequence analyses regarding taxa identity, query cover and percent identity in regards to ciliates were compared to existing sequences on GenBank using blastn.

Results

Blow observations and weight

After correcting for the 2 mL of filtered seawater added to prevent the blow from drying, the amount of material collected per plate (3 blows) ranged from 0.23 to 0.56 g, (Table 1), with M1 consistently producing the greatest amount of mucus. We observed dysteriid ciliates of the genera *Kyaroikeus* and *Planilamina* in the blow (Figures 1A–C; Supplementary Video SV1) in M1 and F1, but not in F2 or in a concentrated volume of habitat water.

Bacterial enumeration

Bacterial concentrations ranged between about 10^7 and 10^8 per mL of diluted blow (includes sample plus 2 mL filtered seawater). M1 had the highest levels of bacteria on four of the six sampling dates, while F2 had the lowest.

Carbon substrates

The prokaryotic microbiome community showed growth on various sole carbon source media in the EcoPlates, with the whale assemblage consistently using a higher number of carbon substrates than that of the habitat water (Table 1). Of all 31 carbon sources tested, the total number used, summed across all 3 replicates, was 5 (aquarium water), 11 (M1), 20 (F1), and 25 (F2). Differences were found among the whales (one-way ANOVA; $P<0.001$), with F2 having the highest carbon source richness (average 22 substrates; n=3), and M1 having the lowest (8 substrates) (Table 1).

TABLE 1 Carbon substrate use and blow weight.

	Habitat	M1	F1	F2
Mean number of substrates used	2.67	8	14	21.67
95% CI	1.43	4.3	8.61	3.79
Mean blow wt per sample (g)	–	0.560	0.328	0.230
SD	–	0.303	0.158	0.063

Taxonomic diversity

Eukaryotic Microbiota. Although the intended targets for Anti-metazoan 18S rRNA primers were non-metazoan eukaryotes, many sequences were classified as unintended bacterial targets. For remaining sequences, the majority of eukaryotes matched ciliates (family Dysteriidae), dinoflagellates (family Thoracosphaeraceae), diatoms (class Bacillariophyceae), and green algae (family Chlorodendraceae) (Supplementary Table S1). No ciliate sequences were identified in the Habitat, which was instead dominated mostly by dinoflagellates (59.85% relative frequency), diatoms (27.10%), and green algae (4.24%). While the primers were intended to amplify only non-metazoan targets, some sequences in the Habitat were found to match the class Insecta (0.27%). For all whale samples, most sequences were bacteria (avg. relative frequencies for M1: 78%, F21: 87%, F2:94%), while the remainder was dominated by ciliates (M1: 22%, F1: 12%, F2: 6%), although some individuals were found to also have dinoflagellates (F1: 1%), and green algae (M1: 0.2%, F1:0.6%). All Dysteriidae ciliate sequences matched *Planilamina ovata* under the accession MN830169 or *Kyaroikeus paracetarius* under the accession number MN830168, and were present in all blow samples from all whales.

Prokaryotic microbiota

Among the prokaryotic 16S rRNA gene sequences found in M1, F1, and F2, 11 ASVs comprised what we consider to be the core prokaryotic microbiota for these three animals (i.e. ASVs shared by all three individuals in 100% of samples for each group) (Figures 2, 3; Supplementary Table S2). In addition, 8 ASVs were also found in Habitat water samples (i.e. 19 ASVs found in all three whales, 8 of which were found in the habitat water and may be of environmental origin). A total of 39 ASVs were found to be unique to M1 and F1samples (Figure 3; Supplementary Table S2), while only 4 ASVs were unique between M1 and F2, and only a single ASV was reported as unique between F1 and F2. For an ASV to be considered it had to be present in all samples for each group, but samples were not pooled due to the possibility for high individual variability within groups. In all three whales, we found a core of three primary phyla at high relative frequency, Proteobacteria, (M1: avg. 37.1± 3.6; F1: avg. 53.9 ± 1.51; F2: avg. 76.7 ± 1.32) Bacteroidota (synonym Bacteroidetes) (M1: avg. 24.2 ± 5.29; F1: avg. 19.3 ± 2.85; F2: avg. 21.3 ± 1.6), and Campylobacterota (M1: avg. 14.35 ± 4.14; F1: avg. 9.6 ± 1.75; F2: avg. 1.23 ± 0.82) (Figure 2). In addition, whales M1 and F1 shared a high relative frequency of the five phyla Actinobacteriota, Patescibacteria, Spirochaetota, Fusobacteriota and Bacillota (Figure 2). Although present in F2, the relative frequency of Bacillota was substantially reduced (avg. 0.17%; range 0-0.4%) in comparison to all other samples from M1 and F1 (avg. 8.8%; range 5.6-12.4%) (Figure 2). Conversely, Bacteroidota remained similar for all three whales, resulting in substantial differences in the Bacillota to Bacteroidota ratio (Bl:B). Expressed via ratio of relative frequency for each phylum averaged for each whale, the Bacillota to Bacteroidota ratio (Bl:B) was 0.52 (12.6: 24.2) in M1, 0.52 (12.6: 24.2) in F1, and 0.01 (0.17: 21.3) in F2. The habitat water samples contained seven phyla that were not observed in the whale samples: WPS-2 (Eremiobacterota), Planctomycetota, Cyanobacteria, Bdellovibrionota, Dependentiae, Crenarchaeota, and MBNT15 (candidate phylum) (Figure 2). Further, habitat water was found to contain the only member of the Archaea identified in this

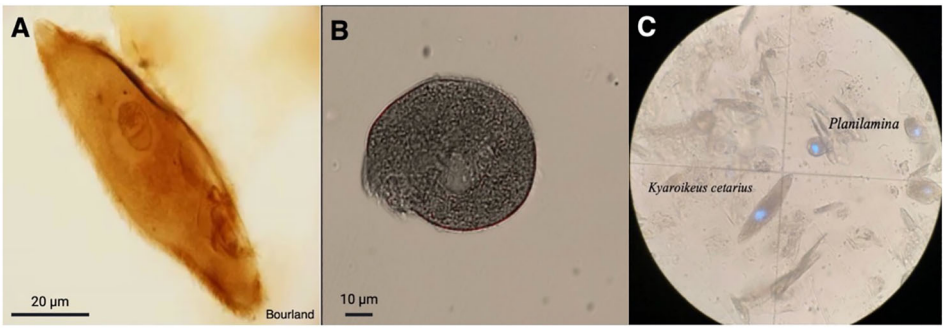


FIGURE 1
Ciliates observed in respiratory exhalate of *Delphinapterus leucas*. **(A)** *Kyaroikeus* post-silver impregnation using protargol stain; **(B)** *Planilamina* under bright field illumination; **(C)** *Kyaroikeus* and *Planilamina* surrounding mucosal exudate collected via blow exhalate on petri dish from M1, imaged in dark field.

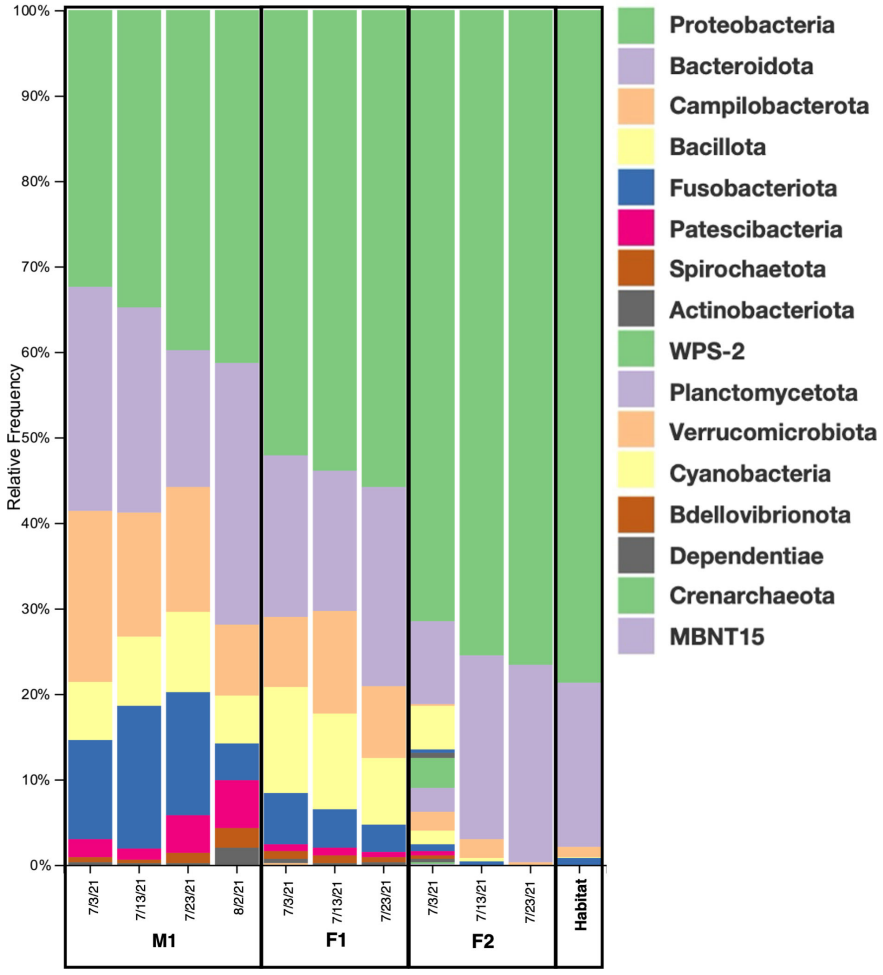


FIGURE 2
Taxonomic diversity. Relative frequencies of the bacterial taxa present in the blow microbiome of all three animals and habitat water on all sampling dates (M1=male 1; F1= female 1; F2= female 2). *Phylum level:* 17 total bacterial phyla were detected among all samples.

study, which were represented solely by the family Nitrosopumilaceae (genus *Candidatus Nitrosotenuis*, 0.2% relative frequency). Habitat samples also contained a relatively high frequency (36.5%) of the class Alphaproteobacteria, compared to an average of 0.18% (0-0.5%) in the whale samples (Supplementary Figure S1). There were no variations in the presence/absence of phyla found within each whale microbiome between sampling days; the same prokaryotic communities were present for each whale at all time points.

Significant differences were found in alpha diversity, measured by ASV richness, between F2 and M1 (Kruskal-Wallis; p-value 0.034), and F2 and F1 (p-value 0.049) (Figure 4; Supplementary Tables S3-S5), with whale F2 maintaining the lowest diversity among all samples (Figure 4). Nearly significant differences were found in alpha diversity among all four groups (i.e. M1, F1, F2, Habitat) (Kruskal-Wallis for all groups; p-value 0.0579) (Supplementary Table S6). Alpha rarefaction analyses indicated appropriate sampling depth was achieved post-rarefaction via Shannon index analysis (Supplementary Figures S2A, C), but Faith's PD (phylogenetic diversity) index suggested sampling depth may not have been adequate for the full sequencing of the habitat water microbiome (Supplementary Figure S2B).

Principal Component Analyses (PCA) indicated a clustering among samples for each whale, as well as compositional differences between whales and habitat water (based on Bray-Curtis values) (Figure 5). ANOSIM comparisons via pairwise Bray-Curtis distances indicated significant differences between M1 and F1 (p-value = 0.029) as well as M1 and F2 (p value 0.035) (Figures 4, 6; Supplementary Table S3).

Taxa of Interest

There were 11 ASVs found in all whale samples (for all M1, F1, and F2 samples) that were not present in the Habitat water. These were represented by the genera *Ganjinia, Suttonella, Porphyromonas, Psychrobacter*, *Buytrivibrio*, *Shewanella*, *Marinifilum, Paracoccus*, and genera within the Weeksellaceae and Cardiobacteriaceae families. There was a higher relative frequency of the genus *Vibrio* (order Vibrionales, family Vibrionaceae) in F2, with an average of 16.1% relative frequency (12.3- 20.5%), as compared to all other samples (average 1.39%; 0.01-3.54%) (Supplementary Table S2). Also found in F2 was a high relative frequency of ASVs affiliated to the species

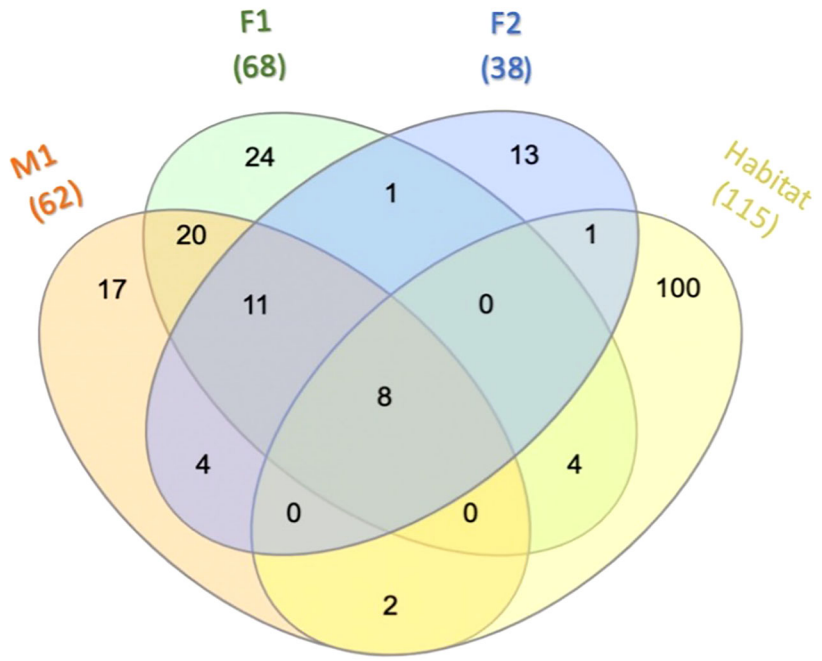


FIGURE 3
Core microbiota. Venn diagram depicting ASVs among M1, F1, F2, and Habitat that are present in 100% of the samples for each overlapping group. For additional details on ASVs see Supplementary Table S2.

Shewanella putrefaciens (avg. 30.8%; 22.3-40.7%), as compared to all other samples (avg. 2.6%; 0.1-13.3%). Two species within the genus *Myroides*, *Myroides* sp. (avg. 7.47%; 4.8-11.6%) and *M. odoratimimus* (avg. 6.5%; 4-10%), were found exclusively in F2. In the two whales frequently found to be harboring ciliate symbionts in their blow microbiota (M1 and F1), there was a high relative frequency (avg. 9.5%; 3-16.7%) of the genus *Oceanivirga*, when compared to F2 and Habitat, which had a relative frequency range of 0-0.7%. Whales M1 and F1 also had a high relative frequency of the family Arcobacteraceae (avg. 14.4% and 6.8%, respectively), as opposed to the rest of samples, which had an average of 0.52% (0.1-1.6%).

Discussion

Respiratory illness (e.g. pneumonia) is a primary cause of mortality in marine mammals (Venn-Watson et al., 2012), contributing to the vulnerability of wild populations (Waltzek et al., 2012). In cetaceans, blowhole microbiome sampling can not only identify the presence of infections, but also enables researchers to track physiological changes, especially those reflected in the respiratory system. However, the inherent difficulty in accessing these animals has led to a dearth of baseline information on their respiratory microbiota. Unique access to beluga whales, *Delphinapterus leucas*, which were trained to carry out behaviors that facilitate biological sampling, enabled us to use respiratory exhalate (blow) to quantify the respiratory microbiome.

Although microbiome studies generally exclude eukaryotic taxa (e.g. Apprill et al., 2017), the presence and sustained high concentrations of the ciliate genera *Kyaroikeus* and *Planilamina* in blow samples reflect a need to expand protist microbiome studies. Further, the limited research to date has led to suggestions that these ciliates are pathogenic or parasitic (Sniezek et al., 1995; Ma et al., 2006), whereas our continued observation of these protists in healthy animals suggests that they may instead be commensal or mutualistic members of a normal *D. leucas* respiratory microbiome.

Among the ASVs unique to all whale samples (i.e. not present in Habitat samples), the genera *Ganjinia, Suttonella, Porphyromonas, Weeksellaceae, Porphyromonas*, and genera within the Cardiobacteriaceae family had all previously been identified in the oral, respiratory, or mucosal microbiomes of various wild cetaceans while the genera *Shewanell*a and *Paracoccus* had only been associated with skin samples, and Buytrivibrio solely with the gut biome of wild pygmy sperm whales (see Supplementary Table S7). Interestingly, *Paracoccus* sp. had only been reported in the skin of killer whales and dolphins under human care by Chiarello et al., 2017, who suggested that the presence of these microbes in cetaceans might indicate a zoonotic transfer from human husbandry specialists to the animals, since *Paracoccus* sp. is a primary member of the human skin microbiome (Cosseau et al., 2016). While these data are only indicative of a small sample size, and reflect animals under professional human care in a circulating seawater system, our results are similar to published work on cetaceans (Apprill et al., 2017; Apprill et al., 2020; Van Cise et al., 2020), suggesting the existence of "core taxa," or conserved compositional taxonomy, as we observed in the respiratory exhalate of all three whales, which shared 19 core ASVs in their blow microbiome regardless of sampling date.

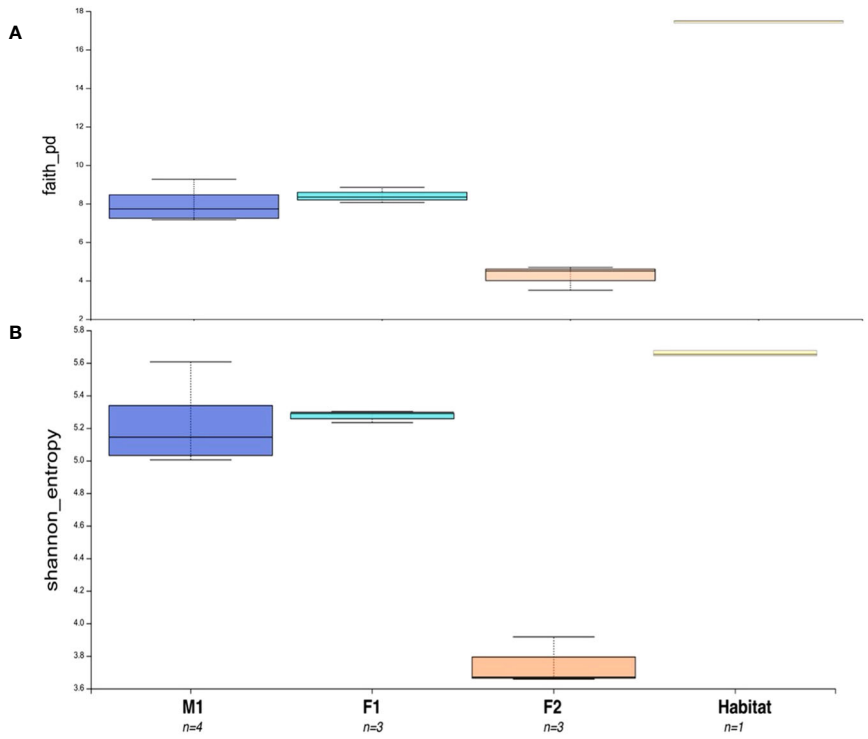


FIGURE 4
Alpha Diversity Boxplot comparisons between three animals and habitat water. **(A)** Faith's Phylogenetic Diversity (PD) Test representing differences in PD between M1, F1, F2, and Habitat microbiome samples. **(B)** Shannon Entropy for all sample groups. Number of samples for each group = n. Boxes represent the interquartile range (IQR) between the first and third quartiles (25th and 75th percentiles, respectively), and the horizontal line inside the box defines the median. Whiskers represent the lowest and highest values within 1.5 times the IQR from the first and third quartiles, respectively.

While a reduced sample size prevents these data from providing a representative microbiome community for *D. leucas*, microbiome analyses among cohabiting individuals can be useful in the study of respiratory microbiota by establishing a baseline for healthy animals, which can be used comparatively if dysbiosis occurs (e.g. due to antibiotics, change in immune health, chronic external stress, viral infection, etc.). Individual microbiome communities were also apparent between whales, with significant differences among M1, F1

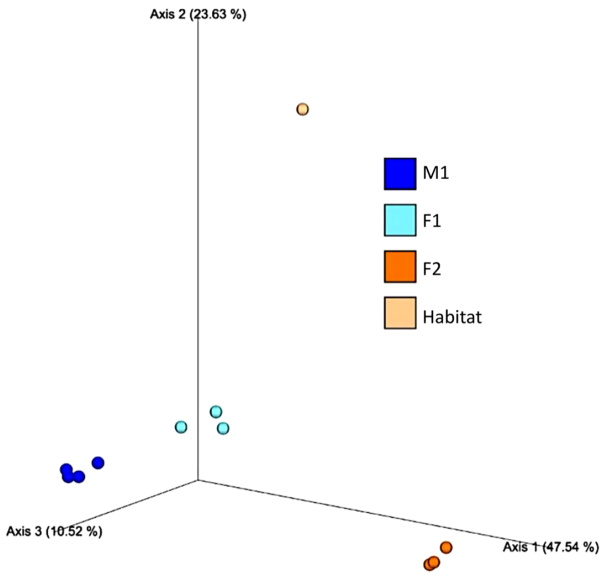


FIGURE 5
Principal Component Analysis (PCA) based on Bray-Curtis distances among whale blow (M1, F1 and F2) and Habitat water samples.

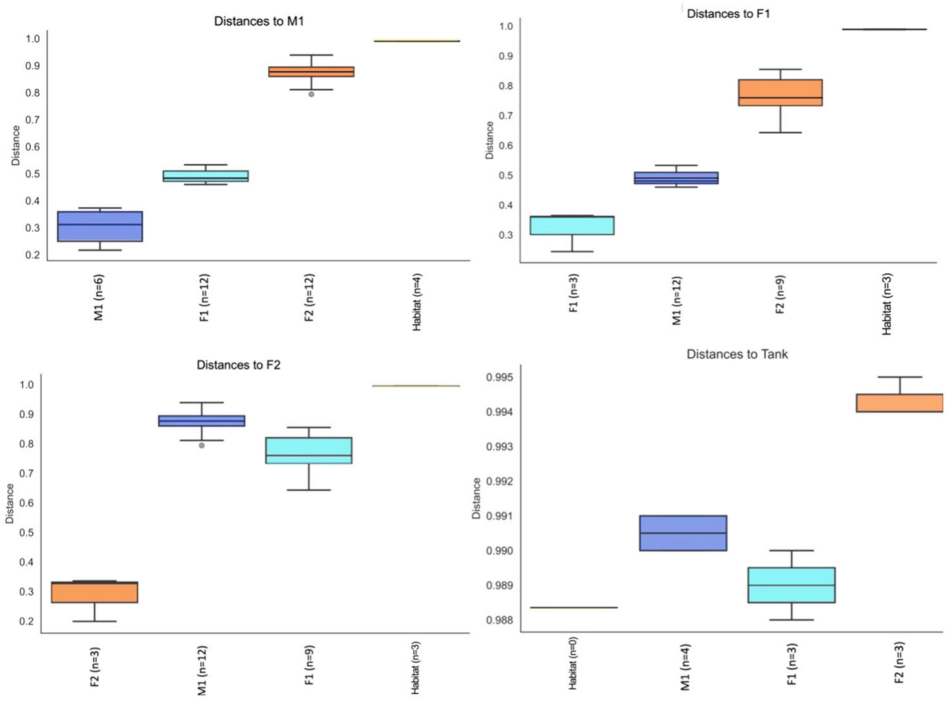


FIGURE 6
Pairwise Group Significance Comparisons (ANOSIM) between M1, F1, F2, and Habitat microbiome samples (n=number of pairwise comparisons).

and F2, including a substantial diversity reduction in F2. Veterinary assessment of all three animals determined the whales to be clinically healthy. F2 was being administered the gastroprotectants sucralfate and omeprazole. Omeprazole is a proton pump inhibitor (PPI), commonly used to decrease stomach acid production (Mishiro et al., 2018). Human studies have identified that PPIs can cause severe dysbiosis with only weeks of administration, which can lead to a sustained disruption to the microbial composition in esophageal microbiomes (Castellani et al., 2017; Mishiro et al., 2018). Further, similar studies have shown that regular doses can substantially alter the relative ratios of microbiome taxa, including an alteration in the Bacillota to Bacteroidota ratio (Naito et al., 2018). In human and non-human terrestrial animals, this ratio is regularly used as a reliable marker in which a high number (i.e. Bl:B) is correlated to the storage of fat (implicated with obesity diagnoses in humans), while a low number is generally linked to inflammatory bowel disease (DeGruttola et al., 2016; Hufnagl et al., 2020). In F2, we saw a substantially reduced Bacillota community in comparison to all other samples from M1 and F1 (Figure 2), while Bacteroidota remained similar for all three whales (Figure 2). This reflected a Bl:B ratio of 0.01 for F2, in comparison to 0.52 and 1.61, for M1 and F1, respectively. These data indicate a need for regular microbiome monitoring, and may warrant an investigation into probiotics or a change in diet for animals requiring PPIs. On the other hand, F2 also had ASVs unique to F2 including the genera *Peptostreptococcaceae* and *Alcaligenes*, both of which had previously only been identified in the gut microbiome of marine mammals (Rothenberg et al., 2021; Bai et al., 2022). The presence of these gut-dominant genera in the blow exhalate of F2 may suggest stomach material or acid reflux in this animal, possibly warranting the need for gastroprotectants.

Identifying and quantifying sustained presence of pathogen-associated taxa is complicated; many animals regularly coexist with pathogenic microbiota without showing clinical signs of an affected health assessment (Nelson et al., 2015; Palmer et al., 2020). In some marine mammals, the conserved presence of disease-associated microbes like some *Vibrio* species and *Helicobacter pylori* among different populations can exist as a background "pathogen core community" (Apprill et al., 2017; Godoy-Vitorino et al., 2017; Raverty et al., 2017; Li et al., 2019) that is found in the majority of healthy populations (Nelson et al., 2015; Apprill et al., 2017; Ochoa et al., 2018). The frequency of *Vibrio* sequences identified in F2 may warrant investigation, but may also be an indirect result of the reduced diversity caused by other factors (e.g. medicine, dietary change, etc.). Similarly, M1 and F1 had a higher frequency of the genus *Oceanivirga*, within which the species *O. salmonicida* is believed to be a pathogen of Atlantic salmon (*Salmo salar*) and has been found in oral microbiomes of several marine mammals, including dolphins, humpback whales, and seals (Eisenberg et al., 2016; Palmer et al., 2020). However, both *Vibrio* and *Oceanivirga* were identified in all samples, including Habitat water samples, suggesting they may be part of an ambient environmental background biome. This serves as an example where microbiome studies can supplement and inform clinical health assessments, including in decisions regarding quarantine procedures triggered by pathogen presence/absence tests, which may be unnecesary if there is a consistent background of a specific pathogenic taxon in clinically-healthy animals or their environment. Further work

including multiple samples taken from various microbiome regions (e.g. oral, skin, fecal, etc.) over time would greatly enhance this dataset and constrain implications regarding core taxa, within-group variation, and pathogenic potential.

The microscopic observations showed prokaryotic cell concentrations about an order of magnitude higher than typical ocean values of c. 10^6 cells l^{-1} and the presence of commensal ciliates in two of the whales. Whale F2 had the lowest bacterial concentrations and we did not observe ciliates in her blow during this sampling period. This is surprising, since the ciliates would presumably crop bacterial populations via grazing, so we might expect higher numbers in F2. On the other hand, F2 had the highest metabolic diversity in her microbiome, 1.5-2x higher than the other two whales and about eight-fold higher than that of the habitat water. We know of no other respiratory microbiome studies that looked at metabolic diversity, but there have been observations on changes, for example, in fish gut microbiome metabolic types related to salinity gradients and diet (Mouchet et al., 2012). Probably the most striking result of the metabolic diversity observations is the large difference between the whale respiratory tract and the habitat water. Given the well-known dominance of unculturable bacteria in the ocean, it is perhaps not surprising that the technique we used, which requires incubation over several days, would yield lower diversity in the aquatic habitat. Apparently, the whale respiratory tract is much more similar to a rich culture-medium environment and hence allows for higher diversity of culturable metabolic types to exist there.

The cetacean respiratory microbiome is an important yet understudied system that allows for the identification of infection, disease, and dysbiosis. Respiratory microbiota can maintain host health, support immune system function, and serve as both an indicator and source of illness. In cetaceans, blow microbiome sampling can identify physiological changes in individuals under professional human care, and serve as a non-invasive health index for wild populations. The sampling methods explored in this study represent the utility of non-invasive means to inform clinicians on the health of cetaceans under professional human care, and serve as an ideal method to monitor wild and endangered populations.

Data availability statement

The data presented in the study are deposited in the Sequence Read Archive repository (SRA; https://www.ncbi.nlm.nih.gov/sra) under the Bioproject accession number PRJNA1026968 (BioSample accession numbers: SAMN37769875, SAMN37769876, SAMN37769877, SAMN37769878, SAMN37769879, SAMN37769880, SAMN37769881, SAMN37769882, SAMN37769883, SAMN37769884). In addition, these data are publicly available on FigShare at https://doi.org/10.6084/m9.figshare.24328462 (Smith et al., 2023).

Ethics statement

The animal study was approved by the Mystic Aquarium Institutional Animal Care and Use Committee (IACUC) (protocol # 18018). The study was conducted in accordance with the local legislation and institutional requirements.

Author contributions

SS, DR, TR, and GM conceived and designed study. SS and LS analyzed data. All authors contributed to writing the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This research was funded by a National Science Foundation (NSF) Research Experiences for Undergraduates (REU) grant to Mystic Aquarium (Award # 1950480) and the University of Connecticut (Award # 1950415), and NSF Award # OCE-1924527.

Acknowledgments

We thank the staff of Mystic Aquarium for assistance with sampling, W. Bourland (Boise State University) for the protargol staining shown in Figure 1, and R. Cole for encouraging us to sample the whales for ciliates. This constitutes scientific contribution #365 from the Sea Research Foundation, Inc.

Conflict of interest

Authors SS, DR and TR were employed by Sea Research Foundation, Inc. d/b/a Mystic Aquarium.

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmars.2023.1168623/full#supplementary-material

References

Acevedo-Whitehouse, K., Rocha-Gosselin, A., and Gendron, D. (2010). A novel non-invasive tool for disease surveillance of free-ranging whales and its relevance to conservation programs. *Anim. Conserv.* 13, 217–225. doi: 10.1111/j.1469-1795.2009.00326.x

Apprill, A. (2017). Marine animal microbiomes: toward understanding host–microbiome interactions in a changing ocean. *Front. Mar. Sci.* 4, 222. doi: 10.3389/fmars.2017.00222

Apprill, A., Miller, C. A., Moore, M. J., Durban, J. W., Fearnbach, H., and Barrett-Lennard, L. G. (2017). Extensive core microbiome in drone-captured whale blow supports a framework for health monitoring. *MSystems* 2 (5), e00119–e00117. doi: 10.1128/mSystems.00119-17

Apprill, A., Miller, C. A., Van Cise, A. M., U'Ren, J. M., Leslie, M. S., Weber, L., et al. (2020). Marine mammal skin microbiotas are influenced by host phylogeny. *R. Soc. Open Sci.* 7 (5), 192046. doi: 10.1098/rsos.192046

Bai, S., Zhang, P., Zhang, X., Yang, Z., and Li, S. (2022). Gut microbial characterization of melon-headed whales (Peponocephala electra) stranded in China. *Microorganisms* 10 (3), 572.

Bass, D., and del Campo, J. (2020). Microeukaryotes in animal and plant microbiomes: Ecologies of disease? *Eur. J. Protistol.* 76, e125719. doi: 10.1016/j.ejop.2020.125719

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., et al. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* 37 (8), 852–857. doi: 10.1038/s41587-019-0209-9

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 (7), 581–583. doi: 10.1038/nmeth.3869

Castellani, C., Singer, G., Kashofer, K., Huber-Zeyringer, A., Flucher, C., Kaiser, M., et al. (2017). The influence of proton pump inhibitors on the fecal microbiome of infants with gastroesophageal reflux-a prospective longitudinal interventional study. *Front. Cell Infect. Microbiol.* 7. doi: 10.3389/fcimb.2017.00444

Chiarello, M., Villéger, S., Bouvier, C., Auguet, J. C., and Bouvier, T. (2017). Captive bottlenose dolphins and killer whales harbor a species-specific skin microbiota that varies among individuals. *Sci. Rep.* 7 (1), 15269. doi: 10.3389/fevo.2023.1197323

Cosseau, C., Romano-Bertrand, S., Duplan, H., Lucas, O., Ingrassia, I., Pigasse, C., et al. (2016). Proteobacteria from the human skin microbiota: species-level diversity and hypotheses. *One Health* 2, 33–41.

DeGruttola, A. K., Low, D., Mizoguchi, A., and Mizoguchi, E. (2016). Current understanding of dysbiosis in disease in human and animal models. *Inflammation Bowel Dis.* 22 (5), 1137–1150. doi: 10.1097/MIB.0000000000000750

del Campo, J., Pons, M., Herranz, M., Wakeman, K. C., Del Valle, J., Vermeij, M. J. A., et al. (2019). Validation of a universal set of primers to study animal-associated microeukaryotic communities. *Environ. Microbiol.* 21, 3855–3861. doi: 10.1111/1462-2920.14733

Eisenberg, T., Kämpfer, P., Ewers, C., Semmler, T., Glaeser, S. P., Collins, E., et al. (2016). *Oceanivirga salmonicida* gen. nov., sp. nov., a member of the Leptotrichiaceae isolated from Atlantic salmon (Salmo salar). *Int. J. Systematic Evolutionary Microbiol. 66*, 2429–2437. doi: 10.1099/ijsem.0.001050

Fraune, S., and Bosch, T. C. (2010). Why bacteria matter in animal development and evolution. *Bioessays* 32 (7), 571–580. doi: 10.1002/bies.200900192

Garland, J. L., and Mills, A. L. (1991). Classification and characterization of heterotrophic microbial communities on the basis of patterns of community-level sole-carbon-source utilization. *Appl. Environ. Microbiol.* 57, 2351–2359. doi: 10.1128/aem.57.8.2351-2359.1991

Godoy-Vitorino, F., Rodriguez-Hilario, A., Alves, A. L., Gonçalves, F., Cabrera-Colon, B., Mesquita, C. S., et al. (2017). The microbiome of a striped dolphin (Stenella coeruleoalba) stranded in Portugal. *Res. Microbiol.* 168 (1), 85–93. doi: 10.1016/j.resmic.2016.08.004

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

Hufnagl, K., Pali-Schöll, I., Roth-Walter, F., and Jensen-Jarolim, E. (2020). Dysbiosis of the gut and lung microbiome has a role in asthma. *Semin. Immunopathol.* 1), 75–93. doi: 10.1007/s00281-019-00775-y

Hugerth, L. W., Muller, E. E., Hu, Y. O., Lebrun, L. A., Roume, H., Lundin, D., et al. (2014). Systematic design of 18S rRNA gene primers for determining eukaryotic diversity in microbial consortia. *PloS One* 9 (4), e95567. doi: 10.1371/journal.pone.0095567

Li, C., Tan, X., Bai, J., Xu, Q., Liu, S., Guo, W., et al. (2019). A survey of the sperm whale (Physeter catodon) commensal microbiome. *PeerJ* 7, e7257. doi: 10.7717/peerj.7257

Lima, N., Rogers, T., Acevedo-Whitehouse, K., and Brown, M. V. (2012). Temporal stability and species specificity in bacteria associated with the bottlenose dolphins respiratory system. *Environ. Microbiol. Rep.* 4, 89–96. doi: 10.1111/j.1758-2229.2011.00306.x

Lombardo, M. (2008). Access to mutualistic endosymbiotic microbes: an underappreciated benefit of group living. *Behav. Ecol. Sociobiol.* 62, 479–497. doi: 10.1007/s00265-007-0428-9

Ma, H., Overstreet, R. M., Sniezek, J. H., Solangi, M., and Coats, W. D. (2006). Two new species of symbiotic ciliates from the respiratory tract of cetaceans with establishment of the new genus *Planilamina* n. gen.(Dysteriida, Kyaroikeidae). *J. Eukaryotic Microbiol.* 53 (6), 407–419. doi: 10.1111/j.1550-7408.2006.00124.x

Mishiro, T., Oka, K., Kuroki, Y., Takahashi, M., Tatsumi, K., Saitoh, T., et al. (2018). Oral microbiome alterations of healthy volunteers with proton pump inhibitor. *J. Gastroenterol. Hepatol.* 33, 1059–1066. doi: 10.1111/jgh.14040

Mouchet, M. A., Bouvier, C., Bouvier, T., Troussellier, M., Escalas, A., and Mouillot, D. (2012). Genetic difference but functional similarity among fish gut bacterial communities through molecular and biochemical fingerprints. *FEMS Microbiol. Ecol.* 79 (3), 568-580.

Naito, Y., Kashiwagi, K., Takagi, T., Andoh, A., and Inoue, R. (2018). Intestinal dysbiosis secondary to proton-pump inhibitor use. *Digestion* 97, 195–204. doi: 10.1159/000481813

Nelson, T. M., Apprill, A., Mann, J., Rogers, T. L., and Brown, M. V. (2015). The marine mammal microbiome: current knowledge and future directions. *Microbiol. Aust.* 36 (1), 8–13. doi: 10.1071/MA15004

Ochoa, J. L., Sanchez, L. M., Koo, B. M., Doherty, J. S., Rajendram, M., Huang, K. C., et al. (2018). Marine mammal microbiota yields novel antibiotic with potent activity against Clostridium difficile. *ACS Infect. Dis.* 4 (1), 59–67. doi: 10.1021/acsinfecdis.7b00105

O'Dwyer, D. N., Dickson, R. P., and Moore, B. B. (2016). The lung microbiome, immunity, and the pathogenesis of chronic lung disease. *J. Immunol.* 196 (12), 4839–4847. doi: 10.4049/jimmunol.1600279

Palmer, R., Fleming, G. T., Glaeser, S., Semmler, T., Flamm, A., Ewers, C., et al. (2020). Marine mammals are natural hosts of *Oceanivirga salmonicida*, a bacterial pathogen of Atlantic salmon. *Dis. Aquat. Organisms* 139, 161–174. doi: 10.3354/dao03478

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: machine learning in python. *J. Mach. Learn. Res.* 12, 2825–2830.

Pierce, M. L., Ward, J. E., and Dobbs, F. C. (2014). False positives in Biolog EcoPlatesTM and MT2 MicroPlatesTM caused by calcium. *J. Microbiological Methods* 97, 20–24. doi: 10.1016/j.mimet.2013.12.002

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 (D1), D590–D596. doi: 10.1093/nar/gks1219

Raverty, S. A., Rhodes, L. D., Zabek, E., Eshghi, A., Cameron, C. E., Hanson, M. B., et al. (2017). Respiratory microbiome of endangered southern resident killer whales and microbiota of surrounding sea surface microlayer in the Eastern North Pacific. *Sci. Rep.* 7 (1), 394. doi: 10.1038/s41598-017-00457-5

Rothenberg, S. E., Sweitzer, D. N., Rackerby, B. R., Couch, C. E., Cohen, L. A., Broughton, H. M., et al (2021). Fecal methylmercury correlates with gut microbiota taxa in Pacific walruses (Odobenus rosmarus divergens). *Front. Microbiol.* 12, 648685.

Shukla, S. D., Budden, K. F., Neal, R., and Hansbro, P. M. (2017). Microbiome effects on immunity, health and disease in the lung. *Clin. Trans. Immunol.* 6 (3), e133. doi: 10.1038/cti.2017.6

Smith, Susan (2003). The respiratory microbiota of three cohabiting beluga whales (Delphinapterus leucas) under human care. FigShare. Dataset. doi. 10.6084/m9.figshare.24328462.v1

Sniezek, J. H., Coats, D. W., and Small, E. B. (1995). *Kyaroikeus cetarius* ng, n. sp.: a parasitic ciliate from the respiratory tract of odonticete cetacea. *J. Eukaryotic Microbiol.* 42 (3), 260–268.

Thompson, L. A., Goertz, C. E., Quakenbush, L. T., Burek Huntington, K., Suydam, R. S., Stimmelmayr, R., et al. (2022). Serological detection of marine origin Brucella exposure in two Alaska beluga stocks. *Animals* 12 (15), 1932. doi: 10.3390/ani12151932

Van Cise, A. M., Wade, P. R., Goertz, C. E., Burek-Huntington, K., Parsons, K. M., Clauss, T., et al. (2020). Skin microbiome of beluga whales: spatial, temporal, and health-related dynamics. *Anim. Microbiome* 2 (1), 1–16. doi: 10.1186/s42523-020-00057-1

Venn-Watson, S., Daniels, R., and Smith, C. (2012). Thirty year retrospective evaluation of pneumonia in a bottlenose dolphin Tursiops truncatus population. *Dis. Aquat. Organ.* 99, 237–242. doi: 10.3354/dao02471

Walters, W., Hyde, E. R., Berg-Lyons, D., Ackermann, G., Humphrey, G., and Parada, A. (2016). Improved bacterial 16S rRNA gene (V4 and V4-5) and fungal internal transcribed spacer marker gene primers for microbial community surveys. *Msystems* 1 (1), e00009–e00015. doi: 10.1128/mSystems.00009-15

Waltzek, T. B., Cortés-Hinojosa, G., Wellehan, J. J.F.X., and Gray, G. C. (2012). Marine mammal zoonoses: a review of disease manifestations. *Zoonoses Public Health* 59 (8), 521–535. doi: 10.1111/j.1863-2378.2012.01492.x

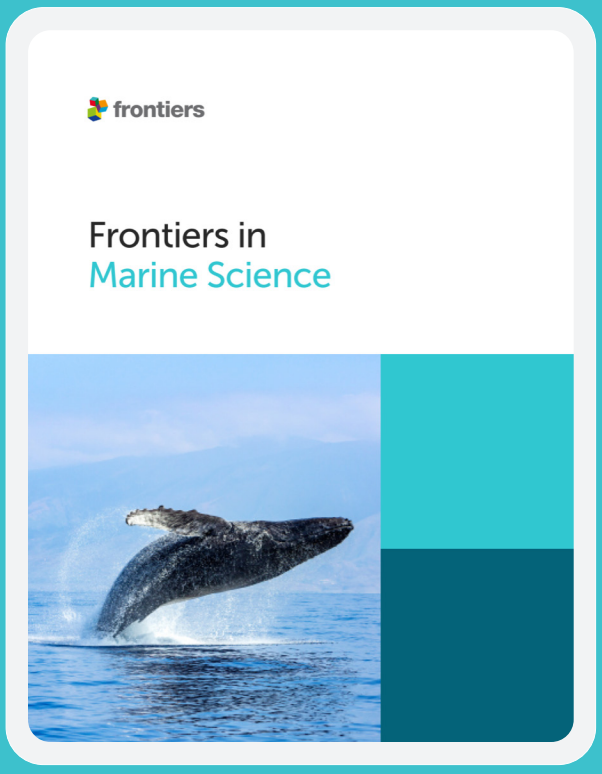
Frontiers in Marine Science

Explores ocean-based solutions for emerging global challenges

The third most-cited marine and freshwater biology journal, advancing our understanding of marine systems and addressing global challenges including overfishing, pollution, 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