Water Research X 20 (2023) 100194

Contents lists available at ScienceDirect

Water Research X

journal homepage: www.sciencedirect.com/journal/water-research-x

Aquaculture bacterial pathogen database: Pathogen monitoring and screening in coastal waters using environmental DNA

Linus Shing Him Lo [a,b], Xuan Liu [b,c], Hongbin Liu [b,c], Minhua Shao [c,d], Pei-Yuan Qian [b,c], Jinping Cheng [a,b,*]

[a] Department of Science and Environmental Studies and State Key Laboratory of Marine Pollution, The Education University of Hong Kong, New Territories, Hong Kong, China
[b] The Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou), Guangzhou 511458, China
[c] Department of Ocean Science, The Hong Kong University of Science and Technology, Hong Kong, China
[d] Department of Chemical and Biological Engineering and Energy Institute, The Hong Kong University of Science and Technology, Kowloon, Hong Kong, China

ARTICLE INFO

Keywords:
Bacterial pathogen database
Pathogen surveillance
Environmental DNA metabarcoding
Aquaculture
Coastal monitoring

ABSTRACT

Increasingly diverse pathogen occurrence in coastal and mariculture areas demands improved monitoring platforms to prevent economic and public health implications. Accessible databases with up-to-date knowledge and taxonomy are critical for detecting and screening environmental pathogens. Condensed from over 3000 relevant reports in peer reviewed articles, we constructed an aquaculture bacterial pathogen database that provides specialized curation of over 210 bacterial pathogenic species impacting aquaculture. Application of the aquaculture bacterial pathogen database to environmental DNA metabarcoding monitoring data in Hong Kong coastal and mariculture waters effectively characterized regional pathogen profiles over a one-year period and improved identification of new potential pathogen targets. The results highlighted the increase in potential pathogen abundance related to aquaculture activity and the associated inorganic nitrogen load, which was chiefly due to the enrichment of *Vibrio* during the atypical dry winter season. The value of the aquaculture bacterial pathogen database for empowering environmental DNA-based approaches in coastal marine pathogen surveillance benefits water resource management and aquaculture development on a global scale.

1. Introduction

Coastal marine waters are important development zones, habitats, and resources for food production. In 2018, global aquaculture production attained over 80 million metric tons with estimated farm sales of over US$250 billion (FAO, 2020). The intrinsic value of coastal waters demands sustainable utilization; however, coastal marine waters are also among the most vulnerable water bodies to elevated levels of contamination from anthropogenic activities (Adyasari et al., 2021). Compared to traditional chemical pollutants enrichment of primary or opportunistic pathogenic agents is a form of microbial pollution that leads to disease outbreaks in aquatic ecosystems and challenges the expansion of aquaculture production and management of water resource quality (Pandey et al., 2014). The annual economic loss associated with aquaculture diseases worldwide was estimated to be more than US$6 billion per year (World Bank, 2014). Opportunistic bacterial pathogens can contact and infect humans through exposure routes such as contaminated seafood and recreational water activities (Stec et al., 2022), with an annual economic cost of recreation-associated waterborne illness in the United States to be between US$2.2 – $3.7 billion (DeFlorio-Barker et al., 2018). Thus, conducting effective and efficient surveillance to identify and monitor pathogen occurrence and distribution in coastal waters is of critical importance to serve warning purposes for disease outbreak and food safety, with implications for water quality management plans and mediating public health risks.

The environmental DNA (eDNA) metabarcoding approach has gained traction in recent years as a promising biomonitoring tool that reinforces established cultivation-based species identification methods for applications including pathogen monitoring (Farrell et al., 2021). Sampling environmental DNA has advantages in broadscale coverage while being cost, time, and labor-effective (Ruppert et al., 2019), and capable of covering unculturable or foreign species. The effectiveness of such pathogen surveillance, however, relies heavily on updated knowledge and reference databases to identify target pathogens. While

* Corresponding author.
E-mail address: jincheng@eduhk.hk (J. Cheng).

https://doi.org/10.1016/j.wroa.2023.100194
Received 9 June 2023; Received in revised form 2 August 2023; Accepted 7 August 2023
Available online 7 August 2023

waterborne human pathogens have been frequently prioritized and studied with extensive knowledge, a comprehensive and up-to-date pathogen database providing a summary of the critically important bacterial pathogen species causing diseases in aquaculture organisms is largely lacking for addressing the risks to the environment and related industries. Existing databases are largely specialized in only bacterial pathogens of fish (e.g., Austin and Austin, 2016) or only provide genus level summary of the wider marine biota pathogens (e.g., Jurelevicius et al., 2021), despite their important impact on the aquatic food chain and the environment. Continuous efforts to develop applicable monitoring platforms and databases in light of fast advancing molecular discoveries are greatly beneficial for protecting water resources and the aquaculture industry.

This study aimed to develop and provide an updated database of potential aquaculture bacterial pathogens for effective pathogen monitoring targeted at all aquatic systems. Subsequently, we applied the database to a model urbanized coastal region to demonstrate the effective and sensitive detection of established pathogen groups in Hong Kong coastal waters. Within the study region, vibriosis and streptococcosis are common and well-recognized zoonotic diseases for fish and even humans and were key detection targets (Wang et al., 2020). The results obtained not only validated and supported eDNA as a globally applicable monitoring tool for pathogen screening, but also revealed spatiotemporally elevated pathogen occurrence and abundance in mariculture sites associated with seasons and human influence, forming a beneficial reference for preventing disease outbreaks in coastal aquaculture areas and maintaining a sustainable coastal ecosystem.

2. Results and discussion

2.1. Building the aquaculture bacterial pathogen database

The process of the literature research in constructing the aquaculture bacterial pathogen database (ABPD) is summarized in Fig. 1**a,** and currently catalogs more than 210 bacterial species from 65 unique genera and 47 families (Fig. 1**b**). Most potential aquaculture bacterial pathogen species belong to four of the six listed phyla (Proteobacteria, ~57%; Actinobacteria, ~14%; Bacteroidota, ~14%; Firmicutes, ~15%; Campylobacterota and Mycoplasmatota, < 1%).

For Proteobacteria, the class Gammaproteobacteria contributes approximately 96% of the proteobacterial entries and 55% of the total entries in ABPD. We curated 30 new potential bacterial pathogen species entries, and the ABPD now contains more than 120 listed potential Proteobacteria pathogen species. A notable proportion of them were established *Vibrio* pathogens of bivalves among recent novel fish bacterial pathogen discoveries in key pathogenic genera such as *Acinetobacter, Edwardsiella,* and *Pseudomonas* (**Supplementary Table 1** and references therein).

For Actinobacteria, we presented a total of 30 bacterial pathogen species among nine potential pathogen genera. Notably, the Mycobacteriaceae and Nocardiaceae species are known as key aquatic pathogens responsible for mycobacteriosis and nocardiosis, respectively. The database was expanded by the recent addition of *Mycobacterium paragondonae* by Machida et al. (2021) to the existing wide range of pathogenic species within *Mycobacterium*, as well as a new potential pathogenic genus *Kocuria* as opportunistic pathogens of brown trout (Pękala et al., 2018).

For Bacteroidota, *Flavobacterium, Tenacibaculum*, and *Chryseobacterium* received monitoring attention as the major pathogenic genera; they made up approximately 87% of the Bacteroidota entries. Accordingly, we reported six new potential fish pathogen species to the database, namely *Flavobacterium bernadetii, Tenacibaculum finnmarkense, Chryseobacterium aquaticum, C. cucumeris, C. gleum,* and *Chryseobacterium* sp. PLI_2, as well as *Tenacibaculum mesophilum*, recently reported to infect Akoya pearl oysters. With the recent report of *Ichthyobacterium seriolicida* proposed as a novel pathogen for yellowtail (Takano et al., 2016), the catalog for Bacteroidota pathogens have rapidly expanded to around 30 species over the past two years in comparison to the other major phyla with a historically larger pool of potential pathogens.

For Firmicutes, we curated a total of 32 bacterial pathogen species among a comparatively diverse range of 14 genera. The Firmicutes pathogens notably consisted of various *Bacillus* and streptococcosis-inducing bacteria (e.g., *Enterococcus, Lactococcus*, and *Streptococcus*). Compared to the many other standalone genus entries, the streptococcosis-inducing pathogens are relatively well evidenced and studied but have not seen substantial revision or addition of many new pathogenic species or infection reports. Finally, for Campylobacterota and Mycoplasmatota, the sole potentially pathogenic members are from genera *Arcobacter* and *Mycoplasma*, respectively. This database highlights the addition of 49 bacterial pathogen species for aquaculture animals, which is approximately 23% of the ABPD, or a 30% increase of initial entries from the fish-specific baseline work (Austin and Austin, 2016). The complete list of species in the ABPD and corresponding further information and references are shown in **Supplementary Table 1**.

2.2. eDNA monitoring of total bacterial community succession in coastal and fish farm waters

The composition and succession of the total bacterial community in Hong Kong coastal waters are summarized at the phylum and genus levels in Fig. 2. Coastal sampling sites were broadly classified into two categories, including fish farms (FF) and reference sites (RS) without any aquaculture activities, with subtle differences in physiochemical parameters (**Supplementary Table 2**). Based on the collected data on environmental conditions, fish farm samples in June had characteristically low total inorganic nitrogen (TIN) concentration measurements (**Supplementary Table 3 and 4**) that were comparable to the background concentrations recorded in the nearest local Environmental Protection Department of Hong Kong monitoring station in open waters (annual average value at 10 μg/L for eastern Hong Kong waters based on the Marine Water Quality Annual Report for 2019 – 2020), likely indicating minimal nutrient input from fish farming. These samples were acknowledged as potential inactive fish farms (IFF; n=7) as opposed to active fish farms (AFF; n=42) in downstream analysis.

At the phylum level, Cyanobacteria and Proteobacteria are the dominant bacteria groups that inhabit coastal waters, recorded with a mean relative abundance of 34.6% and 35.9%, respectively (Fig. 2**a, c**). At the genus level, the composition of the top 20 dominant genera with mean relative abundance > 1% is shown in Fig. 2**b** and **d**. Notable temporal patterns identified among FF samples in June revolved around the dominant *Synechococcus* bacteria group. The photosynthetic *Synechococcus* with the highest mean relative abundance at 30.3% showed significantly higher dominance in June (49.7%, on average) compared to the average site in any other months (28.1%, on average; Mann-Whitney U test, $P < 0.05$). Conversely, February marked a major reduction in *Synechococcus*'s relative abundance.

The results also identified two notable pathogenic genera. *Vibrio* had significantly higher relative abundance in AFF than IFF (Mann-Whitney U test, $P < 0.05$), while *Pseudoalteromonas* showed no significant difference in relative abundance among sampling sites. Temporal patterns in bacterial community composition highlighted the February fish farm samples, where *Vibrio* saw uniform enrichment in relative abundance (10.9%, on average) compared to all other months (0.24%, on average). A core group of recurrent *Vibrio* in active fish farms with presumable high feed and nutrient availability represents a potential threat to reared organisms and the surrounding environment in the studied region, in line with the context of common aquaculture diseases in Southeast Asia (Möller et al., 2020; Zhu et al., 2020). According to the database, these *Vibrio* encompass species such as *Vibrio mediterranei* and *Vibrio fortis*, which can be trout pathogens among other unclassified *Vibrio* bacteria.

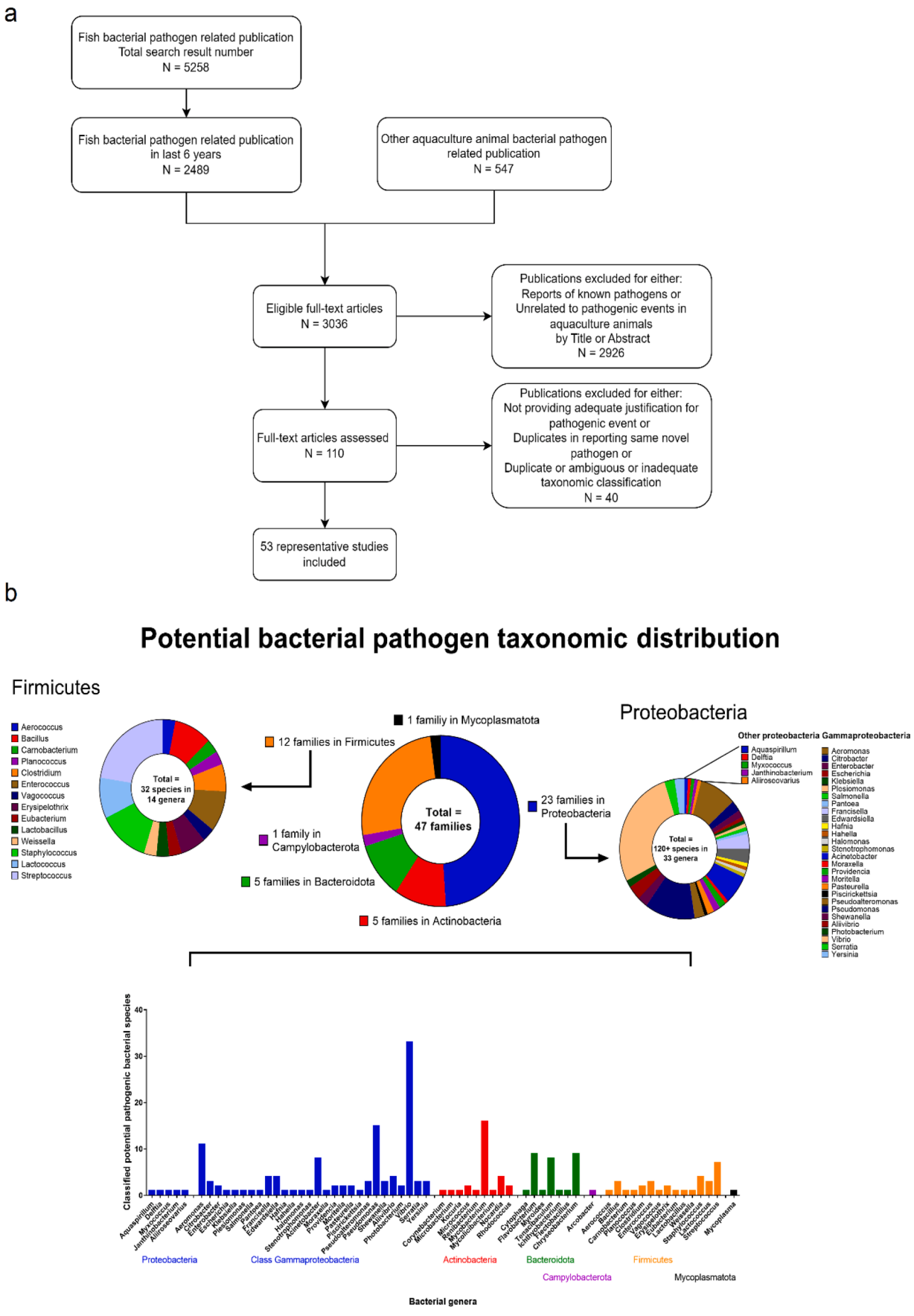

Fig. 1. Summary of (a) the literature survey and construction process and (b) the taxonomy represented in the ABPD.

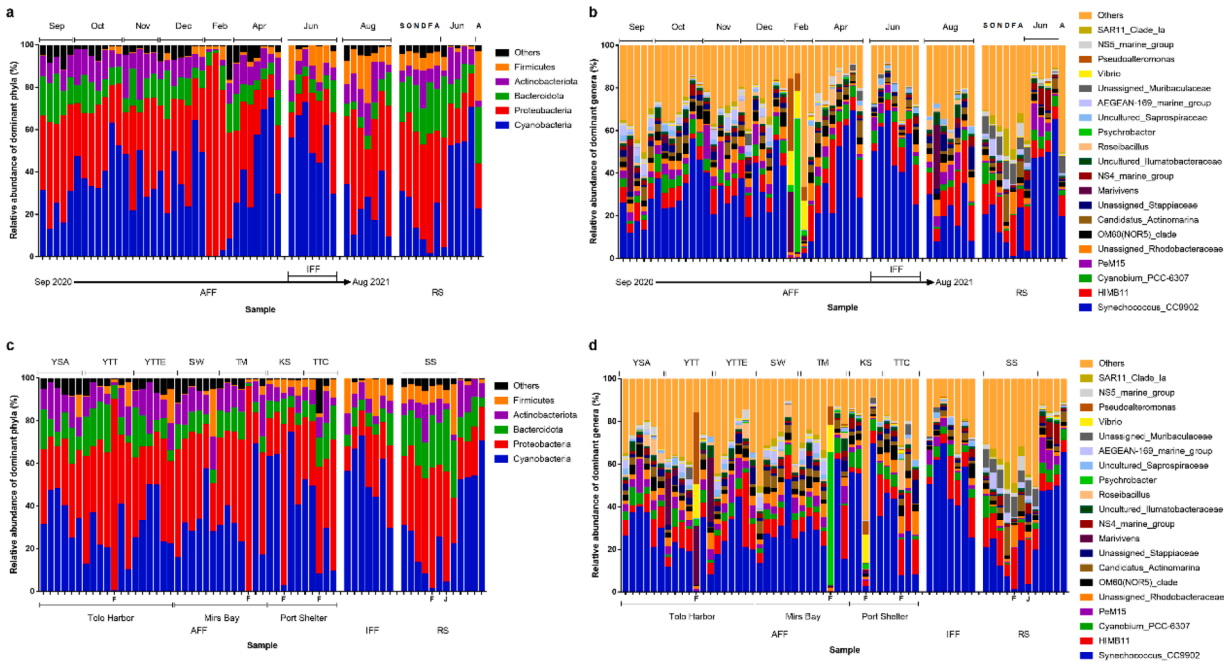

Fig. 2. Temporal succession and spatial variation in dominant bacterial community compositions in Hong Kong coastal waters. (a-b) Succession of bacterial community composition at phylum and genus levels, respectively. (c-d) Spatial variation in bacterial community succession at phylum and genus levels, respectively. For (a-b) and (c-d), grouped samples were sub-ordered by spatial distribution and chronological sampling time, respectively. AFF: active fish farm; IFF: inactive fish farm; RS: reference site. Letter codes indicate the sampling month. S: September; O: October; N: November; D: December; F: February; A: April, August; J: June.

Compared to more routinely screened pathogenic *Vibrio* species in Hong Kong waters, such as *V. alginolyticus, V. parahaemolyticus, V. vulnificus,* and *V. fluvialias*, the dispersal of these less noticed species along the water column may represent a currently overlooked risk. *Vibrio* are ubiquitous bacteria typically known to have enhanced growth with increased temperature and nutrient availability (Stabili et al., 2022). Thus, historical aquaculture disease outbreaks such as vibriosis were often seasonal in the warmer wet season (Yang et al., 2021), while here the results show that dry seasons in particular pose elevated risks for these opportunistic pathogens to bloom in abundance. The relatively rare case of higher relative abundances of *Vibrio* during winter season has been previously reported by Kolda et al. (2020), in which the detection in winter was associated with a time lag leading to a vibriosis disease outbreak that subsequently occurred in spring. The ability for *Vibrio* to bloom into abundance under previously considered 'unfavorable' conditions is their potential resilience which has not been frequently reported and thus represents an ecological risk to coastal waters. This also demonstrates that environmental surveillance for high *Vibrio* abundance could serve as an early warning signal. It is possible that prolonged sunlight duration in the sub-tropical summer of Hong Kong mediated the inactivation of Vibrio microorganisms (Busse et al., 2019). Another explanation for common cases where high *Vibrio* abundance, especially during summer, did not cause disease outbreaks is the co-occurrence of probiotic bacteria such as *Bacillus*, which can regulate nitrogen levels and reduce the abundance of potential pathogens (Dawood and Koshio, 2016). This, however, was not observed in the present study. Instead, *Pseudoalteromonas* was found to frequently co-occur with *Vibrio* during its bloom, which may be indicative of competition as species such as *Pseudoalteromonas piscicida* has *Vibrio*--killing effects (Richards et al., 2017). However, this may also reflect a case of co-existence, as their abundance was observed to increase uniformly during winter. Such co-occurrence has been previously associated with diseased shrimp feces reared from October to November of 2016 (Alfiansah et al., 2020) and could serve as a warning sign.

2.3. *Database-informed screening of aquaculture bacterial pathogens in Hong Kong coastal waters*

Coastal bacterial communities were then screened for potential pathogens using the ABPD. The phylogenetic distribution of major potential aquaculture bacterial pathogens and their relative abundance in collected samples are highlighted in Fig. 3**a**. A summary of the normalized read abundance of potential pathogenic genera is shown in **Supplementary Table 5 and** Fig. 3**b**. In total, we detected over 84,000 amplicon sequence variants (ASVs) from 37 potential pathogen genera in Hong Kong coastal waters, which is approximately 3.5% of the total reads. We identified potential pathogenic ASVs from 19 Proteobacteria genera, 4 Actinobacteria genera, 2 Bacteroidota genera, 14 Firmicutes, and 1 Campylobacterota genus. No Mycoplasmatota was identified. This coverage showed that approximately 57% of the genus entries in the ABPD were detected in Hong Kong coastal waters. However, in terms of abundance, the majority of all detected potential pathogenic ASVs belonged to Proteobacteria (80.2%) followed by Firmicutes (18.2%), requiring further specialization. Other potential pathogens identified in Hong Kong coastal zones were generally rare taxa (<1% abundance) that appear to be more susceptible to environmental changes and present lower pathogenic risks. Only 1.53%, 0.09%, and 0.05% of ASVs belonged to Bacteroidota, Actinobacteria, and Campylobacterota, respectively. The most dominant families were Vibrionaceae and Pseudoalteromonadaceae followed by Enterobacteriaceae under Proteobacteria. *Vibrio* and *Photobacterium* were potential pathogen genera of Vibrionaceae family which had significantly higher abundance observed in AFF sites (Kruskal-Wallis test, $P < 0.002$). Similarly, *Pseudoalteromonas* also had significantly higher abundance in AFF sites (Kruskal-Wallis test, P=0.024). For Enterobacteriaceae pathogens, *Escherichia-Shigella* was only observed to have higher abundance in IFF, while *Klebsiella* and *Enterobacter* were both detected in low abundance at all sites. *Lactobacillus* was the only genus occurring in high abundance in RS. The general lull in pathogen occurrence among samples from inactive fish farms in June may suggest pathogens in this study to be mostly host-associated, such that the removal of fish decreased pathogen

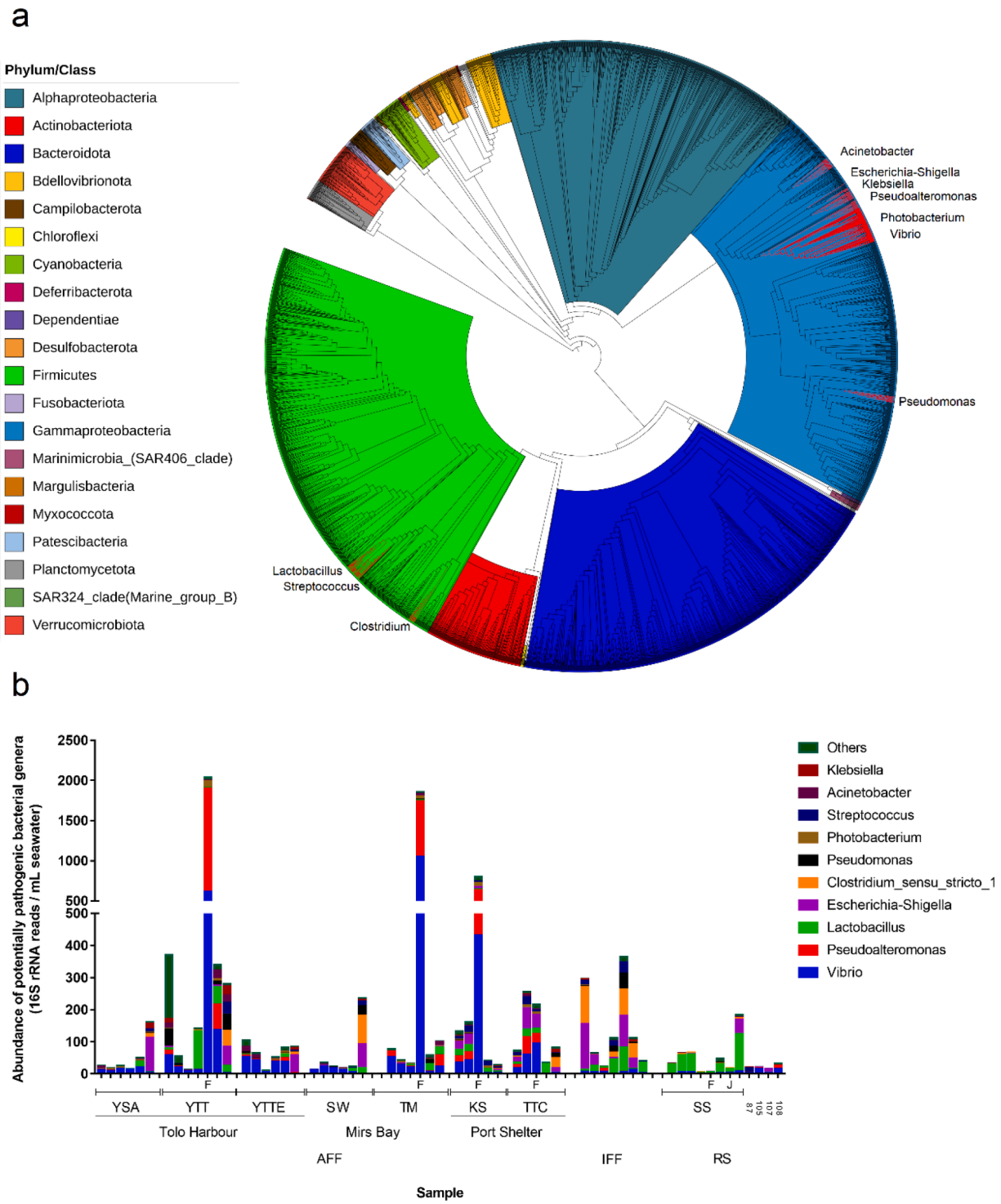

Fig. 3. (a) Phylogenetic heatmap distribution of potential pathogenic genera in Hong Kong coastal waters. Phylogenetic branches of major potential pathogenic genera were highlighted in red. (b) Spatial variation in potential pathogen normalized read abundance. Letter codes indicate key sampling months. F: February; J: June.

prevalence.

2.4. *Environmental drivers of potential pathogenic bacteria enrichment in coastal aquaculture sites*

The possible enrichment of screened potential bacterial pathogens among Hong Kong coastal waters was explored with available environmental parameters. Spearman's correlation analyzed the relationship between normalized ASV abundances for the top 10 dominant pathogen genera and environmental parameters. The TIN concentration displayed a moderate positive correlation with pathogen ASV abundance, namely *Vibrio* and *Pseudoalteromonas* (Fig. 4; Spearman's $r > 0.3$, $P < 0.05$) after accounting for potential outliers. Other recorded environmental parameters such as temperature displayed no significant correlation with pathogen ASV abundances.

The distribution and potential enrichment of pathogenic bacteria

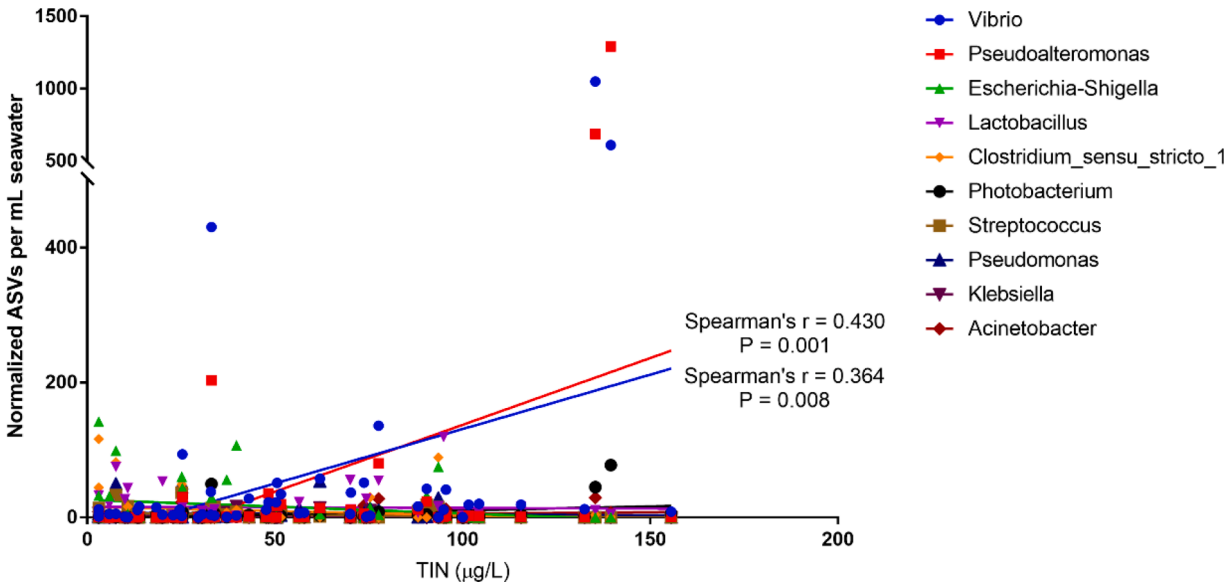

Fig. 4. Spearman's correlation between total inorganic nitrogen and normalized potential bacterial pathogen ASVs in Hong Kong coastal waters.

inhabiting Hong Kong coastal waters were further visualized in respective sampling months (Fig. 5). Among all identified potential pathogenic bacterial genera, *Vibrio* and *Pseudoalteromonas* were the most dominant with a mean relative abundance of 1.10% and 1.07% among FF samples, respectively (in RS, 0.27% and 0.07%, respectively). *Vibrio* and *Pseudoalteromonas* had peak relative abundances observed at 15.9% and 33.7%, respectively, during February in the YTT sample site of the Tolo Harbour, indicating the potential for such pathogenic genera to bloom under favorable conditions. *Escherichia-Shigella* hybrids were further detected in most Hong Kong fish farms during the wet seasons between May to August, regardless of the presence of ongoing fish farm activities. Their relative abundance in Tolo Harbour sampling sites (YSA, YTT, YSA) tended to surpass 2% during May to August. Warmer conditions may have uniformly favored these potential pathogens in a managed aquaculture site setting. These results showed that *Escherichia-Shigella* mainly occurred in the Tolo Harbour to Mirs Bay Channel with rarer occurrence in the Port Shelter sampling sites (KS and TTC; <1%), which may be attributed to limited dispersal from the Tolo Harbour semi-enclosed bay.

2.5. Quantification of bacterial and pathogen abundance

The absolute abundance of total bacteria, potential pathogenic genera, and *Vibrio* was quantified through qPCR assays (Fig. 6). Results first identified a significant difference in mean absolute abundance between FF and RS for total bacteria over the study duration. The temporal influential factor identified from the previous analysis was used to further group samples into the dry or wet season. *Vibrio* was selected for genera-specific quantification analysis to investigate the abundance of potential key pathogens. A significant difference was observed between AFF sites sampled in dry and wet seasons (Mann Whitney U test, $P < 0.05$), with an enriched *Vibrio* occurrence during the dry season. This result again reveals a consistently higher *Vibrio* abundance in Hong Kong coastal waters notably in the colder winter season, which may not be expected as current knowledge tends to suggest that higher temperatures are more optimal conditions for *Vibrio* prevalence (Baker-Austin et al., 2013; Sheikh et al., 2022). According to the methods and guidelines documented in a screening-level risk assessment conducted for bath waters (Gyraite et al., 2020), the active aquaculture sites in Hong Kong could be considered at a low to intermediate integrated risk as contributed by *Vibrio* and *Enterococcus* fecal pathogen abundance. However, the cultivation-based risk assessment guidelines may underestimate environmental pathogens in viable but non-culturable (VBNC) states; alternatively, the current enumeration process was conservative in assuming detected pathogens to have the mean 16S rRNA gene copy.

3. Conclusion

The ecological and economic importance of coastal areas and aquaculture zones on a global scale demands a better understanding of pathogenic microbial components in the environment to properly inform seafood safety, public health, water quality and management concerns. We developed an updated ABPD to address the dire need for accessible information summaries on current knowledge of the major disease-causing aquatic bacteria. This study demonstrated that an environmental DNA metabarcoding approach employing a coastal model with the established database displayed excellent potential for diverse coastal pathogen monitoring and revealed the spatiotemporal prevalence mechanisms associated with intense aquaculture. This study provides a much-needed, globally applicable database, demonstrating its systematic application and presenting its advantages for better coastal monitoring and pathogen risk surveillance. This has implications for water quality management on a global scale.

4. Materials and methods

4.1. Establishing the aquaculture bacterial pathogen database

The construction of the database of potential aquaculture bacterial pathogens began with desktop research of the literature, including reviews. At the time of the search phase (August 2022), notable work by Austin and Austin (2016) provided a representative list of bacterial fish pathogens, but no substantial update had been made since its publication. In the present study, we extended the scope to both update and further include pathogenic bacterial taxa relevant to a broader range of aquacultural animals. The updated database, referred to as the Aquaculture Bacterial Pathogen Database (ABPD), was constructed as follows:

A literature survey on publications describing potential bacterial pathogens and diseases associated with aquaculture animals (including fish and shellfish) was conducted on Web of Science using a combination of search keywords, including but not limited to the following: "bacterial pathogens", "diseases", "infection", "aquaculture", "fish", and

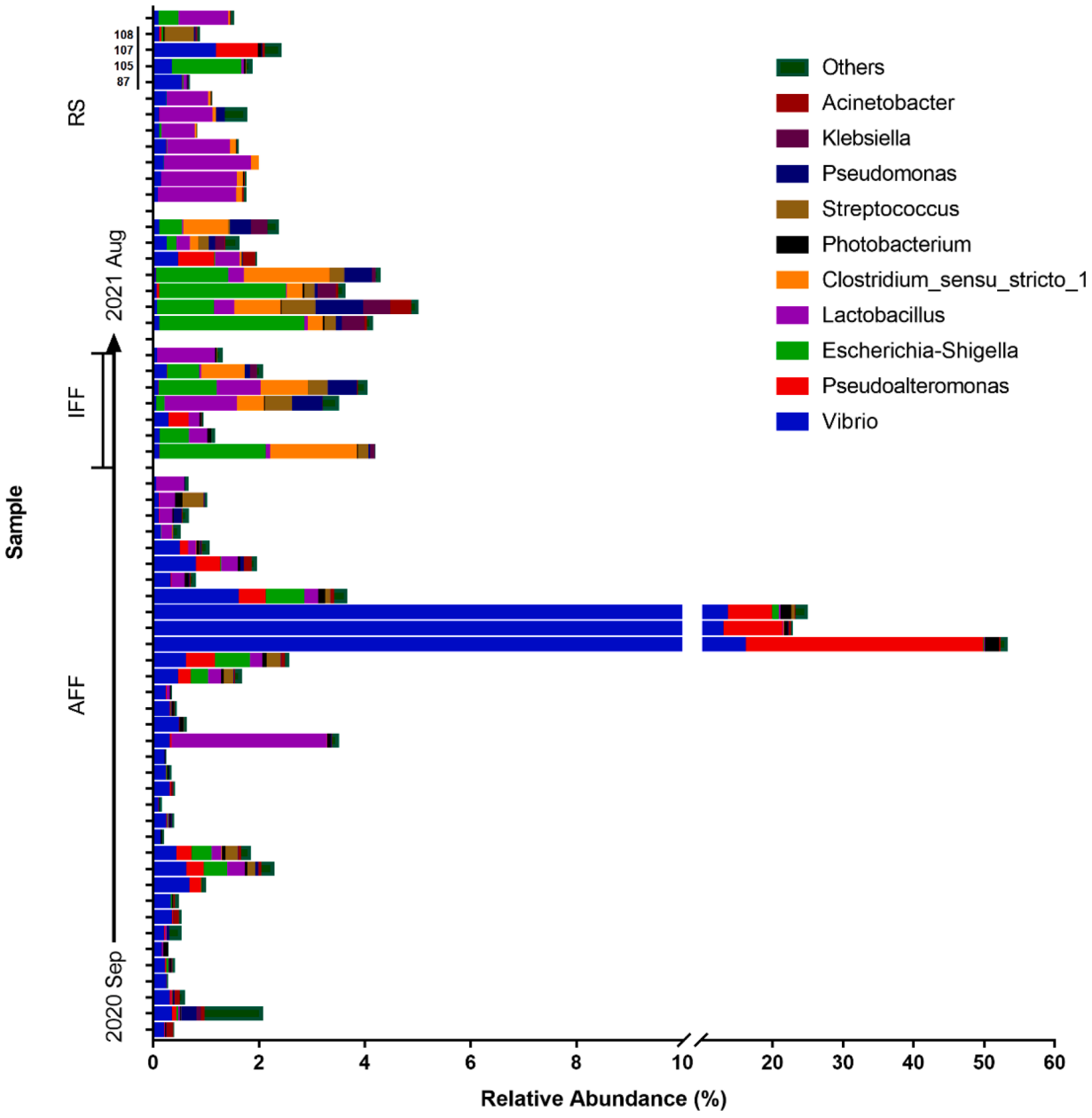

Fig. 5. Summary of temporal succession in relative abundance of potential pathogenic bacterial genera in Hong Kong coastal waters. AFF: active fish farm; IFF: inactive fish farm; RS: reference site.

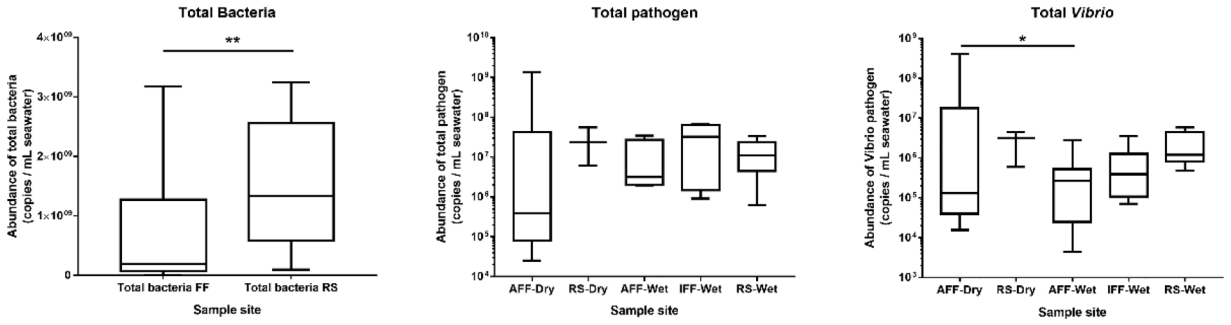

Fig. 6. Box and whisker plot summary of absolute abundance of total bacteria, total pathogen, and total *Vibrio* between different coastal sites of various fish farming activity and precipitation seasons. Asterisks indicate statistically significant difference (Mann-Whitney U test or Kruskal-Wallis test; * $P < 0.033$, ** $P < 0.002$). AFF: active fish farm; IFF: inactive fish farm; RS: reference site; Dry: sampled during November to February; Wet: sampled during May to August.

"shellfish". We first separated the search into two parts concerning different major aquaculture animal hosts. For fish pathogens, the search was shortlisted to papers published after the year 2015, as Austin and Austin (2016) have comprehensively summarized the documentations of fish bacterial pathogens prior to 2016. Among the search results, publications that provided novel updates to either reported hosts and/or associated symptoms and diseases of established bacterial pathogens, or report new bacterial pathogens associated with a disease or clinical symptoms in fish hosts, were selectively included in the ABPD based on their novelty. Publications that explored novel bacterial pathogens at a species level (either new infection reports of established species or novel bacterial species isolated from diseased fish) or novel host groups (e.g., new freshwater fish hosts among reported marine fish hosts) were prioritized over reports of new diseases and symptoms due to the scope of this database. Pathogen entries from publications in which the authors suggest major uncertainty or provide inadequate evidence supporting pathogenicity were carefully examined before exclusion.

In the second part, i.e., the search for publications on pathogens for aquaculture animals other than fish (such as oysters and shrimp), the same shortlisting and selection logic was applied, except no time range was imposed during the search. However, we selected and condensed publications primarily from the last decade. Where appropriate; key reports for diseases that occurred further in the past were included. A total of 53 representative publications were considered for either adding new pathogen information to the baseline or providing revision or supplementary information to described bacterial pathogens. The corresponding taxonomy and nomenclature status for bacterial pathogen entries were verified and revised for any existing synonyms, corrections, or reassignments with reference to the LPSN database (Parte et al., 2020) before inclusion into the ABPD. The formated table of this database (**Supplementary Table 1**) details the family, genus (and species level taxonomy, where available) of potential aquaculture bacterial pathogens, as well as examples of infection targets and illnesses or clinical symptoms.

4.2. Site description and eDNA sample collection

This study examined surface water samples collected from a total of 12 coastal sites to the east of Hong Kong, including seven fish farms and five open water reference sites, between September 2020 and August 2021, as shown in **Supplementary Figure 1**. Water samples from all seven fish farms (n=49) and the Silverstrand reference site (SS; n=8) were collected at monthly to bimonthly intervals, depending on accessibility (n=57). The seven selected local fish farm sites were located within dedicated fish culture zones (FCZs); they were Yim Tin Tsai (YTT), Yim Tin Tsai East (YTTE), Yung Shue Au (YSA), Tap Mun (TM), Sham Wan (SW), Kau Sai (KS), and Tai Tau Chau (TTC). An additional four open water reference sites, numbered as stations 87, 105, 107, and 108, were sampled during June 2021 (n=4) to complement the study. A total of 61 samples were collected and used in this study. Detailed sample information can be found in **Supplementary Table 4 and 6**.

At each sampling site, 1 L of surface seawater sample was collected at 1 m depth. In fish farms, seawater samples were collected from waters in the proximity of fish cages. Seawater samples were stored in a sterile bottle at low temperature on ice in a cooler box and transported back to the laboratory to be processed within six hours of collection to minimize degradation. Individual seawater samples (in four aliquots of 250mL each) were filtered through four 0.22μm pore size polycarbonate membranes (Millipore Corporation, USA) using a vacuum pump. The filter membranes were then labeled and stored at −80°C until subsequent DNA extraction.

4.3. Environmental parameters of coastal seawater

At each sampling site, hydrological parameters of seawater such as temperature, dissolved oxygen (DO), and pH were measured in situ using an EXO2 multiparameter water quality sonde (YSI, USA). Other parameters, such as chlorophyll-a level, were measured using Trilogy Laboratory Fluorometer (Turner Designs, USA); total inorganic nitrogen (TIN) was measured using a San++ Continuous Flow Analyzer (Skalar, Netherlands). Environmental parameters from government monitoring stations in the proximity of each sampling site away from fish farm operations, set up by the Environmental Protection Department (EPD) of the Hong Kong SAR Government (https://cd.epic.epd.gov.hk/EPICRIVER/marine/), were used as a reference for background water conditions near fish farms and the reference site in the design of the study. We employed these available data with a query to the database for marine water records of all sampling stations within the Water Control Zones Port Shelter, Tolo Harbour and Channel, and Mirs Bay, which date between 2019/01/01 to 2021/12/31. The recorded parameters available were 5-day biochemical oxygen demand, ammonia, chlorophyll-a, dissolved oxygen, *E. coli*, fecal coliform, nitrate, nitrite, pH, salinity, Secchi disk depth, suspended solids, temperature, total inorganic nitrogen, total nitrogen, total phosphorus, and turbidity, and are accessible online through https://cd.epic.epd.gov.hk/EPICRIVER/marine.

4.4. DNA extraction, amplification, and 16S rRNA gene sequencing

Total DNA extraction from stored filter membranes was carried out following a modified SDS-based method described by Zhang et al. (2008). The quality and concentration of extracted DNA were verified using a BioDrop μLITE spectrophotometer (Biochrom, United Kingdom). The V3 - V4 variable region of the 16S rRNA gene was amplified using the 341F/806R primer pair (Lu et al., 2015; **Supplementary Table 7**) following the methods detailed by Xu et al. (2020). Sequencing was performed by Novogene (Beijing, China) on the Illumina NovaSeq platform (Illumina, USA) using paired-end sequencing. The details of the DNA extraction and sequencing process are detailed in the Supplementary Materials. The sequence reads generated were deposited in the National Center for Biotechnology Information (NCBI) Sequence Read Archive under the accession numbers (BioProject PRJNA1003239).

4.5. Bioinformatics and screening of aquaculture bacterial pathogens

The sequencing data was processed and analyzed using QIIME2 (version 2021.11; Bolyen et al., 2019). Sequences were demultiplexed, and reads with low (< 25) PHRED scores, short reads and ambiguous bases were removed. Quality filtering and denoising process were conducted following the Deblur (Amir et al., 2017) pipeline integrated in QIIME2 using default parameters. Sequences were trimmed to 400bp, dereplicated, and with chimeric reads removed before subsequent generation of the feature table. The amplicon sequence variants (ASVs) were assigned to taxonomy using the SILVA database (release 138; Quast et al., 2012), and the feature-classifier QIIME2 plugin at a confidence threshold of 0.7. Sequences that could not be assigned beyond the domain level or were classified as chloroplast and mitochondria were removed from downstream analysis.

The occurrence of aquaculture bacterial pathogens was analyzed by screening the taxonomically assigned ASVs using the ABPD at the genus level. ASVs that could not be classified beyond the family level or were not taxonomically classified to a genus listed in the ABPD, were subsequently omitted before subjected to downstream analysis. Due to the limitation in current taxonomic resolution, the species level annotation was only interpreted where available and not considered in the quantification study, i.e., all ASVs assigned to genera identified by the ABPD were considered potentially pathogenic. This provides a more aggressive estimate of pathogen prevalence in downstream analysis, aiming to serve as a preliminary warning in pathogen screening.

4.6. Quantification of pathogen abundance

The absolute abundance of total bacteria and the dominant pathogen

genus *Vibrio* were quantified using qPCR as described by Xu et al. (2020). *Vibrio* is considered and reported as an established pathogen group of concern in the study region according to the Agriculture, Fisheries and Conservation Department of Hong Kong SAR Government, and thus was selected for further quantification analysis as the representative pathogen group. All primers used are summarized in **Supplementary Table 3**. In particular, primers 567F and 680R targeting the *Vibrio* genus 16S rRNA gene (Thompson et al., 2004; Vezzulli et al., 2012), and primers 967F and 1046R (Sogin et al., 2006; Vezzulli et al., 2012) specific to the Bacteria domain were used to quantify the total *Vibrio* and bacteria, respectively. PCR standard curves were prepared from genomic DNA extracted from lab cultures of *Escherichia coli* and *Vibrio parahaemolyticus*, conducted on LightCycler 480II Real-Time PCR system (Roche Life Science, Switzerland). The limit of detection for bacteria and Vibrio were estimated to be 0.094 and 6.936ng/mL. LightCycler 480 SYBR Green I Master (Roche Life Science, Switzerland) was used as the reaction mix, and the reaction mixture and cycling conditions were modified from Liang et al. (2019). In brief, each 15-μl reaction mixture contained 0.4μM forward and reverse primer and 0.6 μL DNA template. The qPCR cycling conditions were as follows: an initial denaturation at 95 °C for 5min; followed by 35 cycles of the three-step reaction of: denaturation at 95 °C for 15s, annealing at 55 °C for 30s, and extension at 72 °C for 30s. All qPCR amplifications were performed in triplicate.

The inference and quantification of pathogen taxa abundance from qPCR was conducted using the method described by Jian et al. (2020) as follows:

$$IA_P = RA_P \times AA_i \quad (1)$$

where IA_P is the inferred abundance of bacterial pathogen taxon *P*, RA_P is the relative abundance of pathogen taxon *P* obtained based on the 16S rRNA gene sequencing reads, and AA_i is the absolute abundance of bacteria in sample *i*. Gene copy normalization was conducted based on the mean 16S rRNA gene copy count according to the rrnDB database (Stoddard et al., 2015). Taxa without a direct match were assumed and assigned to have a gene copy number of 1.

4.7. Statistical analysis

Differences in physicochemical properties and concentrations were analyzed using independent-samples *t*-test. The normality of data was tested with the Shapiro-Wilk normality test. Differences in microbial relative abundance were analyzed using the Mann-Whitney U test or Kruskal-Wallis test for comparison between two groups or among three or more groups, respectively. Spearman's correlation was used to investigate the relationship between environmental parameters and bacterial pathogen abundance. Prism 9 (GraphPad, USA) software and R (R Core Team, 2020) were used for statistical analyses and data visualization. Differences were considered significant when $P < 0.05$.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request

Acknowledgments

This research was supported by a General Research Fund of the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. 16101821) and The Education University of Hong Kong.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.wroa.2023.100194.

References

Adyasari, D., Pratama, M.A., Teguh, N.A., Sabdaningsih, A., Kusumaningtyas, M.A., Dimova, N., 2021. Anthropogenic impact on Indonesian coastal water and ecosystems: current status and future opportunities. Mar. Pollut. Bull. 171, 112689.

Alfiansah, Y.R., Peters, S., Harder, J., Hassenrück, C., Gärdes, A., 2020. Structure and co-occurrence patterns of bacterial communities associated with white faeces disease outbreaks in Pacific white-leg shrimp Penaeus Vannamei aquaculture. Sci. Rep. 10 (1), 11980.

Amir, A., McDonald, D., Navas-Molina, J.A., Kopylova, E., Morton, J.T., Zech Xu, Z., Knight, R, 2017. Deblur rapidly resolves single-nucleotide community sequence patterns. MSystems 2 (2), e00191. -16.

Austin, B., Austin, D.A., 2016. Bacterial Fish Pathogens: Disease of Farmed and Wild Fish. Springer, Dordrecht, The Netherlands.

Baker-Austin, C., Trinanes, J.A., Taylor, N.G., Hartnell, R., Siitonen, A., Martinez-Urtaza, J., 2013. Emerging Vibrio risk at high latitudes in response to ocean warming. Nat. Clim. Chang. 3 (1), 73–77.

Bolyen, E., Rideout, J.R., Dillon, M.R., Bokulich, N.A., Abnet, C.C., Al-Ghalith, G.A., Caporaso, J.G., 2019. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat. Biotechnol. 37 (8), 852–857.

Busse, M.M., Becker, M., Applegate, B.M., Camp Jr, J.W., Blatchley III, E.R, 2019. Responses of Salmonella typhimurium LT2, Vibrio harveyi, and Cryptosporidium parvum to UVB and UVA radiation. Chem. Eng. J. 371, 647–656.

Dawood, M.A., Koshio, S., 2016. Recent advances in the role of probiotics and prebiotics in carp aquaculture: a review. Aquaculture 454, 243–251.

DeFlorio-Barker, S., Wing, C., Jones, R.M., Dorevitch, S., 2018. Estimate of incidence and cost of recreational waterborne illness on United States surface waters. Environ. Health 17 (1), 1–10.

FAO (2020). The State of World Fisheries and Aquaculture 2020. Sustainability in Action. Rome.

Farrell, J.A., Whitmore, L., Duffy, D.J., 2021. The promise and pitfalls of environmental DNA and RNA approaches for the monitoring of human and animal pathogens from aquatic sources. Bioscience 71 (6), 609–625.

Gyraite, G., Kataržytė, M., Overlingė, D., Vaičiūtė, D., Jonikaitė, E., Schernewski, G., 2020. Skip the dip—Avoid the risk? Integrated microbiological water quality assessment in the South-Eastern Baltic Sea coastal waters. Water (Basel) 12 (11), 3146.

Jian, C., Luukkonen, P., Yki-Järvinen, H., Salonen, A., Korpela, K., 2020. Quantitative PCR provides a simple and accessible method for quantitative microbiota profiling. PLoS One 15 (1), e0227285.

Jurelevicius, D., Cotta, S.R., Montezzi, L.F., Dias, A.C., Mason, O.U., Picão, R.C., Seldin, L., 2021. Enrichment of potential pathogens in marine microbiomes with different degrees of anthropogenic activity. Environ. Pollut. 268, 115757.

Kolda, A., Gavrilović, A., Jug-Dujaković, J., Ljubešić, Z., El-Matbouli, M., Lillehaug, A., Kapetanović, D., 2020. Profiling of bacterial assemblages in the marine cage farm environment, with implications on fish, human and ecosystem health. Ecol. Indic. 118, 106785.

Liang, J., Liu, J., Wang, X., Lin, H., Liu, J., Zhou, S., Zhang, X.H., 2019. Spatiotemporal dynamics of free-living and particle-associated Vibrio communities in the northern Chinese marginal seas. Appl. Environ. Microbiol. 85 (9), e00217–e00219.

Lu, Y.Z., Ding, Z.W., Ding, J., Fu, L., Zeng, R.J., 2015. Design and evaluation of universal 16S rRNA gene primers for high-throughput sequencing to simultaneously detect DAMO microbes and anammox bacteria. Water Res. 87, 385–394.

Machida, Y., Tang, B.C.C., Yamada, M., Sato, S., Nakajima, K., Matoyama, H., Kato, G., 2021. Mycobacteriosis in cultured koi carp Cyprinus carpio caused by Mycobacterium paragordonae and two Mycolicibacterium spp. Aquaculture 539, 736656.

Möller, L., Kreikemeyer, B., Luo, Z.H., Jost, G., Labrenz, M., 2020. Impact of coastal aquaculture operation systems in Hainan island (China) on the relative abundance and community structure of Vibrio in adjacent coastal systems. Estuar. Coast. Shelf Sci. 233, 106542.

Pandey, P.K., Kass, P.H., Soupir, M.L., Biswas, S., Singh, V.P., 2014. Contamination of water resources by pathogenic bacteria. AMB Expr. 4 (1), 1–16.

Parte, A.C., Sardà Carbasse, J., Meier-Kolthoff, J.P., Reimer, L.C., Göker, M., 2020. List of prokaryotic names with standing in nomenclature (LPSN) moves to the DSMZ. Int. J. Syst. Evol. Microbiol. 70, 5607–5612.

Pękala, A., Paździor, E., Antychowicz, J., Bernad, A., Głowacka, H., Więcek, B., Niemczuk, W., 2018. Kocuria rhizophila and Micrococcus luteus as emerging opportunist pathogens in brown trout (Salmo trutta Linnaeus, 1758) and rainbow trout (Oncorhynchus mykiss Walbaum, 1792). Aquaculture 486, 285–289.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Glöckner, F.O., 2012. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic. Acids. Res. 41 (D1), D590–D596.

R Core Team, 2020. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Richards, G.P., Watson, M.A., Needleman, D.S., Uknalis, J., Boyd, E.F., Fay, J.P., 2017. Mechanisms for Pseudoalteromonas piscicida-induced killing of vibrios and other bacterial pathogens. Appl. Environ. Microbiol. 83 (11), e00175. -17.

Ruppert, K.M., Kline, R.J., Rahman, M.S., 2019. Past, present, and future perspectives of environmental DNA (eDNA) metabarcoding: a systematic review in methods, monitoring, and applications of global eDNA. Glob. Ecol. Conserv. 17, e00547.

Sheikh, H.I., Najiah, M., Fadhlina, A., Laith, A.A., Nor, M.M., Jalal, K.C.A., Kasan, N.A., 2022. Temperature upshift mostly but not always enhances the growth of Vibrio species: a systematic review. Front. Mar. Sci. 9, 959830.

Sogin, M.L., Morrison, H.G., Huber, J.A., Welch, D.M., Huse, S.M., Neal, P.R., Herndl, G. J., 2006. Microbial diversity in the deep sea and the underexplored "rare biosphere. Proc. Natl Acad. Sci. 103 (32), 12115–12120.

Stabili, L., Di Salvo, M., Alifano, P., Talà, A., 2022. An integrative, multiparametric approach for the comprehensive assessment of microbial quality and pollution in aquaculture systems. Microb. Ecol. 83 (2), 271–283.

Stec, J., Kosikowska, U., Mendrycka, M., Stępień-Pyśniak, D., Niedźwiedzka-Rystwej, P., Bębnowska, D., Grywalska, E., 2022. Opportunistic Pathogens of Recreational Waters with Emphasis on Antimicrobial Resistance-A Possible Subject of Human Health Concern. Int. J. Environ. Res. Public Health 19 (12), 7308.

Stoddard, S.F., Smith, B.J., Hein, R., Roller, B.R., Schmidt, T.M., 2015. rrn DB: improved tools for interpreting rRNA gene abundance in bacteria and archaea and a new foundation for future development. Nucleic. Acids. Res. 43 (D1), D593–D598.

Takano, T., Matsuyama, T., Sakai, T., Nakamura, Y., Kamaishi, T., Nakayasu, C., Iida, T., 2016. Ichthyobacterium seriolicida gen. nov., sp. nov., a member of the phylum 'Bacteroidetes', isolated from yellowtail fish (Seriola quinqueradiata) affected by bacterial haemolytic jaundice, and proposal of a new family, Ichthyobacteriaceae fam. nov. Int. J. Syst. Evol. Microbiol. 66 (2), 580–586.

Thompson, J.R., Randa, M.A., Marcelino, L.A., Tomita-Mitchell, A., Lim, E., Polz, M.F., 2004. Diversity and dynamics of a North Atlantic coastal *Vibrio* community. Appl. Environ. Microbiol. 70 (7), 4103–4110.

Vezzulli, L., Brettar, I., Pezzati, E., Reid, P.C., Colwell, R.R., Höfle, M.G., Pruzzo, C., 2012. Long-term effects of ocean warming on the prokaryotic community: evidence from the *Vibrios*. ISME J. 6 (1), 21–30.

Wang, Q., Ji, W., Xu, Z., 2020. Current use and development of fish vaccines in China. Fish Shellfish Immunol. 96, 223–234.

World Bank Group, 2014. Reducing Disease Risks in Aquaculture. World Bank Group, Washington, DC. World Bank Report Number 88257-GLB.

Xu, W., Gong, L., Yang, S., Gao, Y., Ma, X., Xu, L., Luo, Z., 2020. Spatiotemporal dynamics of *Vibrio* communities and abundance in Dongshan Bay, South of China. Front. Microbiol. 11, 575287.

Yang, B., Zhai, S., Li, X., Tian, J., Li, Q., Shan, H., Liu, S., 2021. Identification of Vibrio alginolyticus as a causative pathogen associated with mass summer mortality of the Pacific Oyster (Crassostrea gigas) in China. Aquaculture 535, 736363.

Zhang, R., Thiyagarajan, V., Qian, P.Y., 2008. Evaluation of terminal-restriction fragment length polymorphism analysis in contrasting marine environments. FEMS Microbiol. Ecol. 65 (1), 169–178.

Zhu, Z., Duan, C., Dong, C., Weng, S., He, J., 2020. Epidemiological situation and phylogenetic relationship of Vibrio harveyi in marine-cultured fishes in China and Southeast Asia. Aquaculture 529, 735652.