Environmental Research 190 (2020) 109982

Contents lists available at ScienceDirect

Environmental Research

journal homepage: www.elsevier.com/locate/envres

High-throughput diagnosis of human pathogens and fecal contamination in marine recreational water

Xin-Li An [a], Jia-Ying Wang [a,b], Qiang Pu [a,b], Hu Li [a], Ting Pan [a,b], Huan-Qin Li [a,c], Fu-Xia Pan [d], Jian-Qiang Su [a,*]

[a] Key Lab of Urban Environment and Health, Institute of Urban Environment, Chinese Academy of Sciences, Xiamen, 361021, China
[b] University of Chinese Academy of Sciences, Beijing, 100049, China
[c] College of the Environmental & Ecology, Xiamen University, 361102, China
[d] Jinan Environmental Research Institute, Jinan, 250100, China

ARTICLE INFO

Keywords:
Microbial contamination
Microbial source tracking
Environmental protection
One health

ABSTRACT

Waterborne pathogens and their associated diseases are major threats to public health, and surveillance of pathogens and identification of the sources of pollution are imperative for preventing infections. However, simultaneously quantitative detection of multiple pathogens and pollution sources in water environments is the major challenge. In this study, we developed and validated a highly sensitive (mostly >80%) and highly specific (>99%) high-throughput quantitative PCR (HT-qPCR) approach, which could simultaneously quantify 68 marker genes of 33 human pathogens and 23 fecal markers of 10 hosts. The HT-qPCR approach was then successfully used to investigate pathogens and fecal pollution in marine recreational water samples of Xiamen, China. Totally, seven pathogenic marker genes were found in 13 beach bathing waters, which targeted *Acanthamoeba* spp., *Clostridium perfringens*, enteropathogenic *Escherichia coli*, *Klebsiella pneumoniae*, *Vibrio cholera/V. parahaemolyticus* and *Legionella* spp.. Fecal markers from human and dog were the most frequently detected, indicating human and dog feces were the main contamination in the recreational waters. Nanopore sequencing of full-length 16S rRNA gene revealed that 28 potential human pathogens were detected and electrical conductivity, salinity, oxidation-reduction potential and dissolved oxygen were significantly correlated with the variation in bacterial community. Our results demonstrated that HT-qPCR approach had the potential rapid quantification of microbial contamination, providing useful data for assessment of microbial pathogen associated health risk and development of management practices to protect human health.

1. Introduction

Waterborne pathogens, including pathogenic bacteria, virus, parasitic protozoon, a few of fungi and helminth species, have caused a high proportion of diseases, representing a major public health burden worldwide (Aw and Rose, 2012). World Health Organization (WHO) reported 842,000 deaths worldwide caused by water, sanitation and hygiene-attributable diarrheal diseases in 2012, accounting for 1.5% of the global disease burden measured as Disability-Adjusted Life Years (DALYs) (World Health Organization, 2014). Deterioration and contamination of water environment are directly related to the occurrence of these pathogens and their infectious diseases. The fecal contamination through the fecal-oral route accounts for a large proportion of pathogen infections in water environments (Raj et al., 2012). Therefore, there is an urgent need to develop powerful, sensitive and specific diagnostic tools for monitoring pathogens and tracking pollution sources in the water environments, with the final aim of improving or safeguarding human, animal and environmental health emphasized in the 'One Health'.

One major challenge for surveillance of waterborne pathogens is to simultaneously quantify the abundance of multiple pathogens, which is essential for comprehensive evaluation of microbial contamination level and health risk assessment. Although traditional culture dependent methods are extensively used for detection of pathogens, these methods are generally time-consuming and some pathogens are uncultivable (Aw and Rose, 2012). With the development of molecular tools, PCR-based method is the most commonly adopted technique for pathogen detection with the advantages of quick and quantitative analysis (Stedtfeld

* Corresponding author. 1799 Jimei Road, Xiamen, 361021, China.
E-mail address: jqsu@iue.ac.cn (J.-Q. Su).

https://doi.org/10.1016/j.envres.2020.109982
Received 26 February 2020; Received in revised form 17 July 2020; Accepted 18 July 2020
Available online 26 July 2020

et al., 2008; Ramalingam et al., 2010). Studies have reported that qPCR (multiplex qPCR and microfluidic qPCR) can simultaneously quantify several virulence factor genes of several pathogens (Wolf et al., 2007; Ishii et al., 2013). Several other DNA-based methods, e.g. DNA microarray (Inoue et al., 2015) and next generation sequencing (Ibekwe et al., 2013), have the high-throughput capacity for monitoring pathogens. However, quantitative information of targeted pathogens is mostly unavailable and the specificity of 16S rRNA gene-based DNA microarray and sequencing fails in distinguishing pathogens from non-pathogens (Ishii et al., 2013). Metagenomic Phylogenetic Analysis (MetaPhlAn) tool has been recently established to investigate the profile of bacterial pathogens (Li et al., 2015). Nevertheless, the cost of this approach is currently high and may not be able to meet the need for analysis of a large amount of samples for pathogen surveillance. Other techniques, such as Fluorescence in Situ Hybridization (Garcia-Armisen et al., 2004) and microfluidic card (Stedtfeld et al., 2015), are designed for simultaneous and rapid detection of multiple pathogens, while they still have challenges in sensitive or quantitative analysis.

Another major challenge is to identify the origin of microbial pollution in water body, which is paramount to project management practice to remedy the problem. Microbial source tracking (MST) has been widely adopted to identify fecal pollution sources through specific signature markers (mostly DNA sequences of hosts or their associated microorganisms). These host species, which have been implicated as potential pollution sources, generally include human, livestock (cow, pig, sheep, horse and rabbit), pets (dog and cat), poultries (duck and chicken) and wild animals (geese and gull). The source of fecal contamination in natural water bodies could be attributed to more than one host. Thus, identification of unknown sources of fecal contamination in aquatic environments may need to test multiple traits used as targets to determine the sources, which is laborious and time-consuming. PCR-based approach is one main mean for MST with high sensitivity and accuracy, whereas it is still required for developing a PCR-based approach to simultaneously detect multiple markers.

Recreational use of coastal environment represents a major pathway for human exposure to waterborne contamination in coastal cities. Sewage, human/animal feces and coastal sediments carried various pathogenic microorganisms into marine water, possibly causing community wide-spread of these waterborne pathogens and infectious diseases as well as significant economic losses (de Oliveira et al., 2008; Aw et al., 2012). Consequently, monitoring pathogens and tracking fecal pollution sources in marine recreational waters should be given more attention for protecting public health. In this study, a qPCR based high-throughput (HT-qPCR) approach was developed for detecting and quantifying human pathogens and tracking fecal pollution. The HT-qPCR method enables 5184 reactions with the reaction system of 100 nL in one run on the WaferGen Smart-Chip Real-Time PCR platform. In the HT-qPCR system, qPCR conditions and fluorescent probes for each assay were optimized and validated to improve the performance of this approach. Subsequently, the developed HT-qPCR method was used to evaluate the contamination level of pathogens and fecal contamination sources in marine recreational waters. We also performed nanopore sequencing for full-length 16S rRNA genes to investigate the bacterial community composition and potential pathogens in the marine recreational water.

2. Material and methods

2.1. Collection, verification, and optimization of marker gene primers for HT-qPCR assays

We collected 69 reported maker genes targeting 34 human pathogens to develop the HT-qPCR approach, most of which are waterborne pathogens and can cause intestinal illness (24 species) and respiratory infections, keratitis and other diseases. Owing to the differences in the sensitivity and specificity of these marker primers, we collected more than one marker for each pathogen. These marker genes mainly encoded the pathogenicity/toxicity-associated functions or ribosomes of the target pathogens (Table S1). The specificity of the primer and probe sequences of these marker genes was validated and optimized by using BLASTn against the NCBI nucleotide database (nr) according to our previous study (Zhuang et al., 2017). First, the primers/probes would be retained in the final primer list, if the sequences did match 100% to the marker gene of the target pathogen by aligning against the NCBI nr database. Otherwise, the primers/probes were removed or modified when the sequences had a low match with target gene or aligned to other host-associated markers. Second, DNA from more than 10 targeted pathogens obtained from CDC was used as DNA template, ensuring that there was no non-specific amplification by using their corresponding primers and probes (data not shown), we should note that we did not test all the pathogens owing to the difficulty in collecting all the pathogen DNA. Third, a composite standard plasmid mixture containing all targeted DNA fragment was used to further check the occurrence of non-specific amplification. DNA samples of partial pathogenic strains were gifts from Zhoushan Centre for Disease Prevention and Control and Zhoushan Hospital (Zhoushan, China) and remaining DNA sequences were synthesized from Thermo Fisher Scientific Inc. (Shanghai, China). Standard plasmids carrying the corresponding marker genes were constructed and equivalent copies of plasmids for each gene were combined to generate a composite standard plasmid mixture. Standard curves for marker genes were established using ten-fold serial diluted composite standard plasmid mixture.

We also collected 40 previously published MST markers for 11 hosts (Table S2). The specificity of these target marker primers was validated by using BLASTn against NCBI nr database. The sensitivity was evaluated by testing the primers and probes from 155 fecal samples of 13 types of hosts. These 155 feces were from humans (n = 25), pigs (n = 17), chickens (n = 28), ducks (n = 9), poultry (fecal mixture from chicken and duck, n = 3), dogs (n = 23), cattle (n = 16), sheep (n = 15), deer (n = 1), horses (n = 8), cats (n = 5), geese (n = 2), and rabbits (n = 3) and were collected from August 2015 to June 2016 (Table S3). Fecal collection was approved from the Institutional Human and Animal Ethics Committee of Institute of Urban Environment, Chinese Academy of Sciences. All experiments were performed in accordance with the approved guidelines. All fecal samples were transported on ice and stored at −20 °C until use. Genomic DNA was extracted using FastDNA SPIN kits for feces (MP Biomedicals, USA) and DNA concentrations were measured using a Qubit™ dsDNA Assay Kit (Invitrogen, USA).

Quantification and detection of microbial markers for human pathogens and MST were performed using TaqMan probe-based HT-qPCR on a WaferGen Smart-Chip Real-Time PCR system platform (WaferGen Inc. USA) as described previously (Zhao et al., 2018). The workflow of SmartChip Real-Time PCR system was provided in Fig. 1. All TaqMan® probes were uniformly optimized and modified using dual-labeled hydrolysis probes with a fluorophore (typically 6-carboxifluorescein FAM] at the 5′-end and a quencher (black hole quencher 1 BHQ1 or minor groove binder-non fluorescent quencher MGB-NFQ) at the 3′-end (Table S1 and Table S2). HT-qPCR reactions were performed in triplicate and sterilized water was included as the no template control. The universal primer pair and probe of bacterial 16S rRNA gene (*Bact2*) were included in the HT-qPCR assay as a reference gene (Table S2).

HT-qPCR assay was conducted in a 100-nL reaction system and the optimal reaction system listed in the TaqMan® Gene Expression Master Mix kit was applied, including 1 × TaqMan Gene Expression Master Mix (Applied Biosystems, USA), 1 × ROX (6-carboxyl-X-rhodamine) reference dye (Invitrogen, USA), bovine serum ampere (1 mg/mL, Sigma, USA), forward and reverse primers (0.9 μM/L), probe (0.25 μM/L), DNA (5 ng/μL) and nuclease-free water. The length of these targeted genes ranged from 50 bp to 250 bp. To enable the optimal amplification of all targeted genes, this study applied the optimal reaction system listed in the TaqMan® Gene Expression Master Mix kit. The amplification conditions were optimized and set as follows:

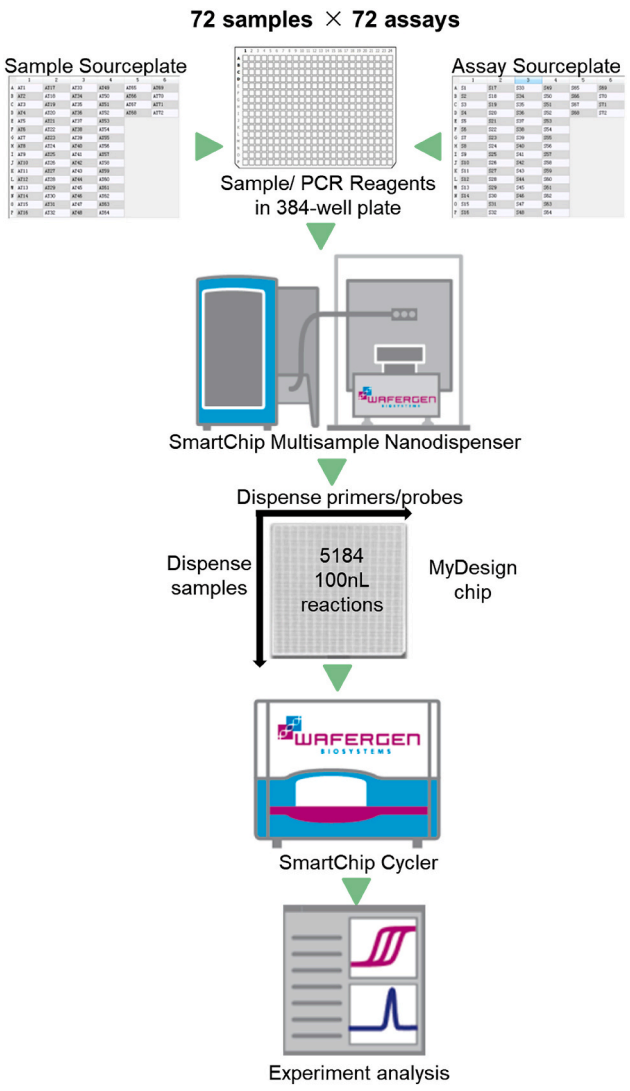

Fig. 1. Sampling sites of marine recreational beach water in this study.

a) 50 °C 2 min, 95 °C 10 min, 40 cycles (95 °C 15s, 60 °C 30s, 72 °C 30s);
b) 50 °C 2 min, 95 °C 10 min, 45 cycles (95 °C 15s, 55 °C 30s, 72 °C 30s);
c) 50 °C 2 min, 95 °C 10 min, 40 cycles (95 °C 15s, 60 °C 1 min);
d) 50 °C 2 min, 95 °C 10 min, 45 cycles (95 °C 15s, 60 °C 1 min);

Based on the amplification efficiency, qPCR and HT-qPCR conditions were finally set as following: 50 °C for 2 min and 95 °C for 10 min, followed by 40 cycles of 95 °C for 15 s and 60 °C for 1 min (Zhuang et al., 2017).

HT-qPCR data were analyzed using SmartChip qPCR software (V 2.7.0.1) according to our previous study with minor modification (Zhao et al., 2018). Reactions were discarded if melt curve had multiple peaks or amplification efficiency was beyond the acceptable range (90%–110%). A threshold cycle (Cq = 29) was set as cutoff value and a gene detected in three technical replicates was considered positive. One pathogen was considered to be truly present only if any one of its marker genes was positively detected. Absolute abundance (copies/L) of microbial markers was calculated according to standard curves.

2.2. Marine recreational water sample collection, DNA extraction and water chemistry

Water samples (surface waters) were collected in sterilized containers from 13 marine recreational beaches around Xiamen city (China) (Fig. S1). Water samples were taken within three consecutive days (July 2018) without recent rainfalls, resulting in a total of 39 samples. Samples were kept on ice and processed within 6 h after sample collection. Microbial cells from 600 mL of each sample were collected by filtering on 0.22-μm mixed cellulose esters filters (Millipore, USA). Filters were cut into 1 cm^2 pieces using sterile forceps and kept at −20 °C until genomic DNA extraction.

Temperature (Tem), electrical conductivity (EC), salinity (Sal), pH, oxidation-reduction potential (ORP) and dissolved oxygen (DO) were measured *in situ* using a multiparameter water quality analyzer (Hydrolab DS5, Hach Company, Loveland, USA). Phosphate phosphorus (PO_4^--P), ammonium-nitrogen (NH_4^+-N) and nitrate-nitrogen (NO_3^-/NO_2^-) were measured by a flow injection analyzer (QC8500, Lachat Instruments, USA) following standard methods (Greenberg et al., 1992).

2.3. Nanopore sequencing for 16S rRNA gene amplicons

Bacterial community was profiled by nanopore sequencing of full-length 16S rRNA gene amplicons with a barcoded primer set 27F (5′-AGAGTTTGATCMTGGCTCAG-3′) and 1492R (5′-TACGGYTACCTTGTTACGACTT-3′) (16S Barcoding Kit, SQK-RAB204, Oxford Nanopore Technologies, UK). PCR reactions were conducted in triplicate using 25 μL of 2 × ExTaq premix (Takara, Japan), 19 μL of nuclease-free water, 5 μL of DNA (20 ng/μL) and 1 μL of barcoded primers with the following thermal cycle: 98 °C for 1 min; 25 cycles of 98 °C for 20 s, 55 °C for 30 s, 72 °C for 2 min; followed by 72 °C for 5 min. Amplicons of each sample were pooled and gel purified using universal DNA purification kit (Tiangen, China) and KAPA Pure Beads (Roche, Switzerland) and quantified using a Qubit™ dsDNA Assay Kit (Invitrogen, USA).

Equivalent amount of PCR products for each sample were pooled for amplicon DNA library preparation. The Pooled amplicons (10 μL) was incubated with 1 μL of RAP adapter at room temperature for 5 min and then used as input for generation of MinION-compatible libraries (MinION™, Oxford Nanopore Technologies Ltd, UK). Each nanopore loading library for sequencing was constructed as follows: 34 μL of sequencing buffer, 25.5 μL of loading beads, 4.5 μL of nuclease-free water and 11 μL of DNA library. Sequencing libraries were loaded on a FLO-MAP 106 flow cell after flowcell equilibration and QC analysis. A 10-h sequencing for each library was then initiated using the MinION control software MinKNOW (version 1.15.4, GUI 2.2.15). Base-calling of raw FAST/HDF files was performed using the Metrichor™ agent v2.29.1.

2.4. Phylotype analysis

Raw 1D read data were analyzed using the FASTQ 16S workflow from the Desktop Agent EPI2ME. Sequences were removed if q-score was below 7, vim score was 0 or taxID was 0 or #N/A. Sequences with a length between 1300 bp and 1600 bp were retained. Eukarya and Archaea sequences were discarded from the final sequence files. The resulting sequences were demultiplexed and adapters were trimmed by matching unique barcode into each sample using *porechop* and *Albacore*. Taxonomy assignment was then performed using BLASTn against the SILVA database with an E-value of 10^{-5}. Based on the similarity of BLAST-based best hits, sequences were discarded if the similarity between query and reference <80%, Qcover <80% and length of overlap sequence <1000 bp. Chloroplast sequences and singletons were also removed.

Human pathogen database, including full-length 16S rRNA gene sequences of 684 human pathogens, was constructed by updating our previous database (containing 557 pathogenic species) (Dataset 1)

(Chen et al., 2016). Some non-human pathogenic sequences were deleted from our previous database and new human pathogenic sequences were added basing on the updated Virulence Factors of Pathogenic Bacteria (2019) (VFPB, 2019, http://www.mgc.ac.cn/cgi-bin/VFs/v3_ai_main.cgi?ID=F06) and the database of pathogen-associated genes (http://www.pathogenomics.sfu.ca/pathogen-associated/2014/). Based on BLASTn sequence similarity search (E-value = 10^{-5}, qcover ≥ 90% and identity ≥ 90%), potential human pathogens were retrieved from the newly constructed human pathogen database.

2.5. Statistical analyses

The sensitivity and specificity of MST molecular markers were calculated and evaluated as previously described by Zhuang et al. (2017). Briefly, sensitivity was defined as the fraction of actual positive [true positive (TP)] host samples divided by all expected positive host samples, including both false negative (FN) and TP (Equation (1)). Specificity was calculated as the fraction of actual negative [true negative (TN)] host samples divided by all expected negative host samples, containing unexpected positive [false positive (FP)] and TN (Equation (2)).

$$\text{Sensitivity} = \text{TP}/(\text{FN} + \text{TP}) \quad (1)$$

$$\text{Specificity} = \text{TN}/(\text{FP} + \text{TN}) \quad (2)$$

Bacterial alpha-diversity (Shannon index), non-metric multidimensional scaling (NMDS) for beta-diversity analysis, redundancy analysis (RDA) and heatmap plotting were performed in R 3.2.3 (R Foundation for Statistical Computing: Vienna, Austria, 2014) with vegan 2.0–10 (Oksanen et al., 2007) and pheatmap package (Kolde, 2013). The mean concentrations represent the arithmetic averages, and errors stand for standard deviations of the mean.

3. Results

3.1. Physicochemical characteristics of marine water

The physicochemical characteristics of recreational water in these beaches showed that water temperature (from 29.4 to 31.3 °C), salinity (from 27.7‰ to 34.4‰) and pH (from 7.8 to 7.9) had no distinct difference among these samples (Table S4). Electrical conductivity ranged from 43,001 to 46,637 μS/cm and the highest electrical conductivity was detected in the sample HC. Oxidation-reduction potential ranged from 391 to 493 mV and dissolved oxygen ranged from 6.0 to 6.5 mg/L. Nitrate nitrogen contents were found to be from 0.1 to 0.7 mg/L and ammonium ion concentrations were lower than nitrate concentration, with the range from 0.1 to 0.4 mg/L. In most of samples, phosphate concentration was lower than detection limit, except HLS, BC and GLY. According to the National Standard of Environmental Quality for Surface Water (China, GB3838-2002), almost all waters were at level II except WYW (level III) based on DO contents. In terms of NH_3^+-N concentration, five samples were at level I and the others were at level II.

3.2. Optimization of pathogens and MST maker genes for HT-qPCR assay

Primer sets for *invA* of *Salmonella enterica* and *cdtA* of *Clostridium difficile* were modified due to the low match with the target genes (Table S1). *Bacillus* spp. 16S rRNA gene marker was removed, since it completely aligned with other species of Staphylococcaceae, Listeriaceae, Planococcaceae, Carnobacteriaceae, Proteobacteria, Micrococcales and so on. The diarrheal enterotoxin gene (*bceT*) of *Bacillus cereus* also completely matched with that of *Bacillus thuringiensis*, thus it was modified to detect *Bacillus cereus/B. thuringiensis*. Finally, 68 pathogenic microbial markers targeting 33 pathogens were used for HT-qPCR assay in this study (Table S1), including 24 pathogens causing gastrointestinal disease (*Helicobacter pylori*, *Vibrio cholera*, *Salmonella*, *Clostridium difficile*, pathogenic *Escherichia coli*, *Giardia lamblia*, *Cryptosporidium* and *Staphylococcus aureus*, etc) and nine pathogens causing pneumonia or keratitis (*Acanthamoeba* spp., *Legionella pneumophila*, *Klebsiella pneumoniae*, *Streptococcus pneumoniae*, *Mycobacterium tuberculosis* and *Pseudomonas aeruginosa*, etc). To further verify the homogeneity of HT-qPCR reactions, standard curves for all target genes were generated under the same recreation system and the same amplification condition. Amplification efficiency of these pathogenic marker genes were in the range from 90.67% to 109.13% with the coefficients (R^2) of the standard curves from 0.992 to 0.999 (Table S5).

We also evaluated the sensitivity and specificity of MST markers (Table 1 and Fig. 2). Finally, 17 markers were removed and primers of 5 marker genes were modified, including Cow mtDNA, Swine mtDNA, Pig mtDNA, Sheep mtDNA and Cat mtDNA (Table S2). Sensitivity and specificity of BoBac showed that the marker gene could be used for detecting the fecal pollution of ruminants including cow, sheep and deer. Thus, the eventual HT-qPCR assays for MST contained 23 marker genes targeting 10 fecal sources (human, poultry [chicken and duck], ruminant [cattle, sheep, and deer], dog, pig, horse and cat) (Table S2). Among these 23 markers, the sensitivity of 17 markers was over 50%, and 11 markers were with a high sensitivity of more than 80%; 21 markers had a high specificity of above 90% (Table 1). Since some of the hosts were targeted by more than two markers, the sensitivity of all HT-qPCR assays was >80% (except duck and horse) and the specificity was >99%.

3.3. Pathogenic microorganisms in marine recreational water

Seven pathogens, including *Acanthamoeba* spp., *Clostridium perfringens*, enteropathogenic *E. coli*, *Klebsiella pneumoniae*, *Vibrio cholerae*-*Vibrio parahaemolyticus*, *Legionella* spp., were detected in 9 of 13 bathing beach waters (Fig. 3 and Table 2). The abundance of these detected marker genes for pathogens ranged from 1.4×10^3 to 1.1×10^5 copies/L. *Legionella* spp. genes were the most prevalent with 61.5% of samples positive and the abundance ranged from 7.0×10^3 to 2.0×10^4 copies/L. *Klebsiella pneumoniae* genes were positive in three beaches (23.1%), with the abundance from 1.3×10^3 to 1.1×10^5 copies/L. The detection frequencies of both *Acanthamoeba* spp. and *V. parahaemolyticus* genes were 15.4%, with the abundance from 4.0×10^3 to 7.2×10^3 copies/L and from 1.4×10^3 to 1.7×10^3 copies/L, respectively. Among 13 bathing beaches, all detected pathogens were positive in GZH and three pathogens (*Acanthamoeba* spp., *Klebsiella pneumoniae* and *Legionella* spp.) were found in ZZW. Total bacterial 16S rRNA genes were determined with the abundance ranging from 3.2×10^9 to 2.4×10^{10} copies/L and GZH harbored the most abundant bacteria.

3.4. Microbial community and potential human pathogens

To characterize the distribution of microbial community and potential human pathogens, full-length 16S rRNA gene amplicons were sequenced using the MinION™ platform. Totally, 537,231 raw 1D reads were obtained with the length ranging from 62 bp to 6129 bp (Fig. S2). After quality control and filtration of low-quality reads, 122,267 clean reads (2202–5169 reads per sample, mean = 3135) were assigned into 746 bacterial species. Alpha diversity of microbial community showed no significant difference among sampling sites. Bacterial community structures were heterogeneous among sampling sites, but regional clustering was not observed (Fig. S3 and Fig. S4). Proteobacteria, Bacteroidetes and Actinobacteria were the dominant phyla in these water samples, accounting for 93.7%–98.1% of total sequences (Fig. S5). *Bacterium* WHC1_1, *Arcobacter cryaerophilus* and *Methylophilaceae bacterium* NB0076_01 were the top 3 species (Fig. S6). RDA analysis revealed that nine environmental factors explained 22.9% of variation in microbial community and electrical conductivity (EC), salinity (Sal), oxidation-reduction potential (ORP) and dissolved oxygen (DO) were significantly correlated with the variation of microbial community

Table 1
Sensitivity and specificity of 40 collected MST markers tested with 155 fresh fecal samples. The final 23 marker genes with relatively high sensitivity and specificity were marked as shades of gray.

Hosts	Markers	Sensitivity	Specificity	Fecal samples (the number of actual positive/total host samples)												
				Human	Pig	Chicken	Duck	Poultry	Dog	Cattle	Sheep	Deer	Horse	Cat	Goose	Rabbit
Human	Human mtCytb	100% (25)	100% (130)	25/25	0/17	0/28	0/9	0/3	0/23	0/16	0/15	0/1	0/8	0/5	0/2	0/3
	HumM2	76% (25)	94.6% (130)	19/25	0/17	0/28	0/9	0/3	2/23	0/16	4/15	0/1	1/8	0/5	0/2	0/3
	HF183-BacR287	36% (25)	94.6% (130)	9/25	0/17	5/28	0/9	0/3	0/23	0/16	1/15	0/1	0/8	0/5	0/2	1/3
	HF183-BFDrev	36% (25)	92.3% (130)	9/25	0/17	5/28	1/9	0/3	0/23	0/16	1/15	0/1	2/8	0/5	0/2	1/3
	B. fragilis gyrB	36% (25)	91.5% (130)	9/25	2/17	3/28	2/9	0/3	2/23	0/16	1/15	0/1	0/8	0/5	0/2	1/3
	B. theta α-mannanase	0% (25)	100% (130)	0/25	0/17	0/28	0/9	0/3	0/23	0/16	0/15	0/1	0/8	0/5	0/2	0/3
	BacHuman	100% (25)	34.6% (130)	25/25	11/17	24/28	5/9	2/3	19/23	4/16	5/15	1/1	4/8	5/5	2/2	3/3
	M. smithii nifH	0% (25)	100% (130)	0/25	0/17	0/28	0/9	0/3	0/23	0/16	0/15	0/1	0/8	0/5	0/2	0/3
	HPyV	0% (25)	100% (130)	0/25	0/17	0/28	0/9	0/3	0/23	0/16	0/15	0/1	0/8	0/5	0/2	0/3
	JCPyV	0% (25)	100% (130)	0/25	0/17	0/28	0/9	0/3	0/23	0/16	0/15	0/1	0/8	0/5	0/2	0/3
Pig	Pig-2-Bac	100% (17)	100% (138)	0/25	17/17	0/28	0/9	0/3	0/23	0/16	0/15	0/1	0/8	0/5	0/2	0/3
	Swine mtDNA	88.2% (17)	99.3% (138)	1/25	15/17	0/28	0/9	0/3	0/23	0/16	0/15	0/1	0/8	0/5	0/2	0/3
	PGprobe	11.8% (17)	100% (138)	0/25	2/17	0/28	0/9	0/3	0/23	0/16	0/15	0/1	0/8	0/5	0/2	0/3
	PAdV	0% (17)	100% (138)	0/25	0/17	0/28	0/9	0/3	0/23	0/16	0/15	0/1	0/8	0/5	0/2	0/3
	Pig mtDNA	5.9% (17)	100% (138)	0/25	1/17	0/28	0/9	0/3	0/23	0/16	0/15	0/1	0/8	0/5	0/2	0/3
Poultry	Chicken and duck *ND5*	97.5% (40)	99.1% (115)	0/25	0/17	25/28	7/9	3/3	0/23	1/16	0/15	0/1	0/8	0/5	0/2	0/3
	Chicken and duck *cytb*	87.5% (40)	99.1% (115)	0/25	0/17	28/28	8/9	3/3	0/23	1/16	0/15	0/1	0/8	0/5	0/2	0/3
	Av43	57.5% (40)	96.5% (115)	0/25	0/17	18/28	4/9	1/3	4/23	0/16	0/15	0/1	0/8	0/5	0/2	0/3
	PLprobe	40% (40)	94.8% (115)	0/25	0/17	13/28	3/9	0/3	2/23	0/16	1/15	0/1	0/8	3/5	0/2	0/3
	Chicken/Duck-Bac	20% (40)	95.6% (115)	0/25	4/17	6/28	2/9	0/3	0/23	1/16	0/15	0/1	0/8	0/5	0/2	0/3
	CL	0% (40)	100% (115)	0/25	0/17	0/28	0/9	0/3	0/23	0/16	0/15	0/1	0/8	0/5	0/2	0/3
Duck Dog	Duck-Bac	22.2% (9)	100% (146)	0/25	0/17	0/28	2/9	0/3	0/23	0/16	0/15	0/1	0/8	0/5	0/2	0/3
	Dog mtDNA	95.6% (23)	100% (132)	0/25	0/17	0/28	0/9	0/3	22/23	0/16	0/15	0/1	0/8	0/5	0/2	0/3
	DG37	78.3% (23)	97.7% (132)	2/25	0/17	0/28	0/9	0/3	18/23	0/16	0/15	0/1	0/8	1/5	0/2	0/3
	BacCan-UCD	65.2% (23)	87.1% (132)	3/25	1/17	4/28	1/9	0/3	15/23	0/16	8/15	0/1	0/8	0/5	0/2	0/3
	DG3	43.5% (23)	100% (132)	0/25	0/17	0/28	0/9	0/3	10/23	0/16	0/15	0/1	0/8	0/5	0/2	0/3
	DG72	0% (23)	100% (132)	0/25	0/17	0/28	0/9	0/3	0/23	0/16	0/15	0/1	0/8	0/5	0/2	0/3
Cattle	BoBac cattle-associated	93.8% (16)	82.0% (139)	0/25	0/17	1/28	1/9	0/3	8/23	15/16	14/15	1/1	0/8	0/5	0/2	0/3
	BoBac[a] ruminant-associated	93.8% (32)[a]	91.9% (123)[a]													
	Cow mtDNA	93.8% (16)	100% (139)	0/25	0/17	0/28	0/9	0/3	0/23	16/16	0/15	0/1	0/8	0/5	0/2	0/3
	CowM3	50.0% (16)	100% (139)	0/25	0/17	0/28	0/9	0/3	0/23	8/16	0/15	0/1	0/8	0/5	0/2	0/3
	CowM2	18.8% (16)	100% (139)	0/25	0/17	0/28	0/9	0/3	0/23	3/16	0/15	0/1	0/8	0/5	0/2	0/3
	BacCow-UCD	93.8% (16)	68.3% (139)	1/25	11/17	13/28	4/9	0/3	0/23	15/16	10/15	1/1	4/8	0/5	0/2	0/3
	BacBovine	100% (16)	57.6% (139)	14/25	2/17	4/28	5/9	0/3	15/23	16/16	14/15	1/1	0/8	3/5	0/2	1/3
	CWprobe	43.8% (16)	78.64% (139)	0/25	5/17	4/28	0/9	2/3	12/23	7/16	4/15	1/1	1/8	0/5	0/2	1/3
Sheep	Sheep mtDNA	80% (15)	100% (140)	0/25	0/17	0/28	0/9	0/3	0/23	0/16	12/15	0/1	0/8	0/5	0/2	0/3
Horse	HorseBact-UCD	62.5% (8)	100% (147)	0/25	0/17	0/28	0/9	0/3	0/23	0/16	0/15	0/1	5/8	0/5	0/2	0/3
Cat	Cat mtDNA	100% (5)	100% (150)	0/25	0/17	0/28	0/9	0/3	0/23	0/16	0/15	0/1	0/8	5/5	0/2	0/3
Muskrat	MuBa01	0% (0)	100% (155)	0/25	0/17	0/28	0/9	0/3	0/23	0/16	0/15	0/1	0/8	0/5	0/2	0/3
White-tailed deer	White-tailed deer mtCytb	0% (0)	100% (155)	0/25	0/17	0/28	0/9	0/3	0/23	0/16	0/15	0/1	0/8	0/5	0/2	0/3
Gull	Gull4	0% (0)	100% (155)	0/25	0/17	0/28	0/9	0/3	0/23	0/16	0/15	0/1	0/8	0/5	0/2	0/3

[a] Sensitivity and specificity of BoBac when the marker gene indicates ruminants (including cow, sheep and deer fecal pollutions).

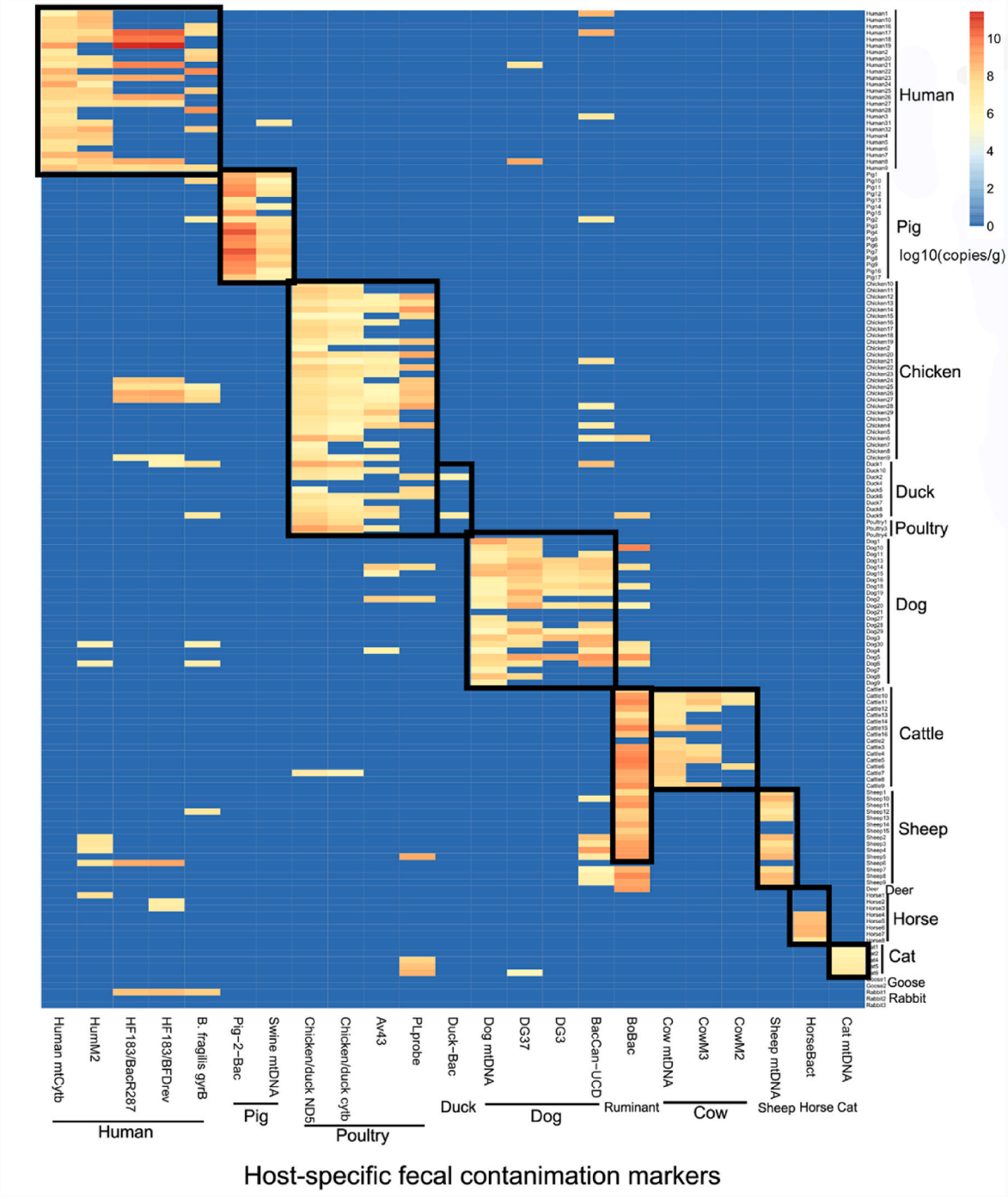

Fig. 2. Sensitivity and specificity of marker genes for tracking fecal contaminant by using HT-qPCR assays. The values presented were log transformed abundance of pathogenic marker genes.

(Fig. 4).

Totally, 28 potential human pathogens (2–9 pathogens for each sample) were detected in 39 samples with the relative abundance ranging from 0.04% to 0.25% (Fig. 5). *Arcobacter butzleri*, *Escherichia coli* and *Klebsiella pheumoniae* were the predominant human pathogenic bacteria. *Aeromonas*, *Burkholderia*, *Staphylococcus*, *Pseudomonas* and *Vibrio* were also detected in these recreational beach waters. RDA analysis revealed that environmental factors explained more than 34.3% of the variation in potential bacterial pathogen community (Fig. S7) and water salinity, electrical conductivity, ammonium and temperature significantly associated with the potential pathogen structure ($P < 0.01$).

3.5. Microbial source tracking (MST) of fecal contaminants in recreational waters

Fecal pollution was one of the main sources of human pathogens. Fecal source markers were used in the HT-qPCR approach to track the source of fecal contamination in recreational waters. Four MST markers for human and dog feces were detected, in which human feces contamination was detected in nearly all samples (except MH) and two beaches (HC and GZH) were contaminated by dog feces (Table 3).

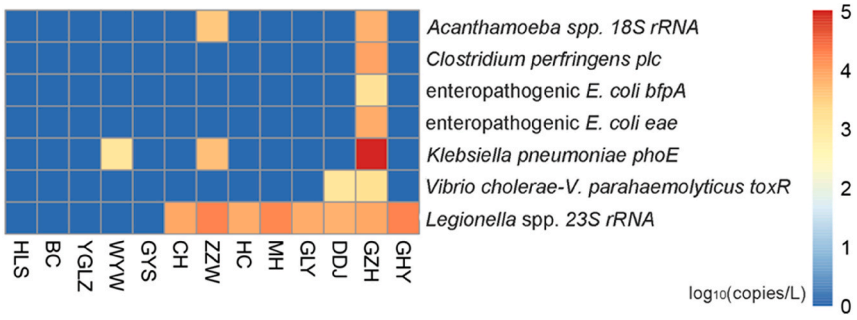

Fig. 3. Abundance of human pathogen marker genes detected in marine recreational beach water using HT-qPCR assay. The values presented were log transformed abundance of pathogen marker genes.

4. Discussion

4.1. Assessment of HT-qPCR for analyzing human pathogens and fecal contamination

The HT-qPCR approach allows the rapid and quantitative detection of multiple genes from multiple samples (5184 reactions each run). HT-qPCR have been widely applied in detection and quantification of various microbial functional genes, such as genes for antibiotic resistance (Looft et al., 2012), nutrient cycling (the transformation of carbon, nitrogen, phosphorus and sulfur) (Zheng et al., 2018) and arsenic metabolism (Zhao et al., 2018). In this study, we evaluated and optimized the HT-qPCR-based assay for pathogens and fecal pollution based on the sensitivity and specificity of selected primers and probes. The optimized HT-qPCR assay could simultaneously target 68 markers for 33 pathogens and 23 fecal markers for 11 hosts, enabling comprehensive and quantitative profiles of microbial pollution. The number of markers for HT-qPCR could be further expanded for identifying more pathogens/fecal sources of interest from different environmental contexts. The established HT-qPCR assay covered several pathogens (e.g. *Pseudomonas aeruginosa*, *Salmonella* species and *Staphylococcus aureus*), which were listed as antibiotic-resistant priority pathogens by WHO (2017), which provided useful environmental reference data when comparing with clinical diagnosis of resistant pathogens. This approach was highly sensitive (mostly >80%), highly specific (>99%) and reliable, and could be adopted to evaluate the presence of selected pathogens and track the source of fecal contaminants by detecting specific pathogenic genes or small-subunit ribosomal RNA gene even if DNA templates were at low concentrations. We performed qPCR reactions using a series 10-fold diluted plasmid as positive control with a minimum concentration of 1 copy per well (100 nL per well). The results showed at the concentration of 1 copy per well, the average Cq is about 31 (most fall within 29–33), which could be clearly discriminated from NTC (Cq 0 or >33), thus Cq 31 was used as a threshold for LOD. At a concentration of 10 copies per well, the average Cq is about 29 ± 0.3, thus Cq 29 was used as a threshold for LOQ. Since the DNA concentration of template used in the platform was adjusted to be 50 ng/μL, LOD and LOQ of this method were corresponded to 2 and 20 copies/ng DNA, respectively. Moreover, most of pathogens/host sources were targeted by using more than one pathogen/host-associated marker, considerably strengthening the robustness of the HT-qPCR array for detection of pathogen/fecal contamination. However, HT-qPCR approach still has its limitations. Firstly, HT-qPCR analysis is unable to capture the full profiles of all potential pathogens owing to the limitation in primer design. Secondly, HT-qPCR is not designed for characterizing pathogenicity of pathogens and thus the actual infection risk to a population is limited owing to its incapacity of viability determination (Ramírez-Castillo et al., 2015).

4.2. Pathogens in recreational water of bathing beaches

By applying the developed HT-qPCR approach, we found that at least one human pathogen was detected in nine recreational marine waters (69.2%). This indicated that human exposure (for example, aerosol inhalation, dermal, eye or ear contact and oral intake) to these recreational bathing waters might represent a potential health risk, despite that these samples (classified at level III) were qualified for direct contact. However, the concentrations of these detected human pathogens were generally lower than those reported in soil, wastewater, human and animal gastrointestinal tract (Table S6). To evaluate the human exposure doses to these pathogens, we collected the information of infectious dose and copy number of the detected marker gene in one pathogen cell (Table 2) (Pathogen safety data sheets and risk assessment, 2014; Walderhaug, 2014). It has been reported that *Clostridium perfringens plc*, enteropathogenic *E. coli eae* and *Vibrio cholerae-V. parahaemolyticus toxR* are the single copy chromosomal genes, and enteropathogenic *E. coli bfpA* and *Klebsiella pneumoniae phoE* are the single copy plasmid genes (Van Der et al., 1987; Tobe et al., 1999; Skånseng et al., 2006; Liu et al., 2012; Luedtke et al., 2014). The average amount of water swallowed by adults was approximately 16 mL during at least 45-min swimming activity (Dufour et al., 2006; Mora et al., 2016). Therefore, we estimated that the oral ingestion doses of *Clostridium perfringens*, EPEC, *Vibrio cholerae- V. parahaemolyticus* and *Klebsiella pneumoniae* were 3.4, 5.3×10^2–2.7×10^3, 5.0×10^2–6.0×10^2 and 4.6×10^2–3.9×10^4 CFU/min, respectively, which were significantly lower than the infectious doses even during 6 h-swimming activity (Table 2). This suggested that the pathogen concentrations of Xiamen beach water were deemed acceptable and their potential human exposure risk was low when swimming in these areas.

In these recreational waters, *Legionella* spp. marker was the most prevalent with a detection frequency of 61.5%. Previous studies revealed that *Legionella* spp. can grow within protozoan hosts (for example, amoebae) and the close symbiotic relationship might contribute to its wide distribution, since intracellular *Legionella* spp. might be highly resistant to adverse conditions and could propagate rapidly under favorable conditions (Richards et al., 2013). *Klebsiella pneumoniae* was another commonly found pathogen (detection frequency 23.1%), which is a coliform bacterium and mainly causes infections of respiratory tract and intestinal tract. *Klebsiella pneumoniae* was frequently detected in the final effluent, although wastewater treatment could significantly reduce its concentration (Shannon et al., 2007). Direct discharge of effluent provided a possible source of *Klebsiella pneumoniae* in the marine recreational water. Remarkably, *Klebsiella pneumoniae* are Gram-negative bacteria and most isolates confer resistance to carbapenem through carbapenem-hydrolysing β-lactamase (Munoz-Price et al., 2013). Plasmid-mediated transfer of bla_{KPC} (gene encoding the carbapenemases) had been observed between *Klebsiella pneumoniae* and *Enterobacteriaceae* (Mathers et al., 2011), indicating that exposure to the “superbugs"-inhabiting environments (e.g. WYW, ZZW

Table 2
Number and abundance of marker genes for human pathogens detected in marine recreational water using HT-qPCR analysis.

Abundance (copies/L)	HLS	BC	YGLZ	WYW	GYS	CH	ZZW	HC	MH	GLY	DDJ	GZH	GHY	Detection frequency (%)	Infectious dose (cell number)[f]	Pathogenicity/Toxicityf	Host range[f]
Number of detected pathogens	0	0	0	1	0	1	3	1	1	1	2	7	1				
Acanthamoeba spp. 18S rRNA	0	0	0	0	0	0	4.0×10^3	0	0	0	0	7.2×10^3	0	15.4	unkown	Keratitis/granulomatous amoebic encephalitis	human/animals
Clostridium perfringens plc[a]	0	0	0	0	0	0	0	0	0	0	0	9.5×10^3	0	7.7	$>10^8$	Food poisoning/Gas Gangrene/Enteritis necroticans	humans/dogs/pigs/goats
enteropathogenic *E. coli bfpA*[b]	0	0	0	0	0	0	0	0	0	0	0	1.5×10^3	0	7.7	$>10^6$–10^8	Diarrhea	human/cattle/horses
enteropathogenic *E. coli eae*[c]	0	0	0	0	0	0	0	0	0	0	0	7.6×10^3	0	7.7			
Klebsiella pneumoniae phoE[d]	0	0	0	1.3×10^3	0	0	5.2×10^3	0	0	0	0	1.1×10^5	0	23.1	$>10^8$		
Vibrio cholerae-V. parahaemolyticus toxR[e]	0	0	0	0	0	0	0	0	0	0	1.4×10^3	1.7×10^3	0	15.4	$>10^6$–10^{11} for *Vibrio cholerae*; $>10^7$–10^8 for *V. parahaemolyticus*	Cholera; gastroenteritis (diarrhea)/wound infections/septicaemia	Humans/water birds/shellfish/fish/herbivores
Legionella spp. 23S rRNA	0	0	0	0	0	8.7×10^3	1.9×10^4	7.4×10^3	1.8×10^4	8.2×10^3	7.0×10^3	9.3×10^3	2.0×10^4	61.5	unkonown	Legionnaire's disease/Pontiac fever	Humans/amoeba, and experimental infection in guinea pigs, mice, rats, embryonated chicken eggs, human and animal cell lines
Bact2	4.3×10^9	4.5×10^9	3.2×10^9	4.3×10^9	7.0×10^9	1.3×10^{10}	1.9×10^{10}	8.5×10^9	1.2×10^{10}	7.3×10^9	1.2×10^{10}	2.4×10^{10}	4.8×10^9	100			

[a] *plc* gene is a single copy number chromosomal gene (Skånseng et al., 2006). [b] *bfpA* is a single copy number gene on enteroadherent factor (EAF) plasmid (Tobe et al., 1999). [c] *eae* is a single copy number chromosomal gene (Luedtke et al., 2014). [d] *phoE* is a single copy number gene on plasmid (Van et al., 1987). [e] *toxR* is a single copy number chromosomal gene. [f] Data was collected from the study reported by Walderhaug (2014) and Public Health Agency of Canada (Pathogen safety data sheets and risk assessment, 2014).

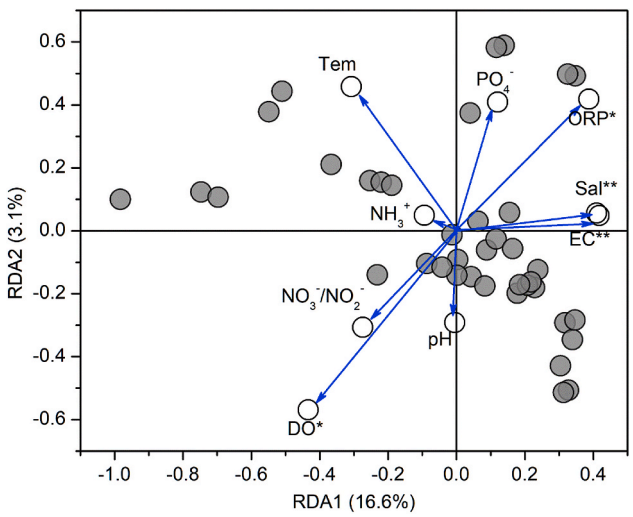

Fig. 4. RDA analysis revealed that electrical conductivity (EC), salinity (Sal), oxidation-reduction potential (ORP) and dissolved oxygen (DO) significantly correlated with the variation in bacterial community basing on nanopore sequencing data of full-length 16s rRNA genes. $^{*}P < 0.05$; $^{**}P < 0.01$.

and GZH) might pose a higher health risk to the public.

Clostridium perfringens and enteropathogenic *E. coli* were only observed in the sample GZH. *C. perfringens* is an anaerobe, while it was detected in the GZH (dissolved oxygen, 6.3 ± 0.3 mg/L), indicating that *C. perfringens* possibly kept a dormant, survival spore state. Enteropathogenic *E. coli* (EPEC) is the most prevalent diarrheagenic *E. coli* pathotype worldwide, which mainly causes diarrhea of children (Lanata et al., 2013). The co-occurrence of the chromosomal gene *eae* (encoding intimin) and the plasmid gene *bfpA* (encoding bundle-forming pilus) was observed in the GZH marine water, indicating that the EPEC subtype, typical EPEC (tEPEC, carrying both *eae* and *bfp*), was definitely present. Therefore, recreational activities in the GZH marine water possibly put human beings at the risk of the diarrhea infection (Lanata et al., 2013; Mora et al., 2016; Fuhrmeister et al., 2019).

By analyzing full length 16S rRNA gene sequences, we detected a low abundance (0.04%–0.25% of total sequences) of potential pathogens. The data was consistent with one recent metagenomic evidence that low relative abundance of opportunistic pathogens in both potable and reclaimed water, ranging from −4.35 to −1.65 log hits/reads (Garner et al., 2018). Beach HC was found to contain the most potential pathogens (0.24%, 9 potential pathogens), which was different from the result from HT-qPCR assay showing that the most abundant pathogens were detected in beach GZH ($\sim 10^5$ copies/L, 7 pathogens). Additionally, only two potential pathogens *E. coli* and *Klebsiella pneumoniae* were simultaneously detected by using both nanopore sequencing analysis and HT-qPCR assay. The inconsistency between the results of HT-qPCR and nanopore sequencing could be explained by the following discrepancies. The established HT-qPCR detected 33 human pathogens that represent a small fraction of all human pathogens in environments. The gene markers of bacteria, virus and protozoon were contained in the HT-qPCR assay. Nanopore sequencing of full-length 16S rRNA gene (targeting bacteria) enables capturing full profile of the potential bacterial pathogens by blasting against bacterial databases. Additionally, high sensitivity of HT-qPCR allows the detection of pathogens with low

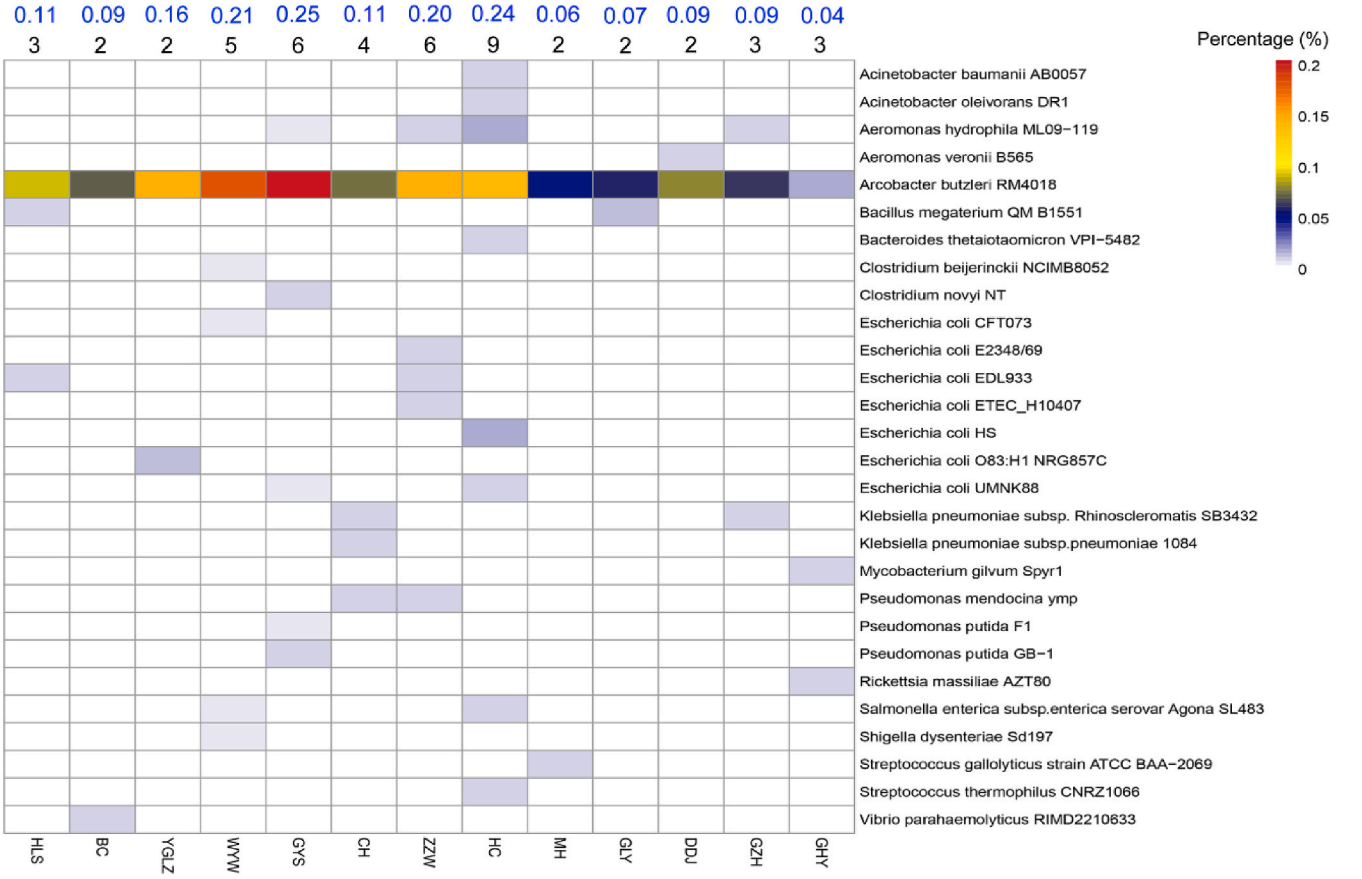

Fig. 5. Potential pathogens detected by BLASTn 16S rRNA gene sequences against human pathogen database (Dataset1). The blue number represents the relative abundance (percentage, %) of potential pathogens in each site. The black number stands for the number of detected pathogens. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Table 3
Microbial pollution sources in marine recreational water detected by HT-qPCR assays.

Source	Functions	Marke gene ID	HC	MH	GLY	DDJ	GZH	GHY	HLS	BC	YGLZ	WYW	GYS	CH	ZZW
Human	*Bacteroides fragilis* DNA gyrase subunit B	B. fragilis *gyrB*	✓	×	×	×	✓	×	×	×	×	×	×	×	×
Dog	Bacteroidales 16S rRNA	BacCan-UCD	✓	×	×	×	✓	×	×	×	×	×	×	×	×
Human	Bacteroides 16S rRNA	HF183-BacR287	✓	×	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Human	Bacteroidales 16S rRNA	HF183-BFDrev	✓	×	×	×	✓	×	✓	✓	✓	✓	✓	✓	✓

concentrations in environments, whereas these pathogens might be undetectable by using nanopore sequencing due to the sequencing depth and coverage. These results indicated that sequencing method could be more robust in capturing the profiles of potential pathogens in the environments, while HT-qPCR provided more accurate and quantitative detection of targeted pathogens.

4.3. Factors influencing the occurrence of human pathogens

Studies have reported that environmental factors interacting with microbial community considerably affect the occurrence of certain pathogens (Wilson and Salyers, 2003; Locatelli et al., 2013). RDA analysis revealed that environmental factors, e.g. EC, Sal and ORP significantly correlated with the variation of microbial community, and they also significantly shaped the structure of potential bacterial pathogen community (totally explained 34.3% of variation of pathogen community). Consequently, the accumulation and elimination of human pathogens could be affected by these environmental factors (Hernroth et al., 2002).

Human and dog feces-specific markers were detected, and human fecal markers were found in almost all samples, indicating that the marine recreational waters had been contaminated by human and dog feces and human feces was the major contamination source. This was consistent with previous studies, which reported that human sources and canine were the influential fecal sources to suburban coastal beaches (Ervin et al., 2014). Therefore, fecal contamination inspection should be strengthened in recreational water bodies. Based on 16S rRNA gene analysis, we identified a few taxonomic groups such as *Bacteroidales* and *Lachnospiraceae*, some of which were considered as alternative indicators and 'core' community members of wastewater treatment plants and domestic sewage (McLellan et al., 2013; Su et al., 2017). This indicated that human sewage contamination possibly also contributed to the occurrence of human pathogens in the recreational water.

The beach GZH harbored the most abundant pathogens with seven gene markers detected, representing a higher illness risk. GZH was located in Gulangyu islet, a popular tourist spot in Xiamen with more bathers (data not shown). Re-suspension of bottom sediments by bathers could cause elevated levels of pathogens (Graczyk et al., 2010; Sinigalliano et al., 2010), thus the bather density might contribute to the prevalence of pathogens in the coastal recreational water. Additionally, the total bacterial biomass might be responsible for the high prevalence of human pathogens, since GZH was detected with the most abundant bacteria. Hence, limiting the number of bathers in recreational waters is recommended to reduce the risk of illness.

5. Conclusions

In the present study, an accurate, high-throughput quantitative PCR method was established to facilitate the detection and quantification of multiple human pathogens and fecal contaminants. This approach was then successfully applied for the surveillance of microbial contaminants in recreational waters and showed its potentials for broader application in other environments (e.g. waste waters, soils, air and feces), finally aiding to draw real-time maps of human pathogens distribution or fecal contamination in the future. Further studies including large scale epidemiologic investigation should be focused to determine the correlation of pathogenic microbes and human illness.

Credit author statement

Xin-Li An: Methodology, Writing - original draft, Formal analysis, Software. Jia-Ying Wang: Sample collection, Data curation. Qiang Pu: Visualization, Formal analysis. Hu Li: Formal analysis. Ting Pan: Sample collection, Methodology. Huan-Qin Li: Sample collection, Methodology. Fu-Xia Pan: Sample collection. Jian-Qiang Su: Conceptualization, Writing- Reviewing and Editing, Supervision.

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.

Acknowledgements

This study was financially supported by the National Key Research and Development Program of China-International collaborative project from Ministry of Science and Technology (2017YFE0107300), the National Key Research and Development Program of China (2016YFD0800205), the National Natural Science Foundation of China (41907217, 31722004 and 41801089), Xiamen Science and Technology Bureau (WES&WQGE201904) and K.C.Wong Education Foundation.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envres.2020.109982.

References

Aw, T.G., Rose, J.B., 2012. Detection of pathogens in water: from phylochips to qPCR to pyrosequencing. Curr. Opin. Biotechnol. 23 (3), 422–430.
de Oliveira, A.J.F.C., Pinhata, J.M.W., 2008. Antimicrobial resistance and species composition of *Enterococcus* spp. isolated from waters and sands of marine recreational beaches in Southeastern Brazil. Water Res. 42 (8–9), 2242–2250.
Dufour, A.P., Evans, O., Behymer, T.D., Cantu, R., 2006. Water ingestion during swimming activities in a pool: a pilot study. J. Water Health 4 (4), 425–430.
Ervin, J.S., Van De Werfhorst, L.C., Murray, J.L., Holden, P.A., 2014. Microbial source tracking in a coastal California watershed reveals canines as controllable sources of fecal contamination. Environ. Sci. Technol. 48 (16), 9043–9052.
Fuhrmeister, E.R., Ercumen, A., Pickering, A.J., Jeanis, K.M., Ahmed, M., Brown, S., Arnold, B.F., Hubbard, A.E., Alam, M., Sen, D., Islam, S., Kabir, M.H., Kwong, L.H., Islam, M., Unicomb, L., Rahman, M., Boehm, A.B., Luby, S.P., Colford, J.M., Nelson, K.L., 2019. Predictors of enteric pathogens in the domestic environment from human and animal sources in rural Bangladesh. Environ. Sci. Technol. 53 (17), 10023–10033.
Garcia-Armisen, T., Servais, P., 2004. Enumeration of viable *E. coli* in rivers and wastewaters by fluorescent in situ hybridization. J. Microbiol. Methods 58, 269–279.
Garner, E., McLain, J., Bowers, J., Engelthaler, D.M., Edwards, M.A., Pruden, A., 2018. Microbial ecology and water chemistry impact regrowth of opportunistic pathogens in full-scale reclaimed water distribution systems. Environ. Sci. Technol. 52 (16), 9056–9068.
Graczyk, T.K., Sunderland, D., Awantang, G.N., Mashinski, Y., Lucy, F.E., Graczyk, Z., Chomicz, L., Breysse, P.N., 2010. Relationships among bather density, levels of human waterborne pathogens, and fecal coliform counts in marine recreational beach water. Parasitol. Res. 106 (5), 1103–1108.
Greenberg, A., Clesceri, L., Eaton, A., 1992. Standard Methods for the Examination of Water and Wastewater. American Public Health Association, Washington DC, USA.
Hernroth, B.E., Conden-Hansson, A.C., Rehnstam-Holm, A.S., Girones, R., Allard, A.K., 2002. Environmental factors influencing human viral pathogens and their potential

indicator organisms in the blue mussel, *Mytilus edulis*: the first Scandinavian report. Appl. Environ. Microbiol. 68 (9), 4523–4533.
Ibekwe, A.M., Leddy, M., Murinda, S.E., 2013. Potential human pathogenic bacteria in a mixed urban watershed as revealed by pyrosequencing. PloS One 8, e79490.
Inoue, D., Hinoura, T., Suzuki, N., Pang, J., Malla, R., Shrestha, S., Chapagain, S.K., Matsuzawa, H., Nakamura, T., Tanaka, Y., Ike, M., Nishida, K., Ike, M., 2015. High-throughput DNA microarray detection of pathogenic bacteria in shallow well groundwater in the Kathmandu Valley, Nepal. Curr. Microbiol. 70 (1), 43–50.
Ishii, S., Segawa, T., Okabe, S., 2013. Simultaneous quantification of multiple food-and waterborne pathogens by use of microfluidic quantitative PCR. Appl. Environ. Microbiol. 79 (9), 2891–2898.
Kolde, R., 2013. Pheatmap: Pretty Heatmaps. R Package, version 0.7. 7.
Lanata, C.F., Fischer-Walker, C.L., Olascoaga, A.C., Torres, C.X., Aryee, M.J., Black, R.E., 2013. Global causes of diarrheal disease mortality in children <5 years of age: a systematic review. PloS One 8, e72788.
Li, B., Ju, F., Cai, L., Zhang, T., 2015. Profile and fate of bacterial pathogens in sewage treatment plants revealed by high-throughput metagenomic approach. Environ. Sci. Technol. 49 (17), 10492–10502.
Liu, J., Gratz, J., Maro, A., Kumburu, H., Kibiki, G., Taniuchi, M., Howlader, A.M., Sobuz, S.U., Haque, R., Talukder, K.A., Qureshi, S., Zaidi, A., Haverstick, D.M., Qureshi, S., 2012. Simultaneous detection of six diarrhea-causing bacterial pathogens with an in-house PCR-luminex assay. J. Clin. Microbiol. 50 (1), 98–103.
Locatelli, A., Spor, A., Jolivet, C., Piveteau, P., Hartmann, A., 2013. Biotic and abiotic soil properties influence survival of Listeria monocytogenes in soil. PloS One 8 (10), e75969.
Looft, T., Johnson, T.A., Allen, H.K., Bayles, D.O., Alt, D.P., Stedtfeld, R.D., Sul, W.Jun, Stedtfeld, T.M., Chai, B., Cole, J.R., Hashsham, S.A., Tiedje, J.M., Stanton, T.B., Hashsham, S.A., 2012. In-feed antibiotic effects on the swine intestinal microbiome. Proc. Natl. Acad. Sci. Unit. States Am. 109 (5), 1691–1696.
Luedtke, B.E., Bono, J.L., Bosilevac, J.M., 2014. Evaluation of real time PCR assays for the detection and enumeration of enterohemorrhagic *Escherichia coli* directly from cattle feces. J. Microbiol. Methods 105, 72–79.
Mathers, A.J., Cox, H.L., Kitchel, B., Bonatti, H., Brassinga, A.K.C., Carroll, J., Scheld, W. M., Hazen, K.C., Sifri, C.D., 2011. Molecular dissection of an outbreak of carbapenem-resistant *Enterobacteriaceae* reveals intergenus KPC carbapenemase transmission through a promiscuous plasmid. mBio 2 (6), e00204–e00211.
McLellan, S.L., Newton, R.J., Vandewalle, J.L., Shanks, O.C., Huse, S.M., Eren, A.M., Sogin, M.L., 2013. Sewage reflects the distribution of human faecal *Lachnospiraceae*. Environ Microbial 15 (8), 2213–2227.
Mora, F.X., Avilés-Reyes, R.X., Guerrero-Latorre, L., Fernández-Moreira, E., 2016. Atypical enteropathogenic *Escherichia coli* (aEPEC) in children under five years old with diarrhea in Quito (Ecuador). Int. Microbiol. 19, 157–160.
Munoz-Price, L.S., Poirel, L., Bonomo, R.A., Schwaber, M.J., Daikos, G.L., Cormican, M., Cornaglia, G., Garau, J., Gniadkowski, M., Hayden, M.K., Kumarasamy, K., Livermore, D.M., Maya, J.J., Nordmann, P., Patel, J.B., Paterson, D.L., Pitout, J., Villegas, M.V., Wang, H., Woodford, N., Kumarasamy, K., 2013. Clinical epidemiology of the global expansion of *Klebsiella pneumoniae* carbapenemases. Lancet Infect. Dis. 13 (9), 785–796.
Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., 2014. vegan: community Ecology Package. R package version 2.2-0. http://CRAN.R-project.org/package=vegan.
Pathogen Safety Data Sheets and Risk Assessment, 2014. Public Health Agency of Canada. http://www.phac-aspc.gc.ca/lab-bio/res/psds-ftss/index-eng.php.
Ramalingam, N., Rui, Z., Liu, H.B., DaiM, C.C., Kaushik, R., Ratnaharika, B., Gong, H.Q., 2010. Real-time PCR-based microfluidic array chip for simultaneous detection of multiple waterborne pathogens. Sensor. Actuator. B Chem. 145, 543–552.
Ramírez-Castillo, F., Loera-Muro, A., Jacques, M., Garneau, P., Avelar-González, F., Harel, J., Guerrero-Barrera, A., 2015. Waterborne pathogens: detection methods and challenges. Pathogens 4 (2), 307–334.
Raj, V.S., Smits, S.L., Pas, S.D., Provacia, L.B., Moorman-Roest, H., Osterhaus, A.D., Haagmans, B.L., 2012. Novel hepatitis E virus in ferrets, The Netherlands. Emerg. Infect. Dis. 18 (8), 1369.
RCoreTeam, 2014. A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org.
Richards, A.M., Von Dwingelo, J.E., Price, C.T., Abu Kwaik, Y., 2013. Cellular microbiology and molecular ecology of Legionella–amoeba interaction. Virulence 4 (4), 307–314.
Shannon, K.E., Lee, D.Y., Trevors, J.T., Beaudette, L.A., 2007. Application of real-time quantitative PCR for the detection of selected bacterial pathogens during municipal wastewater treatment. Sci. Total Environ. 382 (1), 121–129.
Sinigalliano, C.D., Fleisher, J.M., Gidley, M.L., Solo-Gabriele, H.M., Shibata, T., Plano, L. R., Elmir, S.M., Wanless, D., Bartkowiak, J., Boiteau, R., Withum, K., Abdelzaher, A. M., He, G., Ortega, C., Zhu, X., Wright, M.E., Kish, J., Hollenbeck, J., Scott, T., Backer, L.C., Withum, K., 2010. Traditional and molecular analyses for fecal indicator bacteria in non-point source subtropical recreational marine waters. Water Res. 44 (13), 3763–3772.
Skånseng, B., Kaldhusdal, M., Rudi, K., 2006. Comparison of chicken gut colonisation by the pathogens *Campylobacter jejuni* and *Clostridium perfringens* by real-time quantitative PCR. Mol. Cell. Probes 20 (5), 269–279.
Stedtfeld, R.D., Baushke, S.W., Tourlousse, D.M., Miller, S.M., Stedtfeld, T.M., Gulari, E., Gulari, E., Tiedje, J.M., Hashsham, S.A., 2008. Development and experimental validation of a predictive threshold cycle equation for quantification of virulence and marker genes by high-throughput nanoliter-volume PCR on the OpenArray platform. Appl. Environ. Microbiol. 74 (12), 3831–3838.
Stedtfeld, R.D., Liu, Y.C., Stedtfeld, T.M., Kostic, T., Kronlein, M., Srivannavit, O., Khalife, W.T., Tiedje, J.M., Gulari, E., Hughes, M., Etchebarne, B., Etchebarne, B., 2015. Static self-directed sample dispensing into a series of reaction wells on a microfluidic card for parallel genetic detection of microbial pathogens. Biomed. Microdevices 17 (5), 89.
Su, J.Q., An, X.L., Li, B., Chen, Q.L., Gillings, M.R., Chen, H., Zhang, T., Zhu, Y.G., 2017. Metagenomics of urban sewage identifies an extensively shared antibiotic resistome in China. Microbiome 5 (1), 84.
Tobe, T., Hayashi, T., Han, C.G., Schoolnik, G.K., Ohtsubo, E., Sasakawa, C., 1999. Complete DNA sequence and structural analysis of the enteropathogenic *Escherichia coli* adherence factor plasmid. Infect. Immun. 67 (10), 5455–5462.
Van Der Ley, P., Bekkers, A., Van Meersbergen, J., Tommassen, J., 1987. A comparative study on the *phoE* genes of three enterobacterial species: implications for structure-function relationships in a pore-forming protein of the outer membrane. Eur. J. Biochem. 164 (2), 469–475.
Walderhaug, M., 2014. Bad Bug Book: Foodborne Pathogenic Microorganisms and Natural Toxins Handbook. BrainFeed Press. http://www.fda.gov/Food/FoodSafety/FoodborneIllness/FoodborneIllnessFoodbornePathogensNaturalToxins/BadBugBook/default.htm.
Wilson, B.A., Salyers, A.A., 2003. Is the evolution of bacterial pathogens an out-of-body experience? Trends Microbiol. 11 (8), 347–350.
Wolf, S., Williamson, W.M., Hewitt, J., Rivera-Aban, M., Lin, S., Ball, A., Scholes, P., Greening, G.E., 2007. Sensitive multiplex real-time reverse transcription-PCR assay for the detection of human and animal noroviruses in clinical and environmental samples. Appl. Environ. Microbiol. 73 (17), 5464–5470.
World Health Organization, 2014. Preventing Diarrhoea through Better Water, Sanitation and Hygiene: Exposures and Impacts in Low-And Middle-Income Countries. WHO.
World Health Organization, 2017. Prioritization of Pathogens to Guide Discovery, Research and Development of New Antibiotics for Drug-Resistant Bacterial Infections, Including Tuberculosis. WHO (No. WHO/EMP/IAU/2017.12).
Zhao, Y., Su, J.Q., Ye, J., Rensing, C., Tardif, S., Zhu, Y.G., Brandt, K.K., 2018. AsChip: a high-throughput qPCR chip for comprehensive profiling of genes linked to microbial cycling of arsenic. Environ. Sci. Technol. 53 (2), 798–807.
Zheng, B., Zhu, Y., Sardans, J., Peñuelas, J., Su, J., 2018. QMEC: a tool for high-throughput quantitative assessment of microbial functional potential in C, N, P, and S biogeochemical cycling. Sci. China Life Sci. 61 (12), 1451–1462.
Zhuang, F.F., Li, H., Zhou, X.Y., Zhu, Y.G., Su, J.Q., 2017. Quantitative detection of fecal contamination with domestic poultry feces in environments in China. Amb. Express 7 (1), 80.