

RESEARCH ARTICLE

# A Comparison of Microbial Water Quality and Diversity for Ballast and Tropical Harbor Waters

Charmaine Ng<sup>1</sup>✉, Thai-Hoang Le<sup>1</sup>✉, Shin Giek Goh<sup>1</sup>, Liang Liang<sup>1</sup>, Yiseul Kim<sup>2</sup>, Joan B. Rose<sup>2</sup>, Karina Gin Yew-Hoong<sup>1,3\*</sup>

**1** National University of Singapore, Department of Civil and Environmental Engineering, Singapore, Singapore, **2** Michigan State University, Department of Microbiology and Molecular Genetics, East Lansing, Michigan, United States of America, **3** National University of Singapore Environmental Research Institute (NERI), Singapore, Singapore

✉ These authors contributed equally to this work.

\* [ceeginyh@nus.edu.sg](mailto:ceeginyh@nus.edu.sg)



**OPEN ACCESS**

**Citation:** Ng C, Le T-H, Goh SG, Liang L, Kim Y, Rose JB, et al. (2015) A Comparison of Microbial Water Quality and Diversity for Ballast and Tropical Harbor Waters. PLoS ONE 10(11): e0143123. doi:10.1371/journal.pone.0143123

**Editor:** Adelaide Almeida, University of Aveiro, PORTUGAL

**Received:** July 3, 2015

**Accepted:** October 31, 2015

**Published:** November 17, 2015

**Copyright:** © 2015 Ng et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All 16S rRNA gene sequence data files are available from the NCBI SRA the SRA archive with the accession number SRP065174.

**Funding:** Funding was provided by the Public Utilities Board (PUB) administered under the Environment & Water Industry Program Office Incentive for Research & Innovation Scheme (EWI IRIS) scheme (reference: 1102-IRIS-12-02), and administrative support from the National University of Singapore Environmental Institute (NERI).

## Abstract

Indicator organisms and antibiotic resistance were used as a proxy to measure microbial water quality of ballast tanks of ships, and surface waters in a tropical harbor. The survival of marine bacteria in ballast tanks appeared to diminish over longer water retention time, with a reduction of cell viability observed after a week based on heterotrophic plate counts. Pyrosequencing of 16S rRNA genes showed distinct differences in microbial composition of ballast and harbor waters. The harbor waters had a higher abundance of operational taxonomic units (OTUs) assigned to Cyanobacteria (*Synechococcus spp.*) and  $\alpha$ -proteobacteria (SAR11 members), while marine hydrocarbon degraders such as  $\gamma$ -proteobacteria (*Oceanothraustes spp.*, *Thiotrichales spp.*) and Bacteroidetes (*Flavobacteriales spp.*) dominated the ballast water samples. Screening of indicator organisms found *Escherichia coli* (*E. coli*), *Enterococcus* and *Pseudomonas aeruginosa* (*P. aeruginosa*) in two or more of the ballast and harbor water samples tested. *Vibrio spp.* and *Salmonella spp.* were detected exclusively in harbor water samples. Using quantitative PCR (qPCR), we screened for 13 antibiotic resistant gene (ARG) targets and found higher abundances of *sul1* (4.13–3.44 x 10<sup>2</sup> copies/mL), *dfra* (0.77–1.80 x 10<sup>2</sup> copies/mL) and *cfr* (2.00–5.21 copies/mL) genes compared to the other ARG targets selected for this survey. These genes encode for resistance to sulfonamides, trimethoprim and chloramphenicol-florfenicol antibiotics, which are also known to persist in sediments of aquaculture farms and coastal environments. Among the ARGs screened, we found significant correlations ( $P < 0.05$ ) between *ereA*, *ermG*, *cfr* and *tetO* genes to one or more of the indicator organisms detected in this study, which may suggest that these members contribute to the environmental resistome. This study provides a baseline water quality survey, quantitatively assessing indicators of antibiotic resistance, potentially pathogenic organisms and a broad-brush description of difference in microbial composition and diversity between open oceans and tropical coastal environments through the use of next generation sequencing technology.

**Competing Interests:** The authors have declared that no competing interests exist.

## Introduction

The city-state of Singapore has one of the busiest ports in the world receiving an average arrival of 140,000 vessels and 2.33 billion gross tonnes of ballast water per year [1]. These commercial ships use ballast water to provide stability and maneuverability, taking up waters when cargo is being unloaded and discharging ballast waters at ports of call [2]. This ballast water exchange can contain a diverse community of organisms, including invasive species (e.g. toxic dinoflagellates, *Vibrio spp.*) and potentially impose an environmental and health threat to coastal regions, where ballast waters are released [3–7]. Drake & Lodge (2004) identified regions in Southeast Asia, which may function as hotspots for invasion due to high shipping activity, which in turn facilitates microbial transport [8]. Asexual reproduction and dormancy are strategies, which aid the survival of non-indigenous invasive microbes under environmental stressed conditions. These traits help preserve invasive microbes in the nutrient depleted, light scarce environments of ballast water tanks, and upon discharge to coastal regions where conditions are more optimal, enables successful colonization by rapid population growth [2].

The deballasting of waters from ships acts as a vehicle for the global distribution of pathogens (and possibly antibiotic resistant forms) and waterborne diseases, which may have an adverse impact on humans, marine animals and the aquatic ecosystem as a whole [2]. Thomson et al. (2003) also showed that horizontal gene transfer of antibiotic resistant genes (ARGs) might occur due to the closed system and long water retention time within ballast tanks [9]. Cell death of antibiotic resistant bacteria (ARB) results in the release of intracellular ARGs, which may eventually be transferred into other viable cells within the microbial community imparting resistances in recipient cells. ARGs are increasingly being recognized as an emerging contaminant hence there is an urgency towards early detection in an effort to prevent widespread dissemination of antibiotic resistance [10].

In a study of bacteriological assessment of ballast waters of six ships docked at Singapore harbors, Joachimsthal et al. (2004) used fluorescence in situ hybridization (FISH) coupled with flow cytometer to enumerate total bacteria, and fluorescent tags to differentiate between Enterobacteria, *Vibrio spp.*, and *E. coli*. Using these techniques, their findings showed that Enterobacteria (range: 0– $1.45 \times 10^5$  cells/mL) and *Escherichia coli* (*E. coli*) (range: 0– $7.87 \times 10^5$  cells/mL) cells were not always present in all samples, but *Vibrio spp.* (range:  $1.45 \times 10^4$ – $1.07 \times 10^5$  cells/mL) cells were [11, 12].

In this study, we investigated the microbial content of mid-ocean exchanged (MOE) ballast waters obtained from grab samples from ballast tanks calling at the port of Singapore, and harbor waters surrounding these ships to evaluate differences in microbial diversity between sample types. We expected the ballast waters to represent the microbiological content of the waters at the sites of ballast water uptake, taking into consideration the possibility of the contribution of biofilms, which may develop on the surfaces of ballast water tanks during the voyage. Collectively, we applied pyrosequencing of tagged 16S rRNA gene amplicons as a semi-quantitative method to elucidate differences in microbial diversity, and a culturing and quantitative PCR (qPCR) based approach to survey for the presence of indicator organisms, ARB and associated resistance genes. Among known invasive species, we were interested in the pathogens rather than phytoplankton. Specifically, we focused on monitoring known opportunistic pathogens including *E. coli*, *Enterococcus spp.*, *Vibrio spp.*, and *Pseudomonas aeruginosa* to determine if levels measured complied with the International Maritime Organization (IMO) guidelines.

## Materials and Methods

### Sampling of ballast and harbor waters

Ballast water samples were collected from three vessels with voyages within the Middle East and Asian regions. Each ballast water sample was acquired in the morning, on-board the docked vessel at the Port of Singapore. Seawater next to the hull of the docked ships was collected on the same morning of ballast water sampling for comparative analysis. Ballast waters 1, 2 and 3, (abbreviated as BW1, BW2 and BW3) were taken from the South China Sea, Gulf of Aden and to the west of South Korea in the East China Sea accordingly. The retention time of BW1, 2 and 3 (from the time of collection from the source ports to docking at the Port of Singapore) were 4, 107 and 11 days, respectively. Harbor waters corresponding to BW1, BW2 and BW3 were abbreviated as HW1, HW2 and HW3. Permission for water sampling was granted by the Maritime Port Authority of Singapore and did not involve endangered and protected species.

Environmental parameters including temperature, pH, salinity and turbidity were measured on site using a YSI probe, model 63 (YSI incorporated, Yellow Springs, Ohio, USA). Using a bucket, approximately 60 L of surface water was hauled up through a manhole leading to the ballast tank. Likewise, 60 L of surface water surrounding the docked vessels was sampled. The water samples were concentrated through hollow-fiber ultrafiltration (Hemodialyzer Rexeed<sup>TM</sup>-25S, Asahi Kasei Medical Co., Ltd) to a final volume of 500 mL according to the method of Hill et al. 2005. A volume of 20 mL of concentrated water samples (equivalent to 2.4 L of original water sample), were used for total DNA extraction. Whole water samples were taken for culturing work.

### DNA extraction

The bacterial DNA from the concentrated samples was extracted with the PowerWater DNA Isolation kit (Mo Bio Laboratories, Inc, Carlsbad, CA). DNA yields and purity were measured using a nanodrop spectrophotometer (NanoDrop 2000, Thermo Scientific, Wilmington, USA). The extracted DNA was used for quantification of ARG and *P. aeruginosa* and *Salmonella* species.

### Detection of ARG by PCR

A total of 14 primer pairs were used to detect the presence of ARGs, with one primer pair targeting 16S rRNA gene sequences to detect the total bacterial population (Table 1). These primers were also designed to quantify the target genes in real time or quantitative Polymerase Chain Reactions (PCR). Among these, 10 primers were selected from published work, reporting a good detection and quantification of relevant ARGs in qPCR, while 4 primers were designed in this study to target and quantify the specific ARGs, i.e. *ereA*, *ereB*, *sul1*, and *cfr* genes. These gene sequences were downloaded from the GeneBank Database (<http://www.ncbi.nlm.nih.gov/>) and primers were designed using PRIMER 3 tool [13]. PCR amplicon sizes were specified in the range of 100–300 base pairs (bp) for qPCR analyses.

All PCR assays were performed on a Veriti 96-Well Fast Thermal Cycler (Applied Biosystem, CA, USA). Each PCR reaction included 2  $\mu$ L 10 $\times$  Faststart PCR buffer (Roche, IN, USA), 0.25 mM dNTPs (Promega, Singapore), 0.0375 U/ $\mu$ L Taq Faststart, DNA polymerase (Roche, IN, USA), 0.5 mM forward and reverse primers each (AITBiotech, Singapore), 20–200 ng genomic DNA template, with the final volume adjusted to 20  $\mu$ L using diethylpyrocarbonate (DEPC)-treated distilled deionized water (PureLab Option, Elga, Singapore). The thermo cycles for PCR amplifications were carried out under the following settings: 95°C for 5 min;

**Table 1. Primer sets used in the detection of ARGs and indicator organisms in ballast and harbor water samples.**

Name	Sequence of forward and reverse primers (5'-3')	Target gene/species	Amplicon size (bp)	Annealing temperature (°C)	References
<i>16S EUB-f</i>	ACTCCTACGGGAGGCAGCAG	16s ribosomal RNA/all bacteria	180	64	[20, 21]
<i>16S EUB-r</i>	ATTACCGCGGCTGCTGG				
<i>ereA-f</i>	CTTCACATCCGGATTGCTCGA	erythromycin esterase type I/all bacteria	296	60	In this study
<i>ereA-r</i>	ATGGACGCCAACAAGTGAGT				
<i>ereB-f</i>	GAATGCAGGGCGATATGGGT	erythromycin esterase type II/all bacteria	171	60	In this study
<i>ereB-r</i>	TTTGCCCATTTGGTAGGCAA				
<i>ermA-f</i>	CAAGAACAATCAATACAGAGTCTAC	erythromycin resistance methylase/all bacteria	157	60	[22]
<i>ermA-r</i>	AGTCAGGCTAAATATAGCTATC				
<i>ermB-f</i>	GGTTGCTCTTGACACTCAAG	erythromycin resistance methylase/all bacteria	191	62	[22]
<i>ermB-r</i>	CAGTTGACGATATTCTCGATTG				
<i>ermC-f</i>	AATCGTGGAAATACGGGTTTGC	erythromycin resistance methylase/all bacteria	293	62	[22]
<i>ermC-r</i>	CGTCAATTCTGCATGTTTTAAGG				
<i>ermF1-f</i>	TCGTTTTACGGGTCAGCACTT	erythromycin resistance methylase/all bacteria	182	62	[23]
<i>ermF1-r</i>	CAACCAAAGCTGTGTCGTTT				
<i>ermG-f</i>	GTGAGGTAACCTGTAATAAGCT	erythromycin resistance methylase/all bacteria	255	61	[22]
<i>ermG-r</i>	CCTCTGCCATTAACAGCAATG		255		[22]
<i>ermT1-f</i>	GTTCACTAGCACTATTTTTAATGACAGAAGT	erythromycin resistance methylase/all bacteria	124	60	[24]
<i>ermT1-r</i>	GAAGGGTGTCTTTTTAATACAATTAACGA				
<i>sul1-f</i>	CACCGTTGGCCTTCCTGTAA	sulfonamide resistance gene type I/all bacteria	180	61	In this study
<i>sul1-r</i>	TCTAACCCCTCGGTCTCTGGC				
<i>sul2-f</i>	GCGCTCAAGGCAGATGGCATT	sulfonamide resistance gene type II/all bacteria	293	60	[25]
<i>sul2-r</i>	GCGTTTGATACCGGCACCCGT				
<i>dfrA1-f</i>	ACGGATCCTGGCTGTTGGTTGGACGC	trimethoprim resistance gene/all bacteria	237	60	[26]
<i>dfrA1-r</i>	CGGAATTCACCTCCGGCTCGATGTC				
<i>tetM-f</i>	CC[TA]AC[AT]GTCATTTATATGGA[GA]AGACC	tetracycline resistance gene/all bacteria	304	62.5	[27]
<i>tetM-r</i>	CGAAAATCTGCTGG[CGA]GTA[GA]ACAGGGC				
<i>tetO-f</i>	AAGAAAACAGGAGATTCCAAAACG	tetracycline resistance gene/all bacteria	75	60	[28]
<i>tetO-r</i>	CGAGTCCCCAGATTGTTTTTAGC				
<i>cfr-f</i>	GCTCTAGCCAACCGTCAAGT	Linezolid/chloramphenicol resistance gene/all bacteria	195	61	In this study
<i>cfr-r</i>	TCAATTTGCTGCGTTCCTCAC				
<i>tdh-f (toxin)</i>	GTAAGGTCTCTGACTTTTGGAC	<i>tdh (toxin)/V. parahaemolyticus</i>	269	60	[29, 30]
<i>tdh-r (toxin)</i>	TGGAATAGAACCCTTCATCTTCACC				
<i>cth-f</i>	TTCCAACCTCAAACCGAACTATGAC	<i>cth/V. vulnificus</i>	205	65	[29, 31]
<i>cth-r</i>	ATCCAGTCGATGCGAATACGTTG				

(Continued)

Table 1. (Continued)

Name	Sequence of forward and reverse primers (5'-3')	Target gene/species	Amplicon size (bp)	Annealing temperature (°C)	References
16S-23S rDNA ISR-f	TTAAGCGTTTTTCGCTGAGAATG	16S-23S rDNA ISR/ <i>V. cholerae</i>	295	60	[29, 32, 33]
16S-23S rDNA ISR-r	AGTCACTTAACCATACAACCCG				
<i>regA-f</i>	TGCTGGTGGCACAGGACAT	<i>regA/P. aeruginosa</i>	65	60	[34]
<i>regA-r</i>	TTGTTGGTGCAGTTCCTCATTG				
<i>regA-p</i>	FAM-CAGATGCTTTGCCTCAA-BHQ				
<i>invA-f</i>	CGTTTCCTGCGGTACTGTTAATT	<i>invA/Salmonella spp.</i>	67	60	[34]
<i>invA-r</i>	AGACGGCTGGTACTGATCGATAA				
<i>invA-p</i>	FAM-CCACGCTCTTTCGTCT-BHQ				

doi:10.1371/journal.pone.0143123.t001

repeated 40 times with a cycle of denaturation at 95°C for 30 s, at annealing temperatures specific for each primer set for 30 s, elongation at 72°C for 30 s, and finally 72°C for 10 min [14]. ARGs were chemically synthesized (AITBiotech, Singapore) and the gene products served as a positive control. For a negative control, the genomic DNA of a susceptible *E. coli* strain (not containing these ARGs) was used. PCR products were visualized using DNA electrophoresis on a 2% agarose gel stained with GelRed™ (Biotium, Inc., Hayward, CA). The gel was visualized under UV light with an E-Gel® Imager System (Life Technologies, Carlsbad, CA).

### Quantification of 16S rRNA genes and ARGs by Real-time qPCR

The calibration curves for qPCR quantification for 16S rRNA genes and each ARG were built using the ten-fold dilution series of ARG standard DNA. Triplicate samples were used for each dilution. Each qPCR reaction contained 10 µL SYBR Select Master Mix (Applied Biosystem, CA, USA), 0.25 µL 10 mM forward primers and reverse primers each, 2 µL DNA template, and 7.6 µL DEPC-treated distilled deionized water. qPCR reactions were performed in a StepOne-Plus Real-Time PCR system (Applied Biosystem, CA, USA) with the following settings: holding at 50°C for 2 min and 95°C for 2 min; 40 cycle amplification under conditions of 95°C for 15 s with annealing and elongating at 60°C for 1 min, and finally, a melting curve analysis at 95°C for 15 s, 60°C for 1 min, and 95°C for 15 s. The amplification profiles for each series of DNA dilutions during 40 cycles were recorded. A threshold cycle (Ct) value was determined as the cycle number where fluorescence data crossed the threshold line, which is manually defined within the logarithmic increase phase. The gene copies of 16S rRNA genes and ARGs were calculated as previously described [15]. The molecular weight of a gene was determined by multiplying the size of a gene in base pairs and the average molecular weight of a double-stranded DNA base pair (649 Da). Absolute gene abundance was calculated as gene copies normalized to the volume of ballast and harbor water samples, and relative gene abundance was calculated as percentage of ARGs normalized to the abundance of 16S rRNA gene.

### Determining the phenotypic concentrations of ARB

Marine agar (BD Diagnostics, Maryland, USA) was used to estimate the concentrations of viable and culturable heterotrophic bacteria. Marine agar plates were supplemented with antibiotics to determine the concentrations of ARB. The antibiotics (purchased from Sigma-Aldrich, USA) of the following concentrations were used in the screening process; trimethoprim (4 µg/mL), sulphanilamide (300 µg/mL), norfloxacin (10 µg/mL), lincomycin (100 µg/mL), tetracycline

(35 µg/mL), erythromycin (100 µg/mL), kanamycin (80 µg/mL), and linezolid (8 µg/mL). Concentrations used for screening were higher than published minimum inhibitory concentration (MIC) breakpoints in the 2013 Clinical and Laboratory Standards Institute (CLSI) performance standards for Antimicrobial Susceptibility Testing report. Raw water samples were serially diluted in M9 minimal media and 100 µL of samples were spread plated and incubated at 35°C for 24 h and subsequently at room temperature for an additional 24 h. Plates containing colonies within the range of 30–100 cells were counted.

### Measuring indicator organisms using IDEXX kits and qPCR

Colilert™ and Enterolert™ (IDEXX Laboratories, Inc., Westbrook, Maine, USA) were used to quantify the cell numbers of *E. coli* and *Enterococcus*, respectively. 100 mL of raw water and a 10 fold diluted raw water sample were tested to allow a wider range of detection due to high fluctuations in bacterial concentration. The water sample was mixed with the Colilert™ or Enterolert™ reagent in a sterile bag and poured into a multi-well tray (Quanti-Tray/2000). The results were expressed as the most probable number (MPN) according to the presence of fluorescence in individual wells when visualized under long-wave ultraviolet light (365 nm) after 24 h of incubation at 37°C (for Colilert™) and 41°C (for Enterolert™).

### Enumeration of pathogenic bacteria

*Vibrio spp.* were enumerated according to the method of Dickinson et al. (2013) with some modifications. Briefly, ballast and harbor water samples at different dilutions (0.1mL, 1mL, 10mL and 100mL) were used to determine the most probable number (MPN) by enrichment in alkaline peptone water (Sigma-Aldrich, St Louis, MO). The samples were incubated overnight at 37°C. A volume of 1 mL of enriched sample was centrifuged at 10 000 g for 10 min to pellet the culture, which was subsequently resuspended in 200 µL of nuclease-free water and heated to boiling for 10 min to release the DNA for the PCR analysis. The primers used for the detection of *Vibrio spp.* are listed in Table 1. PCR was carried out in a thermal cycler (Eppendorf, Germany) and performed in 25 µL volumes consisting 12.5 µL of 2x GoTaq® master mix (Promega, Madison, WI), 1.25 µL of each primer (10 µM), 7.5 µL of nuclease-free water and 2.5 µL of extracted DNA from the *Vibrio spp.* cultures. PCR annealing temperatures and durations for different target *Vibrio spp.* are shown in Table 1. The conditions for PCR amplification was an initial denaturing cycle at 94°C for 2 min, followed by 30 cycles of denaturation at 94°C for 30 s, at an annealing temperature as stated in Table 1 for 1 min and extension at 72°C for 1 min. A final extension step at 72°C for 5 min was also included at the end of 30 cycles. The amplified products were separated by electrophoresis on a 1.5% agarose gel stained with GelRed™ (Biotium, Inc., Hayward, CA). The separated fragments on the agarose gel were visualized under UV light with the E-Gel® Imager System (Life Technologies, Carlsbad, CA).

*P. aeruginosa* and *Salmonella spp.* were measured with real-time qPCR method. Primer and probe sets for both *P. aeruginosa* and *Salmonella spp.* are summarized in Table 1. A 20 µL qPCR reaction was prepared which contained 10 µL of FastStart Universal Probe Master (ROX) (Roche, Germany), 2 µL of forward and reverse primer each, 0.5 µL of probe, 0.5 µL of nuclease-free reagent-grade water and 5 µL of DNA template. Each set of reactions included both positive (recombinant plasmid DNA) and negative controls (double-deionized water). qPCR amplification for both targets was performed on StepOnePlus™ Real-time PCR system at 50°C for 2 min, 95°C for 10 min, and 45 cycles of 95°C for 15 s and 60°C for 1 min. The threshold cycle values (Ct) obtained from qPCR analysis were converted to gene copies (GC) based on the equations generated from standard curves [*Salmonella spp.*:  $Ct = -3.4651 \log(GC/\mu l) + 40.416$ ; *P. aeruginosa*:  $Ct = -3.341 \log(GC/\mu l) + 42.466$ ].

## 16S rRNA tag encoded gene sequencing and data analysis

To determine microbial community composition, bacterial tag-encoded FLX amplicon pyrosequencing (bTEFAP) using primers 515F (5'-CACGGTTCGKCGGCCATT-3') and 806R (5'-GGACTACHVGGGTWTCTAAT-3') targeting the V4 region of the 16S rRNA gene was performed using the Roche 454 FLX Titanium platform [16], at the Molecular Research LP (Texas, USA). Datasets obtained from the analyses were processed with using version 1.33.3 MOTHUR [17]. Sequences were trimmed to remove barcodes and primers, quality checked to have an average Phred quality score of 35 or above, with sequences exceeding a homopolymer stretch of more than 8 bases being excluded from the dataset. Sequences were aligned by MOTHUR against the SILVA bacterial reference database (14,956 sequences), screened, filtered and de-noised. UCHIME was used to identify and remove chimeric sequences. Datasets for each sample were normalized by subsampling each dataset to 3,230 reads [18]. Operational taxonomic units (OTUs) were assigned at a distance of 3% using an average neighbor-clustering algorithm. Community diversity between samples was assessed using the observed number of OTUs, Shannon and Simpson indices. Rarefaction curves were generated using MOTHUR and taxonomy of OTUs were assigned using the SILVA database with abundances expressed as percentage of sequences assigned to specific taxa at the phyla, class and family level over the total number of sequenced reads for each sample.

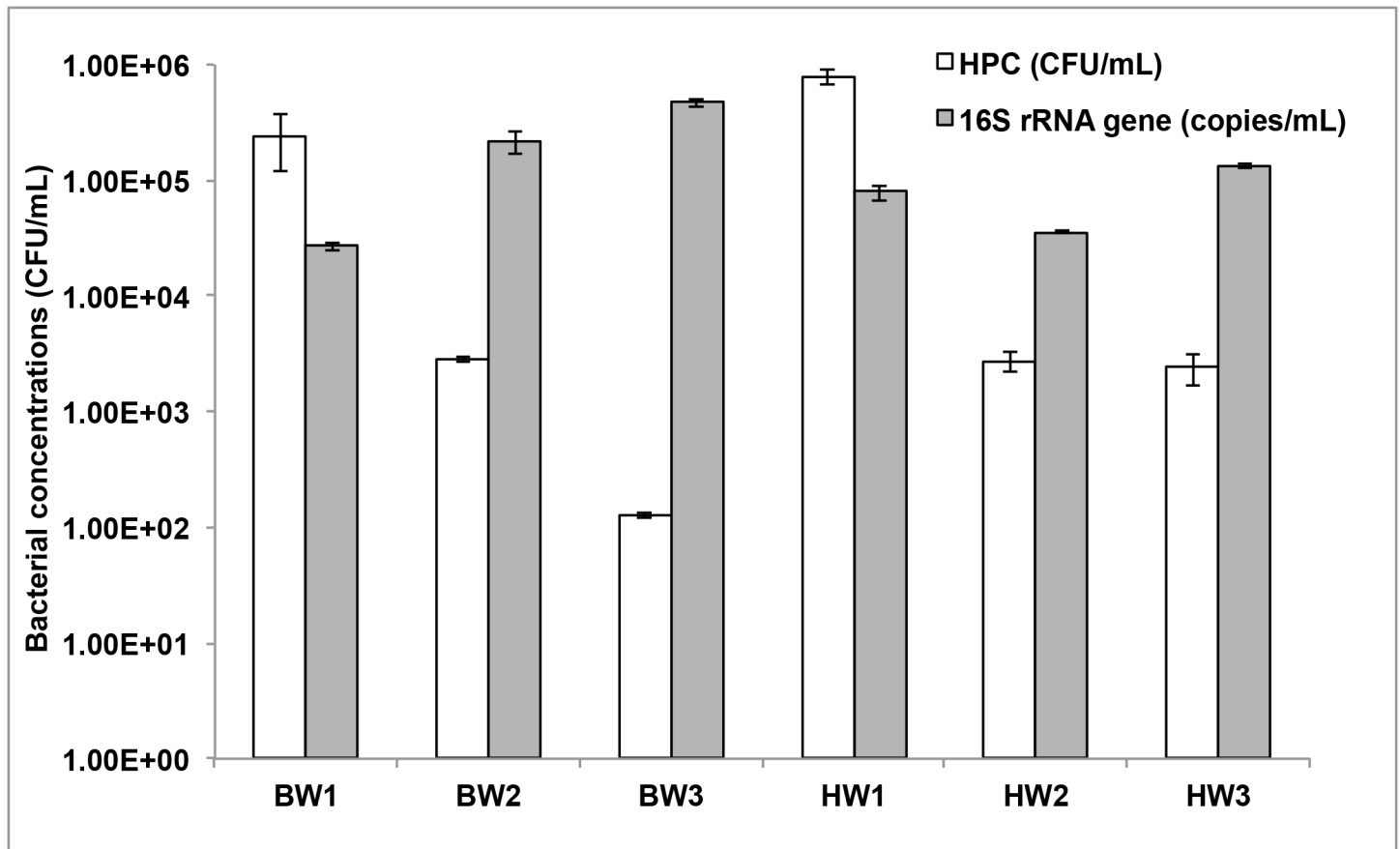
## Statistical analysis

Environmental and biological datasets were analyzed using the statistical software package Primer version 6 [19]. A square-root transformation was applied to the OTU abundance, ARB, ARG, indicator organisms, pathogenic bacteria and environmental datasets. Bray-Curtis similarity values were calculated for the OTU dataset to determine the extent of similarity of microbial community composition between samples and visualized as a dendrogram. A SIMPER analysis was used to determine the taxa, which defined difference in the harbor and ballast water samples. An ANOSIM test was used to verify the significance of observed differences in the microbial community structure. A Principal component analysis (PCA) was used to determine the relationship between the microbial community structure (based on bray-curtis similarity indices), and other biological and environmental parameters. Only parameters that had a Pearson correlation coefficient of  $>0.7$ , with a significance of  $<0.05$  were represented in the plot.

## Results

### Bacterial cell and gene counts

To estimate the concentrations of viable bacterial cells in the two sample types, whole water fractions were plated on marine agar. The average ( $n = 3$ ) heterotrophic plate counts (HPC) in ballast waters ( $4.31 \times 10^3$  CFU/mL) was approximately a magnitude lower than harbor waters ( $1.06 \times 10^4$  CFU/mL, Fig 1). When determining concentrations by qPCR for total bacterial populations including both the culturable and non-culturable bacteria, the average 16S rRNA gene copies in ballast samples ( $1.40 \times 10^5$  CFU/mL) were higher compared to the harbor waters ( $7.24 \times 10^4$  CFU/mL). Discrepancies in the counts of viable HPC and total bacterial populations within ballast water samples could be due to long retention time and thus, nutrient depletion and reduction of viable cells in the ballast tank over the voyage. This is evident in BW2 and BW3 samples where the ballast waters had a longer retention time (107 and 11 days, respectively), and showed higher 16S rRNA gene copy concentrations and lower HPC compared to BW1 samples, which were held in the ballast water tanks for 4 days. Overall, two of the ballast



**Fig 1. Concentrations of culturable bacteria and total 16S rRNA genes in ballast and harbor waters.**

doi:10.1371/journal.pone.0143123.g001

waters in our study had significantly lower concentrations of viable HPC than harbor waters (with exceptions of BW2 and HW2).

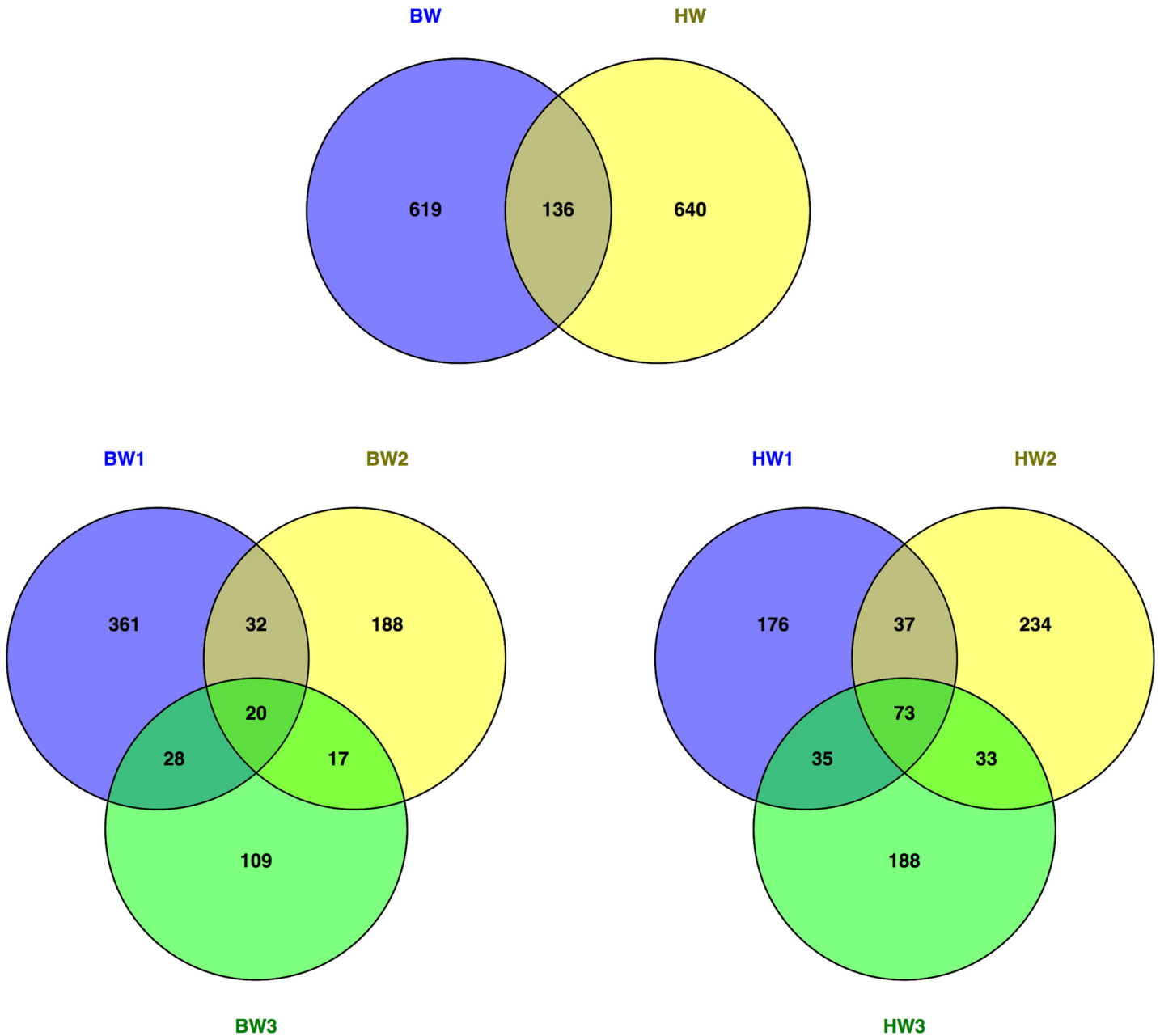
### Microbial diversity of ballast and harbor waters

A total of 57,576 sequences assigned to 1,395 unique OTUs were identified at the genus level (97% similarity cut off) from all samples. Sequences were deposited into the NCBI SRA archive website (<http://www.ncbi.nlm.nih.gov/>) and available under SRA accession number SRP065174. There were 136 OTUs shared by microbial communities, with 619 and 640 OTUs exclusive to the ballast and harbor communities, respectively (Fig 2). Within the ballast and harbor samples, the harbor waters were more similar in microbial composition (73 common OTUs) in comparison to the ballast waters (20 common OTUs).

Rarefaction curves indicated that a sufficient number of sequences were sampled to cover the majority of species in the microbial communities (S1 Fig). Shannon and Simpson indices showed that BW1 had highest species diversity (Shannon= 4.315, Simpson= 0.9501) and richness (441 OTUs) compared to the two ballast water samples which had values lower than those of the HW samples (Table 2). In general, the HW samples appeared to be more diverse than the BW samples with the exception of BW1, which was taken up at from the South China Sea.

Dominant taxa representing an average of >1% of total assigned sequences across all sample types included Proteobacteria (49.36 ± 22.54), Cyanobacteria (25.34 ± 27.84), Bacteroidetes (11.78 ± 6.25), Planctomyces (4.43 ± 2.93), Chloroflexi (3.39 ± 7.78), Actinobacteria





**Fig 2. Total OTUs identified in ballast (BW) and harbor waters (HW) from community sequencing analysis.**

doi:10.1371/journal.pone.0143123.g002

( $1.80 \pm 0.88$ ) and Aquificae ( $1.59 \pm 2.32$ ) (Fig 3). A SIMPER analysis showed an average dissimilarity of 52% between the ballast and harbor water sample types. The harbor waters were enriched in Cyanobacteria sequences (~30 fold) while ballast waters had more Proteobacteria (~2 fold), Bacteroidetes (~2 fold) and Chloroflexi (~18 fold) sequence. A SIMPER analysis showed that these taxa accounted for 90% of the total dissimilarity between the microbial communities of the two sample types.

At the family level, among the dominant phyla, Sphingomonadales, Oceanospirillales, Thio-trichales and Flavobacteriales were more represented in the ballast waters (Table 3). The higher relative abundance of Chloroflexi differentiated BW1 among the other water samples, and

Table 2. Diversity of microbial community composition between samples.

Sample	No. of OTUs	No. of sequences	Shannon diversity	Simpson index
BW1	441	3230	4.315	0.9501
BW2	257	3230	2.752	0.7060
BW3	174	3230	2.670	0.8153
HW1	321	3230	3.522	0.9076
HW2	377	3230	3.590	0.9032
HW3	329	3230	3.984	0.9499

doi:10.1371/journal.pone.0143123.t002

deeper inspection of the sequence reads showed that Caldilineales was the main contributor to the abundance. One of the striking differences between the ballast and harbor water sample types was the prevalence of SAR11 and *Synechococcus spp.* in the harbor waters.

### Quantification of indicator organisms

*Enterococcus* and *E.coli* counts were higher in harbor water samples, with the exception of BW3. Viable indicator organisms were detected in the harbor water samples and together with BW3 (from the Gulf of Aden) exceeded the acceptable levels of (100 CFU/100 mL) of intestinal

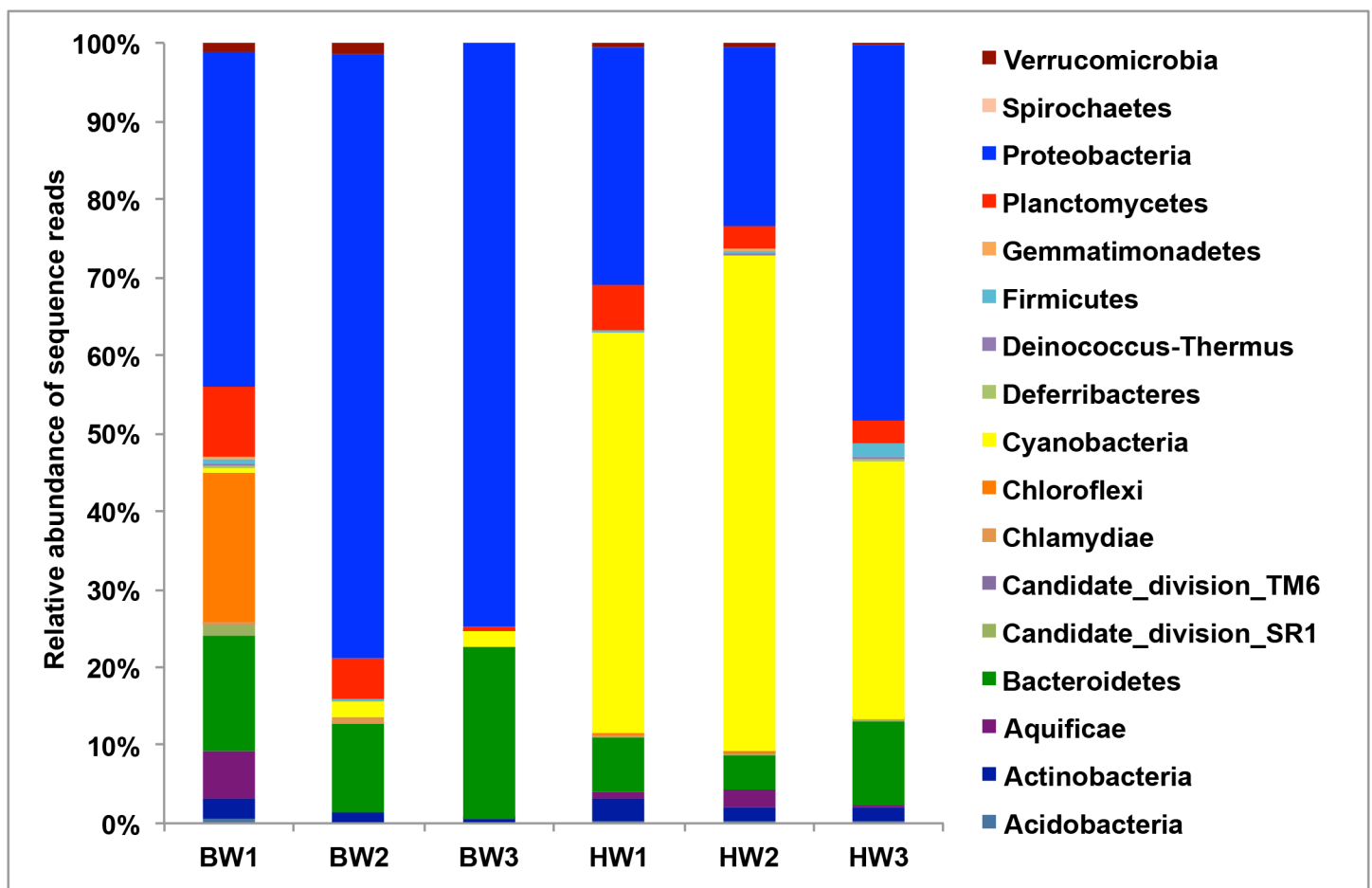


Fig 3. Relative abundance of total sequences assigned at the phyla level.

doi:10.1371/journal.pone.0143123.g003

**Table 3. Percentage of total sequences assigned to most dominant taxa. Data presented include the top 4 abundant phyla/classes across samples.**

Phyla/Class	Family/Genera	BW1	BW2	BW3	HW1	HW2	HW3
α-proteobacteria	Sphingomonadales	7.99	5.20	1.42	0.06	1.27	3.41
	SAR11	1.89	0.00	2.41	16.78	8.14	6.01
	Rhodobacterales	1.39	3.37	1.86	0.50	1.21	4.80
γ-proteobacteria	Alteromonadales	5.14	2.57	3.96	2.51	4.86	16.28
	Oceanospirillales	13.10	1.67	25.11	2.11	1.46	3.59
	Pseudomonadales	0.28	0.06	0.09	2.35	0.40	2.85
	Thiotrichales	1.49	54.46	36.16	0.09	0.09	0.09
Bacteroidetes	Flavobacteriales	3.62	10.68	21.20	6.16	3.38	10.40
	Sphingobacteriales	11.18	0.71	0.93	0.62	1.08	0.40
Chloroflexi	Caldilineales	18.95	0.00	0.00	0.25	0.25	0.03
Cyanobacteria	Synechococcus	0.03	0.06	1.18	42.94	46.01	29.07
	Unclassified	0.37	1.83	0.06	7.89	10.65	3.59

Family/genera representing >1% of total sequences assigned were selected for comparison.

doi:10.1371/journal.pone.0143123.t003

*Enterococci* according to IMO guidelines. *E. coli* was detected in BW1 but at levels within the acceptable range. Indicator organisms were present in all the harbor water samples at levels higher than those in the ballast samples. *Vibrio* species (including toxin-producing species) were below the detection limit in all ballast water samples, but were detected in two of the harbor water samples (HW1 and HW2, Table 4). *V. parahaemolyticus* was absent in all samples, however *V. cholerae* and *V. vulnificus* were detected in two of the harbor water samples. MPN-PCR results showed that BW1 and BW2 were positive for *P. aeruginosa* with a relatively low concentration of 19–20 MPN/100 mL in comparison to two harbor samples, HW1 and HW2, which had much higher concentrations (1,322–19,393 MPN/100 mL, Table 4). Higher levels of *P. aeruginosa* gene copies were detected in two of the harbor water samples (HW1 and HW2) with lower levels (100–1,000 fold less) in the ballast waters (BW1 and BW2, Table 4).

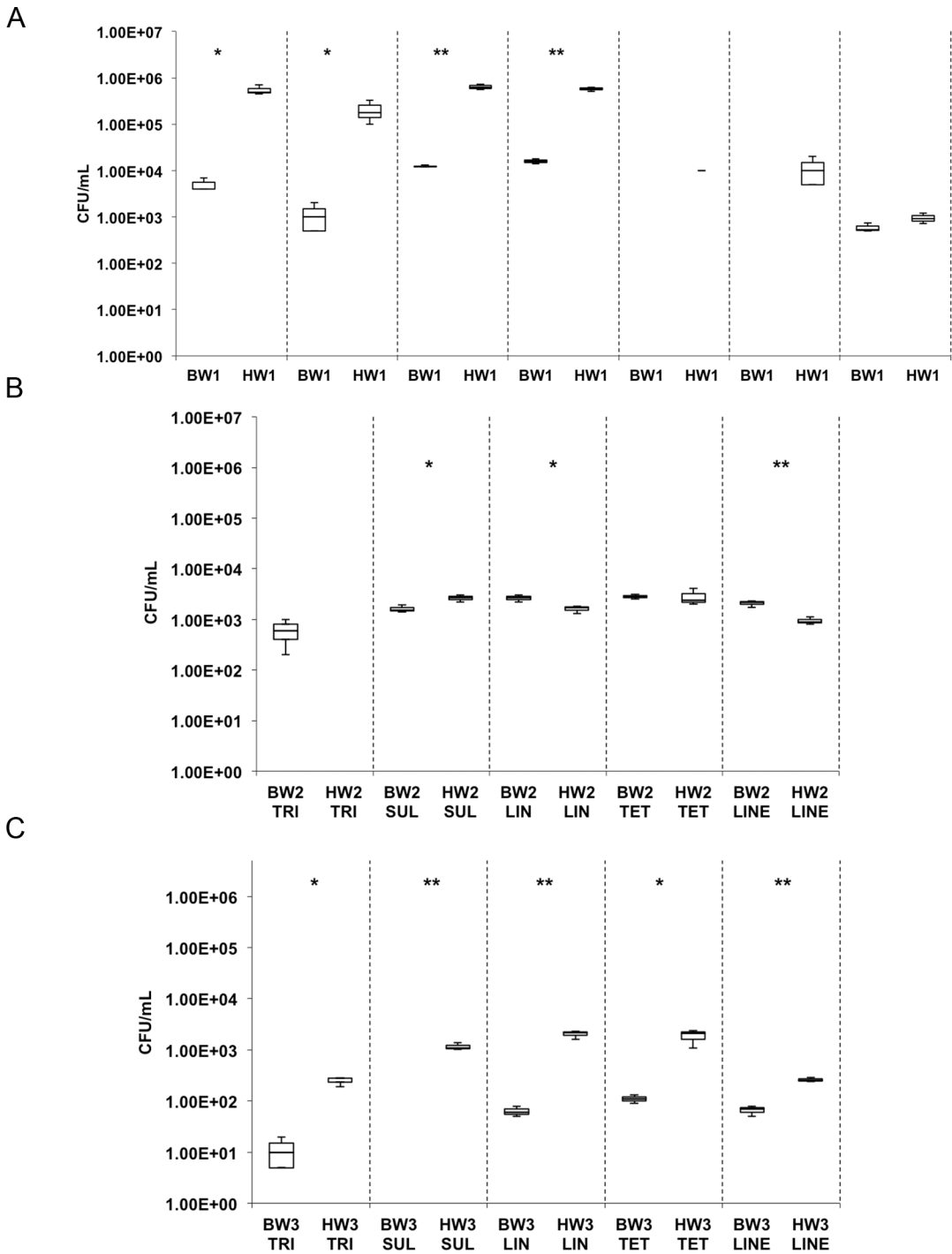
### Detection of ARBs and ARGs

There was no colony growth on norfloxacin-supplemented plates across all samples. BW1 and HW1 samples showed resistance to the remaining 7 antibiotics while BW2 and BW3 as well as HW2 and HW3 showed resistance to 5 antibiotics (Fig 4). In most cases, the ballast water samples had significantly lower counts of ARB compared to the harbor water samples (Fig 4). Among the harbor water samples, HW1 had the highest resistant counts to trimethoprim, sulfanilamide and lincomycin (1.8–7.0 x 10<sup>5</sup> CFU/mL).

**Table 4. Indicator organism measurements in ballast and harbor waters using IDEXX kits and qPCR methods.**

Method of detection	Target	BW1	BW2	BW3	HW1	HW2	HW3
IDEXX kit (MPN/100mL)	<i>E. coli</i>	140	6	0	2,603	893	471
IDEXX kit (MPN/100mL)	<i>Enterococcus</i>	45	51	251	1,908	222	135
PCR (MPN/100mL)	<i>V. cholera</i>	0	0	0	110	24	0
PCR (MPN/100mL)	<i>V. parahaemolyticus</i>	0	0	0	0	0	0
PCR(MPN/100mL)	<i>V. vulnificus</i>	0	0	0	110	110	0
qPCR(GC/100mL)	<i>P. aeruginosa</i>	19	20	0	1,322	19,393	19
qPCR (GC/100mL)	<i>Salmonella spp.</i>	0	0	0	17	0	0

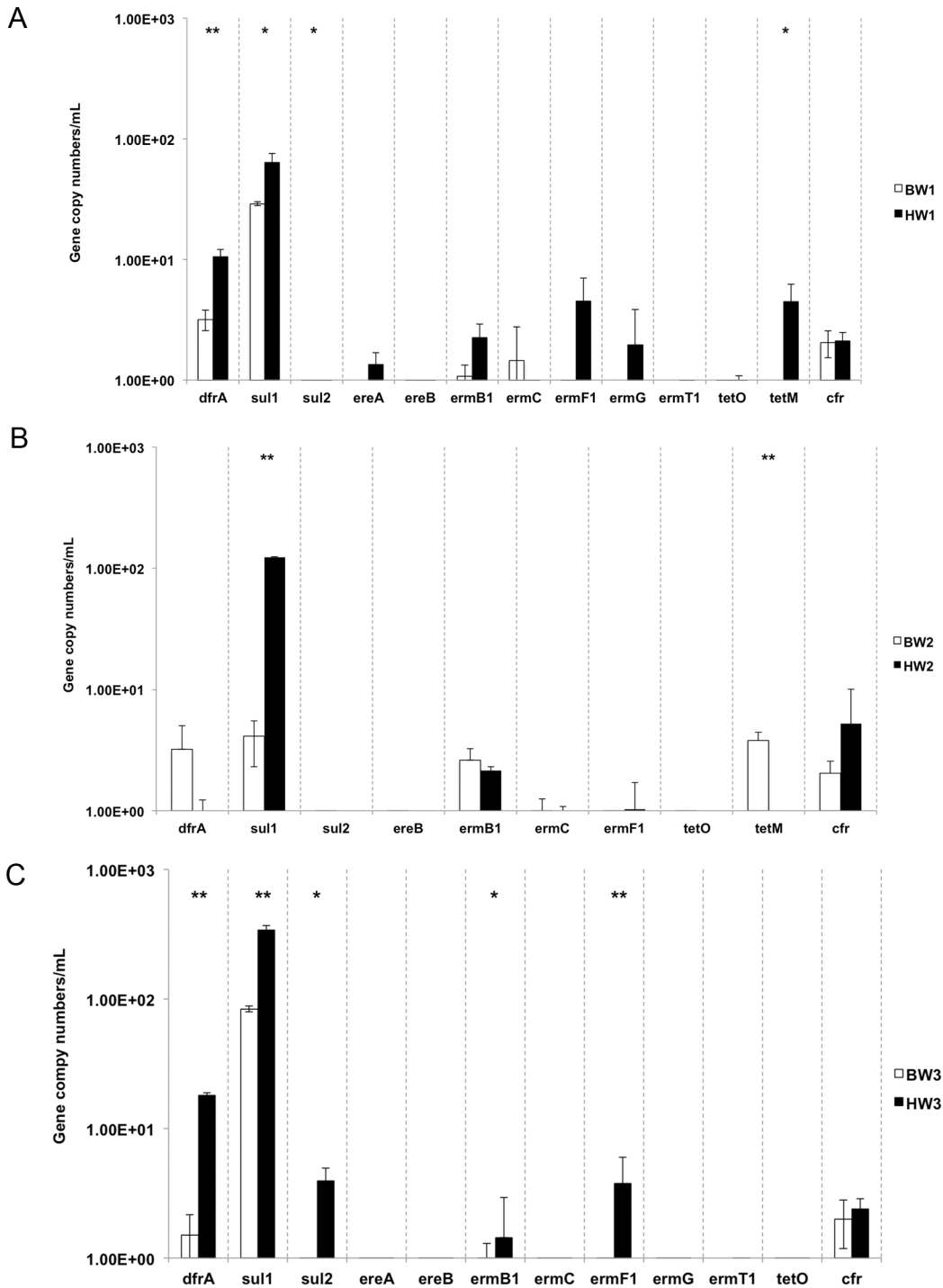
doi:10.1371/journal.pone.0143123.t004



**Fig 4. The concentrations of heterotrophic bacteria and their resistance to 8 antibiotics in ballast and harbor waters (A) BW1 and HW1, (B) BW2 and HW2, (C) BW3 and HW3.** TRI: Trimethoprim; SUL: Sulfanilamide; NOR: Norfloxacin; LIN: Lincomycin; TET: Tetracycline; ERY: Erythromycin; KAN: Kanamycin; LINE: Linezolid. Two-tailed T-test, \* $P < 0.05$ , \*\* $P < 0.01$ .

doi:10.1371/journal.pone.0143123.g004

Genes linked to resistance of sulfanamide (*sul1*), trimethoprim (*dfrA*) and chloramphenicol-florfenicol (*cfr*) were most abundant across all samples with levels ranging from  $4.13\text{--}3.44 \times 10^2$ ,  $0.77\text{--}1.80 \times 10^1$ ,  $2.00\text{--}5.21$  gene copies/mL respectively (Fig 5). *Sul1* genes were at



**Fig 5. The abundance of the ARGs in ballast and harbor waters. (A) BW and HW1, (B) BW2 and HW2, (C) BW3 and HW3.** Two-tailed T-test, \*P<0.05, \*\*P<0.01. ARGs with non-detects in both BW and HW samples were excluded.

doi:10.1371/journal.pone.0143123.g005

higher concentrations in harbor waters, while *sul2* and *dfrA* concentrations were higher in the harbor waters with the exception of HW2. In general, harbor water samples had higher measurements of ARGs in comparison to the ballast waters.

**Table 5. Physical characteristics of ballast and harbor waters.**

Sample	Water residence time (days)	Sampling date	Temperature (°C)	pH	Turbidity (NTU)	Salinity(ppt)
BW1	4	14/5/2014	29.9	6.9	0	14
BW2	107	21/5/2014	30.0	7.85	0	20.7
BW3	11	25/5/2014	29.8	7.71	4	18.9
HW1	N/A	14/5/2014	30.3	7.63	10	11.8
HW2	N/A	21/5/2014	30.5	5.98	10	27.6
HW3	N/A	25/5/2014	30.2	7.87	0	30.9

doi:10.1371/journal.pone.0143123.t005

### Correlations between microbial community composition, environmental variables, indicator organism loads and antibiotic resistance

The pH of water samples ranged between 5.98–7.87, with HW2 having lowest value measured (pH 5.98, Table 5). HW1 and HW2 samples also showed slightly higher temperatures and turbidity measurements than the rest of the samples. HW1 had the lowest salinity levels, which could be attributed to dilution due to rainfall prior to sampling. A Pearson correlation of environmental variables to ARG, ARB and indicator organisms showed negative correlations between pH and *crf* gene copies, linezolid resistant counts and *P. aeruginosa* gene copy numbers ( $P < 0.05$ , Table 6). Temperature and turbidity measurements were positively correlated to *V. vulnificus* ( $P < 0.05$ , Table 6). Significant correlations were found between the ARGs *eraA*, *ermG*, *cfr* and *tetO* and indicator organisms ( $P < 0.05$ , Table 7).

A clustering analysis showed that the microbial community in the harbor waters was more similar in microbial composition than ballast waters at a similarity cut off of 75% (Fig 6). At the same level of similarity, BW2 and BW3 showed more resemblance compared to BW1, and this clustering pattern observed on a Principal Component Analysis (PCA) plot was statistically verified using an ANOSIM analysis ( $P < 0.05$ ).

Using the PERMANOVA analysis in the Primer v6 software, microbial community patterns in the harbor samples were positively correlated (using a multivariate analysis where Pearson correlation coefficient of  $> 0.7$ ) to measurements of *E. coli*, *V. cholerae*, *P. aeruginosa*, *V. vulnificus*, temperature, *tetO* and *ermF1* genes (Fig 7). The PCO1 and PCO2 explained 60.3% and 26.4% of the variation. The ARG *ermC* was positively correlated to microbial composition of BW1, while BW2 and 3 correlated to 16S rRNA gene sequence (Fig 7).

**Table 6. Pearson correlation analysis of environmental variables to ARG, ARB and indicator organism measurements from ballast and harbor waters.**

Variable (units)	Temperature (°C)	pH	Turbidity (NTU)	Salinity (ppt)
<i>cfr</i> gene (gene copies/mL)	0.755 (0.082)	<b>-0.854*(0.030)</b>	0.576 (0.231)	0.530 (0.280)
LINE ARB concentrations (CFU/mL)	0.544 (0.265)	<b>-0.893*(0.017)</b>	0.467 (0.350)	0.122 (0.818)
<i>E. coli</i> (MPN/100mL)	0.627 (0.182)	-0.053 (0.920)	0.756 (0.082)	-0.347 (0.501)
<i>Enterococcus spp.</i> (MPN/100mL)	0.366 (0.475)	0.171 (0.746)	0.666 (0.148)	-0.541 (0.267)
<i>V. cholerae</i> (MPN/100mL)	0.506 (0.306)	0.009(0.986)	0.746 (0.089)	-0.491 (0.322)
<i>V. vulnificus</i> (MPN/100mL)	<b>0.832*(0.040)</b>	-0.536 (0.273)	<b>0.949*(0.004)</b>	-0.099 (0.853)
<i>P. aeruginosa</i> (gene copies/mL)	0.743 (0.090)	<b>-0.875*(0.022)</b>	0.648 (0.165)	0.422 (0.405)
<i>Salmonella spp.</i> (gene copies/mL)	0.340 (0.509)	0.201 (0.703)	0.600 (0.208)	-0.580 (0.227)

P-values are represented in parenthesis with

\* $P < 0.05$ .

doi:10.1371/journal.pone.0143123.t006

**Table 7. Pearson correlation analysis of ARGs to indicator organisms from ballast and harbor waters.**

Variable	<i>ereA</i>	<i>ereB</i>	<i>ermB1</i>	<i>ermC</i>	<i>ermF1</i>	<i>ermG</i>	<i>sul1</i>	<i>sul2</i>	<i>dfrA</i>	<i>cfr</i>	<i>tetO</i>
<i>E. coli</i> (cor. coefficient)	<b>0.998*</b>	-0.181	0.415	-0.285	0.785	0.993	-0.002	-0.037	0.342	0.113	<b>0.957*</b>
<i>E. coli</i> (p value)	<b>(0.037)</b>	(0.732)	(0.413)	(0.584)	(0.064)	(0.075)	(0.997)	(0.945)	(0.507)	(0.831)	<b>(0.003)</b>
<i>Enterococcus</i> (cor. coefficient)	0.959	-0.110	0.331	-0.412	0.712	0.994	-0.142	-0.116	0.293	-0.149	<b>0.981*</b>
<i>Enterococcus</i> (p value)	(0.182)	(0.835)	(0.522)	(0.416)	(0.113)	(0.070)	(0.789)	(0.826)	(0.574)	(0.778)	<b>(0.001)</b>
<i>V. cholerae</i> (cor. coefficient)	0.974	-0.166	0.430	-0.301	0.694	<b>0.999*</b>	-0.165	-0.176	0.233	0.012	<b>0.958*</b>
<i>V. cholerae</i> (p value)	(0.145)	(0.754)	(0.395)	(0.562)	(0.126)	<b>(0.033)</b>	(0.755)	(0.739)	(0.658)	(0.981)	<b>(0.003)</b>
<i>V. vulnificus</i> (cor. coefficient)	0.974	0.028	0.509	-0.109	0.459	<b>0.999*</b>	-0.091	-0.278	-0.063	0.624	0.653
<i>V. vulnificus</i> (p value)	(0.145)	(0.958)	(0.303)	(0.838)	(0.360)	<b>(0.033)</b>	(0.864)	(0.593)	(0.905)	(0.186)	(0.160)
<i>P. aeruginosa</i> (cor. coefficient)	0.977	0.240	0.312	0.181	-0.081	<b>0.999*</b>	0.048	-0.241	-0.377	<b>0.991*</b>	-0.079
<i>P. aeruginosa</i> (p value)	(0.137)	(0.646)	(0.547)	(0.731)	(0.879)	<b>(0.025)</b>	(0.928)	(0.646)	(0.462)	<b>(0.000)</b>	(0.881)
<i>Salmonella</i> spp. (cor. coefficient)	0.974	-0.218	0.359	-0.339	0.708	<b>0.999*</b>	-0.175	-0.122	0.314	-0.205	<b>0.970*</b>
<i>Salmonella</i> spp. (p value)	(0.145)	(0.678)	(0.485)	(0.511)	(0.116)	<b>(0.033)</b>	(0.741)	(0.818)	(0.545)	(0.697)	<b>(0.001)</b>

P-values are represented in parenthesis with

\*P<0.05.

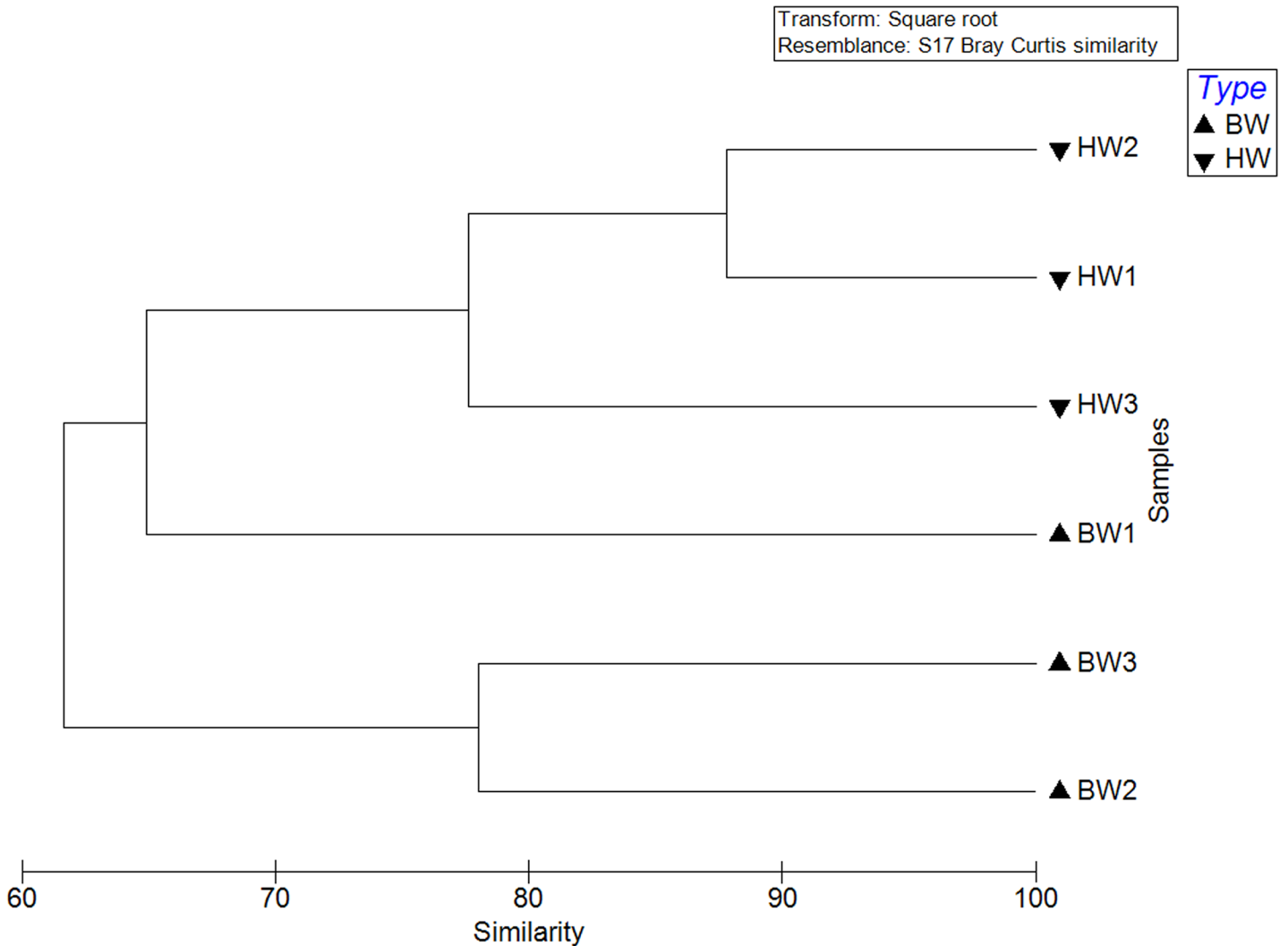
doi:10.1371/journal.pone.0143123.t007

## Discussion

Mid ocean exchange (MOE) of ballast waters has been implemented by IMO to prevent the dissemination of invasive aquatic species. This method requires ships to exchange ballast water at least 200 nautical miles away from shore at an ocean depth of at least 200 m. It is proposed that replacement of ballast water with mid ocean waters (containing oligotrophic marine bacteria) flushes or dilutes any potential coastal dwelling invasive organisms, which may have been picked up during the voyage of ships. In a few other studies, this method of water replacement has shown no significant decrease in bacterial counts (pre- versus post-exchange). Rather, a decline in bacterial abundance in ballast waters was mainly attributed to longer water retention time in ballast tanks [8, 35, 36]. Overall, two of the ballast waters in our study had a significant lower abundance of viable HPC than harbor waters (with exceptions of BW2 and HW2). In an earlier study using flow cytometry, Joachimsthal et al. (2003) reported bacterial concentrations of ballast waters in Singapore ranging from  $2.35 \times 10^6$ – $5.87 \times 10^7$  cells/mL, with differences between harbor and ballast waters attributed to differences in location and season of ballast water uptake and open ocean water samplings [37, 38].

In our study, the levels of indicator organisms measured in ballast waters were within limits administered by IMO, with the exception of one ballast water sample where *Enterococcus* concentrations exceeded guideline values. In comparison to another study using FISH techniques, levels of *Vibrio* spp. and *E. coli* in our study was generally much lower than previously reported [12]. We found a correlation between higher temperatures, turbidity and *V. vulnificus* abundance in harbor waters (P<0.05). Studies have shown that marine pathogens such as *Vibrio* spp. may preferentially thrive in coastal areas with increased water temperatures and reduced salinity (NaCl <25 ppb, [39]) due to terrestrial runoff.

*Pseudomonas* spp. and *Salmonella* spp. are not listed as ballast water indicator organisms of concern according to IMO guidelines. However, we selected these bacteria for testing based on detection in other ballast water studies. For example, Burkholder et al. (2007) reported the presence of other pathogens including *P. aeruginosa* in 48% of ballast waters surveyed from vessels in the east and west coast of the United States [40]. In another study using chromogenic agar, Altug et al. (2012) were able to detect viable *Salmonella* spp. (300–500 CFU/100 mL) in



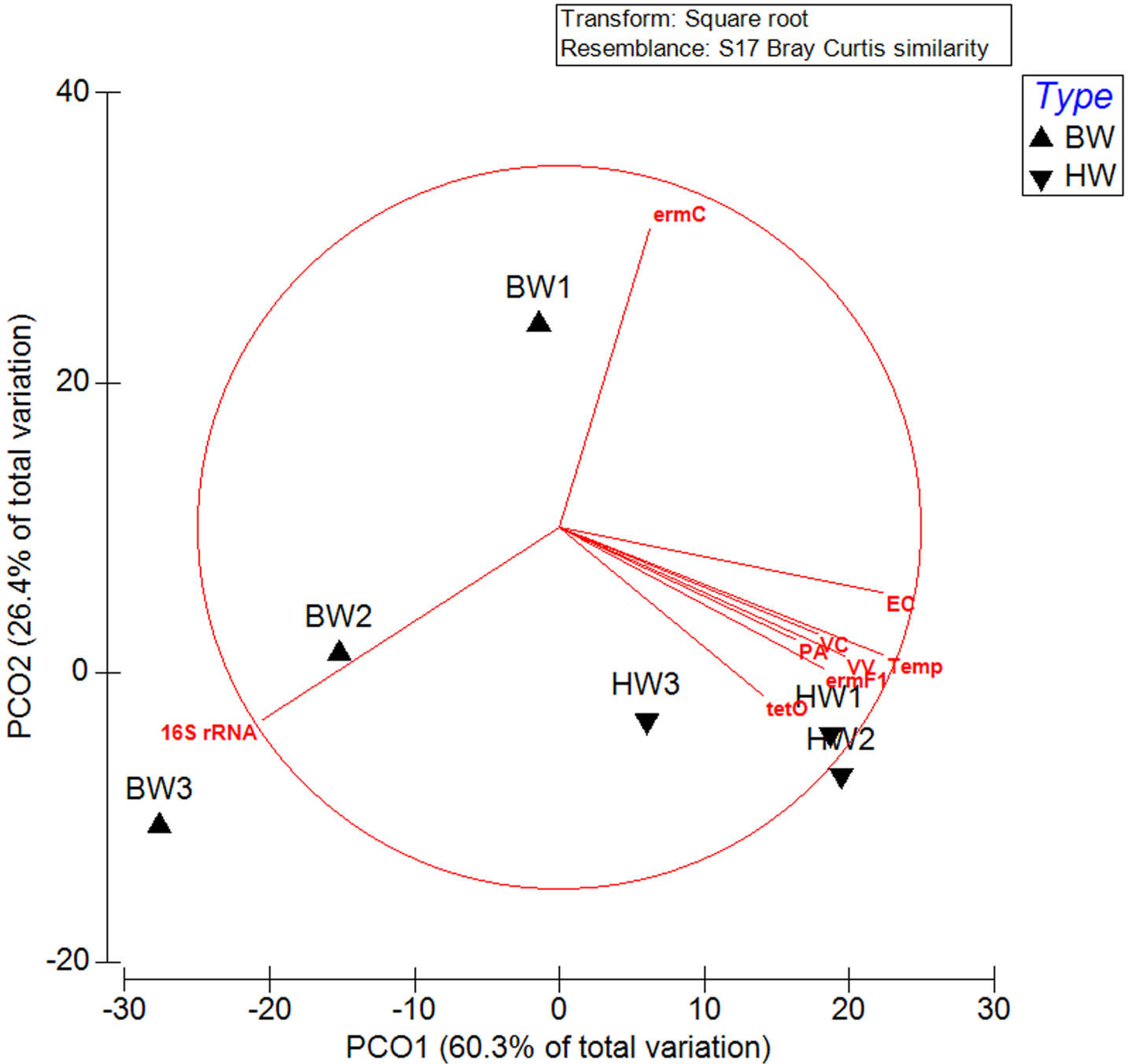
**Fig 6. Clustering analysis of microbial composition in ballast and harbor waters based on Bray-Curtis similarity index.**

doi:10.1371/journal.pone.0143123.g006

ballast waters coming from different regions in the world but not in the Sea of Marmara, Turkey where the ships docked [41]. For this study, although *Pseudomonas spp.* in ballast waters were detected at levels lower (0–20 GC/mL) relative to the other indicator organisms, their ability to form biofilms on the surfaces of tanks and resistance to environmental stress make them important targets to monitor [42].

Compared to soil, antibiotic resistance of bacteria in marine environments has not been well documented. The ocean is a natural habitat of antibiotic-producing bacteria with marine aquaculture introducing more antibiotics to treat infections and boost aquaculture production. Plasmids are mobile vectors for the transfer of resistance genes that confer an advantage to their host's adaptation to different ecological niches [43]. In this study, samples were analyzed for the presence of ARB and associated ARGs, which are increasingly being recognized as xenobiotics, and emerging contaminants in the aquaculture industry. As a whole, the culturable marine microbial communities in ballast and harbor waters were most resistant to trimethoprim, sulfanilamide, lincomycin and tetracycline. Shah et al. (2014) reported the same patterns of ARB (by proxy of pheno- and genotypic screening) in Chilean marine sediments of





**Fig 7. Principal component analysis (PCA) of the microbial community structure of ballast and harbor waters in relationship with pathogen, ARB and ARG data.** Only variables with a Pearson correlation coefficient of >0.7 were included in the PCA plot. EC—*E. coli*, VC—*V. cholera*, PA—*P. aeruginosa*, VV—*V. vulnificus*, Temp—temperature.

doi:10.1371/journal.pone.0143123.g007

salmon fish farms, which suggests that aquaculture activity likely contributes to antibiotic resistance patterns in the marine environment [44]. Interestingly, BW1 had the lowest levels of phenotypic resistance to all antibiotics tested, which could be attributed to its more complex and diverse microbial community structure resulting in its distinct characteristics compared to the other samples. The ARB population in BW1 samples could be more fastidious in nature,

incapable of growing on marine media supplemented by antibiotics. This reasoning is supported by comparing the pheno- and genotypic data which shows the identification of a wide array of ARG targets in BW1 compared to the other samples. This highlights the importance of using both culturing and screening of antibiotic resistance gene markers for a more thorough assessment of antibiotic resistance given that pheno- and genotypic results may contradict each other due to confounding effects of cell viability and other growth impediments.

In most of the samples for which ARGs were detected, higher concentrations were found in the harbor water samples than a non-impacted freshwater body in Singapore, where ARG concentrations were at least half or those reported in the harbor samples (non-impacted site: *sul1*– $1.11 \times 10^2$  gene copies/mL, *cfr*–below detection limit). The persistence of *sul1* and *dfrA* genes together with the relatively high levels of viable ARB counts (for most but not all of the samples), suggests successive pressure exerted by sulfamethoxazole, trimethoprim or co-trimoxazole, a combination of the former two antibiotics. Sulfonamide and tetracycline are commonly used to treat bacterial and protozoan infections in aquaculture leading to the prevalent problem of resistant bacteria in these environments [45–47]. Even without selective pressure, sulfonamide-resistant bacteria can remain stable in the aqueous environment for 5 to 10 years [48].

Other genes including *tetM*, *crf*, *ermB1*, *C*, *F1* and *G* were also identified at lower concentrations in all samples except for BW3 and HW3. Most marine environmental surveys, particularly in fish farms, have identified tetracycline, chloramphenicol and macrolide resistant gene markers [45, 49–52]. Strong positive correlations were shown between *cfr* and *P. aeruginosa* gene copies (in HW2), and *E. coli* and *ereA* gene copies (in HW1) while ARGs *ermG* and *tetO* correlated with a combination of at least four of the indicator organisms tested in this study. These correlation patterns suggest that these indicator organisms may be contributing to the resistome of coastal environments. Plasmids carrying the *cfr* gene mediating resistance to linezolid and chloramphenicol have only been described in *Staphylococcus* and *Enterococcus* [53] but not in other taxa of pathogens. Likewise, there has been no reported association of resistance to macrolides in *Vibrio* spp. encoded by the *ermG* gene. In aquaculture sources, it has been observed that genera of *Pseudomonas* carry florfenicol resistant genes [49]. It would be worthwhile to taxonomically identify a variety of antibiotic resistant phenotypes and test for the presence of these ARGs to verify this correlation. Determining this would enable us to estimate the risk of dissemination of ARGs between the environmental reservoir and human pathogens.

The PCA analysis showed a tighter clustering of the microbial composition of harbor waters indicating higher similarities in structure as compared to the ballast waters which are more dispersed. Although there are some distinguishing bacterial populations among the ballast water samples, it is apparent that diversity is largely dependent on the site of water uptake. The harbor waters were more similar in composition and showed significant positive correlations to indicator organisms (*E. coli*), and pathogens (*V. cholera*, *P. aeruginosa*, *V. vulnificus*), temperature and ARGs (*tetO* and *ermF1* genes).

By comparing microbial communities of ballast and harbor waters, we found low levels of similarity at the species level, which is expected given that the harbor samples are more terrestrially impacted than off-shore oligotrophic seawater samples. At the phyla level, Proteobacteria dominated across all samples, with SAR11 showing notably higher levels of abundance in harbor waters. Members of the SAR11 clade are ubiquitous, and highly adapted to nutrient poor oligotrophic environments such as surface waters of open oceans [54]. Open ocean SAR11 populations, unlike their coastal counterparts, exhibit diel oscillations at the transcript level and may regulate light-dark cycles of transcriptional activity by tracking surrounding environmental conditions [55]. Steindler et al. (2011) found that in light-deprived, nutrient starved conditions, *Candidatus Pelagibacter ubique*, a member of the SAR11 clade, are able to utilize carbon reserves, shrinking to a volume of  $\sim 0.014 \mu\text{m}^3$  [56]. The marine cyanobacteria,

Synechococcus, were also detected at higher levels in harbor waters. The dominance of phototrophs in the harbor waters versus ballast waters (where illumination is absent in the tanks) is influenced by sampling in the late morning at the surface, within the photic zone of the water column. In contrast, nutrient and light-stress in ballast tanks have been shown to take a toll on autotrophic nanoflagellates and mixotrophic populations, with the loss of photosynthetic pigments and dramatic decreases in counts over time [35]. Together, these findings offer a rational explanation for observed low levels of phototrophic cyanobacteria in the ballast water samples and also may suggest a selection of open ocean SAR11 phenotypes.

Taxa associated with marine hydrocarbon degradation (Oceanspirillales, Thiotrichales, Flavobacteria) were enriched in the ballast water samples. Members of these taxa are capable of hydrocarbon oxidation, and have been identified in sediments of marine hydrocarbon seeps from 16S rRNA sequence clone libraries of environmental samples [57]. BW1 had the highest microbial diversity with high abundances of Caldilineales and Sphingobacteriales, which are commonly found in activated sludge in wastewater treatment plants [58, 59]. Over summer periods, Zhang et al. (2014) detected the prevalence of Caldilineales in the subtropical Pearl River Estuary sediments which receives large volumes of inflow from the South China Sea and anthropogenic inputs from surrounding industries [60]. Interestingly, the last point of ballast water uptake for BW1 was in the South China Sea which possibly suggests that a combination of an upwelling event, sediment dispersal and water circulating patterns facilitated the hitchhiking of these microbes in ballast water from this point of origin to our local ports.

The harbor waters showed more similarity in microbial community structure while the ballast waters had more variability. This is expected given that ballast water of the ships surveyed were situated at different locations during uptake. Incorporating methods of assessing microbial community composition by next-generation sequencing (NGS) techniques, in tandem with monitoring indicator organisms and nutrient inputs with the effect on microbial growth and potential harmful algal provides a comprehensive approach of understanding ecological impact of shipping activity and ballast water dispersal.

In conclusion, we observed differences in the bacterial community composition, indicator organisms (*E. coli*, *Enterococcus*, *P. aeruginosa*, *Salmonella spp.*, *Vibrio spp.*), and the concentrations of ARB and ARGs in ballast water of 3 ships and harbor waters at their port of call. Harbor waters exhibited a general trend of higher concentrations of indicator organisms, however we detected the presence of *Enterococcus* in all ballast water samples, with one sample, in particular, exceeding levels recommended by IMO. We found correlations between the concentrations of indicator organisms (*E. coli*, *V. cholera*, *V. vulnificus*, *P. aeruginosa*) and ARGs, which suggests that these species in ballast and harbor waters may be carriers of ARGs. Selective culturing, isolation and antibiotic susceptibility testing of these marine species will be required to confirm if these potential pathogens are indeed resistant and carriers of ARGs. We applied NGS, different molecular assays and conventional microbiological methods to provide detailed microbial measurements and other predictive parameters which may dictate the threat to microbiological safety standards in ballast waters.

## Supporting Information

**S1 Fig. Rarefaction curves of species richness (97% cut off) in ballast and harbor waters.** (TIF)

## Acknowledgments

We would like to thank the Public Utilities Board (PUB) for funding support administered under and the Environment & Water Industry Program Office Incentive for Research &

Innovation Scheme (EWI IRIS) scheme (reference: 1102-IRIS-12-02), and administrative support from the National University of Singapore Environmental Institute (NERI). We also thank Aurore Trottet, Guillaume Drillet and Rui Shan Ker from the Danish Hydraulic Institute (DHI), and Genevieve Gabrielle Rose Valdes Vergara, Fang Haoming and Laurence Haller for assistance with sampling and engaging in scientific discussion.

## Author Contributions

Conceived and designed the experiments: CN THL SGG LL YK JBR KGYH. Performed the experiments: CN THL SGG LL YK. Analyzed the data: CN THL SGG LL. Contributed reagents/materials/analysis tools: JBR KGYH. Wrote the paper: CN THL JBR KGYH.

## References

1. Singapore MPA. Annual report. 2013.
2. Ruiz GM, Rawlings TK, Dobbs FC, Drake LA, T. M, A. H, et al. Global spread of microorganisms by ships. *Nature*. 2000; 408:2.
3. Hallegraef GM, Steffensen DA, Wetherbee R. Three estuarine Australian dinoflagellates that can produce paralytic shellfish toxins. *Journal of plankton research*. 1988; 10:9.
4. McCarthy SA, Khambaty FM. International dissemination of epidemic *Vibrio cholerae* by cargo ship ballast and other non potable waters. *Applied and Environmental Microbiology*. 1994; 60:5.
5. McMinn A, Hallegraef GM, Thomson P, Jenkinson AB, Heijnen H. Cyst and radionucleotide evidence for the recent introduction of the toxic dinoflagellate *Gymnodinium catenatum* into Tasmanian waters. *Marine Ecology Progress Series*. 1997; 161:8.
6. Knight I, Wells CS, Wiggins B, Russell H, Reynolds KA. Detection and enumeration of fecal indicators and pathogens in the ballast water of transoceanic vessels entering the Great Lakes. *General Meeting of American Society for Microbiology American Society for Microbiology*. 1999.
7. Mimura H, Katakura R, Ishida H. Changes of microbial populations in a ship's ballast water and sediments on a voyage from Japan to Qatar. *Marine Pollution Bulletin*. 2005; 50:7.
8. Drake JM, Lodge DM. Global hot spots of biological invasions: evaluating option for ballast-water management. *The royal society of publishing*. 2004; 271:6.
9. Thomson FK, Heinemann SA, Dobbs FC. Patterns of antibiotic resistance in cholera bacteria isolated from ships' ballast water. *Proceedings of the Third International Conference on Marine Bioinvasion; La Jolla, California: Old Dominion University*; 2003.
10. Pruden A, Pei R, Storteboom H, Carlson K. Antibiotic Resistance Genes as Emerging Contaminants: Studies in Northern Colorado. *Environmental Science & Technology*. 2006;(40):6.
11. Joachimsthal EL, Ivanov V, Tay ST, Tay JH. Bacteriological examination of ballast water in Singapore Harbour by flow cytometry with FISH. *Mar Pollut Bull*. 2004; 49(4):334–43. doi: [10.1016/j.marpolbul.2004.02.036](https://doi.org/10.1016/j.marpolbul.2004.02.036) PMID: [15341828](https://pubmed.ncbi.nlm.nih.gov/15341828/).
12. Joachimsthal EL, Ivanov V, Tay ST-L, Tay J-H. Bacteriological examination of ballast water in Singapore Harbour by flow cytometry with FISH. *Marine Pollution Bulletin* 2004; 49:10.
13. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, et al. Primer3—new capabilities and interfaces. *Nucleic Acids Research*. 2012; 240(15).
14. Le TH, Sivachidambaram V, Yi X, Li X, Zhou Z. Quantification of polyketide synthase genes in tropical urban soils using real-time PCR. *J Microbiol Methods*. 2014; 106:135–42. doi: [10.1016/j.mimet.2014.08.010](https://doi.org/10.1016/j.mimet.2014.08.010) PMID: [25181693](https://pubmed.ncbi.nlm.nih.gov/25181693/).
15. Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, et al. Short protocols in molecular biology. *Trends in Biotechnology*. 1992; 10(0):456. doi: [http://dx.doi.org/10.1016/0167-7799\(92\)90304-E](http://dx.doi.org/10.1016/0167-7799(92)90304-E).
16. Dowd SE, Callaway TR, Wolcott RD, Sun Y, McKeenan T, Hagevoort RG, et al. Evaluation of the bacterial diversity in the feces of cattle using 16S rDNA bacterial tag-encoded FLX amplicon pyrosequencing (bTEFAP). *BMC Microbiology*. 2008; 8.
17. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, et al. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and environmental microbiology*. 2009; 75:5.
18. Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics*. 2011; 27(16):7.

19. Clarke KR, Warwick RM. Change in marine communities: an approach to statistical analysis and interpretation. 2nd ed2001.
20. Amann RI, Binder BJ, Olson RJ, Chisholm SW, Devereux R, Stahl DA. Combination of 16S rRNA-targeted oligonucleotide probes with flow cytometry for analyzing mixed microbial populations. *Applied and Environmental Microbiology*. 1990; 56(6):7.
21. Muyzer G, de Waal EC, Uitterlinden AG. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Applied and Environmental Microbiology*. 1993; 59(3):6.
22. Koike S, Aminov RI, Yannarell AC, Gans HD, Krapac IG, Chee-Sanford JC, et al. Molecular ecology of macrolide-lincosamide-streptogramin B methylases in waste lagoons and subsurface waters associated with swine production. *Microbial Ecology*. 2010; 59(3):12.
23. Knapp CW, Zhang W, Sturm BS, Graham DW. Differential fate of erythromycin and beta-lactam resistance genes from swine lagoon waste under different aquatic conditions. *Environ Pollut*. 2010; 158(5):1506–12. doi: [10.1016/j.envpol.2009.12.020](https://doi.org/10.1016/j.envpol.2009.12.020) PMID: [20053492](https://pubmed.ncbi.nlm.nih.gov/20053492/).
24. Zhu YG, Johnson TA, Su JQ, Qiao M, Guo GX, Stedtfeld RD, et al. Diverse and abundant antibiotic resistance genes in Chinese swine farms. *Proceedings of the National Academy of Sciences*. 2013; 110(9):3435–40. doi: [10.1073/pnas.1222743110](https://doi.org/10.1073/pnas.1222743110)
25. Dahmen S, Mansour W, Boujaafar N, Arlet G, Boualle'gue O. Distribution of Cotrimoxazole Resistance Genes Associated with Class 1 Integrons in Clinical Isolates of Enterobacteriaceae in a University Hospital in Tunisia. *MICROBIAL DRUG RESISTANCE*. 2010; 16:5.
26. Lee JC, Oh JY, Cho JW, Park JC, Kim JM, Seo SY, et al. The prevalence of trimethoprim-resistance-conferring dihydrofolate reductase genes in urinary isolates of *Escherichia coli* in Korea. *Journal of Antimicrobial Chemotherapy*. 2001; 47:6.
27. Rapid Detection of Bacterial Antibiotic Resistance: Preliminary Evaluation of PCR Assays Targeting Tetracycline Resistance Genes [Internet]. smithAustralia: Human Protection and Performance Division DSTO Defence Science and Technology Organisation; 2007
28. Smith MS, Yang RK, Knapp CW, Niu Y, Peak N, Hanfelt MM, et al. Quantification of Tetracycline Resistance Genes in Feedlot Lagoons by Real-Time PCR. *Applied and Environmental Microbiology*. 2004; 70(12):7372–7. doi: [10.1128/aem.70.12.7372-7377.2004](https://doi.org/10.1128/aem.70.12.7372-7377.2004) PMID: [15574938](https://pubmed.ncbi.nlm.nih.gov/15574938/)
29. Dickinson G, Lim K-y, Jiang SC. Quantitative Microbial Risk Assessment of Pathogenic Vibrios in Marine Recreational Waters of Southern California. *Applied and Environmental Microbiology*. 2013; 79:9.
30. Panicker G, Call DR, Krug MJ, Bej AK. Detection of Pathogenic *Vibrio* spp. in Shellfish by Using Multiplex PCR and DNA Microarrays. *Applied and Environmental Microbiology*. 2004; 70:9.
31. Brasher CW, DePaola A, Jones DD, Bej AK. Detection of Microbial Pathogens in Shellfish with Multiplex PCR. *Curr Microbiol*. 1998; 37:7.
32. Chun J, Huq A, Colwell RR. Analysis of 16S-23S rRNA Intergenic Spacer Regions of *Vibrio cholerae* and *Vibrio mimicus*. *Applied and Environmental Microbiology*. 1999; 65:7.
33. Vezzulli L, Pezzati E, Moreno M, Fabiano M, Pane L, Pruzzo C, et al. Benthic ecology of *Vibrio* spp. and pathogenic *Vibrio* species in a coastal Mediterranean environment (La Spezia Gulf, Italy). *Microbial Ecology*. 2009; 58:11.
34. Lee D-Y, Shannon K, Beaudette LA. Detection of bacterial pathogens in municipal wastewater using an oligonucleotide microarray and real-time quantitative PCR. *Journal of Microbiological Methods*. 2006; 65:15.
35. Tomaru A, Kawachi M, Demura M, Fukuyo Y. Changes in Microbial Communities, Including both Uncultured and Culturable Bacteria, with Mid-Ocean Ballast-Water Exchange during a Voyage from Japan to Australia. *PLoS One* 2014; 9(5).
36. J.M. S, C.J. W, R.B. R. Microbial hitchhikers: dynamics of bacterial populations in ballast water during a trans-Pacific voyage of a bulk carrier. *Aquatic Invasions* 2010; 5:10.
37. Joachimsthal EL, Tay STL, Ivanov V, Tay JH. Monitoring the bacteriological quality of seawater by flow cytometry: development of a rapid method for the control of ballast water. *Proceedings of the 2nd International Conference on Marine Waste Water Discharges; Istanbul, Turkey2002*.
38. Joachimsthal EL, Ivanov V, Tay STL, Tay JH. Quantification of whole-cell in situ hybridisation with oligonucleotide probes by flow cytometry of *Escherichia coli* cells. *World Journal of Microbiology and Biotechnology*. 2003; 19:7.
39. Baker-Austin C, Trinanés JA, Taylor NGH, Hartnell R, Siitonen A, Martínez-Urtaza J. Emerging *Vibrio* risj at high latitudes in response to ocean warming. *Nature Climate Change*. 2013; 3:5.

40. Burkholder JM, Hallegraef GM, Melia G, Cohen A, Bowers HA, Oldach DW, et al. Phytoplankton and bacterial assemblages in ballast water of U.S. Military ships as a function of port origin, voyage time, and ocean exchange practices. *Harmful Algae*. 2007; 6:33.
41. Altug G, Gurun S, Cardak M, Ciftci PS, Kalkan S. The occurrence of pathogenic bacteria in some ships' ballast water incoming from various marine regions to the Sea of Marmara, Turkey. *Marine Environmental Research*. 2012; 81:8.
42. Landini P. Cross-talk mechanisms in biofilm formation and responses to environmental and physiological stress in *Escherichia coli*. *Research in Microbiology*. 2009; 160:8.
43. Brown Kav A, Sasson G, Jami E, Doron-Faigenboim A, Benhar I, Mizrahi I. Insights into the bovine rumen plasmidome. *Proceedings of the National Academy of Sciences*. 2012; 109:6.
44. Shah SQ, Cabello FC, L'abée-Lund TM, Tomova A, Godfrey HP, Buschmann AH, et al. Antimicrobial resistance and antimicrobial resistance genes in marine bacteria from salmon aquaculture and non-aquaculture sites. *Environmental Microbiology*. 2014; 16(5):20.
45. Tamminen M, Karkman A, Lohmus A, Muziasari WI, Takasu H, Wada S, et al. Tetracycline resistance genes persist at aquaculture farms in the absence of selection pressure. *Environmental Science & Technology*. 2011; 45:6.
46. P.T.P. H, L. N, P.H. V, S. S. Detection of the *sul1*, *sul2*, and *sul3* genes in sulfonamide-resistant bacteria from wastewater and shrimp ponds of north Vietnam. *Science of the Total Environment*. 2008; 405:8.
47. Kim SR, Nonaka L, Suzuki S. Occurrence of tetracycline resistance genes *tet(M)* and *tet(S)* in bacteria from marine aquaculture sites. *FEMS Microbiology Letters*. 2004; 237:10.
48. Gao P, Mao D, Luo Y, Wang L, Xu B, Xu L. Occurrence of sulfonamide and tetracycline-resistant bacteria and resistance genes in aquaculture environment. *Water Research*. 2012; 46:10.
49. Akinbowale OL, eng H, Barton MD. Antimicrobial resistance in bacteria isolated from aquaculture sources in Australia. *J Appl Microbiol*. 2006; 100:11.
50. Akinbowale OL, Peng H, Barton MD. Diversity of tetracycline resistance genes in bacteria from aquaculture sources in Australia. *Journal of Applied Microbiology*. 2007; 103:10.
51. Di Cesare A, Vignaroli C, Luna GM, Pasquaroli S, Biavasco F. Antibiotic-resistant enterococci in seawater and sediments from a coastal fish farm. *Microbial Drug Resistance*. 2012.
52. Buschmann AH, Tomova A, Lopez A, Maldonado MA, Henriquez LA, Ivanova L. Salmon aquaculture and antimicrobial resistance in the marine environment. *PLoS ONE*. 2012; 7:e42724. doi: [10.1371/journal.pone.0042724](https://doi.org/10.1371/journal.pone.0042724) PMID: [22905164](https://pubmed.ncbi.nlm.nih.gov/22905164/)
53. L. D, P. K, R.E. M, D. P, K.V. S, C.A. A. Transferable plasmid-mediated resistance to linezolid due to *cf* in a human clinical isolate of *Enterococcus faecalis*. *Antimicrobial Agents and Chemotherapy*. 2012; 56(7).
54. Tripp HJ. The Unique Metabolism of SAR11 Aquatic Bacteria. *Journal of Microbiology*. 2013; 51:5.
55. Ottesen EA, Young CR, Gifford SM, Eppley JM, Marin R, Schuster SC, et al. Multispecies Diel Transcriptional Oscillations in Open Ocean Heterotrophic Bacterial Assemblages. *Science* 345, 207–212. *Science*. 2014;245:6. doi: [10.1126/science.1252476](https://doi.org/10.1126/science.1252476) PMID: [25013074](https://pubmed.ncbi.nlm.nih.gov/25013074/)
56. Steindler L, Schwabach MS, Smith DP, Chan F, Giovannoni SJ. Energy Starved *Candidatus Pelagibacter Ubique* Substitutes Light-Mediated ATP Production for Endogenous Carbon Respiration. *PLoS One*. 2011; 6:e19725. doi: [10.1371/journal.pone.0019725](https://doi.org/10.1371/journal.pone.0019725) PMID: [21573025](https://pubmed.ncbi.nlm.nih.gov/21573025/)
57. Redmond MC, Valentine DL, Sessions AL. Identification of Novel Methane-, Ethane-, and Propane-Oxidizing Bacteria at marine Hydrocarbon Seeps by Stable Isotope Probing. *Applied and Environmental Microbiology*. 2010; 76:11.
58. Sekiguchi Y, Yamada T, Hanada S, Ohashi A, Harada H. *Anaerolinea thermophila* gen. nov., sp. nov. and *Caldilinea aerophila* gen. nov., sp. nov., novel filamentous thermophiles that represent a previously uncultured lineage of the domain Bacteria at the subphylum level. *International Journal of Systematic and Evolutionary Microbiology*. 2003; 153:8.
59. Ye L, Zhang T. Bacterial communities in different sections of a municipal wastewater treatment plant revealed by 16S rDNA 454 pyrosequencing. *Applied Microbiology and Biotechnology*. 2012; 97:9.
60. Zhang W, Bougouffa S, Wang Y, Lee OO, Yang JY, Chan C, et al. Towards Understanding the Dynamics of Microbial Communities in an Estuarine System. *PLoS One*. 2014; 9(4).