Dynamic Subspecies Population Structure of Vibrio Cholerae in Dhaka, Bangladesh

Mohammad Islam
Department of Biological Sciences, University of Alberta, Edmonton, Alberta T6G 2E9, Canada
https://orcid.org/0000-0002-9113-1430

Tania Nasreen
Department of Biological Sciences, University of Alberta, Edmonton, Alberta T6G 2E9, Canada

Kevin Y. H. Liang
Department of Biological Sciences, University of Alberta, Edmonton, Alberta T6G 2E9, Canada

Fatema-Tuz Johura
Infectious Diseases Division, International Centre for Diarrhoeal Disease Research, Bangladesh (icddr,b), Dhaka, Bangladesh

Paul C. Kirchberger
3Department of Integrative Biology, University of Texas at Austin, Austin, Texas 78712, USA

Marzia Sultana
Infectious Diseases Division, International Centre for Diarrhoeal Disease Research, Bangladesh (icddr,b), Dhaka, Bangladesh

Rebecca J. Case
Department of Biological Sciences, University of Alberta, Edmonton, Alberta T6G 2E9, Canada

Munirul Alam
Infectious Diseases Division, International Centre for Diarrhoeal Disease Research, Bangladesh (icddr,b), Dhaka, Bangladesh

Yann F. Boucher (ephyb@nus.edu.sg)
Department of Biological Sciences, University of Alberta, Edmonton, Alberta T6G 2E9, Canada

Research Article

Keywords: Cholera. Endemic. Pandemic. qPCR. Population. Diversity

Posted Date: April 1st, 2021

DOI: https://doi.org/10.21203/rs.3.rs-364004/v1

Abstract

Cholera has been endemic to the Ganges delta for centuries. Although the causative agent, *Vibrio cholerae,* is autochthonous to coastal and brackish water, cholera occurs continually in Dhaka, the inland capital city of Bangladesh which is surrounded by fresh water. Despite the persistence of this problem, little is known about the environmental abundance and distribution of lineages of *V. cholerae*, the most important being the pandemic generating lineage (PG) consisting mostly of serogroup O1 strains. To understand spatial and temporal dynamics of PG and other lineages belonging to the *V. cholerae* species in surface water in and around Dhaka city, we used qPCR and high throughput amplicon sequencing. Seven different freshwater sites across Dhaka were investigated for six consecutive months and physiochemical parameters were measured *in situ*. Total abundance of *V. cholerae* was found to be relatively stable throughout the six months sampling period, with 2×10^5 to 4×10^5 genome copies/L at six sites and around 5×10^5 genome copies/L at the site located in the most densely populated part of Dhaka city. PG O1 *V. cholerae* was present in high abundance during the entire sampling period and composed between 24-92% of the total *V. cholerae* population, only showing occasional but sudden reductions in abundance. In instances where PG O1 lost its dominance, other lineages underwent a rapid expansion while the size of the total *V. cholerae* population remained almost unchanged. Intraspecies richness of *V. cholerae* was positively correlated to salinity, conductivity and total dissolved solids (TDS), while it was negatively correlated to dissolved oxygen (DO) concentration in water. Interestingly, negative correlation was observed specifically between PG O1 and salinity, even though the changes in this variable were minor (0-0.8 ppt). Observations in this study suggest that at the subspecies level, population composition of naturally occurring *V. cholerae* can be influenced by fluctuations in environmental factors, which can lead to altered competition dynamics among the lineages.

Importance

Although *Vibrio cholerae* has been studied for decades as a human pathogen, little is known about its abundance and population structure in the environment before an outbreak. In this study, we have employed high-throughput amplicon sequencing of a species-specific region within a protein-coding gene (*viuB*) to track subspecies lineages of *V. cholerae*, including the PG lineage responsible for cholera in an urban region endemic for cholera. Coupled with real-time qPCR, this method provided subspecies level resolution of the abundance and lineage composition of *V. cholerae* populations. This is key to understanding how intraspecies diversity could play a role in regulating the abundance of pandemic *V. cholerae*, with environmental factors contributing to the relative fitness of different lineages.

Introduction

Vibrio cholerae is a normal inhabitant of aquatic environments such as rivers, estuaries and coastal waters and has been detected in diverse geographic locations worldwide[1]. Toxigenic strains of *V. cholerae* are capable of causing cholera, an acute life-threatening diarrheal disease [2, 3], which is a major public health concern because of its high morbidity and mortality with an estimation of 1.3 to 4

million cases and 21, 000 to 143, 000 deaths worldwide each year [4]. Seven cholera pandemics have struck human civilization so far; the first pandemic of cholera started in 1817 and was followed by six others in the next two hundred years, leaving a devastating human death toll [5]. In the environment, *V. cholerae* is a diverse species, with more than 200 serogroups being identified based on their surface polysaccharide O antigen [6]. However, only O1 serogroup dominates the pandemic generating (PG) lineage [7, 8], which has been responsible for all seven cholera pandemics, and other serogroups mostly represent *V. cholerae* environmental strains which are generally non-pathogenic [5, 9]. Serogroup O1 strains of the PG lineage are further classified into two biotypes, the Classical biotype that was shown to cause the fifth and sixth pandemic and believed to be associated with the earlier pandemics, and the El Tor biotype which is the causative agent for the seventh pandemic of cholera [5, 8].

Cholera remains as an emerging and reemerging disease today [5]. The seventh pandemic of cholera started in 1961 and is still ongoing, known as the world's longest and most persistent pandemic [4, 8]. Despite being a global disease, the world's worst cholera epidemics can be traced back to Ganges delta, hence that area has been thought to serve as a reservoir for the disease [10, 11]. High population density and improper sewage disposal along with uncontrolled industrialization constantly pollute the water sources in this region. Dhaka, the capital city of Bangladesh, is considered as an hyper-endemic region for cholera, the disease continually exists at a low incidence but exhibits biannual seasonal outbreaks [12]. Consequently, Dhaka has been a center for clinical research on cholera for many years. Recent advancement on cholera surveillance intended to increase prevention, preparedness, intervention and awareness of the disease has been chiefly established based on analysis of clinical cases and clinical samples [13]. However, little attention has been paid to dynamics of the natural population of *V. cholerae*, which could play an important role in the epidemiological outcome of cholera. As *V. cholerae* is an autochthonous member of the aquatic environments, understanding its life cycle and ecology requires a knowledge of the temporal and spatial variations in abundance of various genotypes in its natural habitat.

Most studies looking at the environmental aspects of the disease have been dependent on culture-based isolation of pathogenic *V. cholerae* [14]. Culture-based studies underestimate *V. cholerae* abundance and diversity, because of important limitations including enrichment biases and the viable but non-culturable (VBNC) state [15–17]. Real-time qPCR analyses of environmental DNA targeting *V. cholerae* or O1 serogroup-specific genes are quantitative but do not provide information on the subspecies composition of the populations analyzed [18]. Additionally, culture-independent studies, including 16S rRNA sequencing, can be helpful for genus identification but do not usually provide resolution at the species or subspecies level [19]. A fluorescent antibody staining method has been used to enumerate viable *V. cholerae* cells in water samples [20], but only targets the O1 or O139 serogroups and cannot distinguish between O1 strains belonging to the pandemic generating (PG) lineage and strains from other lineages bearing that antigen.

To overcome these limitations, we have applied high-throughput sequencing of a species-specific, highly variable region of a gene, encoding the vibriobactin utilization protein, *viuB* [21]. This protein releases iron

captured by the siderophore vibriobactin inside the cell [22], which is a housekeeping function for *V. cholerae*. In combination with using qPCR [23], amplification and sequencing of the partial *viuB* gene allowed for the elucidation of the spatio-temporal abundance of the PG and other lineages belonging to the *V. cholerae* species in Dhaka's water reservoirs during six consecutive months from October 2015 to March 2016, roughly encompassing after-monsoon, fall, winter and spring seasons in Bangladesh. The study reveals the continual presence of the PG lineage and occasional reduction of its abundance in conjunction with increases in other lineages, providing insights on the influence of environmental factors in subspecies level population dynamics of *V. cholerae*.

Materials And Methods

Study site

Surface water samples were collected from seven different locations (site 1 to site 7) in Dhaka (23.8103^{o} N, 90.4125^{o} E), Bangladesh (Fig. 1), biweekly for six consecutive months from October 2015 to March 2016. Dhaka is the capital city of Bangladesh surrounded by a river system mostly composed of four rivers: Turag, Buriganga, Shitalakshya and Balu. A population of >21 million was recorded for this area in 2020, with a density of 23,234 people/ km^2 within a total area of 300 km^2 (Fig. 1) (World population review, 2020)). The physical distance between site 1 and site 7 is shorter (9.9 km) than the distance between site 1 and site 7 to the other five sites (approximately 21 to 25 km). The climate of Dhaka is categorized as tropical wet and dry with a distinct monsoon season. The average water temperature recorded is 26.1 ^{o}C (19.1 ^{o}C in Jan and 29.1 ^{o}C in June). The visual inspection of the study areas indicated that local markets surrounded these sites, and human intrusion such as bathing, swimming, washing household utensils dishes as well as bathing domestic animals were frequent. Occasional direct defecation in the water at several study sites were also noticeable.

Sample collection and processing

Water samples (200 ml) were collected directly from the sources in sterile Nalgene bottles. Using 50 ml sterile polypropylene syringes 50 ml water sample filtered through 0.22 µm Sterivex filters (Millipore). Total DNA extraction from the biomass on the filters was done through the following three consecutive steps: cell lysis and digestion, DNA extraction, and DNA concentrating and washing according to the protocol developed by Wright et al. [24]. To reduce impurities that can act as PCR inhibitors during amplification, all extracted DNA samples were further treated with One step PCR inhibitory removal kit (ZYMO Research) by following the user manual instructions with 90-180 µl of yield achieved from 100-200 µl of the eluted extracted DNA sample. Treated samples were kept at -20° C for further analysis.

Physicochemical parameters

Surface water quality was measured in situ at the sampling sites. EXO2 multiparameter sonde (YSI, Xylam Brand, USA) allowed for simultaneous measurement of pH, dissolved oxygen (DO), conductivity,

total dissolved solids (TDS), salinity and water temperature. Properly calibrated sensors attached to the instrument were placed in each site while sampling and data were recorded for analysis.

PCR amplification and Illumina sequencing

A touchdown PCR was performed to amplify a 293bp region of the *viuB* gene from DNA extracted from biomass. Master mix for PCR contained 5 µl of 5× Phire Buffer (ThermoFisher), 0.4 µl of 10 mM dNTP mix (ThermoFisher), 0.4 of µl Phire Hot Start II DNA Polymerase (ThermoFisher), 0.5 µl of Molecular Biology Grade Bovine Serum Albumin (20 mg/mL, New England Biolabs), 0.5 µl each of 10 pmol forward and reverse primers (for *viuB*: *viuB*2f 5'-CCGTTAGACAATACCGAGCAC-3' and *viuB*5r 5'-TTAGGATCGCGCACTAACCAC-3') [21] and 2 µl of template DNA. The PCR reaction was performed as follows: initial denaturation at 98 °C for 4 min, followed by 10 cycles of denaturation at 98 °C for 10 sec, annealing at 60 °C for 6 sec (reduced by 1°C per cycle), and extension 72 °C for 1 sec; followed by 23 cycles of denaturation at 98 °C for 10 sec, annealing at 50 °C for 6 sec (reduced by 1°C per cycle), and extension at 72 °C for 1 sec; and a final extension at 72 °C for 1 min. Dual-indexed sequences using indices developed by Kozich et al. [25] were used to prepare amplified *viuB* products for sequencing. Amplified *viuB* products (2 µl) were used as template for the tagging PCR reaction with the same reagents as above with a set of forward and reverse primers that contained appropriate Illumina-adapters, a sample-specific 8 nucleotide index sequence, a 10 nucleotide pad, 2 nucleotide linker, and the gene specific sequence described above, for a total of 70 and 65 bp [21]. This tagging PCR reaction was performed as follows: initial denaturation at 98 °C for 30 sec; followed by two cycles of denaturation at 98 °C for 10 sec, annealing at 55 °C for 6 sec, and extension at 72 °C for 1 sec; and final extension at 72 °C for 1 min. Use of gene-specific primers during amplification and subsequent tagging to create dual-indexed PCR products facilitated improved yield of amplicons and prevented biased amplification due to unexpected interaction of non-primer sequences with the template.

Additionally, eight tagging reactions were done for each sample to obtain an adequate concentration of amplicon DNA for further analysis. All eight reactions of the same sample were pooled together and run on a 2% agarose gel in 1× Tris-Acetate-EDTA buffer where two bands of very similar size, a smaller band (around 360 bp) representing only half-tagged PCR products, and a slightly bigger band (428 bp) of the fully tagged product were visualized. The 428bp bands were cut out of the gel. PCR products were purified using Wizard SV Gel and PCR Clean-Up System (Promega) according to the instructions by the manufacturer. The concentration of cleaned PCR products was then measured using a Qubit Fluorometer (ThermoFisher) with a Qubit dsDNA HS Assay Kit (ThermoFisher) and pooled together in equal concentrations (>10 ng/µl). The pooled samples were then concentrated using a Wizard SV Gel and PCR Clean-Up System (Promega) according to the instructions by the manufacturer. Quality control of the pooled and the concentrated sample was performed using an Agilent 2100 Bioanalyzer system. Sequencing was performed using a v3 (600 cycles) illumina sequencing reagent kit and in Illumina MiSeq platform.

Sequence Analysis

Processing of amplicon sequence reads was performed following the procedure described by Kirchberger et al [21]. Briefly, de-multiplexed raw reads were processed in R [26] using the DADA2 pipeline [27]. Forward and reverse reads were trimmed due to a drop-off in read quality in the first 10 bp as well as after 240 bp and 160 bp for forward and reverse reads, respectively. Assembled overlapping forward and reverse reads were therefore 272 bp in length, 11 bp shorter than the fully sequenced region. Reads with a maximum expected error rate > 1% were also discarded based on DADA2 analysis. After this procedure, 1072 unique sequences remained in the dataset. Chimera detection implemented in DADA2 was then performed on pooled samples, leaving a total of 460 unique sequences. Sequence reads were assigned to *viuB* alleles, 25 of which were composed of more than 1,000 reads (with an average of 100,000 reads per sample) and of with the 12 most abundant alleles representing >99% of all reads were considered for further analysis.

Real-time qPCR amplification

A real-time qPCR assay was performed to determine the abundance of total *V. cholerae* and toxigenic *V. cholerae* O1[28]. For total *V. cholerae*, the *viuB* gene encoding vibriobactin utilization protein B was targeted using the probe, 5'-/56-FAM/TCA TTT GGC/ZEN/CAG AGC ATA AAC CGG T/3IABkFQ/-3',; forward and reverse primers, 5'-TCG GTA TTG TCT AAC GGT AT-3' and 5'-CGA TTC GTG AGG GTG ATA-3', respectively, for a 77-bp product. For *V. cholerae* O1 serogroup, *rfbO1* gene was targetted; Probe, 5'-/5HEX/AGAAGTGTG/ZEN/

TGGGCCAGGTAAAGT/3IABkFQ/-3', forward 5'- GTAAAGCAGGATGGAAACATATTC-3' And reverse, 5'-TGGGCTTACAAACTCAAGTAAG-3' primers were used for a 113bp product[23]. For the qPCR reaction, Dynamite qPCR Mastermix which is a proprietary mix, developed and distributed by the Molecular Biology Service Unit (MBSU) at University of Alberta, Canada was used. The volume of the PCR reaction was 10 µl containing 5 µl of 2× Dynamite qPCR master mix, 1 µl of each of 500 nM primer-250 nM probe mix, 1 µl of molecular grade water and 2 µl of DNA template. Real-time quantitative PCR was performed using the following conditions: initial primer activation at 95 °C for 2 min followed by 40 cycles of 95 °C for 15 s, 60 °C for 1 min in Illumina Eco Real-Time PCR system.

Statistical and multivariate analyses

Two-dimensional visualizations of the overall *V. cholerae* community structure was performed using NMDS (Non-metric multidimensional scaling) with Bray-Curtis distance at operational taxonomic unit (OTU) levels based on unique *viuB* alleles. 'bioenv'. function in vegan package was used to find a non-parametric monotonic relationship between the dissimilarities in the samples matrix and for plotting the location of each site in a two dimensional space [29]. Abundance and environmental variables data were analyzed using R.Multiple subplots were generated for the environmental variables against *viuB* allele richness, and a correlation test was performed to determine any significant relationships between them.

Hierarchical clustering analysis was done with the R function hclust to identify clusters of *viuB* alleles with the most similar distribution based on the abundance data.

Phylogenetic analysis and determination of genome similarity

For phylogenetic analysis, two representative strains (Table S1) were chosen for each major *viuB* allele. Reference genome sequences were obtained from NCBI database. Whole genome analysis was done using mugsy v1.2.3 with default parameters [30]. Mugsy outputs were analyzed using galaxy web server (Galaxy: a web-based genome analysis tool for experimentalists) (https://usegalaxy.org), and the phylogenetic tree was constructed using Raxml v8.2.11 under the GTRGAMMA model with 100 bootstrap replicates [31]. Defaults were chosen for all other parameters. The phylogenetic tree was visualized using iTOL [32]. *in-silico* DNA-DNA Hybridization (dDDH), which is a proxy for traditional DDH values were used to calculate genomic similarities [33]. All pairwise comparisons for dDDH were calculated using GGDC with default parameters [34]. Allelic differences were also used to determine genomic similarities. To determine allelic differences each genome was first annotated using RAST [35]. A set of 2443 genes common in most *V. cholerae* strains were then identified using Usearch [36]. Allele designations and identification were subsequently performed using automated scripts made available by BIGSdb [37]. Finally, pairwise allelic differences for all isolates were calculated using an in-house script and only loci present in both isolates of a pair are considered. Sequences for all gene alleles are available on https://pubmlst.org/vcholerae under the cgMLST scheme.

Results And Discussion

The pandemic generating lineage is the most abundant *Vibrio cholerae* genotype in Dhaka's water system

Dhaka, one of the most densely populated cities in the world (>21 million residents as of June 2020), is located within the Ganges delta. It is an inland city, with a water system primarily consisting of freshwater rivers and canals. Cholera is endemic there and shows biannual peaks in reported cases during the spring and fall, i.e. before and after the monsoons [38]. However, little is known about the abundance and distribution of *V. cholerae* in natural waterbodies and if it correlates with environmental and human factors. To track the abundance of total *V. cholerae* and pandemic generating (PG) lineage *V.* cholerae, we used culture-based detection as well as qPCR analysis. A 272 bp hypervariable stretch of the *viuB* marker gene, which is present in a single copy in *V. cholerae* [21] was amplified and sequenced from fortnightly samples taken in seven different water reservoirs in and around Dhaka city from October 2015 to Mach 2016 (Fig. 1). Presence of *V. cholerae* could be detected by both culture-based isolation and qPCR throughout the sampling period. Abundance of total *V. cholerae* did not show large fluctuations with fortnightly sampling, ranging from 2×10^5 to 5×10^5 genome copies/L (Figure 2). *V. cholerae* O1 could only be detected in ~8% of the samples using culture-based isolation followed by identification with conventional PCR (Table S2). However, qPCR for the *rfbO1* and *ctxA* genes using DNA extracted from water detected toxigenic PG *V. cholerae* O1 throughout the sampling period, with monthly average abundance of this genotype of 1.2×10^5 to 2.5×10^5 genome copies/L (Figure 2).

Using *viu*B amplicon sequencing, twenty-five *viuB* alleles were found in total, differing from each other by two or more single nucleotide polymorphisms. Each unique allele represents a specific *V. cholerae* lineage, roughly the equivalent of a multilocus sequence typing (MLST) clonal complex [21, 39]. Amongst them, 12 *viuB* alleles (each with > 20,000 sequence reads) were included for further analysis of the *V. cholerae* community composition, while the other 13 alleles representing < 1% of the total population were excluded following the strategy adopted previously [21]. The *viuB*-73 allele corresponding to the PG lineage [21], and specifically to 7th pandemic O1 El Tor strains in the Dhaka environment, was present at all seven sites throughout the six months of sampling (Figure 2). That this allele dominates *V. cholerae* communities in the water bodies of Dhaka is not surprising, given the endemicity of cholera in the city and the contamination of water bodies with human waste. Interestingly, *viuB*-73 displayed high abundance based on *viuB* sequencing throughout the six months sampling period, as opposed to only occasional detection of strains corresponding to this genotype using culture-based techniques [15, 38]. Sustained presence of PG *V. cholerae* detected by both qPCR and amplicon sequencing underscores the importance of culture-independent methods of tracking pathogenic *V. cholerae* in nature. Although regular cholera episodes occur year-round, the bacterium is usually only culturable during the two seasonal peaks [38], presumably being present in its viable but nonculturable (VBNC) state at other times [17]. In subspecies-level population analysis, *viu*B-73 remained the predominant allele in Dhaka, representing 24% to 92% of the total *V. cholerae* population. In a study of the natural *V. cholerae* population in cholera-free Oyster Pond on the US east coast, *viuB*-73 was found sporadically at lower relative abundance and other *viu*B alleles dominated the *V. cholerae* population [40]. This study shows that PG *V. cholerae* was present even in the so called 'non-epidemic' months of the year (December-February) and was the predominant genotype throughout the sampling period in the surface water of cholera endemic Dhaka.

Subspecies level diversity correlates with variation in environmental parameters

Studies conducted to look into populations of *V. cholerae* in natural habitats have shown that different environmental factors might influence their abundance [41, 42]. But how variation in these parameters influence subspecies level diversity of *Vibrio cholerae* is not known. In this study, several environmental variables were measured *in situ* (pH, dissolved oxygen, conductivity, total dissolved solids, salinity and water temperature) to gain insight on their effect on the absolute abundance of the species and relative abundance of the different *viuB* alleles found in the water reservoirs in Dhaka. Salinity, pH and temperature showed little variation at the seven sites sampled, ranging from 0-0.8 ppt, 6.5-8.0 and 30-32 °C, respectively. Conductivity (134.5 -1608 µs/cm) and TDS (67.2-804 mg/L) changed noticeably with sites and time of sampling. Dissolved oxygen (DO) also changed across the time and space ranging from 0.1-5.05 mg/L. Correlation of the environmental parameters with the diversity (richness) and population composition were assessed to see which parameter(s) could influence those. Diversity (richness) of *V. cholerae* alleles was significantly correlated with salinity, conductivity, total dissolved solids (TDS) and dissolved oxygen (DO) (Fig. 4). Salinity, conductivity and TDS were found to be positively correlated with allele richness (Pearson correlation coefficient 0.44, 0.41 and 0.40 respectively; $P < 0.01$), while DO (mg/L) showed a negative correlation (Pearson correlation -0.30; $P<0.05$) with that population

characteristic. In previous studies, various physicochemical variables have been found to be associated with the abundance and persistence of *V. cholerae* in aquatic environments and with the risk of cholera outbreaks [43, 44]. Temperature and salinity have been observed to influence planktonic populations, which is a well-known habitat for *V. cholerae* in the aquatic ecosystem [41]. Adaptation to a wide range of salinity levels also facilitates *V. cholerae*'s survival in various aquatic environments (from coastal to inland water) [42, 43]. Previous studies suggest that abundance of *V. cholerae* decreases with increasing salinity and that they are most abundant in salinities ranging from 0 to 10 ppt [42, 43].

Changes in subspecies level population composition were also correlated with variation in environmental factors. Even though no strong correlation was found with the abundance of individual *V. cholerae* lineages and environmental conditions, there were moderate correlations observed in case of some *viuB* genotypes. Based on Pearson correlation co-efficient values, abundance of *viuB*-73, *viuB*-06 and *viuB*-07 were positively correlated with salinity, conductivity and TDS, whereas that of *viuB*-39 was negatively correlated with these variables (Supplementery sheet 1). This suggests that changes in the abundance of these four genotypes drove the changes in population diversity. In a study of the *V. cholerae* population in cholera free Oyster Pond coastal ecosystem (Falmouth, MA, USA), the *viuB*-73 allele was rarely found in the ocean [21], whereas it was present in the brackish pond and lagoon water connected to the ocean, suggesting that high salinity might represent an environmental barrier to the dispersal and range of cholera. In this study, salinity appeared to be the most significant parameter to explain the distribution of *viuB* alleles (Figure 3). In the NMDS plot (stress 0.13 and $P < 0.01$), generated by plotting the abundance of *viuB* alleles against the direction of change in the environmental variables, It appears that the environmental gradient of salinity significantly impacts the community structure. Among the major alleles, salinity gradient was positively associated with *viuB*-07 and negatively associated with *viuB*-73.

The negative correlation between *viuB*-73 abundance and salinity (Pearson correlation coefficient -0.31) observed in this study also suggest that PG *V. cholerae* (*viuB*-73) might be adapted to low salinity . The PG lineage represented by *viuB*-73 might have higher tolerance of rapid variations at low salinities, as it was consistently predominating in all but one site of Dhaka where salinity levels fluctuated between 0-0.8 ppt. The only site at which *viuB*-73 was not dominant was site 7 (Fig. 5), in which salinity was more stable and only varied from 0.4 ppt to 0.5 ppt, and never dropped below 0.4 ppt (Fig. 3). This suggests that even at the subspecies level, small environmental fluctuations could have an influence on population composition of *V. cholerae*. In the low-lying Ganges delta, salinity intrusion is considered a major threat due to the changing climate. Reduced upstream discharge, sea level rise and other catastrophic events such as cyclones can lead to increase in the salinity of inland water bodies [43]. A salinity increase of about 26% was recorded in coastal regions in Bangladesh over the last 35 years [45]. Such a shift in salinity could be affecting the composition of *V. cholerae* populations, possibly changing the distribution and abundance of various lineages, some of which could pose a threat to human health.

Dissolved oxygen (DO), a measure of free non-compound oxygen present in an aquatic system, is another variable influencing microbial communities [46]. A study in Hood Canal, Washington, USA, demonstrated that there was a strong negative correlation between bacterial richness and DO [47]. So far, there are no

studies describing the correlation between the intraspecies diversity of *V. cholerae* and DO, but it was demonstrated earlier that *V. cholerae* is most abundant in low DO environments [48, 49]. In our analysis, DO was found to impact the *V. cholerae* diversity negatively in the Dhaka environment (Fig. 4). DO could also be one of the factors contributing to the differential adaptation at the subspecies level. Among the sampling locations, site 7 had the lowest average DO (0.90 mg/L). Average DO in site 1 was also low at 0.95 mg/L, but the other five sites (site 2-6) had noticeably higher average DO concentration ranging from 2.7 to 3.0 mg/L. Conductivity, salinity and TDS usually have a negative relationship with DO, which might have caused the lower DO observed in site 7 and site 1. Site 7 was the only one where *viuB* -73 was outcompeted in abundance by other *viuB* alleles (*viuB*-05 and *viuB*-06). *V. cholerae* being a facultative anaerobe, can adapt to the low-oxygen conditions by utilizing alternative energy-producing pathways (i.e. nitrate utilization) [50]. It is possible that some lineages are advantaged at lower oxygen concentrations, giving them the ability to co-exist with other lineages that usually outcompete them, leading to increased diversity at lower DO.

Temperature is another factor affected by climate change that is known to have an impact on *V. cholerae* populations. While this species is found at a wide range of temperatures (10 to 30 °C), the highest abundance is observed at >20 °C [42]. However, no significant correlation between temperature and *viuB* allele diversity (Pearson correlation coefficient 0.1) were observed in Dhaka (Fig. 4). Water temperatures remained within the range of 27.4 °C to 30.8 °C throughout the sampling period, a pattern that differs dramatically from regions where shifts in temperature are noticeable during summer and winter [51]. In a temperate region, attachment to particles and hosts has been shown to increase when temperature increases above 22 °C, contributing to changes in the lineage composition of a *V. cholerae* population when a seasonal change occurs in a temperate climate [21]. It is likely that at conditions with consistent high temperatures such as those found in a tropical region like Bangladesh, *V. cholerae* is not overly responsive to this parameter, either in terms of its growth rate or particle attachment behavior.

High human population density correlates with changes in the *V. cholerae* population of Dhaka reservoirs

Dhaka is one of the most densely populated areas in the world, with a density of 23,234 people per square kilometer within a total area of 300 square kilometers. This huge burden of human population has significant impact on the ecology and evolution of *V. cholerae* and consequently epidemiology of cholera in this region [52, 53]. Based on the demographic records of Dhaka city, the area surrounding sampling site 7 (Kamrangir char) is the most densely populated (100,000 people/km^2) among the seven sites studied, whereas population density at the other six sites ranged from 10,000/km^2 to 60,000/km^2 (Fig. 1) [54]. Thus, human impacts on the water reservoir in this area is expected to be much higher than other sites, with a higher level of fecal contamination and industrial waste mostly from tannery industrial units. This area is mostly inhabited by a dense low-income population where people use shared hygiene facilities such as showers and toilets [55]. Open defecation has been reported from poorly maintained shared facilities. Additionally, high population density in this area frequently causes an overload of septic tanks, which results in the overflow of untreated effluent to the water reservoir [56]. The open drainage system commonly causes mixing between sewage and fresh water, increasing the possibility of *V.*

cholerae transmission between the water reservoir and local population. Amongst the seven different sites studied, most of them had TDS concentration of < 300mg/L on average, except sites 1 and 7, where TDS varied from 308 to 472 mg/L and from 476 to 575 mg/L, respectively (Table S3). TDS value comes from the combination of the disassociated electrolytes and other compounds such as dissolved organic matter [57]. Typically, natural bodies of water have dissolved solids due to the dissolution and weathering of rocks and soil. However, human activities also influence the concentration of TDS in water reservoirs, which is especially likely in Dhaka's inland water bodies because of urban runoff as well as wastewater discharge. This heavy influence of human population at Kamrangir char likely led to different population dynamics of *V. cholerae* at this site compared to other locations in the city. However, in our sampling sites, conductivity, salinity and TDS were found to be tightly correlated indicating inorganic ion might be the major contributor of the variation in TDS values. Hence, major cause of higher conductivity, salinity and TDS values observed in site 7 can also be indicators of increased chemical pollution there in comparison to other sites. Noticable difference in the diversity and population composition in site 7, might suggest that human population density can have a direct or indirect impact on the *V. cholerae* population at the subspecies level.

Although absolute abundance of total *V. cholerae* remained stable at all seven sites throughout the six months sampling period, striking differences appeared in case of site 7, where total *V. cholerae* abundance was around 28% to 43% higher than any other site (Fig. 2). It is unclear if this higher abundance at site 7 is due to a more constant input of *V. cholerae* from human waste, an indirect increase in numbers from nutrients linked to this waste, or the existence of a niche allowing the expansion of a particular lineage that does not compete with others (sympatry). Unlike sites 1 to 6, at which *viuB*-73 was more abundant than all other alleles combined (Fig. 5), three *viuB* alleles (*viuB*-05, *viuB*-06 and *viuB*-07) had greater combined abundance than *viuB*-73 at site 7. The trend is most pronounced during December 2015 to March 2016, coinciding with a decrease of water quality in the reservoir during December to April [58]. These lineages have so far not been found in environmental surveys outside of Dhaka. Interestingly, these three alleles (*viuB*-05, *viuB*-06 and *viuB*-07) have been found to be associated with strains representing a basal long branch clade in whole-genome phylogeny of known global *V. cholerae* strains [59]. Representative isolates of this clade were found to be indistinguishable from typical *V. cholerae* based on conventional phenotypic tests, but they are phylogenetically and genotypically divergent [60].

Allelic differences and *in-silico* DDH of whole genome sequences show that organisms represented by these three alleles (*viuB*-05, *viuB*-06 and *viuB*-07) are not more closely related to each other on average than most pairs of *V. cholerae* strains would be (Fig. S1). Despite this lack of genetic similarity, their presence and absence are strongly correlated, as they are usually co-occurring (Fig. 6 and Fig. S1). Surprising dominance of the *viuB*-05,06 and 07 lineages over *viuB*-73 exclusively at the Kamrangir char location suggests that this group of lineages is specifically adapted to the environmental conditions found at this site. As the site also stood out from a human population density point of view, and that these lineages were most abundant in this site, it suggests a potential human link to the ecology of these lineages in Dhaka.. Indeed, strains phylogenetically related with these lineages have been isolated from

human samples in different parts of the world [59]. Abundance of those lineages was positively correlated with TDS, conductivity and salinity, which can be considered as indicators of the impact of human intrusion in the water. Among the seven sampling sites, these three related parameters were highest on average at site 7. These three alleles (*viuB*-05, 06, 07) were most abundant at site 7 (54% of total *V. cholerae*), where *viuB*-73 had the lowest abundance among all sites (24%). This suggests that some form of competition could be taking place between different *V. cholerae* lineages and/or that the lineages respond differently to environmental factors present at this site. Possible link of the human population to the abundance and distribution of these interrelated lineages makes them compelling candidates for future studies looking into the ecology of human adapted genotypes of *V. cholerae*.

Intraspecies interaction could influence relative abundance of PG *V. cholerae* O1

Even though overall *viuB*-73 was the predominant allele over the six months sampling period, spatial and temporal analysis indicate that direct or indirect intraspecies competition among *V. cholerae* genotypes could play an important role in the population dynamics observed in Dhaka. Whenever *viuB*-73 displayed a drop in abundance, another lineage carrying a *viuB*-05, *viuB*-06, *viuB*-07, *viuB*-25, *viuB*-45, *viuB*-51 or *viuB*-79 allele underwent a rapid expansion. Although the abundance of these other alleles was not directly quantified by qPCR, the absolute abundance measurement of the total *V. cholerae* and PG O1 *V. cholerae* by qPCR (Fig. 2) supported the observation coming from amplicon sequencing data. Furthermore, this shift in abundances is repeated multiple times within our dataset. For example, *viuB*-25 increased in relative abundance when the *viuB*-73 abundance decreased at sites 1 and 7 (Fig. 5). Notably, a new *viuB* allele, *viuB*-79 (for which no cultured isolates have been found), also appeared when *viuB*-73 abundance was reduced at sites 6 and 7 (Fig. 5). All of these alleles showed an inverse correlation with *viuB*-73 relative abundance (Fig. 6). Hierarchical clustering analysis of *viuB* amplicon sequencing data (Fig. 6) showing statistically significant clustering (Hopkins statistics [H] < 0.5) only occurred between *viuB* alleles representing strains from the basal divergent lonb branch clade (*viuB*-05, *viuB*-06 and *viuB*-07) (H <0.5). *viuB*-73 showed significant negative correlation with most other *viuB* alleles (H >0.5) [61]. The only *viuB* alleles positively correlated with *viuB*-73 were *viuB*-10 and *viuB*-39 (Fig. 6), the latter being a ubiquitous allele present at low abundance across sampling sites as well as other geographical locations (Fig. 5).

The cause of these shifts in abundance of some alleles is unclear. It could be due to a differential response to environmental factors or trophic interactions. In the Oyster Pond (MA, USA) *V. cholerae* population, divergent responses of different lineages were observed in ocean, lagoon and pond, where ecosystem parameters varied substantially [40]. It is plausible that *V. cholerae* lineages responded differently to even the slight environmental variations observed at the sites sampled in Dhaka. Differential response of lineages to phage predation can also be a cause of shifts in the community composition. Effect of predation by bacteriophages on the *V. cholerae* population composition can be modulated by the environmental factors i.e. nutrient availability [61], which in turn can give advantage to certain lineages to outcompete others under certain environmental conditions.

Another factor influencing this differential response could be the difference in the ability to avoid predation. The abundance of *V. cholerae* is influenced by grazing from heterotrophic protists [41]. To overcome the grazing pressure *V. cholerae* executes different strategies such as morphological shift, i.e. from smooth to rugose, resulting in the production of VPS (*Vibrio* polysaccharide) that helps to encase themselves in biofilm and resist predation [62]. *V. cholerae* can also survive predation by becoming intracellular in a range of amoeba [63]. They are also able to kill grazers using T6SS [64]. This system encodes a syringe-like structure that can pierce cellular envelopes of other bacteria and some eukaryotes, injecting effector proteins that can kill the recipient if it does not process the cognate immunity protein [65]. This phenomenon could influence the population composition of *V. cholerae*, with different lineages having varying effectors and immunity proteins providing varying predatory success [65, 66]. Another possibility is that incompatibility between subspecies likely plays a role in the diversity and dynamic of the *V. cholerae* populations in Dhaka. Most of these genotypes T6SS effector and immunity protein profiles suggest they are incompatible and can kill each other on contact (Hussain et al., Unpublished). Even though the PG lineage carrying *viuB*-73 could seemingly outcompete strains represented by all other alleles at sites 1-6, it was outcompeted by the three co-occurring alleles corresponding to the *viuB*-05, 06 and 07 group at site 7. This site differs from the other six sites in terms of environmental parameters and surrounding human population density. These observations suggest that environmental conditions can play an important role in shaping the intra-species competition of *V. cholerae* in their natural environment to impact diversity and subspecies level population dynamics of the *V. cholerae*.

Conclusion

For centuries, the Ganges delta has been a reservoir for pandemic and non-pandemic *V. cholerae* lineages. Cholera endemicity is usual in this area, and Dhaka is one of the most densely populated megacities in the world, with cholera epidemics occurring biannually before and after the rainy season. This study revealed the consistent presence of PG *V. cholerae* lineage in the natural water bodies of Dhaka at a considerable proportion of the total *V. cholerae* population, suggesting that the pathogenic bacterium is circulating year-round in this urban aquatic environment, or is constantly shed by human carriers.

Moreover, population analysis with subspecies level resolution revealed that other *V. cholerae* lineages were coexisting with PG *V. cholerae* O1 in that environment. Intraspecies niche specialization and potentially subspecies interactions with these other lineages could decrease PG *V. cholerae* O1 abundance occasionally, especially when environmental parameters, such as consistently higher TDS, salinity or lower dissolved oxygen concentration favored other lineages. Variability of these parameters, even on a small scale, correlated with changes in *V. cholerae* population composition and diversity, with more temporal stability of salinity and TDS at one particular site correlating with a dramatically different lineage composition. The same site also displayed the highest human population density among the seven sampling locations, indicating humans as a possible contributing source of organic and inorganic materials, or of *V. cholerae* themselves, altering *V. cholerae* population composition. Consistent human interaction and frequent leaking of sewage into the water being linked to a substantially different lineage

composition would suggest that human gut may also serves as a potential reservoir for PG O1 and other lineages of *V. cholerae,* resulting in year-long persistence of *V. cholerae* belonging to these lineages through the transmission cycle between human and aquatic environments. A study of the microbiomes of individuals living near water reservoirs is essential to identify if humans play a direct or indirect role in the differences observed between the aquatic *V. cholerae* populations found at different Dhaka sites.

Declarations

Funding

This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Integrated Microbial Biodiversity program of the Canadian Institute for Advanced Research to YB. We also acknowledge the support of graduate student scholarships from NSERC to TN, Alberta Innovates Technology-Futures and Faculty of Graduate Studies and Research to TN, MTI and KYHL.

Acknowledgements

YB, TN, MTI and MA designed the project. TN, MTI and YB wrote the manuscript. MTI, FTJ, MS and MA helped performed sample collection and sample processing during field trips in Dhaka, Bangladesh. TN performed the qPCR and MTI did the amplicon sequencing. TN and MTI did the data analysis. KYHL helped in bioinformatics analysis. MTI, YB, FTJ, MS, RJC and MA reviewed the manuscript. MA and YB supervised the project. MA of icddr, b thanks the government of Bangladesh, Canada, Sweden and United Kingdom for providing unrestricted core support.

Conflicts of interest

The authors declare that there are no conflicts of interest.

Availability of data and material

All the genome sequences and relevant genomic and epidemiological data related to the isolates used are publicly available on PubMLST (https://pubmlst.org/vcholerae/)

References

1. Colwell RR (2004) Infectious disease and environment: Cholera as a paradigm for waterborne disease. Int. Microbiol.
2. Vezzulli L, Pruzzo C, Huq A, Colwell RR (2010) Environmental reservoirs of Vibrio cholerae and their role in cholera. Environ. Microbiol. Rep.
3. Lipp EK, Huq A, Colwell RR (2002) Effects of global climate on infectious disease: The cholera model. Clin. Microbiol. Rev.

4. Ali M, Lopez AL, Ae You Y, et al (2012) The global burden of cholera. Bull World Health Organ. https://doi.org/10.2471/blt.11.093427
5. Faruque SM, Albert MJ, Mekalanos JJ (1998) Epidemiology, Genetics, and Ecology of ToxigenicVibrio cholerae. Microbiol Mol Biol Rev. https://doi.org/10.1128/mmbr.62.4.1301-1314.1998
6. Chatterjee SN, Chaudhuri K (2003) Lipopolysaccharides of Vibrio cholerae: I. Physical and chemical characterization. Biochim Biophys Acta - Mol Basis Dis 1639:65–79. https://doi.org/10.1016/j.bbadis.2003.08.004
7. Chun J, Grim CJ, Hasan NA, et al (2009) Comparative genomics reveals mechanism for short-term and long-term clonal transitions in pandemic Vibrio cholerae. Proc Natl Acad Sci U S A 106:15442–15447. https://doi.org/10.1073/pnas.0907787106
8. Islam MT, Alam M, Boucher Y (2017) Emergence, ecology and dispersal of the pandemic generating Vibrio cholerae lineage. Int. Microbiol.
9. Chun J, Grim CJ, Hasan NA, et al (2009) Comparative genomics reveals mechanism for short-term and long-term clonal transitions in pandemic Vibrio cholerae. Proc Natl Acad Sci U S A. https://doi.org/10.1073/pnas.0907787106
10. Mutreja A, Kim DW, Thomson N, et al (2013) Evidence for multiple waves of global transmission within the seventh cholera pandemic. Nature. https://doi.org/10.1038/nature10392.Evidence
11. Boucher Y, Orata FD, Alam M (2015) The out-of-the-delta hypothesis: Dense human populations in low-lying river deltas served as agents for the evolution of a deadly pathogen. Front Microbiol. https://doi.org/10.3389/fmicb.2015.01120
12. Longini IM, Yunus M, Zaman K, et al (2002) Epidemic and endemic cholera trends over a 33-year period in Bangladesh. J Infect Dis. https://doi.org/10.1086/341206
13. Deen J, Mengel MA, Clemens JD (2020) Epidemiology of cholera. Vaccine
14. Jutla A, Whitcombe E, Hasan N, et al (2013) Environmental factors influencing epidemic cholera. Am J Trop Med Hyg. https://doi.org/10.4269/ajtmh.12-0721
15. Alam M, Sultana M, Balakrish Nair G, et al (2007) Viable but nonculturable Vibrio cholerae O1 in biofilms in the aquatic environment and their role in cholera transmission. Proc Natl Acad Sci U S A. https://doi.org/10.1073/pnas.0705599104
16. Alam M, Hasan NA, Sultana M, et al (2010) Diagnostic limitations to accurate diagnosis of cholera. J Clin Microbiol 48:3918–3922. https://doi.org/10.1128/JCM.00616-10
17. Colwell RR (2000) Viable but nonculturable bacteria: A survival strategy. J. Infect. Chemother.
18. Rashid R Bin, Ferdous J, Tulsiani S, et al (2017) Development and validation of a novel real-time assay for the detection and quantification of Vibrio cholerae. Front Public Heal. https://doi.org/10.3389/FPUBH.2017.00109
19. Chun J, Huq A, Colwell RR (1999) Analysis of 16S-23S rRNA intergenic spacer regions of Vibrio cholerae and Vibrio mimicus. Appl Environ Microbiol. https://doi.org/10.1128/aem.65.5.2202-

2208.1999

20. Chowdhury MAR, Xu B, Montilla R, et al (1995) A simplified immunofluorescence technique for detection of viable cells of Vibrio cholerae O1 and O139. J Microbiol Methods. https://doi.org/10.1016/0167-7012(95)00066-6
21. Kirchberger PC, Orata FD, Nasreen T, et al (2020) Culture-independent tracking of Vibrio cholerae lineages reveals complex spatiotemporal dynamics in a natural population. Environ Microbiol. https://doi.org/10.1111/1462-2920.14921
22. Wyckoff EE, Mey AR, Payne SM (2007) Iron acquisition in Vibrio cholerae. In: BioMetals
23. Nasreen T, Hussain NAS, Islam MT, et al (2019) Simultaneous quantification of Vibrio metoecus and Vibrio cholerae with its O1 serogroup and toxigenic subpopulations in environmental reservoirs. bioRxiv 822551. https://doi.org/10.1101/822551
24. Wright JJ, Lee S, Zaikova E, et al (2009) DNA extraction from 0.22 μM Sterivex filters and cesium chloride density gradient centrifugation. J Vis Exp. https://doi.org/10.3791/1352
25. Kozich JJ, Westcott SL, Baxter NT, et al (2013) Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the miseq illumina sequencing platform. Appl Environ Microbiol. https://doi.org/10.1128/AEM.01043-13
26. Ripley BD (2001) The R Project in Statistical Computing. MSOR Connect. https://doi.org/10.11120/msor.2001.01010023
27. Callahan BJ, McMurdie PJ, Rosen MJ, et al (2016) DADA2: High-resolution sample inference from Illumina amplicon data. Nat Methods. https://doi.org/10.1038/nmeth.3869
28. Nasreen T, Hussain NAS, Islam MT, et al (2020) Simultaneous quantification of vibrio metoecus and vibrio cholerae with its o1 serogroup and toxigenic subpopulations in environmental reservoirs. Pathogens. https://doi.org/10.3390/pathogens9121053
29. Torondel B, Ensink JHJ, Gundogdu O, et al (2016) Assessment of the influence of intrinsic environmental and geographical factors on the bacterial ecology of pit latrines. Microb Biotechnol. https://doi.org/10.1111/1751-7915.12334
30. Angiuoli S V., Salzberg SL (2011) Mugsy: Fast multiple alignment of closely related whole genomes. Bioinformatics. https://doi.org/10.1093/bioinformatics/btq665
31. Stamatakis A (2014) RAxML. Manual/tutorial
32. Letunic I, Bork P (2007) Interactive Tree Of Life (iTOL): An online tool for phylogenetic tree display and annotation. Bioinformatics. https://doi.org/10.1093/bioinformatics/btl529
33. Auch AF, Klenk HP, Göker M (2010) Standard operating procedure for calculating genome-to-genome distances based on high-scoring segment pairs. Stand Genomic Sci. https://doi.org/10.4056/sigs.541628
34. Meier-Kolthoff JP, Auch AF, Klenk HP, Göker M (2013) Genome sequence-based species delimitation with confidence intervals and improved distance functions. BMC Bioinformatics. https://doi.org/10.1186/1471-2105-14-60

35. Aziz RK, Bartels D, Best A, et al (2008) The RAST Server: Rapid annotations using subsystems technology. BMC Genomics. https://doi.org/10.1186/1471-2164-9-75
36. Edgar CRC (2013) USEARCH. J Natl Cancer Inst
37. Jolley KA, Maiden MCJ (2010) BIGSdb: Scalable analysis of bacterial genome variation at the population level. BMC Bioinformatics. https://doi.org/10.1186/1471-2105-11-595
38. Alam M, Islam A, Bhuiyan N, et al (2011) Clonal transmission, dual peak, and off-season cholera in Bangladesh. Infect Ecol Epidemiol. https://doi.org/10.3402/iee.v1i0.7273
39. Kirchberger PC, Orata FD, Barlow EJ, et al (2016) A small number of phylogenetically distinct clonal complexes dominate a coastal Vibrio cholerae population. Appl Environ Microbiol. https://doi.org/10.1128/AEM.01177-16
40. Kirchberger PC, Orata FD, Nasreen T, et al (2020) Culture-independent tracking of Vibrio cholerae lineages reveals complex spatiotemporal dynamics in a natural population. Env Microbiol. https://doi.org/10.1111/1462-2920.14921
41. Lutz C, Erken M, Noorian P, et al (2013) Environmental reservoirs and mechanisms of persistence of Vibrio cholerae. Front. Microbiol.
42. Takemura AF, Chien DM, Polz MF (2014) Associations and dynamics of vibrionaceae in the environment, from the genus to the population level. Front. Microbiol.
43. Huq A, Sack RB, Nizam A, et al (2005) Critical factors influencing the occurrence of Vibrio cholerae in the environment of Bangladesh. Appl Environ Microbiol. https://doi.org/10.1128/AEM.71.8.4645-4654.2005
44. Stine OC, Alam M, Tang L, et al (2008) Seasonal cholera from multiple small outbreaks, rural Bangladesh. Emerg Infect Dis. https://doi.org/10.3201/eid1405.071116
45. Mahmuduzzaman M, Ahmed ZU, Nuruzzaman AKM, Ahmed FRS (2014) Causes of Salinity Intrusion in Coastal Belt of Bangladesh. Int J Plant Res
46. Aldunate M, De la Iglesia R, Bertagnolli AD, Ulloa O (2018) Oxygen modulates bacterial community composition in the coastal upwelling waters off central Chile. Deep Res Part II Top Stud Oceanogr. https://doi.org/10.1016/j.dsr2.2018.02.001
47. Spietz RL, Williams CM, Rocap G, Horner-Devine MC (2015) A dissolved oxygen threshold for shifts in bacterial community structure in a seasonally hypoxic estuary. PLoS One. https://doi.org/10.1371/journal.pone.0135731
48. Blackwell KD, Oliver JD (2008) The ecology of Vibrio vulnificus, Vibrio cholerae, and Vibrio parahaemolyticus in North Carolina Estuaries. J Microbiol. https://doi.org/10.1007/s12275-007-0216-2
49. León Robles A, Acedo Félix E, Gomez-Gil B, et al (2013) Relationship of aquatic environmental factors with the abundance of Vibrio cholerae, Vibrio parahaemolyticus, Vibrio mimicus and Vibrio vulnificus in the coastal area of Guaymas, Sonora, Mexico. J Water Health. https://doi.org/10.2166/wh.2013.160

50. Bueno E, Pinedo V, Cava F (2020) Adaptation of Vibrio cholerae to Hypoxic Environments. Front. Microbiol.
51. Jiang SC, Fu W (2001) Seasonal abundance and distribution of Vibrio cholerae in coastal waters quantified by a 16S-23S intergenic spacer probe. Microb Ecol. https://doi.org/10.1007/s00248-001-0029-5
52. Islam MT, Alam M, Boucher Y (2017) Emergence, ecology and dispersal of the pandemic generating Vibrio cholerae lineage. Int Microbiol 20:106–115. https://doi.org/10.2436/20.1501.01.291
53. Boucher Y, Orata FD, Alam M (2015) The out-of-the-delta hypothesis: dense human populations in low-lying river deltas served as agents for the evolution of a deadly pathogen. Front Microbiol 6:1120. https://doi.org/10.3389/fmicb.2015.01120
54. Hafiza Khatun, Nishat Falgunee MJRK (2017) Analyzing urban population density gradient of Dhaka Metropolitan. Geogr OnlineTM Malaysian J Soc Sp 11:1–13
55. Nelson KB, Karver J, Kullman C, Graham JP (2014) User perceptions of shared sanitation among rural households in Indonesia and Bangladesh. PLoS One. https://doi.org/10.1371/journal.pone.0103886
56. Mansour G, Esseku H (2017) Situation analysis of the urban sanitation sector in Ghana. WSUP - Water Sanit Urban Poor
57. Rusydi AF (2018) Correlation between conductivity and total dissolved solid in various type of water: A review. In: IOP Conference Series: Earth and Environmental Science (2018)
58. Hasan MK, Shahriar A, Jim KU (2019) Water pollution in Bangladesh and its impact on public health. Heliyon. https://doi.org/10.1016/j.heliyon.2019.e02145
59. Islam MT, Liang K, Im MS, et al (2018) Draft Genome Sequences of Nine Vibrio sp. Isolates from across the United States Closely Related to Vibrio cholerae. Microbiol Resour Announc. https://doi.org/10.1128/mra.00965-18
60. Islam MT, Liang K, Im MS, et al (2018) Draft Genome Sequences of Nine Vibrio sp. Isolates from across the United States Closely Related to Vibrio cholerae. Microbiol Resour Announc 7:. https://doi.org/10.1128/MRA.00965-18
61. Lawson RG, Jurs PC (1990) New Index for Clustering Tendency and Its Application to Chemical Problems. J Chem Inf Comput Sci. https://doi.org/10.1021/ci00065a010
62. Matz C, Kjelleberg S (2005) Off the hook - How bacteria survive protozoan grazing. Trends Microbiol. https://doi.org/10.1016/j.tim.2005.05.009
63. Abd H, Weintraub A, Sandström G (2004) Interaction between Vibrio cholerae and Acanthamoeba castellanii. Microb Ecol Health Dis. https://doi.org/10.1080/08910600410029190
64. MacIntyre DL, Miyata ST, Kitaoka M, Pukatzki S (2010) The Vibrio cholerae type VI secretion system displays antimicrobial properties. Proc Natl Acad Sci U S A. https://doi.org/10.1073/pnas.1012931107

65. Unterweger D, Miyata ST, Bachmann V, et al (2014) The vibrio cholerae type VI secretion system employs diverse effector modules for intraspecific competition. Nat Commun. https://doi.org/10.1038/ncomms4549
66. Kirchberger PC, Unterweger D, Provenzano D, et al (2017) Sequential displacement of Type VI Secretion System effector genes leads to evolution of diverse immunity gene arrays in Vibrio cholerae. Sci Rep. https://doi.org/10.1038/srep45133

Figures

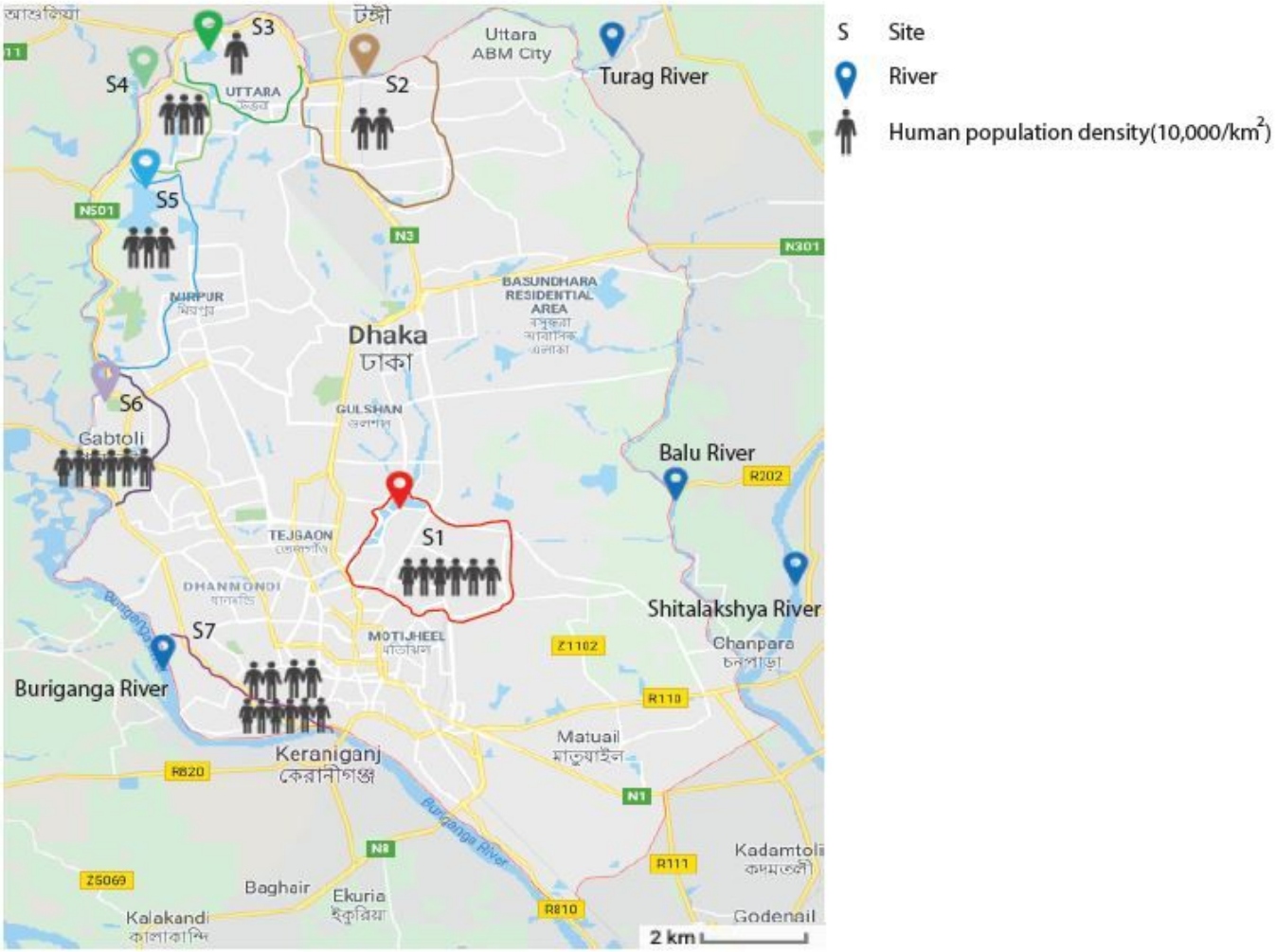

Figure 1

Location of sampling sites in the study. Dhaka, the capital city of Bangladesh, is surrounded by the river system of four rivers including Turag river, Buriganga river, Shitalakshya river and Balu river, as indicated by blue pins on the map. Seven different sampling sites are indicated with 'S' (from S1 to S7) along with the approximate human population density corresponding to each site (marked by pins with distinct

colors). Information on the human population density in this figure was adapted from Khatun et al., 2017 [54]. Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

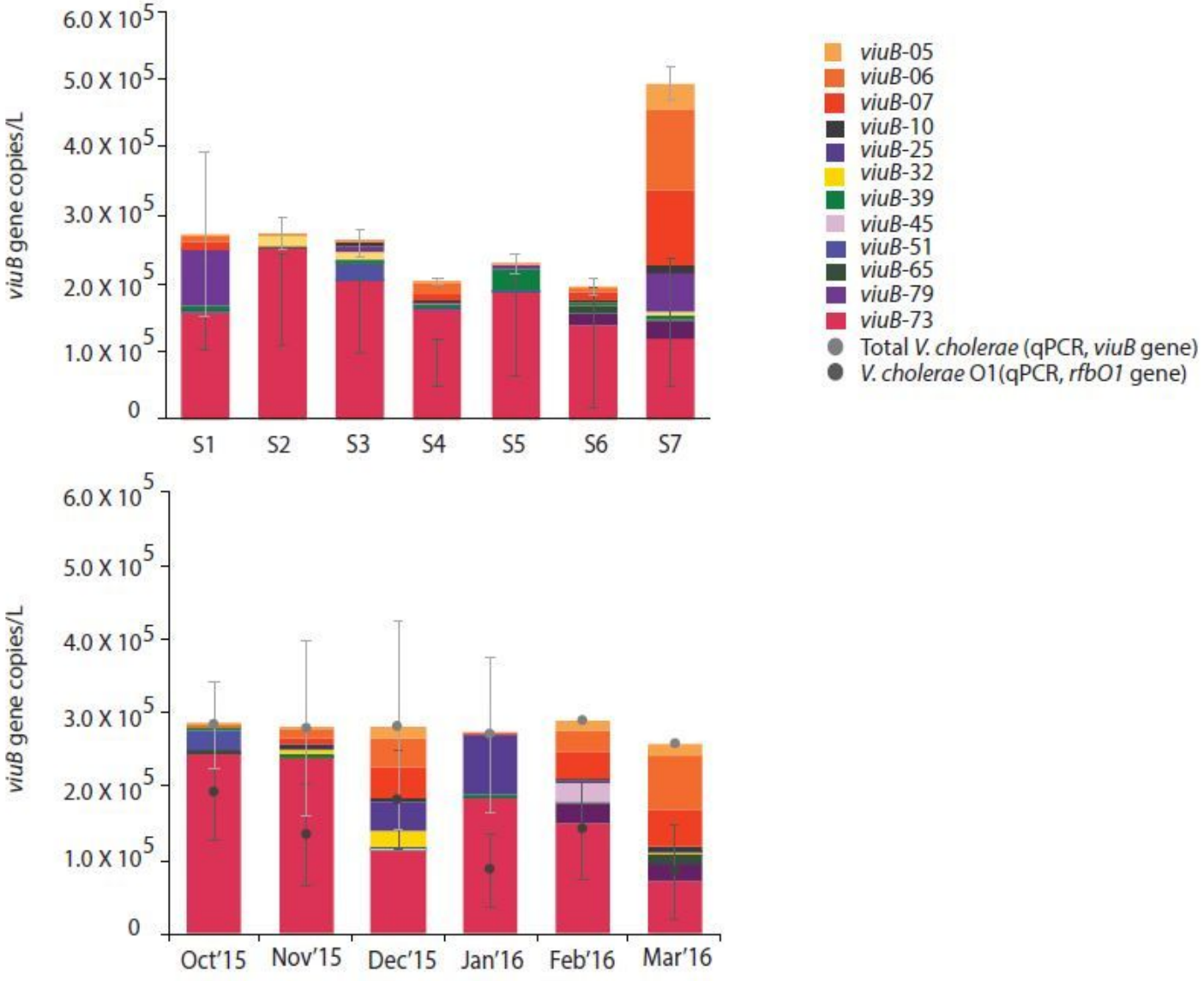

Figure 2

Spatial and temporal distribution of viuB alleles found in Dhaka, Bangladesh. A) Spatial variation of viuB alleles relative abundance. Each of which corresponds to a different V. cholerae lineage (Fig. S1). Error bars represent standard deviation from the average abundance at each site across the entire time series; B) Temporal variation of viuB alleles relative abundance. Error bars represent standard deviation from the average monthly abundance across all sites. Aquatic biomass was extracted from water samples collected from seven different sites around Dhaka, Bangladesh in six consecutive months (Oct 2015 to

Mar 2016). Total V. cholerae abundance was determined by qPCR amplification of the viuB gene marker and used to normalize the number of the 12 most abundant viuB alleles sequences representing >99% of all reads. Light grey dots denote the average count of total V. cholerae genomes copies at each site enumerated by qPCR. Proportions of different viuB alleles are denoted with colors specific to each allele. The relative abundance of viuB-73, a proxy for PG V. cholerae (mostly composed by O1 serogroup strains in Dhaka), was confirmed by qPCR with the amplification of rfbO1 gene (specific for the O1 serogroup), which is represented by dark grey dots.

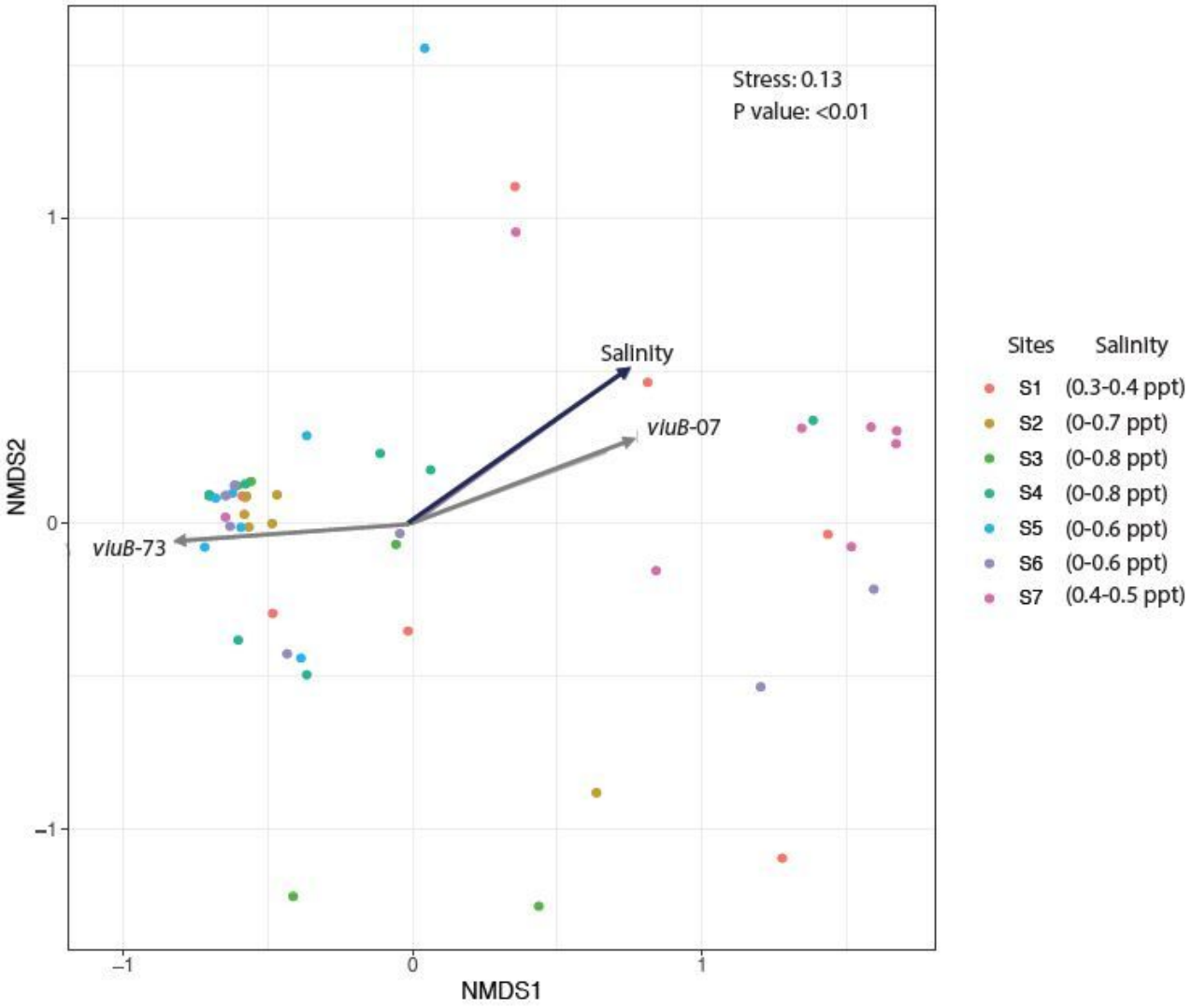

Figure 3

Correlation of environmental variables with V. cholerae population composition. Unique viuB allele sequences from aquatic biomass over six consecutive months collected from seven different sites in Dhaka were taken to represent twelve V. cholerae lineages. The vegan envfit function (stress = 0.13 and $p < 0.01$) was used to calculate the dissimilarity of V. cholerae populations based on their lineage composition and their correlation with recorded physicochemical parameters. Sites are indicated in different colors. Significant correlations are indicated by vectors, with a positive correlation observed

between abundance of viuB-07 and increasing salinity, whereas viuB-73 showed a negative correlation with this parameter.

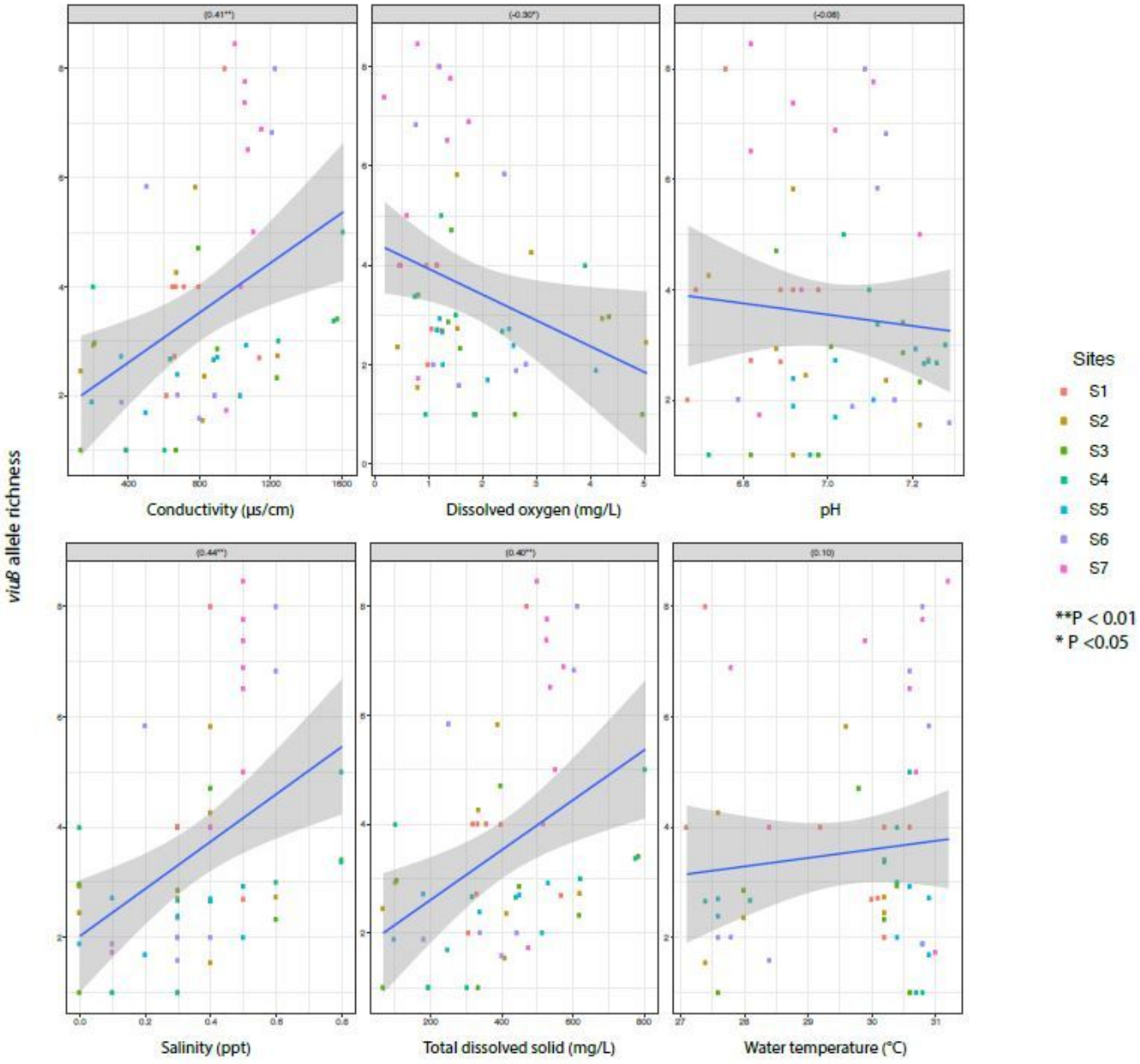

Figure 4

Correlation of environmental variables and viuB allele richness. A correlation test was performed between viuB allele richness and each environmental variable measured in R (vegan library's bioenv function). Each scatter plot shows a different environmental variable (x axis) and the corresponding viuB allele richness (y axis). Pearson's correlations are given in parentheses, and significance is indicated (*P<0.05 and **P<0.01) below each panel. Sites are indicated by different colors.

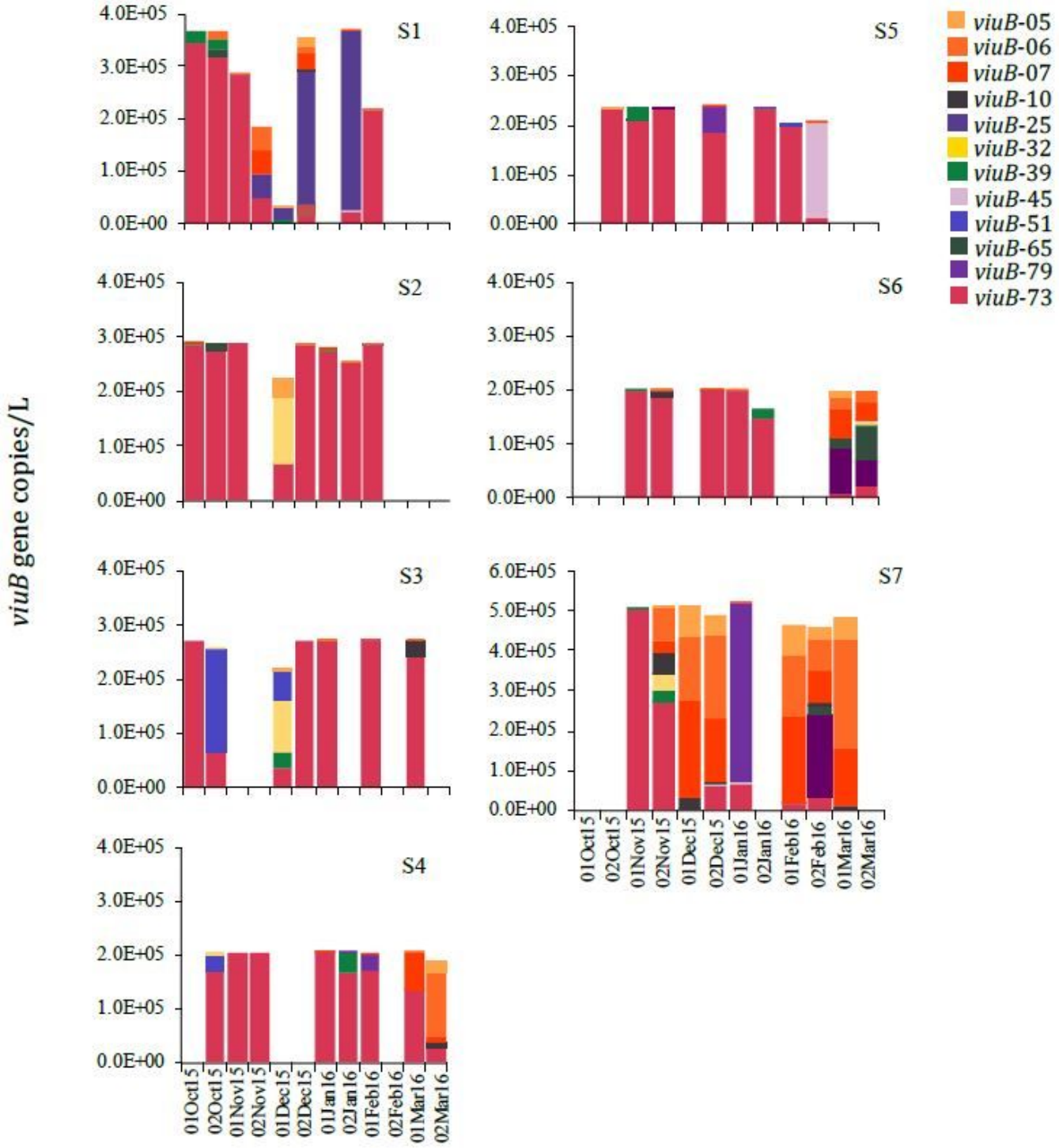

Figure 5

Abundance of V. cholerae lineages represented by different viuB alleles present at various water reservoirs in Dhaka, Bangladesh. Water samples were collected from seven different sites (S1-S7) in six consecutive months (Oct 2015 to Mar 2016). The partial V. cholerae specific viuB gene was amplified and sequenced from DNA extracted from the biomass of these samples. Relative abundance of viuB alleles representing various lineages is presented on a scale of absolute abundance of V. cholerae genome copies determined by qPCR of the viuB gene. Blank spaces indicate missing samples.

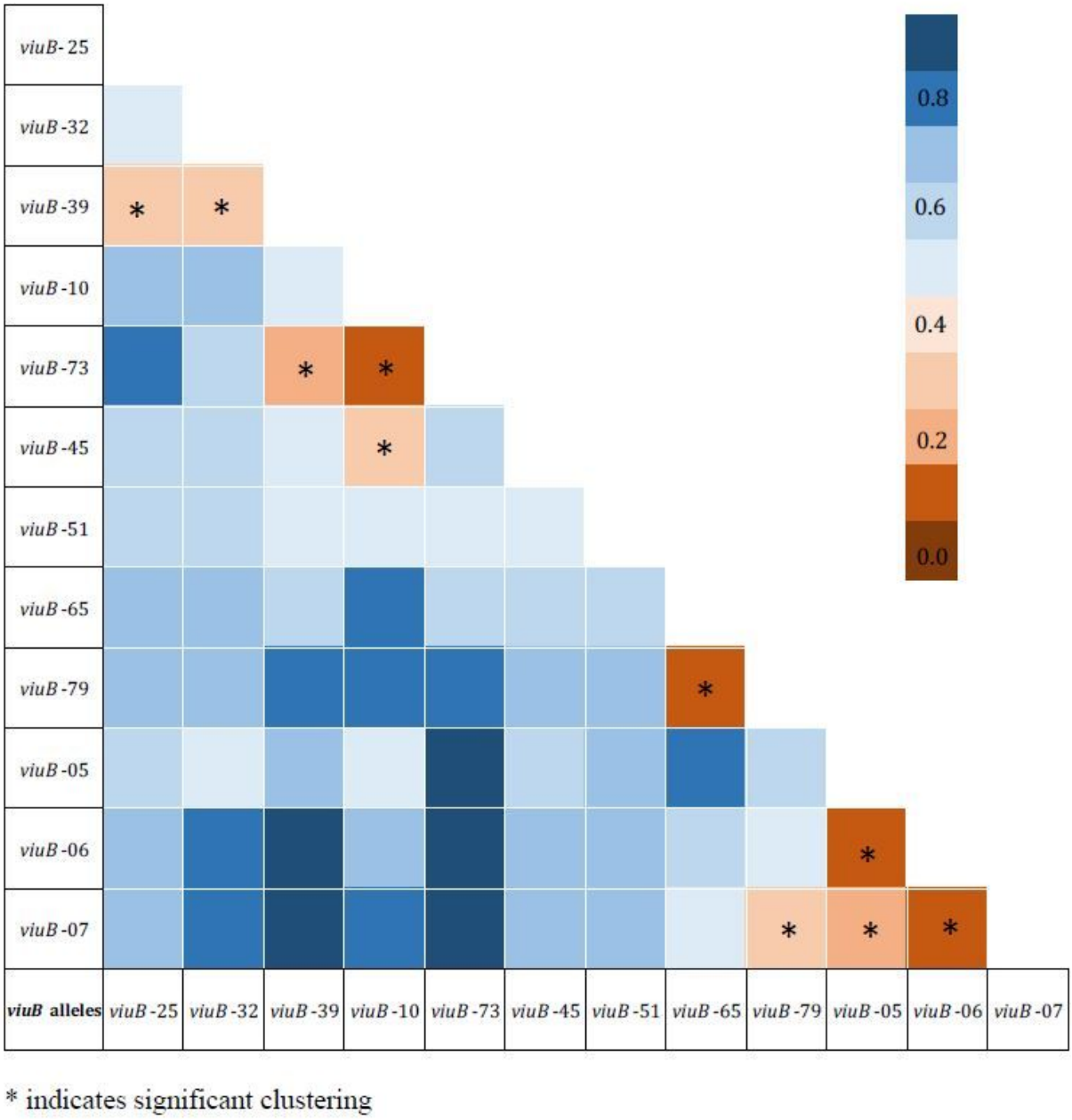

Figure 6

Co-occurrence of V. cholerae lineages in Dhaka water reservoirs. The Hopkins statistics (H) was used to assess the clustering tendency of various viuB allele sequences in the dataset. The threshold value was 0.5, meaning that if H < 0.5, data showed significant clustering. Visualization of analyzed viuB amplicon sequence data is presented with blue and brown color gradients (H < 0.5, indicating that the data is highly clusterable which is shown with a brown color gradient and H > 0.5, indicating that the data is not clusterable which is described by blue color gradient). Significant clustering (H < 0.5) is observed between viuB alleles 05, 06 and 07.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementarydata.docx
- Supplementerysheet1.xlsx