보건학석사 학위논문

High-throughput sequencing based analysis of bacterial diversity as well as *bla*$_{\text{CTX-M}}$ gene variants from river and hospital wastewater

High-throughput sequencing 을 기반으로 한 강물과 병원 폐수 내 박테리아 다양성과 *bla*$_{\text{CTX-M}}$ 유전자 변이 분석

2021 년 8 월

서울대학교 보건대학원

환경보건학과 환경보건학 전공

Bhuiyan Mohammad Imtiaj Uddin

High-throughput sequencing based analysis of bacterial diversity as well as *bla*$_{\text{CTX-M}}$ gene variants from river and hospital wastewater

High-throughput sequencing 을 기반으로 한 강물과 병원 폐수 내 박테리아 다양성과 *bla*$_{\text{CTX-M}}$ 유전자 변이 분석

지도교수 Naomichi Yamamoto

이 논문을 보건학석사 학위 논문으로 제출함

2021 년 8 월

서울대학교 보건대학원

훤경보건학과 환경보건학 전공

Bhuiyan Mohammad Imtiaj Uddin

Bhuiyan Mohammad Imtiaj Uddin 의 보건학석사 학위논문을 인준함

2021 년 8 월

위 원 장 최경호

부 위 원 장 조경덕

위 원 Naomichi Yamamoto

Abstract

High-throughput sequencing based analysis of bacterial diversity as well as *bla*$_{\text{CTX-M}}$ gene variants from river and hospital wastewater

Bhuiyan Mohammad Imtiaj Uddin

Department of Environmental Health

Graduate School of Public Health, Seoul National University

The aquatic environment of river water is the most intensively human influenced ecosystem in the world as it receives huge amount of urban, hospital, animal and industrial effluents and create a reservoir of diverse type of bacteria. This contaminated ecosystem provides optimum conditions to antibiotic resistant bacteria from various sources to mix and transfer their resistant gene(s) to clinically important bacteria for development of human pathogens with novel resistance mechanisms. Hospital wastewater, in

particular, is expected to contain high abundances of antibiotic resistance genes (ARGs) as it contains human enteric bacteria that may include antibiotic-resistant organisms originating from hospital patients, and can also have high concentrations of antibiotics and antimicrobials that facilitate the environmental spread of antibiotic resistance. Therefore, in this study, we applied 16S rRNA gene sequencing, a culture-independent method to analyze bacterial diversity present in river as well as hospital wastewater. River and hospital wastewater were collected from Dhaka, Bangladesh and sequenced for V3 and V4 region of 16S rRNA gene resulting in a total of 36 sequence libraries. In total 4,13,297 sequence reads were obtained where the number of sequence reads ranged from 1333 to 38110 reads per library. The taxonomic analysis revealed 8 phyla, 12 classes, 13 orders, 22 families and 102 genera present in the river and hospital wastewater samples. Additionally, permutational multivariate analysis of variance (PERMANOVA) of bacterial membership in terms of Bray-Curtis, based on operational taxonomic unit (OTUs) at a 97% sequence similarity threshold showed significant bacterial diversity between river and hospital wastewater. Moreover, compositions and relative abundance of the bacterial genera showed the presence of most pathogenic bacteria, mostly in the hospital wastewater. We performed another study to analyze $bla_{\text{CTX-M}}$ variants for their extensively emerging clinical importance. The presence of $bla_{\text{CTX-M}}$ gene in the bacteria producing extended spectrum beta lactamases (ESBLs) has made it resistant to the broad-spectrum

third-generation cephalosporin antibiotics that are used in hospitals to treat infections of the skin and respiratory tract caused by both Gram negative and Gram positive bacteria. The river and hospital wastewater were sequenced for ARG, resulting in a total of 33 sequence libraries. Out of 172 variants of *bla*$_{\text{CTX-M}}$ reported so far, our study found 124 variants, around 62% of which are present in hospital wastewater. The non-metric multidimensional scaling (NMDS) showed that there is close association among the *bla*$_{\text{CTX-M}}$ variants present in hospital wastewater whereas the variants for river water is scattered through the sampling sites. Overall, the findings of our study indicated that river and hospital wastewater contains diverse types of bacteria where the bacterial community from river water significantly differs from that of hospital wastewater. Both of the sampling site contains multidrug resistant bacterial genera that may pose serious threat to the aquatic environment of river water resulting in a serious public health burden. Moreover, the presence of significant number of *bla*$_{\text{CTX-M}}$ variants may facilitate the global burden of antimicrobial resistance.

Key words: Wastewater, Bacterial diversity, Antimicrobial resistance, Gene variants, Permutational multivariate analysis of variance, Non-metric multidimensional scaling.

Student Number: 2019-27155

Contents

List of Figures

List of Tables

1. INTRODUCTION

1.1 Global burden of antimicrobial resistance

Antimicrobial resistance (AMR) is a growing public health burden (Bokhary et al., 2021) that threatens the very core of modern medicine and the sustainability of an effective, global public health response to the enduring threat from infectious diseases. According to the World Health Organization (WHO), antimicrobial resistance (AMR) is one of the major public health hazards of the 21st century, as it is associated with both serious health issues and economic consequences (WHO, 2015). AMR has become an endemic and widespread problem that affects both high-income countries (HICs) and low- and medium-income countries (LMICs) (Hay et al., 2018). The World Bank Group suggests that AMR could cause low-income countries to lose more than 5% of their GDP and push up to 28 million people, mostly in developing countries, into poverty by 2050 (Group, 2016). AMR infections currently result in 700,000 global deaths every year with associated mortality estimated to claim 10 million lives per year by 2050 (Butler & Buss, 2006).

The development of antimicrobial resistance, which is partially attributable to the overuse and/or misuse of antibiotics in health care, is one of the greatest global public health challenges (Belachew et al., 2021). The uncontrolled use of antibiotics exposes bacteria to antibiotics, which leads to subsequent acquirement of resistance through mutations or by eliminating

nonresistant bacteria (Cantón & Morosini, 2011). The overuse of antibiotics and their subsequent poorly managed release to the environment have been linked with the development of such undesired characteristics (Sykes, 2010).

1.2 Antimicrobial resistance in Bangladesh

The southeast Asia is considered to have the highest risk of AMR among all the WHO regions (Chereau et al., 2017). Bangladesh is a southeast Asian developing country where antimicrobials are widely available as over-the-counter(OTC) drug (Ahmed & Hossain, 2007) and it results in a high degree of AMR that subsequently leads to a regional and global threat. Multiple studies have demonstrated that irrational antibiotics are being prescribed by physicians, general people have a habit of self-medication, and there is also an indiscriminate use of antibiotics in agriculture and farming in different parts of the country (Biswas et al., 2014; Mostafa Shamsuzzaman & Kumar Biswas, 2012). All of these events have made Bangladesh super vulnerable to Antimicrobial resistance. Furthermore, the scarcity of newer drugs means resistance must be contained before we runout of options to battle it.

Although Ahmed et al. (2019) reported that the developing countries like Bangladesh are more affected to Antibiotic resistance because of the widespread misuse of antibiotics, non-human antibiotic use, poor quality of drugs, inadequate surveillance, and factors associated with individual and

national poverty (poor healthcare standards, malnutrition, chronic and repeated infections, unaffordability of more effective and costly drugs), another finding of Al Salah et al. (2020) made Bangladesh more vulnerable because they reported that the situation of antibiotic resistance worsens under tropical conditions compared to temperate climates because Horizontal Gene Transfer (HGT) occurs at a higher rate under tropical conditions. Bangladesh falls under tropical climate zone and the rate of plasmid conjugation at such temperatures is much higher than under temperate conditions. Subsequently, the proliferation of antibiotic resistance among bacteria is expected to rise (Al Salah et al., 2020; Devarajan et al., 2017; Devarajan et al., 2016).

1.3 Dhaka: The hotspot of antimicrobial resistance

Dhaka, the capital of Bangladesh accommodates around 163,046,000 populations (Countries, 2020) and in order to ensure primary healthcare of this huge population, so many government and private hospitals has been established each and everywhere of Dhaka. Due to improper surveillance, hospitals wastes are being disposed to open environment without any treatment or with inadequate treatment and also being mixed with municipal sewage systems that are ultimately being discharged to the surrounding rivers (i.e. Buriganga, Turag, Balu and Shitalakkhya). River aquatic environment, in general, is the most intensively human influenced ecosystems in the world (Tejerina-Garro et al., 2005) as it receives huge amount of urban, hospital,

animal and industrial effluents and create a reservoir of bacteria (Adelowo et al., 2018; Sibanda et al., 2015). This contaminated ecosystem provides optimum conditions to antibiotic resistant bacteria from various sources to mix and transfer their resistant gene(s) to clinically important bacteria for development of human pathogens with novel resistance mechanisms (Baquero et al., 2008; Maravić et al., 2016; Tacão et al., 2012).

Hospital wastewater, in particular, is expected to contain high abundances of antibiotic resistance genes (ARGs) compared to municipal wastewater because it contains human enteric bacteria that may include antibiotic-resistant organisms originating from hospital patients, and can also have high concentrations of antibiotics and antimicrobials relative to municipal wastewater (Petrovich et al., 2020). Several studies reported highly pathogenic MDR strains from infected patients in tertiary hospitals of Dhaka which can also contribute to other hospital acquired infections. Ahmed et al. (2019) found a high prevalence of resistance in most tested pathogens, and many of the common first-line drugs were mostly found ineffective. Safain et al. (2020) reported that, in Dhaka, there was a year-wise gradual increase of MDR isolates from 2015-2018 and by 2019 the increase in MDR isolates became almost 2-fold compared to 2015.

Although several studies reported MDR strains from clinical samples of the tertiary hospitals in Dhaka, the drug resistance scenario of the rivers surrounding Dhaka is still unclear. Recently Islam et al. (2017) reported

multidrug resistant *New Delhi metallo-β-lactamase-1* from wastewater of Hospital Adjacent Areas (HAR). It indicates that there is high possibility of finding other multidrug resistant bacterial genera as well as associated gene from the hospital wastewater. The rivers of Dhaka are interconnected to one another and since there is direct link between hospital wastewater and river, it is highly likely that the resistance in hospital wastewater might contaminate the river water. But the role of hospitals adjacent to the rivers in spreading AMR is also still unreported from Bangladesh as well as neighboring countries.

1.4 The emergence of *bla*$_{\text{CTX-M}}$ gene variants and their role in fostering the global burden of antimicrobial resistance (AMR)

The broad-spectrum third-generation cephalosporin antibiotic are used in hospitals to treat infections of the skin and respiratory tract caused by both Gram negative and Gram positive bacteria (Klein & Cunha, 1995). But the presence of *bla*$_{\text{CTX-M}}$ gene in the bacteria producing ESBLs has made it resistant to this third generation antibiotics (Ehlers et al., 2009). To date, over 172 *bla*$_{\text{CTX-M}}$ types have been identified and described (https://www.lahey.org/studies/other.asp) which have been grouped into five clusters (CTX-M-1, CTX-M-2, CTX-M-8, CTX-M-9, and CTX-M-25), named after the first member of each group (Boyd et al., 2004; Naseer & Sundsfjord, 2011).

In the clinical isolates collected from the tertiary hospitals of Dhaka, CTXM-1 was found as the most prevalent Extended Spectrum Beta Lactamase (Safain et al., 2020). The number is rapidly increasing because of point mutations in the gene (Ejaz et al., 2021). Moreover, their capability of Horizontal Gene Transfer mediated by mobile genetic elements can also lead to serious threat to the environment surrounding the hospitals.

1.5 Objective of this study

This study aims to explore the diversity of bacteria from water sample collected from 5 major rivers as well as hospitals surrounding Dhaka, Bangladesh through High-throughput sequencing- a culture-independent approaches that brought a renewed perspective of the bacterial diversity in water habitats, in which < 0.1% of bacteria can be cultivated (Amann et al., 1995; Simon & Daniel, 2011; Vaz-Moreira et al., 2014). The same approach will be followed to study *bla*$_{\text{CTX-M}}$ gene variants present in the river and hospital wastewater.

Since Dhaka represents a megacity surrounded by many rivers and most importantly, the tropical climate has made Dhaka super vulnerable to AMR, the status of MDR strain and their diversity will lead us to important insights that might help Bangladesh as well as other tropical countries to revise their antimicrobial surveillance guidelines. It might also help the health professionals to ensure the best possible treatment that needs antibiotics to be prescribed.

2. MATERIALS AND METHODS

2.1 Collection of river water

Water sample was collected from 5 rivers namely- Buriganga, Shitalakkhya, Turag, Tongi Khal and Balu. Five replicates of samples (from upstream to downstream) were collected following upstream to downstream manner from each river. 200 mL of water was collected for each sample using Water Sample Collection Bottle (AEMTEK Laboratories).

2.2 Collection of hospital wastewater

Hospital wastewater was collected from the discharge line of 5 major hospitals of Dhaka following upstream to downstream manner. The hospitals are Sir Salimullah Medical College Hospital, National Institute of Cardiovascular Disease and Hospital, National Institute of Cancer Research and Hospital, Aichi Medical College and Hospital and Narayangang 300 Bedded Hospital. 200mL sample was collected for each sample using Water Sample Collection Bottle (AEMTEK Laboratories). The bottles were then kept inside an insulated box containing ice packs to maintain the quality of the water. The water samples were immediately shipped to the Environmental Microbiology Lab, Department of Microbiology, University of Dhaka, Dhaka-1000, Bangladesh for further processing.

2.3 Processing of water samples

100mL water from each sample was filtered using 0.45μm filter paper (Whatman filter, Sigma-Aldrich, USA). Rocker 300C Vacuum Filtration System (STERLITECH, USA) was used to filter the water. The filter paper was stored at -20ºC until it was shipped to the Graduate School of Public Health, Seoul National University, Seoul, South Korea for further analyses. While shipping, cold chain was properly maintained to ensure the quality of the sample.

2.4 DNA extraction and PCR amplification

¼ of the filter paper was used for DNA extraction. DNA extraction was performed using the PowerMax Soil DNA Isolation Kit (Mobio Laboratory, Carlsbad, CA, USA) with the modification of adding 0.1mm diameter glass beads (300 mg) and 0.5mm diameter beads (100 mg) to the microcentrifuge tube (Yamamoto et al., 2012). The samples were first homogenized for 4 minutes by a bead beater (BioSpec. Inc.Bartlesville, OK, USA) and then proceeded to the DNA extraction following the protocol. The samples were eluted with 50 ul 10 μM Tris buffer and kept at -20 ºC before amplification.

2.5 Next-generation DNA sequencing and data analysis

2.5.1 16S rRNA gene sequencing library preparation

Library preparation was performed following the Illumina 16S Metagenomic Sequencing Library Preparation protocol. The 16S rRNA genes were amplified using the following Forward and Reverse Primer: Forward: 'TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCTACGGGNGGCWGCAG3', and Reverse: 'GTCTCGTGGGCTCGGAGATGTGTATAAGAG ACAGGACTACHVGGGTATCTAATCC3'.

Each PCR reaction contained 50μl, including 25μl of 2× KAPA HiFi HotStart ReadyMix, 1μl of each 10pM primer and 2μl of the DNA template. PCR was performed using the following condition: initial denaturation at 95ºC for 5min, 35 cycles of 95ºC dissociation for 30 seconds, annealing at 55ºC for 30 seconds and extension at 72ºC for 1 min, followed by final extension at 72ºC for 10 min. Then PCR clean-up was conducted using AMPure XP beads to purify the 16S V3 and V4 amplicon away from free primers and primer dimer species. Dual indices were attached by Index PCR following the thermal cycler conditions: 95ºC for 3 min, 8 cycles of 95ºC for 30 seconds, 55ºC for 30 seconds and 72ºC for 30 seconds, followed by 72ºC for 5 min. The second PCR clean-up was conducted after Index PCR.

Library quantification was performed using PicoGreen method. The quantified libraries were pooled and denatured with NaOH, diluted with

hybridization buffer before Miseq sequencing. 30% PhiX were added to serve as an internal control for these low-diversity libraries. Libraries were then loaded onto a MiSeq reagent cartridge and then onto the instrument. Then the 16S rDNA next generation sequencing was performed by using MiSeq sequencing system, which relies on the fluorescence generated by the incorporation of fluorescently labeled nucleotides into the growing strand of DNA (Quail et al., 2012).

MiSeq Reporter version 2.5 (Illumina) was used to trim the primer and multiplexing barcode sequences and remove reads with quality scores below 20. Next, USEARCH version 11.0.667 (Edgar, 2010) was used to join the forward and reverse reads and then remove low quality reads with > 1.0 expected errors and/or those with lengths less than 200 bp. After quality trimming, the UPARSE-OTU algorithm (Edgar, 2013) was used to cluster unique sequences into 97% operational taxonomic units (OTUs) and remove chimeric reads. Using the SINTAX algorithm (Edgar, 2016), each OTU was taxonomically assigned against RDP training set v16 (rdp_16s_v16.fa.gz) (Cole et al., 2014) with a cutoff confidence value of 0.5(Edgar, 2018). For diversity analyses, libraries were rarefied to 1333 sequence reads using the "MicrobiomeAnalyst" package (Dhariwal et al., 2017). Statistical package R 3.6.2 (McMurdie & Holmes, 2013) was used to calculate α-diversity measures and distance matrices for β-diversity analyses.

2.5.2 ARG sequencing library preparation

Library preparation was performed following the ARG Sequencing Library Preparation protocol. The primer pair CTX-MU1 (5′-ATGTGCAGY ACCAGTAARGT-3′) and CTX-MU2 (5′-TGGGTRAARTARGTSACCAG A-3′) (Pagani et al., 2003) were designed based on the conserved regions of *bla*$_{\text{CTX-M}}$ genes and it targets the amplification of a 593-bp internal region of the *bla*$_{\text{CTX-M}}$ genes. The Illumina overhang adapter sequences (Forward overhang: 5′-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG-[locus-specific sequence], and Reverse overhang: 5′- GTCTCGTGGGCTCG GGAGATGTGT-[locus-specific sequence]) were added to CTX-MU1 and CTX-MU2 and constructed our target primer that resulted in the amplification of 660 bp amplicon.

Each PCR reaction contained 50µl, including 25µl of 2× KAPA HiFi HotStart ReadyMix, 1µl of each 10pM primer and 2µl of the DNA template. PCR was performed using the following condition: initial denaturation at 94°C for 7 min; denaturation at 94°C for 50 s, annealing at 50°C for 40 s, and elongation at 72°C for 60 s, repeated for 35 cycles; final extension at 72°C for 5 min. Then PCR clean-up was conducted using AMPure XP beads to purify the amplified gene away from free primers and primer dimer species. Dual indices were attached by Index PCR following the thermal cycler conditions: 95ºC for 3 min, 8 cycles of 95ºC for 30 s, 55ºC for 30 s and 72ºC

for 30 s, followed by 72ºC for 5 min. The second PCR clean-up was conducted after Index PCR.

The library quantification was performed using PicoGreen method. The quantified libraries were pooled and denatured with NaOH, diluted with hybridization buffer before Miseq sequencing. 30% PhiX were added to serve as an internal control for these low-diversity libraries. Libraries were then loaded onto a MiSeq reagent cartridge and then onto the instrument. Then the *bla*$_{\text{CTX-M}}$ ARG sequencing was performed by using MiSeq sequencing system, which relies on the fluorescence generated by the incorporation of fluorescently labeled nucleotides into the growing strand of DNA (Quail et al., 2012).

The raw sequence reads were quality trimmed, and the adapter sequences were removed using Trimmomatic v0.33 (Bolger et al., 2014). To assess and quantify the relative abundance of the antibiotic resistant genes (ARGs) in our data, we used Short Read Sequence Typing (SRST2) program (Inouye et al., 2014). The SRST2 program mapped the quality-filtered sequence reads and cluster similar sequences against the antibiotic resistance gene database (ARG-ANNOT) that incorporated all sequences of known antibiotics resistance genes (ARGs) (Gupta et al., 2014).

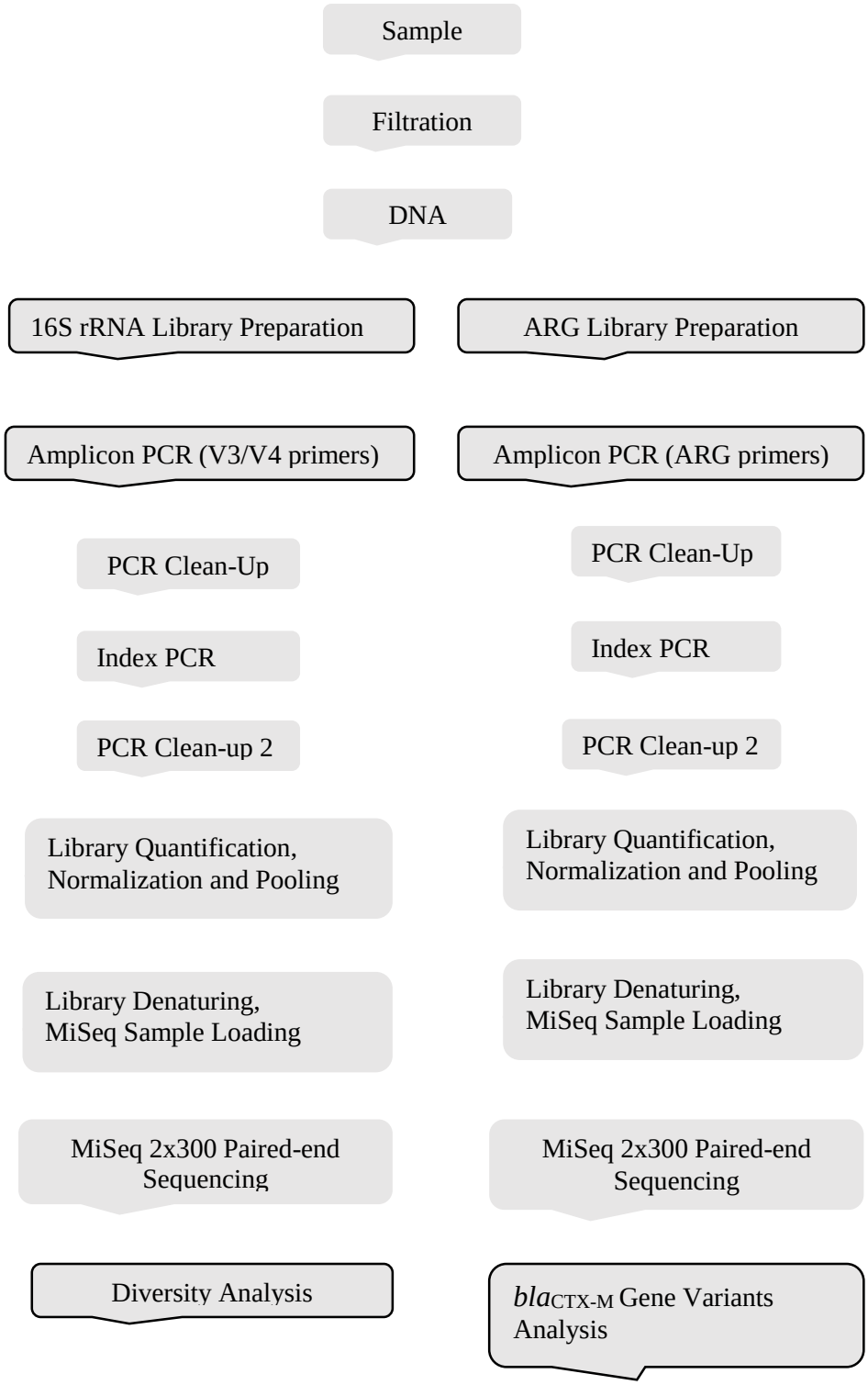

Figure 1. Flowchart of experimental process

3. RESULTS

3.1 Sample statistics:

In total 36 water samples, 25 river water samples from 5 rivers (Buriganga, Turag, Tongi Khal, Balu and Shitalakkhya) 11 hospital wastewater samples from 5 hospitals (Sir Salimullah Medical College Hospital, National Institute of Cardiovascular Disease, Aichi Medical College and Hospital, National Institute of Cancer Research and Hospital, and Narayanganj 300 Bed Hospital) were sequenced for V3 and V4 region of 16S rRNA gene (Table S1) resulting in a total of 36 sequence libraries. In total 4,13,297 sequence reads were obtained where the number of sequence reads ranged from 1333 to 38110 reads per library.

Table 1. Sample statistics of river and hospital wastewater.

Sample Name	Sample Type	Sample Collection Date (yymmdd)	Number of Reads
B1	River water	20191125	8837
B2	River water	20191125	1333
B3	River water	20191125	14833
B4	River water	20191125	3085
B5	River water	20191125	4044
B6	River water	20191125	1960
B7	River water	20191125	2401
B8	River water	20191125	2393
B9	River water	20191125	18841
B10	River water	20191125	4760
B11	River water	20200222	9699

Sample Name	Sample Type	Sample Collection Date (yymmdd)	Number of Reads
B12	River water	20200222	38110
B13	River water	20200222	15312
B14	River water	20200222	1944
B15	River water	20200222	20238
B16	River water	20200222	2118
B17	River water	20200222	1879
B18	River water	20200222	2768
B19	River water	20200222	4238
B20	River water	20200222	3629
B21	River water	20200118	21040
B22	River water	20200118	8621
B23	River water	20200118	11159
B24	River water	20200118	36866
B25	River water	20200118	23198
B26	Hospital Wastewater	20191125	1413
B27	Hospital Wastewater	20191125	25961
B28	Hospital Wastewater	20191125	2869
B29	Hospital Wastewater	20191125	5972
B31	Hospital Wastewater	20200314	2622
B32	Hospital Wastewater	20200102	2408
B33	Hospital Wastewater	20200102	3293
B34	Hospital Wastewater	20200102	31442
B35	Hospital Wastewater	20200222	35497
B36	Hospital Wastewater	20200314	28382
B37	Hospital Wastewater	20200314	10132

3.2 Sample collection site:

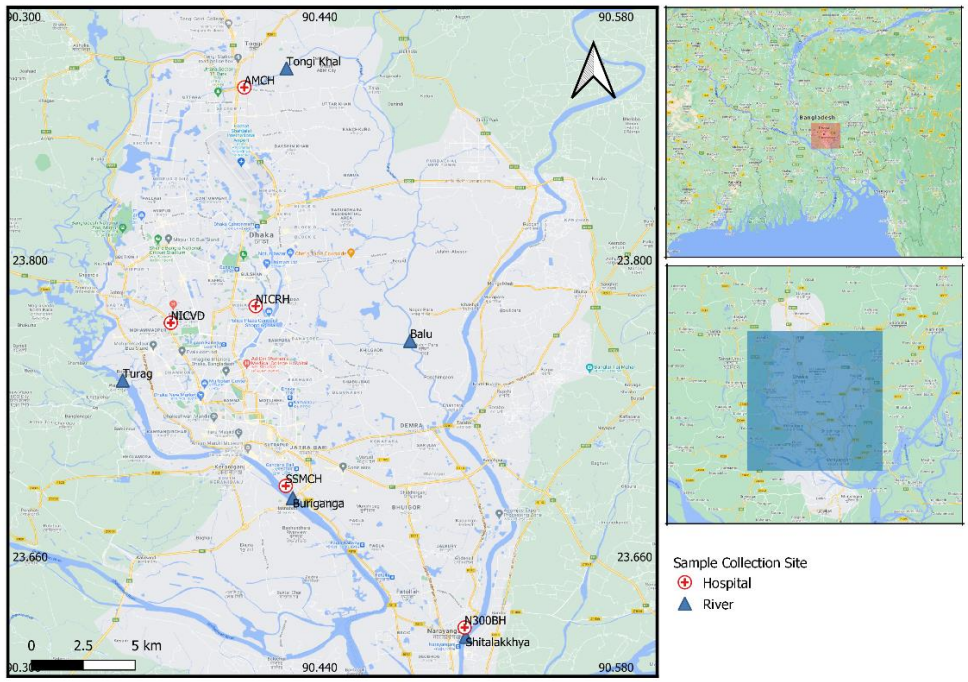

Figure 2. Sample collection sites in the map of Dhaka city. Inset, location of Dhaka district within map of Bangladesh and location of Dhaka city within Dhaka district. The red pointers indicate the locations of hospitals, and the blue pointers indicate the locations of river water collection sites. SSMCH indicates Sir Salimullah Medical College and Hospital, Buriganga indicates Buriganga river, NICVD indicates National Institute of Cardiovascular Diseases, Turag indicates Turag river, AMCH indicates Aichi Medical College and Hospital, Tongi Khal indicates a small canal connecting Tongi River and Balu River, NIRCH indicates National Institute of Cancer Research and Hospital, Balu indicates Balu river, N300BH indicates Narayangang 300

Bed Hospital and Shitalakkhya indicates Shitalakkhya river. Custom made base map was used to construct the map using QGIS 3.8 software.

For data analysis, all 36 samples were categorized into 5 groups A, B, C, D and E.

Table 2. Sample categorization.

Group	Sample ID	Sampling site
A	B1, B2, B3, B4, B5	Buriganga River
	B26, B27, B28, B29	Sir Salimullah Medical College and Hospital
B	B11, B12, B13, B14, B15	Turag River
	B31	National Institute of Cardiovascular Disease
C	B16, B17, B18, B19, B20	Tongi Khal
	B32, B33, B34	Aichi Medical College and Hospital
D	B21, B22, B23, B24, B25	Balu River
	B35	National Institute of Cancer Research and Hospital
E	B6, B7, B8, B9, B10	Shitalakkhya River
	B36, B37	Narayanganj 300 Bed Hospital

3.3 Bacterial diversity analysis

3.3.1 Taxonomy analysis

Taxonomy analyses were done through MicrobiomeAnalyst (Dhariwal et al., 2017) where 8 phyla, 12 class, 13 order, 22 family and 102 genera were identified.

3.3.2 Hierarchical cluster analysis

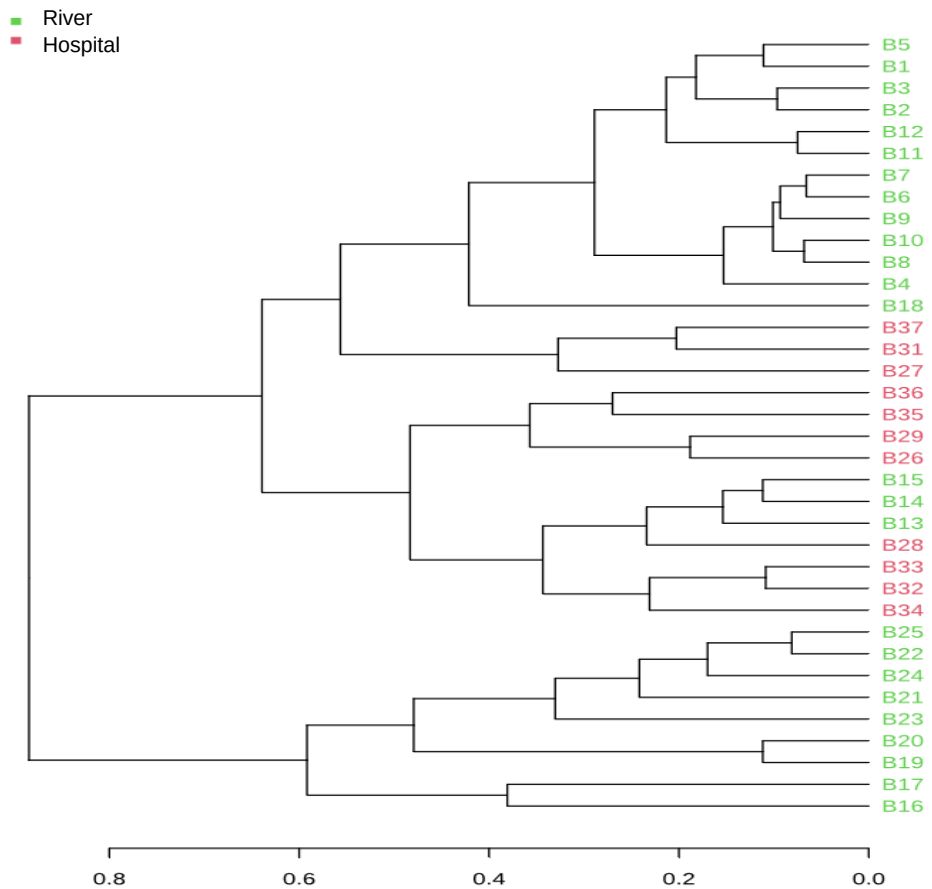

Figure 3. Clustering of samples based on sample type. Green color indicates the samples collected from river whereas pink color indicates the samples collected from hospital wastewater.

3.3.3 Community profiling

To further estimate the diversity and richness of different samples, R version 3.6.2 was used. Species richness was characterized in terms of the numbers of observed OTUs based on the HTS sequences at 97% similarity.

Table 3. Diversity indices for bacteria

Sample ID	Sample Type	Shannon	Simpson
B1	River	3.22	0.89
B2	River	2.82	0.86
B3	River	2.85	0.85
B4	River	2.93	0.86
B5	River	3.26	0.89
B6	River	2.91	0.88
B7	River	2.83	0.88
B8	River	2.96	0.89
B9	River	2.82	0.87
B10	River	3.00	0.89
B11	River	3.28	0.93
B12	River	3.09	0.92
B13	River	3.23	0.92
B14	River	2.61	0.86
B15	River	3.03	0.90
B16	River	3.24	0.91
B17	River	2.86	0.86
B18	River	2.32	0.76
B19	River	2.36	0.74
B20	River	2.18	0.71
B21	River	2.96	0.86

Sample ID	Sample Type	Shannon	Simpson
B22	River	2.96	0.88
B23	River	2.87	0.88
B24	River	2.53	0.82
B25	River	2.73	0.86
B26	Hospital Wastewater	3.18	0.90
B27	Hospital Wastewater	2.82	0.87
B28	Hospital Wastewater	3.25	0.90
B29	Hospital Wastewater	3.16	0.92
B31	Hospital Wastewater	3.18	0.88
B32	Hospital Wastewater	3.52	0.94
B33	Hospital Wastewater	3.44	0.93
B34	Hospital Wastewater	2.83	0.88
B35	Hospital Wastewater	3.18	0.93
B36	Hospital Wastewater	3.14	0.90
B37	Hospital Wastewater	3.24	0.90

Alpha diversity

Alpha diversity measurement was done based on Shannon and Simpson diversity index. Shannon index measures the average degree of uncertainty in predicting as to what species an individual will belong. The value increases as the number of species increases and as the distribution of individuals among the species becomes even (Legovi 1991). Simpson index indicates species dominance and reflects the probability of two individuals that belong to the same species being randomly chosen. It varies from 0 to 1 and the index increases as the diversity decreases (Simpson, 1949).

Shannon and Simpson diversity index (Table 3) based on Pairwise Comparison using Wilcoxon Rank Sum Exact Test indicate that there was significant difference ($P \leq 0.05$) observed (Fig 4) between the samples from river and hospital, that means river and hospital share different bacterial community.

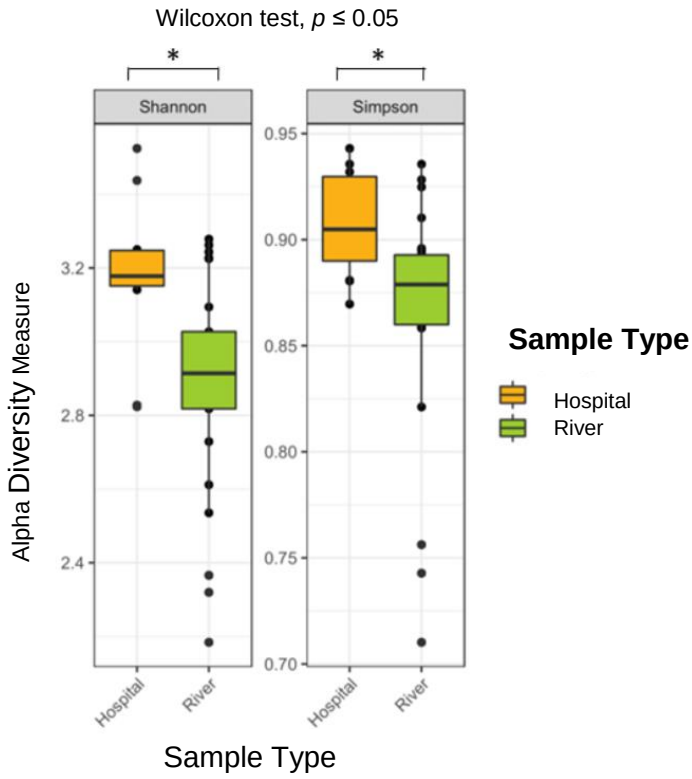

Figure 4. Alpha diversity analysis of river and hospital wastewater. Alpha diversity measured by Shannon and Simpson diversity. Index is plotted for sampling type, River and Hospital. The line inside the box represents the median, while the whiskers represent the lowest and highest value within the interquartile range (IQR). Outliers and individual sample values are shown as dots. Statistical testing based on pairwise comparisons using Wilcoxon rank sum exact test showed significant difference for Shannon (pShannon≤ 0.05), and Simpson (pSimpson≤ 0.05) diversity index.

Beta diversity

Beta diversity measurement was done by Bray-Curtis method. Significance test was done by Permutational multivariate analysis of variance (PERMANOVA). The Nonmetric Multidimensional Scaling (NMDS) showed that there is significant diversity ($P \leq 0.05$) between hospital and river sample, irrespective of sampling site and sample type.

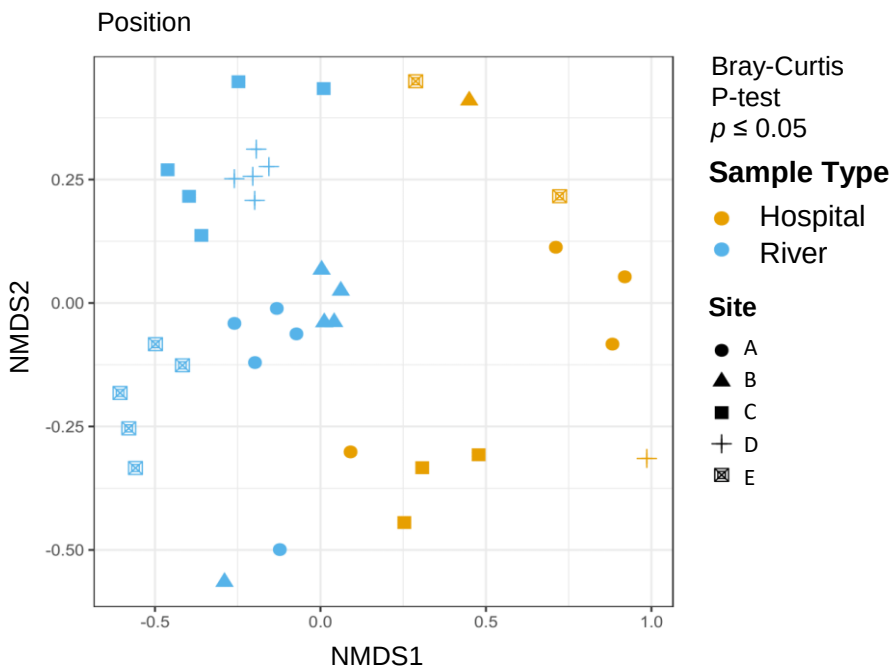

Figure 5. Nonmetric Multidimensional Scaling (NMDS) of river and hospital wastewater. NMDS is an unconstrained, distance based ordination method which was performed with Bray-Curtis dissimilarity. Samples that are more similar to one another are ordinated closer together. River samples are plotted as sky blue dots, and hospital samples are represented as orange dots. The river and hospital wastewater samples show significant differences in similarity tested by PERMANOVA (pPERMANOVA≤ 0.05).

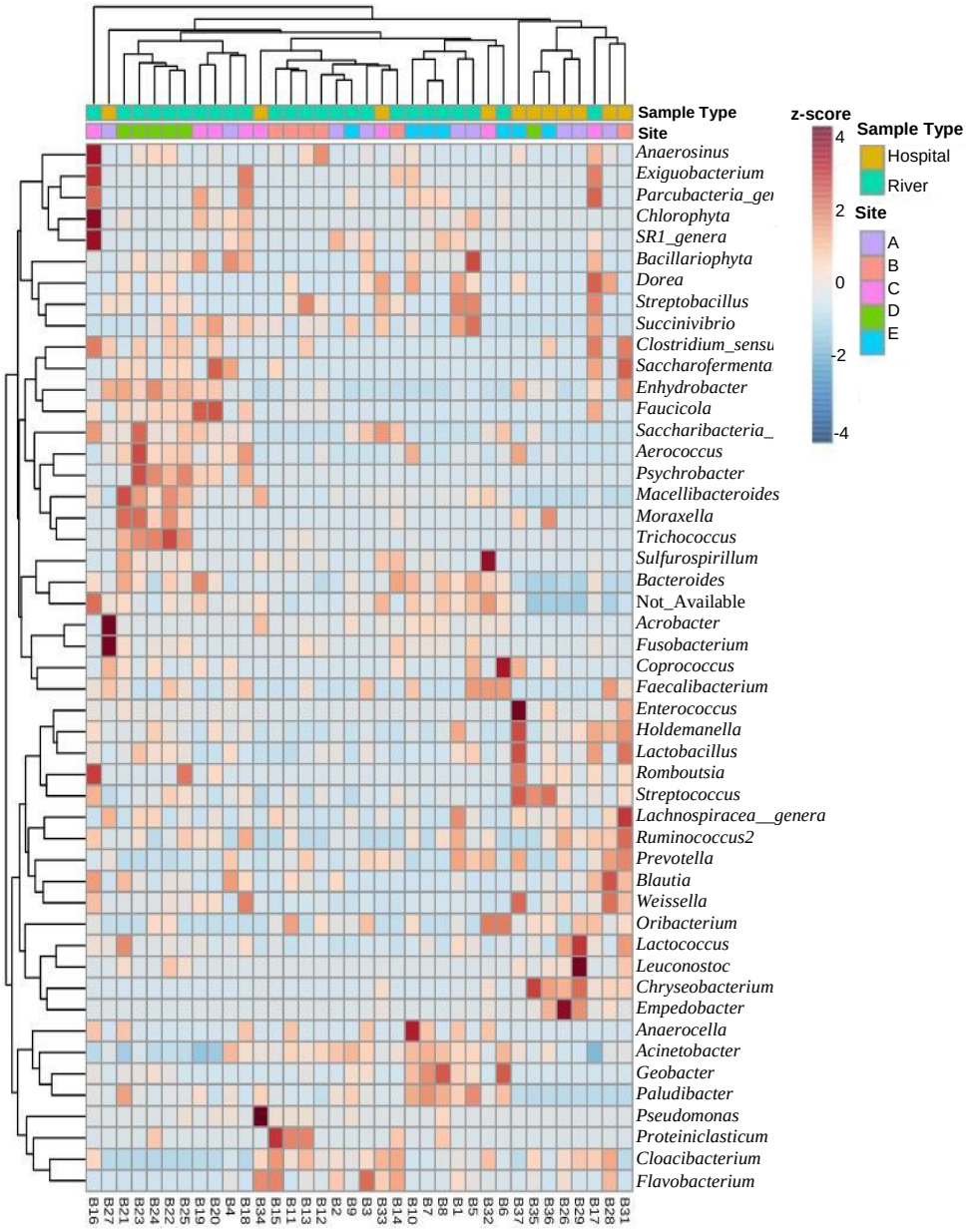

Figure 6. Compositions and relative abundance of the most abundant bacterial genera. Mean relative abundances across all samples are calculated for each genus. Euclidean distances based on log-transformed relative abundance values are shown.

3.4 Detection of potential pathogenic bacterial genera

The US Center for Disease Control and Prevention (CDC) has categorized the most pathogenic and multidrug resistant bacteria into 4 groups namely 1. Urgent threats 2. Serious threats 3. Concerning threats and 4. Watch list (https://www.cdc.gov/drugresistance/biggest-threats.html) that contains 21 bacterial genera. Out of 102 bacterial genera detected from our analyses from river and hospital wastewater, 7 genera matched with US CDC list of most pathogenic bacteria namely *Acinetobacter* (urgent threat), *Enterococci*, *Pseudomonas*, *Staphylococcus* and *Streptococcus* (serious threat), Erythromycin resistant group A *Streptococci* and Clindamycin resistant group B *Streptococcus* (concerning threat).

3.5 ARG variant analysis

3.5.1 *bla*$_{\text{CTX-M}}$ gene variants analysis

For *bla*$_{\text{CTX-M}}$ variant analysis, Short Read Sequence Typing (SRST2) program (Inouye et al., 2014) was used. The SRST2 program mapped the quality-filtered sequence reads and cluster similar sequences against the antibiotic resistance gene database (ARG-ANNOT) that incorporated all sequences of known antibiotics resistance genes (ARGs) (Gupta et al., 2014).

To further estimate the diversity of different variants between river and hospital, R version 3.6.2 was used. Beta diversity measurement was done by Bray-Curtis method. Significance test was done by Permutational multivariate analysis of variance (PERMANOVA).

3.5.2 Community profiling

Species richness was characterized in terms of the numbers of observed OTUs based on the HTS sequences at 97% similarity.

Table 4. Diversity indices for blaCTX-M variants.

Sample ID	Sample Type	Shannon	Simpson
B1	River	4.04	0.98
B2	River	4.02	0.98
B3	River	4.02	0.98
B4	River	4.12	0.98
B5	River	4.11	0.98
B6	River	4.17	0.98
B8	River	4.52	0.98
B10	River	4.35	0.98
B11	River	4.06	0.98
B12	River	4.16	0.98
B13	River	4.12	0.98
B14	River	4.14	0.98
B15	River	4.11	0.98
B16	River	4.54	0.98
B17	River	4.56	0.98
B18	River	4.54	0.98
B19	River	4.54	0.98
B20	River	4.51	0.98
B21	River	4.19	0.98
B22	River	4.38	0.98
B23	River	4.42	0.98
B24	River	4.27	0.98

Sample ID	Sample Type	Shannon	Simpson
B25	River	4.48	0.98
B26	Hospital Wastewater	4.00	0.97
B28	Hospital Wastewater	4.35	0.98
B29	Hospital Wastewater	3.98	0.97
B31	Hospital Wastewater	4.38	0.98
B32	Hospital Wastewater	3.95	0.97
B33	Hospital Wastewater	3.93	0.97
B34	Hospital Wastewater	3.97	0.97
B35	Hospital Wastewater	3.92	0.97
B36	Hospital Wastewater	3.94	0.97
B37	Hospital Wastewater	4.11	0.98

Alpha diversity

Alpha diversity measurement was done based on Shannon and Simpson diversity index. Shannon index measures the average degree of uncertainty in predicting as to what species an individual will belong. The value increases as the number of species increases and as the distribution of individuals among the species becomes even (Legovi 1991). Simpson index indicates species dominance and reflects the probability of two individuals that belong to the same species being randomly chosen. It varies from 0 to 1 and the index increases as the diversity decreases (Simpson, 1949).

Shannon and Simpson diversity index (Table 3) based on Pairwise Comparison using Wilcoxon Rank Sum Exact Test indicate that there was significant difference ($P \leq 0.05$) observed (Fig 7) between the samples from

river and hospital, that means river and hospital share different types of *bla*$_{\text{CTX-M}}$ variants.

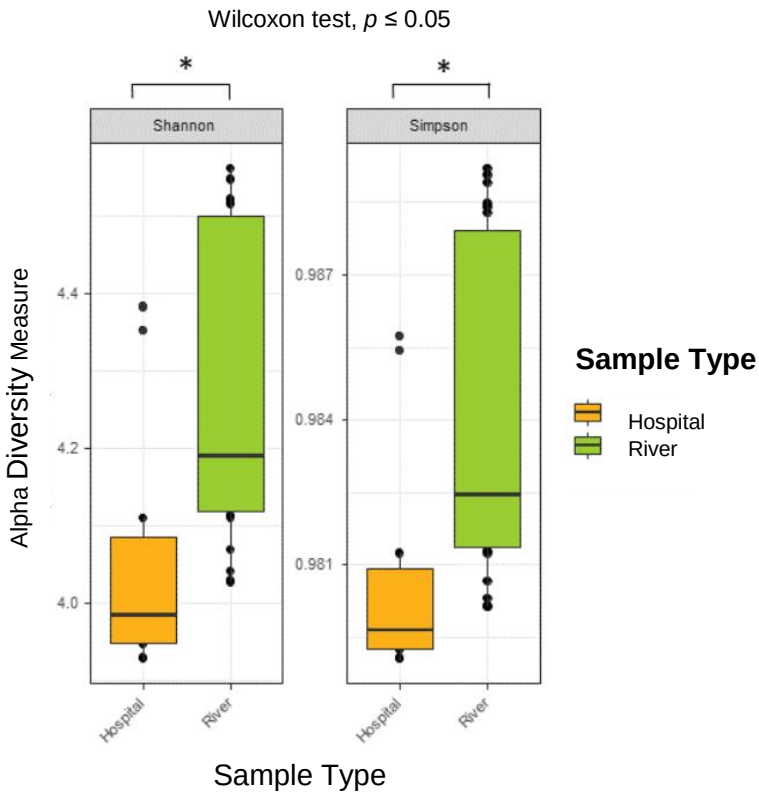

Figure 7. Alpha diversity analysis of blaCTX-M variants. Alpha diversity measured by Shannon and Simpson diversity. Index is plotted for sampling type, River and Hospital. The line inside the box represents the median, while the whiskers represent the lowest and highest value within the interquartile range (IQR). Outliers and individual sample values are shown as dots. Statistical testing based on pairwise comparisons using Wilcoxon rank sum exact test showed significant difference for Shannon (pShannon≤ 0.05), and Simpson (pSimpson≤ 0.05) diversity index.

Beta diversity

Beta diversity measurement was done by Bray-Curtis method. Significance test was done by Permutational multivariate analysis of variance (PERMANOVA). The Nonmetric Multidimensional Scaling (NMDS) showed that there is significant diversity ($P \leq 0.05$) between hospital and river sample, irrespective of sampling site and sample type.

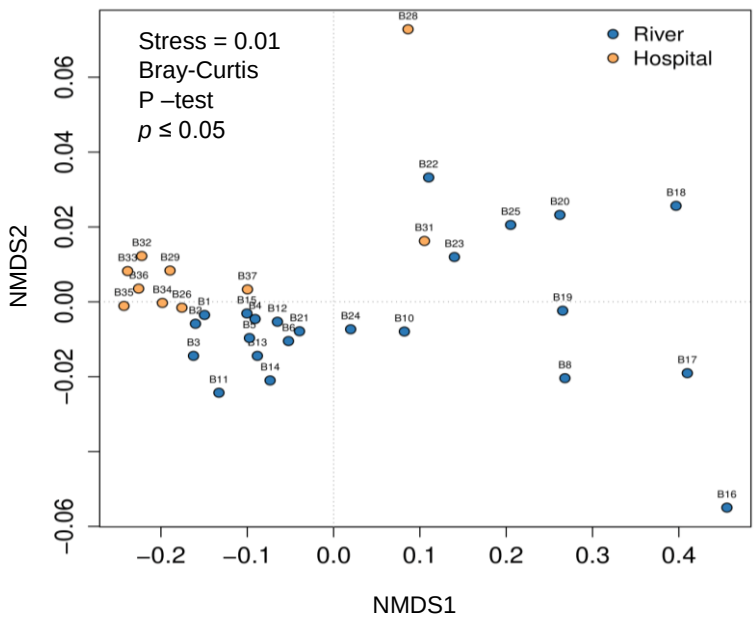

Figure 8. Nonmetric Multidimensional Scaling (NMDS) of blaCTX-M variants. NMDS is an unconstrained, distance based ordination method which was performed with Bray-Curtis dissimilarity. Samples that are more similar to one another are ordinated closer together. River samples are plotted as sky blue dots, and hospital samples are represented as orange dots. The river and hospital wastewater samples show significant differences in similarity tested by PERMANOVA (pPERMANOVA≤ 0.05).

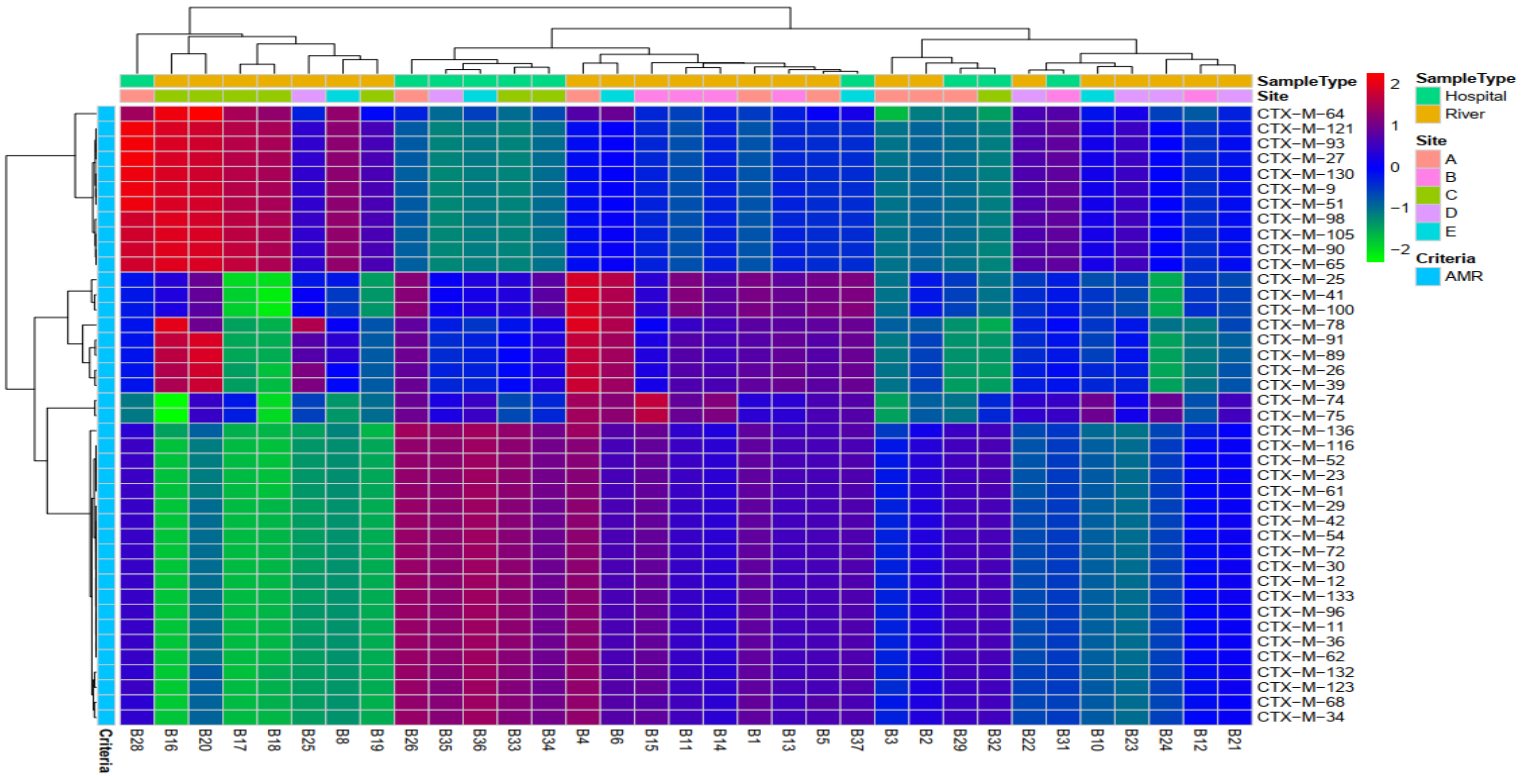

Figure 9. Composition and relative abundance of 40 variants of blaCTX-M gene. Mean relative abundances across all samples are calculated for each variants. Euclidean distances based on log-transformed relative abundance values are shown.

4. Discussion

The rapidly increasing number of studies applying high-throughput sequencing technologies are revealing tremendous diversity within the microbial communities inhabiting all types of environments (Andersson et al., 2010; Eiler et al., 2012; Galand et al., 2009; Peura et al., 2012; Sogin et al., 2006). In our study, we used 16S rRNA gene sequencing to study bacterial diversity as well as ARG Sequencing to study *bla*$_{\text{CTX-M}}$ gene variants from river and hospital wastewater samples owing to its emerging public health importance.

River and hospital wastewater were collected from Dhaka, Bangladesh and sequenced for V3 and V4 region of 16S rRNA gene resulting in a total of 36 sequence libraries. In total 4,13,297 sequence reads were obtained where the number of sequence reads ranged from 1333 to 38110 reads per library. The taxonomic analysis revealed 8 phyla, 12 class, 13 order, 22 families and 102 genera in the river and hospital wastewater samples.

Among 102 bacterial genera identified, 3 most abundant genera were *Acinetobacter* (42%), *Coalcibacterium* (14%), *Enhydrobacter* (10%). The genus *Acinetobacter* is frequently isolated in hospital-acquired infections, especially in intensive care units. *A. baumannii* is a frequent cause of nosocomial pneumonia and other infections, such as skin and wound infections, infective endocarditis, bacteremia, urinary tract infections, and meningitis(Dent et al., 2010). Several studies have demonstrated that *A.*

baumannii is resistant to many of the antibiotics used in the hospital context (Fournier et al., 2006; Hu et al., 2011). Other human pathogens or opportunistic pathogens, such as members of the genera *Enterococcus*, *Pseudomonas*, *Coalcibacterium*, *Enhydrobacter*, and *Streptococcus* were also prevalent in bacterial communities in the hospital wastewaters(Cao et al., 2010; Jani et al., 2019; Oluseyi Osunmakinde et al., 2019; Pammi et al., 2020). All of these bacteria were resistant to more than three classes of clinical antibiotics, in addition to their intrinsic resistance (Wang et al., 2018). Given the presence of ESBL-producing bacteria in these environments, a recurring concern is the transfer of conjugative plasmids, which also carry genes of resistance to other antimicrobial agents, given the bacterial multiresistance patterns (Heuer et al., 2002; Paterson, 2006)

Several reports have indicated that hospital wastewater, without or even with treatment, presents a high risk to human health, with the release of ARB into aquatic environments because Wastewater Treatment Plant provides conditions conducive to the establishment and propagation of ARB(Kim et al., 2007; Marcinek et al., 1998). Therefore, compared with other pollution sources, hospital wastewater presents a greater environmental risk not only because they contain high concentrations of antibiotics but also due to antibiotic-resistant pathogen and/or opportunistic pathogen release.

Although, it is important to monitor the occurrence of ARGs in highly polluted environments where high selection pressure exerted by antimicrobial contaminants contribute toward the spread and persistence of antibiotic-resistant bacteria, it is critical not to ignore the emergence of ARGs in relatively less polluted environments (Baquero et al., 2008; Martinez, 2009; Williams, 2001). Identifying sources of ARGs and their dissemination in relatively less polluted environments will help form strategies to reduce their spread. From this point of view, our study holds huge importance as it is the first study with river and hospital wastewater from Dhaka, Bangladesh and it revealed 102 bacterial genera where most pathogenic multidrug resistant genera are also present. This information has huge public health importance to tackle the burden of frequently emerging pathogenic bacterial population.

We further extended our study to see if there is any possible interconnection between the bacterial genera of river and hospital wastewater. We used hierarchical cluster analysis, alpha diversity analysis and beta diversity analysis to investigate our queries.

We applied hierarchical cluster analysis (Fig 3) based on Bray-Curtis index to check if the hospital wastewater shares totally different bacterial taxonomy than that of river water. The samples from Tongi khal and Balu river showed a distinct taxonomical cluster from other samples. Among 102 bacterial genera, 48 were most abundant in our samples where Tongi khal contained highest number of bacterial genera. Since the geographical location

of Tongi khal is in the upstream whereas Balu river is in the downstream, it is quite possible for the most abundant bacterial genera from Tongi khal to transfer to downstream Balu river due to the regular flow of river water.

For community profiling, we tested Shannon and Simpson index following categorization of the sample based on sample type. Shannon and Simpson diversity index (Table 3) based on Pairwise Comparison using Wilcoxon Rank Sum Exact Test indicate that there was significant difference between the bacterial community present in river and hospital wastewater. This is a clear indication that samples within rivers and hospitals differ significantly from each other.

Beta diversity measurement was done by Bray-Curtis method. Permutational multivariate analysis of variance (PERMANOVA) was used for significance test. Nonmetric Multidimensional Scaling (NMDS) showed that there is significant diversity ($P \leq 0.05$) between hospital and river sample, irrespective of sampling site and sample type. Since hospital wastewater contains increased amount of pathogenic bacteria originated from medical devices, atmosphere and water used in hospital practice and released mainly in the form of excreta of pathogens (Nuñez & Moretton, 2007) whereas river water receives huge amount of urban, hospital, animal and industrial effluents (Sibanda et al., 2015), and represent two different types of reservoir for bacteria, our finding regarding no association between river and hospital wastewater bacterial community makes sense.

Based on US CDC list of most pathogenic bacteria, we also aimed to check if our samples contain any pathogenic bacteria and found that 7 bacterial genera were present in our samples. Among 4 bacterial genera that falls under CDC urgent threats group, only *Acinetobacter* was present (100%) in both river and hospital wastewater. Among 11 bacterial genera under serious threats group, *Enterococci, Pseudomonas, Staphylococcus* and *Streptococcus* were present both in river water as well as hospital wastewater but the percentage was different. River water contains 16% of *Enterococci* whereas it is 100% for hospital wastewater. For *Pseudomonas*, river water contains 28% whereas hospital wastewater contains 54%. *Staphylococcus* is present 40% of river samples whereas the percentage is 45% for hospital wastewater. Another drug resistant genera *Streptococcus* was present 100% in both river and hospital wastewater samples. The US CDC Concerning threat group contains Erythromycin resistant group A *Streptococci* and Clindamycin resistant group B *Streptococcus*. Although our 16S rRNA Sequencing could not reveal species level, it could reach up to genus level and Streptococcus was present in 100% samples of river as well as hospital wastewater. This is a clear threat to human health as the pathogenic bacteria has the potential to jeopardize the existing health facility and impose serious public health burden.

Another major goal of our study was to explore the ARG variants of $bla_{\text{CTX-M}}$ gene owing to its emerging clinical importance (Borgogna et al.,

2016; Botica et al., 2013; Gupta et al., 2003; Gurmessa et al., 2021; Korzeniewska & Harnisz, 2013). We applied ARG Sequencing that generated 33 sequence libraries for *bla*$_{\text{CTX-M}}$ gene. Out of 172 variants of *bla*$_{\text{CTX-M}}$ reported so far, mostly from clinical samples, our study based on Short Read Sequence Typing (SRST2) revealed 123 variants. Previously Safain et al. (2020) reported *bla*$_{\text{CTX-M-1}}$ from clinical isolates from tertiary hospitals of Dhaka but detail variant analysis of *bla*$_{\text{CTX-M}}$ was not done. As far our knowledge, our study is the first one to report this huge number of variants from river and hospital wastewater of Bangladesh.

For community profiling, we tested Shannon and Simpson index following categorization of the sample based on sample type; exactly same as we did for bacteria. Shannon and Simpson diversity index (Table 4) based on Pairwise Comparison using Wilcoxon Rank Sum Exact Test indicate that there was significant difference between the bacterial community present in river and hospital wastewater. This is a clear indication that the *bla*$_{\text{CTX-M}}$ variants within rivers and hospitals differ significantly from each other.

The Nonmetric Multidimensional Scaling (NMDS) was done to check if there is any association between the variants found in the river and hospital wastewater. The river and hospital wastewater samples showed significant differences in similarity tested by PERMANOVA (pPERMANOVA$\leq$ 0.05). There is close association among the *bla*$_{\text{CTX-M}}$ variants present in hospital wastewater whereas the variants for river water is scattered through the

sampling sites. From this point of view, it could be concluded that, although most of the variants are present in hospital wastewater samples, river water is more prone to the spread of the variants to the neighboring environments.

5. Conclusion

The development of resistance toward antibiotics has been observed worldwide and has challenged both public and animal health. Our study based on High-throughput sequencing revealed that the river and hospital wastewater of Dhaka, Bangladesh are not only containing pathogenic bacteria as well as antimicrobial resistant genes with different diversity, but also having the potential to emergence and spread of multidrug resistant bacteria. Unfortunately, the surveillance of its spread and prevalence in environment is limited and has to be expanded more due to the broad impact of antibiotic resistance on human health. However, further planning and implementation of strategies, policies, and experimental approaches have to be done by collaboration of scientific community and public authorities to limit the use of antibiotics, surveillance on disposal of medical wastes to open environment, detection of microbial communities (resistant and/or sensitive), and mapping resistance mechanisms to clearly understand the role of river and hospital wastewater to environmental spread of global antimicrobial resistance.

References

Adelowo, O. O., Caucci, S., Banjo, O. A., Nnanna, O. C., Awotipe, E. O., Peters, F. B., Fagade, O. E., & Berendonk, T. U. (2018). Extended Spectrum Beta-Lactamase (ESBL)-producing bacteria isolated from hospital wastewaters, rivers and aquaculture sources in Nigeria. *Environtal Science and Pollution Research*, *25*(3), 2744-2755. https://doi.org/10.1007/s11356-017-0686-7

Ahmed, I., Rabbi, M. B., & Sultana, S. (2019). Antibiotic resistance in Bangladesh: A systematic review. *International Journal of Infectious Diseases*, *80*, 54-61. https://doi.org/10.1016/j.ijid.2018.12.017

Ahmed, S. M., & Hossain, M. A. (2007). Knowledge and practice of unqualified and semi-qualified allopathic providers in rural Bangladesh: Implications for the HRH problem. *Health Policy*, *84*(2-3), 332-343. https://doi.org/10.1016/j.healthpol.2007.05.011

Al Salah, D. M. M., Ngweme, G. N., Laffite, A., Otamonga, J. P., Mulaji, C., & Poté, J. (2020). Hospital wastewaters: A reservoir and source of clinically relevant bacteria and antibiotic resistant genes dissemination in urban river under tropical conditions. *Ecotoxicology and Environtal Safety*, *200*, 110767. https://doi.org/10.1016/j.ecoenv.2020.110767

Amann, R. I., Ludwig, W., & Schleifer, K. H. (1995). Phylogenetic identification and in situ detection of individual microbial cells

without cultivation. *Microbiological Reviews*, *59*(1), 143-169. https://doi.org/10.1128/mr.59.1.143-169.1995

Andersson, A. F., Riemann, L., & Bertilsson, S. (2010). Pyrosequencing reveals contrasting seasonal dynamics of taxa within Baltic Sea bacterioplankton communities. *The ISME Journal*, *4*(2), 171-181. https://doi.org/10.1038/ismej.2009.108

Baquero, F., Martínez, J.-L., & Cantón, R. (2008). Antibiotics and antibiotic resistance in water environments. *Current Opinion in Biotechnology*, *19*(3), 260-265. https://doi.org/10.1016/j.copbio.2008.05.006

Belachew, S. A., Hall, L., & Selvey, L. A. (2021). Non-prescription dispensing of antibiotic agents among community drug retail outlets in Sub-Saharan African countries: a systematic review and meta-analysis. *Antimicrobial Resistance and Infection Control*, *10*(1), 13. https://doi.org/10.1186/s13756-020-00880-w

Biswas, M., Roy, D. N., Tajmim, A., Rajib, S. S., Hossain, M., Farzana, F., & Yasmen, N. (2014). Prescription of antibiotics for outpatients in Bangladesh: a cross-sectional health survey conducted in three cities. *Annals of Clinical Microbiology and Antimicrobials*, *13*, 15. https://doi.org/10.1186/1476-0711-13-15

Bokhary, H., Pangesti, K. N. A., Rashid, H., Abd El Ghany, M., & Hill-Cawthorne, G. A. (2021). Travel related antimicrobial resistance: A

systematic review. *Tropical Medicine and Infectious Disease*, *6*(1), 11. https://doi.org/10.3390/tropicalmed6010011

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114-2120. https://doi.org/10.1093/bioinformatics/btu170

Borgogna, T. R., Borgogna, J.-L., Mielke, J. A., Brown, C. J., Top, E. M., Botts, R. T., & Cummings, D. E. (2016). High diversity of CTX-M Extended-Spectrum β-Lactamases in municipal wastewater and urban wetlands. *Microbial Drug Resistance*, *22*(4), 312-320. https://doi.org/10.1089/mdr.2015.0197

Botica, M. V., Botica, I., Stamenić, V., Andrasević, A. T., Kern, J., & Spehar, S. S. (2013). Antibiotic prescription rate for upper respiratory tract infections and risks for unnecessary prescription in Croatia. *Collegium Antropologicum*, *37*(2), 449-454. https://pubmed.ncbi.nlm.nih.gov/23940988/

Boyd, D. A., Tyler, S., Christianson, S., McGeer, A., Muller, M. P., Willey, B. M., Bryce, E., Gardam, M., Nordmann, P., & Mulvey, M. R. (2004). Complete nucleotide sequence of a 92-kilobase plasmid harboring the CTX-M-15 Extended-Spectrum Beta-Lactamase involved in an outbreak in Long-Term-Care facilities in Toronto, Canada. *Antimicrobial Agents and Chemotherapy*, *48*(10), 3758-3764. https://doi.org/10.1128/aac.48.10.3758-3764.2004

Butler, M. S., & Buss, A. D. (2006). Natural products--the future scaffolds for novel antibiotics? *Biochemical Pharmacology*, *71*(7), 919-929. https://doi.org/10.1016/j.bcp.2005.10.012

Cantón, R., & Morosini, M. I. (2011). Emergence and spread of antibiotic resistance following exposure to antibiotics. *FEMS Microbiology Reviews*, *35*(5), 977-991. https://doi.org/10.1111/j.1574-6976.2011.00295.x

Cao, S. J., Deng, C. P., Li, B. Z., Dong, X. Q., & Yuan, H. L. (2010). Cloacibacterium rupense sp. nov., isolated from freshwater lake sediment. *International Journal of Systematic and Evolutionary Microbiology*, *60*(Pt 9), 2023-2026. https://doi.org/10.1099/ijs.0.017681-0

Chereau, F., Opatowski, L., Tourdjman, M., & Vong, S. (2017). Risk assessment for antibiotic resistance in South East Asia. *Bmj*, *358*, j3393. https://doi.org/10.1136/bmj.j3393

Cole, J. R., Wang, Q., Fish, J. A., Chai, B., McGarrell, D. M., Sun, Y., Brown, C. T., Porras-Alfaro, A., Kuske, C. R., & Tiedje, J. M. (2014). Ribosomal Database Project: data and tools for high throughput rRNA analysis. *Nucleic Acids Research*, *42*(Database issue), D633-642. https://doi.org/10.1093/nar/gkt1244

Countries, E. A. (2020). World Population Review. In. https://worldpopulationreview.com/

Dent, L. L., Marshall, D. R., Pratap, S., & Hulette, R. B. (2010). Multidrug resistant *Acinetobacter baumannii*: a descriptive study in a city hospital. *BMC Infectious Diseases*, *10*(1), 1-7. https://doi.org/10.1186/1471-2334-10-196

Devarajan, N., Köhler, T., Sivalingam, P., van Delden, C., Mulaji, C. K., Mpiana, P. T., Ibelings, B. W., & Poté, J. (2017). Antibiotic resistant *Pseudomonas spp.* in the aquatic environment: A prevalence study under tropical and temperate climate conditions. *Water Research*, *115*, 256-265. https://doi.org/10.1016/j.watres.2017.02.058

Devarajan, N., Laffite, A., Mulaji, C. K., Otamonga, J.-P., Mpiana, P. T., Mubedi, J. I., Prabakar, K., Ibelings, B. W., & Poté, J. (2016). Occurrence of antibiotic resistance genes and bacterial markers in a tropical river receiving hospital and urban wastewaters. *PLOS ONE*, *11*(2), e0149211. https://doi.org/10.1371/journal.pone.0149211

Dhariwal, A., Chong, J., Habib, S., King, I. L., Agellon, L. B., & Xia, J. (2017). MicrobiomeAnalyst: a web-based tool for comprehensive statistical, visual and meta-analysis of microbiome data. *Nucleic Acids Research*, *45*(W1), W180-w188. https://doi.org/10.1093/nar/gkx295

Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*, *26*(19), 2460-2461. https://doi.org/10.1093/bioinformatics/btq461

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

Edgar, R. C. (2016). *UNOISE2: improved error-correction for Illumina 16S and ITS amplicon sequencing*. Cold Spring Harbor Laboratory. https://dx.doi.org/10.1101/081257

Edgar, R. C. (2018). Accuracy of taxonomy prediction for 16S rRNA and fungal ITS sequences. *PeerJ*, *6*, e4652. https://doi.org/10.7717/peerj.4652

Ehlers, M. M., Veldsman, C., Makgotlho, E. P., Dove, M. G., Hoosen, A. A., & Kock, M. M. (2009). Detection of blaSHV, blaTEM and blaCTX-M antibiotic resistance genes in randomly selected bacterial pathogens from the Steve Biko Academic Hospital. *FEMS Immunology andl Medical Microbiology*, *56*(3), 191-196. https://doi.org/10.1111/j.1574-695X.2009.00564.x

Eiler, A., Heinrich, F., & Bertilsson, S. (2012). Coherent dynamics and association networks among lake bacterioplankton taxa. *The ISME Journal*, *6*(2), 330-342. https://doi.org/10.1038/ismej.2011.113

Ejaz, H., Younas, S., Abosalif, K. O. A., Junaid, K., Alzahrani, B., Alsrhani, A., Abdalla, A. E., Ullah, M. I., Qamar, M. U., & Hamam, S. S. M. (2021). Molecular analysis of blaSHV, blaTEM, and blaCTX-M in extended-spectrum β-lactamase producing Enterobacteriaceae

recovered from fecal specimens of animals. *PLOS ONE*, *16*(1), e0245126. https://doi.org/10.1371/journal.pone.0245126

Fournier, P. E., Richet, H., & Weinstein, R. A. (2006). The epidemiology and control of *Acinetobacter baumannii* in health care facilities. *Clinical Infectious Diseases*, *42*(5), 692-699. https://doi.org/10.1086/500202

Galand, P. E., Casamayor, E. O., Kirchman, D. L., & Lovejoy, C. (2009). Ecology of the rare microbial biosphere of the Arctic Ocean. *Proceedings of the National Academy of Sciences, 106(52),* 22427-22432. https://doi.org/10.1073/pnas.0908284106

Group, W. B. (2016). *Kingdom of Lesotho: Country partnership framework 2016-2020*. World Bank. https://documents.worldbank.org/

Gupta, S. K., Padmanabhan, B. R., Diene, S. M., Lopez-Rojas, R., Kempf, M., Landraud, L., & Rolain, J. M. (2014). ARG-ANNOT, a new bioinformatic tool to discover antibiotic resistance genes in bacterial genomes. *Antimicrobial Agents and Chemotherapy*, *58*(1), 212-220. https://doi.org/10.1128/aac.01310-13

Gurmessa, B., Ashworth, A. J., Yang, Y., Savin, M., Moore, P. A., Ricke, S. C., Corti, G., Pedretti, E. F., & Cocco, S. (2021). Variations in bacterial community structure and antimicrobial resistance gene abundance in cattle manure and poultry litter. *Environmental Research*, *197*, 111011. https://doi.org/https://doi.org/10.1016/

Hay, S. I., Rao, P. C., Dolecek, C., Day, N. P. J., Stergachis, A., Lopez, A. D., & Murray, C. J. L. (2018). Measuring and mapping the global burden of antimicrobial resistance. *BMC Medicine*, *16*(1), 78. https://doi.org/10.1186/s12916-018-1073-z

Heuer, H., Krãogerrecklenfort, E., Wellington, E. M. H., Egan, S., Elsas, J. D., Overbeek, L., Collard, J. M., Guillaume, G., Karagouni, A. D., Nikolakopoulou, T. L., & Smalla, K. (2002). Gentamicin resistance genes in environmental bacteria: prevalence and transfer. *FEMS Microbiology Ecology*, *42*(2), 289-302. https://doi.org/10.1111/j.1574-6941.2002.tb01019.x

Hu, Q., Hu, Z., Li, J., Tian, B., Xu, H., & Li, J. (2011). Detection of OXA-type carbapenemases and integrons among carbapenem-resistant *Acinetobactor baumannii* in a Teaching Hospital in China. *Journal of Basic Microbiology*, *51*(5), 467-472. https://doi.org/10.1002/j

Inouye, M., Dashnow, H., Raven, L. A., Schultz, M. B., Pope, B. J., Tomita, T., Zobel, J., & Holt, K. E. (2014). SRST2: Rapid genomic surveillance for public health and hospital microbiology labs. *Genome Medicine*, *6*(11), 90. https://doi.org/10.1186/s13073-014-0090-6

Islam, M. A., Islam, M., Hasan, R., Hossain, M. I., Nabi, A., Rahman, M., Goessens, W. H. F., Endtz, H. P., Boehm, A. B., & Faruque, S. M. (2017). Environmental spread of New Delhi Metallo-β-Lactamase-1-

producing Multidrug-Resistant bacteria in Dhaka, Bangladesh. *Applied and Environmental Microbiology*, *83*(15), AEM.00793-00717. https://doi.org/10.1128/aem.00793-17

Jani, K., Bandal, J., Rale, V., Shouche, Y., & Sharma, A. (2019). Antimicrobial resistance pattern of microorganisms isolated and identified from Godavari River across the mass gathering event. *Journal of Biosciences*, *44*(5). https://doi.org/10.1007/s12038-019-9941

Kim, S., Jensen, J. N., Aga, D. S., & Weber, A. S. (2007). Tetracycline as a selector for resistant bacteria in activated sludge. *Chemosphere*, *66*(9), 1643-1651. https://doi.org/10.1016/j.chemosphere.2006.07.066

Klein, N. C., & Cunha, B. A. (1995). Third-generation cephalosporins. *The Medical Clinics of North America*, *79(4)*, 705-719 https://doi.org/10.1016/s0025-7125(16)30034-7

Korzeniewska, E., & Harnisz, M. (2013). Extended-spectrum beta-lactamase (ESBL)-positive Enterobacteriaceae in municipal sewage and their emission to the environment. *Journal of Environmental Management*, *128*, 904-911. https://doi.org/10.1016/j.jenvman.2013.06.051

Legovi , T. (1991). Statistical ecology. A primer on methods and computing: J.A. Ludwig and J.F. Reynolds. Wiley-Interscience Publication, New York, 1988. 337 pp. ISBN 0-471-61315-0. *Ecological Modelling*, *54*, 143-144. https://dl.acm.org/doi/book/10.5555/61827

Maravić, A., Skočibušić, M., Fredotović, Ž., Šamanić, I., Cvjetan, S., Knezović, M., & Puizina, J. (2016). Urban riverine environment is a source of multidrug-resistant and ESBL-producing clinically important *Acinetobacter spp.* *Environmental Science and Pollution Research International*, *23*(4), 3525-3535. https://doi.org/10.1007/s11356-015-5586-0

Marcinek, H., Wirth, R., Muscholl-Silberhorn, A., & Gauer, M. (1998). *Enterococcus faecalis* gene transfer under natural conditions in municipal sewage water treatment plants. *Applied and Environmental Microbiology*, *64*(2), 626-632. https://doi.org/10.1128/aem.64.2.626-632.1998 .

Martinez, J. L. (2009). Environmental pollution by antibiotics and by antibiotic resistance determinants. *Environmental Pollution*, *157*(11), 2893-2902. https://doi.org/10.1016/j.envpol.2009.05.051

McMurdie, P. J., & Holmes, S. (2013). phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLOS ONE*, *8*(4), e61217. https://doi.org/10.1371/journal.pone.0061217

Mostafa Shamsuzzaman, M., & Kumar Biswas, T. (2012). Aqua chemicals in shrimp farm: A study from south-west coast of Bangladesh. *The Egyptian Journal of Aquatic Research*, *38*(4), 275-285. https://doi.org/10.1016/j.ejar.2012.12.008

Naseer, U., & Sundsfjord, A. (2011). The CTX-M conundrum: dissemination of plasmids and *Escherichia coli* clones. *Microbial Drug Resistance*, *17*(1), 83-97. https://doi.org/10.1089/mdr.2010.0132

Nuñez, L., & Moretton, J. (2007). Disinfectant-resistant bacteria in Buenos Aires city hospital wastewater. *Brazilian Journal of Microbiology*, *38*(4), 644-648. https://doi.org/10.1590/s1517-83822007000400012

Oluseyi Osunmakinde, C., Selvarajan, R., Mamba, B. B., & Msagati, T. A. M. (2019). Profiling bacterial diversity and potential pathogens in wastewater treatment plants using High-throughput sequencing analysis. *Microorganisms*, *7*(11). https://doi.org/10.3390/microorganisms7110506

Pagani, L., Dell'Amico, E., Migliavacca, R., D'Andrea, M. M., Giacobone, E., Amicosante, G., Romero, E., & Rossolini, G. M. (2003). Multiple CTX-M-Type Extended-Spectrum β-Lactamases in nosocomial isolates of Enterobacteriaceae from a hospital in Northern Italy. *Journal of Clinical Microbiology*, *41*(9), 4264-4269. https://doi.org/10.1128/jcm.41.9.4264-4269.2003

Pammi, M., Thapa, S., Balderas, M., Runge, J. K., Venkatachalam, A., & Luna, R. A. (2020). Microbiome signatures in neonatal central line associated bloodstream infections. *PLOS ONE*, *15*(1), e0227967. https://doi.org/10.1371/journal.pone.0227967

Paterson, D. L. (2006). Resistance in gram-negative bacteria: enterobacteriaceae. *The American Journal of Medicine*, *119*(6 Suppl 1), S20-28; discussion S62-70. https://doi.org/10.1016/j.amjmed.2006.03.013

Petrovich, M. L., Zilberman, A., Kaplan, A., Eliraz, G. R., Wang, Y., Langenfeld, K., Duhaime, M., Wigginton, K., Poretsky, R., Avisar, D., & Wells, G. F. (2020). Microbial and viral communities and their antibiotic resistance genes throughout a hospital wastewater treatment system. *Frontiers in Microbiology*, *11*. https://doi.org/10.3389/fmicb.2020.00153

Peura, S., Eiler, A., Bertilsson, S., Nykänen, H., Tiirola, M., & Jones, R. I. (2012). Distinct and diverse anaerobic bacterial communities in boreal lakes dominated by candidate division OD1. *The ISME Journal*, *6*(9), 1640-1652. https://doi.org/10.1038/ismej.2012.21

Quail, M., Smith, M. E., Coupland, P., Otto, T. D., Harris, S. R., Connor, T. R., Bertoni, A., Swerdlow, H. P., & Gu, Y. (2012). A tale of three next generation sequencing platforms: comparison of Ion torrent, pacific biosciences and illumina MiSeq sequencers. *BMC Genomics*, *13*(1), 341. https://doi.org/10.1186/1471-2164-13-341

Safain, K. S., Bhuyan, G. S., Tasnim, S., Hasib, S. H., Sultana, R., Islam, M. S., Mahmud-Un-Nabi, M. A., Sarker, S. K., Noor, F. A., Rahat, A., Bhuiyan, M. A. M., Islam, M. T., Manzoor, F., Anwar, S., Leung, D.,

Qadri, S. S., Qadri, F., & Mannoor, K. (2020). Situation of antibiotic resistance in Bangladesh and its association with resistance genes for horizontal transfer. *Cold Spring Harbor Laboratory*. https://dx.doi.org/10.1101/2020.04.06.027391

Sibanda, T., Selvarajan, R., & Tekere, M. (2015). Urban effluent discharges as causes of public and environmental health concerns in South Africa's aquatic milieu. *Environtal Science and Pollution Research International*, *22*(23), 18301-18317. https://doi.org/10.1007/s11356-015-5416-4

Simon, C., & Daniel, R. (2011). Metagenomic analyses: past and future trends. *Applied and Environmental Microbiology*, *77*(4), 1153-1161. https://doi.org/10.1128/aem.02345-10

Simpson, E. H. (1949). Measurement of diversity. *Nature*, *163*(4148), 688-688. https://doi.org/10.1038/163688a0

Sogin, M. L., Morrison, H. G., Huber, J. A., Welch, D. M., Huse, S. M., Neal, P. R., Arrieta, J. M., & Herndl, G. J. (2006). Microbial diversity in the deep sea and the underexplored "rare biosphere". *Proceedings of the National Academy of Sciences*, *103*(32), 12115-12120. https://doi.org/10.1073/pnas.0605127103

Sykes, R. (2010). The 2009 Garrod lecture: the evolution of antimicrobial resistance: a Darwinian perspective. *Journal of Antimicrobial*

Chemotherapy, *65*(9), 1842-1852. https://doi.org/10.1093/jac/dkq217

Tacão, M., Correia, A., & Henriques, I. (2012). Resistance to broad-spectrum antibiotics in aquatic systems: anthropogenic activities modulate the dissemination of bla(CTX-M)-like genes. *Applied and Environmental Microbiology*, *78*(12), 4134-4140. https://doi.org/10.1128/aem.00359-12

Tejerina-Garro, F. L., Maldonado, M., Ibañez, C., Pont, D., Roset, N., & Oberdorff, T. (2005). Effects of natural and anthropogenic environmental changes on riverine fish assemblages: a framework for ecological assessment of rivers. *Brazilian Archives of Biology and Technology*, *48*(1), 91-108. https://doi.org/10.1590/s1516-89132005000100013

Vaz-Moreira, I., Nunes, O. C., & Manaia, C. M. (2014). Bacterial diversity and antibiotic resistance in water habitats: searching the links with the human microbiome. *FEMS Microbiology Reviews*, *38*(4), 761-778. https://doi.org/10.1111/1574-6976.12062

Wang, Q., Wang, P., & Yang, Q. (2018). Occurrence and diversity of antibiotic resistance in untreated hospital wastewater. *Science of the Total Environment*, *621*, 990-999. https://doi.org/https://doi.org/10.1016/j.scitotenv.2017.10.128

WHO. (2015). Global action plan on antimicrobial resistance. *WHO Report*. https://www.who.int/publications/i/item/9789241509763

Williams, R. J. (2001). Globalization of antimicrobial resistance: epidemiological challenges. *Clinical Infectious Diseases*, *33*(Supplement_3), S116-S117. https://doi.org/10.1086/321835

Yamamoto, N., Bibby, K., Qian, J., Hospodsky, D., Rismani-Yazdi, H., Nazaroff, W. W., & Peccia, J. (2012). Particle-size distributions and seasonal diversity of allergenic and pathogenic fungi in outdoor air. *The ISME Journal*, *6*(10), 1801-1811. https://doi.org/10.1038/ismej.2012.30

국문초록

High-throughput sequencing 을 기반으로 한 강물과 병원 폐수 내 박테리아 다양성과 bla_{CTX-M} 유전자 변이 분석

Bhuiyan Mohammad Imtiaj Uddin

환경보건학과 환경보건학 전공

서울대학교 보건대학원

강물의 수생 환경은 인간에게 가장 집중적인 영향을 받는 생태계이다. 인간 사회에서 배출되는 가정, 병원, 축산 및 공장 폐수 등은 강으로 유입되어 강물 속에 다양한 박테리아가 존재하게 만든다. 이와 같이 오염된 수생 환경은 다양한 항생제 내성 박테리아가 혼합될 수 있는 최적의 조건을 제공하고 항생제 내성 유전자가 전달되어 강력한 병원성 박테리아를 생성할 수 있다. 특히, 병원으로부터 배출된 폐수는 높은 농도의 항생제와 항미생물제를 포함하고 있으며 항생제 처방을 받은 환자들의 장내 미생물도

포함하고 있다. 따라서, 본 연구에서는 16S rRNA 유전자 시퀀싱을 사용하여 방글라데시 다카의 병원 폐수와 강물에 존재하는 박테리아의 다양성을 분석하였다. 이를 위하여 36 개의 DNA 라이브러리를 제작하였으며 최소 1,333 개에서 최대 38,110 개에 달하는 총 4,13,297 개의 시퀀스를 획득하였다. 이를 활용하여 강물과 병원 폐수에 존재하는 8 개의 문, 12 개의 강, 13 개의 목, 22 개의 과, 102 개의 속으로 분류되는 박테리아의 분류학적 정보를 획득하였다. 그리고 Permutational multivariate analysis of variance 를 통해 강물과 병원 폐수에 존재하는 각각의 박테리아의 구성이 통계적으로 유의한 차이를 보임을 확인하였다. 또한, 병원 폐수에서 병원성을 띄는 박테리아 속의 비율이 강물보다 높은 것을 확인하였다. 추가적으로, 강물과 병원 폐수에서 임상적으로 중요한 $bla_{\text{CTX-M}}$ 유전자 변이를 분석하였다. 그람 음성 박테리아와 그람 양성 박테리아에 의해 야기되는 피부와 호흡기의 감염을 치료하기 위해 세팔로스포린 항생제를 사용하는데, $bla_{\text{CTX-M}}$ 유전자를 지닌 박테리아는 세팔로스포린에 대한 내성을 띄게 된다. 강물과 병원 폐수 내 항생제 내성 유전자를 확인하기 위하여 33 개의 $bla_{\text{CTX-M}}$ 라이브러리를 제작하였다. 현재까지 보고된 172 개의 $bla_{\text{CTX-M}}$ 유전자 변이 중 124 개를 발견하였으며 이 중 62%는 병원 폐수에 존재하는 것을 확인하였다. Non-metric multidimensional scaling 은

병원 폐수에 존재하는 *bla*$_{\text{CTX-M}}$ 변이 간에 밀접한 연관이 있는 반면, 강물에 존재하는 *bla*$_{\text{CTX-M}}$ 변이는 샘플링이 이루어진 장소에 따라 변하는 것을 확인하였다. 결론적으로, 본 연구는 강물과 병원 폐수에 다양한 종류의 박테리아가 존재하고 강물과 병원 폐수 내의 박테리아 집단은 통계적으로 유의하게 다른 것을 보여주었다. 또한, 본 연구에서 확보한 샘플에서 항생제 내성 박테리아 속을 포함하고 있는 것을 확인하였고, 이는 강물의 수생 환경과 공중 보건에 심각한 위협을 줄 수 있음을 의미한다. *bla*$_{\text{CTX-M}}$의 다양한 변이를 확인하였고, 이는 항생제 내성에 대한 전지구적 부담을 촉진할 수 있음을 의미한다.

중심어: 폐수, 박테리아 다양성, 항생제 내성, 유전자 변이, Permutational multivariate analysis of variance, Non-metric multidimensional scaling.

학번: 2019-27155

Acknowledgement

Firstly, I would like to express the earnest gratitude from the core of my heart to my supervisor, Professor Naomichi Yamamoto for his kind mentoring, guidance and invaluable support since the very first day of my journey to this program. From experimental design to data analyses, draft preparation to final thesis writing, he was there with kind suggestions and comments. My thesis is nothing but the accumulation of all these tiny but precious inputs.

Secondly, I would like to express kind gratitude to Professor Anowara Begum, my undergraduate supervisor from University of Dhaka, Bangladesh. I have always been interested to working with samples from Bangladesh. When Professor Yamamoto agreed to my proposal regarding experiments with water samples from Bangladesh, amid COVID-19 pandemic, she could manage to send samples from Bangladesh. She also served as my co-supervisor and helped me with sampling site selection, sample processing and other technical issues.

Thirdly, I would like to thank Dr. Priyanka Kumari, Research Assistant Professor and Mr. Cheolwoon Woo, PhD Candidate from our lab. Dr. Kumari taught me each and every detail of next generation sequencing. Whenever I needed any help, she was always there with her smiling face. Mr. Cheolwoon was my another mentor who was always there throughout my

journey to this MPH program. From research experiments to academic issues, he always helped me just like an elder brother. As a foreign student, what else could anyone expect!

Finally, I would like to thank all the professors of our school who have ever helped me. The knowledge they taught in the class and the inspiration they gave during the conversation made me become better and better. My sincere gratitude to the members of my thesis review committee, Professor Kyungho Choi and Professor Kyung-Duk Zoh. Their invaluable comments and inputs have made my thesis more impactful.