Isolation of Lytic Bacteriophages from Environmental water samples specific to *Vibrio cholerae;* software-based analysis of whole-genome using bioinformatics tools

Submitted by

Safwana Murshed Arthi

Student ID: 15336009

A THESIS SUBMITTED TO THE DEPARTMENT OF MATHEMATICS AND NATURAL SCIENCES, BRAC UNIVERSITY IN PARTIAL FULFILLMENT OF THE REQUIREMENT FOR THE DEGREE OF BACHELOR OF SCIENCE IN BIOTECHNOLOGY

Department of Mathematics and Natural Sciences

Biotechnology Program

Brac University
[August] [2020]

Declaration

It is hereby declared that

1. The thesis submitted is my own original work while completing degree at Brac University.
2. The thesis does not contain material previously published or written by a third party, except where this is appropriately cited through full and accurate referencing.
3. The thesis does not contain material which has been accepted, or submitted, for any other degree or diploma at a university or other institution.
4. I have acknowledged all main sources of help.

Safwana Murshed Arthi
Student ID- 15336009

Approval

The thesis/project titled "Isolation of Lytic Bacteriophages from Environmental Water Samples Specific to *Vibrio cholerae*; software-based analysis of whole-genome using bioinformatics tools" is submitted by Safwana Murshed Arthi of Summer,2015, has been accepted as satisfactory in partial fulfillment of the requirement for the degree of Bachelor of Science in Biotechnology on 27th August, 2020

Examining Committee:

Supervisor:
(Member)

_____

Dr. Iftekhar Bin Naser
Assistant Professor
Department of Mathematics and Natural Sciences
BRAC University

Program Coordinator:
(Member)

_____

Dr. Iftekhar Bin Naser
Assistant Professor
Department of Mathematics and Natural Sciences
BRAC University

Departmental Head:
(Chair)

_____

A F M Yusuf Haider
Professor and Chairperson
Department of Mathematics and Natural Sciences
BRAC University

Ethics Statement

In this research, surface water samples were collected from different lakes and rivers. Biosafety and Bioethics were properly maintained in the laboratory, no animal model was used.

Abstract

Vibrio cholerae is an opportunistic, gram-negative bacterium that is commonly found in fresh and saltwater, soil, food, or in the intestine of humans and animals. It is the most causative agent of cholera; a chronic diarrheal disease with epidemic and pandemic potentials that might lead to death if untreated. Cholera has been recognized as a significant public health problem around the world and already faced seven cholera pandemics. Lytic bacteriophages specific to *V. cholerae* are the most potent triggers of causing cholera outbreaks by killing bacteria present in the infected individuals and environmental reservoir. This study includes molecular isolation of bacteriophage DNA from the environmental water samples and bioinformatics software analysis as well. However, a very unique bacteriophage (WT1774), was isolated from this study which showed the least similarities among other phages. Although the whole genome sequencing might not be done properly, all relevant computational analyses (including genomic and phenotypic characteristics) couldn't be performed as expected. To retrieve all essential information for higher research, further analysis is needed.

Keywords: *Vibrio cholerae;* bacteriophages; cholera; pandemic; genome; sequencing

Dedication

"Dedicated to my family and my well-wishers who were always there for me and I am blessed to have them all"

Acknowledgement

Primarily I would like to express my highest gratitude to Almighty Allah for giving me blessings, strength, mercies, and the ability to complete this thesis with success. Moreover, I would like to devote my deepest appreciation and endless love to my beloved mother who always gives courage, support, and prayers for me for my greatest success. In this research, many people provided motivation, support, advice even remarks that helped me a lot. So it would be an honor for me to express my gratitude and appreciation to all of them.

First and foremost, I would like to express my deepest gratitude and appreciation to my thesis supervisor **Dr. Iftekhar Bin Naser**, (Assistant Professor, Coordinator, Biotechnology Program, Brac University), who provided me all the facilities for this project. Moreover, his constant effort and encouragement towards my research allowed me to grow as a researcher. It was a great honor to finish this work under his supervision.

I also hereby express my sincere gratitude towards **Professor A F M Yusuf Haider**, Chairperson of MNS Department, and **Professor Mahboob Hossain**, Coordinator, Microbiology Program for allowing me and encouraging me to complete my undergraduate thesis.

I would like to offer my special thanks to Tushar Ahmed Shishir, Arka Roy, Teaching Assistant, BRAC University, and Arafat Khan Antu and Farzana Mansur Priyanka who were always there to guide me with their valuable advice throughout my thesis project. I would also like to thank the Lab Officers and other staff for their constant support and expert guidance during laboratory work.

I extend my special thanks to my seniors and my batch mates in the laboratory for providing me with a good work environment by their advice and encouragement to make me feel at home in my hard times and lending me a hand in my work whenever I needed it.

Last but not the least, I especially thank my dearest friend, Jahid Hasan for his constant motivation with the technological support whenever I faced difficulties regarding this dissertation and for always being there for me.

Safwana Murshed Arthi

Table of Contents

List of Tables

List of Figures

List of Acronyms

V.cholerae	*Vibrio cholerae*
E. coli	*Escherichia coli*
DNA	Deoxy ribonucleotide acid
RNA	Ribonucleic acid
DNAse Solution	Deoxy ribonuclease Solution
RNAse Solution	Ribonuclease Solution
ds	Double stranded
ss	Single stranded
Icddr,b	International Centre for Diarrhoeal Disease Research, Bangladesh
CRISPR	Clustered regularly interspaced short palindromic repeats
rpm	Rotation per minute
EDTA	Ethylenediaminetetraacetic acid
RAST	Rapid Annotation using Subsystem Technology
NCBI	National Center for Biotechnology Information
BLAST	Basic Local Alignment Search Tool
FASTA	FAST-All
bp	Base pair

Chapter 1

[Introduction]

1.1 Introduction

Bacteriophages or phages are the most abundant bacterial viruses on earth. They are diverse group of viruses which are easily manipulated and therefore they have potential uses in biotechnology, research, and therapeutics. Basically, the term “phage” is derived from a Greek word called “phagein”, meaning “to devour”. The genetic material of phages is either DNA or RNA which is encapsulated by a protein coat. Four to hundred genes may be encoded by these genomes. Phages infect bacteria to propagate and after vegetative multiplication they kill their hosts by integrating the phage genome for many generations (Inal JM, 2003). For each bacteria, the estimated number of phage viruses are 5 to 10 (Weinbauer, 2004).Locations which are populated with bacterial hosts,for example soil, environmental waste water or the intestine of animals are largely distributed by phages. In the International Nucleotide Sequence Database Consortium (INSDC), more than 25000 phage nucleotide sequences have been submitted from the year1997 to 2017 (Evelien M Adriaenssens&J. Rodney Brister, 2017) and it is predicted that more phages are waiting to be discovered in upcoming years.

Cholera is an acute watery diarrheal syndrome that is caused by the bacterium *Vibrio cholerae*; though a small portion of released toxins are responsible of causing cholera. Pandemic cholera is caused by two *V. cholerae* serogroups such as O1 and O139. Based on the phenotypic and genotypic characteristics, *V. cholerae* O1 is subdivided into two biotypes, classical and El Tor. However,*V. cholerae*O1, classical biotype, was the presumed etiologic agent of the first six pandemics, while the El Tor lineage is the dominant strain in the current seventh pandemic (7P).

1.2 Research objective & specific aim

Objectives:

The main objective of my research was to isolate environmental bacteriophages against

Vibrio cholerae strain specific for non O1; DNA sequencing and analysis through different bioinformatics software.

Specific aims:

1. Isolation of lytic bacteriophage from the numerous surface water samples toward *Vibrio cholerae* strain.

2. The whole DNA Sequencing of the bacteriophage

3. Software-based Analysis had been focused on

   i. Review of the complete genome sequence (BLASTn, annotation, genomic organization)

   ii. Study of Hypothetical Proteins

   iii. Detection of CRISPRs

   iv. Construction of phylogenetic tree

Chapter 2

[Literature Review]

2.1 History of Bacteriophages

In early 2011, Alexander Sulakvelidze defined bacteriophages as “the most ubiquitous organisms on Earth, playing a significant role in maintaining microbial balance on this planet” in the introduction part of the new *Bacteriophage* journal. The therapeutic implication of bacteriophage which is also called ‘the phage therapy’ was first conducted by Félix d’Herelle. However, phage research continued at a fundamental level in the west where the study of phage played a major role in some momentous discoveries in biological science. It led to the identification of DNA as genetic material (Van Valen et al., 2012), understanding of genetic code and phenomenon of restriction-modification and to the development of molecular recombinant technology. Because of the specificity of cellular target hosts, phage therapy is proposed to treat acute and chronic infections with the disciplines of dermatology, ophthalmology, urology, stomatology, pediatrics, otolaryngology, and surgery. One the other hand, Phage derived proteins are now being used as diagnostics agents (Smith et al., 2001), therapeutic tools (Loeffler et al., 2001; Schuch et al., 2002) and for discovering new drug (Liu et al., 2004).

In 1980, American biochemist George P. Smith developed a technology known as phage display, which allowed for the generation of engineered proteins. On the other hand, British biochemist Gregory P. Winter subsequently refined phage display technology for the development of human antibody proteins. Compared with previous therapeutic antibodies derived from animals,their new proteins could be used to treat diseases in humans with less risk of dangerous immune reactions.In 2002, U.S. Food and Drug Administration (FDA)approved the first fully human antibody made via phage display, introduced as

Adalimumab (Humira) for the treatment of rheumatoid arthritis. In 2018, Smith and Winter were awarded a share of the Nobel Prize in Chemistry for their discoveries relating to phage display.

2.1.1 Characteristics and Early History of Bacteriophages:

Bacteriophages are the foremost numerous organic structures within the biosphere and the biggest infection known. Followed phages constitute an estimated 96% of known phages. Thousands of assortments of phages exist, each of which may be sully as if it were one sort or many sorts of microbes or archaea. Phages are categorized in a number of infection families; for occurrence three lytic phages (PhiRP1, PhiRP2, and PhiRP3) decide for Robinia Pseudoacacia rhizobia were disengaged from the soil beneath dark locust. Some others which are found in our nature are Inoviridae, Microviridae, Rudiviridae, and Tectiviridae. Like all infections, phages are straightforward life forms that comprise a center of hereditary substance (nucleic corrosive) encompassed by a protein capsid. The nucleic corrosive may be either DNA or RNA and may be double-stranded or single-stranded. There are three fundamental efficient shapes of phage: an icosahedral (20-sided) head with a tail, an icosahedral head without a tail, and a filamentous form.

An English bacteriologist, Earnest HanburyHankin first stated that some sort of antibacterial properties was found that helped against Cholera disease from the water sample of Ganges and Yamuna rivers in India (Hankin, 1896).Another British bacteriologist, Fredrick Twort, also discovered an unknown agent that could kill bacteria but due to World War I, his research had been postponed (Twort, 1915). Finally, in 1917, Félix d'Herelle (1917) independently discovered bacteriophages which are small viruses displaying the ability to kill bacteria while they do not affect cell lines from other organisms (Xavier Wittebole, 2014). Based on this fundamental work, in 1922, he published "The Bacteriophage " which caught the attention of some commercial companies. These companies were interested in the commercial production

of phages and one of the most renowned companies was established; known as L'Oréal in today's world.

2.1.2 Classification of Bacteriophages:

In the early 1940s, the viral nature of bacteriophages was recognized under the electron microscope. On the basis of morphology and the type of nucleic acid, Lwoff, Horne and Tournier published the order *Urovirales* for tailed phages, *Microviridae* for filamentous, families for *Inoviridae* and ΦX-type phages as well. In 1967, Bradley recognized six basic phage types: tailed, filamentous, icosahedral phages with single-stranded DNA (ssDNA) or single-stranded RNA (ssRNA) and still now it is the basis of phage classification. Later on, the International Committee on Taxonomy of Viruses (ICTV) classified phages in six genera based on the morphology and the types of nucleic acid.

Table 1: ICTV classification of prokaryotic (bacterial and archaeal) viruses [Mc Grath and van Sinderen, 2007]

Order	Family	Morphology	Nucleic Acid	Example
Caudovirales	*Myoviridae*	*Nonenveloped, Contractile tail*	*Linear dsDNA*	*T4 phage, Mu,*
	Siphoviridae	*Nonenveloped, Contractile tail(long)*	*Linear dsDNA*	*λ phage, T5 phage*
	Podoviridae	*Nonenveloped, Noncontractile tail(short)*	*Linear dsDNA*	*T7 phage, T3 phage*
Ligamenvirales	*Lipothrixviridae*	*Enveloped, rodshaped*	*Linear dsDNA*	*T7 phage, T3 phage*
	Rudiviridae	*Nonenveloped, rod-shaped*	*Linear dsDNA*	*Acidianus*

				Filamentous virus 1
	Ampullaviridae	*Enveloped, bottle-shaped*	*Linear dsDNA*	
	Bicaudaviridae	*Nonenveloped, lemon-shaped*	*Circular dsDNA*	
	Clavaviridae	*Nonenveloped, rod-shaped*	*Circular dsDNA*	
	Corticoviridae	*Nonenveloped,*	*Circular dsDNA*	
Unassigned	*Cystoviridae*	*Enveloped, spherical*	*Segmented dsRNA*	
	Fuselloviridae	*Nonenveloped, lemon-shaped*	*Circular dsDNA*	
	Globuloviridae	*Enveloped, isometric*	*Linear dsDNA*	
	Guttaviridae	*Nonenveloped, ovoid*	*Circular dsDNA*	
	Inoviridae	*Nonenveloped, filamentous*	*Circular ssDNA*	*M13*
	Leviviridae	*Nonenveloped, isometric*	*Linear ssRNA*	*MS2, Qβ*
	Microviridae	*Nonenveloped, isometric*	*Circular ssDNA*	*ΦX174*
	Plasmaviridae	*Enveloped, pleomorphic*	*Circular ssDNA*	
	Tectiviridae	*Nonenveloped, isometric*	*Linear dsDNA*	

2.1.3 Structure and Infection:

By infecting the host cell, bacteriophages replicate. Steps involved in this process are referred to as the life cycle of the phage. Via the lytic stage, phages replicate, but others switch between the lytic and lysogenic cycles. The lambda phage is in a position to infect bacteria and the rotation between lytic cells and the lysogenic phase.

1 Virus Particle Structure:

The structures of bacteriophages are complex, but most of them share some similar features. The bacteriophage DNA is protected by the icosahedral capsid, which is attached to the contractile sheath, a highly specialized and extremely efficient phage component required for infecting its host. The hexagonally shaped baseplate is situated at the distal end of the contractile sheath, and coordinates the movement of the tail fibers that initially sense the presence of the host, the short tail fibers that unfold from underneath the baseplate to firmly anchor on its bacterial host surface, and the spiral contractile sheath surrounding a core pipe that contracts, ejecting DNA into the bacterial host.

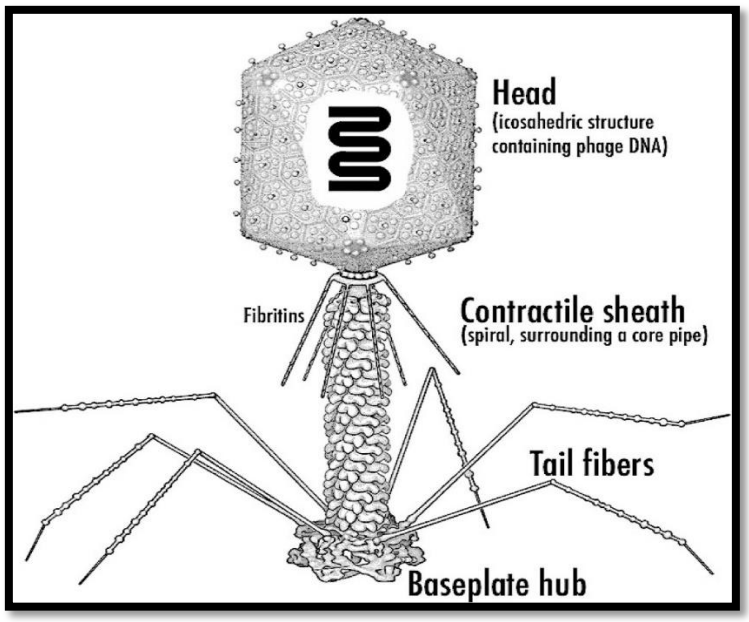

Figure 1: Structure of a prototypical bacteriophage (Rossmann et al, 2005)

Based on genetic materials, morphology and characteristics, bacteriophage particles can be divided into four major groups such as single stranded DNA phages (ssDNA), double stranded DNA phages (dsDNA), single stranded RNA phages (ssRNA), and double stranded RNA phages (dsRNA).

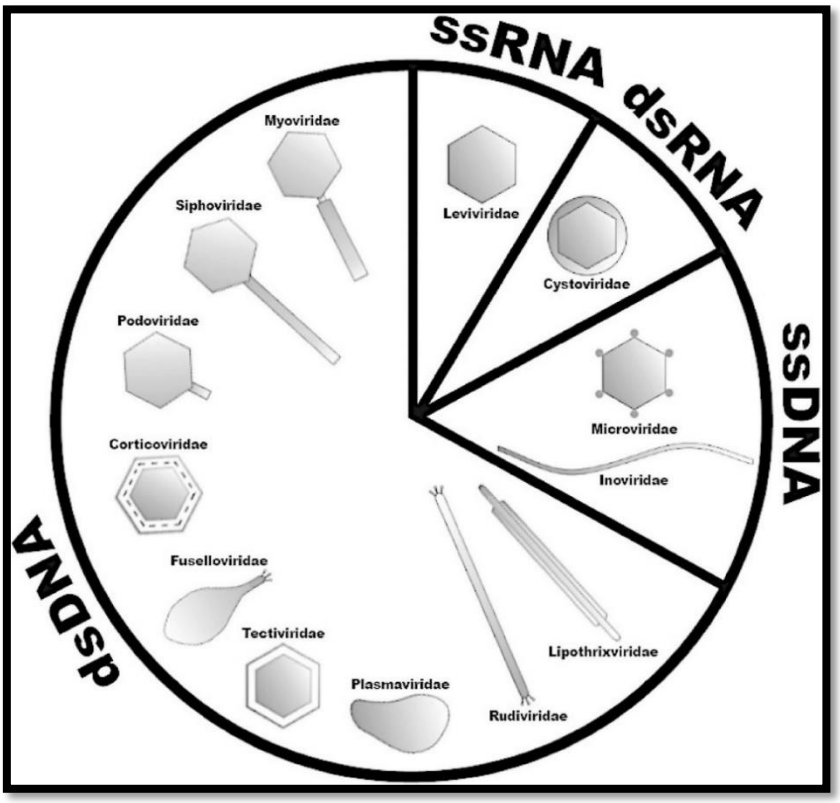

Figure 2: Classification of bacteriophage (Lilium K, 2018)

2 Mode of actions: Lytic & Lysogenic

Depending on the type of bacteriophage, either lytic or lysogenic cycle took place for the bacterial infection. Lytic bacteriophages are commonly termed as virulent phages; attacked host's metabolic activities by replicating viral genetic materials into the host's cytoplasm. As a result, more lytic phages are produced and subsequently lead to cell lysis of the bacterium. On the other hand, lysogenic pages are termed as temperature bacteriophages; replicates at a later stage without lysing the host.

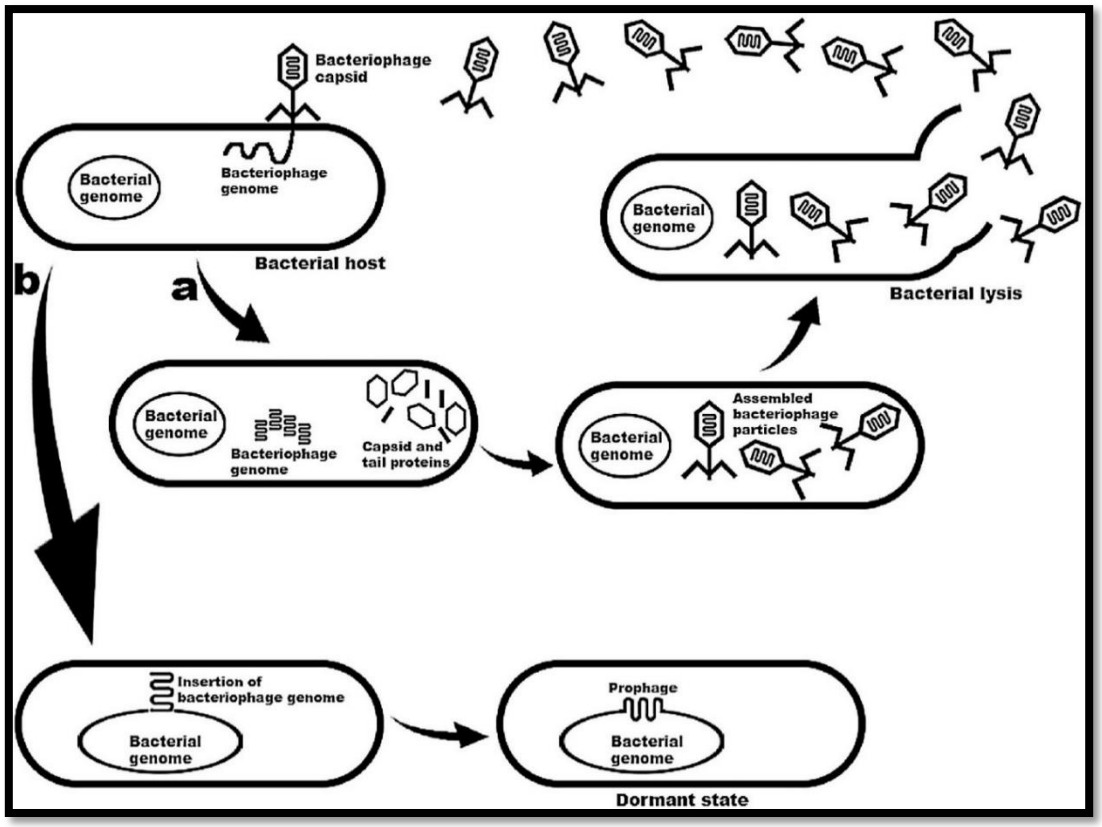

Figure 3: Lytic (a) & Lysogenic (b) infections of Bacteriophage (Lilium K, 2018)

2.2 History of Cholera & Cholera Pandemics

Cholera is a severe diarrheal disease, caused by the infectious *Vibrio cholerae* bacterium. Cholera has been recognized as a significant public health problem around the world and already faced seven cholera pandemics. Before cholera had been recognized as a disease, numerous hints were explained by Gaspar Correa (Portuguese historian and author of Legendary India) as an epidemic called “moryxy” in 1543 in the Ganges Delta.

From the past fifty years, seven cholera pandemics were identified in all over the world [Fig 4]. The first six pandemics were originated in the Indian subcontinent and then spread to South Asia, Africa, Middle East, Southern Europe and finally reached in the South and Central America [WHO Cholera Annual Report, 2017]. The seven pandemics (7P) have occurred in the past two centuries which was mainly originated in Indonesia; from 1961 to the present time.

The more recent outbreaks in Haiti and Yemen are also included in the ongoing pandemic.

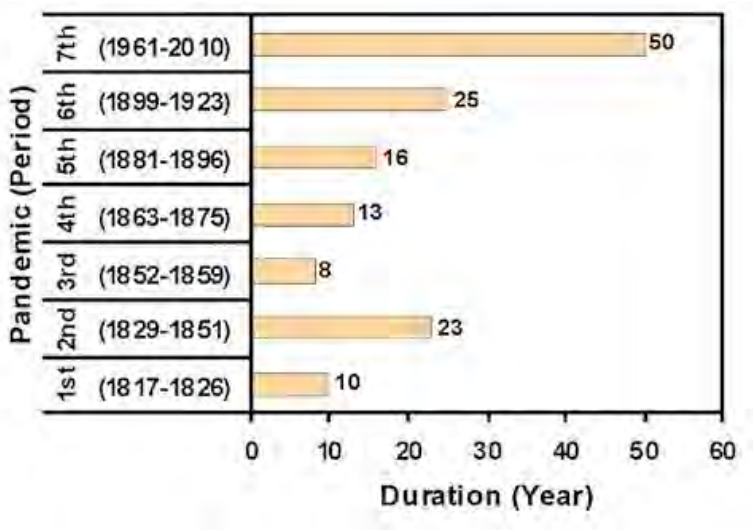

Figure 4: Duration & Period of Seven Cholera Pandemics [6]

2.3 Usefulness of Phages

Phages have been studied as model organisms to gain insights into basic genetic concepts, such as viral gene expression. Research on bacteriophage gave us much of our understanding of viruses and many fundamental concepts of molecular biology.

2.3.1 Phage Therapy:

Bacteriophages are very specific to the host organism and mostly it is harmless to the host and to other bacteria. To consider some extent, bacteriophage could be used as therapeutic agents over antibiotics. Antibiotic has poor penetration capacity with a spread of antibiotic resistant genes where phages could be much more specific against pathogens. The idea of using phages clinically was first convicted by Felix d'Herelle in 1919 to treat bacterial dysentery cases. Unfortunately, those trials were not successful because of poor understanding of phages activity in human body. Later on, those medical trials were stopped after introduction of pharmaceutical antibiotics in the early 1940's [Praveena Devi; et al 2019]. Recently, phage

treatment is reconsidered due to the current upward trend of bacterial resistance and availability of necessary molecular techniques and tools to precisely assess the safety and efficacy of using phages. Moreover, FDA (Food and Drug Administration) also approved the first US clinical trial of intravenously administered bacteriophage therapy in 2019.

2.3.2 Applications in Biotechnology and Research:

In modern Biotechnology, bacteriophages are considered as natural killers of bacteria and an alternative of antibiotics. Phages can be used in different ways such as agriculture, petroleum industry, food industry and most importantly in the clinical research. For the implementation in several sectors, bacteriophages were commercially started producing in larger scale through some renowned companies. However, necessary precautions are taken before applying phages in the food or drugs. Recently, phages were used in the treatment of localized infections in human [10]. "Pyophage & Sekstaphage" technology was applied for the treatment of wound infections. Moreover, several phage cocktail trials were conducted for three years in France, Belgium and Switzerland for the treatment of burn wound infections. Although phages are used for the treatment of ulcers but no specific cocktails are found for this treatment and this application is very new and under processing. Bacteriophages have been used as transport for vaccine delivery in two ways: 1) vaccinating with phages expressing vaccine antigens on their surface and 2) by incorporating a DNA vaccine expression cassette into phage genome and using the phage particle to deliver that DNA cassette (Clark and March, 2004). Last but not the least, recently the uses and application of 'PhagoBioDerm' reduces the microbial infections in wounds [Jikia et al., 2005]. Several clinical trials were performed and it worked significantly.

Chapter 3

[Material and Method]

3.1 Place of Study

The laboratory research was conducted at the Department of Mathematics and Natural Sciences Biotechnology and Microbiology Laboratory, BRAC University, Dhaka, Bangladesh.

After DNA extraction, a sample was sent to the International Centre for Diarrhoeal Disease Research, Bangladesh (icddr,b) for DNA sequencing.

Further Software-Based Analysis was done through online and offline bioinformatics tools.

3.2 Sample Collection

To collect water samples, four areas inside Dhaka city were covered. They are Gulshan Lake, Dhanmondi Lake, Bosila and Turag River; using caution without direct contact with these samples as they contain human pathogens mostly.

3.3 Preparation of Culture Media, Reagents and Solutions in the laboratory

a) Preparation of Luria-Bertani Broth (LB):

According to the instruction, to prepare 1000ml of liquid broth, 20gms of LB powder needs to be added with 1000ml distilled water. The powder was measured using an electronic balance machine and then added distilled water to a flask. Next, it needs to be stirred well and heated on a Bunsen burner until it is clear.

Then the flask was covered with aluminum foil and autoclaved for about an hour at 121°C. After sterilization, it was stored at 4°C.

b) Preparation of Luria-Bertani Agar (LA):

20gms of LB and 15gms of Nutrient agar powder were added to prepare 1000ml of LA media with distilled water. Next, it was heated and autoclaved as same as LB broth preparation. After sterilization, the media was poured onto sterilized Petri dishes on a vertical laminar flow cabinet and kept them to solidify. After solidification, the required amount of plates were used and rest were stored at 4°C.

c) Preparation of 0.7 % Soft Nutrient Agar:

To perform double layer plaque assay, 0.7% soft nutrient agar was prepared as well. For preparing this, within 500ml distilled water 2gms of LB powder was mixed well in a flask. Then heated and autoclaved it at 121°C for 15 minutes and finally after cooling down stored it at 4°C.

d) SM buffer:

To prepare 1000ml of SM buffer, there was added 5.8 gm NaCl, 2 gm $MgSO4 \cdot 7H2O$, 50 ml Tris-Cl (1M, pH 7.4), and 5 ml Gelatin (2 % w/v) respectively. Then the buffer was autoclaved at 121°C for 15 minutes and finally was stored at room temperature.

e) 70 % ethanol:

For avoiding microbial contamination 70% rubbing ethanol was prepared. To prepare 1000ml of 70% ethanol, 737 ml of 95 % ethanol was added to deionized water.

3.3.1 Bacterial Culture:

In this research, all bacterial cultures were used from the stock of Biotechnology and Microbiology laboratory of BRAC University. At first, the bacterial samples were streaked on a fresh LA plate and incubated for overnight at 37°C to get single bacterial colonies. After checking the growth, plates were stored at 4°C for further use. Before each experiment,

bacterial samples were freshly sub-cultured and 24 hours cultures were used to maintain the purity and viability of the organisms.

3.3.2 Bacteriophage Isolation:

For the isolation of lytic bacteriophages specific to Vibrio cholerae, a total of 20 surface water samples were collected. To perform this research, sewage water was considered as an ideal source to get lytic bacteriophages since it contains high numbers of diverse bacteria as well as different which maximize the chance of finding bacteriophages specific to *Vibrio cholerae*.

1. **Collection and Processing of water sample:**

During collecting sewage water sample, an autoclaved water collection bottle was taken for pouring water in it. Then the sample was filtered through Whattman filter paper followed by syringe filtration (0.22 μm) using a syringe filter.

2. **Bacterial Cocktail Preparation:**

At first, five different strains of the same bacterial species (specific for *V. cholerae*) were added in a vial containing 3ml of LB along with a single colony of the freshly cultured bacterial species. The inoculated broth was mixed well and incubated for 2.5 hours in the shaker incubator in order to keep the young culture consistent in its lag phase.

3. **Phage Enrichment:**

The water samples were first well shaken for a couple of minutes then they were centrifuged at 10000 rpm for 10 minutes to remove bacteria and other debris. After that the culture was kept for 2.5 hours of incubation. Next, the filtered water was added to the cocktail in the ratio of 1:1 and again incubated for 4 hours in the shaker incubator. Then again the samples were centrifuged at 13000 rpm for 15 minutes.

The clear supernatant was then filtered by passing through 0.22 μm low protein binding PES syringe filter and collected in another sterile falcon tube and stored at 4°C. It was consecutively tested for the presence of bacteriophage by spot test in several dilutions. For this study, the filtrate was diluted up to 10^-7.

4. **Preparation of Young Bacterial Culture:**

Those specific strains of *Vibrio cholerae* which were previously used for preparing cocktail were individually inoculating in LB broth and incubated for 2 hours in a shaker incubator.

5. **Double Layer Assay:**

The double layer assay was performed to create an even lawn of bacterial culture in an LA media containing plate. In an autoclaved vial, 0.3ml prepared young bacterial culture was mixed with 3ml of 0.7% soft agar and poured it immediately onto LA plate and quickly rotated the plate to let the soft agar spread evenly over the plate and kept for sometimes to let it dry.

6. **Plaque Assay:**

In order to isolate phage from collected water, 5µL of the diluted sample was added dropwise on the marked position of the plate and kept for overnight incubation. JSF 35 was used as a positive control and Saline was used here as a negative control.

3.3.2 Stock Preparation of Phage:

After overnight incubation, the plaque was collected and stored in the mixture of SM Buffer and chloroform in the ratio of 10:1 and stored at 4ºC. The next day in the same phage isolation process phage was enriched and centrifuged at 13000 rpm for 15 minutes. Then the clear supernatant was filtered through 0.22 μm low protein binding PES syringe filter membrane and finally stored at 4ºC.

3.4 DNA Isolation & Sequencing of Bacteriophage whole genome

Several steps were performed in this stage:

1. Previously the single plaque was stored in the SM buffer along with Chloroform and this solution was used for phage DNA Isolation.
2. 300 µl of specific host bacteria along with 1mL of the specific bacteriophage was poured on LA plates to perform double layer assay and then incubated overnight at 37ºC.
3. The next day, LB was added to the plates. By using sterilized spreader, phage was collected in a sterilized falcon tube.
4. Then it was centrifuged at 10,000 rpm for 10 minutes.
5. Next the supernatant was collected and passed through a 0.22 µm syringe filter.
6. DNase was added to the supernatant at a concentration of 1unit/100µl. Then again 10X buffer in the solution is added at 1:10 ratio to facilitate DNase. After mixing the solution properly, it was incubated at 37 ºC overnight.
7. After that, Proteinase-K was added to the solution at the conc. of 5mg/ml. To create a working solution, 10X Proteinase-K buffer is added at 1:10 ratio. The solution is then mixed well and incubated at 37ºC overnight.
8. By several gentle inversion, Phenol-Chloroform and Iso-amyl Alcohol was added at 25:24:1 ratio.
9. Later on, the solution was again centrifuged at 14,000 rpm for 15 minutes and the upper aqueous layer was collected in sterilized falcon tubes.
10. Again Phenol-Chloroform and iso-amyl Alcohol was added in the same ratio and volume and centrifuged under same conditions.

11. Double vol. of absolute ethanol was added to the supernatant and again incubated at -20ºC overnight.
12. Centrifugation was done again at 14,000 rpm for 15 minutes at 4ºC.
13. After removing the supernatant, the pellet was collected and washed with cold 70% ethanol.
14. The pellet was the desired DNA and finally it was stored in a Tris-EDTA buffer solution.

In order to genome sequencing, later on the DNA sample was sent to the International Centre for Diarrhoeal Disease Research, Bangladesh (icddr,b)

3.5 The whole genome sequence Analysis:

The complete genome of bacteriophage was obtained as a FASTA format and then completely annotated by RAST (Version 2.0). By using BLAST, each predicted gene was analyzed from the NCBI database. SnapGene was used to scaffold the contigs of this genome perfectly.

3.6 The Genomic Reconstruction of Bacteriophage:

To generate a circular DNA map, DNA plotter was used. DNA sequence (.gbk) was inserted to this software and generated the genomic construction. It also represents the regions and the graphical view of GC plot and GC skew as well.

3.7 Hypothetical Protein Analysis:

For the analysis of hypothetical proteins from the sequence, RAST annotation was performed online.

3.8 CRISPRs Finder:

For the detection of CRISPR array or cas protein from the DNA sequence, CRISPR finder program online was used here.

3.8 Phylogenetic Analysis:

To identify the evolutionary relationship between similar phages, phylogenetic tree was generated by using BLAST from NCBI database.

Chapter 4

[Result]

This total study was split into 2 parts. First of all, water samples were collected to isolate bacteriophage and then DNA was extracted from it. Obtaining the whole genome sequence, several genome analysis were performed in the second part.

4.1 Bacteriophage Isolation

For the isolation of *Vibrio cholerae* specific bacteriophages, twenty environmental water samples were collected. After that the samples were enriched with 10 bacterial strains for screening vibrio phages such as WT-324 V11.16;006-V1015;WT333V11.16; WT1667;WT1774;WT373 VO1.17;WT335 V1116;WT376 VO1.17; WT371 VO1.17;WT375 VO1.17. During this isolation process, 6 *Vibrio cholerae* specific bacteriophage were obtained. Among all of the isolated results there were 4 of them that gave clear lytic zones.

Table 2: Isolation of Bacteriophages from environmental water samples

Date for Water Collection	**Location**	**Bacteriophages isolated from the specific strains of *Vibrio-chlorae***
28th January'19	Gulshan Lake	-
05th February'19	Bosila	-
10th February'19	Dhanmondi Lake	WT-324 V11.16, WT 1774, WT1667, 006-V1015
16th February'19	Turag	-
10th March'19	Gulshan Lake	-
12th March'19	Bosila	WT 327
13th March'19	Dhanmondi Lake	WT 327
17th March'19	Turag	-
2nd April '19	Gulshan Lake	-

20th May '19	Gulshan Lake	-
21st May '19	Turag River	-
17th June '19	Bosila	-
18th June '19	Dhanmondi Lake	-
9th july '19	Gulshan Lake	-
10th July'19	Dhanmondi Lake	-
20th July'19	Turag River	-
1st August '19	Bosila	-
2nd August '19	Gulshan Lake	-
17th August '19	Gulshan Lake	-

The process of sampling commenced from January 2019-August 2019. Although most of the samples (7 in total) were collected from Gulshan Lake, no bacteriophages were found. On the other hand, 5 lytic zoned phages were collected from the sample of Dhanmondi Lakes.

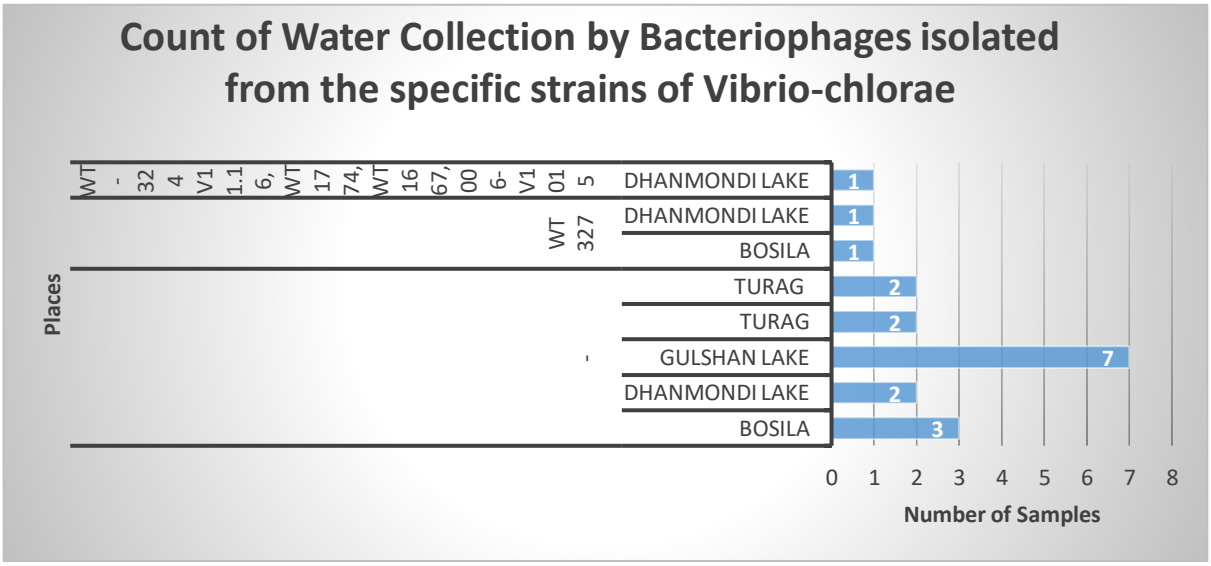

Figure 5: Bacteriophages Isolation

The analysis of the enriched sample water as carried out by double layer assay, concluded by the conformation of finding individual plaque in different dilutions.

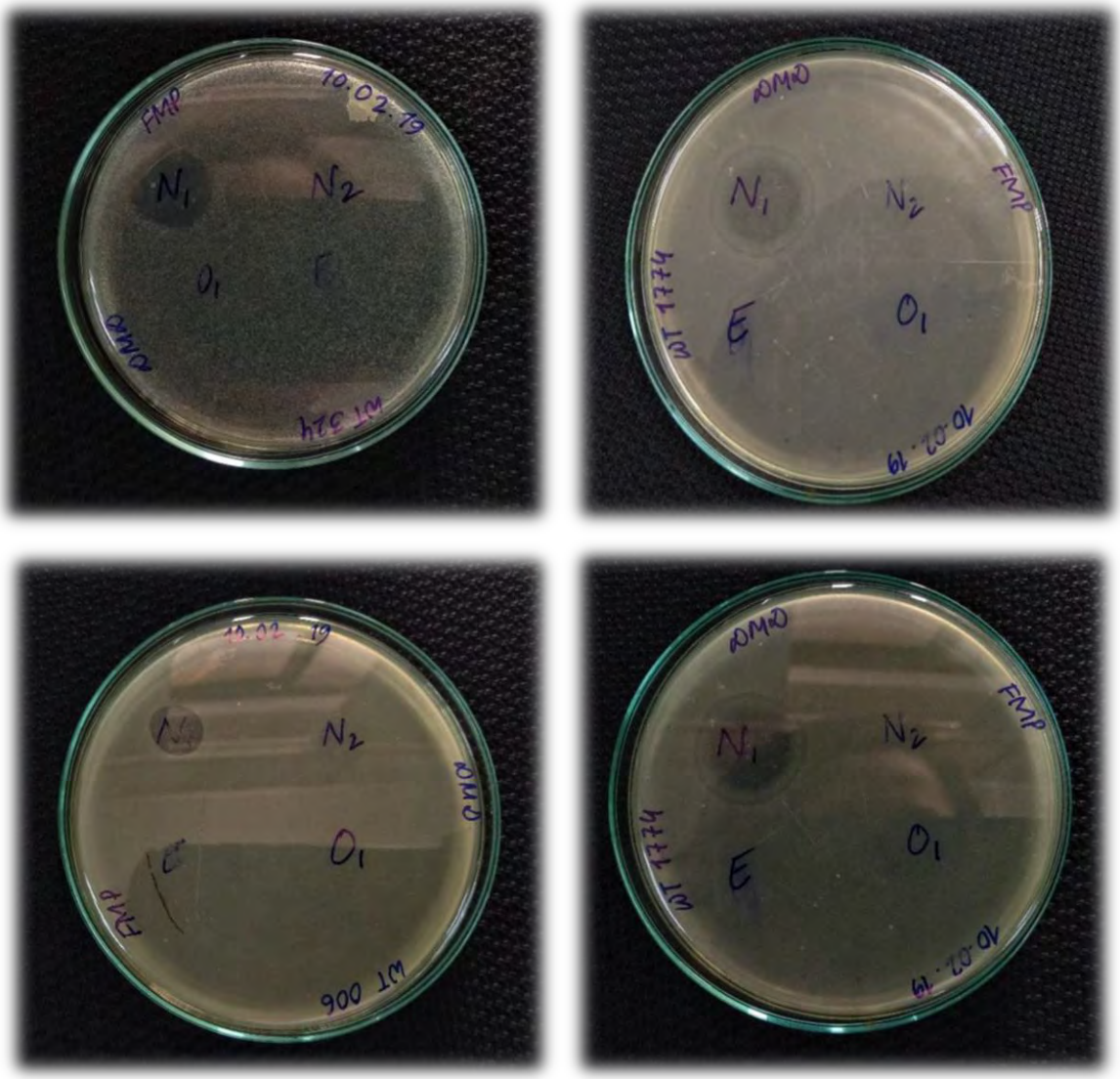

Figure 6: Bacterial Plaque Assay from the water sample

The isolated plaque that gave the clearest and isolated individual bacteriophage (WT1774), its DNA was isolated and then sent for whole genome sequencing in icddr,b. Later on, different bioinformatics methods were applied to analyze the genetic materials.

4.2 The Genomic Reconstruction of Bacteriophage

The genome map was done through DNA Plotter. The innermost circles (Figure 6) coloured in dark yellow and purple indicate the positive and negative GC skew respectively. Gray circle indicates GC content. The functional categories of ORFs are indicated by specific colours; maroon ORFs represent hypothetical proteins, sky blue ORFs represent ribonucleotide reductase, black ORF represents tRNA-Met-CAT, light yellow, light green and pink ORF represents tRNA-Arg-TCT, tRNA-Arg-CCT, tRNA-pseudo respectively. Cyan colour represents the RND efflux system. Navy blue, brown and bottle green ORFs represent ribose-phosphate, nicotinamide phosphoribosyl transferase and phage major capsid protein of Caudovirales respectively. Scale units are base pairs.

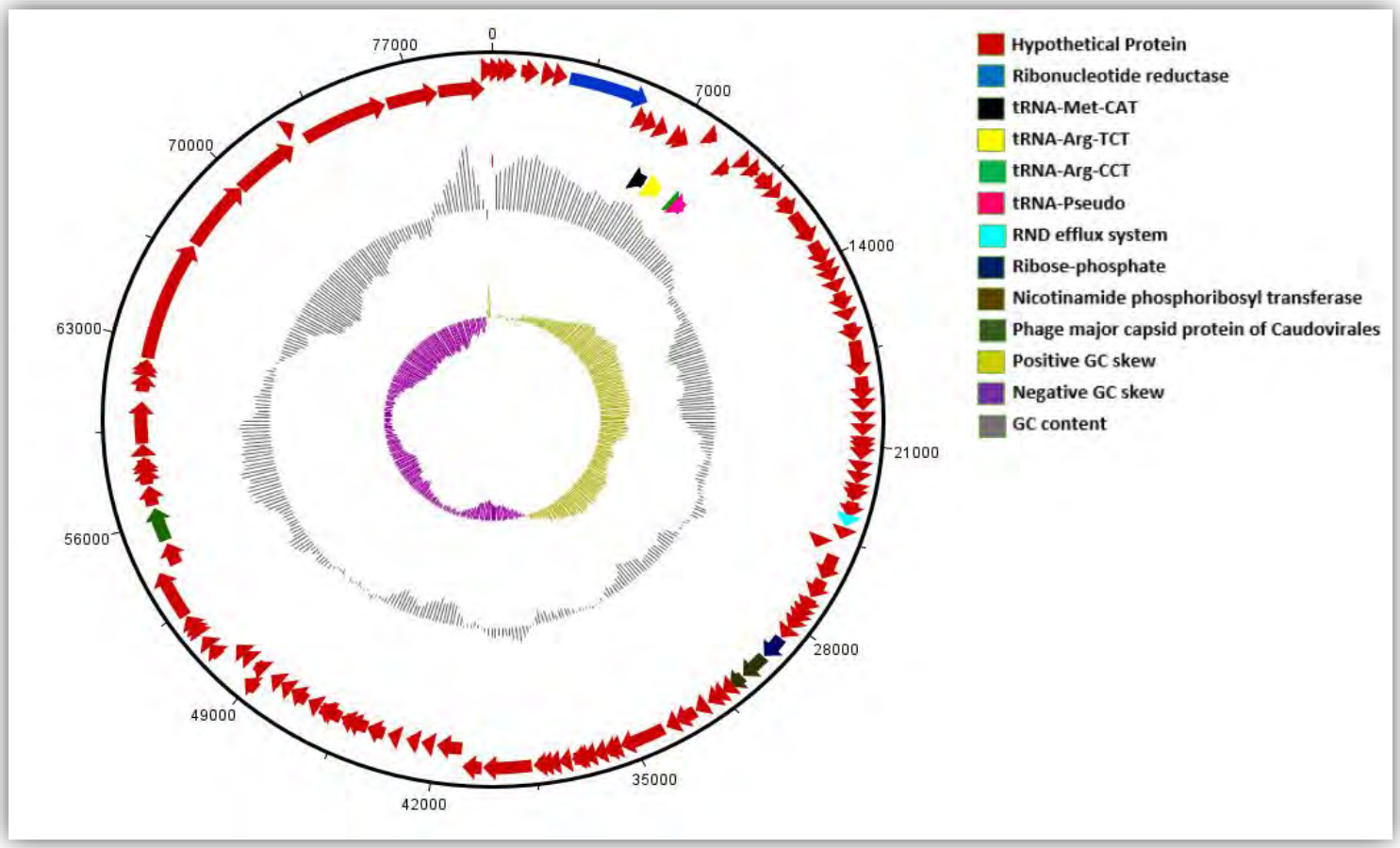

Figure 7: The genome map of bacteriophage

4.3 Hypothetical Protein Analysis

i. From the whole genome (306,455bp):

Rapid Annotation Subsystem Technology (RAST) annotation [Version 2] was used to analyze the whole genomic characterization of bacteriophage (Fig 7). The total length of the genome was 306,455 (bp), contigs were 231 with 45.5% of GC content, the number of coding sequences were 437. The coding density of bacteriophage is about 6% in the subsystem, and a total of 25 open reading frames (ORFs) were predicted. However, most ORFs were annotated as hypothetical proteins (261 in total), while some ORFs were predicted as repeated regions and functional proteins, such as tRNA-Met-CAT, Ribonucleotide reductase of class Ia, and so on.

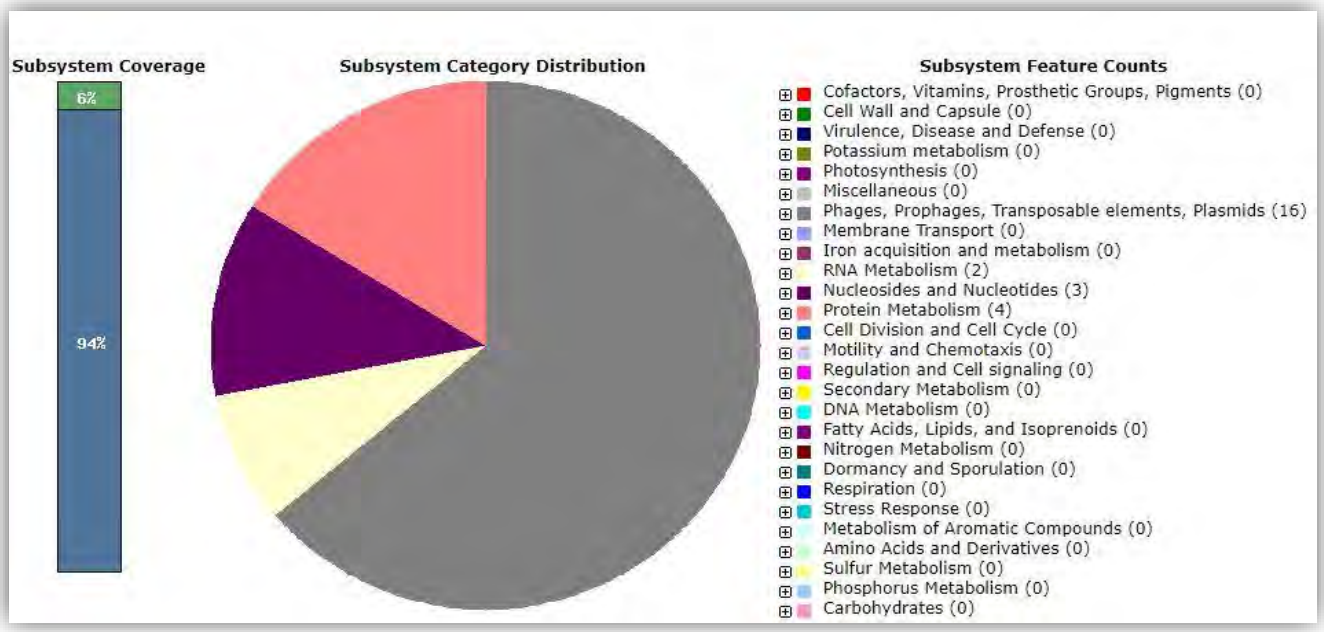

Figure 8: Analysis of Genomic Characterization

ii. From the contig 1 (79,931 bp):

The scaffold number 1 was annotated by RAST annotation which length was 79,931 (bp) with 114 contigs of 46.7% GC content. It contains in total 105 hypothetical proteins out of 114 contigs. For the graphical representation, hypothetical proteins containing upper than 500bp were selected with the start and stop region.

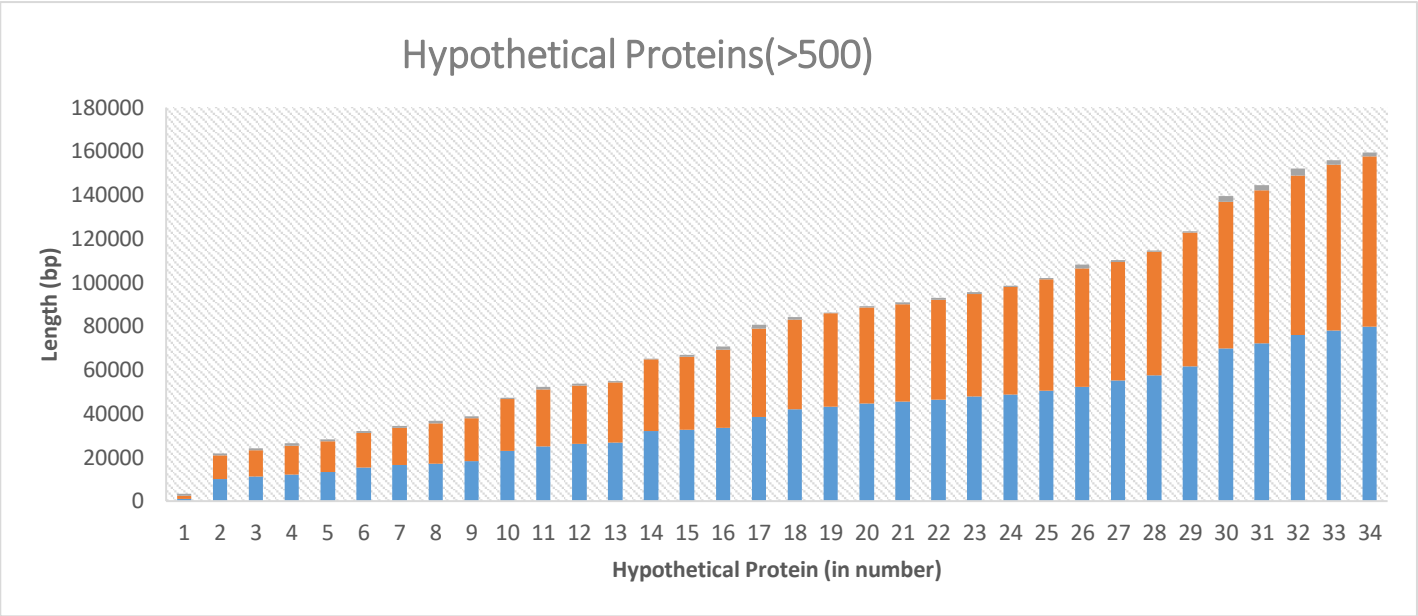

Figure 9: Graphical representation of Hypothetical Proteins **(>500)**

4.4 CRISPRs Analysis

To identify the CRISPR array or cas protein from the DNA sequence, CRISPR finder program online was used here. At first FASTA sequence was uploaded as input, but against this sequence no CRISPR orientation was found [Fig 10].

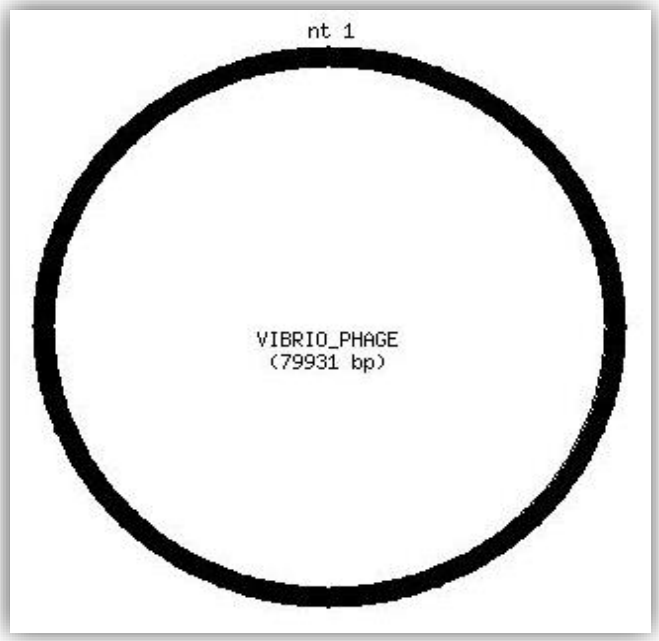

Figure 10: Image developed from CRISPR Finder online

4.5 Phylogenetic Analysis

BLASTn was used to identify homologous and after that phylogenetic analysis was performed to obtain the evolutionary relationship between different types of phages. Figure 10 was generated using BLAST; NCBI database (http://greengene.uml.edu/programs/NCBI_Blast.html).

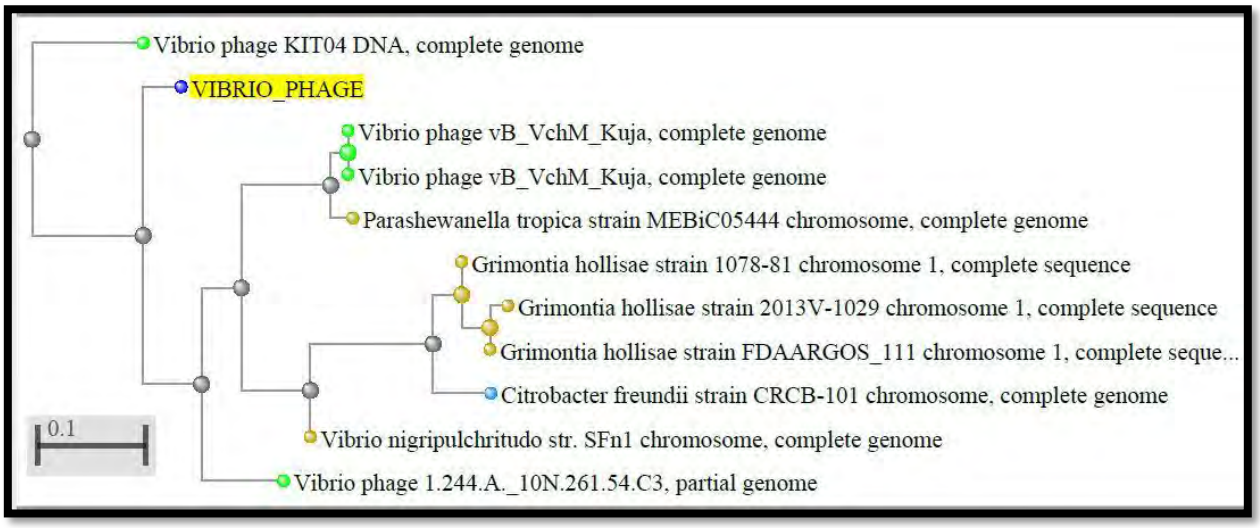

Figure 11: Phylogenetic Tree of the whole genome sequence; the yellow color indicates the main bacteriophage sequence

Chapter 5

[Discussion]

5.1 Bacteriophage Isolation

In this research, bacteriophages specific for *V. cholerae* were isolated from sewage water. These water samples were collected from different areas of Dhaka city. To collect water samples, autoclaved bottles were used to prevent the contamination with other microbes that might be present before in the bottle. Then, 0.22 μm syringe filtration was done to remove bacteria primarily from the water sample. Isolation of bacteriophages was the most crucial part because clear lytic zoned phages were needed for the further process of this study. The existence of bacteriophages were determined by the appearance of a clear plaque in the culture plate. After 24hrs of incubation, clear plaque formation was observed in Luria-Bertani Agar plates. Phage isolation can be affected by season, climate and water sources.

However, for the sampling process, water samples were collected from different places of Dhaka city which cover two lakes (Gulshan & Dhanmondi) and two rivers (Turag & Bosila). Among them, most of the clear zoned lytic phages were found from the Dhanmondi lakes. On the other hand, the study was conducted from January 2019 to August 2019 (Table 2). Moreover, cholera outbreaks occur in Bangladesh during two seasons; typically in spring and autumn season [4]. So, the ideal time for phage isolation is within these two seasons whereas the abundance of phages is less during winter. In this study, most of the phages were isolated in between February to March which means spring season. This also indicates that phages have a seasonal influence within the regular interval of two time periods; this leads to seasonal outbreaks of the epidemic of cholera.

5.2 DNA Isolation and Sequencing

DNA isolation was necessary for the DNA sequencing of this research. During this isolation process, different steps were followed by the protocol [3.4]. Several centrifugation processes were performed very carefully because at first supernatant was collected but at the last pellet was collected to extract the DNA. Later on, after adding phenol-chloroform:isoamyl alcohol, the aqueous phase was carefully removed to ensure full recovery of the DNA. In this process, ice cold ethanol and TE buffer was used to prevent the enzymatic reaction of DNA and to maintain the pH balance of the DNA solution respectively. Finally, to prevent degradation, DNA samples were stored at -20ºC in the freezer.

In order to do whole genome sequencing, the DNA sample was sent to icddr,b on the sanger sequencing platform. They sent the sequenced FASTA file at the middle of January 2020. After that the further bioinformatics analysis was performed.

5.3 The Genomic Reconstruction of Bacteriophage

The complete genome sequence was at first annotated by the RAST server that provided initial assessment of total gene calls, gene function and metabolic reconstruction of the whole genome. The total length of the genome was about 306,755 (bp) along with 231 contigs. Then, this gene was analyzed by BLASTn against the NCBI database to find out similarities between other organisms. In these studies, the sequenced genome was found as a very unique type of phage with 1% similarities with other organisms. Next, for the genomic reconstruction 1st contig was selected and with the DNA plotter software circular image was constructed [Fig 7]. From this structure, different genomic regions and features were identified easily.

5.4 Hypothetical Protein Analysis

It is mentioned before that Rapid Annotation Subsystem Technology (RAST) annotation was used to analyze the genomic characterization of bacteriophage [Fig 8]. The number of coding sequences of bacteriophage was 437 with 45.5% of G+C content. No putative RNAs gene were predicted in the genome, and most of ORFs were annotated as hypothetical proteins, while some ORFs were predicted as repeated regions and functional proteins, such as tRNA-Met-CAT, Ribonucleotide reductase of class Ia, and so on. During genome identification, a large number of hypothetical proteins are predicted by bioinformatics tools. Further research is needed to align hypothetical proteins with known protein sequences to reconstruct 3D structure and to find out the function.

5.5 CRISPR Analysis

CRISPR is a new, simple yet powerful technology that allows genome editing and modification, correction of genetic defects to prevent genetic disease or to improve crops. By maintaining ethical concerns, CRISPR sequence was identified from lower organisms such as bacteria or archaea and then edited genomes; finally transferring into more complex organisms for the purpose of gene editing. CRISPRs contain DNA with short palindromic repeats and spacers (interspaced among the repeated sequences). In this research, a CRISPR sequence was expected but unfortunately nothing was found from the sequence. Here, bacteriophage which previously attacked the *V. cholerae* bacterium; acted as a bank of memories for this bacteria that could be recognized to prevent further attacks. It is interpreted that the whole genome sequencing process might not be performed properly. As a result no CRISPR region was observed [Fig 10] through the online software analysis.

5.6 Phylogenic Analysis

Phylogenetic analysis is the study of the evolutionary development of a species or a group of organisms or a particular characteristic of an organism. Construction of phylogenetic trees is very important which provides in-depth learning about biological diversity, genetic classification or evolutionary manipulation. After getting the whole genome sequence, it was performed in BLASTn against NCBI to find out similarities and an automated phylogenetic tree was constructed through this website [Fig 11]. Although, further the tree was tried to construct by Maximum-likelihood method. Unfortunately, it couldn't be performed because DNA sequencing might not be done perfectly.

References

[1] Centers for Disease Control and Prevention. 2010. Update: cholera outbreak—Haiti, 2010.MMWRMorb. Mortal. Wkly. Rep. 59:1473–1479.

[2] Ackermann, H. W. (2006). Classification of bacteriophages. The bacteriophages, 2, 8-16

[3] Van Belleghem, J. D., Dąbrowska, K., Vaneechoutte, M., Barr, J. J., & Bollyky, P. L. (2018). Interactions between Bacteriophage, Bacteria, and the Mammalian Immune System. Viruses, 11(1)

[4] Liliam K. Harada, Erica C. Silva, Welida F. Campos, Fernando S. Del Fiol, Marta Vila, Krystyna Dąbrowska, Victor N. Krylov, Victor M. Balcão, Biotechnological applications of bacteriophages: State of the art, Microbiological Research, Volumes 212–213, 2018

[5] Md Taufiqul Islam, John D Clemens, Firdausi Qadri, Cholera Control and Prevention in Bangladesh: An Evaluation of the Situation and Solutions, The Journal of Infectious Diseases, Volume 218, Issue suppl_3, 15 November 2018, Pages S171–S172

[6] Griffiths, A. J.F. (2012, July 5). DNA sequencing. Encyclopedia Britannica. https://www.britannica.com/science/DNA-sequencing

[7] Shyamapada Mandal, Manisha Deb Mandal, Nishith Kumar Pal, Cholera: a great global concern, Asian Pacific Journal of Tropical Medicine, Volume 4, Issue 7, 2011, Pages 573-580

[8] Soomro AL, Junejo N. Vibrio cholerae in the environment. J Coll Physicians Surg Pak. 2004 Aug;14(8):509-12. PMID: 15321050.

[9] Waldman, R. and Claeson, . Mariam (2019, August 9). Cholera. Encyclopedia Britannica. https://www.britannica.com/science/cholera

[10] Morozova, V. V., Vlassov, V. V., & Tikunova, N. V. (2018). Applications of bacteriophages in the treatment of localized infections in humans. *Frontiers in microbiology*, *9*, 1696.