A STUDY ON THE EFFECT OF GENOME SIZE ON THE MULTIPLICATION RATE OF BACTERIOPHAGES

By

Tushar Ahmed Shishir
ID: 18376003

A thesis submitted to the Department of Mathematics and Natural Sciences in partial fulfillment of the requirements for the degree of Master of Science in Biotechnology

Department of Mathematics and Natural Sciences
BRAC University
September 2019

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.
5. I would like to request the embargo of my thesis for 12M from the submission date due to publication issue in international journal.

Tushar Ahmed Shishir

Tushar Ahmed Shishir
ID: 18376003

Approval

The thesis "A STUDY ON THE EFFECT OF GENOME SIZE ON THE MULTIPLICATION RATE OF BACTERIOPHAGES" submitted by **Tushar Ahmed Shishir** (18376003) of Summer, 2018 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of Master of Science in Biotechnology on 26th September 2019.

Examining Committee:

Supervisor:
(Member)

_____
Iftekhar Bin Naser
Assistant Professor, MNS Department
BRAC University

Program Coordinator:
(Member)

_____
Iftekhar Bin Naser
Assistant Professor, MNS Department
BRAC University

External Expert Examiner:
(Member)

_____

Departmental Head:
(Chair)

_____
A F M Yusuf Haider
Professor, MNS Department
BRAC University

Ethics Statement

The idea of this thesis was unique and mine therefore there was no conflict of interests. Moreover, no human and animal model was involved in the experiments.

Abstract

Bacteriophages encapsidated by a protein coat have the competency of infecting bacteria, are the most abundant living entities on earth whose genome in general falls into three size ranges and 30-50kbp genetic material containing bacteriophages are most copious encompassing 50% of the whole population whereas bacteriophages containing genome more than 100kbp are only 6%. This uneven distribution of bacteriophages based on genome size points toward the hypothesis, if genome size has any effect on the multiplication rate of the bacteriophages. In response to this proposition, certain tests were put into operation following different parameters to determine their burst-out time, adsorption rate, multiplication rate when both genome size containing phages are co-cultivated, multiplication rate when similar genome size containing phages are co-cultivated, multiplication rate when they are cultivated in absence of other competitors and multiplication rate in presence of other non-specific bacteriophages. The results from these experiments put on show that in spite of genome size, their adsorption rate by hosts is more or less similar but their burst out time varies near around 10 minutes. In addition, their concentration varies three log times since smaller genome containing bacteriophages increases higher in number when co-cultivated however; this variation is negligible the minute when similar genome size containing bacteriophages are cultivated together. This multiplication scheme remains comparable in presence of other non-specific phage as well as when cultured individually. Therefore, this thesis provides an insight that intensification of bacteriophages are greatly sculpted by their genome size, as well as leads to another uncertainty, can we reduce reproduction time of higher animals by chopping out non-coding genome sequences?

Keywords: Bacteriophage, Genome, Multiplication, Replication, *Vibrio cholerae*, Vibrio phage

Dedications

Dedicated to my family and friends for their love, motivations and continuous support…

Acknowledgement

First of all, I would like to express my deepest gratefulness to Almighty Allah for everything he has blessed me with in my whole life and for giving me patience and strength to make this project a success.

I would like to thank Professor A F M Yusuf Haider, Ph.D., Professor and Chairperson of Department of Mathematics and Natural Sciences, BRAC University for allowing me to continue my work at the BRAC University microbiology laboratory.

I am gratified to my supervisor Assistant Professor Dr.Iftekhar Bin Naser, Department of Mathematics and Natural Sciences, BRAC University for believing in me and giving me support, constant encouragement and guidance throughout this project whenever I needed.

I would like to especially thank Professor Dr. Aparna Islam, Department of Mathematics and Natural Sciences, BRAC University for the inspiration and guidance she provided me as a student.

I express my sincere gratitude to Tokee Md. Tarek, Lecturer, Department of Mathematics and Natural Sciences, BRAC University for his constant inspiration, encouragement, care and guidance in my bioinformatics work.

I would like to extend my deepest gratitude and regards to Asma Binte Afzal and Nazrul Islam for their constant guidance, support and wise words throughout the work in the lab.

Moreover, my regards to Ashiqe-E-Khuda, Tanzila Ahmmed Bonna, Shilpi Akter, Nadira Begum, Md. Furkan Mia, Md. Morshed-Al-Mamun for helping me out each time I asked.

I especially thank my friends and classmates Arafat Khan Antu, Nabil, Priyanka, Arthi, Arpon, Sumaiya, Suzana, Naima, Ayesha, Progga and all other labmates for giving me company and helping me throughout my thesis.

Table of Contents

List of Tables

List of Figures

List of Acronyms

AMR	Antimicrobial Resistance
CPS	Capsular Polysaccharide
DLA	Double Layer Agar
DNA	Deoxyribonucleic acid
ds	double stranded
ETT	endotracheal tubes
FDA	Food and Drug Administration
ICDDRB	International Centre for Diarrhoeal Disease Research, Bangladesh
ICTV	International Committee on Taxonomy of Viruses
INSDC	International Nucleotide Sequence Database Consortium
MDR	Multidrug Resistance
ml	milliliter
PFU	plaque forming unit
PLA	Pyogenic Liver Abscess
RNA	Ribonucleic acid
rpm	rotations per minute
ssDNA	single stranded DNA
USA	United States of America
UV	Ultraviolet
μl	microliter
μm	micrometer

Chapter 1: Introduction

Many types of viruses those have the capability of infecting bacteria and well-known to exist in essentially every potential niche where bacteria belong to are termed as Bacteriophages, encapsidated by a protein coat and have genetic materials in the form of either DNA or RNA (Clark & March, 2006). These are one of the most copious entities on earth, are appended with bacterial cell surface using capsid, joined to the tail those are mostly polyhedral in nature (Ackermann, 2003). Bacteriophages infect bacteria and proliferate by either lytic cycle or lysogenic cycle that is killing the hosts for replication or incorporating the genome and being replicated with host for several generations respectively (Siringan, Connerton, Cummings, & Connerton, 2014). “Kill-the-winner” hypothesis best harmonizes bacteriophage activity, enlightens microbial diversity and changes in the environment and act as the reservoirs of significant uncharacterized genetic diversity on Earth (Koskella & Brockhurst, 2014). Since the discovery of bacteriophages in early 20th century, they are used in the many fields of biotechnology and molecular biology. Moreover, complete phage genomes simplify the studies of evolutionary antiquity and relationships, biodiversity, and biogeography. Even though their importance and omnipresent abundance, so far too little is known about their diversity in natural ecosystems (Clark & March, 2004). Therefore, in recent time many studies are focused on the isolation and characterization of bacteriophages to study their potentiality towards many field of studies. One of these studies, isolation and characterization bacteriophages since 2001 to 2017 has revealed that small genome containing bacteriophages are more abundant in environment than large genome containing phages and the ratio is almost 3:1. This result triggers a hypothesis, if the genome size has any effect on the replication rate of the bacteriophages (Naser et al., 2017).

Bacteriophage genome size show a discrepancy enormously, ranging Leuconostoc phage L5 (2,435bp) to the almost 500 kbp genome of Bacillus megaterium phage G (Hatfull & Hendrix, 2011). However, there is no uniform distribution of genome sizes and there is a prevalence of genome sizes in three separate ranges. First of all, the utmost is a peak of genomes in the 30−50kbp interval, corresponding to nearly 50% of all phages (Hatfull, 2008). Then the second group about 20% of total is those smaller than 10kbp, and the third group contains those in the 100−200 kbp interval containing 6% of the total (Hatfull, 2008; Hatfull & Hendrix, 2011). These genomes are packaged at parallel compactness into their capsids therefore, the size of the capsid differs as a function of genome size and effect virion infectivity caused by the amount of DNA packaged within any given capsid since either scanty or excessive genetic material leads to loss of virion stability (Orlova, 2012). Since the density of genetic materials are equivalent for all bacteriophages and the capsid size varies, therefore more time might be required for the formation of larger capsid for large genome containing bacteriophages during replication compared with small genome containing bacteriophages.

In this thesis, I tried to study the effect of genome size of bacteriophages on replication rate in different parameters including their growth pattern when co-cultivated, cultivation in absence of other bacteriophages, cultivation in presence of other non-specific bacteriophages, their burst out time as well as absorption rate. However, the outcome has triggered further question, what if we reduce genome size?

Objective:

The main aim of this project was to demonstrate the effect of genome on the replication rate of bacteriophages.

Specific Aims:

1. Observation of growth rate difference between large and small genome containing bacteriophages when co-cultivated.
2. Observation of growth rate difference between large genome containing bacteriophages when co-cultivated.
3. Observation of growth rate difference between small genome containing bacteriophages when co-cultivated.
4. Observation of growth rate difference between large and small genome containing bacteriophages when cultivated individually.
5. Observation of growth rate difference between large and small genome containing bacteriophages when co-cultivated in presence of another non-specific bacteriophages.
6. Observation of absorption rate of large and small genome containing bacteriophages when co-cultivated.
7. Determination of burst out time of large and small genome containing bacteriophages.

Chapter 2: Background Information

2.1 Bacteriophages:

Bacteriophages are naturally occurring bacterial viruses which infect bacterial cells with higher specificity and have the ability to proliferate inside bacterial cell (Abedon, 2012; Clark and March, 2006; Hagens and Loessner, 2007; Hanlon, 2007; Nishikawa et al., 2008; Viazis et al., 2011). They are deemed as natural killer of bacteria due to their capability to infect and lyse the host organism (Abuladze et al., 2008; Nishikawa et al., 2008). In recent times, it has been far and wide accepted that bacteriophages are highly abundant in the atmosphere and influence the biosphere extensively and estimated to kill between 20-40 % of oceanic bacteria every day. They are found to play a key role in nutrient and energy cycle of the ecosystem and forms the pool of most genetically diverse 'life form' on earth (Suttle, 2005).

2.1.1 History of bacteriophages:

A British bacteriologist, Ernest Hankin first witnessed the bacteriophages in 1896 in the water of Ganga and Jumna River in India, presence of an antibacterial activity against bacteria *Vibrio Cholerae* (Ackermann, 2012). He suggested that this unknown agent, which was heat sensitive and could pass through porcelain filter, producing the bactericidal activity is responsible for preventing the spread of cholera disease. Years later, while working with *Bacillus subtilis,* a Russian bacteriologist witnessed similar phenomenon and in 1901 Emmerich and Löw informed that sample from a culture which demonstrated autolysis was able to lyse different culture, was capable of curing experimentally induced infection (Summers, 2004; Sulakvelidze et al., 2001). Later on, Frederick William Twort, a British pathologist observed a "glassy transformation" of Micrococcus colonies grown on solid agar media. He hypothesized that the unknown substance causing the watery transformation of the bacterial colonies could be a virus. Two years later of Twort's documentation, Felix d'Herelle a French Canadian

microbiologist observed similar kind of incidence. He proposed that it was “ultravirus” that was causing lysis of bacterial cell in liquid media and created clear patches on the bacterial lawn which he primarily called it taches, then taches vierges, and later plaques. Felix d’Herelle also named the virus responsible for these phenomena as “bacteriophage” which derived from “bacteria” and Greek word “phagein” which means to “eat” or “devour”. Felix d’Herelle concluded bacteriophage as “exogenous agents of immunity” following the observation of an increase in phage titer in the stool sample of recovering patients suffering from dysentery and typhoid (Deresinski, 2009). He also conducted trials of therapeutic use of phages at the Hospital Des Enfants-Malades in Paris in 1919 by administering anti-dysentery phage preparation to a 12-year boy with severe dysentery and observed consecutive cease of symptoms and full recovery within a few days (Sulakvelidze et al., 2001; Summers, 1999). However, the discovery of broad range antibiotics played the major role in declining the interest of producing phage commercially. The lack of understanding of phage biology and inadequacies in the diagnostic bacteriology techniques available at the time aided the shift in interest from phage in the western world. On the other hand, phage study 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. Phage resultant 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).

2.1.2 Classification of bacteriophages:

Phages are enormously diverse and vary from one another based on structural, physicochemical, and biological properties. In 1933, Burnet showed heterogeneity among enterobacterial phages and in 1943, Ruska observed three morphological types of

bacteriophage, which evoked the necessity of proper classification of phages. The International Committee on Taxonomy of Viruses (ICTV) classified phages based on nucleic acid and gross morphology and grouped them into six genera (Ackermann, 2004).

Table 2.1: Overview of phage families (Ackermann, 2007)

Shape	Nucleic acid	Virus group	Particulars	Example
Tailed	DNA	Myoviridae	tail contractile	T4
		Siphoviridae	tail long, noncontractile	λ
		Podoviridae	tail short	T7
Polyhedral	DNA, 1, C	Microviridae	conspicuous capsomers	φX174
	2, C, S	Corticoviridae	complex capsids, lipids	PM2
	2, L	Tectiviridae	inner lipid vesicle, pseudotail	PRD1
	2, L	SHI, group*	inner lipid vesicle	SH1
	2, C	STV1 group*	turret-shaped protrusion	STIV
	RNA, 1, L	Leviviridae	poliovirus-like	MS2
	2, L, seg	Cystoviridae	envelope, lipids	Φ6
Filamentous	DNA, 1, C	Inoviridae	a. long filaments	fd
			b. short rods	MVL1
	2, L	Lipothrixviridae	envelope, lipids	TTV1
	2, L	Rudiviridae	TMV-like	SIRV-1
Pleomorphic	DNA, 2, C, S	Plasmaviridae	envelope, lipids, no capsid	L2
	2, C, S	Fuselloviridae	same, lemon-shaped	SSV1
	2, L, S	Salterprovirus	same, lemon-shaped	His1
	2, C, S	Guttaviridae	droplet-shaped	SNDV

	2, L	Ampullaviridae∗	bottle-shaped	ABV
	2, C	Bicaudaviridae∗	two-tailed, growth cycle	ATV
	2, L	Globuloviridae∗	paramyxovirus-like	PSV

C Circular; L linear; S superhelical; seg segmented; 1 single-stranded; 2 double-stranded

The current classification of bacteriophage by the International Committee on Taxonomy of Viruses (ICTV) contains 1 order, 14 families, 37 genera (Ackermann, 2009). Among them, over 96 % of all phages are tailed double-stranded (ds) DNA phage and belong to the order Caudovirales and further classified into three main large families, Siphoviridae, Myoviridae, and Podoviridae, differentiated by their tail length and contractile ability (Ackermann, 2004). Having long flexible tail, 61 % of the phage under Caudovirales falls in Siphoviridae; 25 % are Myoviridae having double-layered contractile tails; and with short noncontractile tails, 15 % are Podoviridae. Other types of phages, Polyhedral, filamentous, and pleomorphic phages comprising less than 4 % of observed phages (Ackermann, 2007).

2.1.3 Bacteriophage abundance in the environment:

Bacteriophages are considered to be the most prominent biological entities with an estimated population size of 1030 or more (Chibani-Chennoufi et al., 2004). These phages are present in different environmental setting such as acidic hot springs (higher than 80°C with pH=3.0), solar salterns (10 times saltier than the ocean), alkaline lakes (pH=10), in the terrestrial subsurface (greater than 2000 m deep), below 30 m of ice in polar lakes (Breitbart and Rohwer, 2005), from soil (Ashelford et al., 2003), sewage sludge (Carey-Smith et al., 2006) and mammalian feces (O'Flynn et al., 2004). Marine water is one of the major resources of bacteriophage and contains a greater variance in overall phage prevalence in these ecosystems and roughly, 70 % of aquatic bacteria are infected by those phages (Ackermann et al., 2012). Phage abundance across aquatic system varies between less than 104 m1-1 to more than 108 ml-1 and this

variation in number is generally correlated with the variation in associated host organism, which ultimately depends on the productivity of the system (Wommack and Colwell, 2000). However, it has also been demonstrated that the phage abundance decreases along the movement from coastal to offshore (Weinbauer, 2004). On the other hand, bacteriophage prevalence in freshwater is higher than in marine water and in sea ice, phage abundance was reported 10 to 100 times higher than in surrounding water (Maranger et al., 1994). Yet, the phage abundance variation also exists in the fresh water system and concentration decrease with sediment depth (Weinbauer, 2004).

2.1.4 Bacteriophage impact on bacterial population:

Studies on marine biodiversity have shown that bacteriophages influence their host bacterial organism in a density-dependent manner, targeting and infecting a few bacterial species at any one time (Ventura et al., 2011). This is in harmony with "kill-the-winner" model where the predation is directed towards "winner" (abundant) bacterial population in that environment (Rodriguez-Valera et al., 2009). As a result, the nutritional resources are more accessible to other bacterial species and provides the opportunity for a new bacterial species to become abundant. Studies involving genome size distribution have demonstrated that over a period of time specific virus become abundant, then reduce to an undetectable level and then again become abundant (Wommack et al., 1999). Bacteriophages also play a key role in bacterial population inside animal intestinal tract. A Study involving horse feces have shown that diversity and abundance of E.coli strains in horse gut are directly correlated to the relative abundance of specific coliphages (Golomidova et al., 2007). Moreover, the presence of prophage in bacteria residing in gut flora provides a competitive advantage to the host, which makes the pathogenic organism hard to outcompete commensal organism hence, maintain the stability of human gut microbiome (Ventura et al., 2011).

2.2 Genomics of bacteriophage:

Bacteriophage give a special aspect on the origin, diversity and evolution of virus depending on their tremendous abundance and distant of origins. Enormous bacteriophages and their diverseness engage in horizontal genetic material exchange results in continual evolution with time (Keen et al., 2017).

2.2.1 Bacteriophage genome:

This is somewhat surprising that having high diversity of the phage population and its abundance, they are only approximately 1% the size the bacterial chromosomes in the sequenced chromosome list in Genebank. However, there are a large number of resident prophages in sequenced bacterial genomes (Fouts, 2006; Lima-Mendez, Van Helden, Toussaint, & Leplae, 2008), therefore it is expected a greater shift towards phage genome sequences and will outnumber sequenced phage genomes, although the proportion of these retaining potential for lytic growth is unclear. The sequenced phage genomes vary considerably in size from Leuconostoc phage L5 (2,435bp) to Pseudomonas phage 201phi2−1 (316,674b) but there is no uniform distribution of genome sizes (Erkus et al., 2013). These genome sizes are predominant in three separate ranges. The largest is a peak contains 50% of whole population having genome size in the 30−50kbp interval, whereas a second group containing 20% of total population is those smaller than 10kbp, and the third group contains those in the 100−200 kbp interval, which are 6% of the total phages (Hatfull, 2008; Hatfull & Hendrix, 2011). This overall distribution influence the phage isolation methods and sequencing technologies but is not a true reflection of size distributions of phages in the environment since smaller genome containing phages were characterized prior to automated sequencing methodologies but larger phage genomes are not readily detected due to lack of technologies as well as due to formation of very tiny plaques (Serwer, Hayes, Thomas, & Hardies, 2007).

Therefore, the use of alternative microbiological substrates and electron microscopy suggests there are many large phages in the environment that have yet to be explored.

2.2.1 Genetic diversity:

The genetic diversity of the bacteriophage population is notable since the nucleotide sequences of genomes derived from phages with non-overlapping host ranges rarely share sequence similarity. On the other hand, bacteriophages infecting a common bacterial host are in genetic contact with each other and they sometimes share common nucleotide sequences (Nale et al., 2012). In general, phages with different virion morphotypes have different genome organizations and greater sequence diversity; genome architecture may therefore impose constraints on genetic exchange (Hatfull et al., 2006). It is noteworthy that the high genetic diversity among phages of a common host coupled with the still limited number of available sequences suggests there is an abundance of new viral genome sequences yet unidentified.

2.2.3 Bacteriophage genome mosaicism:

One of the most striking features of bacteriophage genomes is their apparent mosaic structure where each genome can be considered as a unique combination of modules that are exchangeable among the population. The size of the modules, their rates of exchange, and the phage genomes carrying them all vary greatly, with phages of different virion morphology, size, and host-range all participants in an orgy of recombination as a result this mosaicism is by no means unique to the phage population. Phage genome mosaicism can be viewed at two levels including DNA heteroduplex mapping and subsequent sequence comparison (Juhala et al., 2000). At the resolution of DNA sequence information, precise junctions can be observed corresponding to the boundaries of two DNA segments that clearly have distinct evolutionary origins and the most notable features of these functions is that they predominantly correspond to the boundaries of open reading frames (Mavrich & Hatfull, 2017). However, most of the

progeny would not be viable, except for those that have an appropriate genome size and that retain gene functions, thus accounting for the correlation with gene borders (Klyczek et al., 2017). As a result, productive genetic exchange may require multiple recombination events, and while the overall process is expected to occur at low frequency, this process is potentially highly creative, yielding new combinations of genes, as well as new combinations of protein domains. On the other hand, comparing the predicted amino acid sequences of phage gene products generates an alternative manifestation of mosaicism (Liu, Glazko, & Mushegian, 2006; Hatfull et al., 2006; Pedulla et al., 2003) where many groups of phages – including those that infect common hosts – may not share any nucleotide sequence information; protein sequence data reveals genes that share much older ancestry. Sequence comparison and phylogenetic reconstruction shows those different genes, groups of genes, or segments of genes, all share different ancestry, and thus represent modules within a mosaic genome. The extent of phage genomic mosaicism thwarts the simple determination of phylogenetic relationships of whole phage genomes as units of evolution, and their histories can best be considered as the totality of the evolutionary routes taken by each constituent module (Casjens & Thuman-Commike, 2011).

2.3 Bacteriophage replication:

Bacteriophages most commonly replicate through two major phage life cycles such as the lytic and lysogenic though they have multiple possible life cycles, which determine their role in bacterial, or archaeal biology. Both the lytic and lysogenic life cycle includes 2 common steps: i) absorption of phage i) penetration of genetic material (Salmond and Fineran, 2015). Phage at first interact with the receptors expressed on the surface of the bacterial and get absorbed followed by the injection of its genome into the bacterial cell involving mechanism specific for each phage (Guttman et al., 2004).

The replication strategies following the successful penetration of phage genetic material dictate the virulent (lytic) or temperate nature of bacteriophage. In lytic cycle, injected phage genome controls of the bacterial replication machinery and produce necessary components for new progeny phage. When all the essential components are manufactured, the phage particles are assembled into infective virions, followed by lysis of the host cell, where the new progeny phages are liberated from the bacterial cell via disruption of the cell wall and cell membrane. The lytic life cycle results in the destruction of the host cell (Guttman et al., 2004).

On the other hand, lysogenic cycle is another phage proliferation technique where new phage particles are not immediately produced and the host cell is not destroyed. Rather, in this process, phage genome either is integrated into the host genome or prevail as a plasmid within the cell. In lysogenic relationship, the phage gene product termed a repressor inhibits lytic genes. Due to lytic inhibition, the phage gene will remain integrated into the host chromosome, which is called a prophage, and will replicate along with the host replication process (Little, 2005).

However, lysogenic phase is stable it can be switched to initiate the lytic cycle. Temperate phages such as λ phages are able to proliferate via both lytic and lysogenic cycle (Little, 2005). Various physical and chemical agents such as poly-aromatic hydrocarbons, mitomycin C, hydrogen peroxide, temperature, pressure and UV radiation (Williamson et al., 2001) can induce this introduction of lytic cycle in prophage. In order for virulent phages to replicate and survive in the environment, the rate of phage-host encounter required to exceed the virus decay and inactivation rate. However, the temperate phages are not dependent on the host cell density, rather requires a small number of lysogenic carrier cell and occasional induction of lytic cycle and release of free phages (Wommack and Colwell, 2000). Regeneration of progeny phages requires time ranging from around 30 minutes to 55 minutes depending on their complexity.

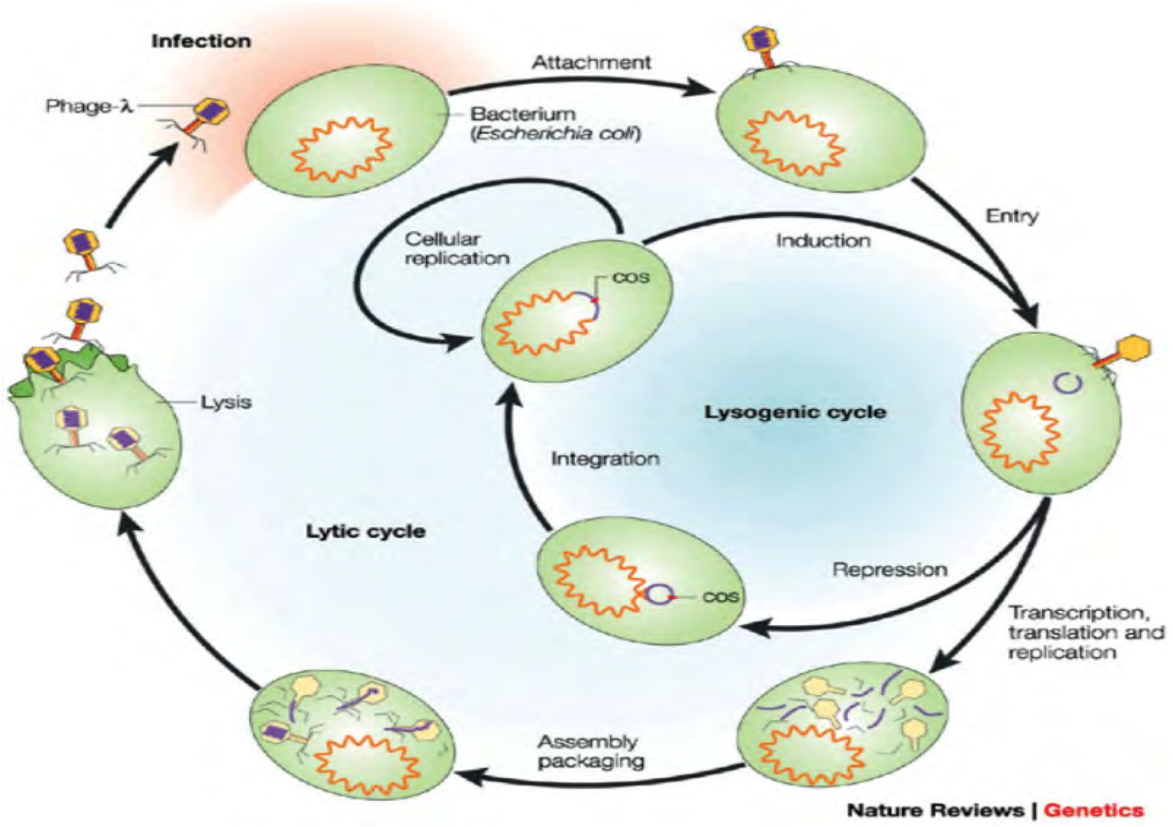

Figure 2.1: Lytic and lysogenic life cycles of bacteriophage (Weigel & Seitz, 2006)

2.4 Bioinformatics' Data:

Prior to carrying out experiments in the laboratory, real world data have been explored consisting of GeneBank reference phage sequences and phage metagenomics.

2.4.1 GeneBank reference phages data:

During completion of this thesis, until now 9575, complete genome sequences of bacteriophages have been submitted to GeneBank database and analysis revealed that almost 55% phages contain genome size shorter than 50 kbp and 68.2% phages contain genome size shorter than 60 kbp.

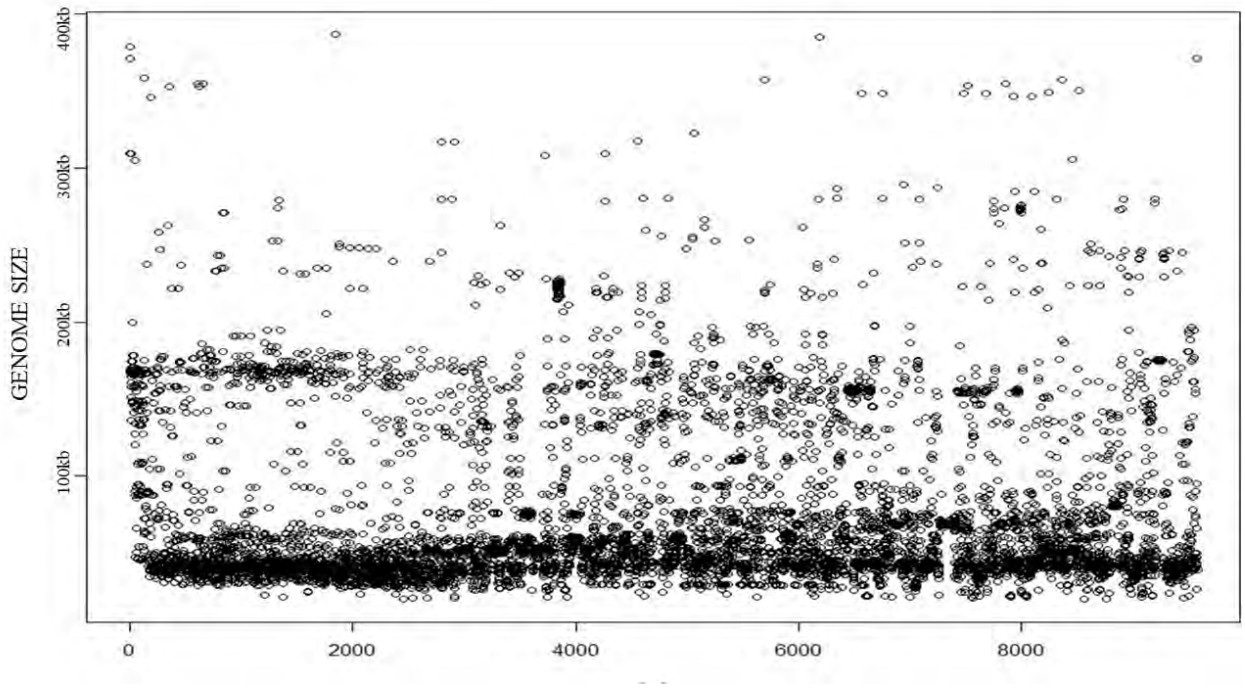

Figure 2.2: Distribution of Genebank bacteriophages according to size

Using command line, 9575 phage genomes were shuffled and their genome size distribution occurred as follows taking randomly selected 150 phages.

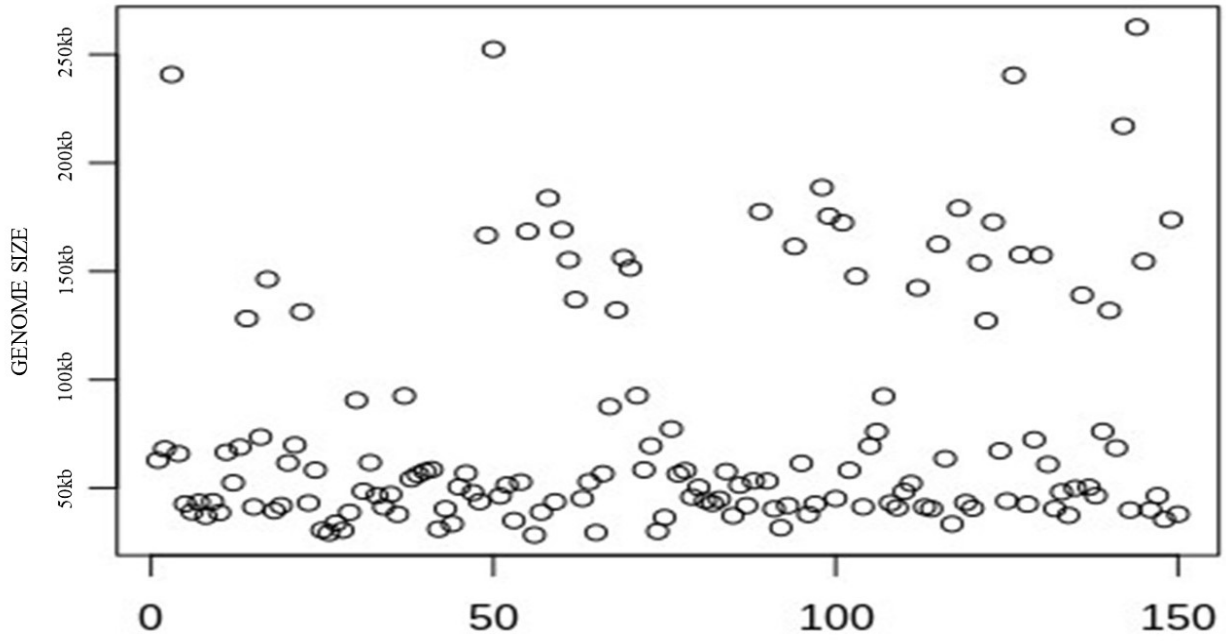

Figure 2.3: Distribution of randomly selected 150 NCBI bacteriophages according to size

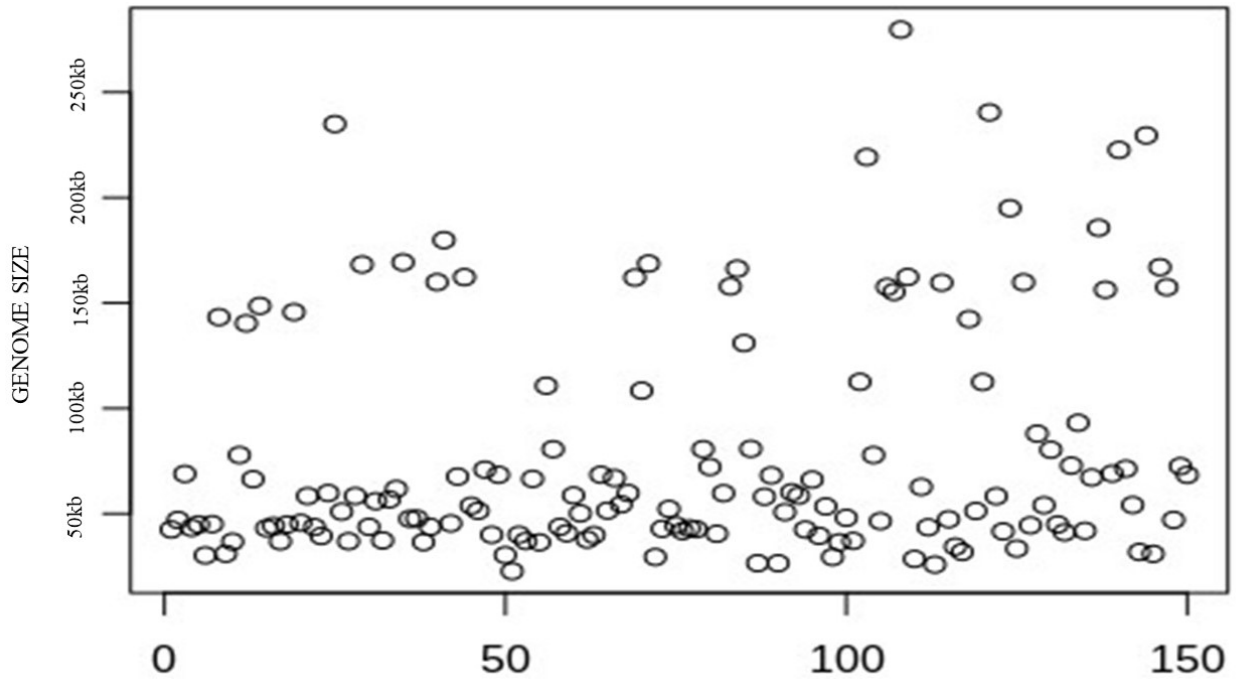

Figure 2.4: Distribution of randomly selected 150 NCBI bacteriophages according to size

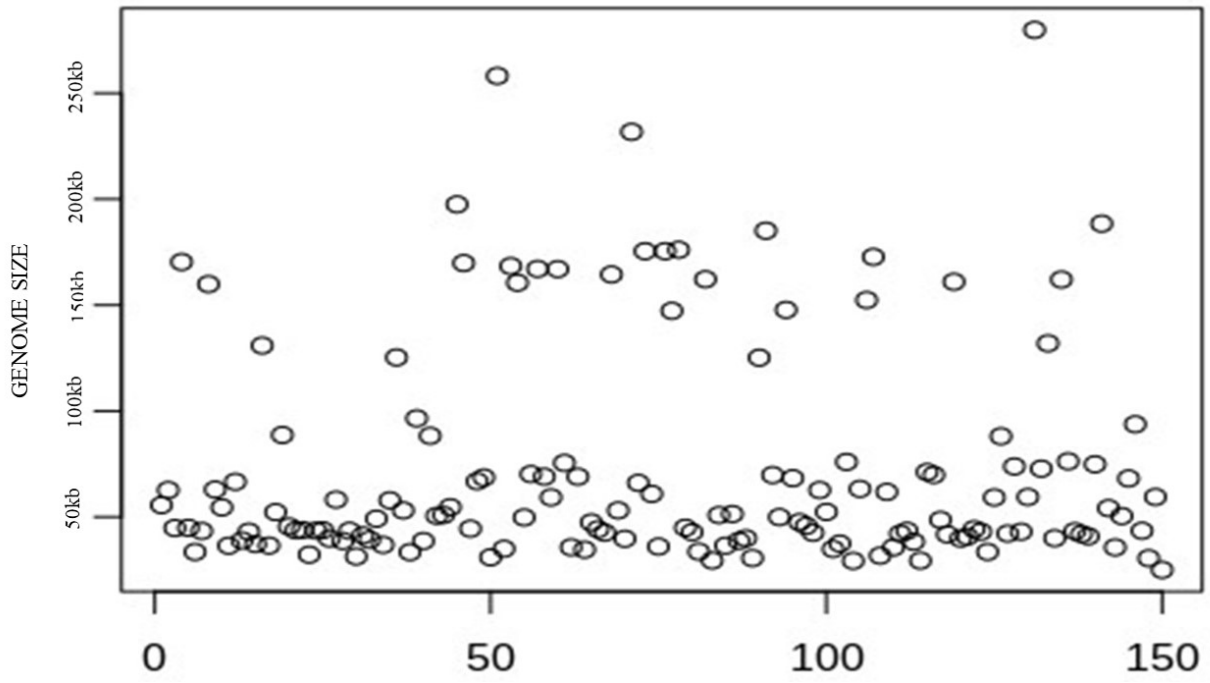

Figure 2.5: Distribution of randomly selected 150 NCBI bacteriophages according to size

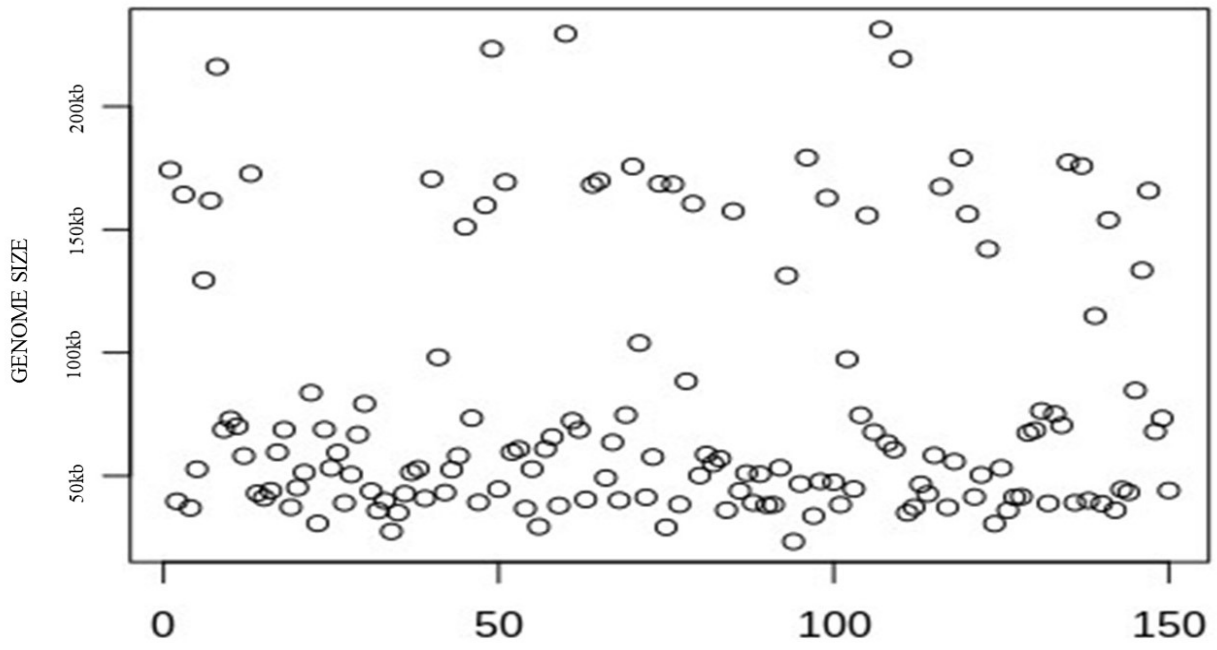

Figure 2.6: Distribution of randomly selected 150 NCBI bacteriophages according to size

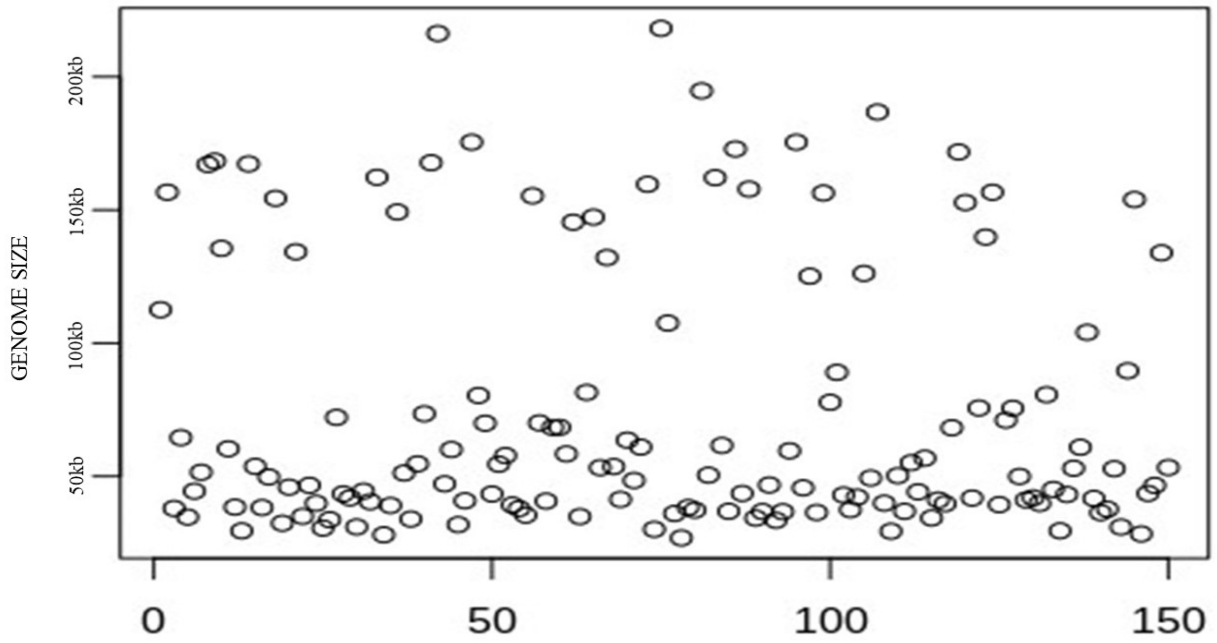

Figure 2.7: Distribution of randomly selected 150 NCBI bacteriophages according to size

From the graphs generated R program by randomly selected phages, we can see larger amount of phages have genome size smaller or close to 50 kbp.

2.4.2 Metagenome data:

Historically, the study of microbes has focused on single species in pure culture, so understanding of these complex communities lags behind understanding of their individual members. On the other hand, though the science of metagenomics is only a few years old, will make it possible to investigate microbes in their natural environments, the complex communities in which they normally live. Metagenomics is the field of research that circumvents the unculturability and genomic diversity of most microbes, the biggest roadblocks to advances in clinical and environmental microbiology depending on computational methods that maximize understanding of the genetic composition and activities of communities so complex that they can only be sampled, never completely characterized. Metagenomics includes cultivation-independent genome-level characterization of communities or their members, high-throughput gene-level studies of communities with methods borrowed from genomics, and other "omics" studies which are aimed at understanding transorganismal behaviors and the biosphere at the genomic level (Haynes, 2013). We know bacteriophages are ubiquitous and numerous parasites of bacteria and play a critical evolutionary role in virtually every ecosystem, yet our understanding of the extent of the diversity and role of phages remains inadequate for many ecological niches, particularly in cases in which the host is unculturable. During the past years, the emergence of the field of viral metagenomics has drastically enhanced our ability to analyse the so-called viral 'dark matter' of the biosphere. Here in this thesis, reconstructing 19,673 bacteriophage sequences and their size distribution revealed that smaller genome containing bacteriophages are highly abundant in the environment. 52.76% genomes were smaller than 40 kb whereas 72.09% genomes were smaller than 50 kbp.

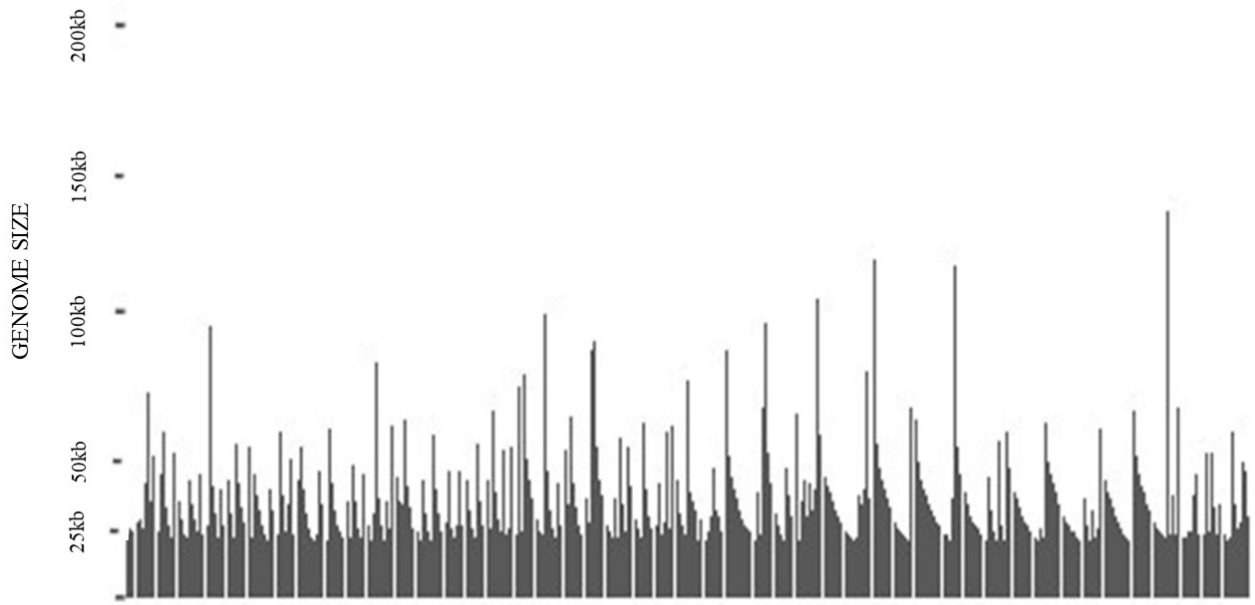

Figure 2.8: Distribution of metagenome bacteriophages according to size

Using command line, metagenomic phage sequences were shuffled and their genome size distribution occurred as follows taking randomly selected 150 phages.

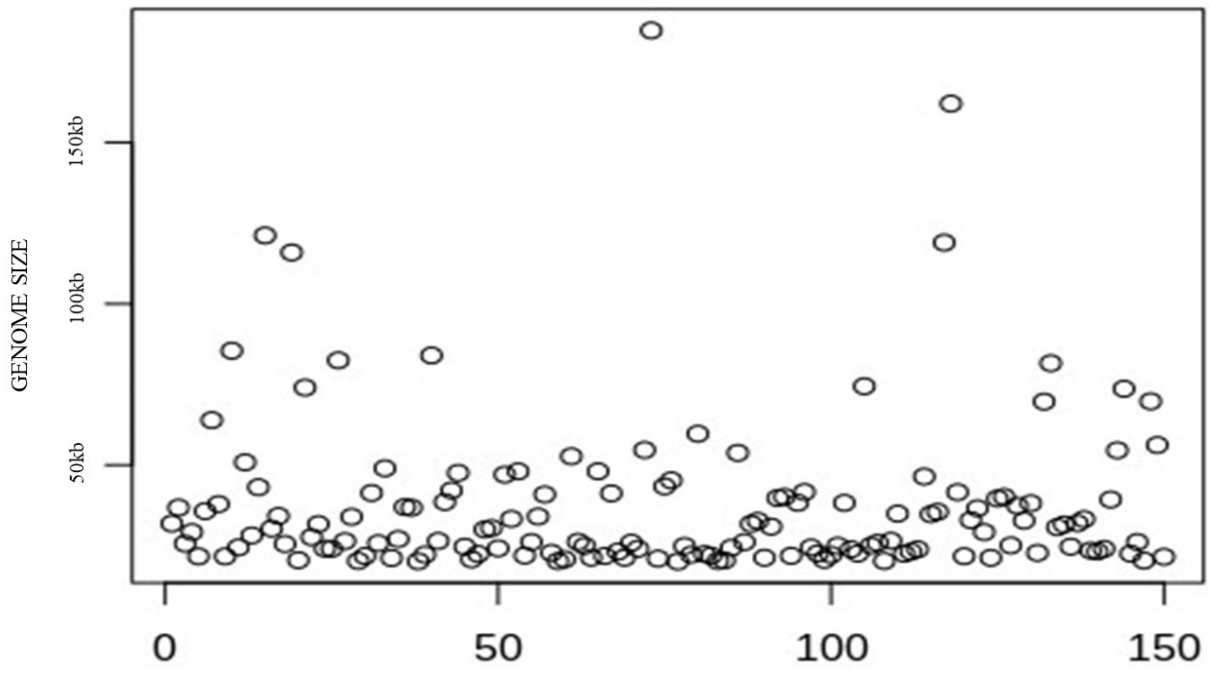

Figure 2.9: Distribution of randomly selected 150 metagenome bacteriophages

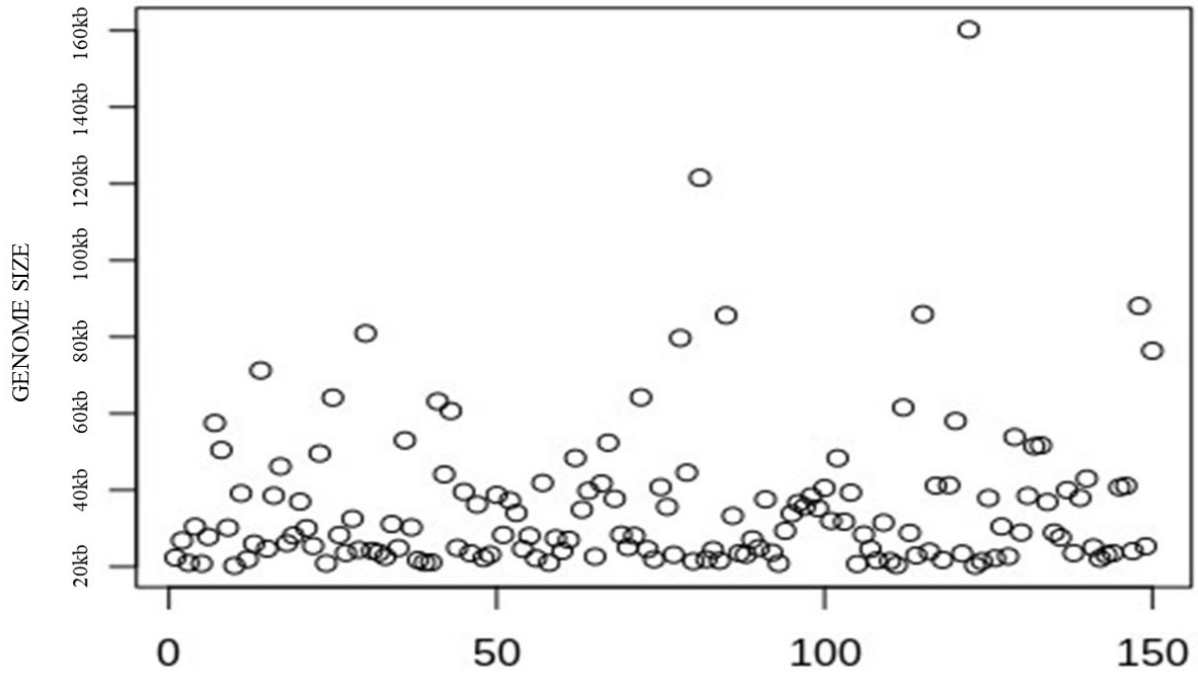

Figure 2.10: Distribution of randomly selected 150 metagenome bacteriophages

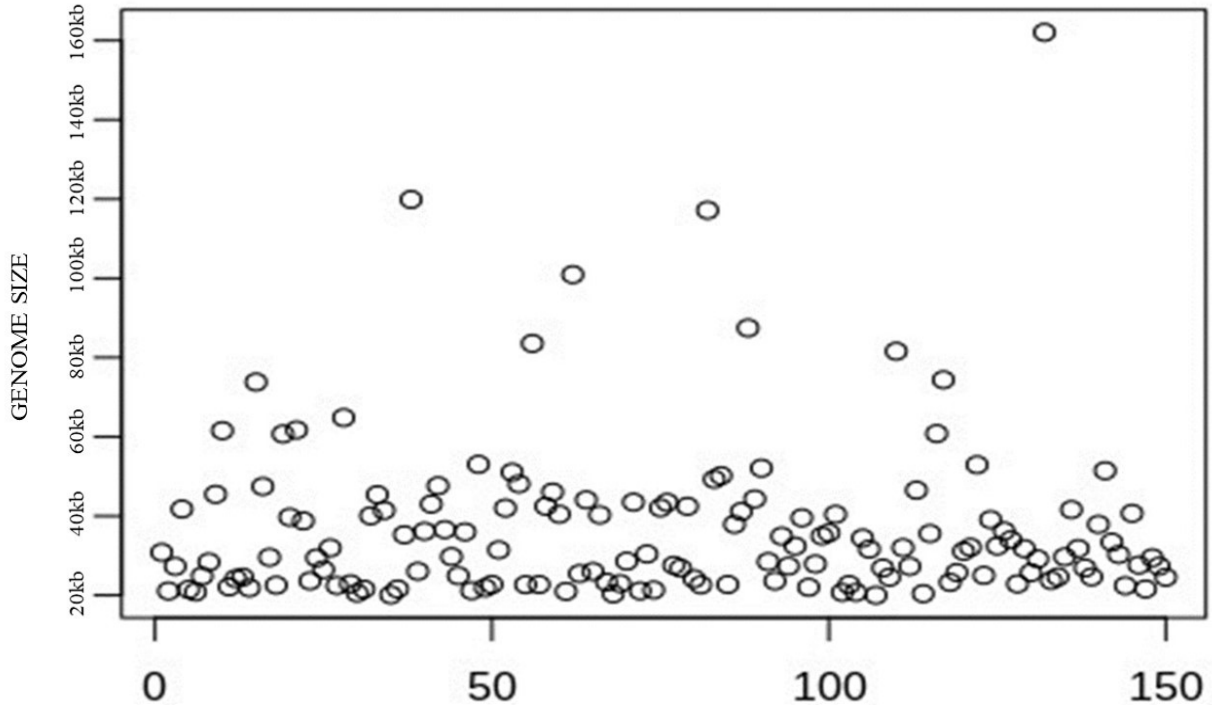

Figure 2.11: Distribution of randomly selected 150 metagenome bacteriophages

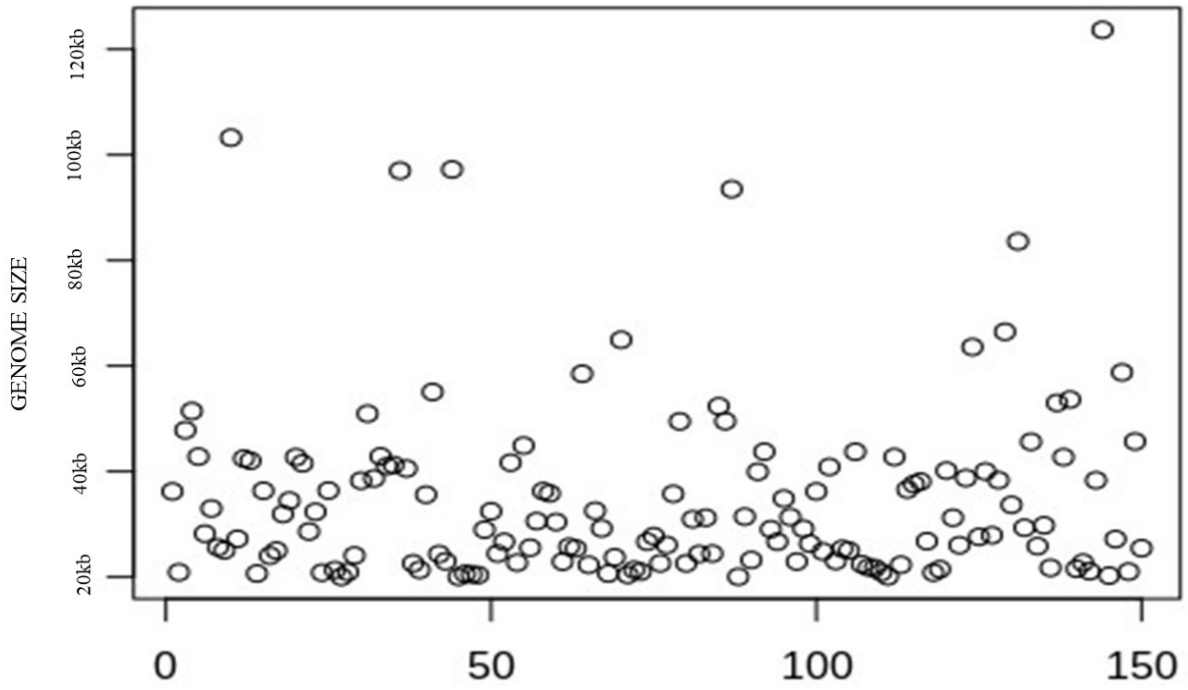

Figure 2.12: Distribution of randomly selected 150 metagenome bacteriophages

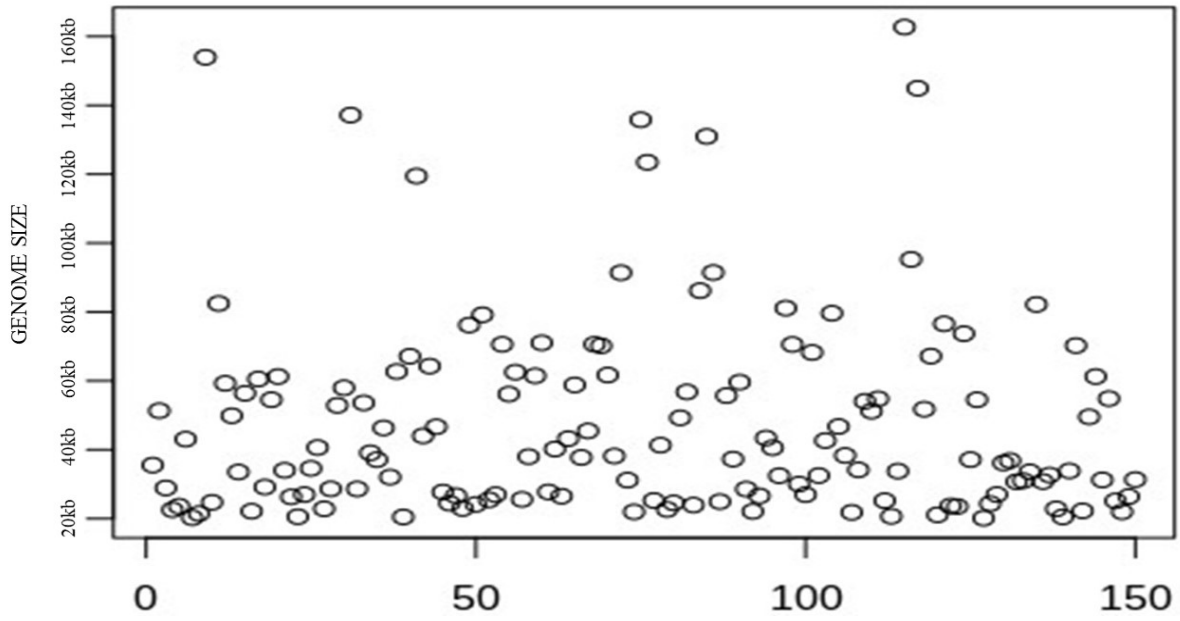

Figure 2.13: Distribution of randomly selected 150 metagenome bacteriophages

These results from Genebank and Metagenome data reveals that smaller genome containing bacteriophages are highly abundant in the environment which triggers the hypothesis of this thesis, does genome size has any effect on the replication rate of bacteriophages?

Chapter 3: Materials and Methods

The very first approach of conducting research is designing methodology and identifying the required materials. This research was carried out in Environmental Life Science Laboratory in BRAC University following different approaches to observe the variation of bacteriophages' multiplication rate based on their genomic size in different settings with the use of different materials. Quite a few experiments include experimenting multiplication competition rate of bacteriophages between different genomic size, similar genomic size of both small and large bacteriophages, individual multiplication, their adsorption rate, multiplication rate in presence of external bacteriophages and multiplication rate in microcosm.

3.1 Materials Required:

Table 3.1: Materials used

Materials	Justification
Bacteriological Agar	Culture plate preparation
Luria Broth	Growth medium
Sodium Chloride	Dilution medium
TCBS agar	Selective media for *Vibrio cholerae*

LB and agar were mixed in different ratio to prepare solid medium and soft medium for plate preparation and lawn culturing respectively. 1.5% agar was mixed with LB in solid medium whereas 0.4% agar was used for soft agar medium preparation. Components of the agar and LB are mentioned in appendix. Last of all, 0.22 µm syringe filter was used to separate bacteriophages from the bacteria in the medium.

3.2 Methodology

3.2.1 Bacteriophage typing:

Suitable bacteriophages and their hosts were identified in the very beginning to design the research. Sixteen O1 *Vibrio cholerae* strains were tried out against twenty-two bacteriophages to sort out hosts and bacteriophages those would serve the purposes of this research. All the bacteria and bacteriophages used were reference strains, which are mentioned in the appendix section. From results, three sets of bacteriophages were chosen based on their genomic size, one set contains two bacteriophages of large and small genomic size, another set contains two bacteriophages of large genomic size and last set contains two bacteriophages of small genomic size. On the other hand, host bacteria were selected based on criteria, those could be infected by the selected bacteriophages and those could not be affected by the bacteriophages to provide selection advantages as well as count. First of all, all the O1 *Vibrio cholerae* strains were grown on LA (Luria agar) plate by streaking method. From each plate, one single colony was taken in separate vials containing LB ((Luria broth) medium where those single colonies were cultured for 3 hours to obtain the young bacterial culture. 300 µl LB medium containing young bacterial cultures were then mixed with fresh soft agar (0.4%) and poured onto LA plates followed by drying for 15 minutes to get the soft agar attached on the plate properly. All the plates were then marked alike and every single bacteriophages of same volume (5 µl) were put onto the plates containing bacterial culture in soft agar. The plates were at that point dried for 30 minutes to get the bacteriophages absorbed and plates were incubated overnight at 37 ^{o}C. Plaques caused by bacteriophages were then looked for in the next day to identify the suitable hosts and phages.

3.2.2 Bacteriophage titer determination

Titer or concentration of the appropriate bacteriophages (JSF2, JSF5, JSF24, and JSF35) were determined prior to the experiment to hold a clear idea about the count of bacteriophage incorporated in the tests. First of all, one single colony of a suitable host was taken in vial containing LB medium and cultured for 3 hours to obtain the young bacteria. 300-µl LB medium containing young bacterial cultures were then mixed with fresh soft agar and poured onto LA plates followed by drying for 15 minutes. Separate plates were then utilized for different phages and spotted according to their dilution factors. All the bacteriophages were then diluted until 10^{-8} and equal volume (5 µl) of phages of all dilution were put onto the plates containing bacterial culture in soft agar in the marked spots. The plates were at that point dried for 30 minutes to get the bacteriophages absorbed and incubated overnight at 37^{o}C. Plaques caused by bacteriophages were then counted at different dilution to determine the titer where a single and small plaque represent one bacteriophage.

3.2.3 Growth competition between large and small genome sizes bacteriophages

Two bacteriophages were selected named JSF5 and JSF35 representing large genomic size and small genomic size respectively. An appropriate common host bacteria (WT-346) that is lysed both the phages was selected based on the phage typing result. A single colony of the host bacteria was cultured in LB medium for 3 hours to get the young culture and 300 µl LB medium was passed to another vial containing fresh LB. Equal count of bacteriophages (65000) of both large and small genomic sizes were added into the vial containing young bacteria in LB. This mixture was then cultured for 4 hours in shaker incubator at 37^{o}C and followed by centrifugation for 5 minutes at 13000 rpm to settle down the bacterial cells as well as releasing the phages. The supernatant was filtered through 0.22 µm syringe filter to obtain the pure bacteriophage solution containing phages of both large and small genomic size. On the other hand, two specific host bacteria (040 specific for JSF5 and WT-333 specific for JSF35) to

determine the growth of separate phages were selected based on phage typing result. These specific bacterial strains are not lysed by both the phages, rather they are lysed by one phage and remain unaffected by other present in the mixed solution, therefore provide selection advantage. One single colony of a both specific hosts were taken in vials containing LB medium and cultured for 3 hours to obtain the young bacteria. 300-µl LB medium containing young bacterial cultures were then mixed with fresh soft agar and poured onto LA plates followed by drying for 15 minutes. Afterward, bacteriophage solution was diluted until 10^{-10} and equal volume (5 µl) of phages of all dilution were put onto the plates containing specific host bacteria in soft agar in the marked spots. The plates were at that point dried for 30 minutes to get the bacteriophages absorbed and incubated overnight at 37ºC. Plaques caused by bacteriophages at different plates were then counted at different dilution to determine the titer against specific hosts.

3.2.4 Growth competition between large genome size bacteriophages

Two bacteriophages named JSF2 and JSF5 representing large genomic load were selected depending on their ability to infect a common host bacteria (WT-346) based on the phage typing result. A single colony of the host bacteria was cultured in LB medium for 3 hours to get the young bacteria and 300 µl LB medium was taken into another vial containing fresh LB. Equal count of both bacteriophages (100,000) of large genomic size were appended into the vial containing young common host bacteria in LB. This mixture was then cultured for 4 hours in shaker incubator at 37ºC and followed by centrifugation for 5 minutes at 13000 rpm to settle down the bacterial cells as well as releasing the phages in supernatant. The supernatant was filtered through 0.22 µm syringe filter to obtain the pure solution containing both bacteriophages of large genomic size. In the meantime, due to lack of availability of specific hosts for both phages, one host (040 specific for JSF5) was selected which could be lysed by one phage and remain unaffected by other. In this case, plaques caused by one phage in specific

strain was subtracted from plaques created by both phages in common host to obtain the titer of another phage. One single colony of a both specific host and common host were taken in vials containing LB medium and cultured for 3 hours to obtain the young bacteria. 300-µl LB medium containing young bacterial cultures were then mixed with fresh soft agar and poured onto LA plates followed by drying for 15 minutes. Afterward, solution containing mixed bacteriophages was diluted until 10^{-8} and equal volume (10 µl) of phages of all dilution were put onto the plates containing specific host and common host bacteria in soft agar in the marked spots. The plates were at that point dried for 30 minutes to get the bacteriophages absorbed and incubated overnight at 37°C. Plaques caused by bacteriophages in different plates were then counted at different dilution to obtain the titer.

3.2.5 Growth competition between small genome size bacteriophages

Two bacteriophages named JSF24 and JSF35 representing small genomic size were selected depending on their ability to infect a common host bacteria (WT-346) based on the phage typing result. On the other hand, due to lack of availability of specific hosts for both phages, one host (040 specific for JSF24) was selected which could be lysed by one phage and remain unaffected by other. At this point, plaques caused by one phage in specific strain was subtracted from plaques created by both phages in common host to obtain the titer of other phage. In the beginning, a single colony of the host bacteria was cultured in LB medium for 3 hours to get the young bacteria and 300 µl LB medium was transferred into another vial containing fresh LB. Equal count of both bacteriophages (100,000) of both small genomic size were appended into the vial containing young common host bacteria in LB. This mixture was then cultured for 4 hours in shaker incubator at 37°C and followed by centrifugation for 5 minutes at 13000 rpm to settle down the bacterial cells as well as releasing the phages in supernatant. The supernatant was filtered through 0.22 µm syringe filter to obtain the pure solution containing both bacteriophages of large genomic size. During this time, one single colony of a both specific

host and common host were taken in vials containing LB medium and cultured for 3 hours to obtain the young bacteria. 300-µl LB medium containing young bacterial cultures were then mixed with fresh soft agar and poured onto LA plates followed by drying for 15 minutes. Afterward, solution containing mixed bacteriophages was diluted until 10^{-8} and equal volume (10 µl) of phages of all dilution were put onto the plates containing specific host and common host bacteria in soft agar in the marked spots. The plates were at that point dried for 30 minutes to get the bacteriophages absorbed and incubated overnight at 37ºC. Plaques caused by bacteriophages in different plates were then counted at different dilution to determine the titer.

3.2.6 Adsorption rate difference between large and small genome sizes bacteriophages

JSF5 and JSF35 two bacteriophages representing large genomic size and small genomic size respectively were chosen for this experiment to test their insertion capacity inside host bacterial cells. For this test, an appropriate common host bacteria (WT-346) that is lysed both the phages was selected based on the phage typing result and a single colony of the host bacteria was cultured in LB medium for 3 hours to get the young bacteria and 300 µl LB medium was passed to another vial containing fresh LB. 10,00,000 of both bacteriophages of both large and small genomic sizes were added into the vial containing young bacteria in LB. This mixture was then cultured for only 20 minutes in shaker incubator at 37ºC and followed by centrifugation for 5 minutes at 13000 rpm to settle down the bacterial cells. The supernatant was filtered through 0.22 µm syringe filter to obtain the pure bacteriophage solution containing phages of both large and small genomic size. On the other hand, two specific host bacteria (040 specific for JSF5 and WT-333 specific for JSF35) to determine the growth of individual phages were chosen based on phage typing result. One single colony of a both specific host were taken in vials containing LB medium and cultured for 3 hours to obtain the young bacteria and 300 µl LB medium containing young bacterial cultures were then mixed with fresh soft agar and poured onto LA plates followed by drying for 15 minutes. Afterward, bacteriophage solution was

diluted until 10^{-10} and equal volume (5 µl) of phages of all dilution were put onto the plates containing specific host bacteria in soft agar in the marked spots. The plates were at that point dried for 30 minutes to get the bacteriophages absorbed and incubated overnight at 37ºC. Plaques caused by bacteriophages were then counted at different dilution to obtain the titer.

3.2.7 Individual growth rate of large and small genome size bacteriophages

Two bacteriophages named JSF5 and JSF35 representing large genomic size and small genomic size respectively were chosen to determine their growth rate when they are not in competition with other type of phages for host cells. For this test, an appropriate host bacteria (WT-346) was selected based on the phage typing result and a single colony of the host bacteria was cultured in LB medium for 3 hours to get the young bacteria and 300 µl LB medium was passed to another two separate vials containing fresh LB. Afterward, 100,000 bacteriophages of both large and small genomic sizes were added into separate vials containing young bacteria in LB. These mixtures were then cultured for only 4 hours in shaker incubator at 37ºC and followed by centrifugation for 5 minutes at 13000 rpm to settle down the bacterial cells. The supernatant was filtered through 0.22 µm syringe filter to obtain the pure bacteriophage solution containing phages of both large and small genomic size in separate vials. On the other hand, two specific host bacteria (040 specific for JSF5 and WT-333 specific for JSF35) to determine the growth rate of individual phages were chosen based on phage typing result. One single colony of a both specific host were taken in vials containing LB medium and cultured for 3 hours to obtain the young bacteria and 300 µl LB medium containing young bacterial cultures were then mixed with fresh soft agar and poured onto LA plates followed by drying for 15 minutes. Afterward, solution of both bacteriophages were diluted until 10^{-10} and equal volume (5 µl) of phages of all dilution were put onto the plates containing specific host bacteria in soft agar in the marked spots. The plates were at that point dried for 30 minutes to get the

bacteriophages absorbed and incubated overnight at 37ºC. Plaques caused by bacteriophages were then counted at different dilution to obtain the titer.

3.2.8 Bacteriophage growth rate in presence of non-affecting bacteriophage

This experiment was carried out using bacteriophage JSF35, JSF5 and another O1 *Vibrio cholerae* non-affecting but *E. coli* specific phage collected from the environment. First of all a host bacteria (WT-346) that is lysed by both JSF35 and JSF5 was selected based on the phage typing result. A single colony of the host bacteria was cultured in LB medium for 3 hours to get the young bacteria and 300 µl LB medium was passed to another vial containing fresh LB. Equal amount of bacteriophages (100,000) of JSF35, JSF5 and *E. coli* phage were added into the vial containing young bacteria in LB. This mixture was then cultured for 4 hours in shaker incubator at 37ºC and followed by centrifugation for 5 minutes at 13000 rpm to settle down the bacterial cells as well as releasing the phages. The supernatant was filtered through 0.22 µm syringe filter to obtain the pure bacteriophages solution containing JSF35, JSF5 and *E. coli* phage. Then one single colony of specific hosts (040 specific for JSF5 and WT-333 specific for JSF35) were selected based on phage typing result and cultured for 3 hours to obtain the young bacteria. 300-µl LB medium containing young bacterial cultures were then mixed with fresh soft agar and poured onto LA plates followed by drying for 15 minutes. Afterward, bacteriophage solution was diluted until 10^{-10} and equal volume (5 µl) of phages of all dilution were put onto the plates containing specific host bacteria in soft agar in the marked spots. The plates were at that point dried for 30 minutes to get the bacteriophages absorbed and incubated overnight at 37ºC. Plaques caused by bacteriophages at different plates were then counted at different dilution to determine the titer.

3.2.9 Burst out time of large and small genome size bacteriophages

Set of bacteriophages incorporating JSF5 and JSF35 were chosen for analyzing burst out time of bacteriophages having different genomic size. Bacterial strain WT-346 were used as the host for both of the bacteriophages then a single colony of the host bacteria was cultured in LB medium for 3 hours to get the young bacteria and 300 µl LB medium was passed to 10 separate vials containing fresh LB. Afterward, 100,000 bacteriophages of both large and small genomic sizes were added into separate vials containing young bacteria in LB. These mixtures were then cultured for only 20, 25, 30, 35, 40, 45 minutes in shaker incubator at 37ºC and followed by centrifugation for 5 minutes at 13000 rpm to settle down the bacterial cells. The supernatant was filtered through 0.22 µm syringe filter to obtain the pure bacteriophage solution containing phages of both large and small genomic size in separate vials. In the meantime, two specific host bacteria (040 specific for JSF5 and WT-333 specific for JSF35) to determine the growth rate of individual phages were chosen based on phage typing result. One single colony of a both specific host were taken in vials containing LB medium and cultured for 3 hours to obtain the young bacteria and 300 µl LB medium containing young bacterial cultures were then mixed with fresh soft agar and poured onto LA plates followed by drying for 15 minutes. Afterward, solution of both bacteriophages of different times were diluted till 10^{-6} and equal volume (5 µl) of phages of all dilution were put onto the plates containing specific host bacteria in soft agar in the marked spots. The plates were at that point dried for 30 minutes to get the bacteriophages absorbed and incubated overnight at 37ºC. Plaques caused by bacteriophages were then counted at different dilution to obtain the titer.

3.3 Bioinformatic Analysis

3.3.1 NCBI data analysis:

To resolve this hypothesis, I started with analysis the Genebank data to get an idea about the real world data regarding bacteriophage existence according to their size distribution. Up to

date, 9575 complete sequences of bacteriophages are deposited in the GeneBank database from different researches throughout the world. I began with downloading all the sequences from the database searched using the term ‘phage’ then sorted by using “grep” command in linux command line using two keywords ‘phage ’ and ‘complete sequences. All the accession IDs were then used to extract the desired sequences of complete phage genomes using the tools seqtk, using the program subseq. Complete phage genomes were then indexed using samtools, under the program faidx (Li et al., 2009) that provided the sequence length of the genome and this length were then plotted in R program to generate graph.

3.3.2 Metagenome data analysis

Although GeneBank phages have revealed that smaller genome containing phages are more abundant in nature, however question might be raised that, Genebank sequences are free from biasness but environment is highly biased towards to fittest organisms. In response to this question, I have analyzed extensive bacteriophage metagenomes. To begin with, more than 6,00,00,000 metagenomic sequences from 241 samples of 18 different projects were downloaded having 89GB data volume. Those sample contained sequences of every organisms present at that environment, as a result it was necessary to identify only the bacteriophage sequences. To do so, PPR-Meta tool was used which is an open source deep learning based program available at GitHub (Fang et al., 2019). The predicted phage sequences were further validated using another novel bacteriophage prediction tool, MARVEL (Amgarten, Braga, da Silva, & Setubal, 2018). Since the metagenome contains many short sequences, therefore it was indispensable to carry out binning of the sequences to reconstruct the genomes. Aimed at this purpose, predicted sequences were binned using MetaBat2 and further checked by CheckM program (Kang et al., 2019; Parks, Imelfort, Skennerton, Hugenholtz, & Tyson, 2015). Upon the completion of all these procedures, I was left with 19,673 bacteriophage sequences, which were indexed by faidx program under samtools and plotted usning R program.

Chapter 4: Results

All the experiments were conducted three times to ensure the accuracy of data obtained from the analysis and their average value was considered as the final result.

4.1 Bacteriophage typing

Twenty-two bacteriophages were used to lyse sixteen *Vibrio Cholerae* O1 strains to select the suitable candidates for this explorative research in order to find out the effects of genomic contents

Table 4.1: Bacteriophage typing results

Strain name	JSF 1	JSF 2	JSF 4	JSF 5	JSF 6	JSF 7	JSF 10	JSF 11	JSF 13	JSF 17	JSF 18
WT – 346	+	+	+	+	+	+	+	+	+	+	+
WT– 324	–	–	–	+	+	–	+	+	+	+	+
WT– 333	–	–	–	–	–	–	+	+	–	+	+
048	–	–	–	+	–	–	+	+	–	+	+
049	–	–	–	–	–	–	+	+	+	+	+
042	–	–	–	–	–	–	+	+	+	+	+
050	–	+	+	+	+	–	+	+	+	+	+
WT– 334	–	+	–	+	+	–	+	+	+	+	+
033	–	+	–	+	+	–	+	+	–	+	+
004	–	+	–	+	+	–	+	+	–	+	+
036	–	+	–	+	+	–	+	+	–	+	+
040	–	–	–	+	+	–	+	–	–	+	+
035	+	+	+	+	+	–	+	+	–	+	+
031	–	–	–	–	–	–	+	+	–	+	+
005	–	+	–	–	–	–	+	+	–	+	+
041	–	+	–	+	+	–	+	+	–	+	+
006	–	–	–	–	–	–	+	+	–	+	+
1667	+	+	+	+	+	+	+	+	–	+	+

Strain name	JSF 20	JSF 23	JSF 24	JSF 25	JSF 27	JSF 28	JSF 30	JSF 31	JSF 32	JSF 33	JSF 35
WT – 346	+	+	+	+	+	+	+	+	+	+	+
WT– 324	+	+	+	+	–	+	+	+	+	+	+
WT– 333	+	+	+	+	+	+	+	+	+	+	+
048	+	+	+	+	+	+	+	+	+	+	+
049	+	+	+	+	+	+	+	+	+	+	+
042	+	+	+	+	+	+	+	+	+	+	+
050	+	+	+	+	+	+	+	+	+	+	+
WT– 334	+	+	+	+	+	+	+	+	+	+	+
033	+	+	+	+	+	+	+	+	+	+	+
004	+	+	+	+	+	+	+	+	+	+	+
036	+	+	+	+	+	+	+	+	+	+	+
040	–	+	–	+	+	+	+	+	+	+	–
035	+	+	+	+	+	+	+	+	+	+	+
031	+	+	+	+	+	+	+	+	+	+	+
005	+	+	+	+	+	+	+	+	+	+	+
041	+	+	+	+	+	+	+	+	+	+	+
006	+	+	+	+	+	+	+	+	+	+	+
1667	+	+	+	+	+	+	+	+	+	+	+

on the rate of multiplication of bacteriophages. Sixteen bacterial strains used are isolated from the environment under different projects at different times and some of them are reference strain indexed in Genebank. On the other hand, all the bacteriophages used are reference strains indexed in Genebank database and from phage typing result; it is observed that almost 64% bacteriophages contain smaller genome whereas around 36% bacteriophages contain larger genome. Interestingly, from the result I found that smaller genome containing has broader host range compared with the larger genome containing phages. Moreover, larger genome

containing phages varies in host selection, for example JSF1 and JSF2 has similar genome size but their host range varies reportedly whereas this variation is very much absent is smaller genome containing phages. Analyzing the lytic capability three sets of phages and hosts were preferred which served the best.

1. For smaller and larger genome size bacteriophage

Table 4.2: set of bacteriophages and host of small and large genome size

Strain name	JSF 5 (132 kb)	JSF 35 (38.5 kb)
Common host	050	
Specific host	040	WT-333

2. For both larger genome size bacteriophages

Table 4.3: set of bacteriophages and host of small and large genome size

Strain name	JSF 2 (126 kb)	JSF 5 (132 kb)
Common host	WT-346	
Specific host	005	WT-324

3. For both smaller genome size bacteriophages

Table 4.4: set of bacteriophages and host of small and large genome size

Strain name	JSF 27 (48.6 kb)	JSF 35 (38.5 kb)
Common host	WT-346	
Specific host	040	WT-324

4.2 Titer of suitable bacteriophages

Phages lyse bacterial lawns on plates to produce clear zones known as plaques, which are used to determine the titer or concentration of the phages. PFUs (plaque-forming units) is the worldwide-accepted expression of phage titer in a given volume. Titer determination of the bacteriophages used was a must prior to the experiments to hold clear idea about the incorporation of number of bacteriophages during different studies to ensure that number of phages brought into play were not biased. Ahead of titer determinations, all the bacteriophages were enriched several times to increase their concentrations in the stock, as well as in working solutions. Since the titer estimation

Table 4.5: Titer of bacteriophages

Strain name	Dilution	PFU/5µl	Concentration/ml
JSF 2	10^{-7}	3	6000000000
JSF 5	10^{-7}	8	16000000000
JSF 27	10^{-4}	4	8000000
JSF 35	10^{-4}	3	6000000
E. Coli phage	10^{-4}	11	22000000

was carried out using PFU at different dilutions, the actual number might vary but provide a strong idea about the amount of solution to be employed to carry out different tests. Stock solution of JSF 2 and JSF 5 were highly enriched therefore they were diluted according to the number of phages required whereas JSF27, JSF35 and E. coli phages were less in count, therefore raw phage solutions were devoted to the tests.

4.3 Growth competition between large and small genome size bacteriophages

Table 4.6: Multiplication rate difference between large and small genome containing bacteriophages

Test number	Small genome size (JSF 35)				Large genome size (JSF 5)			
	Culture time (hours)	Input	Output		Culture time (hours)	Input	Output	
			Dilution factor	Count (PFU/5µl)			Dilution factor	Count (PFU/5µl)
Test 1	4	30000/ml	10^{-6}	8	4	30000/ml	10^{-3}	7
Test 2	4	30000/ml	10^{-6}	6	4	30000/ml	10^{-3}	11
Test 3	4	30000/ml	10^{-6}	11	4	30000/ml	10^{-3}	10

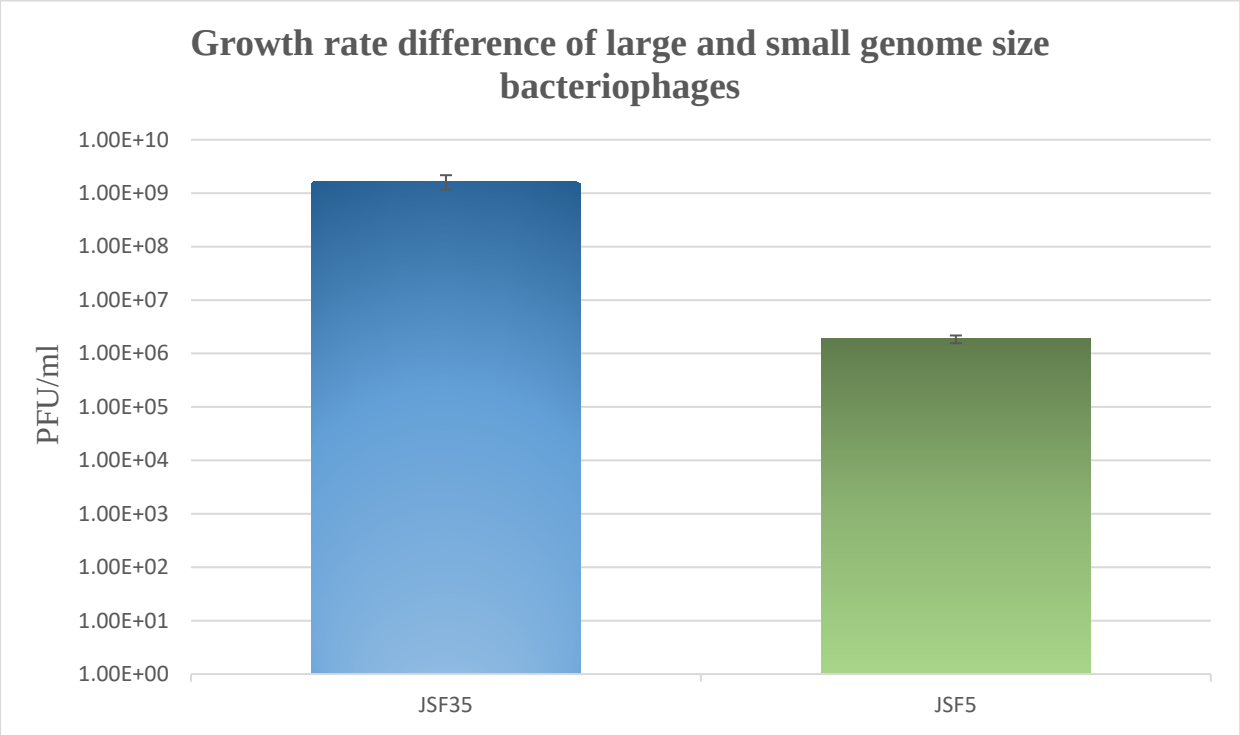

Figure 4.1: Multiplication rate difference between large and small genome containing bacteriophages

The core hypothesis of this research was to examine the bacteriophage multiplication rate difference based on their genomic size. From three different sets of tests incorporating equal number of bacteriophages as input, I observed significant difference between their multiplication rates. Bacteriophages were introduced into the culture medium with abundant

amount of host bacteria as 30000/ml concentration for both of the phages. The outcome revealed the impact of their genomic content since I observed on average 9 plaques at 10^{-2} dilution for large genomic bacteriophage which results in 1800000/ml final concentration whereas on average 8 plaques were witnessed for small genome containing bacteriophage at 10^{-5} dilution leads to the final concentration of 1650000000/ml after 4 hours of co-cultivation. Here we can see, large genome containing bacteriophage JSF5 increased 6 times in count whereas small genome containing bacteriophage JSF35 increased around 5000 times which is 900 times higher than JSF5 increment. Concisely, small genome containing bacteriophage JSF35 increased in number by three logarithmic value higher than large genome containing bacteriophage JSF5.

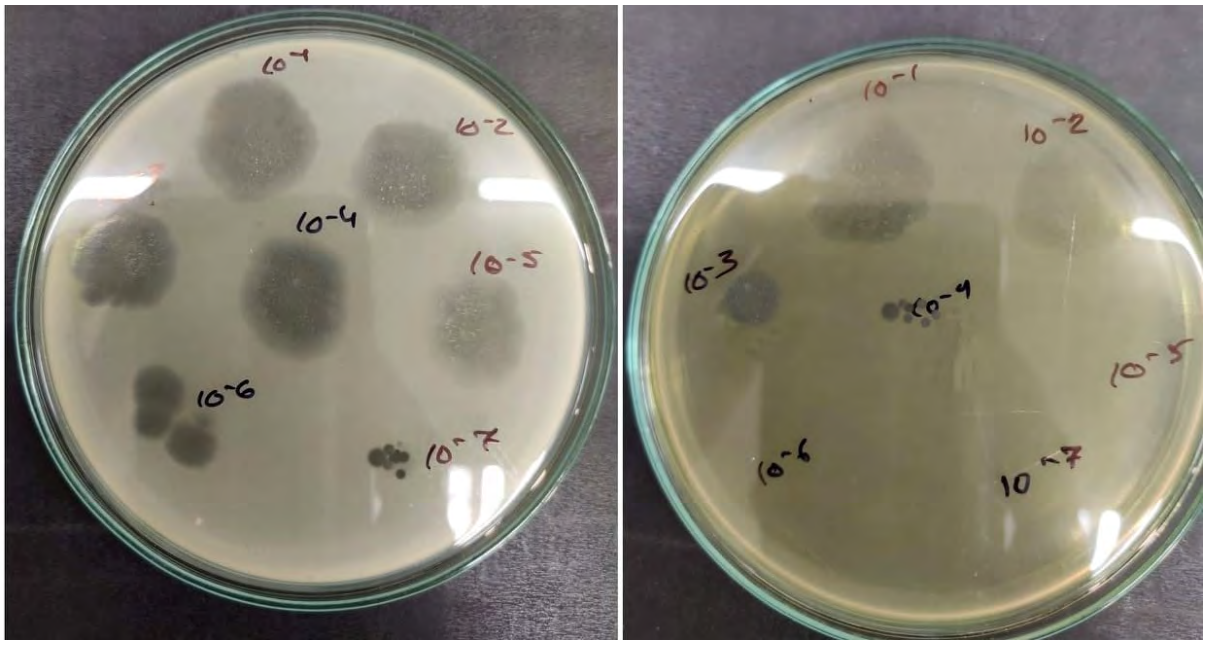

Figure 4.2: Plaques formed by JSF5 and JSF35 when co-cultured

4.4 Growth competition between large genome size bacteriophages

Table 4.7: Multiplication rate difference between large genome containing bacteriophages

Test number	Small genome size (JSF 5)				Large genome size (JSF 2)			
	Culture time (hours)	Input	Output		Culture time (hours)	Input	Output	
			Dilution	Count (PFU/5µl)			Dilution	Count (PFU/5µl)
Test 1	4	50000/ml	10^{-3}	16	4	50000/ml	10^{-4}	2
Test 2	4	50000/ml	10^{-4}	5	4	50000/ml	10^{-4}	6
Test 3	4	50000/ml	10^{-3}	19	4	50000/ml	10^{-4}	5

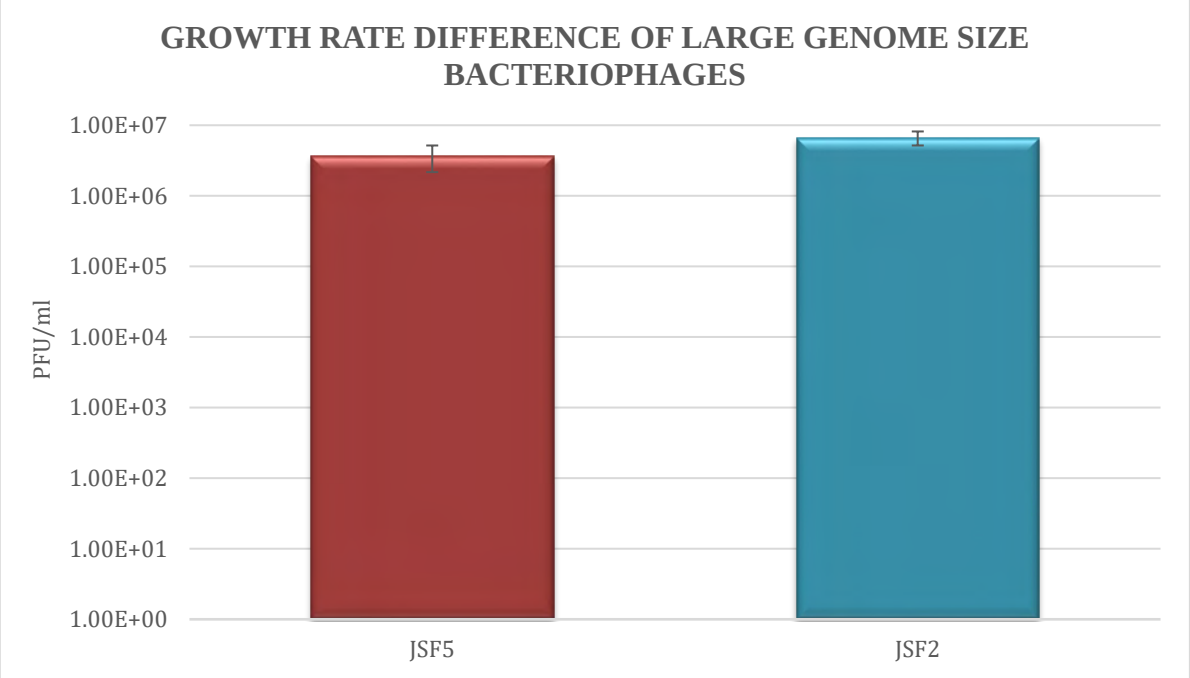

Figure 4.3: Multiplication rate difference between large genome containing bacteriophages

In the very next step, I tried to observe the multiplication rates of similar genome size bacteriophages to understand the significance of dissimilar genome size. Therefore, two bacteriophages of large genome size JSF2 and JSF5 was added into the culture medium as 50000/ml concentration. After 4 hours of successful co-cultivation, on average 5500000/ml JSF5 was found whereas JSF2 was present as 8000000/ml concentration. These increments

disclose that JSF5 increased 110 times and JSF2 increased around 160 times compared with the input value. Increment of JSF2 was 1.45 times higher than JSF5, which unveil no logarithmic value difference between their multiplication rates.

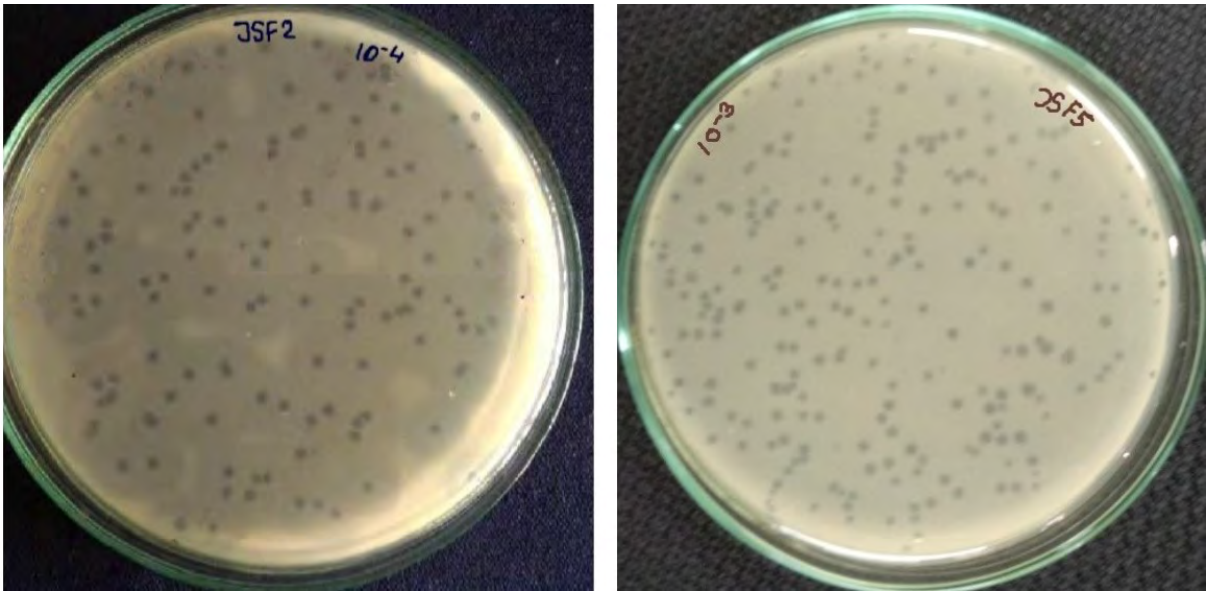

Figure 4.4: Plaques formed by JSF5 and JSF2 when co-cultured

4.5 Growth competition between small genome size bacteriophages

Table 4.8: Multiplication rate difference between small genome containing bacteriophages

Serial number	Small genome size (JSF 35)				Large genome size (JSF 27)			
	Culture time (hours)	Input	Output		Culture time (hours)	Input	Output	
			Dilution	Count (PFU/5µl)			Dilution	Count (PFU/5µl)
Test 1	4	50000/ml	10^{-6}	4	4	50000/ml	10^{-6}	6
Test 2	4	50000/ml	10^{-6}	9	4	50000/ml	10^{-6}	11
Test 3	4	50000/ml	10^{-6}	7	4	50000/ml	10^{-6}	6

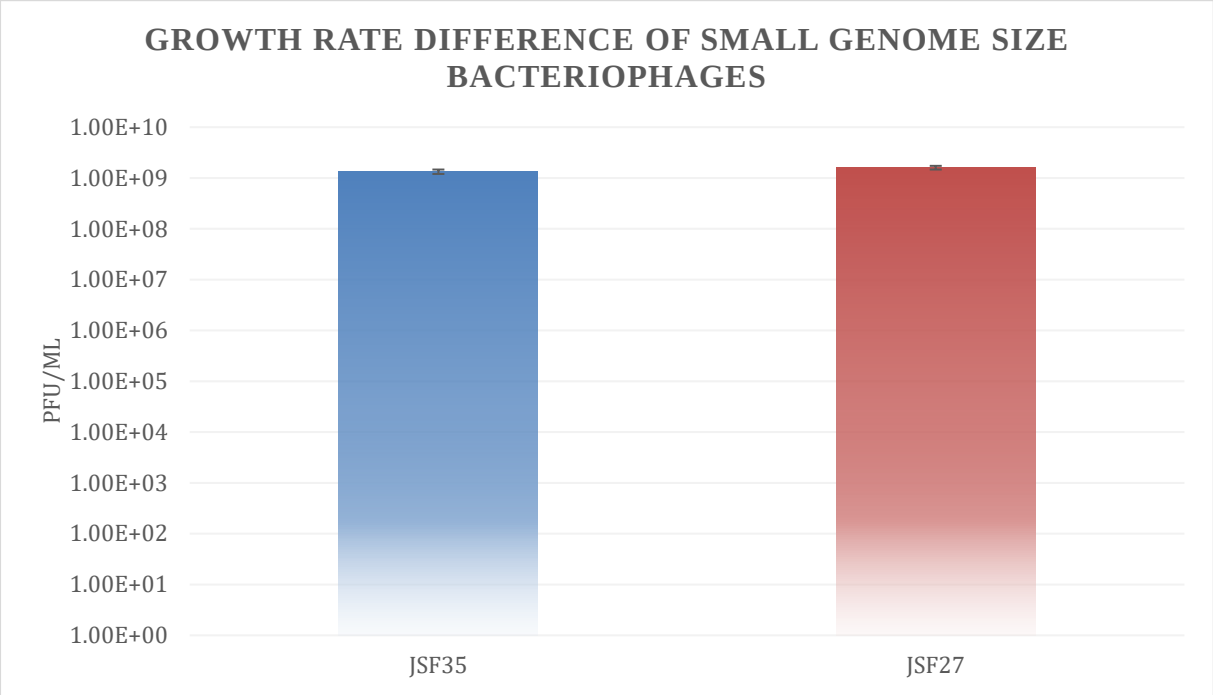

Figure 4.5: Multiplication rate difference between small genome containing bacteriophages

After observing the growth rate of large genome containing bacteriophages, similar tests were carried out for small genome containing bacteriophages JSF27 and JSF35. 50000/ml phages of both JSF27 and JSF35 were supplemented into the culture medium with sufficient bacteria for 4 hours of successful co-cultivation. From the final output, I observed on average 7 plaques

were present at 10^{-6} dilution for JSF35 and this value was 8 plaques for JSF27 at similar dilution. The augmentation rate of JSF35 was 26500 times with final concentration on average 1130000000/ml while JSF27 increased 32000 times than initial input with 1330000000/ml final concentration. Here, these results indicate that JSF27 increased 1.17 times higher than JSF35 that reveals no logarithmic value difference.

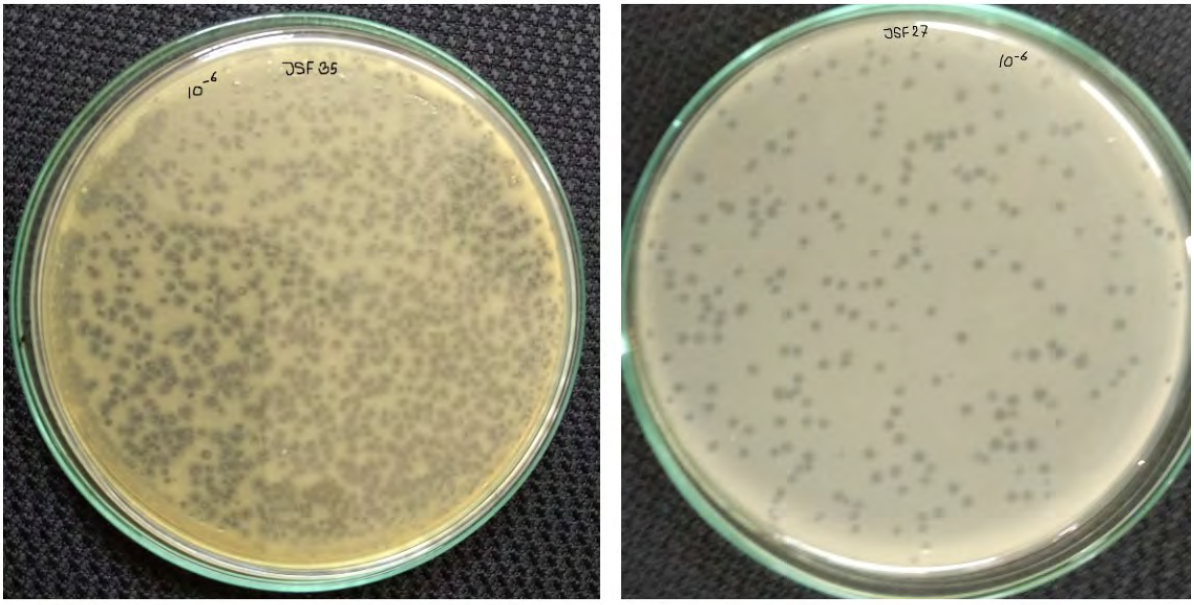

Figure 4.6: Plaques formed by JSF27 and JSF35 when co-cultured

4.6 Adsorption rate difference between large and small genome sizes bacteriophages

Table 4.9: Adsorption rate difference between large and small genome sizes bacteriophages

Serial number	Small genome size (JSF 5)				Large genome size (JSF 35)			
	Culture time (mins)	Input	Output		Culture time (mins)	Input	Output	
			Dilution	Count (PFU/5µl)			Dilution	Count (PFU/5µl)
Test 1	20	500000/ml	10^{-2}	2	20	500000/ml	10^{-2}	3
Test 2	20	500000/ml	10^{-2}	5	20	500000/ml	10^{-2}	7
Test 3	20	500000/ml	10^{-2}	4	20	500000/ml	10^{-2}	4

Afterwards, observing the multiplication rate difference among the bacteriophages, I tried to study the adsorption rate since the multiplication rate can be highly affected by adsorption pattern. If the adsorption rate is different then the multiplication rate will be different because highly absorbed bacteriophages will multiply faster inside the hosts. In this study, 500000/ml bacteriophages JSF5 and JSF35 were appended into the culture medium and co-cultivated for 20 minutes to observe the adsorption rate since adsorption time for most of the bacteriophages are 15 to 20 minutes. After 20 minutes of co-cultivation, on averages 4 to 5 plaques were present at 10^{-2} dilution for JSF35 and 3 to 4 plaques were obtained for JSF5 at similar dilution resulting in 140000/ml and 80000/ml final concentration, which were not absorbed until then. These results specifies that 78% JSF35 bacteriophages were absorbed within 20 minutes when this value is almost 84% for JSF5.

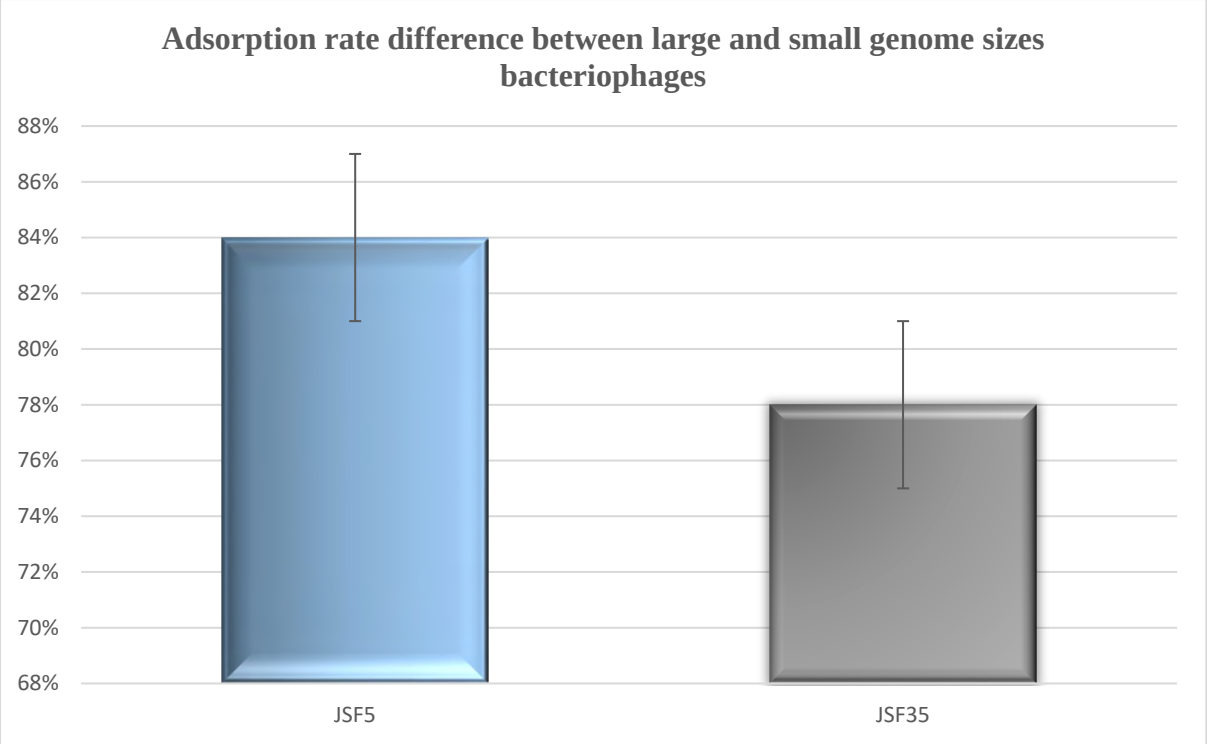

Figure 4.7: Adsorption rate difference between large and small genome sizes bacteriophages

4.7 Growth rate of large and small genome size bacteriophages individually

Table 4.10: Growth rate of large and small genome size bacteriophages individually

Serial number	Small genome size (JSF 35)				Large genome size (JSF 5)			
	Culture time (hours)	Input	Output		Culture time (hours)	Input	Output	
			Dilution	Count (PFU/5µl)			Dilution	Count (PFU/5µl)
Test 1	4	50000/ml	10^{-7}	6	4	50000/ml	10^{-4}	9
Test 2	4	50000/ml	10^{-7}	8	4	50000/ml	10^{-4}	4
Test 3	4	50000/ml	10^{-7}	5	4	50000/ml	10^{-3}	10

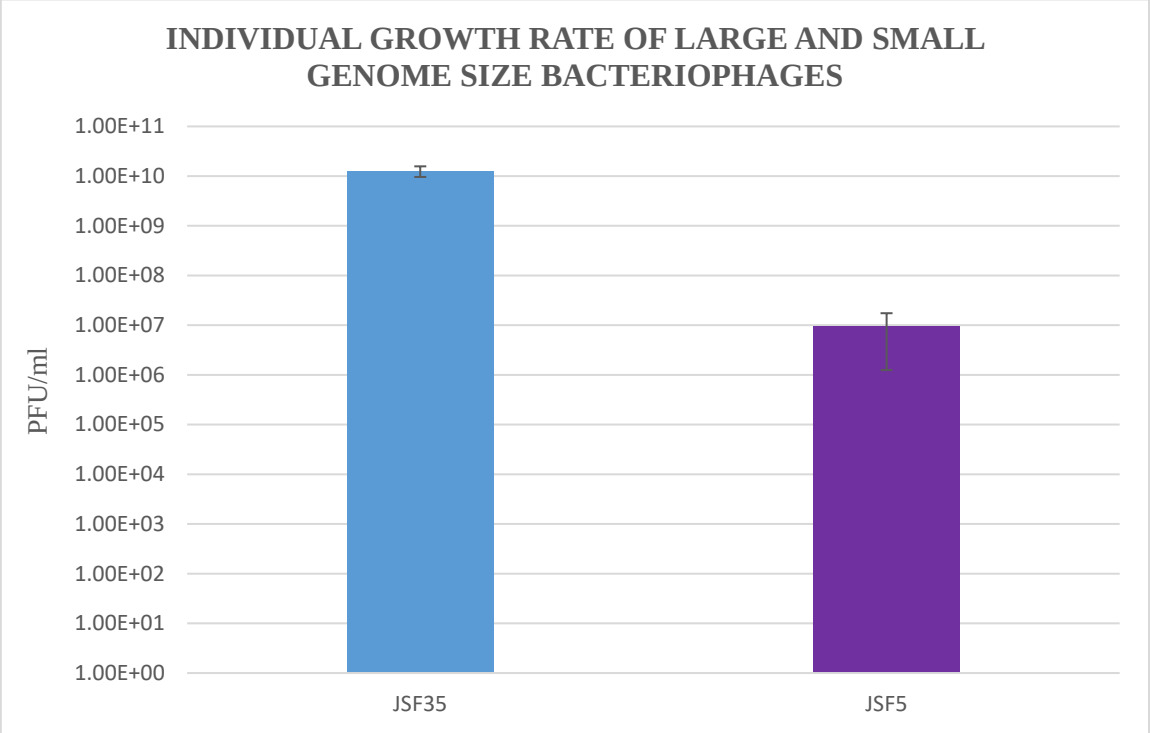

Figure 4.8: Growth rate of large and small genome size bacteriophages individually

After all these studies, the next question came into play was, what would be the multiplication rate of bacteriophages depending on their genome size if they are grown individually without any competition with other bacteriophages? To investigate this, 50000/ml bacteriophages were incorporated and cultivated for 4 hours without any competition. The final outcome was around

6 plaques at 10-7 dilution for JSF35 and 8 to 9 plaques for JSF5 at 10-3 dilution. Reverse calculation confirmed that 12600000000/ml small genome containing JSF35 bacteriophages were present, which was 252000 times higher than the initial bacteriophage number. On the other hand, final concentration of large genome containing bacteriophage JSF5 was 9500000/ml, which was 190 times higher than preliminary bacteriophage number. Therefore, JSF35 augmentation rate was 1260 higher than JSF5, which reveals three logarithmic value difference.

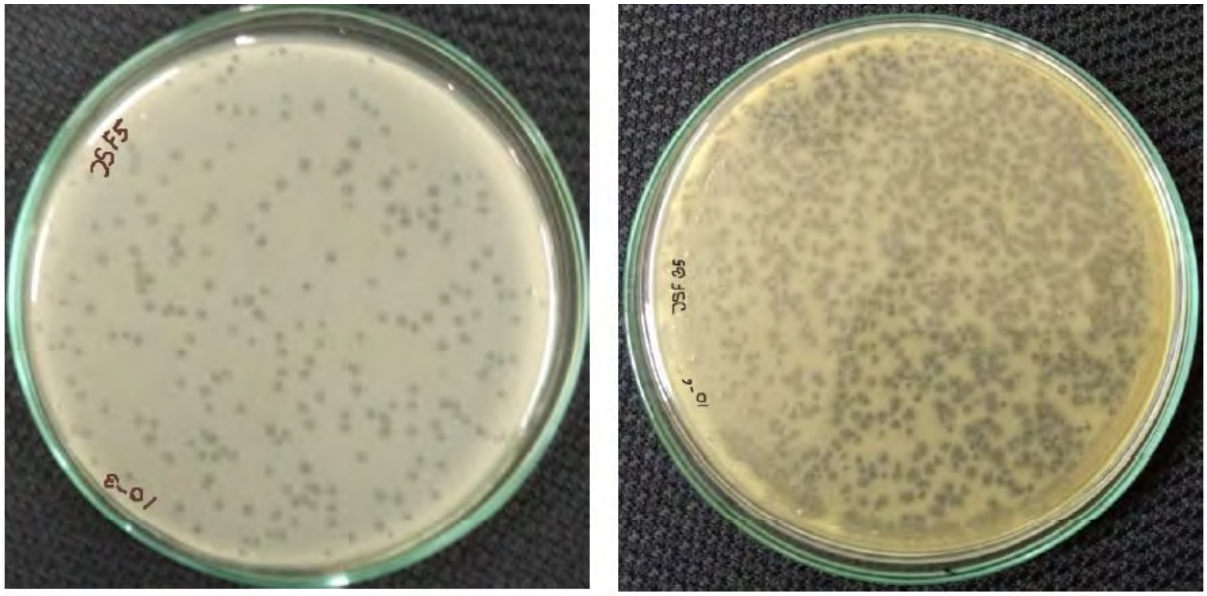

Figure 4.9: Plaques formed by JSF5 and JSF35 when grown individually

4.8 Bacteriophage growth rate in presence of non-affecting bacteriophage

Table 4.11: Bacteriophage growth rate difference in presence of non-affecting bacteriophage

Serial number	Small genome size (JSF 35)				Large genome size (JSF 5)			
	Culture time (hours)	Input	Output		Culture time (hours)	Input	Output	
			Dilution	Count (PFU/5µl)			Dilution	Count (PFU/5µl)
Test 1	4	40000/ml	10^{-6}	4	4	40000/ml	10^{-3}	8
Test 2	4	40000/ml	10^{-6}	7	4	40000/ml	10^{-3}	12
Test 3	4	40000/ml	10^{-6}	4	4	40000/ml	10^{-4}	2

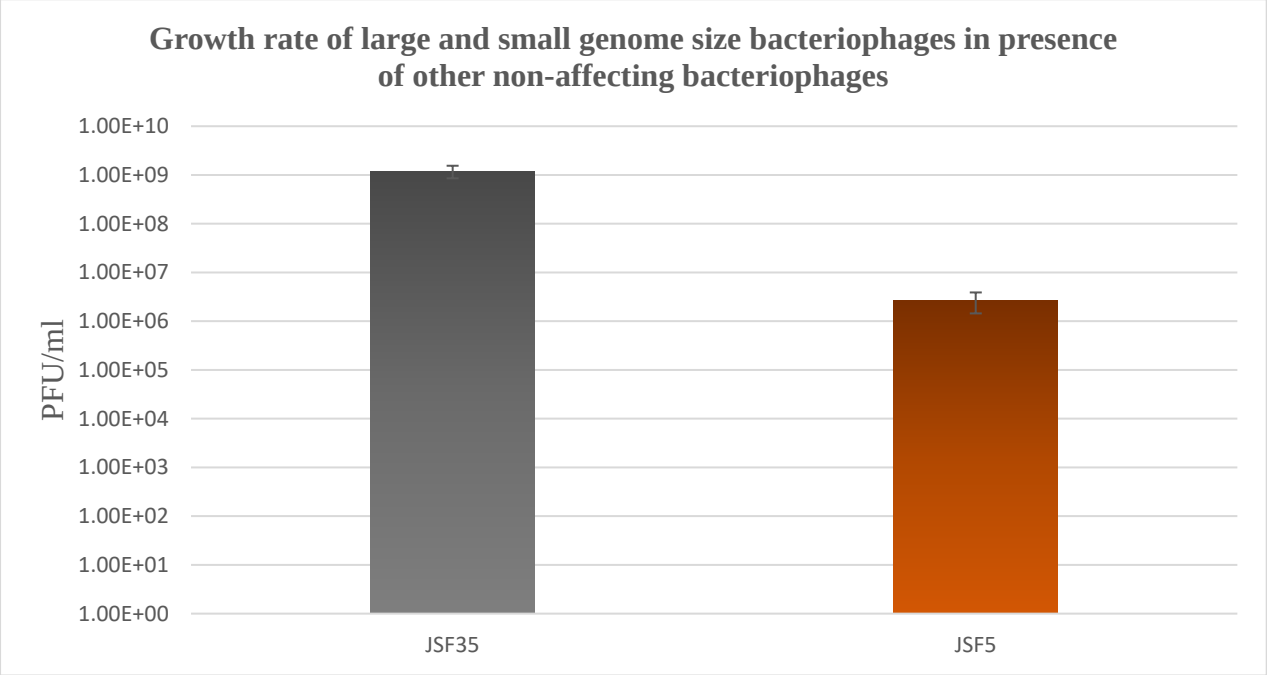

Figure 4.10: Bacteriophage growth rate difference in presence of non-affecting phage

Taking all the results into account, the next concern was the presence of other phages which are abundant in our natural system but do not lyse the host, do they have any effect on the multiplication rates of those lytic phages? In response to this question, 40000/ml JSF5, JSF35 and 500000/ml E. coli bacteriophages were transferred into the medium containing plentiful host bacteria. After 4 hours of co-cultivation of these three bacteriophages, 6 PFU was found

to be present of JSF35 at 10^{-6} dilution that leads to the final concentration 1200000000/ml. On JSF5 formed on average 16 plaques at 10^{-2} dilution resulted in the ultimate concentration 2600000/ml. Titer of E. coli bacteriophage was not determined since it was non-affecting. JSF35 augmented 30000 times than initial count at the same time as JSF5 increased 65 times higher than preliminary concentration. The ratio of JS35 intensification was 460 times higher than JSF5, which is almost 2.5 logarithmic difference.

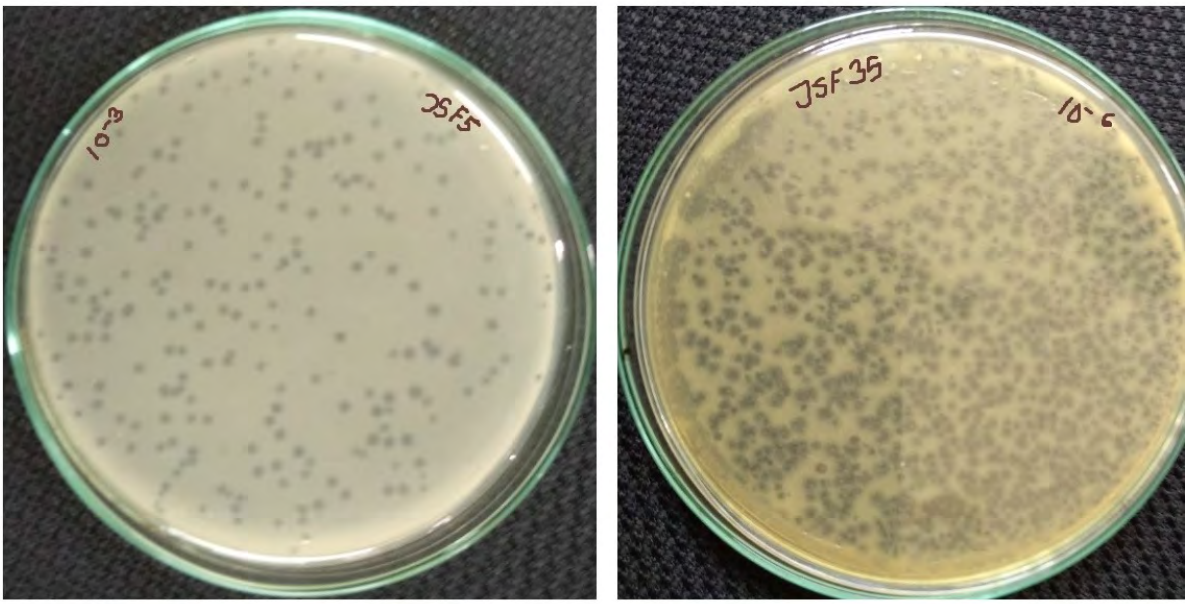

Figure 4.11: Plaques formed by JSF5 and JSF35, grown in presence of non-specific phage

4.9 Burst out time of large and small genome size bacteriophages

Table 4.12: Burst out time of large and small genome size bacteriophages

Serial number	Small genome size (JSF 35)				Large genome size (JSF 5)			
	Input	Culture time	Output		Culture time (hours)	Input	Output	
			Dilution	Count (PFU/5µl)			Dilution	Count (PFU/5µl)
Test 1	200000/ml	25	10^{-1}	2	25	200000/ml	10^{-1}	1
		30		18	30			1
		35		16	35			2
		40		19	40			3
		45		25	45			14
		50		26	50			18
Test 2	200000/ml	25	10^{-1}	3	25	200000/ml	10^{-1}	0
		30		21	30			3
		35		20	35			1
		40		23	40			2
		45		18	45			15
		50		21	50			21
Test 3	200000/ml	25	10^{-1}	3	25	200000/ml	10^{-1}	0
		30		15	30			0
		35		20	35			4
		40		22	40			2
		45		23	45			17
		50		28	50			20

The last question I tried to address in my thesis is the burst out time of bacteriophages depending on their genome size that is the regeneration time of a new generation of bacteriophages from parents. These results basically sets the core of this hypothesis because

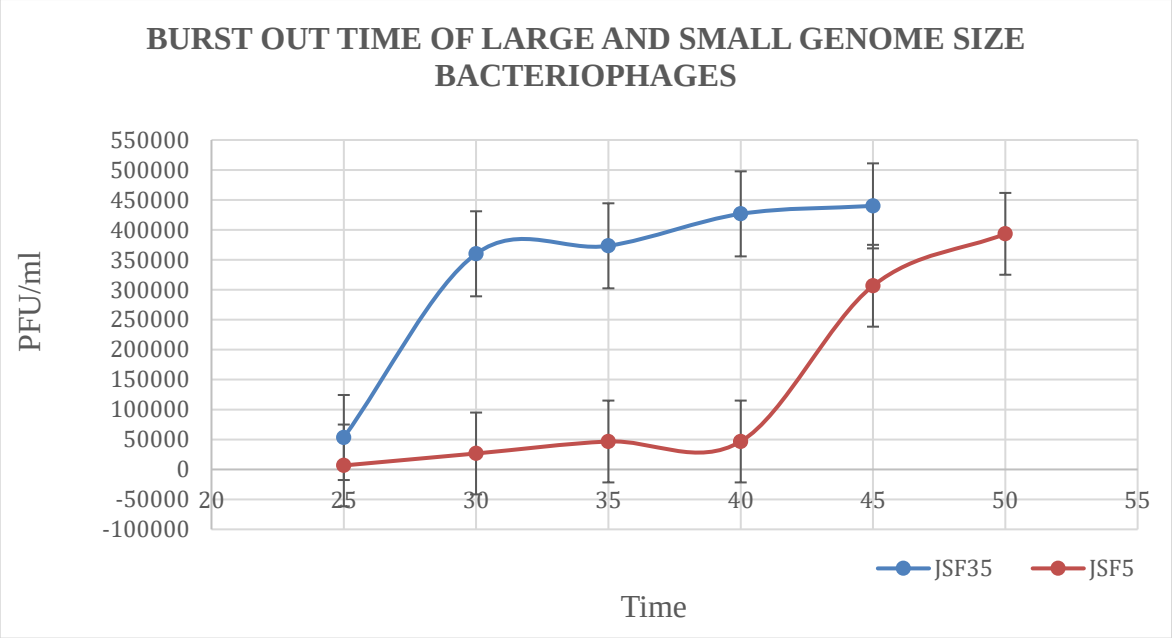

Figure 4.12: Burst out time of large and small genome size bacteriophages

this test was carried out using same amount (200000/ml) of both large and small genome containing bacteriophages and the results revealed the difference of their regeneration time. Bacteriophage titer was determined on every 5 minutes interval from 20 minutes to 45 minutes since adsorption time was found to be 20 minutes in our previous tests. Bacteriophage titer at 25 minute showed that concentration of JSF35 had become 360000/ml with average 18 PFU when 200000/ml bacteriophages were added at the initial phase of the study. Therefore, regeneration time of bacteriophage JSF35 was in between 20 minutes to 25 minutes. In the meantime, large genome containing bacteriophage JSF5 were found to regenerate in between 35 to 40 minutes since the concentration increased from 200000/ml to 305000/ml with average 16 PFU. Therefore, these results demonstrated that, there was around ten minute time difference for regeneration of new progenies depending on the genome size.

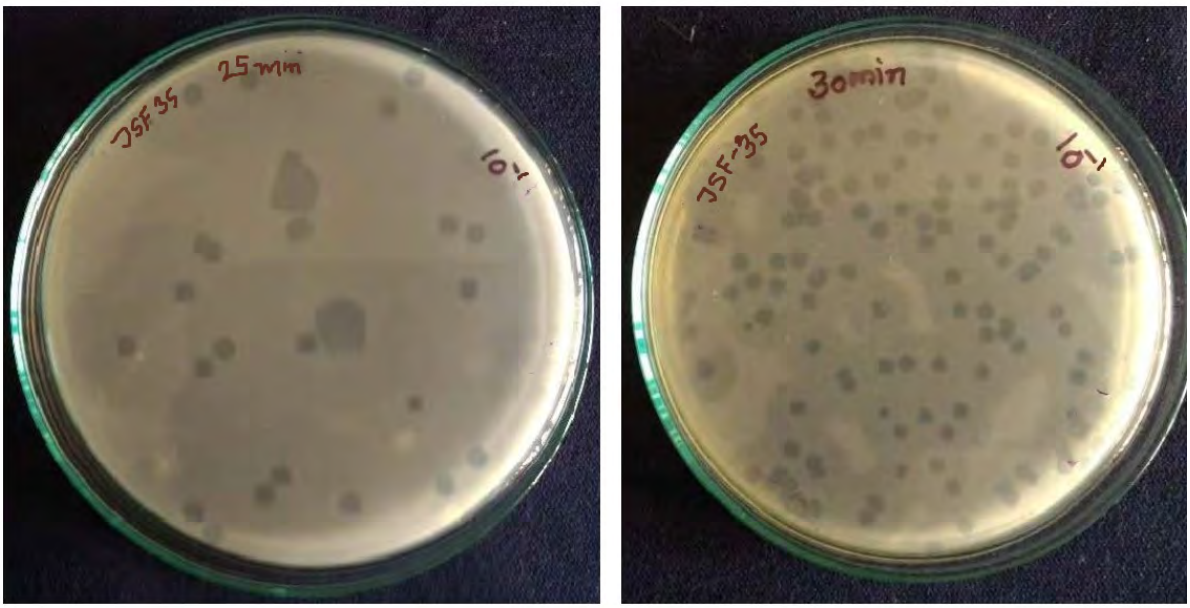

Figure 4.13: Burst out time of JSF 35

Chapter 5: Discussion

The prodigious molecule DNA forms genome, which is essential for building and maintenance of every single organism from single cell prokaryotes to million cells containing eukaryotes. This genome is encompassed of both coding and non-coding DNA that determines the characteristics of particular organism. Size of the genome is the total number of DNA base pairs and the size varies from species to species, for example, human genome is 3.2 billion base pairs while E. coli genome is 4.6 million base pairs (Blattner et al., 1997; Bodmer, 2013). Genome size varies depending on the complexity of the organisms but does this genome size have any direct effect on the regeneration time of organisms?

In this thesis, I have found depending on the genome size, bacteriophage regeneration time greatly varies and their increment number is highly influenced by the genome size within a specified time. First of all, when similar number of large and small genome containing bacteriophages JSF5 and JSF35 respectively were co-cultivated, the augmentation rate of small genome containing bacteriophages were three log higher than large genome containing bacteriophages. These results provided an insight about the effect of genome size. Then I found regeneration time of small genome containing bacteriophages were less than 25 minutes when this time was more than 35 minutes for large genome size bacteriophages that clearly demonstrated the effect of genome size on regeneration time or multiplication rate.

In support of substantiation, several other tests were carried out those also specified the consequence of genomic load on multiplication rate of bacteriophages. First of all, bacteriophages of similar genomic contents were co-cultivated to observe if any significant

changes took place between their multiplication rates. Large genome containing two bacteriophages JSF2 and JSF5 revealed that their multiplication rates were more or less similar and growth rate of JSF2 was 1.45 times higher than JSF5. On the other hand, small genome containing bacteriophages JSF27 and JSF35 also exhibited the similar results since JSF35 grown 1.17 times faster than JSF27. Considering no change of growth rate in logarithmic scale, I would like to conclude that, genome size affects the multiplication rate. Aimed at further evidence, I tried to find out if the adsorption pattern of the selected bacteriophages were different that actually affected growth rates. The results from adsorption pattern showed the evidence of 78% adsorption of small genome containing JSF35 while 84% large genome containing JSF5 were absorbed within the equal specified time period. As a result, I would like to state that, since both the bacteriophages were absorbed at same rate but their multiplication rate was greatly different, it might be due to their variation in genome size. Intended for more clarification, I cultured JSF5 and JSF35 separately to avoid competition for host bacteria in order to ascertain the growth rate changes between them. These experiments presented as similar results as previous tests since growth rate of JSF35 was 1260 times higher than JSF5, which is more than three log value difference. Last of all, I tried to figure out if other bacteriophages present in the environment which are not lytic to the specifics hosts, have any effect on the multiplication rates. In presence of non-affecting *E. coli* bacteriophage, the multiplication rate of JSF35 was 460 times higher than JSF5 which is almost 2.5 times in logarithmic scale. Therefore, bearing all the results in mind, I would like to conclude that there is clear evidence of bacteriophage multiplication rate difference based on their genome size.

However, since the concentrations were determined based on plaques formed at different dilution and their subsequent back calculation, it is not possible to declare the accurate final concentration of bacteriophages. When there is very few number of bacteriophages are present

in large volume of solution, there is higher chance of not getting accurate count. Since the accurate counts could not be determined, therefore I preferred to conclude the results based on the variation of growth rate in logarithmic scale.

Before I finish, this thesis sets off few other questions like, does genome size affect the growth rate evenhandedly or show biasness depending on the proportion of coding and non-coding sequences? What if we shorten the genome size of the organisms? Moreover, does this similar growth rate pattern followed in higher eukaryotes? Can we reduce the breeding time of domestic animals to increase their production rate in order to obtain higher profit by chopping down non-coding sequences from their genome? If we reduce the breeding time, how the morphological and physiological characteristics of those animals will be affected and how likely they are going to acquire mutation in their genome due to enormous stress conditions surrounding them? Further advance researches are required to answer all these questions and the results may lead to fruitful aftermath.

Chapter 6: References

Abedon, S.T., 1989. Selection for bacteriophage latent period length by bacterial density: A theoretical examination. Microb. Ecol. 18, 79–88. https://doi.org/10.1007/BF02030117

Abuladze, T., Li, M., Menetrez, M.Y., Dean, T., Senecal, A., Sulakvelidze, A., 2008. Bacteriophages reduce experimental contamination of hard surfaces, tomato, spinach, broccoli, and ground beef by Escherichia coli O157:H7. Appl. Environ. Microbiol. 74, 6230–6238. https://doi.org/10.1128/AEM.01465-08

Ackermann, H. W. (2003). Bacteriophage observations and evolution. Research in Microbiology. https://doi.org/10.1016/S0923-2508(03)00067-6

Ackermann, H., In, M., Part, A., Diego United, 2012. Bacteriophage Electron Microscopy.

Ackermann, H.-W., 2004. Bacteriophage Classification, in: Kutter, E., Sulakvelidze, A. (Eds.), Bacteriophages: Biology and Applications. CRC Press, pp. 68–90.

Ackermann, H.-W., 2007. 5500 Phages examined in the electron microscope. Arch. Virol. 152, 227–243. https://doi.org/10.1007/s00705-006-0849-1

Ackermann, H.-W., 2009. Phage Classification and Characterization, in: Bacteriophages, Methods in Molecular BiologyTM. Humana Press, pp. 127–140. https://doi.org/10.1007/978-1-60327-164- 6_13

Amgarten, D., Braga, L. P. P., da Silva, A. M., & Setubal, J. C. (2018). MARVEL, a tool for prediction of bacteriophage sequences in metagenomic bins. Frontiers in Genetics. https://doi.org/10.3389/fgene.2018.00304

Ashelford, K.E., Day, M.J., Fry, J.C., 2003. Elevated Abundance of Bacteriophage Infecting Bacteria in Soil. Appl. Environ. Microbiol. 69, 285–289. https://doi.org/10.1128/AEM.69.1.285- 289.2003

Blattner, F. R., Plunkett, G., Bloch, C. A., Perna, N. T., Burland, V., Riley, M., … Shao, Y. (1997). The complete genome sequence of Escherichia coli K-12. Science. https://doi.org/10.1126/science.277.5331.1453

Bodmer, W. (2013). Human Genome Project. In Brenner's Encyclopedia of Genetics: Second Edition. https://doi.org/10.1016/B978-0-12-374984-0.00746-4

Breitbart, M., Rohwer, F., 2005. Here a virus, there a virus, everywhere the same virus? Trends Microbiol. 13, 278–284. https://doi.org/10.1016/j.tim.2005.04.003

Carey-Smith, G.V., Billington, C., Cornelius, A.J., Hudson, J.A., Heinemann, J.A., 2006. Isolation and characterization of bacteriophages infecting Salmonella spp. FEMS Microbiol. Lett. 258, 182– 186. https://doi.org/10.1111/j.1574-6968.2006.00217.x

Casjens, S. R., & Thuman-Commike, P. A. (2011). Evolution of mosaically related tailed bacteriophage genomes seen through the lens of phage P22 virion assembly. Virology. https://doi.org/10.1016/j.virol.2010.12.046

Chibani-Chennoufi, S., Bruttin, A., Dillmann, M.-L., Brüssow, H., 2004. Phage-host interaction: an ecological perspective. J. Bacteriol. 186, 3677–3686. https://doi.org/10.1128/JB.186.12.3677- 3686.2004

Clark, J. R., & March, J. B. (2006). Bacteriophages and biotechnology: vaccines, gene therapy and antibacterials. Trends in Biotechnology. https://doi.org/10.1016/j.tibtech.2006.03.003

Clark, J.R., March, J.B., 2004. Bacterial viruses as human vaccines? Expert Rev. Vaccines 3, 463– 476. https://doi.org/10.1586/14760584.3.4.463

Erkus, O., De Jager, V. C. L., Spus, M., Van Alen-Boerrigter, I. J., Van Rijswijck, I. M. H., Hazelwood, L., … Smid, E. J. (2013). Multifactorial diversity sustains microbial community stability. ISME Journal. https://doi.org/10.1038/ismej.2013.108

Fang, Z., Tan, J., Wu, S., Li, M., Xu, C., Xie, Z., & Zhu, H. (2019). PPR-Meta: a tool for identifying phages and plasmids from metagenomic fragments using deep learning. GigaScience. https://doi.org/10.1093/gigascience/giz066

Fouts, D. E. (2006). Phage_Finder: Automated identification and classification of prophage regions in complete bacterial genome sequences. Nucleic Acids Research. https://doi.org/10.1093/nar/gkl732

Golomidova, A., Kulikov, E., Isaeva, A., Manykin, A., Letarov, A., 2007. The diversity of coliphages and coliforms in horse feces reveals a complex pattern of ecological interactions. Appl. Environ. Microbiol. 73, 5975–5981. https://doi.org/10.1128/AEM.01145-07

Guttman, B., Raya, R., Kutter, E., 2004. Basic Phage Biology, in: Kutter, E., Sulakvelidze, A. (Eds.), Bacteriophages: Biology and Applications. CRC Press, pp. 29–66.

Hagens, S., Loessner, M.J., 2007. Application of bacteriophages for detection and control of foodborne pathogens. Appl. Microbiol. Biotechnol. 76, 513–519. https://doi.org/10.1007/s00253- 007-1031-8

Hanlon, G.W., 2007. Bacteriophages: an appraisal of their role in the treatment of bacterial infections. Int. J. Antimicrob. Agents 30, 118–128. https://doi.org/10.1016/j.ijantimicag.2007.04.006

Hatfull, G. F. (2008). Bacteriophage genomics. Current Opinion in Microbiology. https://doi.org/10.1016/j.mib.2008.09.004

Hatfull, G. F., & Hendrix, R. W. (2011). Bacteriophages and their genomes. Current Opinion in Virology. https://doi.org/10.1016/j.coviro.2011.06.009

Hatfull, G. F., Pedulla, M. L., Jacobs-Sera, D., Cichon, P. M., Foley, A., Ford, M. E., … Hendrix, R. W. (2006). Exploring the mycobacteriophage metaproteome: Phage genomics as an educational platform. PLoS Genetics. https://doi.org/10.1371/journal.pgen.0020092

Haynes, M. (2013). Metagenomics. In Brenner's Encyclopedia of Genetics: Second Edition. https://doi.org/10.1016/B978-0-12-374984-0.00933-5

Juhala, R. J., Ford, M. E., Duda, R. L., Youlton, A., Hatfull, G. F., & Hendrix, R. W. (2000). Genomic sequences of bacteriophages HK97 and HK022: Pervasive genetic mosaicism in the

lambdoid bacteriophages. Journal of Molecular Biology. https://doi.org/10.1006/jmbi.2000.3729

Kang, D. D., Li, F., Kirton, E., Thomas, A., Egan, R., An, H., & Wang, Z. (2019). MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. PeerJ. https://doi.org/10.7717/peerj.7359

Keen, E. C., Bliskovsky, V. V., Malagon, F., Baker, J. D., Prince, J. S., Klaus, J. S., & Adhya, S. L. (2017). Novel "superspreader" bacteriophages promote horizontal gene transfer by transformation. MBio. https://doi.org/10.1128/mBio.02115-16

Koskella, B., & Brockhurst, M. A. (2014). Bacteria-phage coevolution as a driver of ecological and evolutionary processes in microbial communities. FEMS Microbiology Reviews. https://doi.org/10.1111/1574-6976.12072

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., … Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics. https://doi.org/10.1093/bioinformatics/btp352

Lima-Mendez, G., Van Helden, J., Toussaint, A., & Leplae, R. (2008). Prophinder: A computational tool for prophage prediction in prokaryotic genomes. Bioinformatics. https://doi.org/10.1093/bioinformatics/btn043

Little, J.W., 2005. Lysogeny, Prophage Induction, and Lysogenic Conversion, in: Waldor, M.K., Adhya, S.L., Friedman, D.I. (Eds.), Phages. American Society of Microbiology, pp. 37–54. https://doi.org/10.1128/9781555816506.ch3

Liu, J., Glazko, G., & Mushegian, A. (2006). Protein repertoire of double-stranded DNA bacteriophages. Virus Research. https://doi.org/10.1016/j.virusres.2006.01.015

Loeffler, J.M., Nelson, D., Fischetti, V.A., 2001. Rapid killing of Streptococcus pneumoniae with a bacteriophage cell wall hydrolase. Science 294, 2170–2172. https://doi.org/10.1126/science.1066869

Maranger, R., Bird, D.R., Juniper, S.K., 1994. Viral and bacterial dynamics in Arctic sea ice during the spring algal bloom near Resolute, N.W.T., Canada. Mar. Ecol. Prog. Ser. 111, 121–127. https://doi.org/http://dx.doi.org/10.3354/meps111121

Mavrich, T. N., & Hatfull, G. F. (2017). Bacteriophage evolution differs by host, lifestyle and genome. Nature Microbiology. https://doi.org/10.1038/nmicrobiol.2017.112

Nale, J. Y., Shan, J., Hickenbotham, P. T., Fawley, W. N., Wilcox, M. H., & Clokie, M. R. J. (2012). Diverse temperate bacteriophage carriage in Clostridium difficile 027 strains. PLoS ONE. https://doi.org/10.1371/journal.pone.0037263

Naser, I. Bin, Hoque, M. M., Nahid, M. A., Tareq, T. M., Rocky, M. K., & Faruque, S. M. (2017). Analysis of the CRISPR-Cas system in bacteriophages active on epidemic strains of Vibrio cholerae in Bangladesh. Scientific Reports. https://doi.org/10.1038/s41598-017-14839-2

O'Flynn, G., Ross, R.P., Fitzgerald, G.F., Coffey, A., 2004. Evaluation of a cocktail of three bacteriophages for biocontrol of Escherichia coli O157:H7. Appl. Environ. Microbiol. 70, 3417– 3424. https://doi.org/10.1128/AEM.70.6.3417-3424.2004

Orlova, E. V. (2012). Bacteriophages and Their Structural Organisation. In Bacteriophages. https://doi.org/10.5772/34642

Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., & Tyson, G. W. (2015). CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Research. https://doi.org/10.1101/gr.186072.114

Rodriguez-Valera, F., Martin-Cuadrado, A.-B., Rodriguez-Brito, B., Pasić, L., Thingstad, T.F., Rohwer, F., Mira, A., 2009. Explaining microbial population genomics through phage predation. Nat. Rev. Microbiol. 7, 828–836. https://doi.org/10.1038/nrmicro2235

Salmond, G.P.C., Fineran, P.C., 2015. A century of the phage: past, present and future. Nat. Rev. Microbiol. 13, 777. https://doi.org/10.1038/nrmicro3564

Schuch, R., Nelson, D., Fischetti, V.A., 2002. A bacteriolytic agent that detects and kills Bacillus anthracis. Nature 418, 884–889. https://doi.org/10.1038/nature01026

Serwer, P., Hayes, S. J., Thomas, J. A., & Hardies, S. C. (2007). Propagating the missing bacteriophages: A large bacteriophage in a new class. Virology Journal. https://doi.org/10.1186/1743-422X-4-21

Siringan, P., Connerton, P. L., Cummings, N. J., & Connerton, I. F. (2014). Alternative bacteriophage life cycles: The carrier state of campylobacter jejuni. Open Biology. https://doi.org/10.1098/rsob.130200

Smith, D.E., Tans, S.J., Smith, S.B., Grimes, S., Anderson, D.L., Bustamante, C., 2001. The bacteriophage straight phi29 portal motor can package DNA against a large internal force. Nature 413, 748–752. https://doi.org/10.1038/35099581

Summers, W., 2004. Bacteriophage Research: Early History, in: Kutter, E., Sulakvelidze, A. (Eds.), Bacteriophages: Biology and Applications. CRC Press, pp. 5–28

Summers, W.C., 1999. Félix d'Herelle and the origins of molecular biology. New Haven.

Suttle, C.A., 2005. Viruses in the sea. Nature 437, 356–361. https://doi.org/10.1038/nature04160

Van Valen, D., Wu, D., Chen, Y.-J., Tuson, H., Wiggins, P., Phillips, R., 2012. A Single-Molecule Hershey-Chase Experiment. Curr. Biol. CB 22, 1339–1343. https://doi.org/10.1016/j.cub.2012.05.023

Ventura, M., Sozzi, T., Turroni, F., Matteuzzi, D., van Sinderen, D., 2011. The impact of bacteriophages on probiotic bacteria and gut microbiota diversity. Genes Nutr. 6, 205–207. https://doi.org/10.1007/s12263-010-0188-4

Viazis, S., Akhtar, M., Feirtag, J., Brabban, A. d., Diez-Gonzalez, F., 2011. Isolation and characterization of lytic bacteriophages against enterohaemorrhagic Escherichia coli. J. Appl. Microbiol. 110, 1323–1331. https://doi.org/10.1111/j.1365-2672.2011.04989.x

Waldor, M.K., Mekalanos, J.J., 1996. Lysogenic Conversion by a Filamentous Phage Encoding Cholera Toxin. Science 272, 1910–1914. https://doi.org/10.1126/science.272.5270.1910

Weigel, C., & Seitz, H. (2006). Bacteriophage replication modules. FEMS Microbiology Reviews. https://doi.org/10.1111/j.1574-6976.2006.00015.x

Weinbauer, M.G., 2004. Ecology of prokaryotic viruses. FEMS Microbiol. Rev. 28, 127–181. https://doi.org/10.1016/j.femsre.2003.08.001

Williamson, S.J., McLaughlin, M.R., Paul, J.H., 2001. Interaction of the ΦHSIC Virus with Its Host: Lysogeny or Pseudolysogeny? Appl. Environ. Microbiol. 67, 1682–1688. https://doi.org/10.1128/AEM.67.4.1682-1688.200

Wommack, K.E., Colwell, R.R., 2000. Virioplankton: Viruses in Aquatic Ecosystems. Microbiol. Mol. Biol. Rev. 64, 69–114. https://doi.org/10.1128/MMBR.64.1.69-114.2000

Wommack, K.E., Ravel, J., Hill, R.T., Chun, J., Colwell, R.R., 1999. Population Dynamics of Chesapeake Bay Virioplankton: Total-Community Analysis by Pulsed-Field Gel Electrophoresis. Appl. Environ. Microbiol. 65, 231–240.

Appendixes

Appendix- I

List of bacteriophages used for candidate selection

Phage	Year of isolation	GenBank Accession no.	Genomic content
JSF1	2001	KY883636	126,082 base pair
JSF2	2001	KY883637	126,082 base pair
JSF4	2002	KY065147	124,261 base pair
JSF5	2002	KY883634	132,142 base pair
JSF6	2002	KY883635	133,685 base pair
JSF7	2005	KY065149	46,318 base pair
JSF10	2007	KY883654	111,671 base pair
JSF11	2007	KY883641	39341 base pair
JSF13	2009	KY883638	128,814 base pair
JSF17	2012	KY883640	125174 base pair
JSF18	2012	KY883650	38570 base pair
JSF20	2012	KY883651	39378 base pair
JSF23	2012	KY883657	50325 base pair
JSF24	2012	KY883652	39378 base pair
JSF25	2013	MF574151	38,593 base pair
JSF27	2013	KY883658	48690 base pair
JSF28	2013	KY883643	38489 base pair
JSF30	2013	KY883644	38056 base pair
JSF31	2014	KY883645	38581 base pair
JSF32	2014	KY883646	38581 base pair
JSF33	2014	KY883647	39661 base pair
JSF35	2015	KY883648	38688 base pair

Appendix – II

Metagenome projects list

PRJEB26733	PRJEB4347	PRJNA489338
PRJEB28724	PRJEB8767	PRJNA489560
PRJEB29849	PRJEB4347	PRJNA513235
PRJEB29847	PRJEB8767	PRJNA524695
PRJEB29848	PRJEB4347	PRJNA523640
PRJEB29849	PRJEB8767	PRJNA547647
PRJNA506462	PRJEB4347	PRJNA533517
PRJNA481090	PRJEB8767	PRJNA523640
PRJNA488089	PRJEB4347	PRJNA555586
PRJNA257236	PRJEB8767	PRJNA560061
PRJNA421889	PRJEB4347	PRJNA266489
PRJNA257236	PRJEB8767	PRJNA379596
PRJNA506462	PRJEB4347	PRJNA248702
PRJEB10139	PRJEB8767	PRJNA217803
PRJNA167559	PRJEB4347	PRJNA78957
PRJEB22508	PRJEB8767	PRJNA274005
PRJEB24372	PRJEB4347	PRJNA174998
PRJEB29852	PRJEB8767	PRJNA268262
PRJNA485385	PRJEB4347	PRJNA380388
PRJNA338198	PRJEB8767	PRJNA306614
PRJEB7657	PRJEB4347	PRJNA421062
PRJNA237728	PRJNA307600	PRJNA437180
PRJEB8767	PRJNA394609	PRJNA486378
PRJEB4347	PRJNA391102	PRJNA505081
PRJEB8767	PRJNA394609	PRJNA490743
PRJEB4347	PRJNA348753	PRJNA516046
PRJEB8767	PRJNA417962	PRJNA516047
PRJEB4347	PRJNA412344	PRJNA516048
PRJEB8767	PRJNA432305	PRJNA498669
PRJEB4347	PRJNA472208	PRJNA555145
PRJEB8767	PRJNA417962	PRJNA236731
PRJEB4347	PRJNA487675	PRJNA168652
PRJEB8767	PRJNA417962	PRJNA284264

PRJNA201032
PRJNA201035
PRJNA300507
PRJNA171225
PRJNA175412
PRJNA175409
PRJNA279928
PRJNA342010
PRJNA303138
PRJNA348753
PRJNA509280
PRJNA488702
PRJNA513496
PRJNA527183
PRJNA560163
PRJNA52839
PRJNA390162
PRJNA381836
PRJNA236437
PRJNA360856
PRJNA369606
PRJNA294351
PRJNA170978
PRJNA238427
PRJNA230432
PRJNA413549
PRJNA413656
PRJNA488156
PRJNA528176
PRJEB9229
PRJNA369263
PRJNA167262
PRJNA450161
PRJNA485372
PRJNA507471
PRJEB30151
PRJNA245794
PRJDB4394
PRJNA323729
PRJNA289671
PRJNA359498
PRJNA394949
PRJNA509633
PRJEB30150
PRJNA241481
PRJNA256899
PRJNA248703
PRJNA267225
PRJNA33371
PRJNA61401
PRJNA293769
PRJEB22550
PRJNA61255
PRJNA283243
PRJNA348753
PRJNA513496
PRJNA533514
PRJEB30149
PRJEB30156
PRJEB30155
PRJNA310790
PRJNA384071
PRJNA270383
PRJNA175407
PRJEB30152
PRJNA238420
PRJNA448460
PRJDB2871
PRJDB2872
PRJNA315143
PRJNA230567
PRJNA245054
PRJDB4022
PRJNA343219
PRJEB30153
PRJEB30154
PRJEB30170
PRJDB4019
PRJDB4021
PRJNA228949
PRJNA395025
PRJNA393817

Appendix – III

Instruments used

Autoclave	Wisd Laboratory Instruments Made in Korea
Electronic Balance	Model: WTB 200 RADWAG Wagi ELEktroniczne
Incubator	Model: DSI 3000 Digisystem Laboratory Instruments Inc. Made in Taiwan
Microcentrifuge	Model: MC-12 Benchmark Scientific
Refrigerated microcentrifuge	Model: ScanSpeed 1730R Labogene
Shaking Incubator	Model: JSSI-1000C JS RESEARCH INC. Made in Rep. of Korea
Spectrophotometer	Model: UVmini-1240 UV-VIS spectrophotometer SHIMADZU Corp
Syringe filter	MS® MCE Syringe Filter Membrane Solutions, LLC
Vortex Mixer	Model: VM-2000 Digisystem Laboratory Instruments Inc. Made in Taiwan
Water Bath WiseBath®	Wisd Laboratory Instruments DAIHAN Scientific Co., Made in Korea