MOLECULAR ASSESSMENT OF BACTERIOPHAGE GENOME AS A POTENTIAL THERAPEUTIC AGENT IN TREATING MULTI-DRUG RESISTANT BACTERIAL INFECTION

By

Mehnaz Tabassum

16136011

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

Department of Mathematics and Natural Sciences
BRAC University
December 2019

Declaration

It is hereby declared that

1. The thesis submitted is my/our own original work while completing a 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/We have acknowledged all main sources of help.

Student's Full Name & Signature:

_____

Mehnaz Tabassum
16136011

Approval

The thesis **Molecular Assessment of Bacteriophage genome as a potential therapeutic agent in treating multi-drug resistant bacterial infection** submitted by **Mehnaz Tabassum, 16136011** of Spring 2016 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of Bachelor of Science in Biotechnology.

Examining Committee:

Supervisor:
(Member)

_____
Tokee Mohammad Tareq, Lecturer
MNS Department, BRAC University

Program Coordinator:
(Member)

_____
Iftekhar Bin Naser, Assistant Professor
MNS Department, BRAC University

Departmental Head:
(Chair)

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

Abstract

Bacteriophage therapy sheds a fresh light upon the critical growth of antibiotic resistance. Antimicrobial resistance, one of the major public health concerns, requires a novel therapeutic strategy to be optimized. Phage therapy is being renovated in this regard. Few of the main challenges include shifting the dynamics of the microbiome, selecting the most adequate phages, integrating the clinical phages as prophage, probable immune responses and regulatory requirements. The major drawback lies within the phage genome. Though phages carry a plethora of genes that confer fitness advantages to the bacterial community, these genes are highly variable from phage to phage. The presence of fitness factors with the possibility of recombination or horizontal gene transfer complicates the application of phages. Here, we assess the potential of phage therapy from a clinical perspective. The analysis will include the possibility of using phages with no biosafety concern. Considering the isolation techniques, the number of phages per bacterial host vary a lot, which will also be considered for individual bacterial species level. Such evaluation will help us to primarily evaluate the potential of using existing phages as an alternative of treating infectious diseases with multiple drug resistant bacteria.

Keyword: Bacteria, Bacteriophage, Phage therapy, Anti-biotic Resistance, Drug-resistance

Dedication

To
the Counseling Unit of BRAC University

Acknowledgement

I acknowledge the blessings of being alive and willing, to perform my role as a student.

I would like to thank *Tokee Mohammad Tareq, Lecturer* for inducing genuine interest by sharing his work experience. I am grateful to *Dr. Iftekhar Bin Naser, Assistant Professor* for his time and patience to my academic needs. I gently remember Dr. *A F M Yusuf Haider, Professor* for his affectionate gesture during the work.

My gratitude belongs to my friends, Adree Rahman and Tamanna Siddique, who believed me.
I appreciate Tamanna Asgar Tusty for her openness and acceptance in last moments. I pleasantly recall Obaidullah Shoaib for his kindness in all forms. I am grateful to Marjina Apa, of 19th floor, who managed to keep an eye on my wellbeing, always.
I acknowledge the strangers who smiled at me.

Finally,

I am grateful to myself for holding the courage to stay one other day and see where it takes me.

Chapter 1

Introduction

1.1 Introduction

Viruses which infect bacteria are Bacteriophages. They are the most abundant biological entities on earth[1]. Back in 1917, Felix d'Herelle discovered the potential of phages and proposed for treatment and prevention of infectious diseases.[2] However, discovery of the first antibiotic interrupted its growth.[3] One of the biggest threats of health security is bacterial infectious diseases.[4] It is crucial to develop new approaches to the undeniable demand of antibacterials.[5]

Phages are being used to kill bacteria long before antibiotics has been discovered. In its premature years, phage therapy resulted in mixed success due to poor understanding on how phages work on bacterial. [6]

Last three decades has changed the scenario in terms of genetic features of phages.[7] The symbiotic relationship between phages and their bacterial hosts has given new insights. Phages are called phage morons since they carry genes that are not required for their own life cycle but provides fitness to bacterial hosts.[8]

The objective of this study is to explore phage therapy as a viable option.

---

[1] Hendrix et al., "Evolutionary Relationships among Diverse Bacteriophages and Prophages."
[2] Publications service, "On an Invisible Microbe Antagonistic toward Dysenteric Bacilli."
[3] Tan and Tatsumura, "Alexander Fleming (1881–1955)."
[4] "Effect of Antibiotic Prescribing in Primary Care on Antimicrobial Resistance in Individual Patients: Systematic Review and Meta-Analysis | The BMJ."
[5] Czaplewski et al., "Alternatives to Antibiotics—a Pipeline Portfolio Review."
[6] Publications service, "On an Invisible Microbe Antagonistic toward Dysenteric Bacilli"; Chanishvili, "Bacteriophages as Therapeutic and Prophylactic Means."
[7] Keen, "A Century of Phage Research."
[8] Taylor et al., "Chapter One - The Diverse Impacts of Phage Morons on Bacterial Fitness and Virulence."

Chapter 2

Materials and Methods

2.1 Creating the platform

The Linux operating system has been installed. It is an open source system and thus, most widely used platform of choice for scientific computing.[9]

Using linux terminal, the conda environment has been created that manages common dependencies for bioinformatic tools.[10] Among anaconda, conda and miniconda: miniconda has been chosen for its feasibility.

Tool	Function
UNIX	Platform
conda	Environment
prokka	Annotation
BRIG	Circular comparison
seqtk	Parses FASTA and FASTQ files
R Studio (ggplot2, dplyr, tidyr, tidyverse, qgraph, genoplotR)	Statistical Analysis
blast	Search

Table 2.1: List of platform, environment and tools

[9] "What Is Linux?"
[10] "Conda — Conda Documentation."

2.2 Phage Genome Assemblies

All available phage genome assemblies has been downloaded from NCBI.

2.3 Genome Annotation

Prokka is a tool for genome annotation.[11]It is applied on the downloaded phage genomes.

2.4 Few other tools

1. **Terminal**

```
mehnaz@UB21901PC-23:~/Phage/NCBI_PHAGES$ sh grep.sh
>AF503408.1 Enterobacteria phage P7, complete genome
>AY657002.1 Streptococcus phage phi1207.3, complete genome
>FO818745.1 Escherichia phage RCS47, complete genome
>KF030445.1 Escherichia phage 1720a-02, complete genome
>KT336320.1 Streptococcus phage phiNJ3, complete genome
>KT336321.1 Streptococcus phage phiSC070807, complete genome
>KT429160.1 Staphylococcus phage SPbeta-like, complete genome
>KU238067.1 Stx converting phage vB_EcoS_P27, complete genome
>KU238068.1 Stx converting phage vB_EcoS_P32, complete genome
>KU238069.1 Stx converting phage vB_EcoS_P22, complete genome
>KU238070.1 Stx converting phage vB_EcoS_ST2-8624, complete genome
>KU760857.1 Salmonella phage SJ46, complete genome
>KX077896.1 Streptococcus phage phiJH1301-2, complete genome
>KY065497.1 Streptococcus phage IPP61, complete genome
>MF172979.1 Erysipelothrix phage phi1605, complete genome
>MK448997.1 Streptococcus phage Javan630, complete genome
>NC_029119.1 Staphylococcus phage SPbeta-like, complete genome
>NC_031129.1 Salmonella phage SJ46, complete genome
>NC_042128.1 Escherichia phage RCS47, complete genome
```

Figure 2.1: Terminal View

[11] Seemann, “Prokka.”

2. R Studio[12]

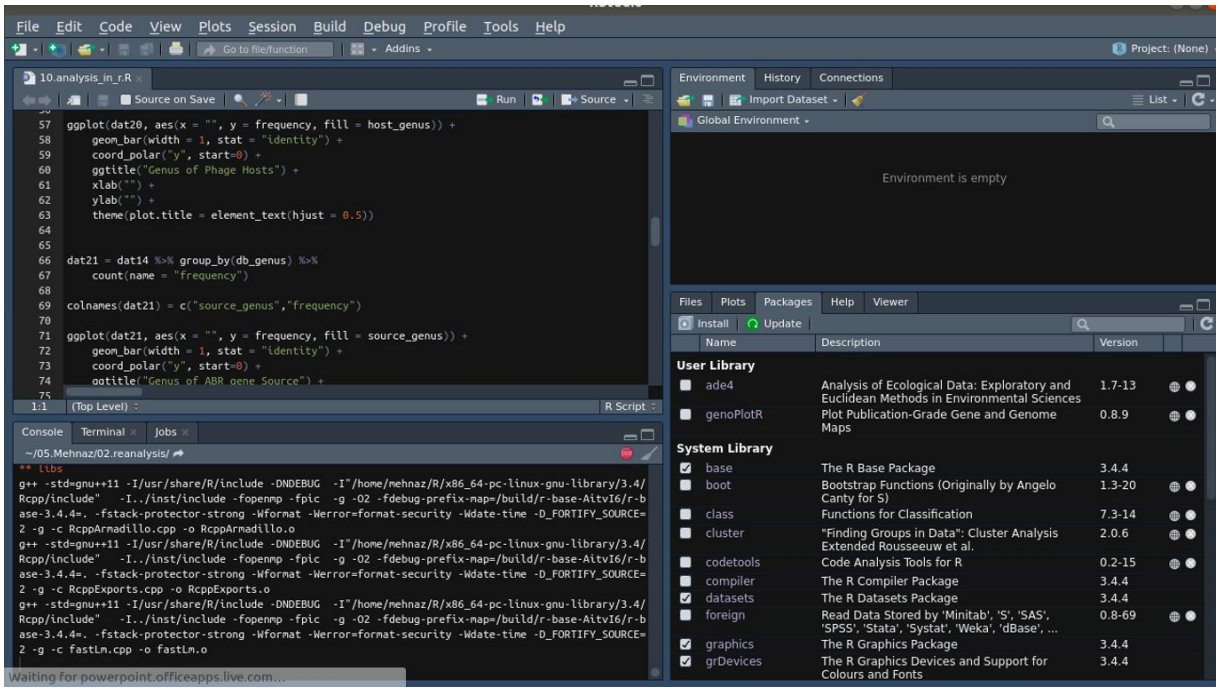

Figure 2.2: Inside R Studio

Graphical comparison and visualisation has been created using R Studio.

3. Roary

Roary has been used for pan genome analysis.[13]

4. iTol

iTol has been used to generate the phylogenetic tree.[14]

---

[12] "RStudio | Open Source & Professional Software for Data Science Teams."
[13] Page et al., "Roary."
[14] Letunic and Bork, "Interactive Tree Of Life (ITOL) V4."

Chapter 3

Result

1.1 Analytical Findings

All available phage genome sequences in NCBI database which counts around 9,000 phage genomes have been analyzed. Among these, 64 phages have been recognized as hazardous for phage therapy. This 64 phages carry antibiotic resistance genes and must be avoided for any clinical application.

It was observed in the phylogenetic tree that though the phage hosts are not closely related, these phages share similarity.

Synteny means finding similarity in linear comparison of the genomic organizations of phage genomes. A circular comparison has also been provided.

In section 1.5, the chart of phage host composition, Nitrincola comprises half of the population. Then comes Escherechia Coli. Lowest is the presence of Vibrio, Streptococcus and Staphylococcus.

The sources of Antibiotic resistance genes are presented in a chart of Section 1.6. It is visible that Klebsiella, Mycobacterium, Listeria, Nocardia with few other genus contributes half of it.

1.2 Phylogenetic Tree

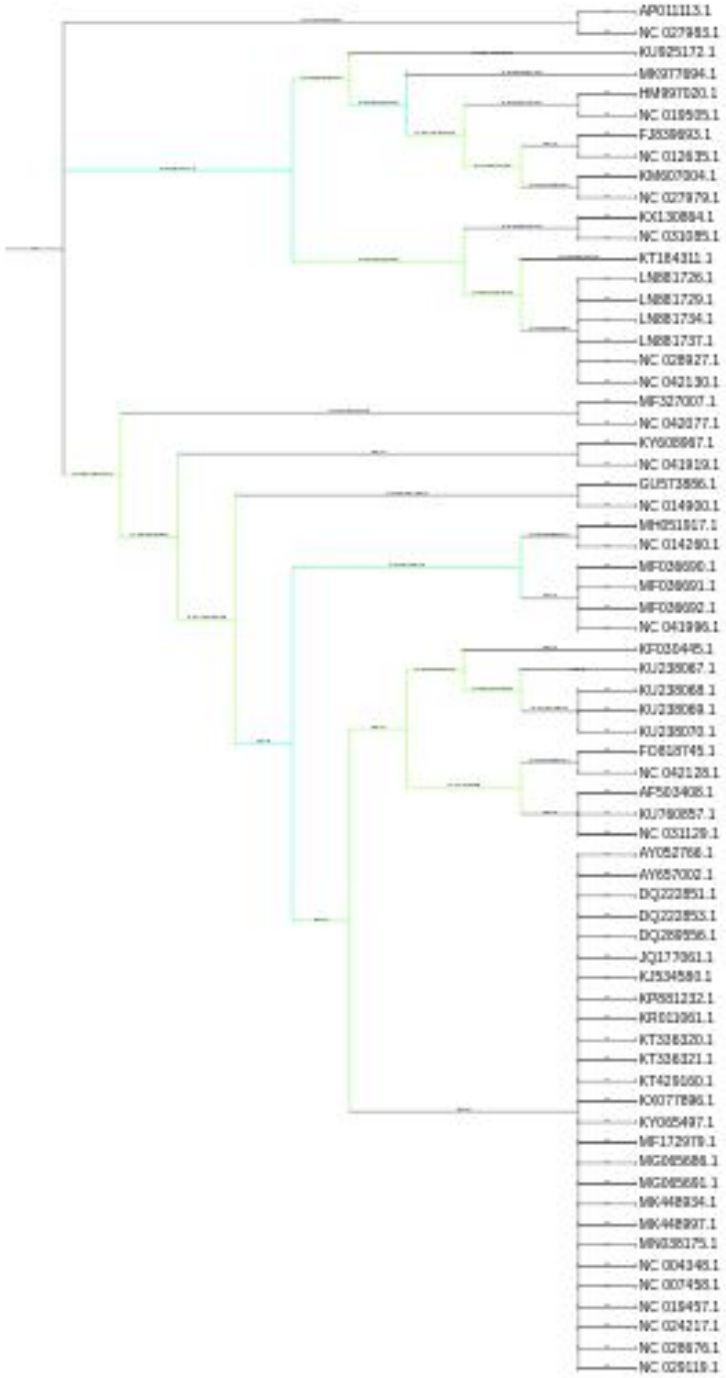

Figure 3.1: Phylogenetic Tree of phages carrying Antibiotic resistance gene/s

1.3 Synteny

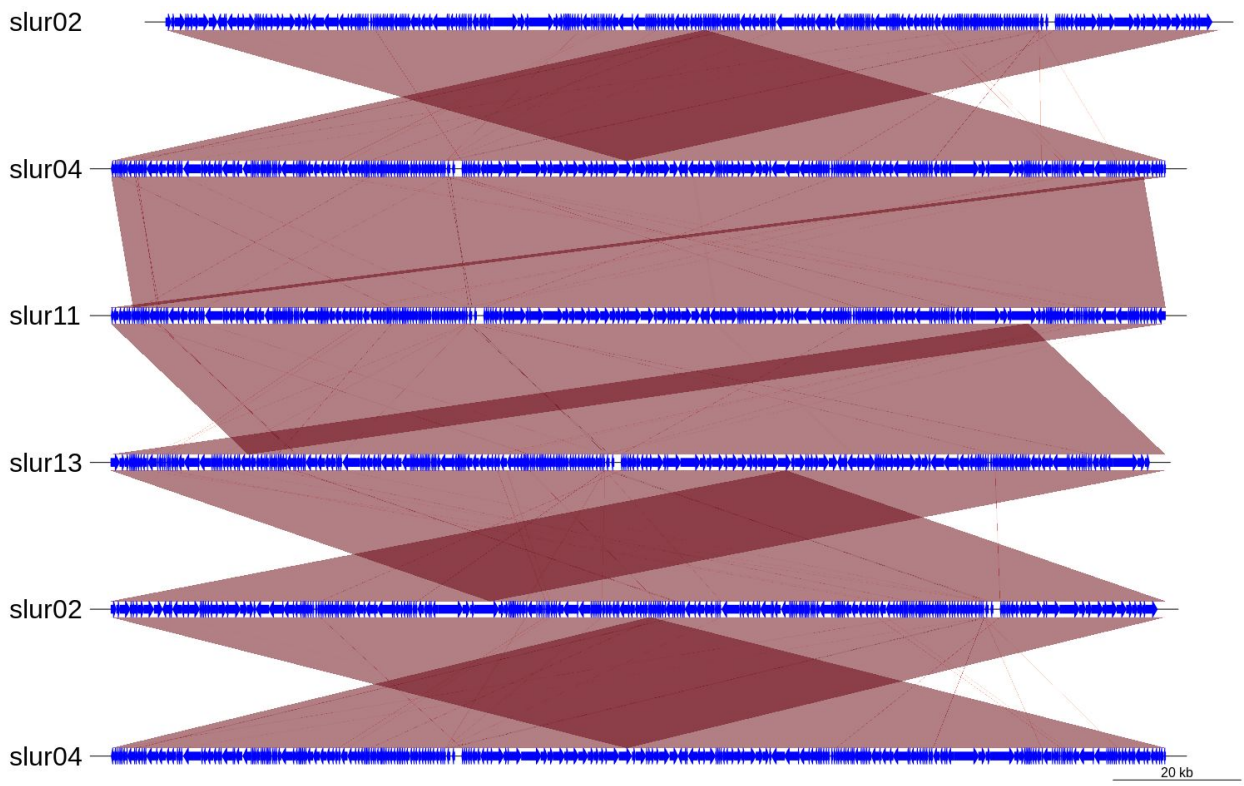

Figure 3.2: Phylogenetic Tree of 6 phages selected using the above tree that are closely related to each other

1.4 Ring

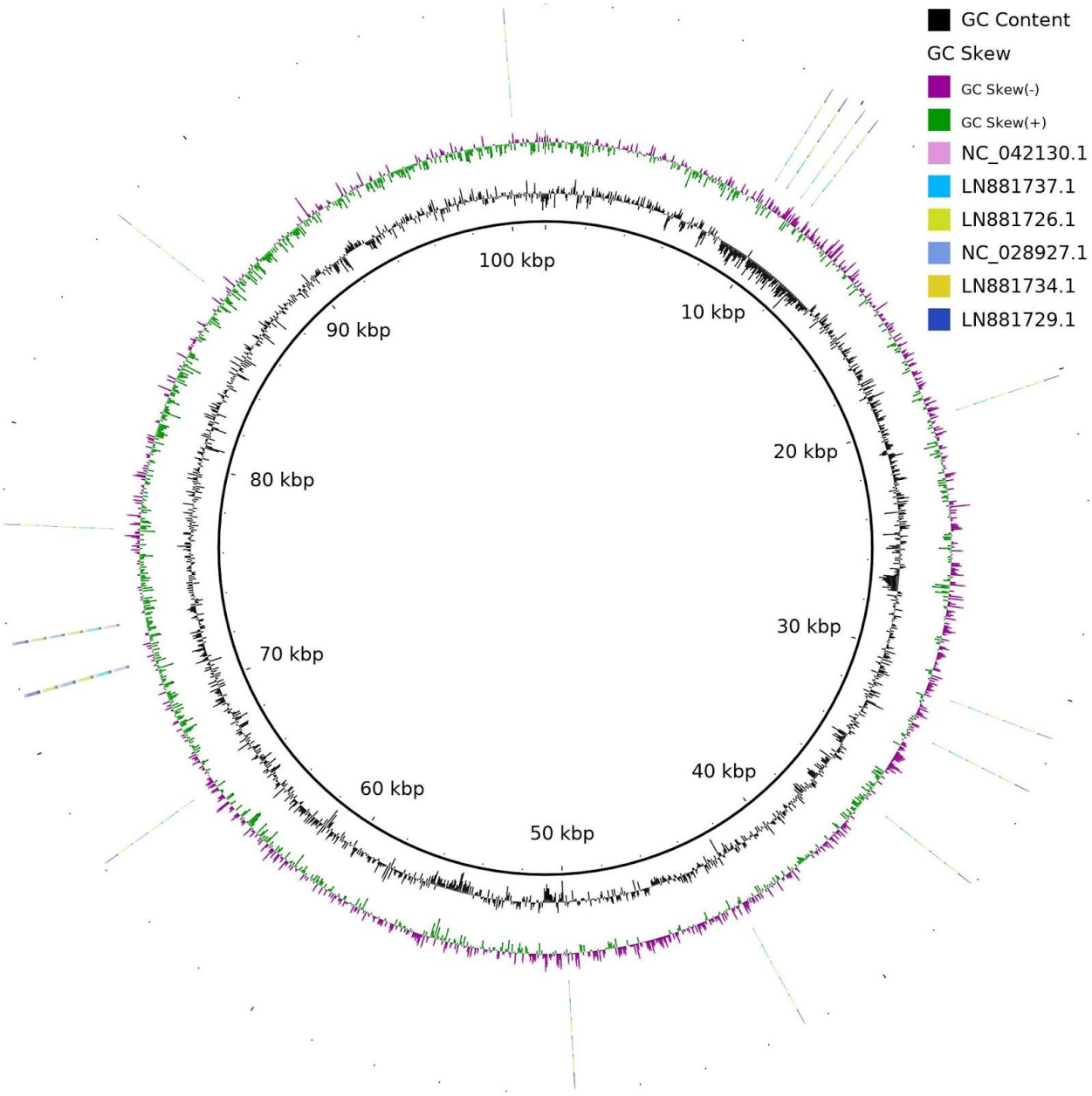

Figure 3.3: Comparison of those six phage genome sequences in circular form

1.5 Host Composition

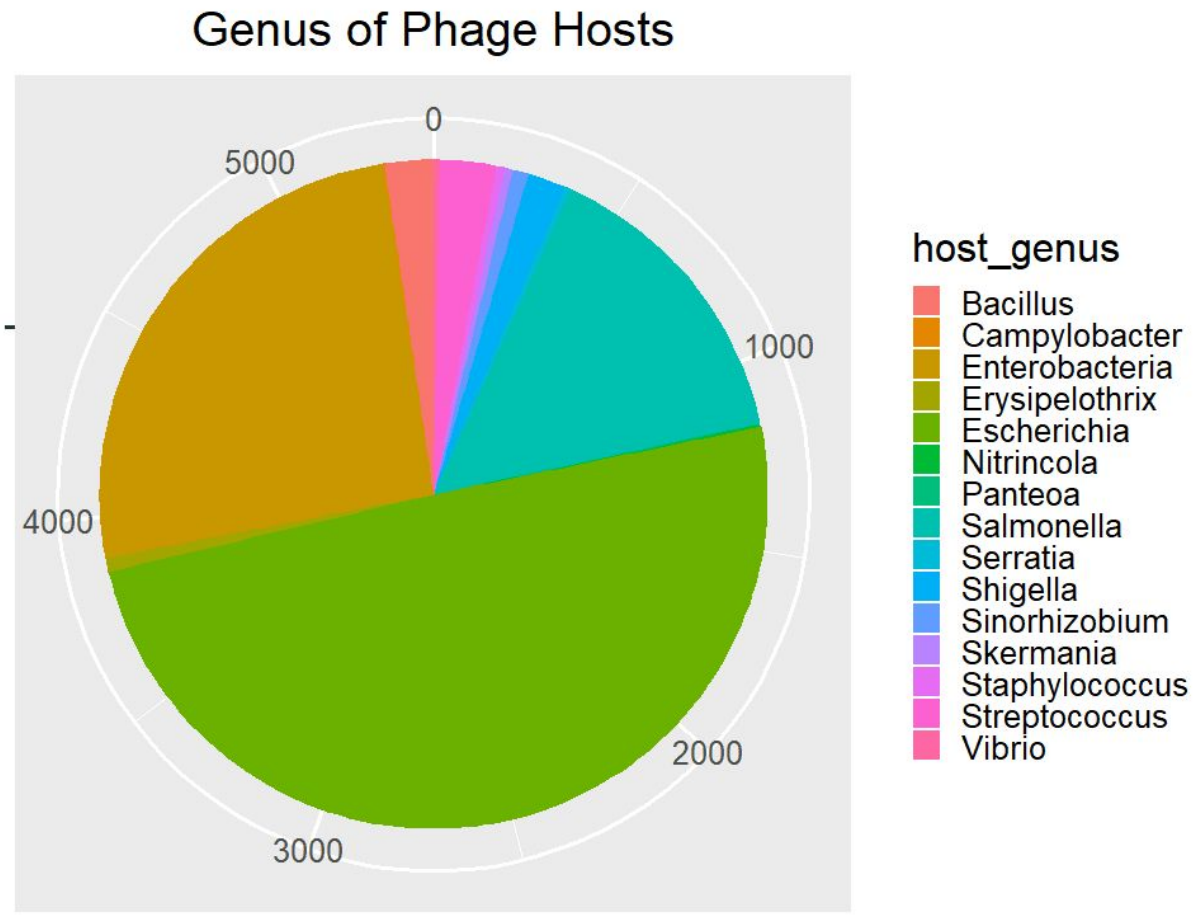

Figure 3.4: The composition of phage host at genus level

1.6 Contributing Sources

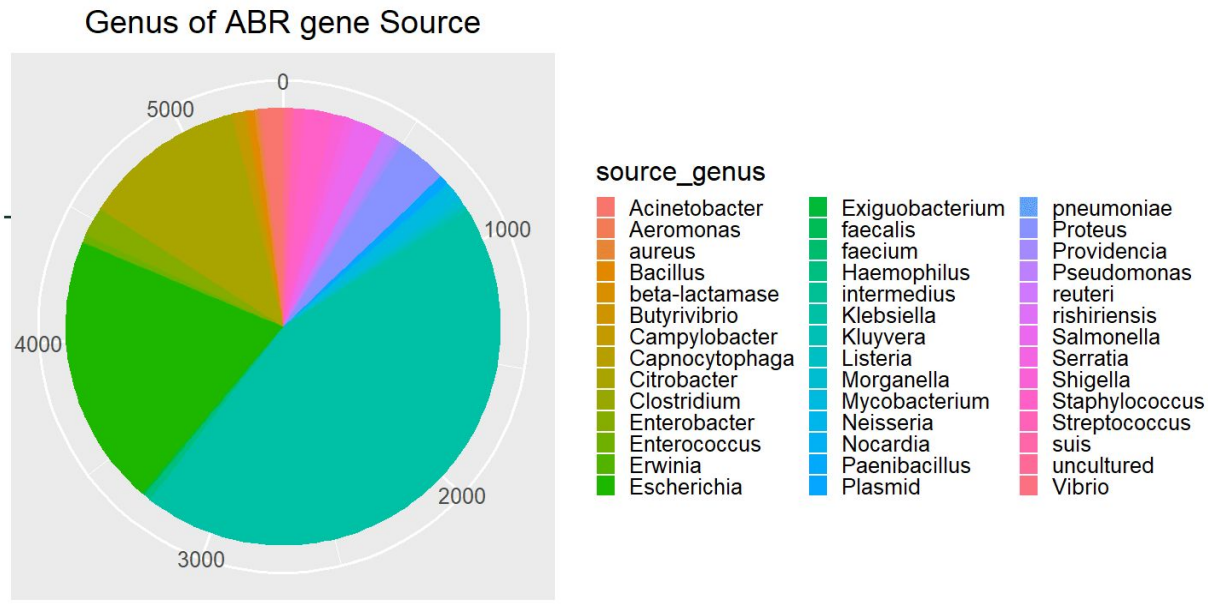

Figure 3.5: The composition of sources of Antibiotic resistant genes at genus level

1.7 Networking Beyond Border

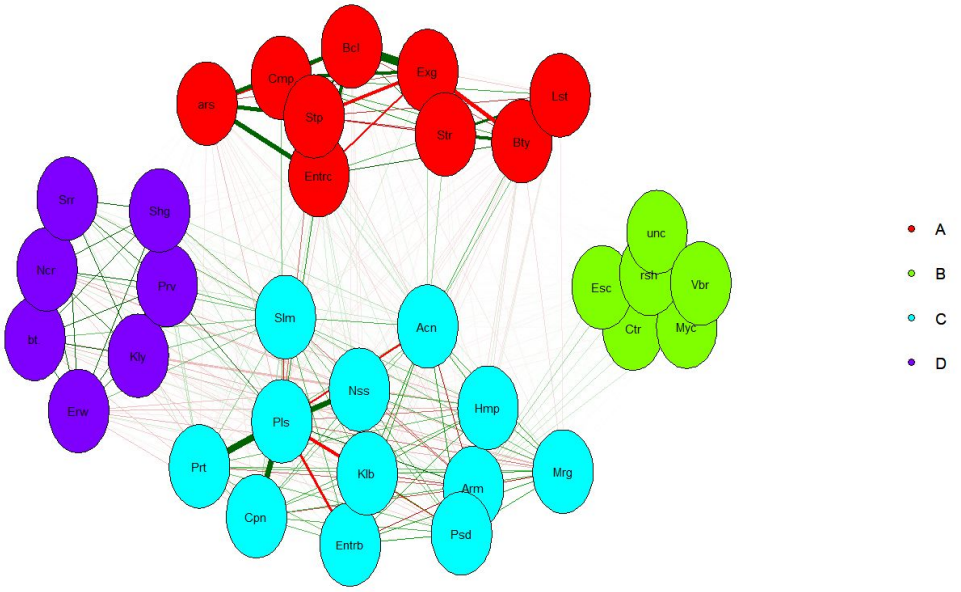

Figure 3.6: Group Network

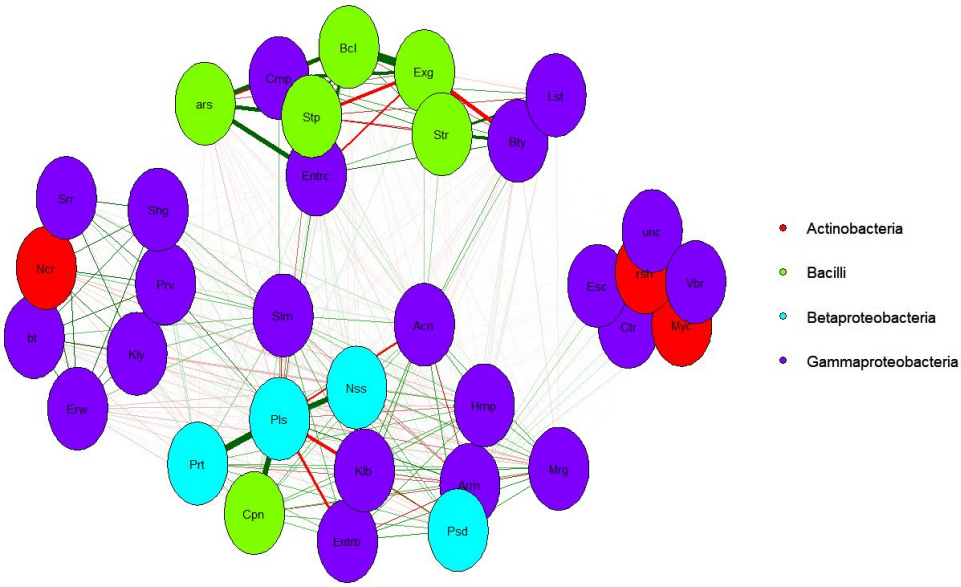

Figure 3.7: Class Network

Chapter 4

Discussion

The initial concern to start with is Antibiotic Resistance of infectious disease causing agents. Bacterial infections that are highly resistant to antibiotics are fatal. Each year, more than 2.8 million antibiotic-resistant infections occur in the U.S. according to the report of the Centers for Disease Control and Prevention(CDC).[15] As an alternative approach, bacteriophage focused treatment of infectious diseases, in other words, Phage Therapy is one of the top notch candidates. Under this light, molecular assessment has been performed in this study to figure out the pitfalls and dangers that comes with it. Consequently, we evaluated all the existing phage genomes in NCBI database to assess the possibility of spreading Antibiotic Resistance by phage therapy.

Analytical deduction confirms 64 phages carry antibiotic resistance genes. A phylogenetic tree has been created using these phages. The finding suggests, though the phage hosts are not closely related, phages share closeness and similarity. Next to that, 6 phages has been selected from the tree and both linear and circular comparison of the genomic sequence has been performed. The linear comparison, synteny represents mosaicism of phages. Different fragment of phages can interchange and results in genetic mosaicism. With the same 6 phages, a circular comparison is also performed. Correspondingly, we have checked source organisms of antibiotic resistant gene carrying phages along with the host organisms of phages. The clustering is done upon co-abundance. For instance, the phages that have analysed may come from the same host yet can acquire resistance from various genus. If horizontal gene transfer occurred in same phylum, same colors would cluster. However, the case is not true. Antibiotic resistance genes has been shared

[15] CDC, “The Biggest Antibiotic-Resistant Threats in the U.S.”

in and across phylum.

In this study, the possibility of phage carrying antibiotic resistance genes has been explored. The outcome is phage is gaining antibiotic resistance genes by suboptimal host switching. Phage is gaining antibiotic resistance genes other than corresponding genus. There are four major groups. Closely related genus should share more Antibiotic resistance compared to distinctively related genus. It has been observed that these are not equally distributed at class level. The acquisition of antibiotic resistance genes among phages and the pattern of spread among sources indicates suboptimal host selection as horizontal gene transfer occurrs in closely related species compared to distinctly related species.

References

CDC. “The Biggest Antibiotic-Resistant Threats in the U.S.” Centers for Disease Control and Prevention, November 14, 2019. https://www.cdc.gov/drugresistance/biggest-threats.html.

“Conda — Conda Documentation.” Accessed December 23, 2019. https://docs.conda.io/en/latest/.

Letunic, Ivica, and Peer Bork. “Interactive Tree Of Life (ITOL) v4: Recent Updates and New Developments.” *Nucleic Acids Research* 47, no. W1 (July 2, 2019): W256–59. https://doi.org/10.1093/nar/gkz239.

Page, Andrew J., Carla A. Cummins, Martin Hunt, Vanessa K. Wong, Sandra Reuter, Matthew T. G. Holden, Maria Fookes, Daniel Falush, Jacqueline A. Keane, and Julian Parkhill. “Roary: Rapid Large-Scale Prokaryote Pan Genome Analysis.” *Bioinformatics* 31, no. 22 (November 15, 2015): 3691–93. https://doi.org/10.1093/bioinformatics/btv421.

“RStudio | Open Source & Professional Software for Data Science Teams.” Accessed December 23, 2019. https://rstudio.com/.

Seemann, Torsten. “Prokka: Rapid Prokaryotic Genome Annotation.” *Bioinformatics (Oxford, England)* 30, no. 14 (July 15, 2014): 2068–69. https://doi.org/10.1093/bioinformatics/btu153.

Linux.com. “What Is Linux?” Accessed December 23, 2019. https://www.linux.com/what-is-linux/.