Safety Assessment of *Citrobacter freundii* Bacteriophages for Use in Phage Therapy in the Face of Antibiotic Resistance

By

Sayema Syeda Hossain
13226005

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 Microbiology

Department of Mathematics and Natural Sciences
Brac University
February 2020

© 2020. Brac University
All rights reserved.

Declaration

It is hereby declared that

1. The thesis submitted is my own original work while completing my 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 that has been accepted or submitted, for any other degree or diploma at a university or other institution.
4. I have acknowledged all of the main sources of help.

Student's Full Name & Signature:

____

Sayema Syeda Hossain
13226005

Approval

The thesis/project titled "Safety Assessment of *Citrobacter freundii* Bacteriophages for Use in Phage Therapy in the Face of Antibiotic Resistance" submitted by

1. Sayema Syeda Hossain (13226005)

of Spring, 2020 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of BS in Microbiology on 27.02.2020.

Examining Committee:

Supervisor: ________________________________

Professor Mahboob Hossain, PhD
Professor, Department of Mathematics and Natural Sciences
Brac University

Program Coordinator: ________________________________

Mahbubul Hasan Siddiqee
Senior Lecturer, Department of Mathematics and Natural Sciences
Brac University

Departmental Head: ________________________________

Professor A F M Yusuf Haider, PhD
Chairperson, Department of Mathematics and Natural Sciences
Brac University

Ethics Statement

Abstract

Citrobacter freundii is a species of bacteria in the Enterobacteriaceae family that is known to cause hospital-acquired infections in the respiratory tract, urinary tract, blood, and other normally sterile sites in immunocompromised patients and is growing resistant to multiple antibiotics. In the current state of growing antibiotic resistance, phage therapy is one of the alternatives being explored. Two of the biggest concerns regarding phage therapy are the virus' ability to transfer antibiotic resistance or virulence genes through phage morons. After comparison using Linux based BLAST between complete *Citrobacter* phage genomes downloaded from the NCBI website and the Comprehensive Antibiotic Resistance Database and the Virulence Factor Database, it was determined that none of the 21 phages showed signs of carrying virulence or antibiotic resistance genes. Through phylogenetic tree construction using Linux based Roary and web-based ITOL, it was determined that there is a possibility of lysogeny in the phages.

Keywords: Bacteriophages; Phage Therapy; Antibiotic Resistance; Virulence Factors; *Citrobacter freundii*

Dedication

To *Abbu*, *Ammu,* and *Bon.*

Acknowledgement

I would like to thank Chairperson Professor A F M Yusuf Haider, Ph.D. and Microbiology Program Coordinator Mahbubul Hasan Siddiqee for their kind cooperation throughout this study.

I wish to wholeheartedly express my gratitude and sincere indebtedness to my thesis supervisor Professor Mahboob Hossain, Ph.D. for providing me with guidance, encouragement, generous help and unfailing enthusiasm at every stage of this study, and every stage of my education at this university. It has been a great pleasure for me to work with him.

I would also like to extend my gratitude to Tokee Mohammad Tareq for introducing me to and helping me navigate the wonderful world of bioinformatics.

My heartfelt thanks to the Department of Mathematics and Natural Sciences at Brac University for providing me with the opportunity and facilities to complete my thesis.

A special thanks to my father without whom I could not have finished this thesis, to my mother for her love and care of me through thick and thin, and to my sister for her support and encouragement.

Finally, I wish to thank my friends for making me believe I could when I thought I could not.

Table of Contents

Declaration ii

Approval iii

Ethics Statement iv

Abstract v

Dedication vi

Acknowledgement vii

Table of Contents viii

List of Tables xi

List of Figures xii

List of Acronyms xiv

Chapter 1 Introduction 1

1.1 Antibiotic Resistance 1

1.2 Reasons behind Antibiotic Resistance 3

1.3 The Burden of Resistant Infections 4

1.4 Possible Remedies and Way Forward 5

1.4.1 Generic Approaches 5

1.4.2 The Environmental Antibiotic Resistome 6

1.4.3 Use of Bacterial Virulence 7

1.4.4 Recruiting Innate Immune System 7

1.4.5 Phage therapy 7

1.5 Bacteriophage 8

1.5.1 Types of phages 8

1.5.2 The lytic cycle 9

1.5.3 Bacteriophage therapy 9

1.5.4 Advantages and disadvantages 10

1.6 *Citrobacter freundii* (*C. freundii*) 11

1.6.1 Scientific Classification of *C. freundii* 11

1.6.2 Common sources of *C. freundii* 13

1.6.3 Multidrug-resistance of *C. freundii* 13

1.7 Objectives of the Study 15

1.8 Organization of the Thesis 15

Chapter 2 Data and Methodology 16

2.1 Data used for the study 17

2.2 Basic Concepts Relating To the Methodology 18

2.2.1 Antibiotic Resistance 18

2.2.2 Virulence Factors 21

2.2.3 Synteny 21

2.2.4 Phylogenetics 23

2.2.5 Lysogeny 24

2.3 Tools and Algorithms 26

2.3.1 The Basic Local Alignment Search Tool (BLAST) 26

2.3.2 The Comprehensive Antibiotic Resistance Database (CARD) 26

2.3.3 The Virulence Factor Database (VFDB).. 27

2.4 Conceptual flow of the methodologies 28

2.4.1 Isolating Full Genome Sequences from NCBI Using BLAST......................... 28

2.4.2 Annotation of Sequence FASTA using Prokka.. 28

2.4.3 Phylogenetic Analysis with Roary.. 28

2.4.4 Phylogenetic Tree from iTOL 28

2.4.6 Synteny and Genomic Organization Using genoPlotR and BLAST 29

2.4.7 Checking for Antibiotic Resistance Genes Using CARD and BLAST 29

2.4.8 Checking for Virulence Factors Using VFDB and BLAST............................. 29

2.4.8 Checking for Lysogeny Using BLAST, Prokka, PhiSpy, Roary and iTOL 29

Chapter 3 Results and Discussion.. 31

3.1 Complete genomes of Unique Citrobacter Phages 32

3.2 Annotation of sequence FASTA from NCBI 33

3.3 The Phylogenetic tree 33

3.4 Synteny and Genomic Organization.. 34

3.5 Checking for Antibiotic Resistance Genes in the Citrobacter Phage 40

3.6 Checking for virulence genes in the phage DNA 45

3.7 Checking phages for possible lysogenic property 47

Chapter 4 Conclusions 49

References.. 50

List of Tables

Table 1 Taxonomy of Citrobacter freundii, Source: Werkman and Gillen (1932) 12

Table 2 Genome sizes of 21 Citrobacter phages 32

Table 3 Matches between Citrobacter phage DNA and CARD with 100% identity............. 41

Table 4 Matches between Citrobacter phage DNA and VFDB with > 90% identity............. 46

List of Figures

Figure 1 Events in the Age of Antibiotic 2

Figure 2 Deaths per year (estimated for 2050) attributable to antimicrobial resistance compared to other major causes of death.. 4

Figure 3 Microscopic image of *Citrobacter freundii*,. 12

Figure 4 Percentage of isolated *Citrobacter spp.* resistant to different antibiotics................ 14

Figure 5 Ways Antibiotics Affect Bacterial Cells. 19

Figure 6 Mechanisms of Antibiotic Resistance. 19

Figure 7 How antibiotic resistance develops. 20

Figure 8 Chromosomal rearrangement processes leading to disruption of synteny. 22

Figure 9 Terminology frequently used in phylogenetic trees. 24

Figure 10 The lysogenic and lytic cycles of bacteriophage. 25

Figure 11 Logical framework of the study 30

Figure 12 Phylogenetic tree of Citrobacter phages produced using iTOL 33

Figure 13 Synteny diagrams for Citrobacter phages SH3, SH4, and SH5 34

Figure 14 Synteny diagrams for Citrobacter phages SH1, SH2, and phiCFP-1 35

Figure 15 Synteny diagrams for Citrobacter phages CF1-ERZ-2017, Merlin, and Moon 36

Figure 16 Synteny diagrams for Citrobacter phages IME-CF2, Miller, Margaery, and Maroon 37

Figure 17 Synteny diagrams for Citrobacter phages CF1-DK-2017, Sazh, and Stevie 38

Figure 18 Synteny diagrams for Citrobacter phages Mijalis, Moogle, Mordin, Maleficent, and Michonne 39

Figure 19 Histogram of the percentage identity in matches between the Citrobacter phage genomes and CARD 40

Figure 20 Histogram of the percentage identity in matches between the Citrobacter phage genomes and VFDB 45

Figure 21 Phylogenetic tree of the combined set of both Citrobacter bacteriophages and prophages.. 47

List of Acronyms

AMR	Antimicrobial resistance
ARO	Antibiotic Resistance Ontology
BLAST	The Basic Local Alignment Search Tool
C. freundii	*Citrobacter freundii*
CARD	Comprehensive Antibiotic Resistance Database
DNA	Deoxyribonucleic acid
ESBL	Extended-spectrum beta-lactamase
FASTA	FAST-All
HGT	Horizontal Gene Transfer
iTOL	Interactive Tree Of Life
NCBI	National Center for Biotechnology Information
RGI	Resistance Gene Identifier
RNA	Ribonucleic acid
SNP	Single-Nucleotide Polymorphism
UTI	Urinary Tract Infections
VFDB	Virulence Factor Database

Chapter 1

Introduction

The management of microbial infections has been historically well-documented starting from the time of ancient Egypt, Greece, and China. It was Sir Alexander Fleming in 1928, who has originated the modern era of antibiotics by discovering penicillin. Antibiotics, since then, have transformed modern medicine and helped save millions of lives and to extend expected life spans of millions more. Many of the major advances in medicine and surgery were successful due to the role of antibiotics in changing the outcome of bacterial infections. Recently, the efficacy of antibiotics has been threatened by the rapid emergence of resistant bacteria occurring worldwide.

1.1 Antibiotic Resistance

Antibiotic resistance occurs when infecting bacteria change in response to the use of these medicines. It is the bacteria, not humans or animals, which become antibiotic-resistant. The infections the antibiotic-resistant bacteria cause are harder to treat than those caused by non-resistant bacteria. Bacteria can inherit genes or take them from other bacteria through mobile genetic material such as plasmids. This horizontal gene transfer (HGT) can transfer antibiotic resistance among different species of bacteria. Selection pressure from antibiotic use and mutations in microbes and cause resistance to bacterial changes. Suboptimal antibiotic dose helps to select step by step resistance. Chromosomal and increasingly transmissible extrachromosomal components give birth to resistance genes.

The overuse and misuse of antibiotics and lack of new drug development are often blamed for the antibiotic resistance issue, however, historically, it has evolved from several dimensions. Since the introduction of sulfonamides in 1937, specific mechanisms of resistance have developed and afflicted their therapeutic use. While the resistance reported originally was in

the late 1930s, the same mechanisms operate some 70 years later. From the late 1960s through the early 1980s, the pharmaceutical industry introduced many new antibiotics to solve the resistance problem, but a similar course of events has ensued. As a result, even at present, many decades after the first patients were treated with antibiotics, bacterial infections have again become a threat. Many of the bacterial pathogens associated with epidemics of human disease have evolved multidrug-resistant (MDR) forms subsequent to antibiotic use. For example, MDR *M. tuberculosis* is a major pathogen found in both developing and industrialized nations and became the 20th-century version of an old pathogen. Figure 1 shows the sequence of discovery and resistance development for the major classes of antibiotics.

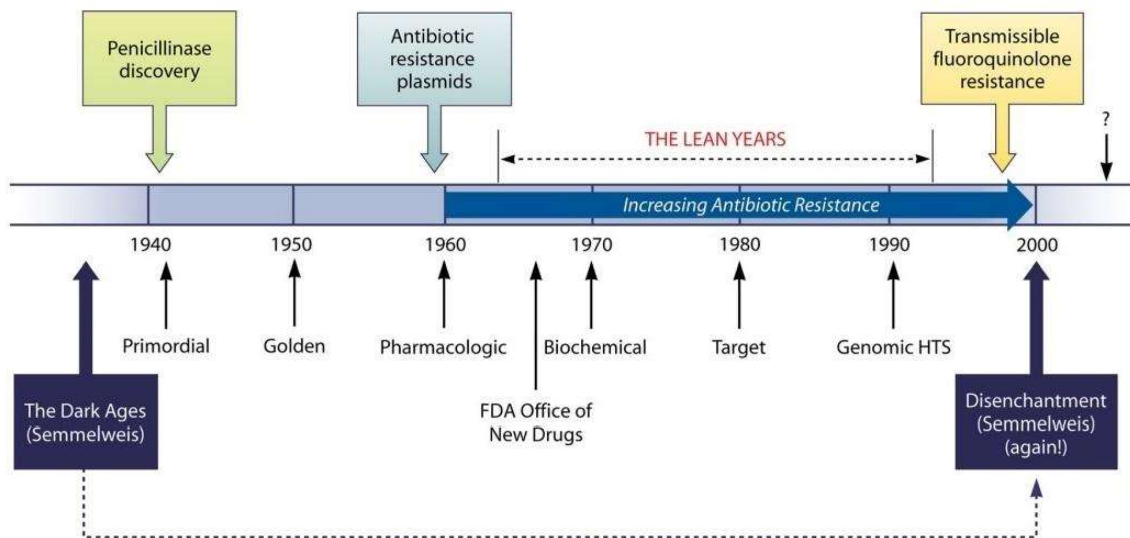


Figure 1 Events in the Age of Antibiotic

1.2 Reasons behind Antibiotic Resistance

The overuse of antibiotics is attributed to the evolution of resistance. A direct relationship between antibiotic consumption and the emergence of resistant bacteria strains have been demonstrated in many epidemiological studies.

For more than half a century, antibiotics have been used in livestock and poultry to control and treat diseases and to promote growth and improve the production of animal products (Stokstad and Jukes, 1949; Page and Gautier, 2012). Antibiotic-resistant bacteria have also been developing as a consequence of this phenomenon (Lu, 2004). Animal-to-human transfers of resistance genes have been confirmed in recent studies using whole genome sequencing (Harrison et al., 2013). The life cycle of the resistant antibiotic in human goes in the following three steps (Laxminarayan et al., 2013)

1. Antibiotic use in food-producing animals kills or suppresses susceptible bacteria, allowing antibiotic-resistant bacteria to thrive;
2. Resistant bacteria are transmitted to humans through the food supply;
3. These bacteria can cause infections in humans that may lead to adverse health consequences.

1.3 The Burden of Resistant Infections

Antibiotic resistance is posing a global burden that may be concentrated in three major categories (D'Costa, et al., 2006; DeLeo and Chambers, 2009):

1. Longer duration of illness and higher rates of mortality in patients with resistant infections,
2. Increasing costs of treatment (longer hospitalisation and costlier medication) for resistant infections, and
3. Inability to do procedures that rely on effective antibiotics to prevent infection.

Globally, a forecasted number of deaths per year estimated for 2050, shows that about 10 million deaths will be attributable to antimicrobial resistance which would be way higher than the number deaths due to Cancer (O'Neill J., 2016). A comparative view of deaths by different major causes (estimated for 2050) is given in Figure 2. The same article reported the economic impact for the US alone as 20 billion USD in excess costs each year.

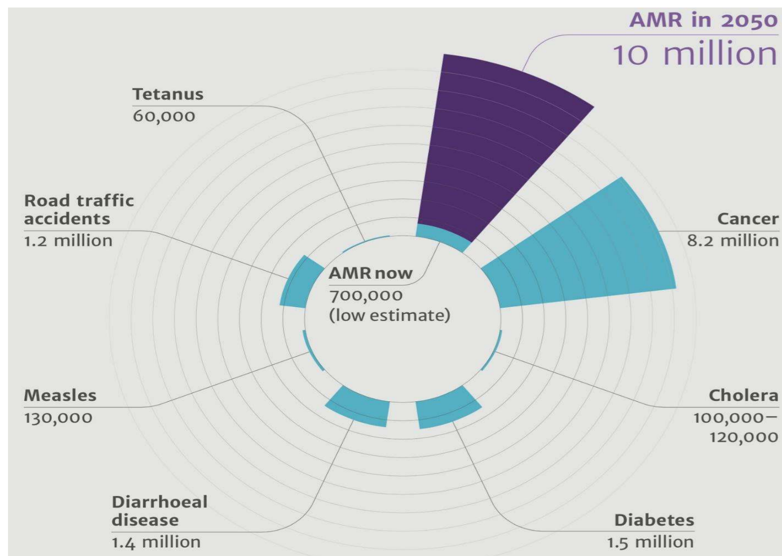


Figure 2 Deaths per year (estimated for 2050) attributable to antimicrobial resistance compared to other major causes of death. Reprinted from AMR Review, by J. O'Neill, 2016, Retrieved from: https://amr-review.org/sites/default/files/160525_Final%20paper_with%20cover.pdf.

1.4 Possible Remedies and Way Forward

While there is a perception of the unavailability of newer antibacterial drugs, many pioneering analyses indicate that many potential drug targets remain to be exploited in antimicrobial discovery. Predicting resistance reliably—and acting appropriately—would be a valuable approach to extending antibiotic lifetimes (Davies and Davies, 2010). The approaches to combat the resistance including multi-drug resistance is a subject of research all along. Some of them (Jassim and Limoges, 2014) are discussed in the following:

1.4.1 Generic Approaches

Some behavioral change, policy, and practices may help reduce antibiotic resistance, for example,

a. National approaches and commitment: Many countries have implemented comprehensive national strategies on producing, prescribing and use of antibiotics for controlling resistance. Some of them have been very successful in this regard.
b. Rational antibiotic use in hospitals and the community: A comprehensive antibiotic stewardship at a global level where set activities and policies to improve the rational use of antibiotics should be advocated. A programme on rational antibiotic use or antibiotic stewardship in the community should cover a wide range of settings, such as ambulatory care facilities, pharmacies, etc.
c. Education and changing social norms: Irrational use of antibiotics becomes a social norm as it happens repeatedly among the public and health professionals. Education and awareness at a mass level can help to mend such norms.
d. Infection control: Prevention is better than treatment, and from a resistance perspective, prevention of infection reduces antibiotic use and the spread of resistant bacteria.

Although prevention may be one of the strategies to control resistance, it must be complemented by the controlled use of antibiotics.

e. Role of diagnostics: Efforts to improve microbiological laboratories are underway. The speed of testing and laboratory automation has been the focus of recent developments. These together can help to reduce the excessive use of antibiotics resulting in reduced resistance.

f. Beyond use in human beings: Bolder interventions outside hospitals and a move to ecological antibiotic stewardship are needed. Strategies should be focused on control of non-human sources of antibiotics, resistant bacteria, and resistance genes, such as agriculture and wastewater from the pharmaceutical industry.

1.4.2 The Environmental Antibiotic Resistome

Isolation of the antibiotic-resistant bacterial strains can be done by plating environmental bacteria on antibiotic-containing media in the laboratory. In several cases, the resistance mechanisms have been identified and shown to be specific enzymatic modifications of the antibiotics. The natural population of r-genes is named as the environmental antibiotic resistome (Baysarowich et al., 2008; Wright, 2007). In a recent approach to quantifying the r-genes/ phenotype density in the environment, a collection of morphologically distinct spore-forming actinomycetes for resistance to 21 different antibiotics were screened (D'Costa et al., 2006). Different environments would be expected to vary in the number and type of resistances. Novel resistance mechanisms, as well as many mechanisms related to those found in pathogens, are being identified which is the best evidence available for the presence of a vast resistome with the potential to be captured and expressed as resistance determinants for any overused inhibitor. However, more studies are necessary to establish a strong environment-clinic connection (Canton, 2009).

1.4.3 Use of Bacterial Virulence

The use of inhibitors of bacterial virulence has also been proposed to arrest the disease process. This smart solution would do away with the requirement for antibiotics and appears to have the advantage over antibiosis since the growth of the infecting organism would not be impaired and hence selection for resistance (survival in the host) might not occur. The idea has demonstrated some success in small-animal models, but for further validation extensive research on possible therapies is required (Balaban et al., 1998).

1.4.4 Recruiting Innate Immune System

Among non-antibiotic approaches to the treatment of bacterial diseases, conscription of the innate immune system of the host involving stimulation is also considered (Finlay and Hancock, 2004). The roles of the human gut microbiome in innate immunity may lead to other therapeutic options (Qin et al., 2010).

1.4.5 Phage therapy

Many researchers are considering bacteriophages as an alternative to antibiotics in infection treatment. These can be used for human therapy to control pathogens as well as prevention, biocontrol, and therapy in animal agriculture. Phage therapy and biocontrol have yet to fulfill their promise or potential, largely due to several key obstacles to their performance which the researchers are working on.

1.5 Bacteriophage

A bacteriophage is a virus that infects and replicates within bacteria and archaea. The Greek word “phagein” meaning "to devour" is suffixed with the word "bacteria" to obtain the term. It is also termed informally as a phage. Bacteriophages are composed of proteins of simple or elaborate structures that encapsulate a DNA or RNA genome. Their genomes may encode as few as four genes and as many as hundreds of genes.

Phages are very common and diverse in the biosphere (McGrath and Van Sinderen, 2007) and are ubiquitous viruses, found wherever bacteria exist. The estimated number of phages on the planet is about 10^{31}.

1.5.1 Types of phages

Phages replicate within the bacterium following the injection of their genome into its cytoplasm. Each of the thousands of varieties of phage evolved to infect only one type or a few types of bacteria. Phages cannot replicate by themselves like other viruses, they must capture the reproductive machinery of bacteria by attaching to a bacterium and insert their genetic material.

I. Lytic phages

Lytic phages then destroy the cell, splitting it open to release new viral particles to continue the process. As such, phages could be considered the only “drug” capable of multiplying; when their job is done, they are excreted by the body.

II. Lysogenic phages

Lysogenic phages incorporate their nucleic acid into the chromosome of the host cell and then replicate with it as a unit, without destroying the cell. Under certain conditions, lysogenic phages can be induced to follow a lytic cycle.

1.5.2 The lytic cycle

The phage depends on the host for its propagation. The process is influenced by temperature, nutrients, light and other environmental forces (Jassim and Limoges, 2013). When a host bacterium is infected by a virulent phage, the replication of the phage goes much faster than that of the host cell. The whole cycle of the reproduction takes 30–40 min which the phage completes by utilizing the host machinery and by subverting the host's biological function. At the end of the lysis, the host bacterium dies and simultaneously liberates a large number of progeny phages. The liberated phages are each then ready to start another cycle by infecting new neighbouring bacterial cells. This cycle is known as a lytic 'virulent' cycle. The lytic cycle or 'virulent phages' are being exploited in various areas of biotechnology, including rapid bacterial detection (Stewart et al., 1993; McIntyre et al., 1996), food bioprocessing (Jassim et al., 2012) and removal of bacterial biofilm (Hibma et al., 1997; Jassim et al., 2012).

1.5.3 Bacteriophage therapy

The idea of using the bacteriophage therapeutically is not new, even in the 1920s and 1930s, phage therapy was popular to treat multiple types of infections and conditions. However, the results of those therapies were subjected to inconsistency and lack of scientific validation. Phage therapy had been pushed back in the 1940s when the use of antibiotics had emerged. But in parts of Eastern Europe and the former Soviet Union, it remained a topic of active research.

1.5.4 Advantages and disadvantages

Sulakvelidze et al. (2001) and LocCarrillo and Abedon (2011) reported are known to have some advantages of phages associated with human therapy over the use of antibiotics. Despite being ubiquitous in food, intestinal microbiomes, and the environment, phage particles were not found to cause diseases in humans, animals or plants (D'Costa, et al., 2006). Phages are thus reflected as safe and generally well tolerated by mammals, although a major concern remains that phage therapy will either adversely affect the patient.

The risk of bacteria acquiring resistance remains a possibility in phage therapies (Guerin, et al., 2009; Guilfoile and Hutchinson, 1991). Phage resistance most commonly develops through the down-regulation, shielding, or modification of bacterial cell surface receptors required for viral attachment (Hacker and Kaper, 2000; Hakenbeck, 1998). Evolved phages tend to bind to highly conserved surface receptors, such as those required for virulence (Hawkey and Jones, 2009; Hodgkin, 1949). Bacteria can also evolve or be acquired through lateral transfer, a diverse array of antiviral so-called immune mechanisms.

1.6 *Citrobacter freundii* (*C. freundii*)

C. freundii is a species of facultative anaerobic gram-negative bacteria of the Enterobacteriaceae family (*Citrobacter spp*. 2012). They have a long rod-like shape (see Figure 3), a typical length of 1–5 μm and generally have several flagella used for locomotion. *C. freundii* is found mainly in soil, but their presence in water, sewage, food and in the intestinal tracts of animals and humans has also be observed (Wang et al., 2000). Werkman and Gillen (1932) first discovered the genus Citrobacter. In the same year, its cultures were isolated and identified.

C. freundii is a common component of the gut microbiome of healthy humans (Schloissnig et al., 2013). While most strains are beneficial, there are significant phenotypic variations among strains, even those that share >99% of their genome (Morowitz et al., 2011). Some rare strains of *C. freundii* have been associated with opportunistic nosocomial infections of the respiratory tract, urinary tract, blood, and many other normally sterile sites in immunocompromised patients (Whalen et al., 2007).

1.6.1 Scientific Classification of *C. freundii*

Citrobacter is a common urinary pathogen in hospitalized patients. The genus Citrobacter was discovered in 1932 by Werkman and Gillen. These organisms are found in soil, water, the intestinal tract of animals, and in human clinical samples. Members of the genus Citrobacter usually utilize citrate as a sole carbon source. Table 1 gives the scientific classification of *C. freundii*.

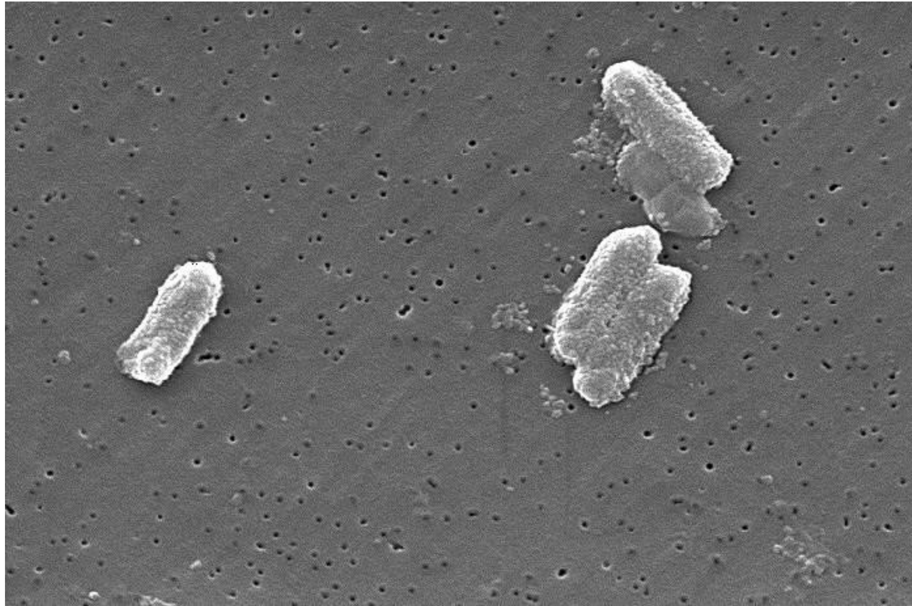


Figure 3 Microscopic image of Citrobacter freundii, Wardell, P. (n.d) Citrobacter freundii [JPG]. Retrieved from https://phil.cdc.gov/Details.aspx?pid=256.

Table 1 Taxonomy of Citrobacter freundii, Source: Werkman and Gillen (1932)

Kingdom:	*Bacteria*
Phylum:	*Proteobacteria*
Class:	*Gammaproteobacteria*
Order:	*Enterobacterales*
Family:	*Enterobacteriaceae*
Genus:	*Citrobacter*
Species:	*C. freundii*
Binomial name	*Citrobacter freundii*

1.6.2 Common sources of *C. freundii*

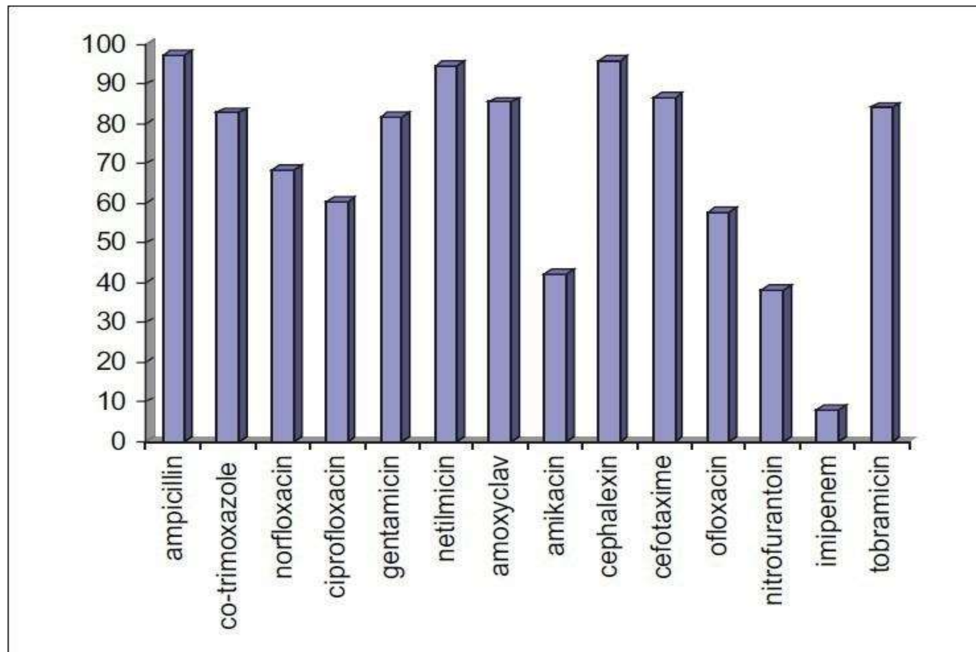
Commonly, a member of the soil microbiome, *C. freundii* plays an important role in the nitrogen cycle in the environment by reducing nitrate to nitrite in the environment (Puchenkova, 1996).

The genus Citrobacter consists of 11 separable genomospecies, out of them, *C. freundii* has been associated with gastroenteritis, neonatal meningitis, and septicemia (Murray et al., 2010). *C. freundii* causes infections of the urinary tract, respiratory tract, blood, and other normally sterile sites in the body. A weak immune system makes the body more vulnerable and predisposed to *C. freundii*, thus triggering UTI or intestinal infection or meningitis. Patients who have been hospitalized for a prolonged period, are more vulnerable to *C. freundii* infections.

Basavaraj, Jyothi, and Basavaraj (2013) have conducted a retrospective study and found that the isolation of *Citrobacter spp.* was associated with catheterization, genitourinary instrumentation, or obstructive uropathy and is the third most common urinary pathogen accounting for 9.4% of the total isolates. A study in Nepal (Baral et al., 2012) also found *Citrobacter spp.* as the second most common urinary pathogen.

1.6.3 Multidrug-resistance of *C. freundii*

A serious matter of concern associated with *Citrobacter spp.* is its multidrug-resistance. This reduces the therapeutic options and further complicate the situation. A retrospective study was done by Davies and Davies (2010) reported that most of the isolates were resistant to the commonly prescribed drugs in UTI including penicillins, cephalosporins, aminoglycosides, and fluoroquinolones. Figure 4 shows the findings of the study.



Source: Basavaraj C. M., P. Jyothi, and Basavaraj V. P. (2013), Antibiotic resistance in Citrobacter spp. isolated from urinary tract infection, *Urol Ann.*, 5(4), 312–313.

Figure 4 Percentage of isolated Citrobacter spp. resistant to different antibiotics.

Shobha et al. (2007) concluded that *Citrobacter spp.* was the third most common urinary pathogen and 30% of the isolates were extended-spectrum beta-lactamase (ESBL) producers. Rizvi et al. (2010) also reported that 62.2% of Citrobacter isolates were producing ESBL. Due to a combination of *Klebsiella pneumoniae* carbapenemase-2 (KPC-2) production, *C. freundii* showed high-level carbapenem resistance (Zhang et al., 2008).

Carbapenems are important antibiotics for the treatment of healthcare-associated infections and have a special role in treating infections with ESBL-producing organisms. The emergence and spread of resistance to carbapenems will end all the treatment options available for treating multidrug-resistant pathogens. Potential Phage therapies as an alternative to treatment of *C. freundii* is, therefore, a research field to be explored. In this thesis, a preliminary assessment of a probable bacteriophage for *C. freundii* is carried out to reconnoiter its viability.

1.7 Objectives of the Study

The main objective of this study is to assess the viability of the described in the bacteriophage listed under the National Center for Biotechnology Information (NCBI) while the specific objectives included:

a. To examine the explicit features of the NCBI listed bacteriophage using The Basic Local Alignment Search Tool (BLAST);
b. To assess the antibiotic resistance of *C. freundii* using the Comprehensive Antibiotic Resistance Database (CARD);
c. To assess the lysogeny or the type of life cycle that takes place when a bacteriophage infects *C. freundii*.

1.8 Organization of the Thesis

Chapter 1 of the thesis elucidates the background and rationale of the study including literature review and objectives of the study. Adopted methodologies are categorized in detail in Chapter 2 while the results and their interpretations are manifested in Chapter 3. Chapter 4 completes the thesis with plausible concluding remarks along with due recommendations.

Chapter 2

Data and Methodology

This study is to be conducted exclusively based on available secondary resources of information. The National Center for Biotechnology Information (NCBI) contributes to the advances of science and health by providing access to biomedical and genomic information and has compiled on their website many resources for the field of bioinformatics. Side by side, recent developments in the field of bio-informatics contributed to making available many protocols and algorithms for analyzing such databases using high dimensional statistical methods. Fortunately, many of these algorithms, with a little help of R programming, can be utilized to produce customized analyses of many of these databases to answer specific research questions. In the following sections, a brief description of the databases and algorithms that are employed in this study along with a step by step description of the activities carried out as the methodology of the study is given.

2.1 Data used for the study

National Center for Biotechnology Information (NCBI) maintains a group of multi-disciplinary researchers who are studying fundamental biomedical problems at the molecular level using mathematical and computational methods including gene organization, sequence analysis, and structure prediction. NCBI has the responsibility for the GenBank DNA sequence database which is built sequences submitted by individual laboratories and by data exchange with the international nucleotide sequence databases, European Molecular Biology Laboratory (EMBL) and the DNA Database of Japan (DDBJ). NCBI also supports and distributes a variety of databases for the medical and scientific communities. Integrated access to sequence, mapping, taxonomy, and structural data are provided for the user on a search and retrieval system using Entrez. Data on all Citrobacter bacteriophage genomes are available on the NCBI database. For this study, only complete genomes were planned to be retrieved for further processing.

2.2 Basic Concepts Relating To the Methodology

2.2.1 Antibiotic Resistance

Antibiotics are used to treat bacterial infections by destroying the bacteria themselves or preventing their growth. It is important to keep in mind that antibiotics are designed to destroy bacteria, not viruses. Antibiotics affect bacterial cells in four ways

- Disrupt cell wall synthesis
- Inhibit RNA synthesis
- Inhibit protein synthesis
- Inhibit DNA replication

Figure 5 depicts the different ways that antibiotics can affect bacterial cells when antibiotics work correctly. The process may be disrupted and antibiotic resistance may be developed in a bacterial cell by following different ways, illustrated in Figure 6:

- Natural resistance due to an impermeable membrane or lack of a target
- Modify antibiotic target
- Produce an enzyme that inactivates antibiotics
- Pump antibiotics out

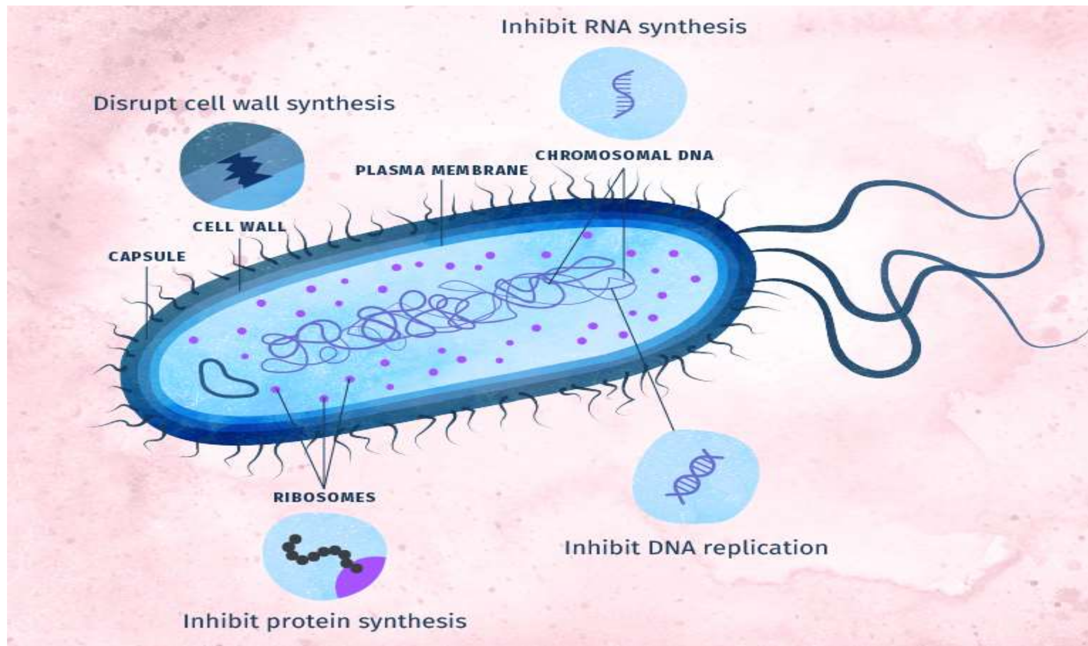


Figure 5 Ways Antibiotics Affect Bacterial Cells. Reprinted from GW Public Health Online, by MPH@GW Staff, 2017, Adapted from https://publichealthonline.gwu.edu/blog/antibiotic-resistance-at-cellular-level/.

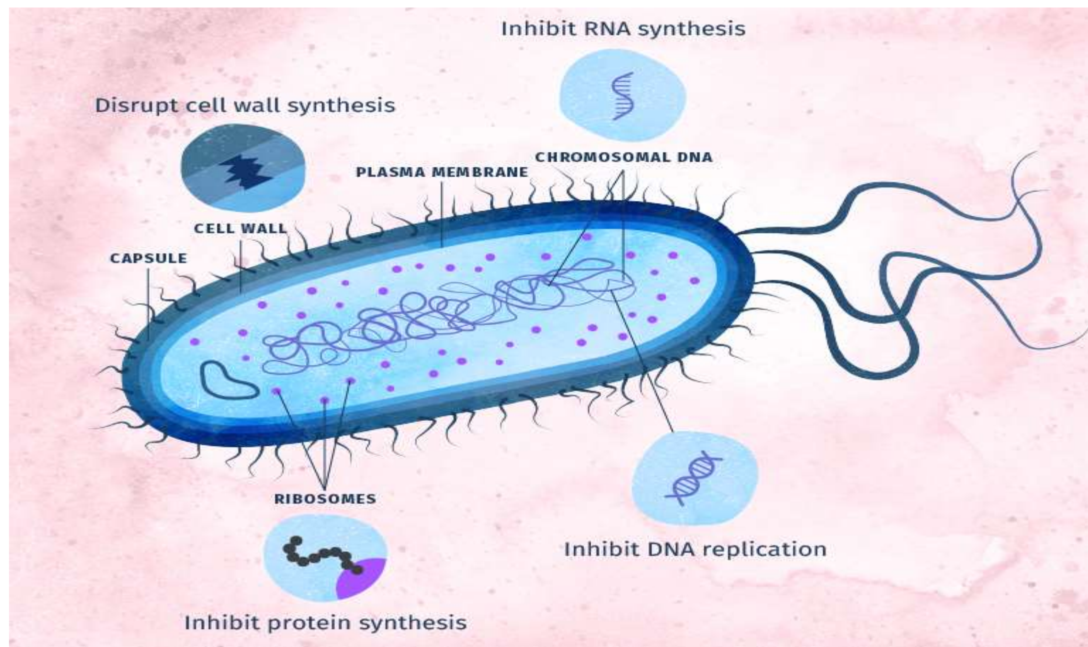


Figure 6 Mechanisms of Antibiotic Resistance. Reprinted from GW Public Health Online, by MPH@GW Staff, 2017, Adapted from https://publichealthonline.gwu.edu/blog/antibiotic-resistance-at-cellular-level/.

Bacteria can acquire resistance by one of the following processes.

- Mutation: Some bacteria develop mutations in the process of cell replication in a way that makes them resistant to antibiotics. The chance of antibiotics being able to destroy the bacteria with the resistance mutation is low. Even when exposed to antibiotics, resistant bacteria can continue to multiply.
- Horizontal Gene Transfer: By either transformation, transduction or conjugation the genetic material of antibiotic-resistance is transferred between different bacteria cells.

The process of how antibiotic resistance develops in bacteria may be summarized in Figure 7.

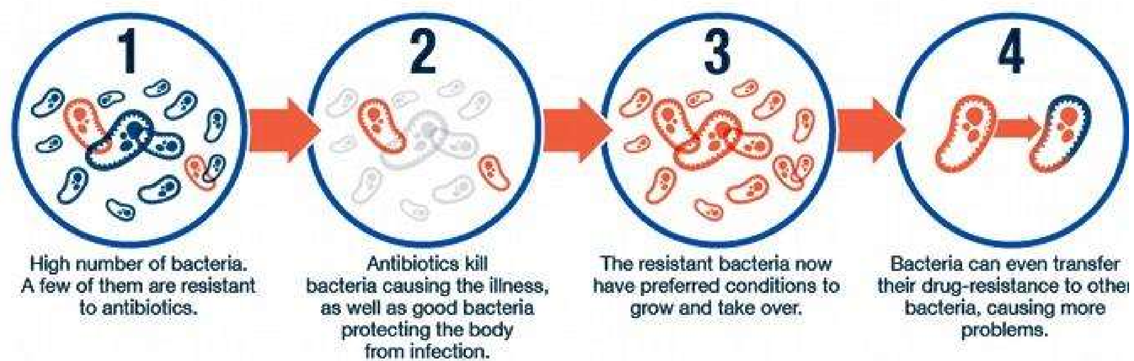


Figure 7 How antibiotic resistance develops, Adapted from Antibiotic Resistance: Key Facts, Retrieved from https://www.assignmentpoint.com/science/health/antibiotic-resistance-key-facts.html.

2.2.2 Virulence Factors

Bacteria, viruses, fungi, and protozoa can produce molecules that add to their effectiveness. These molecules are known as virulence factors. Virulence factors can enable them to achieve:

a. Colonization of a niche in the host (attaching to cells)
b. Immunoevasion, evasion of the host's immune response
c. Immunosuppression, inhibition of the host's immune response
d. Entry into and exit out of cells (if the pathogen is an intracellular one)
e. Obtain nutrition from the host.

Virulence factors are featured in different types of pathogens. Chromosomally encoded virulence factors and intrinsic to the bacteria and some are generated from plasmids and bacteriophages. Virulence factors encoded on mobile genetic elements spread through HGT and can convert harmless bacteria into dangerous pathogens.

2.2.3 Synteny

The distribution pattern of genes on a chromosome is known as the Synteny. The conservation of this pattern of gene locations is important because it is likely that genes positioned near each other on the genome in one species would be close to each other on a single chromosome in evolutionarily related species. Synteny of a chromosomal can be disrupted due to rearrangement or mutation processes such as translocation, inversion, chromosome fusion, and breakage. Figure 8 illustrates the different kinds of synteny.

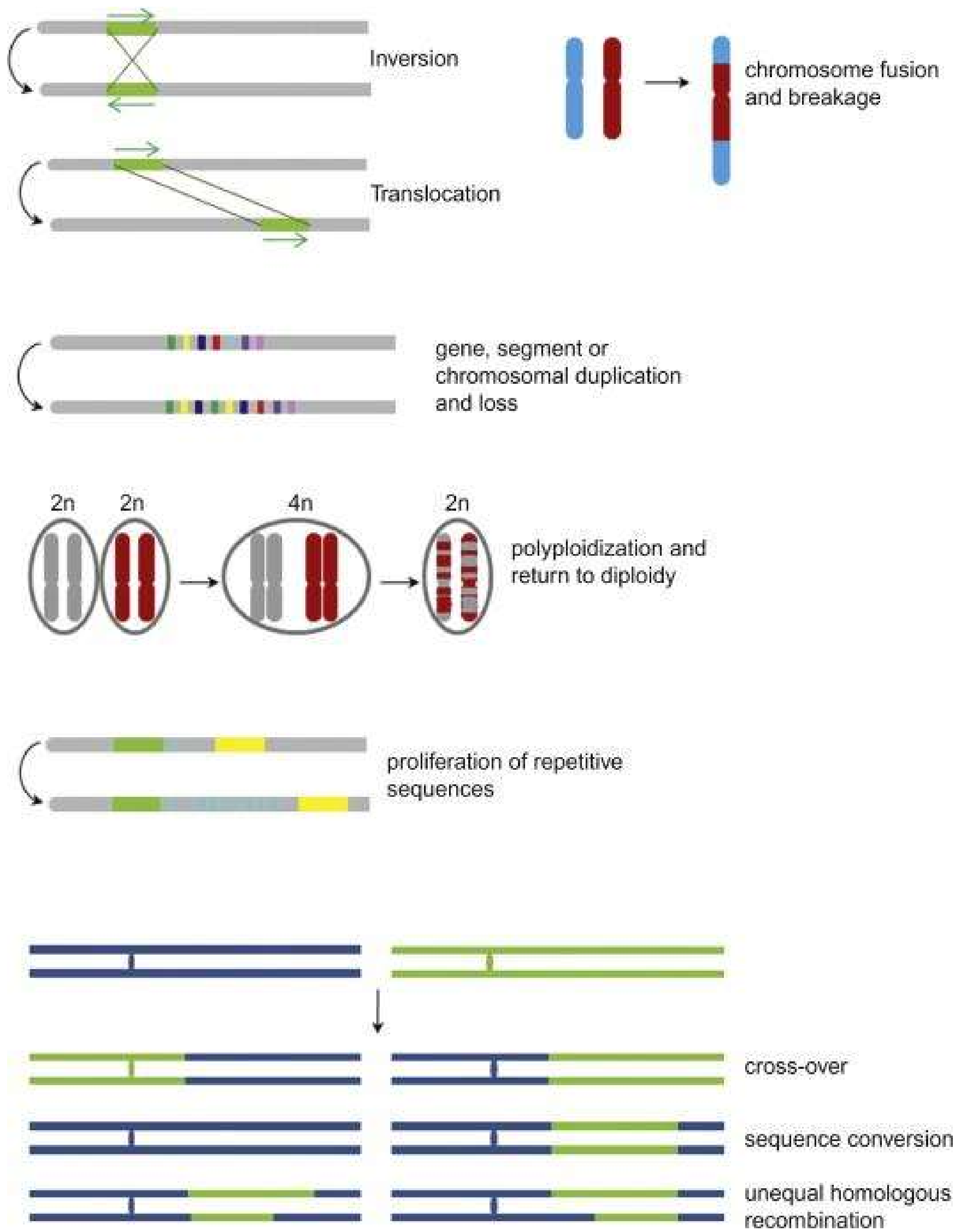


Figure 8 Chromosomal rearrangement processes leading to disruption of synteny. Reprinted from Plant Biotechnology and Agriculture , D. by Fleury, U. Baumann, and P. Langridge, 2012, Elsevier.

2.2.4 Phylogenetics

The study of evolutionary relatedness among groups of organisms is known as Phylogenetics. At the molecular level, to infer these relationships for both organisms and the genes, sequence data is used in phylogenetics. While in modern research, a large volume of sequence data is available publicly, this type of inference has become more prevalent in all fields of biology. The basis for comparative genomics, which is a core demand in recent years is also laid upon by phylogenetics resulting in a term comparative genomics. One of the many of the applications of these studies is to evaluate the evolutionary relationships among genomes. In comparative genomics, similarity and differences among genomes are identified by making the comparison at any of the levels of

- Whole-genome sequences
- Genome sequences involving blocks of conserved synteny
- Number of protein-coding genes
- Regulatory sequences
- Other focused features

A diagrammatic representation of the evolutionary relationships among various taxa is called a phylogenetic tree or evolutionary tree (see, for example, Figure 9). The branching pattern of a tree is called the topology of the tree. The nodes represent taxonomic units, such as species (or higher taxa), populations, genes, or proteins. A branch is called an edge and represents the time estimate of the evolutionary relationships among the taxonomic units.

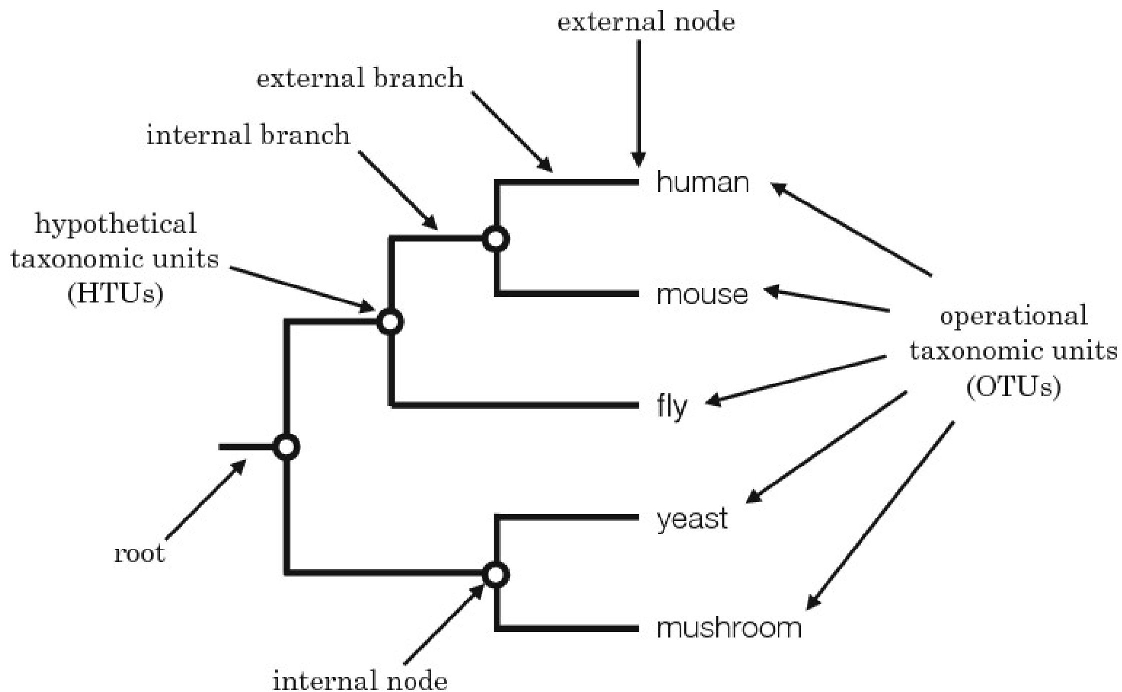


Figure 9 Terminology frequently used in phylogenetic trees. Reprinted from "Molecular Phylogenetics: Concepts for a Newcomer," by P. Ajawatanawong, 2016, Advances in biochemical engineering/biotechnology.

2.2.5 Lysogeny

As already been discussed in section 1.5, phages are the most abundant biological entity and play critical roles in controlling bacterial populations through phage-mediated killing. In some situations and forms of the phages, its survival depends on the survival of the bacterial host, thus, it becomes advantageous for phages to encode genes that may help bacterial fitness.

The lysogenic cycles of phages in comparison with the lytic cycles (see section 1.5) is demonstrated in Figure 10. As a result of the lysogeny, the phages benefit the bacteria when they are infecting a human.

The genes that encode these fitness factors, known as "phage morons" or simply "morons" (Brüssow, Canchaya, and Hardt, 2004). The morons increase bacterial fitness through a wide range of mechanisms and play important roles in bacterial diseases. Genome analyses of

double-stranded DNA tailed bacteriophages argue that they evolve by recombination of genes and by the acquisition of novel genes as simple genetic elements termed morons (Hendrix, et al., 2000).

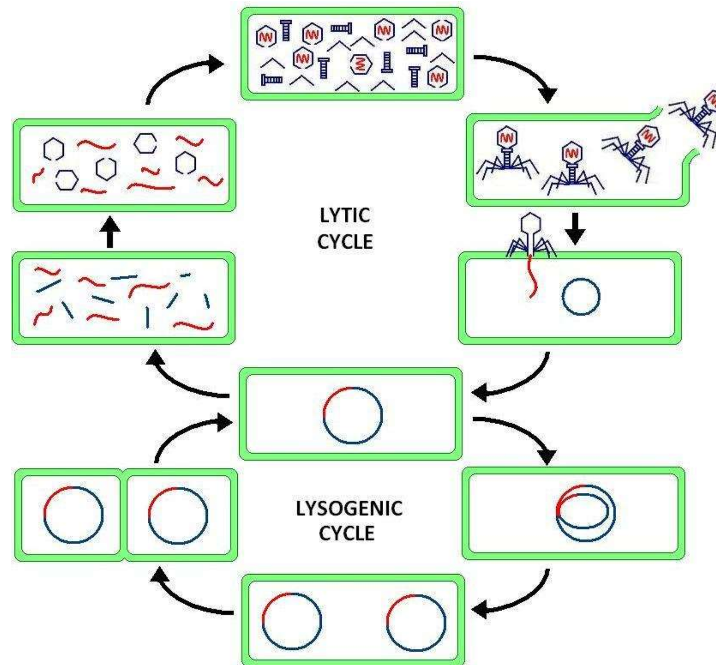


Figure 10 The lysogenic and lytic cycles of bacteriophage. Reprinted from Wikipedia: Lysogenic cycle (n.d) by Suly12. Retrieved from https://en.wikipedia.org/wiki/Lysogenic_cycle

2.3 Tools and Algorithms

2.3.1 The Basic Local Alignment Search Tool (BLAST)

In bioinformatics, BLAST (Casey, 2005) is a search and match algorithm for comparing primary biological sequence information, such as the amino-acid sequences of proteins or the nucleotide sequences of DNA and RNA. BLAST is one of the most widely used bioinformatics programs for sequence searching. The algorithm enables a researcher

- To compare a subject protein or nucleotide sequence (called a query) with a library or database of sequences, and identify library sequences that resemble the query sequence above a certain threshold.
- To find regions of local similarity between sequences.
- To compare nucleotide or protein sequences to sequence databases and calculates the statistical significance of matches.
- To infer functional and evolutionary relationships between sequences
- To identify members of gene families.

Different types of BLAST searches are available, for example, BLASTn for nucleotide sequences and BLASTp for amino acid sequences in protein.

2.3.2 The Comprehensive Antibiotic Resistance Database (CARD)

Organized by the Antibiotic Resistance Ontology (ARO) and Antimicrobial Resistance (AMR) gene detection models, the Comprehensive Antibiotic Resistance Database (CARD) (Alcock et al., 2020) is a rigorously curated database of resistance determinants and associated antibiotics which is characterized and peer-reviewed. It is built entirely using open source software and tools. The CARD

- provides data, models, and algorithms relating to the molecular basis of antimicrobial resistance;
- provides curated reference sequences and SNPs organized via the ARO;
- can be browsed on the website or downloaded in several formats;
- are additionally associated with detection models, in the form of curated homology cut-offs and SNP maps, for prediction of resistome from molecular sequences;
- models can be downloaded or can be used for the analysis of genome sequences using the Resistance Gene Identifier (RGI), either online or as a stand-alone tool.

2.3.3 The Virulence Factor Database (VFDB)

Similar to the CARD, the Virulence Factor Database (VFDB) (Chen et.al., 2005) is an integrated and comprehensive online resource for curating information. This database collects, curates characterizes peer-review data on virulence factors of bacterial pathogens. The VFDB provides

- in-depth coverage of major virulence factors of the best-characterized bacterial pathogens, with the structure features, functions and mechanisms used by these pathogens
- provide current knowledge of the wide variety of mechanisms used by bacterial pathogens for researchers to elucidate pathogenic mechanisms in bacterial diseases.

2.4 Conceptual flow of the methodologies

Using the above-mentioned tools and resources, the following sequential steps were performed to address the broad and specific objectives of this study:

2.4.1 Isolating Full Genome Sequences from NCBI Using BLAST

All available Citrobacter bacteriophage genomes were downloaded from the NCBI database. From the downloaded data, only complete genomes were further processed. Using Linux based BLAST to create a database from the compiled FASTA files, and running BLASTn against the aforementioned compilation, duplicates were identified and removed, leaving only complete genomes of unique Citrobacter phages.

2.4.2 Annotation of Sequence FASTA using Prokka

Whole genome annotation is the process of identifying features of interest in a set of genomic DNA sequences and labelling them with useful information. Prokka (Seemann, 2014) is a software tool to annotate bacterial, archaeal and viral genomes quickly and produce standards-compliant output files. The 21 unique Citrobacter phage genomes were annotated using Prokka.

2.4.3 Phylogenetic Analysis with Roary

Roary is a high-speed standalone pan genome pipeline, which takes annotated assemblies in GFF3 format (produced by Prokka) and calculates the pan genome. The GFF3 files from the output from Prokka were fed into Linux based Roary.

2.4.4 Phylogenetic Tree from iTOL

Interactive Tree Of Life is an online tool for the display, annotation, and management of phylogenetic trees. From the resulting Roary output files, the file in .fa.newick format was uploaded to iTOL online.

2.4.6 Synteny and Genomic Organization Using genoPlotR and BLAST

The phages adjacent to each other in the phylogenetic tree were run through BLASTn to gain "comparison" files. The gbk format files from Prokka and the comparison files from BLAST were then converted into datasets in R.

genoPlotR is an R package that draws gene or genome maps and comparisons between these, in a publication-grade manner. genoPlotR was then used to plot the genome and comparison datasets to show synteny.

2.4.7 Checking for Antibiotic Resistance Genes Using CARD and BLAST

The phage genomes were compared to the Comprehensive Antibiotic Resistance Database (CARD) using BLAST to check for matches. Percentage identity and length of alignment were important to note.

2.4.8 Checking for Virulence Factors Using VFDB and BLAST

The phage genomes were compared to the Virulence Factor Database (VFDB) using BLAST to check for matches. Percentage identity and length of alignment were important to note.

2.4.8 Checking for Lysogeny Using BLAST, Prokka, PhiSpy, Roary and iTOL

PhiSpy is a tool used to identify prophages in Bacterial genomes. Given an annotated genome it will use several approaches to identify the most likely prophage regions. (Akhter, Aziz, and Edwards, 2012).

All available *Citrobacter freundii* genomes were downloaded from the NCBI database and were annotated using Prokka. The annotated files were then fed through PhiSpy which identified possible prophages. These potential prophages were then fed to Roary, alongside the 21 prophages being studied. A phylogenetic tree was produced using iTOL.

The logical framework of the methodology can be summarized in Figure 11.

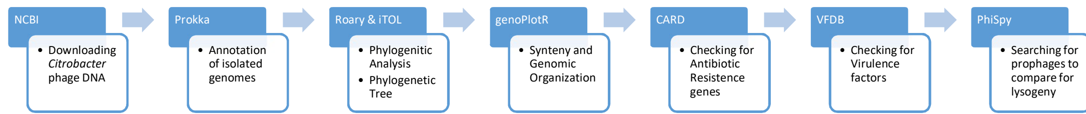


Figure 11 Logical framework of the study

Chapter 3

Results and Discussion

In this chapter, the results of the algorithms and analyses conducted are discussed and interpreted focusing on the study objectives. Outcomes of each of the procedures are summarized and presented in the following sections.

3.1 Complete genomes of Unique Citrobacter Phages

Complete genomes of unique Citrobacter phages were compiled from the FASTA files downloaded NCBI database. Using the BLAST duplicate entries were identified and removed, resulting in 21 complete Citrobacter phage genomes being identified. The names and genome sizes are listed in Table 2.

Table 2 Genome sizes of 21 Citrobacter phages

Name of Phage	Size of Genome (bp)
phiCFP-1	38625
SH2	39158
SH4	39274
SH1	39434
SH3	39444
SH5	39832
Sazh	49665
Stevie	49816
CF1-DK-2017	50339
Mijalis	87998
Moogle	87999
Maleficent	89570
Mordin	89596
Michonne	90000
Moon	170341
CF1-ERZ-2017	171911
Merlin	172733
IME-CF2	177688
Miller	178171
Margaery	178182
Maroon	178830

3.2 Annotation of sequence FASTA from NCBI

The tool Prokka was used for annotation of the genes of the *Citrobacter freundii* phages, the process identified genes in the FASTA sequences and added names for the genes whose function is recognized. The program produced several files that were then used for later processes.

3.3 The Phylogenetic tree

The phylogenetic tree of Citrobacter phages was produced by Roary and the Interactive Tree Of Life (iTOL) was used to visualize this data. The resulting tree is presented in Figure 12 and demonstrates 6 major clades.

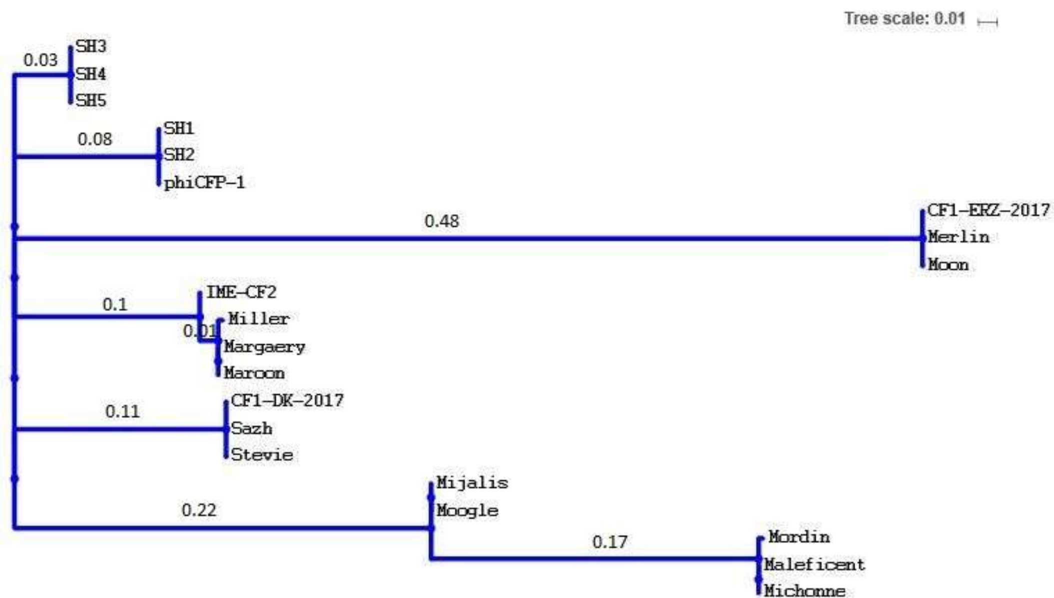


Figure 12 Phylogenetic tree of Citrobacter phages produced using iTOL

3.4 Synteny and Genomic Organization

The phages adjacent to each other in the phylogenetic tree were compared through BLASTn to gain "comparison" files. Using genoPlotR these files were used to generate the synteny diagrams in Figures 13 through 18. In each of these diagrams, each blue arrow is representative of a gene, and each colored block extending between the genomes is a conserved block.

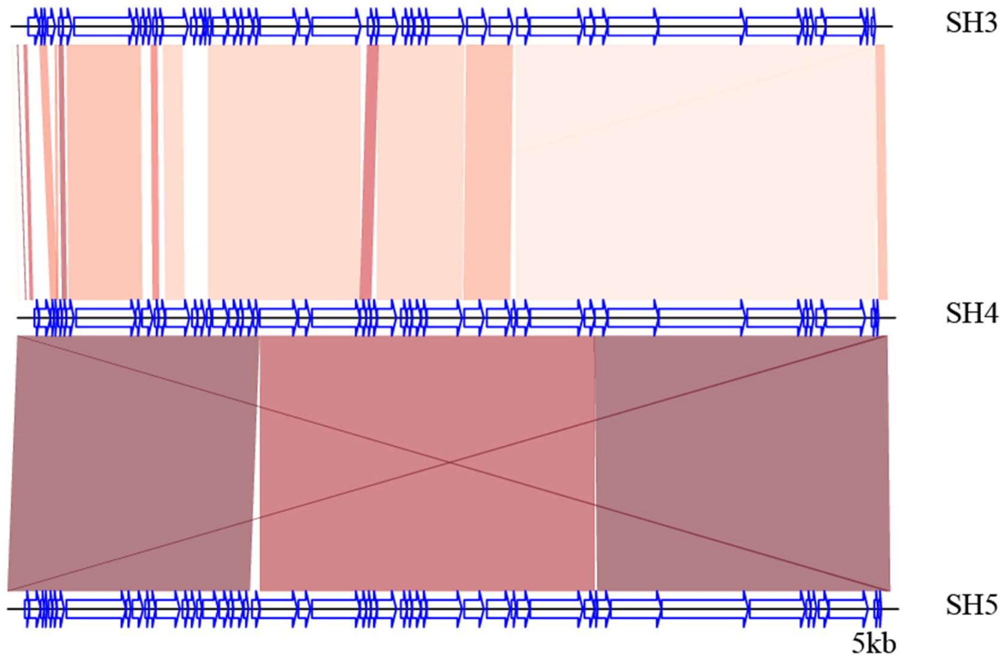


Figure 13 Synteny diagrams for Citrobacter phages SH3, SH4, and SH5

From Figure 13 it can be seen that phages SH3 and SH4 show conserved synteny, similarly phages SH4 and SH5 also have highly similar genomes.

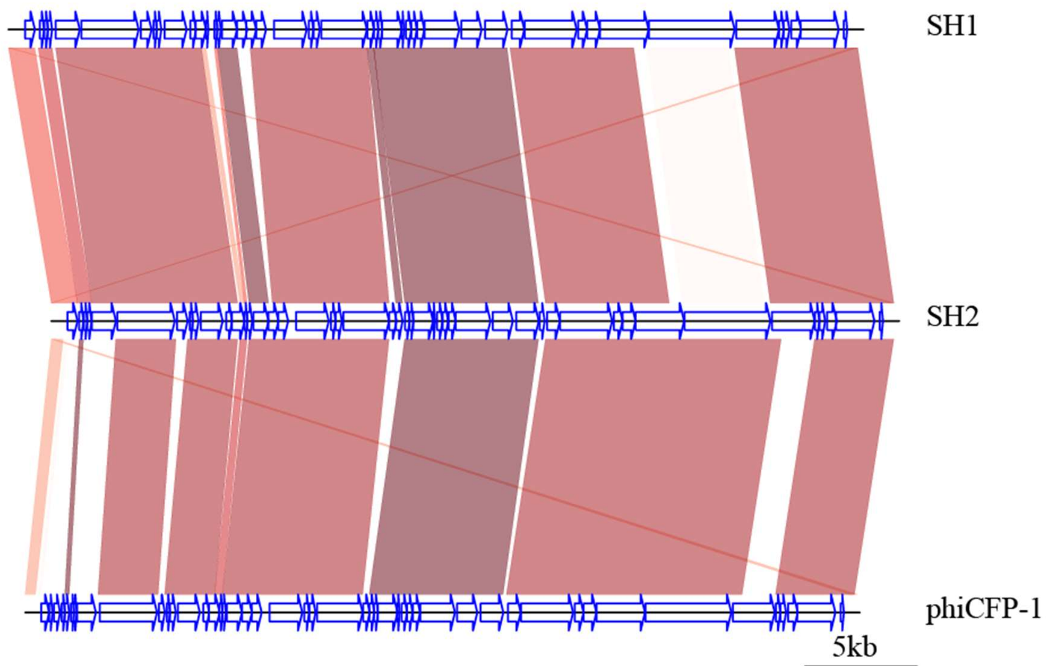


Figure 14 Synteny diagrams for Citrobacter phages SH1, SH2, and phiCFP-1

From Figure 14 it can be seen that phages SH1 and SH2 show conservation of synteny with a slight shift in structure, however, a large block is notably dissimilar. Phages SH2 and phiCFP-1 have largely conserved synteny with a shift in genomic structure in the opposite direction.

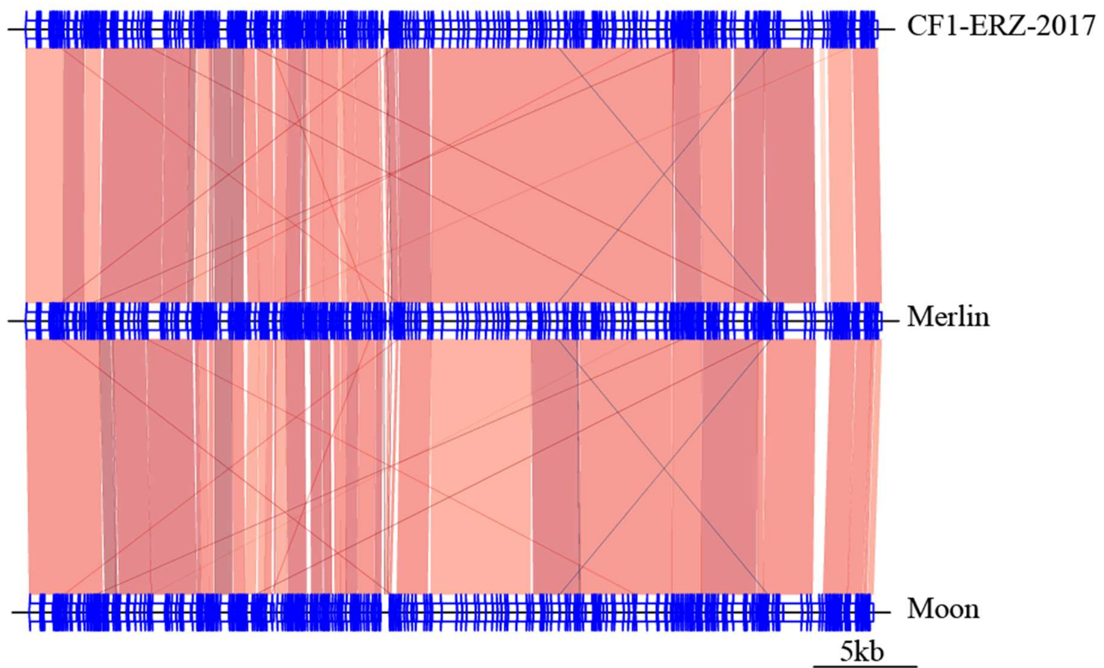


Figure 15 Synteny diagrams for Citrobacter phages CF1-ERZ-2017, Merlin, and Moon

From Figure 15 it can be seen that phages CF1-ERZ-2017 and Merlin show largely conserved synteny, and similarly phages Merlin and Moon also show highly similar genomes.

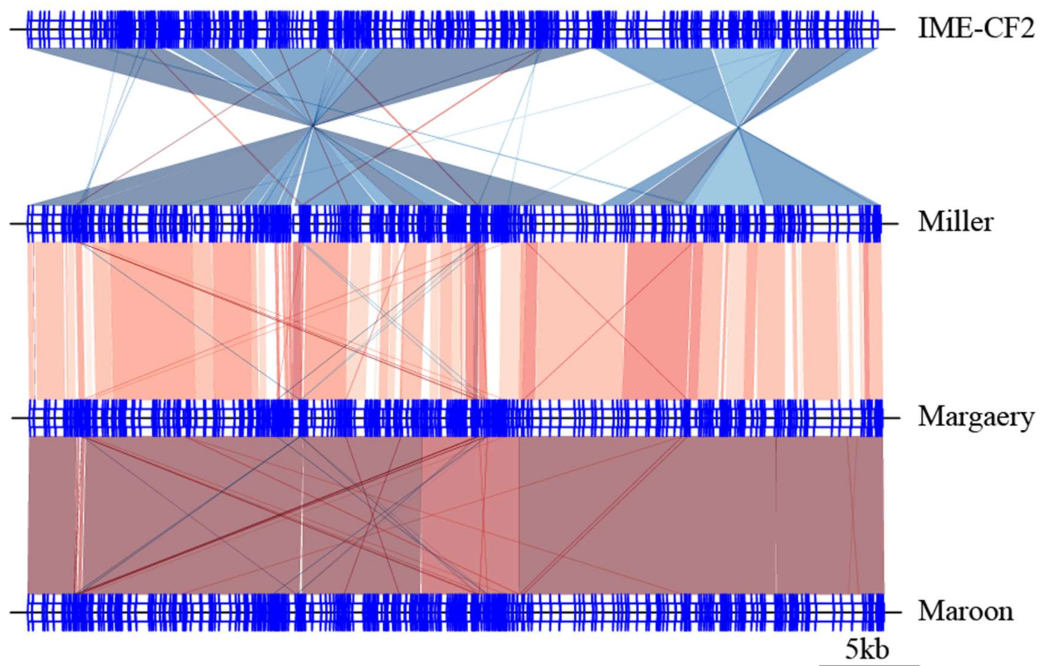


Figure 16 Synteny diagrams for Citrobacter phages IME-CF2, Miller, Margaery, and Maroon

From Figure 16 it can be seen that phages IME-CF2 and Miller have two distinct inverted gene blocks, while phages Miller and Margaery show several smaller detectable blocks of genes. Similarly Margaery and Maroon show highly conserved synteny. This means phages Miller and Maroon also have conserved synteny.

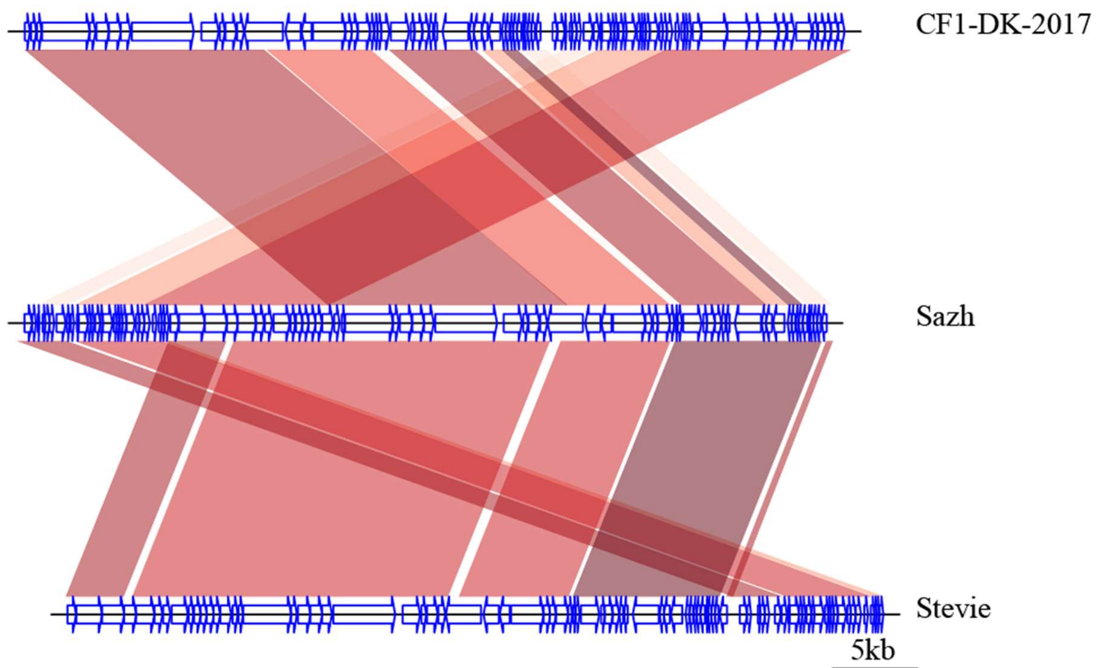


Figure 17 Synteny diagrams for Citrobacter phages CF1-DK-2017, Sazh, and Stevie

From Figure 17 it can be seen that phages CF1-DK-2017 and Sazh show a shift of the conserved gene blocks, while phages Sazh and Stevie also show a shift, but in the opposite direction, and smaller in size.

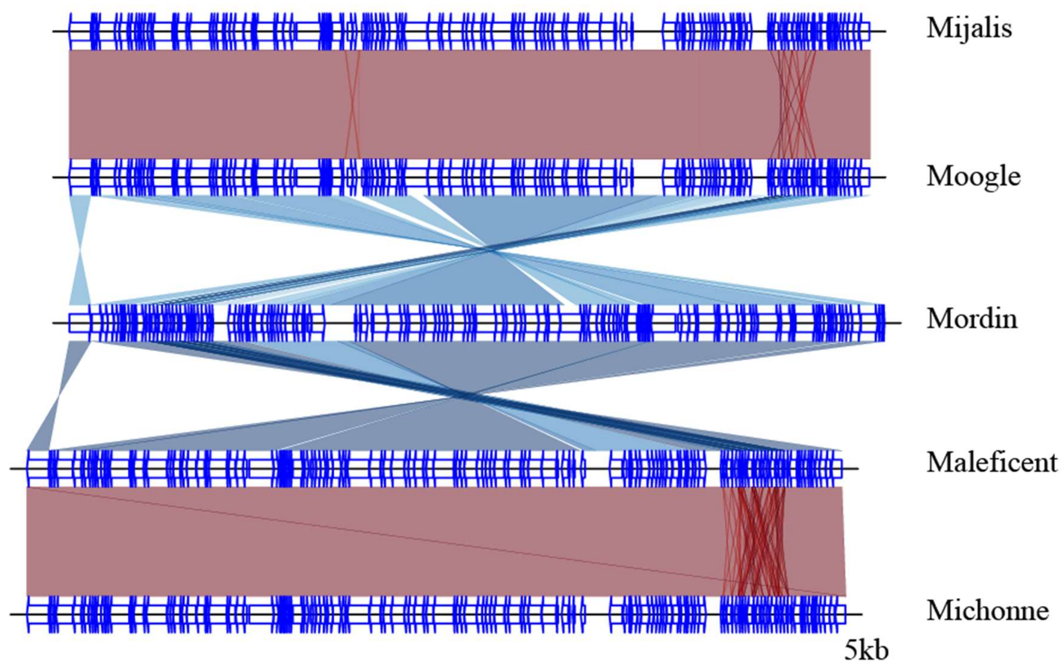


Figure 18 Synteny diagrams for Citrobacter phages Mijalis, Moogle, Mordin, Maleficent, and Michonne

From Figure 18 it can be seen that phages Mijalis and Moogle have conserved synteny, with almost entirely similar genomes. Similarly phages Maleficent and Michonne also show conserved synteny. However, the phage Mordin shows the inversion of almost the entire genome with both phages Moogle and Maleficent, except the terminal gene block, which remains in the same position but is inverted within that spot. This means that phages Moogle and Maleficent also have conserved synteny and also show high similarity with phages Mijalis and Michonne.

3.5 Checking for Antibiotic Resistance Genes in the Citrobacter Phage

By comparing the Citrobacter phage DNA data with the Comprehensive Antibiotic Resistance Database (CARD), it can be found whether the phage is carrying any antibiotic resistance genes that it can then pass on to *Citrobacter freundii* that it attacks. A total of 21223 hits were returned by the BLASTn comparison. The percentage occurrence of matches at each level of percentage identity is summarized in a histogram and is demonstrated in Figure 19.

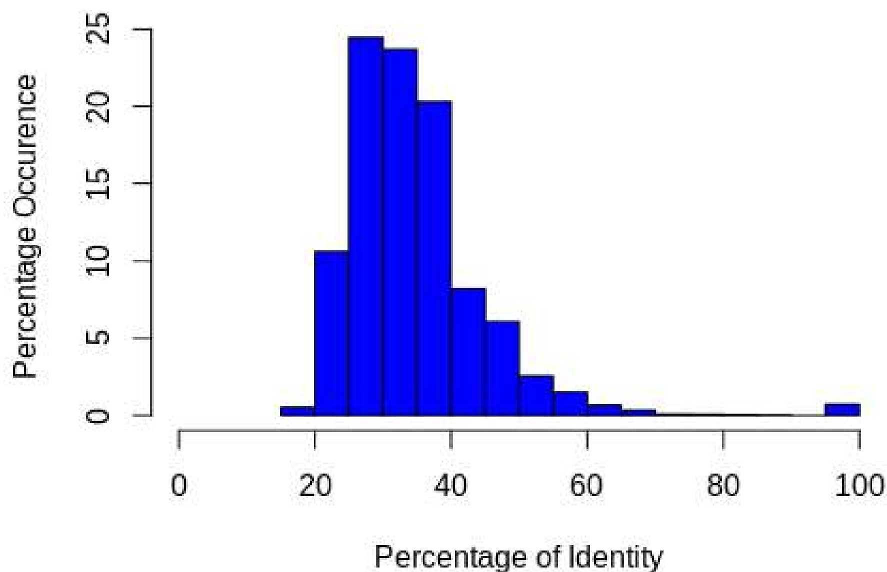


Figure 19 Histogram of the percentage identity in matches between the Citrobacter phage genomes and CARD

From Figure 19, it can be observed that almost all the 21223 hits show less than 70% identity. These matches are not significant enough to warrant concern. However, 149 hits (0.702% of the total number of matches) showed 100% identity. That is to say, certain portions of antibiotic resistance genes were found in the phage genomes.

However, as can be seen in Table 3, the maximum alignment length of any of these matches is 7 nucleotides long. This match length can be considered insignificant.

Table 3 Matches between Citrobacter phage DNA and CARD with 100% identity

Phage Sequence ID	CARD Sequence ID	Percentage Identity (%)	Length of Alignment (bp)
IME-CF2_00177	gb|CAA83855.1|ARO:3000380|FosC	100	7
Miller_00022	gb|CAA83855.1|ARO:3000380|FosC	100	7
SH3_00048	gb|AAK55330.1|ARO:3001768|OXA-118	100	6
SH3_00048	gb|AAN41427.1|ARO:3001775|OXA-119	100	6
SH3_00048	gb|AAC41449.1|ARO:3001398|OXA-3	100	6
SH3_00048	gb|CAA71699.2|ARO:3001416|OXA-21	100	6
SH3_00048	gb|ACM67635.1|ARO:3001800|OXA-226	100	6
SH3_00048	gb|AIG94927.1|ARO:3001606|OXA-415	100	6
SH3_00048	gb|ACT09125.1|ARO:3001799|OXA-161	100	6
SH3_00048	gb|AAK15582.1|ARO:3001428|OXA-34	100	6
SH3_00048	gb|AAK58418.1|ARO:3001426|OXA-32	100	6
SH3_00048	gb|AAB59082.1|ARO:3001397|OXA-2	100	6
SH3_00048	gb|ABQ15112.1|ARO:3001802|OXA-141	100	6
SH3_00048	gb|AAG24866.1|ARO:3001430|OXA-36	100	6
SH3_00048	gb|AAP43641.1|ARO:3001810|OXA-53	100	6
SH3_00048	gb|AAB05874.1|ARO:3001410|OXA-15	100	6
SH3_00048	gb|AEE61368.1|ARO:3001487|OXA-210	100	6
SH3_00048	gb|AFU91598.1|ARO:3001788|OXA-204	100	6
SH3_00048	gb|AAG33665.1|ARO:3001431|OXA-37	100	6
SH3_00048	gb|AHF71363.1|ARO:3001774|OXA-370	100	6
SH3_00048	gb|YP_001844885.1|ARO:3001415|OXA-20	100	6
SH3_00048	gb|AGC60013.1|ARO:3001787|OXA-245	100	6
SH3_00048	gb|AAP70012.1|ARO:3001782|OXA-48	100	6
SH3_00048	gb|AGC60012.1|ARO:3001786|OXA-244	100	6
SH3_00048	gb|ADY06444.1|ARO:3001783|OXA-163	100	6
SH3_00048	gb|ADG27454.1|ARO:3001776|OXA-162	100	6
SH3_00048	gb|AEP16366.1|ARO:3001784|OXA-181	100	6
SH3_00048	gb|AGD91915.1|ARO:3001778|OXA-232	100	6
SH3_00048	gb|AAR89917.1|ARO:3001812|OXA-54	100	6
SH3_00048	gb|AGC70814.1|ARO:3001791|OXA-247	100	6
SH3_00048	gb|AFC95894.1|ARO:3001814|OXA-199	100	6
SH3_00048	gb|AOQ26572.1|ARO:3004362|OXA-535	100	6
SH3_00048	gb|AAA93528.1|ARO:3001409|OXA-14	100	6
SH3_00048	gb|AAB97924.1|ARO:3001411|OXA-16	100	6
SH3_00048	gb|ACO72579.1|ARO:3001801|OXA-147	100	6
SH3_00048	gb|AAF72942.1|ARO:3001423|OXA-28	100	6
SH3_00048	gb|CAP69660.1|ARO:3001811|OXA-129	100	6
SH4_00048	gb|ACM67635.1|ARO:3001800|OXA-226	100	6
SH4_00048	gb|AAC41449.1|ARO:3001398|OXA-3	100	6
SH4_00048	gb|AAK55330.1|ARO:3001768|OXA-118	100	6
SH4_00048	gb|AAN41427.1|ARO:3001775|OXA-119	100	6

Table Continued

SH4_00048	gb|CAA71699.2|ARO:3001416|OXA-21	100	6
SH4_00048	gb|AIG94927.1|ARO:3001606|OXA-415	100	6
SH4_00048	gb|ACT09125.1|ARO:3001799|OXA-161	100	6
SH4_00048	gb|AAK58418.1|ARO:3001426|OXA-32	100	6
SH4_00048	gb|ABQ15112.1|ARO:3001802|OXA-141	100	6
SH4_00048	gb|AAB59082.1|ARO:3001397|OXA-2	100	6
SH4_00048	gb|AAK15582.1|ARO:3001428|OXA-34	100	6
SH4_00048	gb|AEE61368.1|ARO:3001487|OXA-210	100	6
SH4_00048	gb|AAB05874.1|ARO:3001410|OXA-15	100	6
SH4_00048	gb|AAG24866.1|ARO:3001430|OXA-36	100	6
SH4_00048	gb|AAG33665.1|ARO:3001431|OXA-37	100	6
SH4_00048	gb|YP_001844885.1|ARO:3001415|OXA-20	100	6
SH4_00048	gb|AAP43641.1|ARO:3001810|OXA-53	100	6
SH4_00048	gb|AFU91598.1|ARO:3001788|OXA-204	100	6
SH4_00048	gb|AHF71363.1|ARO:3001774|OXA-370	100	6
SH4_00048	gb|AGC60013.1|ARO:3001787|OXA-245	100	6
SH4_00048	gb|AAP70012.1|ARO:3001782|OXA-48	100	6
SH4_00048	gb|ADG27454.1|ARO:3001776|OXA-162	100	6
SH4_00048	gb|AGC60012.1|ARO:3001786|OXA-244	100	6
SH4_00048	gb|AEP16366.1|ARO:3001784|OXA-181	100	6
SH4_00048	gb|ADY06444.1|ARO:3001783|OXA-163	100	6
SH4_00048	gb|AAR89917.1|ARO:3001812|OXA-54	100	6
SH4_00048	gb|AGD91915.1|ARO:3001778|OXA-232	100	6
SH4_00048	gb|AGC70814.1|ARO:3001791|OXA-247	100	6
SH4_00048	gb|AFC95894.1|ARO:3001814|OXA-199	100	6
SH4_00048	gb|AOQ26572.1|ARO:3004362|OXA-535	100	6
SH4_00048	gb|AAA93528.1|ARO:3001409|OXA-14	100	6
SH4_00048	gb|AAB97924.1|ARO:3001411|OXA-16	100	6
SH4_00048	gb|AAC15074.1|ARO:3001412|OXA-17	100	6
SH4_00048	gb|AMB19637.1|ARO:3001555|OXA-368	100	6
SH4_00048	gb|AER57903.1|ARO:3001500|OXA-251	100	6
SH4_00048	gb|AHC31001.1|ARO:3001807|OXA-246	100	6
SH4_00048	gb|ABY79006.1|ARO:3001803|OXA-142	100	6
SH4_00048	gb|CAA80304.1|ARO:3001406|OXA-11	100	6
SH4_00048	gb|AFN20670.1|ARO:3001499|OXA-240	100	6
SH4_00048	gb|CCE94500.1|ARO:3001502|OXA-256	100	6
SH4_00048	gb|AAG45720.1|ARO:3001405|OXA-10	100	6
SH4_00048	gb|BAR45714.1|ARO:3003610|OXA-454	100	6
SH4_00048	gb|AID67109.1|ARO:3001495|OXA-233	100	6
SH4_00048	gb|AAF72942.1|ARO:3001423|OXA-28	100	6
SH4_00048	gb|ACO72579.1|ARO:3001801|OXA-147	100	6
SH4_00048	gb|CAH69530.1|ARO:3001798|OXA-74	100	6
SH4_00048	gb|AAD02245.1|ARO:3001414|OXA-19	100	6
SH4_00048	gb|AAR32651.1|ARO:3001795|OXA-56	100	6
SH4_00048	gb|CAA53242.1|ARO:3001402|OXA-7	100	6
SH4_00048	gb|CAP69660.1|ARO:3001811|OXA-129	100	6

Table Continued

SH4_00048	gb|ACN85419.1|ARO:3001804|OXA-145	100	6
SH4_00048	gb|CAL85435.1|ARO:3001439|OXA-101	100	6
SH4_00048	gb|ADV41666.1|ARO:3001475|OXA-183	100	6
SH4_00048	gb|AAK49460.1|ARO:3001429|OXA-35	100	6
SH4_00048	gb|CAA41211.1|ARO:3001400|OXA-5	100	6
SH4_00048	gb|AAC46344.1|ARO:3001408|OXA-13	100	6
SH5_00051	gb|ACM67635.1|ARO:3001800|OXA-226	100	6
SH5_00051	gb|AAC41449.1|ARO:3001398|OXA-3	100	6
SH5_00051	gb|AAK55330.1|ARO:3001768|OXA-118	100	6
SH5_00051	gb|AAN41427.1|ARO:3001775|OXA-119	100	6
SH5_00051	gb|CAA71699.2|ARO:3001416|OXA-21	100	6
SH5_00051	gb|AIG94927.1|ARO:3001606|OXA-415	100	6
SH5_00051	gb|ACT09125.1|ARO:3001799|OXA-161	100	6
SH5_00051	gb|AAK58418.1|ARO:3001426|OXA-32	100	6
SH5_00051	gb|ABQ15112.1|ARO:3001802|OXA-141	100	6
SH5_00051	gb|AAB59082.1|ARO:3001397|OXA-2	100	6
SH5_00051	gb|AAK15582.1|ARO:3001428|OXA-34	100	6
SH5_00051	gb|AEE61368.1|ARO:3001487|OXA-210	100	6
SH5_00051	gb|AAB05874.1|ARO:3001410|OXA-15	100	6
SH5_00051	gb|AAG24866.1|ARO:3001430|OXA-36	100	6
SH5_00051	gb|AAG33665.1|ARO:3001431|OXA-37	100	6
SH5_00051	gb|YP_001844885.1|ARO:3001415|OXA-20	100	6
SH5_00051	gb|AAP43641.1|ARO:3001810|OXA-53	100	6
SH5_00051	gb|AFU91598.1|ARO:3001788|OXA-204	100	6
SH5_00051	gb|AHF71363.1|ARO:3001774|OXA-370	100	6
SH5_00051	gb|AGC60013.1|ARO:3001787|OXA-245	100	6
SH5_00051	gb|AAP70012.1|ARO:3001782|OXA-48	100	6
SH5_00051	gb|ADG27454.1|ARO:3001776|OXA-162	100	6
SH5_00051	gb|AGC60012.1|ARO:3001786|OXA-244	100	6
SH5_00051	gb|AEP16366.1|ARO:3001784|OXA-181	100	6
SH5_00051	gb|ADY06444.1|ARO:3001783|OXA-163	100	6
SH5_00051	gb|AAR89917.1|ARO:3001812|OXA-54	100	6
SH5_00051	gb|AGD91915.1|ARO:3001778|OXA-232	100	6
SH5_00051	gb|AGC70814.1|ARO:3001791|OXA-247	100	6
SH5_00051	gb|AFC95894.1|ARO:3001814|OXA-199	100	6
SH5_00051	gb|AOQ26572.1|ARO:3004362|OXA-535	100	6
SH5_00051	gb|AAA93528.1|ARO:3001409|OXA-14	100	6
SH5_00051	gb|AAB97924.1|ARO:3001411|OXA-16	100	6
SH5_00051	gb|AAC15074.1|ARO:3001412|OXA-17	100	6
SH5_00051	gb|AMB19637.1|ARO:3001555|OXA-368	100	6
SH5_00051	gb|AER57903.1|ARO:3001500|OXA-251	100	6
SH5_00051	gb|AHC31001.1|ARO:3001807|OXA-246	100	6
SH5_00051	gb|ABY79006.1|ARO:3001803|OXA-142	100	6
SH5_00051	gb|CAA80304.1|ARO:3001406|OXA-11	100	6
SH5_00051	gb|AFN20670.1|ARO:3001499|OXA-240	100	6
SH5_00051	gb|CCE94500.1|ARO:3001502|OXA-256	100	6

Table Continued

SH5_00051	gb|AAG45720.1|ARO:3001405|OXA-10	100	6
SH5_00051	gb|BAR45714.1|ARO:3003610|OXA-454	100	6
SH5_00051	gb|AID67109.1|ARO:3001495|OXA-233	100	6
SH5_00051	gb|AAF72942.1|ARO:3001423|OXA-28	100	6
SH5_00051	gb|ACO72579.1|ARO:3001801|OXA-147	100	6
SH5_00051	gb|CAH69530.1|ARO:3001798|OXA-74	100	6
SH5_00051	gb|AAD02245.1|ARO:3001414|OXA-19	100	6
SH5_00051	gb|AAR32651.1|ARO:3001795|OXA-56	100	6
SH5_00051	gb|CAA53242.1|ARO:3001402|OXA-7	100	6
SH5_00051	gb|CAP69660.1|ARO:3001811|OXA-129	100	6
SH5_00051	gb|ACN85419.1|ARO:3001804|OXA-145	100	6
SH5_00051	gb|CAL85435.1|ARO:3001439|OXA-101	100	6
SH5_00051	gb|ADV41666.1|ARO:3001475|OXA-183	100	6
SH5_00051	gb|AAK49460.1|ARO:3001429|OXA-35	100	6
SH5_00051	gb|CAA41211.1|ARO:3001400|OXA-5	100	6
SH5_00051	gb|AAC46344.1|ARO:3001408|OXA-13	100	6

3.6 Checking for virulence genes in the phage DNA

By comparing the Citrobacter phage DNA data with the Virulence Factor Database (VFDB), the phage genomes can be searched for any known virulence which it can then pass on to *Citrobacter freundii* that it attacks. A total of 25148 hits were returned by the BLASTn comparison. The percentage occurrence of matches at each level of percentage identity is summarized in a histogram and is demonstrated in Figure 20.

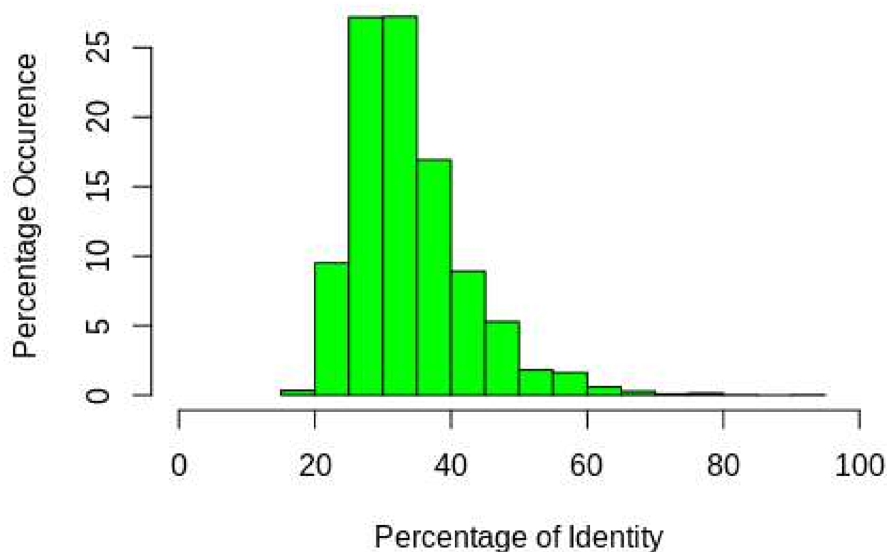


Figure 20 Histogram of the percentage identity in matches between the Citrobacter phage genomes and VFDB

From Figure 20, it can be observed that almost all of the 25148 hits show less than 70% identity. These matches are therefore not considered significant. However, 6 hits (0.024% of the total number of matches) showed above 90% identity. However, as can be seen in Table 4, these 6 hits all had an alignment length of 12 nucleotides. This is not a long enough alignment to consider these matches significant.

Table 4 Matches between Citrobacter phage DNA and VFDB with > 90% identity

Phage Sequence ID	VFDB Sequence ID	Percentage Identity (%)	Length of Alignment (bp)
Maleficent_00161	VFG014422(gi:152987829)	91.667	12
Maleficent_00161	VFG001252(gb|NP_249793)	91.667	12
Michonne_00162	VFG014422(gi:152987829)	91.667	12
Michonne_00162	VFG001252(gb|NP_249793)	91.667	12
Mordin_00007	VFG014422(gi:152987829)	91.667	12
Mordin_00007	VFG001252(gb|NP_249793)	91.667	12

3.7 Checking phages for possible lysogenic property

Prophages for *Citrobacter freundii* are identified using PhiSpy from *C. freundii* genomes. The phylogenetic tree shown in Figure 21 was generated from the combined set of both prophages and bacteriophages using Roary and iTOL.

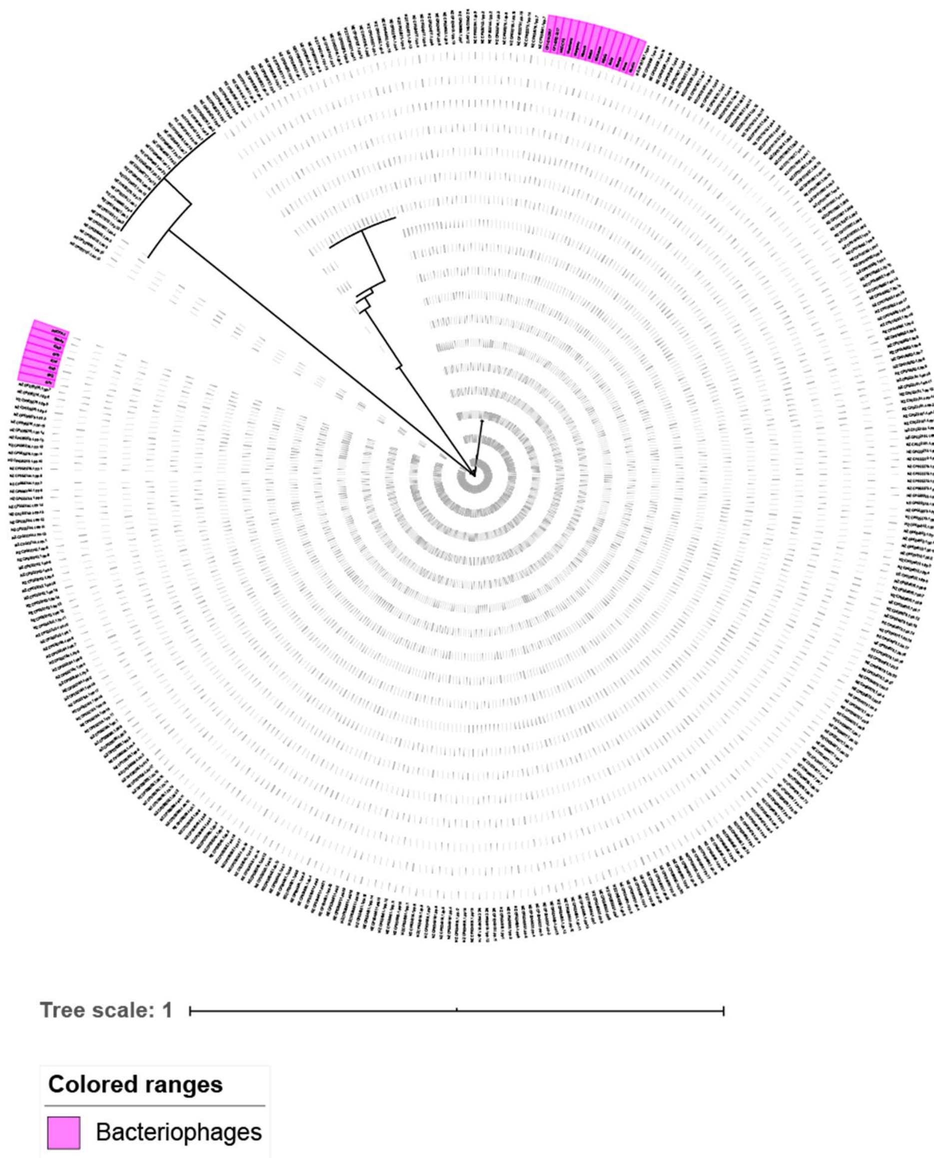


Figure 21 Phylogenetic tree of the combined set of both Citrobacter bacteriophages and prophages

From Figure 21, it can be seen that a clade with 294 leaves is formed, in which there are 273 prophages and all of the 21 bacteriophages studied in this thesis. This demonstrates a similarity of the bacteriophages with those prophages. Therefore possible lysogeny cannot be ruled out.

Chapter 4

Conclusions

Of the *C. freundii* phages listed in the NCBI database, 21 complete and unique genomes were found. Assessing these *C. freundii* phages according to the results of the BLASTn algorithm, the following conclusions can be made:

All the 21 phages studied have the feature that they attack *C. freundii* bacteria.

The comparison with the VFDB showed that none of the *C. freundii* phages under study indicates any evidence of the presence of virulence factor as per the VFDB.

When assessed against the CARD, the phages are found not to prevail any features of antibiotic resistance either.

A combined interpretation of the above conclusions indicates that the *C. freundii* phages studied are viable candidates for a potential alternative to antibiotics.

Possible lysogeny of the bacteriophages is suspected.

However, this study has been conducted solely on a database analysis approach, further lab-based and clinical research would be required before capitalizing the above-mentioned potentials. Moreover, further research in this area should also keep into consideration the possible adverse effect and/or bacterial resistance.

References

1 Akhter S., Aziz R. K., Edwards R. A., (2012) PhiSpy: a novel algorithm for finding prophages in bacterial genomes that combines similarity- and composition-based strategies. *Nucleic Acids Research* 40 (16): e126.

2 Alcock *et al.* 2020. CARD 2020: antibiotic resistome surveillance with the Comprehensive Antibiotic Resistance Database. *Nucleic Acids Research*, 48, D517-D525.

3 Balaban, N., T. Goldkorn, R. T. Nhan, L. B. Dang, S. Scott, R. M. Ridgley, A. Rasooly, S. C. Wright, J. W. Larrick, R. Rasooly, and Carlson, J. R., (1998), Autoinducer of virulence as a target for vaccine and therapy against *Staphylococcus aureus*. Science 280:438.

4 Baral P, Neupane S, Marasini BP, Ghimire KR, Lekhak B, Shrestha B., (2012) High prevalence of multidrug resistance in bacterial uropathogens from Kathmandu, Nepal. *BMC Res Notes,* 5(38)

5 Basavaraj C. M., P. Jyothi, and Basavaraj V. P. (2013), Antibiotic resistance in Citrobacter spp. isolated from urinary tract infection, *Urol Ann.*, 5(4), 312–313.

6 Baysarowich, J., K. Koteva, D. W. Hughes, L. Ejim, E. Griffiths, K. Zhang, M. Junop, and Wright, G. D. (2008), Rifamycin antibiotic resistance by ADP-ribosylation: structure and diversity of Arr., *Proc. Natl. Acad. Sci. U. S. A.*, 105, 4886-4891.

7 Brüssow H., Canchaya C., Hardt W.D., (2004) Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion, *Microbiology and Molecular Biology Reviews*, 68(3), 560-602

8 Canton, R. (2009), Antibiotic resistance genes from the environment: a perspective through newly identified antibiotic resistance mechanisms in the clinical setting. *Clin. Microbiol. Infect.* 15 (1), 20-25.

9 Casey, R. M. (2005), BLAST Sequences Aid in Genomics and Proteomics, *Business Intelligence Network*.

10 Davies, J. and Davies, D. (2010), Origins and Evolution of Antibiotic Resistance, *Microbiology And Molecular Biology Reviews*, 74 (3), 417–433.

11 D'Costa, V. M., K. M. McGrann, D. W. Hughes, and Wright, G. D. (2006), Sampling the antibiotic resistome. *Science*, 311, 374-377.

12 DeLeo, F. R., and Chambers, H. F. (2009), Reemergence of antibiotic-resistant *Staphylococcus aureus* in the genomics era., *J. Clin. Invest.*, 119, 2464-2474.

13 Finlay, B. B., and R. E. Hancock (2004), Can innate immunity be enhanced to treat microbial infections?, *Nat. Rev. Microbiol.*, 2, 497-504.

14 Guerin, E., G. Cambray, N. Sanchez-Alberola, S. Campoy, I. Erill, S. Da Re, B. Gonzalez-Zorn, J. Barbé, M. C. Ploy, and D. Mazel (2009), The SOS response controls integron recombination, *Science*, 324:1034.

15 Guilfoile, P. G., & Hutchinson, C. R. (1991). A bacterial analog of the mdr gene of mammalian tumor cells is present in Streptomyces peucetius, the producer of daunorubicin and doxorubicin. *Proceedings of the National Academy of Sciences of the United States of America*, *88*(19), 8553–8557.

16 Hacker, J., and J. B. Kaper (2000), Pathogenicity islands and the evolution of microbes. *Annu. Rev. Microbiol*, 54, 641-679.

17 Hakenbeck, Regine (1998), Mosaic genes and their role in penicillin-resistant Streptococcus pneumoniae (minireview), *Electrophoresis,* 19 (4), 597 – 601.

18 Harrison EM, Paterson GK, Holden MT *et al* (2013), Whole genome sequencing identifies zoonotic transmission of MRSA isolates with the novel mecA homologue mecC, *EMBO Mol Med*, 5, 509–515.

19 Hawkey, P. M., and A. M. Jones (2009), The changing epidemiology of resistance. *J. Antimicrob. Chemother.*, 64(1), i3-i10.

20 Hendrix, R. W., Lawrence, G. J., Hatfull G. F., and Casjens S., (2000), The origins and ongoing evolution of viruses, *Trends in Microbiology*, 8(11), 504-508

21 Hibma, A. M. Jassim, S. A.A., Griffiths, M. W. (1997), Infection and removal of L-forms of Listeria monocytogenes with bred bacteriophage, *Food Microbiology*, 34, 197-207.

22 Hodgkin D. C. (1949) The X-ray analysis of the structure of penicillin. *Adv Sci.*, 6(22):85–89.

23 Jassim, G., Skrede, S., Vázquez, M. J., Wergedal, H., Vik-Mo, A. O., Lunder, N., … Fernø, J. (2012). Acute effects of orexigenic antipsychotic drugs on lipid and carbohydrate metabolism in rat. *Psychopharmacology*, *219*(3), 783–794.

24 Jassim, S. A. A. and Limoges, R. G. (2014) Natural solution to antibiotic resistance: bacteriophages 'The Living Drugs' Review paper in *World J Microbiol Biotechnol* 30, 2153–2170.

25 Jia, B., Raphenya, A. R., Alcock, B., Waglechner, N., Guo, P., Tsang, K. K., Lago, B. A., Dave, B. M., Pereira, S., Sharma, A. N., Doshi, S., Courtot, M., Lo, R., Williams, L. E., Frye, J. G., Elsayegh, T., Sardar, D., Westman, E. L., Pawlowski, A. C., Johnson, T. A., … McArthur, A. G. (2017). CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. *Nucleic acids research*, *45*(D1), D566–D573.

26 Laxminarayan, R., Duse, A., Wattal, C., Zaidi, A.K.M., Wertheim, H. F. L., Sumpradit, N., Vlieghe, E., Hara, G. L., Gould, I. M., Goossens, H., Greko, C., D So, A., Bigdeli, M., Tomson, G., Woodhouse, W., Ombaka, E., Peralta, A. Q., Qamar, F. N., Mir, F., Kariuki, S., Bhutta, Z. A., Coates, A., Bergstrom, R., Wright,

G. D., Brown, E. D. and Cars, O. (2013). Antibiotic resistance—the need for global solutions. *The Lancet Infectious Diseases*, 13(12), 1057–1098.

27 Lihong Chen, Jian Yang, Jun Yu, Zhijian Yao, Lilian Sun, Yan Shen, Qi Jin (2005), VFDB: a reference database for bacterial virulence factors, *Nucleic Acids Research*, 33 (1), D325–D328.

28 Loc-Carrillo C, Abedon S. T., (2011) Pros and cons of phage therapy. *Bacteriophage*. 1(2), 111–114.

29 Lu L (2004). Autoinducer 2-based quorum sensing response of Escherichia coli to subtherapeutic tetracycline exposure. Retrieved from: http://repository.tamu.edu/bitstream/handle/1969.1/4198/etd-tamu2004B-FSTC-Lu.pdf?sequence=1

30 MacIntyre, H., Geider, R., & Miller, D. (1996). Microphytobenthos: The Ecological Role of the "Secret Garden" of Unvegetated, Shallow-Water Marine Habitats. I. Distribution, Abundance and Primary Production. *Estuaries, 19*(2), 186-201.

31 McGrath S and van Sinderen D (editors). (2007). Bacteriophage: Genetics and Molecular Biology *(1st ed.). Caister Academic Press.* ISBN 978-1-904455-14-1.

32 Morowitz, Michael J.; Denef, Vincent J.; Costello, Elizabeth K.; Thomas, Brian C.; Poroyko, Valeriy; Relman, David A.; Banfield, Jillian F. (2011-01-18). Strain-resolved community genomic analysis of gut microbial colonization in a premature infant, *Proceedings of the National Academy of Sciences*. 108 (3): 1128–1133.

33 Murray PR, Holmes B, Aucken HM. London: Edward Arnold (2010), Citrobacter, Enterobacter, Klebsiella, Plesiomonas, Serratia, and other members of the enterobacteriaceae, *Topley and Wilson's microbiology and microbial infections*.

34 Page SW, Gautier P (2012) Use of antimicrobial agents in livestock. *Rev Sci Tech Off Int Epiz* 31(1):145–188.

35 Puchenkova, S. G. (1996). "Enterobacteria in areas of water along the Crimean coast". *Mikrobiolohichnyi Zhurnal,* 58 (2): 3–7

36 Qin, J., R. Li, J. Raes, M. Arumugam, K. S. Burgdorf, C. Manichanh, T. Nielsen, N. Pons, F. Levenez, T. Yamada, D. R. Mende, J. Li, J. Xu, S. Li, D. Li, J. Cao, B. Wang, H. Liang, H. Zheng, Y. Xie, J. Tap, P. Lepage, M. Bertalan, J.-M. Batto, T. Hansen, D. L. Paslier, A. Linneberg, H. B. Nielsen, E. Pelletier, P. Renault, T. Sicheritz-Ponten, K. Turner, H. Zhu, C. Yu, S. Li, M. Jian, Y. Zhou, Y. Li, X. Zhang, S. Li, N. Qin, H. Yang, J. Wang, S. Brunak, J. Doré, F. Guarner, K. Kristiansen, O. Pedersen, J. Parkhill, J. Weissenbach, M. Consortium, P. Bork, S. D. Ehrlich, and J. Wang (2010), A human gut microbial gene catalogue established by metagenomic sequencing. *Nature,* 464:59-65.

37 Rizvi M, Fatima N, Shukla I, Malik A. (2010), Epidemiology of extended spectrum β-lactamases in *Serratia* and *Citrobacter* species in North India. *Indian J Pathol Microbiol.*, 53, 193-4.

38 Schloissnig, Siegfried; Arumugam, Manimozhiyan; Sunagawa, Shinichi; Mitreva, Makedonka; Tap, Julien; Zhu, Ana; Waller, Alison; Mende, Daniel R.; Kultima, Jens Roat (2013). Genomic variation landscape of the human gut microbiome, *Nature*. 493 (7430): 45–50.

39 Seemann T., (2012)Prokka: rapid prokaryotic genome annotation, *Bioinformatics,* 15;30(14):2068-9. PMID:24642063

40 Shobha KL, Gowrish Rao S, Rao S, Sreeja CK. (2007), Prevalence of extended spectrum Beta-Lactamases in urinary isolates of *Escherichia coli, Klebsiella* and *Citrobacter species* and their antimicrobial susceptibility pattern in a tertiary care hospital, *Indian J Practising Doctor*, 3.

41 Stewart, G. R., R. N. Perry, and D. J. Wright (1993) Studies on the amphid specific glycoprotein gp32 in different life-cycle stages of Meloidogyne species, *Parasitology*, 107 (5): 573-578.

42 Stokstad ELR, Jukes TH (1949) Further observations on the animal protein factor. *Proc Soc Biol Exp Med* 73:523–528

43 Sulakvelidze, A., Alavidze, Z., & Morris, J. G., Jr (2001). Bacteriophage therapy. *Antimicrobial agents and chemotherapy*, *45*(3), 649–659. doi:10.1128/AAC.45.3.649-659.2001

44 Wang, J. T.; Chang, S. C.; Chen, Y. C.; Luh, K. T. (2000). Comparison of antimicrobial susceptibility of Citrobacter freundii isolates in two different time periods, *Journal of Microbiology, Immunology, and Infection,* 33 (4): 258–262.

45 Werkman, C.H. and Gillen, G.F. (1932), Bacteria producing trimethylene glycol. *Journal of Bacteriology*, 23, 167-182.

46 Whalen, J. G.; Mully, T. W.; English, J. C. (2007). Spontaneous Citrobacter freundii Infection in an Immunocompetent Patient, *Archives of Dermatology*. 143 (1): 124–125.

47 Wright, G. D. (2007), The antibiotic resistome: the nexus of chemical and genetic diversity, *Nat. Rev. Microbiol.*, 5, 175-186.

48 Zhang R, Yang L, Cai JC, Zhou HW, Chen GX. (2008), High-level carbapenem resistance in a *Citrobacter freundii* clinical isolate is due to a combination of KPC-2 production and decreased porin expression, *J Med Microbiol.*, 57, 332–7.