ISOLATION AND CHARACTERIZATION OF BACTERIOPHAGE FROM RAW SEWAGE SPECIFIC FOR *Escherichia coli* O157:H7

SITI FARIZA BT JUHARUL ZAMAN

UNIVERSITI SAINS MALAYSIA
2014

ISOLATION AND CHARACTERIZATION OF BACTERIOPHAGE FROM RAW SEWAGE SPECIFIC FOR *Escherichia coli* O157:H7

by

SITI FARIZA BT JUHARUL ZAMAN

Thesis submitted in fulfillment of the requirements for the degree of Master of Science

September 2014

ACKNOWLEDGEMENT

First of all, my utmost gratitude and appreciation go to my main supervisor Associate Professor Dr. Yahya Mat Arip for his patience, concern, moral support, encouragement, assistance and immense knowledge in the completion of this research work. This dissertation would not have been possible without his steadfast guidance in all the time of research and writing of this thesis.

My sincere thanks also goes to my fellow laboratory colleagues from Lab 218 for the stimulating discussions, help, guidance and scientific advice during research period. My appreciation also extends to my close friends. Encouragement and numerous supports they gave have been valuable in my difficult times.

I place on record, my thanks to all the staffs of School of Biological Sciences, USM especially from Electron Microscope Unit, Ms. Faizah and Mr. Masrul for their assistance in electron microscopy. Besides, I would like to thank the Institute of Postgraduate Studies staffs who directly or indirectly have lent their helping hand in this venture.

I would like to express my deep gratitude to my parents and relatives for giving me the strength, unceasing encouragement, endless love and support in the pursuit of these studies throughout my life.

This study was financially supported by research grant, USM graduate assistant and MyMaster.

TABLE OF CONTENTS

CHAPTER 3 – MATERIALS AND METHODS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

APS	Ammonium persulfate
ATCC	American Type Culture Collection
BLASTn	Basic Local Alignment Search Tool-nucleotide
bp	Base pair
ddH_2O	Double distilled water
DNA	Deoxyribonucleic acid
DNase I	Deoxyribonuclease I
dNTP	Deoxynucleotide triphosphates
dsDNA	Double stranded deoxyribonucleic acid
E. coli	*Escherichia coli*
EB	Elution buffer
EDTA	Ethylenediaminetetraacetic acid
EHEC	Enterohaemorrhagic *E. coli*
FDA	Food and Drug Administration
ICTV	International Committee on Taxonomy of Viruses

kb	Kilobase pair
kDa	Kilodalton
LB	Luria-Bertani
$MgCl_2$	Magnesium chloride
NaCl	Sodium chloride
$NaOAc.3H_2O$	Sodium acetate trihydrate
NCBI	National Center for Biotechnological Information
NEB	New England BioLabs
nm	Nanometer
OD	Optical density
ORFs	Open reading frames
PCR	Polymerase chain reaction
pfu	Plaque forming unit
RNA	Ribonucleic acid
RNase A	Ribonuclease A
rpm	Revolutions per minute
SDS	Sodium dodecyl sulphate

SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
ssDNA	Single stranded deoxyribonucleic acid
STEC	Shiga toxin- producing *E. coli*
Taq	*Thermus aquaticus*
TBE	Tris-Borate-EDTA
TEM	Transmission electron microscope
TEMED	Tetramethylethylenediamine
TMS	Tris-Magnesium-Sodium
Tris base	Tris (hydroxymethyl)-aminomethane
Tris-HCl	Tris hydrochloric acid
tRNA	Transfer ribonucleic acid
UTIs	Urinary tract infections
VTEC	Verotoxin-producing *E. coli*
w/v	Weight/volume

LIST OF APPENDICES

LIST OF SYMBOLS

Φ	Phi
®	Registered trademark
™	Trademark
β	Beta

PENGASINGAN DAN PENCIRIAN BAKTERIOFAJ DARI SISA KUMBAHAN KHUSUS PADA *Escherichia coli* O157:H7

ABSTRAK

Faj khusus pada *E. coli* O157:H7 telah berjaya diasingkan untuk pertama kalinya di Malaysia dari kemudahan sisa kumbahan dalam kampus Universiti Sains Malaysia di Pulau Pinang. Berdasarkan kajian morfologi, faj ini dipercayai adalah faj-menyerupai T4 yang tergolong dalam keluarga *Myoviridae*; begitu juga seperti faj khusus pada *E. coli* O157:H7 lain yang pernah diasingkan sebelum ini. Ciri fizikokimia faj ini menunjukkan ia dapat menjangkiti bakteria pada julat suhu daripada 10 °C kepada 37 °C, julat pH dari pH 5 hingga pH 10 dan julat kepekatan garam dari 0.17 M kepada 0.3 M. Faj khusus pada *E. coli* O157:H7 yang telah diasingkan ini mempunyai spektrum tuan rumah yang sempit kerana ia hanya dapat menjangkiti satu hanya satu strain *E. coli* (*E. coli* ATCC 13706), daripada dua belas bacteria yang berbeza (*Enterobacteriaceae* dan bukan *Enterobacteriaceae*) yang diuji. Kajian separa genomik menunjukkan ia mempunyai perkongsian identiti yang tinggi dengan Enterobakteria faj RB69, dan HX01 yang masing-masing telah diasingkan dari sisa kumbahan di Amerika Syarikat dan najis itik di China. Yang menghairankan, sel rumah bagi kedua-dua faj adalah bukan *E. coli* O157:H7 iaitu *E. coli* strain B untuk RB69 dan avian patogenik *E. coli* (APEC) untuk HX01. Perbandingan genomik selanjutnya antara faj yang diasingkan dengan RB69 (sama dengan kebanyakan urutan klon) menunjukkan corak profail enzim penghadaman yang berbeza walau pun kedua-duanya adalah faj-menyerupai T4 yang tergolong

dalam keluarga *Myoviridae*. Di samping itu, analisis protein separa menunjukkan bahawa faj yang diasingkan ini mempunyai profail protein yang berbeza daripada faj T4 dan T7, dua faj lazim berekor. Oleh itu, kajian ini menyediakan potensi pertambahan kepada faj yang terasing, khususnya faj khusus kepada *E. coli* O157:H7 dari sisa kumbuhan daripada Malaysia. Kajian berkenaan ciri-ciri faj ini berkemungkinan menyumbang kepada pengetahuan yang boleh digunakan untuk pembangunan agen kawalan bio terhadap *E. coli* O157:H7.

ISOLATION AND CHARACTERIZATION OF BACTERIOPHAGE FROM RAW SEWAGE SPECIFIC FOR *Escherichia coli* O157:H7

ABSTRACT

E. coli O157:H7-specific phage was successfully isolated for the first time in Malaysia, from a sewage facility of Universiti Sains Malaysia campus in Penang. Based on morphological study, the isolated phage was suggested to be a T4-like phage belonging to *Myoviridae* family; similar to other *E. coli* O157:H7-specific phages previously isolated. Physicochemical properties of the isolated phage indicate infective (able to replicate) at temperature range from 10 °C to 37 °C, pH range from pH 5 to pH 10 and salt concentration range from 0.17 M to 0.3 M. The isolated *E. coli* O157:H7-specific phage had a narrow host range as it was able to infect only one strain of *E. coli* (*E. coli* ATCC 13706), out of twelve different bacteria (*Enterobacteriaceae* and non-*Enterobacteriaceae*) tested. Partial genomic studies demonstrated high degree of identity sharing with Enterobacteria phage RB69 and HX01 which was isolated from raw sewage in the U.S. and duck faeces in China, respectively. The host for both phages are non *E. coli* O157:H7 which is *E. coli* B strain for RB69 and avian pathogenic *E. coli* (APEC) strains, for HX01. Further genomic comparison between the isolated phage and RB69 (similar with most of clone sequences) showed different restriction enzyme pattern profiling though both of them are T4-like phage in the same family, *Myoviridae*. Besides, partial protein analysis revealed that the isolated phage displayed distinctive protein profile compared with phage T4 and T7. Hence, this study provides a potential addition to

the growing number of phages discovered, specifically *E. coli* O157:H7-specific phages from raw sewage from Malaysia. The studies on its characterizations may provide knowledge that could be useful for the development of bio-control agent against *E. coli* O157:H7.

CHAPTER 1

INTRODUCTION

Bacteriophages or phages for short are viruses infecting specific bacteria. Phages are among the most common biological entities on earth and are found in all habitats in the world where bacteria and archaea proliferate (Clokie *et al*., 2011). Being the most widely distributed biological entity in the biosphere, phage population is greater than 10^{31} or approximately 10 million per cubic centimeter (Kwiatek *et al*., 2012). Recent estimates suggest that there exist globally ~100 million phage species; however, only a small fraction of phages have so far been characterized with around 6000 have been identified and reported towards the end of last century (Ackermann, 2000). Thus, this means, many phages are waiting to be discovered.

The notorious *E. coli* O157:H7 is an enterohaemorrhagic strain of *E. coli* (EHEC) recognized as the most important EHEC causing hemorrhagic diarrheal and kidney failure via food contamination (Goncuoglu *et al*., 2010). The bacteria could be found in the lower intestinal tracts of human, free-living animals and warm-blooded organisms (Vogt & Dippold, 2005). The bacterium is also found in water, foods and soil due to contamination of faecal or during animal slaughter (Schroeder *et al*., 2002).

Among the discovered phages, they are phages specific to *E. coli* O157. Up till now, there are more than fifty *E. coli* O157-specific phages have been discovered by previous researchers and twenty four of them are highly specific against *E. coli* O157:H7. However, only six of the *E. coli* O157:H7-specific phages have been

isolated from Asia regions and the rest are from North America countries. Majority of the isolated *E. coli* O157:H7-specific phages are from faecal sample with one from salt water sample and two from industrial wastewater. Currently, there is no record of *E. coli* O157:H7-specific phage ever been isolated from Southeast Asia region. Therefore, an attempt was made to isolate *E. coli* O157:H7-specific phage from raw sewage sample of sewage treatment facility in Penang, Malaysia.

Every *E. coli* O157:H7-specific phages isolated so far shows variations, as well as, similarities among them that contribute to phage diversities. Hence, the isolated *E. coli* O157:H7-specific phage from raw sewage in Penang, Malaysia could as well possibly show variations and similarities to previously isolated *E. coli* O157:H7-specific phages and might have the potential as an addition to the ICTV database. The basic understanding of phage biology of the isolated *E. coli* O157:H7-specific phage could be useful for the development of bio-control agent against *E. coli* O157:H7. Due to the emergence of antibiotic resistant bacteria, natural control strategies have received growing demand and attention including the application of phages as bio-control agents (Coffey *et al*., 2011).

Thus, the main purposes of this project were to isolate and characterize *E. coli* O157:H7-specific phage from raw sewage sample. The specific objectives of this work were:

1) To isolate *E. coli* O157:H7-specific phage from raw sewage.
2) To characterize the isolated *E. coli* O157:H7-specific phage based on:
    a) morphological study.
    b) physical chemical attributes (temperature, pH and salinity).
    c) phage-host interaction specificity.
    d) partial molecular identification using genomic and proteomic approaches.

CHAPTER 2

LITERATURE REVIEW

2.1 Viruses in general

The word *virus* came from the Latin meaning "slimy liquid" or "poison" referring to poisonous and lethal substance (Pelczar *et al*., 2010; Black, 2012). Viruses are often defined as obligate intracellular parasites that can only replicate dependently inside the host organisms (Koonin *et al*., 2006). Viruses could have only one type of genetic material, either DNA or RNA, which depend upon hosts to carry out their replication cycles for the production of new virions. They would inject their genomes into suitable living host cells via inhalation, direct contact and ingestion (Madigan *et al*., 2010). Since viruses have no ability to metabolize on their own, they have the capabilities of becoming parasites on the host cells for almost all of their life-sustaining functions. Once they are inside, they would gain control of the hosts and produce all necessary molecules before assembling and releasing new virions that lead to the disruption in cell functions (Rybicki, 1990; Clark & March, 2006).

Viruses are thought to be the smallest form of entities on earth and they do not respire, grow or divide. They are measured in nanometer (nm) compare to bacteria which is in micrometer (μm) size. Suffice to say, viruses are 100 times smaller than bacteria (Shors, 2013). By reason of their sizes, viruses cannot be observed with a basic optical microscope, hence, scanning and transmission electron microscopes are the only way to visualize them (Collier, 2011). Overall, majority of

viruses fall in the range of 30 to 90 nm in measurement. However, the largest known virus is *Mimivirus* with the size of could be up to 400 nm while *Parvovirus*, considered as one of the smallest viruses, could be measured as small as 18 nm in dimension (Dimmock *et al*., 2007; Shors, 2013).

The kinds of genomes separate the viruses into two main groups which are DNA viruses and RNA viruses. Each group is further topologically divided into single-stranded or double-stranded, linear or circular forms (Metzler & Metzler, 2001; Madigan *et al*., 2010). These genome types would depend on the viruses, which made them unique and different from other organisms. The basic structure of a virus is shown in Figure 2.1.

In viral taxonomy, viruses are grouped according to their equivalence properties such as size, nucleic acid type and topology, capsid structure and symmetry, presence or absence of an envelope, host range and immunological characteristics (Christian, 2002). They are classified into two complementary systems for standardize identification purposes. In 1996, the International Committee on Taxonomy of Viruses (ICTV) has established a single comprehensive scheme for classification of all viruses into order, family, genera and species based on Linnaean hierarchy system with current standing at 7 orders and 96 families (Hurst, 2000; Delwart, 2007; King *et al*., 2011). On the other hand, the Baltimore system provides a helpful guide in virus classification based on the unique method of viral genome replication strategy (Christian, 2002; Hogan *et al*., 2005) that categorize viruses into seven different classes based on virus's nucleic acid type and topology (Dimmock *et al*., 2007).

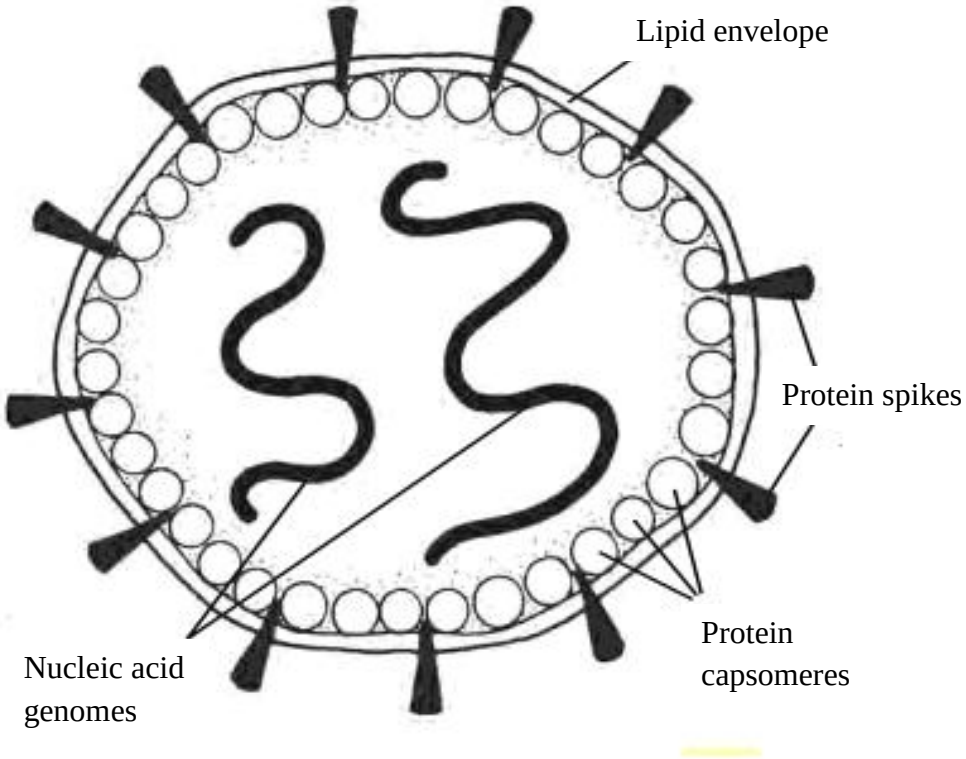

Figure 2.1: Basic structure of a virus. The nucleic acid genomes could be either DNA or RNA. The nucleic acid genome is protected by protein coat or capsid that is made up of a finite number of protein subunits called capsomeres. A lipid membrane or envelope provides additional protection to the nucleic acid genome. The presence of protein spikes embedded in the envelope serve as attachment point to the host cell (Williams, 2002).

2.2 Bacteriophages

Bacteriophages or phages for short are bacterial viruses that are highly specific in their host-cell recognition infecting only targeted bacteria species or strains (Clark & March, 2006; Hagens & Loessner, 2007; Hanlon, 2007; Nishikawa *et al*., 2008; Viazis *et al*., 2011). They are also considered as natural predators of bacteria that cause lysis of the infected host cells (Abuladze *et al*., 2008; Nishikawa *et al*., 2008).

ICTV presently classifies viruses into 7 orders and 96 families. Within this system, phage is placed into only one order, *Caudovirales*, 13 families and 30 genera (Ackermann, 2003; Ackermann, 2011). The prominent members of the *Caudovirales* are grouped into three large families: *Myoviridae, Siphoviridae* and *Podoviridae*. All phages constituted in these families have non-enveloped icosahedral heads but differ in their tail length and contractile ability (Ackermann, 1998). Up to now, most of the identified phages are tailed phages with isometric heads containing double-stranded DNA (Ackermann, 2003; Hagens & Loessner, 2007; Ackermann; 2011).

Phages belong to *Myoviridae* family are characterized by their long contractile tails consisting of a sheath (Ackermann, 2003; O'Flaherty *et al*., 2004; Ackermann, 2011). Examples of phages in this family are T4, P1, P2, SP01 and Mu-like viruses (Ackermann, 2003; O'Flaherty *et al*., 2004; Lavigne *et al*., 2009). The genome size of these phages distinctly varies but a complete genome sequence has

revealed that the T4-related phages represent one of the largest phages (Lavigne *et al*., 2009).

Among the tailed phages, 61% have long and non-contractile tails which belong to *Siphoviridae* (Ackermann, 2003). Examples of phages in this family are lambda (λ) and T5-like viruses (Ackermann, 1998; Grabow, 2001; Ackermann, 2003; Ackermann, 2011). Besides, the majority of the known tailed phages belong to this family (Ackermann, 2003).

Unlike the other families, *Podoviridae* phages have short and non-contractile tails (Ackermann, 1998; Grabow, 2001; Ackermann, 2003; Ackermann, 2011) such as T7-like viruses.

Figure 2.2 shows the comparison in structure of these three families *Myoviridae*, *Siphoviridae* and *Podoviridae*. Based on the ICTV classification, the phages are placed according to their respective order, families, genome type and size as shown in Table 2.1.

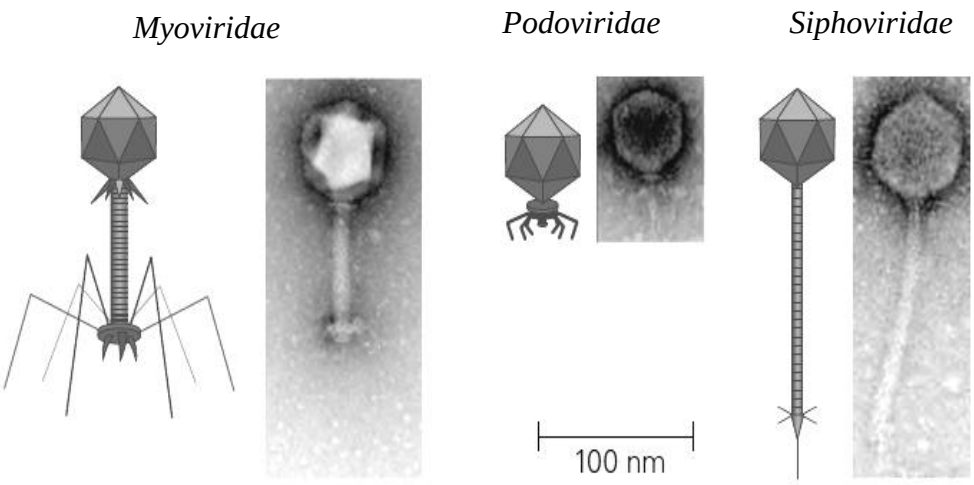

Figure 2.2: Comparison of three family members of *Caudovirales*; *Myoviridae*, *Podoviridae* and *Siphoviridae* families (Harper, 2011).

Table 2.1 ICTV classification of phages (Harper, 2011).

Virus family	Genome type	Genome size (kb)	Structure	Example
Caudovirales				
Myoviridae	dsDNA	33.6-170	Non-enveloped, icosahedral head (50-110 nm, may be elongated) with long contractile tail	*Enterobacteria* phage T4
Podoviridae	dsDNA	40-42+	Non-enveloped, icosahedral head (60 nm) with short, non-contractile tail	*Enterobacteria* phage T7
Siphoviridae	dsDNA	48.5	Non-enveloped, icosahedral head (60 nm) with long, non-contractile tail	*Enterobacteria* phage λ
Other families				
Tectiviridae	dsDNA	147-157	Icosahedral, contains lipid, 63 nm with 20 nm spikes	*Enterobacteria* phage PRD1
Corticoviridae	dsDNA	9-10	Icosahedral, contains lipid 60 nm+	*Pseudoalteromonas* phage PM2
Plasmaviridae	dsDNA	12	Enveloped, spherical/pleomorphic, 80 nm	*Acholeplasma* phage L2
Inoviridae	ssDNA	4.4-8.5	Non-enveloped, filamentous, 6-8 nm x 760-1950 nm	*Enterobacteria* phage M13
Microviridae	ssDNA	4.4-5.4	Non-enveloped, icosahedral, 25-27 nm	*Enterobacteria* phage ϕX174
Leviviridae	ssDNA	3.4-4.2	Non-enveloped, icosahedral, 26 nm	*Enterobacteria* phage MS2
Cystoviridae	dsRNA (segmented)	13.4 (3segments)	Enveloped, spherical, 86 nm with 8 nm spikes	*Pseudomonas* phage ϕ6

2.2.1 The lytic and lysogenic cycle

Different bacteriophage populations undergo different life cycles depending on the kind of infection cycle and mode of replication they use to carry their genome into the host (Marsh & Wellington, 1994; Rao, 2006; Courchesne *et al*., 2009). Following the initial infection, there are two categories of bacteriophages; lytic (virulent) or lysogenic (temperate). Lytic bacteriophages lyse the cells they infect and produce phage progeny for further infection while lysogenic bacteriophages establish an unapparent and continual infection without killing the host cell (Rao, 2006; Chaudari, 2014). Furthermore, virulent phages can only replicate by means of lytic cycle, while temperate phages are able to replicate in both lytic and lysogenic cycles. A key difference between lytic and lysogenic cycles is that the lytic phage multiplies the viral DNA by a production of infectious individual phage progeny and infects other cells while the lysogenic phage reproduces the viral DNA by prokaryotic production (Lodish *et al*., 2008).

The lytic cycle is one of the two reproductive cycles in which phage multiplies and ultimately ends in the death of the infected host cell by bursting and releasing virions. Lytic phages only undergo virulent infection and destroy the host cells as a normal part of their life cycle (Mayer, 2010). Subsequent to infecting the host cell, the virulent phages typically proceed with immediate replication of the virion prior to produce large numbers of new viruses (Rao, 2006).

As in Figure 2.3, the first stage of lytic infection is the penetration in which phage enters the host cell and culminating in the mRNA biosynthesis (Hanlon, 2007). The attachment of phage usually occurs through the interaction of the phage tails with variety of cell membrane surface components (Kropinski, 2006; Dimmock *et al*., 2007; Hanlon, 2007). After infection, the viral nucleic acids are copied by the host cell to produce necessary proteins (Kropinski, 2006). Basically, early mRNA is produced by transcription of viral genome using host cell RNA polymerase (Hanlon, 2007). The synthesized mRNAs are then translated by host cell ribosomes into proteins such as the capsid or tail proteins. In general, lytic phages take over the cell biosynthetic machinery by destroying the host genome and utilizing nucleotides in phage DNA replication (Kropinski, 2006). As soon as the nucleic acid is injected, the phage cycle is followed by the synthesis of phage components, late proteins, assembly and mature phage (Rao, 2006). Due to the accumulation of the phage particles within the host, the cell capacity is full and consequently bursts open the cell wall (Rao, 2006; Chaudari, 2014). Hence, this process is known as lysis and release phase (Rao, 2006; Mayer, 2010).

Similar to that of lytic cycle, lysogenic (temperate) phages begin the cycle with the adsorption of nucleic acids upon entering the host cell (Campbell & Reece, 2005; Fortuna *et al*., 2008). In this cycle alternatively, phages do not necessarily enter a lytic cycle but instead results in the integration of the phage DNA into the host chromosome forming a non-infectious phage genetic material called prophage (Figure 2.3) (Grabow, 2001; Hanlon, 2007, Mayer, 2010; McNair *et al*., 2012). Most of the phage genomes are capable of maintaining their chromosome in stable, dormant or silent within host cell during this period (Mayer, 2010). Furthermore, in

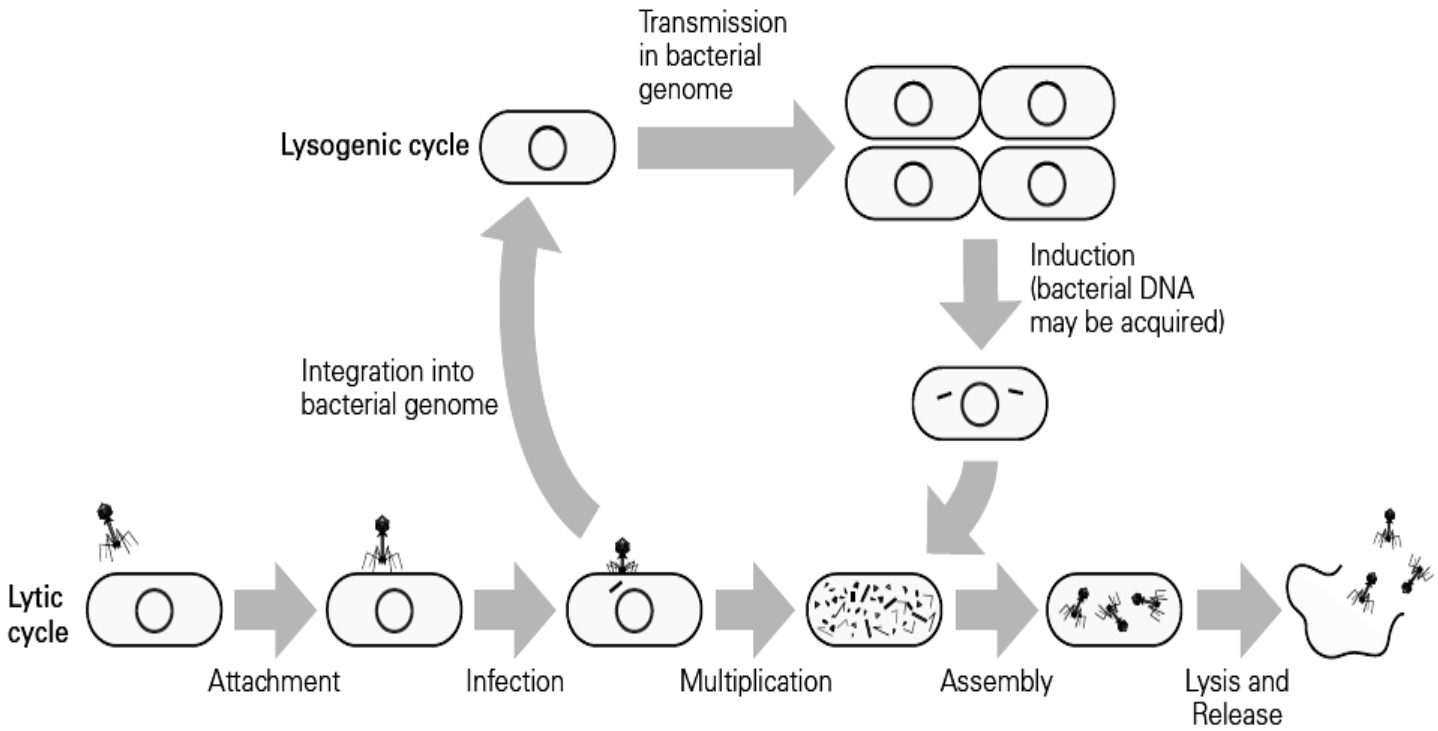

Figure 2.3: The lytic and lysogenic pathways of bacteriophage (Harper, 2011).

this quiescent state, the genetic material is not transcribed but instead replicated simultaneously with the bacterial DNA in the cytoplasm of host cell without killing it (Grabow, 2001; Hanlon, 2007; Fortuna *et al*., 2008; Mayer, 2010). As the host cell reproduces, the prophage is copied and this integrated genetic material is transmitted to the daughter cells accordingly to each successive cell division (Mayer, 2010). Subsequently, each daughter cell may continue several rounds of replication for many generations with the prophage existing in every cell (Hanlon, 2007).

Occasionally, these lysogens are able to remain in dormant state until they become active through induction (Campbell & Reece, 2005). Lysogenic phages can be spontaneously directed to the lytic cycle by subjecting them to adverse conditions or stress such as dessication, ultraviolet light (UV) irradiation, mutagenic agent exposure and environmental stressors (Rao 2006; Fortuna *et al*., 2008; McNair *et al*., 2012). These conditions trigger the termination of lysogenic state which eventually causes cell lysis and initiates release of progeny phages (Grabow, 2001; Rao, 2006; Hanlon, 2007).

2.2.2 Phage history

The discovery of phages could be traced back to the late 1910's. In 1915, Frederick William Twort, a British pathologist was the first one who independently discovered the antibacterial potential of phages and later by the French-Canadian microbiologist, Felix d'Herelle in 1917 at the Pasteur Institute, Paris. Both pioneer researchers had given an account of a filterable and transmissible entity which able to kill bacteria culture and claimed that specific bacterial growth could be inhibited by the addition of bacteria-free filtrates (Grabow, 2001; Gravitz, 2012; Lavigne & Robben, 2012).

It was d'Herelle who named the virus as "bacteriophage" or "bacteria eater", derived from the Greek word "phagein" meaning "to eat" (Gravitz, 2012). In addition, he was the first scientist to apply bacteriophage against bacterial infections and this concept is also known as phage therapy. Since then, phage therapy was extensively developed in many places (Kutateladze & Adamia, 2008). Regardless of the intensive use, this treatment and clinical applications were not completely accepted and subsequently abandoned in the West due to the emergence of antibiotics in the 1940s (Nishikawa *et al*., 2008).

2.2.3 Phage distribution

Phages are the most numerous entities in the biosphere (McGrath & Sinderen, 2007; Fortuna *et al*., 2008; Liao *et al*., 2011). It is conservatively estimated that the total number of phages worldwide to be in the range of 10^{30} to 10^{31}, that is equal to

100 million to 1 billion phage particles exist globally (Kropinski, 2006; Hanlon, 2007; Courchesne *et al*., 2009; McNair *et al*., 2012). Thus, they are approximately ten times more diverse than bacteria making them the most abundant in microbial communities (Marsh & Wellington, 1994; Kropinski, 2006; Hanlon, 2007; McNair *et al*., 2012). Out of this estimation, only a small fraction which is less than ten thousands of them has been identified so far (Courchesne *et al*., 2009; McNair *et al*., 2012). Therefore, there are enormous numbers of phages have yet to be discovered (Hanlon, 2007).

2.2.4 Phage morphology

The simplest morphology seen in phages is similar to other viruses that they have capsids protecting the nucleic acids (Hanlon, 2007). As seen in other viruses, certain phages could have protrusion proteins on the surface as well. Yet, there are phages with long tails and present of appendages (Mayer, 2010; Chaudari, 2014). A typical head and tail phage is shown in Figure 2.4 with size in the range of 20-200 nm in length and 80- 100 nm in width (Rao, 2006; Mayer, 2010).

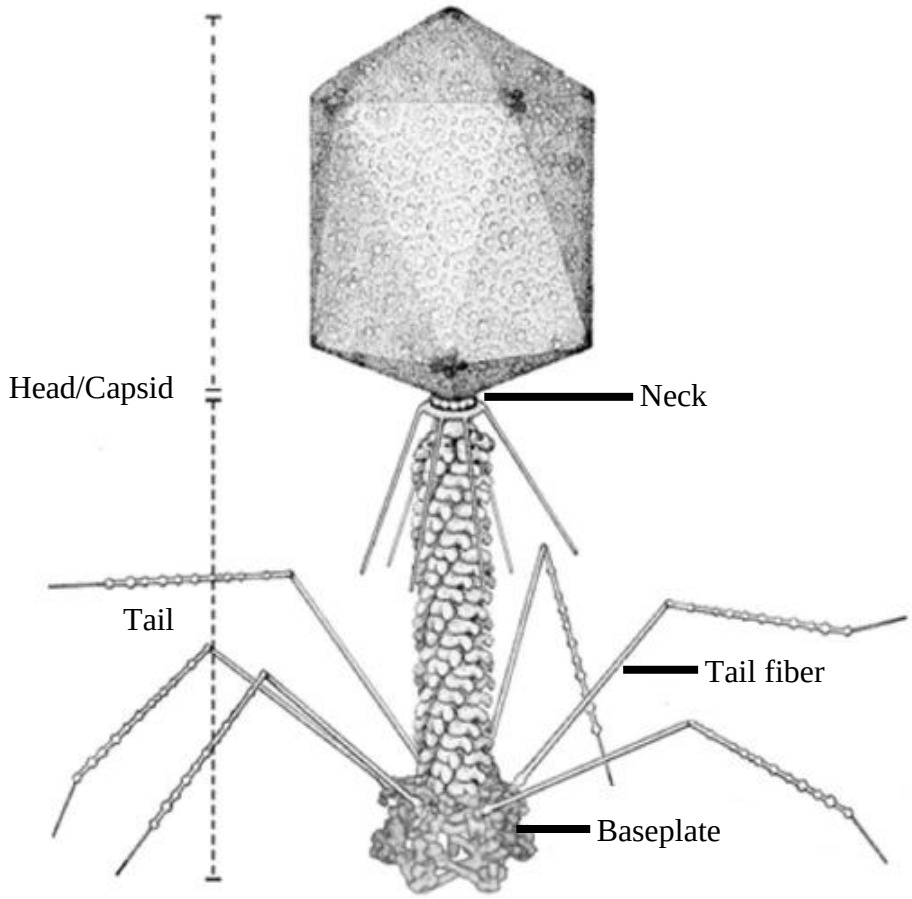

Figure 2.4: A typical phage structure (Miller et al., 2003).

2.2.5 Phage as biological control agent

Following the discovery of phages, the first known antibacterial potential of bacteriophage was recognized by Felix d'Herelle since 1919, against dysentery, cholera and bubonic plague (Clark & March, 2006; Kutateladze & Adamia, 2008; Nishikawa *et al*., 2008). Since then, the use of phages had generated a flurry of interest in modern medical industry in Europe (Clark & March, 2006; Dublanchet, 2007).

One primary application of phage is as bio-control agent. The biological control application of phage is generally referred to the process of applying lytic phages for the treatment of infectious diseases caused by pathogenic bacteria or also known as phage therapy (Clark & March, 2006; Dublanchet, 2007; Uchiyama *et al*., 2008). Phages are the natural enemies of bacteria which selectively attacks their specific hosts (Hagens & Loessner, 2007). This unique characteristic is essentially important as a bio-control of bacterial infections to target and kill diseases-causing bacteria without damaging the natural bacterial flora (Capparelli *et al*., 2005; Hagens & Loessner, 2007; Uchiyama *et al*., 2008).

However, since the implementation of antibiotics in the 1940s, the research and clinical application of phage therapy were largely abandoned by most western scientists after World War II (Tanji *et al*., 2005; Clark & March, 2006; Hanlon, 2007; Fortuna *et al*., 2008; Kutateladze & Adamia, 2008; Nishikawa *et al*., 2008; Vinodkumar *et al*., 2008).

Due to recent increases in antibiotic-resistant bacterial strains, the therapeutic exploitation of phages has once again received renewed interest as alternative treatment (Goodridge *et al*., 2003; Tanji *et al*., 2005; Clark & March, 2006; Kropinski, 2006; Dublanchet, 2007; Hanlon, 2007; Nishikawa *et al*., 2008; Vinodkumar *et al*., 2008; Courchesne *et al*., 2009) and/or synergistic approach to battle against bacterial infections (Ryan *et al*., 2012). Thus, many pharmaceutical companies are putting a lot of efforts into phage technology through investment, rigorous research and development activities in favor of therapeutic phage preparations (Clark & March, 2006; Hanlon, 2007).

In addition, with the recent advances in molecular biology and gradually improved knowledge of phage biology have created more opportunities for second-time success in phage therapy (Kudva *et al*., 1999; Courchesne *et al*., 2009). Furthermore, it has become apparent that phages offer numerous unique advantages over the use of conventional antibiotic therapy (Hanlon, 2007), such as, phage specificity in destroying drug-resistant bacteria that minimally cause disturbance to normal beneficial flora, quickly producing new phages in response to the appearance of phage-resistant bacteria compared to inability of antibiotics to respond to bacteria resistant and lower production cost since phages are easily discovered from various environments (Courchesne *et al*., 2009).

Eliava Institute of Bacteriophage, Microbiology and Virology, located in Tbilisi, the former Soviet Union has been and still the primary manufacturer of phage products in the world. Besides, the main focus area of Eliava Institute appears to be

the world authority in research and development of phages for pathogenic bacteria control (Hanlon, 2007).

2.3 *Escherichia coli* bacteria host

Escherichia coli is a Gram-negative, robust and rod-shaped bacterium from the family *Enterobacteriaceae* (O'Flynn *et al*., 2004; Naylor *et al*., 2005; Vogt & Dippold, 2005). This bacterium was previously discovered in 1885 by a German paediatrician, Theodor Escherich (Goodridge *et al*., 2003; Naylor *et al*., 2005). This species is the most abundant facultative anaerobe that is usually found in the lower intestinal tracts of human, free-living animals and warm-blooded organisms (Schroeder *et al*., 2002; Goodridge *et al*., 2003; Naylor *et al*., 2005; Vogt & Dippold, 2005). The bacterium is also found in water, foods and soil due to contamination by fecal or during animal slaughter (Schroeder *et al*., 2002).

2.3.1 Significance of *E. coli* O157:H7 infections

Studies have shown that food borne diseases in humans are caused by certain serotypes of *E. coli* strains producing Shiga toxin, for examples *E. coli* O157:H7 and *E. coli* O104:H4. Serotypes are the group of cells distinguished by their shared cell surface antigens. The "O" in the name refers the cell wall (somatic) antigen number, while the "H" refers the flagella antigen (Baron, 1996). These antigens are essential for phage infection as phage recognizes them prior to attachment (Kropinski, 2006). These *E. coli* strains are also described as 'Shiga toxin-producing' *E. coli* (STEC) by producing Shiga-like toxins (Stx) I and II (Tanji *et al*., 2005; Liao *et al*., 2011). Shiga

toxin is the most important *E. coli* pathogenic factor that is responsible for the bacterial infection and pathogenicity. Moreover, these harmful strains are also known as the primary etiologic agent of urinary tract infections (UTIs) in humans and animals. These infections are one of the most common bacterial diseases in humans (Nishikawa *et al*., 2008).

The spread of infectious diseases caused by food borne bacterium such as *Campylobacter, Salmonella*, *E. coli* and *Listeria* remains as problems to public health (Hagens & Loessner, 2007). In fact, the numbers of cases of food borne diseases have been increasing dramatically including diseases caused by *E. coli* O157:H7 (Currie *et al*., 2007). This notorious *E. coli* O157:H7 is also referred as an enterohaemorrhagic strain of *Escherichia coli* (EHEC).

E. coli O157:H7 has been a main food safety concern due to its low infective dose in humans with only one hundred cells (Tanji *et al*., 2004; Raya *et al*., 2006; Liao *et al*., 2011). This low infectious dose of high virulence of *E. coli* O157:H7 could cause severity of infections that may seriously result in death due to hemorrhagic colitis with highest incidence of reported cases occurring mostly in children aged less than 15 years and elderly (Galland *et al*., 2001; Nishikawa *et al*., 2008). Meanwhile, the World Health Organization (WHO) estimates that five millions children die each year due to acute diarrhea. Indeed, *E. coli* O157:H7 has been claimed as one of major cause of childhood diarrhea in developing and threshold countries (Hanlon, 2007).

The Centers for Disease Control and Prevention (CDC) estimated that there were approximately 265,000 STEC infections occur each year in the U.S.A and out of this estimation, 36% were caused by *E. coli* O157:H7 with 73500 illnesses, 2100 hospitalizations and 60 deaths (Schroeder *et al*., 2002; National Institute of Allergy and Infectious Diseases, 2011). CDC has claimed that multiple food borne diseases outbreaks of *E. coli* O157:H7 have been primarily associated with consumption of undercooked ground beef and contaminated bovine products such as unpasteurised milk (Goodridge *et al*., 1999; Kudva *et al*., 1999; Schroeder *et al*., 2002; O'Flynn *et al*., 2004; Capparelli *et al*., 2005; Naylor *et al*., 2005; Abuladze *et al*., 2008; Viazis *et al*., 2011). Other food products that have epidemiologically implicated in the outbreaks include fruits, fresh vegetables, salads, and salami contained with preserved ready-to-eat beef (Goodridge *et al*., 1999; Capparelli *et al*., 2005; Abuladze *et al*., 2008). For examples, the unintentional outbreaks in the U.S between 1992 and 1993 were linked to the undercooked ground beef consumption at fast food outlets (Goodridge *et al*., 1999). Apart from that, several outbreaks have associated with lettuce which was one of the sources of contamination (Kudva *et al*., 1999). In addition, according to Abuladze *et al*. (2008), the outbreak of 2006 in the U.S. has been linked to contaminated spinach whereas in Japan; radish sprouts was the main source of contamination in the massive 1996 outbreak (Kudva *et al*., 1999). Abuladze *et al*. (2008) has also revealed that the contaminated radish sprouts were in fact served in school lunches and thus largely affected 8,000 children.

E. coli O157:H7 infections of have serious complications in humans such as thrombotic thrombocytopenic purpura (TTP), acute renal diseases and fatal bloody diarrhea which develops to a range of potentially life-threatening conditions from

hemorrhagic colitis (HC) occasionally to a type of kidney failure known as hemolytic-uremic syndrome (HUS) (Tanji *et al*., 2004; Capparelli *et al*., 2005; Hagens & Loessner, 2007). Besides, current treatment of *E. coli* O157:H7 human infections showed high prevalence of resistance towards standard antibiotics example, ampicilin, tetracycline, cephalothin and sulfamethoxazole (Schroeder *et al*., 2002). In fact, the use of some antibiotics such as fluoroquinolones for this infection is not recommended in the U.S. as it may potentially induce Shiga-toxin encoding bacteriophages *in vivo* and release Shiga toxin in the intestinal tract (Galland *et al*., 2001; Schroeder *et al*., 2002). Due to the emergence and raising cases of antibiotic resistance of *E. coli* O157:H7, natural control strategies have received growing demand and attention including the application of phage (Coffey *et al*., 2011; Park *et al*., 2012). Hence, *E. coli* O157:H7-specific phages could be used in phage therapy to deal with this resistance by infecting and lysis the pathogen.

The transmission of *E. coli* O157:H7 may occur from bovine feces onto meat during slaughter or milking as direct fecal contact may contaminate food, water and person-to-person (Kudva *et al*., 1999; O'Flynn *et al*., 2004; Naylor *et al*., 2005). Tracing the principal source of food borne outbreaks, reveals that the gastrointestinal tracts of ruminants particularly cattle and sheep have been discovered as major asymptomatic reservoirs of this pathogen (Kudva *et al*., 1999; O'Flynn *et al*., 2004; Tanji *et al*., 2004; Capparelli *et al*., 2005; Naylor *et al*., 2005; Raya *et al*., 2006).

2.3.2 *E. coli* O157:H7-specific virulent phages

Previous studies have discovered over fifty *E. coli* O157-specific phages that efficiently infect and cause lysis to *E. coli* O157 cells (Table 2.2) (Kudva *et al*., 1999; Raya *et al*., 2006; Villegas *et al*., 2009; Liao *et al*., 2011; Kim *et al*., 2013; Kropinski *et al*., 2013; Shahrbabak *et al*., 2013). Among these *E. coli* O157-specific phages, only twenty four of them (Kropinski *et al*., 2013) were found to be highly effective against *E. coli* O157:H7 cells (studied from previous literatures). However, the available information related to the biology, molecular biology and other characteristics of most of these phages are still lacking (Kropinski *et al*., 2013).

Table 2.2 List of *E. coli* O157 and *E. coli* O157:H7 -specific phages.

Bacteria	Phage	References
E. coli O157	38, 39, 41, 42, AR1, Bo-21, Av-05, SP21, Av-06, Av-08, CA933P, CA911, MFA933P, CA9311 MFA45D, wV8, CBA65, CEV1, CEV2, CSLO157, DC22, e4/1c, e11/2, ECA1, ECB7, ECML-4, ECML-117, ECML-134, JK06, KH1,KH4, KH5, LG1, φV10, ϕD, PBECO 4, PhaXI, PP01, PP17, Rv5, SFP10, SH1, SP15, SP21, SP22, vB_EcoM_CBA120(CBA120), bV_EcoS_AKFV33(AKFV33), and vB_EcoS_Rogue1 (Rogue1)	Kudva *et al*., 1999; Raya *et al*., 2006; Villegas *et al*., 2009; Liao *et al*., 2011; Kim *et al*., 2013; Kropinski *et al*., 2013; Shahrbabak *et al*., 2013.
E. coli O157:H7	AKFV33, AR1, CBA120, CEV1, ECML-4, ECML-117, ECML-134, e4/1c, e11/2, KH1, KH4, KH5, LG1, ϕD, PBECO 4, PhaXI, PP01, PP17, Rogue1, Rv5, SH1, SFP10, SP15 and wV8	Shahrbabak *et al*., 2013.

2.4 Comparison of phages infecting *E. coli* O157:H7

Among the listed *E. coli* O157:H7-specific phages (Table 2.2), only a few of them were well-studied (Kropinski *et al*., 2013) previously and the information on their sources of isolation, morphological and genome characteristics, and their lytic activities were obtained from previous literatures. Thus, this information was described and compared in the following subsections.

2.4.1 Sources and regions of isolation

Phages are remarkably abundant in our environment, circulating among human population. They are ubiquitous and reside in all reservoirs occupied by bacteria including intestines, food or soil. Examples of their natural sources are sewage, water, and feces from animals or humans. Therefore, these sources are principally used for phage isolation (Morita *et al*., 2002).

E. coli O157:H7-specific phages were isolated from different types of samples collected at various locations. Table 2.3 shows the collected samples and their original locations for each phage. From Table 2.3, most of phages infecting *E. coli* O157:H7 had been isolated from fecal and sewage samples. However, the pattern of prevalence showed the abundance of phages was highest in feces compared to sewage. This is due to the fact that feces of ruminants are considered as a rich source of phage infecting *E. coli* O157:H7 because ruminants are the natural niche for EHEC (Viazis *et al*., 2011).