博 士 論 文
(Doctoral Thesis)

Screening, isolation and characterization of large bacteriophages for use in biocontrol of a wide-range of pathogenic bacteria

広範囲の病原菌を対象とした生物防除に有用な大型バクテリオファージのスクリーニング、単離および特徴付け

Alaa Eldin Mohamed Saad Meslhi

広島大学大学院先端物質科学研究科
分子生命機能科学専攻

Department of Molecular Biotechnology
Graduate School of Advanced Sciences of Matter
Hiroshima University

March 2020

目 次
(Table of Contents)

1. 主論文 （**Main Thesis**）

Screening, isolation and characterization of large bacteriophages for use in biocontrol of a wide-range of pathogenic bacteria

（広範囲の病原菌を対象とした生物防除に有用な大型バクテリオファージのスクリーニング、単離および特徴付け）

Alaa Eldin Mohamed Saad Meslhi

2. 公表論文（Articles）

(1) Systemic method to isolate large bacteriophages for use in biocontrol of a wide-range of pathogenic bacteria.
Alaaeldin Mohamed Saad, Ahmed Mahrous Soliman, Takeru Kawasaki, Makoto Fujie, Hirofumi Nariya, Tadashi Shimamoto and Takashi Yamada.
Journal of Bioscience and Bioengineering, **127** (1), 73-78 (2019).

(2) Full genome sequence of a polyvalent bacteriophage infecting strains of *Shigella*, *Salmonella*, and *Escherichia*
Alaaeldin Mohamed Saad, Ahmed Askora, Ahmed Mahrous Soliman, Hirofumi Nariya, Takeru Kawasaki, Makoto fujie, Tadashi Shimamoto and Takashi Yamada.
Archives of Virology, **163** (11), 3207-3210 (2018).

主論文
(Main thesis)

CONTENTS

CHAPTER I

GENERAL INTRODUCTION

The continuous decline in the effectiveness of antibiotics and increased rates of bacterial resistance represents a public health challenge universally (Lin *et al*., 2017). It has been suggested that by 2050, around 10 million people could die per year due to antimicrobial resistance (O'neill, 2014). Focus must be turned towards alternative approaches for treating infections. One of the most promising strategy is the use of bacterial viruses, the natural predators of bacteria known as bacteriophages or phages (Ghosn *et al*., 2019). Bacteriophages represent the major biological organisms on the earth with an estimated 10^{31} bacteriophages (Hendrix *et al*., 1999; Keen, 2015). They are capable of killing bacteria as effectively as antibiotics and are far more flexible in their capabilities (Drulis-Kawa *et al*., 2012), as they can overcome almost all disadvantages of antibiotics.

Bacteriophages can be classified according to their genome sizes into small and large phages. To date reported sizes of all isolated and characterized phages were varied between 3,300 base of ssRNA viruses of *Escherichia coli* (small phages) to almost 500 kbp of dsDNA genome phage G of *Bacillus megaterium* (jumbo phages) (Hatfull and Hendrix, 2011; Yuan and Gao, 2017). Interestingly, jumbo phages with genome sizes more than 200 kbp are characterized by their wider host range (Bhunchoth *et al*., 2015), sustainability and long-lasting lytic infection (Yamada *et al*., 2010).

Unfortunately, there was no specific method to isolate large phages. As almost all of large phages have been isolated by chance (Yuan and Gao, 2017) by using classical methods for screening of bacteriophages. My first goal in this study is to develop a specific method for selective screening of large phages. To investigate the host range of isolated large phages, we prepared several hosts including Gram positive and Gram negative bacteria with focus on members of the *Enterobacteriaceae*. As gastrointestinal infections caused by pathogenic strains of *Enterobacteriaceae* are considered as serious economic and public health problem worldwide (Logan and Weinstein, 2017). They are leading causes of death among human beings especially young children because of severe diarrhea (Humphrey *et al*., 2015). Moreover, they were reported in the list of global priority pathogens (Tacconelli *et al*., 2018).

The second purpose of my research is to genetically characterize the isolated large phages completely to confirm their ideality in biocontrol applications (Shende *et al*., 2017) and to understand evolutionary changes in large phage genome.

This thesis specifically focuses on screening, isolation and characterization of large bacteriophages for use in biocontrol of pathogenic bacteria especially the most common pathogenic strains of *Enterobacteriaceae.*

1.1. General characteristics of pathogenic *Enterobacteriaceae*

Enterobacteriaceae is a family of Gram-negative, non-spore-forming bacteria and is one of the most popular groups of bacteria known to man. This family includes a number of important foodborne pathogens such as *Salmonella*, pathogenic *Escherichia coli*, *Shigella* spp., *Yersinia enterocolitica* and *Cronobacter* spp. Other members of the family are regarded as opportunistic pathogens (e.g., *Klebsiella* spp, *Serratia* spp. and *Citrobacter* spp.). Pathogenic strains of *Enterobacteriaceae* use the intestinal tract as main reservoir for causing foodborne intestinal diseases and are known as the most common cause of intestinal upset. They are responsible for a variety of human illnesses, including urinary tract infections, gastroenteritis, septicemia, and pneumonia. Members of the *Enterobacteriaceae* are ubiquitous in nature. Although strains of some species are commensals, others are critical human and animal pathogens, and some are pathogenic to other creatures such as plants and insects. Their wide distribution indicates the higher possibility of food chain contamination (Bridier, 2019).

In addition to their etiology as foodborne diseases, some of the family members are associated with food spoilage and contribute to significant economic losses for agriculture and food industries. Bacterial food-borne diseases represent a growing public health concern for the entire world, including both developed and developing countries (Fung *et al*., 2018). According to the estimates of World Health Organization (WHO), about 2.2 million deaths around the world each year are caused by food-borne or water-based diarrhea (Johnson *et al*., 2011). For example; nontyphoidal *Salmonella* serotypes are responsible for around 1.4 million cases and 270,000 cases are caused by pathogenic *Escherichia coli* (Mead *et al*., 1999). According to Global burden of disease (GBD) study of 2015; foodborne diarrheal illness was estimated as the sixth leading cause of global disability-adjusted life years (DALYs) (Vos *et al*., 2016).

Antimicrobials have saved many of human lives and also have been used routinely to maintain animal health and productivity (Silbergeld *et al*., 2008). However, the majority (73%) of

these antimicrobials are used in animals raised for food (Van Boeckel *et al*., 2017) which give growing evidence that the majority of emerging infectious diseases have been associated with drug-resistant pathogens of zoonotic origins (Jones *et al*., 2008). Infections with antibiotic-resistant bacteria are more serious with highly public health concern due to increased morbidity and mortality with more social and economic costs (Cosgrove, 2006). The continuous increase of multi drug resistant strains (Ansari *et al*., 2015) and the restrictions of antibiotics usage in many countries (Barton, 2000) entail the development of effective and safe alternatives to these synthetic antimicrobials. Bacteriophages are a promising alternative to antibiotics (Drulis-Kawa, 2012).

1.2. Bacteriophages characteristics:

Bacteriophages are viruses those can infect and replicate within bacteria. Phages are widely distributed in nature and they are harmless to humans, animals and plants. Each bacteriophage is between 24-200 nm in size. Bacteriophages are completely parasitic and depend on the presence of bacterial host to reproduce. They infect specific groups of bacteria. Bacteriophages were officially discovered 100 years ago by d'Hérelle (1917). This discovery initiated a new branch of science in which, phages were utilized as biocontrol agents for the most of pathogenic bacteria (Summers, 1999). The research in phage therapy have been enthusiastic within just eastern Europe and the Soviet Union (Abedon *et al*., 2011), while this interest decreased for the rest of the world due to the discovery of Penicillin by Alexander Fleming in 1929. From this time, almost all the world started imprudent usage of antibiotics which led to the development of antimicrobial resistance and its serious consequences (Landers *et al*., 2012). Bacteriophages are promising alternative to antibiotics (Drulis-Kawa, 2012) especially in our recent era (Wittebole *et. al*., 2014; Bao *et. al*., 2015), for several reasons: (i) high specificity to target their host only leaving the remaining microbiota untouched unlike other antimicrobials that can cause damage to beneficial microbiota (Campbell, 2003); (ii) self-replicating, meaning that low dosages will multiply as long as there is still a host present (Johnson *et al*., 2008); (iii) low inherent toxicity, as they consist mostly of nucleic acids and proteins; (iv) phage isolation is relatively simple, fast and inexpensive; (v) they have proved to have prolonged shelf life. Most of these phage characteristics enable them to be used as efficient antimicrobial agents (Endersen *et al*., 2014).

1.2.1. Bacteriophages as biocontrol agents:

Recently, phage therapy is extensively used only on three countries, Georgia, Russia and Poland. For example, "Intestiphage" is phage cocktail targets around 20 different pathogenic

enteric bacteria that is available to the public in Georgia and Russia. Another cocktail is "Pyophage" that is employed routinely in the treatment of various wound infections and targeting *Streptococcus, Staphylococcus, Proteus, Pseudomonas* and *E. coli*. Unfortunately, there has been little transfer of phage therapy expertise from Georgia and Russia to other world countries, due to very rare primary publication in journals with English language.

Bacteriophages can also be added to variable types of food products. The most advantage of bacteriophages application on foods is the controlled selectivity and absence of impact on the food organoleptic and physiochemical properties (Pérez Pulido *et al*., 2015). Nowadays, there are several phage products already approved to be used in foods as food preservatives to control specific food-borne pathogens such as; ''ListShield™'' (Intralytix, Inc., Baltimore, MD, USA), a cocktail of six phages targeting *L. monocytogenes*, that was first approved by the United States Food and Drug Administration (FDA) and Food Safety and Inspection Service (FSIS) for applications in foods in 2006. Also ''Listex™ P100'' a single phage targeting *L. monocytogenes*, was approved as GRAS status by FDA in 2006. The Intralytix company has also commercialized other phage preparations (SalmoFresh™ and SalmoLyse™) for controlling *S. enterica* (Woolston *et al*., 2013). In addition, a cocktail of three phages (ECP-100) targeting *E. coli* O157:H7 "EcoShield™", was also approved by FDA and FSIS for food application in 2011. These phage products are allowed to use in foods as preservatives to control specific food-borne pathogens. In addition to direct food applications of phages, they can be used to prevent cross contamination of pathogens in food-contact materials as well as food processing facilities (Sulakvelidze, 2013). Furthermore, phages can be used to sanitize human hands and utensils. Therefore, phage applications would also be useful for extension of food preservation periods and food safety.

1.2.2. Bacteriophage obstacles:

The most important factors of phages usage as effective biocontrol agents are their host range and lasting infection (Ross *et al*., 2016). As the most of bacteriophages are highly specific for their hosts with little or no interaction other than their host strain even within the same species (Welkos *et al*., 1974). This characteristic may be limiting the ability of phage products to control bacterial infections. As for phage therapy, good candidate is a phage that can infect wide range of pathogenic bacteria. To assure the broad host range of phage formulations against target bacteria, numerous types of phages are represented into mixtures called "phage cocktails". Recently number of phage types increasing in the phage cocktails because insufficient outcomes have been obtained

from evaluating single phage preparations (Skurnik and Strauch, 2006). Unfortunately, there are many limitations of phage cocktail application in spite of sustainable demand of continuous levels of medical and commercial formulations. A cocktail containing many phages might have severe effects on beneficial bacteria. Moreover, formulations containing too many phages will result in higher developing and manufacturing costs (Chan *et al*., 2013). So, the ideal solution for this drawback is using jumbo phages (Hendrix, 2009).

1.2.3. Jumbo phages:

Jumbo phages are tailed phages characterized by large genome size over 200 kbp and most of them are related to family *Myoviridae* (Hendrix, 2009). The most characteristic features of jumbo phages are their large phage particles mainly big capsids (Fig. 1.1) that can package larger genomes in comparison with smaller phages. The advantage of large genome size is to help jumbo phages to acquire many extra genes. As, the most of jumbo phages have multiple genes for expressing DNA polymerase and RNA polymerase (RNAP) (Hertveldt *et al*., 2005; Thomas *et al*., 2007) and it is reported that expression of jumbo phage genes independent from their host RNAPs and may be only dependent on their own RNAPs (Ceyssens *et al*., 2014; Leskinen *et al*., 2016). Furthermore, jumbo phages have extra proteins , like endolysin, chitinase, glycoside hydrolase and many other similar proteins for the lysis of the host cell wall to facilitate the ability of phage infection (Yuan and Gao, 2016). Moreover, almost all jumbo phages have many tRNA genes thought to correspond to abundant codons in phage genes mainly for structural proteins and to enhance the efficiency of phage specific genes translation (Kiljunen *et al*., 2005). It is thought that jumbo phages that own additional proteins may reduce their dependence on specific hosts (O'Donnell *et al*., 2013). As a result of all of these jumbo phages' genome characteristics, it is thought that jumbo phages would have broad host ranges due to reduced dependence on their host bacterium (Yuan and Gao, 2017). Moreover, the previously mentioned characteristics of jumbo phage genomes attracted scientists for more genomic analysis to provide deep investigation of the phage-host interaction and greater understanding of the origin and the evolution of the jumbo phages (Yuan and Gao, 2017). Although, some research groups discovered jumbo phages against highly pathogenic bacteria, they didn't make more exploration for this advantage and directed their inclinations for more genomic analysis (Abbasifar *et al*., 2014; Burkal'tseva *et al*., 2002; Kim *et al*., 2013 and Simoliunas *et al*., 2012). While other groups took more advanced step in the characterization from the point of biocontrol through discovering the wide host range of their

jumbo phages (Bhunchoth *et al*., 2015; Kim *et al*., 2012 and Miller *et al*., 2003a). Also, the genomic information of some of the jumbo phages revealed the reason for their stability and ability for long-lasting lytic infection (Yamada *et al*., 2010).

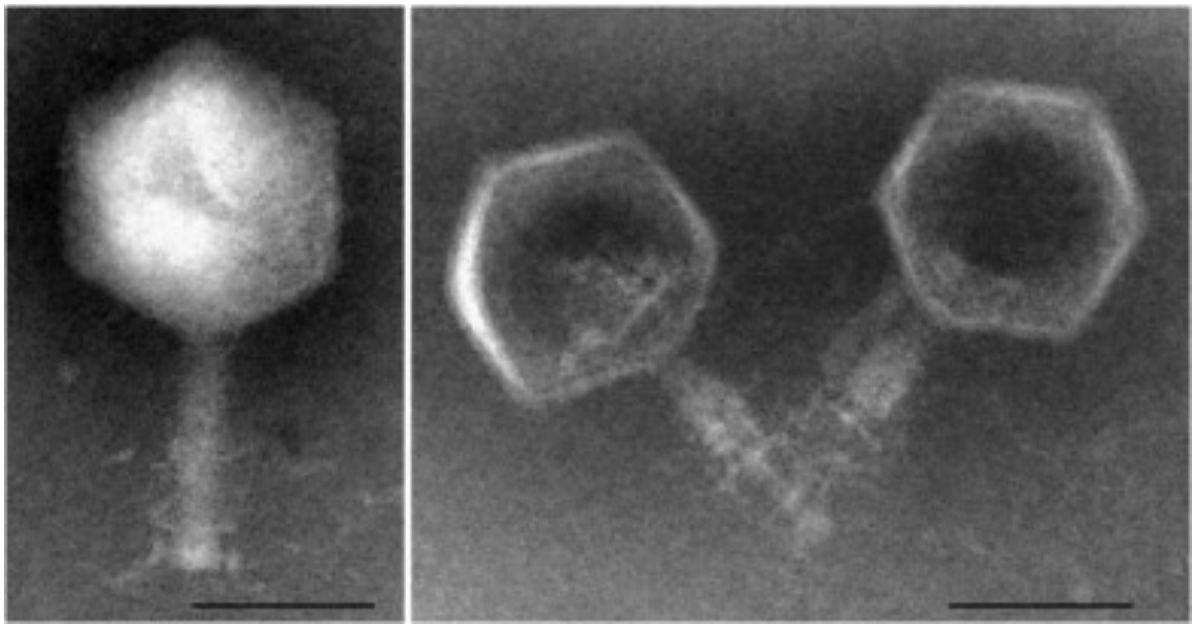


Fig. 1.1. Electron micrographs of jumbo phage ϕRSL1 that infect to the phytopathogen *Ralstonia solanacerum* (Yamada *et al*., 2010). The ϕRSL1 particle has an icosahedral head of 150 nm in diameter and a contractile tail of 138 nm in length and 22.5 nm in diameter.

1.3. Method for jumbo phages isolation:

In spite of the amazing characteristics of jumbo phages, they are rare in isolation (very few systemic reviews currently available on jumbo phages) (Hendrix, 2009; Van Etten *et al*., 2010; Yuan and Gao 2017), only around 100 ones have been isolated since the first discovery of phages at the beginning of last century (Yuan and Gao 2017). Although jumbo phages have high frequency and abundance in our environment, there is no approved method for isolation of jumbo phages until now. Most of the jumbo phages were isolated by chance using the standard method for bacteriophage propagation. For example, *Bacillus megaterium* phage G was accidentally found during preparation of another bacteriophage (Donelli, 1968). Recently, it is obvious that the inaccessibility to jumbo phages detection is attributed to the highly dependence on the classical

procedures of phage isolation (Serwer *et al*., 2007; Yuan and Gao 2017). Jumbo phages are hardly making plaques in the 0.4-0.7% agarose gel due to their large virions which hinder their diffusion in top agar medium (Carlson, 2005). Also, liquid enrichment cultures are usually ineffective for detection of oversized phages due to absence of visible lysis in liquid culture (Carlson, 2005). Moreover, most of the jumbo phages are lost during filtration for removing bacteria due to their large size, i. e. jumbo phages might hardly pass through the filter pores (Williamson *et al*., 2005; Yuan and Gao 2017). Proper modifications of classical methods and developing specific methods for large phage isolation will pave the way for interesting discoveries in the field of molecular biotechnology and biocontrol applications.

1.4. Genomic characterization of phages:

Bacteriophages are highly diverse and infect essentially all bacteria on earth (Catalao *et al*., 2013**)**. Even if they are isolated from a single habitat, their diversity is huge (Jurczac-Kurek *et al*., 2016). Phages affect bacteria into two ways, as they can control bacterial population numbers, and on the other hand contribute to move genes from one bacterium to another. Complete genomic characterization of phages is mightily essential to understand phages' evolutionary history including relationships, biodiversity and biogeography of each phage to its close members (Hambly *et al*., 2001). Moreover, phage genomes encode not only proteins that have been useful for biotechnology but also for biocontrol applications including food additives, biotherapy of infections caused by antibiotic-resistant bacterial strains, DNA delivery vehicles and many more relevant technologies (Harada *et al*., 2018). That's why deep understanding of phage genetics must be exploited to generate a broad application spectrum like novel nanotechnologies, bacterial detection strategies and biological control of pathogenic bacteria on an industrial scale (Petty *et al*., 2007).

There is no evidence to date that bacteriophages exhibit any harmful effects on humans or even on animals (Abedon *et al*., 2011). However, guaranty measures in the use of bacteriophages for biocontrol applications must be taken through whole-genome sequencing to ensure the safety of that genome and it is free from genes encoding bacterial virulence factors and or resistance genes of bacteria. In my work, full genome sequencing and characterization was carried out for a novel lytic phage (EcS1) that can infect to broad range of pathogenic bacterial strains.

1.5. Tail fibers and their roles in host range difference:

The first step in phage infection is adsorption of the phage to the host cell (Chatterjee and Rothenberg 2012). It was revealed that the tail structures are the key determinants for phage specificity to certain host and infection process (Hu *et al*., 2015 and González-García *et al*., 2015). Phages diffuse randomly until they encounter a bacterial cell. Then, they will continue to diffuse on the cell surface until they bind to a receptor that initiates the infection process. Nonspecific phages fall off from the cell surface (Fig 1. 2**)** and continue their free motion (Le *et al*., 2013).

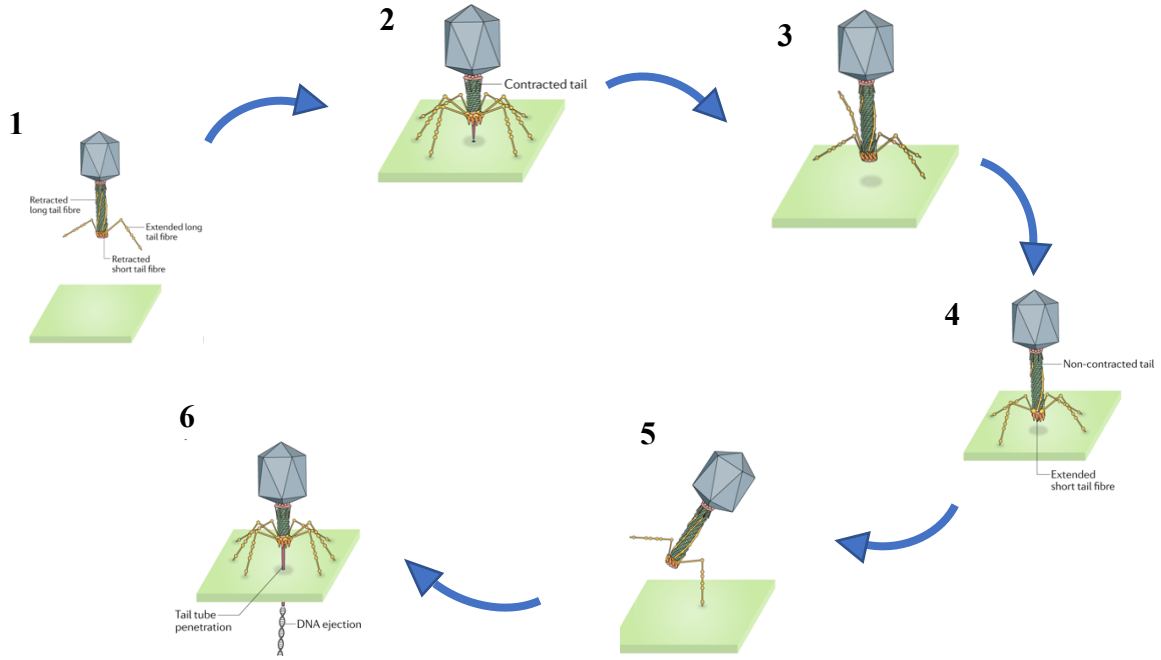
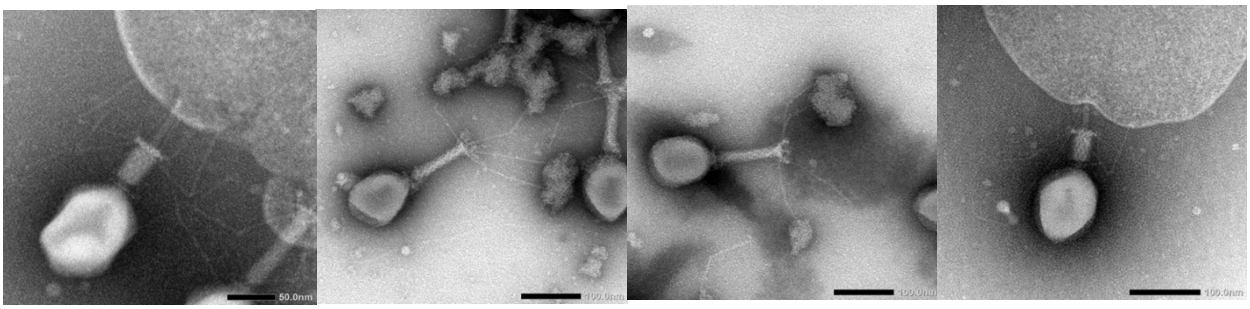


Fig 1. 2. A model of adsorption of phage T4 to its host surface (Nobrega *et al*., 2018).

1) In unattached state, the six long tail fibers of phage T4 extend and retract. 2) An extended long tail fiber initiates a trial to contact with the bacterial cell surface. 3) Then the transient binding of extended long tail fibers allows phage T4 to move on the cell surface to find an optimal site for irreversible adsorption. 4) Binding of any tail fibers (two or three of the long tail fibers) to their cellular receptor changes the baseplate conformation and consequently releases the short tail fibers, which irreversibly bind to the host receptor. 5) Irreversible binding causes the baseplate for triggering contraction of the tail 6) Ejection of the phage DNA into the cell.

Almost all tailed phages use tail fibers/tail spikes with receptor-binding proteins (RBPs) emerging from the base plate to interact with receptors on the bacterial cell surface (such as lipopolysaccharide, teichoic acids and porins). The initial reversible attachment is followed by irreversible adsorption and ejection of the phage genome into the specific host cytoplasm. On the other hand, only reversible attachment occurs in case of non-specific host (Fig 1. 3).

A)



B)

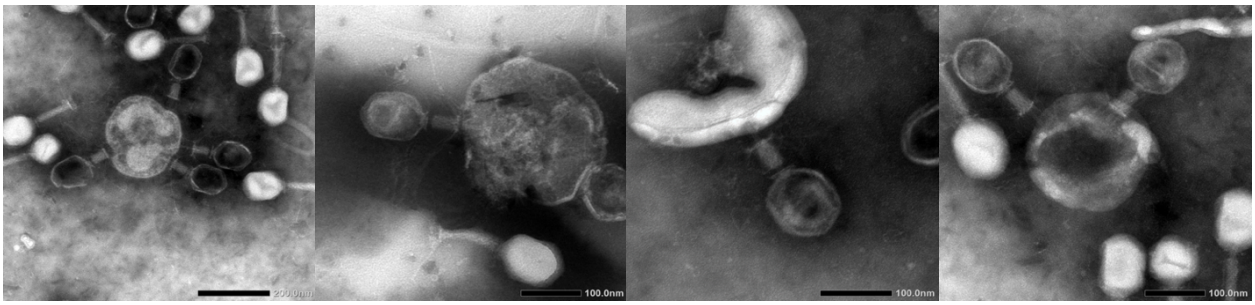


Fig 1. 3. (A) Reversible binding of EcS1 phage long tail fibers to *Serratia marcescens* outer membrane vesicles OMVs (B) Irreversible binding of EcS1 to *E. coli* OMVs (Current study).

Interestingly, under selective pressure, tail fiber genes seem to evolve faster than other phage genes, giving it the advantage to gain new specificities to infect different hosts as well as to enter other ecological niches (Casjens, 2005). Morphologies of tailed phages are unique and different from the morphogenesis of other viruses. However, many tailed phages show high similarity to one another when observed with an electron microscope, in fact they have very different genomes (Leiman *et al*., 2010). Thus, nucleotide sequence information (preferably whole genome sequence) is required to understand the relationships among the members of any set of phages being compared. Annotation and BLAST analyses were performed for whole genome of the newly

isolated phage EcS1. Further homology analyses were done for two tail fiber protein genes to understand the relationship between genetic evolution in tail fibers gene products of the isolated phage and its broad host range.

1.6.The scope of this study:

The outbreak of multidrug-resistant pathogenic bacteria (mainly *Enterobacteriaceae*) made the novel control strategies necessary. Phages have regained attention for their specific lytic activity against pathogenic bacterium, especially jumbo phages those are characterized by sustainability and long-lasting infection. However, almost all methods for screening phages were traditional and methods for screening and isolating large phages have not been reported.

The main purpose of this study is to develop a new systemic method to isolate large phages and to evaluate their host ranges followed by complete genetic characterization to confirm their novelty and evolutionary relationship to other bacteriophages. Furthermore, full genome sequencing ensures the absence of any harmful genes such as bacterial virulence genes and/or antibiotic resistance genes. Moreover, it is important to study tail fibers proteins of isolated phages because they are responsible for the first step in host-recognition and binding. Characterization of receptor binding proteins (RBPs) is essential to investigate how polyvalent phages can recognize and bind wide range of hosts utilizing its specific tail fibers proteins.

Chapter II

Systemic method to isolate large bacteriophages for use in biocontrol of a wide-range of pathogenic bacteria

2.1. SUMMARY

Large phages with a genome around 200 kbp or more have attracted great interests for biocontrol because they can infect wide host ranges of bacteria and maintain long-lasting infection. However, there is no standard method for selective isolation of large phages. I developed in this study a systemic method to isolate large phages and succeeded in isolating 11 large phages, named *Escherichia* phage E1∼E11. By electron microscopic observations, they were characterized as typical *Myoviridae* phages with big capsids and long contractile tails. Genome sizes of the isolated phages were determined by Pulsed-field gel electrophoresis and found to be in two groups, those around 200 kbp for E1, E2, E5, E6, E7, E9 and E10 phages, and others of approximately 450 kbp for E3, E4, E8 and E11 phages. The isolated large phages showed wide host ranges: for example, E9 was effective against *Shigella sonnei* SH05001, *Shigella bydii* SH00007, *Shigella flexneri* SH00006, *Salmonella enterica* serovar Enteritidis SAL01078 and *Escherichia coli* C3000 (K-12 derivative), as well as its original host *E. coli* BL-21. For initial screening of these jumbo phages non-pathogenic *E. coli* strains were used as hosts. Therefore, this method opens a way to isolate large phages infecting wide ranges of pathogenic bacteria in a typical laboratory with standard laboratory strains as the hosts. The isolated large phages will be good candidates for biocontrol of various pathogens.

2.2. Introduction:

Bacteriophages (abbreviated as phages) have been used for clinical applications to kill pathogenic bacteria since their first discovery at the beginning of the 20th century (Twort, 1915 and d'Herelle, 1917). The most important factors of phages as effective biocontrol agents for practical use are their host ranges and ability to execute lasting infection (Ross *et al*., 2016). Most bacteriophages have specific host ranges, usually limited to strains of the same bacterial species (Ross *et al*., 2016 and Welkos *et al*., 1974). This is a big disadvantage of phages compared with antibiotics. Therefore, good candidates for phage therapy can infect a broad range of pathogenic bacteria. To assure a broad host range of phages formulations against target bacteria, different types of phages are mixed as so called 'phage cocktails'. Unfortunately, there are many limitations of phage cocktail application associated with sustainable and continuous supply of medical and commercial formulations. Also, formulations containing too many phages can result in high manufacturing and development costs (Chan *et al*., 2013).

"Jumbo phages" are tailed phages characterized by a large genome (>200 kbp) and most of them are members of the family *Myoviridae* (Hendrix, 2009). Most jumbo phages have multiple genes for expressing functional DNA polymerases (DNAPs) and RNA polymerases (RNAPs) (Hertveldt *et al*., 2005; Thomas *et al*., 2007). It was reported that the expression of jumbo phage genes is independent of the host RNAP and may be driven solely by the phage own RNAPs (Ceyssens *et al*., 2014; Leskinen *et al.,* 2016). Moreover, jumbo phages have genes for multiple enzymes involved in host cell lysis, such as endolysin, chitinase, glycoside hydrolase and lyases, that facilitate expanded infection (Yuan and Gao, 2017). Many other genes encoded by jumbo phages may function to reduce their dependence on bacterial hosts and to activate self-establishment (O'Donnell *et al*., 2013). Because of these special characteristics of jumbo phages, it is thought that most have a broad host range because of their reduced dependence on their host bacterium (Yuan and Gao, 2017). In addition to this, some jumbo phages were reported to execute sustainable, long-lasting infection (Yamada *et al*., 2010; Fujiwara *et al*., 2011).

Despite the amazing characteristics of jumbo phages, only100 jumbo phages have been isolated and reported since the first discovery of phage at the beginning of the last century. However, they may be more frequent and abundant in our environment, with a high impact on microbial ecology. Most jumbo phages have been isolated by chance by normal methods for bacteriophage propagation. An example is bacteriophage G, which was accidentally found during

preparation of another bacteriophage (Donelli, 1968). The inaccessibility of jumbo phages is attributable to our dependence on the classical procedures for phage detection and isolation (Yuan and Gao, 2017; Serwer *et al*., 2007). For example, using high concentrations (0.6~0.8%) of top agar in plate assays hinders the diffusion of jumbo phages and hence limits detection of their plaques (Carlson, 2005). Removal of bacterial contamination by filtration through a membrane filter with fine pores (~0.2 μm) could eliminate most jumbo phages coexisting in the sample (Williamson *et al*., 2005). Unfortunately, there is still no standard method for isolation of jumbo phages. The aim of this study was to develop an effective method for selective isolation of large phages such as jumbo phages (with a genome larger than 200 kbp) and phages with a genome around 200 kbp including T4-like phages.

2.3. MATERIALS AND METHODS

2.3.1. Bacterial strains, bacteriophages and culture conditions

Bacterial strains used in this work are listed in Table 2.1. They included three strains of Gram-positive bacteria from two genera (*Bacillus* and *Staphylococcus*) and 16 strains of Gram-negative bacteria from nine genera (*Acinetobacter*, *Escherichia*, *Klebsiella*, *Proteus, Providencia*, *Pseudomonas*, *Salmonella, Shigella* and *Vibrio*). Bacteria were cultivated in LB medium containing 1% hipolypepton, 0.5% yeast extract and 1% sodium chloride (Sambrook and Russel, 2001) at appropriate temperatures for each strain (37^{o}C for overnight). For animal pathogenic bacterial strains, all experiments were performed in biosafety level 2 laboratories.

For phage propagation and to collect sufficient phage particles, for each phage, a total of 2L of bacterial culture was grown. When OD_{600} of the cultures reached 0.1 U, the phage was added at a multiplicity of infection (moi) of 0.01–0.05. After further incubation for 16–18 h, (culture lysis), the remnant cells and debris were removed by centrifugation with a R12A2 rotor in a Hitachi himac CR21E centrifuge (Hitachi Koki Co. Ltd., Tokyo, Japan), at 5000 ×g for 10 min at 4 °C. The supernatant was passed through a 0.45 μm membrane filter followed by precipitation of the phage particles. The pellet collected by centrifugation in a Hitachi himac CR21E centrifuge at 15,000 ×g for 1 hour at 4 °C was dissolved in SM buffer (50 mM Tris–HCl at pH 7.5 containing 100 mM NaCl, 10 mM MgSO4). Phage preparations were stored at 4 °C until use.

Table 2.1. Bacterial strains used to identify the host range of the isolated phages:

Host (Bacterial species)	
Acinetobacter baumannii 395	*Salmonella enterica* serovar Enteritidis SAL01078
Escherichia coli BL21	*Shigella sonnei* SH05001
E. coli C3000 (C strain)	*Shigella boydii* SH00007
E. coli K12	*Shigella flexneri* SH00006
E. coli O157:H7	*Serratia marcescens* D601
Enterohaemorrhagic *E. coli* EHEC 03064	*Serratia marcescens* (ATCC 13880)
E. coli 7	*Vibrio cholerae* O139 MDO6
Klebsiella pneumoniae 63	*Vibrio cholera* P1418O1 E1
Proteus mirabilis 59	*Vibrio fluvialis* AQ0005
Providencia stuartii 50	*Vibrio parahaemolyticus* ACA339
Pseudomonas aeruginosa 16	*Bacillus subtilis* 168
Pseudomonas aeruginosa PAO1	*Staphylococcus aureus* ATCC 6538
Salmonella enterica serovar Paratyphi B SAL 04100	*Staphylococcus epidermidis* ATCC 35984

2.3.2. Selective detection of large phages:

To facilitate detection of jumbo or large phages, standard plaque assays were modified as follows: Bacterial cultures (*E. coli* strains BL-21, XL-1-Blue, DH5α or Mach-1) were prepared in 4.5 ml LB broth adjusted to $OD_{600\ nm}$ = 0.25. A 100 µl phage sample was mixed with 250 µl bacterial culture and stood for 10 min at room temperature to allow adsorption. With 4.5 ml of molten LB top agar (0.35%), the mixture was poured onto an LB agar plate (1.5% agar). The plates were incubated for 24 h at 23°C (or 18°C if strain Mach-1 was the host). Small plaques that appeared on the plates were selectively picked for further purification and enrichment. Single plaque purification was repeated three times to confirm the plaque was derived from only one kind of phage. Enriched phage preparation was obtained from plate lysates. The scheme of this systemic detection and isolation of large phages is summarized in Fig. 2.1.

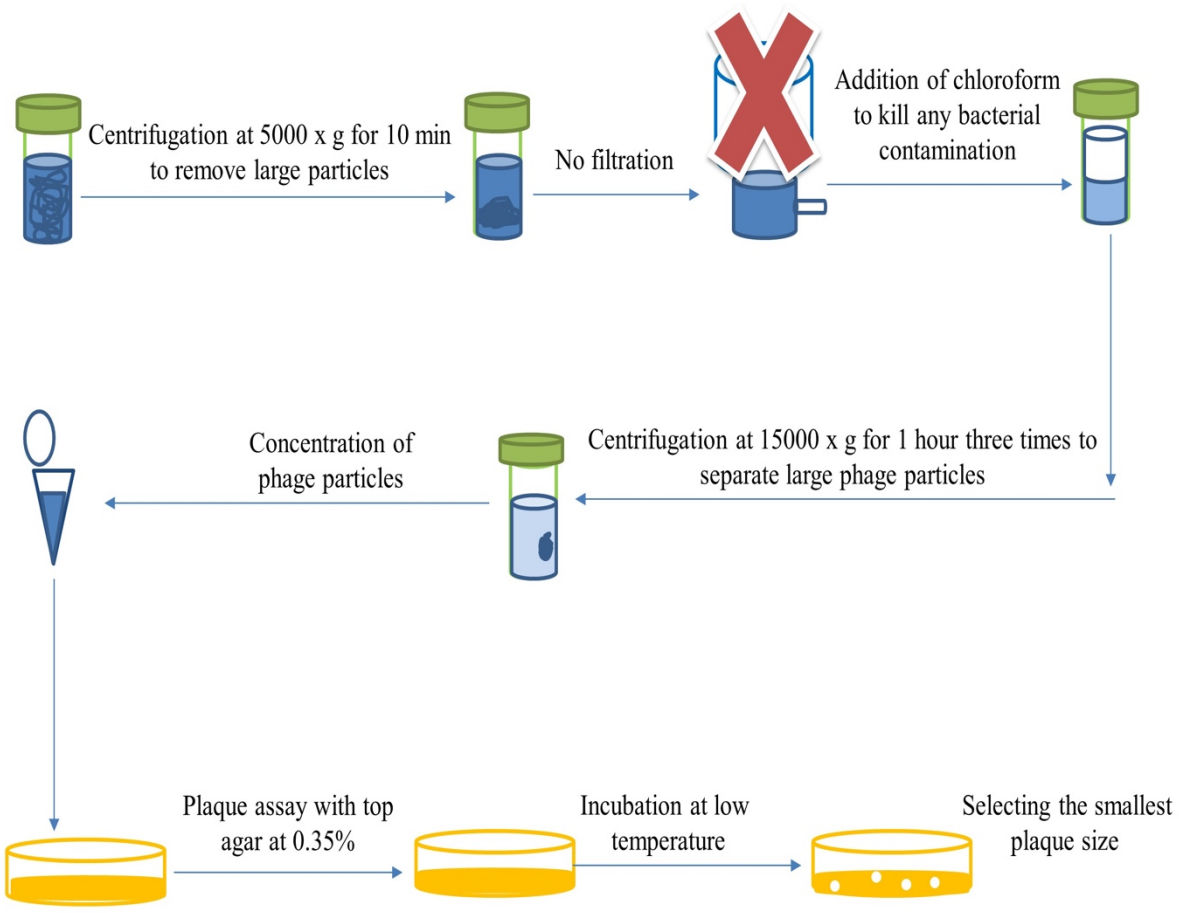


Fig. 2.1. Systemic detection and isolation of large phages. After treating with chloroform (without filtration), relatively large phage particles were separated by differential centrifugation. Plaque assays using low concentrations of top agar (0.35%) allow large phage particles to diffuse easily. Plaque assays are conducted at lower temperatures to delay host growth, resulting in larger plaque formation by phages. Smaller plaques are picked, as large phages always form very small plaques.

2.3.3. Host range assays

The host range of phages was determined initially by spot tests and then standard plaque-forming assays (Kutter, 2009). Test strains were mainly chosen from pathogenic bacteria (Table 2.1). Overnight cultures of bacterial strains prepared in LB medium were sub cultured in 500 µl LB broth and adjusted to OD_{600} = 1. An aliquot (100 µl) of each bacterial subculture was mixed with 3 ml of molten LB top agar (0.35%), poured onto the surface of an LB agar plate (1.5% agar) and left to dry for 10 min. Then, 1 µl of each phage preparation (with titer ~10^8 plaque forming units/ml) was spotted onto the bacterial overlay (Kutter, 2009), left for 15 min to dry, and then incubated at 30°C for 24 h. When a lysis zone appeared, efficiency of plating (EOP) was determined for each strain as the host (Kutter, 2009).

2.3.4. Electron microscopy for initial phage characterization

High titers of phages were obtained from liquid cultures of *E. coli* BL21 as the host. Phage particles were purified by sucrose gradient (20–40%) centrifugation at 40,000 × *g* for 1 h. Then, the concentrated phages were subjected to electron microscopic observation according to Ackermann (Ackermann, 2009). Briefly, phage particles were spotted onto a copper grid coated with formvar-carbon, left to adsorb for 2 min, stained with 1% Na-phosphotungstate, and observed under a Hitachi H600A electron microscope.

2.3.5. Pulsed-field gel electrophoresis for genomic analysis

The phage genome size was determined by pulsed-field gel electrophoresis (PFGE), as described by Higashiyama and Yamada (1991). Briefly, after purification of phage particles by sucrose gradient (20–40%) centrifugation at 40,000 × *g* for 1 h, phage particles were embedded in 1.7% low-melting-point agarose (InCert agarose; FMC Corp., Philadelphia, PA, USA). Phage-containing plugs were treated with proteinase K (1 mg/ml; Merck Ltd., Tokyo, Japan) and 1% sarkosyl at 55°C for 2 h, and subjected to PFGE by a using a CHEF Mapper electrophoresis apparatus (Bio-Rad, Hercules, CA, USA).

2.3.6. Isolation and characterization of genomic DNA from phage particles

Standard techniques for DNA isolation and digestion with restriction enzymes and other nucleases, including nuclease S1 were followed according to Sambrook and Russell (2001). Phage DNA was isolated from the purified phage particles using phenol-chloroform method. In brief, high-titre lysates of purified phage particles were incubated with Proteinase K (20 mg/mL) and 10% N-lauroylsarcosinate-Na for 2 h at 55°C. To the solution 500 µL of phenol: chloroform:

isoamyl alcohol (25:24:1) was added and centrifuged at 12,000 x g for 10 min (repeated twice). The aqueous phase was extracted into a new tube and an equal volume of 100% isopropanol was added and allowed DNA to precipitate. The precipitated DNA was washed with 70% ethanol and the purified phages DNA was visualized on 1% agarose gels.

For comparison of DNA fragment patterns, all isolated Escherichia phages DNAs (eleven genomic DNAs) were digested with *Eco*RV and genomic DNAs of E3, E4, E8, and E11 were digested with *Hae*III according to the instructions of the manufacturer (Roche). Restriction fragments were separated by electrophoresis (1.5h, 50V) on 1.0% agarose (Sigma, USA) gel stained with ethidium bromide. DNA molecular weight marker (Lambda-*Sty*I fragments) was used for size determination of DNA fragments.

2.4. RESULTS

2.4.1. Isolation of *Escherichia* large phages

From our experience in isolation and characterization of jumbo phages infecting *Ralastonia solanacaerum* (Yamada, 2010; Bhunchoth, 2015), I chose five conditions to select large phages infecting pathogenic enteric bacteria: **(i) Contaminating microbial cells were killed by chloroform without membrane filtration. (ii) Relatively large phage particles were selectively precipitated by differential centrifugation. (iii) Plaque assays used a low concentration of top agar (0.35%). (iv) Small plaques were selectively picked since large phages always form very small plaques. (v) Plaque assays were conducted at lower temperatures to delay host growth.** The last step was important to make minute plaques clearly visible. The effectiveness of this method is shown in Fig. 2.2 compared with a standard method (with 0.7% top agar and incubated at 37°C).

By using these approaches as described in Materials and Methods and Fig. 2.1, I started isolation of large phages from sewage samples collected from a large wastewater treatment plant in Higashi-Hiroshima. Eleven candidate phages (E1~E11) were finally obtained, which stably formed very small plaques (<0.1 mm on 0.35% top agar). *E. coli* BL21 was the ideal host strain for screening E1~E11 phages. Around 30 ± 5 plaques appeared on the first assay plates and the level of small plaques was about 16%. For comparison, no such small plaques were obtained by a conventional method with filtration and plaque assays with 0.7% top agar.

(A) **(B)**

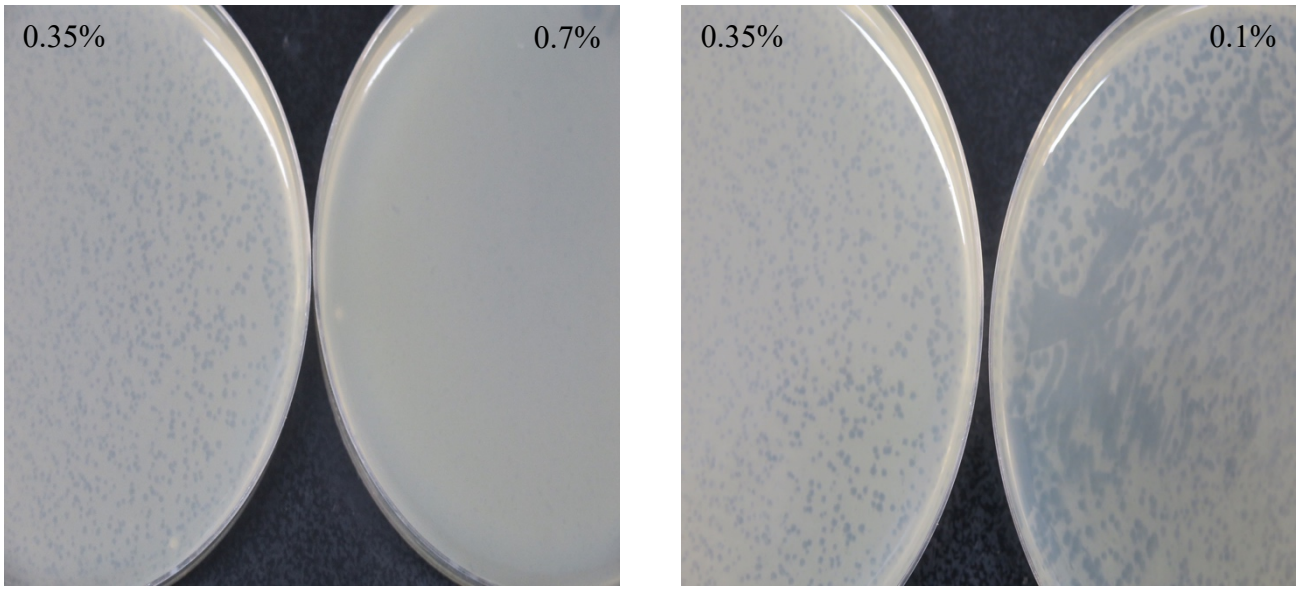


Fig. 2.2. Comparison of phage plaques formed by the new method (schematically shown in Fig. 2.1) compared with those by a standard method. (A) Relatively large clear plaques are visible for phage E8 with *Escherichia coli* BL21 as the host by the new method but plaques are very obscure by a standard method (the same phage and host with 0.7% top agar and incubated at 37°C). (B) However, the fragility of top agar at 0.1% hindering plaque handling and phage purification.

2.4.2. Electron microscopy observation

Electron microscopic observation of Escherichia phages E1~E11 revealed characteristic features of the family *Myoviridae* (Fig. 2.3). There are two types of particles: (i) T4 like particles characterized by a prolate head, a contractile tail, and long tail fibers (E1, E2, E5, E7, E9 and E10); (ii) large particles characterized by an icosahedral head (diameter 110 ± 5.5 nm) and a long contractile tail (length 145 ± 7.25 nm) (E3, E4, E6, E8 and E11).

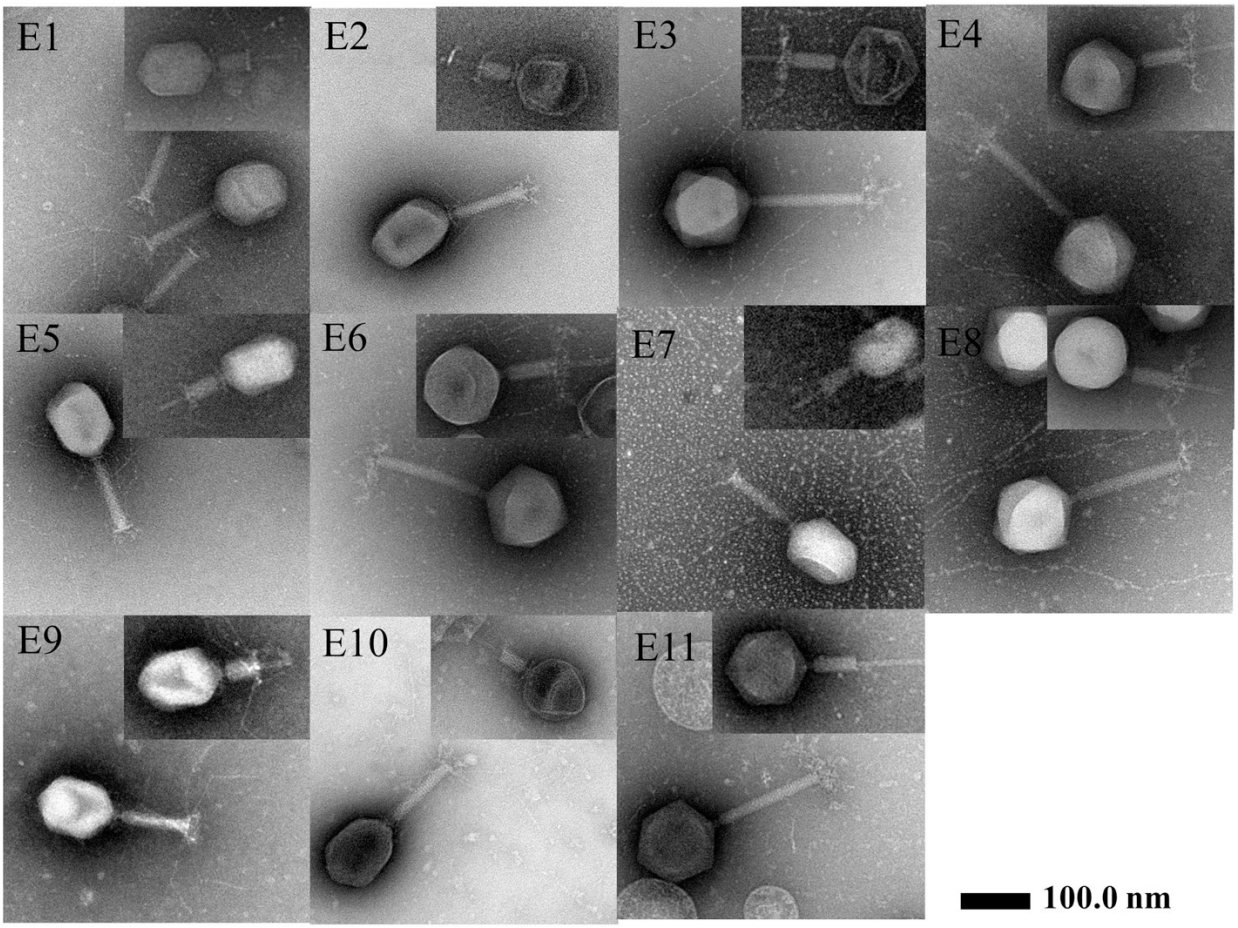


Fig. 2.3. Morphology of eleven Escherichia phages (E1~E11) stained with 1% Na-phospho tungstate and examined by transmission electron microscopy. All phages are members of the family *Myoviridae*. Six phages **(**E1, E2, E5, E7, E9, and E10**)** are typical Caudovirales phages characterized by a prolate head and contractile tail. The remaining five phages (E3, E4, E6, E8, and E11) showed large particles characterized by an icosahedral head and long contractile tail without visible tail fibers. A contracted tail is shown in each inlet.

2.4.3. Genomic analysis by PFGE

To determine genomic DNA sizes of the phages, I performed in-gel PFGE with purified phage particles. Fig. 2.4 shows the DNA separation patterns of these phages. The genomes of phages E1, E2, E5, E6, E7, E9 and E10 were around 200 kbp in size. While genome sizes of E3, E4, E8 and E11 were about 450 kbp. These sizes in PFGE were not affected by S1 nuclease digestion (Bhunchoth et al., 2015), indicating linear DNA forms of these genomes.

2.4.4. Digestion by restriction enzymes

*Eco*RV digestion patterns of genomic DNAs of Escherichia phages E1∼E11 are shown in Fig. 2.5A, where E1, E2, E6, E7, and E10 showed digested banding patterns, whereas E3, E4, E5, E8, E9, and E11 DNAs were not digested. Although large fragments for E1, E2, and E6 DNAs were not well separated, the digested fragmentation patterns were different from each other among the ∼200-kbp genome phages. As for ∼450-kbp genome phages, all E3, E4, E8, and E11 DNAs were digested with *Hae*III (Fig. 2.5B), giving banding patterns similar to each other among E3, E4, and E8 with some differences. E11 DNA gave a unique banding pattern. Other restriction enzymes we tried, failed to digest these large DNA genomes, suggesting some modification of the genomic DNA. Although exact genomic size for these phages could not be determined, 11 phages isolated in this study were large (putative jumbo) phages by the criterion of genome size ∼200 kbp.

2.4.5. Host range of the phages

Host ranges of phages E1∼E11 were determined with various bacterial strains (Table 2.2). E9 phage has a wide host range, infecting *Shigella sonnei* SH05001, *Salmonella enterica* serovar Enteritidis SAL01078 and *E. coli* C3000 (K-12 derivative), as well as its original host *E. coli* BL21. The most of other Escherichia phages could lyse multiple bacterial hosts, including *S. sonnei* SH05001, *Shigella boydii* SH00007, *Shigella flexneri* SH00006, *E. coli* C3000 and *E. coli* BL21. These results indicate that the large phages isolated using *E. coli* laboratory strains as hosts have a wide host range including important pathogens. This is the first demonstration that large phages infecting wide ranges of pathogenic bacteria can be screened in normal laboratories using nonpathogenic bacterial strains as the hosts.

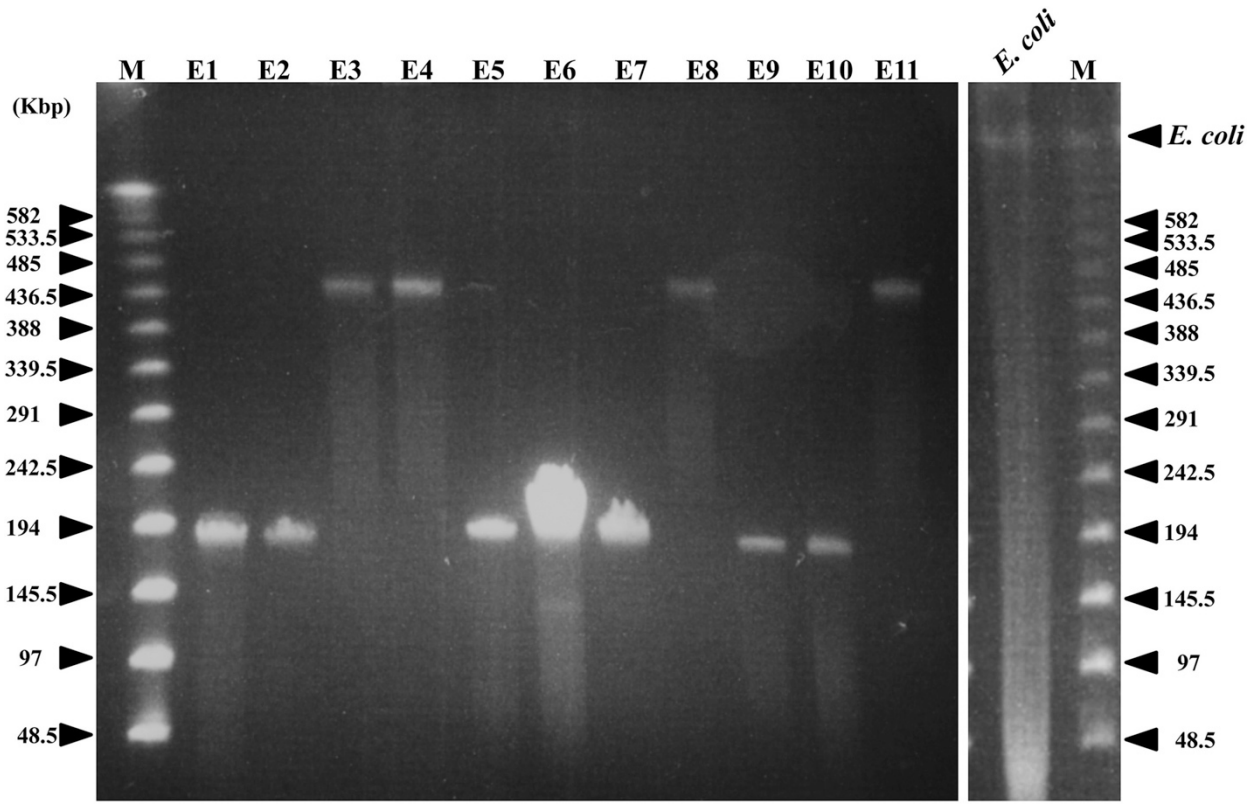


Fig. 2.4. Pulsed-field gel electrophoresis (CHEF) separation patterns of the genomic DNAs of 11 Escherichia phages (E1~E11). The genomic size of seven phages (E1, E2, E5, E6, E7, E9, and E10) is around 200 kbp, whereas that of the other four phages (E3, E4, E8 and E11) is around 450 kbp. For comparison, the genomic DNA of host *E. coli* cells (strain BL21) separated under the same condition is shown on the right side. M, lambda DNA ladder for size markers.

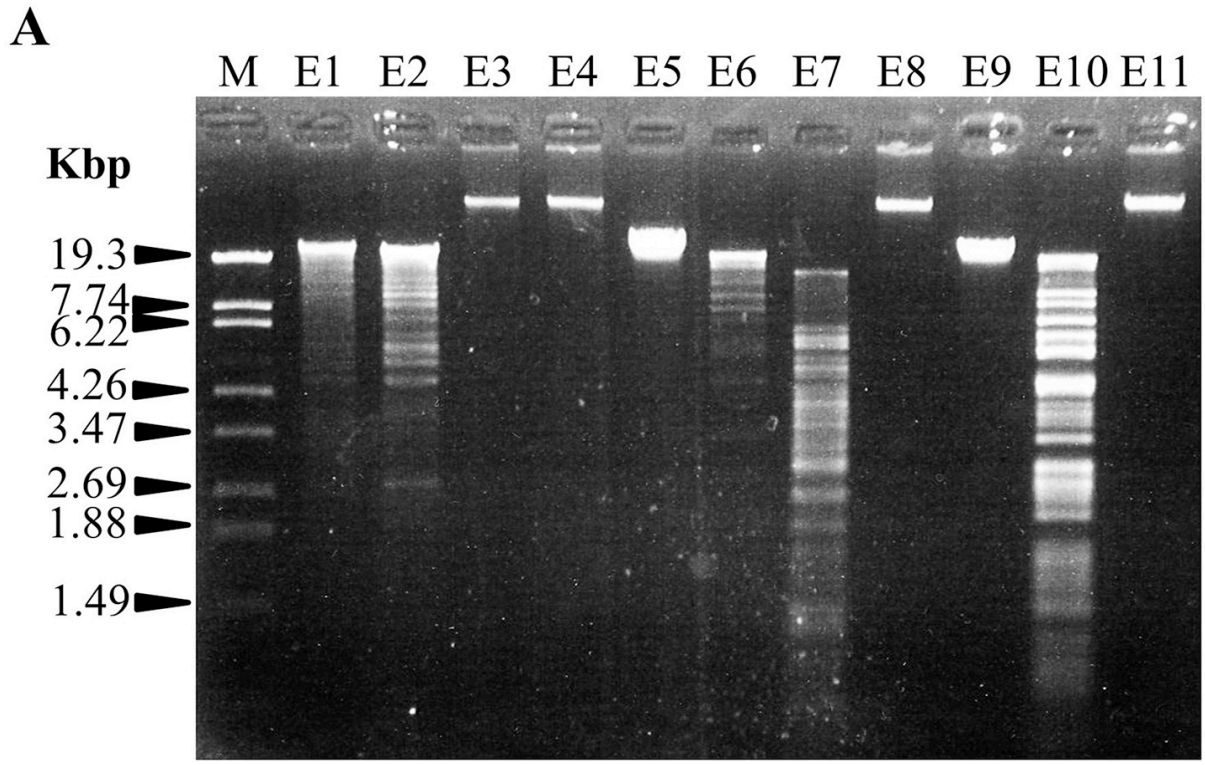


Fig. 2.5A. Restriction enzyme digestion patterns of genomic DNA of Escherichia phages. Eleven genomic DNAs were digested with *Eco*RV. M, molecular marker (Lambda-*Sty*I fragments).

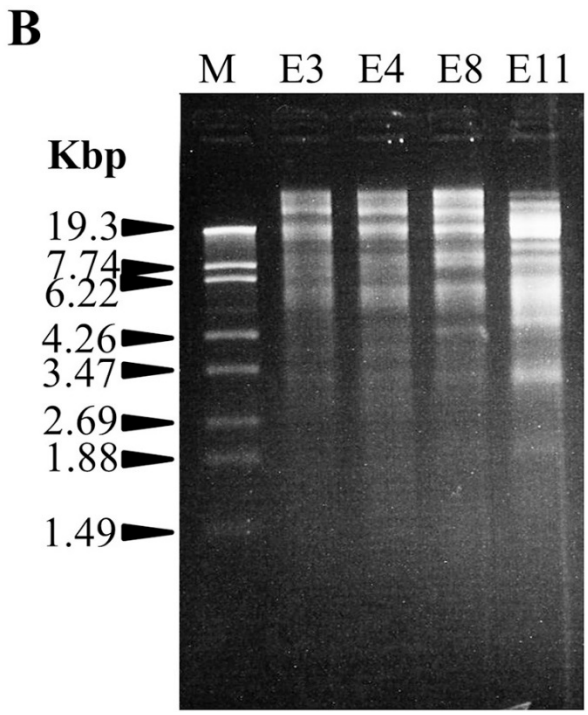


Fig. 2.5B. Restriction enzyme digestion patterns of genomic DNAs of E3, E4, E8, and E11. Genomic DNAs were digested with *Hae*III. M, molecular marker (Lambda-*Sty*I fragments).

Table 2.2. Host range of *Escherichia* phages on Gram negative and Gram-positive bacterial strains

Host (Bacterial species)	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11
Acinetobacter baumannii 395	-	-	-	-	-	-	-	-	-	-	-
Escherichia coli BL21	+	+	+	+	+	+	+	+	+	+	+
E. coli C3000 (C strain)	+	+	-	-	+	+	+	+	+	-	-
E. coli K12	+	+	-	-	-	-	+	-	-	-	-
E. coli O157:H7	-	-	-	-	-	-	-	-	-	-	-
Enterohaemorrhagic *E. coli* EHEC 03064	-	-	-	-	-	-	-	-	-	-	-
E. coli 7	-	-	-	-	-	-	-	-	-	-	-
Klebsiella pneumoniae 63	-	-	-	-	-	-	-	-	-	-	-
Proteus mirabilis 59	-	-	-	-	-	-	-	-	-	-	-
Providencia stuartii 50	-	-	-	-	-	-	-	-	-	-	-
Pseudomonas aeruginosa 16	-	-	-	-	-	-	-	-	-	-	-
Pseudomonas aeruginosa PAO1	-	-	-	-	-	-	-	-	-	-	-
Salmonella enterica serovar Paratyphi B SAL 04100	-	-	+	+	-	-	-	-	-	-	-
Salmonella enterica serovar Enteritidis SAL01078	-	-	-	+	-	-	-	-	+	-	-
Shigella sonnei SH05001	+	+	-	-	+	+	+	-	+	+	+
Shigella boydii SH00007	+	+	-	-	-	-	+	-	+	-	+
Shigella flexneri SH00006	+	-	-	-	-	-	+	-	+	+	+
Serratia marcescens D601	-	-	+	+	-	-	-	+	-	-	-
Serratia marcescens (ATCC 13880)	-	-	-	-	-	-	-	-	-	-	-
Vibrio cholerae O139 MDO6	-	-	-	-	-	-	-	-	-	-	-
Vibrio cholera P1418O1 E1	-	-	-	-	-	-	-	-	-	-	-
Vibrio fluvialis AQ0005	-	-	-	-	-	-	-	-	-	-	-
Vibrio parahaemolyticus ACA339	-	-	-	-	-	-	-	-	-	-	-
Bacillus subtilis 168	-	-	-	-	-	-	-	-	-	-	-
Staphylococcus aureus ATCC 6538	-	-	-	-	-	-	-	-	-	-	-
Staphylococcus epidermidis ATCC 35984	-	-	-	-	-	-	-	-	-	-	-

Sensitivity: +, sensitive (EOP>10^{-2} PUF/plate); -, resistant (EOP<10^{-6} PFU/plate). An EOP of 1 was equivalent to 408, 374, 280, 265, 355, 334, 369, 240, 530, 439, and 382 PFU per plate for E1 ~ E11 with *Escherichia coli* BL21 as the host, respectively.

2.5. DISCUSSION

In this work, I could selectively obtain many large phages by making several modifications to the classical methods for detection and isolation of bacteriophages. The modifications included: (i) Initial selection by differential centrifugation of samples to precipitate relatively large particles, and (ii) killing of contaminating bacteria with chloroform without filtration. This allowed us to remove small phages but retain large phage particles without contamination from small bacterial cells. (iii) Plaque assays using low concentrations of top agar (0.35%) and (iv) selective picking of smaller plaques, as large phages always form very small plaques. In the standard method applying top agar of ~0.7%, large phage particles cannot form visible plaques because of limited diffusion. Lowering the top agar concentration means that phage particles can diffuse more easily. However, even in this condition, plaques of large phages are usually small. (v) Only after plaque assays were conducted at lower temperatures to delay host growth, larger plaque formation was possible for these phages. With *E. coli* strains as the host, my preliminary tests of plaque formation by lambda phage gave the largest plaques at 23°C. Vigorous growth of the host strain at higher temperatures (e.g., 37°C for *E. coli*) leads to rapid spread of cells over the plate surface inactivating the host cells. Small plaques of large phages might be caused by other factors such as small burst size, low activity of lytic enzymes etc, but by using this systemic method, large phages easily missing from ordinary screening can be efficiently detected and isolated as shown in Fig. 2.2. However, even by using this method, some large phages may not be obtained because of their sensitivity to chloroform (Kęsik-Szeloch *et al*., 2013; Roszniowski *et al*., 2017). For such phages, some further refinements including extended differential centrifugation without chloroform treatment should be given to the method.

I used laboratory strains of *E. coli* as the hosts and obtained large phages (including T4-like phages) that showed wide-host ranges, including important pathogens such as strains of *Shigella* and *Salmonella*. Phages are usually species-specific and even strain-specific (Ross *et al*., 2016). Even though, some polyvalent phages infecting strains of either different genera or species have been reported, predominantly among phages of *Enterobacteria* (Hamdi et al., 2017; Doore *et al*., 2018) and *Staphylococcus* (Haddad *et al*., 2014). For example, Shigella phage SH7 was reported to infect strains of *E. coli*, *Salmonella* Paratyphi, and *Shigella dysenteriae* in addition to its original host *Shigella flexeri* (Hamdi *et al*., 2017). In respect of host strains ranging over three different genera including *Escherichia*, *Shigella*, and *Salmonella*, phage E9 found in this work is

similar to SH7. Phage E4 found in this work also showed a unique host range covering three genera such as *Escherichia*, *Salmonella*, and *Serratia*. Such a rare case of wide host range was also reported for bacteriophage χ (Iino and Mitani, 1967). These phages may be useful for effective biocontrol of pathogenic enteric bacteria used solely or in phage cocktails.

Handling of usual pathogenic bacterial strains in biosafety level 2 or 3 laboratories requires special experimental facilities, skills, and care. However, the screening method adapted here for large phages actually used only biosafety level 1 hosts. This is a big advantage of this method. It is thought that large phages with many adaptive genes generally have wide host ranges (Bhunchoth *et al*., 2015; Chowdhury and Sawer, 1976; Miller *et al*., 2003a; Kim *et al*, 2012), so that large phages will be good candidates for various biocontrol purposes. Long-lasting phage effects also increase the potential value of large phages (Fujiwara *et al*., 2011). I hope this method will open a new door to extend jumbo phage isolation and their use in biocontrol of a wide range of pathogenic bacteria.

CHAPTER III

Full genome sequence of a polyvalent bacteriophage infecting strains of *Shigella*, *Salmonella*, and *Escherichia*

3.1. SUMMARY

Escherichia phage EcS1 is a large, lytic myovirus isolated from sewage samples in Japan. The whole genome of EcS1 was found to be 175,437 bp in length with a mean G+C content of 37.8%. A total of 295 genes were identified along the genome and contained structural, functional, and hypothetical genes. BLAST analysis of the EcS1 genomic sequence revealed the highest identity (79%; query cover of 73–74%) to three T4-related phages that infect *Serratia* sp. ATCC 39006. Host range experiments revealed that EcS1 has lytic effects on three pathogenic strains of *Shigella* spp. and a pathogenic strain of *Salmonella enterica* as well as on *E. coli* strains. However, two strains of *Serratia marcescens* showed resistance to this phage. Phylogenetic trees for phage tail fiber protein sequences revealed that EcS1 is closely related to *Enterobacteriaceae*-infecting phages. Thus, EcS1 is a novel phage that infects several pathogenic strains of the family *Enterobacteriaceae*.

3.2. INTRODUCTION

Foodborne illness is a common public health problem all over the world. Bacterial organisms represent 66% of all food borne diseases etiology (Addis and Sisay, 2015). The bacterial family *Enterobacteriaceae* includes many important foodborne pathogens. Moreover, multidrug resistant strains of *Enterobacteriaceae* represent 71.5% of all infections with drug resistant strains (Picot-Guéraud *et al.,* 2015). Non-typhoidal Salmonella (NTS) are a leading cause of bacterial food-borne disease outbreaks and human gastroenteritis in both developed and developing countries (Aarestrup *et al*., 2007). They are responsible for considerable fatality rate of 20–25% globally, particularly in immune-compromised people (Hohmann, 2001; Evans *et al*., 2004). *S.*

enterica serovar Enteritidis (*S.* Enteritidis) emerged as the most prevalent serovar of (NTS) in many countries (Herikstad *et al*., 2002; Schroeder *et al*., 2005). Notably, *Shigella* was the second leading cause of enteric deaths globally in 2015, and, among the four *Shigella* subgroups, *S. sonnei* and *S. flexneri* are the leading causes of dysentery worldwide (Starling, 2017). *Shigella* spp. are reported to be increasingly resistant to various antimicrobial drugs (Nikfar *et al*., 2017), which necessitates the development of effective alternative choices for the control of these multidrug-resistant (MDR) strains. Phage therapy research to utilize bacteriophages for killing pathogenic bacteria has recently gained high interest due to the continued increase in MDR bacterial strains (Ansari *et al*., 2015). Bacteriophages could be an ideal alternative to antibiotics because they are ubiquitous in nature, self-replicating, host-specific, and able to overcome MDR strains, and they have potential for not only treating but also preventing bacterial infections (Bao *et al*., 2015).

Each isolated phage is new, and can be named by its owner, the sense of ownership in their discovery is a big motivation to explore the secrets of obtained phage by isolating genomic DNA, determining its sequence, annotating gene predictions, and comparing the sequence to that of other known viruses (Hatfull *et al*., 2006). Moreover, phages are genetically diverse, and their genomes are mosaic, driven by horizontal gene transfer with other phages and host genomes (Pedulla *et al*., 2003). As a consequence, phage evolution is complex and their genomes are composed of genes with distinct and varied evolutionary histories (Lawrence *et al*., 2002). It is highly important, chosen phages to be approved and released for use need to be well characterized, sequenced and the attendant phage biology understood (Merril *et al*, 2006).

Bacteriophage sequencing is highly essential for phages to be considered as biocontrol tools in the food industry and in the medical field by regulatory agencies such as the Food and Drug Administration (FDA). As detecting any of the following: genes that encode virulence factors, antimicrobial resistance, toxins, or transducing elements by genetic screening would immediately disqualify a phage for therapeutic use (Skurnik et al., 2007). Genome sequencing presents one of the most complete technique to study phage encoded proteins, however, the main bottleneck for functional overall genomics, including phage genomics, is the low number of available genomes and described genes (open reading frames or ORFs). Many of phage predicted proteins represent “hypothetical proteins” with none described function. This shows how low we know about genome and demonstrates the need of much more studies concerning genome

sequencing (Oliveira *et al*., 2012). Regarding to genetics, phages with broad host ranges are usually isolated from natural microbial communities that enhance genetic exchanges (Eisen, 2000). Among all phage genes, tail fiber genes appear to evolve faster than other, presumably because it seems advantageous to gain new specificities to infect different hosts and to enter other ecological niches (Casjens, 2005; Haggård-Ljungquist *et al*., 1992). The bacteriophage tail is essential molecular machine used for infection to recognize the host and ensure efficient genome delivery to the cell cytoplasm, as tail fibers/tail spikes in the phage tail facilitate the initial binding of the phage to the bacterial host and have roles in host specificity determination (Vinga *et al*., 2012). It is well known that *Myoviridae* phages have the most complex tail structures with the greatest number of proteins involved in the tail assembly and function (Leiman *et al*., 2010).

The main aim of this study is to make complete genetic characterization of the novel polyvalent bacteriophage EcS1 infecting strains of *Shigella*, *Salmonella*, and *Escherichia.* It is also to confirm phage EcS1 novelty and safety for biocontrol applications with identification of evolutionary changes of its tail fiber proteins.

3.3. MATERIALS AND METHODS

3.3.1. Isolation and expansion of phage EcS1:

Escherichia phage EcS1 was isolated from sewage samples collected from a sewage treatment plant in Higashi-Hiroshima, Japan, as previously described (Saad *et al*., 2019); briefly a 400 ml portion of the sewage sample was centrifuged at 5,000 × *g* for 10 min at room temperature to remove most particulates, including unwanted debris and microbial cells. After recentrifugation of the supernatant at 15,000 × *g* for 1 h at 4°C, all the pellets were suspended by vortexing in 5 ml of SM buffer (50 mM Tris-HCl [pH 7.5], 100 mM NaCl, 10 mM $MgSO_4$, and 0.01% gelatin). To selectively remove small phages, this centrifugation-resuspension cycle was repeated three times. An equal volume of chloroform was added to the final suspension to kill any residual contaminating bacteria. A plaque assay was done as follows: Bacterial culture of *E. coli* BL21 strain was prepared in 4.5 ml LB broth adjusted to $OD_{600\ nm}$ = 0.25. A 100 µl phage sample was mixed with 250 µl bacterial culture and stood for 10 min at room temperature to allow adsorption. With 4.5 ml of molten LB top agar (0.35%), the mixture was poured onto an LB agar plate (1.5% agar). The plates were incubated for 24 h at 23°C. Small plaques that appeared on the plates were

selectively picked for further purification and enrichment. Single plaque purification was repeated three times to confirm the plaque was derived from only one kind of phage. Enriched phage preparation was obtained from plate lysates.

3.3.2. Host range determination:

The host range of phage EcS1 against test strains [mainly chosen from pathogenic bacteria (Table 3.1)], was determined initially by spot tests and then standard plaque-forming assays (Kutter, 2009). Briefly; overnight cultures of bacterial strains prepared in LB medium were sub cultured in 500 µl LB broth and adjusted to OD600 = 1. An aliquot (100 µl) of each bacterial subculture was mixed with 3 ml of molten LB top agar (0.35%), poured onto the surface of an LB agar plate (1.5% agar) and left to dry for 10 min. Then, 1 µl of each phage preparation (with titer $\sim 10^8$ plaque forming units/ml) was spotted onto the bacterial overlay, left for 15 min to dry, and then incubated at 30°C for 24 h. When a lysis zone appeared, efficiency of plating (EOP) was determined for each strain as the host (Kutter, 2009). In detail, preparing cultures of positive bacterial strains in the host-range spot-test assay, and also for the host strain (*E. coli* BL21). Adding a series of 10-fold dilutions of phage EcS1 to each tested bacterial strain, going down to a predicted titer of 100 phage per ml. Spotting the last 6 dilutions of phage EcS1 and duplicating plates for each set check the plates and plaques can be counted for the next day.

3.3.3. Morphological characteristics of phage EcS1:

High titer of phage EcS1 was obtained from liquid culture of *E. coli* BL21 as a host. Phage particles were purified by sucrose gradient (20–40%) centrifugation at 40,000 × *g* for 1 h. Then, the concentrated phages were subjected to electron microscopic observation according to Ackermann (Ackermann, 2009). Briefly, phage particles were spotted onto a copper grid coated with formvar-carbon, left to adsorb for 2 min, stained with 1% Na-phosphotungstate, and observed under a Hitachi H600A electron microscope.

3.3.4. Phage genome isolation, sequencing, annotation and analyses

Phage DNA was isolated from the purified phage particles using phenol-chloroform method. In brief, high-titer lysates of purified phage particles were incubated with Proteinase K (20 mg/mL) and 10% N-lauroylsarcosinate-Na for 2 h at 55°C. To the solution 500 µL of phenol: chloroform: isoamyl alcohol (25:24:1) was added and centrifuged at 12,000 x g for 10 min (repeated twice). The aqueous phase was extracted into a new tube and an equal volume of 100% isopropanol was added and allowed DNA to precipitate. The precipitated DNA was washed with 70% ethanol.

The whole genome sequencing of Phage EcS1 was performed using an Illumina Miseq System (Fasmac System Science Co., Ltd. ,Kanagawa, Japan). The high-quality sequenced reads were de novo assembled using SPAdes Genome Assembler software, version 3.7.1 with sequence depth of 214. The complete genome sequence was submitted to DDBJ website and accession number was generated (LC371242). Functional annotation of the EcS1 genome and identification of open reading frames (ORFs) were performed using PHAST at http://phast.wishartlab.com/ (Zhou et al., 2011). The transfer-RNA (tRNA) sequences were predicted using tRNAscan-SE (version 1.3.1). The predicted genes of EcS1 were investigated for homologs on the NCBI using the BLAST program (https://blast.ncbi.nlm.nih.gov/) by applying an E-value cutoff le-^{5}. BLASTn analysis was performed to the whole EcS1 genome to detect any similarity to previously isolated phages on the NCBI. Followed by dot plot analyses using the MAFFT version 7 program (https://mafft.cbrc.jp/alignment/server/) for comparison of the genomic nucleotide sequences between EcS1 and Serratia myoviruses, including CH14 (accession no. MF036690), CBH8 (MF036691) and T4 (AF15810). A search by BLASTp was also done for all gene sequences of EcS1 phage.

3.3.5. Phylogenetic analysis of tail fibers gene products:

Amino acid sequences of tail fiber proteins (Gp198 and Gp270) were chosen by tail fiber gene products of BlastP for phylogenetic analysis. As Gp198 formed from 445 amino acids and putative for extra phage tail fibers and Gp270 formed from 977 amino acids and putative for long tail fiber distal subunit. Two phylogenetic trees were constructed using version 7.0 of Molecular Evolutionary Genetic Analysis (MEGA) (Kumar et al., 2016). The sequences were aligned with other similar phage sequences according to BlastP analyses. Phylogenetic trees were inferred using the Maximum—Likelihood method. The trees rooted and the bootstrapping supported scores using 1000 replicates. Protein alignments between EcS1 and Serrattia phages for AsiA (encoded from *asiA* gene) were conducted using the Clustal Omega program (Sievers *et al*., 2011).

3.4. RESULTS

3.4.1. Host range of EcS1 phage:

A novel lytic phage was isolated from sewage samples collected from a sewage treatment plant in Higashi-hiroshima, Japan using the non-pathogenic host strain *E. coli* BL21. After several rounds of purification, the phage was obtained and designated as "Escherichia phage EcS1". The host range of phage EcS1 was then tested on 19 bacterial strains by spot test and efficiency of plating (EOP) of diluted phage lysate. Of the 19 strains tested, phage EcS1 was able to infect a wide range of pathogenic bacterial strains, including *Shigella sonnei* SH05001, *Shigella boydii* SH00007, *S. flexneri* SH00006, and *Salmonella enterica* serovar Enteritidis (SAL 01078) (Table 1). The lytic nature of EcS1 was judged from its clear plaques stably formed with every sensitive host strain and EOP>10^{-2}pfu/plate was considered as sensitive.

3.4.2. EcS1 morphological characteristics:

The morphological characteristics of EcS1 were examined under a Hitachi H600A electron microscope. It was revealed that EcS1 particles have a prolate head (111 ± 5.5 nm long and 81 ± 4 nm wide, n = 10) and a long tail of 110 ± 5.5 nm in length, n = 10 (Fig. 1). These morphological characteristics as well as the following genomic information indicate that this phage belongs to the group of T4-like phages (genus *T4virus*) in the family *Myoviridae*.

3.4.3. Genomic analysis of EcS1 phage:

The final assembled size of the EcS1 genome was 175,437 bp in length with a mean G+C content of 38.2%. A total of 295 genes were identified along the EcS1 genome (Table 2). Of these genes, 117 had predicted known functions which encoded functional proteins required for DNA replication, repair, recombination, processing, transcription, nucleotide metabolism, host and/or phage interaction, homing endonuclease and lysis; 46 genes of those with predicted functions were identified as structural genes that encoded proteins such as phage capsid and scaffold, phage collar, tail, sheath and phage tail fibers; the remaining 178 genes lacked known functions as hypothetical proteins (Fig. 3.1) (Table. 3.3). A search by BLASTp of all gene sequences revealed that 242 shared highest identities with Serratia phages CHI14 and/or X20 while 53 gene products with highest identity to gene products of phages other than Serratia phages (Table 3.3). Within the EcS1 genome, there were 18 tRNA sequences detected by tRNAscan-SE (version 1.3.1) (Table 3.4). BLASTn analyses revealed that the sequence of phage EcS1 was most similar to three myophages ϕX20 (accession no. MF036692), ϕCHI14

(MF036690), and ϕCBH8 (MF036691), that infect *Serratia* sp. ATCC 39006 with a query cover of 73–74%. The highest identity between these sequences was 79%. In addition to those three phages; EcS1 also had 76% identity and a query coverage of 5% to T4 (GenBank accession number AF158101) (Fig. 3.2). Moreover, phage EcS1 had 12% query cover and 76% identity to Enterobacteria phage vB_EcoM_VR5 (accession number KP007359) and had 8% query cover and 75% identity to Shigella phage SP18 (accession number GQ981382) that isolated from *S. sonnei* KCTC 2518 (Kim et al., 2010a). And also, with 7% query cover and 74% identity with *Salmonella* phage (accession number JX181825).

3.4.4. Evolutionary changes in EcS1 tail fibers proteins:

The relationship between phage EcS1 and other members of enterobacteria phages was revealed by two phylogenetic trees (Fig. 3.3) based on the amino acid sequences of tail fiber proteins (Gp198 and Gp270). The constructed tree using phage tail fiber protein (Gp 198; 445 aa) showed that phage EcS1 formed 99% bootstrap value cluster with Klebsiella phage PMBT1 SCO64804. Moreover, the cluster was also closely related to Shigella phage Sf16, Escherichia phage vB_EcoM_AYO145A, and Salmonella phage BPS17W1 (Fig. 3.3A). On the other hand, the tree for long tail fiber distal subunit (Gp270; 977 aa) formed a 55% bootstrap value cluster with *Yersinia* phage vB YenM TG1, which was part of larger cluster that included *Citrobacter* phage CF1 and Salmonella phages vBSenMs 16 and Salmonella phages STML-198 (Fig. 3.3B).

There were other 4 functional gene products showing relative identity to other phages rather than Serratia phages as following; as (Gp12) that predicted to encode modifier of suppressor tRNAs with 96% coverage and 73% identity to Citrobacter phage Merlin, (Gp 194) encode head vertex protein have 99% coverage and 70% identity to Edwardsiella phage Pei20, (Gp 199) encoding hoc large outer capsid protein showed 98% coverage and 57% identity to Enterobacteria phage RB16, (Gp 252) that encode I-TevI homing endonuclease with 96% coverage and 65% identity to Yersinea phage phiR1.

3.4.5. Alignment of asiA genes (another evidence of genetic difference between EcS1 and Serratia phages):

For more declaration of the novelty of my phage EcS1; Alignment of the primary sequence of conserved asiA gene in EcS1 phage and other wild types of Serratia phages CH14, CBH8 and X20 revealed the presence of mutations in both C-terminal and N-terminal domains of phage EcS1 gene (Fig. 3.4).

Table 3.1: Host range of EcS1.

Host (Bacterial species)	EcS1
Acinetobacter baumannii 395 (carbapenem-resistant clinical isolate)	-
Escherichia coli BL21 (B strain)	+
E. coli C3000 (C strain)	+
E. coli K12	-
E. coli O157:H7	-
Enterohaemorrhagic *E. coli* EHEC 03064	-
E. coli 7 (carbapenem-resistant clinical isolate)	-
Klebsiella pneumoniae 63 (carbapenem-resistant clinical isolate)	-
Proteus mirabilis 59 (clinical isolate)	-
Providencia stuartii 50 (carbapenem-resistant clinical isolate)	-
Pseudomonas aeruginosa 16 (carbapenem-resistant clinical isolate)	-
Pseudomonas aeruginosa PAO1	-
Salmonella enterica serovar Paratyphi B SAL 04100	-
Salmonella enterica serovar Enteritidis SAL01078	+
Shigella sonnei SH05001	+
Shigella boydii SH00007	+
Shigella flexneri SH00006	+
Serratia marsencens D601	-
Serratia marcescens (ATCC 13880)	-

Sensitivity: +, sensitive (EOP>10^{-2} PUF/plate); -, resistant (EOP<10^{-6} PFU/plate). An EOP of 1 was equivalent to 2.12 x 10^{9}PFU/ml with *Escherichia coli* BL21 as the host, respectively.

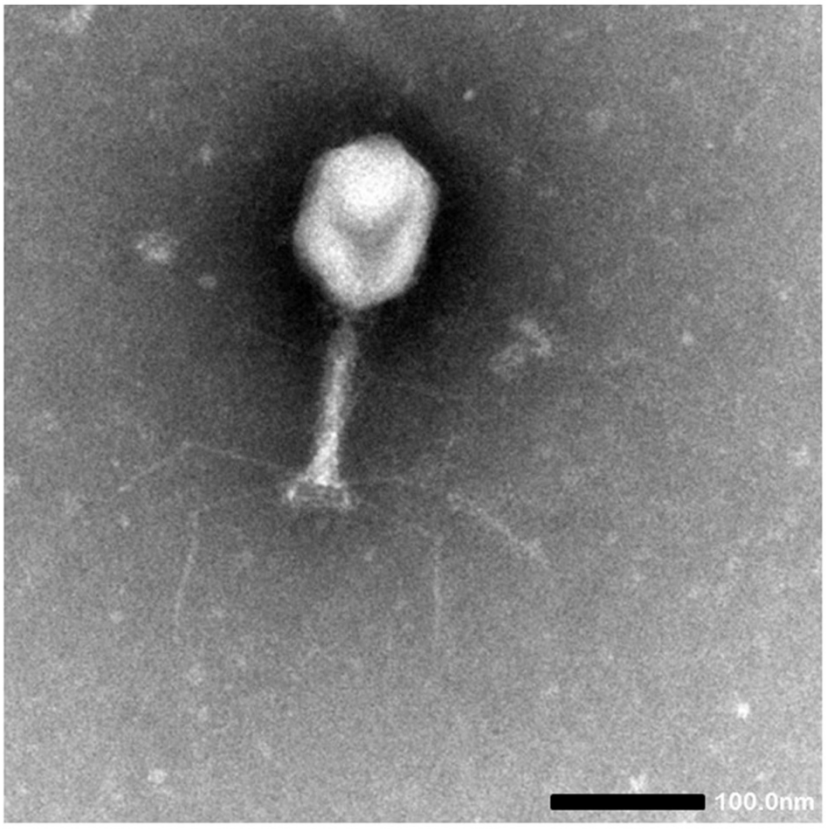


Fig. 3.1. Electron micrograph of a phage EcS1 particle stained with 2% phosphotugstate. Bar=100 nm.

Table 3.2: Predicted genes of EcS1

Gene	Position	GC (%)	Length of protein	MW (kDa)	Amino acid sequence identity/similarity to best homologs (% amino acid identity)	BLAST score (E-Value)	Accession no
g001	1 ← 2169	36	722	668	Protector from phage-induced early lysis [Serratia phage X20]	0.0	ARW57974
					Protector from prophage-induced early lysis [Serratia phage CHI14]	0.0	ARW57424
g002	2180 ← 2383	35	67	62	Hypothetical protein [Serratia phage X20]	4e-28	ARW57976
					Hypothetical protein [Serratia phage CHI14]	1e-27	ARW57425
g003	2370←2702	34	110	102	Hypothetical protein [Serratia phage CHI14]	6e-55	ARW57426
					Hypothetical protein CPT_Merlin3 [Citrobacter phage Merlin]	0.031	YP_009203717
g004	2769←2921	45	50	46	Hypothetical protein [Serratia phage CHI14]	2e-20	ARW57427
					Hypothetical protein CPT_Merlin3 [Citrobacter phage Merlin]	3e-11	YP_009203718
g005	2967←4805	37	612	566	DNA topoisomerase II large subunit [Serratia phage CHI14]	0.0	ARW57428
					DNA topoisomerase subunit [Serratia phage X20]	0.0	ARW57979
g006	4863←5087	35	74	69	Hypothetical protein PM2_004 [Pectobacterium bacteriophage PM2]	3e-04	YP_009211425
					LEAFY-like protein 1 [Fragaria x ananassa]	4.3	AHI62726
g007	5087←5335	40	82	76	gp39.1 hypothetical protein [Enterobacter phage CC31]	2e-37	YP_004009866
					Hypothetical protein [Serratia phage X20]	2e-34	ARW57980
g008	5332←5448	35	38	36	Hypothetical protein		
g009	5414←5608	37	64	59	FmdB family regulatory protein [Citrobacter phage Merlin]	1e-15	YP_009203725
					Hypothetical protein CPT_Moon11 [Citrobacter phage Moon]	4e-15	YP_009146444
g010	5614←6051	36	145	134	Hypothetical protein [Serratia phage X20]	1e-45	ARW57983
					Hypothetical protein [Serratia phage CHI14]	1e-32	ARW57432
g011	6051←6422	38	123	114	Hypothetical protein [Serratia phage X20]	4e-32	ARW57975
					Hypothetical protein [Serratia phage CHI14]	5e-32	ARW57433
g012	6422←6619	35	65	122	Modifier of suppressor tRNAs [Citrobacter phage Merlin]	1e-24	YP_009203727
					Modifier of suppressor tRNA [Citrobacter phage Moon]	1e-21	YP_009146446
g013	6688←7086	38	132	123	Hypothetical protein [Serratia phage CHI14]	6e-41	ARW57435
					NADPH dehydrogenase [Hyphopichia burtonii NRRL Y-1933]	7.4	XP_020073963
g014	7097←7780	41	227	211	Exonuclease A [Serratia phage X20]	2e-159	ARW57986
					Exonuclease A [Serratia phage CHI14]	7e-158	ARW57437
g015	7773←8012	35	79	73	Hypothetical protein [Serratia phage CHI14]	2e-26	ARW57438
					Hypothetical protein [Serratia phage X20]	5e-26	ARW57987
g016	8020←9345	38	441	408	DNA helicase [Serratia phage X20]	0.0	ARW57988
					DNA helicase [Serratia phage CHI14]	0.0	ARW57439
g017	9345←9641	35	98	91	Hypothetical protein [Serratia phage X20]	1e-44	ARW57989
					Hypothetical protein [Serratia phage CHI14]	8e-43	ARW57440
g018	9641←10351	37	236	218	Putative antisigma factor [Serratia phage CHI14]	5e-138	ARW57441
					Anti-sigma factor, putative [Serratia phage X20]	1e-137	ARW57990
g019	10396←10992	36	198	184	ADP-rybosylase [Serratia phage CHI14]	5e-64	ARW57442
					ADP-rybosylase [Serratia phage X20]	4e-63	ARW57991
g020	11051←11233	41	60	56	Hypothetical protein [Serratia phage X20]	2e-11	ARW57992
					Hypothetical protein [Serratia phage CHI14]	7e-11	ARW57443
g021	11233←11733	37	166	153	Hypothetical protein [Klebsiella phage KPV15]	1e-12	APD20401
						8e-12	YP_009288697

					Hypothetical protein kpv477_021 [Klebsiella phage vB_KpnM_KpV477]		
g022	11693←11908	36	71	66	Hypothetical protein [Serratia phage CHI14]	2e-08	ARW57445
					Hypothetical protein CPT_Merlin23 [Citrobacter phage Merlin]	8e-08	YP_009203737
g023	11865←12224	40	119	110	Hypothetical protein EpBp7_0113 [Enterobacteria phage Bp7]	7e-10	YP_007004154
					Hypothetical protein VR20_025 [Escherichia phage vB_EcoM_VR20]	8e-10	YP_009207204
g024	12297←12533	35	78	73	Small outer capsid protein [Salmonella phage STP4-a]	5e-18	YP_009126230
					Capsid and scaffold protein [Salmonella phage STML-198]	1e-17	YP_009148016
g025	12568←13089	37	173	160	gp56 dCTPase [Escherichia phage JS98]	2e-89	YP_001595154
					dCTPase [Escherichia phage QL01]	7e-89	YP_009202758
g026	13129←13428	34	99	92	Hypothetical protein [Serratia phage CHI14]	1e-38	ARW57450
					Hypothetical protein [Serratia phage X20]	1e-38	ARW57999
g027	13428←14450	38	340	308	DNA primase [Serratia phage CHI14]	0.0	ARW57452
					DNA primase [Serratia phage X20]	0.0	ARW58001
g028	14497←14964	39	155	143	Hypothetical protein CPT_Moon37 [Citrobacter phage Moon]	7e-60	YP_009146470
					Hypothetical protein [Escherichia phage ST0]	8e-60	ASD53796
g029	14981←15412	34	143	132	Hypothetical protein		
g030	15409←15651	36	80	75	Conserved hypothetical protein [Edwardsiella phage PEi20]	1e-09	YP_009190191
					gp61.4 hypothetical protein [Escherichia phage vB_EcoM_VR7]	4e-06	YP_004063718
g031	15648←15971	40	106	99	Hypothetical protein [Serratia phage X20]	3e-41	ARW58003
					Spackle periplasmic protein [Serratia phage CHI14]	1e-39	ARW57454
g032	15996←17360	40	454	420	DNA primase-helicase subunit [Serratia phage X20]	0.0	ARW58004
					Putative DNA primase/helicase [Serratia phage CHI14]	0.0	ARW57455
g033	17437←17661	36	74	69	Head vertex assembly chaperone [Serratia phage CHI14]	7e-43	ARW57456
					Head vertex assembly chaperone [Pectobacterium bacteriophage PM2]	1e-29	YP_009211460
g034	17777←18952	39	391	362	RecA-like recombination protein [Serratia phage CHI14]	0.0	ARW57457
					RecA-like recombination protein [Escherichia phage vB_EcoM_VR20]	0.0	YP_009207222
g035	19037←19579	33	180	167	Hypothetical protein [Serratia phage CHI14]	6e-95	ARW57459
					Hypothetical protein PM2_041 [Pectobacterium bacteriophage PM2]	2e-78	YP_009211462
g036	19567←20478	36	303	281	Hypothetical protein [Serratia phage CHI14]	0.0	ARW57460
					Hypothetical protein EpJS98_gp038 [Escherichia phage JS98]	1e-126	YP_001595167
g037	20478←21176	39	232	215	Putative thymidylate synthase [Serratia phage CHI14]	8e-154	ARW57461
					Hypothetical protein QL01_42 [Escherichia phage QL01]	1e-127	YP_009202773
g038	21173←21517	36	114	106	Hypothetical protein [Serratia phage CHI14]	2e-67	ARW57462
					Hypothetical protein fHeYen901_50 [Yersinia phage fHe-Yen9-01]	2e-15	ARB05823
g039	21514←22689	38	389	361	Conserved hypothetical protein [Edwardsiella phage PEi20]	0.0	YP_009190202
					Hypothetical protein [Serratia phage CHI14]	0.0	ARW57463
g040	22739←23098	37	119	111	Hypothetical protein fHeYen901_54 [Yersinia phage fHe-Yen9-01]	1e-80	ARB05827
					Capsule biosynthesis protein [Serratia phage CHI14]	3e-60	ARW57465
g041	23098←24777	37	559	517	Hypothetical protein [Serratia phage CHI14]	0.0	ARW57466
					Hypothetical protein [Serratia phage X20]	0.0	ARW58015
g042	24816←24959	34	47	44	Hypothetical protein [Serratia phage CHI14]	6e-17	ARW57467
					Hypothetical protein YenMTG1_055 [Yersinia phage vB_YenM_TG1]	2e-12	YP_009200316
g043	24959←25585	42	208	193	D-arabinose 5-phosphate isomerase [Serratia phage CHI14]	4e-122	ARW57468
					Conserved hypothetical protein [Edwardsiella phage PEi20]	5e-119	YP_009190207

g044	25585←26001	35	138	128	Hypothetical protein [Serratia phage X20]	9e-62	ARW58018
					Hypothetical protein [Serratia phage CHI14]	1e-61	ARW57469
g045	26050←28761	39	903	836	DNA polymerase [Serratia phage CHI14]	0.0	ARW57470
					DNA polymerase [Shigella phage Shf125875]	0.0	YP_009100595
g046	28837←29199	35	120	111	Translational repressor protein [Serratia phage CHI14]	1e-77	ARW57471
					Endoribonuclease translational repressor of early genes [Klebsiella phage KPV15]	8e-75	APD20422
g047	29202←29780	33	192	178	Clamp loader small subunit [Serratia phage CHI14]	4e-102	ARW57472
					Clamp loader subunit, DNA polymerase accessory protein [Pectobacterium bacteriophage PM2]	4e-101	YP_009211476
g048	29771←30730	35	319	295	Clamp loader small subunit [Enterobacteria phage vB_EcoM_VR5]	0.0	YP_009205746
					Hypothetical protein WG01_53 [Escherichia phage WG01]	0.0	YP_009323252
g049	30788←31477	36	229	213	Sliding clamp protein [Serratia phage CHI14]	7e-149	ARW57474
					Sliding clamp protein [Serratia phage X20]	3e-147	ARW58023
g050	31523←31909	39	128	119	RNA polymerase binding protein [Serratia phage X20]	8e-59	ARW58024
					RNA polymerase binding protein [Serratia phage CHI14]	6e-57	ARW57475
g051	32110←33792	37	560	518	Recombination endonuclease subunit [Pectobacterium bacteriophage PM2]	0.0	YP_009211481
					Recombination endonuclease subunit [Edwardsiella phage PEi20]	0.0	YP_009190217
g052	33789←34004	35	71	66	Hypothetical protein QL01_57 [Escherichia phage QL01]	1e-20	YP_009202788
					46.1 gene product [Enterobacteria phage IME08]	3e-20	
g053	34001←35020	37	339	314	Recombination endonuclease subunit [Serratia phage CBH8]	0.0	ARW57753
					Recombination endonuclease subunit [Serratia phage CHI14]	0.0	ARW57478
g054	35083←35343	38	86	80	46.1 gene product [Enterobacteria phage IME08]	4e-29	YP_003734205
					Hypothetical protein kpv477_055 [Klebsiella phage vB_KpnM_KpV477]	4e-16	YP_009288731
g055	35324←35509	35	61	57	Hypothetical protein [Serratia phage CHI14]	1e-29	ARW57481
					Prohead core protein protease [Citrobacter phage Moon]	2.0	YP_009146494
g056	35494←35619	44	41	38	Hypothetical protein CC31p054 [Enterobacter phage CC31]	2e-14	YP_004009912
					Conserved hypothetical protein [Edwardsiella phage PEi20]	1e-13	YP_009190221
g057	35669←36022	39	117	109	Hypothetical protein [Serratia phage CHI14]	7e-63	ARW57483
					Hypothetical protein CGG41_053 [Salmonella phage vB_SnwM_CGG4-1]	3e-37	YP_009286419
g058	36025←36177	35	50	47	Hypothetical protein [Serratia phage CHI14]	1e-25	ARW57484
					Hypothetical protein [Serratia phage X20]	5e-25	ARW58031
g059	36236←36769	37	177	164	Sigma factor for late transcription [Serratia phage X20]	7e-115	ARW58032
					Sigma factor for late transcription [Serratia phage CHI14]	1e-114	ARW57485
g060	36836←37102	35	88	81	Hypothetical protein [Serratia phage X20]	1e-37	ARW58033
					Hypothetical protein [Serratia phage CHI14]	8e-37	ARW57486
g061	37099←37356	35	85	79	Hypothetical protein		
g062	37353←37583	40	76	71	Hypothetical protein [Serratia phage CHI14]	8e-35	ARW57487
					Hypothetical protein [Serratia phage X20]	2e-34	ARW58034
g063	37576←37824	32	82	76	Hypothetical protein CPT_Merlin69 [Citrobacter phage Merlin]	3e-15	YP_009203783
					Conserved hypothetical protein [Edwardsiella phage PEi20]	3e-13	YP_009190228
g064	38063←38371	36	102	95	Hypothetical protein PM2_071 [Pectobacterium bacteriophage PM2]	2e-17	YP_009211492
					Hypothetical protein CC31p062 [Enterobacter phage CC31]	2e-14	YP_004009920
g065	38368←38700	40	110	102	Hypothetical protein [Serratia phage X20]	8e-56	ARW58035
					Conserved hypothetical protein [Edwardsiella phage PEi20]	8e-56	YP_009190230
g066	38762←38995	30	77	72	Hypothetical protein JS09_0111 [Escherichia phage vB_EcoM_JS09]	8e-08	YP_009037434
						8e-06	YP_009225191
					Hypothetical protein APCEc01_231 [Escherichia phage APCEc01]		
g067	39003←39263	37	86	80	Hypothetical protein, partial [Streptococcus pyogenes]	3e-31	WP_085577434
					Hypothetical protein [Serratia phage X20]	3e-27	ARW58036

g068	39329←39640	41	103	96	Glutaredoxin [Serratia phage CHI14] Glutaredoxin [Serratia phage X20]	1e-61 4e-61	ARW57490 ARW58037
g069	39621←39893	37	90	83	Hypothetical protein [Serratia phage CHI14] Hypothetical protein [Serratia phage X20]	6e-41 1e-39	ARW57491 ARW58038
g070	39895←40137	34	80	75	Hypothetical protein [Serratia phage CHI14]	1e-17	ARW57492
g071	40146←40262	37	38	36	Hypothetical protein [Serratia phage X20]	0.003	ARW58040
g072	40263←40736	38	157	145	Anaerobic NTP reductase small subunit [Serratia phage CHI14] Anaerobic NTP reductase small subunit [Serratia phage X20]	5e-96 2e-95	ARW57493 ARW58041
g073	40723←42546	38	607	563	Anaerobic NTP reductase large subunit [Serratia phage CHI14] Anaerobic NTP reductase large subunit [Serratia phage X20]	0.0 0.0	ARW57494 ARW58042
g074	42543←43016	38	157	146	Recombinase endonuclease VII [Serratia phage X20] Recombinase endonuclease VII [Serratia phage CHI14]	2e-106 2e-105	ARW58043 ARW57495
g075	43047←43349	34	100	93	Hypothetical protein [Serratia phage CHI14] Hypothetical protein [Serratia phage X20]	2e-23 1e-20	ARW57496 ARW58044
g076	43336←43785	35	149	138	Inhibitor of host Lon protease [Serratia phage CHI14] Protease inhibitor [Serratia phage X20]	5e-67 2e-64	ARW57497 ARW58045
g077	43766←43921	40	51	47	Hypothetical protein JS09_096 [Escherichia phage vB_EcoM_JS09] Hypothetical protein [Escherichia phage ST0]	4e-13 4e-13	YP_009037419 ASD53741
g078	43906←44208	34	100	92	Hypothetical protein CPT_Moon81 [Citrobacter phage Moon] Hypothetical protein [Serratia phage X20]	1e-42 4e-42	YP_009146514 ARW58046
g079	44201←44395	37	64	59	Hypothetical protein [Serratia phage X20] Hypothetical protein [Serratia phage CHI14]	3e-37 2e-36	ARW58047 ARW57499
g080	44396←44548	39	50	47	Hypothetical protein [Serratia phage CHI14] Hypothetical protein PKO111_076 [Klebsiella phage PKO111]	2e-24 1e-15	ARW57500 YP_009289477
g081	44545←44841	36	98	91	Hypothetical protein [Serratia phage X20] Hypothetical protein [Serratia phage CHI14]	8e-38 3e-37	ARW58049 ARW57501
g082	44831←45001	31	56	52	Hypothetical protein [Serratia phage X20] Hypothetical protein [Serratia phage CHI14]	4e-23 6e-23	ARW58050 ARW57502
g083	44998←45273	39	91	85	Thioredoxin [Serratia phage X20] Thioredoxin [Serratia phage CHI14]	4e-57 7e-55	ARW58052 ARW57504
g084	45236←45541	36	101	93	Hypothetical protein [Serratia phage X20] Hypothetical protein [Serratia phage CHI14]	3e-57 3e-33	ARW58053 ARW57505
g085	45519←45677	32	52	48	Hypothetical protein CGG41_076 [Salmonella phage vB_SnwM_CGG4-1] Secreted glycosyl hydrolase [Salmonella phage STP4-a]	1e-08 2e-08	YP_009286442 YP_009126288
g086	45674←45961	35	95	88	Conserved hypothetical protein [Edwardsiella phage PEi20] Hypothetical protein PG7_077 [Enterobacter phage PG7]	2e-17 9e-17	YP_009190244 YP_009005341
g087	45961←46467	36	168	156	Hypothetical protein [Serratia phage CHI14] Hypothetical protein [Serratia phage X20]	9e-86 2e-84	ARW57507 ARW58055
g088	46595←46816	34	43	39	Hypothetical protein [Serratia phage CHI14] Hypothetical protein [Serratia phage X20]	2e-12 2e-12	ARW57508 ARW58056
g089	46806←47219	33	73	68	Hypothetical protein [Serratia phage CHI14] Hypothetical protein [Serratia phage X20]	6e-27 8e-27	ARW57508 ARW58056
g090	47216←47452	35	137	127	NrdC.7 conserved hypothetical, predicted membrane protein [Enterobacteria phage RB69] Hypothetical protein PhAPEC2_80 [Escherichia phage vB_EcoM_PhAPEC2]	6e-27 6e-26	NP_861784 YP_009056672
g091	47445←48491	36	348	316	Hypothetical protein [Serratia phage X20] Hypothetical protein [Serratia phage CHI14]	4e-96 8e-95	ARW58058.1 ARW57510.1

g092	48594←49370	39	258	239	Hypothetical protein [Serratia phage CHI14]	2e-158	ARW57511
					Hypothetical protein [Serratia phage X20]	2e-155	ARW58059
g093	49429←49635	38	68	63	Hypothetical protein CGG41_084 [Salmonella phage vB_SnwM_CGG4-1]	3e-08 6e-08	YP_009286450 YP_007501127.1
					Hypothetical protein [Salmonella phage vB_SenMS16]		
g094	49635←49994	38	119	111	Hypothetical protein [Serratia phage X20]	7e-29	ARW58061
					Hypothetical protein [Serratia phage CHI14]	8e-29	ARW57513
g095	49994←50167	35	57	53	Hypothetical protein		
g096	50160←50642	42	160	149	Hypothetical protein [Serratia phage X20]	1e-102	ARW58063
					Hypothetical protein [Serratia phage CHI14]	8e-99	ARW57515
g097	50734←51225	38	163	151	Hypothetical protein [Serratia phage CHI14]	5e-87	ARW57516
					Hypothetical protein [Serratia phage X20]	3e-85	ARW58064
g098	51222←51446	37	74	68	Hypothetical protein [Serratia phage X20]	6e-20	ARW58065
					Hypothetical protein [Serratia phage CHI14]	1e-19	ARW57517
g099	51448←51654	39	68	63	Hypothetical protein [Serratia phage X20]	7e-36	ARW58066
g100	51714←52007	33	97	90	Hypothetical protein		
g101	52014←52991	39	325	301	Hypothetical protein [Serratia phage X20]	0.0	ARW58067
					Hypothetical protein [Serratia phage CHI14]	0.0	ARW57518
g102	53063←53491	34	142	131	Conserved hypothetical protein [Edwardsiella phage PEi26]	6e-62	BAQ23049
					Conserved hypothetical protein [Edwardsiella phage PEi20]	1e-61	YP_009190256
g103	53615←53836	32	73	68	Hypothetical protein [Serratia phage X20]	7e-32	ARW58068
					Hypothetical protein [Serratia phage CHI14]	1e-31	ARW57519
g104	53848←54069	39	64	60	Hypothetical protein [Serratia phage CHI14]	2e-36	ARW57520
					Hypothetical protein CPT_Moon102 [Citrobacter phage Moon]	1e-26	YP_009146535
g105	54047←55060	40	337	312	Hypothetical protein PM2_105 [Pectobacterium bacteriophage PM2]	0.0	YP_009211526
					Thioredoxin [Serratia phage X20]	0.0	ARW58069
g106	55057←55479	35	121	112	Hypothetical protein [Serratia phage CHI14]	7e-60	ARW57522
					Hypothetical protein [Serratia phage X20]	4e-59	ARW58070
g107	55469←55708	34	79	74	Hypothetical protein [Serratia phage CHI14]	2e-24	ARW57523
					Hypothetical protein [Serratia phage X20]	2e-22	ARW58071
g108	55701←56222	36	173	161	Hypothetical protein [Serratia phage X20]	7e-64	ARW58072
					Hypothetical protein [Serratia phage CHI14]	3e-63	ARW57524
g109	56261←56749	35	162	150	Hypothetical protein [Serratia phage X20]	1e-60	ARW58073
					Hypothetical protein [Serratia phage CHI14]	4e-57	ARW57525
g110	56739←56966	34	75	70	Hypothetical protein CPT_Merlin105 [Citrobacter phage Merlin]	2e-06	YP_009203819
					Hypothetical protein CPT_Moon107 [Citrobacter phage Moon]	8e-06	YP_009146540
g111	56975←57238	33	87	81	Hypothetical protein [Serratia phage X20]	3e-25	ARW58075
					Hypothetical protein [Serratia phage CHI14]	4e-20	ARW57527
g112	57228←57410	31	60	56	Hypothetical protein		
g113	57412←57705	42	97	90	Putative C4-type zinc finger protein [Serratia phage X20]	6e-52	ARW58077
					Hypothetical protein PM2_112 [Pectobacterium bacteriophage PM2]	1e-42	YP_009211533
g114	57733←58200	38	155	144	Hypothetical protein [Serratia phage X20]	1e-37	ARW58079
					Hypothetical protein [Serratia phage CHI14]	8e-36	ARW57531
g115	58303←58683	38	126	117	Hypothetical protein [Serratia phage X20]	9e-61	ARW58080
					Hypothetical protein [Serratia phage CHI14]	1e-59	ARW57532
g116	58673←58963	36	96	89	Lysis inhibition regulator membrane protein [Serratia phage CHI14]	9e-54	ARW57533
					Lysis inhibitor regulator [Serratia phage X20]	4e-53	ARW58081
g117	58973←59185	39	70	66	Conserved hypothetical protein [Edwardsiella phage PEi20]	6e-24	YP_009190270
					Hypothetical protein [Serratia phage CHI14]	6e-20	ARW57534
g118	59232←59576	39	114	106	Hypothetical protein CPT_Moon115 [Citrobacter phage Moon]	5e-48	YP_009146548
					Hypothetical protein CPT_Merlin114 [Citrobacter phage Merlin]	6e-46	YP_009203828

g119	59635←59757	35	40	37	Hypothetical protein		
g120	59759←59929	32	56	52	Hypothetical protein [Serratia phage CHI14]	6e-14	ARW57536
					Hypothetical protein AB185_26225 [Klebsiella oxytoca]	5e-04	AKL38916
g121	59913←60506	40	197	182	Thymidine kinase [Serratia phage CHI14]	6e-122	ARW57537
					Thymidine kinase [Serratia phage X20]	1e-121	ARW58085
g122	60506←60718	37	70	65	Hypothetical protein CPT_Merlin118 [Citrobacter phage Merlin]	8e-32	YP_009203832
					Hypothetical protein CGG41_102 [Salmonella phage vB_SnwM_CGG4-1]	1e-31	YP_009286469
g123	60718←60918	40	66	61	Hypothetical protein [Serratia phage X20]	7e-09	ARW58086
					Putative keratin [Serratia phage CHI14]	7e-09	ARW57538
g124	60915←61094	35	59	55	Hypothetical protein		
g125	61094←61240	35	48	45	Hypothetical protein		
g126	61267←61476	29	69	74	Hypothetical protein phAPEC8_0048 [Escherichia phage phAPEC8]	2e-29	YP_007348416
					Hypothetical protein ESCO5_00044 [Escherichia phage ESCO5]	8e-29	AOT23420
g127	61448←61939	36	163	151	Hypothetical protein [Serratia phage CHI14]	3e-94	ARW57539
					Hypothetical protein SHP1_051 [Salmonella phage SHP1]	3e-64	ASJ79373
g128	61939←62289	38	117	108	Valyl-tRNA synthetase modifier [Serratia phage CHI14]	5e-62	ARW57540
					Valyl-tRNA synthetase modifier [Pectobacterium bacteriophage PM2]	6e-35	YP_009211545
g129	62289←62837	38	182	168	Hypothetical protein [Serratia phage CHI14]	2e-114	ARW57541
					Conserved hypothetical protein [Edwardsiella phage PEi20]	4e-99	YP_009190277
g130	62848←63306	38	152	141	Site-specific RNA endonuclease [Serratia phage CHI14]	6e-84	ARW57542
					Site-specific RNA endonuclease [Pectobacterium bacteriophage PM2]	3e-70	YP_009211547
g131	63362←63634	37	90	84	Hypothetical protein YenMTG1_120 [Yersinia phage vB_YenM_TG1]	4e-36	YP_009200381
						4e-36	YP_007235949
					Phage protein [Yersinia phage phiR1-RT]		
g132	63627←63863	36	78	72	Hypothetical protein [Serratia phage CHI14]	2e-31	ARW57544
					Hypothetical protein kpv477_122 [Klebsiella phage vB_KpnM_KpV477]	1e-21	YP_009288798
g133	63845←64183	34	112	104	Hypothetical protein CC31p127 [Enterobacter phage CC31]	1e-38	YP_004009985
					Hypothetical protein PKO111_044 [Klebsiella phage PKO111]	6e-23	YP_009289445
g134	64180←64701	38	173	160	Hypothetical protein [Serratia phage CHI14]	5e-50	ARW57546
					Hypothetical protein CPT_Merlin132 [Citrobacter phage Merlin]	2e-45	YP_009203846
g135	64952←65401	39	149	138	Hypothetical protein		
g136	65416←65829	36	137	127	Endonuclease V, N-glycosylase UV repair enzyme [Edwardsiella phage PEi20]	2e-78	YP_009190283
					Endonuclease V, N-glycosylase UV repair enzyme [Enterobacter phage CC31]	1e-73	YP_004009988
g137	65885←66172	38	95	88	Hypothetical protein [Serratia phage CHI14]	4e-59	ARW57548
					Hypothetical protein PM2_132 [Pectobacterium bacteriophage PM2]	1e-34	YP_009211553
g138	66169←66657	40	162	150	Lysozyme murein hydrolase [Serratia phage CHI14]	6e-102	ARW57549
					Lysozyme murein hydrolase [Enterobacter phage CC31]	3e-95	YP_004009990
g139	66696←67169	44	157	145	NTP pyrophosphohydrolases including oxidative damage repair enzymes [Cronobacter phage Pet-CM3-4]	7e-76	SCN45809
						3e-74	YP_009005393
					Nudix hydrolase [Enterobacter phage PG7]		
g140	67162←67401	33	79	73	Hypothetical protein PG7_130 [Enterobacter phage PG7]	1e-11	YP_009005394
					Hypothetical protein CC31p134 [Enterobacter phage CC31]	1e-11	YP_004009992
g141	67401←67706	36	101	94	Conserved hypothetical protein [Edwardsiella phage PEi26]	5e-30	BAQ23082
					Hypothetical protein CC31p135 [Enterobacter phage CC31]	1e-29	YP_004009993
g142	67703←68041	35	112	104	Hypothetical protein [Serratia phage X20]	4e-36	ARW58100
					Hypothetical protein [Serratia phage CHI14]	1e-35	ARW57552
g143	68038←68544	35	168	156	Hypothetical protein [Serratia phage CHI14]	4e-57	ARW57553
					Hypothetical protein [Serratia phage CBH8]	4e-57	ARW57828

g144	68651←68989		112		Hypothetical protein		
g145	68989←69360	30	123	114	Hypothetical protein [Serratia phage CHI14]	0.72	ARW57557
g146	69360←70172	33	270	250	Hypothetical protein [Serratia phage CHI14]	8e-127	ARW57558
					3-hydroxyacyl-CoA dehydrogenase [Ruminococcus sp. CAG:379]	1.4	CDD54212
g147	70212←70817	42	201	187	Hypothetical protein [Serratia phage CHI14]	2e-115	ARW57559
					Hypothetical protein CPT_Merlin150 [Citrobacter phage Merlin]	1e-85	YP_009203864
g148	70875←71075	33	66	62	Hypothetical protein BN80_132 [Yersinia phage phiR1-RT]	7e-28	YP_007235965
					Hypothetical protein [Salmonella phage vB_SenMS16]	2e-19	YP_007501160
g149	71189←71794	39	201	187	Hypothetical protein [Serratia phage CHI14]	1e-103	ARW57561
					Hypothetical protein EpJS98_gp135 [Escherichia phage JS98]	8e-66	YP_001595264
g150	72355←72525	33	56	52	DUF1837 domain-containing protein [Lactococcus garvieae]	6.8	WP_017371069
g151	72759←72983	32	74	69	Hypothetical protein PG7_148 [Enterobacter phage PG7]	7e-36	YP_009005412
					Hypothetical protein [Cronobacter phage Pet-CM3-4]	5e-28	SCN45824
g152	72990←73436	38	148	137	Hypothetical protein [Serratia phage X20]	6e-89	ARW58112
					Hypothetical protein [Serratia phage CHI14]	6e-88	ARW57563
g153	73528←73782	40	84	78	Hypothetical protein [Serratia phage X20]	5e-20	ARW58114
					Hypothetical protein [Serratia phage CHI14]	5e-20	ARW57565
g154	73776←74114	36	112	104	Hypothetical protein kpv477_146 [Klebsiella phage vB_KpnM_KpV477]	0.034	YP_009288822
						0.18	YP_002854094.1
					Hypothetical protein RB51ORF141 [Enterobacteria phage RB51]		
g155	74134←74304	29	56	52	Hypothetical protein kpv477_140 [Klebsiella phage vB_KpnM_KpV477]	3e-06	YP_009288816
						1e-05	YP_009289437
					Putative vertex head subunit [Klebsiella phage PKO111]		
g156	74459←74743	36	94	88	Hypothetical protein [Serratia phage X20]	4e-35	ARW58115
					Hypothetical protein [Serratia phage CHI14]	8e-23	ARW57566
g157	75423←75668	31	81	75	Hypothetical protein [Serratia phage CHI14]	1e-31	ARW57567
					Conserved hypothetical protein [Edwardsiella phage PEi20]	1e-10	YP_009190310
g158	76530←77015	32	161	149	Hypothetical protein kpv477_148 [Klebsiella phage vB_KpnM_KpV477]	2e-51	YP_009288824
					PHG31p119nc [Aeromonas virus 31]	2e-39	YP_238848
g159	77395←77586	39	63	58	Hypothetical protein [Serratia phage CHI14]	8e-26	ARW57569
					Hypothetical protein [Edwardsiella phage PEi20]	7e-08	YP_009190315
g160	77579←77788	34	69	64	Hypothetical protein [Serratia phage CHI14]	2e-28	ARW57570
					Hypothetical protein [Serratia phage X20]	9e-27	ARW58118
g161	77769←77960	35	63	59	Hypothetical protein [Serratia phage X20]	1e-29	ARW58119
					Hypothetical protein [Serratia phage CHI14]	7e-29	ARW57571
g162	78022←78234	39	70	65	Hypothetical protein WG01_147 [Escherichia phage WG01]	1e-23	YP_009323346
					Hypothetical protein CPT_Merlin166 [Citrobacter phage Merlin]	9e-12	YP_009203880
g163	78231←78686	41	151	141	Hypothetical protein [Serratia phage X20]	4e-105	ARW58121
					Hypothetical protein [Serratia phage CHI14]	2e-104	ARW57573
g164	78686←78934	36	82	76	Chaperone for long tail fiber formation [Serratia phage X20]	3e-25	ARW58122
					57A chaperone for long tail fiber formation [Salmonella phage vB_SnwM_CGG4-1]	7e-07	YP_009286495
g165	78934←79623	36	229	212	dNMP kinase [Serratia phage X20]	8e-141	ARW58123
					dNMP kinase [Serratia phage CHI14]	2e-139	ARW57574
g166	79625←80212	40	195	181	Tail completion and sheath stabilizer protein [Serratia phage CHI14]	2e-133	ARW57575
					gp3 tail completion and sheath stabilizer protein [Enterobacteria phage JS10]	1e-105	YP_002922490

g167	80307←81137	35	276	256	DNA end protector protein [Serratia phage CHI14]	0.0	ARW57576
					DNA end protector protein [Serratia phage X20]	0.0	ARW58126
g168	81137←81586	36	149	138	Head completion protein [Serratia phage CHI14]	7e-88	ARW57577
					Head completion protein [Klebsiella phage PKO111]	1e-70	YP_009289410
g169	81635→82210	37	191	178	Baseplate wedge subunit [Serratia phage CHI14]	2e-113	ARW57578
					Baseplate wedge subunit [Citrobacter phage Merlin]	1e-92	YP_009203887
g170	82210→83937	38	575	535	Baseplate hub + tail lysozyme [Serratia phage CHI14]	0.0	ARW57579
					Baseplate hub + tail lysozyme [Serratia phage X20]	0.0	ARW58129
g171	84003→84452	39	149	139	Hypothetical protein [Serratia phage X20]	6e-70	ARW58130
					Hypothetical protein [Serratia phage CHI14]	2e-68	ARW57580
g172	84452→84745	39	97	90	Hypothetical protein CGG41_137 [Salmonella phage vB_SnwM_CGG4-1]	1e-56	YP_009286504
					Hypothetical protein [Serratia phage X20]	1e-56	ARW58131
g173	84745→86700	37	651	604	Baseplate wedge subunit [Serratia phage X20]	0.0	ARW58133
					Baseplate wedge subunit [Serratia phage CHI14]	0.0	ARW57581
g174	86697→89786	36	1029	955	Baseplate wedge initiator [Serratia phage CHI14]	0.0	ARW57582
					Baseplate wedge initiator [Serratia phage X20]	0.0	ARW58134
g175	89786→90790	39	334	311	Baseplate wedge subunit [Serratia phage X20]	0.0	ARW58135
					Baseplate wedge subunit [Serratia phage CHI14]	0.0	ARW57583
g176	90853→91722	37	289	269	Baseplate wedge tail fiber connector [Serratia phage CHI14]	4e-156	ARW57584
					Baseplate wedge tail finer connector [Serratia phage X20]	3e-155	ARW58136
g177	91722→93548	36	608	565	Baseplate wedge subunit and tail pin [Serratia phage X20]	0.0	ARW58137
					Baseplate wedge subunit and tail pin [Serratia phage CHI14]	0.0	ARW57585
g178	93548→94231	37	227	211	Baseplate wedge subunit and tail pin [Serratia phage X20]	5e-103	ARW58138
					Baseplate wedge subunit and tail pin [Serratia phage CHI14]	7e-103	ARW57586
g179	94231→95775	40	514	478	Short tail fiber protein [Serratia phage X20]	0.0	ARW58139
					Short tail fiber protein [Serratia phage CHI14]	0.0	ARW57587
g180	95785→97608	38	607	564	Fibritin neck whisker protein [Serratia phage CHI14]	0.0	ARW57588
					Fibritin neck whisker protein [Serratia phage X20]	0.0	ARW58140
g181	97673→98614	39	313	292	Neck protein [Serratia phage CHI14]	0.0	ARW57589
					Neck protein [Serratia phage X20]	0.0	ARW58141
g182	98617→99393	34	258	240	Neck protein [Serratia phage X20]	1e-164	ARW58142
					Neck protein [Serratia phage CHI14]	3e-164	ARW57590
g183	99438→100271	38	277	257	Tail sheath stabilizer and completion protein [Serratia phage CHI14]	4e-176	ARW57591
					Tail sheath stabiliser and completion protein [Serratia phage X20]	4e-175	ARW58143
g184	100268→100765	36	165	154	Terminase DNA packaging enzyme small subunit [Serratia phage CHI14]	1e-102	ARW57592
					gp16 terminase DNA packaging enzyme, small subunit [Enterobacter phage CC31]	3e-95	YP_004010026
g185	100749→102584	38	611	568	Terminase DNA packaging enzyme large subunit [Serratia phage CHI14]	0.0	ARW57593
					Terminus DNA packaging enzyme, large subunit [Serratia phage X20]	0.0	ARW58145
g186	102620→104602	41	660	613	Hypothetical protein [Serratia phage CHI14]	0.0	ARW57594
					Tail sheath protein [Enterobacteria phage vB_EcoM_VR5]	0.0	YP_009205860
g187	104714→105205	41	163	152	Tail tube protein [Serratia phage CHI14]	8e-116	ARW57596
					Tail tube protein [Edwardsiella phage PEi20]	9e-112	YP_009190345
g188	105322→106890	37	522	485	Portal vertex protein of head [Serratia phage X20]	0.0	ARW58149

					Portal vertex protein of head [Serratia phage CHI14]	0.0	ARW57597
g189	106890→107165	40	91	85	Prohead core protein [Serratia phage CHI14]	3e-19	ARW57598
					Prohead core protein [Serratia phage X20]	4e-19	ARW58150
g190	107166→107591	40	141	131	Prohead core protein [Serratia phage CHI14]	9e-80	ARW57599
					Prohead core protein [Salmonella phage vB_SenMS16]	7e-64	YP_007501201
g191	107591→108235	43	214	199	Prohead assembly (scaffolding) protein and protease [Serratia phage CHI14]	5e-141	ARW57600
					Prohead assembly (scaffolding) protein [Salmonella phage STML-198]	3e-133	YP_009148153
g192	108268→109074	39	268	249	Prohead core scaffold protein [Serratia phage CHI14]	3e-142	ARW57601
					Prohead core scaffold protein [Serratia phage X20]	3e-141	ARW58154
g193	109096→110655	45	519	482	Major capsid protein [Serratia phage CHI14]	0.0	ARW57602
					Major capsid protein [Citrobacter phage Moon]	0.0	YP_009146626
g194	110756→112036	38	426	402	Head vertex protein [Edwardsiella phage PEi20]	0.0	YP_009190352
					Head vertex protein [Enterobacter phage PG7]	0.0	YP_009005456
g195	12074←113078	38	334	309	RNA ligase 2 [Serratia phage X20]	0.0	ARW58158
					RNA ligase 2 [Serratia phage CHI14]	0.0	ARW57605
g196	113177←113449	34	90	83	Hypothetical protein [Serratia phage X20]	7e-49	ARW58159
					Hypothetical protein CPT_Moon197 [Citrobacter phage Moon]	2e-17	YP_009146630
g197	113451←113702	34	83	77	Hypothetical protein [Serratia phage X20]	7e-34	ARW58160
					Hypothetical protein [Serratia phage CHI14]	5e-32	ARW57607
g198	113787→115124	39	445	412	Phage tail fibers [Klebsiella phage PMBT1]	3e-86	SCO64804
					Whisker protein [Enterobacter phage phiEap-3]	2e-83	ALA45334
g199	115134→115631	38	165	153	Hoc large outer capsid protein [Enterobacteria phage RB16]	7e-48	YP_003858524
					Protein of unknown function [Enterobacteria phage RB43]	7e-48	CCL97698
g200	115662←116375	36	237	220	Inhibitor of prohead protease [Serratia phage X20]	3e-109	ARW58162
					Inhibitor of prohead protease [Serratia phage CHI14]	3e-107	ARW57609
g201	116428→117927	35	499	464	RNA-DNA and DNA-DNA helicase [Serratia phage X20]	0.0	ARW58163
					Helicase [Enterobacteria phage vB_EcoM_VR5]	0.0	YP_009205877
g202	117924→118151	38	75	70	RNA-DNA and DNA-DNA helicase [Serratia phage CHI14]	4e-31	ARW57611
					Putative split helicase [Enterobacter phage PG7]	1e-25	YP_009005466
g203	118213←118377	40	54	50	Hypothetical protein [Serratia phage CHI14]	4e-30	ARW57612
					Hypothetical protein [Salmonella phage vB_SenMS16]	2e-23	YP_007501220
g204	118410←118625	35	71	66	UvsY.-1 hypothetical protein [Enterobacteria phage RB51]	1e-19	YP_002854140
					Hypothetical protein [Clostridioides difficile]	1e-19	WP_074146452
g205	118625←119059	38	144	133	Recombination repair and ssDNA binding protein [Serratia phage CHI14]	9e-96	ARW57613
					UvsY gene product [Enterobacteria phage IME08]	7e-70	YP_003734328
g206	119117←119470	37	112	104	Baseplate wedge subunit [Serratia phage CHI14]	2e-70	ARW57614
					Putative baseplate wedge subunit [Escherichia coli]	1e-57	WP_016039599
g207	119515←120144	36	209	194	Baseplate hub subunit [Serratia phage CHI14]	7e-123	ARW57615
					Putative baseplate hub subunit [Klebsiella phage vB_KpnM_KpV477]	1e-77	YP_009288876
g208	120195→120947	32	250	232	Baseplate hub assembly catalyst [Serratia phage X20]	6e-147	ARW58169
					Baseplate hub assembly catalyst [Serratia phage CHI14]	9e-147	ARW57616
g209	120944→122152	36	402	373	Baseplate hub subunit [Serratia phage CHI14]	0.0	ARW57617
					Baseplate hub subunit [Serratia phage CBH8]	0.0	ARW57892
g210	122199→122606	34	135	126	Baseplate hub distal subunit [Serratia phage X20]	7e-66	ARW58171

					Baseplate hub distal subunit [Serratia phage CHI14]	1e-62	ARW57618
g211	122587→124353	38	588	547	Base plate hub subunit, tail length determinator [Serratia phage CHI14] Baseplate hub subunit tail length determinator [Serratia phage X20]	8e-178 1e-148	ARW57619 ARW58172
g212	124362→125420	37	352	327	Baseplate tail tube cap [Serratia phage X20] Baseplate tail tube cap [Serratia phage CHI14]	0.0 0.0	ARW58173 ARW57620
g213	125420→126385	41	321	299	Baseplate subunit [Serratia phage CHI14] Baseplate subunit [Serratia phage X20]	0.0 0.0	ARW57621 ARW58174
g214	126418←126729	33	1103	96	Hypothetical protein [Serratia phage CHI14] Hypothetical protein [Serratia phage X20]	6e-47 2e-46	ARW57622 ARW58175
g215	126787←128850	37	687	635	RNA polymerase ADP-ribosylase [Serratia phage X20] RNA polymerase ADP-ribosylase [Serratia phage CHI14]	0.0 0.0	ARW58176 ARW57623
g216	128914←129108	34	64	60	Hypothetical protein [Serratia phage X20] Hypothetical protein HX01_0245 [Escherichia phage HX01]	1e-33 2e-23	ARW58177 YP_006907287
g217	129105←130619	38	504	465	DNA ligase [Serratia phage X20] DNA ligase [Serratia phage CHI14]	0.0 0.0	ARW58178 ARW57625
g218	130839←131678	36	279	258	Hypothetical protein JD18_193 [Klebsiella phage JD18] Hypothetical protein kpv477_212 [Klebsiella phage vB_KpnM_KpV477]	1e-94 4e-93	YP_009190774 YP_009288888
g219	131653←131829	38	58	54	Hypothetical protein [Serratia phage CHI14] Hypothetical protein [Serratia phage X20]	2e-13 1e-12	ARW57628 ARW58181
g220	131822←132019	38	65	60	Hypothetical protein [Serratia phage X20] Hypothetical protein [Serratia phage CHI14]	2e-25 1e-24	ARW58182 ARW57629
g221	131994←132137	38	47	44	Hypothetical protein [Serratia phage X20] Hypothetical protein JD18_195 [Klebsiella phage JD18]	8e-21 6e-08	ARW58183 YP_009190776
g222	132197←132364	36	55	51	Hypothetical protein [Serratia phage X20] Hypothetical protein [Serratia phage CHI14]	1e-25 2e-25	ARW58184 ARW57630
g223	132364←132813	39	149	138	Hypothetical protein [Serratia phage CHI14] Hypothetical protein [Serratia phage X20]	8e-46 2e-43	ARW57631 ARW58185
g224	132854←133249	35	131	121	Hypothetical protein CC31p208 [Enterobacter phage CC31] Hypothetical protein PG7_223 [Enterobacter phage PG7]	3e-64 2e-62	YP_004010066 YP_009005487
g225	133280←133783	37	176	154	Conserved hypothetical protein [Edwardsiella phage PEi20] Conserved hypothetical protein [Edwardsiella phage PEi26]	4e-14 4e-14	YP_009190382 BAQ23176
g226	133818←134171	36	117	108	Hypothetical protein CPT_Melville_208 [Salmonella phage Melville] Hypothetical protein [Clostridioides difficile]	2e-51 1e-49	ATN93171 WP_074146635
g227	134233←134427	38	64	59	Hypothetical protein CC31p211 [Enterobacter phage CC31] Hypothetical protein [Cronobacter phage Pet-CM3-4]	4e-05 1e-04	YP_004010069 SCN45904
g228	134476←134655	44	59	58	Gp30.9 conserved hypothetical protein [Yersinia phage phiR1-RT] Hypothetical protein fHeYen901_229 [Yersinia phage fHe-Yen9-01]	2e-20 4e-20	YP_007236043 ARB06002
g229	134977←135225	40	82	77	Lysis inhibition accessory protein rapid lysis phenotype [Serratia phage CHI14] MULTISPECIES: hypothetical protein [Bacteria]	2e-50 1e-35	ARW57637 WP_015969374
g230	135339←135668	36	238	109	Head assembly cochaperone with GroEL [Salmonella phage STP4-a] Head assembly chaperone protein [Salmonella phage STML-198]	2e-49 1e-48	YP_009126151 YP_009147934
g231	135725←136048	38	107	99	Hypothetical protein [Serratia phage CHI14] Conserved hypothetical protein [Edwardsiella phage PEi20]	3e-70 5e-52	ARW57639 YP_009190388
g232	136049←136600	39	183	170	DCMP deaminase [Serratia phage CBH8] DCMP deaminase [Serratia phage CHI14]	8e-104 8e-103	ARW57915 ARW57640

g233	136600←136935	36	111	103	Hypothetical protein [Serratia phage CHI14] Hypothetical protein [Serratia phage X20]	3e-48 1e-46	ARW57641 ARW58195
g234	136916←137926	41	336	311	Phospho-2-dehydro-deoxyheptonate aldolase [Serratia phage X20] Phospho-2-dehydro-3-deoxyheptonate aldolase [Serratia phage CHI14]	3e-174 1e-165	ARW58196 ARW57642
g235	137926←138156	39	76	71	Hypothetical protein [Serratia phage X20] Hypothetical protein [Serratia phage CHI14]	2e-34 1e-33	ARW58197 ARW57643
g236	138158←138619	33	153	142	Conserved hypothetical protein [Edwardsiella phage PEi26] Conserved hypothetical protein [Edwardsiella phage PEi20]	3e-06 4e-06	BAQ23187 YP_009190393
g237	138616←138819	34	67	62	Hypothetical protein SP18_gp234 [Shigella phage SP18] Hypothetical protein VR5_222 [Enterobacteria phage vB_EcoM_VR5]	9e-08 1e-07	YP_003934857 YP_009205916
g238	138819←139718	41	299	277	Polynucleotide kinase [Serratia phage CHI14] Polynucleotide kinase [Serratia phage X20]	0.0 0.0	ARW57645 ARW58199
g239	139715←139927	38	70	65	Hypothetical protein [Serratia phage CHI14] Hypothetical protein vBEcoMUFV13_g221 [Escherichia phage vB_EcoM-UFV13]	4e-18 6e-09	ARW57646 YP_009290486
g240	140096←140335	43	60	56	Hypothetical protein fHeYen901_242 [Yersinia phage fHe-Yen9-01]	0.007	ARB06015
g241	140322←140501	34	79	73	No significant similarity found		
g242	140498←140995	36	59	55	Hypothetical protein QL01_227 [Escherichia phage QL01] Hypothetical protein MX01_220 [Escherichia phage MX01]	2e-06 2e-06	YP_009202958 YP_009324117
g243	141198←141494	39	72	66	Outer membrane lipoprotein [Serratia phage CHI14] Outer membrane lipoprotein Rz1 [Serratia phage X20]	5e-40 3e-39	ARW57648 ARW58202
g244	141491←141823	36	110	102	I-spanin [Citrobacter phage Merlin] putative membrane protein [Serratia phage X20]	5e-27 2e-26	YP_009203963 ARW58203
g245	141835←142341	37	168	156	Inhibitor of host transcription [Salmonella phage vB_SenMS16] Hypothetical protein [Salmonella phage STML-198]	3e-75 3e-75	YP_007501255 YP_009147948
g246	142400←143572	38	390	362	RNA ligase 1 and tail fiber attachment [Serratia phage CHI14] RNA ligase 1 and tail fiber attachment catalyst [Serratia phage X20]	0.0 0.0	ARW57651 ARW58205
g247	143569←143982	33	137	127	Endonuclease II [Serratia phage CHI14] Hypothetical protein WG01_232 [Escherichia phage WG01]	5e-87 5e-75	ARW57652 YP_009323431
g248	144011←145189	39	392	363	Aerobic NDP reductase, small subunit [Serratia phage CHI14] Aerobic NTP reductase, small subunit [Serratia phage X20]	0.0 0.0	ARW57653 ARW58207
g249	145186←145359	33	57	53	Hypothetical protein [Isosphaera pallida]	7.7	WP_013566058
g250	145352←145555	31	67	62	Hypothetical protein [Serratia phage CHI14] Hypothetical protein PM2_251 [Pectobacterium bacteriophage PM2]	2e-34 1e-13	ARW57654 YP_009211672
g251	145594←147855	40	753	697	Aerobic NDP reductase, large subunit [Serratia phage X20] Aerobic NDP reductase, large subunit [Serratia phage CHI14]	0.0 0.0	ARW58208 ARW57655
g252	147908←148666	32	252	234	I-TevI homing endonuclease [Yersinia phage phiR1-RT] Endonuclease I-Tevi [Klebsiella phage PKO111]	6e-98 3e-86	YP_007236067 YP_009289561
g253	148633→149496	40	287	266	dTMP thymidylate synthase [Serratia phage X20] dTMP thymidylate synthase [Serratia phage CHI14]	0.0 0.0	ARW58210 ARW57657
g254	149496←149867	37	123	92	Hypothetical protein [Serratia phage CHI14] Hypothetical protein [Serratia phage X20]	9e-57 1e-56	ARW57658 ARW58211
g255	149848←150189	37	113	105	Hypothetical protein [Serratia phage X20] Hypothetical protein [Serratia phage CHI14]	5e-63 6e-63	ARW58212 ARW57659
g256	150170←150778	39	202	188	Dihydrofolate reductase [Serratia phage CHI14]	6e-93	ARW57660

					Dihydrofolate reductase [Serratia phage X20]	6e-91	ARW58213
g257	150775←151011	33	78	72	Hypothetical protein [Edwardsiella phage PEi26]	0.40	BAQ23206
g258	151008←151211	34	67	62	Hypothetical protein [Serratia phage CHI14]	9e-09	ARW57661
					Hypothetical protein PG7_256 [Enterobacter phage PG7]	5e-06	YP_009005520
g259	151302←151580	40	92	85	Hypothetical protein [Serratia phage CHI14]	4e-25	ARW57662
					Hypothetical protein [Serratia phage X20]	1e-24	ARW58215
g260	151636←151974	35	112	104	Hypothetical protein [Serratia phage CHI14]	2e-16	ARW57663
					Hypothetical protein [Serratia phage X20]	3e-16	ARW58216
g261	151977←152207	32	76	71	Hypothetical protein CGG41_228 [Salmonella phage vB_SnwM_CGG4-1]	1e-24	YP_009286594
					Conserved hypothetical protein [Edwardsiella phage PEi20]	9e-15	YP_009190422
g262	152305←153210	40	301	279	Single-stranded DNA binding protein [Serratia phage X20]	3e-170	ARW58220
					Single-stranded DNA binding protein [Serratia phage CHI14]	1e-168	ARW57667
g263	153296←153949	33	217	201	Loader of DNA helicase [Serratia phage X20]	7e-146	ARW58221
					Loader of DNA helicase [Serratia phage CHI14]	9e-143	ARW57668
g264	153946←154254	38	102	95	Late promoter transcription accessory protein [Serratia phage CHI14]	4e-61	ARW57670
					Late promoter transcription accessory protein [Serratia phage X20]	7e-60	ARW58222
g265	154229←154519	35	96	89	dsDNA binding protein, late transcription [Serratia phage CHI14]	8e-47	ARW57671
					dsDNA binding protein [Serratia phage X20]	1e-46	ARW58223
g266	154528←155469	35	313	290	Ribonuclease [Serratia phage CHI14]	0.0	ARW57672
					Ribonuclease [Serratia phage X20]	0.0	ARW58224
g267	155570←159346	39	1258	1165	Long tail fiber proximal subunit [Serratia phage CHI14]	0.0	ARW57673
					Long tail fiber proximal subunit [Serratia phage CBH8]	0.0	ARW57948
g268	159355→160503	38	382	355	Hinge connector of long tail fiber, proximal connector [Serratia phage CBH8]	0.0	ARW57949
					Hinge connector of long tail fiber, proximal connector [Serratia phage CHI14]	0.0	ARW57674
g269	160565→161212	37	215	199	Tail connector protein [Yersinia phage vB_YenM_TG1]	1e-137	YP_009200510
					Tail fiber protein p36 (protein Gp36) [Yersinia phage phiR1-RT]	1e-134	YP_007236082
g270	161245→164178	37	977	908	Long tail fiber distal subunit [Yersinia phage vB_YenM_TG1]	5e-101	YP_009200511
					gp37 large distal tail fiber subunit [Enterobacteria phage RB16]	3e-44	YP_003858559
g271	164210→164710	37	166	155	Distal long tail fiber assembly catalyst [Edwardsiella phage PEi26]	3e-44	BAQ23225
					Distal long tail fiber assembly catalyst [Serratia phage CHI14]	2e-42	ARW57677
g272	164853→165371	40	172	161	Distal long tail fiber assembly catalyst [Cronobacter phage vB_CsaM_leB]	6e-61	AOG16396
					Distal long tail fiber assembly catalyst [Cronobacter phage vB_CsaM_leN]	8e-61	AOG16682
g273	165404→166063	36	219	204	Holin lysis mediator [Serratia phage CHI14]	7e-139	ARW57679
					Holin lysis mediator [Serratia phage X20]	1e-137	ARW58231
g274	166090→166362	34	90	83	Anti-sigma 70 protein [Serratia phage CHI14]	7e-52	ARW57680
					Anti-sigma 70 protein [Serratia phage X20]	9e-52	ARW58232
g275	166414←166752	37	112	104	Hypothetical protein [Serratia phage X20]	1e-39	ARW58233
					Hypothetical protein [Serratia phage CHI14]	4e-38	ARW57681
g276	166749←167027	33	92	102	Hypothetical protein		
g277	167024←167290	35	88	82	Hypothetical protein S13_134 [Cronobacter phage S13]	0.014	YP_009196518
g278	167280←167405	25	41	39	Hypothetical protein [Serratia phage CHI14]	4e-15	ARW57682
					Hypothetical protein EpJS98_gp248 [Escherichia phage JS98]	1e-07	YP_001595377
g279	167398←167685	34	95	88	Hypothetical protein CPT_Merlin287 [Citrobacter phage Merlin]	2e-35	YP_009204001
					Hypothetical protein CPT_Moon281 [Citrobacter phage Moon]	4e-34	YP_009146714

g280	167685←168026	38	113	105	Hypothetical protein RB32ORF251c [Enterobacteria phage RB32]	1e-48	YP_803193
					Hypothetical protein HY03_0227 [Escherichia phage HY03]	2e-48	YP_009284203
g281	168023←168382	35	119	110	Hypothetical protein [Serratia phage X20]	6e-54	ARW58237
					Hypothetical protein [Serratia phage CHI14]	3e-53	ARW57686
g282	168397←169032	38	211	195	Transcriptional regulator of middle promoters [Serratia phage X20]	3e-137	ARW58238
					Transcriptional regulator of middle promoters [Serratia phage CHI14]	6e-137	ARW57687
g283	169147←169410	32	87	81	Hypothetical protein		
g284	169482←170135	38	217	201	Hypothetical protein [Serratia phage X20]	2e-99	ARW58239
					Hypothetical protein [Serratia phage CHI14]	9e-98	ARW57688
g285	170307←171680	37	457	423	Topoisomerase II medium subunit [Serratia phage CHI14]	0.0	ARW57690
					DNA topoisomerase II, medium subunit [Serratia phage X20]	0.0	ARW58240
g286	171719←171844	37	42	40	Conserved hypothetical protein [Edwardsiella phage PEi20]	1e-06	YP_009190448
					Conserved hypothetical protein [Edwardsiella phage PEi26]	2e-06	BAQ23240
g287	171816←171974	35	52	49	Hypothetical protein		
g288	171977←172114	38	45	42	Acridine resistance [Shigella phage SHSML-52-1]	2.6	YP_009289025
					Hypothetical protein ECTP3_00414 [Escherichia coli O157 typing phage 3]	3.1	AKE45292
g289	172108←172251	43	47	44	Hypothetical protein [Edwardsiella phage PEi20]	4e-07	YP_009190449
					Hypothetical protein [Edwardsiella phage PEi26]	6e-07	BAQ23241
g290	172286←172747	40	153	142	Nucleoid disruption protein [Serratia phage CHI14]	9e-87	ARW57692
					Nucleoid disruption protein [Serratia phage X20]	3e-86	ARW58244
g291	172890←173078	39	62	58	Hypothetical protein [Cronobacter phage Pet-CM3-4]	5e-07	SCN45968
					Conserved hypothetical protein [Edwardsiella phage PEi20]	9e-05	YP_009190451
g292	173240←173812	43	190	176	Endonuclease IV [Serratia phage CHI14]	1e-104	ARW57695
					Endonuclease IV [Serratia phage X20]	8e-104	ARW58247
g293	173802←174131	34	113	104	Hypothetical protein PG7_293 [Enterobacter phage PG7]	1e-15	YP_009005557
					Hypothetical protein VR26_289 [Escherichia phage vB_EcoM_VR26]	8e-15	YP_009214126
g294	174174←174368	36	64	60	Hypothetical protein [Serratia phage X20]	5e-23	ARW58249
					Hypothetical protein [Serratia phage CHI14]	8e-23	ARW57697
g295	174412←175413	38	333	309	Protector from prophage-induced early lysis [Serratia phage CHI14]	0.0	ARW57698
					Protector from prophage-induced early lysis [Serratia phage X20]	0.0	ARW58250

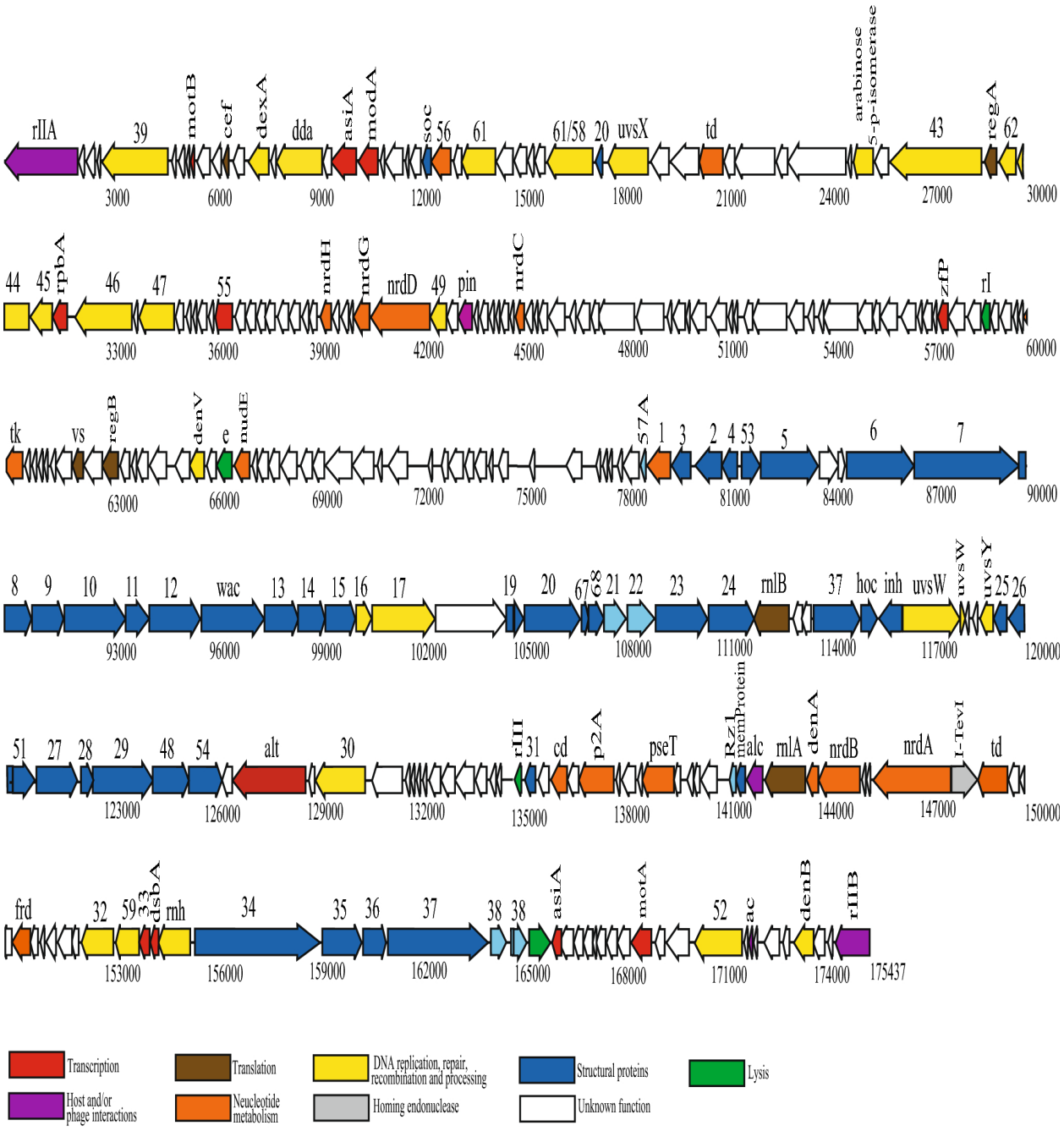


Fig. 3.1. Genome map of Escherichia phage EcS1. Colored arrows indicate the directions and categories of the genes. The gene names and numbers above individual genes are according to the T4 standard [Miller et al., 2003b] (GenBank accession no. AF15810 and NCBI accession no. NC_000866).

Functions numbering of EcS1 gene products according to (Miller et al., 2003b):

Gene	Function of gene product
rIIA	Membrane-associated protein; affect host membrane ATPase
39	DNA topoisomerase subunit; DNA-dependent ATPase; membrane-associated protein
motB	FmdB family regulator protein
cef	Processing of T4 tRNAs
dexA	Exonuclease A
dda	DNA helicase; DNA-dependent ATPase
asiA	Protein that binds to host σ70, inhibits interaction with −35 regions of classical promoters, and facilitates interaction with T4 MotA protein
modA	Adenylribosylating enzyme
soc	Small outer capsid protein
56	dCTPase; dUTPase; dCDPase; dUDPase
61	Primase; requires interaction with gp41 helicase for priming at unique sequence
61 = 58	Primase; requires interaction with gp41 helicase for priming at unique sequence
20	Portal vertex protein of the head
uvsX	RecA-like recombination protein; DNA-ATPase
td	Putative thymidelate synthase
43	DNA polymerase; 3′-to-5′ exonuclease
regA	Translational repressor of several early genes
62	Clamp-loader subunit
44	Clamp-loader subunit
45	Processivity enhancing sliding clamp of DNA polymerase; and mobile enhancer of late promoters
rpbA	RNAP-binding protein
46	Recombination protein and nuclease subunit
47	Recombination protein and nuclease subunit
55	σ factor recognizing late T4 promoters
nrdH	Anaerobic nucleotide reductase subunit
nrdG	Anaerobic nucleotide reductase subunit
nrdD	Anaerobic ribonucleotide reductase subunit
49	Recombination endonuclease VII
pin	Inhibitor of host Lon protease
nrdC	Thioredoxin, glutaredoxin
zfp	Zinc finger protein
rI	Membrane protein
tk	Thymidine kinase
vs	Modifier of valyl-tRNA synthetase
regB	Site-specific RNase
denV	Endonuclease V; *N*-glycosidase
e	Soluble lysozyme; endolysin
nudE	Nudix hydrolase
57A	Chaperone of long and short tail fiber assembly

1	dNMP kinase
3	Head-proximal tip of tail tube
2	Protein protecting DNA ends
4	Head completion protein
53	Base plate wedge component
5	Base plate lysozyme; hub component
6	Base plate wedge component
7	Base plate wedge component
8	Base plate wedge component
9	Base plate wedge component, tail fiber socket, trigger for tail sheath contraction
10	Base plate wedge component, tail pin
11	Base plate wedge component, tail pin, interface with short tail fibers, gp12
12	Short tail fibers
wac	Whiskers, facilitate long tail fiber attachment
13	Head completion
14	Head completion
15	Proximal tail sheath stabilizer, connector to gp3 and/or gp19
16	Terminase subunit, binds dsDNA
17	Terminase subunit with nuclease and ATPase activity; binds single-stranded DNA, gp16 and gp20
19	Tail tube monomer
20	Portal vertex protein of the head
67	Prohead core protein; precursor to internal peptides
68	Prohead core protein
21	Prohead core protein and protease
22	Prohead core protein; precursor to internal peptides
23	Precursor of major head subunit
24	Precursor of head vertex subunit
rnlB	Second RNA ligase
37`	Phage tail fibers
hoc	Large outer capsid protein
inh	Minor capsid protein; inhibitor of gp21 protease
uvsW	RNA-DNA- and DNA-helicase; DNA-dependent ATPase
uvsY	ssDNA binding, recombination and repair protein; helper of UvsX, inhibitor of endoVII
25	Base plate wedge subunit
26	Base plate hub subunit
51	Base plate hub assembly catalyst
27	Base plate hub subunit
28	Base plate distal hub subunit
29	Base plate hub; determinant of tail length
48	Base plate; tail tube associated
54	Base plate; tail tube initiator
alt	Adenosylribosyltransferase
30	DNA ligase
rIII	Unknown

31	Cochaperonin for Groel
cd	dCMP deaminase
p2A	Phospho-2-dehydro-deoxyheptonate aldolase
Pset	Deoxy ribonucleotide 3`phosphatase, 5` polynucleotide kinase
Rz1	Outer membrane lipoprotein
alc	RNA polymerase and DNA binding protein
rnlA	RNA ligase, catalyst of tail fiber attachment
denA	Endonuclease II
nrdB	Ribonucleotide reductase B subunit
nrdA	Ribonucleotide reductase alfa subunit
I-TevI	Intron-homing endonuclease
Td	Thymidylate synthetase
Frd	Dihydrofolate reductase
32	ssDNA-binding protein
59	Loader of gene 41 DNA helicase
33	Late promoter transcription accessory protein
dsbA	dsDNA binding protein
Rnh	RNase H
34	Proximal tail fiber subunit
35	Tail fiber hinge
36	Small distal tail fiber subunit
37	Large distal tail fiber subunit
38	Assembly catalyst of distal tail fiber
motA	Activator of middle promoters
52	DNA topoisomerase subunit
ac	Membrane protein
denB	Endonuclease IV, single strand specific endonuclease
rIIB	Membrane-associated protein

Table 3.3. 53 gene products with highest identity to gene products of phages other than Serratia phages.

Gene product	Length of protein in amino acids	Amino acid sequence identity/similarity to best homologs (% amino acid identity)	Query coverage %	Identity %	BLAST score (E-Value)	Accession no
Gp006	74	Hypothetical protein PM2_004 [Pectobacterium bacteriophage PM2]	71	38	3e-04	YP_009211425
		LEAFY-like protein 1 [Fragaria x ananassa]	58	37	4.3	AHI62726
Gp008	38	Hypothetical protein				
Gp012	65	Modifier of suppressor tRNAs [Citrobacter phage Merlin]	96	73	1e-24	YP_009203727
		Modifier of suppressor tRNA [Citrobacter phage Moon]	96	70	1e-21	YP_009146446
Gp021	166	Hypothetical protein [Klebsiella phage KPV15]	85	29	1e-12	APD20401
		Hypothetical protein kpv477_021 [Klebsiella phage vB_KpnM_KpV477]	85	28	8e-12	YP_009288697
Gp029	143	Hypothetical protein				
Gp030	80	Conserved hypothetical protein [Edwardsiella phage PEi20]	86	42	1e-09	YP_009190191
		gp61.4 hypothetical protein [Escherichia phage vB_EcoM_VR7]	90	35	4e-06	YP_004063718
Gp052	71	Hypothetical protein QL01_57 [Escherichia phage QL01]	100	51	1e-20	YP_009202788
		46.1 gene product [Enterobacteria phage IME08]	100	50	3e-20	
Gp061	85	Hypothetical protein				
Gp063	82	Hypothetical protein CPT_Merlin69 [Citrobacter phage Merlin]	100	45	3e-15	YP_009203783
		Conserved hypothetical protein [Edwardsiella phage PEi20]	100	43	3e-13	YP_009190228
Gp064	102	Hypothetical protein PM2_071 [Pectobacterium bacteriophage PM2]	94	41	2e-17	YP_009211492
		Hypothetical protein CC31p062 [Enterobacter phage CC31]	90	40	2e-14	YP_004009920
Gp066	77	Hypothetical protein JS09_0111 [Escherichia phage vB_EcoM_JS09]	90	40	8e-08	YP_009037434
		Hypothetical protein APCEc01_231 [Escherichia phage APCEc01]	88	35	8e-06	YP_009225191
Gp067	86	Hypothetical protein JS09_096 [Escherichia phage vB_EcoM_JS09]	90	40	4e-13	YP_009037419
		Hypothetical protein [Escherichia phage ST0]	88	35	4e-13	ASD53741
Gp085	52	Hypothetical protein CGG41_076 [Salmonella phage vB_SnwM_CGG4-1]	78	63	1e-08	YP_009286442
		Secreted glycosyl hydrolase [Salmonella phage STP4-a]	78	63	2e-08	YP_009126288
Gp086	95	Conserved hypothetical protein [Edwardsiella phage PEi20]	96	45	2e-17	YP_009190244
		Hypothetical protein PG7_077 [Enterobacter phage PG7]	96	41	9e-17	YP_009005341
Gp089	73	NrdC.7 conserved hypothetical, predicted membrane protein [Enterobacteria phage RB69]	88	45	6e-27	NP_861784.1
		hypothetical protein vBEcoMNBG1_088 [Escherichia phage vB_EcoM_NBG1]	88	45	7e-27	AWM11822.1
Gp093	68	Hypothetical protein CGG41_084 [Salmonella phage vB_SnwM_CGG4-1]	86	47	3e-08	YP_009286450
		Hypothetical protein [Salmonella phage vB_SenMS16]	92	46	6e-08	YP_007501127.1
Gp100	97	Hypothetical protein				

Gp112	60	Hypothetical protein				
Gp119	40	Hypothetical protein				
Gp122	70	Hypothetical protein CPT_Merlin118 [Citrobacter phage Merlin]	97	76	8e-32	YP_009203832 YP_009286469
		Hypothetical protein CGG41_102 [Salmonella phage vB_SnwM_CGG4-1]	95	76	1e-31	
Gp124	59	Hypothetical protein				
Gp125	48	Hypothetical protein				
Gp126	169	Hypothetical protein phAPEC8_0048 [Escherichia phage phAPEC8]	100	69	2e-29	YP_007348416
		Hypothetical protein ESCO5_00044 [Escherichia phage ESCO5]	100	67	8e-29	AOT23420
Gp133	112	Hypothetical protein CC31p127 [Enterobacter phage CC31]	100	56	1e-38	YP_004009985
		Hypothetical protein PKO111_044 [Klebsiella phage PKO111]	100	54	6e-23	YP_009289445
Gp135	149	Hypothetical protein				
Gp141	101	Conserved hypothetical protein [Edwardsiella phage PEi26]	96	53	5e-30	BAQ23082
		Hypothetical protein CC31p135 [Enterobacter phage CC31]	98	50	1e-29	YP_004009993
Gp144	112	Hypothetical protein				
Gp148	66	Hypothetical protein BN80_132 [Yersinia phage phiR1-RT]	84	93	7e-28	YP_007235965
		Hypothetical protein [Salmonella phage vB_SenMS16]	100	59	2e-19	YP_007501160
Gp150	56	DUF1837 domain-containing protein [Lactococcus garvieae]	67	34	6.8	WP_017371069
Gp151	74	Hypothetical protein PG7_148 [Enterobacter phage PG7]	100	78	7e-36	YP_009005412
		Hypothetical protein [Cronobacter phage Pet-CM3-4]	87	75	5e-28	SCN45824
Gp154	112	Hypothetical protein kpv477_146 [Klebsiella phage vB_KpnM_KpV477]	75	32	0.034	YP_009288822
		Hypothetical protein RB51ORF141 [Enterobacteria phage RB51]	47	34	0.18	YP_002854094.1
Gp155	56	Hypothetical protein kpv477_140 [Klebsiella phage vB_KpnM_KpV477]	98	55	3e-06	YP_009288816
		Putative vertex head subunit [Klebsiella phage PKO111]	98	55	1e-05	YP_009289437
Gp158	161	Hypothetical protein kpv477_148 [Klebsiella phage vB_KpnM_KpV477]	92	54	2e-51	YP_009288824
		PHG31p119nc [Aeromonas virus 31]	95	45	2e-39	YP_238848
Gp162	70	Hypothetical protein WG01_147 [Escherichia phage WG01]	97	74	1e-23	YP_009323346
		Hypothetical protein CPT_Merlin166 [Citrobacter phage Merlin]	92	48	9e-12	YP_009203880
Gp194	426	Head vertex protein [Edwardsiella phage PEi20]	99	70	0.0	YP_009190352
		Head vertex protein [Enterobacter phage PG7]	99	70	0.0	YP_009005456
Gp198	445	Phage tail fibers [Klebsiella phage PMBT1]	100	49	3e-86	SCO64804
		Whisker protein [Enterobacter phage phiEap-3]	100	48	2e-83	ALA45334
Gp199	165	Hoc large outer capsid protein [Enterobacteria phage RB16]	98	57	7e-48	YP_003858524
		Protein of unknown function [Enterobacteria phage RB43]	98	56	7e-48	CCL97698
Gp204	71	UvsY.-1 hypothetical protein [Enterobacteria phage RB51]	87	77	1e-19	YP_002854140
		Hypothetical protein [Clostridioides difficile]	87	77	1e-19	WP_074146452

Gp225	176	Conserved hypothetical protein [Edwardsiella phage PEi20]	52	43	4e-14	YP_009190382
		Conserved hypothetical protein [Edwardsiella phage PEi26]	52	43	4e-14	BAQ23176
Gp227	64	Hypothetical protein CC31p211 [Enterobacter phage CC31]	93	34	4e-05	YP_004010069
		Hypothetical protein [Cronobacter phage Pet-CM3-4]	93	34	1e-04	SCN45904
Gp237	67	Hypothetical protein SP18_gp234 [Shigella phage SP18]	100	41	9e-08	YP_003934857
		Hypothetical protein VR5_222 [Enterobacteria phage vB_EcoM_VR5]	94	43	1e-07	YP_009205916
Gp240	60	Hypothetical protein fHeYen901_242 [Yersinia phage fHe-Yen9-01]			0.007	ARB06015
Gp249	57	Hypothetical protein [Isosphaera pallida]	49	50	7.7	WP_013566058
Gp252	252	I-TevI homing endonuclease [Yersinia phage phiR1-RT]	96	65	6e-98	YP_007236067
		Endonuclease I-Tevi [Klebsiella phage PKO111]	93	57	3e-86	YP_009289561
Gp257	78	Hypothetical protein [Edwardsiella phage PEi26]	37	31	0.40	BAQ23206
Gp270	977	Long tail fiber distal subunit [Yersinia phage vB_YenM_TG1]	55	49	5e-101	YP_009200511
		gp37 large distal tail fiber subunit [Enterobacteria phage RB16]	62	29	3e-44	YP_003858559
Gp276	92	Hypothetical protein				
Gp277	88	Hypothetical protein S13_134 [Cronobacter phage S13]	53	45	0.014	YP_009196518
Gp283	87	Hypothetical protein				
Gp286	42	Conserved hypothetical protein [Edwardsiella phage PEi20]			1e-06	YP_009190448
		Conserved hypothetical protein [Edwardsiella phage PEi26]			2e-06	BAQ23240
Gp287	52	Hypothetical protein				
Gp288	45	Acridine resistance [Shigella phage SHSML-52-1]	95	45	2.6	YP_009289025
		Hypothetical protein ECTP3_00414 [Escherichia coli O157 typing phage 3]	95	45	3.1	AKE45292
Gp289	47	Hypothetical protein [Edwardsiella phage PEi20]	97	52	4e-07	YP_009190449
		Hypothetical protein [Edwardsiella phage PEi26]	89	57	6e-07	BAQ23241

Table 3.4: tRNAs, their positions, direction, recognized codons and products:

tRNA	Position in EcS1 genome	Codon	Product
1	Complement(72123..72195)	ACA	Thr
2	Complement(72203..72284)	TTA	Leu
3	Complement(72292..72364)	AGA	Arg
4	Complement(72541..72624)	TGC	Cys
5	Complement(74805..74876)	GGA	Gly
6	Complement(74881..74951)	TGG	Trp
7	Complement(75106..75177)	TTC	Phe
8	Complement(75266..75338)	ATC	Lle
9	Complement(75348..75421)	CCA	Pro
10	Complement(75697..75787)	TCA	Ser
11	Complement(75795..75867)	CAC	His
12	Complement(75971..76043)	CAA	Gln
13	Complement(76049..76121)	ATG	Met
14	Complement(76252..76325)	AAC	Asn
15	Complement(76479..76552)	AAA	Lys
16	Complement(77034..77106)	GAC	Asp
17	Complement(77116..77187)	GAA	Glu
18	Complement(77199..77284)	TAC	Tyr

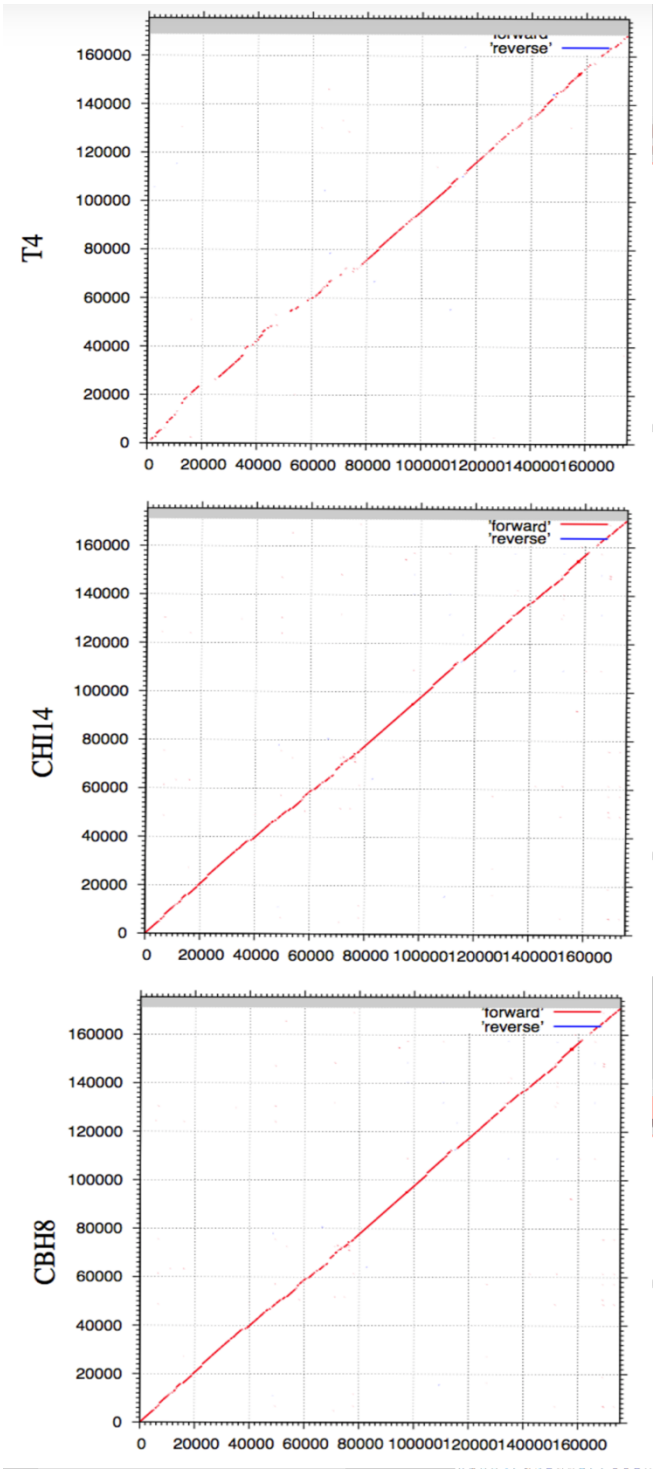


Fig. 3.2. Dot plot comparison of the genomic nucleotide sequences between EcS1 and *Serratia* myoviruses, including CHI14 (accession no. MF036690) and CBH8 (MF036691). The T4 sequence (AF15810) is also included. Nucleotide positions are shown along the genomic sequence. Dot plot analyses were performed using the MAFFT version 7 program (http://mafft.cbrc.jp/alignment/server/) with a threshold score 39 (E=8.4e-11).

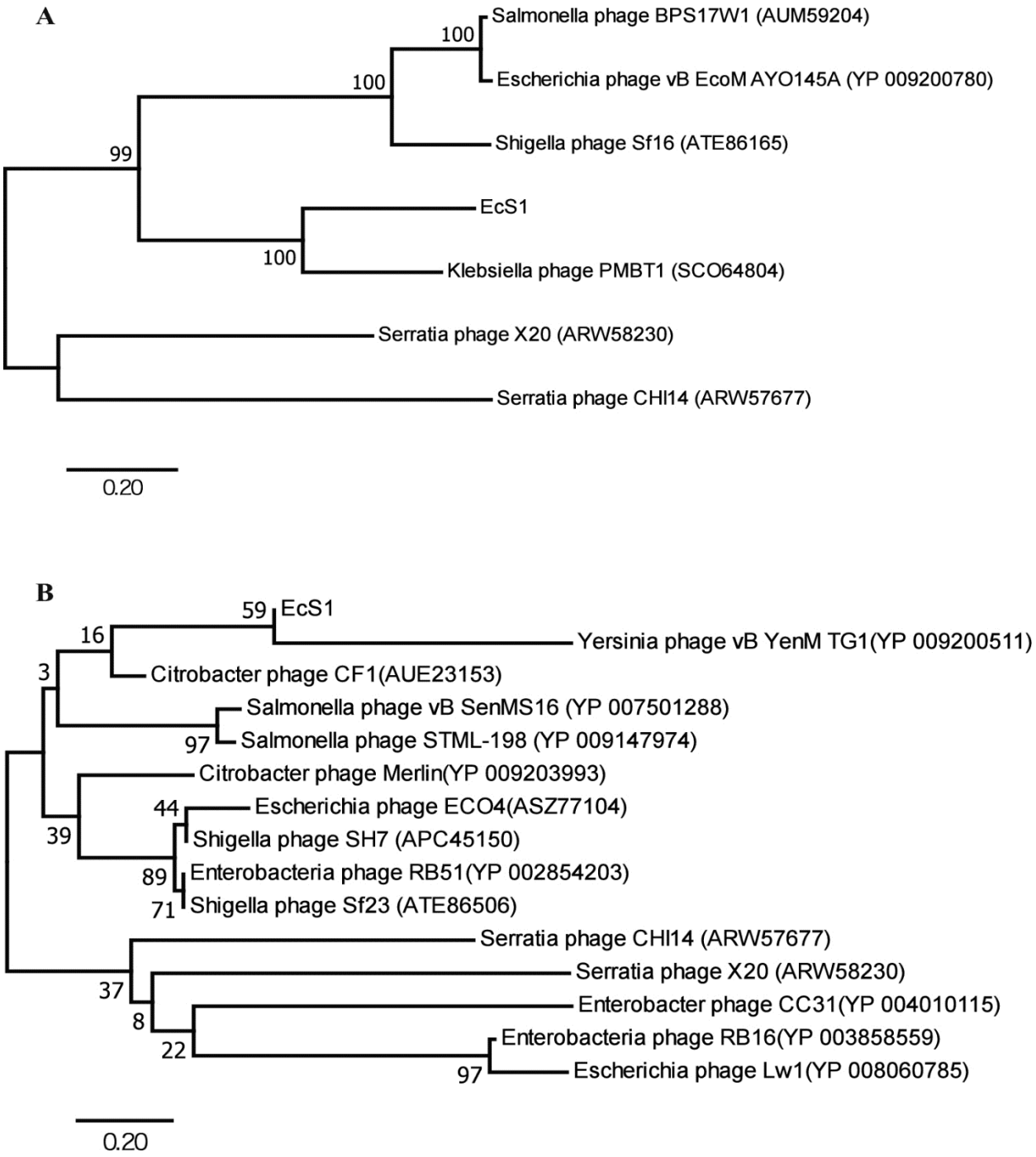


Fig. 3.3. Phylogenetic trees for the amino acid sequences of tail fiber proteins ORF198 (445 aa) (A) and ORF270 (977 aa) (B) showing the relationships between EcS1 and other known myoviruses. The values at the nodes indicate the bootstrap support scores as calculated using 1000 replicates. Numbers in parentheses are the accession numbers of the amino acid sequence in the databases.

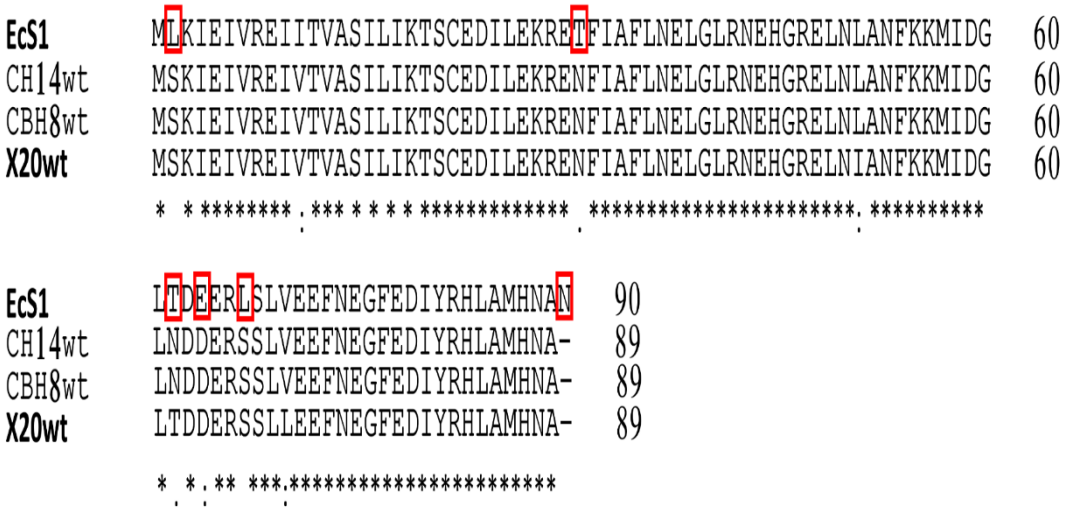


Fig. 3.4: Alignment of the primary sequence of conserved asiA gene in phage EcS1 and other wild types of Serratia phages CH14, CBH8 and X20 reveal presence of mutations in EcS1 phage genome. Identical residues are indicated with asterisks and similar ones with dots.

3.5. Discussion:

Whole genome sequencing of phage EcS1 revealed its advantageous features for biocontrol application, including potential high lytic activites and safety for animals. As for lytic activities, all genes involved in the host lysis by holin-endolysin pathways were found in the phage EcS1 genome. Gp138 and gp243 correspond to lysozyme murein hydrolase (lysin) and outer membrane lipoprotein (holin) respectively and might play major roles to lyse the host cell at the end of the phage life cycle, releasing progeny phage virions (Catalao *et al*., 2013). Moreover, gp244 correspond to I-spanin protein which is essential for the final step in host lysis by crippling the outer membrane after lysin-holin system have permeabilized the inner membrane (Summer et al., 2007). As for safety of EcS1, any lysogenic genes were not detected and there were no known harmful genes identified in the genome of phage EcS1.

In addition to this, complete genetic characterization confirmed the novelty of phage EcS1. Although from the whole genomic sequence, EcS1 was more similar to *Serratia* phage than *E. coli, Shigella* and/or other *Enterobacteriacae* phages, the tail fibers proteins were completely different between EcS1 and Serratia phages. The tail fiber proteins play a key role in the first step of host recognition and are supposed to evolve rapidly more than other genes in most phages to gain novel abilities to infect variable hosts and to enter different ecological niches (Casjens, 2005). The phylogenetic trees of (Gp198) and (Gp270) beside the host range data indicate that EcS1 is rather related to *Enterobacteriacae* phages. It is not surprising because Hamdi et al. (2017) revealed that the major differences among phages infecting *Enterobacteriaceae* are in their tail fiber proteins. Moreover, Su *et al*. (2017) indicated the role of tail fibers genes in differentiation between phages. The extended host range of EcS1 may be related to the acquisition of novel tail fiber genes such as Gp198 that is located outside of the cluster for other tail fiber genes. Phage EcS1 could not show lytic effects on the tested *Serratia* strain (*Serratia marsensens* D601). Because EcS1 Gp274 corresponding to anti-sigma 70 protein (asiA) have several changes in amino acid sequence compared to Serratia phage asiA (Fig. 3.4), this may explain resistance of Serratia cells to EcS1. One mechanism for phage to overcome abortive infection mediated by a type III toxin-antitoxin system of the host is mutations in phage *asiA* gene demonstrated for Serratia phages (Chen *et al*., 2017). All these data, in addition to, phylogenetic analyses of tail fibers proteins (Fig.3.3) indicate that phage EcS1 is a novel lytic member of the genus *T4virus* that can infect *Shigella* and *Salmonella* strains in addition to its *E. coli* host.

The example of EcS1 suggests that there are many similar phages but with different host ranges, which can be isolated in the same way as EcS1 with non-pathogenic *E. coli* as a host. Those phages will serve as promising candidates for biocontrol agents against various pathogenic bacteria (Santos *et al*., 2010).

CHAPTER IV

GENERAL DISCUSSION AND CONCLUSION

Bacteriophages are the most attractive alternative to control bacterial pathogens since they are the most abundant microorganisms on the planet and have no any side effect to eukaryotes especially to human beings. Already, they have great applications for use as biocontrol agents in many fields of life such as phage therapy (Lin *et al*., 2017), as preservatives in food products (Pulido *et al*., 2015) and in agricultural system (Addy *et al*., 2012; Ahmad *et al*., 2014). The exploitation of bacteriophages has consequently become an interesting tool to combat the emergence of multi-drug resistant bacteria. Easiness of application and their low cost made phages as constant part of integrated disease management (Obradovic *et al*., 2005). As host range is the trojan horse for phage therapy; the specificity of phage towards definite host is one of its multiple advantages due to harmless effect to non-pathogenic bacteria (beneficial microbes). However, this same specificity limits the ability of the phage use to a strict set of potential pathogens (Kim *et al*., 2010b; Nilsson, 2014; Mapes *et al*., 2016). Ongoing efforts have been developed to overcome phage selectivity. For example: mixing multiple phage species into a cocktail to treat several different bacteria (Gill and Hyman 2010; Lu and Koeris, 2011; Chan and Abedon, 2012). Unfortunately, phage cocktails are laborious, wasting time and costs (Loc-Carrillo and Abedon, 2011). It is also well known that even small number of phages with broad host range could be more useful than large number of phages with narrow host range (Ross *et al*., 2016; Kutter *et al*., 2010). Moreover, phage-host specificity is almost associated with bacterial phage resistance which can occur at any stage of phage infection (Hyman and Abedon, 2010). Jumbo phages are promising solution for bacteriophages drawbacks.

Jumbo phages are large and tailed phages with genome size more than 200 kbp. The large genome size of jumbo phages enables them to contain many genes responsible for their genome replication and nucleotide metabolism those do not exist in small-genome phages (Mesyanzhinov *et al*., 2002; Kiljunen, 2005; Thomas 2007). Investigating the expression of phage genes by transcriptomic analysis during jumbo phage infection showed that they may be dependent only on their own RNAPs and independent from the host RNAPs (Ceyssens *et al*., 2014) Moreover, jumbo phages have more cell-wall lysis proteins, such as endolysin, glycoside hydrolase, and chitinase. Many of these proteins were helpful for facilitating phage infection ability (Gill *et al*., 2012; Yuan and Gao, 2016). Furthermore, almost all jumbo phages have tRNA in their genomes as in the case

of phage G, Yersinia phage ΦR1-37, and others (Kiljunen *et al*., 2005). The tRNAs in jumbo phage genomes enhance the translation efficiency of phage-specific genes. These additional proteins encoded by jumbo phages may subrogate for the function of the host proteins that are essential for the life cycle of the smaller-genome phages (Kiljunen *et al.*, 2005). All of these notable genomic features reduce the dependence of jumbo phages on their bacterial hosts (O'Donnell *et al*., 2013) which in turn could broaden the phage host range (Yuan and Gao 2017).

In spite of the unrivaled characteristics of jumbo phages, only 93 of them have been isolated since the discovery of bacteriophages within the last 100 years (Yuan and Gao 2017). The possible causes for the scarce isolation of large phages are using of the classical procedures for bacteriophage screening, propagation and classical processing of environmental samples (Serwer *et al*., 2007). Nearly all jumbo phages have been isolated using the fairly standard procedures for phage isolation. For example, the largest phage G was discovered by accident during preparation of another phage (Serwer *et al*., 2007). That is why developing of specific smart method to screen for jumbo phages is urgent need.

The recently used classical method for phage isolation is fairly simple process and could be accomplished by researchers in a very similar manner to that used by the earliest phage biologists (d'Hérelle, and Smith 1926). Briefly, the basic method is to primarily obtain an environmental sample that can be raw fecal matter (Jensen *et al*., 2015), sewage samples (Lin *et al*., 2010), water samples (Uchiyama *et al*., 2008), soil samples (Anand *et al*., 2015) and any ecological materials likely to contain or have been in contact with the targeted host bacteria, followed by adding buffer or broth media to the samples and then they are filtered to remove the bacteria and any other solid material. Then the filtrate is supposed to either spot or plaque testing for screening the existing phages. This procedure was approved as standard method of screening for coliphage in water samples according to International Organization for Standardization 2000.

The main goal of this PhD dissertation was to enhance utilization of large phages to control wide range of pathogenic bacteria, followed by comprehensive investigation for the virological and genetic characteristics of successfully isolated phages.

I started my work by developing systemic, smart and specific method for detection and isolation of large phages for use in biocontrol of a wide-range of pathogenic bacteria by making various modifications to the classical methods. My previous results and those of others inspired me for the following modifications. (i) Initial selection by differential centrifugation of samples to

precipitate relatively large particles. Centrifugation at 15000 x g for 1 hour three times was enough to separate large phage particles. This result was obtained after several trials of different centrifugation conditions for ϕRSL1 previously isolated in our laboratory (Yamada et al., 2010). The established differential centrifugation was also close to that used by Attai *et al*. (2018) to isolate jumbo phage that infect *Agrobacterium tumefaciens*. (ii) Get rid of contaminating bacteria by killing with chloroform without filtration. One of the major causes to lose jumbo phages is using filters to remove bacteria. As big-sized phages are unable to pass through the pores of the filter (Yuan and Gao 2017). (iii) Plaque assays using low concentrations of top agar (0.35%) and picking smaller plaques because large phages always form very small plaques. In the top agar of around 0.7% usually used in the standard method, large phage particles cannot form visible plaques because of limited diffusion. The high dependence of plaque size on the concentration of a supporting agar gel was clearly clarified by Serwer *et al*. (2007), lowering the top agar concentration means large phage particles can diffuse more easily. (iv) Plaque assays were conducted at lower temperatures around 25 °C to delay host growth consequently resulting in increased size of plaques related to large phages. As phage replication is known to depend mainly on the physiological state of the host (Chibani-Chennoufi *et al*., 2004; Denes and Wiedmann, 2014). Vigorous growth of the host strain at higher temperatures (e.g., 37°C for *E. coli*) leads to rapid spread of cells over the plate surface and soon induce inactive states of the cells. With *E. coli* strains as the host, my preliminary tests of plaque formation by lambda phage gave the largest plaques at 23°C. This result is in the line with the previous conclusion of Fister *et al.* (2016); that phage P100 replication was dependent upon the growth of *L. monocytogenes,* and efficacy was higher when bacterial growth was reduced by certain environmental conditions as low temperature.

In spite of all of these modifications, my developed method may show little drawbacks related to chloroform application. As some of large phages may not be obtained because of their sensitivity to chloroform (Kęsik-Szeloch *et al*., 2013; Roszniowski *et al.*, 2017). However, for such phages, further refinement including extended differential centrifugation without chloroform treatment should be given to the method.

By using my method, I succeeded to isolate eleven large phages (E1-E11). The genome sizes of E3, E4, E8 and E11 were around 450 kbp. While the genomes of phages E1, E2, E5, E6, E7, E9 and E10 were about 200 kbp. Although phages are usually host specific and even strain specific (Weinbauer, 2004; Hyman and Abedon, 2010, Ross *et al*., 2016), the obtained large

phages showed wide-host ranges, infecting different genera including important pathogens such as *Shigella* and *Salmonella*. Especially E9 phage could cover variable host strains ranging over three different genera including *Escherichia*, *Shigella*, and *Salmonella* as the host. Similar case of wide host range was previously identified for Shigella phage SH7 (Hamdi *et al*., 2017), which was reported to infect strains of *E. coli*, *Salmonella* Paratyphi, and *Shigella dysenteriae* in addition to its original host *Shigella flexeri.* Moreover, Phage E4 also showed a unique host range covering three genera such as *Escherichia*, *Salmonella*, and *Serratia*. Phage SFP10 could infect strains of *Salmonella enterica* and *Escherichia coli* (Park *et al*., 2012). The wide host ranges of almost all phages obtained in this work support the idea that jumbo or large phages could have broad host ranges (Miller *et al*., 2003a; Hendrix, 2009; Yoshikawa *et al*., 2018; Yuan and Gao 2017; Sharma *et al*., 2019; Matsui *et al*., 2017). It is thought that large phages with many adaptive genes generally could have wide host range (Bhunchoth *et al*., 2015; Kim *et al*., 2012).

I used laboratory strains of *E. coli* BL21 as the initial host to obtain large phages. Usual pathogenic bacterial strains should be handled in biosafety level 2 or 3 laboratories and require special experimental facilities, skills, and care. The use of a nonpathogenic host in the production process would override the risk of accident administering a pathogen (Bielke *et al*., 2007a; Santos *et al*., 2010). Furthermore, it would greatly simplify the process of purification, increase the safety of phage preparations with a consequent reduction in preparation costs, leading to easy and faster approval of phage products (Bielke *et al*., 2007b). Large phages can maintain their long-lasting effects (Fujiwara *et al*., 2011) and it was also observed that the nonpathogenic host could keep the lytic spectrum (Santos *et al*., 2010). I hope this method will open a new door to enhance jumbo phage isolation and their use in biocontrol of a wide range of pathogenic bacteria.

I developed the new method and succeeded to isolate interesting large and jumbo phages with broad host ranges, then I determined the whole genome sequence of E9 (designated as Escherichia phage EcS1) for whole genome characterization.

Escherichia phage EcS1 is a novel lytic phage infecting a wide range of bacterial strains, including *Shigella sonnei* SH05001, *Shigella boydii* SH00007, *Shigella flexneri* SH00006, and *Salmonella enterica* serovar Enteritidis (SAL 01078). Shotgun sequencing of the EcS1 phage genome was performed at Fasmac System Science Co., Ltd. (Kanagawa, Japan) using an Illumina Miseq System. I determined the full genome sequence not only to clarify the novelty and genetic characteristics of phage EcS1 but also to confirm absence of perilous genes (from biocontrol

applicable point of view) such as integrase genes of the lysogenic type phage, antibiotic resistant genes, genes encoding toxins or any genes for other bacterial virulence factors (Vandenheuvel *et al*., 2015; Principi *et al*., 2019).

The size of completely assembled EcS1 genome was 175,437 bp in length. Primarily, this size as well as the morphological characteristics indicate that this phage belongs to the T4-like phages (Miller et al., 2003b), genus *T4virus* in the family *Myoviridae*. A total of 295 genes were identified along the EcS1 genome. Of these genes, 117 had predicted known functions (structural and functional products) and the other 178 lacked known functions (hypothetical proteins). Interestingly, all genes involved in phage lytic activity (holin-endolysin pathways) were annotated. For example, lysozyme murein hydrolase (lysin) and outer membrane lipoprotein (holin) corresponding to gp138 and gp243 respectively. As phage utilize holin to create a lesion in the cytoplasmic membrane of the host cell wall through which the murein hydrolase passes to gain access to the murein layer (Young, 1992). Moreover, a gene encoding I-spanin (gp244), was also detected in my phage genome. This gene has been recently reported as a third functional class of lysis proteins and required for outer membrane disruption (Cahill and Young, 2019). Presence of complete lytic system indicate the strong lytic activity of phage EcS1 (Berry *et al*. 2008; Berry *et al.* 2013; Catalao *et al*., 2013).

I also detected 18 tRNA sequences within the EcS1 genome this is another evidence to ensure optimum lytic activity of EcS1 (Wilson, 1973; Bailly-Bechet *et al*., 2007). The benefits of having tRNA genes are not only associated with better growth in the host (larger burst size and shorter latency) but also the ability to infect more hosts (broader host range), than less or no-tRNA coding phages (Delesalle *et al*., 2016; Howard-Varona *et al.*, 2017).

Fortunately, there were no genes associated with lysogeny such as integrase gene or any other virulence or genes encoding pathogenicity identified in the genome of EcS1 phage. Together with the strong lytic system this clearly indicates that EcS1 is a prominent candidate to be used safely in biocontrol against wide range of pathogenic bacteria such as phage therapy or food preservative (Merabishvili *et al*., 2014).

Comparative genomic analysis of EcS1 revealed that, though EcS1 showed the highest similarity to Serratia phages (ϕX20, ϕCHI14 and ϕCBH8), some of its predicted proteins, especially the tail fiber proteins (Gp198 and Gp270), were similar to those of Enterobacteriaceae phages. As previously reported, the tail fiber proteins especially receptor binding proteins (RBPs)

of *Enterobacteriaceae* infecting phages showed high genetic diversity due to the gene recombination at their C-terminal regions (Hooton *et al*., 2011; Moreno Switt *et al*., 2013; Li *et al.*, 2016; Peng and Yuan, 2018). Phylogenetic study for Gp198 (predicted phage tail fiber protein) revealed high similarity to the corresponding gene of Klebsiella phage PMBT1 SCO64804 which have lytic effect to the pathogenic *Klebsiella pneumoniae* strain 182 (Koberg et al., 2017). Moreover, it was also closely related to Shigella phage Sf16, Escherichia phage vB_EcoM_AYO145A, and Salmonella phage BPS17W1. On the other hand phylogenetic analysis of Gp270 (long tail fiber distal subunit) showed high similarity to Yersinia phage vB YenM TG1 (Leon-Velarde *et al*., 2016) which was part of a larger cluster that included Citrobacter phage CF1 and Salmonella phages vBSenMs 16 and STML-198. These results may explain why EcS1 infected strains of *Shigella*, *Salmonella*, and *Escherichia* but not strains of *Serratia* despite the high similarity of the EcS1 genome to those of Serratia phages. All of these data demonstrate that phage EcS1 is a novel lytic member of the genus T4virus that can infect *Shigella* and *Salmonella* strains in addition to its *E. coli* host. However, further functional analysis of Gp198 and Gp270 proteins and the interaction between EcS1 and its hosts might give deep explanation to the roles of these two proteins in phage host range determination.

In conclusion, firstly, I believe my newly developed method will open the door to extend jumbo or large phage isolation and their use in biocontrol of a wide range of pathogenic bacteria. Secondly, complete genetic characterization of Phage EcS1 not only explained its divergent host specificity against pathogenic strains of *Shigella, Salmonella* and *E. coli* but also ensures its safety in biocontrol applications.

REFERENCES:

Aarestrup, F. M., Hendriksen, R. S., Lockett, J., Gay, K., Teates, K., McDermott, P. F., White, D. G., Hasman, H., Sørensen G., Bangtrakulnonth, A., Pornreongwong, S., Pulsrikarn, C., Angulo, F. J., and Gerner-Smidt, P. (2007). International spread of multidrug-resistant *Salmonella* Schwarzengrund in food products. Emerging Infectious Diseases. 13: 726–31.

Abbasifar, R., Griffiths, M. W., Sabour, P. M., Ackermann, H. W., Vandersteegen, K., Lavigne, R., Noben, J-P., Villa, A. A., Abbasifar, A., Nash, J. H. E., and Kropinski, A. M. (2014). Supersize me: *Cronobacter sakazakii* phage GAP32. Virology. 460, 138–146.

Abedon, S. T., Kuhl, S. J., Blasdel, B. G., and Kutter, E. M. (2011). Phage treatment of human infections. Bacteriophage. 1(2): 66–85.

Ackermann, H.-W. (2009). Basic phage electron microscopy. Methods in Molecular Biology. 501, pp. 113-126.

Addy, H. S., Askora, A., Kawasaki, T., Fujie, M., and Yamada, T. (2012). Loss of virulence of the phytopathogen *Ralstonia solanacearum* through infection by ϕRSM filamentous Phages. Phytopathology. 102, 469–477.

Addis, M., and Sisay, D. (2015). A review on major food-borne bacterial illnesses. Journal of Tropical Diseases. 3:176.

Ahmad, A. A., Askora, A., Kawasaki, T., Fujie, M., and Yamada, T. (2014). The filamentous phage XacF1 causes loss of virulence in *Xanthomonas axonopodis* pv. *citri*, the causative agent of citrus canker disease. Frontiers in Microbiology. 5: 321.

Anand, T., Vaid, R. K., Bera, B. C., Barua, S., Riyesh, T., Virmani, N., Yadav, N., and Malik, P. (2015). Isolation and characterization of a bacteriophage with broad host range, displaying potential in preventing bovine diarrhoea. Virus Genes. 51, 315–321.

Ansari, S., Nepal, H. P., Gautam, R., Shrestha, S., Neopane, P., Gurung, G., and Chapagain, M. L. (2015). Community acquired multi-drug resistant clinical isolates of *Escherichia coli* in a tertiary care center of Nepal. Antimicrobial Resistance and Infection Control. 4:15.

Attai, H., Boon, M., Phillips, K., Noben, J-P., Lavigne, R., and Brown, P.J.B. (2018). Larger than life: Isolation and genomic characterization of a jumbo phage that Infects the bacterial plant pathogen, *Agrobacterium tumefaciens*. Frontiers in Microbiology. 9:1861.

Bailly-Bechet, M., Vergassola, M., and Rocha, E. (2007). Causes for the intriguing presence of tRNAs in phages. Genome research. 17 (10), 1486–1495.

Bao, H., Zhang, P., Zhang, H., Zhou, Y., Zhang, L., and Wang, R. (2015) Bio-control of *Salmonella Enteritidis* in foods using bacteriophages. Viruses. 7:4836–4853.

Barton, M.D. (2000) Antibiotic use in animals feed and its impact on human health. Nutrition Research Reviews. 13, 279–299.

Berry, J., Summer, E.J., Struck, D.K., and Young, R., (2008). The final step in the phage infection cycle: the Rz and Rz1 lysis proteins link the inner and outer membranes. Molecular Microbiology. 70, 341.

Berry, J. D., Rajaure, M., and Young, R., (2013). Spanin function requires subunit homodimerization through intermolecular disulfide bonds. Molecular Microbiology. 88, 35.

Bhunchoth, A., Phironrit, N., Leksomboon, C., Chatchawankanphanich, O., Kotera, S., Narulita E., Kawasaki, T., Fujie, M., and Yamada, T. (2015). Isolation of *Ralstonia solanacearum*-infecting bacteriophages from tomato fields in Chiang Mai, Thailand, and their experimental use as biocontrol agents. Journal of Applied Microbiology 118, 1023–1033.

Bielke, L. R., Higgins, S. Donoghue, A. M., Donoghue, D. J., and Hargis, B. M. (2007b). *Salmonella* host range of bacteriophages that infect multiple genera. Poultry Science. 86**:** 2536-2540.

Bielke, L. R., Higgins, S. E., Donoghue, A. M., Donoghue, D. J., Hargis, B. M., and Tellez, G. (2007a). Use of wide-host-range bacteriophages to reduce *Salmonella* on poultry products. International Journal of Poultry Science. 6**:** 754-757.

Bridier, A. (2019). Foodborne bacterial pathogens; methods and protocols. Methods in Molecular Biology. Book.

Burkal'tseva, M. V., Krylov, V. N., Pleteneva E. A., Shaburova, O. V., Krylov S. V., Volkart G., Sykilinda, N. N., Kurochkina, L. P., and Mesyanzhinov, V. V. (2002). Phenogenetic characterization of a group of giant φ KZ-like bacteriophages of *Pseudomonas aeruginosa*. Genetika. 38, 1470–1479.

Cahill, J., and Young, R. (2019). Phage Lysis: Multiple Genes for Multiple Barriers. Adv Virus Res;103: 33–70

Campbell A. (2003). The future of bacteriophage biology. Nature Reviews Genetics. 4 (6):471-7.

Carlson, K. (2005). Working with Bacteriophages: common techniques and methodological approaches. In: Kutter E, Sulakvelidze A, et al., editors. Bacteriophages, Biology and Applications. Florida: CRC Press. pp. 437–494.

Casjens, S. R. (2005). Comparative genomics and evolution of the tailed-bacteriophages. Current Opinion in Microbiology. 8, 451–458.

Catalao, M. J., Gil, F., Moniz-Pereira, J., Sao-Jose, C., and Pimentel, M. (2013). Diversity in bacterial lysis systems: bacteriophages show the way. FEMS Microbiology Review. 37, 554–571.

Ceyssens, P. J., Minakhin, L., Van den Bossche, A., Yakunina, M., Klimuk, E., Blasdel, B., De Smet, J., Noben, J. P., Bläsi, U., Severinov, K., and Lavigne, R. (2014). Development of giant bacteriophage ϕKZ is independent of the host transcription apparatus. Journal of Virology. 88 (18): 10501-10.

Chan, B. K., Abedon, S. T., and Loc-Carrillo C. (2013). Phage cocktails and the future of phage therapy. Future Microbiology. 8 769–783.

Chan, B. K., and Abedon, S. T. (2012). Phage therapy pharmacology: phage cocktails. Advances in Applied Microbiology. 78, 1–23.

Chatterjee, S., and Rothenberg, E. (2012). Interaction of bacteriophage λ with its *E. coli* receptor, LamB. Viruses. 4: 3162-3178.

Chen, B., Akusobi, C., Fang, X., and Salmond, G. P. C. (2017): Environmental T4-Family bacteriophages evolve to escape abortive infection via multiple routes in a bacterial host employing "Altruistic Suicide" through type III Toxin-Antitoxin systems. Frontiers in Microbiology. (8). 1006.

Chibani-Chennoufi, S., Bruttin, A., Dillmann, M.-L., and Brüssow, H. (2004). Phage-host interaction: an ecological perspective. Journal of Bacteriology. 186, 3677–3686.

Chowdhury, K., and Sawer, S. (1976). S1 nuclease converts relaxed circular closed duplex DNA to open linear DNA of unit length. FEBS Letters. 68, pp. 68-70.

Cosgrove, S. E. (2006). The relationship between antimicrobial resistance and patient outcomes: mortality, length of hospital stay, and health care costs. Clinical Infectious Diseases. 42 (2): S82–S89.

d'Hérelles, F. (1917). "Sur un microbe invisible antagoniste des bacilles dysentériques" (PDF). Comptes rendus Academie des Sciences Paris. 165: 373–5.

d'Hérelle, F., and Smith, G. H. (1926). The bacteriophage and its behavior [English translation]. Baltimore: The Williams &Wilkins Co.

Delesalle, V. A., Tanke, N. T., Vill, A. C., and Krukonis, G. P. (2016). Testing hypotheses for the presence of tRNA genes in mycobacteriophage genomes. Bacteriophage. 6 (3), e1219441.

Denes, T., and Wiedmann, M. (2014). Environmental responses and phage susceptibility in foodborne pathogens: implications for improving applications in food safety. Current Opinion in Biotechnology. 26, 45–49.

Donelli, G. (1968). Isolamento di un batteriofago di eccezionali dimensioni attivo su *B. megatherium.* Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Nat., 44. pp. 95-97

Doore, S. M., Schrad, J. R., Dean, W. F., Dover, J., and Parent K.N. (2018). Shigella phages isolated during a dysentery outbreak reveal uncommon structures and broad species diversity. Journal of Virology. 92, p. 02117-17.

Drulis-Kawa, Z., Majkowska-Skrobek, G., Maciejewska, B., Delattre, A.-S., and Lavigne, R. (2012). Learning from bacteriophages-advantages and limitations of phage and phage-encoded protein applications. Current Protein and Peptide Science. 13: 699–722.

Eisen, J. A. (2000). Horizontal gene transfer among microbial genomes: new insights from complete genome analysis. Current Opinion in Genetics and Development. 10, 606–611.

Endersen, L., O'Mahony, J., Hill, C., Ross, R. P., McAuliffe, O., and Coffey, A. (2014). Phage therapy in the food Industry. Annual Review of Food Science and Technology. 5: 327–349.

Evans, J. A., Adusei, A., Timmann, C., May, J., Mack, D., Agbenyega, T., Horstmann, R. D., and Frimpong, E. (2004). High mortality of infant bacteraemia clinically indistinguishable from severe malaria. Quarterly Journal of Medicine. 97: 591–7.

Fister, S., Robben, C., Witte, A. K., Schoder, D., Wagner, M., and Rossmanith, P. (2016). Influence of Environmental Factors on Phage-Bacteria Interaction and on the Efficacy and Infectivity of Phage P100. Frontiers in Microbiology. 7, 1152.

Fujiwara, A., Fujisawa, M., Hamasaki, R., Kawasaki, T., Fujie, M., and Yamada, T. (2011). Biocontrol of *Ralstonia solanacearum*by treatment with lytic bacteriophages. Applied Environmental Microbiology. 77. pp. 4155-4162.

Fung, F., Wang, H. S., and Menon, S. (2018). Food safety in the 21st century. Biomedical Journal. 41, 88–95.

Ghosh, C., Sarkar, P., Issa, R., and Haldar, J. (2019). Alternatives to Conventional Antibiotics in the Era of Antimicrobial Resistance. Trends in Microbiology. 27, (4). 323-338.

Gill, J. J., and Hyman, P. (2010). Phage choice, isolation, and preparation for phage therapy. Curr. Pharm. Biotechnol. 11, 2–14.

Gill, J. J., Berry, J. D., Russell, W. K., Lessor, L., Escobar-Garcia, D. A., Hernandez, D., Kane, A., Keene. J., Maddox, M., Martin, R., Mohan, S., Thorn, A. M., Russell, D. H., and Young, R. (2012). The Caulobacter crescentus phage phiCbK: genomics of a canonical phage. BMC Genomics. 13:542.

González-García, V. A., Pulido-Cid, M., Garcia-Doval, C., Bocanegra, R., van Raaij, M. J., Martín-Benito, J., Cuervo, A., and Carrascosa, J. L. (2015). Conformational changes leading to T7 DNA delivery upon interaction with the bacterial receptor. The Journal of Biological Biochemistry. 290(16):10038-44.

Haddad, I.E. Abdallah, N.B. Plante, P. L. Dumaresq, J. Katsarava, R. Labrie, S. Corbeil, J. St-Gelais, D., and Moineau, S. (2014). Improving the safety of *Staphylococcus aureus* polyvalent phages by their production on a *Staphulococcus xylosus* strain. PLoS One, 9, p. e102600.

Haggård-Ljungquist, E., Halling, C. and Calendar, R. (1992). DNA sequences of the tail fiber genes of bacteriophage P2: evidence for horizontal transfer of tail fiber genes among unrelated bacteriophages. Journal of Bacteriology. 174, 1462–1477.

Hambly, E., Tetart, F., Desplats, C., Wilson, W. H., Krisch, H. M., and Mann, N. H. (2001). "A conserved genetic module that encodes the major virion components in both the coliphage T4 and the marine cyanophage S-PM2," PNAS. 98, (20), pp. 11411–11416.

Hamdi, S., Rousseau, G. M., Labrie, S. J., Tremblay, D. M., Kourda, R. S., Slama, K. B., and Moineau, S. (2017): Characterization of two polyvalent phages infecting *Enterobacteriaceae*. Scientific Reports. 7-40349.

Harada, L. K., Silva, E.C., Campos, W.F., Del Fiol, F.S., Vila, M., Dąbrowska, K., Krylov, V.N., and Balcão, V.M. (2018). Biotechnological applications of bacteriophages: state of the art Microbiological Research. 212–213, pp. 38-58

Hatfull, G. F., Pedulla, M. L., JacobsSera, D., Cichon, P. M., Foley, A., Ford, M. E., Gonda, R. M., Houtz, J. M., Hryckowian, A. J., Kelchner, V. A., Namburi, S., Pajcini, K. V., Popovich, M. G., Schleicher, D. T., Simanek, B. Z., Smith, A. L., Zdanowicz, G. M., Kumar, V., Peebles, C. L., Jacobs, W. R. Jr., Lawrence, J. G., and Hendrix, R.W. (2006). Exploring the mycobacteriophage metaproteome: phage genomics as an educational platform. PLoS Genetics. 2: e92.

Hatfull, G. F., and Hendrix, R. W. (2011). Bacteriophages and their genomes. Current Opinion in Virology. 1(4): 298-303.

Hendrix R. W. (2009). Jumbo Bacteriophages. Volume 328 of the series Current Topics in Microbiology and Immunology. pp 229-240.

Hendrix, R. W., Margaret, C. M., Neil Burns, S. R., Ford, M. E., and Hatfull, G. F. (1999). Evolutionary relationships among diverse bacteriophages and prophages: All the world's a phage. Proceedings of the National Academy of Sciences. 96 (5): 2192-2197.

Herikstad, H., Motarjemi, Y., and Tauxe, R. V. (2002). *Salmonella* surveillance: a global survey of public health serotyping. Epidemiology and Infection. 129**:**1-8.

Hertveldt K., Lavigne R., Pleteneva E., Sernova N., Kurochkina L., Korchevskii R., Robben, J., Mesyanzhinov, V. Krylov, V.N. and Volckaert, G. (2005). Genome comparison of Pseudomonas aeruginosa large phages. Journal of Molecular Biology. 354, 536–545.

Higashiyama, T., and Yamada, T. (1991). Electrophoretic karyotyping and chromosomal gene mapping of *Chlorella.* Nucleic Acids Research. 19, pp. 6191-6195

Hohmann, E. L. (2001). Nontyphoidal salmonellosis. Clinical Infectious Diseases. 32:263–9.

Hooton, S. P., Timms, A. R., Rowsell, J., Wilson, R., and Connerton, I. F. (2011). *Salmonella Typhimurium*-specific bacteriophage PhiSH19 and the origins of species specificity in the Vi01-like phage family. Virology Journal. 8, 498.

Howard-Varona, C., Hargreaves, K. R., Abedon, S. T., and Sullivan, M. B. (2017). Lysogeny in nature: mechanisms, impact and ecology of temperate phages. The ISME journal, 11(7), 1511–1520.

Hu, B., Margolin, W., Molineux, I. J. and Liu, J. (2015). Structural remodeling of bacteriophage T4 and host membranes during infection initiation. PNAS. 112, E4919–E4928.

Humphrey, J. H., Prendergast, A. J., and Mbuya, M. N. N. (2015). The sanitation hygiene infant nutrition efficacy (SHINE) trial: rationale, design, and methods. Clinical Infectious Diseases. 61 (7): S685–S702.

Hyman, P., and Abedon, S. T. (2010). Bacteriophage host range and bacterial resistance. Advances in Applied Microbiology. 70, 217–248.

Iino, T. and Mitani, M. (1967). Infection of Serratia marcescens by bacteriophage chi. Journal of Virology. 1(2): 445–447.

International Organization for Standardization (2000). Water Quality – Detection and Enumeration of Bacteriophages, Part 2: Enumeration of Somatic Coliphages. ISO10705-2:2000(E). Geneva: International Organization for Standardization. Journal of Virology. 1, pp. 445-447.

Jensen, K. C., Hair, B. B., Wienclaw, T. M., Murdock, M. H., Hatch, J. B., Trent, A. T., White, T. D., Haskell, K. J., and Berges, B. K. (2015). Isolation and Host Range of Bacteriophage with Lytic Activity against Methicillin-Resistant *Staphylococcus aureus* and Potential Use as a Fomite Decontaminant. PLoS One. 10(7): e0131714.

Johnson, L. R., Hannah Gould, L., Dunn, J. R., Berkelman, R., and Mahon, B. E. for the FoodNet Travel Working Group (2011). Salmonella Infections Associated with International Travel: A Foodborne Diseases Active Surveillance Network (FoodNet) Study. Foodborne Pathogens and Disease. 8 (9).

Johnson R. P., Gyles C. L., Huff W. E., Ojha S., Huff G. R., Rath N. C., and Donoghue A. M. (2008). Bacteriophages for prophylaxis and therapy in cattle, poultry and pigs. Animal Health Research Reviews. 9 (2): 201-15.

Jones, K. E., Patel, N. G., Levy, M. A., Storeygard, A. Balk, D. Gittleman, J. L. and Daszak, P. (2008). Global trends in emerging infectious diseases. Nature. 451, 990–993.

Jurczak-Kurek, A., Gąsior, T., Nejman-Faleńczyk, B., Bloch, S., Dydecka, A., Topka, G., Necel, A., Jakubowska-Deredas, M., Narajczyk, M., Richert, M., Mieszkowska, A., Wróbel, B., Węgrzyn, G., and Węgrzyn, A. (2016). Biodiversity of bacteriophages: morphological and biological properties of a large group of phages isolated from urban sewage. Scientific Reports. 6:34338.

Keen, E. C. (2015). A century of phage research: bacteriophages and the shaping of modern biology. Bioessays. 37, 6–9.

Kęsik-Szeloch, A., Drulis-Kawa, Z., Weber-Dąbrowska, B., Kassner, J., Majkowska-Skrobek, G., Augustyniak, D., Lusiak Szelachowska, M., Zaczek, M., Górski, A., and Kropinski. A.M. (2013). Characterising the biology of novel lytic bacteriophages infecting multidrug resistant *Klebsiella pneumoniae.* Virology Journal. 10, p. 100

Kiljunen, S., Hakala, K., Pinta, E., Huttunen, S., Pluta, P., Gador, A., Lönnberg, H., and Skurnik, M. (2005). Yersiniophage phi R1-37 is a tailed bacteriophage having a 270 kb DNA genome with thymidine replaced by deoxyuridine. Microbiology-SGM. 151, 4093–4102.

Kim, J. H., Son, J. S., Choi, Y. J., Choresca, C. H., Shin, S. P., Han, J. E., Jun, J.W., and Park, S.C. (2012). Complete genome sequence and characterization of a broad-host range T4-like bacteriophage phiAS5 infecting *Aeromonas salmonicida* subsp. *Salmonicida*. Veterinary Microbiology. 157, pp. 164-171

Kim, J.H., Gomez, D.K., Nakai, T., and Park, S.C. (2010). Isolation and identification of bacteriophages infecting ayu *Plecoglossus altivelis* altivelis specific *Flavobacterium psychrophilum*. Veterinary Microbiology. 140 (1–2), pp. 109-115

Kim, K. H., Chang, H. W., Nam, Y. D., Roh, S. W. and Bae, J. W. (2010): Phenotypic characterization and genomic analysis of the *Shigella sonnei* bacteriophage SP18. Journal of Microbiology. 48 (2): 213-22.

Kim, M. S., Hong S. S., Park K., and Myung H. (2013). Genomic analysis of bacteriophage PBECO4 infecting *Escherichia coli* O157:H7. Archive of Virology. 158, 2399–2403.

Koberg, S., Brinks, E., Fiedler, G., Hüsing, C., Cho, G. S., Hoeppner, M. P., Heller, K. J., Neve, H., and Franz, C. M. (2017). Genome Sequence of *Klebsiella pneumoniae* Bacteriophage PMBT1 Isolated from Raw Sewage. Genome Announcement. 5(8): e00914-16.

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Molecular Biology and Evolution. 33(7):1870–1874.

Kutter, E. (2009). Phage host range and efficiency of plating. Methods in Molecular Biology. 501, pp. 141-149.

Kutter, E., De Vo,s D., Gvasalia, G., Alavidze, Z., Gogokhia, L., Kuhl, S., and Abedon, S. T. (2010). Phage therapy in clinical practice: treatment of human infections. Current Pharmaceutical Biotechnology. 11: 69–86.

Landers, T. F., Cohen, B., Wittum, T. E., and Larson, E. L. (2012). A review of antibiotic use in food animals: perspective, policy, and potential. Public health reports. 127(1), 4–22.

Lawrence, J. G., Hatfull, G. F., and Hendrix, R. W. (2002). Imbroglios of viral taxonomy: genetic exchange and failings of phenetic approaches. Journal of Bacteriology. 184, 4891–4905.

Le, S., He, X., Tan, Y., Huang, G., Zhang, L., Lux, R., Shi, W., and Hu, F. (2013) Mapping the Tail Fiber as the Receptor Binding Protein Responsible for Differential Host Specificity of Pseudomonas aeruginosa Bacteriophages PaP1 and JG004. PLoS ONE. 8(7): e68562.

Leiman, P. G., Arisaka, F., van Raaij, M. J., Kostyuchenko, V. A., Aksyuk, A. A., Kanamaru, S., and Rossmann, M. G. (2010). Morphogenesis of the T4 tail and tail fibers. Virology Journal. **3; 7:355.**

Leiman, P. G., and Shneider, M. M. (2012) Contractile Tail Machines of Bacteriophages. In: Rossmann M., Rao V. (eds) Viral Molecular Machines. Advances in Experimental Medicine and Biology, vol 726. Springer, Boston, MA. 93-114.

Leon-Velarde, C. G., Happonen, L., Pajunen, M., Leskinen, K., Kropinski, A. M., Mattinen, L., Rajtor, M., Zur, J., Smith, D., Chen, S., Nawaz, A., Johnson, R. P., Odumeru, J. A., Griffiths, M. W., and Skurnik, M. (2016). *Yersinia enterocolitica*-Specific Infection by Bacteriophages TG1 and ϕR1-RT Is Dependent on Temperature-Regulated Expression of the Phage Host Receptor OmpF. Applied and Enviromental Microbiology. 82(17):5340-5353.

Leskinen, K., Blasdel, B. G., Lavigne, R., and Skurnik, M. (2016). RNA-Sequencing reveals the progression of phage-host interactions between phi R1-37 and *Yersinia enterocolitica*. Viruses-Basel. 8:111.

Li, E., Wei, X., Ma, Y., Yin, Z., Li, H., Lin, W., Wang, X., Li, C., Shen, Z., Zhao, R., Yang, H., Jiang, A., Yang, W., Yuan, J., and Zhao, X. (2016). Isolation and characterization of a bacteriophage phiEap-2 infecting multidrug resistant *Enterobacter aerogenes*. Scientific Reports. 6:28338.

Lin, D. M., Koskella, B., and Lin, H. C. (2017). Phage therapy: An alternative to antibiotics in the age of multi-drug resistance. World Journal of Gastrointestinal Pharmacology and Therapeutics. 8(3):162–173.

Lin, N. T., Chiou, P. Y., Chang, K. C., Chen, L. K., and Lai, M. J. (2010). Isolation and characterization of ϕAB2: a novel bacteriophage of *Acinetobacter baumannii*. Research in Microbiology. 161, 308–314.

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

Logan, L. K., and Weinstein, R. A. (2017). The Epidemiology of Carbapenem-Resistant Enterobacteriaceae: The Impact and Evolution of a Global Menace. The Journal of Infectious Diseases. 15,215. S28-S36.

Lu, T. K., and Koeris, M. S. (2011). The next generation of bacteriophage therapy. Current Opinion in Microbiology. 14, 524–531.

Mapes, A. C., Trautner, B. W., Liao, K. S., and Ramig, R. F. (2016). Development of expanded host range phage active on biofilms of multi-drug resistant *Pseudomonas aeruginosa*. Bacteriophage. 6(1): e1096995.

Matsui, T., Yoshikawa, G., Mihara, T., Chatchawankanphanich, O., Kawasaki, T., Nakano, M., Fujie, M., Ogata, H., and Yamada, T. (2017). Replications of Two Closely Related Groups of Jumbo Phages Show Different Level of Dependence on Host-encoded RNA Polymerase. Frontiers in Microbiology. 8:1010.

Mead, P. S., Slutsker, L., Dietz, V., McCaig, L. F., Bresee, J. S., Shapiro, C., Griffin, P. M., and Tauxe, R. V. (1999). Food-related illness and death in the United States. Emerging infectious diseases. 5(5), 607–625.

Merabishvili, M., Vandenheuvel, D., Kropinski, A. M., Mast, J., De Vos, D., Verbeken, G., Noben, J. P., Lavigne, R., Vaneechoutte, M., and Pirnay, J. P. (2014). Characterization of newly isolated lytic bacteriophages active against *Acinetobacter baumannii*. PLoS One. 9 (8): e104853.

Merril, C. R., Scholl, D., and Adhya, S. (2006) Phage therapy. In: Calendar R (ed). The bacteriophages. 2nd edn. New York: Oxford University Press. pp 725–741.

Mesyanzhinov, V. V., Robben, J., Grymonprez, B., Kostyuchenko, V. A., Bourkaltseva, M. V., Sykilinda, N. N., Krylov, V. N., and Volckaert, G. (2002). The genome of

bacteriophage phi KZ of *Pseudomonas aeruginosa*. Journal of Molecular Biology. 317, 1–19.

Miller, E. S., Heidelberg, J. F., Eisen, J. A., Nelson, W. C., Durkin, A. S., Ciecko, A., Feldblyum, T. V., White, O., Paulsen, I. T., Nierman, W. C., Lee, J., Szczypinski, B., and Fraser, C. M. (2003a). Complete genome sequence of the broad-host-range vibriophage KVP40: comparative genomics of a T4-related bacteriophage. Journal of Bacteriology. 185(17):5220-33.

Miller, E. S., Kutter, E., Mosig, G., Arisaka, F., Kunisawa, T., and Rüger, W. (2003b). Bacteriophage T4 genome. Microbiology and molecular biology reviews: MMBR, 67(1), 86–156.

Moreno Switt, A. I., Orsi, R. H., den Bakker, H. C., Vongkamjan, K., Altier, C., and Wiedmann, M. (2013). Genomic characterization provides new insight into Salmonella phage diversity. BMC genomics. 14, 481.

Nikfar, R., Shamsizadeh, A., Darbor, M., Khaghani, S., and Moghaddam, M. (2017). A study of prevalence of *Shigella* species and antimicrobial resistance patterns in paediatric medical center, Ahvaz, Iran. Iranian Journal of Microbiology. 9(5):277–282

Nilsson, A. S. (2014). Phage therapy–constraints and possibilities. Upsala Journal of Medical Sciences. 119, 192–198.

Nobrega, F.L., Vlot, M., Jonge, P.A., Dreesens, L. L., Beaumont, H. J. E., Lavigne, R., Dutilh, B. E., and Brouns, S. J. J. (2018). Targeting mechanisms of tailed bacteriophages. Nature Reviews in Microbiology. 16, 760–773.

O'neill, J. (2014). Antimicrobial resistance: tackling a crisis for the health and wealth of nations. Review on Antimicrobial Resistance.

O, neill, j. (2016). The review on antimicrobial resistance. Tackling drug-resistant infections globally: final report and recommendations. UK. 2016.

O'Donnell, M., Langston, L., and Stillman, B. (2013). Principles and concepts of DNA replication in Bacteria, Archaea, and Eukarya. Cold Spring Harbor Perspective in Biology. 5: a010108.

Obradovic, A., Jones, J. B., Momol, M. T., Olson, S. M., Jackson, L. E., Balogh, B., Guven, K., and Iriarte, F. B. (2005). Integration of biological control agents and systemic acquir ed resistance inducers against bacterial spot on tomato. Plant Disease. 89:712-716.

Oliveira, H., Azeredo, J., Lavigne, R., and Kluskens, L. D. (2012). Bacteriophage endolysins as a response to emerging foodborne pathogens. Trends in Food Science and Technology. 28: 103-115.

Park, M., Lee, J.H., Shin, H., Kim, M., Choi, J., Kang, D.H., Heu, S., and Ryu, S. (2012). Characterization and comparative genomic analysis of a novel bacteriophage, SFP10, simultaneously inhibiting both *Salmonella enterica* and *Escherichia coli* O157:H7. Applied Environmental Microbiology. 78:58–69.

Pedulla, M. L., Ford, M.E., Houtz, J.M., Karthikeyan, T., Wadsworth, C., and Lewis, J.A., Jacobs-sera, D., Falbo, J., Gross, G., Pannunzio, N. R., Brucker, W., Kumar, V., Kandasamy, J., Keenan, L., Bardarov, S.Kriakov, J., Lawrence, J. G., Jacobs, W. R., Hendrix, R. W., and Hatfull, G. F. (2003). Origins of highly mosaic mycobacteriophage genomes. Cell. 113, 171–182.

Peng, Q., and Yuan, Y. (2018). Characterization of a newly isolated phage infecting pathogenic *Escherichia coli* and analysis of its mosaic structural genes. Scientific Reports. 8, 8086.

Pérez Pulido, R., Grande Burgos, M., J., Gálvez, A., and Lucas López, R. (2015). Application of bacteriophages in post-harvest control of human pathogenic and food spoiling bacteria. Critical Reviews in Biotechnology. 36(5): 851–861.

Petty, N. K., Evans, T. J., Fineran, P. C., and Salmond, G. P. C. (2007). "Biotechnological exploitation of bacteriophage research. Trends in Biotechnology. 25 (1), pp. 7–15.

Picot-Gueraud, R., Batailler, P., Caspar, Y., Hennebique, A., and Mallaret, M. R. (2015). Bacteremia caused by multidrug-resistant bacteria in a French university hospital center: 3 years of collection. American Journal of Infection Control. 43, pp. 960-964.

Principi, N., Silvestri, E., and Esposito, S. (2019) Advantages and limitations of bacteriophages for the treatment of bacterial infections. Frontiers in Pharmacology. 10:513.

Ross, A., Ward, S. and Hyman, P. (2016). More Is Better: Selecting for broad host range bacteriophages. Frontiers in Microbiology. 7:1352.

Roszniowski, B., Latka, A., Maciejewska, B., Vandenheuvel, D., Olszak, T., Briers, Y., Holt, G.S., Valvano, M.A., Lavigne, R., Smith, D.L., and Drulis-Kawa, V. (2017). The temperate *Burkholderia* phage AP3 of the *Peduovirinae* shows efficient antimicrobial

activity against *B. cenocepacia* of the IIIA lineage. Applied Microbiology and Biotechnology. 101, pp. 1203-1216

Sambrook, J., and Russel, D. (2001). Molecular cloning: A laboratory manual (3rd ed.), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Santos, S. B., Fernandes, E., Carvalho, C. M., Sillankorva, S., Krylov, V. N., Pleteneva, E. A., Shaburova, O.V., Nicolau, A., Ferreira E. C., and Azeredo, J. (2010). Selection and characterization of a multivalent Salmonella phage and its production in a nonpathogenic *Escherichia coli* strain. Applied Environmental Microbiology. 76:7338–7342.

Schroeder, C. M., Naugle, A. L., Schlosser, W. D., Hogue, A. T., Angulo, F. J., Rose, J. S., Ebel, E. D., Disney, W. T., Holt, K. G., and Goldman, D. P. (2005). Estimate of illnesses from *Salmonella* Enteritidis in eggs, United States, 2000. Emerging Infectious Diseases Journal.11**:**113-115.

Serwer, P., Hayes, S. J., Thomas, J. A., and Hardies, S. C. (2007). Propagating the missing bacteriophages: a large bacteriophage in a new class. Virology journal. 4, 21.

Sharma, R., Pielstick, B. A., Bell, K. A., Nieman, T. B., Stubbs, O. A., Yeates, E. L., Baltrus, D. A., and Grose, J. H. (2019). A Novel, Highly Related Jumbo Family of Bacteriophages That Were Isolated Against Erwinia. Frontiers in Microbiology. 10:1533.

Shende, R. K., Hirpurkar, S. D., Sannat, C., Rawat, N., and Pandey, V. (2017). Isolation and characterization of bacteriophages with lytic activity against common bacterial pathogens. *Vet World*. 10(8):973-978.

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., et al. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Molecular Systems Biology.7:539.

Silbergeld, E. K., Graham, J., and Price, L. B., (2008). Industrial food animal production, antimicrobial resistance, and human health. Annual Review of Public Health. 29, 151–169.

Simoliunas, E., Kaliniene, L., Truncaite, L., Klausa, V., Zajanckauskaite, A., and Meskys R. (2012). Genome of *Klebsiella* sp.-Infecting Bacteriophage vB_KleM_RaK2. Journal of Virology. 86, 5406–5406.

Skurnik, M., and Strauch, E., (2006). Phage therapy: facts and fiction. International Journal of Medical Microbiology. 296 (1), 5–14.

Skurnik, M., Pajunen, M., and Kiljunen, S. (2007). Biotechnological challenges of phage therapy. Biotechnology Letters. 29: 995—1003.

Starling, S. (2017). *Shigella sonnei* has the edge. Nat Rev Microbiol 15:450–451.

Su, J., Sun, H., Liu, J., Guo, Z., Fan, G, Gu, G., and Wang, G. (2017): Complete genome sequence of a novel lytic bacteriophage isolated from *Ralstonia solanacearum*. Archives of Virology. 162: 3919-3923.

Sulakvelidze, A. (2013). Using lytic bacteriophages to eliminate or significantly reduce contamination of food by foodborne bacterial pathogens. Journal of the Science of Food and Agriculture. 93, 3137–3146.

Summer, E. J., Berry, J., Tran, T. A., Niu, L. Struck, D. K., and Young, R. (2007). Rz/Rz1 lysis gene equivalents in phages of Gram-negative hosts. Journal of Molecular Biology. 373. 1098-1112

Summers, W. C. (1999). Felix d'Herelle and the origins of molecular biology. Yale University Press 1999; New Haven, Conn.

Tacconelli, E., Carrara, E., Savoldi, A., Harbarth, S., Mendelson, M., Monnet, D. L., et al. (2018). Discovery, research, and development of new antibiotics: the WHO priority list of antibiotic-resistant bacteria and tuberculosis. Lancet Infectious Diseases. 18, 318–327.

Thomas, J. A., Hardies, S. C., Rolando, M., Hayes, S. J., Lieman, K., Carroll, C. A., Weintraub, S. T., and Serwer, P. (2007). Complete genomic sequence and mass spectrometric analysis of highly diverse, atypical Bacillus thuringiensis phage 0305phi8-36. Virology. 368(2):405-21.

Twort, F. W. (1915). An investigation on the nature of ultra-microscopic viruses. Lancet. 186, pp. 1241-1243.

Uchiyama, J., Rashel, M., Maeda, Y., Takemura, I., Sugihara, S., Akechi, K., Muraoka, A., Wakiguchi, H., and Matsuzaki, S. (2008). Isolation and characterization of a novel Enterococcus faecalis bacteriophage ϕEF24C as a therapeutic candidate. FEMS Microbiology Letters. 278, 200–206.

Van Boeckel, T. P., Glennon, E. E., Chen, D., Gilbert, M., Robinson, T. P., Grenfell, B. T., Levin, S. A., Bonhoeffer, S., and Laxminarayan, R. (2017). Reducing antimicrobial use in food animals. Science. 357(6358):1350-1352.

Van Etten, J. L., Lane, L. C., and Dunigan, D. D. (2010). DNA viruses: the really big ones (Giruses). Annual Review of Microbiology. 64, 83–99.

Vandenheuvel, D., Lavigne, R., and Brüssow, H. (2015). Bacteriophage therapy: advances in formulation strategies and human clinical trials. Annual Review of Virology. 2, 599–618.

Vinga, I., Baptista, C., Auzat, I., Petipas, I., Lurz, R., Tavares, P., and São-José C (2012). Role of bacteriophage SPP1 tail spike protein gp21 on host cell receptor binding and trigger of phage DNA ejection. Molecular Microbiology. 83:289–303

Vos, T., Abajobir, A.A., Abate, K.H. Abbafati, C. Abbas, K.M., Abd-Allah, F. *et al.* (2017). Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet. 390, pp. 1211-1259.

Weinbauer, M. G. (2004). Ecology of prokaryotic viruses. FEMS Microbiology Reviews. 28, 127–181.

Welkos, S., Schreiber, M., and Bae, H. (1974). Identification of *Salmonella* with the O-1 bacteriophage. Applied Microbiology. 28. pp. 618-622.

Williamson, K. E., Radosevich, M., and Wommack, K. E. (2005). Abundance and diversity of viruses in six Delaware soils. Applied Environmental Microbiology. 71:3119–3125

Wilson, J. H. (1973). Function of the bacteriophage T4 transfer RNA's. Journal of molecular biology. 74 (4), 755-757.

Wittebole, X., Roock, S. D., and Opa S. M. (2014). A historical overview of bacteriophage therapy as an alternative to antibiotics for the treatment of bacterial pathogens. Virulence. 5(1): 226–235.

Woolston, J., Parks, A. R., Abuladze, T., Anderson, B., Li, M., Carter, C., Hanna, L. F., Heyse, S., Charbonneau, D., and Sulakvelidze, A. (2013). Bacteriophages lytic for *Salmonella* rapidly reduce *Salmonella* contamination on glass and stainless steel surfaces. Bacteriophage. 3(3): e25697.

Yamada, T., Satoh, S., Ishikawa, H., Fujiwara, A., Kawasaki T., Fujie M., Ogata, H. (2010). A jumbo phage infecting the phytopathogen *Ralstonia solanacearum* defines a new lineage of the *Myoviridae* family. Virology. 398, 135–147.

Yoshikawa, G., Askora, A., Blanc-Mathieu, R., Kawasaki, T., Li, Y., Nakano, M., Ogata, H., and Yamada, T. (2018). *Xanthomonas citri* jumbo phage XacN1 exhibits a wide host range and high complement of tRNA genes. Scientific Reports. 8:4486.

Young, R. (1992). Bacteriophage lysis: mechanism and regulation. Microbiological reviews. 56(3), 430–481.

Yuan, Y. and Gao, M. (2017). Jumbo bacteriophages: an overview, Frontiers in Microbiology. 8, p. 403.

Yuan, Y., and Gao, M., (2016a). Proteomic analysis of a novel bacillus jumbo phage revealing glycoside hydrolase as structural component. Frontiers in Microbiol. 7:745.

Zhou, Y., Liang, Y., Lynch, K. H., Dennis, J. J., and Wishart, D. S. (2011). PHAST: a fast phage search tool. Nucleic Acids Research. 39: W347–W352.

ACKNOWLEDGEMENTS

First of all, I think all my sincere thanks would not fulfill the pot for my research advisor **Professor Dr. Takashi Yamada** for his unlimited help and support not only during my research and graduate career but also throughout all my life aspects. He provided me with lots of insights, the best research ideas, detailed informations in our field and treated me like his son. Really, I consider him not only my sensei but also my second father. Thanks again and again Prof. Dr. Takashi Yamada.

I also would like to thank **Associate professor Dr. Makoto Fujie** and **Dr. Takeru Kawasaki** for their invaluable guidance, assistance and discussion to improve my scientific skills and finishing my thesis. I also appreciate to my lab colleagues and all members of Graduate School of Advanced Sciences of Matter (Adsm), Hiroshima University for their kindness and friendship during my study.

I also have to thank **Dr. Ahmed Askora**, Professor of Microbiology, Faculty of Science, Zagazig University and also **Professor Tadashi Shimamoto**, his lab members and my dear friend Dr. Ahmed Mahrous Soliman of Graduate School of Biosphere for their help and cooperation.

My sincere thanks to **all staff members of Zoonoses Department**, Faculty of Veterinary Medicine, Zagazig University- Egypt, for their continuous support and help.

I would like to acknowledge my gratefulness to the **Ministry of Higher Education, Egypt** for granting me the scholarship for this study.

I would like to thank **my parents,** my uncle **Prof. Dr. Alaa Elmorshdy,** my family, my friends, my brother, my sisters, **my dear wife** and **my little baby Mona** for inspiring me and continuous help and support.

Finally, all praise and glory belong to **Almighty Allah**, the sustainer of the universe for all my life and guidance to finish my research and study successfully.

Alaaeldin Saad, 2020