Isolation and characterization of phage infects to novel soft rot Enterobacter species

グエン, コン, タイン

出版情報：九州大学, 2020, 博士（農学）, 課程博士
バージョン：
権利関係：

Isolation and characterization of phage infects to novel soft rot *Enterobacter* species

NGUYEN CONG THANH

2020

Table of Contents

List of abbreviation

Electrospray ionization	ESI
Enterobacter cloacae complex	Ecc
Lysogeny broth	LB
National Institute of Technology and Evaluation	NITE
NITE Biological Resource Center	NBRC
Pectobacterium carotovorum subsp. *carotovorum*	Pcc
Plaque-forming units	PFU
Polyacrylamid gel electrophoresis	PAGE
Sodium dodecyl sulfate	SDS
Soft Rot *Enterobacteriacae*	SRE
Soft Rot *Pectobacteriacae*	SRP
Superinfection exclusion	Sie
Tadem mass spectrometry	MS/MS
Transmission electron microscopy	TEM

Chapter 1. Introduction

1. Introduction of soft rot

Bacterial soft rot is one of the most important diseases of the agricultural ecosystem. Unlike some other common diseases, bacterial soft rot pathogens lead to serious damage in vegetable or ornamental crop production. It occurs on fleshy vegetables such as the potato, carrot, eggplant, squash, and tomato (Gavrilovic *et al*., 2001). Bacterial soft rot symptoms are caused primarily by secreted pectinases that degrade pectin in the middle lamella and primary plant cell walls. This macerates the plant tissues and causes a wet, often foul-smelling rot of the plant organ (Hugouvieux *et al*., 2014). The bacteria reaches high concentrations in the xylem and can cause necrosis in the vascular tissue. Soft rot is especially observed on plant storage organs, such as tubers, rhizomes, and bulbs (Ma *et al*., 2007). It can also appear in fleshy plant parts, such as succulent stems and leaves, or vegetables such as lettuce (Ma *et al*., 2007). The disease develops quickly after the first symptom observed. When conditions are suitable, it takes two or three days to inoculate tubers, or storage roots rot and takes only a few hours to destroy infected plants after initial wilting symptoms.

Bacterial soft rot appears worldwide and leads to severe loss. It can lead up to 80% loss on Felsina variety of potato (Rahmanl *et al.*, 2010) (Figure 1.1). This disease was first reported in 1891 by Halsted on celery and recorded later in 1898 on cabbage and other crucifers (Bhat *et al*., 2010; Gavrilovic *et al*., 2001). Soft rot on carrots was then described by Jones Harrison (Bhat *et al.,* 2010). In the early 2000s, *Dickeya solani* was identified as a new species in potato, which caused blackleg levels to rise in Europe. This bacterium also led to significant loss of seed potato in European countries in many years (Toth *et al*., 2011), and its appearance caused critical observation of this and related soft rot pathogens in potato. It was also causing disease on monocot bulbs (Chen *et al*., 2015), which are sometimes rotated with potato in Europe. Parkinson *et al.* suggested that it may have been transferable from previous potato crops (2014). In the summer of 2015, there was a similar epidemic caused by *Dickeya* in Maine, the United States. Bacterial soft rot was also reported in forest trees and other vital fruits (Furtado *et al*., 2012; Masyahit *et al*., 2009; Wu *et al*., 2011; García-González *et al*., 2018; Soto *et al*., 2019; Reyes-García *et al*., 2020).

Various bacterial species cause bacterial soft rot. It was reported that there were several bacteria considered as causes of the disease. Both gram-positive and gram-negative bacteria can cause soft rot in plants such as *Psedomonas* (Godfrey and Marshall, 2002; Zhang *et al*., 2016), *Bacillus* (Elbanna *et al*., 2014; Zhong *et al*., 2015), *Burkholderia* (Lu *et al*., 2007), *Pantoea* (Zhou *et al*., 2015; Liao *et al*., 2016), *Pectobacterium* and *Dickeya* (Ma *et al*., 2007; Charkowski *et al*., 2012), *Klebsiella* (Fan *et al*., 2016), *Leuconostoc* (Lampert *et al*., 2017), *Bipolaris* (He PF, 2012), *Clostridia* (Campos, 1982) and *Enterobacter* (Masyahit *et al*., 2009; Wu *et al*., 2011). They all can

cause bacterial soft rot symptoms on plant storage organs. However, *Pectobacterium* sp. and *Dickeya* sp., which previously belonged *Enterobacteriaceae* family, were mainly reported. In 2016, *Enterobacteriaceae* family was reclassified to *Enterobacteriaceae* and *Pectobacteriaceae* (Adeolu *et al*., 2016). Bacterial soft rot caused by *Pectobacterium* and *Dickeya* species were then well known as SRP (Ma *et al*., 2007; Rossmann *et al*., 2018; Salmond, 1992; Adeolu *et al*., 2016). In contrast, SRE which caused by *Enterobacter* species is less reported and understood.

2. Introduction of Soft Rot Enterobacteriacae

SRE was first reported when *Enterobacter cloacae* was determined as a reason for the bacterial disease of papaya fruits in Hawaii based on physiological characteristics (Nishijima, 1987). The causal bacterium was isolated repeatedly from diseased fruit, hot-water treatment tanks, papaya flowers, and the crop and stomach of the oriental fruit fly. This bacterium led the fruit to have soft, yellow, discolored flesh with diffuse margins and an offensive odor (Nishijima, 1987). More seriously, *E. cloacae* was identified as a pathogen causing internal decay of mature onions during hot temperatures, which affected approximately 1 – 5% of 35,000 acres of the crops in California (Bishop, 1990). This disease caused leaves to be discolored and flaccid. In another case, sixteen *Enterobacter* isolates were then collected in an outbreak of soft rot on onion in Washington State, where there has been the same outbreak two years prior. Four of those isolates were identified as *E. cloacae* and pathogenic causes of bulb rot on onion (Schroeder *et al*., 2009). Later on, various species were also determined as the cause of soft rot on chili pepper, konjac, mango, and dragon fruit in Asian countries (Masyahit *et al*., 2009; Wu *et al*., 2011; García-González *et al*., 2018), on *Mabea fistulifera*, a native forest species in Brazil (Furtado *et al*., 2012), dragon fruit in Peru (Soto *et al*., 2019) and on mystax cactus (*Mammillaria mystax*) in the United States. (Reyes-García *et al*., 2020). Interestingly, an isolate causing bacterial wilt on mulberry in China, *Enterobacter mori*, was assigned as a novel species of the genus *Enterobacter* (Wang *et al*., 2010).

3. Introduction of *Enterobacter* genus

Enterobacter was proposed as a genus in 1960 by Hormaeche and Edwards based on the division of the former genus *Aerobacter* into motile, ornithine decarboxylase-positive strains (*Enterobacter*) and nonmotile ODC-negative strains (*Klebsiella*) (Iversen, 2014). Members of *Enterobacter* genus are straight rod-shaped (0.6 – 1.0 × 1.2 – 3.0 μm), motile by flagella, gram-negative, facultatively anaerobic (Joseph, 2005). The species differ much about physiological characteristic. However, in general, the species ferment glucose to produce acid and gas and most strains give a positive Voges-Proskauer reaction, and utilize L-rhamnose (except *Enterobacter asburiae*) (Joseph, 2005). Due to the diversity of this genus, biochemical tests were frequently conducted using API 20E kit for identification of species level (Farmer *et al*., 1985; Pepper *et al*., 2009; Mishra *et al*., 2012). For the thermal condition, while the optimum temperature for growth is 30°C,

some clinical and environmental strains give erratic biochemical reactions at 37°C (Joseph, 2005). The genus *Enterobacter* is also divergent in the genomic term. Although GC content of genomic DNA ranges widely from 52 to 60% (Joseph, 2005), the result of 16S rRNA gene sequences is not reliable to differentiate species in the genus (Naum *et al*., 2008). Therefore, several housekeeping genes such as *rpoB, recA, gyrB* or others were used to classify the species. Consequently, twelve species were assigned, which included *E. cloacae* (Type species)*, Enterobacter amnigenus, Enterobacter asburiae, Enterobacter cancerogenus, Enterobacter cowanii, Enterobacter dissolvens, Enterobacter gergoviae, Enterobacter hormaechei, Enterobacter kobei, Enterobacter nimipressuralis, Enterobacter pyrinus,* and *Enterobacter sakazakii* (current name as *Cronobacter sakazakii*) (Hata *et al*., 2016; Tailliez *et al*., 2010; Fukushima *et al*., 2002; Hong Nhung *et al*., 2007; Joseph, 2005). *Enterobacter* species are found in the natural environment, including water, sewage, vegetables, and soil (Joseph, 2005).

In addition, the result of multilocus sequence analysis and comparative genomic hybridization divided the genus into thirteen genovars (cluster I-XIII): *E. asburiae* (cluster I), *E. kobei* (cluster II), *E. ludwigii* (cluster V), *E. hormaechei* subsp. *oharae* (cluster VI), subsp. *hormaechei* (cluster VII), and subsp. *steigerwaltii* (cluster VIII)*, E. nimipressuralis* (cluster X), *E. cloacae* subsp. *cloacae* (cluster XI) and subsp. *disolvens* (cluster XII), unnamed *E. cloacae* Hoffmann cluster III, IV, and IX, and an unstable *E. cloacae* (cluster XIII). Many of them are human pathogens (Washington *et al*., 1969; Farmer *et al*., 1980; Grimont *et al*., 1983; Brenner *et al*., 1986; Giamarellou, 2010) and some of them are reported as phytopathogens as mentioned above (Table 1.1).

Although *Enterobacter* species were determined as plant pathogens very early, the research of their phytopathogenic genes was few. Related to bacterial soft rot, the plant pathogenic genes of the SRP species were well described (Lee *et al*., 2013; Tanui *et al*., 2017; Narváez-Barragán *et al*., 2020). 16 genes in Pcc mutants were identified and able to be grouped into nutrient utilization (*pyrD, purH, purD, leuA* and *serB*), production of plant cell wall degrading enzymes (*expI, expr* and PCC21_023220 with an unknown function), motility (*flgA, fliA* and *flhB*), biofilm formation (*expI, expr* and *qseC*), susceptibility to antibacterial plant chemicals (*tolC*) and unknown function (ECA2640). Furthermore, due to soft rot infection, pectin, a structural polysaccharide that is required for stability of plant cell walls, was lysed by several extracellular enzymes such as polygalacturonase (Peh), pectate lyase (Pel), pectate lyase (Pel), cellulose (Cel), xylanase and protease (Prt). Accordingly, degradation of pectin leads to plant cell necrosis and tissue maceration (Abbott and Boraston, 2008). In another case, Exl1 genes of *Pectobacterium brasiliense* and *Pectobacterium atrosepticum* were also shown to be virulence factors due to the remodel cell wall component or to alter the barrier properties of plant defense response (Narváez-Barragán *et al*., 2020).

4. Control strategies for bacterial soft rot

Various approaches have been studied to control soft rot, but the level of success was different. The methods to control the bacterial soft rot include using disease-resistant

breeds, physical treatment, chemical treatment, and biological treatment. However, these methods mainly focus on controlling and managing SRP disease.

Using the disease-resistant breed approach, some cultivars showed naturally to get partial resistancy to *Dickeya* and *Pectobacterium* species. The number of researches on the breeding of potato cultivars increased in the early 19th century. However, the research was only partially successful and never resulted in immune cultivars (Czajkowski *et al*., 2011).

The traditional method to store seed tubers is to conserve them in ventilated storages at low temperatures. This method is convenient and cheap. However, it is difficult to control the temperature on a large scale. Hot water treatment has been studied, but this method showed a failure to dry large quantities of the tuber and vegetables rapidly. Ranganna *et al.* (1997) also tested the efficacy of UV irradiation for controlling Pcc in potato tubers. The tubers were inoculated by vacuum infiltration 6 h before irradiation; bacteria were eliminated by a relatively low UV dose of 15 kJ/m^2. However, this method faced a problem of penetration of UV to the pathogens inhabiting inside the tubers.

Chemical control methods used against bacterial diseases are based on the eradication of the pathogen or the disadvantages of environmental conditions for disease development. Antibiotics, inorganic and organic salts or combinations of these compounds are mostly used. For years, streptomycin was considered a potential control agent against soft rot in potato. Immersion of seed tubers in a mixture of “streptomycin and oxytetracycline hypocholoride” or ”streptomycin and mercury compounds” before planting reduced the appearance of bacterial disease in the field and tuber decay in storage (Bonde and de Souza, 1954). Kasugamycin or virginiamycin also showed positive results (Bartz *et al*., 1986; Wyatt and Lund, 1981). Later years, the application of bronopol and 7-chloro-1-methyl-6-fluoro-1,4-dihydro-4-oxo-3-quinolinic carboxylic acid was also studied (Bartz *et al*., 1986). Mills *et al.* (2006) showed that inorganic and organic salts, including aluminum acetate, sodium metabisulphate, propyl parabe, sodium benzoate, potassium sorbate, calcium propionate, sodium hydrochloride, sodium bicarbonate, aluminum chloride, and copper sulphate, could inhibit the growth of Pcc *in vitro*. Some of these salts have already been approved as food preservatives. Consequently, they were used to control bacterial soft rot in registration testing. However, using the chemicals raised the dangers of the transfer of pathogenic bacteria to humans and animals. Alternatively, organic compounds such as hydroxyquinoline and 5-nitro-8-hydroxyquinoline were tested to control disease in potato tubers (Harris, 1979).

Biological control is a trendy method nowadays. Biocontrol strategies include the use of antagonists affecting the pathogens directly, or via antibiosis, competition for nutrients of plant systemic resistance (Howarth, 1991). These studies were restricted to *in vitro* overlay studies, potato slice assays or *in vitro*-cultured potato plants. After nearly a decade, fluorescent *Pseudomonas* spp. was used to eliminate populations of blackleg and soft rot bacteria on potato roots and inside progeny tubers (Kloepper *et al*., 1980; Kloepper, 1983). Later, *Pseudomonas putida* showed its influence on bacterial soft rot on

potato when a bacterial suspension was directly applied to the tuber periderm (Colyer, 1984). The bacterial predators *Badellovibrio* and similar to organisms that prey on gram-negative bacteria were tested for controlling SRP diseases (Rendulic *et al*., 2004; Youdkes *et al*., 2020). However, the interaction with the prey bacterial cells is ruled by a specific predator-prey relationship in which the population of both microorganisms may fluctuate without complete eradication of the target bacteria (Crowley *et al*., 1980).

5. Introduction of bacteriophage

Bacteriophages (phages) are known as bacterial viruses, which infect bacterial cells. It was first described independently by two scientists: Frederick Twort, a British pathologist, in 1915 and Felix d'Herelle, a French-Canadian microbiologist, in 1917. At that time, phages were recognized as filterable, infectious agents that could lyse bacteria (Chanishvili, 2012). In 1940, phages were confirmed as viruses using an electron microscope (Kruger *et al*., 2000; Ackermann, 2003). One year later, Luria and Anderson, described phages in more detail, as containing "a round head" with a straight tail, and a peculiar sperm-like particles (Luria and Anderson, 1942; Luria *et al*., 1943). Just a few years from d'Herelle's first discovery of phage, phage therapy was first proposed by using phages to treat dysentery (Chanishvili, 2012). The phages were ingested by d'Herelle, Hutinel, and several hospital interns to confirm its safety before administering it to a 12-year-old boy. The patient's symptoms ceased after the treatment with the phages and the boy fully recovered within a day. In animals, phage therapy was first used to treat *Escherichia coli* in mice, calves, piglets, and lambs (Smith and Huggins, 1982; Williams Smith and Huggins, 1983). These work's results showed that single phage treatment was more effective than multiple injections of various antibiotics (tetracycline, ampicillin, chloramphenicol, or trimethoprim plus sulfafurazole). The phages persisted for one day in the blood stream, and for many days in the spleen.

The phages exist abundantly in the environment with the estimated population of 10^{30} to 10^{32} in the biosphere (Abedon, 2009). Phage particles vary in size from 20 to 200 nm and consist of two major components: a nucleic acid (DNA or RNA) and a coat protein (capsid) (Ackermann, 2009). Phage genome has various sizes. To date, the smallest is the genome of levivirus coliphage BZ13 containing 3,412 bp, and the biggest is myovirus *Bacillus megaterium* phage G genome consisting of 497,513 bp (Keen *et al*., 2015). Chemical and physical resistance to phage depends on the variety of phage. Generally, phages are degraded by UV light and sunlight. Phage is stable in a wide range of pH and temperature. An *Enterobacter aerogenes* phage, phiEap-2, which was isolated from sewage, kept stable at 4, 25, 37°C and was slightly affected at 50°C. From pH of 6 to 11, the stability of this phage was constant (Li *et al*., 2016). Another example is that phage phiEap-3 of this bacterium was stable in the range of pH from 6 to 7 and up to 37°C. The titer of phiEap-3 was reduced sharply at 50°C after 30-min incubation (Zhao *et al.*, 2019). Especially, in another case, hyperhermophilic phage can survive up to 85^0C and pH of 3 (Nagayoshi *et al*., 2016). Protein denaturing agents also effect phages.

However, the effect of the agents depends on concentration, temperature, and phage strains (Ackermann, 2009).

Phage includes more than 6000 different phages which were discovered and described both morphologically and physiology. In 2017, the ICTV EC49 meeting two new families were added, eight new subfamilies, 34 new genera, and 91 new species (Table 1.2). The two new families are *Ackermanviridae* and *Portogloboviridae* (Adriaenssens *et al*., 2018). Mostly, phages are tailed while the remains are polyhedral, filamentous or pleomorphic. Phages are classifiable based on morphology, their genomic materials, their specific hosts [for example the *Staphylococcal* phage family (Deghorain and Van Melderen, 2012), the *Pseudomonas* phage family (Ceyssens and Lavigne, 2010) and so on], their habitats (marine virus vs. other habitats), and their life cycles. In terms of virion morphology, there are tailed, isometric (generally icosahedral), helical (which are filamentous or rod-shaped), and pleomorphic phages. Based on genomic material, phages are divided into double strand DNA (dsDNA), single strand DNA (ssDNA), single stranded RNA (ssRNA), and double stranded RNA (dsRNA). In fact, the dsDNA phages are the majority, while the remains are relatively rare (Ackermann, 2009). Based on life cycle, phages were divided into lytic, lysogenic, pseudo-lysogenic, and chronic infection (Weinbauer, 2004; Drulis-Kawa *et al*., 2013).

5.1. Life cycles of phage

5.1.1. Lytic life cycle

The lytic phages infect their host cells by inhibition of host protein synthesis and then initiate phage genome replication and protein synthesis. Finally, phage particles are released from host cells and continued to infect other host cells (Figure 1.2). Generally, the lytic cycle of phage contains five stages including adsorption, penetration, latent period, maturation, and lysis (Clokie and Kropinski, 2009).

In the adsorption stage, phage tail fibers or spikes bind to bacterial cell surfaces such as cell wall, flagella, pili, or capsules. The complete tail apparatus is composed of the portal protein, tail spikes, and tail fibers. The tail spike structures play important roles in this stage. Binding sites are recognized by specific receptors on the host cells. Bacteriophage receptors are various and determined by the surface layer structure of host cells. Table 1.3 shows examples of some bacteriophage receptors (Letarov and Kulikov, 2017). The lipopolysaccharide molecule of the bacteria membrane consists of the highly polar phospholipid, the core oligosaccharide (inner and outer cores), and the O-antigen (Figure 1.3). The O-antigens of gram-negative bacteria are often recognized by phages as primary receptors. Whereas, lipopolysaccharide core can also be both primary and secondary receptors (Letarov and Kulikov, 2017). In some cases, phage must bind to the primary receptors and degrade this structure before attaching to the secondary receptors. *E. coli* phage CBA120 is clear evidence for this mechanism. This phage contains four tail spike proteins: TSP1, TSP2, TSP3, and TSP4 (*orf213*, *orf 212*, *orf 211*, and *orf210*

products). While TSP4 binds to primary receptors, TSP1, TSP2, and TSP3 degrade O-antigen polysaccharides (secondary receptors) from *E. coli* strains O157, O77, and O78, respectively (Greenfield *et al*., 2019; Plattner *et al*., 2019).

After attaching to the host cell surfaces, the DNA of phages is injected into the host cell to the next process, so-called penetration (Paucker, 1977). Phages with different morphology have various modes of injection. In the *Bacillus subtilis* ϕ29, a phage belongs to the family *Podoviridae*, six tail knob gene products 9 (gp9) form a hexameric tube structure with six flexible hydrophobic loops blocking one end of the tube before DNA ejection. During infection, the loops turn and come out, submerging with their hydrophobic α-helices into the cytoplasmic membrane of the cell to span the lipid layer and allow the release of the bacteriophage genomic DNA (Xu *et al*., 2016). Siphoviruses are also supposed to have tape-measure proteins, which have peptidoglycan-hydrolase activity and cause membrane fusion (Casjens and Molineux, 2012). With T4, a *Myoviridae* phage, the contractile tails are changed in its structure to push hollow inner tube into the bacterial cell. The baseplate at the distal end of the tail changes from a hexagonal to a star shape. This causes the sheath around the tail tube to contract and the tail tube to protrude from the baseplate and pierce the outer cell membrane and the the cell wall before reaching the inner cell membrane for subsequent viral DNA injection (Crowther *et al*., 1977; Leiman *et al*., 2004, 2010).

As soon as the injection of the genome into the bacterial cell, the latent phase has started. Phage latent period impacts on plaque enlargement in a manner that is analogous to the impact of phage adsorption (Clokie and Kropinski, 2009). The less time phage spends infecting bacteria, the more time they can spend as diffusing virions, and thus the potentially faster plaques can enlarge. An experimental characterization on the wild-type phage T7 and its derived mutants showed a negative correlation between the shorter latent period with plaque size, phage population, and propagation rate (Yin, 1993). During this period, phages are used cellular machinery to synthesis proteins and new phage particles. In the maturation period, new phage particles are synthesized. The phage DNA is encapsulated by preassembled protein, namely procapsids.

Finally, new phages are released from the host cell using enzymes, endolysin, and/or holing produced from phage genes. Endolysin is a lytic enzyme which is responsible for degrading peptidoglycan layer of the host cell. Holin is the protein which forms pores in the cytoplasmic membrane of the host cell. These pores stimulate endolysins access to the peptidoglycan layer (Pelzek *et al*., 2013). An alternative method for promoting the lysis event is that an N-terminal transmembrane domain mediates the export of endolysin to the membrane (Xu *et al*., 2004). The degradation of peptidoglycan by the action of endolysins leads to lysis of the host cells. After that, releasing from the host cells progenies quickly infect the new cells and start new life cycles.

5.1.2. Lysogenic life cycle

In the lysogenic cycle, the lysogenic phage interacts reversibly with the host components, which does not lead to multiplication but allows the viral genome to

replicate (Figure 1.2). Lysogenic phages incorporate their DNAs into host genomes. In some cases, the phage genome can appear independently in the bacterial cytoplasm as a circular or linear plasmid (Ceyssens and Lavigne, 2010). In the case of the phage genome remains in the host cell, it can exist as either as a plasmid or integrated into the host chromosome as a prophage (Clokie and Kropinski, 2009). The prophage is replicated using host enzymes until it is triggered by chemical or physical stimulation (Brunner *et al*., 1969; Müller *et al*., 2012). Prophages consist of 10% to 20% of the bacterial genome. It contributes to the diversity of bacterial species (Cheetham and Katz, 1995; Casjens, 2003). They provide the infected bacteria with immunity against additional phages' infection by modifying the genome structure, transferring of virulence genes. The prophage genes can be induced and expressed virulence genes of pathogenic host (Paucker, 1977; Ceyssens and Lavigne, 2010; Pelzek *et al*., 2013). Evans *et al.* reported that *P. atrosepticum* prophages ECA41 and ECA29 improved the swimming motility of the bacterial host (2010). Prophage may contain toxic genes such as shiga, cholera and diphtheria toxins (Abedon and LeJeune, 2005). In another case, some phage genes can be transferred their host genes into their own genomes (Coetzee, 1966). This phenomenon also happens by lytic phages. A report of an accidental pack of bacterial DNA into their own capsid heads during the later phase of lytic cycle is clear evidence for this conclusion (Klumpp *et al*., 2008).

5.2. Phage therapy for human disease

Phage therapy was first applied very early after the first description of phage by d'Herelle (Chanishvili, 2012). The therapy study conducted at the Hospital des Enfants-Malades to confirm its safety before administering it to a 12-year-old boy. The patient's symptoms ceased after treatment of the phages, and the boy fully recovered within a few days. Afterward, phage was also applied to treating staphylocooccal skin disease. Phages were injected into and around surgical lesions, and the authors reported regression of the infection within 24 to 48 h. Many similarities studies were then experimented (Chanishvili, 2012). Encouraged by these early results, d'Herelle and collaborators continued studies of the therapeutiec use of phages for thousands of people having cholera and/or bubonic plaque in India (Chanishvili, 2012). During period of 1916 to 1930, d'Herelle and collaborators undertook a lot of expeditions to China, Laos, India, Vietnam, and African countries to control epidemics caused by cholera. According to the estimation of publications in this time, due to phage treatment, the mortality was reduced to 10% in India (Chanishvili, 2012). Later, because of the emergence of broad-spectrum antibiotics, phage therapy was less concerned, and the virus-related researches were directed into rather theoretical aspects in Western Europe. Nevertheless, in Georgia and Poland, phage therapeutical treatments were continued and widely practiced (Chanishvili, 2012).

In addition, several companies began to be interested in commercial products of phages against various bacterial pathogens. The first commercial product was produced by the d'Herelle's laboratory with five phage preparation against various bacterial

diseases (Sulakvelidze *et al*., 2001; Chanishvili, 2012). Afterward, the Eli Lilly Company produced seven phage products for human use, which combat for Staphylococci, Streptococci, *E. coli*, and other pathogenic bacteria (Sulakvelidze *et al*., 2001). Nowadays, several emergency cases infected by ten pathogens and eleven infection types were reported as the successful cures. These case studies were developed to PhageBank therapy (http://www.aphage.com/case-studies/).

5.3. Phage therapy for plant disease

The population of the world is expected to reach 9.6 billion by 2050 and this raising alarm of food supply demands (https://news.un.org/en/story/2013/12/456912#.Vvxj0uIrLIU). It might require an increase in crop supply as much as 80-110% (Ray *et al*., 2013). Therefore, it is important to reduce crop diseases. It has been estimated that at least 10% of global food production is lost by phytopathology (Strange and Scott, 2005). Among pathogens, over 200 bacteria cause plant diseases, which belong to the genera of *Pseudomonas, Ralstonia, Agrobacterium*, *Xanthomonas*, *Erwinia*, *Xylella*, *Pectobacterium*, and *Dickeya* mainly (Mansfield *et al*., 2012).

5.3.1. History of phages and phage therapy against to plant diseases

Phage for plant disease was first reported by Mallmann and Hemstreet that the filtrate of decomposing cabbage was able to inhibit the cabbage rot organism *Xanthomonas campestris* pv. *Campestris* growth (Mallmann *et al*., 1924). A year later, an experiment with phage to prevent soft rot causing by *P. atrosepticum* and *P. carotovorum* subsp. *carotovorum* on slices of potato tuber and carrot was performed (Coons, 1925). Ten years later, Thomas *et al.* made the first trial in the field for controlling Stewart's wilt disease caused by *Pantoea stewartii* (Thomas, 1935). However, phage research for plant disease has neglected due to being poor at reliability and efficacy (Okabe and Goto, 1963). The appearance of antibiotics and multiple antibiotic resistant bacteria has pushed several researches for phage biocontrol on significant plant pathogens. After over 30 years from the earliest report on using phage for controlling phytopathogens, Stonier *et al.* showed that fewer than ten phage particles present at the beginning of a 21-hour induction period were able to completely inhibit tumor induced by *Agrobacterium tumefaciens* (1967). Just 2 years later, treatment with phage to peach foliage caused by *Xanthomonas campestris* pv. *pruni* (Xcp) showed that 22% of leaves infected with threat pathogens in one hour treatment with phage and 58% for control plants. Foliage was treated 24 h prior to inoculation with Xcp resulted in 29% infected leaves (Coetzee, 1966). After that, there were two reports on phage therapy for tomato disease caused by *A. tumefaciens* and rice disease caused by *Xanthomonas oryzae* (Boyd *et al*., 1971; Kuo *et al*., 1971).

5.3.2. Recent research of phage therapy for plant diseases

Since the years of 2000s, phage therapy has been found to be effective for the

inhibition of variation of plant bacteria such as *Xanthomonas* spp. (bacterial spot of tomato, peach, geranium and citrus, onion blight, walnut blight, and citrus canker) (Balogh *et al*., 2003, 2008; Byrne *et al*., 2005; Balogh, 2006; Lang *et al*., 2007; Ahern *et al*., 2014), *P. carotovorum* (Soft rot of potato) (Lim *et al*., 2013) and *Ralstonia solanacearum* (bacterial wilt of tomato, potato, tobacco, and eggplant) (Yamada, 2013).

Ideally phages for biocontrol should be lytic, broad host range, stable in chemicals for pesticide formulation and disadvantageous physical condition. Furthermore, for applying them to control epidemic, phages should be able to lyse the host quickly. It requires phages to have high burst size. However, phage ϕRSL1 of *R. solanacearum* showed great biocontrol effect despite it is not highly lytic (Fujiwara *et al*., 2011). Phage ϕRSL1 also kept inhibit the bacterial growth over long period (140 h) (Fujiwara *et al*., 2011). Pretreatment with tomato seedlings, the treated tomato seedlings showed no symptom of wilting during the experimental time, whereas all untreated plants had wilted by 18 days post-infection.

In another case study, Balogh *et al.* showed the results of three phages of *Xanthomonas citri* pv. *citri* exhibiting lytic activity in plate assay. Two of the three phages were unable to lyse their host bacterium on grapefruit leaves (2006). The data also described that they also did not suppress citrus canker in greenhouse trials. The most updated phages for plant disease control were likely "jumbo phages", which are tailed phages with genome larger than 200 kb. Two Asian jumbo phages, ϕRSL2 and ϕRSF1, infecting *R. solanacearum* were isolated in Thailand and Japan, respectively (Bhunchoth *et al*., 2016). These two phages can inhibit broad host range comparing to various *R. solanacearum* strains on tobacco, tomato, eggplant, and potato. *X. citri* jumbo phage XacN1 showed wide host range to various *X. citri* strains (Yoshikawa *et al*., 2018). In the same year, eight *Dickey* phages, which belong to two families *Myoviridae* and *Podoviridae*, were isolated and tested with six different *Dickeya* species (Day *et al*., 2018). *Erwinia amylovora* is a phytopathogen from the *Erwiniaceae* family for agricultural severe disease "fire blight". This bacterium was inhibited by eight novel phages. Although these phages were closely related to *Pseudomonas* and *Ralstonia* phages rather than *Enterobacteriales* phages, they showed the inhibition ability to *Erwinia* only (Sharma *et al*., 2019).

5.3.3. Challenges for phage therapy

Using phage therapy has been facing to difficulties because the phage attaches to its host before it is destroyed (Goodridge, 2004) and the probability of phage – bacterium contact depends on many factors: initial phage titer, rate of virion decay, phage's burst size, concentration and location of target bacteria, and presence of equivalent water as a solvent for phage diffusion (Gill and Abedon, 2003). Furthermore, other conditions may affect the effectiveness, such as the timing of phage application, relative fitness of phage resistant mutant, and the environment (Gill and Abedon, 2003). Several other factors can limit the efficiency of disease inhibition in the rhizosphere. The

rate of diffusion through the heterogeneous soil matrix is low and changes as the function of available free water (Gill and Abedon, 2003). In addition, biofilms or soil clay with low pH also can trap and inactivate phage particles. Physical refuges can protect bacteria from coming into contact with phages. Due to the low possibility of phage diffusion and high rates of phage inactivation, only a low number of viable phages is available to lyse target bacteria. Another issue is the requirement of a high titrate of phage and bacterium to start the chain reaction of bacterial lysis (Sykes *et al*., 1981; Storey, 2001; Gill and Abedon, 2003).

Out of the mentioned above factors, sunlight also facilitates the inactivation process of phage. An experiment in laboratory with phage CB 38phi and CB 7phi showed a significant decrease of plaque forming unit after 30 and 60 h (Wommack *et al*., 1996). In the field, sunlight UV was also evaluated for detrimental effects on phage survival. The intensity of UV irradiation positively correlated with phage titrate decrease when the test was taken place with *Xanthomonas* phage ϕXacm 2004-16 and ϕXV3-16 (Iriarte *et al*., 2007).

Also, to discover bacterial diversity is another challenge for phage therapy. The development of phage resistance can be occurred in various modes such as preventing phage attachment, blocking phage DNA entry, restriction or modification systems, abortive infection, and assembly interference (Seed, 2015). *E. coli* strains produce K1 polysaccharide capsule. This capsule blocks the infection by T7, a phage recognizes lipopolysaccharide as the primary receptor (Scholl *et al*., 2005). Otherwise, expression of the $ltp_{\text{TP-J34}}$ gene of temperate *Streptococcus thermophilus* phage TP-J23 is another evidence for the diversity of phage resistance mode (Bebeacua *et al*., 2013). The gene interferes with phage infection at the period of DNA injection into the host cell by targeting the phage's tape measure protein.

6. Research objective

The purpose of this study is to determine the potential of phages as natural antibacterial agents for biocontrol of soft rot. Up to now, the leading causes of soft rot are defined as *Pectobacterium* and *Dickeya.* However, some researchers reported that *Enterobacter* species are causes for the disease on vegetables and fruits in subtropical and tropical areas. Also, diversity and genome information about these *Enterobacter* species causing soft rot have not studied. Furthermore, the usage of phage as an alternative method for chemical pesticides against soft rot *Enterobacter* has not been examined yet. Thus, the aim of this study is to evaluate the possibility to create the phage therapy for the disease on vegetables and fruits in the subtropical and tropical areas.

The content in each chapter of this thesis is as follows.

Chapter 1 describes a research background about soft rot and phage therapy for biological control.

Chapter 2 describes the identification and characterization of isolates from the soft rot in Vietnam.

Chapter 3 describes whole genome sequence analysis of *Enterobacter* sp. M4-VN isolated from potatoes with soft rot and its phylogenetic position in genus *Enterobacter*.

Chapter 4 describes the identification and characterization of virulent phage EspM4VN to control *Enterobacter* sp. M4-VN isolated from soft rot. The result provides basic information for phage therapy against soft rot caused by *Enterobacter* species.

Chapter 5 describes the conclusion of this thesis.

Table 1.1. *Enterobacter* species and recently identified hosts

Bacterial species	Host plant species
Enterobacter cloacae subsp. *dissolvens*	*Mammillaria mystax* (Reyes-García *et al*., 2020)
Enterobacter cloacae	*Hylocereus undatus* (Soto *et al*., 2019) *Capsicum annuum* L. (García-González *et al*., 2018) *Amorphophallus konjac* (Wu *et al*., 2011) *Hylocereus* spp. (Masyahit *et al*. 2009) Onion (Nishijima, 1987; Schroeder *et al*., 2009)
Enterobacter sp.	*Solanum tuberosum* (Razanakoto *et al*., 2015)
Enterobacter cowanii	*Mabea fistulifera* (Furtado *et al*., 2012)
Enterobacter mori	*Morus alba* (Wang *et al*., 2010)
Enterobacter asburiae; *Enterobacter* sp.	*Morus alba*

Table 1.2. Taxonomy proposals proposing new taxa submitted to the ICTV Executive Committee in 2017

Family	Subfamily	Genus	Type species	No. of species in genus***
Ackermannviridae	*Aglimvirinae*	*Ag3virus*	*Shigella virus AG3*	1 (2)
Ackermannviridae	*Aglimvirinae*	*Limestonevirus*	*Dickeya virus Limestone*	1 (2)
Ackermannviridae	*Cvivirinae*	*Cba120virus*	*Escherichia virus CBA120*	4 (9)
Ackermannviridae	*Cvivirinae*	*Vi1virus**	*Salmonella virus ViI*	-5
Ackermannviridae	unassigned	unassigned	*Erwinia virus Ea2809, Serratia virus MAM1, Serratia virus IME250, Klebsiella virus 0507KN21*	4
*Myoviridae**		*Arvunavirus*	*Arthrobacter virus ArV1*	2
*Myoviridae**		*Eah2virus*	*Erwinia virus EaH2*	2
*Myoviridae**		*Machinavirus*	*Erwinia virus Machina*	1
*Myoviridae**		*Ntreusvirus*	*Salmonella virus SPN3US*	1
*Myoviridae**		*Svunavirus*	*Geobacillus virus GBSV1*	2
*Myoviridae**	*Ampvirinae*	*Chippewavirus*	*Arthrobacter virus BarretLemon*	1
*Myoviridae**	*Ampvirinae*	*Jawnskivirus*	*Arthrobacter virus Jawnski*	2
*Myoviridae**	*Ampvirinae*	*Sonnyvirus*	*Arthrobacter virus Sonny*	3
*Podoviridae**		*Dfl12virus*	*Dinoroseobacter virus DFL12phi1*	1
*Podoviridae**		*Jwalphavirus*	*Achromobacter virus JWAlpha*	2
*Podoviridae**		*P22virus**	*Salmonella virus P22*	1 (5)
*Podoviridae**		*Sp58virus*	*Salmonella virus SP058*	3
Portogloboviridae		*Alphaportoglobovirus*	*Sulfolobus alphaportoglobovirus 1*	1
*Siphoviridae**		*Anatolevirus*	*Propionibacterium virus Anatole*	2
*Siphoviridae**		*Attisvirus*	*Gordonia virus Attis*	1
*Siphoviridae**		*Doucettevirus*	*Propionibacterium virus Doucette*	4

*Siphoviridae**		*Hk97virus*	*Escherichia virus HK97**	9 (11)
*Siphoviridae**		*Lambdavirus**	*Escherichia virus Lambda*	3 (4)
*Siphoviridae**		*Pfr1virus*	*Propionibacterium virus PFR1*	1
*Siphoviridae**		*Tp84virus*	*Geobacillus virus TP84*	1
*Siphoviridae**		*Trigintaduovirus*	*Mycobacterium virus 32HC*	1
*Siphoviridae**		*Wizardvirus*	*Gordonia virus Wizard*	2
*Siphoviridae**	*Chebruvirinae*	*Brujitavirus*	*Mycobacterium virus Brujita*	-2
*Siphoviridae**	*Chebruvirinae*	*Che9cvirus**	*Mycobacterium virus Che9c*	1 (2)
*Siphoviridae**	*Dclasvirinae*	*Hawkeyevirus*	*Mycobacterium virus Hawkeye*	1
*Siphoviridae**	*Dclasvirinae*	*Plotvirus*	*Mycobacterium virus PLot*	1
*Siphoviridae**	*Mccleskeyvirinae*	*Lmd1virus*	*Leuconostoc virus Lmd1*	6
*Siphoviridae**	*Mccleskeyvirinae*	*Una4virus*	*Leuconostoc virus 1A4*	6
*Siphoviridae**	*Nclasvirinae*	*Buttersvirus*	*Mycobacterium virus Butters*	2
*Siphoviridae**	*Nclasvirinae*	*Charlievirus*	*Mycobacterium virus Charlie*	2 (3)
*Siphoviridae**	*Nclasvirinae*	*Redivirus*	*Mycobacterium virus Redi*	3 (4)
*Siphoviridae**	*Nymbaxtervirinae*	*Baxtervirus*	*Gordonia virus BaxterFox*	2
*Siphoviridae**	*Nymbaxtervirinae*	*Nymphadoravirus*	*Gordonia virus Nymphadora*	3
*Cystoviridae**		*Cystovirus**	*Pseudomonas virus phi6*	6 (7)
*Tectiviridae**		*Alphatectivirus***	*Pseudomonas virus PRD1*	1 (2)
*Tectiviridae**		*Betatectivirus*	*Bacillus virus Bam35*	2 (4)

*taxon established, **previously known as *Tectivirus,* ***Number in parenthesis indicates the total number of viral species in this genus

Table 1.3. Examples of phage receptors (*to be continued*)

Bacterio-phage	Family	Genus/ group	Host	Primary receptor	Secondary receptor	Note
T1	S	T1-like	*E. coli*	?	FhuA (requires TonB)	it is possible to obtain mutants of both a host not needing the energy transduced by TonB and a phage not depending on the FhuA protein
T4	M	T4-like	*E. coli, Shigella*	OmpC	LPS core	inner membrane phospholipids can be a tertiary receptor; the phage interacts with mucin of the mucus via the Hoc protein
T5	S	T5-like	*E. coli*	LPS O-antigen (polyman-nose) – optionally	FhuA	
BF23	S	T5-like	*E. coli*	LPS?	BtuB	
λ	S	lambdoids (λ-like)	*E. coli*	OmpC	LamB	
P22	P	lambdoids (P22-like)	*E. coli*	LPS O-antigen	LPS?	reversible and irreversible adsorption; depends on LPS
Sf6	P	?	*Shigella flexneri*	LPS	OmpA, OmpC	2,4 OmpA loops

Table 1.3. Examples of some bacteriophage receptors (*to be continued*)

Sf6	P	?	*Shigella flexneri*	LPS	OmpA, OmpC	2,4 OmpA loops
N4	P	N4-like	*E. coli*	?	NfrA	low representation of receptor; interaction via noncontractile tail protein (gp65)
G7C	P	N4-like	*E. coli* 4s	LPS O-antigen O22-like	unknown (OmpA and ?)	LPS acetylation is necessary; receptor-binding protein of the phage – enzyme deacetylase
Alt63	P	N4-like	*E. coli* 4s	LPS O-antigen	unknown (OmpA and ?)	the relative of G7C; depolymerase activity of receptor-binding protein instead of deacetylase activity abolishes the acute dependence on acetylation
CP81 and related phages	M	?	*Campylobacter jejuni* NCTC12658	exopolysaccharide; modification of the MeOPN type is important for some phages	?	phage F341 (CP81-like) can infect a host mutant without the exopolysaccharide, being adsorbed to flagella
CP220 and related phages	M	?	*Campylobacter jejuni* NCTC12658	motile flagellum	?	motility is necessary
NCTC12673			*Campylobacter jejuni*	glycosylated flagellin	?	

Table 1.3. Examples of some bacteriophage receptors

SPC35	S	T5-like	*Salmonella enterica* serovar *Typhimurium*	LPS O-antigen	BtuB	glycosylation of O-antigen blocks phage adsorption, but the presence of BtuB is sufficient in the absence of O-antigen
SPN10H (and 6 other isolates)	S	T5-like	*S. enterica* serovar *Typhimurium*	LPS?	BtuB	
SPN2T (and 10 other isolates)	S	?	*S. enterica* serovar *Typhimurium*	flagellum	?	
SPN1S (and 6 other isolates)	P	?	*S. enterica* serovar *Typhimurium*	LPS	?	
phiA1122	P	T7-like	*Yersinia pestis*, *Y. pseudotuberculosis*	?	Hep/Glc-Kdo/Ko regions of LPS core	
phiCb13 and phiCbK	S	?	*Caulobacter crescentus*	flagellum	pili portal	fibrillar structures located on the phage head interact with the motile flagellum; the virion is "screwed" on the flagellum and transported to the cell pole, where it interacts with pilus portal

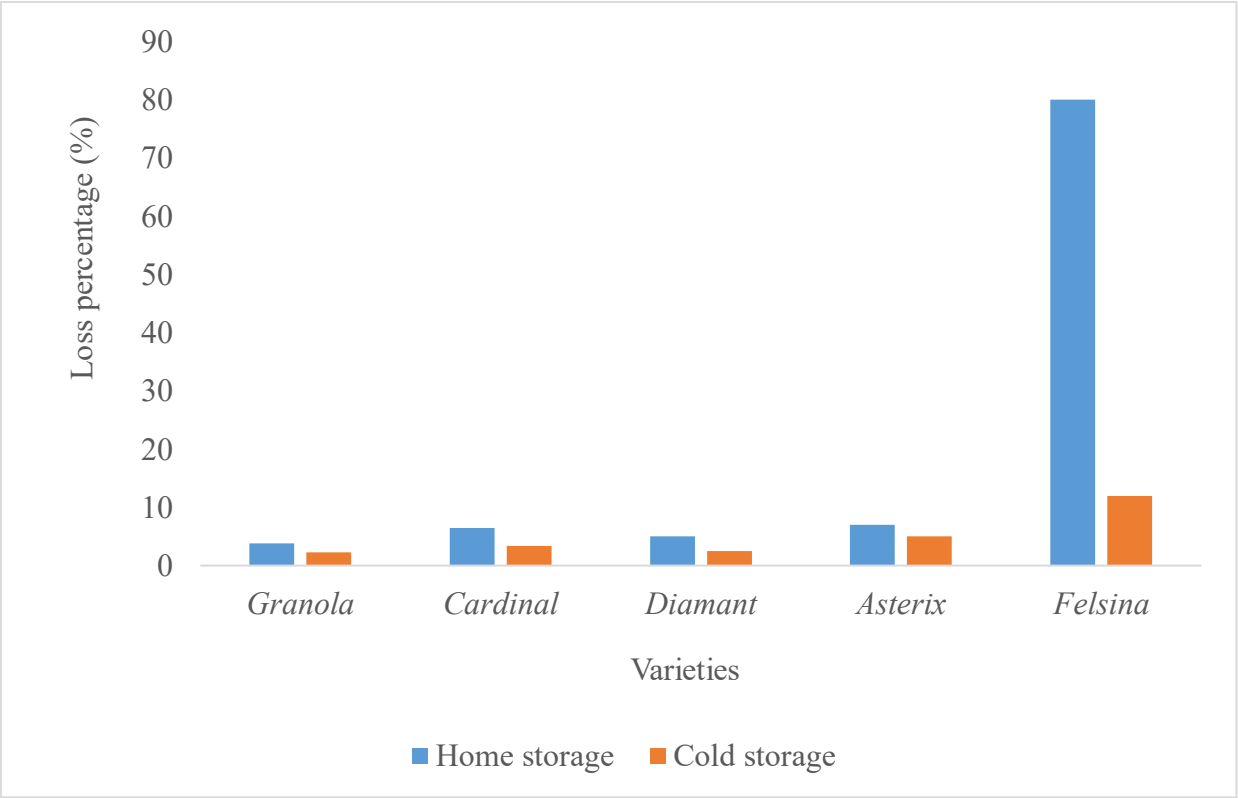


Figure 1.1. Loss of different potato varieties due to soft rot at farmer's level in Rangpur district, Bangladesh in two year 2009 and 2010

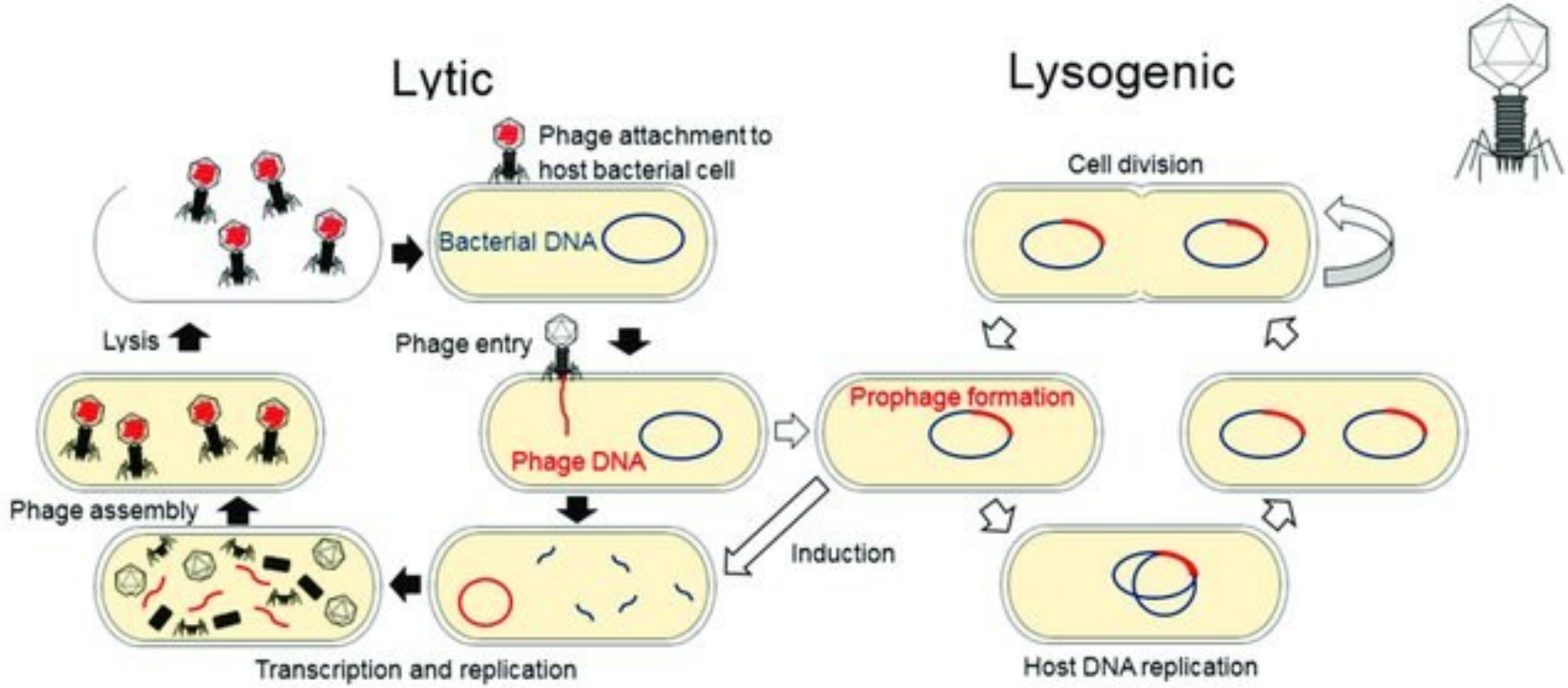


Figure 1.2. Bacteriophage life cycles (Batinovic *et al*., 2019).

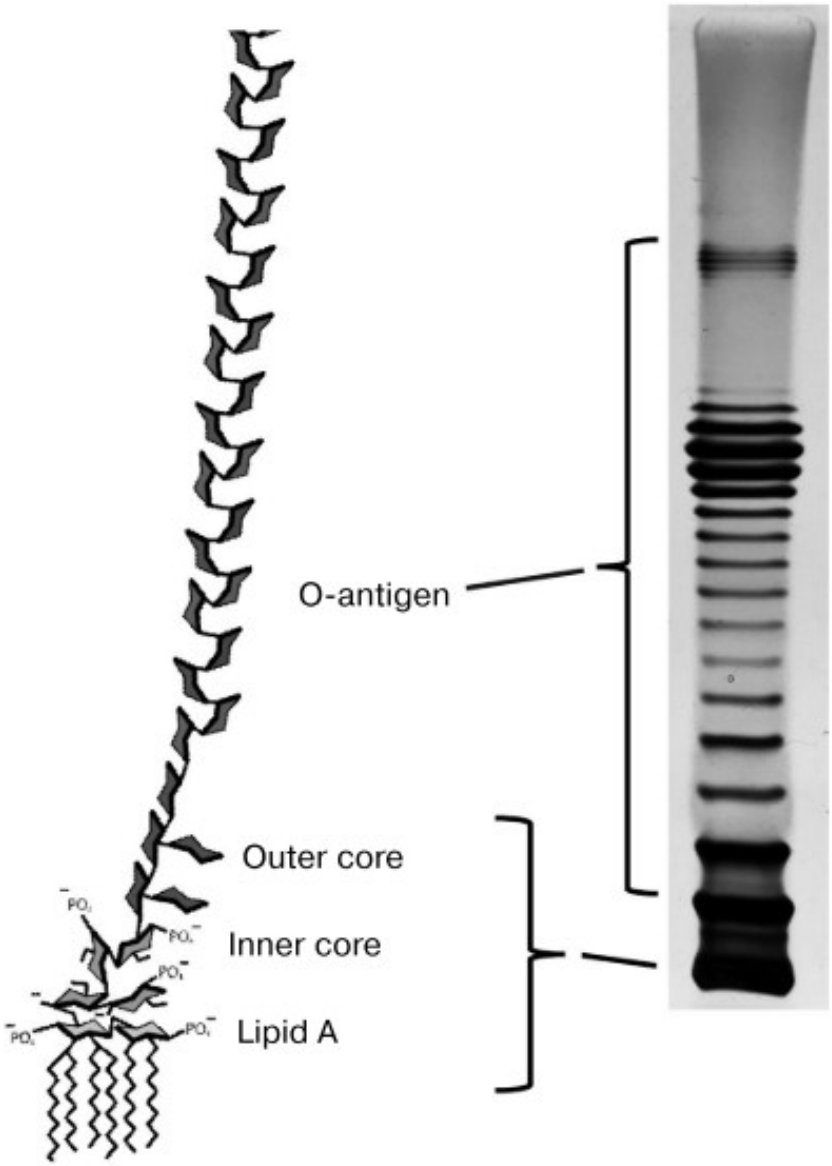


Figure 1.3. Structure of lipopolysaccharide molecules (Letarov and Kulikov, 2017).

Chapter 2. Identification and characterization of isolates from the soft rot in Vietnam

1. Introduction

Bacterial soft rot is well known as a disease of fleshy plant such as potato tubers, fruits, and flower bulbs. Bacteria which cause soft rot usually perforate the plant body through wounds and multiply in the intercellular spaces. Once bacteria have entered to plant cells at suitable environmental conditions, bacterial cells will multiply quickly and produce toxins by the pectolytic enzymes and other metabolic products. Potato soft rot can arise in the field or as a post-harvest biodeterioration of a stored crop. The latter can be a more serious problem due to the long storage times between harvest and processing. The infected crops will eventually lose their production ability or die (Barnes, 1979).

This infection is caused by many kinds of bacteria, as mentioned in chapter 1. However, *Pectobacterium* and *Dickeya* species were identified as the main causes. These genera were previously belonged to *Enterobacteriaceae* family and reordered to *Pectobacteriaceae* in 2016. From this year, soft rot caused by *Pectobacterium* and *Dickeya* is known as SRP. In the early 21st century, soft rot by *D. solani* caused a considerable loss of seed potato in European countries (Toth *et al*., 2011). Another species of *Dickeya, Dickeya. dianthicola* was also determined to be the reason for disease outbreak in potatoes in North America and other states of America in 2014 and 2016, respectively. This species is distributed largely in Europe, Asia and Oceania, and North America (https://gd.eppo.int/taxon/ERWICD/distribution). Additionally, *Pectobacterium* species were also defined as a contribution to the loss of plant production during the vegetative growing and post-harvested stage. Pcc, which is the main cause of soft rot in sweet potato, was distributed largely in Africa, Asia, Europe, North America, South America and Oceania (https://www.cabi.org/isc/datasheet/21913). While in Poland, *P. wasabie* was identified as one of the most important soft rot pathogens.

Interestingly, another case in one of the tropical areas, Malaysia, *E. cloacae* was the earliest revealed as a causative pathogen of soft rot on dragon fruit (Masyahit *et al*., 2009). This species was also reported to be a pathogen on chili pepper (*Capsicum annuum* L.) and dragon fruit (*Hylocereus* spp.) (García-González et al., 2018; Soto et al., 2019). Another *Enterobacter* species, *Enterobacter asburiae*, has been reported as the reason for soft rot symptom on Konjac (*Amorphophallus konjac*) in China (Wang *et al*., 2010; Wu *et al*., 2011). Both of these species belong to the family *Enterobacteriaceae.* Up to date, SRE has not yet been widely understood. Especially in Vietnam, the identification of bacterial soft rot species was mostly based on morphological research. To the best of my knowledge, there was no study to characterize bacteria causing soft rot symptom disease on the above crops. This study aims to identify and characterize bacterial strains in soft rot of plants in Vietnam in various aspects, including morphology, physiology, and genetic

analysis.

2. Methodology

2.1. Bacterial strains

Soft rot symptom samples from soft-rot infected cabbage, potato, and dragon fruit were collected in Hanoi, Vietnam. Infected samples were peeled and slid before sterilizing the surface by using ethanol 70%. Bacterial cultures were isolated from the margin of infected tuber samples using sterile loops before dissolving in sterilized distilled water, followed by streaking on LB agar plates. The plates were incubated overnight at 37°C. Single colonies were then picked up and purified by re-streaking on LB plates for three times.

Table 2.1. LB agar

Reagents	Final concentration (% w/v)
NaCl	1
Tryptone	1
Yeast extract	0.5
Agar	2

LB agar medium was prepared by combining the first three reagents and shaking until the solutes have dissolved. After that, the pH was adjusted to 7.0 if it was necessary. The agar was added into the mixture and autoclaved for 20 min at 121°C. After autoclaving, the medium was poured into petri dishes.

2.2. Microscopy

To conduct gram staining, colonies were flood air-dried, heat-fixed with a smear of cells for 1 min with crystal violet staining reagent (Jones, 1981). The slides were washed in a gentle and indirect stream of sterilized water for 2 s before flooding with gram's iodine. After 1 min, the slides were decolored gently with autoclaved water twice.

For the morphological study, 1µL of the overnight culture was negatively stained with 3 µL of 0.125% phosphotungstic acid before placing on carbon film-coated copper grids hydrophilized using an ion bombarder (PIB-10, Vacuum Device). The ion bombarder was set 50 mA for 30 s. Cells were examined by transmission electron microscopy (Hitachi, model HT7700) at an acceleration voltage of 80 kV. The sizes of bacterial cells were determined from at least three times.

2.3. Growth curve of bacteria

Twenty mL fresh LB liquid media were added 10µL overnight culture, and bacteria growth were recorded using Real-Time Cell Growth Logger (Biosan Corp., Riga, Latvia). The cultures were incubated at 30 or 37°C at 2000 rpm. Optical densities (600

nm) were recorded for an interval of 30 min.

2.4. 16S rRNA gene sequence analysis

LB cultures of bacteria incubated at 37°C with overnight shaking were used for extraction and purification of genome DNAs. Chromosomal DNA of isolates was isolated using MightyPrep reagent for DNA (TaKaRa Bio Inc., Shiga, Japan). The 16S rRNA gene was amplified using polymerase chain reaction (PCR) with Tks Gflex DNA Polymerase (TaKaRa Bio Inc.) and primers 27f (5'-AGAGTTTGATCCTGGCTCAG-3') and 1492r (5'-ACGGCTACCTTGTTACGACCT-3').

Table 2.2. PCR mixture

Reagents	Amount/Concentration
Tks Gflex DNA Polymerase (1.25 units/μL)	1μL
2X Gflex PCR Buffer (Mg2＋, dNTP plus)	25 μL
Template DNA	< 500 ng
Primer 27f	0.2 - 0.3 μM (final conc.)
Primer 1492r	0.2 - 0.3 μM (final conc.)
Sterilized distilled water	up to 50 μL

In the PCR reaction, the target gene was initially denatured at 94°C for 1 minnute before repeating 30 cycles of 10 s denaturation at 94°C, 15 s annealing at 55°C and followed by an extension for 1 minnute at 68°C.

The PCR products were applied to electrophoresis in 2.0% agarose gels in TAE buffer (40 mM Tris-acetate; 1 mM EDTA) at100V for 50 minnutes, and were then recovered using a Fast Gene Gel/PCR Extraction Kit (NIPPON Genetics Co. Ltd., Tokyo, Japan).

Table 2.3. 50X TAE stock solution (per 1 L)

Reagents	Amount/Concentration
Tris base	242 g
Acetic acid (glacial)	57.1 mL
0.5 M EDTA (pH 8.0)	100 mL

Tris-base was firstly entirely dissolved in 500 mL of H2O before mixing with acetic acid and EDTA. The mixture was then fill up to 1 L of distilled H_2O. The 1X working solution consist with 40 mM Tris-base and 1 mM EDTA.

pTA2 vector was used to attach the purified target DNAs with using a TArget clone-plus-kit (TOYOBO, Osaka, Japan). Resulted recombinant plasmid was transformed into *E. coli* DH5α, followed by spreading onto LB agar plates added IPTG

(40 µg/mL) and X-Gal (150 µg/mL). The plasmid DNAs from the transformants were then extracted by Xprep Plasmid DNA Mini Kit (AS ONE Corporation, Osaka, Japan).

The nucleotide of inserted DNA in the plasmid was sequenced using a BigDye Terminator v3.1 Cycle Sequencing Kit (Life Technologies, Carlsbad, CA, USA) with M13 Primer M3 (5'-GTAAAACGACGGCCAGT-3') or M13 primer RV (5'-CAGGAAACAGCTATGAC-3').

PCR reaction mixture was prepared as follow:

Table 2.4. PCR mixture was prepared as follow:

Reagents	Amount/Concentration
BigDye Terminator	0.5 µL
2X Big Dye Terminator Buffer	1.75 µL
DNA Template	< 500 ng
Primer	0.16 µM (final conc.)
Sterilized distilled water	up to 10 µL

Sequences were identified using an Applied Biosystems Gene Analyzer 3130xl (Life Technologies). In the PCR reaction, the sample was initially denatured at 95°C for 1 min before repeating 30 cycles of 10-second denaturation at 94°C, 15-second annealing at 42°C, and followed by a final extension for 4 min at 72°C. Obtained 16S rDNA sequence was then searched for homology in the Nucleotide Sequence Data Library using the BLAST program (https://blast.ncbi.nlm.nih.gov/Blast.cgi). Phylogenetic tree was then constructed using the neighbor-joining method with the program of GENETYX software (GENETYX, Tokyo, Japan).

2.5. Biochemical analysis

Biochemical tests were performed using API 20E kit (bioMérieux, Marcy-l'Étoile, France) with the manufacturer's instruction. Positive and negative tests were scored after observation of the changing color of substrates. Three-time replication was taken, and analytical index records were defined after 23 h and 48 h of incubation at 30°C. Identification was carried out with comparison with other previous publications. The identification of species based on physiology was then made by using API test finder platform (https://bacdive.dsmz.de/api-test-finder).

3. Results

3.1. Growth curve

The seven isolates kept growing in a range of temperatures from 15 to 45°C (Data not shown) and completely deactivated at 10 and 50°C. Figure 2.1 showed strains B5, BC and M4-VN had the same lag phase of 2.5 and 1.5 h at 30°C and 37°C, respectively and the same log phase of 7 h at both thermal conditions. While, strain B3

had lag phase of 2 h at 30°C and 1.5 h at 37°C, log phase of 6 h. The remains depicted another pattern while their lag phase time differentiation between 37 and 30°C was bigger than the two strains'. At 37°C, strains KT and TL3 got an early log phase at OD of 0.1 after 2 h, but at 30°C, they did not reach to the same OD until the 4th hour. But for TL5, the lag phase appeared in the first 30 min, which is different from the others. Notably, TL3 reached to stationary phase with OD of 1 after 7 h at 37°C, yet, the OD was dropped to 0.9 at the 8th hour.

3.2. Morphology

Gram staining showed that all strains were gram-negative. Colonies grown on LB agar were slightly yellow, smooth, translucent and convex, adhering to the agar surface. Electron micrograph of seven strains showed that most strains have rod-shaped with various length ranges from 1.1 μm to 2.9 μm, width ranges from 0.6 μm to 1.2 μm (Figure 2.2). Namely, at 37°C, the strains B3, B5, BC, M4-VN, KT, TL3 and TL5 sized 2.6 × 1.0 μm, 1.6 × 1 μm, 1.2 × 0.8 μm, 1.3 × 0.8 μm, 2.3 × 1.3 μm, and 1.4 × 0.9 μm, respectively.

3.3. Classification of seven isolates based on 16S rRNA gene sequencing analysis and biochemical test

The results of sequencing analysis showed that 16S rDNA sequence of strain B5 and BC had no difference. Sequences of the five strains B3, B5, BC, M4-VN and KT shared similarity to *Enterobacter* species with 99% of identity. On the other hand, the analysis of strains TL3 and TL5, revealed a 99% identity to *Pantoea dispersa* and *Acinetobacter baumannii*, respectively.

The fermentative characterization of the seven strains were accomplished using API 20E kit (Table 2.5). The outcomes showed that all of the strains were indole, hydro sulfur, L-lysine, gelatin and oxidase negative. There were similarities between four strains B3, B5, M4-VN and KT in positive reactions with 2-nitrophenyl-β-D-galactopyranodie, L-arginine, L-orthine, trisodium citrate, L-tryptophane. Acid was produced from the fermentation of D-glucose, D-manitol, D-sacharose, D-melibiose, amygdaline, L-arabinose. These strains also produced nitrogen dioxide. Strains B3, B5, BC, M4-VN and KT were Voges – Proskauer negative. Strain BC was slightly different from the other four strains. It had positive results with L-tryptophane, D-glucose, D-manitol, inositol, rhamnose, melibiose, amygdaline, L-arabinose, β-galactosidase, arginine dihydrolase, lysine decarboxylase, orthinine decarboxylase,

The remaining strains, TL3 and TL5, shared numerous similar characteristics. β-galactosidase and tryptophan deaminase were produced from ortho-nitrophenyl-β-D-galactopyranoside (OPNG) and L-tryptophane. Trisodium citrate, sodium pyruvate were utilized and acid were generated from D-glucose, D-manitol, inositol, sorbitol, D-sacharose, amygdaline, arabinose. Both strains digested nitrate and formed nitrogen gas. While strain TL3 produced acid with L-rhamnose and D-melibiose, strain TL5 had negative reactions.

3.4. Accession numbers

The 16S rDNA sequences of the isolates have been deposited in the DDBJ under accession numbers to LC415135, LC415612, LC416590, LC423530, LC498100, LC498101, and LC498102.

4. Discussion

In this study, the 16S rDNA sequences of the strains B3, B5, BC, M4-VN and KT shared 99% identification to *E. asburiae, Enterobacter tabaci, Enterobacter roggenkanpii, E. ludwigii* and *E. kobei.* While the gene sequence is very useful for bacteria classification but it has low phylogenetic ability (Janda and Abbott, 2007), the phylogenetic tree showed the two strains BC and KT to be closed to *Enterobacter* species, strains TL3 and TL5 to be related closely to *P. dispersa* and *A. baumannii*, respectively. The remaining strains are closed to various *Enterobacter* species (Figure 2.3). Strain TL3 and the closely related strain, *P. dispersa* DSM 32899 is placed in a different cluster from *Pantoea cypripedii*, *P. dispersa* DSM30073, *P. dispersa* AS18, and *P. dispersa* LMG2603. Furthermore, the biochemical characteristics of the strain TL3 are more similar to *P. dispersa* than *P. cypripedii* due to D-melibiose reaction (Table 2.5). *P. dispersa* differs from *P. cypripedii* by utilization of melibiose (Brady *et al*., 2010). While *P. dispersa* gives negative reaction with melibiose, *P. cypripedii* shows positive result. Thus, strain TL3 is basically distinguishable and classified into *P. dispersa*. Phylogenetic tree showed TL5 to be related closely to only *A. baumannii*. Thus, this evidence suggests that the strain TL5 belongs to *A. baumannii*. The remaining isolates are also differentiated by physiological tests. B3, B5, BC, M4-VN, and KT shared the most important characteristic: negative to indole, which has 100% positive reaction by *E. cloacae* and completely negative reaction by *E. asburiae* (Washington *et al*., 1969; Brenner *et al*., 1986). However, strains B3, B5 and KT are negative to L-rhamnose, which is utilized by all *Enterobacter* strains except *E. asburiae* (Ristuccia and Cunha, 1985). In addition, while most *Enterobacter* species with the exception of *E. kobei* give a positive Voges-Proskauer reaction, the four strains BC7, B3, KT and M4-VN are negative to this test (Ristuccia and Cunha, 1985; Kosako *et al*., 1996). Using the API test finder, the results did not show any matched results. This suggests that the species might be a biovar or novel species of the genus *Enterobacter*.

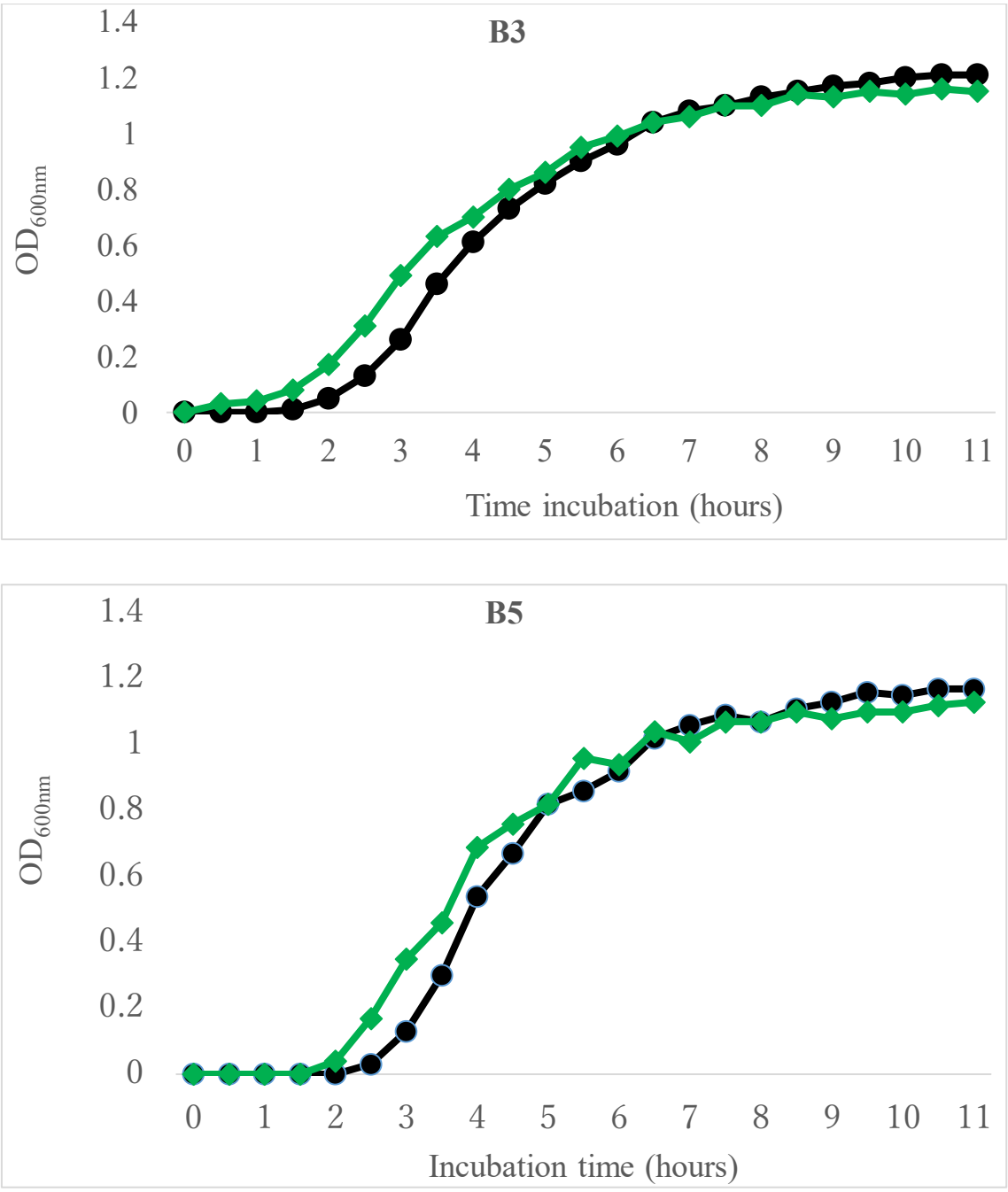
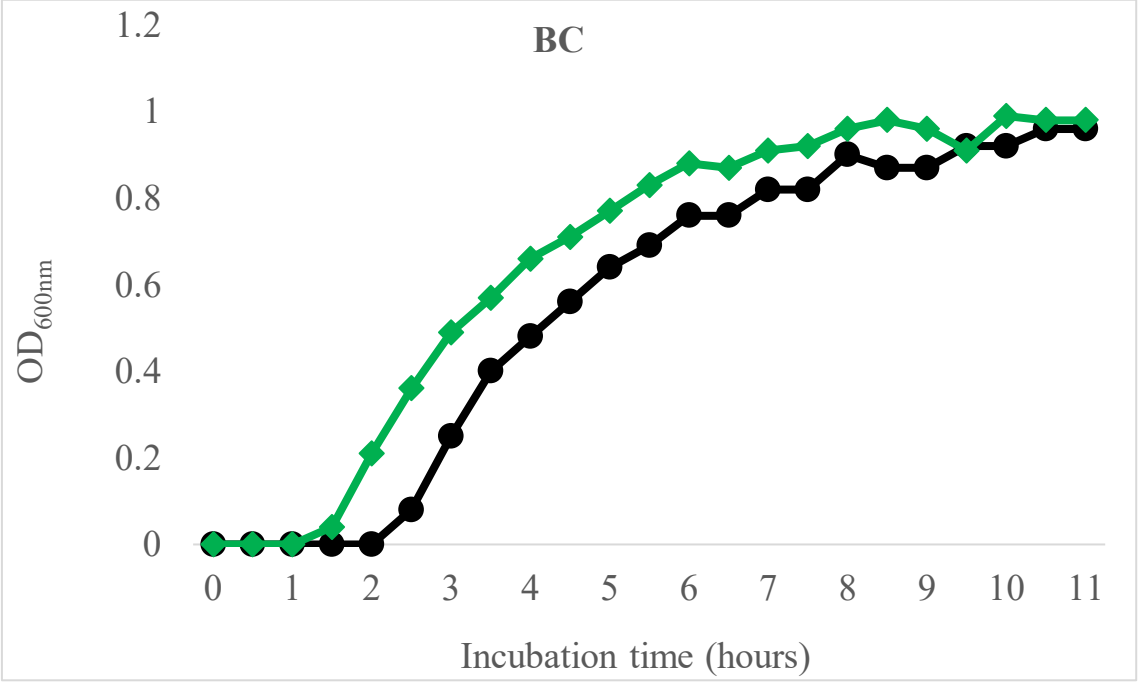


Figure 2.1. Growth curve of the seven strains (B3, B5, BC, M4-VN, KT, TL3 and TL5) at 30 and 37°C (black and green lines respectively). (*To be continued)*



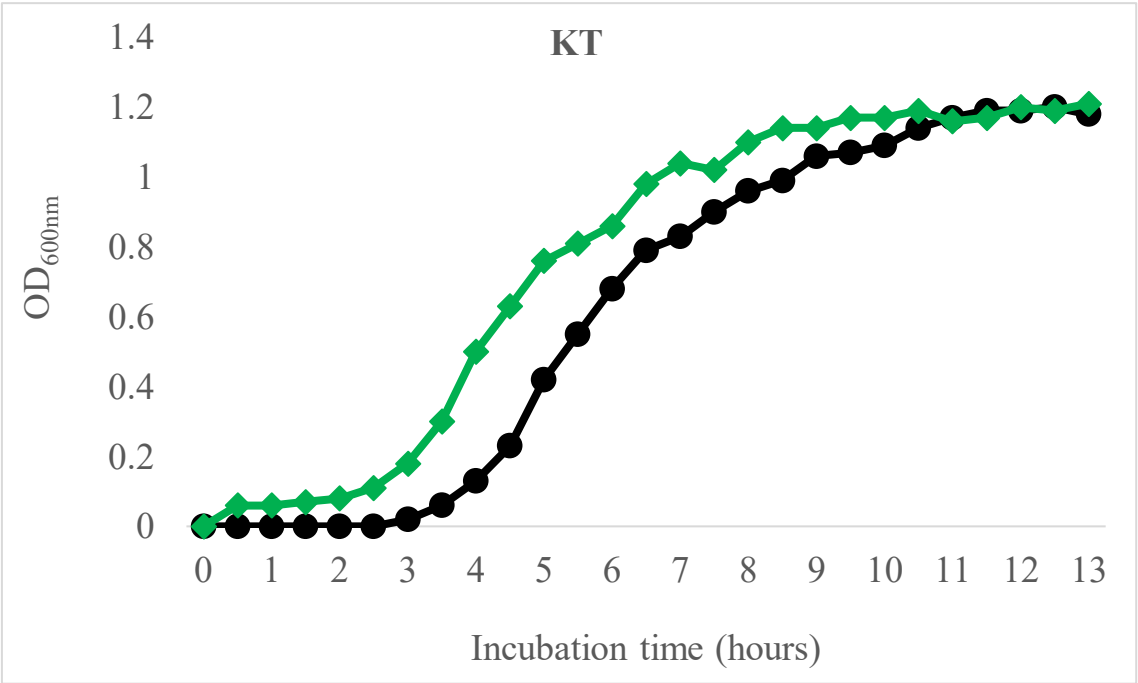


Figure 2.1. Growth curve of the seven strains (B3, B5, BC, M4-VN, KT, TL3 and TL5) at 30 and 37°C (black and green lines respectively). (*To be continued)*

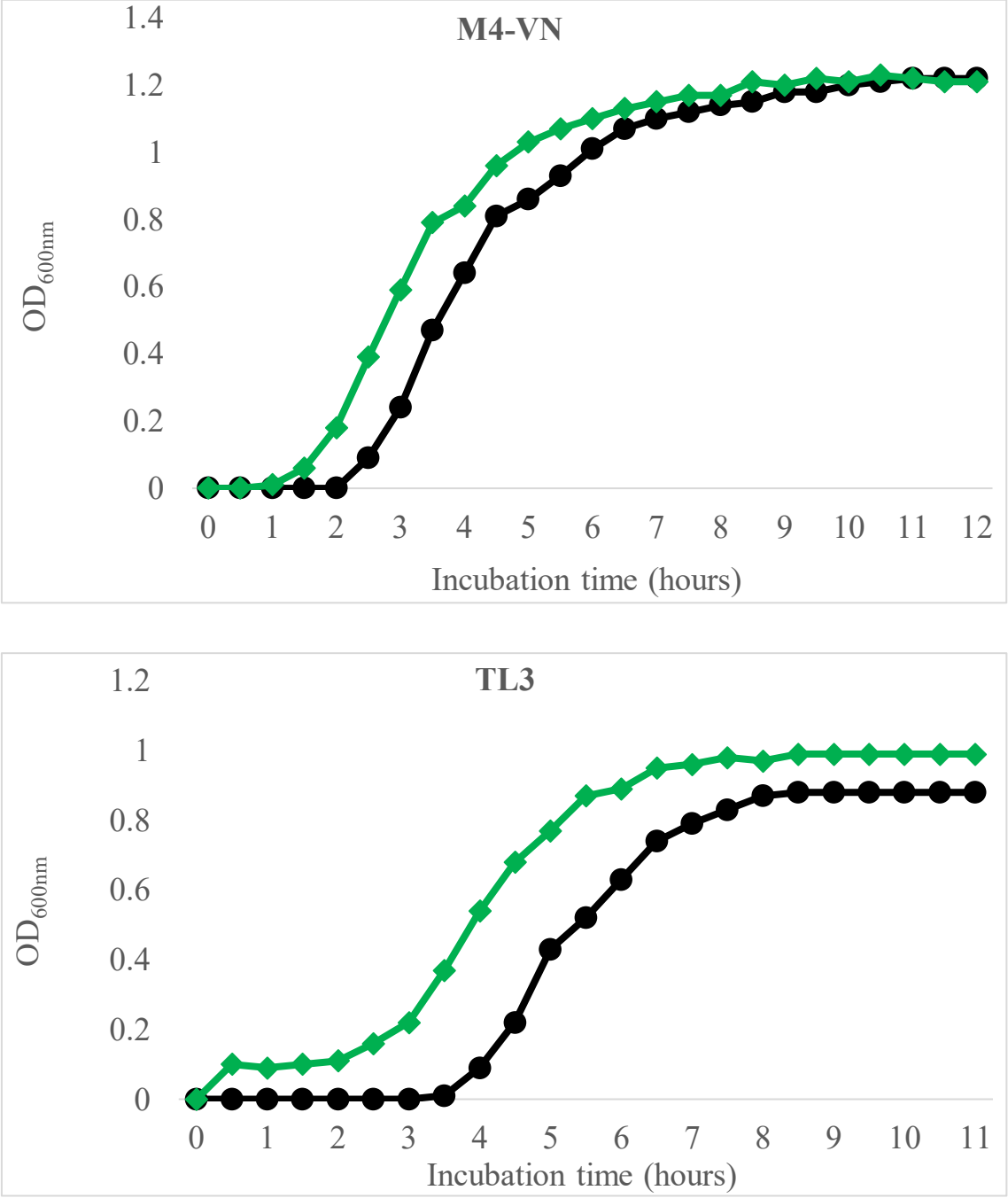


Figure 2.1. Growth curve of the seven strains (B3, B5, BC, M4-VN, KT, TL3 and TL5) at 30 and 37°C (black and green lines respectively). (*To be continued)*

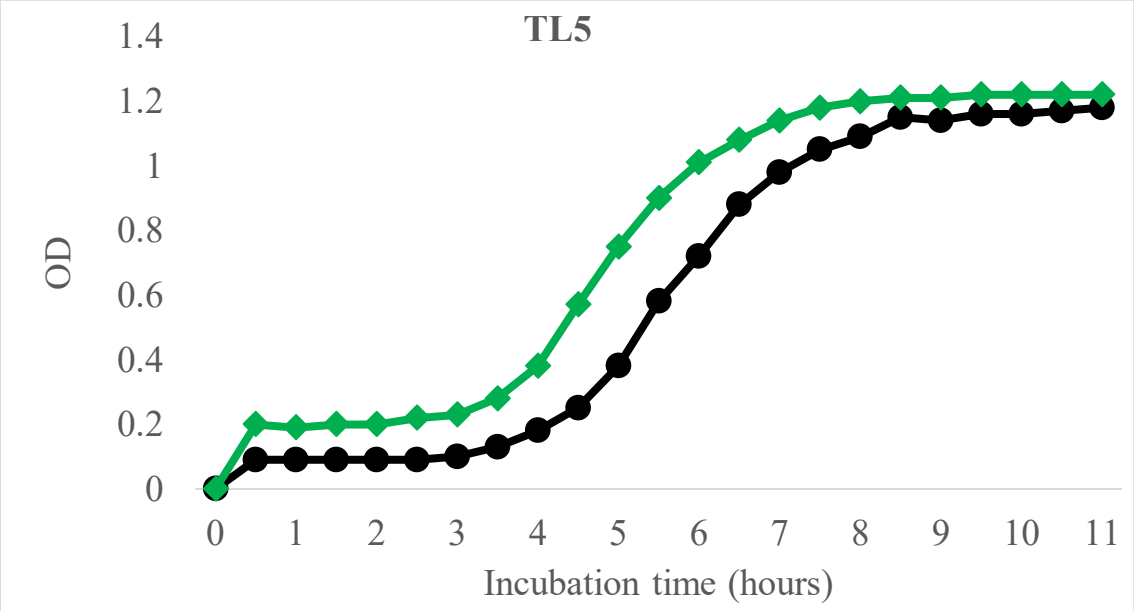


Figure 2.1. Growth curve of the seven strains (B3, B5, BC, M4-VN, KT, TL3 and TL5) at 30 and 37°C (black and green lines respectively).

Table 2.5. Biochemical characteristics of the seven strains B3, B5, BC, M4-VN, KT, TL3 and TL5.

Substrates	B3	B5	BC	M4-VN	KT	TL3	TL5	1	2	3	4	5
OPNG (Ortho-nitrophenyl-β-$_D$-galactopyranoside)	+	+	-	+	+	+	+			+	+	+
L-arginine	+	+	-	+	+	-	-	+	+	+	-	-
L-lysine	-	-	-	-	-	-	-		-	+		
L-orthine	+	+	-	+	-	-	-	+	+	+		
Trisodium citrate	+	+	-	+	+	+	+	+	+	+		
Natriumthiosulfate	-	-	-	-	-	-	-					
Ure	-	-	-	-	-	-	-	+	+	-	-	-
L-tryptophane	+	+	+	+	+	+	+			-		
Indole	-	-	-	-	-	-	-	+	-	-	-	-
Voges-Proskauer	-	-/+	-	-	-	+	+	-	-	-	+	+
Gelatin	-	-	-	-	-	-	-	-	-	-		
$_D$-glucose	+	+	+	+	+	+	+			+	+	+
D-manitol	+	+	+	+	+	+	+			+	+	+
Inositol	-	-	+	-	-	+	+			+	+	+
Sorbitol	-	-	-	+	+	+	+			+	-	-
L-Rhamnose	-	-	+	+	-	+	-	+	-	+	+	+
$_D$-Saccharose/succrose	+	+	-	+	+	+	+			+	+	+
$_D$-Melibiose	+	+	+	+	-	-	-			+	-	+
Amygdaline	+	+	+	+	+	+	+			+		
L-Arabinose	+	+	+	+	+	+	+	+	+	+	+	+
Cytochorome-oxydase	-	-	-	-	-	-	-			-		

NO_2	+	+	-	+	+	-	-		+		
N_2			-			+	+				
Growth at 50°C	-	-	-	-	-	-	-	-	-	-	-
Growth at 10°C	-	-	-	-	-	-	-	-	-	-	-

(+), (-), (-/+) indicate positive, negative and weak reactions. 1, *E. cloacae*; 2, *E. asburiae*; 3, *E. kobei; 4, P. dispersa*; 5, *P.cypripedii*
Biochemical characteristics of references were retrieved from (*bacdive.dsmz.de*) website.

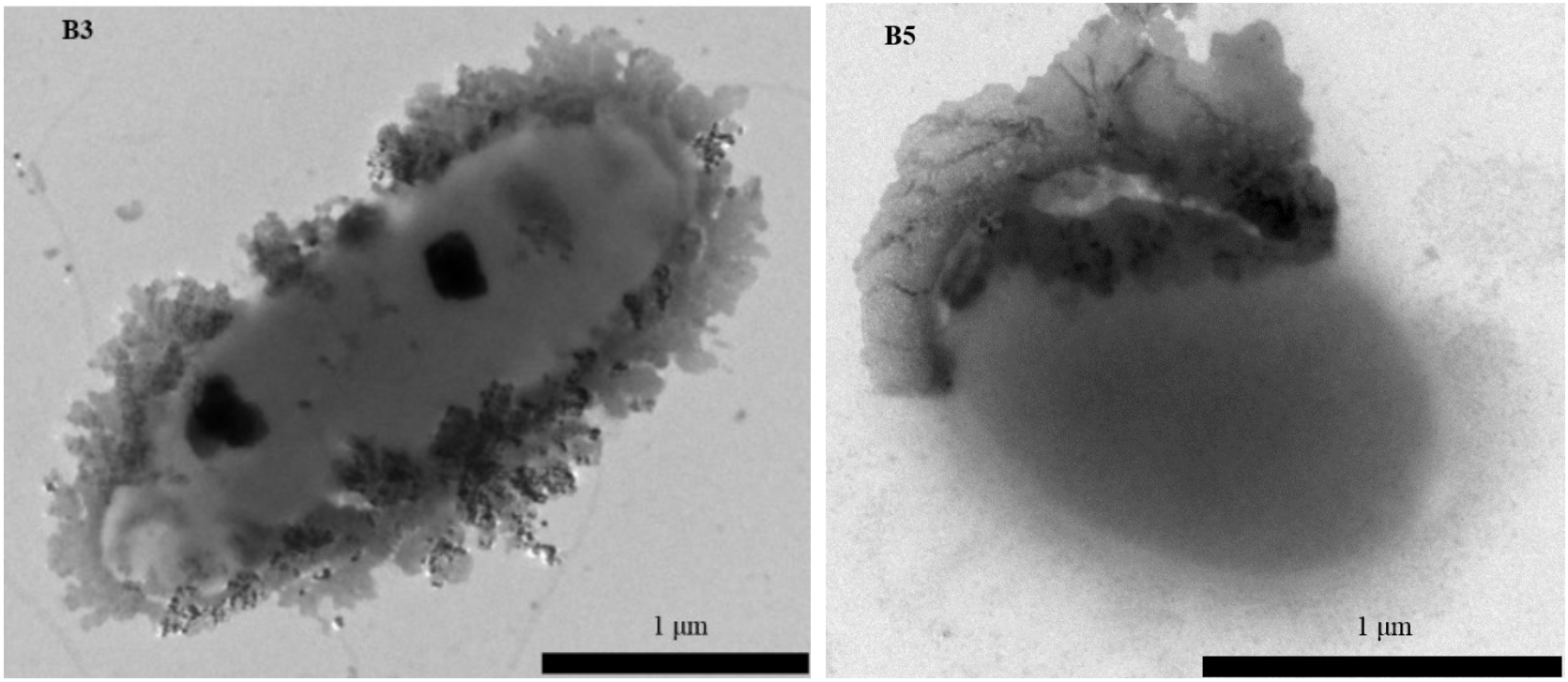


Figure 2.2. Transmission electron microscopy of B3, B5, BC, M4-VN, KT, TL3 and TL5. Bars indicate 1 µm (B3, B5, M4, KT, TL3, TL5), and 2 µm (BC). (*To be continued*)

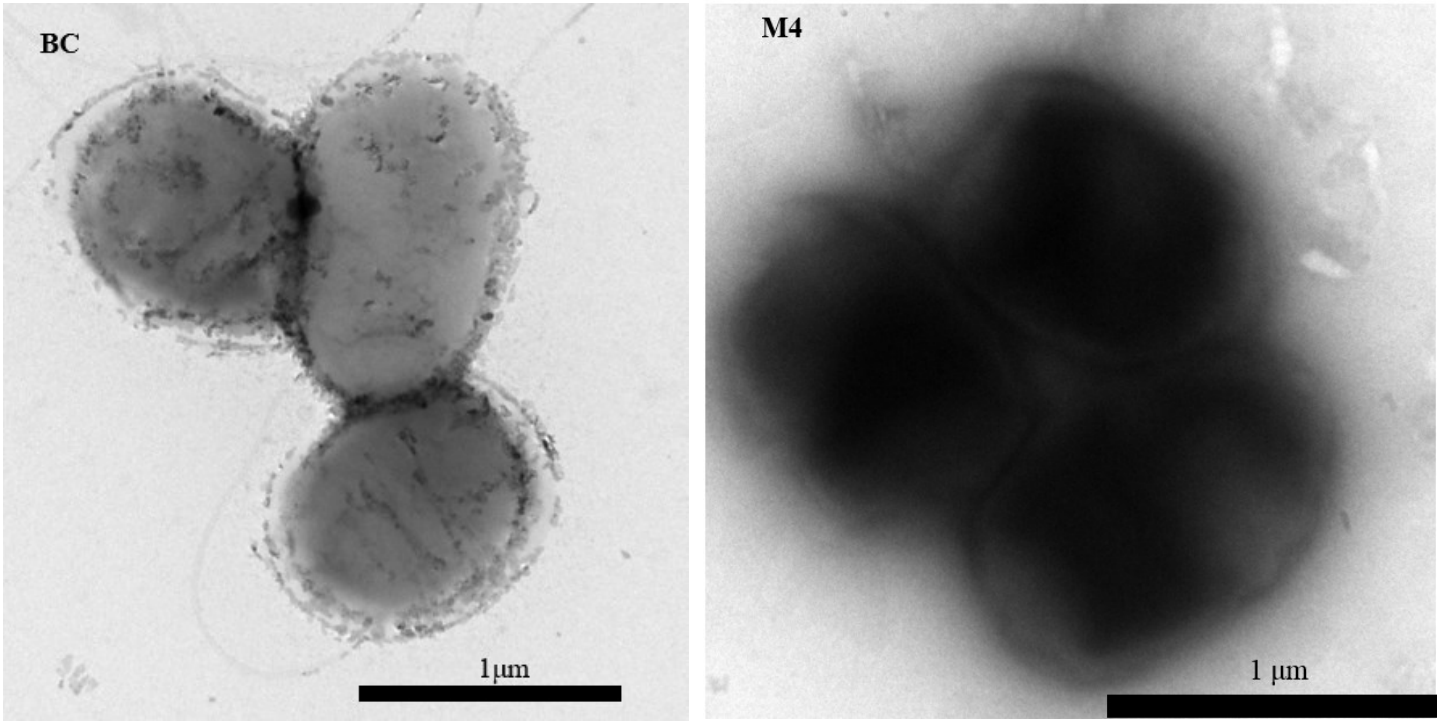


Figure 2.2. Transmission electron microscopy of B3, B5, BC, M4-VN, KT, TL3 and TL5. Bars indicate 1 µm (B3, B5, M4, KT, TL3, TL5), and 2 µm (BC). (*To be continued*)

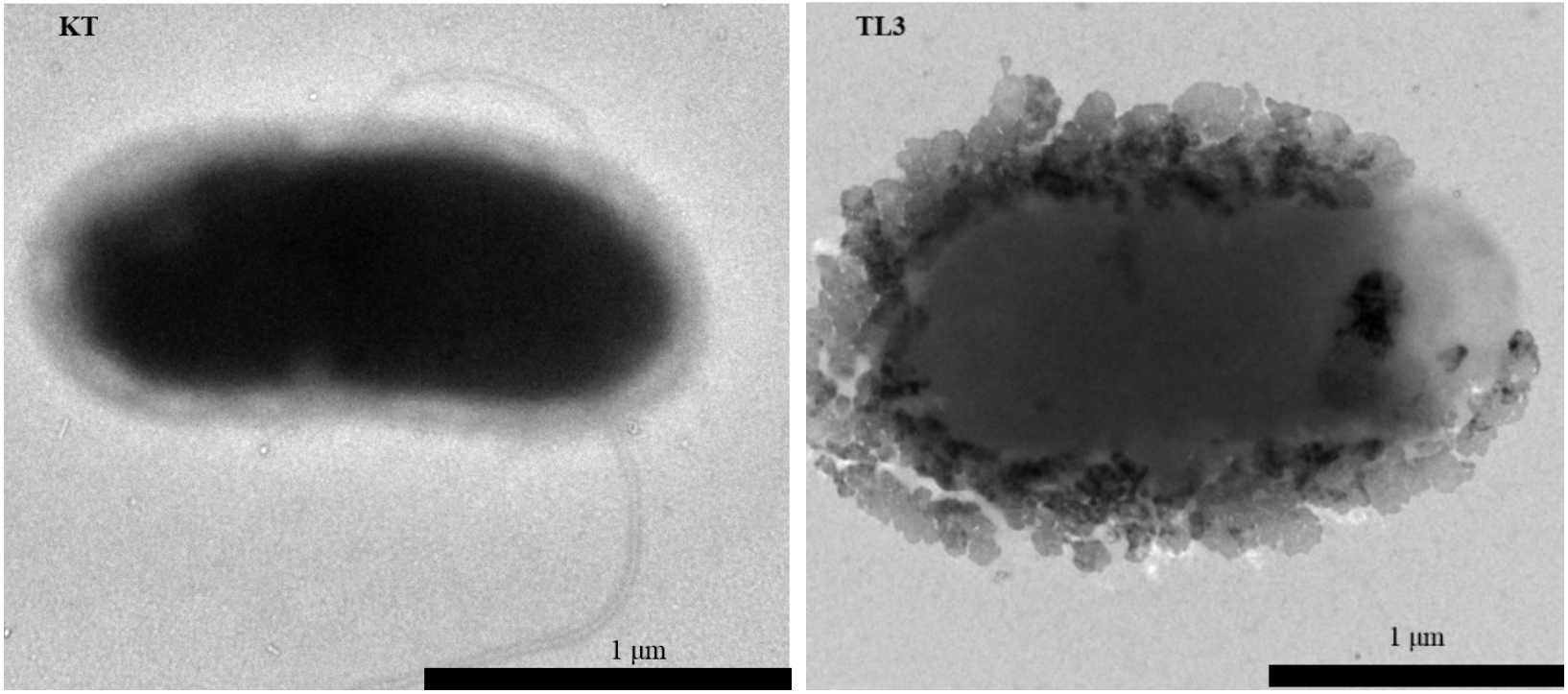


Figure 2.2. Morphology of B3, B5, BC, M4-VN, KT, TL3 and TL5. Bars indicate 1 µm (B3, B5, M4, KT, TL3, TL5), and 2 µm (BC). (*To be continued*).

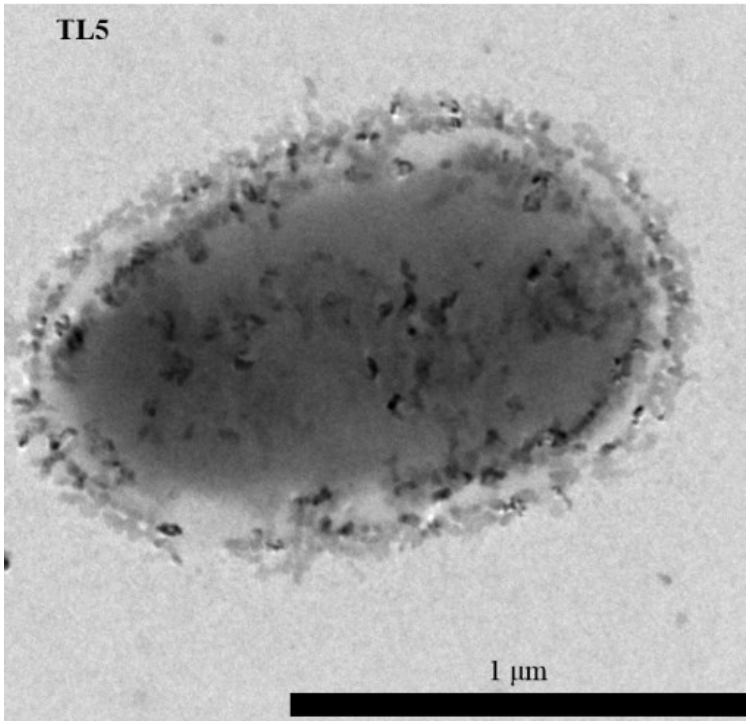


Figure 2.2. Transmission electron microscopy of B3, B5, BC, M4-VN, KT, TL3 and TL5. Bars indicate 1 µm (B3, B5, M4, KT, TL3, TL5), and 2 µm (BC).

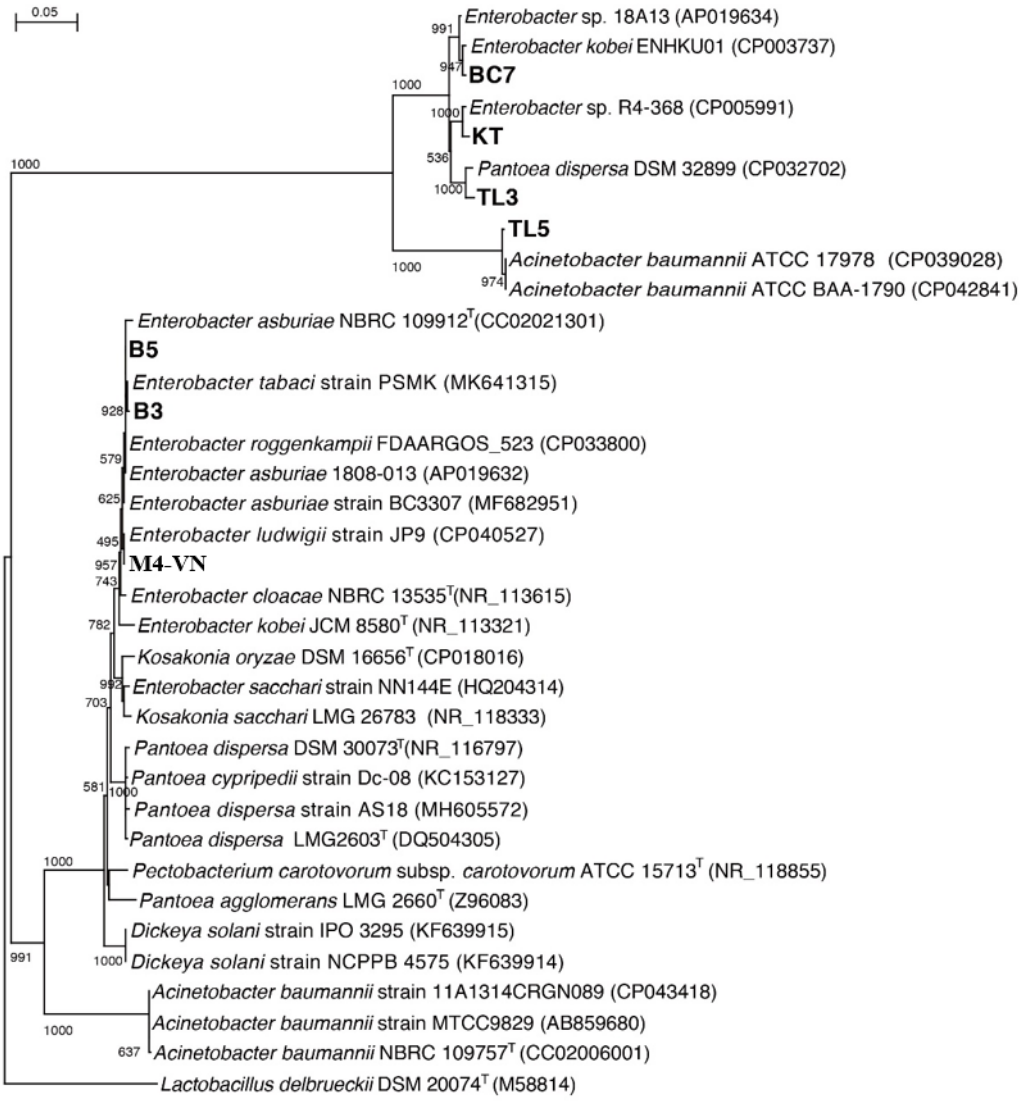


Figure 2.3. Phylogenetic tree of isolated strains and type strains of related genera. Their nucleotide sequences were compared using ClustalW, and phylogenetic trees were generated using the neighbor-joining method. Bar indicates the number of substitutions per sequence position.

Chapter 3. Genome analysis of *Enterobacter* sp. M4-VN isolated from soft rot

1. Introduction

Enterobacter species are gram-negative facultatively anaerobic bacteria belonging to the family *Enterobacteriaceae* (Brenner *et al*., 2005; Adeolu *et al*., 2016). In most cases, the species have been reported to be nosocomial pathogens due to a resistance to disinfectants and antimicrobial chemicals (Joseph, 2005). However, some of them have been reported as phytopathogens, such as *E. asburiae* on Konjac (*A. konjac*) in China and *E. cloacae* on chili pepper (*C. annuum* L.) and dragon fruit (*Hylocereus* spp.) (Wang *et al*., 2010; Wu *et al*., 2011; García-González *et al*., 2018; Masyahit *et al*., 2009).

Identification of the order *Enterobacteriales* based on the 16S rRNA genes has been difficult due to low discriminatory power (Adeolu *et al*., 2016). A number of alternative genes have been used in the phylogenetic analysis of the order *Enterobacteriales* to gain additional insight into the relationships of the members of the order, such as *gyrB* (Fukushima *et al*., 2002), *dnaJ* (Hong Nhung *et al*., 2007), *oriC* (Roggenkamp, 2007) and *recA* (Tailliez et al., 2010). More recently, multiple gene-based multilocus sequence analysis studies have been conducted to further clarify the phylogeny of the order *Enterobacteriales* including studies based on the genes *tuf* and *atpD* (Paradis *et al*., 2005), the genes *atpD, carA,* and *recA* (Young and Park, 2007), the genes *gapA, gyrA* and *ompA* (Naum *et al*., 2011), the genes *rpoB, gyrB, dnaJ* and *recA* (Hata *et al*., 2016), the genes *fusA, pyrG, rplB, rpoB* and *sucA* (Ee *et al*., 2016), and the genes *gyrB, rpoB, atpD,* and *infB* (Brady *et al*., 2008; Glaeser and Kämpfer, 2015; Zhang and Qiu, 2015). These reports have led to a considerable number of reclassification within the order *Enterobacteriales.* A deeper resolution for the identification is digital DNA-DNA hybridization (dDDH) (Auch *et al*., 2010). The hybrid DDH value is usually specified relative to the DDH value obtained by hybridizing a reference genome with itself. DDH values $\leq 70\%$ are considered as an indication that the tested organism belongs to a different species (Auch *et al*., 2010). Recently, Type Strain Genome Server (TYGS) has been developed. The server provides comprehensive access to nomenclature and highly reliable taxonomy (Meier-Kolthoff and Göker, 2019).

Therefore, this study aims to identify the strain M4-VN based on sequences of the whole genome. Furthermore, based on the genome, plant pathogenic genes, and phage-resistant genes might be identified.

2. Methodology

2.1. Isolation of strain

The strain M4-VN used in this study was isolated from the potato having the soft

rot at Hanoi, Vietnam, in 2013. Potatoes were washed with steriled water and 70% alcohol to remove the surface contaminants, rinsed with sterile distilled water, and cut into specimens. The specimens that have the disease were selected and streaked on to LB plates and inoculated at 37°C for 24–48 h. The bacterial colonies were purified with serial streaking.

2.2. Genome extraction

A single colony of strain M4-VN was then cultivated anaerobically overnight at 37°C in 50 mL LB broth for the sequencing purpose. Next, the total genomic DNA was extracted and purified according to the method of Saito and Miura (1963) with TE saturated phenol (pH 8.0) (NIPPON GENE CO., LTD., Tokyo, Japan) and phenol-chloroform. The genomic DNA was precipitated with 3 M sodium acetate to get a final concentration of 0.3 M, then two times the acquired volume of 100% ethanol was added. The precipitate was washed twice with 70% ethanol, air-dried, and resuspended in TE buffer (0.1 M Tris-HCl; 1 mM EDTA (pH 8.0)). The prepared genome DNA was finally sequenced using the Illumina HiSeq platform (Illumina, San Diego, CA, USA) by GeneWiz Inc. (South Plainfield, NJ, USA).

1 M Tris-HCl was prepared by dissolving 6.05 g of Tris base in 30 mL of H_2O. After that, pH was adjusted to 7.5 with 5 M HCl. Finally, the mixture was filled up to 50 mL with H_2O. 0.1 M Tris-HCl was diluted ten times from 1 M Tris-HCl by H_2O.

2.3. Sequencing method

Next generation sequencing library preparations were constructed following the instruction of the VAHTS Universal DNA Library Prep Kit for Illumina V3. For each sample, 100 ng genomic DNA was randomly fragmented to <500 bp by sonication S220 (Covaris, INC). The fragments were treated with End Prep Enzyme Mix for end repairing, 5’ Phosphorylation and dA-tailing in one reaction.

Table 3.1. 5’ Phosphorylation and dA-tailing

Component	Volume
Input DNA	x µl (100 ng genomic DNA acquired)
End Prep Mix 4	15 µl
ddH_2O	fill up to 65 µl
Total volume	65 µl

The treated fragments were then mixed thoroughly by gently pipetting up and down before putting in a PCR instrument and running the following PCR program (Hot lid: 105°C).

Table 3.2. PCR program

Temperature	Time
20°C	15 min
65°C	15 min
4°C	Hold

The PCR product was ligated to adapters using ligation buffer and DNA ligase and incubating at 20°C for 15 min with the heated lid off (105°C).

Table 3.3. Ligation reaction

Component	Volume
End Preparation Products	65 µl
Rapid Ligation buffer 2	25 µl
Rapid DNA ligase	5 µl
DNA Adapter X	5µl
Total volume	100 µl

The ligated sample was then cleanup with VAHTS DNA Cleanup Beads. After that, the purified product was amplified by PCR reaction using P5 (5'-AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT-3') and P7 (5′-AGATCGGAAGAGCACACGTCTGAACTCCAGTCAC-3′) primers.

Table 3.4. PCR reaction mixture

Component	Volume
Size-selected Adapter-ligated Library	20 µl
P5 primer	2.5µl
P7 primer	2.5µl
VAHTS HiFi Amplification Mix	25µl
Total volume	50µl

Table 3.5. PCR cycling condition

Temperature (°C)	Time	Cycles
95	3 min	1
98	10 s	
60	75 s	8
72	30 s	

72	5 min	1
4	Hold	

The PCR products were cleaned up with VAHTS DNA Clean Beads, validated using an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA, USA), and quantified by Qubit 3.0 Fluorometer (Invitrogen, Carlsbad, CA, USA). Then libraries with different indices were multiplexed and loaded on an Illumina HiSeq instrument according to manufacturer's instructions (Illumina, San Diego, CA, USA). Sequencing was carried out using a 2x150 paired-end (PE) configuration; image analysis and base calling were conducted by the HiSeq Control Software, Offline Base Caller 1.9.3 and GAPipeline-1.6 (Illumina) on the HiSeq instrument. The reads were analyzed and assembled using Velvet, gap-filled with SSPACE and GapFiller (Zerbino and Birney, 2008; Zerbino *et al*., 2009; Boetzer *et al*., 2011; Boetzer and Pirovano, 2012; Hunt *et al*., 2014).

2.4. Genome analysis

The determination of closely related type strains and the phylogenetic tree analysis were done by the TYGS platform (https://tygs.dsmz.de/user_requests/new) (Ondov *et al*., 2016; Meier-Kolthoff and Göker, 2019). The pairwise distance (d) between genomes was calculated to evaluation of dDDH as the following formula:

$$d = 1 - \left(2 * \frac{I_{XY}}{H_{XY} + H_{YX}}\right)$$

In this formula: I_{XY} denote the sum of identical base pairs over all HSPs and H_{XY} denote the total length of all HSPs found by BLASTing genome X against genome Y, whereas H_{YX} and I_{YX} are obtained by sing Y as the query and X as the subject sequence.

The phylogenetic tree was constructed based on Genome Blast Distance Phylogeny (GBDP) distances (Meier-Kolthoff *et al*., 2013; Lefort *et al*., 2015). The GBDP distances were calculated from genome sequences as the following formula:

$$\text{GBDP distance} = -\log\left(2 * \frac{I_{XY}}{H_{XY} + H_{YX}}\right)$$

The average nucleotide identity (ANI) was calculated using ANI calculator platform (https://www.ezbiocloud.net/tools/ani) (Seok Hwan *et al*., 2017). Annotation was performed using the DFAST tool (https://dfast.nig.ac.jp/) (Tanizawa *et al*., 2016, 2018).

2.5. Screening of prophage-like sequences in strain M4-VN

Viral sequences were discovered in the bacterial genome and were annotated with PHASTER platform (http://phaster.ca/) using setting described in a report by Arndt (2016), followed by manual l inspection of the sequences for the presence of signature genes: attachment sites (*att*), genes encoding integrase, terminase, transposase, genes coding for structural viral proteins and the sequences of prophage integration sites, as

suggested by others (Boyd and Brüssow, 2002).

3. Results

3.1. Genomic analysis of *Enterobacter* sp. M4-VN

The assembled genome of *Enterobacter* sp. M4-VN contained 18 contigs with a total of 4,754,309 bp (G+C content, 55.1%), an N_{50} value of 636,975 bp, and minimum and maximum contig lengths of 1,812 and 949,261 bp, respectively. Annotation using DFAST revealed 4,424 predicted coding regions, 65 tRNA genes, seven rRNA genes, and one CRISPR region.

The results of dDDH (Table 3.6) showed that the strain M4-VN is homologous (92.6%) with *E. kobei* DSM 13645^T. The percentage is remarkably higher than the second closest the dDDHs (42.5%) calculated with *Enterobacter bugandensis* EB-247^T and *Enterobacter chengduensis* WCHECl-C4^T. The phylogenetic tree (Figure 3.1) created using related type strains' whole genome depicted that only strain M4-VN and *E. kobei* DSM 13645^T created separated cluster. Furthermore, ANI results (Table 3.6) also remarked the genetically close relationship between the strain M4-VN and *E. kobei* DSM 13645^T with a score of 99.07. These ANIs computed with other strains were 91.08 (with *E. chengduensis* WCHECl-C4^T), 90.97 (with *E. bugandensis* EB-247^T) 90.91 (with *Enterobacter chuandaensis* 090028^T), and lower (with the others).

3.2. Accession numbers

The *Enterobacter* sp. M4-VN genome sequence and annotation data were deposited in the DDBJ/Genbank under BioProject number PRJDB9609, BioSample number SAMD00218344 and the accession numbers BLVN01000001-BLVN01000018.

3.3. Presence of pathogenic genes in genome of *Enterobacter* sp. M4-VN

The annotation showed a gene encoding pectinesterase, which plays a role in macerating plant tissue (Gainvors *et al*., 1994; Gavrilovic *et al.*, 2001). In addition, numerous virulence factors of plant pathogenic bacteria were found in M4-VN's genome. These include a gene coding for the TonB protein (ORF2681), four genes encoding the TonB-dependent receptor (FusA) (ORF263, ORF856, ORF2542, and ORF3703), and a gene encoding the M16 proteases (FusC) (ORF3359) (Mosbahi *et al*., 2018), genes related to susceptibility to antibacterial plant chemicals (*tolC*) (ORF257) (Lee *et al*., 2013).

3.4. Presence of prophage-like sequences in genome of *Enterobacter* sp. M4-VN

The analyses of the 18 scaffolds of the strain M4-VN showed that four intact prophages were presumed (Figure 3.2 & Table 3.7). The four intact prophages (Pro1, Pro2, Pro3 and Pro4) have the lengths of 24.1, 51.3, 40.8 and 22.6 kb which were found in scaffolds 1, 3, 4 and 16t, respectively. The remained scaffolds did not harbor any prophage-like elements. Although predicted sequences differ greatly from lengths, the G+C content of the four prophage ranges from 50.24% (Pro3) to 55.28% (Pro1) (Table

3.7). In Pro1 genome, totally, 30 genes were found. The gene number were 53, 38 and 33 in the phage Pro2, Pro3 and Pro4, respectively. The predicted gene products include head proteins, fiber proteins, integrase, phage-like protein, plate protein, protease, portal protein, terminase, tail protein and hypothetical proteins (Figure 3.3). Phylogenetic tree based on whole genome of the four phages with the most common phage showed that Pro1, Pro3 and Pro4 created a cluster with two myovirus, *Escherichia* phage 186 (Xue and Egan, 1995) and *Salmonella* phage SP_004 (Moreno Switt *et al*., 2013), while Pro2 formed a cluster with two siphovirus, *Klebsiella* phage phiKO2 (Casjens *et al*., 2004) and *Cronobacter* phage phiES15 (Lee *et al*., 2012) (Figure 3.4). This result suggests that Pro1, Pro3 and Pro4 belong to *Myoviridae* family and Pro2 is a *Siphoviridae* member.

4. Dicussion

It was reported that the dDDH value was highly reliable estimator for the relatedness of genomes, which have several advantages over the various ANI implementations (Meier-Kolthoff and Göker, 2019). However, in this study, the results show high homology between strain M4-Vn and *E. kobei* DSM 13645^{T} with exceptional values of dDDH (92.6%) and ANI (99.07). These scores are much higher than the second highest scores of other species. These are strong evidences to conclude that the strain M4-VN belongs to *E. kobei* species. *E. kobei* was first described in 1996 and mainly isolated from clinical specimens (Kosako *et al*.). It differed from other *Enterobacter* species based on the negative reaction with Voges-Proskauer test. Thus, the result of conclusion of *E. kobei* M4-VN in this chapter is compatible with that of chapter 2.

Up to date, there is no report on the plant pathogenicity of *E. kobei.* However, genes encoding the pectinesterase and virulence factors were found in *E. kobei* M4-VN. Pectinesterase in one of the degradative enzymes, including other pectinases, cellulases, and proteases secreted by SRP species (Plastow, 1988a; Gavrilovic *et al.*, 2001). The initial interaction of the pathogen with the plant involves degradation of the cell wall and enables penetration and colonization of host tissue. Pectin is an important structural component of the cell wall and the degradation is required for initial stage of infection. Pectinesterase was found in various species of *Pectobacterium* family such as Pcc and *Dickeya chrysanthemi* (formerly known as *Erwinia chrysanthemi*) (Plastow, 1988a; Gavrilovic *et al.*, 2001). Furthermore, during the invasion process into host, pathogenic bacteria encounter different environmental conditions, of which iron limitation and reactive oxygen species produced by the host plant are major factors limiting the ability of the bacteria to spread and colonize the host (Plastow, 1988b; Lee *et al.*, 2013). Thus, bacteria must tightly regulate iron uptake and deal with changes occurring during redox conditions (Lee *et al.*, 2013). In *Pectobacterium* species, FusC, an M16 protease, displays a highly specific proteolytic activity against plant ferredoxin and enhances the species to acquired irons from ferredoxin imported via the TonB-dependent receptor FusA. On transport to the periplasm, FusC cleaves ferredoxin and consequently releases 2Fe-2S cluster, which may be transported to the cytoplasm by a putative transporter (Mosbahi *et*

al., 2018). In Pcc, *tolC* was reported as a novel gene associated with the pathogenicity due to the different sensitivity of the *tolC* mutant to plant-derived chemicals (Lee *et al*., 2013). Therefore, containing these genes highly suggests that *E. kobei* M4-VN is a plant pathogen.

Using phage therapy for plant disease control is trendy. However, this method has been facing various challenges as mentioned in chapter 1. Phage resistance is one reason. Resistant mutants can appear via interaction between phages and bacteria (Wang *et al*., 2019). Phages that integrate into bacterial chromosomes are called prophages, which are important gene elements of bacterial chromosomes and enable horizontal gene transfer between phages and bacteria. Some prophages provide the host bacteria beneficial traits such as protection from infection by other related phages based on Sie mechanism (Hofer *et al.*, 1995; Mahony *et al*., 2008). Four intact prophages were detected in *E. kobei* M4-VN chromosomes. However, no phage-resistant genes were found. This result explores that Sie mechanism is not acquired in *E. kobei* M4-VN.

Table 3.6. Pairwise comparisons of *Enterobacter* sp. M4-VN genome with related type strain *Enterobacter* genomes using TYGS platform

Subject strain	dDDH (d, in %)	C.I. (d, in %)	ANI	G+C content difference (in %)
Enterobacter kobei DSM 13645^T	92.6	[90.6 - 94.2]	99.07	0.14
Enterobacter bugandensis EB-247^T	42.5	[40.0 - 45.0]	90.97	0.95
Enterobacter chengduensis WCHECl-C4^T	42.5	[40.0 - 45.1]	91.08	0.69
Enterobacter chuandaensis 090028^T	42.4	[39.8 - 44.9]	90.91	0.63
Enterobacter asburiae ATCC 35953^T	42.1	[39.6 - 44.6]	90.97	0.42
Enterobacter roggenkampii DSM 16690^T	40.9	[38.4 - 43.4]	90.48	0.99
Enterobacter sichuanensis WCHECL1597^T	38.6	[36.1 - 41.1]	89.64	0.19
Enterobacter mori LMG 25706^T	35.3	[32.9 - 37.8]	88.47	0.25
Enterobacter cloacae subsp. *dissolvens* ATCC 23373^T	34.6	[32.2 - 37.1]	88.16	0.11
Enterobacter ludwigii DSM 16688^T	34.1	[31.7 - 36.6]	87.97	0.45
Enterobacter taylorae NCTC 12126^T	31.1	[28.7 - 33.6]	86.27	0.68
Enterobacter cancerogenus ATCC 33241^T	31.1	[28.7 - 33.6]	86.28	0.63

d: pairwise distance of genomes
dDDH: digital DNA-DNA Hybridization
C.I: Confidence Intervals
ANI: Average Nucleotide Identity

Table 3.7. Prophage regions identified in *Enterobacter* sp. M4-VN genome

Region	Pro1	Pro2	Pro3	Pro4
Region Length	24.1 kb	51.3 kb	40.8 kb	22.6 kb
Completeness(score)	intact(120)	intact(150)	intact(150)	intact(110)
Specific Keyword	tail, lysin, plate, head, protease	integrase, lysis, terminase, portal, head, capsid, tail	integrase, portal, terminase, capsid, head, tail, lysin, plate	tail, capsid, portal, terminase, head
Region Position	scaffold1:208153-232269	scaffold3:562608-613941	scaffold4:882580-923400	scaffold16:331-23028
GC %	55.28	50.24	52.49	51.36
# tRNA	1	0	0	0
# Total Proteins	30	53	38	33
# Phage Hit Proteins	27	43	34	27
# Hypothetical Proteins	3	10	4	6
Phage + Hypothetical Protein %	100%	100%	100%	100%

# Bacterial Proteins	0	0	0	0
Attachment Site	no	yes	yes	no
Most Common Phage Name/Accession number/(hit proteins count)	*Escherichia* phage 186 (NC_001317)(19)	*Klebsiella* phage phiKO2 (NC_005857)(15)	*Salmonella* phage SP_004 (NC_021774)(22)	*Cronobacter* phage phiES15 (NC_018454)(9)

Region: The number assigned to the region.
Region Length: The length of the sequence of that region.
GC %: The percentage of GC nucleotides of the region.
Completeness: A prediction of whether the region contains an intact or incomplete prophage based on the above criteria.
Score: The score of the region based on the above criteria.
Total Proteins: The number of ORFs present in the region.
Region Position: The start and end positions of the region on the bacterial chromosome.
Most Common Phage: The phage(s) with the highest number of proteins most similar to those in the region.

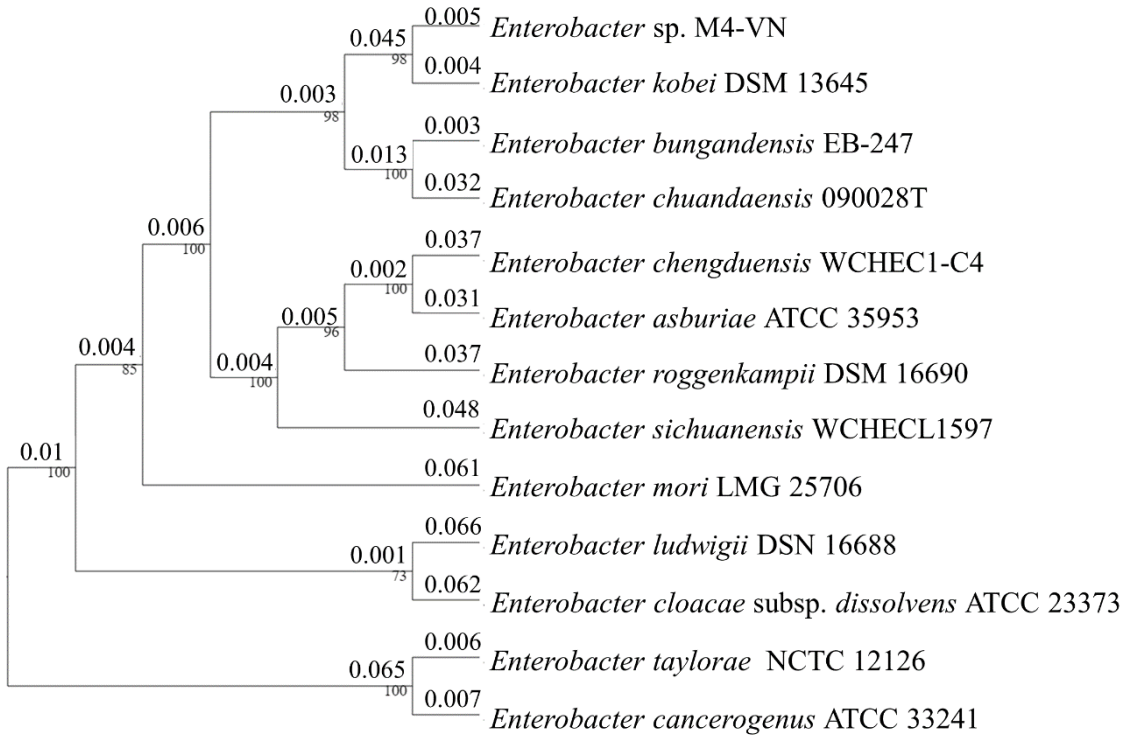


Figure 3.1. Phylogenetic tree of *Enterobacter* sp. M4-VN with related type strains based on GBDP distances using TYGS platform (https://tygs.dsmz.de/user_requests/new). The number on the branches indicates branch length values. The number under the branches indicates confidence intervals.

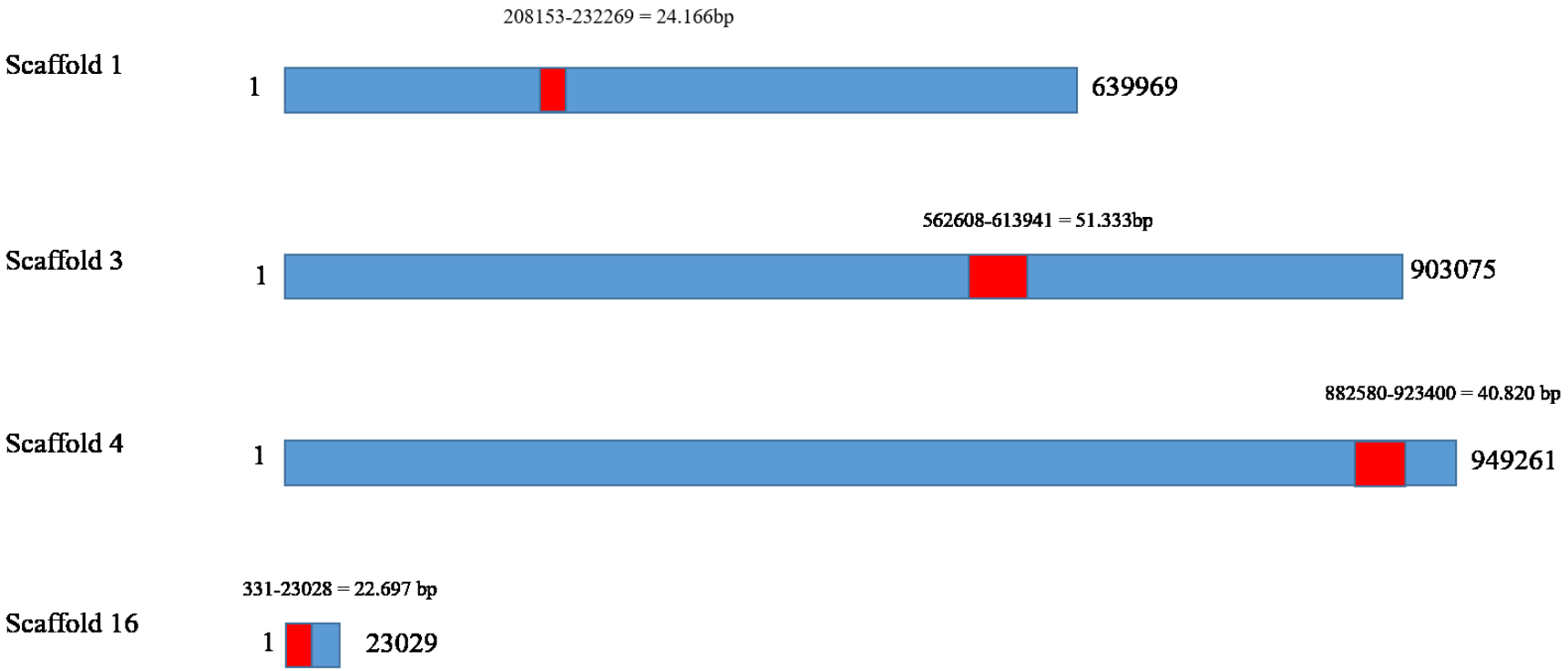


Figure 3.2. Distribution of prophage regions in 4 different scaffolds (the blue color bar). The red color indicates the prophage regions. The prophage sequences were detected using PHASTER.

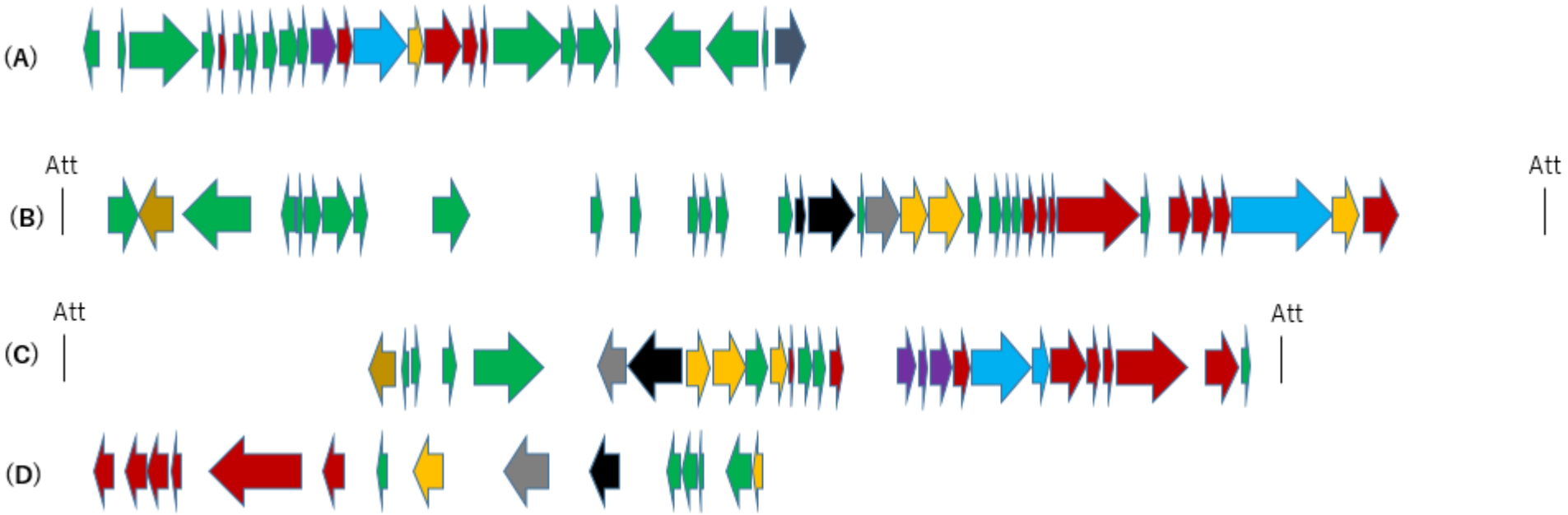


Figure 3.3. Genome map of the four prophage. (A)Pro1, (B) Pro2, (C) Pro3, (D) Pro4. Att: Attachment site; Head protein (Yellow); Fiber protein (Blue); integrase (Dark yellow); Phage-like protein (Green); Plate protein (Purple); Protease (Dark blue); portal protein (Grey); Terminase (Black); Tail protein (Red). Hypothetical proteins are not shown.

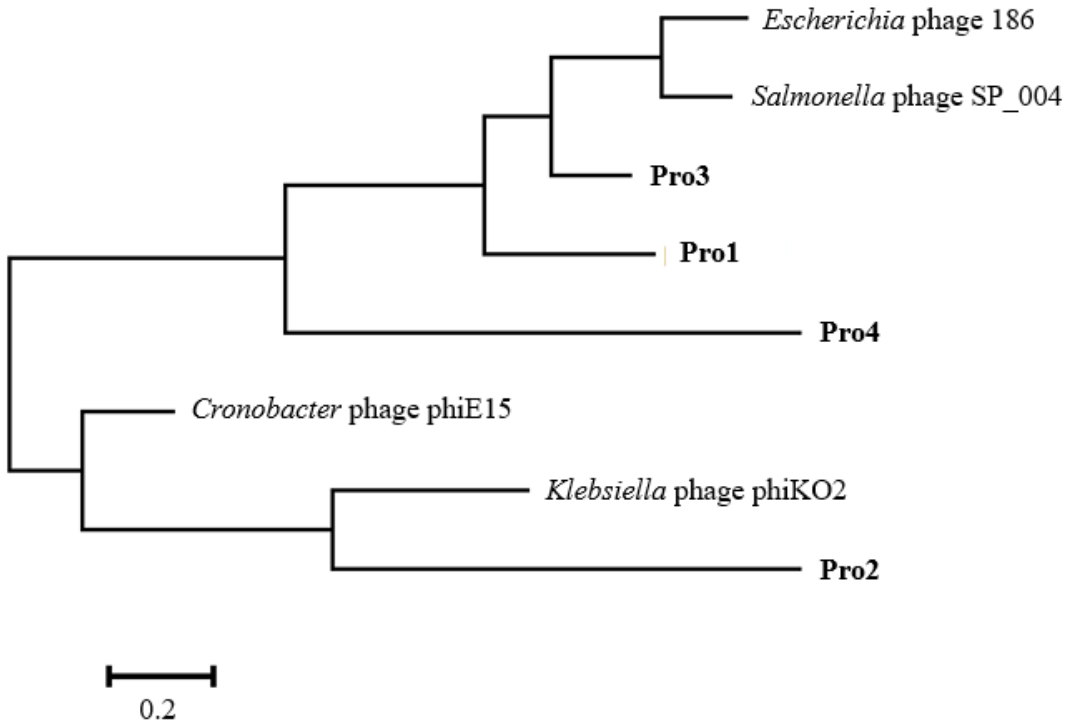


Figure 3.4. Phylogenetic trees generated based on whole genome of prophages and the most common protein phages in the Table 3.7. Nucleic acid sequences were compared using ClustalW, and phylogenetic trees were generated using the neighbor-joining method. Bar indicates the number of substitutions per sequence position.

Chapter 4. Identification and characterization of virulent phage EspM4VN to control *Enterobacter kobei* M4-VN isolated from soft rot

1. Introduction

Soft rot is a serious disease of worldwide economic significance that occurs on fleshy vegetables such as the potato, carrot, eggplant, squash, and tomato and is caused by SRP species (Ma *et al*., 2007; Masyahit *et al*., 2009; Thanh *et al*., 2009; Lim *et al*., 2013; George *et al*., 2018; Rossmann *et al*., 2018). Recently, various *Enterobacter* species were reported as causative pathogens as mentioned in chapter 1. Interestingly, among the seven strains isolated from soft rot in Vietnam, five strains were identified as *Enterobacter* species. Also, strain M4-VN was determined as *E. kobei* based on TYGS platform.

E. kobei is a gram-negative, facultative anaerobic, oxidase negative, non-motile, and non-pigmented rod-shaped species isolated from soil, water, and clinical specimens (Kosako *et al*., 1996). *E. kobei* was classified into *Enterobacteriacae.* The GC contents of its DNA ranges from 52.7 to 53.6% (Kosako *et al*., 1996). This species resembles *E. cloacae* but differs in giving negative Voges-Proskauer reactions. In addition, *E. kobei* was reported to resist various antibiotics such as piperacillin, piperacillin, tazobactam, cefotaxime and ceftazidime. This resistance is raising a concern of using higher level of antibiotics and threatening to human health, environment and biodiversity in clinical and agricultural sectors.

In Southeast Asian countries such as Vietnam, chemical agents have been a dominant strategy for controlling plant diseases. The use of pesticides in the region has been increasing annually and has caused severe issues for human health and the natural environment due to residual concentration in products, land, air, and water. Roughly ten thousand tons of pesticides were used annually in the 1980s, and this number reached forty thousand tons in the early 2000s (Engineering and Consulting Firms Association, 2006). In 2015, averagely, it was estimated that 100,000 tons of active ingredients of pesticides were imported into Vietnam (www.thoibaotaichinhvietnam.vn). In addition to pesticides, integrated pest management involving cultivation and pest-resistant plants have been applied widely. However, various bacterial organisms have developed pesticide resistance such as kasugamycin resistance tested in *Erwinia* sp. or carbapenem and cephalosporin resistance reported in *Enterobacteriaceae* (McGhee and Sundin 2011). Despite a growing need for innovative solutions to this problem, use of biological control agents such as phages has not been significantly considered.

Therefore, phage control may provide a promising solution for environmental and human friendly treatment of bacteria-related plant diseases. However, the utility of phage therapy often presents a challenge because of specificity to their host bacteria. Titer of *E. aerogenes* phage phiEap-3 was decreased sharply at 50°C after 30-min incubation (Zhao *et al.*, 2019). Two broad-host lysates φPD10.3 and φPD23.1 have been used for effectively controlling the strains of *Pectobacterium* and *Dickeya* spp. (Czajkowski *et al*.,

2015). Gašić *et al*. reported *Xanthomonas euvesicatoria* phage Kφ1 had potential for biological control against pepper bacterial spot (2018). Additionally, phage typing also widely used in bacterial characterization and differentiation. For example, it was reported that nine of twenty-two *Enterobacter* phages investigated were isolated and used for distinguishing *Enterobacter* species (Loessner *et al*., 1993). Biological control using phages for soft rot have been applied to SRE such as *P. carotovorum* (Muturi *et al*., 2019), *D. solani* (Carstens *et al*., 2018) and *P. atrosepticum* (Buttimer *et al*., 2018) causing potato soft rot. However, there is no report studying phages infecting *E. kobei* and *Enterobacter* species such as *E. asburiae, E. cloacae, E. intermedium*, and related species. Understanding the biological and genetic characteristics of these phages is beneficial for applying phage therapy against emerging SRE infectious diseases and might be useful for screening profiles of *E. kobei* phages and other *Enterobacter* species phages.

Therefore, the aim of this study is to isolate and characterize a phage infecting *Enterobacter* species to be used against soft rot in infected crops.

2. Methodology

2.1. Bacterial strains and growth conditions

Bacterial strains used in this study is listed in Table 4.3. The strains B3, B5, BC7, M4-VN, KT, TL3, and TL5 were isolated and cultivated as described in chapter 2. Type strains were purchased by NBRC (Biological Resource Center, NITE; Chiba, Japan) and the *Pectobacterium* species were donated from Laboratory of Plant Pathology, Kyushu University for references. All bacteria were cultured in LB broth with shaking. Bacterial strains were cultivated at their optimum tempratures. Grown cells were suspended in 30% glycerol and stored at −80°C until use.

2.2. Isolation of the phage

Soil samples were collected from cabbage or potato fields in Bac Ninh province, Vietnam. These samples were suspended with SM buffer (Van Twest and Kropinski, 2009). Soil suspensions were added to a log-phase culture of *E. kobei* M4-VN for an enrichment culture. After shaking culture (37°C, 180 rpm, 24 h), the culture was centrifuged (10,000 ×*g*, 10 min, 4°C). The supernatant was filtered through a membrane filter with pore size of 0.45 μm (Advantec, Tokyo, Japan). The phage was assayed using the LB soft agar overlap technique (Adams *et al.*, 1959).

After overnight incubation, typical plaques were suspended in SM buffer and purified through five rounds of single-plaque isolation.

Table 4.1. SM buffer

Reagents	Final concentration
NaCl	100 mM
$MgSO_4 \cdot 7H_2O$	8 mM

Tris-HCl (1 M, pH 7.5)	50 mM
H_2O	Up to 1 Liter
Gelatin	0.002% (w/v)

2.3. Purification of phage

Phage particles were purified using cesium chloride (CsCl) gradient ultracentrifugation according to a previously reported protocol (Sambrook *et al*., 2001).

Table 4.2. CsCl solution

Density (g/mL)	CsCl	SM buffer (mL)
1.45	60	85
1.50	67	82
1.70	95	75

Enriched phage lysate (5×10^{10} PFU) was treated with DNase I (1 µg/mL), RNase A (1 µg/mL) at 37°C for 1 h. Phage lysates were centrifuged (10,000 ×*g*, 10 min, 4°C). The supernatant was mixed with 1 M NaCl and 10% (w/v) polyethylene glycol 8000, and the mixture was stored at 4°C for 2 h. After centrifugation (12,000 ×*g*, 30 min, 4°C), the pellets were suspended in 4 mL of SM buffer. The resulting suspension was gently mixed with 4.5 g of CsCl and loaded onto a step-layered CsCl gradient (1 mL of 0.5 g/mL CsCl, 2.5 mL of 0.775 g/mL CsCl, 2.5 mL of 1 g/mL CsCl) prepared in a 38.5 mL Ultra-clear tube (Beckman Coulter, Indianapolis, IN, USA). The gradient was centrifuged at 22,000 rpm in the swinging-bucket SW28 rotor for 2 h at 4°C in Optima XE-90 ultracentrifuge (Beckmann Coulter). The phage band was recovered and phage suspension was dialyzed against dialysis buffer (10 mM NaCl, 50 mM Tris-HCl (pH8.0), 10 mM $MgCl_2$) for 4 h. Purified phage particles were resuspended SM buffer with 25% glycerol and then stored at −80°C.

2.4. Transmission electron microscopy

Phage morphology was determined using transmission electron microscopy (TEM) to observe negatively stained preparations (Luo *et al*., 2012). Purified phage were suspended to 2% (wt/vol) phosphotungstic acid (pH 7.2) and then applied to the surface of a glow-discharged carbon/Formavar-coated grid (EM Japan Co., Ltd, Tokyo, Japan). Negatively stained phage particles were examined using Hitachi H-7500 transmission electron microscope operated at 80 kV (Hitachi High-Technologies Corp., Tokyo, Japan). The phage size was determined from at least 10 measurements.

2.5. Host range test

The determination of host range was followed by Addy's description with minor modifications (Addy *et al*., 2019). Briefly, 10 µL of serial dilutions of phage solution (10^8 PFU/mL) was dropped on double-layer agar plates. The top layer was prepared with 100

μL of each strain (OD_{660} of 0.4) mixed with 7 mL of LB. The plates were incubated overnight at bacterial optimum temperature (Table 4.3). The phage sensitivity of the bacteria was confirmed by observation of clear zones or plaques. And then further experiment was performed with plaque essay using the dilutions to confirm if they were affected by the phage.

2.6. Determination of the optimal MOI

Optimal multiplicity of infection (MOI) was determined according to the report of Czajkowski *et al*. (2015). The indicator strain (4×10^8 CFU/mL) at early log phase was infected with phage at different MOI as 0.1, 1, and 10. After a 15 min incubation at 37°C, the mixtures were centrifuged (12,000 ×*g*, 10 min, room temperature). Supernatants were then filtered through a 0.45 μm pore size filter and titrated as described above. The highest titter was considered the optimal MOI. All experiments were averaged from results of triplicate experiments.

2.7. Thermal and pH stability

Phage stability tests were carried out as described in our previous report (Luo *et al*., 2012). Briefly, to examine of thermal stability, phage stocks (1.0×10^7 PFU/mL in SM buffer) were incubated separately for 60 min at temperature ranges 10°C–60°C. For pH stability, 1.0×10^7 PFU/mL phage suspension was added to SM buffer at different pH (pH 3–11), and then incubated at 25°C for 24 h. After incubation, the surviving phages were enumerated by the double-layer agar method. All tests were performed in triplicate.

2.8. One-step growth curve

One-step growth experiments were performed to determine the latent period and burst size as described previously (Addy *et al*., 2019). *E. kobei* M4-VN cells were harvested by centrifugation of early log phase culture (4×10^8 CFU/mL) and resuspended in LB broth. Phage was added at an MOI of 1 and allowed to adsorb for 15 min at 37°C. Cell pellet was washed with 1 mL of fresh LB broth three times. The resuspension was added to 100 mL LB broth and cultured at 37°C in condition of aerobic and anaerobic Samples were taken at 10 min intervals (up to 120 min), and phage titers were then determined by the double-layered agar plate method to obtain one-step curves. The average burst size of the phage was calculated according to the report of Bolger-munro *et al*. (2013).

2.9. Stability with surfactants

Phage stability was tested with various concentrations of five surfactants: Tween 20, SDS, ethanol, skim milk (Becton, Dickinson and Company, Franklin Lakes, NJ, USA), and sucrose. Tween 20 and SDS were chosen a non-ionic detergent and anionic surfactant, respectively. Ethanol was examined as a standard disinfectant. Phage particles were suspended in each solution at various concentrations for 2 h at room temperature. The

resistant capability was investigated by plaque assay. Experiments were performed in triplicate.

2.10. Phage DNA extraction and genome analysis

Purified phage particles were used for phage DNA extraction with TE saturated phenol (pH 8.0) (NIPPON GENE CO., LTD., Tokyo, Japan) and phenol-chloroform and were precipitated with sodium acetate and ethanol. The precipitate was washed twice with 70% ethanol, air dried, and resuspended in TE buffer.

Phage whole genome was sequenced using the PacBio RSII platform (Pacific Biosciences of California, Inc., Menlo Park, CA, USA). The filtered reads were assembled using HGAP version 2.3.0 (Chin *et al*., 2013) and resulted in a 1-contig scaffold. Potential open reading frames (ORFs) were predicted by the Microbial Genome Annotation Pipeline (MiGAP; http://www.migap.org/). The BLAST algorithm at the National Center of Biotechnology Information (NCBI) was used to search similarities. Analysis of the nucleotide sequences was performed with Geneious Prime (Biomatters Ltd, Auckland, New Zealand). Searching for tRNA genes were performed with tRNAscan-SE (Schattner *et al*., 2005) and ARAGORN ver. 1.2.38 (Laslett and Canback, 2004). Phylogenetic analysis was performed with ClustalW ver. 2.1 (https://clustalw.ddbj.nig.ac.jp), and phylogenetic trees were generated using the neighbor-joining method. Prediction of promoter regions was carried out with Neural Network Promoter Prediction (https://www.fruitfly.org/seq_tools/promoter.htmL).

2.11. Structural protein ientification by mass spectroscopy

To analyze virion proteins, phage particles purified by ultracentrifugation were mixed with lysis buffer (62.5 mM Tris-HCl, pH 6.8, containing 5% 2-mercaptoethanol, 2% SDS, 10% glycerol, and 0.01% bromophenol blue (Appendix 10)) and boiled for 10 min. Prepared proteins were then separated by electrophoresis on precast SDS 15% polyacrylamide gels (e-PAGEL E-T/R15L; ATTO Corporation, Tokyo, Japan).

2.12. In-gel digestion

Excised gel pieces were de-stained with 100 µL of 50% acetonitrile containing 25 mM ammonium bicarbonate solution for 1 h at room temperature with gentle agitation. De-stained gel pieces were treated of reduction and alkylation using 100 µL of 10 mM DTT in 25 mm ammonium bicarbonate for 45 min at 56°C and 100 µL of freshly prepared 10 mM iodoacetamide in 25 mM ammonium bicarbonate in the dark for 30 min at 37°C. After gel plugs were dried, 400 ng of sequencing-grade trypsin (Trypsin Gold, Promega, Madison, WI) in 20 µL of 25 mM ammonium bicarbonate were added and incubated for 12 h at 37°C. Digested peptides were recovered from the gel plugs using 50 µL of 50% acetonitrile in 5% formic acid (FA) for 30 min at 25°C. The extracted peptides were concentrated in a speed vacuum concentrator and added to 20 µL of 5% acetonitrile in 0.1% FA.

2.13. Tandem mass spectrometry based proteomics

A Nano-HPLC system (nanoADVANCE, Bruker-Michrom, Billerica, MA, USA) was used to identify proteins automatically using a micro-column switching device coupled to an autosampler and a nanogradient generator. Peptide solution (5 µL) was loaded onto a C18 reversed-phase capillary column (100 µm ID × 30 cm, Zaplous α X Pep C18; AMR, Tokyo, Japan) in conjunction with a Magic C18AQ trapping column (300 µm ID × 10 mm; Bruker-Michrom). The peptides were separated using a nanoflow linear acetonitrile gradient of buffer A (0.1% FA) and buffer B (0.1% FA, 99.9% acetonitrile), going from 5% to 45% buffer B over 50 min at a flow rate of 500 nL/min. The column was then washed in 95% buffer B for 5 min. Hystar 3.2 system-control software (Bruker Daltonics Inc., Billerica, MA, USA) was used to control the entire process. The eluted peptides were ionized through a CaptiveSpray source (Bruker Daltonics) and introduced into a Maxis 3G Q-TOF mass spectrometer (Bruker Daltonics) set up in a data-dependent MS/MS mode to acquire full scans (m/z acquisition range from 50 to 2200 Da). The four most intense peaks in any full scan were selected as precursor ions and fragmented using collision energy. MS/MS spectra were interpreted and peak lists were generated using DataAnalysis 4.1 and Biotools 3.2.

2.14. Protein identification

The filtered data were searched on the Mascot 2.2 server (Matrix Science) using the NCBInr (NCBI 201805) database and custom expected protein databases for EspM4VN. Fixed modification was set on cysteine with carbamidomethylation. Variable modification was based on methionine with oxidation and asparagine/glutamine with deamidation. Maximum missed cleavage was set to two and limited to trypsin cleavage sites. Precursor mass tolerance (MS) and fragment mass tolerance (MS/MS) were set to 100 ppm and ±0.6 Da, respectively. Positive protein identifications using a threshold of 0.05 were used. Peptides scoring <20 were automatically rejected, ensuring all protein identifications were based on the reliable peptide identifications. Protein identification was set to require at least two unique peptides. Homology searches against matched protein sequences were conducted using the BLASTp program (http://www.ncbi.nlm.nih.gov/).

2.15. Accession number

The whole genome sequence of phage EspM4VN was submitted to DDBJ under accession number LC373201.

3. Results

3.1. Isolation and morphology of EspM4VN

EspM4VN formed large semi-turbid plaque on a lawn of the strain M4-VN. The

plaques were approximately 0.5 mm in diameter with a halo effect (2.5 mm) (Figure 4.1A). Purified phage particles from the plaques were applied to TEM analysis. The phage particles were composed of an icosahedral-shaped head approximately 100 nm in diameter, a neck, and a contractile sheath 100 nm in length and 18 nm in width (Figure 4.1B). Intermediate structures with both prongs and stars appeared as have been observed in some *Ackermannviridae* phages such as *Agtrevirus* Ag3, *Limestonevirus* JA15, and XF4 (Anany *et al*., 2011, Day *et al.*, 2017). Thus, the phage morphology of EspM4VN is that of the genus *Agtrevirus* belonging to the *Ackermannviridae* family according to ICTV rule 3.12 (Adriaenssens *et al*., 2018).

3.2. Host range of EspM4VN4

The host range of EspM4VN was determined with six strains of *Enterobacter,* nine strains of *Pectobacterium*, two *Acinetobacter* species, *A. baumannii* TL5, *P. dispersa* TL3, and *E. coli* DHα5 (Table 4.3). EspM4VN formed clear plaques only on *E. kobei* M4-VN. A phage suspension of 4×10^6 PFU/mL was applied to strain M4-VN at a low density level for a similar host range experiment used by others (Adriaenssens *et al*., 2012). It showed a quite narrow host range even in the same *Enterobacter* cluster.

3.3. Thermal and pH stability of EspM4VN

EspM4VN kept stable infectivity at temperatures 10°C–50°C (Fig. 4.2A) and pH 4–10 (Figure 4.2B). The phage was deactivated at 70°C and pH 3 and 12. Related phages infecting *Dickeya* strains were reported to be stable at temperatures from 4°C to 37°C but not stable at 50°C for 24 h, and they showed stability at pH 5–12 (Czajkowski *et al*., 2014).

3.4. One-step growth of EspM4VN

At MOI = 1, a one-step growth curve showed that the latent phase of EspM4VN was 20 min, rise period was 10 min, and an average of 131 and 122 phage particles were released from each absorbed cell in aerobic and anaerobic conditions, respectively (Figure 4.3 (A) and (B)). The latent phase of EspM4VN was much shorter than LIMEstone 1 and 2 (60 min and 65 min, respectively) or soilborne lytic phages (40 min) infecting *Dickeya* strains (Adriaenssens *et al*., 2012; Czajkowski *et al*., 2014).

3.5. Stability of EspM4VN within various chemicals

The stability of EspM4VN was analyzed against detergents and surfactants (Figure 4.4). Phage viability in sterile water was approximately 50% or less compared with that in SM buffer. Up to 1.5% (v/v) Tween 20 and ethanol did not inhibit phage infection. By contrast, the phage showed resistance (>70%) to 0.1% to 1.5% SDS. EspM4VN was stable in skim milk at all concentrations examined.

3.6. Genomic analysis of EspM4VN

Nucleotide sequence analysis of the EspM4VN showed the genome size to be 160,766 bp with a G + C content of 50.7%, and 219 ORFs were annotated (Figure 4.5 and Appendix 1). BLAST search revealed that the EspM4VN genome showed high similarity to *Agtrevirus* SKML-39 (JX181829; identity 95.3%), *Limestonevirus*, Coodle (MH807820; 95.2%), PP35 (MG266157; 95.2%), JA15 (KY942056; 95.1%), LIMEstone1 (93%), phiDP23.1 (93%), φD3 (93%) and *Agtrevirus* Ag3 (90%) (Appendix 2). EspM4VN as a *Ackermannviridae* infects *Enterobacter*, and its genome has significant homology with other *Enterobacter* phages such as myPSH1140 (MG999954; Manohar *et al*., 2019), CC31 (GU323318; Petrov *et al*., 2010), and vB_EaeM_φEap-3 (KT321315; Zhao *et al.*, 2019). Analysis demonstrated that EspM4VN tRNA species and their codons were as follows: $tRNA^{Ser}$ (TCA), $tRNA^{Asn}$ (AAC), $tRNA^{Tyr}$ (TAC), $tRNA^{Ser}$ (AGC), and $tRNA^{Asp}$ (GAT). Two unique regulatory motifs, TTCAAT $[N_{14}]$TATAAT and CTAAATAcCcc, were found in their entireties in EspM4VN genome (positions 7219–7244, 143195–143220, 152828–152853, 158832–158857, 61226–61251 and 65377–65387, 70604–70614, 98591–98601, 101772–101782, 53611–53621). Phylogenetic trees of structural proteins (major capsid protein and baseplate) and enzymes (DNA polymerase and ligase) showed that EspM4VN formed an independent cluster among *Ackermannviridae* phages (Figure 4.6). Based on the phylogenetic analysis of ligase and major capsid protein, EspM4VN was positioned in the same cluster of *Agtrevirus*, *Shigella* phage Ag3, and *Salmonella* phage SKML-39 and not in the same cluster of *Limestonevirus* and *Kuttervirus* belonging to *Ackermannviridae* (Figure 4.6(B) and (C)). The baseplate protein of EspM4VN was not positioned in *Agtrevirus* but in *Limestonevirus* (Figure 4.6D).

3.7. Protein analysis of EspM4VN

For identification and further characterization of bacteriophage EspM4VN protein, SDS-PAGE and MS analysis of the purified phage were performed. Ten major protein bands were found in the gel and band patterns including major band around 50 kDa (band 6 in Figure 4.7) were similar to those of *Dickeya* spp. bacteriophage (Czajkowski *et al*., 2015), *Shigella boydii* phage Ag3 (Anany *et al.*, 2011) and broad host range phages, φPD23.1 and φD10.3 (Czajkowski *et al*., 2015).

Ten protein bands excised from the SDS-PAGE gel were digested by trypsin. Proteomics analysis was performed using quadrupole time-of-flight tandem mass spectrometer. Peptides were identified by matching the identified peaks to the predicted protein sequence library of the custom EspM4VN gene including 219 ORFs or NCBInr viruses database using in-house Mascot server. In all 10 gel samples, proteins matched against EspM4VN proteins showed much higher Mascot scores than other bacteriophage proteins including *Dickeya* and *Shigella*, indicating that the purified bacteriophage was from EspM4VN gene.

A total of 18 EspM4VN proteins were identified and are listed in Figure 4.7 along with gene number, calculated molecular mass, corresponding protein sequence

coverage, and predicted annotation by BLASTp protein similarity search. Because 18 identified protein bands were excised from stained bands, these proteins were suggested to be highly expressed proteins. Of these, 16 were identified as structural proteins and two were classified as uncharacterized or unknown proteins. Genes 105 (VriC), 114 (gp18), 116 (gp20), and 121 (gp23) were identified from *Dickeya* phage φD10.3 (Czajkowski *et al.*, 2015).

Some genes from identified proteins were sequentially located in the genome sequence including genes 25–26, 101–104, and 114–115. BLASTp search indicated that proteins of genes 25–26 were both baseplate subunits and those of 101–104 were all tail-spike-related proteins. Proteins from genes 114–115 were also identified as tail proteins. These proteins were multimeric proteins for function, and all required high expression levels. Analysis of the EspM4VN genome sequence suggested that similar functions or structural proteins were clustered at certain regions as is the case in other phages.

Besides protein identification, Mascot analysis also provides label-free, relative quantitation of proteins, the exponentially modified protein abundance index (emPAI) in a mixture based on protein coverage by the peptide matches in a database search result (Arike and Peil, 2014). The protein of gene 121 yielded the highest emPAI values at gel #6 and #7 (emPAI: 17.9 and 11.2, respectively) and identified as a major capsid protein, gp23. The gp23 protein nicely migrated at 50 kDa in the gel according to the calculated molecular weight and showed the highest density of the coomassie brilliant blue stain. Other SDS-PAGE analyses showed that the migrations and band densities of EspM4VN were highly similar to those of the *Dickeya* phage φD10.3 sample (Czajkowski *et al.*, 2015).

Tryptic fragments from the products of genes 92, 100, 114, and 116 included plausible N-terminal sequences (Appendix 3). MS/MS analysis showed that all of the peptides started from the amino acid following Met (next to Met) because N-terminal Met could be co-translationally cleaved by methionine aminopeptidase. The gene 116 product included not only N-terminal end Ala^{2} but also a C-terminal end Glu^{563}. The calculated molecular weight of this protein (63.23 kDa without Met^{1}) closely matched with band migration (65 kDa) and was identified as a portal vertex protein, gp20.

Among the plausible 219 genes of EspM4VN, gene analysis indicated that 18 proteins did not start from Met at the N-terminal end but Val or Leu. This suggests that EspM4VN uses only the prokaryote expression system as the start codon in the eukaryote expression system is almost exclusively Met (AUG). Gene115 was found in the gel band #10 sample and included Val as its plausible N-terminal sequence. The expected N-terminal tryptic peptide VTVNFPAFVAGSDTIR was detected seven times via MS/MS with a high Mascot score of 80. In the genome DNA sequence between the stop codon (TGA) of gene 114 and the start codon (GTG) of gene 115, there were no possible Met (AUG) codons found. Based on DNA sequence analysis and MS/MS analysis, I concluded that gene 115 was expressed from Val at the N-terminal amino acid (Figure 4.8). BLASTp analysis and quantification analysis showed that gene 3115 was identified

as a tail tube protein, gp19, at high expression because its emPAI was 4.5.

Gene 101 was found at gel band #2 and had highest emPAI (2.21) in the gel sample. The molecular weight of gene 101 product was 119,592 Da including 1,109 amino acids. Of the sequence, 52% from the N-terminus to the C-terminus was observed by MS/MS analysis. BLASTp analysis suggested that gene 101 has two unique domains: N-terminal and C-terminal structures. The N-terminal region (1–413) was closely related to other bacteriophage tail protein (Figure 4.9B). Interestingly, the C-terminal region (436–1109) was similar to a hypothetical protein from bacteria including *Enterobacter* (WP_117582064.1) or *E. coli* (APK16060.1) (Figure 4.9C). BLASTp analysis showed that genes 101, 103, and 104 had pectin lyase fold motif, a typical motif in tail-spike proteins, including the beta helix structure. Gene 101 has a conserved identical sequence region (L^{371}-D-X-K-T-V-I-Y-D^{379}) with gene 102 (42–50) and gene 103(42–50). Gene 104 has a part of the identical sequence (V^{46}-I-Y-D^{49}) and has a conserved motif of the tail-spike (N-terminal domain at 97–153).

4. Discussion

Bacterial soft rot is one of the most common diseases of important crops in the world. Therefore, several control strategies including physical methods and chemical treatments using copper sulfate, dimethylammonium chloride, sodium hypochlorite, formaldehyde, and antibiotics have been applied. However, these approaches have the risk of increasing environmental burden and health hazards (Azaiez *et al*., 2018). The genus *Enterobacter* contains 14 species and two subspecies (Zhu *et al*., 2017). Among these, Ecc including *E. cloacae*, *E. asburiae*, *E. hormaechei*, *E. kobei*, *E. ludwigii* and *E. nimipressuralis* have genomic heterogeneity and are therefore difficult to identify (Paauw *et al*., 2008). In this study, I isolated soft rot pathogenic bacteria for usage as phage hosts. Phylogenetic analysis of the isolates with their 16S rDNA sequences showed that they located in the Ecc cluster but had unconfirmed identification in the species level (Figure 2.3). It is not enough to prove that isolated strains, including the M4-VN strain, are the main villain for soft rot; therefore, the function of the strains for maceration will be examined with both *in vitro* and in-filed experiments. If so, phage therapy with phage EspM4VN might be a potential solution for soft rot.

With morphological observation, EspM4VN belongs to *Ackermannviridae* family and resembles *Dickeya* phages φXF4 (Day *et al*., 2017), but it does not possess short tail spikes as observed in φXF4. Instead of the gene cluster encoding short-tail-spike proteins of φXF4 (TSP genes), four genes (genes 101–104) consisted of gene cluster for tail-spike proteins in EspM4VN (Figure 4.9A). The difference in gene organization may cause morphological differentiation of tail structures between EspM4VN and φXF4. In other words, EspM4VN tails with unfolded tail entities, displaying an umbrella-like structure (Fig. 4.1B). Based on alignment of thymidylate synthase genes, EspM4VN possess the thymidylate synthase gene, a homologue of deoxyuridylate hydroxymethyltransferase gene of *Escherichia* phage vB_EcoM Sa157lw (AYC62411)

with 71.43% homology.

Phage therapy had been increasingly researched and applied as biological control agents against plant pathogens since the early 1920s (Balogh *et al*., 2010; Gill and Abedon 2003, Jones *et al*., 2007). While the application of phages to an integrated plant disease management strategy is relatively simple, cheap, and effective, phage efficacy is highly dependent on general environmental factors and the susceptibility of the target bacterium. Symptomized soft rot is a soil-borned infection; hence phage was isolated from farmland, where the above mentioned plants have been cultivated. In this study, EspM4VN isolated from soil sample was capable of inhibiting only *Enterobacter* strain M4-VN, indicating that EspM4VN has a narrow host range. This is quite similar to other phages such as WS-EP 19, WS-EP 13, WS-EP 20, WS-EP 26, WS-EP 28, WS-EP 57, WS-EP 32, and WS-EP 94, which infect specific *Enterobacter* strains isolated from milk powder and other foods (Loessner *et al*., 1993). Therefore, EspM4VN may be used in phage cocktails and classification as phage typing for isolated *Enterobacter* strains. One of the reasons for its restricted host range is the characteristic tail structure of the EspM4VN (Figs. 1B and 8). One-step growth curve demonstrated that EspM4VN has a short latent phase and quickly reaches burst size. This phage could be a potential epidemic inhibitor for immediate treatment after an outbreak of plant disease. Furthermore, thermal and pH stability tests indicate that it may be applied in tropical and subtropical zones. In the production industry, pesticides are normally formulated in the presence of surface active agents, which can change pesticide adsorption in the soil-water system. Addition of Tween 20 at low and high concentrations increased the adsorption of diazinon and atrazine into soil (Iglesias-Jiménez *et al*., 1996). Phage EspM4VN resisted Tween 20, ethanol, SDS, and sucrose at various concentrations as well. Skim milk is one of the best substances for product formulation and long-term storage. In addition to some chemicals used in this study, various materials to extend the life of the phage following exposure to various physical factors may be examined (Jones *et al*., 2007).

Whole genome analysis of EspM4VN revealed that five tRNAs (Ser1, Ser2, Tyr, Asn, and Asp) were identified. Those tRNAs may correspond to the short latent phage and large burst size because of their positive influences on reproduction in host cells. However, the homologous phage genome *S. boydii* phage Ag3 had four similar tRNAs and a 52 min latent phase. This contrast may be explained by ampleness of the tRNAs presented in the hosts. In this case, the tRNAs of the phage would have less of an effect on the phage environment during infection (Bailly-Bechet *et al*., 2007).

Like other phage genome sequences, similar functions or structural proteins were clustered at certain regions in the EspM4VN genome. These results suggested that gene locations were closely related with protein expression. However, genes and gene products of EspM4VN exhibit distinctive features (Figure 4.9). Genes 101 to 104 comprised a gene cluster for tail-spike proteins. Similar cluster structures comprising four genes were observed in *Agtrevirus* (Ag3) and *Limestonevirus* (RC-2014) genomes (Figure 4.9 (A)). In the genome of *Dickeya* phage XF4 (taxid:1983656), whose tail-spike structure is

similar to that of EspM4VN, there are three genes in the tail spike protein cluster. Phage XF4 has a gene structure in which three tail spike genes (TSP1-3) have been sequentially located at the region (118454–122334) (Day *et al*., (2017). TSP1, TSP2, and TSP3 have sequence similarities with those of genes 104, 103, and 101. In the gene cluster of the tail spike protein of *Escherichia* virus CBA120 (taxid:1987159), TSP1 gene (*orf213*) and TSP3 gene (*orf211*) show 68% and 52% similarity with genes 101 and 103, respectively (Figure 4.9A). TSP2 and TSP4 of CBA120 have 52% similarity with the product of gene 104. CBA120-TSP2 binds its substrate and functions to recognize the host receptor (Plattner *et al*., 2019). These findings suggest that EspM4VN tail-spike proteins might form heteromeric structure in the phage. Only the product of gene 102 shows high homology with the hypothetical proteins of *Enterobacter mori* (WP_157929994.1), *E. coli* (OJN39208.1), *E. cloacae* (WP_063860747.1), and *Klebsiella aerogenes* (WP_063963681.1). Therefore, the unique structure of the product of gene 102 might be caused by the narrow host range of EspM4VN and the possibility that some gene origin of EspM4VN could be from bacterial chromosomes.

Also, gene 101 might be fused with phage-derived gene and bacterial genome-derived gene (Figure 4.9B, C). Homology analysis of the gene 101 product indicates that its N-terminal region closely resembles a right-handed beta helix region of phage proteins SKML-39 (identity; 94%, 1–423 aa), Ag3 (identity; 94%, 1–423 aa), RC-2014 (identity; 90%, 1–423 aa), phiDP23.1 (identity; 89%, 1–423 aa), and LIMEstone1 (identity; 84%, 1–426 aa), respectively (Figure 4.9 B). The C-terminal region of the gene 101 product has high homology with enterobacterial gene products described as follows: *Enterobacter* sp. AM17-18 (identity; 57%, 436–1106 aa), *E. coli* D4 (identity; 56%, 439–1106 aa), *E. cloacae* subsp. *cloacae* GN2616 (identity; 56%, 439–1106 aa), *Serratia* sp. Ag2 (identity; 38%, 439–1106 aa) and *E. kobei* GN02266 (identity; 36%, 439–1106 aa) (Figure 4.9B). From gene locations in their chromosome, some of the homologous genes might be inserted in the glycan metabolic, cell wall, and amino acid biosynthesis pathways in *Enterobacter* sp. AM17-18, *E. coli* D4, and *E. cloacae* subsp. *cloacae* GN2616 (Figure 4.9C). Matilla *et al*. suggested φXF1, φXF3, and φXF4 are very efficient generalized transducers capable of transducing chromosomal markers at high frequencies (Matilla *et al*., 2014). The above mentioned heterogeneous gene structure is reasonable for phage function, as corresponding genes in *Serratia* sp. Ag2 and *E. kobei* GN02266 genomes locate within phage tail gene clusters. This suggests the genes 101 and 102 of EspM4VN existed partially from lysogenic phage genomes or was fused after phage transduction. However, the question remains why a tail spike protein, which is a necessary accessory protein, would start from a rare start codon and whether transcripts longer than 10 kb could be polycistronic-synthesized. These and other genetic regulation characteristics of EspM4VN transcription are topics for future studies. Lysogenic related genes such as *int*, *xis*, and *xerD* were not observed in the EspM4VN genome; this is a beneficial property of the phage for biocontrol of the plant pathogenic bacteria (Álvarez *et al*., 2019).

Table 4.3. Bacterial strains used in this work and their sensitivities against EspM4VN

Strain	Source	Accession number	Plaque formation	Reference
Pantoea dispersa TL3	Dragon fruit	LC498100	-	This work
Enterobacter sp. B3	Cabbage	LC498101	-	This work
Enterobacter sp. B5	Cabbage	LC498102	-	This work
Enterobacter sp. BC7	Cabbage	LC415135	-	This work
Enterobacter kobei M4-VN	Potato	LC415612	+	This work
Enterobacter sp. KT	Potato	LC416590	-	This work
Acinetobacter baumannii TL5	Dragon fruit	LC423530	-	This work
Acinetobacter baumannii NBRC 109757	Urine	Z93435	-	Bouvet and Grimont (1986)
Acinetobacter junii NBRC 109759	Urine	Z93438	-	Bouvet and Grimont (1986)
Enterobacter asburiae NBRC 109912	Human, lochia exudate	CP011863	-	Hoffmann *et al*. (2005)
Escherichia coli DHα5			-	RBC Bioscience
Pectobacterium carotovorum subsp. *carotovorum* 489-4	Cabbage		-	Plant Pathology Laboratory, Kyushu University
Pectobacterium carotovorum subsp. *carotovorum* 489-5	Cabbage		-	Plant Pathology Laboratory, Kyushu University
Pectobacterium carotovorum subsp. *carotovorum* 493-1	Potato		-	Plant Pathology Laboratory, Kyushu University
Pectobacterium carotovorum subsp. *betavasculorum* ATCC 43762^T	Sugar beet	U80198	-	Gardan *et al.* (2003)
Pectobacterium carotovorum subsp. *wasabiae* ATCC 43316^T	Japanese horseradish	U80199	-	Gardan *et al.* (2003)
Pectobacterium carotovorum subsp. *atroseptica* ATCC 33260^T	Solanum tuberosum	Z96090	-	Gardan *et al.* (2003)
Pectobacterium carotovorum subsp. *carotovorum* ATCC 15713^T	Potato	U80197	-	Hauben *et al.* (1998)

Dickeya chrysanthemi pv. *Zeae* 511-3 S2	Corn	-	Plant Pathology Laboratory, Kyushu University
Pectobacterium milletiae 1		-	National Institute of Agrobiological Sciences, Japan

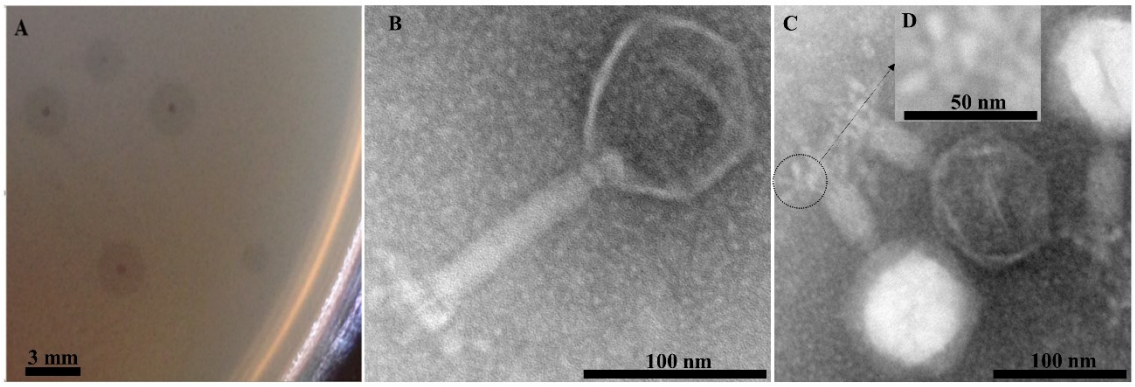


Figure 4.1. Plaques formed in the top agar layer (A), morphology of the extended tail phage EspM4VN (B), contracted tail phage EspM4VN (C), and displaying an umbrella-like structure (D). Bars indicate 3 mm (A), 100 nm (B) and (C), and 50 nm (D).

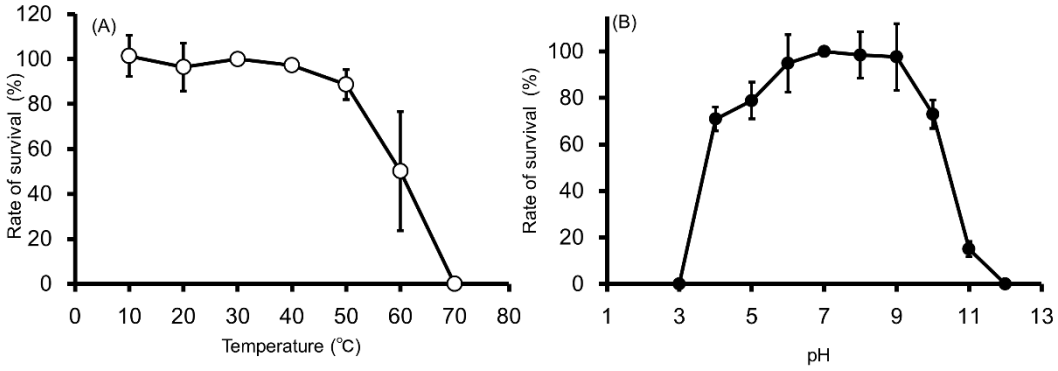


Figure 4.2. Thermal and pH stability of EspM4VN. Course of survival ratio (%) of EspM4VN particles during heat (A) and pH (B) treatment are plotted. The mean titer is shown from triplicate assays.

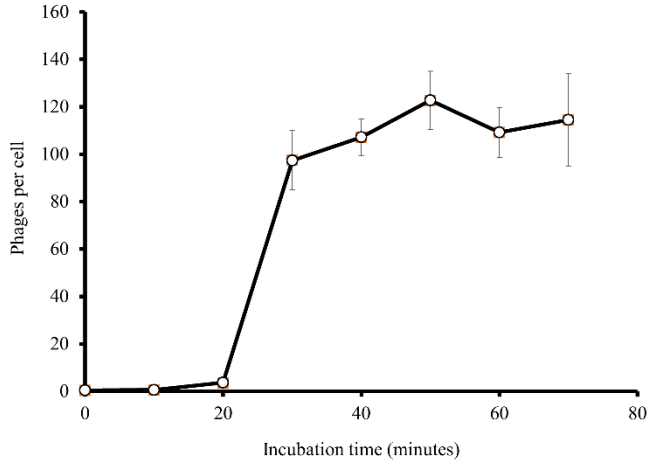


Figure 4.3. (A) One-step growth curve of the EspM4VN in aerobic condition.

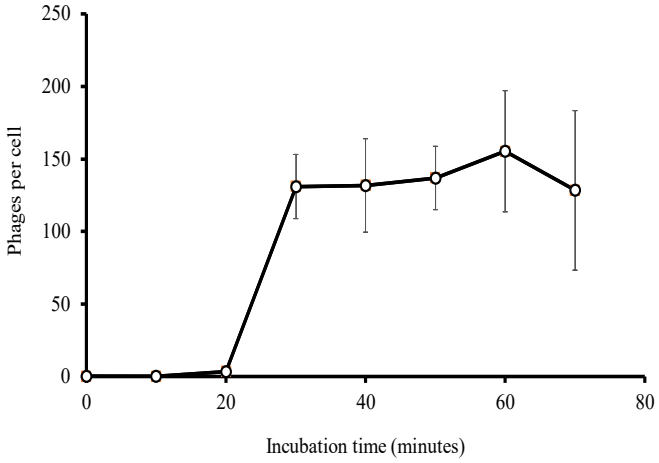


Figure 4.3. (B) One-step growth curve of the EspM4VN in anaerobic condition.

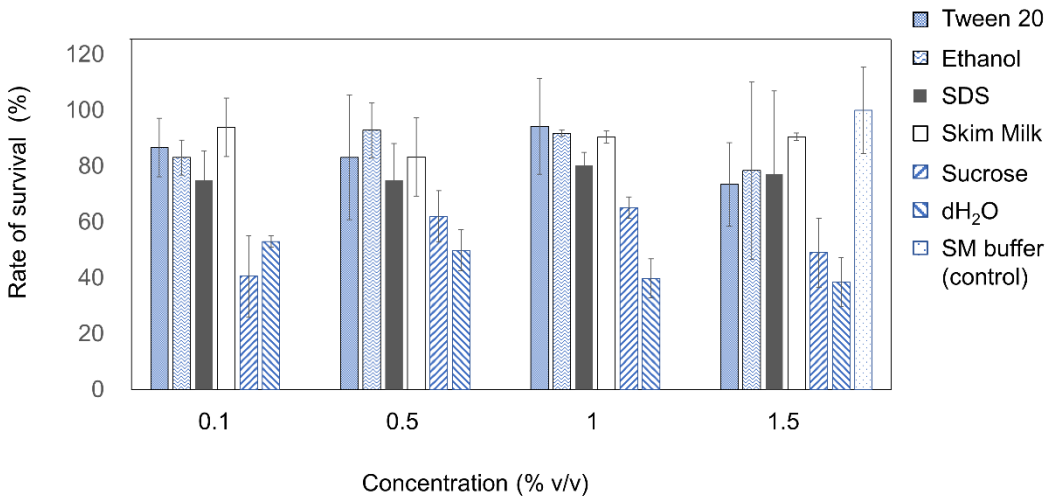


Figure 4.4. Chemical sensitivity of EspM4VN. Phage particles were inoculated in LB containing different concentrations of Tween 20, ethanol, SDS, skim milk, and sucrose from 0.1% to 1.5%. SM buffer (outlined bars) was used as a control. Quantification was achieved by plaque assay. The plates were incubated at 37°C for 12 h, and the phage growth was assayed by PFU. Experiments were performed in triplicate on three different occasions and means ±SD are shown.

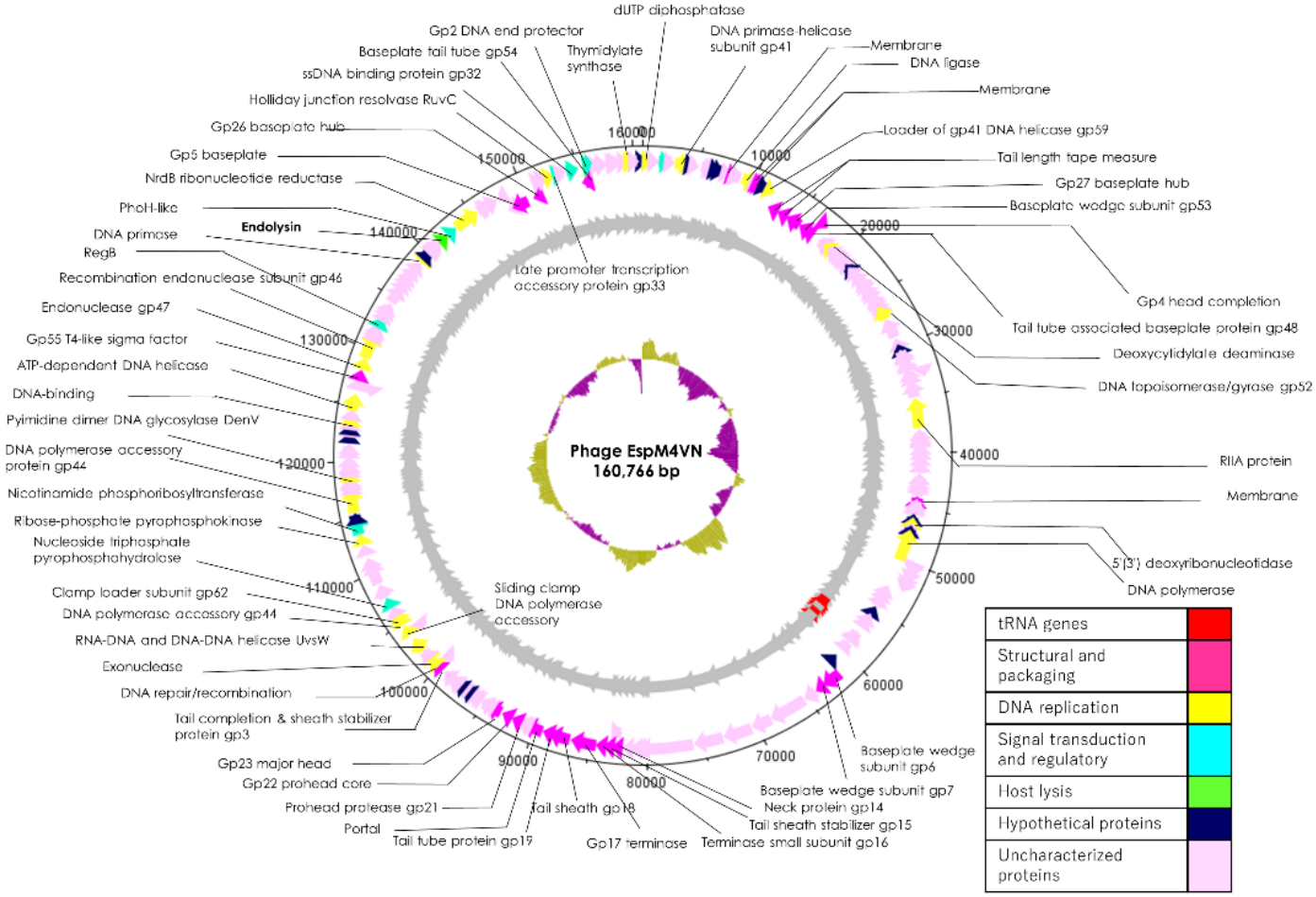


Figure 4.5. Genome map of EspM4VN. The outer lane represents genes in the plus strand. The next lane illustrates genes on the minus strand. The lane with black peaks and valleys indicate the GC content, while the innermost lane (green and violet) shows a GC skew analysis.

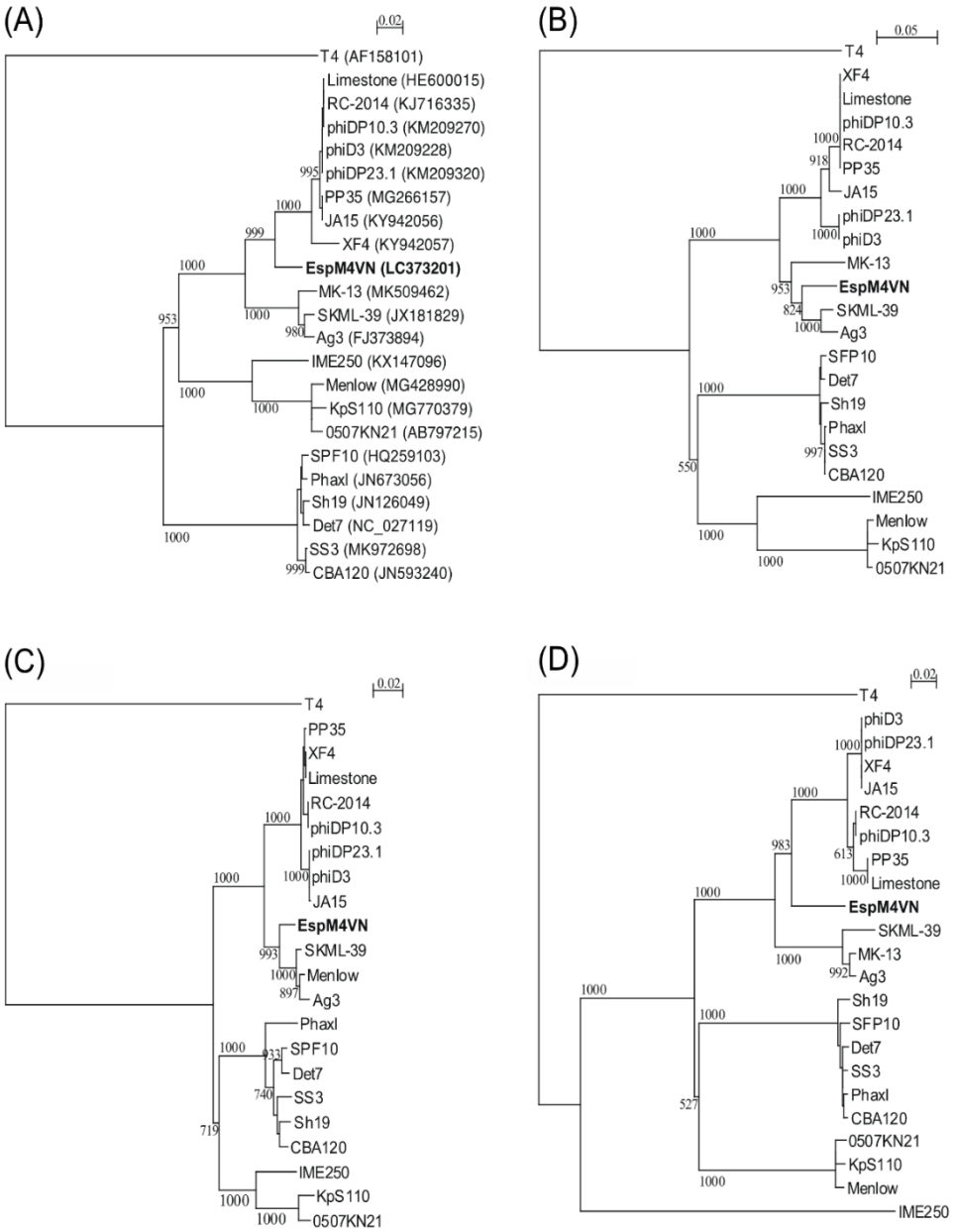
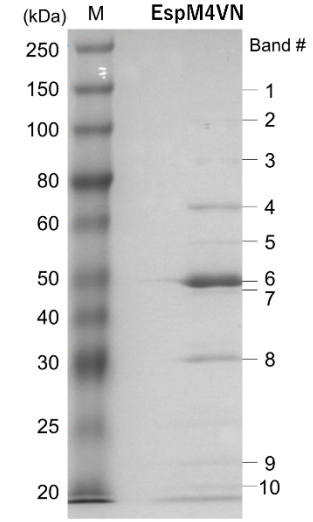


Figure 4.6. Phylogenetic trees generated based on (A) DNA polymerases, (B) DNA ligase, (C) major capsid protein, and (D) baseplate hub subunit of EspM4VN and homologous proteins from other phage members of the *Ackermannviridae* family. Nucleic acid sequences were compared using ClustalW, and phylogenetic trees were generated using the neighbor-joining method. Bar indicates the number of substitutions per sequence position. Numbers in brackets show gene ID.

(A)



(B)

Band #	Gene #	Predicted gene product	ESI-MS/MS analysis		
			MW	Coverage (%)	emPAI
1	105	VriC protein	174760	40	1.35
2	101	Phage tail fiber protein	119592	52	2.21
3	104	Tail fiber protein	92762	50	1.12
	103	Tail fiber protein	93065	35	1.05
4	114	Putative tail sheath protein gp18	68949	61	3.35
	116	Portal vertex protein gp20	63475	53	2.22
	102	Tail fiber protein	63798	61	1.88
	98	Putative baseplate wedge subunit gp6	64822	31	0.98
5	182	Hypothetical protein	60019	59	5.38
6	121	Major head protein gp23	48013	74	17.9
7	100	Unknown structural protein	42806	18	0.32
8	121	Major head protein gp23	48013	55	2.29
9	56	Putative tail tube protein	27677	17	0.71
	26	Putative baseplate wedge subunit gp53	21408	23	0.51
10	92	Phage tail needle protein	18424	86	4.76
	115	Putative tail tube protein gp19	17163	54	4.47
	134	Putative tail completion & sheath stabilizer protein gp3	18502	39	1.21
	160	Uncharacterized protein	17269	13	0.4

Figure 4.7. Bacteriophage EspM4VN structural proteins separated in SDS-PAGE gel (A) and identification of structural proteins with ESI-MS/MS (B). (A) For SDS-PAGE, EspM4VN particles (ca. 10^{12} PFU) were mixed with lysis buffer and then boiled for 10 min. Phage proteins were separated in a 15% acrylamide SDS-PAGE gel for approximately 19 h at 50 V at 22°C. The bands were stained with Coomasie Brilliant Blue according to the protocol provided by the manufacturer. For ESI-MS/MS analysis of phage structural proteins, protein bands obtained from SDS-PAGE were excised from the gel with a sterile scalpel.

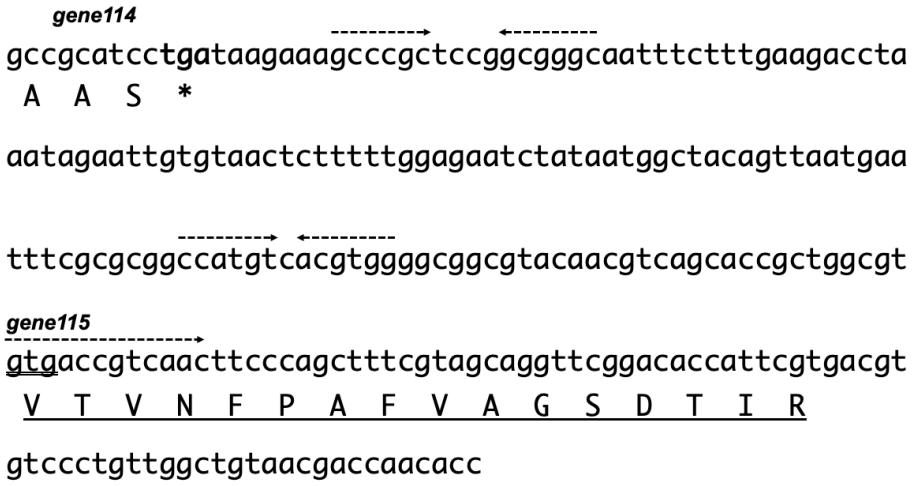


Figure 4.8. The nucleotide sequence of the intergeneric spacer region and partial encoding regions of genes 114 and 115. Dotted arrows represent inverted repeats, and an asterisk denotes the stop codon. The detected amino acid residues from MS analysis are underlined. The nucleotide sequence is from position 88621 to 68627 of GenBank Database, accession no. LC373201.2.

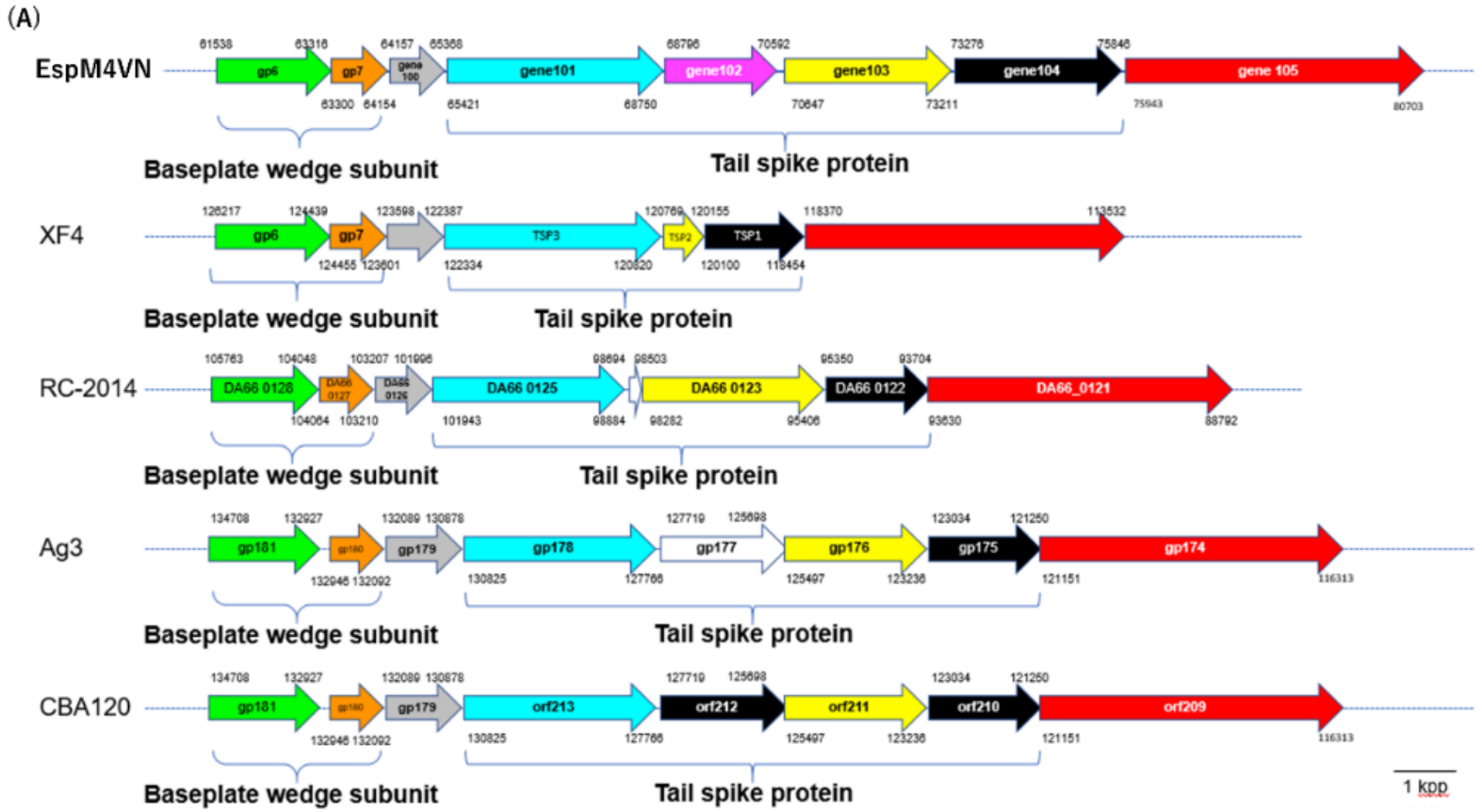


Figure 4.9A. Gene clusters of phage tail proteins in EspM4VN and *Ackermannviridae* family phages. Same colored arrows indicate homologous genes, respectively. Blanked arrows show no homology with gene 102 in EspM4VN. Numbers indicate positions of ORFs in the genome.

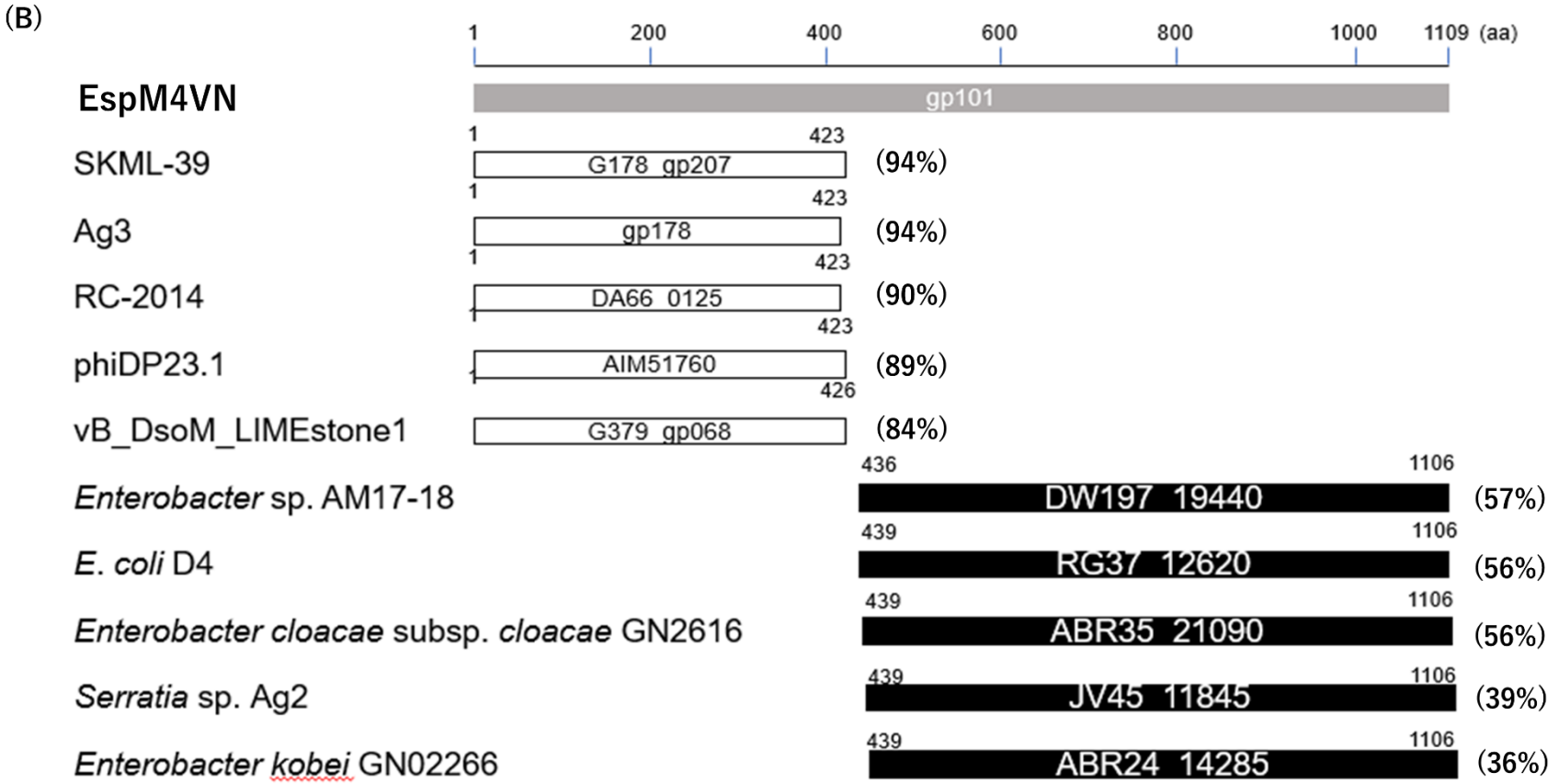


Figure 4.9B. Homologous amino acid sequence distribution of gene 101 product (gray bar). White and black bars show partial regions of phage gene products and bacterial chromosomal gene products, respectively. Characters in white and black bars show locus tag of proteins. Numbers indicate amino acid positions in each protein. Percentages in brackets indicate identity of protein sequences.

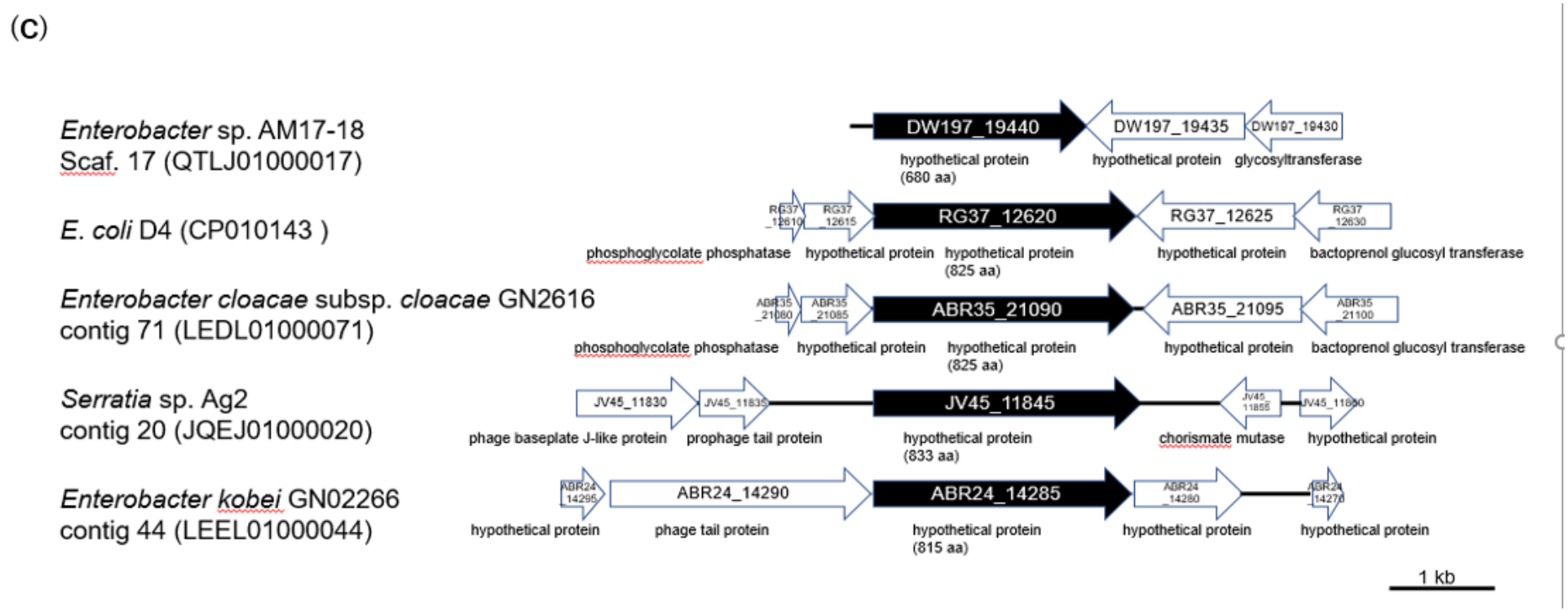


Figure 4.9C. Gene organization adjacent to gene 101 homologues in bacterial chromosomes. Black arrows indicate gene 101 homologous genes. Predicted gene products are shown below each arrow.

Chapter 5. Conclusions

Bacterial soft rot is a disease that severely affects agricultural ecosystems. It has been recorded as a real issue in vegetable storage and on the field of various vegetables and fruits in very early history. Several causative pathogens that belonged to different families have been reported. Among them, SRP species were mainly concerned while SRE species were less reported and understood. SRE was first reported when *E. cloacae* was identified as a causative pathogen of soft rot papaya. Later on, *E. cloacae* subsp. *dissolvens, E. cowanii, E. asburiae, E. mori,* and other *Enterobacter* species were reported as plant pathogens in tropical fruits and forest trees. Many methods including using disease-resistant breeds, physical treatment, chemical treatment, and biological treatment have been applied to control soft rot. However, these methods mainly focus on controlling and managing SRP disease. In addition, the prevalence of chemical application is threatening human health, environment, and biodiversity. Therefore, phage therapy provides an alternative application as a biocontrol strategy. Phages are bacterial viruses that can infect specific bacterial cells. The advantages of phages are free toxicity to human, animals, and environment, and abundant population while the drawbacks come from its selective infection, and environmental factors. These are raising critical questions about the identity of pathogenic bacteria, phages' characteristics.

Chapter 2 describes identification and characterization of seven bacterial isolates from soft rot samples in Vietnam. All strains are gram-negative, colony grown on LB agar were slightly yellow, smooth, translucent and convex, adhering to the agar surface. These strains have rod-shape with various length ranges from 1.1 μm to 2.9 μm, width ranges from 0.6 μm to 1.2 μm. The seven isolates were able to grow in a range of temperatures from 15 to 45°C and eased the growth at 10 and 50°C. Among the seven strains, B3, B5, BC, and M4-VN have the shorter lag phase and log phase than the remains. The lag phases of strains B3, B5, BC and M4-VN lasted from 1.5 to 2.5 h, and their log phases ranged from 6-7 h at 30°C and 37°C. However, at 37°C, strains KT and TL3 got an early log phase after 2 h, but at 30°C, this stage took 4 h. On the other hand, strain TL5, took only 30 min for the lag phase at 30°C and 37°C and lift to stationary phase after 8 h. Genome analysis showed closed relationship between TL3 and *P. dispersa,* TL5 and *A. baumannii,* and the remains and *Enterobacter* species. Phylogenetic tree showed *P. cypripedii* to be different cluster from TL3. Furthermore, biochemical test revealed high similarity physiological characteristics between TL3 and *P. dispersa.* Thus, TL3 was identified as *P. dispersa* species. The phylogenetic tree also depicted TL5 to be closed with only *A. baumannii.* Strains TL5, *A. baumannii* ATCC 17978 and *A. baumannii* ATCC BAA-1790 created a different cluster from the others. This evidence proves that TL5 belongs to *A. baumannii* species. Whereas, B3, B5, BC, KT and M4-VN were not distinguishable due to low power resolution and unclear physiology.

Chapter 3 describes whole genome sequence analysis of *Enterobacter* sp. M4-VN isolated from potatoes with soft rot and its phylogenetic position in genus

Enterobacter. In chapter 2, the results revealed that the identification at the species level of the isolated strains was questionable. Therefore, whole genome sequencing showed an accurate resolution for the identification based on the scores of dDDH and ANI. Surprisingly, the data showed a high similarity between *Enterobacter* sp. M4-VN and *E. kobei* DSM 13645 with the remarkable values of dDDH (92.6%) and ANI (99.07). This is strong evidence to conclude that the strain M4-VN belongs to *E. kobei*. In addition, based on genomic annotation, genes encoding for plant pathogenicity were predicted. Genes encoding the pectinesterase and virulence factors which were found in SRP species assumed the phytotoxicity of *E. kobei* M4-VN. The annotation also revealed that M4-VN's chromosomes contained four prophages. Some prophages provide the host bacteria phage-resistant traits from infection by other phages based on Sie mechanism. However, Sie system's genes were not found in *E. kobei* M4-VN.

Chapter 4 describes the identification and characterization of virulent phage EspM4VN to control *E. kobei* M4-VN isolated from soft rot. The phage EspM4VN was isolated from soil in Bac Ninh province, Vietnam. The phage has an icosahedral head (100 nm in diameter) and a contractile tail (100 nm in length, 18 nm in width), and prongs and stars-liked tail spikes. These morphological characteristics were similar to those of *Shigella* phage Ag3 (*Agtrevirus*) and *Dickeya* phage JA15 and φXF4 (*Limestonevirus*) of the *Ackermannviridae* family. In addition, EspM4VN showed narrow host range due to forming clear plaques only on the strain M4-VN among the 20 tested including 7 isolates and existing strains. EspM4VN was stable in a range of 10°C to 50°C and pH from 4 to 10, and was resistant to ethanol and Tween 20 up to 1.5% (v/v). One-step growth curve described that EspM4VN had a latent phase of 20 min, a rise period of 10 min, and the average released virion numbers of 131 and 122 in aerobic and anaerobic conditions, respectively. Thus, EspM4VN is applicable in tropical and subtropical areas and useful for phage typing. Nucleotide sequence analysis of the phage EspM4VN revealed the genome size to 160,766 bp and 219 ORFs were predicted. The BLAST search showed similarity to those of *Shigella* phage Ag3, *Salmonella* phage SKML-39, *Dickeya* phage Coodle, PP35, JA15, and Limestone. The phylogenetic analysis of major capsid protein and DNA ligase showed EspM4VN to be positioned *Agtrevirus*, while the phylogenetic tree of baseplate proteins and DNA polymerase placed EspM4VN into the cluster of *Limestonevirus*. Moreover, the cluster genes encoding short tail spike proteins of EspM4VN are different from *Limestonevirus* φXF4. Based on these results, the *E. kobei* phage EspM4VN was classified into the *Ackermannviridae* family, *Aglimvirinae* subfamily, and *Agtrevirus* genus.

Acknowledgement

First and foremost, I would like to express my sincerely gratitude to my supervisor Prof. Dr. Katsumi Doi, Laboratory of Microbial Genetic Resources, Department of Bioscience and Biotechnology, Faculty of Agriculture, Kyushu University, for his painstaking and indispensable advices, guidance, inspiration, and support throughout my studies. He teach me not only how to conduct research, to write scientific reports but also inspired me to research.

I would like to thank deeply to our laboratory's assistant research, Dr. Yasuhiro Fujino, Laboratory of Microbial Genetic Resources, Department of Bioscience and Biotechnology, Faculty of Agriculture, Kyushu University, for his valuable advices, care, support, and encouragements throughout my academic life.

I am grateful to Prof. Yasuaki Hiromasa, Attached Promotive Center for International Education and Research of Agriculture, Faculty of Agriculture, Kyushu University, Fukuoka, Japan; Assoc. Prof. Kazuhiro Iiyama and Prof. Naruto Furuya, Plant Pathology Laboratory, Faculty of Agriculture, Kyushu University; Prof. Takeo Iwamoto Core Research Facilities, Research Center for Medical Sciences, Jikei University School of Medicine; Prof. Takahisa Miyamoto, Food Hygienic Chemistry, Food Science and Biotechnology, Faculty of Agriculture, Kyushu University, Fukuoka, Japan; and Prof. Katsuya Fukami, Material Management Ceter, Kyushu University, Fukuoka, Japan for their helpful advices, reviewing my thesis and supports during my research.

To all my lab mates at Laboratory of Microbial Genetic Resources, Department of Bioscience and Biotechnology, Faculty of Agriculture, Kyushu University, my sincerely thanks for helping me and memorial student life.

I am thankful for the Japanese Grant Aid for Human Resource Development Scholarship (JDS) and Japanese Government Scholarship (MEXT) program for supporting financial aid.

Finally, specially thanks to my parents, my wife, and my daughters for unconditional love, support, and motivation.

Reference

Abbott, D. W., and Boraston, A. B. (2008). Structural biology of pectin degradation by *Enterobacteriaceae*. *Microbiol. Mol. Biol. Rev.* 72, 301–316. doi:10.1128/mmbr.00038-07.

Abedon, S. T. (2009). *Phages, ecology, evolution*. Cambridge University Press doi:10.1017/cbo9780511541483.004.

Abedon, S. T., and LeJeune, J. T. (2005). Why bacteriophage encode exotoxins and other virulence factors. *Evol. Bioinforma.* 1, 117693430500100. doi:10.1177/117693430500100001.

Ackermann, H.-W. (2009). Phage classification and characterization. *Methods Mol. Biol.* 501, 287–292. doi:10.1007/978-1-60327-164-6.

Ackermann, H. W. (2003). Bacteriophage observations and evolution. *Res. Microbiol.* 154, 245–251. doi:10.1016/S0923-2508(03)00067-6.

Adeolu, M., Alnajar, S., Naushad, S., and Gupta, R. S. (2016). Genome-based phylogeny and taxonomy of the '*Enterobacteriales*': Proposal for enterobacterales ord. nov. divided into the families *Enterobacteriaceae*, *Erwiniaceae* fam. nov., *Pectobacteriaceae* fam. nov., *Yersiniaceae* fam. nov., *Hafniaceae* fam. nov., Morgane. *Int. J. Syst. Evol. Microbiol.* 66, 5575–5599. doi:10.1099/ijsem.0.001485.

Adriaenssens, E. M., Wittmann, J., Kuhn, J. H., Turner, D., Sullivan, M. B., Dutilh, B. E., *et al*. (2018). Taxonomy of prokaryotic viruses: 2017 update from the ICTV Bacterial and Archaeal Viruses Subcommittee. *Arch. Virol.* 163, 1125–1129. doi:10.1007/s00705-018-3723-z.

Ahern, S. J., Das, M., Bhowmick, T. S., Young, R., and Gonzalez, C. F. (2014). Characterization of novel virulent broad-host-range phages of *Xylella fastidiosa* and *Xanthomonas*. *J. Bacteriol.* 196, 459–471. doi:10.1128/JB.01080-13.

Arndt, D., Grant, J. R., Marcu, A., Sajed, T., Pon, A., Liang, Y., *et al*. (2016). PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res.* 44, W16–W21. doi:10.1093/nar/gkw387.

Auch, A. F., von Jan, M., Klenk, H. P., and Göker, M. (2010). Digital DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison. *Stand. Genomic Sci.* 2, 117–134. doi:10.4056/sigs.531120.

Bailly-Bechet, M., Vergassola, M., and Rocha, E. (2007). Causes for the intriguing presence of tRNAs in phages. *Genome Res.* 17, 1486–1495. doi:10.1101/gr.6649807.

Balogh, B. (2006a). Characterization and use of bacteriophages associated with citrus. *Techniques*, 112. Available at: http://plaza.ufl.edu/bbalogh/balogh_b_dissertation.pdf.

Balogh, B. (2006b). Characterization and Use of Bacteriophages Associated With Citrus. *Techniques*, 112. Available at: http://plaza.ufl.edu/bbalogh/balogh_b_dissertation.pdf [Accessed May 26, 2020].

Balogh, B., Canteros, B. I., Stall, R. E., and Jones, J. B. (2008). Control of citrus canker and citrus bacterial spot with bacteriophages. *Plant Dis.* 92, 1048–1052. doi:10.1094/PDIS-92-7-1048.

Balogh, B., Jones, J. B., Momol, M. T., Olson, S. M., Obradovic, A., King, P., *et al*. (2003). Improved efficacy of newly formulated bacteriophages for management of

bacterial spot on tomato. *Plant Dis.* 87, 949–954. doi:10.1094/PDIS.2003.87.8.949.
Balogh, B., Jones, J., Iriarte, F., and Momol, M. (2010). Phage therapy for plant disease control. *Curr. Pharm. Biotechnol.* 11, 48–57. doi:10.2174/138920110790725302.
Barnes, E. H. (1979). "Bacterial Soft Rot," in atlas and manual of plant pathology (Boston, MA: Springer US), 35–54. doi:10.1007/978-1-4684-3495-8_5.
Bartz, J. A., and Kelman, A. (1986). Reducing the potential for bacterial soft rot in potato tubers by chemical treatments and drying. *Am. Potato J.* 63, 481–493. doi:10.1007/BF02852943.
Batinovic, S., Wassef, F., Knowler, S. A., Rice, D. T. F., Stanton, C. R., Rose, J., *et al.* (2019). Bacteriophages in natural and artificial environments. *Pathogens* 8, 100. doi:10.3390/pathogens8030100.
Bebeacua, C., Lorenzo Fajardo, J. C., Blangy, S., Spinelli, S., Bollmann, S., Neve, H., *et al*. (2013). X-ray structure of a superinfection exclusion lipoprotein from phage TP-J34 and identification of the tape measure protein as its target. *Mol. Microbiol.* 89, 152–165. doi:10.1111/mmi.12267.
Bhat, K. A., Masood, S. D., Bhat, N. A., Bhat, M. A., Razvi, S. M., Mir, M. R., *et al.* (2010). Current status of post harvest soft rot in vegetables: A review. *Asian J. Plant Sci.* 9, 200–208. doi:10.3923/ajps.2010.200.208.
Bhunchoth, A., Blanc-Mathieu, R., Mihara, T., Nishimura, Y., Askora, A., Phironrit, N., *et al*. (2016). Two asian jumbo phages, φRSL2 and φRSF1, infect *Ralstonia solanacearum* and show common features of φKZ-related phages. *Virology* 494, 56–66. doi:10.1016/j.virol.2016.03.028.
Bishop, A. L. (1990). Internal decay of onions caused by *Enterobacter cloacae*. *Plant Dis.* 74, 692. doi:10.1094/pd-74-0692.
Boetzer, M., Henkel, C. V., Jansen, H. J., Butler, D., and Pirovano, W. (2011). Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics* 27, 578–579. doi:10.1093/bioinformatics/btq683.
Boetzer, M., and Pirovano, W. (2012). Toward almost closed genomes with GapFiller. *Genome Biol.* 13, R56. doi:10.1186/gb-2012-13-6-r56.
Bolger-Munro, M., Cheung, K., Fang, A., and Wang, L. (2013). T4 bacteriophage average burst size varies with *Escherichia coli* B23 cell culture age. *J. Exp. Microbiol. Immunol.* 17, 115–119.
Bonde, R., and de Souza, P. (1954). Studies on the control of potato bacterial seed-piece decay and blackleg with antibiotics. *Am. Potato J.* 31, 311–316. doi:10.1007/BF02861599.
Boyd, E. F., and Brüssow, H. (2002). Common themes among bacteriophage-encoded virulence factors and diversity among the bacteriophages involved. *Trends Microbiol.* 10, 521–529. doi:10.1016/S0966-842X(02)02459-9.
Boyd, R. J., Hildebrand, A. C., and Allen, O. N. (1971). Retardation of crown gall enlargement after bacteriophage treatment. *Plant Dis. Report.* 55, 145–148. Available at: http://agris.fao.org/agris-search/search.do?recordID=US201301153915 [Accessed May 25, 2020].
Brady, C., Cleenwerck, I., Venter, S., Vancanneyt, M., Swings, J., and Coutinho, T. (2008). Phylogeny and identification of *Pantoea* species associated with plants, humans and the natural environment based on multilocus sequence analysis (MLSA). *Syst. Appl. Microbiol.* 31, 447–460. doi:10.1016/J.SYAPM.2008.09.004.

Brady, C. L., Cleenwerck, I., Venter, S. N., Engelbeen, K., De Vos, P., and Coutinho, T. a. (2010). Emended description of the genus *Pantoea*, description of four species from human clinical samples, *Pantoea septica* sp. nov., *Pantoea eucrina* sp. nov., *Pantoea brenneri* sp. nov. and *Pantoea conspicua* sp. nov., and transfer of *Pectobacterium cypripedii* (Hor. *Int. J. Syst. Evol. Microbiol.* 60, 2430–2440. doi:10.1099/ijs.0.017301-0.

Brenner, D. J., Krieg, N. R., Staley, J. T., Garrity, G. M., Boone, D. R., De Vos, P., *et al*. (2005). The *Proteobacteria*, The *Gammaproteobacteria - Enterobacteriales*. *Bergey's manual® Syst. Bacteriol.* 2, 587–850. doi:10.1007/0-387-28022-7.

Brenner, D. J., McWhorter, a. C., Kai, a., Steigerwalt, a. G., and Farmer, J. J. (1986). *Enterobacter asburiae* sp. nov., a new species found in clinical specimens, and reassignment of *Erwinia* dissolvens and *Erwinia nimipressuralis* to the genus *Enterobacter* as *Enterobacter dissolvens* comb. nov., and *Enterobacter nimipressuralis* comb. nov. *J. Clin. Microbiol.* 23, 1114–1120.

Brunner, M., and Pootjes, C. F. (1969). Bacteriophage release in a lysogenic strain of agrobacterium tumefaciens1. *J. Virol.* 3, 181–186. doi:10.1128/jvi.3.2.181-186.1969.

Byrne, J. M., Dianese, A. C., Ji, P., Campbell, H. L., Cuppels, D. A., Louws, F. J., *et al*. (2005). Biological control of bacterial spot of tomato under field conditions at several locations in North America. *Biol. Control* 32, 408–418. doi:10.1016/j.biocontrol.2004.12.001.

Campos, E. (1982). Relationship of pectolytic clostridia and *Erwinia carotovora* strains to decay of potato tubers in storage . *Plant Dis.* 66, 543. doi:10.1094/pd-66-543.

Casjens, S. (2003). Prophages and bacterial genomics: What have we learned so far? John Wiley & Sons, Ltd doi:10.1046/j.1365-2958.2003.03580.x.

Casjens, S. R., Gilcrease, E. B., Huang, W. M., Bunny, K. L., Pedulla, M. L., Ford, M. E., *et al*. (2004). The pKO2 linear plasmid prophage of *Klebsiella oxytoca*. *J. Bacteriol.* 186, 1818–1832. doi:10.1128/JB.186.6.1818-1832.2004.

Casjens, S. R., and Molineux, I. J. (2012). Short noncontractile tail machines: Adsorption and DNA delivery by podoviruses. *Adv. Exp. Med. Biol.* 726, 143–179. doi:10.1007/978-1-4614-0980-9_7.

Ceyssens, P.-J., and Lavigne, R. (2010). Bacteriophages of *Pseudomonas*. *Future Microbiol.* 5, 1041–1055. doi:10.2217/fmb.10.66.

Chanishvili, N. (2012). Phage therapy-history from Twort and d'Herelle through Soviet experience to current approaches". *Advances in Virus Research* (Academic Press), 3–40. doi:10.1016/B978-0-12-394438-2.00001-3.

Charkowski, A., Blanco, C., Condemine, G., Expert, D., Franza, T., Hayes, C., *et al*. (2012). The role of secretion systems and small molecules in soft rot *Enterobacteriaceae* pathogenicity . *Annu. Rev. Phytopathol.* 50, 425–449. doi:10.1146/annurev-phyto-081211-173013.

Cheetham, B. F., and Katz, M. E. (1995). A role for bacteriophages in the evolution and transfer of bacterial virulence determinants. John Wiley & Sons, Ltd doi:10.1111/j.1365-2958.1995.mmi_18020201.x.

Chen, X. F., Zhang, H. L., and Chen, J. (2015). First report of *Dickeya solani* causing soft rot in imported bulbs of *Hyacinthus orientalis* in China . *Plant Dis.* 99, 155–155. doi:10.1094/pdis-09-14-0916-pdn.

Choffnes, E. R., Relman, D. a, and Mack, A. (2010). Antibiotic resistance: implications for global health and novel intervention strategies. *Workshop*.

Clokie, M. R. J., and Kropinski, A. M. (2009). Bacteriophages : methods and protocols. doi:10.1007/978-1-60327-164-6.

Coetzee, J. N. (1966). Transduction in *Proteus morganii* [37]. *Nature* 210, 220. doi:10.1038/210220a0.

Colyer, P. D. (1984). Bacterization of potatoes with *Pseudomonas putida* and its influence on postharvest soft rots. *Plant Dis.* 68, 703. doi:10.1094/pd-69-703.

Coons, G. H., and Kotila, J. E. (1925). The transmissible lytic principle (bacteriophage) in relation to plant pathogens. *Phytopathology* 15, 357–70.

Crowley, P. H., Straley, S. C., Craig, R. J., Culin, J. D., Fu, Y. T., Hayden, T. L., *et al*. (1980). A model of prey bacteria, predator bacteria, and bacteriophage in continuous culture. *J. Theor. Biol.* 86, 377–400. doi:10.1016/0022-5193(80)90013-2.

Crowther, R. A., Lenk, E. V., Kikuchi, Y., and King, J. (1977). Molecular reorganization in the hexagon to star transition of the baseplate of bacteriophage T4. *J. Mol. Biol.* 116, 489–523. doi:10.1016/0022-2836(77)90081-X.

Czajkowski, R., Ozymko, Z., De Jager, V., Siwinska, J., Smolarska, A., Ossowicki, A., *et al*. (2015). Genomic, proteomic and morphological characterization of two novel broad host lytic bacteriophages PdblPD10.3 and PdblPD23.1 infecting pectinolytic *Pectobacterium* spp. and *Dickeya* spp. *PLoS One* 10, e0119812. doi:10.1371/journal.pone.0119812.

Czajkowski, R., Pérombelon, M. C. M., van Veen, J. A., and van der Wolf, J. M. (2011). Control of blackleg and tuber soft rot of potato caused by *Pectobacterium* and *Dickeya* species: a review. *Plant Pathol.* 60, 999–1013. doi:10.1111/j.1365-3059.2011.02470.x.

Day, A., Ahn, J., and Salmond, G. P. C. (2018). Jumbo bacteriophages are represented within an increasing diversity of environmental viruses infecting the emerging phytopathogen, Dickeya solani. *Front. Microbiol.* 9, 2169. doi:10.3389/fmicb.2018.02169.

Deghorain, M., and Van Melderen, L. (2012). The *Staphylococci* phages family: An overview. *Viruses* 4, 3316–3335. doi:10.3390/v4123316.

Dong, L., Lv, L. B., and Lai, R. (2012). Molecular cloning of *Tupaia belangeri chinensis* neuropeptide Y and homology comparison with other analogues from primates. 2nd ed. Cold Spring Harbor Laboratory Press doi:10.3724/sp.j.1141.2012.01075.

Drulis-Kawa, Z., Majkowska-Skrobek, G., Maciejewska, B., Delattre, A.-S., and Lavigne, R. (2013). Learning from bacteriophages - advantages and limitations of phage and phage-encoded potein applications. *Curr. Protein Pept. Sci.* 13, 699–722. doi:10.2174/138920312804871193.

E.L., C., and H.L., K. (1969). Inhibition of bacterial spot of peach foliage by *Xanthomonas pruni* bacteriophage. *Phytopathology* 59, 1966–1967. Available at: https://agris.fao.org/agris-search/search.do?recordID=US201301227121 [Accessed May 27, 2020].

Ee, R., Madhaiyan, M., Ji, L., Lim, Y. L., Nor, N. M., Tee, K. K., *et al*. (2016). Chania multitudinisentens gen. nov., sp. nov., an N-acyl-homoserine-lactone-producing bacterium in the family *Enterobacteriaceae* isolated from landfill site soil. *Int. J.*

Syst. Evol. Microbiol. 66, 2297–2304. doi:10.1099/ijsem.0.001025.
Elbanna, K., Elnaggar, S., and Bakeer, A. (2014). Characterization of *Bacillus altitudinis* as a New Causative Agent of Bacterial Soft Rot. *J. Phytopathol.* 162, 712–722. doi:10.1111/jph.12250.
Engineering and Consulting Firms Association, J. (2006). Study on strengthening food standards and the certification system in the Socialist Republic of Vietnam study report February 2006 Engineering and Consulting Firms Association , Japan Overseas Merchandise Inspection Co ., Ltd . Available at: http://www.ecfa.or.jp/japanese/act-pf_jka/H17/renkei/renkei_Vietnam_Food_Eng.pdf.
Evans, T. J., Coulthurst, S. J., Komitopoulou, E., and Salmond, G. P. C. (2010). Two mobile *Pectobacterium atrosepticum* prophages modulate virulence. *FEMS Microbiol. Lett.* 304, 195–202. doi:10.1111/j.1574-6968.2010.01901.x.
Fan, H. C., Zeng, L., Yang, P. W., Guo, Z. X., and Bai, T. T. (2016). First report of banana soft rot caused by *Klebsiella variicola* in China. *Plant Dis.* 100, 517–517. doi:10.1094/PDIS-05-15-0586-PDN.
Farmer, J. J., Asbury, M. A., Hickman, F. W., and Brenner, D. J. (1980). *Enterobacter sakazakii*: A new species of "*Enterobacteriaceae*" isolated from clinical specimens. *Int. J. Syst. Bacteriol.* 30, 569–584. doi:10.1099/00207713-30-3-569.
Farmer, J. J., Davis, B. R., Hickman-Brenner, F. W., McWhorter, A., Huntley-Carter, G. P., Asbury, M. A., *et al*. (1985). Biochemical identification of new species and biogroups of *Enterobacteriaceae* isolated from clinical specimens. *J. Clin. Microbiol.* 21, 46–76. doi:10.1128/jcm.21.1.46-76.1985.
Frampton, R. A., Taylor, C., Holguín Moreno, A. V., Visnovsky, S. B., Petty, N. K., Pitman, A. R., *et al*. (2014). Identification of bacteriophages for biocontrol of the kiwifruit canker phytopathogen *Pseudomonas syringae* pv. *actinidiae*. *Appl. Environ. Microbiol.* 80, 2216–2228. doi:10.1128/AEM.00062-14.
Fujiwara, A., Fujisawa, M., Hamasaki, R., Kawasaki, T., Fujie, M., and Yamada, T. (2011). Biocontrol of *Ralstonia solanacearum* by treatment with lytic bacteriophages. *Appl. Environ. Microbiol.* 77, 4155–4162. doi:10.1128/AEM.02847-10.
Fukushima, M., Kakinuma, K., and Kawaguchi, R. (2002). Phylogenetic analysis of *Salmonella*, *Shigella*, and *Escherichia coli* strains on the basis of the *gyrB* gene sequence. *J. Clin. Microbiol.* 40, 2779–2785. doi:10.1128/JCM.40.8.2779-2785.2002.
Furtado, G. Q., Guimarães, L. M. S., Lisboa, D. O., Cavalcante, G. P., Arriel, D. A. A., Alfenas, A. C., *et al*. (2012). First report of *Enterobacter* cowanii causing bacterial spot on *Mabea fistulifera* , a native frest species in Brazil. *Plant Dis.* 96, 1576–1576. doi:10.1094/pdis-02-12-0160-pdn.
Gainvors, A., Frézier, V., Lemaresquier, H., Lequart, C., Aigle, M., and Belarbi, A. (1994). Detection of polygalacturonase, pectin-lyase and pectin-esterase activities in a *Saccharomyces cerevisiae* strain. *Yeast* 10, 1311–1319. doi:10.1002/yea.320101008.
García-González, T., Sáenz-Hidalgo, H. K., Silva-Rojas, H. V., Morales-Nieto, C., Vancheva, T., Koebnik, R., *et al*. (2018). *Enterobacter cloacae*, an emerging plant-pathogenic bacterium affecting chili pepper seedlings. *Plant Pathol. J.* 34, 1–10. doi:10.5423/PPJ.OA.06.2017.0128.

Gavrilovic, *et al* (2001). Plant pathogenic bacteria. doi:10.1007/978-94-010-0003-1.
Giamarellou, H. (2010). Multidrug-resistant gram-negative bacteria: how to treat and for how long. *Int. J. Antimicrob. Agents* 36, S50–S54. doi:10.1016/J.IJANTIMICAG.2010.11.014.
Gill, J., and Abedon, S. T. (2003). Bacteriophage ecology and plants. *APSnet Featur. Artic.* doi:10.1094/APSnetFeature-2003-1103.
Glaeser, S. P., and Kämpfer, P. (2015). Multilocus sequence analysis (MLSA) in prokaryotic taxonomy. *Syst. Appl. Microbiol.* 38, 237–245. doi:10.1016/j.syapm.2015.03.007.
Godfrey, S. A. C., and Marshall, J. W. (2002). Identification of cold-tolerant *Pseudomonas viridiflava* and *P. marginalis* causing severe carrot postharvest bacterial soft rot during refrigerated export from New Zealand. *Plant Pathol.* 51, 155–162. doi:10.1046/j.1365-3059.2002.00679.x.
Goodridge, L. D. (2004). Bacteriophage biocontrol of plant pathogens: Fact or fiction? *Trends Biotechnol.* 22, 384–385. doi:10.1016/j.tibtech.2004.05.007.
Greenfield, J., Shang, X., Luo, H., Zhou, Y., Heselpoth, R. D., Nelson, D. C., *et al.* (2019). Structure and tailspike glycosidase machinery of ORF212 from *E. coli* O157:H7 phage CBA120 (TSP3). *Sci. Rep.* 9, 1–11. doi:10.1038/s41598-019-43748-9.
Grimont, P. A. D., Farmer, J. J., Grimont, F., Asbury, M. A., Brenner, D. J., and Deval, C. (1983). *Ewingella americana* gen. nov., sp.Nov., a new *Enterobacteriaceae* isolated from clinical specimens. doi:10.1016/0769-2609(83)90102-3.
Harris, R. I. (1979). Chemical control of bacterial soft-rot of wounded potato tubers. *Potato Res.* 22, 245–249. doi:10.1007/BF02357357.
Hata, H., Natori, T., Mizuno, T., Kanazawa, I., Eldesouky, I., Hayashi, M., *et al.* (2016). Phylogenetics of family *Enterobacteriaceae* and proposal to reclassify *Escherichia hermannii* and *Salmonella subterranea* as *Atlantibacter hermannii* and *Atlantibacter subterranea* gen. nov., comb. nov. *Microbiol. Immunol.* 60, 303–311. doi:10.1111/1348-0421.12374.
He, P. F. (2012). Bipolaris cactivora causing fruit rot of dragon fruit imported from Vietnam. *Plant Pathol. Quar.* 2, 31–35. doi:10.5943/ppq/2/1/5.
Hemstreet, C. (1924). Isolation of an inhibitory substance From. *J. Agric. Res. "Washington, D. C* XXVIII, 599–602. Available at: https://naldc.nal.usda.gov/download/IND43966880/PDF [Accessed May 26, 2020].
Hofer, B., Ruge, M., and Dreiseikelmann, B. (1995). The superinfection exclusion gene (sieA) of bacteriophage P22: Identification and overexpression of the gene and localization of the gene product. *J. Bacteriol.* 177, 3080–3086. doi:10.1128/jb.177.11.3080-3086.1995.
Hong Nhung, P., Ohkusu, K., Mishima, N., Noda, M., Monir Shah, M., Sun, X., *et al.* (2007). Phylogeny and species identification of the family *Enterobacteriaceae* based on dnaJ sequences. *Diagn. Microbiol. Infect. Dis.* 58, 153–161. doi:10.1016/j.diagmicrobio.2006.12.019.
Howarth, F. G. (1991). Environmental impacts of classical biological control. *Annu. Rev. Entomol. Vol. 36* 36, 485–509. doi:10.1146/annurev.ento.36.1.485.
http://www.aphage.com/case-studies/ Available at: http://www.aphage.com/case-studies/ [Accessed June 25, 2020].
Hugouvieux C.P., N., Condemine, G., and Shevchik, V. E. (2014). Bacterial pectate

lyases, structural and functional diversity. *Environ. Microbiol. Rep.* 6, 427–440. doi:10.1111/1758-2229.12166.

Hunt, M., Newbold, C., Berriman, M., and Otto, T. D. (2014). A comprehensive evaluation of assembly scaffolding tools. *Genome Biol.* 15, R42. doi:10.1186/gb-2014-15-3-r42.

Iglesias-Jiménez, E., Sánchez-Martín, M. J., and Sánchez-Camazano, M. (1996). Pesticide adsorption in a soil-water system in the presence of surfactants. *Chemosphere* 32, 1771–1782. doi:10.1016/0045-6535(96)00094-X.

Iriarte, F. B., Balogh, B., Momol, M. T., Smith, L. M., Wilson, M., and Jones, J. B. (2007). Factors affecting survival of bacteriophage on tomato leaf surfaces. *Appl. Environ. Microbiol.* 73, 1704–1711. doi:10.1128/AEM.02118-06.

Iversen, C. (2014). *Enterobacter* view all topics electrical techniques. *Encyclopedia of Food Microbiolog.* 653-658. https://doi.org/10.1016/B978-0-12-384730-0.00095-1

Janda, J. M., and Abbott, S. L. (2007). Miniview 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: Pluses, Perils, and Pitfalls. 45, 2761–2764. doi:10.1128/JCM.01228-07.

Jones, D. (1981). Terminology, etiology,. *J. Clin. Pathol.* 34, 185–189. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC494262/.

Joseph, A. M.-C. & S. W. (2005). Bergey's Manual® of Systematic Bacteriology. *Syst. Bacteriol.* Volume Two, 556–578. doi:10.1007/0-387-28022-7.

Keen, E. C., Adhya, S. L., and Wormser, G. P. (2015). Phage therapy: Current research and applications. *Clin. Infect. Dis.* 61, 141–142. doi:10.1093/cid/civ257.

Kloepper, J. W. (1983). Effect of seed piece inoculation with plant growth promoting *Rhizobacteria* on populations of *Erwinia carotovora* on potato roots and in daughter tubers . *Phytopathology* 73, 217. doi:10.1094/phyto-73-217.

Kloepper, J. W., Leong, J., Teintze, M., and Schroth, M. N. (1980). Pseudomonas siderophores: A mechanism explaining disease-suppressive soils. *Curr. Microbiol.* 4, 317–320. doi:10.1007/BF02602840.

Klumpp, J., Dorscht, J., Lurz, R., Bielmann, R., Wieland, M., Zimmer, M., *et al*. (2008). The terminally redundant, nonpermuted genome of *Listeria* bacteriophage A511: A model for the SPO1-like myoviruses of gram-positive bacteria. *J. Bacteriol.* 190, 5753–5765. doi:10.1128/JB.00461-08.

Kosako, Y., Tamura, K., Sakazaki, R., and Miki, K. (1996). *Enterobacter kobei* sp. nov., a new species of the family *Enterobacteriaceae* resembling *Enterobacter cloacae*. *Curr. Microbiol.* 33, 261–265. doi:10.1007/s002849900110.

Kotila, J. (1920). Investigations of the blackleg disease of the potato. *Agri. Exp. Station technical bulletin,* Volume 67.

Kou, T.-T., Chang, L.-C., Yang, C.-M., and Yang, S.-E. (1971). Bacterial leaf blight of rice plant. IV. Effect of bacteriophages on the infectivity of *Xanthomonas oryzae*. *Bot. Bull. Acad. Sin.* 12, 1–8. Available at: https://agris.fao.org/agris-search/search.do?recordID=US201302342275 [Accessed May 27, 2020].

Kruger, D. H., Schneck, P., and Gelderblom, H. R. (2000). Helmut ruska and the visualisation of viruses. *Lancet* 355, 1713–1717. doi:10.1016/S0140-6736(00)02250-9.

Lampert, Y., Dror, B., Sela, N., Teper-Bamnolker, P., Daus, A., Sela (Saldinger), S., *et al*. (2017). Emergence of *Leuconostoc mesenteroides* as a causative agent of oozing in carrots stored under non-ventilated conditions. *Microb. Biotechnol.* 10,

1677–1689. doi:10.1111/1751-7915.12753.
Lang, J. M., Gent, D. H., and Schwartz, H. F. (2007). Management of *Xanthomonas* leaf blight of onion with bacteriophages and a plant activator. *Plant Dis.* 91, 871–878. doi:10.1094/PDIS-91-7-0871.
Lee, D. H., Lim, J. A., Lee, J., Roh, E., Jung, K., Choi, M., *et al*. (2013). Characterization of genes required for the pathogenicity of *Pectobacterium carotovorum* subsp. *carotovorum* Pcc21 in Chinese cabbage. *Microbiol. (United Kingdom)* 159, 1487–1496. doi:10.1099/mic.0.067280-0.
Lee, J.-H., Choi, Y., Shin, H., Lee, J., and Ryu, S. (2012). Complete genome sequence of *Cronobacter sakazakii* temperate bacteriophage phiES15. *J. Virol.* 86, 7713–7714. doi:10.1128/jvi.01042-12.
Lefort, V., Desper, R., and Gascuel, O. (2015). FastME 2.0: A comprehensive, accurate, and fast distance-based phylogeny inference program: Table 1. *Mol. Biol. Evol.* 32, 2798–2800. doi:10.1093/molbev/msv150.
Leiman, P. G., Arisaka, F., Van Raaij, M. J., Kostyuchenko, V. A., Aksyuk, A. A., Kanamaru, S., *et al*. (2010). Morphogenesis of the T4 tail and tail fibers. *Virol. J.* 7, 355. doi:10.1186/1743-422X-7-355.
Leiman, P. G., Chipman, P. R., Kostyuchenko, V. A., Mesyanzhinov, V. V., and Rossmann, M. G. (2004). Three-dimensional rearrangement of proteins in the tail of bacteriophage T4 on infection of its host. *Cell* 118, 419–429. doi:10.1016/j.cell.2004.07.022.
Letarov, A. V., and Kulikov, E. E. (2017). Adsorption of bacteriophages on bacterial cells. *Biochem.* 82, 1632–1658. doi:10.1134/S0006297917130053.
Li, E., Wei, X., Ma, Y., Yin, Z., Li, H., Lin, W., *et al*. (2016). Isolation and characterization of a bacteriophage phiEap-2 infecting multidrug resistant *Enterobacter aerogenes*. *Sci. Rep.* 6, 28338. doi:10.1038/srep28338.
Liao, L., Hei, R., Tang, Y., Liu, S., and Zhou, J. (2016). First report of soft rot of peach caused by *Pantoea ananatis* in China. *Plant Dis.* 100, 516–516. doi:10.1094/PDIS-06-15-0620-PDN.
Lim, J. A., Jee, S., Lee, D. H., Roh, E., Jung, K., Oh, C., *et al*. (2013). Biocontrol of *Pectobacterium carotovorum* subsp. *carotovorum* using bacteriophage PP1. *J. Microbiol. Biotechnol.* 23, 1147–1153. doi:10.4014/jmb.1304.04001.
Lin, L., Han, J., Ji, X., Hong, W., Huang, L., and Wei, Y. (2011). Isolation and characterization of a new bacteriophage MMP17 from *Meiothermus*. *Extremophiles* 15, 253–258. doi:10.1007/s00792-010-0354-z.
Loessner, M. J., Neugirg, E., Zink, R., and Scherer, S. (1993). Isolation, classification and molecular characterization of bacteriophages for *Enterobacter* species. *J. Gen. Microbiol.* 139, 2627–2633. doi:10.1099/00221287-139-11-2627.
Lu, S.-E., Henn, R. A., and Nagel, D. H. (2007). First report of ear soft rot of corn (*Zea mays*) caused by *Burkholderia gladioli* in the United States. *Plant Dis.* 91, 1514–1514. doi:10.1094/PDIS-91-11-1514C.
Luria, S. E., and Anderson, T. F. (1942). The identification and characterization of bacteriophages with the electron microscope. *Proc. Natl. Acad. Sci. U. S. A.* 28, 127. doi:10.1073/PNAS.28.4.127.
Luria, S. E., Delbrück, M., and Anderson, T. F. (1943). Electron microscope studies of bacterial viruses1. *J. Bacteriol.* 46, 57–77. doi:10.1128/jb.46.1.57-77.1943.
Ma, B., Hibbing, M. E., Kim, H.-S., Reedy, R. M., Yedidia, I., Breuer, J. J. J. J., *et al*.

(2007). Host range and molecular phylogenies of the soft rot enterobacterial genera *Pectobacterium* and *Dickeya*. *Phytopathology* 97, 1150–1163. doi:10.1094/PHYTO-97-9-1150.

Mahony, J., McGrath, S., Fitzgerald, G. F., and Van Sinderen, D. (2008). Identification and characterization of *lactococcal*-prophage-carried superinfection exclusion genes. *Appl. Environ. Microbiol.* 74, 6206–6215. doi:10.1128/AEM.01053-08.

Mansfield, J., Genin, S., Magori, S., Citovsky, V., Sriariyanum, M., Ronald, P., *et al*. (2012). Top 10 plant pathogenic bacteria in molecular plant pathology. *Mol. Plant Pathol.* 13, 614–629. doi:10.1111/j.1364-3703.2012.00804.x.

Masyahit, M., Sijam, K., Awang, Y., and Satar, M. G. M. (2009). First report on bacterial soft rot on dragon fruit (Hylocereus spp.) caused by *Enterobacter cloacae* in Peninsular Malaysia. Int J Agric Biol 11:659–. *Int. J. Agric. Biol.* 11, 659–666.

McGhee, G. C., and Sundin, G. W. (2011). Evaluation of kasugamycin for fire blight management, effect on nontarget bacteria, and assessment of kasugamycin resistance potential in *Erwinia amylovora*. *Phytopathology* 101, 192–204. doi:10.1094/PHYTO-04-10-0128.

Meier-Kolthoff, J. P., Auch, A. F., Klenk, H. P., and Göker, M. (2013). Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics* 14, 60. doi:10.1186/1471-2105-14-60.

Meier-Kolthoff, J. P., and Göker, M. (2019). TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nat. Commun.* 10, 2182. doi:10.1038/s41467-019-10210-3.

Mills A., Platt H. W., Hutta R. (2006) Sensitivity of *Erwinia* spp. to salt compounds in vitro and their effect on the development of soft rot in potato tuber in storage. *Postharvest Biology and Technology,* 41,2, p208-214. doi: 10.1016/j.postharvbio.2006.03.015

Mishra, A. K., Lagier, J.-C., Robert, C., Raoult, D., and Fournier, P.-E. (2012). Non contiguous-finished genome sequence and description of *Peptoniphilus timonensis* sp. nov. *Stand. Genomic Sci.* 7, 1–11. doi:10.4056/sigs.2956294.

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. doi:10.1186/1471-2164-14-481.

Mosbahi, K., Wojnowska, M., Albalat, A., and Walker, D. (2018). Bacterial iron acquisition mediated by outer membrane translocation and cleavage of a host protein. *Proc. Natl. Acad. Sci. U. S. A.* 115, 6840–6845. doi:10.1073/pnas.1800672115.

Müller, I., Lurz, R., and Geider, K. (2012). Tasmancin and lysogenic bacteriophages induced from *Erwinia tasmaniensis* strains. *Microbiol. Res.* 167, 381–387. doi:10.1016/J.MICRES.2012.01.005.

Nagayoshi, Y., Kumagae, K., Mori, K., Tashiro, K., Nakamura, A., Fujino, Y., *et al.* (2016). Physiological properties and genome structure of the hyperthermophilic filamentous phage φOH3 which infects *Thermus thermophilus* HB8. *Front. Microbiol.* 7, 1–11. doi:10.3389/fmicb.2016.00050.

Narváez-Barragán, D. A., Tovar-Herrera, O. E., Torres, M., Rodríguez, M., Humphris, S., Toth, I. K., *et al*. (2020). Expansin-like Exl1 from *Pectobacterium* is a virulence factor required for host infection, and induces a defence plant response

involving ROS, and jasmonate, ethylene and salicylic acid signalling pathways in Arabidopsis thaliana. *Sci. Rep.* 10, 7747. doi:10.1038/s41598-020-64529-9.
Nathan, A. J., and Scobell, A. (2012). How China sees America. *Foreign Aff.* 91, 1689–1699. doi:10.1017/CBO9781107415324.004.
Naum, M., Brown, E. W., and Mason-Gamer, R. J. (2008). Is 16S rDNA a reliable phylogenetic marker to characterize relationships below the family level in the *Enterobacteriaceae*? *J. Mol. Evol.* 66, 630–642. doi:10.1007/s00239-008-9115-3.
Naum, M., Brown, E. W., and Mason-Gamer, R. J. (2011). Is a robust phylogeny of the enterobacterial plant pathogens attainable? *Cladistics* 27, 80–93. doi:10.1111/j.1096-0031.2010.00313.x.
Nishijima, K. A. (1987). Internal yellowing, a bacterial disease of papaya fruits caused by *Enterobacter cloacae*. *Plant Dis.* 71, 1029. doi:10.1094/pd-71-1029.
Okabe, N., and Goto, M. (1963). Bacteriophages of plant pathogens. *Annu. Rev. Phytopathol.* 1, 397–418. doi:10.1146/annurev.py.01.090163.002145.
Ondov, B. D., Treangen, T. J., Melsted, P., Mallonee, A. B., Bergman, N. H., Koren, S., *et al*. (2016). Mash: Fast genome and metagenome distance estimation using MinHash. *Genome Biol.* 17, 132. doi:10.1186/s13059-016-0997-x.
Paradis, S., Boissinot, M., Paquette, N., Bélanger, S. D., Martel, E. A., Boudreau, D. K., *et al*. (2005). Phylogeny of the *Enterobacteriaceae* based on genes encoding elongation factor Tu and F-ATPase β-subunit. *Int. J. Syst. Evol. Microbiol.* 55, 2013–2025. doi:10.1099/ijs.0.63539-0.
Parkinson, N., Pritchard, L., Bryant, R., Toth, I., and Elphinstone, J. (2014). Epidemiology of *Dickeya dianthicola* and *Dickeya solani* in ornamental hosts and potato studied using variable number tandem repeat analysis. *Eur. J. Plant Pathol.* 141, 63–70. doi:10.1007/s10658-014-0523-5.
Paucker, K. (1977). Antigenic properties. *Tex. Rep. Biol. Med.* Vol. 35, 23–28.
Pelzek, A. J., Schuch, R., Schmitz, J. E., and Fischetti, V. A. (2013). Isolation, culture, and characterization of bacteriophages. *Curr. Protoc. Essent. Lab. Tech.* 2013, 4.4.1-4.4.33. doi:10.1002/9780470089941.et0404s07.
Pepper, I. L., Gerba, C. F., Gentry, T., and Maier, R. M. (2009). Environmental microbiology. doi:10.1016/B978-0-12-370519-8.X0001-6.
Plastow, G. S. (1988a). Molecular cloning and nucleotide sequence of the pectin methyl esterase gene of *Erwinia chrysanthemi* B374. *Mol. Microbiol.* 2, 247–254. doi:10.1111/j.1365-2958.1988.tb00026.x.
Plastow, G. S. (1988b). Molecular cloning and nucleotide sequence of the pectin methyl esterase gene of *Erwinia chrysanthemi* B374. *Mol. Microbiol.* 2, 247–254. doi:10.1111/j.1365-2958.1988.tb00026.x.
Plattner, M., Shneider, M. M., Arbatsky, N. P., Shashkov, A. S., Chizhov, A. O., Nazarov, S., *et al*. (2019). Structure and function of the branched receptor-binding complex of bacteriophage CBA120. *J. Mol. Biol.* 431, 3718–3739. doi:10.1016/j.jmb.2019.07.022.
Rahmanl M., M Ad Khan, M Al Khan (2010). Loss assessment of potato due to bacterial soft rot disease in Bangladesh. *Bangladesh J. Agriculturist*. ISSN 1812-4631, 93-100.
Ranganna, B., Kushalappa, A. C., and Raghavan, G. S. V. (1997). Ultraviolet irradiance to control dry rot and soft rot of potato in storage. *Can. J. Plant Pathol.* 19, 30–35. doi:10.1080/07060669709500568.

Ray, D. K., Mueller, N. D., West, P. C., and Foley, J. A. (2013). Yield trends are insufficient to double global crop production by 2050. *PLoS One* 8, e66428. doi:10.1371/journal.pone.0066428.

Razanakoto, L. M., Massart, S., De Clerck, C., Rabemanantsoa, C., Raherimandimby, M., El Jaziri, M., *et al*. (2015). First report on the occurrence of *Enterobacter* sp. causing leaf dieback and wilt of potato in Madagascar. *New Dis. Reports* 32, 34. doi:10.5197/j.2044-0588.2015.032.034.

RC, Thomas (1935). A bacteriophage in relation to Stewart's disease of corn. *Phytopathology*, Springer publisher, 25, 371.

Rendulic, S., Jagtap, P., Rosinus, A., Eppinger, M., Baar, C., Lanz, C., *et al*. (2004). A Predator Unmasked: Life Cycle of *Bdellovibrio bacteriovorus* from a Genomic Perspective. *Science (80-.).* 303, 689–692. doi:10.1126/science.1093027.

Reyes-García, G., Ortega-Acosta, S. Á., Palemón-Alberto, F., Ramírez, Y. R., Toribio-Jiménez, J., Vargas-Álvarez, D., *et al*. (2020). First report of bacterial soft rot on *Mammillaria mystax* caused by *Enterobacter cloacae* subsp. Dissolvens in Mexico. *Plant Dis.* 104, 1536–1536. doi:10.1094/PDIS-08-19-1662-PDN.

Ristuccia, P. A., and Cunha, B. A. (1985). Enterobacter. *Infect. Control* 6, 124–128. doi:10.1017/S0195941700062810.

Roggenkamp, A. (2007). Phylogenetic analysis of enteric species of the family *Enterobacteriaceae* using the oriC-locus. *Syst. Appl. Microbiol.* 30, 180–188. doi:10.1016/j.syapm.2006.06.004.

Rohan Van Twest and Andrew M. Kropinski (2009). Bacteriophage enrichment from water and soil. *Methods Mol. Biol.* 501, 287–292. doi:10.1007/978-1-60327-164-6.

Salmond, G. P. C. (1992). Bacterial diseases of potatoes: from classical phytobacteriology to molecular pathogenicity. *Netherlands J. Plant Pathol.* 98, 115–126. doi:10.1007/BF01974478.

Scholl, D., Adhya, S., and Merril, C. (2005). *Escherichia coli* K1's capsule is a barrier to bacteriophage T7. *Appl. Environ. Microbiol.* 71, 4872–4874. doi:10.1128/AEM.71.8.4872-4874.2005.

Schroeder, B. K., Du Toit, L. J., and Schwartz, H. F. (2009). First report of *Enterobacter cloacae* causing onion bulb rot in the columbia basin of washington state. *Plant Dis.* 93, 323. doi:10.1094/PDIS-93-3-0323A.

Seed, K. D. (2015). Battling phages: How bacteria defend against viral attack. *PLoS Pathog.* 11. doi:10.1371/journal.ppat.1004847.

Seok Hwan, Y., SM, H., J, L., S, K., and J, C. (2017). A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie Van Leeuwenhoek* 110. doi:10.1007/S10482-017-0844-4.

Sharma, R., Pielstick, B. A., Bell, K. A., Nieman, T. B., Stubbs, O. A., Yeates, E. L., *et al*. (2019). A novel, highly related jumbo family of bacteriophages that were isolated against *Erwinia*. *Front. Microbiol.* 10, 1533. doi:10.3389/fmicb.2019.01533.

Smith, H. W., and Huggins, M. B. (1982). Successful treatment of experimental *Escherichia coli* infections in mice using phage: its general superiority over antibiotics. *J. Gen. Microbiol.* 128, 307–318. doi:10.1099/00221287-128-2-307.

Soto, J., Cadenas, C., Mattos, L., and Trigoso, C. (2019). First report of *Enterobacter cloacae* as a causative agent of soft rot in dragon fruit (*Hylocereus undatus*) stems in Peru. *Peruvian J. Agron.* 3, 144. doi:10.21704/pja.v3i3.1367.

Stonier, T., McSharry, J., and Speitel, T. (1967). *Agrobacterium tumefaciens* conn IV. bacteriophage PB21 and its inhibitory effect on tumor induction. *J. Virol.* 1, 268–273. doi:10.1128/jvi.1.2.268-273.1967.

Storey, M. V., and Ashbolt, N. J. (2001). Persistence of two model enteric viruses (B40-8 and MS-2 bacteriophages) in water distribution pipe biofilms. *Water Sci. Technol.* 43, 133–138. doi:10.2166/wst.2001.0724.

Strange, R. N., and Scott, P. R. (2005). Plant disease: A threat to global food security. *Annu. Rev. Phytopathol.* 43, 83–116. doi:10.1146/annurev.phyto.43.113004.133839.

Sulakvelidze, A., Alavidze, Z., and Morris, J. G. (2001). Bacteriophage therapy. *Antimicrob. Agents Chemother.* 45, 649–59. doi:10.1128/AAC.45.3.649-659.2001.

Sykes, I. K., Lanning, S., and Williams, S. T. (1981). The effect of pH on soil actinophage. *J. Gen. Microbiol.* 122, 271–280. doi:10.1099/00221287-122-2-271.

Tailliez, P., Laroui, C., Ginibre, N., Paule, A., Pagès, S., and Boemare, N. (2010). Phylogeny of *Photorhabdus* and *Xenorhabdus* based on universally conserved protein-coding sequences and implications for the taxonomy of these two genera. Proposal of new taxa: *X. vietnamensis* sp. nov., *P. luminescens* subsp. *caribbeanensis* subsp. nov., P. l. *Int. J. Syst. Evol. Microbiol.* 60, 1921–1937. doi:10.1099/ijs.0.014308-0.

Tanizawa, Y., Fujisawa, T., Kaminuma, E., Nakamura, Y., and Arita, M. (2016). DFAST and DAGA: web-based integrated genome annotation tools and resources. *Biosci. microbiota, food Heal.* 35, 173–184. doi:10.12938/bmfh.16-003.

Tanizawa, Y., Fujisawa, T., and Nakamura, Y. (2018). DFAST: A flexible prokaryotic genome annotation pipeline for faster genome publication. *Bioinformatics* 34, 1037–1039. doi:10.1093/bioinformatics/btx713.

Tanui, C. K., Shyntum, D. Y., Priem, S. L., Theron, J., and Moleleki, L. N. (2017). Influence of the ferric uptake regulator (Fur) protein on pathogenicity in *Pectobacterium carotovorum* subsp. *brasiliense*. *PLoS One* 12, e0177647. doi:10.1371/journal.pone.0177647.

Toth, I. K., van der Wolf, J. M., Saddler, G., Lojkowska, E., Hélias, V., Pirhonen, M., *et al*. (2011). *Dickeya* species: an emerging problem for potato production in Europe. *Plant Pathol.* 60, 385–399. doi:10.1111/j.1365-3059.2011.02427.x.

United Nations News Service (2013). World must sustainably produce 70 per cent more food by mid-century – UN report. *United Nations*, 1–2. Available at: https://news.un.org/en/story/2013/12/456912#.Vvxj0uIrLIU [Accessed May 26, 2020].

Wang, F., Wang, D., Hou, W., Jin, Q., Feng, J., and Zhou, D. (2019). Evolutionary Diversity of Prophage DNA in *Klebsiella pneumoniae* Chromosomes. *Front. Microbiol.* 10, 2840. doi:10.3389/fmicb.2019.02840.

Wang, G. F., Xie, G. L., Zhu, B., Huang, J. S., Liu, B., Kawicha, P., *et al*. (2010). Identification and characterization of the *Enterobacter* complex causing mulberry (*Morus alba*) wilt disease in China. *Eur. J. Plant Pathol.* 126, 465–478. doi:10.1007/s10658-009-9552-x.

Washington, J. A., Yu, P., and Martin, W. J. (1969). Biochemical and clinical characteristics and antibiotic susceptibility of atypical *Enterobacter cloacae*. *Appl. Microbiol.* 17, 843–846. doi:10.1128/aem.17.6.843-846.1969.

Weinbauer, M. G. (2004). Ecology of prokaryotic viruses. *FEMS Microbiol. Rev.* 28,

127–181. doi:10.1016/j.femsre.2003.08.001.
Williams Smith and Huggins, M. B. (1983). Effectiveness of phages in treating experimental *Escherichia coli* diarhoea in calves, piglets and lambs. *J. Gen. Microbiol.* 129, 2659–2675. doi:10.1099/00221287-129-8-2659.
Wommack, K. E., Hill, R. T., Muller, T. A., and Colwell, R. R. (1996). Effects of sunlight on bacteriophage viability and structure. *Appl. Environ. Microbiol.* 62, 1336–1341. doi:10.1128/aem.62.4.1336-1341.1996.
Wu, J., Ding, Z., Diao, Y., and Hu, Z. (2011). First report on *Enterobacter* sp. causing soft rot of *Amorphophallus konjac* in China. *J. Gen. Plant Pathol.* 77, 312–314. doi:10.1007/s10327-011-0330-1.
Wyatt, G. M., and Lund, B. M. (1981). The effect of antibacterial products on bacterial soft rot of potatoes. *Potato Res.* 24, 315–329. doi:10.1007/BF02360369.
Xu, J., Gui, M., Wang, D., and Xiang, Y. (2016). The bacteriophage φ29 tail possesses a pore-forming loop for cell membrane penetration. *Nature* 534, 544–547. doi:10.1038/nature18017.
Xu, M., Struck, D. K., Deaton, J., Wang, I. N., and Young, R. (2004). A signal-arrest-release sequence mediates export and control of the phage P1 endolysin. *Proc. Natl. Acad. Sci. U. S. A.* 101, 6415–6420. doi:10.1073/pnas.0400957101.
Xue, Q., and Egan, J. B. (1995). Tail sheath and tail tube genes of the temperate coliphage 186. *Virology* 212, 218–221. doi:10.1006/viro.1995.1471.
Yamada, T. (2013). Filamentous phages of *Ralstonia solanacearum*: Double-edged swords for pathogenic bacteria. *Front. Microbiol.* 4, 325. doi:10.3389/fmicb.2013.00325.
Yin, J. (1993). Evolution of bacteriophage T7 in a growing plaque. *J. Bacteriol.* 175, 1272–1277. doi:10.1128/jb.175.5.1272-1277.1993.
Yoshikawa, G., Askora, A., Blanc-Mathieu, R., Kawasaki, T., Li, Y., Nakano, M., *et al.* (2018). *Xanthomonas citri* jumbo phage XacN1 exhibits a wide host range and high complement of tRNA genes. *Sci. Rep.* 8, 4486. doi:10.1038/s41598-018-22239-3.
Youdkes, D., Helman, Y., Burdman, S., Matan, O., and Jurkevitch, E. (2020). Potential control of potato soft rot by the obligate predators bdellovibrio and like organisms. *Appl. Environ. Microbiol.* 86. doi:10.1128/AEM.02543-19.
Young, J. M., and Park, D. C. (2007). Relationships of plant pathogenic enterobacteria based on partial atpD, carA, and recA as individual and concatenated nucleotide and peptide sequences. *Syst. Appl. Microbiol.* 30, 343–354. doi:10.1016/j.syapm.2007.03.002.
Zerbino, D. R., and Birney, E. (2008). Velvet algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* 18, 821–829. doi:10.1101/gr.074492.107.
Zerbino, D. R., McEwen, G. K., Margulies, E. H., and Birney, E. (2009). Pebble and rock band: Heuristic resolution of repeats and scaffolding in the Velvet short-read de Novo Assembler. *PLoS One* 4, e8407. doi:10.1371/journal.pone.0008407.
Zhang, C., Lin, T., Li, J., Ma, G., Wang, Y., Zhu, P., *et al.* (2016). First report of the melon stem rot disease in protected cultivation caused by *Pseudomonas fluorescens*. *J. Plant Dis. Prot.* 123, 247–255. doi:10.1007/s41348-016-0030-3.
Zhang, Y., and Qiu, S. (2015). Examining phylogenetic relationships of *Erwinia* and *Pantoea* species using whole genome sequence data. *Antonie van Leeuwenhoek,*

Int. J. Gen. Mol. Microbiol. 108, 1037–1046. doi:10.1007/s10482-015-0556-6.
Zhao, J., Zhang, Z., Tian, C., Chen, X., Hu, L., Wei, X., *et al*. (2019). Characterizing the Biology of Lytic Bacteriophage vB_EaeM_φEap-3 Infecting Multidrug-Resistant *Enterobacter aerogenes*. *Front. Microbiol.* 10, 420. doi:10.3389/fmicb.2019.00420.
Zhong, L., Harijati, N., Ding, Y., Bao, Z. Z., Ke, W. D., and Hu, Z. L. (2015). First report of black rot of *Sagittaria sagittifolia* caused by *Bacillus amyloliquefaciens* in China. *Plant Dis.* 99, 1270. doi:10.1094/PDIS-02-15-0148-PDN.
Zhou, J. N., Liu, S. Y., Chen, Y. F., and Liao, L. S. (2015). First Report of *Pantoea anthophila* Causing Soft rot in *Clausena lansium* (Wampee) in China. *Plant Dis.* 99, 416–416. doi:10.1094/PDIS-10-14-1025-PDN.

Appendix

Appendix 1: Annotation table of phage EspM4VN's genome

Gene	Start	End	Products	Group	Genbank ID
gene_1	2	382	Hypothetical phage protein	Hypothetical protein	BBD52181.1
gene_2	379	933	Putative dUTP diphosphatase	DNA replication	BBD52182.1
gene_3	933	1475	Hypothetical protein	Hypothetical protein	BBD52183.1
gene_4	1460	2551	RecA recombination protein	DNA replication	BBD52184.1
gene_5	2529	2858	Hypothetical protein	Hypothetical protein	BBD52185.1
gene_6	2865	4292	Putative DNA primase-helicase subunit gp41	DNA replication	BBD52186.1
gene_7	4355	4690	Hypothetical phage protein	Hypothetical protein	BBD52187.1
gene_8	4706	5020	Hypothetical protein	Hypothetical protein	BBD52188.1
gene_9	5148	6344	Hypothetical protein	Hypothetical protein	BBD52189.1
gene_10	6457	6657	Putative Hypothetical protein	Hypothetical protein	BBD52190.1
gene_11	6657	6869	Hypothetical phage protein	Hypothetical protein	BBD52191.1
gene_12	6871	7218	Hypothetical phage protein	Hypothetical protein	BBD52192.1
gene_13	7278	7886	Putative Hypothetical protein orf00049	Hypothetical protein	BBD52193.1
gene_14	7889	8470	Hypothetical membrane protein	Structure and packaging	BBD52194.1
gene_15	8470	8655	Hypothetical protein	Hypothetical protein	BBD52195.1
gene_16	8652	8831	Hypothetical protein	Hypothetical protein	BBD52196.1

gene_17	8887	10314	DNA ligase	DNA replication	BBD52197.1
gene_18	10311	10535	Hypothetical protein	Hypothetical protein	BBD52198.1
gene_19	10528	10773	Hypothetical membrane protein	Structure and packaging	BBD52199.1
gene_20	10775	11032	Hypothetical membrane protein	Structure and packaging	BBD52200.1
gene_21	11073	11339	Hypothetical phage protein	Hypothetical protein	BBD52201.1
gene_22	11333	11992	Putative loader of gp41 DNA helicase gp59	DNA replication	BBD52202.1
gene_23	11993	13069	Putative tail length tape measure protein	Structure and packaging	BBD52203.1
gene_24	13000	13944	Putative tail length tape measure protein	Structure and packaging	BBD52204.1
gene_25	13955	15343	Putative baseplate hub subunit gp27	Structure and packaging	BBD52205.1
gene_26	15340	15894	Putative baseplate wedge subunit gp53	Structure and packaging	BBD52206.1
gene_27	15906	16217	Putative tail tube associated baseplate protein gp48	Structure and packaging	BBD52207.1
gene_28	16924	17538	Gp4 head completion protein	Structure and packaging	BBD52208.1
gene_29	17535	17747	Hypothetical protein	Hypothetical protein	BBD52209.1
gene_30	17750	18157	Hypothetical protein	Hypothetical protein	BBD52210.1
gene_31	18168	18674	Putative deoxycytidylate deaminase	DNA replication	BBD52211.1
gene_32	18674	18988	Hypothetical protein	Hypothetical protein	BBD52212.1
gene_33	19059	19391	Hypothetical protein	Hypothetical protein	BBD52213.1
gene_34	19388	19705	Hypothetical protein	Hypothetical protein	BBD52214.1
gene_35	19702	20115	Hypothetical protein	Hypothetical protein	BBD52215.1
gene_36	20181	20366	Hypothetical protein	Hypothetical protein	BBD52216.1

gene_37	20363	20728	Hypothetical protein	Hypothetical protein	BBD52217.1
gene_38	20922	21263	Hypothetical phage protein	Hypothetical protein	BBD52218.1
gene_39	21270	21965	Putative Hypothetical protein	Hypothetical protein	BBD52219.1
gene_40	21965	22450	Hypothetical protein	Hypothetical protein	BBD52220.1
gene_41	22587	22892	Putative Hypothetical protein	Hypothetical protein	BBD52221.1
gene_42	22889	23521	Hypothetical protein	Hypothetical protein	BBD52222.1
gene_43	23508	23903	Hypothetical protein	Hypothetical protein	BBD52223.1
gene_44	23903	24502	Tk.4 protein	DNA replication	BBD52224.1
gene_45	24499	24741	Hypothetical protein	Hypothetical protein	BBD52225.1
gene_46	24792	25202	Putative Hypothetical protein orf00021	Hypothetical protein	BBD52226.1
gene_47	25206	25403	Hypothetical protein	Hypothetical protein	BBD52227.1
gene_48	25623	25808	Hypothetical protein	Hypothetical protein	BBD52228.1
gene_49	25851	27182	Putative DNA topoisomerase/gyrase gp52	DNA replication	BBD52229.1
gene_50	27184	29097	Topoisomerase II large subunit	DNA replication	BBD52230.1
gene_51	29147	29728	Hypothetical protein	Hypothetical protein	BBD52231.1
gene_52	29725	30201	Hypothetical phage protein	Hypothetical protein	BBD52232.1
gene_53	30258	30455	Hypothetical protein	Hypothetical protein	BBD52233.1
gene_54	30501	31001	Hypothetical protein	Hypothetical protein	BBD52234.1
gene_55	31120	31356	Hypothetical protein	Hypothetical protein	BBD52235.1
gene_56	31359	32168	Hypothetical protein	Hypothetical protein	BBD52236.1

gene_57	32147	32539	Hypothetical protein	Hypothetical protein	BBD52237.1
gene_58	32569	32775	Hypothetical protein	Hypothetical protein	BBD52238.1
gene_59	32911	33264	Hypothetical protein	Hypothetical protein	BBD52239.1
gene_60	33319	34881	Hypothetical protein	Hypothetical protein	BBD52240.1
gene_61	34913	37669	RIIA protein	DNA replication	BBD52241.1
gene_62	37776	38123	Hypothetical protein	Hypothetical protein	BBD52242.1
gene_63	38134	38307	Hypothetical protein	Hypothetical protein	BBD52243.1
gene_64	38279	38470	Hypothetical protein	Hypothetical protein	BBD52244.1
gene_65	38479	39147	Hypothetical protein	Hypothetical protein	BBD52245.1
gene_66	39147	39521	Putative Hypothetical protein orf00256	Hypothetical protein	BBD52246.1
gene_67	39572	40018	Hypothetical protein	Hypothetical protein	BBD52247.1
gene_68	40151	40720	Hypothetical protein	Hypothetical protein	BBD52248.1
gene_69	40740	41855	Hypothetical protein	Hypothetical protein	BBD52249.1
gene_70	41852	42238	Hypothetical protein	Hypothetical protein	BBD52250.1
gene_71	42235	42459	Hypothetical protein (Fragment)	Hypothetical protein	BBD52251.1
gene_72	42541	42708	Hypothetical protein	Hypothetical protein	BBD52252.1
gene_73	42745	43980	Hypothetical protein	Hypothetical protein	BBD52253.1
gene_74	43977	44162	Hypothetical membrane protein	Structure and packaging	BBD52254.1
gene_75	44220	44516	Hypothetical protein	Hypothetical protein	BBD52255.1
gene_76	44513	44710	Hypothetical protein	Hypothetical protein	BBD52256.1

gene_77	44713	45627	Hypothetical protein	Hypothetical protein	BBD52257.1
gene_78	45627	45932	Hypothetical phage protein	Hypothetical protein	BBD52258.1
gene_79	45942	46520	Putative 5'(3') deoxyribonucleotidase	Signal transduction and regulatory	BBD52259.1
gene_80	46490	46717	Hypothetical protein	Hypothetical protein	BBD52260.1
gene_81	46714	47055	Hypothetical phage protein	Hypothetical protein	BBD52261.1
gene_82	47117	50116	DNA polymerase	DNA replication	BBD52262.1
gene_83	50196	50768	Hypothetical protein	Hypothetical protein	BBD52263.1
gene_84	50811	51023	Hypothetical protein	Hypothetical protein	BBD52264.1
gene_85	51119	51646	Hypothetical protein	Hypothetical protein	BBD52265.1
gene_86	51695	52246	Hypothetical protein	Hypothetical protein	BBD52266.1
gene_87	52246	52605	Hypothetical protein	Hypothetical protein	BBD52267.1
gene_88	52631	53089	Hypothetical protein	Hypothetical protein	BBD52268.1
gene_89	53100	53588	Hypothetical protein	Hypothetical protein	BBD52269.1
gene_90	53628	54821	Hypothetical protein	Hypothetical protein	BBD52270.1
gene_91	54891	55592	Hypothetical phage protein	Hypothetical protein	BBD52271.1
gene_92	55663	56181	Hypothetical protein	Hypothetical protein	BBD52272.1
gene_93	56190	56426	Hypothetical protein	Hypothetical protein	BBD52273.1
tRNA	56908	56996	tRNA-Ser		
tRNA	57256	57339	tRNA-Tyr		
tRNA	57597	57672	tRNA-Asn		

gene_94	57674	57871	Hypothetical protein	Hypothetical protein	BBD52274.1
gene_95	57908	58135	Hypothetical protein	Hypothetical protein	BBD52275.1
tRNA	58560	58639	tRNA-Ser		
gene_96	58921	59133	Hypothetical protein	Hypothetical protein	BBD52276.1
gene_97	60617	61189	Hypothetical phage protein	Hypothetical protein	BBD52277.1
gene_98	61538	63316	Putative baseplate wedge subunit gp6	Structure and packaging	BBD52278.1
gene_99	63300	64154	Putative baseplate wedge subunit gp7	Structure and packaging	BBD52279.1
gene_100	64157	65368	Hypothetical protein	Hypothetical protein	BBD52280.1
gene_101	65421	68750	Hypothetical protein	Hypothetical protein	BBD52281.1
gene_102	68796	70592	Hypothetical protein	Hypothetical protein	BBD52282.1
gene_103	70647	73211	Hypothetical protein	Hypothetical protein	BBD52283.1
gene_104	73276	75846	Hypothetical protein	Hypothetical protein	BBD52284.1
gene_105	75943	80703	Hypothetical protein	Hypothetical protein	BBD52285.1
gene_106	80832	81083	Uncharacterzed protein	Hypothetical protein	BBD52286.1
gene_107	81064	81402	Hypothetical protein	Hypothetical protein	BBD52287.1
gene_108	81392	82144	Hypothetical protein	Hypothetical protein	BBD52288.1
gene_109	82172	82384	Hypothetical protein	Hypothetical protein	BBD52289.1
gene_110	82444	83094	Putative neck protein gp14	Structure and packaging	BBD52290.1
gene_111	83097	83792	Putative tail sheath stabilizer gp15	Structure and packaging	BBD52291.1
gene_112	83795	84490	Putative terminase small subunit gp16	Structure and packaging	BBD52292.1

gene_113	84471	86681	Gp17 terminase DNA packaging enzyme large subunit	Structure and packaging	BBD52293.1
tRNA	84879	84966	tRNA-Asp		
gene_114	86734	88632	Putative tail sheath protein gp18	Structure and packaging	BBD52294.1
gene_115	88780	89241	Putative tail tube protein gp19	Structure and packaging	BBD52295.1
gene_116	89309	91000	Putative portal protein	Structure and packaging	BBD52296.1
gene_117	91039	91206	Hypothetical protein	Hypothetical protein	BBD52297.1
gene_118	91217	91522	Hypothetical protein	Hypothetical protein	BBD52298.1
gene_119	91533	92198	Putative prohead protease gp21	Structure and packaging	BBD52299.1
gene_120	92244	93119	Gp22 prohead core protein	Structure and packaging	BBD52300.1
gene_121	93211	94533	Gp23 major head protein	Structure and packaging	BBD52301.1
gene_122	94626	94850	Hypothetical protein	Hypothetical protein	BBD52302.1
gene_123	94943	95260	Putative Hypothetical protein	Hypothetical protein	BBD52303.1
gene_124	95318	96175	Hypothetical protein	Hypothetical protein	BBD52304.1
gene_125	96233	96457	Hypothetical protein	Hypothetical protein	BBD52305.1
gene_126	96475	96915	Hypothetical protein	Hypothetical protein	BBD52306.1
gene_127	96924	97160	Hypothetical phage protein	Hypothetical protein	BBD52307.1
gene_128	97264	97560	Hypothetical protein	Hypothetical protein	BBD52308.1
gene_129	97630	97950	Hypothetical phage protein	Hypothetical protein	BBD52309.1
gene_130	97956	98378	Hypothetical protein	Hypothetical protein	BBD52310.1
gene_131	98420	98587	Hypothetical protein	Hypothetical protein	BBD52311.1

gene_132	98625	99353	Hypothetical protein	Hypothetical protein	BBD52312.1
gene_133	99354	99878	Hypothetical protein	Hypothetical protein	BBD52313.1
gene_134	100034	100534	Putative tail completion & sheath stabilizer protein gp3	Structure and packaging	BBD52314.1
gene_135	100577	101029	Putative DNA repair/recombination protein UvsY	DNA replication	BBD52315.1
gene_136	101029	101775	Putative exonuclease	Signal transduction and regulatory	BBD52316.1
gene_137	101804	101977	Hypothetical protein	Hypothetical protein	BBD52317.1
gene_138	101953	103311	RNA-DNA and DNA-DNA helicase UvsW	DNA replication	BBD52318.1
gene_139	103302	103678	Hypothetical protein	Hypothetical protein	BBD52319.1
gene_140	104024	104692	Sliding clamp DNA polymerase accessory protein	DNA replication	BBD52320.1
gene_141	104772	105761	Putative DNA polymerase accessory protein gp44	DNA replication	BBD52321.1
gene_142	105766	106188	Putative clamp loader subunit gp62	Signal transduction and regulatory	BBD52322.1
gene_143	106218	106682	Hypothetical protein	Hypothetical protein	BBD52323.1
gene_144	106699	107562	Nucleoside triphosphate pyrophosphohydrolase	Signal transduction and regulatory	BBD52324.1
gene_145	107631	108845	Hypothetical protein	Hypothetical protein	BBD52325.1
gene_146	108961	110991	Hypothetical protein	Hypothetical protein	BBD52326.1
gene_147	111036	111407	Hypothetical protein	Hypothetical protein	BBD52327.1
gene_148	111849	112643	Hypothetical protein	Hypothetical protein	BBD52328.1
gene_149	112754	113608	Ribose-phosphate pyrophosphokinase	DNA replication	BBD52329.1
gene_150	113605	115269	Nicotinamide phosphoribosyltransferase	DNA replication	BBD52330.1
gene_151	115311	115481	Hypothetical phage protein	Hypothetical protein	BBD52331.1

gene_152	115474	117705	Putative vWa containing protein	DNA replication	BBD52332.1
gene_153	117748	118074	Hypothetical protein	Hypothetical protein	BBD52333.1
gene_154	118198	118437	Hypothetical protein	Hypothetical protein	BBD52334.1
gene_155	118437	118682	Hypothetical protein	Hypothetical protein	BBD52335.1
gene_156	118742	119197	Putative pyimidine dimer DNA glycosylase DenV	DNA replication	BBD52336.1
gene_157	119210	119584	Hypothetical protein	Hypothetical protein	BBD52337.1
gene_158	119581	119868	Hypothetical protein	Hypothetical protein	BBD52338.1
gene_159	119870	120028	Hypothetical protein	Hypothetical protein	BBD52339.1
gene_160	120062	120511	Hypothetical protein	Hypothetical protein	BBD52340.1
gene_161	120614	121276	Hypothetical protein	Hypothetical protein	BBD52341.1
gene_162	121273	121605	Hypothetical protein	Hypothetical protein	BBD52342.1
gene_163	121655	122581	Hypothetical phage protein	Hypothetical protein	BBD52343.1
gene_164	122638	123096	Hypothetical protein	Hypothetical protein	BBD52344.1
gene_165	123029	123289	Hypothetical phage protein	Hypothetical protein	BBD52345.1
gene_166	123291	123740	Hypothetical protein	Hypothetical protein	BBD52346.1
gene_167	123833	124111	Putative DNA-binding protein	Signal transduction and regulatory	BBD52347.1
gene_168	124241	124585	Hypothetical protein	Hypothetical protein	BBD52348.1
gene_169	124471	125958	Putative ATP-dependent DNA helicase	DNA replication	BBD52349.1
gene_170	125955	126722	Hypothetical protein	Hypothetical protein	BBD52350.1
gene_171	126768	127295	Putative ribonuclease HI	DNA replication	BBD52351.1

gene_172	127307	128098	Gp55 T4-like sigma factor involved in late transcription	Signal transduction and regulatory	BBD52352.1
gene_173	128085	129200	Putative endonuclease gp47	Signal transduction and regulatory	BBD52353.1
gene_174	129203	131539	Putative recombination endonuclease subunit gp46	DNA replication	BBD52354.1
gene_175	131542	131706	Hypothetical protein	Hypothetical protein	BBD52355.1
gene_176	131703	132032	Hypothetical protein	Hypothetical protein	BBD52356.1
gene_177	132013	132297	Hypothetical protein	Hypothetical protein	BBD52357.1
gene_178	132397	132397	Putative RegB protein	Signal transduction and regulatory	BBD52358.1
gene_179	133011	133649	Hypothetical protein	Hypothetical protein	BBD52359.1
gene_180	133757	134128	Hypothetical protein	Hypothetical protein	BBD52360.1
gene_181	134151	134462	Hypothetical protein	Hypothetical protein	BBD52361.1
gene_182	134536	136176	Hypothetical protein	Hypothetical protein	BBD52362.1
gene_183	136233	136802	Hypothetical protein	Hypothetical protein	BBD52363.1
gene_184	136812	137294	Hypothetical protein	Hypothetical protein	BBD52364.1
gene_185	137345	137698	Hypothetical protein	Hypothetical protein	BBD52365.1
gene_186	137758	138267	Hypothetical protein	Hypothetical protein	BBD52366.1
gene_187	138267	138878	Hypothetical protein	Hypothetical protein	BBD52367.1
gene_188	138878	139936	DNA primase	DNA replication	BBD52368.1
gene_189	139933	140139	Hypothetical phage protein	Hypothetical protein	BBD52369.1
gene_190	140335	140823	Hypothetical protein	Hypothetical protein	BBD52370.1
gene_191	140885	141073	Hypothetical protein	Hypothetical protein	BBD52371.1

gene_192	141070	141357	Hypothetical protein	Hypothetical protein	BBD52372.1
gene_193	141427	142221	Putative endolysin	Endolysin	BBD52373.1
gene_194	142326	143168	PhoH-like protein	Signal transduction and regulatory	BBD52374.1
gene_195	143254	145533	Ribonucleoside-diphosphate reductase	DNA replication	BBD52375.1
gene_196	145604	146707	NrdB ribonucleotide reductase subunit beta	DNA replication	BBD52376.1
gene_197	146717	146941	Hypothetical protein	Hypothetical protein	BBD52377.1
gene_198	147047	147508	Hypothetical protein	Hypothetical protein	BBD52378.1
gene_199	147515	147904	Hypothetical protein	Hypothetical protein	BBD52379.1
gene_200	147905	148285	Hypothetical protein	Hypothetical protein	BBD52380.1
gene_201	148358	149968	Gp5 baseplate hub subunit and tail lysozyme	Structure and packaging	BBD52381.1
gene_202	150479	151183	Putative baseplate hub subunit gp26	Structure and packaging	BBD52382.1
gene_203	151336	151860	Hypothetical protein	Hypothetical protein	BBD52383.1
gene_204	151838	152329	Hypothetical protein	Hypothetical protein	BBD52384.1
gene_205	152369	152971	Putative holliday junction resolvase RuvC	DNA replication	BBD52385.1
gene_206	153505	153750	Putative late promoter transcription accessory protein gp33	Signal transduction and regulatory	BBD52386.1
gene_207	153758	153997	Hypothetical protein	Hypothetical protein	BBD52387.1
gene_208	154096	155142	Putative ssDNA binding protein gp32	Signal transduction and regulatory	BBD52388.1
gene_209	155169	155501	Putative baseplate tail tube initiator gp54	Structure and packaging	BBD52389.1
gene_210	156165	156542	Putative DNA end protector protein gp2	Signal transduction and regulatory	BBD52390.1
gene_211	156473	156865	Gp2 DNA end protector protein	Signal transduction and regulatory	BBD52391.1

gene_212	156931	157095	Hypothetical protein	Hypothetical protein	BBD52392.1
gene_213	157092	157682	Hypothetical protein	Hypothetical protein	BBD52393.1
gene_214	157889	158557	Hypothetical protein	Hypothetical protein	BBD52394.1
gene_215	158476	158814	Hypothetical protein	Hypothetical protein	BBD52395.1
gene_216	159178	159564	Hypothetical protein	Hypothetical protein	BBD52396.1
gene_217	159805	160071	Thymidylate synthase	DNA replication	BBD52397.1
gene_218	160029	160370	Putative thymidylate synthase	DNA replication	BBD52398.1
gene_219	160354	160515	Thymidylate synthase	DNA replication	BBD52399.1

Appendix 2: General genomic features of phage EspM4VN and related phages

Features	Bacteriophages											
	EM4	**SboM-AG3**	**SKML-39**	**SH19**	**Coodle**	**PP35**	**JA15**	**LIMEstone1**	**phiDP23.1**	**D3**	**D5**	**XF4**
Genome size (bp)	160,766	158,006	159,624	157,785	152,515	152,048	153,757	152,427	188,540	152 308	155,346	151519
% of coding regions in the genome	90.6	92.7	88.6	-	94.4	92.7	90.7	-	83.2	91.2	89.9	91
Number of predicted ORFs (PEGs)	219	216	209	339	202	198	199	201	223	191	196	196
% of PEGs with assigned fuctions	30% with assigned 9% with hypthetical functions 61%	33% with assigned functions 67% with hypothetical	9% with assigned functions 91% with hypothetical	31% with assigned functions 17% with hypothetical	42% with assigned functions 58% with hypothetical	54% with assigned functions 46% with hypothetical	45% with assigned 48% with hypthetical functio	46% with assigned 54% with unknown function	39.5% with assigned 22.4% with cal functions	54.9% with assigned 45.1% with hypthetical functions	25.5% with assigned 49% with hypthetical functions	

	with unknown function	functions	functions	functions 52% with unknown function	functions	functions	ns 7% with unknown function		8% with unknown function		25.5% with unknown function	
Average gene length (bp)	653	731	677	-	709	712	693	-	538-804	730	711	703
Number of PEGs in functional groups	Phage replication (8) Phage tail proteins 2 (1) Phage lysis modules (1) Nucleosides and Nucleotides (2)	-	Phage tail proteins 2 (1) Phage tail fiber proteins (1) Phage lysis modules (1) RNA Metabol	Phage replication (10) Phage tail proteins 2 (1) Phage lysis modules (1) Nucleosides and Nucleotides (2)	Phage replication (9) Phage tail proteins 2 (1) Phage lysis modules (1) RNA Metabolism (1) Nucleos	Phage replication (14) Phage tail proteins 2 (1) Phage tail fiber proteins (1) Phage lysis modules (1)	Phage replication (14) Phage tail proteins 2 (1) Phage tail fiber proteins (1) Phage lysis module	Phage replication (10) Phage tail proteins 2 (1) Phage lysis modules (1) RNA Metaboli sm (1) Nucleosi	-	Phage replication (10) Phage tail proteins 2 (1) Phage lysis modules (1) RNA Metabolism (1) Nucleos	-	Phage replication (14) Phage tail proteins 2 (1) Phage tail fiber proteins (1) Phage lysis module

		ism (1) Nucleosides and Nucleotides (2) DNA Metabolism (1)		ides and Nucleotides (2) DNA Metabolism (1)	RNA processing and modification (1) Nucleosides and Nucleotides (2) DNA Metabolism (1)	s (1) RNA Metabolism (1) Nucleosides and Nucleotides (2) DNA Metabolism (1)	des and Nucleotides (2) DNA Metabolism (1)		ides and Nucleotides (2) DNA Metabolism (1)		s (1) RNA metablolism (1) Nucleosides and Nucleotides (2) DNA Metabolism (1)
Transcription start codon (% of genes of this start codon)	-	ATG (95.4%) GTG (3.2%) CTG (0.9%) TTG (0.5%)	-	-	-	-	-	-	ATG (85.3%) GTG (13.6%) TTG (1.1%)	-	ATG (94.4%) GTG (4.1%) TTG (1.5%)	-

Appendix 3: MS/MS results for phage proteins identified in Figure 4.7

Band	Gene name	Peptide count	Protein score	Ion Score	miss cleavage	± da	Start	End	Modification	Peptide sequence
1	gene 105	35	2305	25	0	0.002	1081	1089		FNTGDNVFR
				24	0	0.0006	989	998		DVTTYTYTDK
				41	0	0.0018	1069	1080		GVFAVPQNETIK
				33	0	0.0021	425	437		DLVVTGDMGTVTK
				43	1	0.0034	830	841	Deamidated (NQ)	NGIAADPFKDFR
				43	0	0.0027	329	341		FVAVMKPGISYVR
				74	0	0.0024	157	172		VGYGVAATVGGSIVAR
				70	1	0.0029	344	357		RIENLGEELVTIDK
				94	0	0.0038	249	263		FGYDEEVADFVELAK
				60	0	0.0045	726	741		IDAVYLADNGVFNVAR
				23	0	0.0039	1020	1034		TNINFEATGLRPFTR

70	0	0.0037	476	491		TLAPTGTIDINYTVLR
32	0	0.0013	966	980		VWEPHGAGGVWWGYR
41	0	0.0029	742	759		GVSSNNLASPAIPDNGMR
92	0	0.0024	610	627		NASGADVTSNFTLDGGQR
73	0	0.0009	37	51		ELNQMQTIFQDQLEK
101	0	-0.0003	396	414		FLNASSVTQATALLISAER
58	0	0.0044	59	78	2 Oxidation (M)	DGSMVIPGGLTITNTAVSMK
77	0	0.0035	842	860		LIDDLSSDWMGSIDTDNGR
44	0	0.0031	289	306		TYETNGDYNVSTHQIDLR
65	0	0.0009	920	936		LNPTTDYWFENYYVAPR
102	0	0.0007	360	380		DTDVLNNNPVAVATGNYLVSK
71	0	0.0039	760	778	Oxidation (M)	LYELMIPPYTANIDDIQIR

64	0	0.0062	999	1018	Oxidation (M)	TTTTLTGEQIVETQVIPYMR
66	0	0.0024	269	288	Oxidation (M)	IQSMVTQSTYNILEDTMAQR
44	1	0.0081	358	380		ARDTDVLNNNPVAVATGNYLVSK
41	0	-0.0009	79	101		FTLAGGSEFTDLEGISELYVLGK
87	0	0.0035	182	204	Oxidation (M)	GMFLDVEAATLIVDNASNSTSHR
29	0	0.0039	1043	1068		DVSPYCTPNGGALGGALNTDANGNIK
87	0	-0.0002	897	919		QDYATTTINVNPYAVFNWEGFLK
100	0	0.0013	209	233		VTETIVTEDDDESLFSNAQGTPNSK
76	0	0.0026	579	604		TVTEVTETVTFTTAASVPLANHDGYK
114	0	0.0025	699	725	Oxidation (M)	ITSGASDTDMVRPNTAIVLDAEYYLPR
102	0	-0.0004	495	524		ITLNAAGAGSISAPLGYSFSPEFSLYSAAK

				139	0	0.0026	797	827	Oxidation (M)	ISNVEYYTSLSQLESSAMTQQVFDPITGNPR
2	gene101	34	2215	27	0	0.0004	521	527		TILDYVK
				37	0	0.0014	600	609		VELDGVEATR
				56	0	0.0027	824	832		LDNLNMDIR
				64	0	0.0038	510	520		SILIYGDGIGK
				52	0	0.0021	305	315		GLGFTYDSDTK
				51	0	0.0033	459	469		ITFTDNTTALK
				64	0	0.0038	499	509		IILDPTTLPYK
				89	0	0.0024	995	1007		VSNQVLSAASPLK
				55	0	0.0069	789	800		LGSVVMNDCSIR
				41	0	0.0023	171	184		GWIVTAGVGLGQGR
				65	0	0.0034	118	130		VGDSVTVCDLYGK
				44	0	0.0023	843	854		TVYTFMPITMAR

60	0	0.0019	470	483	Deamidated (NQ)	EFLLSSENVVNGVK
70	0	0.0026	484	498		YVYIPTGHYGFGGPK
42	0	0.0065	82	97		WEALPHASTATLTEGR
113	0	0.0011	1008	1024		ILTGNLYGNLSMVLSGR
74	0	0.0057	855	872	Deamidated (NQ)	DNNTLVINGGDYLTGSGK
84	0	0.0037	353	370		SANVPVSQVILTSDTITK
65	0	0.0026	441	458		GVADYVGTPLYDGNDQTR
60	0	0.0009	528	547		EDNTGTGNTEQGDNALELIR
102	0	0.0014	98	117		GYLVDNSTNVSTVVFPSPTR
101	0	0.0001	873	893		MNDNTGVASNALFISQVATGR
87	0	0.0021	614	633	Oxidation (M)	CISIDGISDSSPAVTDLFMR
87	0	0.0007	663	685	Deamidated (NQ)	NGTLGVTATGYGLTFSQYCSNIK
69	0	0.0022	894	916	Oxidation (M)	VISDNACYDMQDMVLFGTTGGGR

				55	0	0.0026	971	994	Deamidated (NQ)	VGFGDCNVVADYAFGDLDNGVTFK
				70	0	-0.0014	801	823	Oxidation (M)	YLPGVTQNFYGYNGLNCMASELR
				70	0	0.0014	59	81		YNADSDEFEGFYENGGWLPLGGGGIR
				61	1	0.0074	521	547		TILDYVKEDNTGTGNTEQGDNALELIR
				57	0	0.0034	1075	1103		AIPLTTTTDLAVTEFYLGEITGYDSYTPK
				63	0	0.0068	189	218		TIFTETVTADTAQVTLTSQPSIVDVYVDGK
				54	0	-0.0014	316	345		IISLADELDAGDEVVVVINGDPTLYNQIDR
				54	1	0.0031	189	219		TIFTETVTADTAQVTLTSQPSIVDVYVDGKR
				71	0	0.0084	385	430	Oxidation (M)	MWSLPSGIPTGASIVSVSGSNLTYAPGNVVVSLKPVVGSGTDTDLK
3	gene 104	23	1558	47	0	0.0013	251	260	Oxidation (M)	TMSLLGLEAR
				57	0	0.0024	840	848		TLDEWYSYR
				69	0	0.0021	734	744		VTLLLTIENAR

48	0	-0.0001	558	568	Oxidation (M)	DITYTYMGGPR
73	0	0.0045	112	123		NELVSDGANLYR
35	0	0.0019	169	181		NDITFTTGGVVNK
53	0	0.0024	405	418		GGEFILEAPLSGGR
65	0	0.0012	231	243		AWYAPSSTLDNTK
72	0	0.0016	745	759		SSLSLNLLALPAGSK
52	0	0.0027	271	283		VTGYAQLDLSQCR
90	0	0.0029	786	802		NTVTGTNFGGSPAGSDK
75	0	0.0017	578	592		GDYGQSWEGVIDVSR
78	0	0.0027	284	300		IYIDPNSDDTVSGGGGK
68	0	-0.0004	131	151		TVSAGTSVGASGGVGSSAWVR
95	0	0.0003	93	109		GEFVNLVENFTTGFTLK

				55	0	0.0025	803	823		AYYYQNHITTVNAVAVGTVAK
				61	0	0.0045	786	802		CIIGNLVLGNANSAGEYYLER
				99	0	0.0015	669	692		YPETYLVDNFVVEGVGNNFSFDPR
				76	0	-0.0022	597	624		YGAAAPSGQGCSIVYCEGAYNGAIDLTR
				56	0	-0.0031	355	385	Oxidation (M)	QPQNTYIQASDLFVMDGGTTGLHPDTPALFK
				127	0	0.0063	55	92		SYALPSNLPAGATITSLTDGILVHNTGTVDLGALAVLR
				80	0	-0.0031	301	338		VFLIEDPDTQVLTGQSFATSNIGQVAHVQLSGAAFSNK
				31	0	-0.0001	531	555		CTLMPVLVGGNYQGGISLTGGGLLK
3	gene 103	41	2834	33	0	0.0006	320	328		TTGGAAWIR
				65	0	0.0018	200	209		TVIYDVVAEK
				35	0	0.0019	161	170		QPNISPLFDR

41	0	0.0021	443	453		NVAFYGLPAQR
60	0	0.0003	82	93		ILSPEDYADGIR
80	0	0.001	632	645		DSVSSLALTGPNSR
73	0	0.0023	607	619		LSTPYFMVCSGER
49	0	0.0033	68	81		TLEPEDVVYCLFDK
81	0	0.0023	303	319		AVLDGSSYTDNSGTIIK
112	0	0.0027	823	838	Oxidation (M)	GNVNYTMTAGVPVEFR
107	0	0.0022	329	346		VNDVVSVTDFGAVSDFNR
82	0	0.001	802	822		VSVIANTAGAVINTTGNVYPK
27	0	-0.0023	347	366		DTNVGTDNTTAFQAAIDYCK
50	1	0.0013	801	822		RVSVIANTAGAVINTTGNVYPK
99	0	-0.0014	178	199		STNLADTAVILSTDIGTPLDDK

				26	1	0.0004	251	274		LKEELAEFDGESLVGVCPTVAILR
				81	0	0.001	97	123		FQAVGTETSFTPDFTTYGVQTLYIDGK
				108	1	-0.0001	178	209		STNLADTAVILSTDIGTPLDDKTVIYDVVAEK
				99	0	-0.0042	210	250		IYGLPTLPSNAYINTVSNGQLTYSPGNVTVDLVAIPDSAVK
4	gene 114	24	1924	58	0	-0.0024	296	304		SDGANAYFK
				52	0	0.0012	449	457		SIELAGIYK
				52	0	0.0028	379	388		DTVSFVSPLR
				64	0	0.0012	398	406	Oxidation (M)	EMDDVVAWR
				43	0	0.0002	473	482	Oxidation (M)	MAWSASSDER
				29	1	-0.0007	378	388		KDTVSFVSPLR
				63	0	0.0136	174	186	Deamidated (NQ)	ITNSSGAVGQVER

21	1	-0.0016	396	406	Oxidation (M)	GREMDDVVAWR
49	0	0.0022	435	448		WIPACGGTAGVWAR
85	0	0.0027	538	549		YYLGENNDEFTR
94	0	0.0022	562	575	Oxidation (M)	QLTNMGAIYDGTVK
108	0	0.0026	523	537	Oxidation (M)	GLFIMAEQNIAAIAK
127	0	0.0038	247	263		VDIGPQTAIAVSVEASK
82	0	0.0037	2	17		ATTSFSVAPSVQWTER
95	0	0.0048	110	126		LDFESASPSASITWTGR
116	0	0.005	18	36		DATLQTSPSVVVQGATVGK
116	0	0.0054	264	283		GITIGNVVTTVGTSGSIIEK
39	0	0.0064	152	170		NNFAYAPQAGEYHIVVLDK
57	0	-0.0022	412	428	Oxidation (M)	DSSYFFMDDNWAYVYDK

				98	0	0	488	507		NQINSIVTFSTEGIVLYGDK
				89	0	0.0054	37	57		FQWGEVELPVLVTGGETGLVK
				65	1	0.0015	412	432	Oxidation (M)	DSSYFFMDDNWAYVYDKYNDK
				102	0	0.0032	127	151		YPGSLGNDVAVNICDAASFSTWEFR
				113	0	-0.0013	308	339		DVINDTSNWVYVFTDTLAAGVTELEGGVDDYDVNR
				107	0	-0.007	340	377		VAAIEVLNNAEAYAAKPVFAYCEELIEQQAIIDLSTER
4	gene116	19	1245	39	0	0.0009	381	389	Oxidation (M)	LYEALMIPK
				76	0	0.0032	296	306		AFYLDVGTLGK
				59	0	0.0022	426	435		YSGFFLELLR
				60	0	0.0032	459	468		FEFTSDSFIR
				31	0	0.0054	426	435		NITDQTDWNEK
				35	1	-0.0026	197	209		ELRDGVEVISNVK

	42	0	0.0003	210	220		LQYYYNPNYNR
	63	0	0.0013	548	563		FRPNVIPAPIQPGGEE
	55	0	0.0006	138	152		EVLELMDFDNTAYQK
	100	0	0.0034	2	21	2 Deamidated (NQ)	AFGNGAFGNGFFGLFGTGGK
	49	1	0.0017	22	38		IEAPVDTDKLVTNQEEK
	63	0	0.0017	79	94	Oxidation (M)	QVVEEYQSMAQQPEIR
	52	0	0.002	333	350	Oxidation (M)	VAGNTHLMGIAEDYWLPR
	93	1	0.0022	524	547		DGKYPEVQADESGGFGGGEVSPLK
	133	0	0.001	480	501	Oxidation (M)	MSALGQIEPYIGTLFSIDYAQR
	113	0	0.0026	392	414		LQEEGSINIGGSNLAEITQEELR
	74	0	0.0063	352	378	Oxidation (M)	EGQNATEISNVGGGDQLGQMDHVVYFR
	39	1	0.0072	351	378	Oxidation (M)	REGQNATEISNVGGGDQLGQMDHVVYFR
	71	1	0.002	95	120		KAVDIIVNDVVTCEEDETPVTINLEK

4	gene 102	17	1064	58	0	0.0007	224	232		FTGAVEFLR
				39	0	0.0018	392	402		VDNVTIEMEVK
				33	0	0.0173	306	317	Deamidated (NQ)	NIIATYSNIDNR
				108	0	0.0013	469	482		FQGTNSNISVFNVK
				73	0	0.0053	486	503		AVFVGLDNVTDLTVNWER
				61	0	0.0033	248	266		YHPDVVAGGYGVLLEGCTR
				101	0	0.0014	549	569		NCTVTPASVAGLVVLPEITGR
				78	0	0.0007	575	595		FYNGSGTLVPLSSSIVSADIR
				51	0	-0.0013	24	45		AANVAESAVIFSSETTVNLDSK
				38	0	0.0037	120	142		RPDGYGGSFASVLANGQDVQVNK
				69	0	0.0024	365	391		IIQYDDNAIYGISGGVDNTPATVIGLR
				80	0	0.0025	143	172	Oxidation (M)	DVNLTSPILPGDYQVIQGAGGSVTMTSIAR

				38	0	-0.0731	330	359	Oxidation (M)	CMIDGFDINGSNGGIGLNADNGAIQDFQIR
				84	0	-0.0008	69	100	Oxidation (M)	IISVSGSNLTYTGPVTVPLVSYTGSADEMVAR
				62	0	0.0019	507	539		IVNNNGTITTTDSNSLISSVTTTGTGEITVVFK
				28	1	0.0026	110	142	Deamidated (NQ)	ISDDLNAIDRRPDGYGGSFASVLANGQDVQVNK
				60	0	0	431	468	Oxidation (M)	SHGSSNPILVPVGAMNLVFDGISDFVTAGSATGAALIR
4	gene98	12	660	27	0	0.0002	13	19		TFEFLLK
				25	0	-0.0011	505	514		DVSGVLNMYK
				65	0	0.0011	577	587		IGTVAVLPEVR
				55	0	0.0013	30	43		DYNFEGSGLSAIIR
				78	0	0.0002	72	89		SNVGLAAGFLSYTPDNYK
				64	0	0.0047	487	504		ISGFKPLPAEVDFTYYIR

				87	0	-0.001	127	146	SYNFTVDTPVSATLVDGSYK
				23	1	-0.0049	24	43	ADPTFKDYNFEGSGLSAIIR
				99	0	-0.0008	299	322	LAPLAYQAEGAAVAEMDYAVLTER
				46	0	-0.0004	90	115	AAFLYANVTVTPYDASTAPDTIVIDR
				48	1	0.0073	90	116	AAFLYANVTVTPYDASTAPDTIVIDRR
				43	0	-0.0046	332	363	AYGGDTLSPPDSGYVYIAVIPTTGETLSDAEK
5	gene182	28	1636	35	0	-0.0012	155	161	GYVWLVR
				33	0	0.0005	323	330	ALVNAFYK
				38	0	-0.0002	99	107	QFLATPLGK

29	0	0.0003	215	224	Deamidated (NQ)	SNGIFSSVVK
22	0	-0.0003	366	375		EDDPVVLTVK
78	0	0.0009	404	414		LTAIVQTAIDR
47	0	0.0018	108	118		TIDGYAVNFSK
49	0	0.0013	51	61		IQFYTPQGLGK
36	0	0.0015	446	455		DLDWQIYASR
21	0	0.0009	175	185	Oxidation (M)	MQGITWDDPDK
62	0	0.0012	426	438		TGVDLATYAGSIR
73	0	0.0003	253	264		TQTIAQVTVFNK
32	1	-0.0011	425	438		KTGVDLATYAGSIR
35	1	-0.0027	444	455		GRDLDWQIYASR

77	0	0.003	487	499		ANQYLQTLYQDAK
72	0	0.0002	336	352		GVSSSTMASGSVSFGAK
52	0	0.0027	502	515		GYNPTEPNLMTLAR
96	0	0.0001	141	154	Oxidation (M)	LNVMDLTYAPVPDR
85	0	0.0014	456	469		DGQTLNVDWDITFR
59	0	-0.0001	33	47		NAGDVFFTFIDEDEK
30	1	0.0065	32	47		KNAGDVFFTFIDEDEK
94	0	0.0007	225	243	Oxidation (M)	VVGANDVTLSSNDLAIMVK
73	1	0.0031	33	50		NAGDVFFTFIDEDEKEYR
78	0	-0.001	295	320		IVAYVEFGVGAVAPVQTTAAELTTDR
93	0	0.0017	186	214	Oxidation (M)	VGDVPVQSNAPDVGASSNNPDWNISVMAK

				88	1	-0.0041	295	322		IVAYVEFGVGAVAPVQTTAAELTTDRWR
				73	0	0.0089	516	545	Oxidation (M)	AQQSDEWAASNGESAYSEYEQSLGGNMQIK
				77	1	-0.0015	175	214	2 Oxidation (M)	MQGITWDDPDKVGDVPVQSNAPDVGASSNNPDWNISVMAK
6	gene 121	20	1846	42	0	-0.0003	239	246	Oxidation (M)	TMNFSAVR
				43	0	0.0006	382	391		TQGETTFSPR
				49	0	0.002	65	74	Oxidation (M)	WQPVLIDMAK
				54	0	0.0052	195	205		GLYADYSHELR
				28	0	-0.001	105	117		QGIADGSNTQQSR
				117	0	0.0043	250	265	Deamidated (NQ)	FGANGVVDISTDISGR
				66	1	0.0033	249	265	Deamidated (NQ)	KFGANGVVDISTDISGR
				132	0	0.0042	325	343	Deamidated (NQ)	LAIDPTGQTFAGVLSNGMR

	95	0	0.0129	344	360		VYIDPYAVAEYITLAYK
	101	0	0.0063	161	179		GMPTTDAELLGTTTNPWAR
	87	0	0.0013	273	290		YMTFMLEVEANGIGVDTR
	64	1	0.0031	273	291	2 Oxidation (M)	YMTFMLEVEANGIGVDTRR
	111	0	0.0009	361	381		GATALDAGIFFAPYVPLEMYR
	108	0	0.0043	76	102		LAPINIAMDFFGVQPLSGPDGQIFALR
	124	0	0.0062	297	324	Oxidation (M)	VLCSPNVASALAMSGMLDYAPALQENTK
	107	1	0.0021	75	102	Oxidation (M)	RLAPINIAMDFFGVQPLSGPDGQIFALR
	94	0	-0.0246	206	234	Oxidation (M)	QDMMAIHGEDVDSILSDVMVTEIQAEMNR
	100	0	-0.0013	398	432		YGICANPFVQIPANQDPQVYVTADGIAQDSNPYFR
	164	0	0.0028	119	160	Oxidation (M)	ELFMQEADSGYSGDGTVQAGDPSGFTQAEIEGSGAGVTTIGK
	159	1	0.003	118	160	Oxidation (M)	KELFMQEADSGYSGDGTVQAGDPSGFTQAEIEGSGAGVTTIGK

7	gene 100	4	194	36	0	0.00 06	364	379		STSGTVSFAGSVVSSK
				55	0	0.00 06	175	189		YDTTAFVTSIFNDTR
				47	0	0.00 27	99	118	Oxidation (M)	LSITDEPTVATDGVNLSMLR
				56	0	0.00 14	2	23		ADFLSQYTGQQIDAILGSVDDK
8	gene 121	11	807	111	0	0.00 22	250	265	Deamidated (NQ)	FGANGVVDISTDISGR
				31	0	0.00 72	325	343	Deamidated (NQ); Oxidation (M)	LAIDPTGQTFAGVLSNGMR
				58	0	-0.00 03	344	360		VYIDPYAVAEYITLAYK
				70	0	0.00 06	161	179	Oxidation (M)	GMPTTDAELLGTTTNPWAR
				51	0	0.00 38	273	290	2 Oxidation (M)	YMTFMLEVEANGIGVDTR
				91	0	0.00 05	361	381	Oxidation (M)	GATALDAGIFFAPYVPLEMYR
				115	0	0.00 06	76	102		LAPINIAMDFFGVQPLSGPDGQIFAL R
				30	0	0.01 08	297	324	2 Oxidation (M)	VLCSPNVASALAMSGMLDYAPAL QENTK

				44	0	0.0022	398	432		YGICANPFVQIPANQDPQVYVTADGIAQDSNPYFR
				131	0	0.0046	119	160	Oxidation (M)	ELFMQEADSGYSGDGTVQAGDPSGFTQAEIEGSGAGVTTIGK
				74	1	0.0064	118	160	Oxidation (M)	KELFMQEADSGYSGDGTVQAGDPSGFTQAEIEGSGAGVTTIGK
9	gene 56	4	339	60	0	0.0051	2	14		PTITVLVAPEVVR
				101	0	0.0027	35	49		TSLNQDPDEILTECK
				124	0	0.0045	50	67		GLDALLTQSNLQADGVTK
				56	1	0.0019	35	67		TSLNQDPDEILTECKGLDALLTQSNLQADGVTK
9	gene 26	3	201	30	0	-0.0028	40	57		DIEGLLLPYSVQEGETPR
				76	1	0.0053	38	57		IKDIEGLLLPYSVQEGETPR
				95	0	0.0013	162	184		NIQVLDPDYVSSFVNQLEQELAK
10	gene 92	9	710	52	0	-0.0001	68	80		FHAFPVDSILPVK
				76	0	0.0014	81	94		VNIIGTWSGSTSNR

				68	0	0.0014	48	62		TPHPSSVTGLAPFFK
				127	0	0.0029	95	110		TMLLDFVGSTGNQLSK
				58	0	0.0034	2	17	Oxidation (M)	AILPSPYLGNMLQTHR
				71	0	-0.0017	113	132		DSSLPPPDVLSFITFFSVDK
				103	0	0.0017	24	47		FSGLTAALPAGATGVDLLTLLDGK
				81	1	0.0003	111	132		SRDSSLPPPDVLSFITFFSVDK
				73	0	0.0018	145	172	Oxidation (M)	LNSYGGDFTINEIILIAEQVVPLYMNTI
10	gene115	9	700	40	0	0.0011	43	50		ELPFPGDR
				47	0	0.0015	144	151		YFQAISDR
				67	0	0.0012	86	95		ASANADDYFR
				102	0	0.003	96	110		DVIMELLDANDNVTK
				86	0	0.0061	1	16		VTVNFPAFVAGSDTIR

				100	0	0.0025	96	110	Oxidation (M)	DVIMELLDANDNVTK
				84	0	-0.0052	17	42		DVSLLAVTTNTPTGQLGEILVPWGGR
				134	1	0.0009	86	110	Oxidation (M)	ASANADDYFRDVIMELLDANDNVTK
				40	1	0.0041	1	42		VTVNFPAFVAGSDTIRDVSLLAVTTNTPTGQLGEILVPWGGR
10	gene 134	4	226	84	0	0.0012	21	31		LSIGDLTLVSR
				44	0	0.0005	150	160	Oxidation (M)	FQQMVPQLATK
				36	1	0.0033	8	20		NENPNFAASDKWR
				62	0	0.003	32	62		NIHDFSIPGLYSEGIDGPSPGDVLVSIPSER
10	gene 160	2	76	50	0	0.001	18	27		IPVGNLGLSR
				26	0	0.0021	122	131	Oxidation (M)	VDMPALEFVK