Phage-mediated biological control of *Erwinia amylovora*:

The role of CRISPRs and exopolysaccharide

By

Abdelbaset Yagubi, M.Sc.

Submitted to the Department of Biological Sciences

In conformity with the requirements

for the degree of

Doctor of Philosophy

November, 2016

Faculty of Sciences, Brock University

St. Catharines, Ontario

Abstract

Fire blight, caused by bacterium *Erwinia amylovora,* is a very serious disease affecting apple, pear and other fruit plants. The development of phage-based biopesticides is currently in progress in our lab. Emergence of phage-resistant bacteria is a valid concern. Two attributes of the bacterial host that may contribute to the development of resistance were studied, the Clustered Regularly Interspaced Short Palindromic Repeats/ CRISPR-associated (CRISPR/Cas) system and exopolysaccharide (EPS) interaction with phages.

The structure of *E. amylovora* CRISPR/Cas system was determined in 8 *E. amylovora* isolates from different geographical regions. Three CRISPR-array sets named CR1, CR2 and CR3 were detected in 4 isolates, and only 2 arrays were determined in the rest of the isolates. No significant similarity was found between spacers in any of these systems to phage DNA sequenced in this study or from GenBank. Also the *Cas* level of expression was not stimulated during phage infection. Introduction of extra copies of Cas genes to enhance expression did not result in phage resistance. Nevertheless, *E. amylovora* CRISPR/Cas system was found to be efficient in blocking the transformation of plasmids carrying protospacers matched spacers in CRR1 and CRR2.

Among phages that have been sequenced in this study are ΦEa9-2 and ΦEa35-70. ΦEa9-2 (*Podoviridae)* genome is 75,568 bp, and found to be related to coliphage N4. ΦEa35-70 (*Myoviridae*) genome is 271,084 bp, and found to carry a potential EPS depolymerase gene. Activity of ΦEa35-70 EPS

depolymerase was only detected when cloned and expressed in *E. coli,* but His-tagged purified protein did not exhibit any EPS-depolymerase activities.

This study offers critical information for the design of novel and effective phage-based biopesticides for the control of *E. amylovora*. It provides a new knowledge on the molecular structure and function of CRISPR/Cas system and EPS-phage interaction.

Acknowledgements

I would like to extend my gratitude to the many people who supported and inspired me to bring this research project to completion. Words are not enough to express my deep and sincere appreciation to my supervisors Dr. Antonet Svircev and Dr. Alan Castle for the help and support throughout this project and in the course of the process of writing this thesis. Also, I must express my profound gratefulness to my committee members Dr. Ping Liang, Dr. Yousef Haj-Ahmad and Dr. Charles Després for providing me with valuable guidance.

I would like to thank my lab members Darlene Nesbitt, Kathy Whybourne, David Sjaarda, Steven Gayder and Kinga Bodnar for the good times and continuous encouragement throughout my years of study.

I would like to offer my special thanks to Dr. Andrew Kropinski and Richard Gardiner for their very valuable help with this project.

My special thanks are extended to the staff of AAFC at Vineland Station especially Wendy Pawlina, Kathleen Janson and Neva Greig for the help and support.

This journey would not have been possible without the support of my family. I would like to thank my wife Junko and my boys Malik and Mazen for supporting and inspiring me to pursuit my dreams. Finally, all the love is extended to my parents for their prayers, inspiration and adoration.

Table of Contents

List of Figures

List of Tables

Abbreviations

CRISPRs - Clustered Regularly Interspaced Short Palindromic Repeats

Cas - CRISPR-associated

Cascade - CRISPR-associated complex for antiviral defence

PAM - Protospacer adjacent motif

Amp - Ampicillin

Kan - Kanamycin

ORF - Open reading frame

PCR - Polymerase chain reaction

qPCR - Quantitative real-time PCR

RT-qPCR - Reverse transcription qPCR

RT real-time PCR - Reverse transcription real-time PCR

NTC - No template control

TEM - Transmission electron microscopy

CFU - Colony forming unit

MOI - Multiplicity of infection; the ratio of phages to host cells

PFU - Plaque forming units

EOP - Efficiency of plating

EPS - Exopolysaccharide

lsc - Levansucrase

rcs - Regulator of capsular synthesis

hrp - Hypersensitive reaction and pathogenicity gene cluster

dsp - Disease-specific gene cluster

T3SS - Type three secretion system

HR - Hypersensitive response

HEP - High exopolysaccharide producing bacterial isolate

LEP - Low exopolysaccharide producing bacterial isolate

asm - Amylovoran synthesis

BIM - Bacteriophage insensitive mutant

NA - Nutrient agar

NAS - Nutrient agar with 2% sucrose and 1% sorbitol

NB - Nutrient broth

LB - Luria broth

PB - Phosphate buffer

PBS - Phosphate buffer saline

SOB - Super optimal broth

SOC - Super optimal broth with catabolic repressor

TAE - Tris base, acetic acid, and ETDA buffer

RFLP - Restriction fragment length polymorphism

IPTG - Isopropyl β-D-1-thiogalactopyranoside

OD_{600} - Optical density to light with a wavelength of 600 nm

BLAST - Basic local alignment tool

C_t - Threshold cycle

dsDNA - Double stranded DNA

AAFC - Agriculture Agri-Food Canada

Chapter 1

General Literature Review

Erwinia amylovora (Burrill) Winslow et al. is a pathogen that can infect a wide range of species in the *Rosaceae* including *Pyrus* spp., *Malus* spp., *Rubus* spp., *Sorbus* spp., and *Cydonia* spp. (Bonn and van der Zwet, 2000; Johnson and Stockwell, 1998; van der Zwet and Keil, 1979). The infected tissues generally appear black and collapsed, as if they were scorched by fire (Bonn and van der Zwet, 2000; Coxe, 1817). The common name for the disease is fire blight and the symptoms are generally described by the specific location of the blight on the tree such as blossom, shoot, fruit, leaf, trunk and root (Thomson et al., 1992; van der Zwet and Beer, 1995). .

Fire blight was first observed in New York State in the Hudson River highlands in 1780 (Stapp, 1961). Despite the early discovery of the disease symptoms in the US, the first observation of fire blight in Canada took place in British Columbia in 1865 (Bonn and van der Zwet, 2000), and in England in the 1950s (Crosse et al., 1960). In the next 20 years, the pathogen was reported in Belgium, Denmark, Poland and Netherlands (Borecki and Lyskanowska, 1968; Hockenhull, 1979; Kleijburg, 1979; Meijneke, 1972; van der Zwet and Sugar, 1994; Verhoyen, 1983).

E. amylovora is a Gram negative, motile, rod shaped bacterium that belongs to the family *Enterobacteriaceae.* It is closely related to important human and animal pathogens such as *Escherichia coli* and *Salmonella enterica* (Holt et al., 1994). *E. amylovora* strains are relatively homogeneous in their biochemical and protein electrophoretic characteristics despite their different

geographical origins (Paulin, 2000; Vantomme et al., 1982). To date, there are two published complete genome sequences for *E. amylovora:* the European strain CFBP 1430, accession number of FN434113 (Smits et al., 2010), and the American strain Ea273, accession number of FN666575 (Sebaihia et al., 2010). *E. amylovora* CFBP 1430 strain was isolated in 1972 from a *Crataegus* sp., and Ea273 was isolated from a *Malus* sp. Even though these strains are from two different parts of the world, they have similar genome size of 3.8 Mb and show 99.99% sequence similarity (Sebaihia et al., 2010; Smits et al., 2010). The high similarity between the genomes of these strains indicates minimal evolutionary changes of both strains (Jock et al., 2002; Smits et al., 2010; Zhang and Geider, 1997).

During cell growth *E. amylovora* produces exopolysaccharides (EPS) that are the main components of the bacterial capsule which is loosely attached to the bacterial surface and/or secreted into its surroundings (Bellemann et al., 1994; Cerning, 1990; Durlu-Ozkaya et al., 2007; Kim et al., 2002). The production of the highly hydrated EPS is advantageous for growth and survival of the bacterium because it provides protection against environmental stress and avoidance of recognition by plant defense mechanisms (Leigh and Coplin, 1992). EPS plays a major role in *E. amylovora* pathogenicity and is the main component of the bacterial ooze. Ooze is a yellow to orange-colored substance that is present on the surface of infected host tissues (Bellemann and Geider, 1992; Bennett, 1980; Bennett and Billing, 1978; Koczan et al., 2011; Koczan et al., 2009; Ordax et al., 2010; Vrancken et al., 2013). In addition, EPS production may enhance the ability of the bacteria to resist antibiotic treatments since it

acts as a protection shield (Gander and Gilbert, 1997; Palmer et al., 2007; Ramey et al., 2004).

   E. amylovora EPS is composed of levan (Du and Geider, 2002; Geider et al., 1993; Gross et al., 1992; Vrancken et al., 2013) and amylovoran (Bellemann et al., 1994; Bereswill and Geider, 1997). Koczan et al. (2011) showed that both amylovoran and levan are involved in the formation of the bacterial biofilm (Koczan et al., 2011; Ramey et al., 2004). Levan (Figure 1) is a polyfructan synthesized in the presence of sucrose by levansucrase (Geider et al., 1993; Geier and Geider, 1993), an enzyme encoded by *lsc* gene (Rairakhwada et al., 2010). Levansucrase cleaves sucrose into glucose and fructose, and then polymerizes fructose to levan (Du and Geider, 2002; Geier and Geider, 1993; Gross et al., 1992; Zhang and Geider, 1997). Amylovoran, the most abundant component of the EPS, is a polymer, composed of pentasaccharide repeating units that are generally made of four galactose molecules and one glucuronic acid (Figure 1) (Jumel et al., 1997; Maes et al., 2001; Nimtz et al., 1996). The backbone of the repeating unit of amylovoran consists of three differently linked galactose residues. The major side chain consists of glucuronic acid and a terminal galactose decorated with pyruvyl and acetyl groups. A second side chain is formed when glucose is attached to the branched galactose residue in 10 to 50% of the repeating units (Bellemann et al., 1994). Amylovoran is synthesized by the *ams* operon, a 16 kb region (Bugert and Geider, 1995; Eastgate, 2000; Langlotz et al., 2011). It consists of 12 genes (*ams*A to *ams*L) organized in a large gene cluster and all transcribed in the same direction. The products of ams genes seem be involved in individual steps of amylovoran synthesis (Bugert et al., 1996; Bugert and Geider, 1995).
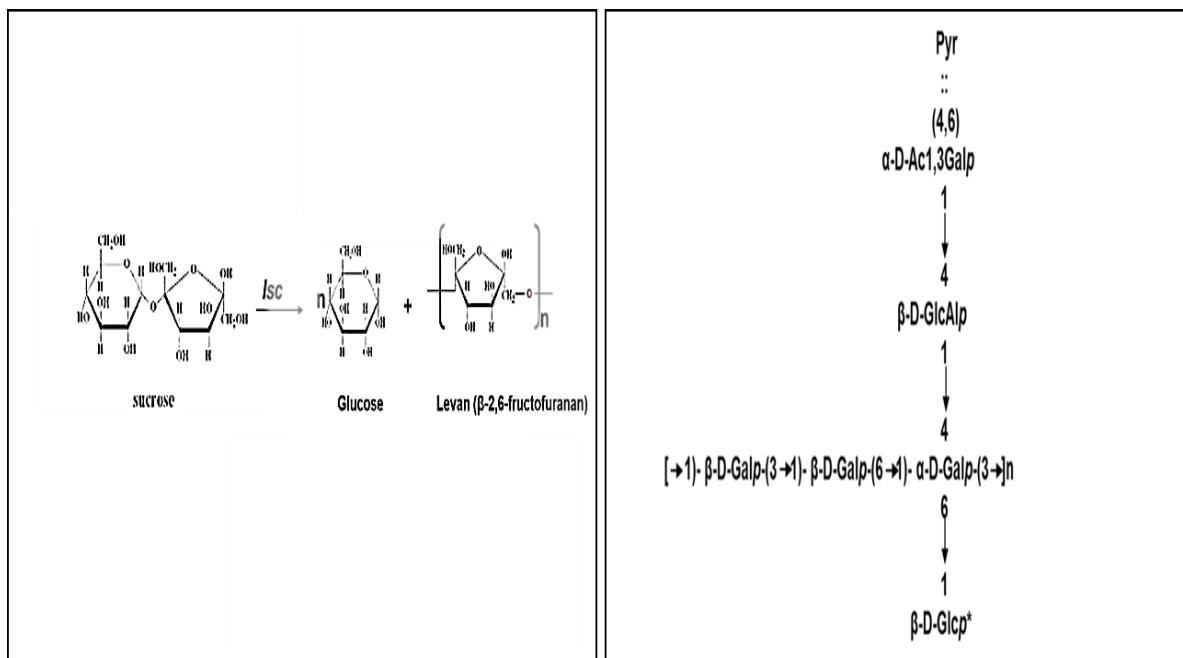
Figure 1: *E. amylovora* levan synthesis and amylovoran structure.

Left: levan synthesis in *E. amylovora*. Sucrose is metabolized by levansucrase releasing glucose and fructose. Fructose polymerizes to form levan. **Right**: amylovoran repeating unit structure. Letter code of the monosaccharide residues is shown: Gal, galactose; GlcA, glucuronic acid; Pyr, pyruval residue; p: pyranoside; n, degree of polymerization. (*) An additional glucose residue found in only 10% of the repeating units. Graphics by (Nimtz et al., 1996).

The sugar residues of the repeating unit are presumably attached sequentially before the unit is transported through two cell membranes and attached to an existing amylovoran molecule (Geider et al., 1996). The role of some individual ams genes in EPS synthesis was determined by characterization of mutants obtained by transposon mutagenesis or of a resistance of mutants obtained by transposon mutagenesis or of a resistance cassette insertion (Aldridge et al., 1998; Bernhard et al., 1990; Bugert et al., 1996; Bugert and Geider, 1995). EPS synthesis of *E. amylovora* is regulated by the activator proteins RcsA and RcsB and by RcsC as a sensor in a two-component system (TCR) (Bereswill and Geider, 1997; Bernhard et al., 1990; Coleman et al., 1990; Poetter and Coplin, 1991). TCR plays critical roles in sensing and responding to environmental conditions and in bacterial pathogenesis. In this system, RcsC acts as a sensor to transfer environmental signals by attaching a phosphate group to RcsB an activator protein. RcsA, in turn, binds to RcsB, converting it to an efficient activator (Bereswill and Geider, 1997; Bernhard et al., 1990; Coleman et al., 1990; Poetter and Coplin, 1991).

Amylovoran is a major pathogenicity factor (Mann et al., 2013; Roach et al., 2013) and the main constituent of the bacterial EPS (Ayers et al., 1979; Billing, 1960; Cerning, 1990; Geier and Geider, 1993; Koczan et al., 2011; Koczan et al., 2009; Ramey et al., 2004). The pathogenicity associated with amylovoran could be due to its role in bypassing the plant defence system and in disturbing and obstructing the vascular system of the plant (Bernhard et al., 1996; Denny, 1995; Koczan et al., 2011; Ordax et al., 2010; Vrancken et al., 2013). *E. amylovora* strains that do not produce amylovoran are non-pathogenic and are unable to spread in plant vessels (Bellemann and Geider, 1992; Geider

et al., 1993; Lee et al., 2010; Oh and Beer, 2005; Roach et al., 2011; Roach et al., 2013; Sjaarda, 2012; Steinberger and Beer, 1988; Vrancken et al., 2013). A mutation in ams gene cluster that involved in the biosynthesis of amylovoran resulted in the lack of bacterial pathogenicity (Geider et al., 1993; Roach et al., 2011; Roach et al., 2013). Moreover the magnitude of amylovoran produced is associated with the degree of *E. amylovora* virulence (Maes et al., 2001; Vrancken et al., 2013). *E. amylovora* isolates that lack levan synthesis exhibit delayed symptom development in the host (Geider et al., 1993; Koczan et al., 2011; Sjaarda, 2012; Vrancken et al., 2013). Levan is also considered a virulence factor in *E. amylovora*, since levan deficient mutants show a decrease in fire blight symptoms (Geider et al., 1993; Koczan and Sundin, 2011; Piqué et al., 2015). Two additional factors are directly involved in the pathogenicity of *E. amylovora.* These include the hrp genes involved in hypersensitivity response in non-host plants and dsp genes, disease specific genes involved in disease development (Oh et al., 2005; Piqué et al., 2015; Steinberger and Beer, 1988; Zhao et al., 2009).

Disease cycle

E. amylovora from the infection events in the previous season survive and overwinter in bacterial cankers which are situated on the woody portions of the plant tissues. In early spring, under optimal temperatures of 24-28°C, *E. amylovora* multiply rapidly and the margins of the cankers discharge ooze. Ooze is a mixture of *E. amylovora* cells and EPS (Biggs, 1994; van der Zwet and Beer, 1995). The bacterial cells in the ooze, serve as the primary inoculum source. The ooze is spread to open blossoms via insects, birds, rain and wind

(Miller, 1929; Thomson, 2000; van der Zwet and Beer, 1995; Van Laere et al., 1980). Primary infection of the host occurs during open bloom when the bacterial cells reach the nutrient rich stigmal surface. Spring conditions at temperatures of 20-28°C and in the presence of nectar result in the multiplication of the pathogen (Johnson and Stockwell, 1998; Thomson, 2000; Vanneste, 2000). The infected blossom looks water-soaked then discoloured (van der Zwet and Keil, 1979). In the presence of high humidity and/or free water the bacterial cells washes down into the hypanthium. The bacterial cells continue to multiply and achieve high population numbers within the nectarthode (Johnson and Stockwell, 1998; Thomson, 2000). The bacteria continue to multiply when invading the peduncle and the spur reaching the stem. EPS produced by the pathogen is accumulated leading to disruption of water flow in the xylem leading to the formation of bacterial aggregates (Sjulin and Beer, 1978). These aggregates cause the xylem vessels to leak forcing the EPS to appear on the plant surface forming ooze, a characteristic of fire blight.

Bacteria can spread from blossom to blossom by pollinating insects initiating more blossom blight (Johnson et al., 1993; Thomson et al., 1998; Thomson et al., 1992). As long as favorable temperatures persist with high humidity, *E. amylovora* secondary infections may continue throughout all spring and early summer (Bonn and van der Zwet, 2000; Thomson, 2000). The secondary infection includes shoot, fruit, and rootstock blight. Shoot blight is a result of the spread of inoculum from infected blossoms to the healthy shoot tissue. Injuries caused by winds and hail can create wounds in host plant tissues creating an entrance for the pathogen, reducing their natural defense mechanisms and facilitating the spread of the infection (Steiner, 2000).

Rootstock blight occurs as a result of the movement of *E. amylovora* systemically from infected shoots and cankers into the rootstock (Roberts et al., 1998; Steiner, 2000; van der Zwet and Beer, 1995).

Fire blight control

Good agricultures practises such as proper pruning methods, suitable irrigation timing and type and timing to apply pesticides are good practises to prevent and manage the disease. To prevent the spread of the disease, previously infected tissue such as branches, cankers, or even whole trees should be removed during winter dormancy.

The difficulty to eradicate *E. amylovora* once it exists in an orchard and the complexity of the disease cycle makes controlling fire blight with a single method very difficult. In addition to good agricultural practises to prevent and manage the disease, there are two main approaches for the control *E. amylovora* in the orchard. In chemical control, bactericides such as antibiotics and copper are used to reduce *E. amylovora* inoculum and prevent the proliferation of the bacterium. In biological control, antagonistic non-pathogenic bacteria compete with the pathogen for nutrients on the stigma surface. The *E. amylovora* populations are suppressed and infection of the flower is avoided (Johnson and Stockwell, 2000; Stockwell et al., 2011).

Despite the relatively low cost, the use of copper as a pesticide comes with drawbacks. The efficacy of copper in preventing fire blight disease is still controversial due to the lack of conclusive results from efficacy studies (Dimova Aziz, 1990; El Nasr et al., 1990; Tsiantos and Psallidas, 1993; Tsiantos et al.,

2003). Copper is bactericidal to *E. amylovora* but it may be phytotoxic to vegetative and reproductive plant tissues. Most importantly, the appearance of copper-resistant plant pathogens has been documented in bacteria (Behlau et al., 2011; Cooksey, 1994).

Streptomycin and kasugamycin are registered antibiotics for fire blight control in Canada (Bonn and Leuty, 1993; Lightner and Steiner, 1993; McGhee and Sundin, 2011; Norelli et al., 2003; Russo et al., 2008; Schnabel and Jones, 1999). While they have a high efficacy for the control *E. amylovora* and other plant pathogenic bacteria, the use of antibiotics as bactericides comes with serious problems. The appearance of antibiotic resistance in *E. amylovora* and the regulatory movement away from antibiotics in certain jurisdictions has resulted in the re-evaluation of control options for fire blight (Breth et al., 2014; Loper et al., 1991; McManus and Jones, 1994; Russo et al., 2008; Schroth et al., 1979; Sholberg et al., 2001; Stockwell, 2014; Walsh et al., 2014). Streptomycin-resistant *E. amylovora* isolates have been discovered in British Columbia (Sholberg et al., 2001), California (Förster et al., 2015), Michigan (Coyier and Covey, 1975), Utah (Nischwitz and Dhiman, 2013) New York state (Russo et al., 2008; Tancos et al., 2016) and New Zealand (Thomson et al., 1993; Vanneste and Voyle, 1998).

Biopesticides are generally composed of antagonistic bacteria that suppress *E. amylovora* populations on susceptible blossoms by competing for nutrients and space (Johnson and Stockwell, 1998; Kabaluk, 2010; Pusey, 2002; Rezzonico et al., 2009; Wilson and Lindow, 1993). This competition reduces the ability of *E. amylovora* to cause disease (Wilson and Lindow, 1993). Biopesticides that have been licenced for use in Canada for the control of fire

blight include Bloomtime™, BlightBan® A506, BlightBan® C9-1 (Nufarm Agricultural Inc.) and Blossom Protect™ (Bio-Ferm). Bloomtime contains the epiphytic *Pantoea agglomerans* strain E325, BlightBan® A506 contains *Pseudomonas fluorescens* strain A506, and the active ingredient of BlightBan® C9-1 is *Pantoea vagans* C9-1 and Blossom protect™ contains *Aureobasidium pullulans*, an epiphytic yeast.

The ability of antagonistic bacteria to supress *E. amylovora* mainly depends upon their ability to colonize open blossoms. Timing of application and quantity applied are critical for successful colonization and high efficacy (Ishimaru et al., 1988; Pusey, 2002). Biologicals need to be applied during the bloom period usually at 30-50 and 70-100% open bloom. The populations of the biological need to be established on the pistil of every opened flower. In this manner the pathogen is excluded from the same environmental niche since it competes for the same nutrients as the biological control agent.

Bacteriophages and their use as biological control agents

Bacteriophages, or phages, were first discovered by Twort in 1915 and d'Herelle in 1917 (Summers, 2005). They are viruses that multiply and replicate inside the bacterial host by manipulating the host's biosynthetic machinery. Phage particles are generally composed of three major components that include the genome, capsid and tail. The phage genome can be linear or circular and either double- or single-stranded DNA or RNA. The viral genome is enclosed in a protective coating or capsid composed of phage-encoded proteins.

E. amylovora phages are icosahedral, tailed double-stranded DNA viruses belong to the *Caudovirale* order (Born et al., 2011; Gill et al., 2003; Lehman, 2007; Lehman et al., 2009; Müller et al., 2011a; Schnabel and Jones, 2001; Weinbauer, 2004; Weinbauer et al., 2007; Yagubi et al., 2014). *Erwinia* spp. phages can be classified into three families: *Myoviridae, Siphoviridae*, and *Podoviridae* (Figure 2) (Gill et al., 2003). The *Myoviridae* phages have contractile tails; *Siphoviridae* phages, which constitute about 61% of tailed phages (Ackermann, 2007; Ackermann and Prangishvili, 2012), have long non-contractile tails and the *Podoviridae* have short tails. *Erwinia* spp. phages genome sizes vary and could range from 45 kbp up to 280 kbp (Born et al., 2011; Lehman et al., 2009; Müller et al., 2011a; Yagubi et al., 2014). The size of the viral capsid can be different from phage to phage and it depends of genome size. The function of phage tail is to correctly identify a host and deliver the phage genome into the host cell interior.
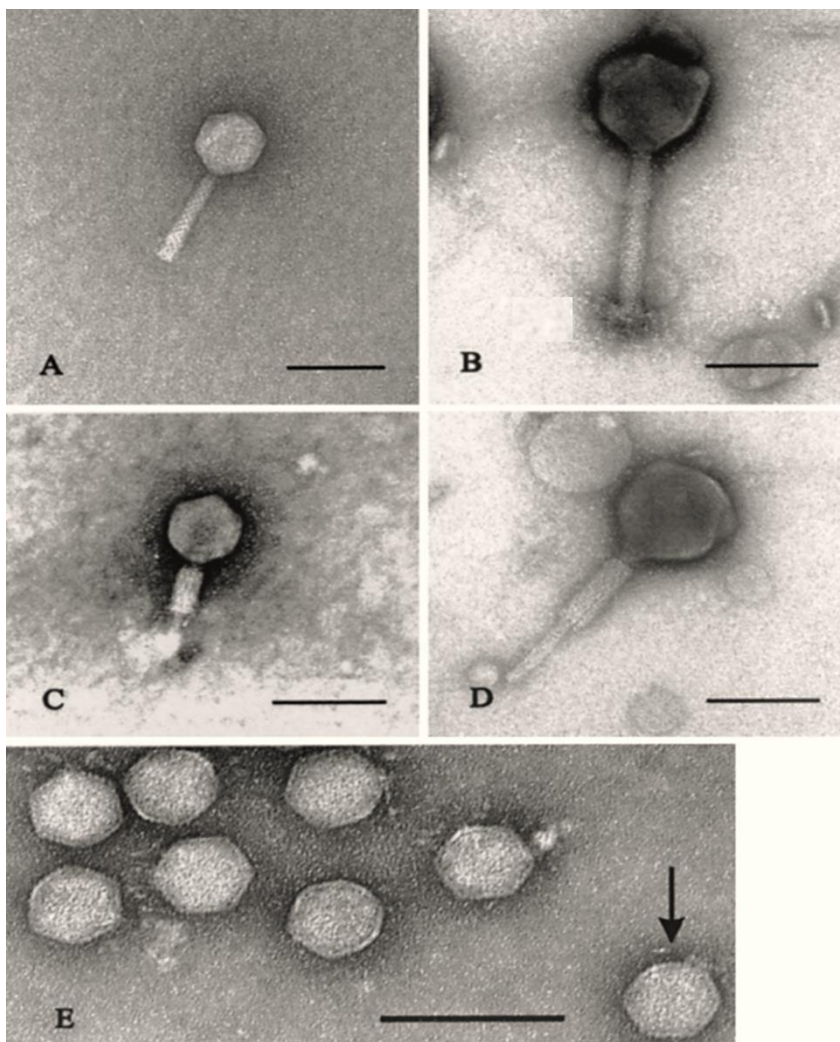
(Gill et al., 2003)

Figure 2 : Electron micrographs of *E. amylovora* phages.

A-D phages belong to the *Myoviridae* family; A and B in uncontracted state; C and D in the contracted state. E phages from the *Podoviridae* family. The arrow points to the tail. Micron marker in each frame represents 100 nm.

Bacteriophage life cycle

Phages can exhibit lytic or temperate life cycles. Lytic phages infect, replicate, and release progeny; this life cycle eventually leads to bacterial cell death. Temperate phages, on the other hand, are phages capable of undergoing both a lytic and lysogenic cycle (Ptashne, 2004; St-Pierre and Endy, 2008; Zeng et al., 2010). In the lysogenic cycle, phage genome integrates into the host chromosome and exists as prophage (Campbell, 2006; Campbell, 2007). This dormant integration state may persist indefinitely without any production of new virions or it may convert to a lytic infection under specific environmental stresses (Campbell, 1961; Stewart and Levin, 1984).

Phage lytic cycle (Figure 3) may be divided into five timely coordinated stages: attachment and genome insertion, synthesis of early proteins, phage DNA replication, synthesis of late proteins, lysis of the host cell and the release of mature viral particles (Abedon, 2008; Lindberg, 1977). Initially, the phage binds to a receptor on the bacterial cell surface. The phage genome is inserted into the bacterial host, and it must be protected from the bacterial nucleases. Once the viral genome enters the host cell, the phage life cycle may undertake either lysogenic or lytic cycle (Figure 3) (Abedon, 2008; Koch, 2007). During the lytic cycle, some viral genes are expressed as mRNAs then translated, using host cell ribosomes and tRNAs to produce early viral proteins. Early proteins mainly allow the phage to control the host biosynthesis machinery and support viral DNA replication. The newly-replicated viral-DNA molecules are used to synthesis late viral proteins. The late proteins are mostly structural proteins that make up the capsid and tail. Following DNA replication and the synthesis of late
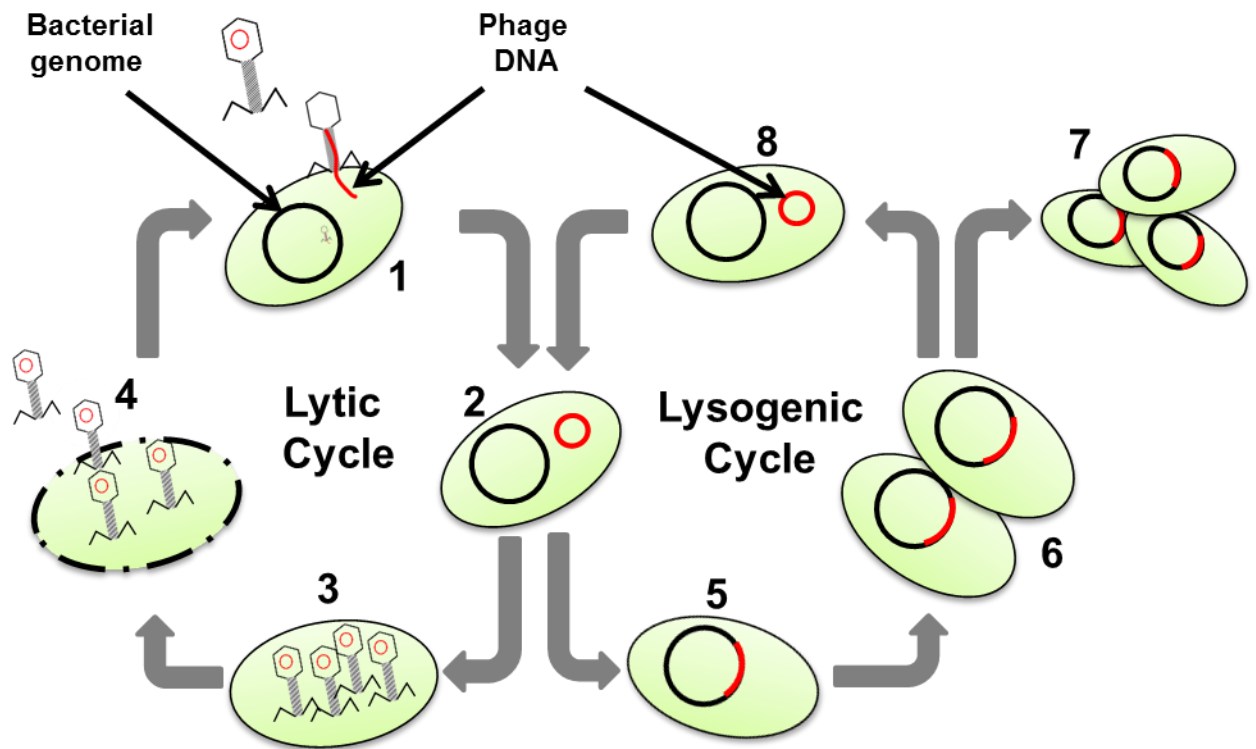
Figure 3: Bacteriophage lytic and lysogenic life cycles.

Phage attaches to a cell (1) and inserts its DNA. (2) Phage DNA circularizes; at this stage, phage can initiate either a lytic cycle or lysogenic cycle. In lytic cycle: (3) Phage DNA replicates and phage protein synthesis followed by phage particles assembly; (4) Cell lysis and release of new phages. In lysogenic cycle: (5) Phage DNA integration into host chromosome (prophage, in red); (6) Phage DNA replicates when host cell divides; (7) daughter cell with prophage. (8) At this stage, a prophage may exit the bacterial chromosome, initiating a lytic cycle. Graphics by A. Yagubi.

proteins, phage particle assembly takes place. Once the assembly is accomplished, the host cell is lysed allowing the release of mature phages. The newly synthesized phages in turn infect nearby host cells and start a new infection cycle (Weinbauer, 2004).

In the lysogenic cycle (Figure 3), temperate phages keep their chromosome stable and silent by integrating their genome into the bacteria host chromosome (Campbell, 2006; Campbell, 2007). The phage genome can silently replicate along with bacterial chromosome during normal cell division (Campbell, 2006; Campbell, 1961; Stewart and Levin, 1984; Zong et al., 2010). Phage DNA integrated into host cells during lysogeny is called prophage. Bacterial cells carrying the prophage are referred to as the lysogenized host (Campbell, 1961; Stewart and Levin, 1984). In the lysogenic state, genes that involved in the lytic cycle are repressed (Oppenheim and Adhya, 2007). If the repressed genes are activated, phage genome can be released from the bacterial chromosome (Dodd et al., 2004; Ptashne, 2004). The excised phage genome then acts as a template for the synthesis of viral proteins involved in the production and assembly of new viral particles. The end point of the lytic cycle results in the lysis of the bacterial host and the release of new phage.

Prophages have been detected in many phytopathogens and plant-associated bacteria. These includes: *Rhizobium* spp*.* (Malek, 1990; Uchiumi et al., 1998), *rhizosphere pseudomonads* (Shaburova et al., 2000), *Ralstonia solanacearum* (Yamada (Addy, 2012a; Addy, 2012b; Murugaiyan et al., 2011; Yamada et al., 2007), *Erwinia carotovora* subsp*. carotovora* (Faltus and Kishko, 1980; Kishko et al., 1982; Ruban et al., 1981; Toth, 1993; Tovkach, 2002a; Tovkach, 2002b; Tovkach, 2002c), *Xanthomonas campestris pv. azadirachtae*

(Borkar, 1997) and *P. agglomerans* (Erskine, 1973; Harrison and Gibbins, 1975). The presence of prophage(s) in bacterial hosts may prevents the infection of bacteria by homologous phage through superinfection immunity (Abedon, 2008). Temperate phages may involve in phage-mediated transfer of genes that can improve bacterial virulence (Addy, 2012a; Brüssow, 2009; Brüssow et al., 2004; Canchaya et al., 2004; Wagner et al., 2002) or decrease virulence (Addy, 2012b). Temperate phages may also contribute to toxin genes and antibiotic resistance genes transfer (Mauro and Koudelka, 2011; Muniesa et al., 2013). Many phages are known to carry lysogenic genes that are involved in establishing lysogeny (Brüssow et al., 2004). The genomes of all fully sequenced *Erwinia* spp*.* phages in this study and others were found to lack the necessary genes for prophage integration (Born et al., 2011; Lehman et al., 2009; Nagy et al., 2012; Yagubi et al., 2014).

Roach et al. (2015) showed that while lysogeny is possible in *E. amylovora*, it is rare and/or absent in natural populations. Real-time PCR primers were developed that detected the presence of *E. amylovora* phages belonging to the *Myoviridae* and *Podoviridae* families in a collection of bacterial isolates belonging to global *E. amylovora* and local isolates of *P. agglomerans*. Prophage were not detected in 161 *E. amylovora* isolates from different geographical areas of the world and 82 *P. agglomerans* isolates from southern Ontario (Roach et al., 2015). The possibility still remains that temperate phage may exist that belong to the *Siphoviridae* and/or unique phage isolates that are not detected by current real-time PCR primers (Roach et al., 2015).

Phages as biological control agents in agriculture

In 1924, phages were proposed as agents for the control of plant disease (Moore, 1926). Phage filtrate collected from decomposed cabbage infected with *Xanthomonas campestris* pv. *campestris* inhibited *in vitr*o growth of the phytopathogen (Mallmann and Hemstreet, 1924) . In 1930s, phages were used to reduce the severity of bacterial blight in cotton (Massey, 1934). Phages were used for effective control of *Pantoea stewartii*, the causal agent of Stewart's wilt disease (Thomas, 1935). The discovery of antibiotics in the 1940s and their resulting high efficacy against bacterial pathogens resulted in the abandonment of bacteriophage based research (Chanishvili, 2012; Summers, 2001).

The appearance of antibiotic resistance in *E. amylovora* (Coyier and Covey, 1975; Miller and Schroth, 1972; Russo et al., 2008; Sholberg et al., 2001), the demand for environmentally friendly pesticides (Dunne et al., 1996; Emmert and Handelsman, 1999; Manso et al., 2010) and the legislated removal of antibiotics in certain jurisdictions contributed to increasing interest in finding alternative solutions for control of plant pathogens. The use of bacteriophages as potential replacements for antibiotics for the control of fire blight has been recently investigated by a number of research groups in Europe and North America (Erskine, 1973; Gill et al., 2003; Lehman, 2007; Schnabel et al., 1999; Svircev et al., 2010). Phages have been used for the control of: bacterial wilt caused by *Ralstonia solanacearum* (Tanaka et al., 1990), bacterial spot of tomato (X*anthomonas* sp.) (Balogh et al., 2003; Flaherty et al., 2000), bacterial blotch of mushrooms *(Pseudomonas* spp.) (Munsch et al., 1995), potato scab (*Streptomyces scabies)* (Mckenna et al., 2001), blight on onion (*Xanthomonas* sp.) (Lang et al., 2007), and citrus bacterial canker (Balogh, 2006)

Challenges in using phages as biological control agents

Phages are highly sensitive to harsh environmental factors when applied to aerial plant tissues (Balogh, 2002; Balogh et al., 2003; Civerolo and Kiel, 1969; Frampton et al., 2012; Iriarte et al., 2007; McNeil et al., 2001). Field and laboratory studies have demonstrated that phages can be degraded and inactivated by exposure to high temperature, extreme pH, and UV light (Ignoffo et al., 1989; Ignoffo and Garcia, 1992; Iriarte et al., 2007). Phage survival in the environment has to be ensured by incorporating protective materials such as charcoal, starch (Balogh, 2002; Balogh et al., 2003; Ignoffo et al., 1997; Obradovic et al., 2004) and/or the use of a protective non-pathogenic bacterial host (Erskine, 1973; Lehman, 2007; Svircev et al., 2010). Field trials with phages failed due to the declined of the phage populations in the blossoms due to sun light (Ritchie and Klos, 1979; Schnabel et al., 1999). Some research groups have proposed the use of avirulent *E. amylovora* mutants to improve phage persistence in the environment (Schnabel et al., 1999; Tharaud et al., 1997). The possible revision of mutants to virulent bacteria prompted the requirement for use a safer alternative. Our research group developed a successful and novel strategy for phage delivery by using an epiphytic bacterium. The epiphyte delivered and maintained a replicating phage population in the blossom during the primary stages of *E. amylovora* infection (Lehman, 2007; Svircev et al., 2010). The epiphytic and non-pathogenic *Pantoea agglomerans* (formerly *Erwinia herbicola*) and the *Erwinia* spp. phages were applied onto the blossoms (Lehman, 2007). *P. agglomerans* Pa21-5 controlled *E. amylovora* in open blossoms by competing with the pathogen for nutrients, protecting phages from the environmental factors and supporting

phage population replication (Lehman, 2007). The dual action biological system uses *P. agglomerans* as a biological control agent and phage carrier (Lehman, 2007; Svircev et al., 2010).

The ability of bacteria to develop resistance to bacteriophages would deter the acceptance of phage therapy for the control of plant pathogens (Marks and Sharp, 2000; Okabe and Goto, 1963; Vidaver, 1976). To avoid the development of bacterial resistance phage cocktails or mixtures of 2-6 phages are commonly used. Generally, the mixture is prepared with individual phages that have dissimilar receptors to the pathogenic bacterium (Abuladze et al., 2008; Balogh et al., 2010; Flaherty et al., 2001; Jackson, 1989). To avoid phage resistance, Jackson (1989) developed a new strategy by preparing mixtures of wild-type and host range mutant phages (h-mutants). The h-mutants were capable of lysing bacterial strains resistant to the parent wild-type phages (Jackson, 1989). The phage mixtures retain a host range wider than the parent phage as such the occurrence of phage-resistant host can be to prevented (Jackson, 1989). The strategy of using h-mutants was evaluated in the control *Xanthomonas campestris* pv. *vesicatoria,* bacterial spot disease on tomato, and *Xanthomonas campestris* pv. *pelargonii,* bacterial blight of geranium (Flaherty et al., 2001; Flaherty et al., 2000). The application of these phage mixtures provided an equivalent disease control of bacterial blight of geranium and bacterial spot disease on tomato comparing to plants treated with the chemical bactericide (Flaherty et al., 2001; Flaherty et al., 2000).

Bacterial host resistance to bacteriophages

The development of bacterial resistance to phages may involve a number of mechanisms such as: the blocking of phage attachment to bacterial cell wall receptors (Nordström and Forsgren, 1974), abortive infection (Allison and Klaenhammer, 1998; Durmaz and Klaenhammer, 2007; Fineran et al., 2009; Smith et al., 1969), CRISPR/Cas systems (Marraffini and Sontheimer, 2010a; Sorek et al., 2008; van der Oost et al., 2009), physical exclusion by the production of extracellular matrix (Andersson and Banfield, 2008; Bernheimer and Tiraby, 1976; Cerning, 1990; Deveau et al., 2002; Looijesteijn et al., 2001; Scholl et al., 2005; Valyasevi et al., 1990; Wilkinson and Holmes, 1979), and restriction-modification of invading phage nucleic acid (Bujnicki, 2004; Pingoud et al., 2005).

Blocking receptors and prevention of phage adsorption

Phage attachment and adsorption to bacterial cell receptors constitutes the initial step in infection. In order to successfully infect a cell, phages need to recognise and attach to a receptor on the bacterial cell. In turn, bacteria have evolved a wide range of mechanisms to prevent phage adsorption (Allison and Klaenhammer, 1998; Bohannan and Lenski, 2000; Labrie et al., 2010; Lenski, 1988). In *E. amylovora*, the most common mechanisms involve mutation in the phage receptors and the production of an extracellular matrix (Born et al., 2014; Hyman and Abedon, 2010; Roach et al., 2011; Roach et al., 2013).

EPS and phage infection

In addition to protecting the bacterial cell against the harsh environment, the bacterial EPS may act as a protective shield for the cells by providing a physical barrier between phage and the receptors on the cell membrane (Andersson and Banfield, 2008; Bernheimer and Tiraby, 1976; Cerning, 1990; Deveau et al., 2002; Looijesteijn et al., 2001; Scholl et al., 2005; Valyasevi et al., 1990; Wilkinson and Holmes, 1979). Bacteria produce different types of EPS depending on the genus. In *Pseudomonas* spp. and *Azotobacter* spp., the EPS is composed of alginate, an anionic copolymer of mannuronic acid and glucuronic acid (Remminghorst and Rehm, 2006; Sutherland, 1995). Hyaluronan (hyaluronic acid) is a high-molecular-mass polysaccharide produced by pathogenic *Streptococcus pyogenes* as a part of their capsule structure (Van de Rijn, 1983). *P. agglomerans* produces a polysaccharide that is composed of arabinose, glucose, galactose and glucuronic acid (Wang et al., 2007).

E. amylovora isolates in the AAFC Vineland Bacterial Collection have been found to produce EPS in varying amounts when grown under similar condition (Roach et al., 2011; Roach et al., 2013). The level of EPS production can be used to roughly classify them as being high-exopolysaccharide-producing (HEP) and being low-exopolysaccharide-producing (LEP) (Roach et al., 2011; Roach et al., 2013).

To successfully infect a bacterial host, phages have to pass the EPS structure of the bacterial host. Many of these phages such as phages belong to *Podoviridae* family have evolved to overcome the EPS by producing virion-associated proteins with EPS-depolymerization activities (Born et al., 2014; Hughes et al., 1998a; Pelkonen et al., 1992). The depolymerases degrade the

EPS structure allowing the phage to attach to the bacterial cell membrane (Hughes et al., 1998a). The role of EPS in phage infection is not fully understood. EPS can act as a barrier which blocks phage infection and/or as a receptor that enhances phage adsorption to the cell. *E. amylovora* ability to produce an EPS layer may be directly involved in phage host interactions (Roach et al., 2011; Roach et al., 2013; Roach, 2011; Sjaarda, 2012). *Podoviridae* phages have an absolute requirement for amylovoran. *Podoviridae* phages cannot infect amylovoran-deficient bacterial mutants (Roach et al., 2013). In contrast, the *Myoviridae* phages do not require the presence of amylovoran on the cell surface. However, *Myoviridae* phages could not infect levan-deficient *E. amylovora* mutants (Sjaarda, 2012).

CRISPR/Cas system

The CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats)/Cas (CRISPR-associated) system was first described in the genome of *E. coli* (Ishino et al., 1987). The term CRISPR was proposed by Jansen et al. (2002) to reflect the specific structure of these sequences (Jansen et al., 2002a). Similar systems were reported in the genomes of other prokaryotes: *Streptococcus* (Hoe et al., 1999), *Methanococcus jannaschii* (Bult et al., 1996), *Archaeoglobus fulgidus* (Klenk et al., 1997), *Thermotoga maritima* (Nelson et al., 1999), and *Pyrococcus horikoshii* (Kawarabayasi et al., 1998). The CRISPR/Cas system has been found in viral genomes and plasmids suggesting that this system can be horizontally transferred between microorganisms (Godde and Bickerton, 2006; Makarova et al., 2006; Seed et al., 2013). This was a novel discovery since bacteria were not anticipated to have immunological memory

that could be invested in the fighting of repetitive viral infection as is the case in vertebrates. The CRISPR/Cas system functions as an immune system for prokaryotes, as it defends prokaryotic cells against invading phages and other exogenous genetic elements such as conjugative plasmids and transposons (Barrangou and Marraffini, 2014; Bolotin et al., 2005; Godde and Bickerton, 2006; Koonin and Makarova, 2009; Mojica et al., 2005; Pourcel et al., 2005; Sorek et al., 2008; Westra et al., 2012a). This system is a highly adaptive and heritable resistance mechanism. In this defensive system some bacterial hosts may acquire resistance against invading phages after survival of a previous exposure to the same or very similar phage (Horvath et al., 2008; Sorek et al., 2008).

CRISPRs are sites containing conserved short direct repeats that typically range in size from 24 to 48 bp (Figure 4) (Grissa et al., 2007). The repeats are separated by sequences termed spacers, which range from 17 to 80 bp and have an average length of 36 nucleotides (Grissa et al., 2007). The spacers are usually unique in a genome and match sequences in phage genomes and/or plasmids (Barrangou et al., 2007; Bolotin et al., 2005; Horvath et al., 2009; Mojica et al., 2009; Mojica et al., 2005; Pourcel et al., 2005). Spacers may be acquired during phage infection or plasmid transformation, and later act as a genetic memory that helps protecting the bacterial cell from subsequent phage infection or plasmid conjugation through recognizing and silencing exogenous genetic elements at the DNA or RNA level (Barrangou et al., 2007; Brouns et al., 2008; Marraffini and Sontheimer, 2008; Marraffini and Sontheimer, 2010a).

The leader sequences and CRISPR-associated (Cas) genes are the additional important components of the CRISPR system (Figure 4) (Haft et al.,

2005; Jansen et al., 2002b; Makarova et al., 2006). Leader sequences are AT-rich regions (approximately 200 bp) located upstream from CRISPR loci (spacers and repeats). This region contains promoters and sequence elements that are essential for spacer acquisition (Barrangou et al., 2007; Díez-Villaseñor et al., 2013; Jansen et al., 2002a; Pourcel et al., 2005; Tyson and Banfield, 2008). *Cas* are conserved group of genes usually located as a cluster in the vicinity of CRISPR loci (Jansen et al., 2002b). Cas genes play a central role in spacer acquisition and silencing of exogenous genetic elements (Barrangou et al., 2007). They are genetically and functionally diverse, exemplifying their different functions and roles in CRISPR/Cas system defensive mechanisms. To date, more than 40 Cas gene families have been identified (Haft et al., 2005; Makarova et al., 2006).

The molecular mechanism of the CRISPR defense system in prokaryotes is not fully understood (Manica and Schleper, 2013; Richter et al., 2013). The proposed mechanism by which the CRISPR/Cas system resists foreign DNA can be summarized in three distinct stages (Figure 5). In the acquisition stage: foreign nucleic acid is recognised and processed by proteins encoded by Cas genes into smaller fragments and inserted between the repeats in the CRISPR locus as spacers. In the subsequent transcription stage: the CRISPR loci are expressed and processed by Cas proteins to small RNAs (crRNAs). Each crRNA is made of an individual spacer with two flanking half repeats. In the final surveillance and interference stage: the crRNAs guide other Cas proteins to recognize invading phages at the DNA or RNA level through detecting specific sequences called protospacers that are complementary to spacer sequences of crRNA. Upon recognition, crRNA hybridized with the protospacers in the target

DNA for degradation by the Cas nucleases or other nucleases (Figure 5) (Haft et al., 2005; Szczepankowska, 2012).
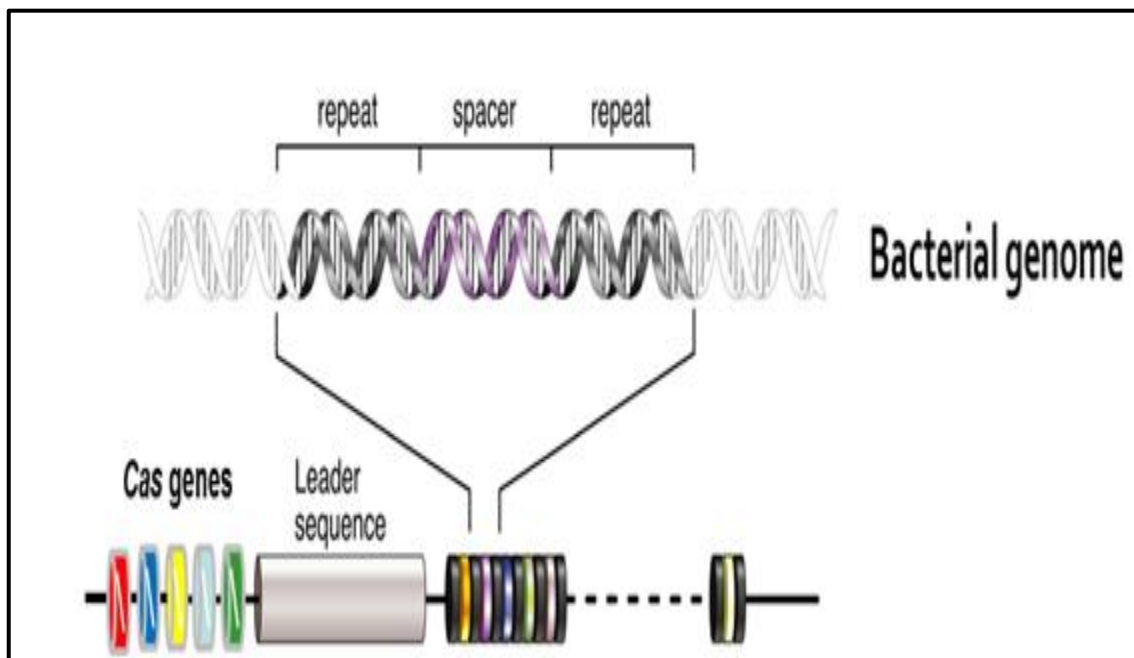
Modified from (Karginov and Hannon, 2010)

Figure 4: CRISPR/Cas system structure in bacteria.

A CRISPR system includes Cas genes and the leader sequence followed by spacers and repeats. The leader sequence is an AT rich region harbouring promoters. Spacer numbers vary depending on bacterial host.
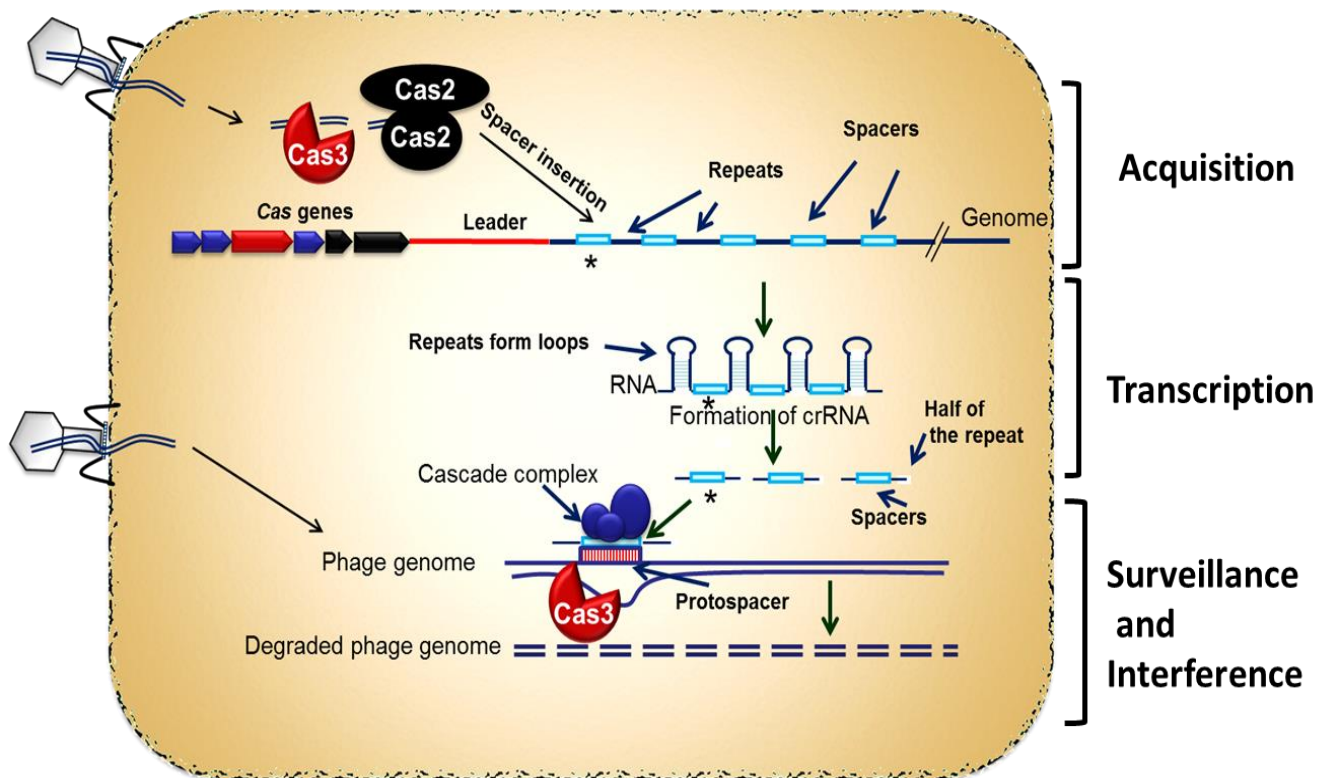
Figure 5: The proposed mechanism of CRISPR/Cas system.

The CRISPR mechanism may be summarized in three stages: **(1) Acquisition:** upon the invasion of foreign DNA, Cas3 digests foreign DNA and inserts it into a CRISPR locus as a spacer with the help of Cas1 and Cas2. **(2) Transcription:** a CRISPR array (spacers and repeats) is transcribed into a single long transcript under the control of a promoter in the leader. Repeats form loops where the large transcript is cleaved into to smaller RNA molecules (crRNA). The crRNA contains one spacer flanked with half of a repeat on each side. **(3) Surveillance and interference:** crRNA guides a Cas protein complex (labelled cascade complex) to recognize invading DNA and digest it with the help of Cas3. Graphics by A. Yagubi.

Study Objectives

The aims of this study are to explore the potential roles of two phage resistance mechanisms in *E. amylovora*, the CRISPR/Cas system and the production of the EPS.

A) In order to investigate the role of CRISPR/Cas system, the following will be carried out:

1) Investigate the existence and structure of CRISPR/Cas system in *E. amylovora* isolates.

2) Study the responses of CRISPR/Cas system to phage infection.

3) Explore the ability of CRISPR/Cas system in prevention of invasion of exogenous elements such as phages and plasmids.

4) Study the ability of CRISPR/Cas system to gain new spacers upon phage infection.

B) To investigate the mechanisms by which *E. amylovora* phages interact with the bacterial EPS, ΦEa35-70 and ΦEa31-3 EPS depolymerase proteins will be characterized.

Chapter 2

Investigating the structure and function of CRISPR/Cas system in *E. amylovora*

Abstract

The CRISPR/Cas system is present in *E. amylovora*, but the functionality and role of this system in phage resistance is still unknown. In this study, the genome of *E. amylovora* isolates Ea6-4, EaD7, Ea110R and E17-1-1 were found to harbour three CRISPR arrays (CRR1, CRR2 and CRR3). A set of Cas genes were determined in the vicinity of CRR1 and CRR2 of EaD7 and Ea6-4; no Cas genes were located near CRR3. Only two CRISPR arrays (CRR2 and CRR3) were detected in the genomes of EaG5, BC29, BC1280 and Ea4-96 isolates. BLAST analysis of spacers detected in all CRISPR arrays in these isolates showed no matches to any existing *Erwinia* spp. phage genomes in GenBank. Monitoring expression level of Cas genes during phage infection revealed that CRISPR/Cas system is not stimulated by phage invasion. The enhancement of *E. amylovora* CRISPR/Cas system by introducing plasmids that carried an artificial set of Cas genes and spacers that target existing *Erwinia* sp. phages did not improve CRISPR/Cas system ability in blocking phages. *E. amylovora* CRISPR/Cas system was tested for its efficiency in blocking plasmid DNA transformation. When CRISPR/Cas system was challenged by transforming *E. amylovora* with plasmids carrying protospacers matching existing spacers, 90% of the plasmids carrying protospacers matching spacers in CRR1 and CRR2 were blocked in comparison to plasmid without protospacers and plasmid carrying protospacers against CRR3. Research data from this study indicate that the CRISPR/Cas system of *E. amylovora* is

not effective and efficient host resistance mechanism. Preliminary data indicate that

the system can partially block plasmid DNA transformation.

Introduction

The CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) /Cas system has been characterized as an adaptive immune system in bacteria; it provides a specific and acquired immunization against repeated phage infection (Barrangou et al., 2007; Jansen et al., 2002b; Rath et al., 2015). Certain bacterial hosts can insert pieces of phage DNA as spacers into a CRISPR array, which later allows these hosts to exhibit immunity against phages that carry homologous sequences to the inserted spacers. The first direct evidence that CRISPR/Cas system could protect bacteria from phages was reported in *Streptococcus thermophilus* (Barrangou et al., 2007). *S. thermophilus* phage resistant mutants were isolated subsequent to phage infection. Phage resistant isolates had incorporated new spacers in the CRISPR system that matched sequences in the phage genome. Elimination of these new spacers restored host sensitivity to phage infection (Barrangou et al., 2007). In addition, CRISPR/Cas systems can prevent plasmid transformation in *Staphylococcus epidermidis* (Marraffini and Sontheimer, 2008). Bacterial CRISPR/Cas systems are highly diverse perhaps reflecting the selective pressures that are enforced by mobile genetic elements. The CRISPR/Cas system significantly contributes to the continuous co-evolution of phages and their hosts (Bondy-Denomy et al., 2013).

Types of CRISPR/Cas systems

Currently, CRISPR/Cas systems have been classified into three major types and more than 12 different subtypes (Makarova et al., 2011). The classification of CRISPR/Cas systems is mainly based on differences in crRNA maturation, interference mechanism and architecture of the CRISPR/Cas loci (Jansen et al., 2002a; Kunin et al.,

2007; Makarova et al., 2011). The key feature for type I, II and III CRISPR/Cas system is the presence of a signature Cas protein such as Cas3, Cas9, or Cas10. In general, six Cas gene families (*Cas*1–*Cas*6) are found in a wide range of CRISPR/Cas types with *Cas*1 and *Cas*2 in particular present in almost all CRISPR types. Cas3 protein is a nuclease/helicase essential in the interference step and in spacer acquisition (Beloglazova et al., 2011; Jackson et al., 2014; Mulepati and Bailey, 2011; Richter et al., 2012b; Sinkunas et al., 2011). Type I and type III systems are slightly related and are found in many bacteria and archaea, however, type II systems are genetically different and found only in bacteria (Fonfara et al., 2014; Haft et al., 2005; Heidrich and Vogel, 2013).

Mature crRNA production and target surveillance and interference

Highly effective CRISPR/Cas systems need to be able to express and process CRISPR array (spacers and repeats) transcripts and use them to assist in the recognition and interference of invaders. CRISPR array transcription initiates at the upstream of the leader sequences at the end of the locus sequence to generate a long single precursor transcript (known as a pre-crRNA) (Figure 6). Transcription is controlled by promoters located in the leader region that possibly serve as binding sites for regulatory proteins (Hale et al., 2012; Lillestøl et al., 2009). The long pre-crRNA is eventually processed into small crRNAs that enable the CRISPR system to recognise and interfere with invaders (Figure 6). This process occurs in dissimilar organisms such as *Archaeoglobus fulgidus* (Tang et al., 2002), *Sulpholobus solfataricus* (Tang et al., 2005), *Escherichia coli* (Brouns et al., 2008), *Pyrococcus furiosus* (Hale et al., 2008), *Sulpholobus acidocaldarius* (Lillestøl et al., 2009) and *Xanthomonas oyrzae* (Semenova

et al., 2009).   In type I and III systems, the pre-crRNA transcript is processed and cleaved by the endoribonucleases Cas6 to generate mature crRNA (Carte et al., 2008; Jore et al., 2011; Sinkunas et al., 2013; Wiedenheft et al., 2011b). The resultant mature crRNA has a complete spacer flanking with a short 5' repeat handle and a repeat sequence that forms a stem-loop at the other end (Haurwitz et al., 2010; Judith et al., 2013b; Nam et al., 2012b; Sashital et al., 2011; Wiedenheft et al., 2012).   Processing of pre-crRNA in the type II system is entirely different from type I and type III. In addition to Cas proteins, CRISPR/Cas system type II encodes a trans-activating crRNA (tracrRNA), a 25 bp stretch that is complementary to the to the repeats within the pre-crRNA (Deltcheva et al., 2011). In addition to the tracrRNA, Cas9 and host RNase III are needed for the processing of pre-crRNA into mature crRNA (Deltcheva et al., 2011; Jinek et al., 2012).

CRISPR surveillance complexes

To facilitate targeting and cleavage of invading nucleic acid, mature crRNAs and some Cas proteins form multi-subunit ribonucleoprotein targeting complexes generally known as crRNA and Cas proteins (crRNP) complexes. These complexes act as a surveillance and interference tool that recognises and binds to invading DNA for elimination. Different types of CRISPR/Cas systems form different crRNP complexes and their structure dictates the terminology used to describe these complexes (Makarova et al., 2011).

In CRISPR/Cas system type I, the targeting and surveillance crRNP complexes are known as Cascade (CRISPR-associated complex for antiviral defence) (Brouns et al., 2008; Judith et al., 2013a; Westra et al., 2012a).   In CRISPR/Cas type I-E of *E. coli*,

the Cascade complex has a size of 405 KDa, and is composed a hexameric backbone

formed by Cas7 bound to Cas5, Cas6e, Cse1 and Cse2 proteins in addition to a single

61-nucleotide crRNA; Cse proteins are also CRISPR-associated proteins found In

CRISPR/Cas system type I. in The 5' end of crRNA in this system, which forms the

stem-loop, is bound by Cas6e and the 3' end is bound by Cse1 (Brouns et al., 2008;

Jore et al., 2011; Wiedenheft et al., 2011b). The Cascade complexes of the rest of type

I systems share the same feature as type I-E; however, they contain different and fewer

Cas proteins (Makarova et al., 2011). In CRISPR system type II, crRNPs are known as

Cas9 complexes which are structurally different from crRNP complex of CRISPR/Cas

system type I (Gasiunas et al., 2012; Jinek et al., 2012; Jinek et al., 2014; Nishimasu et

al., 2014). Cas9 complexes of type II contain only Cas9 associated with a single-guide

RNA (Jinek et al., 2012; Jinek et al., 2014; Nishimasu et al., 2014; Sapranauskas et al.,

2011).

The role of protospacer adjacent motif and seed sequences

In order for CRISPR/Cas systems to interfere with invading DNA, crRNP

complexes have to find and recognize a sequence that is complementary to the

sequence of the scanning crRNA. These sequences in invading DNAs are known as

protospacers. It has been suggested that crRNA forms a loop structure in Cse1 that

facilitates recognition and interaction with a specific short sequence that has been

named protospacer adjacent motif (PAM) (Figure 6) (Jore et al., 2011; Sashital et al.,

2012; Westra et al., 2012a). PAMs are highly conserved sequence motifs of 3 to 6

nucleotides that are immediately located at either side of the protospacer (3' or 5' end)

in the target and help to discriminate between self- and non-self- targeting by the
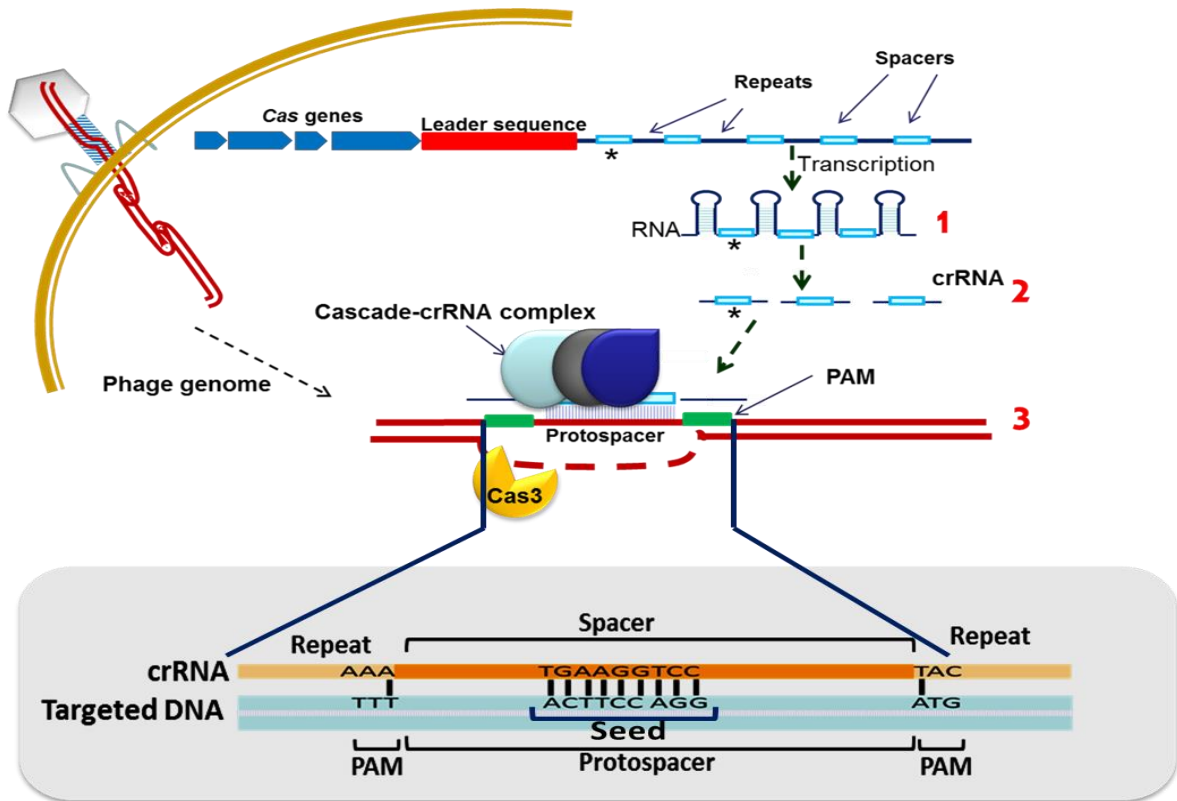
Figure 6: Surveillance and interference by crRNP complexes of type I CRISPR/Cas system.

Steps 1, 2 and 3 show the production of pre-crRNA, crRNA maturation and interference, respectively. In the interference step, Cascade and crRNA form a complex that interferes with invading DNA. The PAM motif promotes the identification of the foreign DNA which is followed by hybridization between the seed sequence of the crRNA and the target DNA. Subsequently, the nuclease Cas3, is recruited and it degrades the target DNA. Graphics by A. Yagubi.

CRISPR system (Deveau et al., 2008; Jore et al., 2011; Mojica et al., 2009; Mulepati et al., 2012; Sashital et al., 2012; Semenova et al., 2009; Shah et al., 2009; van der Ploeg, 2009; Westra et al., 2012a). PAM sequences are replaced by repeats in CRISPR arrays in bacterial genome. The absence of PAM within the CRISPR locus in bacterial genome might explain why the CRISPR locus itself is not recognized by crRNA of CRISPR/Cas system (Marraffini and Sontheimer, 2010b). The recognition of the CRISPR/Cas system to its arrays could lead to the degradation of bacterial genetic material and eventually cell death.

The recognition of PAM motif by crRNP complex through Cse1 seems to denature the invader DNA duplex allowing a secondary alignment between a short sequence termed seed region (first 6 to 10 nucleotides) of the spacer and the complementary sequence in the target protospacer (Sashital et al., 2012; Sorek et al., 2013; Westra et al., 2012a). The seed sequence within the crRNA spacer is critical for binding to target DNA (Semenova et al., 2011; Wiedenheft et al., 2011a; Wiedenheft et al., 2011b). Single nucleotide mutation in the seed sequence results in failure to bind to target DNA. In contrast, mutations in the crRNA spacer sequence outside of the seed sequence do not affect binding efficiency (Semenova et al., 2011). The complementary base pairing of crRNA and target strand leads to the formation of an R‑loop structure within crRNA (Spilman et al., 2013; Wiedenheft et al., 2011a). The formation of the R‑loop structure may also function as a signal to stimulate nucleases such as Cas3 to target the invading DNA (Hochstrasser et al., 2014; Westra et al., 2012b). The helicase and nuclease activity activities of Cas3 facilitate the degradation of the target DNA. After initial cleavage, degradation of the target strand proceeds in the 3′ to 5′ direction (Beloglazova et al., 2011; Sinkunas et al., 2013; Westra et al., 2012b).

In type II systems, the interference tools for the invading DNA are completely different from that of type I and type III systems, and require minimum Cas proteins for immunity. The interference in type II systems is carried out by Cas9–RNP complex which simply consists of Cas9 and a mature crRNA-tracrRNA hybrid (Deltcheva et al., 2011; Garneau et al., 2010; Gasiunas et al., 2012; Jinek et al., 2012).

Selection and acquisition of spacers and protospacer targeting

The first evidence of spacer acquisition from phages was reported in the type II-A system of *S. thermophilus* (Barrangou et al., 2007; Garneau et al., 2010). Spacer acquisition was also reported in CRISPR systems of *E. coli* type I-E (Datsenko et al., 2012; Swarts et al., 2012; Yosef et al., 2012), *P. aeruginosa* type I-F (Cady et al., 2012), *Streptococcus agalactiae* type II-A (Lopez-Sanchez et al., 2012) and *Sulfolobus solfataricus* type I and III-B systems (Erdmann and Garrett, 2012). The details of the spacer acquisition mechanism are only partially understood. Spacer acquisition starts with the recognition and fragmentation of invading DNA. The selection of protospacers, from invading DNA, which will be inserted as spacers in CRISPR arrays, is influenced by the recognition of PAM sequences. The existence of conserved PAM sequences that flank protospacers indicates that protospacers are not randomly selected for insertion in CRISPR loci. The conserved sequences provide a recognition signal for the selection of sequences which will be added as new spacers (Deveau et al., 2008; Mojica et al., 2009; Shah et al., 2013). Upon the selection of the proper sequences for spacer insertion, the CRISPR system processes target DNA sequences into spacer precursors, small fragments with a defined size (Swarts et al., 2012). These small fragments are then inserted at the end of a CRISPR locus in a polarized fashion next to the leader sequence (Figure 5, Chapter 2) (Barrangou et al., 2007; Díez-Villaseñor et al., 2013;

Pourcel et al., 2005; Tyson and Banfield, 2008). Spacer-repeat units can be added to CRISPR loci as a single or as group of four units (Deveau et al., 2008). Upon the insertion of a new spacer, the repeat at the end of the leader sequence duplicates to flank the newly inserted spacer (Yosef et al., 2012). Once a spacer is established in a CRISPR locus, it provides specific immunity against phages or plasmids with matching protospacers.

Cas1 and Cas2 proteins are required for spacer acquisition (Babu et al., 2011; Datsenko et al., 2012; Díez-Villaseñor et al., 2013; Nuñez et al., 2014; Yosef et al., 2012). Cas1 and Cas2 are not important for other CRISPR-related processes such as crRNA biogenesis and targeting (Brouns et al., 2008; Deltcheva et al., 2011; Hatoum-Aslan et al., 2014; Sapranauskas et al., 2011). *Cas*1 and *Cas*2 are conserved genes and exist in most CRISPR/Cas systems (Makarova et al., 2011). Cas1 and Cas2 seem to function together directly since they were found to form a complex (Plagens et al., 2012; Yosef et al., 2012). To date, the exact mechanisms by which these proteins facilitate the acquisition of new spacers are not fully understood. Cas1 is a metal-dependent endonuclease that binds to dsDNA in a sequence-independent manner at the cleavage site (Babu et al., 2011; Beloglazova et al., 2008; Nam et al., 2012a; Wiedenheft et al., 2009). Cas1 cleaves dsDNA into short fragments that might serve as precursors for new spacers (Han et al., 2009; Wiedenheft et al., 2009) Cas2 is also a metal-dependent nuclease that is known to have endonuclease activities against dsDNA, and sometimes is found fused to another Cas protein (Beloglazova et al., 2008; Nam et al., 2012a). In addition to Cas1 and Cas2, other Cas proteins such as Cas3 and Cas4 may play a role in spacer acquisition (Makarova et al., 2006; Makarova et al., 2011; Plagens et al., 2012; van der Oost et al., 2009).

E. amylovora CRISPR/Cas system

Rezzonico et al. (2011) characterized the molecular structure of the CRISPR/Cas system of 37 *E. amylovora* isolates from worldwide, and a total of 18 CRISPR genotypes were defined. The genetic diversity of *E. amylovora* was explored using sequences analysis of CRISPR regions from 85 *E. amylovora* strains from different geographic locations, plant hosts, plasmid content and variable sensitivity to streptomycin (McGhee and Sundin, 2012). A total of 588 unique spacers were determined; however, only 16% of the spacers were found to match sequences in plasmids and 5% match sequences in phage genomes from GenBank (McGhee and Sundin, 2012). The role of CRISPR/Cas system of *E. amylovora* in phage resistance and plasmids transformation blocking has not been investigated.

The objective of the study in this chapter is to determine the existence then the structure of CRISPR/Cas system in the genome of selected *E. amylovora* isolates from different regions of Canada. Also, to study spacer sequences in CRISPR arrays and determined if they match sequence in plasmids and phages in GenBank. The response of *E. amylovora* CRISPR system to phage infection and plasmid transformation will be also investigated.

Materials and Methods

Bacterial isolates and culture conditions

E. amylovora isolates (Table 1) used in this study were originally isolated from apple and pear orchards demonstrating fire blight symptoms. *E. amylovora* was cultured on 2.3% nutrient agar (NA; Difco Laboratories, Sparks, MD, USA). Cultures were incubated at 27°C for 18 h. Bacterial liquid cultures were carried out in 0.8% nutrient broth (NB; Difco Laboratories, Sparks, MD, USA). Liquid cultures were incubated at 27°C on an orbital shaker at 200 rpm for 18 h using New Brunswick™ Innova® 44 incubator shaker (Eppendorf, Mississauga, ON, Canada).

E. coli (used for cloning) was cultured on Luria-Bertani agar (LB agar, plates (1% tryptone, 0.5% yeast extract, 0.5% sodium chloride and 1.5% agar, Difco Laboratories, Sparks, MD, USA) plates, and incubated at 37°C for 16 h. *E. coli* liquid cultures were grown in Luria-Bertani broth (LB, 10 g/L tryptone, 5 g/L yeast extract and 10 g/L NaCl, Difco Laboratories, Sparks, MD, USA), pH 7.5 . Antibiotics were added to broth and agar media at the following concentrations: 50 µg/mL ampicillin and 34 µg/mL kanamycin. Liquid cultures were incubated at 37°C with shaking at 200 rpm for 16 h.

PCR amplification and DNA sequencing of CRISPR arrays and Cas genes

All PCR products in this study were amplified using the *Pfx50*™ DNA polymerase (Invitrogen, Burlington, ON, Canada). PCRs for CRISPR arrays were performed using the following general PCR conditions: 2 min at 94°C, (30 sec at 95°C, 30 sec at 58°C and 3 min at 68°C) × 30 cycles, and final extension for 5 min at 68°C. The primers used for PCR amplification of CRISPR loci and Cas genes of all *E. amylovora* isolates were designed based on sequences of CRISPR regions of *E. amylovora* strain CFBP 1430

(GenBank accession number FN434113) (Smits et al., 2010). Primers used for

amplification can be found in Appendix, Tables 1A and 2A. All PCR products were

tested for size on 1% agarose gel. All PCR products were purified using Norgen Biotek

Corp. PCR purification kit (Norgen Biotek Corp., Thorold, ON Canada). PCR products

were cloned in PCR-Blunt II-TOPO® vector (Invitrogen), and sent for sequencing

(MOBIX Lab McMaster University, Hamilton, ON, Canada). When the whole CRISPR

array or *Cas* genes could not be sequenced with a single sequence read, the complete

sequences of CRISPR array was obtained using primer walking. Primers used for Cas

genes sequencing can be found in Appendix, Tables 3A. Cas genes were identified

using BLASTN against other Cas genes in GenBank data base. CRISPR spacers were

determined using the CRISPR finder tool (available at http://crispr.u-psud.fr/). Spacer

sequences were analysed by BLASTN for homology to phage or plasmid DNAs in the

NCBI nucleotide database. To identify conserved sequences within spacers and

repeats, alignments were done using WebLogo (Crooks et al., 2004).

Plasmid and bacterial genomic DNA isolation

Plasmid DNA was isolated from 2 mL culture using Plasmid MiniPrep DNA Kit

(Norgen Biotek Corp.). Bacterial genomic DNA was extracted from 1 mL culture using

Norgen Bacterial Genomic DNA Isolation Kit.

Plasmid DNA transformation into *E. coli* and *E. amylovora*

Transformation using chemically competent *E. coli* cells

Approximately 5 to 10 μL (10 to 50 ng) of plasmid DNA or of ligation reaction was

transformed into 100 μL *E. coli* DH5α competent cells (Invitrogen). DNA and cells were

mixed very gently and incubated on ice for 15 to 30 min. Cells were heat shocked for

45 sec to 1 min at 42°C. Subsequently, the tubes were placed immediately on ice for 2

min, and 300 to 500 μL of super optimal broth (SOB) medium (2% tryptone, 0.5% yeast

extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl2, 10 mM MgSO4) was added to each

tube. The cells were incubated for 1 h at 37°C and shaken vigorously using

Thermomixer (Eppendorf, Mississauga, ON, Canada). Cells were plated out on LB agar

plates (Difco Laboratories). Appropriate antibiotic for selection was used as needed

followed by incubation for overnight at 37°C.

E. amylovora transformation using electroporation

Preparing *E. amylovora* electrocompetent cells

Frozen glycerol stocks of *E. amylovora* isolates were first streaked on NA plates

and incubated 18 h at 27°C. Two mL of NB was inoculated with a loop full of bacteria as

a starter culture and shaken at 27°C at 200 rpm for 8 h. Next, a 100 mL of NB was

inoculated with 0.5 mL of the starter culture and grown in 27°C shaker until OD_{600}

reading of 0.4 - 0.6 was reached. Once the culture reached the desired density, cells

were placed on ice for 20 to 30 min and then pelleted by centrifugation at 3500 xg for 10

min at 4°C. Cells were washed by resuspension in 50 mL of ice-cold double-distilled

water (ddH_2O). Cells were harvested again by centrifugation at 3500 xg for 10 min at

4°C. The cell pellet was resuspended in 20 mL of ice-cold 10% glycerol in ddH_2O. The

cells were harvest by centrifugation at 3500 xg for 10 min at 4°C. The supernatant was

carefully aspirated and the cell pellet resuspended gently in 1 mL of ice-cold 15%

glycerol. Cells were aliquoted into sterile 1.5 mL microfuge tubes (100 μL each) and

stored in the -80°C freezer.

Transformation using electroporation

BIO-RAD Gene-Pulser (Bio-Rad, Mississauga, ON, Canada) was used for

transformation of *E. amylovora* isolates. The Gene-Pulser was set to capacitance of 25

µF, resistance of 200 Ω, and voltage of 2.5 kV. Competent cells were placed on ice for

15 min, then 5 to 10 µL (10 to 50 ng) of plasmid DNA was added and allowed to rest for

a few minutes. The mixture was transferred to a pre-cooled cuvette (Fisher brand 2 mm

gap). The cuvette was quickly placed into the Gene-Pulser chamber and the pulse was

delivered. Cells were immediately recovered by adding 1 mL of SOC medium (2%

tryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl2, 10 mM MgSO4,

and 20 mM glucose) and shaken for 2 to 4 h at 27°C. After the recovery period, cells

were plated on NAS plates (2.3% nutrient agar, 2% sucrose and 1% sorbitol). 50 µg/mL

ampicillin antibiotic for selection was used as needed followed by incubation for 24 to 48

h at 27°C.

Constructions of plasmids carrying Cas genes for enhancing *E. amylovora*
CRISPR/Cas system

To investigate the efficiency of *E. amylovora* CRISPR/Cas system, four plasmids

were constructed: pSp21-4, pSpCas123, pSpCasg and pSpdel (Figure 13, Results).

pIDTBlue plasmid, encoding ampicillin resistance (Integrated DNA Technologies, Inc.,

Coralville, IA, USA), was used to construct these plasmids. pSp21-4 plasmid carried

CRISPR leader sequences from isolate EaD7 followed by three spacers that were

created to target sequences of structural proteins in phage ΦEa21-4, refer to Figure 1A

in the Appendix for more details about the plasmid and target proteins in ΦEa21-4 . The

spacers were separated by four repeats matching repeat sequences found in EaD7

CRISPR arrays (refer to Figure 13 for repeats and spacers sequences). The second plasmid, pSpCas123, carried the exact sequence as in pSp21-4 plus *Cas*1, *Cas*2 and *Cas*3 of EaD7. To construct pSpCas123, Cas1, 2 and 3 genes from EaD7 were amplified using PCR as described previously. Cas genes were amplified either individually or in groups using primers with overhangs harboring sequences of restriction enzymes needed for cloning. Amplified Cas genes were then cloned in a stepwise manner. For the third plasmid, pSpCasg, *Cas*4, 5 and *Cse*1, 2, 3, 4 from EaD7 were cloned in pSp21-4. A fourth plasmid, pSpdel, carried no insert and was used as a control to test the effects of the plasmid backbone. All the leader sequences, spacers and repeats, and Cas genes in these plasmids were cloned downstream from the T7 promoter to enhance the level of expression, refer to Figure13 in the Results for more details on the plasmids. All plasmids were introduced to electrocompetent *E. amylovora* Ea6-4 cells as described previously.

Construction of plasmids for *E. amylovora* transformation studies

In order to investigate the ability of the *E. amylovora* CRISPR/Cas system in blocking plasmid transformation, three plasmids were constructed pProtoCR1, pProtoCR2 and pProtoCR3, using ampicillin resistant pIDTBlue plasmid (Integrated DNA Technologies, Inc.). Each carries three protospacers complementary to endogenous CRISPR spacers in CRISPR arrays 1, 2, or 3 of Ea6-4, respectively (Figure 18, in Results). pProtoCR1 carried protospacers complementary to target spacers 1, 6 and 21 in CRISPR region 1 arrays (CRR1) of Ea6-4, pProtoCR2 plasmid protospacers target spacers 1, 22 and 26 in CRISPR region 2 arrays (CRR2), and pProtoCR3 for spacers 1, 3 and 4 of CRISPR region 3 arrays (CRR3). Each protospacer cloned in these plasmids was flanked with a PAM sequence to enhance

plasmid targeting by CRISPR/Cas system of the host. PAM sequences used in this experiment were determined based on protospacers alignment results from this study (Figure 18). A fourth plasmid, pProtoN, carried a random sequence insert with similar size (126bp) as inserts in target plasmids, was used as a control. The sequences of the insert in the control plasmid do not complement sequences to any of CRISPR spacers in *E. amylovora*. Plasmids and inserted protospacers sequences are shown in Figure 18. Protospacers cloned in these plasmids were chemically synthesized to carry *Eco*R1 sequences at both ends. To generate these plasmids, the pIDTBlue plasmid was modified to carry *Eco*R1 site. The *Eco*R1 was then used to linearize the plasmid and insert protospacers just downstream of T7 promoter (Figure 18). All cloned plasmids were verified by restriction enzymes analysis and sequencing of the inserts.

For transformation, electrocompetent Ea6-4 cells were thawed on ice, and 5 μL of plasmid DNA (2 ng/μL) was added to the cells. Cells were transformed by electroporation using BIO-Rad Gene-Pulser as explained previously. 50 μL and 100 μL of transformed cells were plated on 90 mm NAS plates supplemented with 50 μg/mL ampicillin. Plates were incubated for 16 h at 27°C. Next day, the colonies were counted and transformation efficiency was calculated. All transformations were carried out in triplicate. Transformation efficiency of each test plasmid was expressed by calculating the number of CFU per 1 μg of plasmid DNA. Plasmids that escaped the CRISPR/Cas system were isolated and sent for sequencing to check for possible deletions or mutation in protospacers.

Generation of *E. amylovora* bacteriophage-insensitive mutants (BIMs) and investigation of CRISPR region for spacer acquisition

A modified secondary culture method was used to generate *E. amylovora* BIMs (Guglielmotti, 2006). BIMs were generated using wild-type Ea6-4 and Ea6-4 transformed with pSpCas123 or pSpCasg plasmids and phagesΦEa31-3, ΦEa9-2, ΦEa35-70 and ΦEa21-4. In this procedure, 3 mL of *E. amylovora* culture (OD_{600} = 0.6; approximately 10^8 cells per mL) in 0.8% NB was prepared. Phages were added to the culture at MOI of 1.0. Then, the culture was incubated at 27°C for 16 h at 150 rpm. 5 mL of fresh NB was added to the culture and further incubated for 18 h to allow secondary growth of bacteria. The culture was then streaked onto NA plates and incubated for 18 h at 27°C. Single colonies were picked and resuspended in 3 mL of 0.8% NB. The liquid bacterial culture was incubated until OD_{600} reached 0.6. The same culture was reinfected with the same phage at MOI of 1.0. The culture was incubated at 150 rpm for 16 h at 27°C, and non-infected cells were plated as described previously. A third round of subculture and infection was carried out as described. The bacteria which survived the continuous infection process with the same phage were termed BIMs. The resistance of the BIMs to phage infection was further confirmed by exposing to the original phage and plated using the double agar overlay plate method (Adams, 1959).

To investigate if CRISPRs were responsible for BIM generation, BIM CRISPR arrays were PCR amplified, cloned and sent for sequencing (MOBIX Lab McMaster University, Hamilton, ON, Canada). The sequence results of BIM CRISPR arrays were analyzed and compared to parental hosts to see if any alteration or addition of new spacers within CRISPR arrays had occurred. An alternative PCR procedure was employed to screen for spacer acquisition. In this protocol, set of primers were designed

to amplify the regions starting from the leader sequences proximal end of the CRISPR array of CRR1, CRR2 and CRR3 and to include the first few spacers of these regions (Figure 7). The forward primers targeting the end of the leader sequences: CR1AcqF (5'CGTGCGTGCTTTAAAGTGGAAA3') is located 59nt upstream of the first repeat of CRR1; the reverse primer CR1AcqR (5'AACACAGTTCCCGAGATGCTCG3') annealing in spacer 4 of this region. For CRR2, forward primer CR2AcqF (5'CGCGGCTTTTGACAGCAAA3') annealing 60 bases upstream of the first repeat, reverse primer CR2AcqR (5'CGGGGAACACGCTGTTTTGT3') located at the junction of spacer 3 and repeat 4 of this locus. To detect spacer acquisition in CRR3, forward primer CR3F (5'AGACTGGCGTTATAGGATGG3') and reverse primer CR3R (5'AGCCCGTGAAGCAAAGTATG3') were used to amplify the whole CRR3 to produce a PCR product of 555bp. All PCR product sizes were evaluated on agarose gels to detect any spacer acquisition. In case of spacer insertion, shift in the PCR product by at least 61 bp. Primers used to detect spacer acquisition in CRR1, CRR2 and CRR3 of *E. amylovora* are listed in Table 5A, Appendix.
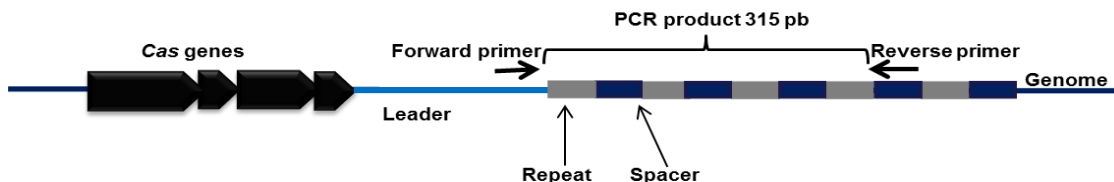
 Figure 7: PCR strategy used to investigate integration of new spacers in CRR1. Leader sequences upstream of CRISPR arrays are shown by light blue line. Grey boxes represent repeats and dark blue boxes represent spacers. The expected PCR product size from wild type is 315 bp. It was expected that the PCR product to be larger in size by at least 61 bp in case of integration of one spacer and one repeat. A similar strategy was used to explore spacer insertion in CRR2. In CRR3, however, the entire CRISPR arrays were amplified due to the small size.

Investigation of Cas genes and crRNA expression

Total RNA extraction and crRNA enrichment

Total RNA from *E. amylovora* was extracted from 1 mL bacterial culture (10^8 CFU) using the Norgen Total RNA Isolation Kit (Norgen Biotek Corp., Thorold, ON Canada). DNase treated RNA was analyzed using a 1% agarose gel electrophoresis to ascertain the quality and integrity of RNA and the absence of DNA contamination. crRNA enrichment was carried out using microRNA Purification Kit (Norgen Biotek Corp.).

RT real-time PCR

One-step RT real-time PCR assay was performed with the final volume of 25 µL using the Roto-Gene Rrobe RT-PCR Kit (Qiagen, Toronto, ON, Canada) and TaqMan Probes (Integrated DNA Technologies, Inc., Coralville, IA, USA). Each reaction was carried out using 2 µL of total RNA. RT real-time PCR program was reverse transcription of RNA at 50°C for 10 min. The PCR performed with an initial activation step at 95°C for 5 min, then 40 cycles of: denaturation at 95°C for 20 sec, annealing at 45°C for crRNA, 54°C for *Cas*1 or *Cas*3 for 20 sec, and amplification at 60°C 20 sec. The fluorescence data were collected at the end of every 60°C step. Real-time multiplex reaction mix was carried out in total reaction volume of 25 µL according to manufacturer's recommendations with final concentration of 200 nM primers and 100 nM of gene of interest probe, and 100 nM primers and 50 nM probes for each housekeeping gene.

Three biological replicates (RNA isolated from three independent cultures) were carried out. The RT real-time PCR was chosen to measure RNA level since it is very

sensitive and more reliable than the commonly used Northern blot protocol. The comparative C_t method $\Delta\Delta CT$ (Livak and Schmittgen, 2001) was used for level of expression data analysis and genes of interest were normalized to two endogenous housekeeping genes (*rec*A and *gro*EL). Level of expression at T_0 was considered as an initial value of one fold. Fold change at a given time is the ratio of the value at a given time to the initial value (T_0). Changes in expression levels were considered significant when the relative expression between a given point and T_0 was ≥ one fold change. Primers and probes used for RT real-time PCR are listed in Table 4A in Appendix.

Phage titre

Plaque assay

Phage titres were determined by the double agar overlay method as described by (Adams, 1959). Phage stock solutions were serially diluted by using 0.1ml phage and 10 tubes contained 0.9 mL sodium phosphate buffer, pH 6.8 to generate 10 fold serial dilutions. 0.1 mL of 10^{-4}, 10^{-6} and 10^{-8} dilutions were mixed with 0.1 mL of bacterial dilution at 10^8 CFU/ml in a test tube. The phage/cell mixes were incubated at room temperature for 10 to 15 min. In the meantime, molten top agar (0.8% nutrient agar, 0.25% yeast extract and 0.5% sucrose) was held at 50°C in a water bath. 4 mL of the molten top agar was added to each phage/bacteria mix. Tubes were swirled gently and poured on to solidified nutrient agar to form a thin top layer. Plates were incubated at room temperature to allow the top agar to solidify. The plates were incubated at 27°C for 18 h and PFU/mL in each sample was determined.

Determining phage titer with quantitative PCR

In this method, the absolute phage titer was determined using the standard curve method. The unknown phage titer was extrapolated based on a standard curve generated using known quantities of plasmid DNA.

Standard curve preparation using plasmid template and phage titer

The pTotalStdA plasmid (Appendix, Figure 2A), developed by Steven Gayder (PhD Candidate, Brock University), was used to create standard curves in order to quantify phages with quantitative PCR assay. pTotalStdA was constructed to contain the full-length PCR amplicon sequences targeted by primers and probes of phages used in this study: ΦEa31-3, ΦEa21-4, ΦEa9-2 and ΦEa35-70. pTotalStdA was propagated in *E. coli* and purified using the Plasmid MiniPrep DNA Kit (Norgen Biotek Corp., Thorold, ON Canada). The concentration of plasmid DNA was measured by at OD_{260}, and the plasmid copy number was calculated based on the concentration and molecular weight of the DNA. The purified plasmid DNA was linearized with *Xho*I and the standard curve was generated (Invitrogen). The standard curve was created using 10 fold serial dilutions of the linearized pTotalStdA ranging from 10^1 to 10^{11} plasmid copy/mL. Standard curve was created using at least four points in duplicate. Additionally, the range of concentrations in the standard curve covered the entire range of unknown phage concentrations. The amplification efficiency of varying plasmid dilutions were determined by obtaining a linear standard curve and $R^2 \geq 0.98$. After amplification of pTotalStdA dilution series, the C_t values generated for each dilution were plotted against the $\log_{10}$ of copy number of the plasmid. The Stratagene™ Mx3005P system analysis software (Agilent Technologies, Santa Clara, CA, USA) was used to determine the unknown phage concentration by comparing C_t values of the unknown samples to the standard curve.

To titer phage solutions, quantitative PCR was performed on the Stratagene™ Mx3005P system (Agilent Technologies) in 96-well plates covered with adhesive sealing (VWR, USA). The quantitative PCR reaction was carried out in a total volume of 20 μL using EVOlution Probe qPCR mix kit (MBI Montreal Biotech Inc., Montreal, QC, Canada) with a final concentration of 200 nM primer and 100 nM probe. The PCR was performed with a two-step PCR protocol: activation of the *Taq* DNA polymerase at 95°C for 10 min, followed by 40 cycles of 95°C for 10 sec and 54°C for 45 sec. All real-time PCRs were performed as triplicates. Primer and probes used to quantify different phage types are listed in Appendix, Table 6A.

Challenging artificial Cas genes with phages

Three *E. amylovora* Ea6-4 mutants were generated to carry plasmids: pSpdel, pSpCas123, and pSpCasg (Figure 13). The specifications for these plasmids were previously described in this chapter. To determine the effect of transformed plasmids on the efficiency of *E. amylovora* CRISPR/Cas system, transformed isolates were challenged with four different phages: *Myoviridae* phages ΦEa21-4 and ΦEa35-70 and *Podoviridae* phages ΦEa31-3 and ΦEa9-2. Transformed *E. amylovora* isolates were grown in liquid culture on a shaker at 27°C until OD_{600} reached 0.1. Bacterial cultures were infected with the previously listed phages at MOI of 0.1. Cultures were incubated for 8 h. Phages titers were determined every 2 h with quantitative PCR as described previously.

Results

Bioinformatics analysis of CRISPR arrays and Cas genes

CRISPR arrays sequences (spacer and repeats)

The sequences of CRISPR arrays of selected *E. amylovora* isolates from different areas in Canada were determined (Table 1). A total of three CRISPR arrays labelled CRR1, CRR2 and CRR3 were identified in *E. amylovora* isolates Ea6-4, EaD7, Ea17-1-1 and Ea110R. Only two CRISPR arrays, CRR2 and CRR3, were detected in *E. amylovora* isolates EaG5, BC29, BC1280 and Ea4-96 (Table 1).

E. amylovora isolates Ea6-4, EaD7, Ea17-1-1 and Ea110R carry 36 spacers in CRR1. CRR1 in these isolate is located upstream of the Cas gene cluster (Figure 8). The average length of the spacers in this region is 32bp and spacers were interspersed by 37 palindromic repeats. Spacer sequences from these three isolates were identical to CRR1 of *E. amylovora* CFBP1430 (Smits et al., 2010). The average length of the repeats in CRR1 of these isolates is 29 bp. The consensus sequence of the repeats is (5'GTGTTCCCCGCG<u>TGA</u>GCGGGGATAAACCG3'). CRR2 of Ea6-4, EaD7 and Ea17-1-1, on the other hand, were found just downstream of the Cas gene cassette (Figure 8), and contained 26 spacers. Spacer sequences of CRR2 of Ea6-4, EaD7 and Ea17-1 isolates were identical and separated by 27 repeats (average 29 bp in size) with a consensus sequences (5'GTGTTCCCCGCG<u>TAT</u>GCGGGGATAAACCG3') which was slightly different from repeats in CRR1 (refer to underlined sequences in both repeat sequences). CRR2 of Ea110R carries only 23 spacers interrupted by 24 palindromic repeats with consensus sequences matching Ea6-4, EaD7 and Ea17-1-1. The 23 spacers of CRR2 of Ea110R match spacers in Ea6-4, EaD7 and Ea17-1-1; however, Ea110R lacks spacer 7 (5'ATCGGAACGACTTAGATTAGCGTCCTTGCACAT 3'), spacer 8 (5'TAACCGCAACCATCGCCGCGATAAATCCACTG3') and spacer 9

(5'GTTGATACGGCTGATTACAATAAAATGTCACT3') that exist in Ea6-4, EaD7 and Ea17-1-1 isolates.

Isolates EaG5, BC29, BC1280 and Ea4-96 have a larger CRR2 than isolates Ea6-4 and EaD7 from Essex County, Ontario. CRR2 of EaG5 has 57 spacers interspersed with 58 repeats, which makes this region as one of the largest CRISPR array as compared to any known CRISPR arrays found in *E. amylovora* isolates worldwide. CRR2 from BC29 and BC1280 were identical and carried 49 spacers. Isolate Ea4-96 has a shorter CRR2 with 30 spacers. Alignments of CRR2 of all isolates revealed that isolates Ea6-4, Ea17-1-1 and EaD7 from apple (Essex County, Ontario) shared the exact spacer sequences in all CRISPR regions. Interestingly, majority of spacers exist in BC29 isolates from pear (Kelowna, British Columbia) and Ea6-4 from apple (Essex County, Ontario) exhibited matches to spacers found in EaG5. Isolate Ea4-96 from raspberry (Kentville, Nova Scotia) carried unique spacers which were not found in any other isolate (Figure 9).

All bacterial isolates in this study contained only five spacers in CRR3 with exception to Ea4-96 which carried 6 spacers. Spacer 1 (5'GGCTAAATGCCTGAATTCAACGTTCTACAAAA3') of Ea4-96 was not found in other isolates. However, all spacers in CRR3 are identical among all isolates. The repeat in CRR3 of Ea4-96 is 24 bp in size (5'GTTCACCGCCGTACAGGCGGCTTA3'), and is shorter than the repeats of CRR3 of other isolates, which is 28 bp (5'GTTCACTGCCGTACAGGCAGCTTAGAAA3').

Bacterial isolates	Origin	CRISPR1		CRISPR2		CRISPR3	
		CRISPR length (bp)	# Spacers	CRISPR length (bp)	# Spacers	CRISPR length (bp)	# Spacers
Ea6-4	Apple (Essex County, ON)	2224	36	1616	26	328	5
EaD-7	Apple (Essex County, ON)	2224	36	1616	26	328	5
Ea17-1-1	Apple (Essex County, ON)	2224	36	1616	26	328	5
Ea110R	Rifampicin resistant-mutant Apple (MI, USA)	2224	36	1432	23	328	5
EaG5	Unknown	Not detected		3508	57	328	5
BC29	Pear (Kelowna, BC)	Not detected		3020	49	328	5
BC1280	Apple (Westbank, BC)	Not detected		3020	49	328	5
Ea4-96	Raspberry (Kentville, NS)	Not detected		1858	30	383	6

Table 1: List of *E. amylovora* isolates selected for CRISPR arrays sequencing. The origin of each isolate is shown, ON: Ontario, BC: British Columbia, NS: Nova Scotia, MI; Michigan. The length of the CRISPR array in base pairs is shown. The number of spacers found in each CRISPR array and the average size of spacers (32 bp) are also indicated.
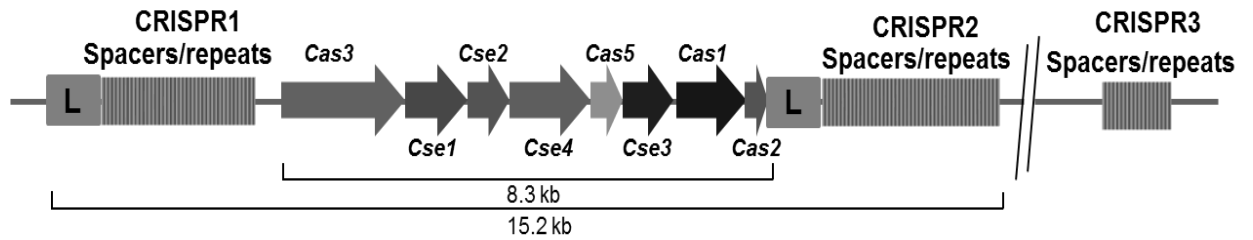
Figure 8: Genetic map of the CRISPR locus of *E. amylovora* Ea6-4 and EaD7. The location of CRISPR-associated genes (*Cas* and *Cse*) and CRISPR arrays 1, 2 and 3 (CRR1, CRR2, and CRR3) are shown. Cas genes occupy a region of 8.3 kbp. Sequences represented by thin gray lines contain genes unrelated to CRISPR function. Leader sequences of CRR1 and CRR2 are shown by letter (L). The distance between CRR2 and CRR3 is not known.

Isolate	Spacer location																							
Ea6-4, EaD7, Ea17-1-1																								
EaG5	1	2	3	4	5	6	7	8	9	10	11			12	13	14	15	16	17	18	19	20		21
BC29, BC1280												1	2	3	4	5	6	7	8	9	10	11	12	13
Ea4-96	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24

Isolate	Spacer location																							
Ea6-4, EaD7, Ea17-1-1					1	2	3	4	5	6	7	8	9	10	11	12		13		14	15			
EaG5	22	23	24	25						26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
BC29, BC1280	14	15	16	17						18	19	20	21	22	23	24	25	26		27	28	29	30	31
Ea4-96	25	26	27	28	29	30																		

Isolate	Spacer location																	
Ea6-4, EaD7, Ea17-1-1								16	17	18	19	20	21	22	23	24	25	26
EaG5	41	42	43	44	45	46	47		48	49	50	51	52	53	54	55	56	57
BC29, BC1280	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49
Ea4-96																		

Figure 9: *E. amylovora* CRR2 spacer genotypes.

Each CRISPR spacer is represented by a box and number that indicates the spacer position in the array of a specific isolate. The first spacers that lie next to the leader sequence are on the left side of the picture and the oldest spacer is on the right. . Identical spacers of different isolates are aligned in the same column and coloured grey. The white boxes with numbers indicate the absence of spacer match among isolates (e.g. spacer numbers 1 to 30 of Ea4-96).

Alignments of spacers by WebLogo representation to determine conservation within spacer sequences from CRISPR arrays CRR1, CRR2, and CRR3 from all isolates showed more than one conserved sequences among each CRISPR arrays. The spacers of CRR1of all isolates showed conversed sequences at nt3-5, nt15-17 and nt24-26. In CRR2 Spacers, conserved sequences located at nt6-15; and in CRR3 spacers at nt13-16. The existence of PAM conserved sequences flanking protospacers was also investigated in this study. The alignment of PAM sequences from phage genomes revealed the presence of short conserved regions. At the 5' end of protospacers, sequences (AAA, GAA, AAC, AAG and GTA) were detected. No conserved PAM sequences were found at the 3' end of protospacers (Figure 10).
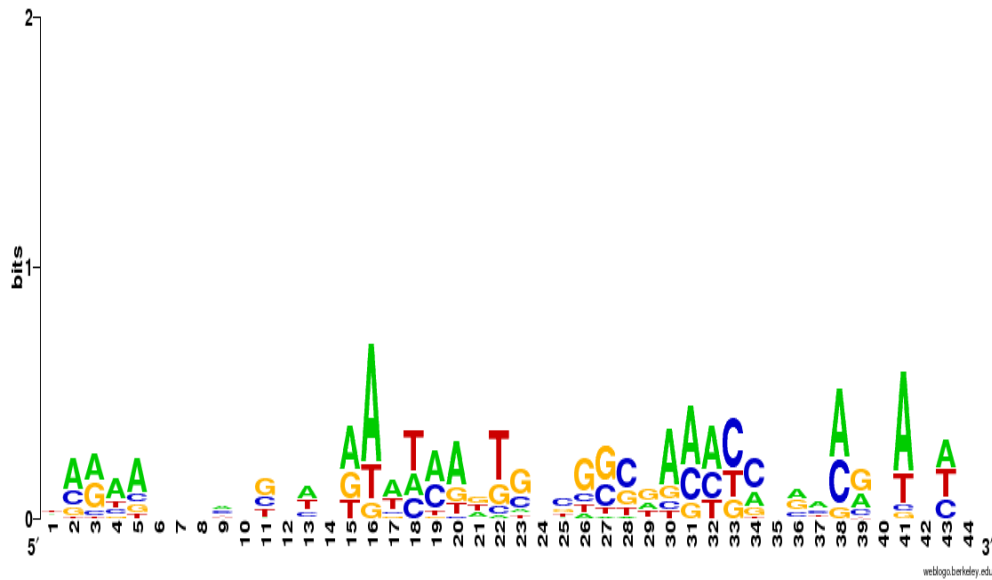
Figure 10: WebLogo representation of the sequence conservation of protospacers.

Sequence conservation of protospacers and PAM sequences from CRISPR arrays CRR1, and CRR2 of *E. amylovora* isolates used in this study. In total, only 19 protospacers were found in phages and plasmids to have exact match to spacers from this study. Protospacer sequences are present in positions 7 to 38, and PAM sequences are in positions 1 to 6 and 39 to 44. There is no continued conserved PAM sequence upstream of protospacers, an indication of insignificance of these sequences in CRISPR system function.

CRISPR spacers blast results

      124 individual spacers were identified within the three CRISPR arrays of the seven bacterial isolates in this study. A spacer was considered unique if it contained 10% or more nucleotide mismatch compared to other spacers. Sequences in phages homologous to spacers were identified using NCBI BLASTn analysis. Each spacer sequence was blasted against *Erwinia* spp. phages in NCBI database. The default values of NCBI BLASTn for short sequences were used: match/mismatch 2(-3), Expect value (10) and with modified word size (7). A majority of spacers showed no significant similarity to sequences of phages in the NCBI database. No spacers exactly matched the sequences of any sequenced *E. amylovora* phages from *Myoviridae* and *Podoviridae* families. However, three spacers exactly matched *Erwinia tasmaniensis* phage Et88 genome sequence (Müller et al., 2011a). Few *Erwinia* spp. phages shared 17 or less nucleotide match to some spacers, and the majority of these sequences are interrupted with a single nucleotide mismatch. Table 2 shows spacers that have 13 or more nucleotide match to *Erwinia* phages ΦEa21-4 (Lehman et al., 2009), ΦEa35-70 and ΦEa9-2 (this study). For example, spacer 2 in EaG5 showed 15 nucleotides match out of 32 nucleotides (47% similarity) to *E. amylovora* phage ΦEa21-4 (Lehman et al., 2009). Eleven spacers showed from 14 to 15 nucleotides match out of 32 nucleotides, approximately 44 to 47% similarity, to sequences in phage ΦEa35-70 (Yagubi et al., 2014) (Table 2). Of the all CRISPR spacers sequences determined, only 13 spacers showed similarity to sequences in plasmids and other bacterial genomes in the database, data not shown.

Spacer sequence	Host	CRISPR region	Spacer no.	Phage	Location on phage genome	Target phage gene
TTATT<u>CATGAGCC</u>**T**<u>TTTTAT</u>CTTCGCGGCATG	Ea6-4	1	1	φEa35-70	177033 to 177047	Hp
AAGATG<u>CTTTGACATTAATTA</u>TCTCCATAAAA	Ea6-4	1	5	φEa35-70	38341 to 38355	Hp
<u>AAGATGCTTTGACA</u>TTAATTATCTCCATAAAA	Ea6-4	1	5	φEa35-70	54602 to 54615	Hp
AAGATGCTT<u>TGACATTAATTATC</u>TCCATAAAA	Ea6-4	1	5	φEa35-70	73984 to 73997	Hp
CCGCGA<u>AAAT</u>**C**<u>CGCAG</u>**T**<u>GAGCTGG</u>CAATGAGC	Ea6-4	1	29	φEa9-2	63339 to 63357	Hp
TTT<u>CTTCACACACAACG</u>GTGAGGGCATTGTCT	Ea6-4	1	35	φEa21-4	61435 to 61449	NA
GCTCGGGGGGACATGAG<u>CTTGTACAAAACAG</u>C	Ea6-4	2	3	φEa21-4	73508 to 73522	Hp
GT<u>TGAT</u>**A**<u>CGGCTGATTACA</u>ATAAAATGTCACT	Ea6-4	2	9	φEa35-70	254734 to 254750	Tail fiber protein
AATGGACGAGA<u>TTTCACAG</u>**A**<u>AAATATCT</u>GTTC	Ea6-4	2	13	φEa21-4	41408 to 41423	NA
GCTACTAC<u>GTGTACGCAC</u>**A**<u>GCCGCT</u>GGCCAGT	Ea6-4	2	14	φEa35-70	100839 to 100855	Hp
CGGGCG<u>AGTTTTT</u>CAGGATGCGTCTCATTGAA	EaG5	2	2	φEa21-4	79202 to 79216	NA
TACGTCAGC<u>GGCATTAT</u>**A**<u>AGGATTTA</u>CACGGA	EaG5	2	6	φEa21-4	20967 to 20983	NA
CAACTG<u>AGCAACTCC</u>**G**<u>CTTTGCCC</u>CATACCAA	EaG5	2	15	φEa35-70	72650 to 72666	Hp
CGACTATTCG<u>GCGAAC</u>**T**<u>GAAAGAGCCA</u>ACCGT	EaG5	2	18	φEa35-70	104335 to 104351	Hp
GAGA<u>GCAACAACAA</u>**A**<u>CTCC</u>GACCGGAGAAACC	EaG5	2	23	φEa35-70	3070 to 3084	Hp
<u>GTTTCGG</u>**C**<u>GAGCGCGTA</u>TATGGTCACGTTCAC	EaG5	2	25	φEa21-4	73578 to 73594	Hp
<u>ACGGT</u>**C**<u>AGATGGTGGCG</u>CTGGTTGCGCTGGCA	EaG5	2	39	φEa35-70	22206 to 22222	Hp
TTTACGTTTGCGTTAA<u>CAGTAA</u>**G**<u>CTCTGCAAC</u>	Ea6-4	3	1	φEa21-4	43706 to 43721	Putative baseplate component
ATCGCACCC<u>CACTGATTGAAGA</u>GCAGCACACT	Ea6-4	3	2	φEa35-70	238369 to 238381	Hp

Table 2: Summary of blast results blast of spacers against available *E. amylovora* phages in GenBank.

Sequences in spacers that show similarity to sequences in phages are underlined. The mismatches are shown in bold and are not underlined. The target phage genes by these spacers are shown in the most right column; Hp, hypothetical protein and NA, not available (genes with unknown function). The location of the matching sequence in the phage genomes is shown.

Cas gene sequences analysis

The sequence analysis of CRISPR regions of bacterial isolates EaD7 and Ea6-4 revealed the presence of eight Cas genes between CRR1 and CRR2 (Figure 8). Cas genes were found to span a region of 8.3 kbp. The start codon of *Cas*3, the first gene in the cluster, is located approximately at 2.5 kbp downstream of the last repeat of CRR1. On the other hand, the stop codon of *Cas*2, the last gene in the group, is located just upstream of the leader sequences of CRR2 (Figure 8). The location of CRR3 on the genome was not determined. It is expected to be thousands of base pairs downstream of CRR2 similar to CRR3 of *E. amylovora* strain ATCC 49946 (Sebaihia et al., 2010). Furthermore, no Cas genes were found associated with CRR3 of Ea6-4 and EaD7. Generally, the amino acid sequence and the orientation of Cas genes of EaD7 and Ea6-4 were found to match Cas genes found in *E. amylovora* strains ATCC 49946 and CFBP 1430 (Sebaihia et al., 2010; Smits et al., 2010). These strains carry CRISPR/Cas type I-E similar to CRISPR/Cas found in *E. coli* (Makarova et al., 2011; Touchon et al., 2011). Cas gene sequences information and potential function are shown in Table 3.

Stage	Protein	Amino acid	Possible function
Acquisition	Cas1	305	Single-stranded DNA nucleases. Could be members of Cascade
	Cas2	116	
Processing RNA	Cas5	218	Cleaves pre-crRNA to form the mature crRNA. Could be a member of Cascade-crRNA complex
	Cse3 (Cas6)	200	
Surveillance and Interference	Cas3	877	Nuclease and helicase involved in interference. May be involved in spacer acquisition
	Cse1	507	Member of Cascade-crRNA complex
	Cse4 (Cas7)	364	
	Cse2	156	

Table 3: List of Cas proteins found in *E. amylovora* Ea6-4 and EaD7 and their potential function.

The stage when these proteins are involved during different CRISPR system mechanisms is shown in the right column. Acquisition means involved in spacer acquisition step, processing RNA: involved in maturation of pre-crRNA to produce crRNA, and surveillance and interference: proteins are involved in finding invading phage DNA and degradation. Collected from (Richter et al., 2012a; Richter et al., 2013; van der Oost et al., 2014). Cas and Cse proteins: are CRISPR-associated proteins found in CRISPR/Cas system type I.

Cas gene expression during phage infection

To obtain an understanding of CRISPR function in *E. amylovora*, the level of expression of three core genes involved in different stages in the CRISPR mechanism were monitored in Ea6-4 and EaD7 during phage infection. *Cas*1 is involved in the acquisition step, *Cas*3 in acquisition and interference and crRNA transcripts are involved in surveillance.

To investigate the response of *E. amylovora* CRISPR system to phage infection, Ea6-4 and EaD7 isolates were infected with phages ΦEa21-4, ΦEa9-2 and /or ΦEa35-70. ΦEa9-2 has one sequence with 17 nucleotides match to one spacer (32 nt) in CRR1 of Ea6-4 and EaD7. ΦEa21-4 has four sequences with 14 to 16 nucleotides match to spacers in Ea6-4 and EaD7, one in CRR1, two in CRR2 and one in CRR3. ΦEa35-70 has four sequences with 14 to 15 nucleotides match to spacers in CRR1, two in CRR2 (each with16 nt match) and one in CRR3 of Ea6-4 and EaD7 (13 nt match). It was anticipated, based on the described similarities, that the relative roles of the three CRISPR loci could be discerned by comparisons of the three phages during infection of Ea6-4 and EaD7, and the role of CRISPR/Cas systems in general by comparison of the results from challenge of the two bacterial hosts. The level of expression of *Cas*1, and *Cas*3 and crRNA were monitored for 2 h post infection. To differentiate between the phage-induced changes in Cas gene expression from changes as a result of *E. amylovora* growth, in addition to phage-infected culture, a non-infected *E. amylovora* culture was used as a control.

Cas gene level of expression was measured during host infection by the *Myoviridae* phage ΦEa21-4 (Figure 11). Results indicated that there were no major changes in the level of expression of Cas genes after 2 h of infection in comparison to non-infected cells. A minor increase in the level of expression in crRNA in Ea6-4

in the middle of the first phage replication cycle was observed, and then the level of

expression declined to a level similar to the control (Figure 11). Each replication

cycle of ΦEa21-4 in the host, takes approximately 60 min (data not shown). The

duration of infection in the experiment allowed two complete cycles of phage

replication. In addition, the growth rate of the *E. amylovora* cells was not affected by

phage infection during this experiment (data not shown). The level of expression of

Cas genes were tested using phages, ΦEa9-2 (*Podoviridae*; accession no.

NC_023579) and ΦEa35-70 (*Myoviridae*; accession no. KF806589) (Yagubi et al.,

2014). Expression of Cas genes and crRNA in EaD7 increased during the first 30 min of

ΦEa9-2 infection (Figure12). After 30 min, expression declined to levels similar to non-

infected cells. Interestingly, no effect on Cas genes and crRNA expression was

detected when Ea6-4 isolate was infected with the same phage with an exception of a

minor increase in *Cas*3 expression in the first 30 min of infection (Figure12). A major

drop in *Cas*1 was also observed late infection cycle (after 60 min) of ΦEa9-2 infection.

On the other hand, ΦEa35-70 didn't elicit any change in expression of Cas genes in

both bacterial hosts (Figure 12).
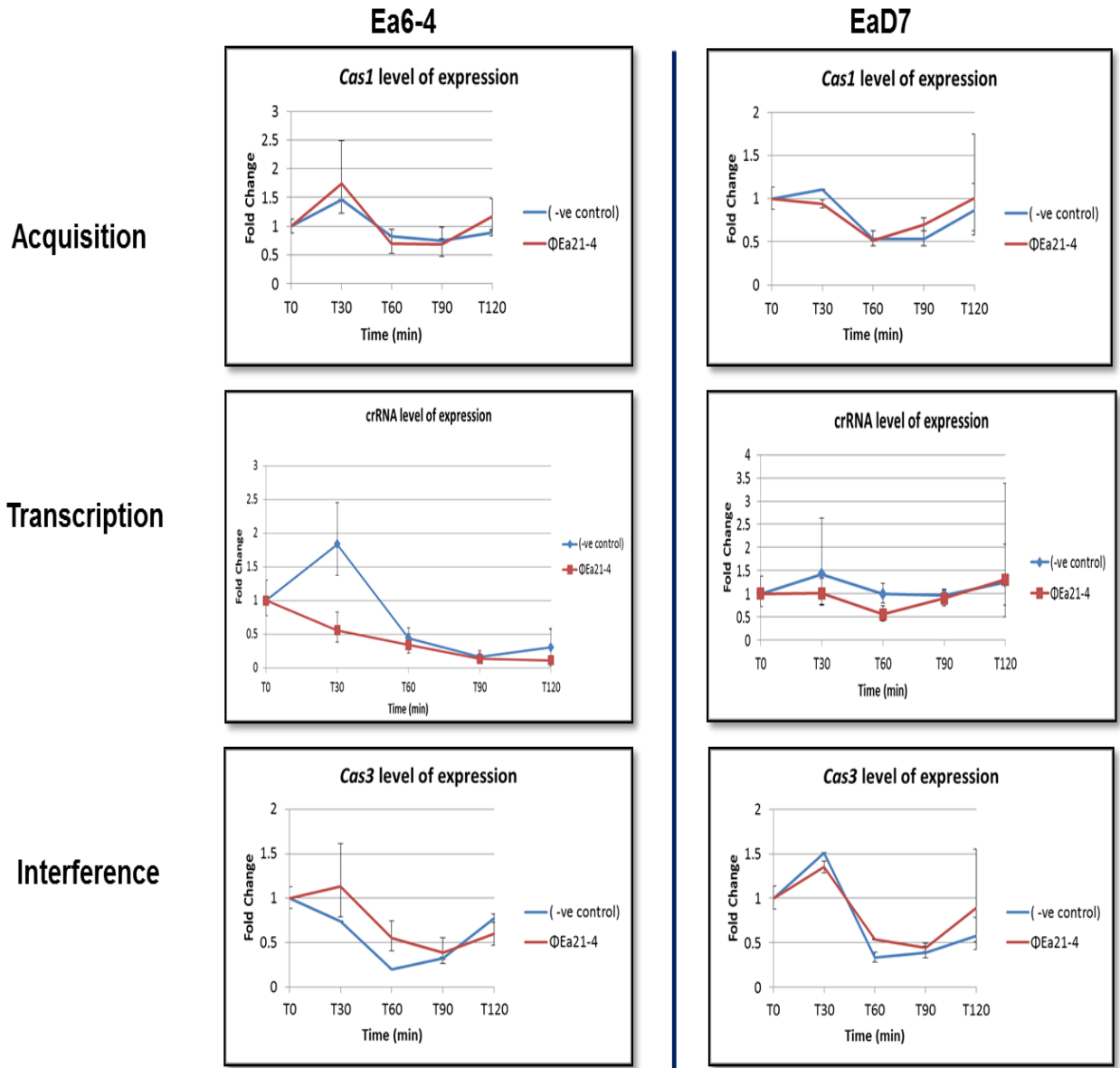
Figure 11: *Cas*1 and *Cas*3 and crRNA level of expression in *E. amylovora* cells during ΦEa21-4 infection.

Two bacterial host isolates were used: Ea6-4 and EaD7. Level of expression was measured every half hour for two hours. Negative control (-ve control) were cells with no phage infection. Steps in which these genes are involved in CRISPR function (acquisition, transcription and interference) are shown. The results were calculated using three biological replicate.
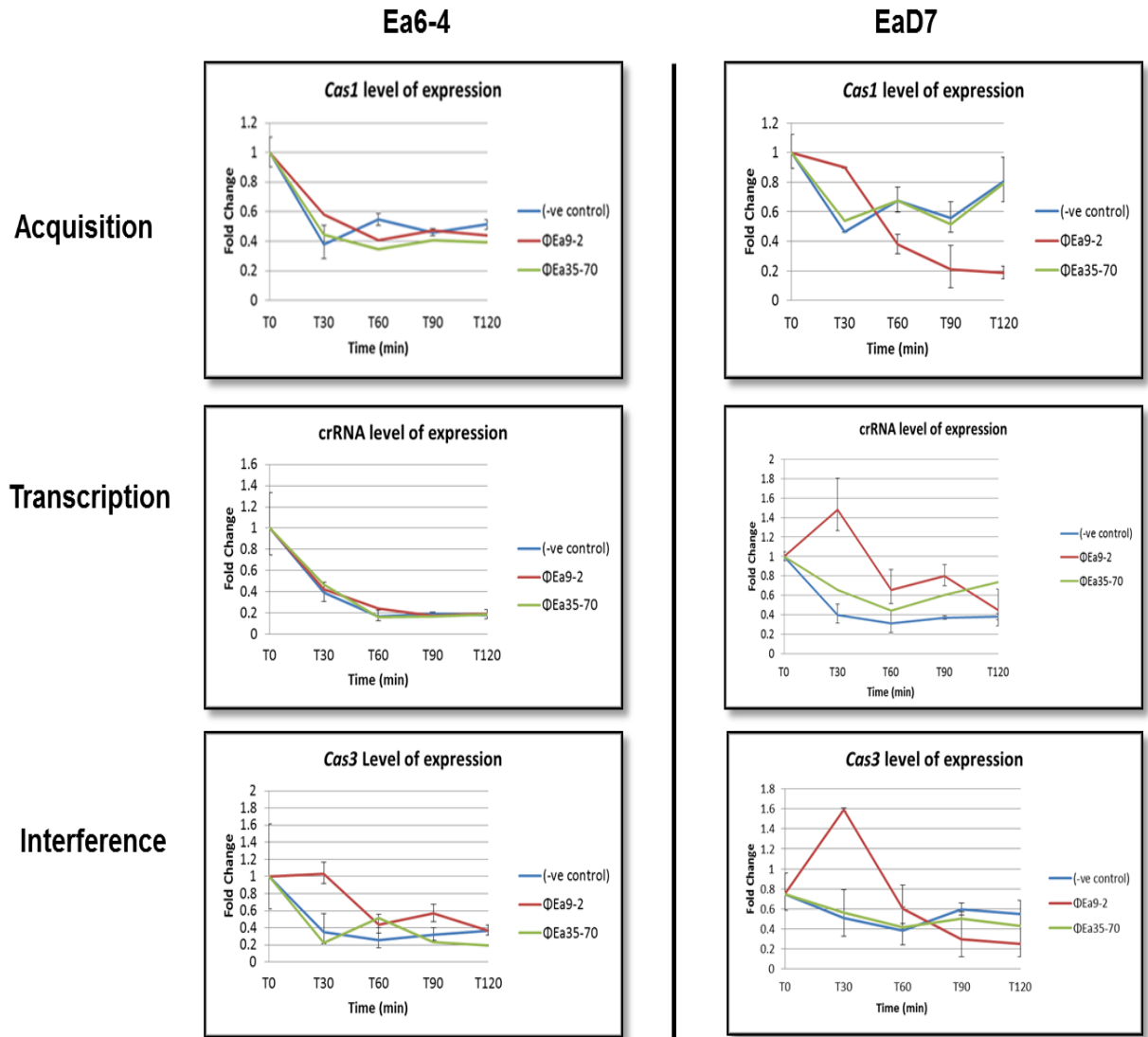
Figure 12: *Cas*1 and *Cas*3 and crRNA level of expression in *E. amylovora* cells during infection by phages ΦEa9-2 and ΦEa35-70.

Two bacterial host isolates were used, Ea6-4 and EaD7. Level of expression was measured every half hour for two hours. Negative control (-ve control; blue line) is cells with no phage infection. Level of expression at T_0 was considered as an initial value of one fold. Fold change at different times is the ratio of the value at a given time to the initial value (T_0). ΦEa9-2 (brown line) and ΦEa35-70 (green line). Steps in which these genes are involved in CRISPR function (acquisition, transcription and interference) are shown. The results were calculated using three biological replicate.

Studying *E. amylovora* Cas gene function by introducing an artificial CRISPR/Cas system into Ea6-4

Despite the existence of a CRIPSR/Cas system in *E. amylovora*, to date the presence of phage resistant bacteria have not been detected among in our collections of phages and *E. amylovora* hosts. The CRISPR/Cas system appears to be naturally inactive in *E. coli* K12 due to the low levels of Cas gene expression (Westra et al., 2010). To test whether a low level of Cas gene expression is the reason for not detecting any phage resistance in *E. amylovora*, and to study the function of these genes, four plasmids were constructed to carry various sets of Cas genes: pSp21-4, pSpCas123, pSpCasg, and pSpdel (Figure 13). The plasmids are described in greater detail in the Materials and Methods. These plasmids were introduced into Ea6-4. First, the level of Cas gene expression and crRNA production were investigated to test the effectiveness of the introduced genes. The transformed Ea6-4 host cells were subsequently challenged with selected *Erwinia* spp. phages to determine if boosting Cas gene production has an effect on phage replication.

Monitoring Cas gene expression

To test if the transformed plasmids would enhance the expression of Cas genes, and to ensure that T7 prompters would sustain a read-through transcription to Cas1 and Cas3 genes which are located distal to T7 promoter in pSpCas123 and pSpCasg plasmids, a quantitative reverse transcription PCR analysis of transcripts levels of *Cas*1 and *Cas*3 was performed. After introducing plasmids pSp21-4, pSpCas123, pSpCasg, and pSpdel into electrocompetent Ea6-4, transformants were incubated on a shaker for 4 h, and total RNA was isolated every 2 h. Then, the level of expression of *Cas*1 and *Cas*3 was evaluated using RT real-time PCR (Figure 14).

Results showed that after 2 h of incubation the transcript levels of *Cas*1 increased seven fold in Ea6-4 with pSpcas123 and 15 fold in Ea6-4 carrying pSpCasg in comparison to wild type cells. The level of expression in Ea6-4 hosts carrying pSpdel or pSp21-4 plasmids didn't increase since these plasmids do not carry an extra copy of Cas genes (Figure 14). After 4 h of incubation, the level of expression of *Cas*1 increased to 32 fold in Ea6-4 harbouring pSpCas123 and pSpCasg compared to wild-type Ea6-4. The level of expression of *Cas*1 in cells that carry only genomic copy of the gene (Ea6-4 harbouring pSpdel or pSp21-4 plasmids) did not change during the 4 h of incubation (Figure 14). On the other hand, the level of expression of *Cas*3 was relatively different from *Cas*1. The level of expression of *Cas*3 was increased by only one fold with pSpCasg and pSpCas123 in two hours incubation (Figure 14). However, in 4 h incubation the expression increase up to four fold in cells carry pSpCasg, and only two fold in cells with pSpCas123. Interestingly, the expression level of *Cas*3 was down regulated in cells with pSpdel or pSp21-4 compared to wild-type Ea6-4 (Figure 14).
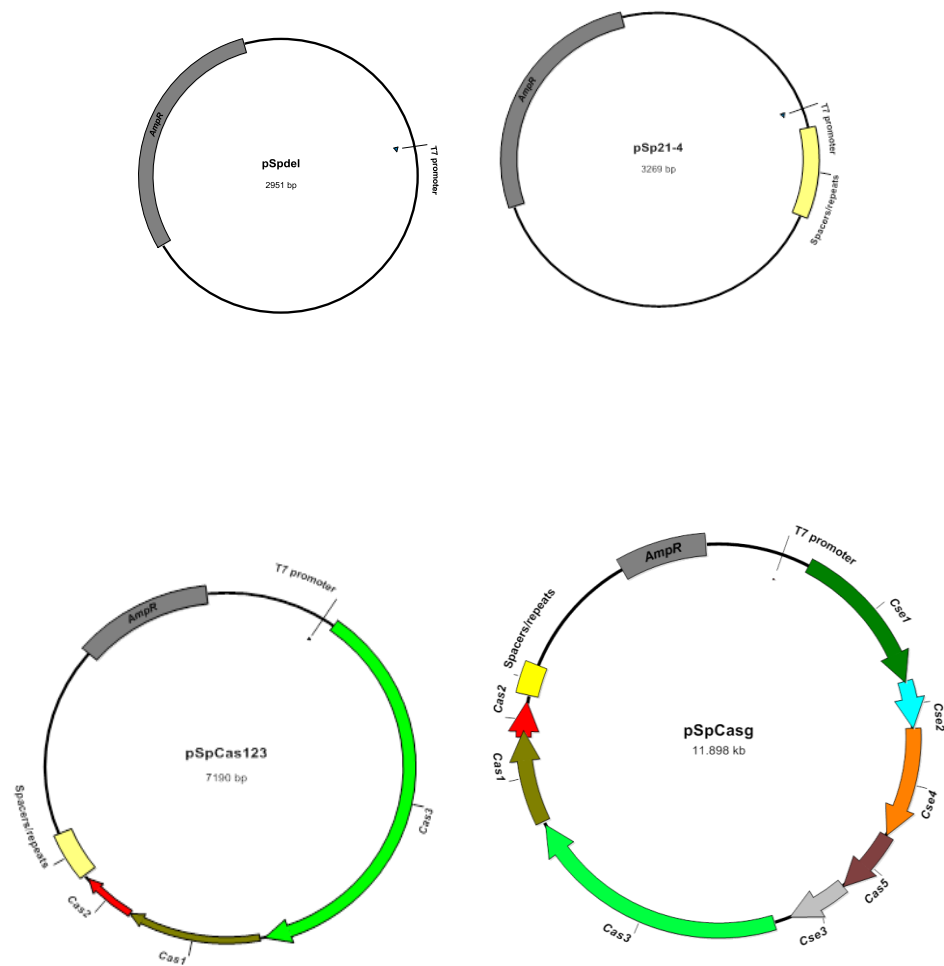
Figure 13: Maps of pSpdel, pSp21-4, pSpCas123 and pSpCasg plasmids.
pSp21-4 carries CRISPR leader sequences followed by three spacers that match
protospacers existing in phage ΦEa21-4 (more information about the protospacers and
target phage genes can be found in Figure 1A, Appendix). pSpCas123 carries the exact
sequences as in pSp21-4 in addition to *Cas*1, *Cas*2 *and Cas*3. pSpCasg carries the
exact sequences cloned in pSp21-4 in addition to all Cas genes found in EaD7 (*Cas*1,
2, 3, 4, 5 and *Cse*1, 2, 3, 4). pSpdel does not carry any inserts and serves as a plasmid
background control. T7 promoter location is shown in all plasmids.
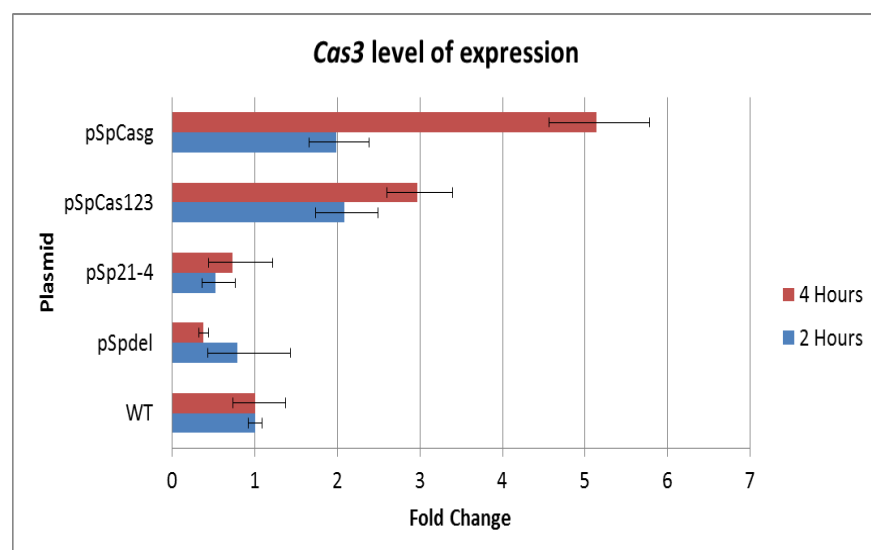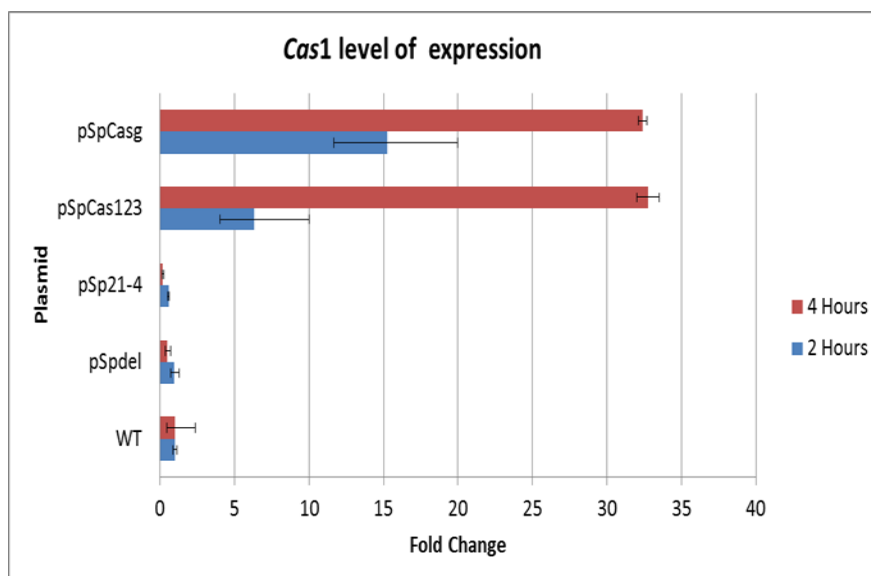
Figure 14: *Cas*1 and *Cas*3 level of expression in Ea6-4 hosts transformed with plasmids carrying Cas genes.

The level of expression of both genes was measured after 2 and 4 h of incubation. Fold changes were calculated in comparison to wild type; the level of expression in wild type host is considered as one fold. The results were calculated using three independent biological replicate.

Monitoring mature crRNA production

After determining that the levels of expression of Cas1 and 3 genes were enhanced in Ea6-4 hosts transformed with plasmids carrying these genes, it was important to assess the efficiency of the introduced Cas genes in processing and production of crRNA. If these genes are functional, an increase of crRNA production should be detected. For this purpose a RT real-time PCR was carried out to evaluate the production of pre-crRNA and mature crRNA.

RNA was isolated from wild type and transformed Ea6-4 after 2 and 4 h of incubation at 27°C. The RNA quality was verified on a Mini-PROTEAN® TBE-Urea Precast Gel (Bio-Rad). In order to separate pre-crRNA (>160 b) from crRNA (<160b), the microRNA Purification Kit (Norgen Biotek Corp.) was used; Figure 15 shows the separation of large RNA and crRNA. Then RT real-time PCR was used to evaluate the level of production of total CRISPR RNA (pre-crRNA and mature crRNA) and crRNA only.

After 2 h of incubation the production of CRISPR RNA was not changed in any Ea6-4 transformant compared to the wild type host. However, after 4 h of incubation, the total CRISPR RNA production increased by at least 13 fold in cells harbouring pSpCasg and by 3 fold in cells with pSpCas123 (Figure 16). On the other hand, the accumulation of crRNA in these hosts was quite different. In 2 h of incubation, the production of crRNA was slightly increased in cells with pSpCas123 and pSpCasg plasmids. After 4 h of incubation, the level of accumulation of mature crRNA increased only 12 fold in Ea6-4 harbouring pSpCasg. No increase was observed in other hosts and particularly in Ea6-4 that harboured pSpCas123 plasmid, which carries *Cas*1, 2 and 3 (Figure 16).
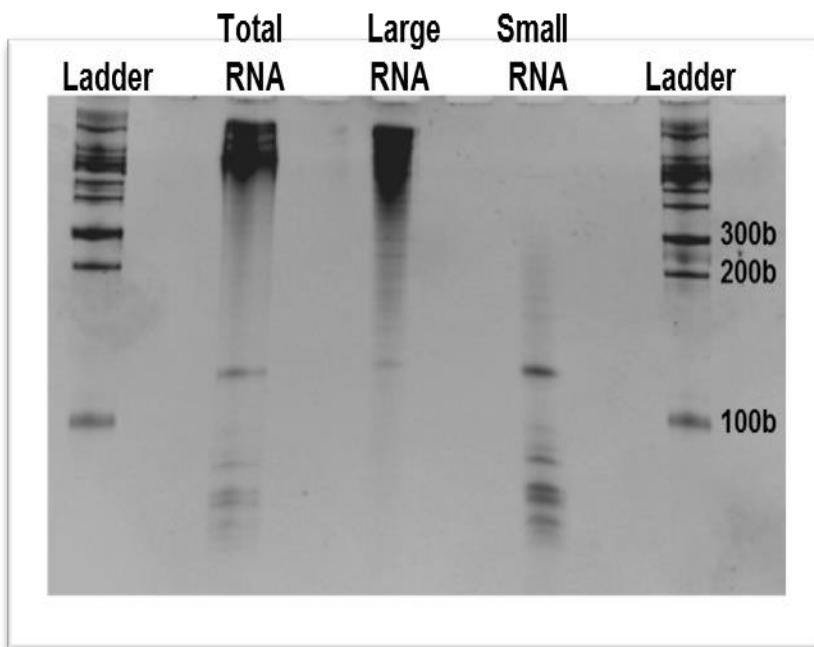
Figure 15: TBE-Urea gel of isolated RNA.

Total RNA before crRNA separation is shown in the second lane from the left. Small

RNA lane shows enriched RNA that is less than 160 bases, including crRNA. Large

RNA lane represents all RNA molecules larger than 160 bases including pre-crRNA.
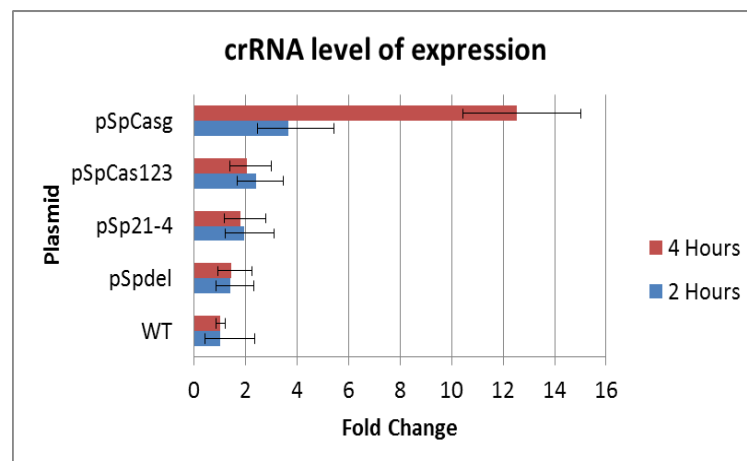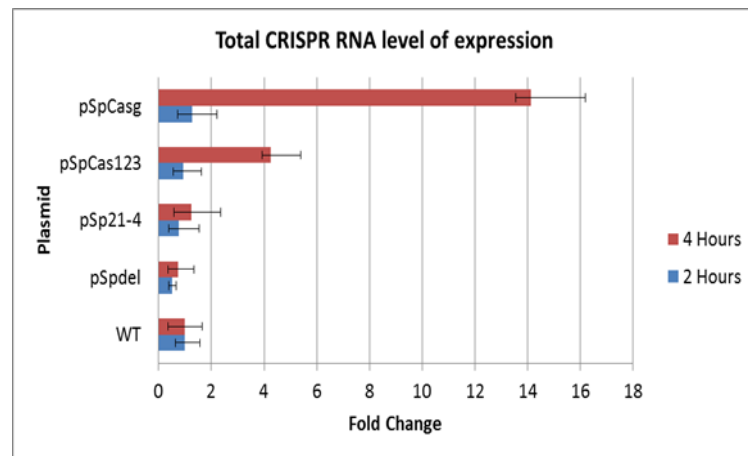
Figure 16: Pre-crRNA and crRNA level of production in Ea6-4 hosts transformed with plasmids carrying Cas genes.

Top Figure shows the level of production of total RNA (pre-crRNA and mature crRNA). Lower Figure shows the production level of only mature crRNA. The level of expression was measured after 2 and 4 h of incubation. Fold changes is calculated in comparison to wild type; the level of expression in wild type host is considered as one fold. The results were calculated using three independent biological replicate.

Phage replication in Ea6-4 transformed with plasmids carrying artificial spacers and Cas genes

Ea6-4 with pSpdel, pSpCas123 or pSpCasg was challenged with *Podoviridae* phages ΦEa31-3 and ΦEa9-2 and *Myoviridae* phages ΦEa35-70 and ΦEa21-4 to determine the effect of the extra copies of Cas genes on phage replication. The Ea6-4 host that transformed with pSpdel plasmid was used as a control since pSpdel carries no insert (no spacers and Cas genes). Transformed Ea6-4 isolates were grown in liquid culture until OD_{600} of 0.1 was reached (approximately 4 h). The cultures were infected with the listed phage at MOI of 0.1 and incubated for an additional 8 h at 27°C. Phage titers were determined every 2 h using a standard curve and real-time PCR. In all hosts, there was no difference in phage titre. Transforming Ea6-4 with artificial Cas genes in addition to existing endogenous CRISPR system did not reduce the replication of phages from *Myoviridae* and *Podoviridae* families (Figure 17A, B, C and D).
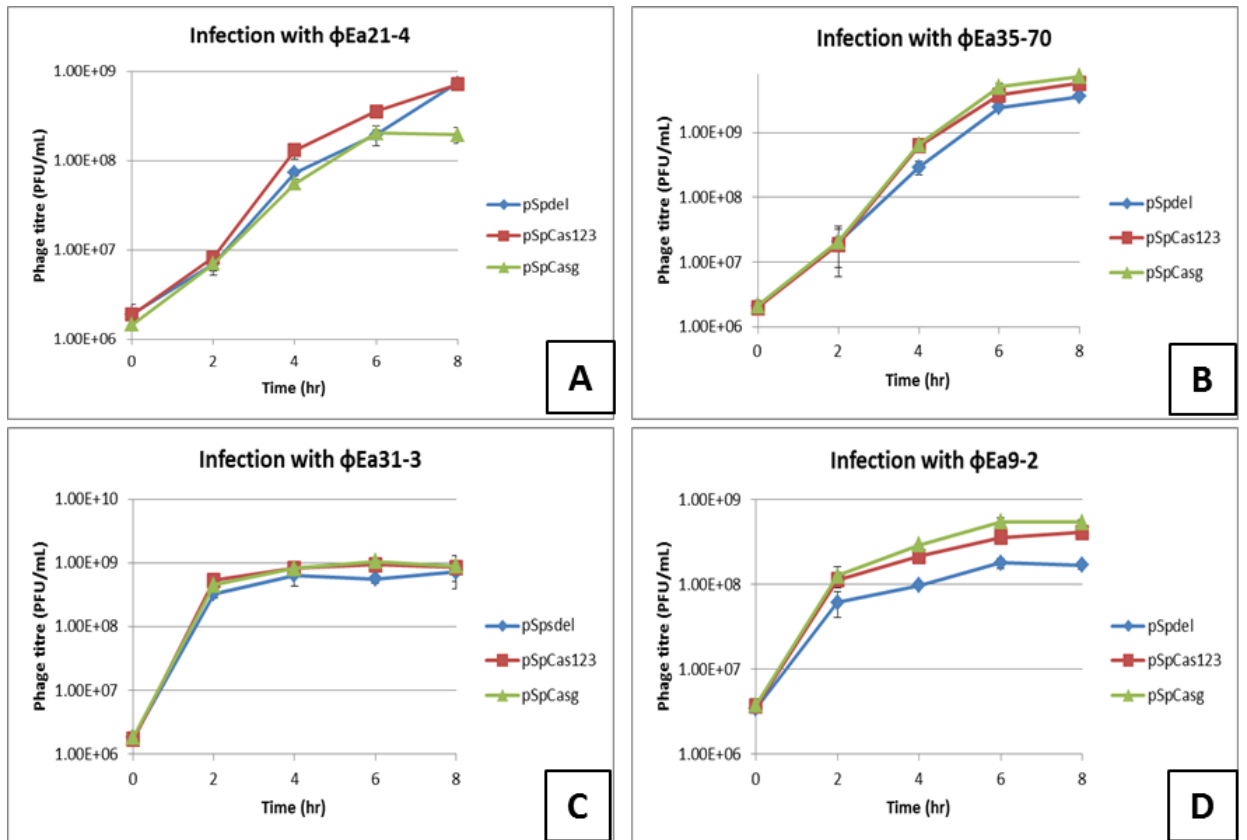
Figure 17: Selected *Myoviridae* and *Podoviridae* phages replication in Ea6-4

transformed with plasmids carry Cas genes.

A: phage ΦEa21-4; B: ΦEa35-70; C: ΦEa31-3; ΦEa9-2. Blue lines represent phages

titre in Ea6-4 transformed with the control plasmid pSpdel (no Cas genes). The results

were obtained using three independent biological replicate.

Generation of *E. amylovora* BIMs and spacer acquisition

In order to investigate the ability of *E. amylovora* to acquire new spacers in CRISPR arrays during phage infection, wild-type Ea6-4 and Ea6-4 transformed with pSpCas123 or pSpCasg were challenged with phages ΦEa31-3, ΦEa9-2, ΦEa35-70 and ΦEa21-4. Following three rounds of *E. amylovora* phage infection, bacterial colonies of Ea6-4 that survived phage infection were investigated for spacer acquisition using PCR on formed colonies. 273 *E. amylovora* colonies formed by cells that survived three rounds of phages infection were screened for spacer acquisition. Spacer acquisitions were not detected in any CRISPR array.

E. amylovora CRISPR/Cas system and plasmid transformation

The efficacy of *E. amylovora* CRISPR/Cas system in blocking plasmid transformation was investigated. Plasmids pProtoCR1, pProtoCR2 and pProtoCR3 were constructed to carry protospacers complementary to CRISPR spacers in CRR1, 2, or 3 of Ea6-4, respectively. Plasmid pProtoN that did not carry protospacers was used as a control (Figure 18).

Transformation efficiency of all plasmids was calculated per 1µg of plasmid DNA (Figure 19). If the plasmids carrying protospacers are blocked by the CRISPR system, low transformation efficiency should be expected comparing to control plasmid that carries random sequences that do not match spacers in Ea6-4. Transformation of plasmids carrying protospacers with similarity to CRR1 and CRR2 was blocked by one log compared to the control plasmid. The transformation efficiency of pProtoCR1 was approximately 4 times less than pProtoCR2 plasmid (Figure 19). The transformation efficiency of plasmid carrying protospacers with similarity to CRR3 was not considerably different from the control plasmid (Figure 19). Transformed plasmids that escaped

CRISPR system were investigated for mutations in protospacer sequences and site methylation. No mutations were detected in all plasmid, and plasmids were digestible with selected endonucleases that are known to be sensitive to methylation.
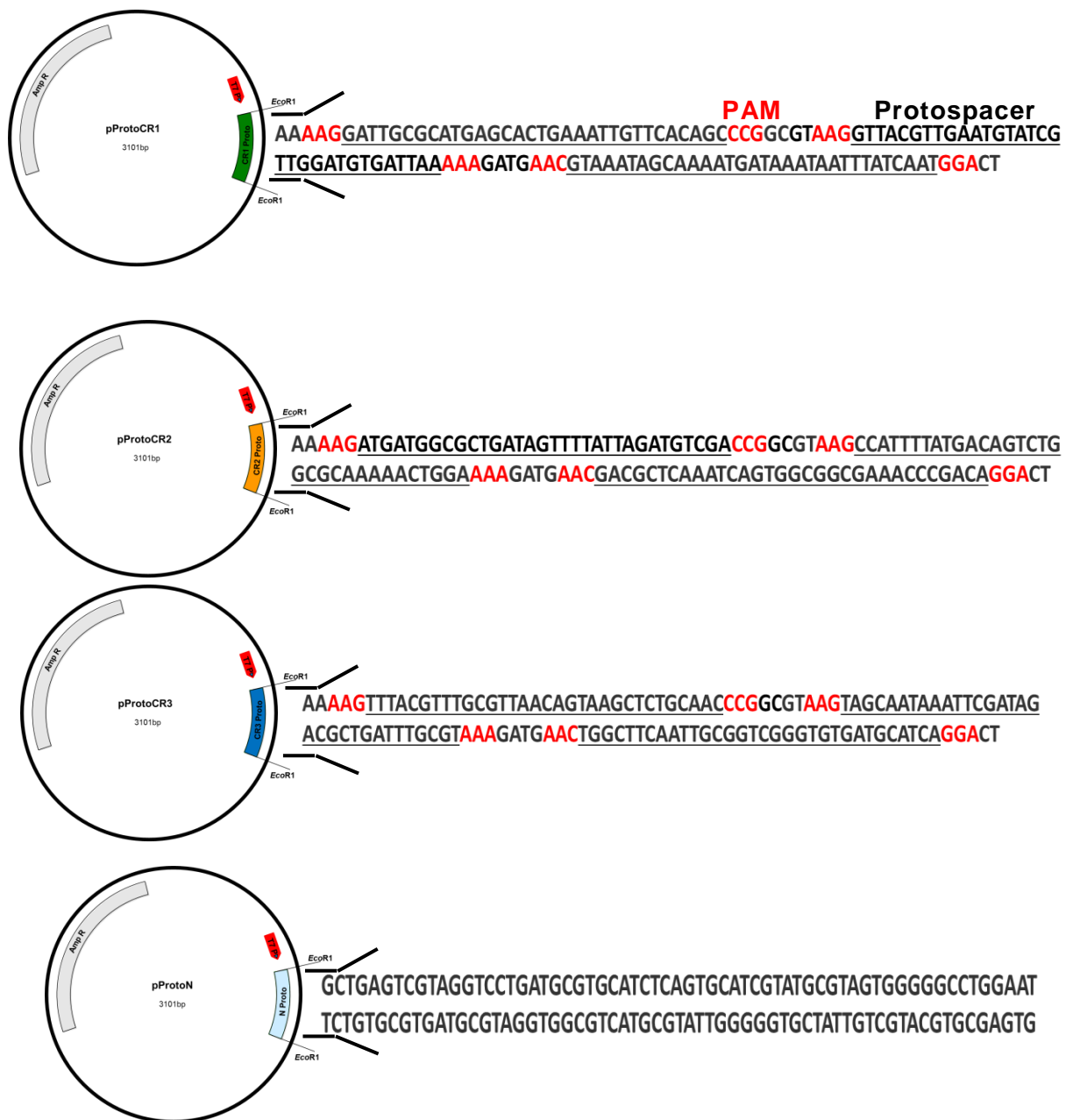
Figure 18: Plasmids used to challenge CRISPR/Cas system in Ea6-4.
pProtoCR1 plasmid carries protospacers targeting CRR1; pProtoCR2 plasmid carries protospacers targeting CRR2; pProtoCR3 plasmid carries protospacers targeting CRR3. pProtoN: control plasmid. PAM sequences are shown in red. Protospacers are underlined.
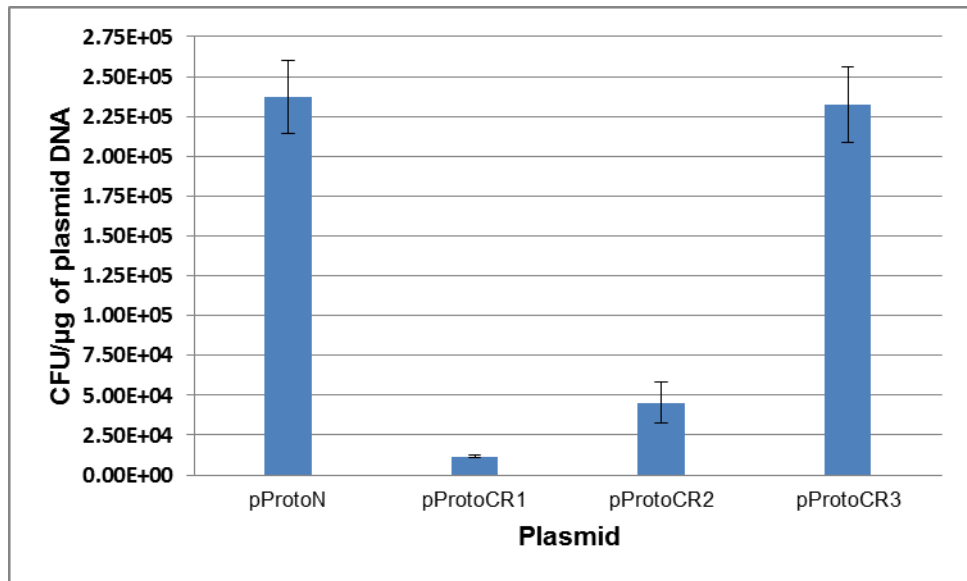
Figure 19: Challenging *E. amylovora* Ea6-4 with plasmids carrying protospacers. Transformation efficiency is expressed in CFU/μg of plasmid DNA. The results were calculated using three biological replicate.

Discussion

Sequences analysis of CRISPR/Cas system

CRISPR arrays sequences

The CRISPR/Cas systems of eight different *E. amylovora* isolates were investigated in this study (Table 1). Earlier, three CRISPR arrays have been identified in the genomes of previously sequenced *E. amylovora* strains, European strain CFBP 1430, isolated from a *Crataegus sp.* (Smits et al., 2010) and the American strain Ea273, isolated from a *Malus sp.* (Sebaihia et al., 2010). The existence of three CRISPR arrays was further confirmed in the genomes of many other *E. amylovora* strains from different geographical regions (McGhee and Sundin, 2012; Rezzonico et al., 2011). Even though the number of isolates included in this study was limited, similar results were obtained from the genomes of *E. amylovora* isolates from different locations in Canada and USA (Table 1).

E. amylovora CRISPR loci sequences showed that spacer content is mainly dictated by plant host and geographical location (McGhee and Sundin, 2012; Rezzonico et al., 2011). In this study, it seems that the influence of the host plant on spacer sequences is more apparent than the geographical location. The spacers of Ea6-4, EaD7 and Ea17-1-1 isolates from infected apple in Essex County, Ontario were found to be identical, despite that these isolates are characterised to be profoundly different based on the criteria of pathogenicity, response to phage infection and EPS production level (Lehman, 2007; Roach et al., 2011; Sjaarda, 2012). Ea110R, rif R mutant, originating from EA110 isolated from infected apple in Michigan USA, also carried similar spacers as Essex County isolates. The observation that isolates from different host plants exhibit variation in spacer sequences was further confirmed from the spacer sequences of CRR2 of EaG5. EaG5 was significantly different from isolates from the

same geographical location, and more related to BC29 from infected pear in British Columbia. Despite the genetic similarity between *E. amylovora* strains, the difference in spacer sequences was more evident between isolates from apple isolates and raspberry isolate Ea4-96. Previously, important differences in spacer sequences were detected between *E. amylovora* isolated from apple and from *Rubus* sp. (McGhee and Sundin, 2012; Powney et al., 2011; Rezzonico et al., 2011). *Rubus* strains are pathogenic on raspberry but not on apple and pear (Asselin et al., 2011; Triplett et al., 2006). The variation in the sequences of in CRR2 of all isolates in study might also indicate that this region is active and allows spacer acquisition upon the bacterial host facing foreign DNA elements. In conclusion, the diversity of CRR2 arrays seems to be mainly based on the host plant rather than on geographical region. Bacterial isolates that inhibit different environmental niche usually encounter different environmental components which may have shaped their evolution independently (Bolotin et al., 2005; Touchon and Rocha, 2010).

The origin of *E. amylovora* CRISPR spacers

Despite vast sequence variability of spacers detected in *E. amylovora* in this study, only a limited number of spacers was found to match phage genomes and plasmids. None of the detected spacers were found to exactly match the sequences of any *E. amylovora* phages in GenBank. Some spacers showed partial similarity to some *E. amylovora* phages from our collection, 13 to 17 nt match out of 32 nt spacer length (approximately 50% similarity), refer to Table 2. The lack of spacers that match phage sequences could be attributed to the small number of phage genomes that are available for comparison or more importantly to the inefficiency of *E. amylovora* CRISPR system.

PAMs are sequences flanking protospacers, and are important elements for phage resistance in at least some of the bacterial CRISPR/Cas systems (Deveau et al., 2008; Mojica et al., 2009). In CRISPR mediated immunity against phages of *S. thermophilus,* mutations in PAM sequences abolished CRISPR-mediated immunity even in the presence of perfect matching between spacer and protospacer (Deveau et al., 2008). Alignments of protospacers detected in this study revealed short conserved sequences at the '5 that could be PAMs (AAA, GAA, AAC, AAG and GTA) (Figure 10). These PAMs seem to be similar to those found in *E. coli* k12. *E. coli* K12, carries type I-E CRISPR/Cas system, contained a PAM consensus sequence 5`-AWG (Mojica et al., 2009). The majority of anti-plasmid spacers acquired by the host were a complement to the protospacers flanked with an AAG PAM. Less spacers were gained with AAA, AGG , GAG, TAG, CGA, AAT, and ATG PAM (Swarts et al., 2012).

E. amylovora CRISPR/Cas system and phage resistance

Earlier studies on *E. amylovora* CRISPR/Cas system have focused on the sequence analysis of Cas genes and CRISPR arrays, origin of spacers, phylogenesis and diversity of *E. amylovora* (McGhee and Sundin, 2012; Rezzonico et al., 2011). This is the first study that addresses the expression and role of *E. amylovora* CRISPR/Cas system in phage resistance and blocking plasmid DNA transformation.

In order to examine the response of *E. amylovora* CRISPR/Cas system to phage infection, the expression profiles of crRNA and *Cas*1 and 3 genes of Ea6-4 and EaD7 were explored during phage infection. *Cas*1 plays an important role in spacer acquisition, and *Cas*3 is involved in spacers acquisition and interference step. *Myoviridae* phages ΦEa21-4 and ΦEa35-70 and *Podoviridae* phage ΦEa9-2 were used to investigates the response of CRISPR/Cas system since these phages showed

variability in host range studies (Steven Gayder, PhD unpublished data). The *Myoviridae* phages did not elicit *E. amylovora* CRISPR system, and did not affect the level of expression of Cas genes during two cycles of phage replication. It was expected that invasion of phages to induce expression of the transcriptional regulators that, in turn, activate the transcription of genes involved in phage resistance mechanisms such as Cas genes. The expressions of Cas proteins and crRNA are likely to respond to phage or plasmid invasion as they function as a defense system for bacteria. The level of expression of *Cas1* and *Cas2* in archaea increased after virus infection (Liu et al., 2015; Quax et al., 2013). Similar results were shown by *S. thermophilus* (Young et al., 2012) and *Thermus thermophilus* (Agari et al., 2010). However, the expression level of several Cas genes and CRISPR arrays in *T. thermophilus* were not up-regulated during the first 100 min of phage infection (Agari et al., 2010). Even though the level of expression of Cas genes in *E. amylovora* in this study was monitored for 120 min, no up/down regulation was observed. It is important to note that phages used in this study were still in the eclipse phase up to 50 min post infection. Phage titer started to increase after 70 min post infection as an indication of the end of phages first replication cycle (data not shown).

In contrast to *Myoviridae* phages, an increase of Cas genes and crRNA level expression was observed with *Podoviridae* phages during the first 30 min in EaD7. However, the level of expression was down regulated 30 min post infection, an indication that the phage were taking over the host machinery. Infection of Ea6-4 with the same phage did not stimulate the level of expression of Cas genes. It is not clear whether the response of the CRISPR/Cas system to the *Podoviridae* ΦEa9-2 invasion is a result of challenging process of CRISPR/Cas system to phage infection, or reflecting an attempt of adaptation through spacer acquisition. EaD7 and Ea6-4 are

sensitive to phage ΦEa9-2 and other phages used in this study. No spacers were found in the CRISPR arrays of these isolates that perfectly match sequences in ΦEa9-2. As a result, one might speculate that the increase in Cas gene level of expression in EaD7 could be related to an attempt for a spacer acquisition process.

E. amylovora CRISPR/Cas system expression profile and response could be streamlined depending on the invading phage and infection mode of phages. The *Podoviridae* phages such as ΦEa9-2 are known to carry EPS depolymerase. The depolymerases facilitate phage entry by degrading the amylovoran, around the bacterial host allowing phages to reach their primary receptor on the cell membrane. Dissolving the EPS capsule creates stress on the bacterial host stimulating the bacteria defence mechanisms including CRISPR/Cas system. The bacterial cell membrane contains signal transduction apparatuses that are responsible for sending signals that respond to environmental stress which leads to change in the bacterial gene expression (Beier and Gross, 2006; Mole et al., 2007). EaD7 isolates produce a higher amount of amylovoran than Ea6-4 (Roach et al., 2013). Dissolving the EPS structure around EaD7 by the depolymerase of ΦEa9-2 might explain the difference in response of EaD7 and Ea6-4 to *Podoviridae* phage infection.

The mechanism by which *Myoviridae* phages are able to escape from or down regulate *E. amylovora* CRISPR/Cas system is not understood and has to be further investigated. The CRISPR/Cas system in *E. amylovora* may be non-functional during phage infection due to low level of Cas genes expression as the case in *E. coli* k12 (Westra et al., 2010). Also, *Erwinia* phages may be equipped with the tools to escape CRISPR/Cas system by down regulating Cas gene expression. Phage proteins with anti-CRISPR function against type I-F CRISPR/Cas and the type I-E CRISPR/Cas systems of *P. aeruginosa* have been identified (Bondy-Denomy et al., 2013; Pawluk et

al., 2014). The finding of anti-CRISPR genes carried by *P. aeruginosa* phages raised the possibility that other phage genomes may harbour similar genes. In this study, 17 *Erwinia* phages have been sequenced and analyzed (Chapter 4); no genes with similar sequences to anti-CRISPR genes of *P. aeruginosa* phages were detected in all known *E. amylovora* phage genomes. Nevertheless, temperate phages entering the lysogenic cycle were able to down regulate the *P. aeruginosa* type I-F CRISPR/Cas system. These anti-CRISPR genes appear to function to permit and establish prophage existence (Bondy-Denomy et al., 2013). To date, no temperate *E. amylovora* phages have been found in *E. amylovora* (Roach et al., 2015).

Spacer acquisition by bacteria upon phage infection has been shown in several bacteria (Erdmann and Garrett, 2012; Erdmann et al., 2014). During phage infection CRISPR loci of *S. thermophilus* acquired new spacers derived from the invaded phage DNA (Barrangou et al., 2007). The acquisition of these spacers led to a high level of resistance to specific phages bearing homologous sequences in subsequent infection (Barrangou et al., 2007; Deveau et al., 2008). *Streptococcus* mutants, the principal aetiological agent of dental cavities, were challenged with phage M102 to study phage therapy for the prevention of tooth decay in rats. BIMs from *Streptococcus* were isolated that had added a new spacer that matched sequences in phage M102 (van der Ploeg, 2009). New acquired spacers are usually inserted in the CRISPR arrays just upstream of the leader sequence and before the first repeat in the cluster (Barrangou et al., 2007; Deveau et al., 2008; Horvath et al., 2008).

In this study, three sets of BIMs were generated using *E. amylovora* isolates wild-type Ea6-4 and Ea6-4 transformed with pSpCas123 or pSpCasg plasmids. Bacterial cells were infected with phages from *Myoviridae* and *Podoviridae* families. CRISPR regions 1, 2 and 3 of all BIMs were investigated for spacer acquisition. No

insertions of new spacers were detected in all CRISPR arrays of the generated BIMs

indicting that the gained phage-resistance feature of these BIMs was not due to spacer

acquisition. The overall mechanism of spacer acquisition and the synthesis of new

repeats are not well understood; however Cas1 and Cas2 play a role. The

overexpression of both *Cas*1 and *Cas*2 in type I systems facilitate spacer acquisition

indicating that both gene are required for this process (Yosef et al., 2012). The

CRISPR/Cas system of *E. coli* K12 gained new phage-derived spacers in many cells at

leader-proximal end of CRISPR following the increase of *Cas* level of expression

(Datsenko et al., 2012). In this study, enhancing in *E. amylovora Cas* expression seem

to have no effect on spacer acquisition. These results suggested that *E. amylovora*

CRISPR/Cas system is not cable of spacer acquisition despite boosting it with extra

copies of Cas genes.

The effect of over expression of Cas genes on phage infection

To better understand the CRISPR/Cas system in *E. amylovora*, the function of

CRISPR system was investigated in-depth by enhancing the level of expression of Cas

genes by introducing an extra set of Cas genes cloned into plasmids. The *E. coli*

CRISPR/Cas system is found to be transcriptionally repressed by the H-NS regulator

protein that strongly inhibited Cas gene promoters (Pul et al., 2010; Westra et al.,

2010). Deletion in hns gene caused an increase in the expression of some of the Cas

genes and increased abundance of small crRNA (Hommais et al., 2001; Pougach et al.,

2010). The approach of mutating hns gene to increase the level of expression of Cas

genes was not the choice in this study. Instead, an artificial CRISPR/Cas system cloned

into a plasmid next to T7 promoter was the chosen tactic. When a plasmid carrying hns

gene was introduced into in *E. amylovora*, the capsular EPS production was reduced

(Hildebrand et al., 2006). Also a mutation in chromosomal hns gene of *E. amylovora*

increased amylovoran production (Hildebrand et al., 2006). Inducing EPS production in

E. amylovora might affect phage adsorption while investigating CRSIPR/Cas system.

Cas genes in *E. coli* K12 provided resistance against λ phage when they co-

overexpressed from a plasmid but only when the cell harbors spacers matching the λ

phage genome is transcribed by T7 RNA polymerase (Brouns et al., 2008). A similar

strategy was carried out in this study; a pSpCasg plasmid was used to enhance

CRISPR/Cas system this plasmid carried all Cas genes found in *E. amylovora*

CRISPR/Cas system (*Cas*1, 2, 3, 4, 5 and *Cse*1, 2, 3, 4) in addition to three spaces

targeting sequences in phage ΦEa21-4 (Figure13).

It was important to investigate the level of expression of Cas genes and crRNA in

E. amylovora transformed with plasmids carried Cas genes prior to challenging with

phages. The level of expression of *Cas*1 and *Cas*3 increased in all hosts that were

transformed with a plasmid carrying a copy of *Cas*1 (Figure14). Interestingly, the level

of expression of *Cas*3 in a host harbouring pSpCas123 was almost two fold less than

pSpCasg (Figure 14). The apparent increase in the expression level of *Cas*3 in the

isolate carrying pSpCasg could be a result in the presence of other Cas genes such as

*Cse*1, 2, 3, and 4 in this plasmid. Cas genes were found to regulate each other; the

overexpression of *Csa*3a in type I-A of the archaeon *Sulfolobus islandicus* increased

the transcriptional level of Cas1 and 2 genes (Liu et al., 2015).

In addition to *Cas*1 and *Cas*3, a high increase in total CRISPR RNA (pre-crRNA

and crRNA) production in hosts harbouring pSpCasg was observed. The accumulation

of total CRISPR RNA (pre-crRNA and crRNA) was less in cells carry pSpCas123

(Figure 16). Interestingly, the amount of processed mature transcript (crRNA) was also dramatically increased in cells harbouring pSpCasg comparing to other hosts. These results indicate that the noticeable increases in crRNA abundance only took place when all Cas genes were overexpressed. The increase of crRNA abundance was most likely due to the increase in intracellular concentration of all Cas genes which help in the expression or stabilizing the produced crRNA. Overexpression of only *Cas*1, 2 and 3, in hosts carrying pSpCas123, was not sufficient to increase pre-crRNA processing. However, the overexpression of *CasE* (*Cse3*) in *E. coli* was sufficient for the increase in the accumulation *of* crRNA (Pougach et al., 2010). The effect of absent of some Cas genes (*Cse*1, and *Cse*2, *Cse*4, *Cas*5 and *Cse*3) on the function of CRISPR region in *E. coli* K12 was investigated (Brouns et al., 2008). Only Cse3 protein is required for pre-crRNA cleavage to produce mature crRNA (Brouns et al., 2008). In conclusion, the increase in processing mature crRNA in hosts carrying pSpCasg indicates that Cas genes of *E. amylovora* are functional. Also other genes rather than *Cas*1, 2, and 3 are needed for pre-RNA processing and may be stabilizing the produced crRNA.

Cas genes in *E. coli* K12 provided resistance against λ phage when they were co-overexpressed from a plasmid but only when the cell harbors spacers matching the λ phage genome (Brouns et al., 2008). Eliminating the Cas gene plasmid from the host permitted λ phage infection suggesting that the levels of expression of chromosomal Cas genes are not enough to deliver phage resistance (Brouns et al., 2008). This was not the case in *E. amylovora*; this work shows that introducing plasmids which carry Cas genes and spacers matching ΦEa21-4 did not affect phage replication when compared to the control (Figure 17). In addition, transformed Ea6-4 carrying extra Cas genes did not reduce the replication of *Podoviridae* phages ΦEa31-3 and ΦEa9-2 and *Myoviridae* phage and ΦEa35-70 (Figure 17). Phages can evade CRISPR immunity by

mutating residues in target sequence (protospacer) or around the target sequence in PAM (Deveau et al., 2008; Semenova et al., 2009). Sequence analysis of 19 phages escaped CRISPR immunity in *S. thermophilus* were found to contain a single or double mutation in the protospacer sequence (Deveau et al., 2008). Mutation in protospacers sequences should not be the reason for the absence of resistance in Ea6-4. PSpCasg and pSpCas123 carry three spacers that target three different genes in ΦEa21-4 genome; mutation in three protospacers at once is doubtful.

In conclusion, the CRISPR/Cas system in *E. amylovora* is inefficient in preventing phage infection of the host. Overexpression of Cas genes and enhancing the production of crRNA, the surveillance and interfering tools, also did not enable CRISPR/Cas system in *E. amylovora* to block phage infection. The CRISPR/Cas system appears to be inactive or partially ineffective in *E. amylovora*. *E. amylovora* may have maintained the existence of an intact CRISPR/Cas system for some beneficial function to the host other than host phage resistance.

CRISPR/Cas system efficiency in blocking plasmid transformation

To establish *E. amylovora* CRISPR/Cas system efficiency in blocking plasmid DNA transformation, four target plasmids were constructed to challenge CRISPR/Cas system of Ea6-4. The partial blocking of pProtoCR1 and pProtoCR2 plasmids, which target CRR1 and CRR2, respectively, might reflect the moderate efficacy of these regions in preventing DNA transformation. The transformation efficiency of pProtoCR3 plasmid which targets CRR3 was not affected when compared to the control. It could be concluded that CRR3 is inactive in comparison to CRR1 and CRR2.

In conclusion, the detection of CRISPR/Cas system in almost 50% of bacteria indicates its significance as defense mechanism against phages and other exogenous elements. However, the *E. amylovora* CRISPR/Cas system seems to have no role in blocking phage infection. Improving the level of expression of Cas genes and inserting spacers to target certain phages also did not enable this system to function as a defences mechanism. Spacers that already exist in *E. amylovora* CRISPR loci might indicate that this system is involved in processing and acquisition of new spacers but not in defending the bacterial host against phages. In this study, it was shown that *E. amylovora* CRISPR/Cas system appears to be more effective in blocking plasmid DNA transformation than preventing phage infection. This might suggest that *Erwinia* phages are capable of escaping CRISPR/Cas system using unknown mechanism.

Chapter 3

Sequences analysis of selected *E. amylovora* phages

Abstract

Limited numbers of *E. amylovora* phage genomes are sequenced in GenBank. Sequencing of more phage genomes will improve our understanding on the relationship between *Erwinia* spp. phages and the CRISPPR/Cas system of their host. The complete genomes of seventeen *E. amylovora* phages from the *Podoviridae* and *Myoviridae* families were sequenced and analyzed in this study. The genome of *E. amylovora* phages is composed of double-stranded linear DNA with a length from 45 to 280 kbp. Bioinformatics analysis shows that *Erwinia* spp. phages ΦEa10-1, ΦEa10-2, ΦEa10-4, ΦEa10-16, ΦEa21-2, ΦEa35-2, ΦEa45-1b, and ΦEa51-7 are 99.9% similar to the published *Myoviridae* phage ΦEa21-4. Phages ΦEa10-6, ΦEa10-7, ΦEa31-3, ΦEa35-3, ΦEa35-10 ΦEa46-1-A-1, and ΦEa46-1-A-2 were found to be very similar to the *E. amylovora Podoviridae* phage ΦEa1h that was previously published. In all sequenced phages, no protospacers were found to match spacers determined in all *E. amylovora* isolates CRISPR system from this study. Also no anti-CRISPR proteins similar to proteins found in *P. aeruginosa* were detected in the genome of all sequenced phages. The genomes of *Podoviridae* phage ΦEa9-2 and the *Myoviridae* phage ΦEa35-70 were found to be unique when compared to the known phage genomes in GenBank. The genome of ΦEa35-70 is 271,084 bp and it encodes 318 putative proteins and one tRNA. Comparative analysis with other *Myoviridae* genomes suggests that ΦEa35-70 is peripherally related to the Phikzlikevirus genus within the *Myoviridae* family. Twenty six percent of the proteins are homologs with *Pseudomonas* phage ΦKZ. The genome of ΦEa9-2 is 75,568 bp in size. The sequence analysis shows that ΦEa9-2

is related to coliphage N4 in that their genomes are syntenic and they share 50 protein homologs (69% identity). Out of 94 CDSs encoded only 23 CDSs were annotated as functional genes encoding structural proteins and proteins involved in nucleic acid modification, an indication to distinctiveness characteristics of this phage. The data extracted from the sequence of these phages expands our knowledge of *Erwinia* spp. phages and their relation to their bacterial host.

Introduction

     E. amylovora phages were first isolated in the 1960s (Billing, 1960; Hendry et al., 1967). The characteristics and properties of *E. amylovora* phages were further described in the 1970s (Erskine, 1973; Ritchie, 1978; Ritchie and Klos, 1979). In recent years a number of *E. amylovora* phages were isolated and sequenced (Born et al., 2011; Boulé et al., 2011; Gill et al., 2003; Lagonenko et al., 2015; Lehman et al., 2009; Müller et al., 2011b; Schnabel and Jones, 2001; Yagubi et al., 2014). *E. amylovora* phages are tailed, double-stranded DNA bacterial viruses that belong to the order *Caudovirales* (Gill et al., 2003; Lehman, 2007; Meczker et al., 2014; Schnabel and Jones, 2001; Yagubi et al., 2014). They are classified into three families: *Myoviridae*, *Siphoviridae*, and *Podoviridae* (Gill et al., 2003). Phages belonging to the *Myoviridae* have contractile tail, *Siphoviridae* have non-contractile tail, and the *Podoviridae* have short tail (Figure 2, Chapter 2). Gill et al. (2003) isolated 42 *E. amylovora* phages mainly from the soil underneath pear and apple trees exhibiting fire blight symptoms in southern Ontario, Canada. These phages were classified into 6 groups based on molecular characterization using PCR, restriction fragment length polymorphisms (RFLPs) and morphology.   Group 1 phages belonged to the *Myoviridae*, Group 2 to the *Siphoviridae* and Groups 3 to 6 were members of *Podoviridae* family (Gill et al., 2003). *E. amylovora* phages have been further studied based on host range, RFLPs (Barionovi et al., 2006; Gill, 2000) and genomic sequence similarities (Born et al., 2011; Lagonenko et al., 2015; Lehman et al., 2009; Meczker et al., 2014; Yagubi et al., 2014).

     Classification of phages based on plaque morphology, restriction map analysis, and transmission electron microscopy has limitations. Using restriction endonuclease analysis for phage classification comes with constraints especially in case of phages

with large genomes and methylated restriction endonuclease sites. Classification of phages based on phage morphology does not discriminate phages belong to the same family, and plaques formed on *E. amylovora* LEP hosts are often faint and difficult to visualise.

Phage genome sequences play a critical role in phage taxonomy and classification. Genomic analysis of phages used as biological control agents needs be carried out in order to expand the knowledge about their evolution and genetic structure. Knowledge of phage genome structure is critical to our understanding of the biological relationship between phages and their bacterial hosts, and for phage detection in the field based biological control trials. Phage genome sequencing enhances our knowledge about the performance of these phages and identification of undesirable genes such as toxin encoding genes and lysogeny-associated genes and/or genes related to different phage resistance mechanism.

The objective of this study is to determine the genome sequence of 17 phages from the *Myoviridae* and the *Podoviridae* families. These phages were chosen from the AAFC Vineland Collection (Table 4), and were selected on the basis of phage family, host range and efficacy as biological control agents. It was important to fully characterize all the phage isolates in the collection on the molecular level in order to be able to quantify and keep track of individual phages in bioassays and field based studies using qPCR protocols (Lehman, 2007). Availability of more information on the sequence of phage genomes will help in the detection of protospacers that might match spacer sequences in *E. amylovora* CRISPR arrays determined in this study.

Materials and Methods

Phage isolation and purification

To obtain a homogenous phage solution which originated from a single plaque, phage isolate were serially diluted and plated using the double agar overlay method (Adams, 1959). The dilution containing low numbers of similar plaques was used to pick off a single plaque using a sterile 1 mL micropipette tips. The plaque was added to an Eppendorf tube that contained 1 mL of sterile NB. The sample was vortexed briefly to release phages from agar. To remove bacterial and agar residues, the phage suspension was filter-sterilised through a 0.45 µm filter. To ensure a single phage-strain population, second and third rounds of phage purification were carried out by repeating the above described protocol three times. Real-time PCR using primers and probes targeting a specific phage family were used to confirm the purity of the phage solution.

Phage liquid culture

Bacterial starter cultures were prepared by inoculating 2 mL of 0.8% NB with a loop full of *E. amylovora* obtained from overnight streaked NA plates. The 2 mL inoculated cultures were shaken for 1 h at 150 rpm at 27°C. Then 25 mL of 0.8% NB was inoculated with 0.5 mL of the starter culture and placed on an orbital shaker at 27°C until the OD_{600} reached 0.5. This resulted in a bacterial concentration of approximately 10^8 CFU/mL. The bacterial culture was infected with the desired phage to an MOI of 0.1 or 1.0. The culture was further incubated and shaken at 150 rpm for 16 h at 27°C. Cell debris were removed by centrifugation at 8000 xg for 25 min at 4°C, the supernatant was filtered through 0.2 µm syringe filter and collected into 50 mL dark glass tubes. Phage titre was determined by plaque assay and the phage stock was

stored at 4°C following the addition of a drop of chloroform to avoid bacterial

contamination.

Double agar overlay method

This method was fully described by Adams (1959). *E. amylovora* from overnight

cultures was resuspended in 0.01 M PB; pH 6.8 to give 1 x 10^8 CFU/mL. Then 100 µL

of resuspended host was mixed with phage in test tube at MOI of 0.1 or 1. The mixture

was incubated at RT for 15 min; 4 mL of molten top agar at 55°C was added to the

phage/host mixture. Tubes were swirled, and the mixture was poured onto a previously

prepared plate containing solidified NA. The top agar was allowed to solidify before

placing the plate in a 27°C incubator for 18 h. Phage plaques were flooded with 5 to 10

mL of 0.8% NB and the liquid was collected into a centrifuge tube. Bacterial debris and

agar were removed by centrifugation for 25 min at 8000 xg. The supernatant was

filtered through 0.45 µm filter and placed into 50 mL dark glass tubes.

Phage concentration by PEG precipitation

One hundred milliliters of overnight *E. amylovora* infected with phage was

centrifuged at 6000 xg for 20 min at 4°C to remove cell debris. The supernatant was

filtered using 0.45 µm bottle top filter. To concentrate phage particles, 5.84 g of NaCl

(Fisher Scientific, Fair Lawn, NJ, USA) was added to 100 mL of culture and swirled at

RT until the salt was dissolved in order to make a final concentration of 1 M NaCl. 10 g

PEG 8000 (Fisher Scientific) was added to make a 10% (w/v) final concentration. The

PEG was dissolved by stirring slowly on a magnetic stir plate at RT. The mixture was

incubated overnight at 4°C in order to precipitate the phage. To recover the precipitated

phage following overnight incubation, the mixture was centrifuged for 15 min at 11,000

xg. The supernatant was removed and the phage pellet was slowly and gently

resuspended in 10 mL of SM buffer (50 mM Tris-HCl, pH7.5; 0.1 M NaCl; 8 mM

MgSO4, pH7.4). The PEG and NaCl were removed by adding and mixing the phage

suspension with 2 volumes of chloroform. The mixture was then centrifuged at 5,000 xg,

for 10 min at 4°C. The upper aqueous layer (containing phage) was transferred to a

fresh tube, and further extracted with 2 volumes of chloroform.

Phage titre

Phage titer using plaque assay and real-time PCR methods were described in materials

and methods in Chapter 3.

Phage DNA extraction

Two methods were used to extract phage DNA, extraction and purification using

the Norgen Phage DNA Isolation Kit (Norgen Biotek Corp.) or using the

phenol/chloroform. In the phenol/chloroform protocol, 2 mL of concentrated phage

suspension was first subjected 1 μL of DNase I (Roche) (1mg/mL) and 1 μL of RNase A

to remove any bacterial genomic DNA and RNA. SDS was then added (0.5% final

concentration) and proteinase K (5 μg/mL final concentration), followed by incubation

for 15 min at 65°C. Phage DNA was then extracted by adding an equal volume of

phenol:chloroform:isoamyl alcohol (25:24:1) saturated with 10 mM Tris, pH 8.0. The

aqueous layer containing phage DNA was collected after centrifugation at 6000 xg for 5

min at room temperature. The phage DNA was precipitated by adding 2X volumes of

95% ethanol and centrifugation for 15 min at 15,000 xg, 4°C. The DNA pellet was air

dried and resuspended into TE buffer. The DNA quality and concentration were tested on 0.9 % agarose gel, and by using spectrophotometer at OD_{260} before storing at -20°C.

Phage DNA sequencing analysis and assembly

A DNA library was constructed using an Illumina Nextera XL kit with an average insert size of 800 bp, and this was sequenced by the McGill University and Genome Quebec Innovation Centre (Montreal, Canada) on an Illumina MiSeq. The phage genome was assembled using Seqman NGen 3 (DNASTAR, Madison, WI, USA). Contigs greater than 1kb were exported, reassembled into a single contig and proofread, using Seqman Pro (DNASTAR). The average coverage for the phage was 713X. The final contig was oriented in the same manner as its nearest homolog (previously sequenced phages). The genome was autoannotated with MyRAST (Overbeek et al., 2014), with final visual analysis in Kodon (Applied Maths, Austin, TX, USA) where the annotation was checked for missed coding sequences (CDSs) and proper starts sites. tRNA genes were identified using tRNAscan-SE 1.21 (Schattner et al., 2005). To determine the level of relatedness to other phages TBLASTX (e-value $\leq 10^{-5}$) was used to compare the predicted protein sequences to the genome sequences of all other phages in NCBI. Phages sequenced in this study are listed in Table 4.

Phage	Genome size in bp	Family
φEa10-1	84568	*Myoviridae*
φEa10-2	84575	
φEa10-4	84666	
φEa10-16	85534	
φEa21-2	84685	
φEa35-2	84581	
φEa45-1b	84651	
φEa51-7	84593	
φ Ea35-70	271085	
φEa10-6	45271	*Podoviridae*
φEa10-7	45177	
φEa31-3	45323	
φEa35-3	44739	
φEa35-10	45151	
φEa46-1-A-1	44923	
φEa46-1-A-2	45138	
φEa9-2	75588	

Table 4: List of phages sequenced in this study.

The size of phage genome is shown in bp. *Myoviridae* phages (white box) are similar to *E. amylovora* phage ΦEa21-4 (Lehman et al., 2009). *Podoviridae* phages (gray box) are similar to *E. amylovora* phage ΦEa1h (Müller et al., 2011a; Ritchie, 1978; Ritchie and Klos, 1977). Phages ΦEa9-2 and ΦEa35-70 have no matches in GenBank.

Results and Discussion

The genomic sequence of 17 phages in this study revealed that 15 of the phages have notable relatedness to previously sequenced phages. Phages ΦEa10-1, ΦEa10-2, ΦEa10-4, ΦEa10-16, ΦEa21-2, ΦEa35-2, ΦEa45-1b, and ΦEa51-7 belong to *Myoviridae* family, and are similar to ΦEa21-4 (Lehman et al., 2009). All these phages were included in Group 1 of Gill's classification based on RFLPS and Transmission electron microscopy (TEM) morphology. On the other hand, *Podoviridae* phages belonging to Gill Group 3A: ΦEa10-6, ΦEa10-7, ΦEa31-3, ΦEa35-3, ΦEa35-10 ΦEa46-1-A-1, and ΦEa46-1-A-2 were found to be similar to *E. amylovora* phage ΦEa1h (Müller et al., 2011a; Ritchie, 1978; Ritchie and Klos, 1977). Two phages namely, ΦEa9-2 and ΦEa35-70, were shown to contain unique genome sequences and gene functions when comparing to sequenced phages in GenBank.

Characteristics and morphological features of ΦEa35-70 and ΦEa9-2

Phages ΦEa9-2 and ΦEa35-70 were isolated from soil samples collected under infected pear trees in southern Ontario, Canada (Gill et al., 2003). ΦEa9-2 and ΦEa35-70 are tailed phages with icosahedral heads. TEM revealed that ΦEa35-70 and ΦEa9-2 to be members of the *Myoviridae* and *Podoviridae* families, respectively. ΦEa35-70 is a large phage with a head size of approximately 150 nm across; and a long contractile tail of approximately 200 nm in length (Figure 20).
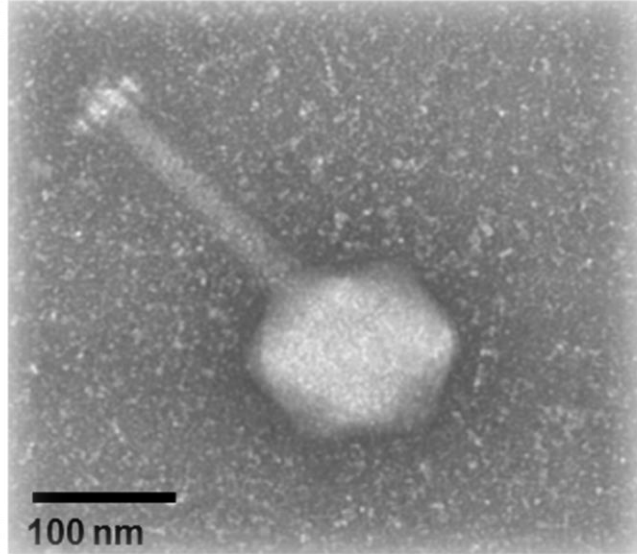
Figure 20: Electron micrograph of ΦEa35-70 with an icosahedral head and extended tail.

Phage was stained with uranyl acetate before visualizing under electron microscope at the integrated microscopy unit in the Biotron, Western University, London, Ontario.

Protospacers in ΦEa35-70 and ΦEa9-2

One of the main purposes of sequencing this group of *E. amylovora* phages was to detect the presence of protospacers with similarity to spacers determined in the CRISPR system of *E. amylovora* isolates in this study (Chapter 3). Investigating ΦEa35-70 and ΦEa9-2 genomes revealed that there were no protospacers in these two phages genomes that perfectly match spacers found in the CRISPR arrays of *E. amylovora* isolates. Only 11 spacers from isolate EaG5 and Ea6-4 were found to have 14 to 15 nucleotide matches (average 48% similarity) to sequences in ΦEa35-70 (Table 5). Among all determined spacers in this study, only spacer 29 in CRR1 of Ea6-4 was found to have 17 nucleotide match (53% similarity) to sequences in a gene that encodes for a hypothetical protein in ΦEa9-2 genome (Table 5).

It is not clear if 50% similarity (16 nt out of 32 nt spacer length) between spacers and their target in phage DNA is sufficient to cause interference by CRISPR system. It has been shown that the seed sequence of the spacers (6 to 10 nucleotides) is the most important region in CRISPR interference (Semenova et al., 2011; Wiedenheft et al., 2011a; Wiedenheft et al., 2011b). Mutations in the spacer sequence out of the seed sequence didn't affect crRNA binding to target; however, a single nucleotide mutation within seed sequence resulted in failure to bind to target DNA (Semenova et al., 2011). It is not confirmed if these sequences found in phage genomes that matched spacer sequences are located within seed sequences.

Spacer sequence	Host	CRISPR region	Spacer no.	Phage	Location on phage genome	Target phage gene
TTATT<u>CATGAGCC</u>**T**<u>TTTTATC</u>TTCGCGGCATG	Ea6-4	1	1	φEa35-70	177033 to 177047	Hp
AAGATG<u>CTTTGACATTAATTA</u>TCTCCATAAAA	Ea6-4	1	5	φEa35-70	38341 to 38355	Hp
<u>AAGATGCTTTGACA</u>TTAATTATCTCCATAAAA	Ea6-4	1	5	φEa35-70	54602 to 54615	Hp
AAGATGCT<u>TTGACATTAATTATC</u>TCCATAAAA	Ea6-4	1	5	φEa35-70	73984 to 73997	Hp
CCGCGA<u>AAATC</u>**C**<u>GCAG</u>**T**<u>GAGCTGG</u>CAATGAGC	Ea6-4	1	29	φEa9-2	63339 to 63357	Hp
GT<u>TGAT</u>**A**<u>CGGCTGATTACA</u>ATAAAATGTCACT	Ea6-4	2	9	φEa35-70	254734 to 254750	Tail fiber protein
GCTACTAC<u>GTGTACGCAC</u>**A**<u>GCCGC</u>TGGCCAGT	Ea6-4	2	14	φEa35-70	100839 to 100855	Hp
CAACTG<u>AGCAACTCC</u>**G**<u>CTTTGCCC</u>CATACCAA	EaG5	2	15	φEa35-70	72650 to 72666	Hp
CGACTATTCG<u>GCGAAC</u>**T**<u>GAAAGAGCCA</u>ACCGT	EaG5	2	18	φEa35-70	104335 to 104351	Hp
GAGA<u>GCAACAACAA</u>**A**<u>CTCC</u>GACCGGAGAAACC	EaG5	2	23	φEa35-70	3070 to 3084	Hp
<u>ACGGT</u>**C**<u>AGATGGTGGCG</u>CTGGTTGCGCTGGCA	EaG5	2	39	φEa35-70	22206 to 22222	Hp
ATCGCACCC<u>CACTGATTGAAGA</u>GCAGCACACT	Ea6-4	3	2	φEa35-70	238369 to 238381	Hp

Table 5: Spacers from EaG5 and Ea6-4 with match to sequences in ΦEa35-70 and ΦEa9-2.

Sequences within each spacer with similarity to sequences in phages are underlined. Mismatches are in bold and not underlined. CRSIPR regions where the listed spacers are located are shown. The target phage genes by these spacers are shown in the most right column; Hp, hypothetical protein. The location of protospacers on phage genomes is shown.

General features and annotation of the ΦEa35-70 genome

The ΦEa35-70 genome is quite different from any other phage in GenBank. The complete genome sequence of ΦEa35-70 has been submitted to GenBank and assigned accession number KF806589.1. The genome of ΦEa35-70 is 271,084 bp long with a G+C content of 49.9%. To determine the level of relatedness of ΦEa35-70 to other phages TBLASTX, with default settings, was used to compare the predicted protein sequences to the genome sequences of all other phages in NCBI. Comparative analysis with other phage genomes suggests that ΦEa35-70 is related to the Phikzlikevirus genus within the *Myoviridae* since 26% of ΦEa35-70 proteins share homology to proteins in *Pseudomonas* ΦKZ (Mesyanzhinov et al., 2002). Bacteriophage ΦKZ has a linear genome and 280,334 bp in size (Mesyanzhinov et al., 2002). ΦEa35-70 is also peripherally related to *Pseudomonas* phage phiPA3 (Monson et al., 2011) and *Erwinia* phage PhiEaH1 (Meczker et al., 2014), refer to Tables 6 and 7 for list of phages with similarity to ΦEa35-70.

The genome of ΦEa35-70 contains 318 CDSs of which 284 CDSs were annotated as encoding hypothetical proteins. 69% of the hypothetical proteins are unique to this phage and have no matches in GenBank. The orientation of genome annotation was chosen so that the majority of genes are encoded by plus strand; 302 CDSs on plus strand and only 16 on minus strand. Only 34 open reading frames (ORFs) were annotated as functional genes encoding for proteins involved in nucleic acid modification and replication and phage DNA packaging. Most of identified genes in ΦEa35-70 were found to be involved in nucleotide metabolism, such as thymidylate kinase, dihydrofolate reductase, and diphosphate reductase ribonucleoside. A putative HNH endonuclease which is involved in phage DNA packaging and common in double-stranded DNA phages was also detected. Interestingly, an EPS-depolymerase gene,

involved in EPS degradation, and gene similar to Concanavalin A-like lectin was also identified. The existence of EPS-depolymerase gene in ΦEa35-70 is interesting since *E. amylovora Myoviridae* phages usually lack this gene; the function of the EPS gene found in ΦEa35-70 was further investigated in this study in Chapter 4. Many important genes such those involved in DNA replication were not identified in ΦEa35-70 genome. This might suggest that ΦEa35-70 is evolutionarily distinctive phage and possibly sustain unique replication mechanism. Also a holin gene was not detected in the genome of ΦEa35-70. Among all genes involved in cells lysis only endolysin was identified.

There were 15 proteins identified as virion structural and tail proteins based on genome sequences, SDS-PAGE electrophoresis and liquid chromatography-tandem mass spectrometry (LC/MS/MS) analysis at Centre de Recherche du CHU de Quebec, Quebec, Canada. In contrast to ΦKZ genome which encodes six tRNAs, ΦEa35-70 is encoding one tRNA despite large genome size. The tRNA is located between ORF 199 and ORF 200 which encode for proteins with unknown function. The existence of a single tRNA suggests that there may be similar codon usage patterns between ΦEa35-70 and its bacterial hosts.

Phage	Family	Head dimension (nm)	Tail dimension (nm)	genome size (Kb)	GC%	Proteins	tRNA
ΦEa35-70	*Myoviridae*	150	200	27108	49.9	318	1
Pseudomonas phage ΦKZ	*Myoviridae*	145	200	280334	37	306	6
Pseudomonas phage phiPA3	*Myoviridae*	100	185	309208	47.7	375	5

Table 6: General features of ΦEa35-70, *Pseudomonas* phage φKZ and

Pseudomonas phage phiPA3.

Phage	number of shared homologs	percentage related proteins	Reference Sequence	Reference
Pseudomonas phage EL	76 homologs	23.9%	NC_007623.1	(Hertveldt et al., 2005)
Pseudomonas phage OBP	81 homologs	25.47%	NC_016571.1	(Cornelissen et al., 2012b)
Pseudomonas phage 201phi2-1	78 homologs	25.47%	NC_010821.1	(Thomas et al., 2008)
Pseudomonas phage ΦKZ	84 homologs	26.42%	AF399011.1	(Mesyanzhinov et al., 2002)
Pseudomonas phage phiPA3	84 homologs	26.42%	NC_028999.1	(Monson et al., 2011)
Erwinia phage PhiEaH1	89 homologs	27.99%	NC_023610.1	(Meczker et al., 2014)

Table 7: Number of shared homologs and the percentage related proteins to

ΦEa35-70.

Generated using CoreGenes 3.0.

General features and annotation of the ΦEa9-2 genome

The genome of ΦEa9-2 is 75,568 bp long with a G+C content of 47% (Figure 21). The complete genome sequence of *E. amylovora* phage ΦEa9-2 has been submitted to GenBank and assigned accession number KF806588. The ΦEa9-2 genome contains 94 CDSs of which 71 (75.53%) were annotated as encoding hypothetical proteins. Based on gene predictions and annotation of the genome, only 23 CDSs were annotated as functional genes encoding structural proteins and proteins involved in nucleic acid modification and replication. Seventy of the CDSs are located on (+) strand and the rest are encoded by (-) strand. Additionally, a single cluster of seven tRNA genes was detected in the genome. Many genes in ΦEa9-2 such as RNA polymerase, DNA polymerase, DNA helicase and HNH endonuclease have been found to have similarity to 'N4 like phages'. Even though at the DNA level ΦEa9-2 and *Escherichia coli* phage N4 only share 19% overall sequence identity, bioinformatics analysis shows that ΦEa9-2 is related to *Escherichia* phage N4 in that their genomes are syntenic and they share 50 protein homologs with at least 69% identity. *Escherichia* phage N4, belongs to *N4likevirus* genus, has a double-stranded DNA genome, and is 70,153 bp in length. Recently, based on protein homology, Wittmann et al. (2015), classified ΦEa9-2 to be a member of subfamily of N4-like phages which tentatively are called the ''*Enquartavirinae*'' (Wittmann et al., 2015), refer to Figure 22 for genome comparison of ΦEa9-2 and other N4-like phages . This subfamily also includes coliphage N4 (Ohmori et al., 1988; Schito et al., 1966; Willis et al., 2002), the E. coli vB_EcoP_G7C phage the member of ''*G7cvirus*'' (Kulikov et al., 2012) and *Achromobacter* phage JWAlpha (Wittmann et al., 2014). Genomic and morphological features of phages related to ΦEa9-2 are included in Table 8.

One of the early genes identified in ΦEa9-2 is a large virion-associated RNA polymerase (vRNA), 3523 amino acids, which is located on the (-) strand of phage genome between gene clusters of unknown function. The existence of large vRNA is a common characteristic of N4-like phages (Gan et al., 2013). Phage-associated RNA polymerases are known to be involved in in the transcription of late-early genes (Willis et al., 2002). ΦEa9-2 also carries a DNA-directed methylase which might be involved in protecting phage DNA from restriction modification of the bacterial host. *E. amylovora* phage ΦEa9-2 is the only N4-like phage that possesses a gene for a DNA-directed methylase (Wittmann et al., 2015). *rIIA and rIIB* genes are also predicted in ΦEa9-2 genome just downstream of thymidylate synthase gene. N4-Like phages are also known to possess *rIIA* and *rIIB* that are involved in host lysis inhibition during phage infection. Delayed in cell lysis during infection can increase total burst size (Paddison et al., 1998).

Some genes that are involved in DNA replication were also identified in ΦEa9-2. Genes such as DNA polymerase, DNA primase, ATP-dependent DNA helicase and 5'-polynucleotide kinase are found just downstream of *rIIA* and *rIIB* genes, a common feature of some N4-Like phage genomes. The identified genes that involved in DNA replication were located among other genes of unknown function. However, these unidentified genes are expected to be involved in DNA replication as well.

All of the predicted structural proteins in ΦEa9-2 such as major capsid protein, EPS depolymerase and phage tail length tape-measure proteins are located on the lower strand at the end of the phage genome. Genes encoding for structural proteins are interrupted by a putative holin gene that is involved in cell lysis.

Seven tRNA genes specific for the codons for Asn (GTT), Asp (GTC), Glu (TTC), Ile (GAT), Lys (TTT), Ser (GCT), Tyr (GTA) were predicted in the beginning of the

second half of the viral genome. Phage tRNAs play a role in phage genes transcription to compensate for the compositional differences between the viral and the bacterial host genome (Bailly-Bechet et al., 2007).

Phage	host	Head dimension (nm)	Tail dimension (nm)	genome size (Kb)	GC%	Proteins	tRNA	GenBank accession number
ΦEa9-2	*E. amylovora*	NA	NA	75.57	47	94	7	NC_023579.1
Escherichia phage N4	*E. coli*	70	20	70.15	41.3	72	0	NC_008720.1
Lit1virus	*P. aeruginosa*	70	30	72.3	54.9	87	0	FN422399
G7C	*E. coli*	70	25	72.92	43.4	80	0	HQ259105
S6	*E. amylovora*	66	-	74.7	41.3	115	0	HQ728266
JWAlpha	*Achromobacter*	59	22	72.3	54.4	92	0	KF787095
DFL12phi1	*Dinoroseobacter shibae*	-	-	75	49.3	86	2	KJ621082
pCB2047-B	*Sulfitobacter sp.*	-	-	74.5	42.9	77	13	HQ317387

Table 8: Genomic and morphological features of ΦEa9-2 and some of N4-like phages.
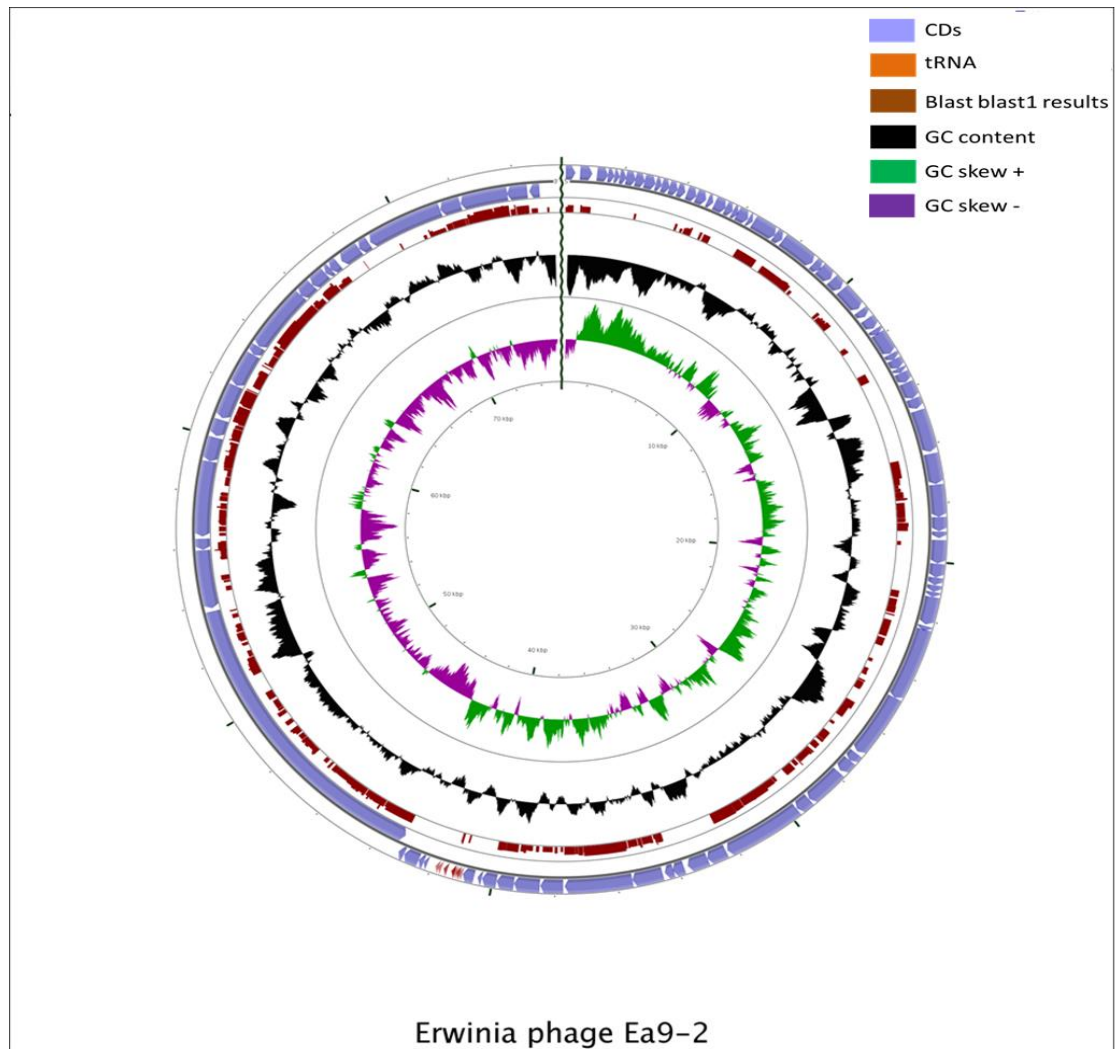
Figure 21: Circular representation of ΦEa9-2 genome.

The outermost circles in purple represent the CDSs on the plus strand (track 1) and the

minus strand (track 2). Also tRNA (maroon) is shown on track 1. The second circle from

outside represents the positions of protein-coding genes as determined by BLASTp.

The third circle shows GC content (black). Innermost ring indicates the GC skew,

minus strand (fuchsia) and positive strand (green).
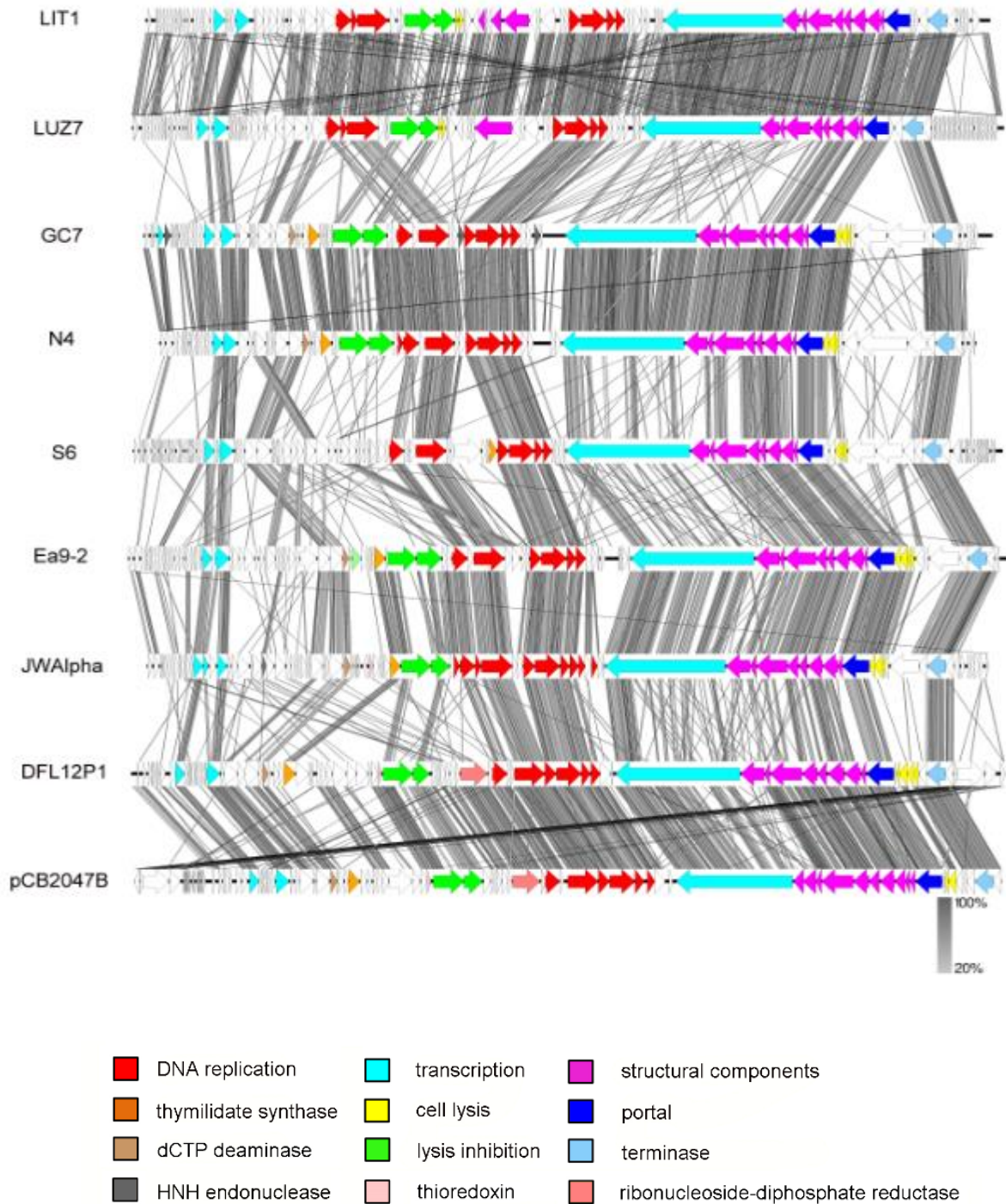
Figure 22: Genome comparison (amino acid level) of ΦEa9-2 and related N4-like phages.

Proposed functional clusters are marked by the same colour. Level of amino acid identity range is shown via the gradient scale. The figure was reproduced from (Wittmann et al., 2015).

Chapter 4

Characterization of depolymerase from a *Myoviridae* phage

Abstract

Phages are currently being investigated for their potential use as biological control agents (BCAs) for the control of *E. amylovora,* the fire blight pathogen. The ability of phages to kill bacteria may be hindered by the presence of a biofilm which is mainly composed of exopolysaccharides (EPS). Phage EPS-depolymerase enzyme degrades amylovoran, an essential bacterial pathogenicity factor, and facilitates phage access to the target bacterial host membrane. The genome sequence of the *E. amylovora Myoviridae* phage ΦEa35-70 revealed the presence of an EPS-depolymerase gene. The existence of depolymerase gene in ΦEa35-70 makes this phage a potential candidate for phage mixtures (or cocktails) used for the control of *E. amylovora.* The goal of this study was to characterize the novel EPS-depolymerase gene from the *Myoviridae* phage and compare it to the EPS depolymerase of *E. amylovora Podoviridae* phage ΦEa31-3 (this study). The *Podoviridae* phages are known to produce depolymerase, and on soft agar overlays their plaques are characterised by large clear halos. In this work, the potential EPS-depolymerase gene from ΦEa35-70 was cloned and expressed in *E. coli* DH5α. When *E. coli* cells harbouring the EPS-depolymerase gene plasmid were plated with a soft agar overlay containing *E. amylovora* cells, a halo was produced around each *E. coli* bacterial colony. In contrast, purified His-tagged EPS depolymerase from ΦEa35-70 did not exhibit EPS-depolymerase activity on purified EPS from *P. agglomerans* and *E. amylovora.* Depolymerase from *Podoviridae* phage ΦEa31-3, which was used as a control, degraded the EPS from both bacteria.

Introduction

E. amylovora capsule is composed of exopolysaccharides (EPS) which serve as the primary bacterial pathogenicity factors. The EPS is the main component of the bacterial biofilm which in turn provides the bacterium with protection against environmental stresses and host plant defense systems (Leigh and Coplin, 1992). *E. amylovora* EPS is composed primarily of levan and amylovoran (Bellemann et al., 1994; Bellemann and Geider, 1992; Bennett and Billing, 1978; Durlu-Ozkaya et al., 2007; Kim et al., 2002; Koczan et al., 2011; Ramey et al., 2004). Levan is made of chains of fructose molecules (Du and Geider, 2002; Du et al., 2004; Geider et al., 1993; Geier and Geider, 1993; Gross et al., 1992; Zhang and Geider, 1997) and amylovoran is a polymer of repeating units made of galactose and glucuronic acid (Bellemann et al., 1994; Bereswill and Geider, 1997; Jumel et al., 1997; Nimtz et al., 1996).

On soft agar overlays, *Podoviridae* phages produce clear plaques with halos with *E. amylovora* host cells that are characterized as high-exopolysaccharide-producing (HEP), and turbid plaques with the low-exopolysaccharide-producing (LEP) host cells. *Myoviridae* phages, on the other hand, do not produce halos around plaques with the LEP or HEP hosts (Roach et al., 2013). The production of clear plaques with halos by the *Podoviridae* phages on HEP bacterial host cells is attributed to phage EPS-depolymerase activity which is not present in most *Myoviridae* phages (Bayer et al., 1979; Born et al., 2014; Lagonenko et al., 2015; Müller et al., 2011a).

Roach et al. (2013) used lab generated amylovoran and levan deficient mutants to study the role of bacterial EPS in phage infection. *Erwinia* spp. *Podoviridae* phages have an absolute requirement for the presence of amylovoran (Roach et al., 2011; Roach et al., 2013; Roach, 2011; Sjaarda, 2012). Amylovoran-deficient bacterial mutants could not be infected by *Podoviridae* phages. In contrast, *Myoviridae* phages

could infect the amylovoran deficient mutants (Roach et al., 2013). Levan, on the other hand, may act as a receptor for the *Myoviridae* phages since levan-deficient *E. amylovora* mutants resulted in reduction of *Myoviridae* phage titers (Sjaarda, 2012). Levan-deficient mutants, however, showed a 29-119% increase in the amylovoran production (Sjaarda, 2012). Therefore, it was not clear if the reduction in *Myoviridae* phage titers using levan-deficient mutants was due to the lack of levan production or the increase in amylovoran production. Amylovoran may create a physical barrier between *Myoviridae* phages and receptors on the cell surface reducing the probability of their binding to the bacterial host. In general, the mechanism by which *E. amylovora* phages penetrate the EPS barrier and infect their bacterial hosts is not fully understood. In fact, it is unclear whether *E. amylovora* EPS plays a role in blocking phage infection, by providing a physical barrier between the phage and cell surface receptors (Deveau et al., 2002), or enhances infection by providing specific receptors for phage attachment.

Certain bacteriophages have evolved to specifically bind and, in some cases, degrade EPS structures (Cornelissen et al., 2012a; Deveau et al., 2002; Glonti et al., 2010; Hsu et al., 2013; Scholl et al., 2005; Stummeyer et al., 2006; Sutherland et al., 2004). Many of these phages, such as phages that belong to *Podoviridae* families; seem to enhance their ability to overcome the EPS by producing virion-associated proteins with EPS-depolymerisation activities (Born et al., 2014; Glonti et al., 2010; Hsu et al., 2013; Hughes et al., 1998a; Kim and Geider, 2000; Pelkonen et al., 1992). The depolymerases function by destroying the EPS physical integrity allowing the phage to gain access to receptors on bacterial cells (Cornelissen et al., 2012a; Hsu et al., 2013; Hughes et al., 1998a; Yan et al., 2014). The mechanism(s) by which non-depolymerase producing phages break down the cell capsule and access the bacterial cell wall is still unknown. *Erwinia* spp. *Podoviridae* phages are also known to produce EPS

depolymerases which may be involved in promoting phage adsorption (Born et al., 2014; Kim and Geider, 2000; Vandenbergh et al., 1985). DNA and amino acids sequences comparison of EPS depolymerases from different *Erwinia* spp. *Podoviridae* phages shows major sequence variations (this study, unpublished data). It is not clear if these variations dictate differences in structure or mode of action of these proteins.

Myoviridae phages from the AAFC Vineland Collection infect *E. amylovora* HEP and LEP isolates, and the orchard epiphyte, *Pantoea agglomerans*. The *P. agglomerans* EPS is composed of arabinose, glucose, galactose, and glucuronic acid (Dong and Fang, 2000). The ability of these phages to penetrate two structurally different EPS molecules and reach the bacterial host without depolymerase activity suggests that alternate mechanism may be present in this family. Certain *E. coli* phages are able to recognize and bind to a serotype-specific sugar "K antigen" a specific component of the bacterium capsular EPS (Scholl et al., 2005; Stirm, 1968). EPS-deficient *E. coli* mutants were shown to be resistant to K antigen-specific phages (Stirm, 1968).

Proteins that bind to specific sugars in plants and/or animals are generally called "lectins", and in mammalian viruses, they are termed "hemagglutinins". Hemagglutinins are found in viruses such as influenza, a part of the virion structure, and can bind to a pacific receptor on the blood cell (Collier et al., 1991; Naim and Roth, 1993; Simpson and Lamb, 1992; Zurcher et al., 1994). Lectins, on the other hand, are a heterogeneous group of proteins with diverse molecular structures that have the ability to bind to carbohydrates. The presence of lectins, as a part of *Myoviridae* phages structure, may enable phages to bind to the receptors on the EPS and allow the penetration to the host cell. However, the existence of these structural proteins in *Myoviridae* phages has yet to be confirmed.

Phage ΦEa35-70 has been classified as a member of the *Myoviridae* family (Yagubi et al., 2014). In contrast to other known *Myoviridae* phages, ΦEa35-70 produces clear plaques on *E. amylovora* HEP. *E. amylovora Myoviridae* phages typically produce hazy plaques on HEP (Roach et al., 2013). Genomic sequence analysis of ΦEa35-70 revealed the presence of a gene, *Depol35-70*, with sequence similarity to EPS-depolymerase genes found in *Podoviridae* phages. However, phage ΦEa35-70 does not produce plaques with halos (Yagubi et al., 2014)

In contrast to the *Podoviridae, Myoviridae* phages showed high efficacies in controlling *E. amylovora* under field, greenhouse and in pear blossom flower bioassay trials (Lehman, 2007). Phage ΦEa35-70 may be a superior candidate for inclusion in phage mixtures since it is a *Myoviridae* phage with potential EPS-depolymerase activity. Studies have demonstrated that phages capable of degrading the EPS are highly effective in targeting pathogenic bacteria that produce biofilms (Hanlon et al., 2001; Hughes et al., 1998b; Tait et al., 2002). The EPS depolymerases act by degrading the host biofilm by digesting the EPS and subsequently exposing bacterial cells to phage infection. In addition, purified phage EPS-depolymerase proteins were found to target and eliminate the colonization of *E. amylovora* in plants (Kim and Geider, 2000). Releasing bacterial cells from biofilm exposes the cells to plant host defence system, environmental influences and reduces the spread of the pathogen. Phages that carry depolymerases could be a promising tool to target *E. amylovora* and other pathogens that are resistant to antibiotics due to the presence of EPS. Bacterial EPS is found to play an important role in resistance mechanisms against antibiotics (Gander and Gilbert, 1997). *E. amylovora* colonies can grow on NA plates supplemented with ampicillin and kanamycin antibiotics; in contrast the bacteria cannot survive in liquid cultures supplemented with the same concentration of antibiotics (Yagubi, personal

observation). Bacterial cells on a solid surface are perhaps surrounded by a thick layer of EPS which might protect them from the antibiotic activity.

The objective of this study is to further investigate the role of *E. amylovora* EPS by characterizing the EPS depolymerase protein (Depol35-70) found in *Myoviridae* ΦEa35-70 and to compare the activity to EPS depolymerase (Depol31-3) found in *Podoviridae* phage ΦEa31-3. This work will provide important information in understanding of the interaction between phages, phage mixtures and the pathogen *E. amylovora*.

Materials and Methods

Bacterial isolates and phages growth conditions

Conditions and media used for growing bacteria and propagating phages were described in Materials and Methods of Chapter 3. Two phages, ΦEa21-4 (*Myoviridae*) and ΦEa31-3 (*Podoviridae*), and three bacterial hosts, Ea6-4, Ea29-7 and Pa13-5 (isolated from pear, Vineland Station) were used in this study.

Depolymerase gene cloning, expression and purification

Depol35-70 of ΦEa35-70 and *Depol31-3* of ΦEa31-3 cloning and expression were carried out using Champion™ pET Directional TOPO® Expression Kits (Invitrogen). All target genes were cloned in pET200/D-TOPO® vector (Invitrogen). This vector contains the IPTG-inducible T7 promoter for high-level expression and N-terminal 6x His-tag for protein detection and purification (Figure 23). All genes of interest were amplified using primers targeting gene sequences including start and stop codons. To clone all genes in the right orientation, all PCR products were amplified by adding four bases (CACC) to the forward primer to allow base pairing with the overhang sequence, GTGG, in pET200/D-TOPO® vector (Figure 23). The forward PCR primers were also designed to place the gene of interest in frame with the N-terminal His-tag (list of primers can be found in table 7A in Appendix).

PCR Reactions were run using 1 μL of 1×10^8 PFU/mL of phage suspension as a template, 5 units of Pfx50™ polymerase (Invitrogen, USA), and final concentration of 1X Pfx50 buffer, 0.4 μM of each primer and 10 mM dNTP Mix. Reactions were run using the following amplification conditions: 94°C for 2 min, (94°C for 10 sec, 45°C for 20 sec, 68°C for 60 sec) x 30 cycles, and a final extension at 68°C for 5 min.

Before ligation of PCR products to the vector, all PCR products were treated with FastAP Thermosensitive Alkaline Phosphatase (ThermoFisher Scientific) according to the supplier recommendations; 5´ phosphate interferes with cloning into the pET200/D-TOPO® vector using topoisomerase ligation. The ligation reaction of the PCR products and pET200/D-TOPO® vector was then transformed into One Shot® TOP10 Chemically Competent *E. coli* (Invitrogen, USA) as described previously in Chapter 3. Transformed cells were then plated on NA plates supplemented with kanamycin and incubated at 37°C for 16 h.

To screen for the correct clones, single colonies were picked using a sterile toothpick and placed in a sterile test tube containing 2 mL of LB supplemented with kanamycin. The tubes were then incubated in a shaker at 37°C for 16 h at 200 rpm. Plasmid DNA isolation was carried out using The Norgen Plasmid DNA MiniPrep kit (Norgen Biotek Corp., Thorold, ON Canada). Clones were confirmed with the appropriate restriction enzymes and digested DNA was run on an agarose gel. The agarose gel was placed in electrophoresis chamber, and covered with TAE buffer (40 M Tris, 20 mM acetic acid, 1 mM EDTA; pH 8.0) and run at 90V for 60-90 min. To visualize the DNA, the gel was stained with GelRed (Biotium, USA). DNA fragments size was estimated using molecular weight ladders (Norgen Biotek Corp., Thorold, ON Canada). Constructed plasmids in this study are listed in Table 9.

```
121   ATAGGCGCCA GCAACCGCAC CTGTGGCGCC GGTGATGCCG GCCACGATGC GTCCGGCGTA GAGGATCGAG ATCTCGATCC
                                     T7 promoter/priming site
              T7 promoter                             lac operator
201   CGCGAAATTA ATACGACTCA CTATAGGGGA ATTGTGAGCG GATAACAATT CCCCTCTAGA AATAATTTTG TTTAACTTTA

         RBS        Nde I                        Polyhistidine region              Nhe I
281   AGAAGGAGAT ATACAT ATG CGG GGT TCT CAT CAT CAT CAT CAT CAT GGT ATG GCT AGC ATG ACT GGT GGA
                       Met Arg Gly Ser His His His His His His Gly Met Ala Ser Met Thr Gly Gly
                              Xpress™ epitope
351   CAG CAA ATG GGT CGG GAT CTG TAC GAC GAT GAC GAT AAG GAT CAT CCC TTC ACC ... ... AAGGGC
      Gln Gln Met Gly Arg Asp Leu Tyr Asp Asp Asp Asp Lys Asp His Pro Phe Thr GGG AAG TGG
                              EK recognition site      EK cleavage site     T7 reverse priming site
411   GAGCTCAACG ATCCGGCTGC TAACAAAGCC CGAAAGGAAG CTGAGTTGGC TGCTGCCACC GCTGAGCAAT AACTAGCATA
```

Figure 23: Cloning site of pET200/D-TOPO®.

Reproduced from Invitrogen, USA website. pET200/D-TOPO® allows expression of recombinant proteins with an N-terminal tag containing the Xpress™ epitope and a 6xHis-tag. The N-terminal tag also includes an enterokinase recognition site to enable removal of the tag after protein purification. The "GTGG" overhang for cloning in the right orientation is shown.

Plasmid	Cloned gene
pET-depolH31-3	Depolymerase gene of ΦEa31-3 (*Podoviridae*)
pET-depolH35-70	Depolymerase gene of ΦEa35-70 (*Myoviridae*)
pET-H26	Tail fiber gene of ΦEa21-4 (*Myoviridae*)

Table 9: List of plasmids constructed to investigate EPS-depolymerase activities of Depol35-70 of ΦEa35-70 and Depol31-3 of ΦEa31-3.

Protein Expression

In order to express the cloned *Depol35-70* and *Depol31-3* genes, pET-depolH31-3 and pET-depolH35-70 plasmids were first transformed into BL21 Star™ (DE3) One Shot® Chemically Competent *E. coli* (Invitrogen, USA). These cells carry the T7 RNA polymerase gene under the control of the lacUV5 promoter that can be induced with isopropylthio-β-galactoside (IPTG). Briefly, 10 ng of plasmid DNA was added to one vial of BL21 Star *E. coli* competent cells, and the mixture was incubated on ice for 30 min. The cells were then heat shocked for 30 sec at 42°C in a water bath. Then cells were placed on ice for 2 min. 250 µL of SOC medium was added, and cells were incubated on a shaker at 300 rpm for 30 min at 37°C for recovery. After incubation, 100 µL of the transformed cells were added to 10 mL of LB supplemented with kanamycin and incubated in shaker at 37°C overnight. To express target proteins, 500 µL of the overnight culture were added to a 10 mL of LB containing kanamycin. Then the cultures were incubated with shaking at 37°C until OD_{600} reached 0.5 in approximately 2 h. IPTG was added to 0.5 mM, and the cells were further grown for 4 h at 22°C (22°C and 4 h incubation were found to give the highest protein yield). Cells were harvested by centrifugation at 8000 xg for 5 min at 4°C, and and the supernatant was aspirated. The cells pellets were immediately frozen at -20°C.

Purification of His-tagged proteins

All His-tagged proteins in this study were purified using Dynabeads® His-Tag Isolation & Pulldown (Invitrogen, USA). Protein purification was carried out according to the supplier's recommendations. In general, the cell pellet from a frozen 10 mL IPTG induced culture was thawed and resuspended in 5 mL of binding/wash buffer (100 mM

Sodium Phosphate, pH 8.0, 600 mM NaCl, 0.02% Tween®-20). DNase I (Roche) was added to avoid a sticky bacterial pellet. Cells were freeze-thawed three times at -80°C and 42°C, respectively. Approximately 4 mg of Dynabeads® was added to each 1 mL aliquot of protein sample in a microcentrifuge tube and mixed thoroughly. The tubes were incubated on a roller for 5 min at room temperature. The tubes were then placed on a magnet until the supernatant was totally separated from the Dynabeads® then supernatant was discarded. The beads were washed four times by resuspension in 2 mL of binding/wash buffer followed by discarding the supernatant after placing the tube on a magnet for at least 2 min. To elute the His-tagged proteins, 100 µL of His-Elution Buffer (300 mM imidazole, 50 mM sodium phosphate pH 8.0, 300 mM NaCl and 0.01% Tween®-2) was added, and the suspension was incubated on a roller for 10 min at room temperature. The mixture was placed on a magnet for 8 min and the supernatant containing the eluted proteins was transferred to a clean tube. Each purified protein was dialyzed using dialysis tubing (MWCO of 6-8 kDa) against 2 changes of 3 L of phosphate buffer saline (PBS) at 4°C for 1 h and 6 h, respectively. All proteins were then analysed on a protein polyacrylamide gel to ensure their integrity.

Depol35-70 and Depol31-3 depolymerase activities on EPS

Two methods were used to analyze the activities of Depol35-70 and Depol31-3 depolymerases.

Agar Overlay

In this method *E. coli* DH5α competent cells (Invitrogen, USA) were transformed with pET-depolH31-3, pET-depolH35-70 or pET-H26 plasmids. One hundred microliters (10^2 CFU/mL) of the three *E. coli* transformants were prepared and mixed

with 100 μL of approximately 10^8 CFU/mL of EaD7 cells in three separate tubes. Then 3 mL of molten top agar was added to the cells and poured on plates containing 25 mL of solidified NA. Plates were then incubated at 27°C for 48 h. After incubation, plates were assessed for the presence of a halo around *E. coli* colonies due to depolymerase activity. The negative control was *E. coli* carrying pET-H26 plasmid with a putative tail fiber protein, a viral protein with no depolymerase activities, from ΦEa21-4.

Incubating Depol35-70 and Depol31-3 proteins with isolated EPS

To test depolymerase activity of Depol35-70 and Depol31-3 proteins, approximately 100 ng of His-tag purified depolymerase proteins were mixed with 200 ng of bacterial EPS. Mixtures were incubated at 27 or 37°C and various pHs (4, 6 and 8) for 1 h. PB (pH 6.8) was mixed with 200 ng of EPS as a negative control in all experiments. After incubation, samples were mixed with protein loading dye and loaded on a polyacrylamide gel. The gel was stained with Alcian blue solution (0.1% Alcian blue in 5% acetic acid, 40% ethanol) and incubated for 2 h at RT in the dark. Gels were de-stained three times in 100 mL of de-stain solution (40% ethanol and 15% acetic acid) for 2 h each time. To ensure the removal of the Alcian blue solution, a final wash with de-stain solution was carried out overnight. Gels were then were placed on a fluorescent light box for visualization.

***E. amylovora* and *P. agglomerans* EPS production and purification**

EaD7 and Pa13-5 cells were streaked on 150 mm plates containing 23 g/L NA (Difco Laboratories) supplemented with 2.5 g/L sucrose to enhance exopolysaccharide production. Plates were incubated at 27°C for 72 h. Cells were suspended in 10 mL of

0.01 M PB (pH 6.8). EPS and the cells were separated by centrifugation at 6,000 xg for 15 min at 4°C. The supernatant was filtered with a 0.45 μm Millipore filter. The collected EPS was dialyzed extensively against of 0.01 M PB (pH 6.8) for overnight at 4°C.   The amount of EPS was estimated calorimetrically using anthrone reagent method (Scott Jr and Melvin, 1953). Anthrone reagent was prepared by dissolving 0.2 g of anthrone (Thermo Fisher Scientific, NJ, USA) in 100 mL of 72% sulphuric acid. Glucose solutions at concentrations of 25, 50, 100, 150 and 200 mg/L were used as a standard. In this procedure, 1 mL of each of the standards or the EPS solution was added to a test tube and placed in ice for 10 min. Five milliliters of anthrone reagent were added to each sample, and placed in 100°C water bath for 15 min. Samples were then placed on ice. The absorbance of samples and standards was measured at 600 nm using a spectrophotometer. Various concentrations of glucose were used to prepare a standard curve. The standard curve was then used to determine the concentration of EPS in samples.

Depolymerase labelling and fluorescence microscopy analysis

To test the ability of Depol35-70 and Depol31-3 to bind to EPS, purified Depol35-70 and Depol31-3 were labelled using Alexa Fluor 350 fluorescent dye (Invitrogen, USA) and mixed with EaD7 and analyzed using fluorescence microscope.

Depolymerase labelling

Proteins were labelled using the Alexa Fluor 350 Protein Labelling Kit according to the supplier's recommendations (Invitrogen, USA). Briefly, purified protein sample (approximately 2 mg in PBS) was added to a vial containing a reactive dye resin and

was stirred at room temperature for 1 h. The resin was pipetted into the column,

allowing excess buffer to drain away into a waste beaker. The protein/resin reaction

mixture was carefully loaded into a column. An elution buffer (PBS, pH 7.2, with 0.2 mM

sodium azide) was slowly added to the column. The column was occasionally

illuminated with a handheld UV lamp to observe the labeled protein movement. The

labelled protein was collected and stored at 4°C for short-term use or at -20°C for long

time storage.

Fluorescence microscopy

An EaD7 cell pellet from a 500 µL culture was washed twice with 1 mL of 0.01 M

PB, pH 6.8. The pellet was then mixed with 4 µg of labelled Depol35-70 or Depol31-3

protein in 500 µL PBS, pH 6.8 and incubated on ice for 10 min. The mixture was

centrifuged and washed three times with 500 µL PBS, pH 6.8 to remove unbound

proteins. A fine fixed smear of cells/protein mix was placed on a microscope slide and

viewed on a Leica DM2000 LED fluorescent microscope equipped with 346 nm

absorption and 442 nm emission filters and digital camera.

Results

Sequence Analysis of ΦEa35-70 depolymerase gene Depol35-70

The *Depol35-70* gene is 1920 bp and encodes 640 amino acids. Blasting of Depol35-70 amino acid sequences against sequences in GenBank showed 99% identity to recently sequenced *Erwinia* sp. phages vB_EamM_Simmy50, vB_EamM_Special G and vB_EamM_RAY (Table 10), and 54% identity (82% similar) to EPS depolymerase of *Erwiana Siphoviridae* phage PhiEaH1 (Meczker et al., 2014) and 28% identity with the putative tail fiber subunit of *E. amylovora* ΦEa 21-4 (Lehman et al., 2009), and 29% identity the tail spike protein in *E. amylovora* phage ΦEa9-2 (this study). Interestingly, Depol 35-70 showed 42% identity with the amylovoran biosynthesis protein AmsF in *Pantoea vagans*, 41% identity with AmsF in *Pantoea ananatis*, and 41% identity with AmsF in *Pantoea stewartii*, *Pantoea agglomeran* and 39 % identity with AmsF in *E. amylovora* (Table 10). Sequence aligmnent of EPS depolymerases of ΦEa35-70 and *Erwiana Siphoviridae* phage PhiEaH1 showed ten conserved regions. These regions are mainly located at the C-terminus of these proteins (Figure 24).

Description	Max score	Total score	Query cover (%)	E value	Identity (%)	Accession
Putative EPS-depolymerase [*Erwinia* phage vB_EamM_Simmy50]	1321	1321	99%	0	99%	ANH51541.1
Putative depolymerase/tail fiber protein [*Erwinia* phage vB_EamM_Special G]	1319	1319	99%	0	99%	ANJ64889.1
Putative EPS-depolymerase [*Erwinia* phage vB_EamM_RAY]	1319	1319	99%	0	99%	ANH51857.1
Putative EPS-depolymerase [*Erwinia* phage PhiEaH1]	686	686	99%	0	54%	YP_009010167.1
Amylovoran biosynthesis protein AmsF [*Pantoea vagans*]	456	456	89%	2E-148	42%	WP_048784190.1
Amylovoran biosynthesis protein AmsF [*Pantoea ananatis*]	441	441	91%	1E-142	41%	WP_019106526.1
Amylovoran biosynthesis protein AmsF [*Pantoea dispersa*]	436	436	91%	5E-141	40%	WP_058780357.1
Amylovoran biosynthesis protein AmsF [*Pantoea stewartii*]	434	434	91%	5E-140	41%	WP_054633476.1
Amylovoran biosynthesis protein AmsF [*Pantoea agglomerans*]	433	433	89%	1E-139	41%	WP_039391519.1
Amylovoran biosynthesis protein AmsF [*Erwinia billingiae*]	425	425	89%	2E-136	42%	WP_053144111.1
Amylovoran biosynthesis protein AmsF [*Erwinia pyrifoliae*]	414	414	89%	2E-132	40%	WP_014538792.1
Amylovoran biosynthesis protein AmsF [*Erwinia amylovora*]	411	411	89%	4E-131	39%	WP_004162445.1

Table 10: A summary of Depol35-70 protein blast results against phage and bacterial proteins in GenBank.

```
ΦEa35-70 1     MSTDVLGQADSLSMFGQVESFEALRKLKPQKEGARISLRGWNKGSCLGGGEFIGHLVTKD    60
               ++ + L   DS +MFG V+SF  LRKLKP      R++LRGWN GS LGGG+FIG +
PhiEaH1  50    VTLEKLSAVDSTAMFGSVQSFADLRKLKPTANNVRVNLRGWNNGSDLGGGQFIGRIGVAK   109

ΦEa35-70 61    LKARPDDGGMIA-AGSGFFWERVCSDCGQMNVTHFGALPGATAATDCVPAMTAMLNWSNA   119
                 + DDGG+IA +G+ ++WER   D   ++VTHFGA+ A   TD +PA+ AM  WS A
PhiEaH1  110   ASMKKDDGGVIATSGTDWYWERAIGDINTLDVTHFGAV--ADGVTDSLPAVLAMFKWSQA   167

ΦEa35-70 120   NAVGLGIQFPAGTFFASSLKFTKEIGYFRIAGNHIN-FGYFANTVLVSDATENFMLDVAA   178
               +     IQFPAG FF S ++  EI  FR+ G   N FGYF  T L SD+ EN +  + A
PhiEaH1  168   SGGFQAIQFPAGKFFLSQFEYLAEIARFRLNGTATNHFGYFNATTLTSDSNENVLFSLNA   227

ΦEa35-70 179   RRVELANIIFRGVSTTEKPNKKGIFNNIVIGGQYFNMTCVRFEKIGGTCVNLIDTLDTKM   238
               R VE++  +F G +  + NKKG + NI+ GGQY  ++ +R+ K+GGTCV+++DTLD+K+
PhiEaH1  228   RWVEISGFVFEGRTDIDG-NKKGFYKNIIAGGQYTRISMMRWSKVGGTCVDMLDTLDSKI   286

ΦEa35-70 239   TQWYASQCTGDVIVNKWSDRQAGGWNHTTAVEMTNFNIQKCVGGKVFDMQRCAQSFITNG   298
                QWYAS CTGDVI   WS R AG W+H TAVE++FN+Q C GGKV ++ RC QSFI NG
PhiEaH1  287   DQWYASGCTGDVIKGYWSGRAAGVWDHMTAVELSNFNVQNCRGGKVLNLPRCTQSFIWNG   346

ΦEa35-70 299   WIEHCDDPGDLSNGQWTVNGLSIEDCK-KPLKASYARFIEIQRNLQSGSRIDYSIDPELE   357
               WIEH D+PGDLSNGQW +NGL+IEDCK  PLK ++AR I  Q  LQSGS + Y +
PhiEaH1  347   WIEHSDNPGDLSNGQWIINGLNIEDCKMSPLKLTFARIINFQLGLQSGSSVSYDVQDGDV   406

ΦEa35-70 358   EWLSEWERGRLDIANHGIFIDGTLEPGTIMSRNKLSNSSDKAKWFRVGTWYSPSEGDSID   417
               EWLS+WERG +DI NHGI+IDGT++    + SRNKLSNSSDKAKWF +G +++PSEGD ID
PhiEaH1  407   EWLSQWERGHVDIDNHGIYIDGTIDAAIVGSRNKLSNSSDKAKWFCLGRFFTPSEGDQID   466

ΦEa35-70 418   INLVGTGNFLSVAPKLDDIDGVRQGGGNTLIRVQLKKDSVGATLEPHGSSPLQAAKIAKV   477
               IN+VGTGNFLS++  + DID  RQGGGNTLIR Q KK + G +L P GSSP++A
PhiEaH1  467   INMVGTGNFLSISTTMADIDDTRQGGGNTLIRYQAKKGAAGLSLMPMGSSPIKAVAYTST   526

ΦEa35-70 478   GAGKFVIYAQLKPYTRNVIPVIVATGATHYEAGVSYYFRPAIASMSDEEIAAVPDAVVIA   537
                 AG   +Y QLKPYTRNVIP+I ATG TH+EAGVS Y+RP   +++DEE+AA+ D   +
PhiEaH1  527   SAGNITLYVQLKPYTRNVIPIITATGLTHFEAGVSVYWRPDRKALTDEEMAAIVDVKPVV   586

ΦEa35-70 538   EQWSLGQMAGVGATNEGNLMLKSKIVNDHLQVQVMVSQAQWAKPELRYIELK          589
               EQWS+GQ+AG+GAT EG L++KSK+ N HL V+V            +++Y++LK
PhiEaH1  587   EQWSIGQLAGIGATQEGYLLVKSKVENGHLAVKV--------NGKVQYLQLK          630
```

Figure 24: Amino acid alignment of Depol35-70 protein with EPS depolymerase of PhiEaH1.

The two proteins shared 54% identiy. Highly conserved regions were found between the two proteins are underlined.

Testing Depol35-70 and Depol31-3 depolymerase activities on EPS

Top agar overlay method

Three different plasmids pET-depolH31-3, pET-depolH35-70 and pET-H26 were

constructed to test depolymerase activities of Depol35-70 and Depol31-3. Incubation of

the transformed *E. coli* with EaD7 on NA plates revealed depolymerase activities for

Depol35-70 and Depol31-3. A clearance (halo) around the *E. coli* colonies harbouring

pET-depolH35-70 that carries *Depol35-70* was detected as a result of Depol35-70

depolymerase activity (Figure 25, B). A similar, but more distinct, halo was observed

around the *E. coli* colonies containing pET-depolH31-3 (Figure 25, A). No halo was

detected around *E. coli* colonies that carry pET-H26 (negative control, Figure 25, C).

Some of the bacterial colonies that exhibited halos around them were picked and

plasmid DNA was isolated and analysed using restriction enzymes to confirm that they

carried the correct plasmid.
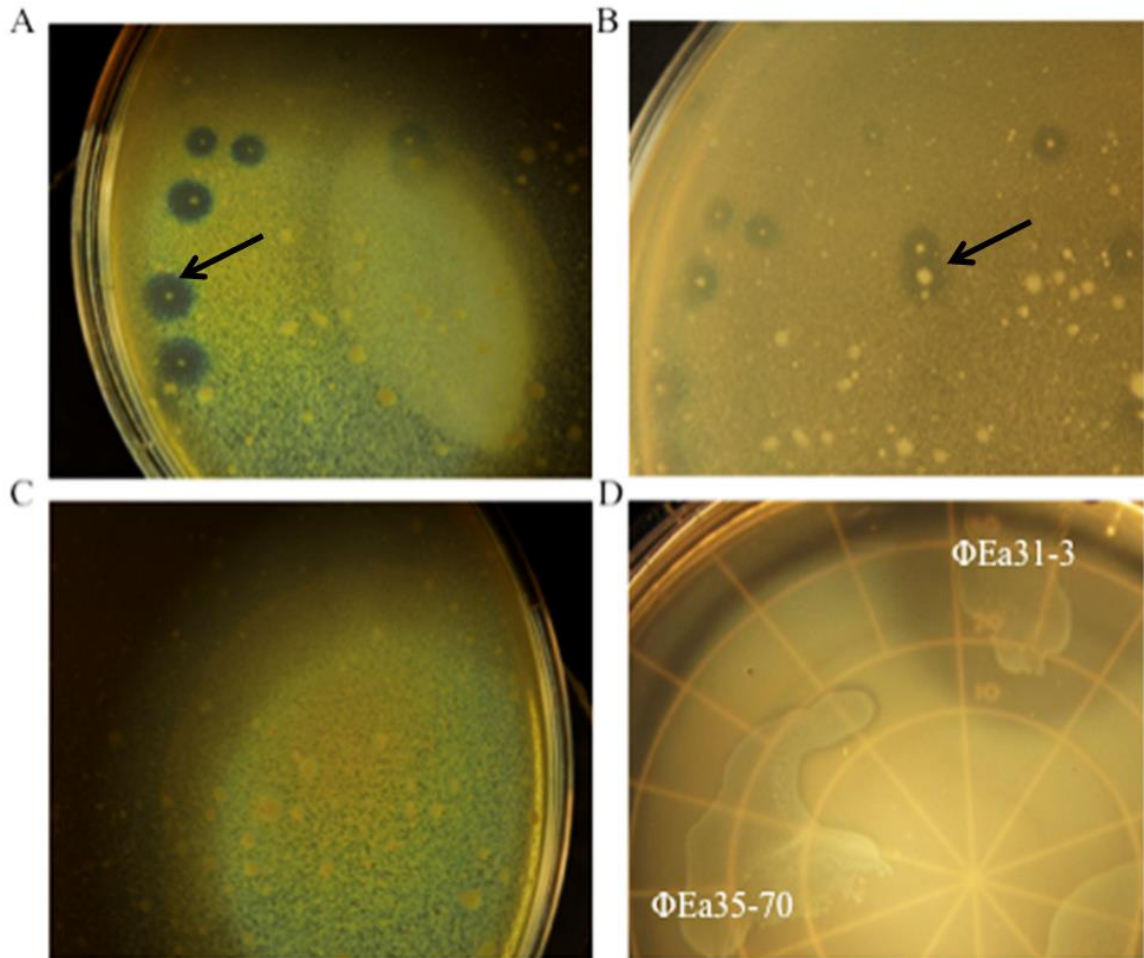
Figure 25: EPS degrading activities of Depol35-70 and Depol31-3 using agar overlay method.

(A) Agar overlay of EaD7 cells mixed with *E. coli* carrying pET-depolH31-3. (B) Agar overlay of EaD7 cells mixed with *E. coli* carrying pET-depolH35-70. (C) Negative control: Agar overlay of EaD7 cells mixed with *E. coli* carrying pET-H26. (D) Spot test of ΦEa31-3 and ΦEa35-70 phages over a lawn of EaD7 cells. Clearances (halos) around *E. coli* colonies due to depolymerase activities are shown by black arrows. No halos are detected around *E. coli* colonies carrying pET-H26 (C). A halo can be seen around ΦEa31-3 phage in the spot test (D); however, no halo was observed with ΦEa35-70 (D).

Exploring purified Depol35-70 and Depol31-3 digestion specificity

To test Depol35-70 and Depol31-3 digestion specificity, two different EPSs purified from *E. amylovora* and *P. agglomerans* were exposed to Depol35-70 and Depol31-3. EPS structure of *P. agglomerans* is slightly different from *E. amylovora* EPS. Phage ΦEa35-70 is known to infect and replicate in *P. agglomerans*; however, ΦEa31-3 can only replicate in certain *P. agglomerans* isolates. EPS was first isolated from *P. agglomerans* and *E. amylovora* as described previously. The isolated EPS from all bacterial samples were incubated at 27°C for 1 h with Depol35-70 and Depol31-3. Samples were then loaded on SDS-PAGE stained with Alcian Blue to visualize EPS digestion. Results showed that Depol35-70 did not degrade EPS isolated from *E. amylovora* or *P. agglomerans.* On the other hand, Depol31-3 digested both types (Figure 26).
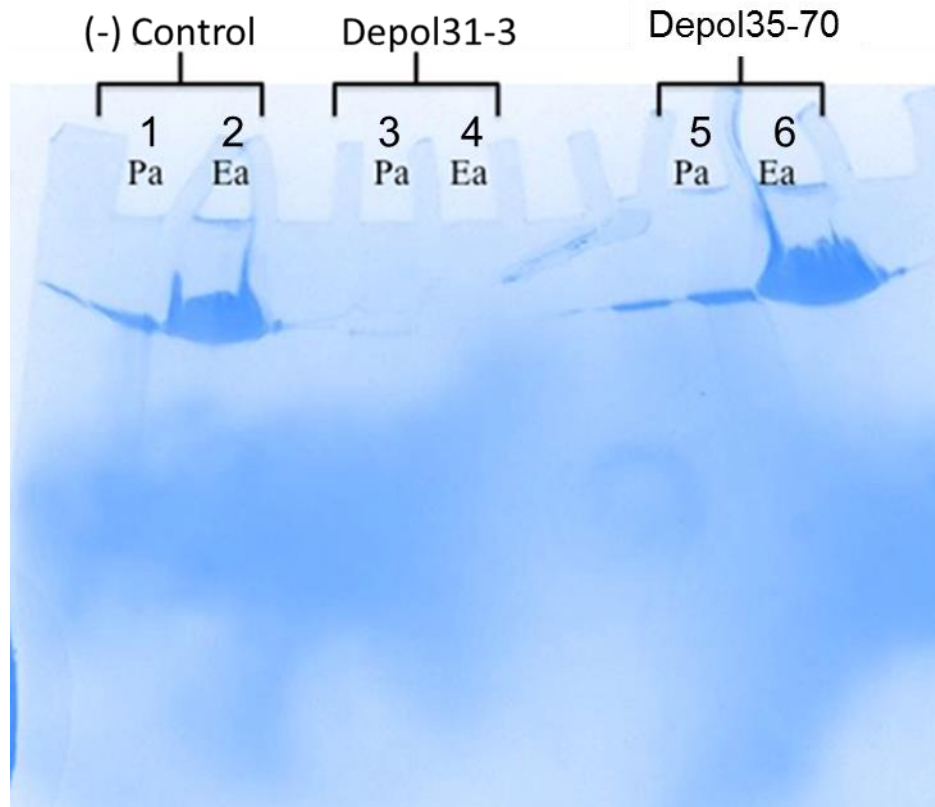
Figure 26: The digestion of EPS from *E. amylovora* and *P. agglomerans* with Depol35-70 and Depol31-3.

SDS-PAGE stained with Alcian Blue shows the digestion of EPS from *E. amylovora* (EaD7) and *P. agglomerans* (Pa13-5) with Depol35-70 and Depol31-3. Lane 1 contains EPS from *P. agglomerans* (Pa), and Lane 2 contains EPS of *E. amylovora* (Ea). Both lanes contain only EPS without adding depolymerase. Lane 3 and 4 contain EPS from both hosts incubated with DepolH31-3. Clearance in the lanes represented EPS degradation. Lanes 5 and 6 contain EPS incubated with DepolH35-70. The intensity of EPS at the top of the lane indicated no discernable digestion activity.

The effect of temperature on Depol35-70 and Depol31-3 activities

Two temperatures were used to test the EPS digestion abilities of the Depol35-70 and Depol31-3. The digestion incubation was carried out at 27°C and 37°C for 1 h. Depol31-3 was able to digest the EPS similarly at 27°C and 37°C. No digestion activity was detected using Depol35-70 at both temperatures (Figure 27).

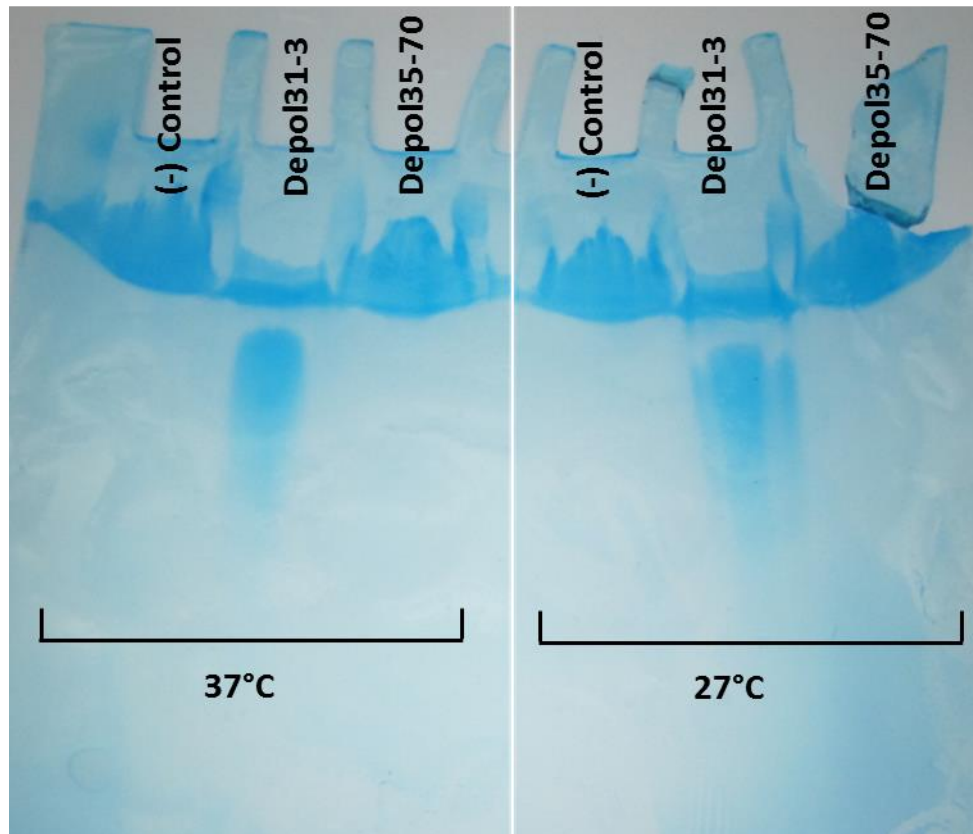Figure 27: The digestion of EPS from *E. amylovora* with Depol35-70 and Depol31-3 under different temperatures.

SDS-PAGE stained with Alcian blue visualizing *E. amylovora* EPS digestion using Depol35-70 and Depol31-3 at 27°C and 37°C. (-) control: EPS incubated with PB, pH 6.8. The depolymerases used for EPS digestion are shown. The digested EPS by Depol31-3 appeared as a smear inside the gel.

The effect of pH on Depol35-70 and Depol31-3 activities

The effects of three different pHs were tested on the EPS-digestion activities of Depol35-70 and Depol31-3. Depol35-70 or Depol31-3 was mixed with *E. amylovora* EPS in three buffers with different pH sets: pH 4, 6, and 8; the mixtures were then incubated at 27°C for 1 h before loading on SDS-PAGE. Results showed that Depol35-70 showed no activity and Depol31-3 degraded the EPS at all pHs. However, pH 6 and 8 slightly improved EPS-digestion activities of Depol31-3 compared to pH 4 (Figure 28).
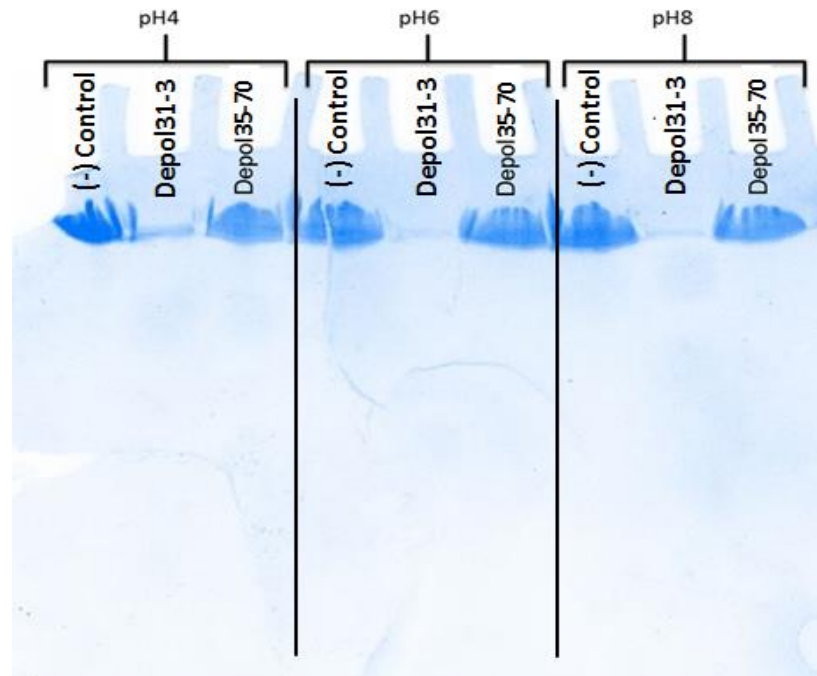
Figure 28: The digestion of EPS from *E. amylovora* with Depol35-70 and Depol31-3 using different pH sets.

Alcian blue stained SDS-PAGE to investigate the effect of pH on depolymerase activities of depol35-70 and depol31-3. Very little EPS is seen at any pH when Depol31-3 is used due to enzymatic degradation. When Depol35-70 was used, EPS staining intensity was indistinguishable from the buffer control, indicating no digestion. (-) control is only EPS. The pHs of the digestion reactions are shown.

Investigating binding properties of Depol35-70 and Depol31-3

The ability of the depolymerases of ΦEa35-70 and ΦEa31-3 to bind to the EPS of EaD7 cells was characterized by confocal imaging. The binding of fluorescent-labeled depolymerases to EaD7 EPS is illustrated in Figure 29. EaD7 was incubated with Depol35-70 and Depol31-3 labelled with Alexa Fluor™ 350 and examined under a Leica DM2000 fluorescent microscope. It is clearly evident that Depol35-70 cannot bind to EaD7 EPS while Depol31-3 can (Figure 29).
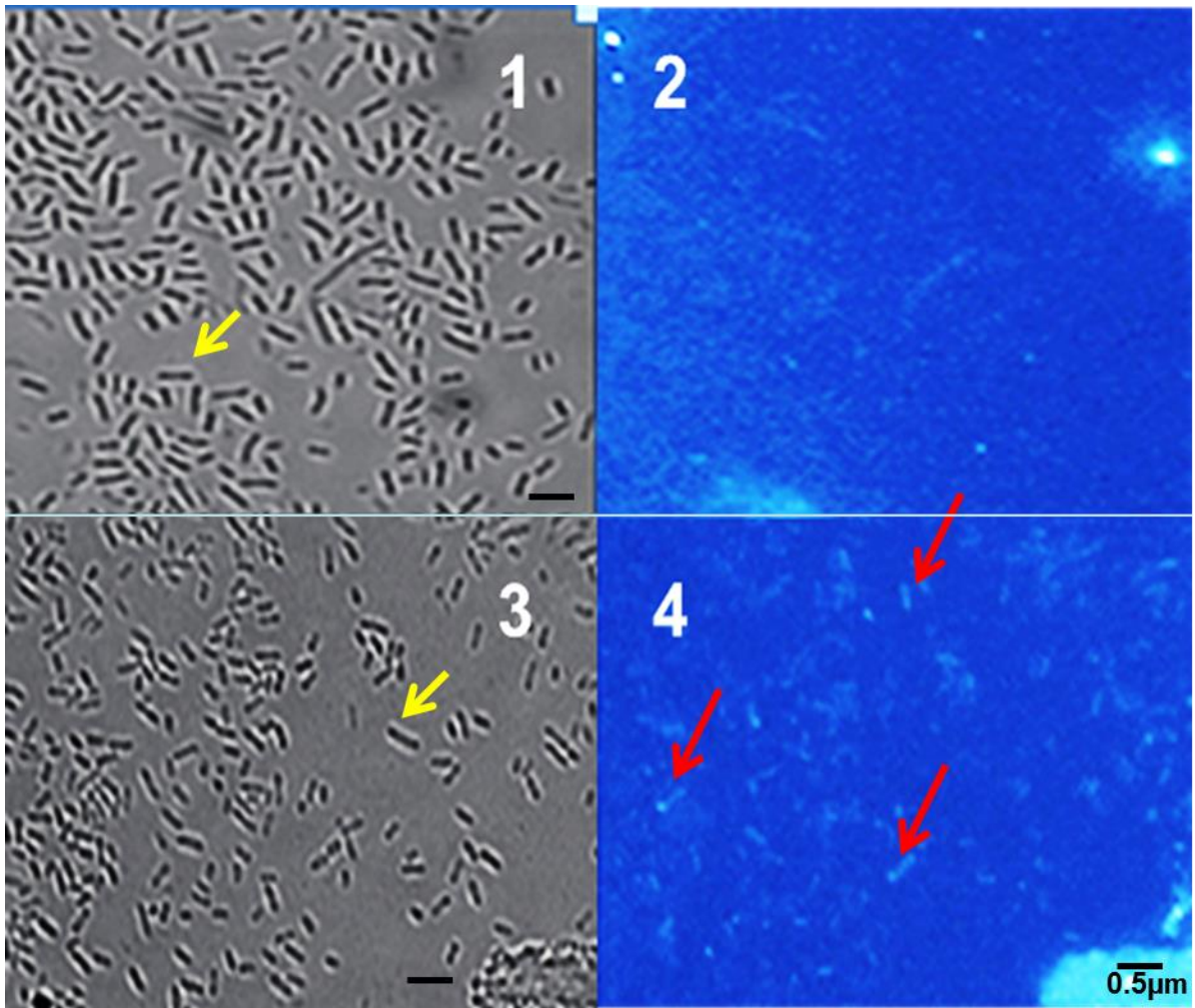
Figure 29: Fluorescent microscope images of EaD7 mixed with Depol35-70 and Depol31-3 fluorescently labelled proteins.

Images taken from Leica DM2000 Fluorescent Microscope at 40X magnification. (1 and 2) Depol35-70 protein mixed with EaD7 cells. (3 and 4) Depol31-3 mixed with EaD7 cells. Images 1 and 3 were taking with regular light. Images 2 and 4 (fluorescence). In image 4, it seems that Depol31-3 is binding and decorating EaD7 (red arrows). No similar observation was detected with Depol35-70 mixed with EaD7. The EPS appears like a halo around cells, and is indicated by yellow arrows.

Discussion

The sequence analyses of *Myoviridae* ΦEa35-70 genome revealed the existence of a potential EPS-depolymerase gene with similar sequences to other EPS depolymerases found in *Podoviridae* phages (Yagubi et al., 2014). The existence of an EPS depolymerase in *E. amylovora* phages is novel in that there is a single report of an *E. amylovora Myoviridae* phage in GenBank with potential EPS depolymerase is phiEa2809 (Lagonenko et al., 2015). In this study, the EPS depolymerases of ΦEa35-70 (Yagubi et al., 2014), a *Myoviridae* phage with a genome size of 271 kbp, was characterized in comparison to EPS depolymerase of ΦEa31-3 (*Podoviridae*), genome size of 45 kbp (this study). The size of the potential depolymerase gene (*Depol35-70*) in ΦEa35-70 is 1961 bp. Depol35-70 showed a 52% identity to *Erwinia Siphoviridae* phage ΦPhiEaH1 EPS depolymerase (Meczker et al., 2014) (Figure 24). Amino acid sequences alignment of Depol35-70 and EPS depolymerase found in ΦPhiEaH1 showed conserved sequences at the C-terminus (Figure 24). It has been shown that the C-terminus of some *Podoviridae* phages EPS depolymerase is responsible for degrading EPS (Born et al., 2014). This suggests that the C-terminus of Depol35-70 may also be responsible for EPS depolymerase activity. Depol35-70 might also have sugar binding abiltiy since this protein showed a 42% and 41% identity with the amylovoran biosynthesis protein AmsF in *P. vegans*, *P. ananasis*, *P. agglomeran* and *E. amylovora, respectively*. AmsF is known to be involved in building up repeating units of the amylovoran and having significant homology to viral EPS depoymerases (Geider, 2000). Langlotz et al. (2011) showed that Ams F protein is involved in processing the newly formed EPS repating units and/or polymerization of the new amylovoran chain (Langlotz et al., 2011).

Phage EPS depolymerases may be structural viral proteins or soluble proteins expressed during phage infection (Drulis-Kawa et al., 2015; Pires et al., 2016; Stummeyer et al., 2006). This study confirmed that the Depol35-70 is expressed as part of the viral particle structural proteins since it was associated with other phage ΦEa35-70 structural proteins (Yagubi, unpublished data). Also, sequences analyses of ΦEa35-70 genome revealed that *Depol35-70* is positioned in the vicinity of other phage structural proteins. The location of Depol35-70 protein on the phage particle was not explored in this study. The sequence similarity of *Depol35-70* to the tail fiber subunit found in ΦEa 21-4 and the tail spike protein in ΦEa9-2 might indicate that Depol35-70 is a part of tail structure of ΦEa35-70. In other *Podoviridae* phages, EPS depolymerases were found on the phage tail (Born et al., 2014; Luna et al., 2013; Ritz et al., 2013).

The depolymerase activity of Depol35-70 was further confirmed when *E. coli* DH5α harbouring a plasmid carried *Depol35-70* gene produced a halo when plated with EaD7 cells (Figure 25). A similar halo was produced by in the control when *E. coli* DH5α harbouring a plasmid carried *Depol31-3*. No halo was observed in the negative control which consisted of *E. coli* DH5α sheltering a plasmid with a phage protein with no depolymerase activity (Figure 25).

To further investigate both Depol31-3 and Depol35-70 proteins, the coding genes of these proteins were cloned in plasmids next to His-tag sequences near the protein N-terminus. The fusion of these genes to a His-tag facilitated purification of the proteins. The EPS-depolymerase activities of His-tagged proteins were investigated using different EPS digestion conditions such as temperature and pH to determine the optimal conditions for EPS digestion. Depol35-70 did not exhibit any EPS-digestion activities with EPS isolated from *E. amylovora* and *P. agglomerans* regardless of using various

pHs and incubation temperatures (Figures 26, 27, 28). The lack of depolymerase activity of the His-tagged Depol35-70 may be attributed to the His-tag interference with the overall folding of the protein. Adding the tag to C-terminus instead of N-terminus of Depol35-70 was not an option since the C-terminus is expected to be responsible for depolymerase activity according to Born et al. (2014). Depol31-3, on the other hand, was shown to be a versatile enzyme since it digested both EPSs from *E. amylovora* and *P. agglomerans.* The ability of Depol31-3 to digest *E. amylovora* and *P. agglomerans* EPS is interesting. Usually EPS degrading enzymes are made specifically to recognize certain EPS structures (Pires et al., 2016). These EPS degrading enzymes might not recognise the diversity of EPS structures that are made by different bacteria (Azeredo and Sutherland, 2008). Depol31-3 was also able to digest EPS at 37°C and 27°C, and at different pH sets (Figures 26, 27, 28). However, pH 8 seemed to be optimum for Depol31-3 depolymerase activities. Maximal activity of the purified protein from *E. amylovora* bacteriophage ΦEa1h, a *Podoviridae,* was observed between pH 4 and 5 at 52°C (Kim and Geider, 2000). Depolymerase protein found in *E. amylovora Podoviridae* phage L1 has been demonstrated to have optimal activity at 50°C, and at a pH 6.0 (Born et al., 2014). The temperature of 50°C was not investigated in this study since it is irrelevant in using phages in biological control.

Since Depol35-70 did not show depolymerase activity, EPS-binding abilities of Depol35-70 and depol31-3 proteins, as a control, were investigated. The His-tagged purified Depol31-3 and Depol35-70 proteins were labeled using Alexa Fluor 350 fluorescent dye (Invitrogen, USA). The labeled depolymerases were mixed with EaD7 cells and viewed under a fluorescent microscope. In case of Depol31-3, EaD7 cells were found to be decorated with fluorescence an indication of protein binding to the EPS. No fluorescent cells were observed with Depol35-70 (Figure 29). Instead, the

Depol35-70 formed aggregations which were observed under the fluorescent microscope. It is not clear that the aggregation of Depol35-70 was a result of proteins interaction or improper folding of Depol35-70. Increasing salt (NaCl or MgCl2) concentration or adding glycerol and EtOH might help prevent aggregation. However, this was not tested in this experiment.

A host range study of selected *E. amylovora* phages was carried out against a total of 30 *P. agglomerans* and 100 *E. amylovora* bacterial isolates from worldwide (S. Gayder, PhD thesis). Phages included in that study: ΦEa10-1, ΦEa10-2, ΦEa21-4, ΦEa45-1b and ΦEa35-70, which belong to *Myoviridae* family, and ΦEa10-7, ΦEa31-3, ΦEa46-1-A-1, ΦEa46-1-A-2, and ΦEa9-2, which belong to *Podoviridae* family. The genome sequence of these phages was determined in this study. The data from the host range study was used to investigate the effect of Depol35-70 on ΦEa35-70 infectivity in comparing to other phages including ΦEa31-3 that carries *Depol31-3*. When *E. amylovora* was used as a host, ΦEa35-70 generally didn't grow well in all hosts. However, it preferred HEP over LEP by producing higher titers of up to 2 logs units on the HEP. *Myoviridae* phages, which do not carrying EPS depolymerase, replicated more efficiently than ΦEa35-70 by 3 to 4 logs in all host including HEP and LEP. *Podoviridae* phages, carry EPS depolymerase, did extremely well in HEP and much less by up to 3 logs in LEP due to EPS depolymerase activates. These results may indicate that Depol35-70 is enhancing ΦEa35-70 infectivity in *E. amylovora* isolates that produce high levels of EPS.

The host range pattern for the same phages in *P. agglomerans* was very different from *E. amylovora.* ΦEa35-70 replicated more efficiently in all *P. agglomerans* isolates when compared to other phages, especially *Podoviridae* phages which

replicated very poorly in all isolates. The ability of ΦEa35-70 to achieve a higher titre in *P. agglomerans* compared to other phages may indicate that ΦEa35-70 is a *Pantoea* phage. As a result, using ΦEa35-70 as a biological control agent against *E. amylovora* might not be effective since it prefers *P. agglomerans* over the pathogen *E. amylovora*. This would mean that when the phage is delivered inside *P. agglomerans* (carrier) as part of a BCA for fire blight control, the phage would kill the epiphyte rather than the pathogen on the flower stigma resulting in failed efficacy under field conditions.

In conclusion, the sequence analysis of EPS-depolymerase gene *Depol35-70* from ΦEa35-70 showed high similarity to genes in three phages, low similarity to other EPS depolymerases found in other *E. amylovora Podoviridae* phages and to the Ams F protein from *P. vegans*, *P. ananatis* and *E. amylovora*. When this protein was cloned in a plasmid and expressed in *E. coli,* it exhibited depolymerase activities by degrading EPS of *E. amylovora*. However, when this protein was expressed using an expression vector and purified using His-tag column, no depolymerase activities were detected under various EPS digestion conditions. It is not understood why Depo35-70 protein was not biologically active after His-tag purification despite it being purified under native conditions and eluted using a mild buffer. Also it is not clear if the His-tag of the protein might have had a negative effect on the protein folding and activity. Depol3-31 of the *Podoviridae* phage ΦEa31-3, which was used as a control, is found to be a versatile protein that can efficiently degrade EPS isolated from *P. agglomerans* and *E. amylovora* at 27°C at 37°C and a range of pH.

General Discussion

Lytic phages are currently being developed as biopesticides for the control of the fire blight pathogen. In order to develop an effective phage-based biopesticide against *E. amylovora,* the resistance mechanisms of the bacterial host against phages need to be explored and understood. The role of CRISPR/Cas systems, an acquired resistance mechanism, of *E. amylovora* in phage resistance has not been investigated. All the previous studies on *E. amylovora* CRISPR/Cas system have concentrated on the exploration of CRISPR/Cas system structure and diversity and CRISPR array spacers origin (McGhee and Sundin, 2012; Rezzonico et al., 2011). To date, resistance mechanism(s) to phages have not been identified in this pathogen.

The role of *E. amylovora* CRISPR/Cas system in response to phage invasion and plasmids transformation were explored in this thesis. Previous research demonstrated that the level of expression of Cas genes was increased during viral infection in *S. thermophilus* (Young et al., 2012) and *Thermus thermophilus* (Agari et al., 2010). The CRISPR/Cas system of *E. amylovora* was anticipated to respond in a similar manner to phage invasion. In this study, the level of expression of innate Cas genes of *E. amylovora* did not react to phage infection when challenged with phages from the *Myoviridae* and *Podoviridae* families. This is an indication that *E. amylovora* CRISPR/Cas system may not be effective in the recognition of invading phages. Only a minor increase was observed in the expression level of Cas genes in the beginning of *Podoviridae* phages infection with EaD7, an isolate that produces high levels of EPS. The genomes of *Podoviridae* phages sequenced in this study showed no protospacers match spacers in CRISPR arrays of bacterial isolates used in this project. Therefore, the initial response of bacterial isolate EaD7 CRISPR/Cas system in the start of *Podoviridae* phages infection could be attributed to the ability of phages form the

Podoviridae family to degrade the EPS structure. Dissolving the EPS capsule, in turn, might create a stress on the bacterial host stimulating the bacteria defence mechanisms.

The tactic used by of *E. amylovora* phages to avoid the host CRISPR/Cas system might be credited to an unidentified anti-CRISPR function used by the *Erwinia* spp. phages. Anti-CRISPR genes have been identified in *P. aeruginosa* phages (Bondy-Denomy et al., 2013; Pawluk et al., 2014). The anti-CRISPR proteins of *P. aeruginosa* phages did not affect the expression level of Cas genes during infection, but prevented CRIPSR/Cas system of the host from recognizing phages during CRISPR interference step (Bondy-Denomy et al., 2013; Pawluk et al., 2014). It is possible that *E. amylovora* phages are equipped with similar anti-CRISPR genes. To investigate this further, 17 *Erwinia* spp. phages belonging to *Myoviridae* and *Podoviridae* families were sequenced and analyzed. None of the phages were found to carry comparable anti-CRISPR genes. None the less, *Erwinia* phages might harbour anti-CRISPR genes with distinctive sequences that differ from the anti-CRISPR genes of *P. aeruginosa* phages. Alternative, *E. amylovora* phages may lack anti-CRISPR genes and use different mechanisms to avoid CRISPR/Cas system recognition.

The alternative mode of action of the CRISPR/Cas system was explored based on the infection of *E. coli* K12 by phages. In this system, CRISPR/Cas provided resistance against λ phage when Cas genes were overexpressed using a plasmid that carried Cas genes (Brouns et al., 2008). Increasing in *Cas* level of expression also led to adding new phage derived spacers to CRISPR arrays of *E. coli* K12 (Datsenko et al., 2012). In this study, a plasmid which carried an artificial CRISPR/Cas system and three spaces targeting sequences in phage ΦEa21-4 was introduced into *E. amylovora* isolate Ea6-4. In contrast to the *E. coli* K12 results, this study showed that improving

Cas genes level of expression in *E. amylovora* did not prevent normal phage infection and replication, and did not enhance the gaining of new spacers from invading phages.

The CRISPR/Cas system of *E. amylovora* did not prevent phage infection; however, it appears to be functional. The existence of intact CRISPR/Cas system, conservation of Cas gene sequences, and existence of various spacers indicate that this system is evolutionarily selected to serve specific functions. This study demonstrated that *E. amylovora* CRISPR/Cas system was capable in blocking plasmid DNA transformation especially with plasmids carried protospacers targeting CRR1 and CRR2. The apparent increase in the accumulation of crRNA in the isolates transformed with plasmids carried an artificial CRISPR/Cas system comparing to wild type isolate. The increase of crRNA accumulation may be a result of the increase in the intercellular concentration of Cas proteins expressed by the artificial CRISPR/Cas system. The overexpression of *CasE* (*Cse3*) in *E. coli* was shown to increase in the accumulation of crRNA (Pougach et al., 2010).

Spacer sequences in *E. amylovora* CRISPR arrays have been examined in bacterial isolates in this study and other worldwide isolates (McGhee and Sundin, 2012; Rezzonico et al., 2011). No protospacers were found in already sequenced *E. amylovora* phages that match spacers detected in the CRISPR arrays of all bacterial isolates. However, a limited number of fully sequenced phages are available in GenBank. Availability of larger number of phage genomes will significantly contribute to the further understanding of the CRISPR/Cas system and other bacterial phage resistance mechanisms. In this thesis, the genomes of 17 *E. amylovora* phages from the *Podoviridae* and *Myoviridae* families were sequenced and annotated. None of the sequenced phages were found to harbour protospacers with a perfect match to spacers detected in *E. amylovora* CRISPR/Cas system. It is important to note, that all of these

phages were originally isolated and enriched on 6 wild type isolates of *E. amylovora* obtained from southern Ontario. Therefore, phages carried protospacers match existing spacers in *E. amylovora* may be excluded during the isolation step from environmental samples due to CRISPR/Cas system.

This thesis progressed the understanding of the CRISPR/Cas system in *E. amylovora*; however, more questions needed to be answered. For example, how the existing spacers in *E. amylovora* CRISPR arrays were gained, and what motivates the gaining of new spacers. Future work on anti-CRISPR genes in *E. amylovora* may explain what the mechanisms allow the phages to escape CRISPR/Cas systems of *E. amylovora*.

The 17 phages studied in this project demonstrated distinct host range and replication efficiency (Steven Gayder unpublished data). Genomic analyses of 15 of these phages revealed that they share 99.9% similarity with existing phages in GenBank. *E. amylovora* phages ΦEa10-1, ΦEa10-2, ΦEa10-4, ΦEa10-16, ΦEa21-2, ΦEa35-2, ΦEa45-1b, and ΦEa51-7 belong to *Myoviridae* family, and are similar to *E. amylovora* ΦEa21-4 (Lehman et al., 2009). Phages ΦEa10-6, ΦEa10-7, ΦEa31-3, ΦEa35-3, ΦEa35-10 ΦEa46-1-A-1, and ΦEa46-1-A-2 were found to be similar to *Podoviridae* phage ΦEa1h (Müller et al., 2011a). It is interesting to note that phage ΦEa1h was originally isolated in Michigan, US (Ritchie, 1978; Ritchie and Klos, 1977). The reason that phages belonging to the same family share 99.9% sequence similarity despite their apparent host range difference is not clear. A closer look at the phage genomes revealed rearrangement in the location of some genes. However, this rearrangement could be due errors in the assembly process of phage genomes. The actuality of genes rearrangement has to be confirmed further using different molecular

biological techniques. The construction of phage genomic library and sequencing analysis of the cloned fragments could be useful in the validation of genes location in the genome.

Among all sequenced phages, only the *Podoviridae* phage ΦEa9-2 and the *Myoviridae* phage ΦEa35-70 genomes were found to be different when compared to other phages in GenBank. ΦEa9-2 genome is 75,568 bp, and found to be is related to coliphage N4 since they share 50 protein homologs (69% identity). ΦEa9-2 can infect *E. amylovora* (pathogen) and *P. agglomerans* (phage carrier), unpublished data. The distinctive sequence of ΦEa9-2 and its ability to infect both the pathogen and carrier potentially makes this phage a good biopesticide candidate. Phages with unique sequences are easier to be detected and monitored in the field when they applied.

ΦEa35-70 genome, on the other hand, is 271,084 bp and it encodes 318 putative proteins and one tRNA. ΦEa35-70 shares 26% protein homologs with *Pseudomonas* phage ΦKZ, and is related to the Phikzlikevirus genus within the *Myoviridae* family. An EPS-depolymerase gene was detected in the genome of the *Myoviridae* phage ΦEa35-70. The EPS-depolymerase gene of the ΦEa35-70 was characterized in comparison to the EPS depolymerase of *E. amylovora Podoviridae* phage ΦEa31-3 (this study). ΦEa35-70 EPS depolymerase degraded the EPS structure of EaD7 when it was cloned and expressed form *E. coli* cells. However, when the EPS-depolymerase protein was His-tag purified, no EPS depolymerase activity was detected. In a host range study carried by Steven Gayder (PhD Candidate, Brock University), ΦEa35-70 infectivity in *E. amylovora* producing high EPS is higher than in low EPS producers. An indication that the EPS depolymerase might play a role in helping the phage to penetrate the EPS capsule of the host. *Myoviridae* phages were

shown to do well in field trails to control *E. amylovora* (Lehman, 2007). *Myoviridae* phages that carry EPS-depolymerase gene could be a promising tool in eliminating *E. amylovora*. The EPS depolymerases degrades the host biofilm by digesting the EPS exposing of the pathogen allowing phage infection and plant host defence system recognition. Despite the characterization of the EPS depolymerase of ΦEa35-70, further confirmation of its ability to degrade EPS structure and enhance phage infectivity have to be explored. Furthermore, ΦEa35-70 efficacy to control *E. amylovora* has to be investigated.

Appendix

Table 1A: Primers used to amplify and sequence CRR1, CRR2 and CRR3 of *E. amylovora* isolates in this study.

Primer	Sequence	Function
CR11Fr CR12Rev	5' ACCGGAGTGCGTAATATGAG 3' 5' GAATAGCAGGGCCATTAACC 3'	Amplify CRISPR region 2 of Ea110R, Ea6-4, Ea17-1-1, EaG5, EaD7, BC1280, BC29 and Ea4-96
CR21Fr CR22Rev	5'AAACTGACGTGCCGTATCTG 3' 5' TCCTGACGTAACGTGGATAG 3'	Amplify CRISPR region 1 of Ea110R, Ea6-4, Ea17-1-1 and EaD7
CR1Ea110F CR1 Ea110R	5' GCAGCGTGACAGTTATGGAG 3' 5' ACACCCGTAAGCAGCCTACC 3'	Amplify CRISPR region 1 of Ea110R
CR1-480 CR1-850 CR1-1201 CR1-1953 CR1 EA110R	5' CGGAGATGCACTGGATATAC 3' 5' ACCATCGCCGCGATAAATCC 3' 5' GGGATAAACCGCAGATGAGG 3' 5' AAATCAGTGGCGGCGAAACC 3' 5' ACACCCGTAAGCAGCCTACC 3'	Sequence CRISPR region 2 of Ea110R, Ea6-4, Ea17-1-1 and EaD7
CR1-609 CR1-973 CR1-1351 CR1-1842 CR1-2155 CR1 EA110F	5' CGAGCATCTCGGGAACTGTG 3' 5' ACCGCAGGTATTTCGGATAG 3' 5' CGGAAATGTCGGCGGAGATG 3' 5' GGCTTTGCGTTGTTACTCAG 3' 5' ATTGCCACCAGATGTGTTCC 3' 5' GCAGCGTGACAGTTATGGAG 3'	Sequence CRISPR region 1 of Ea110R, Ea6-4, Ea17-1-1 and EaD7
CR1-1700 CR1 Ea110R CR1-1201 CR1-850 CR1-1201 CR1-1953 CR1 Ea110R CR1EAG5FR CR1EAG5REV	5' GTTCGCATACGACAATCTCC 3' 5' ACACCCGTAAGCAGCCTACC 3' 5' GGGATAAACCGCAGATGAGG 3' 5' ACCATCGCCGCGATAAATCC 3' 5' GGGATAAACCGCAGATGAGG 3' 5' AAATCAGTGGCGGCGAAACC 3' 5' ACACCCGTAAGCAGCCTACC 3' 5' AGTCGAGAACGTGCCTTATG 3' 5' TTCGGCTGAGGCTTGCTTTG 3'	Sequence CRISPR region 2 of EaG5

G5LaRev	5' CCGGTCGGAGTTTGTTGTTG 3'	
G5LaFr	5' AACCGCAACTGAGCAACTCC 3'	
BC292Fr	5' GATAAACCGCCATTCCCGAG 3'	
BC292Rev	5' TCGCCGCCACTGATTTGAGC 3'	
BC293Fr	5' AAACCGTAACCGCAACCATC 3	Sequence CRISPR region 2 of BC29 and BC1280
CR1-1700	5' GTTCGCATACGACAATCTCC 3'	
CR1-1953	5' AAATCAGTGGCGGCGAAACC 3'	
CR1-1201	5' GGGATAAACCGCAGATGAGG 3'	
Ea4-962Fr	5' TCGCCAAGGTAGTCATCGTG 3'	Sequence CRISPR region 2 Ea4-96
Ea4-962Rev	5' TAACGGGTGTCGTAGTGATG 3'	
CR3F	5' AGACTGGCGTTATAGGATGG 3'	Amplify and sequence RISPR region 3
CR3R	5' AGCCCGTGAAGCAAAGTATG 3'	

Table 2A: Primers used to amplify and clone Cas genes of EaD7 isolate. Bolded-underlined letters indicate overhangs to introduce specific restriction sites for cloning purposes.

Primer	Sequence	Function
Cas1Cas2F	5' CCGGAGGCTTTATGTCGTTTG 3'	Sequence *Cas*1 and *Cas*2
Cas1Cas2R	5' GGTTTGCATGTCAGACGGGA 3'	
Cas12ekxF	5' **GGGCGGCCGCGTGGTACCGACATCTCGAG**CCGGAGGCTTTATGTCGTTTG 3'	Amplify and clone *Cas*1 and *Cas*2
Cas12bR	5' **GGCGTCTCAAACGC**GGTTTGCATGTCAGACGGGA 3'	
Cas3kaaF	5'**CGTGGTACCGGGCCCGTACGGGCGCGCC**CGTAAAGGGAGCTTATGAATTCGC 3'	Amplify and clone *Cas*3
Cas3xR	5' **CGGCTCGAG**CCAGTCATCGGTCAATAAATTCATC 3'	
CasgF2	5' GATGGTACCTGGATGGTCAGCAGCTGG 3'	Amplify and clone Cas genes
CasgR	5' ATCTTGGGCCCCCATCCAGCACATCGATTTGC 3'	

Table 3A: Primers used to sequence a Cas genes of EaD7 and Ea6-4 isolates.

Primer	Sequence (5'-3')	Function
Casg1	TTTGACAGCGAACACGGCAGGG	
Casg2	CGCATGCCAGCAAATCACCT	
Casg3	AGCTGCTGGTCGGGTTAACC	
Casg4	AGG AAGAGGTGATTGCCGGC	
Casg5	GCA ACAGGTCAATAAACAGCAG	
Casg6	GAA AAGGTTGATGCCCCGTA	
Casg7	TTACAGCTGGATCCGGGGAGCAAG	
Casg8	AGGTGTACTGGCCCGGTTTC	
Casg9	AGCATGCGCACCATCCATCT	
Casg10	AATCCCCGGTGCCAAACAGC	
Casg11	CGAGGGATAACAATGACCGA	
Casg12	TGACACGATCCTCACCTGGC	Sequence Cas genes
Casg13	CCGTTAATCCACCTACCACG	
Casg14	TAACGCAGTTTGCCCTGTCGCC	
Casg15	CTGGTTAACGGTATCGGACC	
Casg16	ATAATGCGGTGGATGTGGCG	
Cas1	CGCTGCCGCTAAAAGATCGC	
Cas2	GATGGTGAGAAGATCTCAGGCAGA	
Cas3	TGGAGCAGTGAGGAGACCTTCTG	
Cas4	CGCGCTTTAATCCGGCTTTT	
Cas5	TTATCCGCTGGTCACCTGGC	
Cas6	GGATGTCGATTTCGACTGGCTG	
Cas7	AACGATCATGCCGTACGCGC	

Table 4A: Primers and probes used in this study to monitor the level of expression of

*Cas*1, *Cas*3, CrRNA and house-keeping genes in *E. amylovora.* Probes are in bold.

Fluorophores are attached to the 5' end of the probe and a quencher to the 3' end.

Primer/Probe	Sequence	Function
Cas1ExF2	5' GATATGCACCCGCTATTGGT 3'	Monitor *Cas*1 expression
Cas1ExP3	**/ROXN/TAAACAAATGACAGCGGCTTGC CAC/IABRQSp/**	
Cas1ExR	5' GCTTGCCCAGTATTCTTCCA 3'	
groELExF	5' ATGAACCCAATGGACCTGAA 3'	Monitor *gro*EL expression
groELExP	**/HEX/CGGTATCGACAAAGCAGTCA/IABkFQ/**	
groELExR	5' CACACCTTCTTTGCCCACTT 3'	
recA FWD	5' GGGACGCTGCTGATCTTTAT 3'	Monitor *Rec*A expression
recA PRB	**/FAM/TGATGTTCG/ZEN/GTAACCCGGAAACCA/IABkFQ/**	
recA REV	5' CGATACGGCGAATATCCAGAC 3'	
CrRNA R1 FWD	5' GGGATAAACCGTGAAG 3'	Monitor region 1 CrRNA expression
CrRNA R1 PRB	**/ROXN/ AGCCAGAATCCCATCCG/IBHQ_2/**	
CrRNA R1 REV	5' GGGAACACGATAAAGG 3'	
CrRNA R2 FWD	5' GGGATAAACCGGATGG 3'	Monitor region 2 CrRNA expression
CrRNA R2 PRB	**/ROXN/CGTACCGATGTTTGCGAA/IBHQ_2/**	
CrRNA R2 REV	5' GGAACACGCGAATCT 3'	
Cas3GD FWD	5' CCGTGGGATCACGGTAAAG 3'	Monitor *Cas*3 expression
Cas3GD PRB	**/ROXN/CCGTTTGCTATGCGGTGGTTCATC/IBHQ_2/**	
Cas3GD REV	5' ATACTCACGCACCACATAGC 3'	

Table 5A: Primers used to detect spacer acquisition in CRR1, CRR2 and CRR3 of *E. amylovora.*

Primer	Sequence	Function
CR1AcqF CR1AcqR	5' CGTGCGTGCTTTAAAGTGGAAA 3' 5' AACACAGTTCCCGAGATGCTCG 3'	Detect spacer acquisition in CRISPR region 1
CR2AcqF CR2AcqR	5' CGCGGCTTTTGACAGCAAA 3' 5' CGGGGAACACGCTGTTTTGT 3'	Detect spacer acquisition in CRISPR region 2
CR3F CR3R	5' AGACTGGCGTTATAGGATGG 3' 5' AGCCCGTGAAGCAAAGTATG 3'	Detect spacer acquisition in CRISPR region 3
CR2AcqF M13 Reverse	5' CGCGGCTTTTGACAGCAAA 3' 5´ CAGGAAACAGCTATGAC 3´	Detect spacer acquisition in pSp21-4, pSpCas123 and pSpCasg plasmids

Table 6A: Primers and probes used to monitor phage replication in *E. amylovora.*
Probes are in bold. Fluorophores are attached to the 5' end of the probe and a
quencher to the 3' end.

Primer	Sequence	Function
Phage 9-2 F	5' CATGGGTAATCCCTTTGAG 3'	Monitor ФEa9-2 replication
Phage 9-2 P	**/FAM/TCTGGTGGA/ZEN/CAGAGACGATGTAAT/IABkFQ/**	
Phage 9-2 R	5' GATAGACTGGTTCCCCTGTG 3'	
Phage 35-70 F	5' TGGAAGGTCTTCTTCGAGAC 3'	Monitor ФEa35-70 replication
Phage 35-70 P	**/ROXN/GACGGAAAAGATCACGGTACCTT/IABRQSP/**	
Phage 35-70 R	5' GACTACCTGGGGATGTTCAG 3'	
Pun45-F	5' TTCAGCTTTAGCGGCTTCGAGA 3'	Monitor ФEa21-4 replication
Pun45-P	**/ROX/AGTCGGTACACCTGCAACGTCAAGAT/IABRQ/**	
Pun45-R	5' AGCAAGCCCTTGAGGTAATGGA 3'	
dpo1-F	AGGTGTGGATTGACGTAGGC	Monitor ФEa31-3 replication
dpo1-P	**/FAM/TGGTGGTAAGACGACCATCA/BHQ/**	
dpo1-R	CCTTCGCATCAGAGTCCTTC	
Ea-IscF	5' CGCTAACAGCAGATCGCA 3'	Monitor *E. amylovora* replication
Ea-IscP	**/CY5/CTGATAATCCGCAATTCCAGGATG/IABRQ/**	
Ea-IscR	5' AAATACGCGCACGACCAT 3'	

Table 7A **:** Primers used to amplify and clone depolymerase genes of ФEa35-70 and
ФEa31-3 in pET200/D-TOPO® vector.

Primer	Sequence (5'-3')	Function
HdepolFr	CACCATGACAGATAGCACTACA	Amplify and clone ФEa35-70 depolymerase protein (with a His-tag)
depolRev	GAATCCCCTCCTCCAGTTAGC G	
Hdepol31-3Fr	CACCATGCCGGAAACTAAGATTAAG	Amplify and clone ФEa31-3 depolymerase protein (with a His-tag)
depol31-3Rev	TTACGTGTTAGATACTGTACCA	

Figure 1A: Sequences of the leader and protospacers and repeats inserted in the

plasmid carried the artificial CRISPR system. The leader sequence is in bold. The

protospacers are underlined. The three protospacers are made to target ΦEa21-4 gene

66 (putative baseplate component), gene 69 (putative tail fiber protein) and gene 83

(conserved hypothetical protein).

ACGCATGCAAACCGCGGCTTTTGACAGCAAAAAATCCGGTAGATTTTGACGGGT
AAAAAAAGCCGTTATGGTTCAATGGTTTGTATCTAGAGTGTTCCCCGCGTGAGCG
GGGATAAACCG<u>ATTTGGGACACAAAACCTGCTGGTATTCGTGC</u>GTGTTCCCCGCG
TGAGCGGGGATAAACCG<u>AAGTATGTTACAGCGAAAGTTGTAAACGACC</u>TGTGTTC
CCCGCGTGAGCGGGGATAAACCG<u>CGTCTTAACTCGTTGATTGATAACGCAACTTT</u>
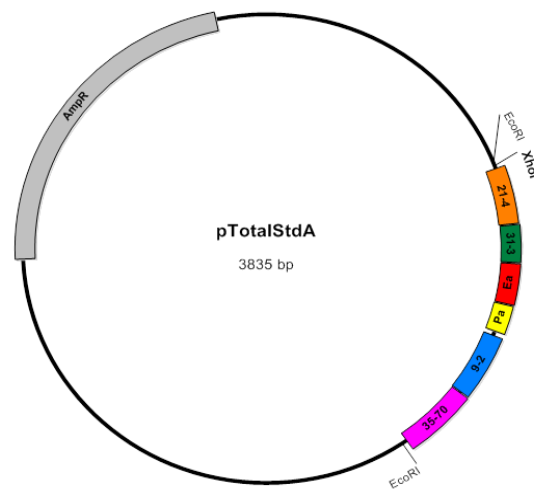GTGTTCCCCGCGTGAGCGGGGATAAACCGTGT

Figure 2A: the map of pTotalStdA plasmid used to create standard curves titer phages

and *E. amylovora* using real-time PCR in this study. This plasmid was constructed by

cloning the PCR target sequence in phages and *E. amylovora* in plasmid vector. Coned

sequences are shown in colors.

Literature Cited

Abedon, S. T. (2008). Bacteriophage Ecology: Population growth, evolution, and impact of bacterial viruses. Advances in Molecular and Cellular Microbiology pp. 508. Cambridge, NY: Cambridge University Press.

Abuladze, T., Li, M., Menetrez, M. Y., Dean, T., Senecal, A., and Sulakvelidze, A. (2008). Bacteriophages reduce experimental contamination of hard surfaces, tomato, spinach, broccoli, and ground beef by *Escherichia* coli O157 : H7. Applied and Environmental Microbiology*, 74*, 6230-6238.

Ackermann, H. W. (2007). 5500 phages examined in the electron microscope. Archives of Virology*, 152*, 227-243.

Ackermann, H. W., and Prangishvili, D. (2012). Prokaryote viruses studied by electron microscopy. Archives of Virology*, 157*, 1843–1849.

Adams, M. H. (1959). Bacteriophages. New York: Interscience Publishers.

Addy, H. S., Askora, A., Kawasaki, T., Fujie, M., Yamada, T. (2012a). The filamentous phage ΦRSS1 enhances virulence of phytopathogenic *Ralstonia solanacearum* on tomato. Phytopathology*, 102*, 244-251.

Addy, H. S., Askora, A., Kawasaki, T., Fujie, M., Yamada, T. (2012b). Loss of virulence of the phytopathogen *Ralstonia solanacearum* through infection by ΦRSM filamentous phages. Phytopathology*, 102*, 469-477.

Agari, Y., Sakamoto, K., Tamakoshi, M., Oshima, T., Kuramitsu, S., and Shinkai, A. (2010). Transcription profile of *Thermus thermophilus* CRISPR systems after phage infection. Journal of Molecular Biology*, 395*, 270-281.

Aldridge, P., Bernhard, F., Bugert, P., Coplin, D. L., and Geider, K. (1998). Characterization of a gene locus from *Erwinia amylovora* with regulatory functions in exopolysaccharide synthesis of *Erwinia* spp. Canadian Journal of Microbiology*, 44*, 657-666.

Allison, G. E., and Klaenhammer, T. R. (1998). Phage resistance mechanisms in lactic acid bacteria. International Dairy Journal*, 8*, 207-226.

Andersson, A. F., and Banfield, J. F. (2008). Virus population dynamics and acquired virus resistance in natural microbial communities. Science*, 320*, 1047-1050.

Asselin, J. E., Bonasera, J. M., Kim, J. F., Oh, C. S., and Beer, S. V. (2011). Eop1 from a *Rubus* strain of *Erwinia amylovora* functions as a host-range limiting factor. Phytopathology*, 101*, 935-944.

Ayers, A. R., Ayers, S. B., and Goodman, R. N. (1979). Extracellular polysaccharide of *Erwinia amylovora*: a correlation with virulence. Applied and Environmental Microbiology*, 38*, 659-666.

Azeredo, J., and Sutherland, I. W. (2008). The use of phages for the removal of infectious biofilms. Current pharmaceutical biotechnology*, 9*, 261-266.

Babu, M., Beloglazova, N., Flick, R., Graham, C., Skarina, T., Nocek, B., Gagarinova, A., Pogoutse, O., Brown, G., Binkowski, A., Phanse, S., Joachimiak, A., Koonin, E. V., Savchenko, A., Emili, A., Greenblatt, J., Edwards, A. M., and Yakunin, A. F. (2011). A dual function of the CRISPR-Cas system in bacterial antivirus immunity and DNA repair. Molecular Microbiology*, 79*, 484-502.

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

Balogh, B. (2002). Strategies of improving the efficacy of bacteriophages for controlling bacterial spot of tomato. MSc. Thesis, Gainesville, Florida USA: University of Florida.

Balogh, B. (2006). Characterization and use of bacteriophages associated with citrus bacterial pathogens for disease control. PhD. Thesis, Gainesville, Florida USA: University of Florida.

Balogh, B., Jones, J. B., Iriarte, F. B., and Momol, M. T. (2010). Phage therapy for plant disease control. Current Pharmaceutical Biotechnology*, 11*, 48-57.

Balogh, B., Jones, J. B., Momol, M. T., Olson, S. M., Obradovic, A., King, P., and Jackson, L. E. (2003). Improved efficacy of newly formulated bacteriophages for management of bacterial spot on tomato. Plant Disease*, 87*, 949-954.

Barionovi, D., Giorgi, S., Stoeger, A. R., Ruppitsch, W., and Scortichini, M. (2006). Characterization of *Erwinia amylovora* strains from different host plants using repetitive-sequences PCR analysis, and restriction fragment length polymorphism and short-sequence DNA repeats of plasmid pEA29. Journal of Applied Microbiology*, 100*, 1084-1094.

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D. A., and Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. Science*, 315*, 1709-1712.

Barrangou, R., and Marraffini, L. A. (2014). CRISPR-Cas systems: prokaryotes upgrade to adaptive immunity. Molecular Cell*, 54*, 234-244.

Bayer, M. E., Thurow, H., and Bayer, M. H. (1979). Penetration of the polysaccharide capsule of *Escherichia coli* (Bi 161/42) by bacteriophage K29. Virology*, 94*, 95-118.

Behlau, F., Canteros, B. I., Minsavage, G. V., Jones, J. B., and Graham, J. H. (2011). Molecular characterization of copper resistance genes from *Xanthomonas citri* subsp. citri and *Xanthomonas alfalfae* subsp. *citrumelonis*. Applied and environmental microbiology*, 77*, 4089-4096.

Beier, D., and Gross, R. (2006). Regulation of bacterial virulence by two-component systems. Current Opinion in Microbiology*, 9*, 143-152.

Bellemann, P., Bereswill, S., Berger, S., and Geider, K. (1994). Visualization of capsule formation by *Erwinia amylovora* and assays to determine amylovoran synthesis. International Journal of Biological Macromolecules*, 16*, 290-296.

Bellemann, P., and Geider, K. (1992). Localization of transposon insertions in pathogenicity mutants of *Erwinia amylovora* and their biological characterization. Journal of General Microbiology*, 138*, 931-940.

Beloglazova, N., Brown, G., Zimmerman, M. D., Proudfoot, M., Makarova, K. S., Kudritska, M., Kochinyan, S., Wang, S., Chruszcz, M., and Minor, W. (2008). A novel family of sequence-specific endoribonucleases associated with the clustered regularly interspaced short palindromic repeats. Journal of Biological Chemistry*, 283*, 20361-20371.

Beloglazova, N., Petit, P., Flick, R., Brown, G., Savchenko, A., and Yakunin, A. F. (2011). Structure and activity of the Cas3 HD nuclease MJ0384, an effector enzyme of the CRISPR interference. The EMBO journal*, 30*, 4616-4627.

Bennett, R. A. (1980). Evidence for two virulence determinants in the fireblight pathogen *Erwinia amylovora*. Journal of General Microbiology*, 116*, 351-356.

Bennett, R. A., and Billing, E. (1978). Capsulation and virulence in *Erwinia amylovora*. Annals of Applied Biology*, 89*, 41-45.

Bereswill, S., and Geider, K. (1997). Characterization of the *rcsB* gene from *Erwinia amylovora* and its influence on exopolysaccharide synthesis and virulence of the fire blight pathogen. Journal of Bacteriology*, 179*, 1354-1361.

Bernhard, F., Poetter, K., Geider, K., and Coplin, D. L. (1990). The *rcsA* Gene from *Erwinia amylovora:* identification, nucleotide sequence, and regulation. Molecular Plant-Microbe Interactions*, 3*, 429-437.

Bernhard, F., Schullerus, D., Bellemann, P., Geider, K. M., D.R., Coplin, D. L., Nimtz, M., and Bonn, W. G. (1996). Genetics and complementation of DNA regions involved in amylovoran synthesis of *Erwinia amylovora* and stewartan synthesis of *Erwinia stewartii.* Acta Horticulturae*, No. 411*, 269-274.

Bernheimer, H. P., and Tiraby, J.-G. (1976). Inhibition of phage infection by pneumococcus capsule. Virology*, 73*, 308-309.

Biggs, A. R. (1994). Characteristics of fire blight cankers following shoot inoculations of three apple cultivars. HortScience*, 29*, 795-797.

Billing, E. (1960). An association between capsulation and phage sensitivity in E*rwinia amylovora*. Nature*, 186*, 819-820.

Bohannan, B. J. M., and Lenski, R. E. (2000). Linking genetic change to community evolution: insights from studies of bacteria and bacteriophage. Ecology Letters*, 3*, 362-377.

Bolotin, A., Quinquis, B., Sorokin, A., and Ehrlich, S. D. (2005). Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. Microbiology*, 151*, 2551-2561.

Bondy-Denomy, J., Pawluk, A., Maxwell, K. L., and Davidson, A. R. (2013). Bacteriophage genes that inactivate the CRISPR/Cas bacterial immune system. Nature*, 493*, 429-432.

Bonn, W. G., and Leuty, T. (1993). An assessment of the MARYBLYT R computer program for the prediction of fire blight in Ontario, Canada. Acta Horticulturae*, 338*, 145-152.

Bonn, W. G., and van der Zwet, T. (2000). Distribution and economic importance of fire blight. Fire blight: the disease and its causative agent, *Erwinia amylovora*, 37-53.

Borecki, Z., and Lyskanowska, K. (1968). *Erwinia amylovora* (Burr.) Winsl. in Poland. Journal of Phytopathology*, 61*, 157-166.

Borkar, S. G. (1997). Lysogeny in *Xanthomonas campestris* pv. *erythrinae* and *X. campestris* pv. *azadirachtae* and the behaviour of harboured temperate phages. Journal of Mycology and Plant Pathology*, 27*, 283-285.

Born, Y., Fieseler, L., Klumpp, J., Eugster, M. R., Zurfluh, K., Duffy, B., and Loessner, M. J. (2014). The tail-associated depolymerase of *Erwinia amylovora* phage L1 mediates host cell adsorption and enzymatic capsule removal, which can enhance infection by other phage. Environmental Microbiology*, 16*, 2168-2180.

Born, Y., Fieseler, L., Marazzi, J., Lurz, R., Duffy, B., and Loessner, M. J. (2011). Novel virulent and broad-host-range *Erwinia amylovora* bacteriophages reveal a high degree of mosaicism and a relationship to *Enterobacteriaceae* phages. Applied and Environmental Microbiology*, 77*, 5945-5954.

Boulé, J., Sholberg, P. L., Lehman, S. M., O'Gorman, D. T., and Svircev, A. M. (2011). Isolation and characterization of eight bacteriophages infecting *Erwinia amylovora* and their potential as biological control agents in British Columbia, Canada. Canadian Journal of Plant Pathology*, 33*, 308-317.

Breth, D. I., Borejsza-Wysocka, E., Kuehne, S., Aldwinckle, H. S., and Cox, K. (2014). The presence of Erwinia amylovora in asymptomatic apple bud wood: A threat to newly established apple plantings. Acta Horticulturae Vol. 1056, pp. 235-238.

Brouns, S. J., Jore, M. M., Lundgren, M., Westra, E. R., Slijkhuis, R. J., Snijders, A. P., Dickman, M. J., Makarova, K. S., Koonin, E. V., and van der Oost, J. (2008). Small CRISPR RNAs guide antiviral defense in prokaryotes. Science*, 321*, 960-964.

Brüssow, H. (2009). The not so universal tree of life or the place of viruses in the living world. Philosophical Transactions of the Royal Society B: Biological Sciences*, 364*, 2263-2274.

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

Bugert, P., Bereswill, S., Geider, K., and Bonn, W. G. (1996). Regulation and structure of the *ams*-region involved in exopolysaccharide synthesis of *Erwinia amylovora*. Acta Horticulturae*, 411*, 275-280.

Bugert, P., and Geider, K. (1995). Molecular analysis of the *ams* operon required for exopolysaccharide synthesis of *Erwinia amylovora.* Molecular Microbiology*, 15*, 917-933.

Bujnicki, J. M. (2004). Molecular phylogenetics of restriction endonucleases. Restriction endonucleases pp. 63-93: Springer.

Bult, C. J., White, O., Olsen, G. J., Zhou, L., Fleischmann, R. D., Sutton, G. G., Blake, J. A., FitzGerald, L. M., Clayton, R. A., and Gocayne, J. D. (1996). Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. Science*, 273*, 1058-1073.

Cady, K. C., Bondy-Denomy, J., Heussler, G. E., Davidson, A. R., and O'Toole, G. A. (2012). The CRISPR/Cas adaptive immune system of *Pseudomonas aeruginosa* mediates resistance to naturally occurring and engineered phages. Journal of Bacteriology*, 194*, 5728-5738.

Campbell, A. (2006). General aspects of lysogeny. The bacteriophages. Oxford University Press, Oxford, 66-73.

Campbell, A. (2007). Phage integration and chromosome structure. A personal history. Annual Review of Genetics*, 41*, 1-11.

Campbell, A. M. (1961). Conditions for the existence of bacteriophage. Evolution*, 15*, 153-165.

Canchaya, C., Fournous, G., and Brussow, H. (2004). The impact of prophages on bacterial chromosomes. Molecular Microbiology*, 53*, 9-18.

Carte, J., Wang, R., Li, H., Terns, R. M., and Terns, M. P. (2008). Cas6 is an endoribonuclease that generates guide RNAs for invader defense in prokaryotes. Genes & Development*, 22*, 3489-3496.

Cerning, J. (1990). Exocellular polysaccharides produced by lactic acid bacteria. FEMS Microbiology Reviews*, 87*, 113-130.

Chanishvili, N. (2012). Phage therapy-history from Twort and d'Herelle through soviet experience to current approaches Vol. 83, pp. 3-40.

Civerolo, E. L., and Kiel, H. L. (1969). Inhibition of bacterial spot of peach foliage by *Xanthomonas pruni* bacteriophage. Phytopathology*, 59*, 1966-1967.

Coleman, M., Pearce, R., Hitchin, E., Busfield, F., Mansfield, J. W., and Roberts, I. S. (1990). Molecular cloning, expression and nucleotide sequence of the *resA* gene of *Erwinia amylovora*, encoding a positive regulator of capsule expression: evidence for a family of related capsule activator proteins. Journal of General Microbiology*, 136*, 1799-1806.

Collier, N. C., Knox, K., and Schlesinger, M. J. (1991). Inhibition of influenza virus formation by a peptide that corresponds to sequences in the cytoplasmic domain of the hemagglutinin. Virology*, 183*, 769-772.

Cooksey, D. A. (1994). Molecular mechanisms of copper resistance and accumulation in bacteria. FEMS Microbiology Reviews*, 14*, 381-386.

Cornelissen, A., Ceyssens, P. J., Krylov, S. N., Noben, J. P., Volckaert, G., and Lavigne, R. (2012a). Identification of EPS-degrading activity within the tail spikes of the novel *Pseudomonas putida* phage AF. Virology*, 434.2*.

Cornelissen, A., Hardies, S. C., Shaburova, O. V., Krylov, V. N., Mattheus, W., Kropinski, A. M., and Lavigne, R. (2012b). Complete genome sequence of the giant virus OBP and comparative genome analysis of the diverse φKZ-related phages. Journal of virology*, 86*, 1844-1852.

Coxe, W. (1817). A view of the cultivation of fruit trees, and the management of orchards and cider. Philadelphia: M. Carey and Son.

Coyier, D., and Covey, R. (1975). Tolerance of *Erwinia amylovora* to streptomycin sulfate in Oregon and Washington [Pears, bacterial diseases]. Plant Disease Reporter (USA).

Crooks, G. E., Hon, G., Chandonia, J.-M., and Brenner, S. E. (2004). WebLogo: a sequence logo generator. Genome Research*, 14*, 1188-1190.

Crosse, J. E., Bennett, M., and Garrett, C. M. E. (1960). Investigation of fire-blight of pear in England. Annals of Applied Biology*, 48*, 541-558.

Datsenko, K. A., Pougach, K., Tikhonov, A., Wanner, B. L., Severinov, K., and Semenova, E. (2012). Molecular memory of prior infections activates the CRISPR/Cas adaptive bacterial immunity system. Nature Communications, *3*, 945.

Deltcheva, E., Chylinski, K., Sharma, C. M., Gonzales, K., Chao, Y., Pirzada, Z. A., Eckert, M. R., Vogel, J., and Charpentier, E. (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. Nature, *471*, 602-607.

Denny, T. P. (1995). Involvement of bacterial polysaccharides in plant pathogenesis. Phytopathology, *33*, 173-197.

Deveau, H., Barrangou, R., Garneau, J. E., Labonté, J., Fremaux, C., Boyaval, P., Romero, D. A., Horvath, P., and Moineau, S. (2008). Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. Journal of Bacteriology, *190*, 1390-1400.

Deveau, H., Van Calsteren, M. R., and Moineau, S. (2002). Effect of exopolysaccharides on phage-host interactions in *Lactococcus lactis*. Applied and Environmental Microbiology, *68*, 4364-4369.

Díez-Villaseñor, C., Guzmán, N. M., Almendros, C., García-Martínez, J., and Mojica, F. J. (2013). CRISPR-spacer integration reporter plasmids reveal distinct genuine acquisition specificities among CRISPR-Cas IE variants of *Escherichia coli*. RNA Biology, *10*, 792-802.

Dimova Aziz, M. (1990). Chemical control of fire blight blossom infection under field conditions in Cyprus. Acta Horticulturae, *No. 273*, 377-382.

Dodd, I. B., Shearwin, K. E., Perkins, A. J., Burr, T., Hochschild, A., and Egan, J. B. (2004). Cooperativity in long-range gene regulation by the λ CI repressor. Genes & Development, *18*, 344-354.

Dong, Q., and Fang, J. (2000). Characterization of polysaccharides from the roots of *Sophora subprostrata*. Zhongguo yao xue za zhi (Zhongguo yao xue hui: 1989), *36*, 85-87.

Drulis-Kawa, Z., Majkowska-Skrobek, G., and Maciejewska, B. (2015). Bacteriophages and phage-derived proteins–application approaches. Current Medicinal Chemistry, *22*, 1757-1773.

Du, Z., and Geider, K. (2002). Characterization of an activator gene upstream of lsc, involved in levan synthesis of Erwinia amylovora. Physiological and Molecular Plant Pathology, *60*, 9-17.

Du, Z., Jakovljevic, V., Salm, H., and Geider, K. (2004). Creation and genetic restoration of *Erwinia amylovora* strains with low levan synthesis. Physiological and Molecular Plant Pathology, *65*, 115-122.

Dunne, C., Delany, I., Fenton, A., and Ogara, F. (1996). Mechanisms involved in biocontrol by microbial inoculants. Agronomie, *16*, 721-729.

Durlu-Ozkaya, F., Aslim, B., and Ozkaya, M. T. (2007). Effect of exopolysaccharides (EPSs) produced by *Lactobacillus delbrueckii* subsp. *bulgaricus* strains to bacteriophage and nisin sensitivity of the bacteria. Lwt-Food Science and Technology, *40*, 564-568.

Durmaz, E., and Klaenhammer, T. R. (2007). Abortive phage resistance mechanism abiz speeds the lysis clock to cause premature lysis of phage-infected *Lactococcus lactis*. Journal of Bacteriology, *189*, 1417-1425.

Eastgate, J. A. (2000). *Erwinia amylovora*: the molecular basis of fireblight disease. Molecular Plant Pathology, *1*, 325-329.

El Nasr, S., Hamdy, I., and Ali, M. H. (1990). Fire blight incidence in pear orchards and its control in Egypt. Acta Horticulturae, *No. 273*, 392.

Emmert, E. A. B., and Handelsman, J. (1999). Biocontrol of plant disease: a (Gram-) positive perspective. FEMS Microbiology Letters, *171*, 1-9.

Erdmann, S., and Garrett, R. A. (2012). Selective and hyperactive uptake of foreign DNA by adaptive immune systems of an archaeon via two distinct mechanisms. Molecular Microbiology, *85*, 1044-1056.

Erdmann, S., Le Moine Bauer, S., and Garrett, R. A. (2014). Inter-viral conflicts that exploit host CRISPR immune systems of *Sulfolobus*. Molecular Microbiology, *91*, 900-917.

Erskine, J. M. (1973). Characteristics of *Erwinia amylovora* bacteriophage and its possible role in the epidemiology of fire blight. [Pears]. Canadian Journal of Microbiology, *19*, 837-845.

Faltus, I. I., and Kishko, Y. G. (1980). Studies on DNA transfection properties of *Erwinia carotovora* virulent and temperate phages. Mikrobiologicheskii Zhurnal, *42*, 339-344.

Fineran, P. C., Blower, T. R., Foulds, I. J., Humphreys, D. P., Lilley, K. S., and Salmond, G. P. C. (2009). The phage abortive infection system, ToxIN, functions as a protein-RNA toxin-antitoxin pair. Proceedings of the National Academy of Sciences of the United States of America, *106*, 894-899.

Flaherty, J. E., Harbaugh, B. K., Jones, J. B., Somodi, G. C., and Jackson, L. E. (2001). H-mutant bacteriophages as a potential biocontrol of bacterial blight of geranium. HortScience, *36*, 98-100.

Flaherty, J. E., Jones, J. B., Harbaugh, B. K., Somodi, G. C., and Jackson, L. E. (2000). Control of bacterial spot on tomato in the greenhouse and field with h-mutant bacteriophages. HortScience, *35*, 882-884.

Fonfara, I., Le Rhun, A., Chylinski, K., Makarova, K. S., Lécrivain, A.-L., Bzdrenga, J., Koonin, E. V., and Charpentier, E. (2014). Phylogeny of Cas9 determines functional exchangeability of dual-RNA and Cas9 among orthologous type II CRISPR-Cas systems. Nucleic Acids Research, *42*, 2577-2590.

Förster, H., McGhee, G. C., Sundin, G. W., and Adaskaveg, J. E. (2015). Characterization of streptomycin resistance in isolates of Erwinia amylovora in California. Phytopathology, *105*, 1302-1310.

Frampton, R. A., Pitman, A. R., and Fineran, P. C. (2012). Advances in bacteriophage-mediated control of plant pathogens. International Journal of Microbiology, 1-11.

Gan, H. M., Sieo, C. C., Tang, S. G. H., Omar, A. R., and Ho, Y. W. (2013). The complete genome sequence of EC1-UPM, a novel N4-like bacteriophage that infects *Escherichia coli* O78: K80. Virology Journal, *10*, 1.

Gander, S., and Gilbert, P. (1997). The development of a small-scale biofilm model suitable for studying the effects of antibiotics on biofilms of gram-negative bacteria. Journal of Antimicrobial Chemotherapy, *40*, 329-334.

Garneau, J. E., Dupuis, M. E., Villion, M., Romero, D. A., Barrangou, R., Boyaval, P., Fremaux, C., Horvath, P., Magadan, A. H., and Moineau, S. (2010). The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. Nature, *468*, 67-71.

Gasiunas, G., Barrangou, R., Horvath, P., and Siksnys, V. (2012). Cas9–crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. Proceedings of the National Academy of Sciences, *109*, E2579-E2586.

Geider, K. (2000). Exopolysaccharides of *Erwinia amylovora*: structure, biosynthesis, regulation, role in pathogenicity of amylovoran and levan, In J. L. Vanneste (ed.). Fire Blight: the disease and its causative agent, *Erwinia amylovora* pp. 117-140. New York: CABI Pub.

Geider, K., Aldridge, P., Bereswill, S., Bugert, P., Langlotz, C., and Bonn, W. G. (1996). Characterization of exopolysaccharide synthesis by *Erwinia amylovora.* Acta Horticulturae, *411*, 259-263.

Geider, K., Geier, G., Bellemann, P., Bernhard, F., Bugert, P., and Metzger, M. (1993). Exopolysaccharides in pathogenicity of *Erwinia amylovora.* Acta Horticulturae, *338*, 255-262.

Geier, G., and Geider, K. (1993). Characterization and influence on virulence of the levansucrase gene from the fireblight pathogen *Erwinia amylovora*. Physiological and Molecular Plant Pathology, *42*, 387-404.

Gill, J. J. (2000). Bacteriophages of *Erwinia amylovra* and their potential use in the biological control of fire blightMSc.: Brock Univeristy.

Gill, J. J., Svircev, A. M., Smith, R., and Castle, A. J. (2003). Bacteriophages of *Erwinia amylovora.* Applied and Environmental Microbiology, *69*, 2133-2138.

Glonti, T., Chanishvili, N., and Taylor, P. (2010). Bacteriophage-derived enzyme that depolymerizes the alginic acid capsule associated with cystic fibrosis isolates of *Pseudomonas aeruginosa*. Journal of Applied Microbiology, *108*, 695-702.

Godde, J. S., and Bickerton, A. (2006). The repetitive DNA elements called CRISPRs and their associated genes: evidence of horizontal transfer among prokaryotes. Journal of Molecular Evolution, *62*, 718-729.

Grissa, I., Vergnaud, G., and Pourcel, C. (2007). The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. BMC Bioinformatics, *8*, 172.

Gross, M., Geier, G., Rudolph, K., and Geider, K. (1992). Levan and levansucrase synthesized by the fireblight pathogen *Erwinia amylovora*. Physiological and Molecular Plant Pathology*, 40*, 371-381.

Guglielmotti, D. M., Reinheimer, J.A., Binetti, AG., Giraffa, G., Carminati, D., Quiberoni, A. (2006). Characterization of spontaneous phage-resistant derivatives of *Lactobacillus delbrueckii* commercial strains. International Journal of Food Microbiology*, 111*, 126-133.

Haft, D. H., Selengut, J., Mongodin, E. F., and Nelson, K. E. (2005). A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. PLoS Computational Biology*, 1*, e60.

Hale, C., Kleppe, K., Terns, R. M., and Terns, M. P. (2008). Prokaryotic silencing (psi) RNAs in *Pyrococcus furiosus*. RNA*, 14*, 2572-2579.

Hale, C. R., Majumdar, S., Elmore, J., Pfister, N., Compton, M., Olson, S., Resch, A. M., Glover, C. V., Graveley, B. R., and Terns, R. M. (2012). Essential features and rational design of CRISPR RNAs that function with the Cas RAMP module complex to cleave RNAs. Molecular Cell*, 45*, 292-302.

Han, D., Lehmann, K., and Krauss, G. (2009). SSO1450–a CAS1 protein from *Sulfolobus solfataricus* P2 with high affinity for RNA and DNA. FEBS Letters*, 583*, 1928-1932.

Hanlon, G. W., Denyer, S. P., Olliff, C. J., and Ibrahim, L. J. (2001). Reduction in exopolysaccharide viscosity as an aid to bacteriophage penetration through *Pseudomonas aeruginosa* biofilms. Applied and Environmental Microbiology*, 67*, 2746-2753.

Harrison, A., and Gibbins, L. N. (1975). The isolation and characterization of a temperate phage, Y46/(E2), from *Erwinia herbicola* Y46. Canadian Journal of Microbiology*, 21*, 937-944.

Hatoum-Aslan, A., Maniv, I., Samai, P., and Marraffini, L. A. (2014). Genetic characterization of antiplasmid immunity through a type III-A CRISPR-Cas system. Journal of Bacteriology*, 196*, 310-317.

Haurwitz, R. E., Jinek, M., Wiedenheft, B., Zhou, K., and Doudna, J. A. (2010). Sequence-and structure-specific RNA processing by a CRISPR endonuclease. Science*, 329*, 1355-1358.

Heidrich, N., and Vogel, J. (2013). Same same but different: new structural insight into CRISPR-Cas complexes. Molecular Cell*, 52*, 4-7.

Hendry, A. T., Carpenter, J. A., and Gerrard, E. H. (1967). Bacteriophage studies of isolates from fire blight sources. Canadian Journal of Microbiology*, 13*, 1357-1364.

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

Hildebrand, M., Aldridge, P., and Geider, K. (2006). Characterization of hns genes from Erwinia amylovora. Mol Genet Genomics*, 275*, 310-319.

Hochstrasser, M. L., Taylor, D. W., Bhat, P., Guegler, C. K., Sternberg, S. H., Nogales, E., and Doudna, J. A. (2014). CasA mediates Cas3-catalyzed target degradation during CRISPR RNA-guided interference. Proceedings of the National Academy of Sciences*, 111*, 6618-6623.

Hockenhull, J. (1979). Fireblight in Denmark from 1968 to 1977. EPPO Bulletin*, 9*, 64-65.

Hoe, N., Nakashima, K., Grigsby, D., Pan, X., Dou, S. J., Naidich, S., Garcia, M., Kahn, E., Bergmire-Sweat, D., and Musser, J. M. (1999). Rapid molecular genetic subtyping of serotype M1 group A *Streptococcus* strains. Emerging Infectious Diseases*, 5*, 254.

Holt, J. G., Krieg, N. R., Sneath, P. H., Staley, J., and Williams, S. (1994). Bergey's Manual of Determinative Bacteriology. Williams and Wilkins, Baltimore, 1-25.

Hommais, F., Krin, E., Laurent-Winter, C., Soutourina, O., Malpertuy, A., Le Caer, J. P., Danchin, A., and Bertin, P. (2001). Large-scale monitoring of pleiotropic regulation of gene expression by the prokaryotic nucleoid-associated protein, H-NS. Molecular Microbiology*, 40*, 20-36.

Horvath, P., Coute-Monvoisin, A. C., Romero, D. A., Boyaval, P., Fremaux, C., and Barrangou, R. (2009). Comparative analysis of CRISPR loci in lactic acid bacteria genomes. International Journal of Food Microbiology*, 131*, 62-70.

Horvath, P., Romero, D. A., Coute-Monvoisin, A. C., Richards, M., Deveau, H., Moineau, S., Boyaval, P., Fremaux, C., and Barrangou, R. (2008).

Diversity, activity, and evolution of CRISPR loci in *Streptococcus thermophilus*. Journal of Bacteriology*, 190*, 1401-1412.

Hsu, C.-R., Lin, T.-L., Pan, Y.-J., Hsieh, P.-F., and Wang, J.-T. (2013). Isolation of a bacteriophage specific for a new capsular type of *Klebsiella pneumoniae* and characterization of its polysaccharide depolymerase. PLoS One*, 8*, e70092.

Hughes, K. A., Sutherland, I. W., Clark, J., and Jones, M. V. (1998a). Bacteriophage and associated polysaccharide depolymerases - novel tools for study of bacterial biofilms. Journal of Applied Microbiology*, 85*, 583-590.

Hughes, K. A., Sutherland, I. W., and Jones, M. V. (1998b). Biofilm susceptibility to bacteriophage attack: the role of phage-borne polysaccharide depolymerase. Microbiology*, 144*, 3039-3047.

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

Ignoffo, C., Rice, W., and McIntosh, A. (1989). Inactivation of nonoccluded and occluded baculoviruses and baculovirus-DNA exposed to simulated sunlight. Environmental Entomology*, 18*, 177-183.

Ignoffo, C. M., and Garcia, C. (1992). Combinations of environmental factors and simulated sunlight affecting activity of inclusion bodies of the heliothis (*lepidoptera: Noctuidae*) nucleopolyhedrosis virus. Environmental Entomology*, 21*, 210-213.

Ignoffo, C. M., Garcia, C., and Saathoff, S. G. (1997). Sunlight stability and rain-fastness of formulations of *Baculovirus heliothis*. Environmental Entomology*, 26*, 1470-1474.

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. Applied and Environmental Microbiology*, 73*, 1704-1711.

Ishimaru, C. A., Klos, E. J., and Brubaker, R. R. (1988). Multiple antibiotic production by *Erwinia herbicola*. Phytopathology*, 78*, 746-750.

Ishino, Y., Shinagawa, H., Makino, K., Amemura, M., and Nakata, A. (1987). Nucleotide sequence of the *iap* gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. Journal of Bacteriology*, 169*, 5429-5433.

Jackson, L. E. (1989). Bacteriophage prevention and control of harmful plant bacteria. Patent. USA.

Jackson, R. N., Lavin, M., Carter, J., and Wiedenheft, B. (2014). Fitting CRISPR-associated Cas3 into the helicase family tree. Current Opinion in Structural Biology, *24*, 106-114.

Jansen, R., Embden, J., Gaastra, W., and Schouls, L. (2002a). Identification of genes that are associated with DNA repeats in prokaryotes. Molecular Microbiology, *43*, 1565-1575.

Jansen, R., van Embden, J. D., Gaastra, W., and Schouls, L. M. (2002b). Identification of a novel family of sequence repeats among prokaryotes. Omics: a Journal of Integrative Biology, *6*, 23-33.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., and Charpentier, E. (2012). A programmable dual-RNA–guided DNA endonuclease in adaptive bacterial immunity. Science, *337*, 816-821.

Jinek, M., Jiang, F., Taylor, D. W., Sternberg, S. H., Kaya, E., Ma, E., Anders, C., Hauer, M., Zhou, K., and Lin, S. (2014). Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. Science, *343*, 1247997.

Jock, S., Donat, V., López, M. M., Bazzi, C., and Geider, K. (2002). Following spread of fire blight in western, central and southern Europe by molecular differentiation of *Erwinia amylovora* strains with PFGE analysis. Environmental Microbiology, *4*, 106-114.

Johnson, K. B., and Stockwell, V. O. (1998). Management of fire blight: a case study in microbial ecology. Annual Review of Phytopathology, *36*, 227-248.

Johnson, K. B., and Stockwell, V. O. (2000). Biological control of fire blight, In J. L. Vanneste (ed.). Fire Blight: the disease and its causative agent, *Erwinia amylovora*, pp. 319-337. New York: CABI Pub.

Johnson, K. B., Stockwell, V. O., McLaughlin, R. J., Sugar, D., Loper, J. E., and Roberts, R. G. (1993). Effect of antagonistic bacteria on establishment of honey bee-dispersed *Erwinia amylovora* in pear blossoms and on fire blight control. Phytopathology, *83*, 995-1002.

Jore, M. M., Lundgren, M., van Duijn, E., Bultema, J. B., Westra, E. R., Waghmare, S. P., Wiedenheft, B., Pul, Ü., Wurm, R., and Wagner, R. (2011). Structural basis for CRISPR RNA-guided DNA recognition by Cascade. Nature Structural & Molecular Biology*, 18*, 529-536.

Judith, R., James, H. N., and Malcolm, F. W. (2013a). CRISPR interference: a structural perspective. Biochemical Journal*, 453*, 155-166.

Judith, R., Richard, D. S., Shirley, G., Huanting, L., James, H. N., and Malcolm, F. W. (2013b). Structure of a dimeric crenarchaeal Cas6 enzyme with an atypical active site for CRISPR RNA processing. Biochemical Journal*, 452*, 223-230.

Jumel, K., Geider, K., and Harding, S. E. (1997). The solution molecular weight and shape of the bacterial exopolysaccharides amylovoran and stewartan. International Journal of Biological Macromolecules*, 20*, 251-258.

Kabaluk, J. T., Svircev, A.M., Goettel, M.S., S.G. Woo. (2010). The use and regulation of of microbial pesticides in representative jurisdictions worldwide. Canada: IOBC Global.

Karginov, F. V., and Hannon, G. J. (2010). The CRISPR system: small RNA-guided defense in bacteria and archaea. Molecular Cell*, 37*, 7-19.

Kawarabayasi, Y., Sawada, M., Horikawa, H., Haikawa, Y., Hino, Y., Yamamoto, S., Sekine, M., Baba, S.-i., Kosugi, H., and Hosoyama, A. (1998). Complete sequence and gene organization of the genome of a hyper-thermophilic archaebacterium, *Pyrococcus horikoshii* OT3. DNA Research*, 5*, 55-76.

Kim, W.-S., and Geider, K. (2000). Characterization of a viral EPS-depolymerase, a potential tool for control of fire blight. Phytopathology*, 90*, 1263-1268.

Kim, W. S., Schollmeyer, M., Nimtz, M., Wray, V., and Geider, K. (2002). Genetics of biosynthesis and structure of the capsular exopolysaccharide from the Asian pear pathogen *Erwinia pyrifoliae*. Microbiology*, 148*, 4015-4024.

Kishko, Y. G., Ruban, V. I., Tovkach, F. I., and Murashchik, I. K. (1982). Electron microscopic structure of temperate phage 59 of *Erwinia carotovora* and its DNA. Soviet Progress in Virology*, 92*, 99-105.

Kleijburg, P. (1979). Present status of fireblight in the Netherlands. EPPO Bulletin, *9*, 69-70.

Klenk, H.-P., Clayton, R. A., Tomb, J.-F., White, O., Nelson, K. E., Ketchum, K. A., Dodson, R. J., Gwinn, M., Hickey, E. K., and Peterson, J. D. (1997). The complete genome sequence of the hyperthermophilic, sulphate-reducing archaeon *Archaeoglobus fulgidus*. Nature, *390*, 364-370.

Koch, A. L. (2007). Evolution of temperate pathogens: the bacteriophage/bacteria paradigm - Art. No. 121. Virology Journal, *4*, 121.

Koczan, J. M., Lenneman, B. R., McGrath, M. J., and Sundin, G. W. (2011). Cell surface attachment structures contribute to biofilm formation and xylem colonization by *Erwinia amylovora*. Applied and Environmental Microbiology, *77*, 7031-7039.

Koczan, J. M., McGrath, M. J., Zhao, Y., and Sundin, G. W. (2009). Contribution of *Erwinia amylovora* exopolysaccharides amylovoran and levan to biofilm formation: implications to pathogenicity. Phytopathology, *99*, 1237-1244.

Koczan, J. M., and Sundin, G. W. (2011). Deletion of *Erwinia amylovora* flagellar motor protein genes *motAB* alters biofilm formation and virulence in apple. Acta Horticulturae, *896*, 203-209.

Koonin, E. V., and Makarova, K. S. (2009). CRISPR-Cas: an adaptive immunity system in prokaryotes. F1000 Biology Reports *1*, 95.

Kulikov, E., Kropinski, A. M., Golomidova, A., Lingohr, E., Govorun, V., Serebryakova, M., Prokhorov, N., Letarova, M., Manykin, A., and Strotskaya, A. (2012). Isolation and characterization of a novel indigenous intestinal N4-related coliphage vB_EcoP_G7C. Virology, *426*, 93-99.

Kunin, V., Sorek, R., and Hugenholtz, P. (2007). Evolutionary conservation of sequence and secondary structures in CRISPR repeats. Genome Biology *8*, R61.

Labrie, S. J., Samson, J. E., and Moineau, S. (2010). Bacteriophage resistance mechanisms. Nature Reviews Microbiology *8*, 317-327.

Lagonenko, A. L., Sadovskaya, O., Valentovich, L. N., and Evtushenkov, A. N. (2015). Characterization of a new ViI-like *Erwinia amylovora* bacteriophage phiEa2809. FEMS Microbiology Letters *362*, 1-6.

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 Disease*, 91*, 871-878.

Langlotz, C., Schollmeyer, M., Coplin, D. L., Nimtz, M., and Geider, K. (2011). Biosynthesis of the repeating units of the exopolysaccharides amylovoran from *Erwinia amylovora* and stewartan from *Pantoea stewartii*. Physiological and Molecular Plant Pathology*, 75*, 163-169.

Lee, S. A., Ngugi, H. K., Halbrendt, N. O., O'Keefe, G., Lehman, B., Travis, J. W., Sinn, J. P., and McNellis, T. W. (2010). Virulence characteristics accounting for fire blight disease severity in apple trees and seedlings. Phytopathology*, 100*, 539-550.

Lehman, S. M. (2007). Development of a bacteriophage-based biopesticide for fire blight. Biological Sciences. PhD. Thesis, St. Catharines, ON Canada: Brock University.

Lehman, S. M., Kropinski, A. M., Castle, A. J., and Svircev, A. M. (2009). Complete genome of the broad-host-range *Erwinia amylovora* phage φEa21-4 and its relationship to Salmonella phage Felix O1. Applied and Environmental Microbiology*, 75*, 2139-2147.

Leigh, J. A., and Coplin, D. L. (1992). Exoploysaccharides in plant-bacterial interactions. Annual Review of Microbiology*, 46*, 307-346.

Lenski, R. E. (1988). Dynamics of interactions between bacteria and virulent bacteriophage Advances in microbial ecology, Vol. 10, pp. 1-44. New York, USA: Plenum Press.

Lightner, G. W., and Steiner, P. W. (1993). An update on version 4.1 of the MARYBLYT R computer program for predicting fire blight. Acta Horticulturae*, 338*, 131-136.

Lillestøl, R. K., Shah, S. A., Brügger, K., Redder, P., Phan, H., Christiansen, J., and Garrett, R. A. (2009). CRISPR families of the crenarchaeal genus *Sulfolobus*: bidirectional transcription and dynamic properties. Molecular Microbiology*, 72*, 259-272.

Lindberg, A. A. (1977). Bacterial surface carbohydrates and bacteriophage adsorption, In I. Sutherland (ed.). Surface carbohydrates of the prokaryotic cell. London: Academic Press.

Liu, T., Li, Y., Wang, X., Ye, Q., Li, H., Liang, Y., She, Q., and Peng, N. (2015). Transcriptional regulator-mediated activation of adaptation genes triggers CRISPR de novo spacer acquisition. Nucleic Acids Research, *43*, 1044-1055.

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C_T}$ method. Methods, *25*, 402-408.

Looijesteijn, P., Trapet, L., de Vries, E., Abee, T., and Hugenholtz, J. (2001). Physiological function of exopolysaccharides produced by *Lactococcus lactis*. International Journal of Food Microbiology, *64*, 71-80.

Loper, J. E., Henkels, H. D., Roberts, R. G., Grove, G. G., Willet, M. J., and Smith, T. J. (1991). Evaluation of streptomycin, oxytetracycline and copper resistance of *Erwinia amylovora* isolated from orchards in Washington State. Plant Disease, *75*, 287-290.

Lopez-Sanchez, M. J., Sauvage, E., Da Cunha, V., Clermont, D., Ratsima Hariniaina, E., Gonzalez-Zorn, B., Poyart, C., Rosinski-Chupin, I., and Glaser, P. (2012). The highly dynamic CRISPR1 system of Streptococcus agalactiae controls the diversity of its mobilome. Molecular Microbiology, *85*, 1057-1071.

Luna, A. J., Wood, T. L., Chamakura, K. R., and Everett, G. F. K. (2013). Complete genome of *Salmonella enterica* serovar *Enteritidis* myophage Marshall. Genome Announcements, *1*, e00867-00813.

Maes, M., Orye, K., Bobev, S., Devreese, B., Van Beeumen, J., De Bruyn, A., Busson, R., Herdewijn, P., Morreel, K., and Messens, E. (2001). Influence of amylovoran production on virulence of *Erwinia amylovora* and a different amylovoran structure in *E. amylovora* isolates from *Rubus*. European Journal of Plant Pathology, *107*, 839-844.

Makarova, K. S., Grishin, N. V., Shabalina, S. A., Wolf, Y. I., and Koonin, E. V. (2006). A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. Biology Direct, *1*, 1-26.

Makarova, K. S., Haft, D. H., Barrangou, R., Brouns, S. J., Charpentier, E., Horvath, P., Moineau, S., Mojica, F. J., Wolf, Y. I., and Yakunin, A. F. (2011). Evolution and classification of the CRISPR–Cas systems. Nature Reviews Microbiology, *9*, 467-477.

Malek, W. (1990). Properties of the transducing phage MI of *Rhizobium meliloti*. Journal of Basic Microbiology*, 30*, 43-50.

Mallmann, W., and Hemstreet, C. (1924). Isolation of an inhibitory substance from plants. Agricultural Research*, 28*, 599-602.

Manica, A., and Schleper, C. (2013). CRISPR-mediated defense mechanisms in the hyperthermophilic archaeal genus *Sulfolobus*. RNA Biology*, 10*, 671-678.

Mann, R., Smits, T. H. M., Buhlmann, A., Blom, J., Goesmann, A., Frey, J., Plummer, K. M., Beer, S. V., Luck, J., Duffy, B., and Rodoni, B. C. (2013). Comparitive genomics of 12 strains of *Erwinia amylovora* identifies a pan-genome with a large conserved core. PLoS One*, 8*, 1-11.

Manso, T., Nunes, C., Raposo, S., and Lima-Costa, M. E. (2010). Carob pulp as raw material for production of the biocontrol agent *P. agglomerans* PBC-1. Journal of Industrial Microbiology and Biotechnology*, 37*, 1145-1155.

Marks, T., and Sharp, R. (2000). Bacteriophages and biotechnology: a review. Journal of Chemical Technology and Biotechnology*, 75*, 6-17.

Marraffini, L. A., and Sontheimer, E. J. (2008). CRISPR interference limits horizontal gene transfer in *staphylococci* by targeting DNA. Science*, 322*, 1843-1845.

Marraffini, L. A., and Sontheimer, E. J. (2010a). CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. Nature Reviews Genetics*, 11*, 181-190.

Marraffini, L. A., and Sontheimer, E. J. (2010b). Self versus non-self discrimination during CRISPR RNA-directed immunity. Nature*, 463*, 568-571.

Massey, R. (1934). Studies on blackarm disease of cotton III. Empire Cotton Growing Rev*, 11*, 188-193.

Mauro, S., and Koudelka, G. B. (2011). Shiga toxin: expression, distribution, and its role in the environment. Toxins*, 3*, 608-625.

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.

McGhee, G. C., and Sundin, G. W. (2012). *Erwinia amylovora* CRISPR elements provide new tools for evaluating strain diversity and for microbial source tracking. PLoS One*, 7*, e41706.

Mckenna, F., Eltarabily, K. A., Hardy, G. E. S., and Dell, B. (2001). Novel *in vivo* use of a polyvalent *Streptomyces* phage to disinfest *Streptomyces scabies*-infected seed potatoes. Plant Pathology*, 50*, 666-675.

McManus, P. S., and Jones, A. L. (1994). Epidemiology and genetic analysis of streptomycin-resistant *Erwinia amylovora* from Michigan and evaluation of oxytetracycline for control. Phytopathology*, 84*, 627-633.

McNeil, D. L., Romero, S., Kandula, J., Stark, C., Stewart, A., and Larsen, S. (2001). Bacteriophages: a potential biocontrol agent against walnut blight (*Xanthomonas campestris* pv. *juglandis*). New Zealand Plant Protection*, 54*, 220-224.

Meczker, K., Domotor, D., Vass, J., Rakhely, G., and Schneider, G. (2014). The genome of the *Erwinia amylovora* phage PhiEaH1 reveals greater diversity and broadens the aplicability of phages for the treatment of fire blight. FEMS Microbiology Letters*, 350*, 25-27.

Meijneke, C. A. R. (1972). Fireblight of pears and its spread. Gewasbescherming*, 6*, 128-136.

Mesyanzhinov, V. V., Robben, J., Grymonprez, B., Kostyuchenko, V. A., Bourkaltseva, M. V., Sykilinda, N. N., Krylov, V. N., and Volckaert, G. (2002). The genome of bacteriophage phi KZ of *Pseudomonas aeruginosa*. Journal of Molecular Biology*, 317*, 1-19.

Miller, P. (1929). Studies of fire blight of apple in Wisconsin. Journal of Agriculture Research*, 39*, 579-621.

Miller, T. D., and Schroth, M. N. (1972). Monitoring the ephiphytic population of *Erwinia amylovora* on pear with a selective medium. Phytopathology*, 62*, 1175-1182.

Mojica, F. J., Diez-Villasenor, C., Garcia-Martinez, J., and Almendros, C. (2009). Short motif sequences determine the targets of the prokaryotic CRISPR defence system. Microbiology*, 155*, 733-740.

Mojica, F. J., Diez-Villasenor, C., Garcia-Martinez, J., and Soria, E. (2005). Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. Journal of Molecular Evolution, *60*, 174-182.

Mole, B. M., Baltrus, D. A., Dangl, J. L., and Grant, S. R. (2007). Global virulence regulation networks in phytopathogenic bacteria. Trends in Microbiology, *15*, 363-371.

Monson, R., Foulds, I., Foweraker, J., Welch, M., and Salmond, G. P. (2011). The *Pseudomonas aeruginosa* generalized transducing phage φPA3 is a new member of the φKZ-like group of 'jumbo'phages, and infects model laboratory strains and clinical isolates from cystic fibrosis patients. Microbiology, *157*, 859-867.

Moore, E. (1926). D'Herelle's bacteriophage in relation to plant parasites. South African Journal of Science, *23*, 306.

Mulepati, S., and Bailey, S. (2011). Structural and biochemical analysis of nuclease domain of clustered regularly interspaced short palindromic repeat (CRISPR)-associated protein 3 (Cas3). Journal of Biological Chemistry, *286*, 31896-31903.

Mulepati, S., Orr, A., and Bailey, S. (2012). Crystal structure of the largest subunit of a bacterial RNA-guided immune complex and its role in DNA target binding. Journal of Biological Chemistry, *287*, 22445-22449.

Müller, I., Kube, M., Reinhardt, R., Jelkmann, W., and Geider, K. (2011a). Complete genome sequences of three *Erwinia amylovora* phages isolated in North America and a bacteriophage induced from an *Erwinia tasmaniensis* strain. Journal of Bacteriology, *193*, 795-796.

Müller, I., Lurz, R., Kube, M., Quedenau, C., Jelkmann, W., and Geider, K. (2011b). Molecular and physiological properties of bacteriophages from North America and Germany affecting the fire blight pathogen *Erwinia amylovora*. Microbial Biotechnology, *4*, 735-745.

Muniesa, M., Colomer-Lluch, M., and Jofre, J. (2013). Potential impact of environmental bacteriophages in spreading antibiotic resistance genes. Future Microbiology, *8*, 739-751.

Munsch, P., Olivier, J. M., and Elliott, T. J. (1995). Biocontrol of bacterial blotch of the cultivated mushroom with lytic phages: some practical considerations. Science and cultivation of edible fungi, volume 2:

Proceedings of the 14th International Congress, Oxford, 17-22 September 1995., 595-602.

Murugaiyan, S., Bae, J. Y., Wu, J., Lee, S. D., Um, H. Y., Choi, H. K., Chung, E., Lee, J. H., and Lee, S. W. (2011). Characterization of filamentous bacteriophage PE226 infecting *Ralstonia solanacearum* strains. Journal of Applied Microbiology*, 110*, 296-303.

Nagy, J., Király, L., and Schwarczinger, I. (2012). Phage therapy for plant disease control with a focus on fire blight. Central European Journal of Biology*, 7*, 1-12.

Naim, H. Y., and Roth, M. G. (1993). Basis for selective incorporation of glycoproteins into the influenza virus envelope. Journal of virology*, 67*, 4831-4841.

Nam, K. H., Ding, F., Haitjema, C., Huang, Q., DeLisa, M. P., and Ke, A. (2012a). Double-stranded endonuclease activity in *Bacillus halodurans* clustered regularly interspaced short palindromic repeats (CRISPR)-associated Cas2 protein. Journal of Biological Chemistry*, 287*, 35943-35952.

Nam, K. H., Haitjema, C., Liu, X., Ding, F., Wang, H., DeLisa, M. P., and Ke, A. (2012b). Cas5d protein processes pre-crRNA and assembles into a cascade-like interference complex in subtype IC/Dvulg CRISPR-Cas system. Structure*, 20*, 1574-1584.

Nelson, K. E., Clayton, R. A., Gill, S. R., Gwinn, M. L., Dodson, R. J., Haft, D. H., Hickey, E. K., Peterson, J. D., Nelson, W. C., and Ketchum, K. A. (1999). Evidence for lateral gene transfer between Archaea and bacteria from genome sequence of *Thermotoga maritima*. Nature*, 399*, 323-329.

Nimtz, M., Mort, A., Domke, T., Wray, V., Zhang, Y. X., Qiu, F., Coplin, D., and Geider, K. (1996). Structure of amylovoran, the capsular exopolysaccharide from the fire blight pathogen *Erwinia amylovora.* Carbohydrate Research*, 287*, 59-76.

Nischwitz, C., and Dhiman, C. (2013). Streptomycin resistance of Erwinia amylovora isolated from apple (Malus domesticus) in Utah. Plant Health Progress.

Nishimasu, H., Ran, F. A., Hsu, P. D., Konermann, S., Shehata, S. I., Dohmae, N., Ishitani, R., Zhang, F., and Nureki, O. (2014). Crystal structure of Cas9 in complex with guide RNA and target DNA. Cell*, 156*, 935-949.

Nordström, K., and Forsgren, A. (1974). Effect of protein A on adsorption of bacteriophages to *Staphylococcus aureus*. Journal of Virology*, 14*, 198-202.

Norelli, J. L., Jones, A. L., and Aldwinckle, H. S. (2003). Fire blight management in the twenty-first century: using new technologies that enhance host resistance in apple. Plant Disease*, 87*, 756-765.

Nuñez, J. K., Kranzusch, P. J., Noeske, J., Wright, A. V., Davies, C. W., and Doudna, J. A. (2014). Cas1–Cas2 complex formation mediates spacer acquisition during CRISPR–Cas adaptive immunity. Nature Structural & Molecular Biology*, 21*, 528-534.

Obradovic, A., Jones, J. B., Momol, M. T., Balogh, B., and Olson, S. M. (2004). Management of tomato bacterial spot in the field by foliar applications of bacteriophages and SAR inducers. Plant Disease*, 88*, 736-740.

Oh, C.-S., and Beer, S. V. (2005). Molecular genetics of *Erwinia amylovora* involved in the development of fire blight. FEMS Microbiology Letters*, 253*, 185-192.

Oh, C.-S., Kim, J. F., and Beer, S. V. (2005). The Hrp pathogenicity island of *Erwinia amylovora* and identification of three novel genes required for systemic infection. Molecular Plant Pathology*, 6*, 125-138.

Ohmori, H., Haynes, L. L., and Rothman-Denes, L. B. (1988). Structure of the ends of the coliphage N4 genome. Journal of Molecular Biology*, 202*, 1-10.

Okabe, N., and Goto, M. (1963). Bacteriophages of plant pathogens. Annual Review of Phytopathology, 397-418.

Oppenheim, A. B., and Adhya, S. L. (2007). A new look at bacteriophage λ genetic networks. Journal of Bacteriology*, 189*, 298-304.

Ordax, M., Marco-Noales, E., López, M. M., and Biosca, E. G. (2010). Exopolysaccharides favor the survival of *Erwinia amylovora* under copper stress through different strategies. Research in Microbiology*, 161*, 549-555.

Overbeek, R., Olson, R., Pusch, G. D., Olsen, G. J., Davis, J. J., Disz, T., Edwards, R. A., Gerdes, S., Parrello, B., and Shukla, M. (2014). The

SEED and the rapid annotation of microbial genomes using subsystems technology (RAST). Nucleic Acids Research, *42*, D206-D214.

Paddison, P., Abedon, S. T., Dressman, H. K., Gailbreath, K., Tracy, J., Mosser, E., Neitzel, J., Guttman, B., and Kutter, E. (1998). The roles of the bacteriophage T4 r genes in lysis inhibition and fine-structure genetics: a new perspective. Genetics, *148*, 1539-1550.

Palmer, J., Flint, S., and Brooks, J. (2007). Bacterial cell attachment, the beginning of a biofilm. Journal of Industrial Microbiology & Biotechnology, *34*, 577-588.

Paulin, J. P. (2000). *Erwinia amylovora*: General characteristics, biochemistry and serology, In J. L. Vanneste (ed.). Fire Blight: the disease and its causative agent, *Erwinia amylovora*, pp. 87-115. New York: CABI Pub.

Pawluk, A., Bondy-Denomy, J., Cheung, V. H., Maxwell, K. L., and Davidson, A. R. (2014). A new group of phage anti-CRISPR genes inhibits the type I-E CRISPR-Cas system of *Pseudomonas aeruginosa*. MBio, *5*, e00896.

Pelkonen, S., Aalto, J., and Finne, J. (1992). Differential activities of bacteriophage depolymerase on bacterial polysaccharide: Binding is essential but degradation is inhibitory in phage infection of K-1 defective *Escherichia coli*. Journal of Bacteriology, *174*, 7757-7761.

Pingoud, A., Fuxreiter, M., Pingoud, V., and Wende, W. (2005). Type II restriction endonucleases: structure and mechanism. Cellular and Molecular Life Sciences, *62*, 685-707.

Piqué, N., Miñana-Galbis, D., Merino, S., and Tomás, J. M. (2015). Virulence factors of *Erwinia amylovora*: a review. International Journal of Molecular Sciences, *16*, 12836-12854.

Pires, D. P., Oliveira, H., Melo, L. D., Sillankorva, S., and Azeredo, J. (2016). Bacteriophage-encoded depolymerases: their diversity and biotechnological applications. Applied Microbiology and Biotechnology, *100*, 2141-2151.

Plagens, A., Tjaden, B., Hagemann, A., Randau, L., and Hensel, R. (2012). Characterization of the CRISPR/Cas subtype IA system of the hyperthermophilic crenarchaeon *Thermoproteus tenax*. Journal of Bacteriology, *194*, 2491-2500.

Poetter, K., and Coplin, D. L. (1991). Structural and functional analysis of the rcsA gene from *Erwinia stewartii*. Molecular and General Genetics MGG*, 229*, 155-160.

Pougach, K., Semenova, E., Bogdanova, E., Datsenko, K. A., Djordjevic, M., Wanner, B. L., and Severinov, K. (2010). Transcription, processing and function of CRISPR cassettes in *Escherichia coli*. Molecular Microbiology*, 77*, 1367-1379.

Pourcel, C., Salvignol, G., and Vergnaud, G. (2005). CRISPR elements in *Yersinia pestis* acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. Microbiology*, 151*, 653-663.

Powney, R., Smits, T. H. M., Sawbridge, T., Frey, B., Blom, J., Frey, J. E., Plummer, K. M., Beer, S. V., Luck, J., Duffy, B., and Rodoni, B. (2011). Genome sequence of an *Erwinia amylovora* strain with pathogenicity restricted to *Rubus* plants. Journal of Bacteriology*, 193*, 785-786.

Ptashne, M. (2004). A genetic switch: phage lambda revisited. New York, USA: Cold Spring Harbor Laboratory Press Cold Spring Harbor.

Pul, Ü., Wurm, R., Arslan, Z., Geißen, R., Hofmann, N., and Wagner, R. (2010). Identification and characterization of *E. coli* CRISPR-cas promoters and their silencing by H-NS. Molecular Microbiology*, 75*, 1495-1512.

Pusey, P. L. (2002). Biological control agents for fire blight of apple compared under conditions limiting natural dispersal. Plant Disease*, 86*, 639-644.

Quax, T. E., Voet, M., Sismeiro, O., Dillies, M.-A., Jagla, B., Coppée, J.-Y., Sezonov, G., Forterre, P., Van Der Oost, J., and Lavigne, R. (2013). Massive activation of archaeal defense genes during viral infection. Journal of Virology*, 87*, 8419-8428.

Rairakhwada, D., Seo, J.-W., Seo, M.-Y., Kwon, O., Rhee, S.-K., and Kim, C. H. (2010). Gene cloning, characterization, and heterologous expression of levansucrase from *Bacillus amyloliquefaciens*. Journal of Industrial Microbiology & Biotechnology*, 37*, 195-204.

Ramey, B. E., Koutsoudis, M., Bodman, S. B. V., and Fuqua, C. (2004). Biofilm formation in plant–microbe associations. Current Opinion in Microbiology*, 7*, 602-609.

Rath, D., Amlinger, L., Rath, A., and Lundgren, M. (2015). The CRISPR-Cas immune system: Biology, mechanisms and applications. Biochimie, *117*, 119-128.

Remminghorst, U., and Rehm, B. H. (2006). Bacterial alginates: from biosynthesis to applications. Biotechnology Letters, *28*, 1701-1712.

Rezzonico, F., Smits, T. H. M., and Duffy, B. (2011). Diversity, evolution, and functionality of clustered regularly interspaced short palindromic repeat (CRISPR) regions in the fire blight pathogen *Erwinia amylovora*. Applied and Environmental Microbiology, *77*, 3819-3829.

Rezzonico, F., Smits, T. H. M., Pelludat, C., Montesinos, E., Frey, J. E., and Duffy, B. (2009). Genotypic comparison of Pantoea agglomerans biocontrol and clinical isolates to address taxonomic and bio-safety questions. IOBC/WPRS Bulletin, *43*, 35-39.

Richter, C., Chang, J. T., and Fineran, P. C. (2012a). Function and regulation of clustered regularly interspaced short palindromic repeats (CRISPR) / CRISPR associated (Cas) systems. Viruses, *4*, 2291-2311.

Richter, C., Gristwood, T., Clulow, J. S., and Fineran, P. C. (2012b). *In vivo* protein interactions and complex formation in the *Pectobacterium atrosepticum* subtype I-F CRISPR/Cas System. PLoS One, *7*, 1-11.

Richter, H., Randau, L., and Plagens, A. (2013). Exploiting CRISPR/Cas: interference mechanisms and applications. International Journal of Molecular Sciences, *14*, 14518-14531.

Ritchie, D. F. (1978). Bacteriophages of *Erwinia amyovora*: Their isolation, distribution, characterization, and possible involvement in the etiology and epidemiology of fire blight. PhD. Thesis, East Lansing, MI USA: Michigan State University.

Ritchie, D. F., and Klos, E. J. (1977). Isolation of *Erwinia amylovora* bacteriophage from aerial parts of apple trees. Phytopathology, *67*, 101-104.

Ritchie, D. F., and Klos, E. J. (1979). Some properties of *Erwinia amylovora* bacteriophages. Phytopathology, *69*, 1078-1083.

Ritz, M. P., Perl, A. L., Colquhoun, J. M., Chamakura, K. R., and Everett, G. F. K. (2013). Complete genome of *Bacillus subtilis* myophage CampHawk. Genome Announcements, *1*, e00984-00913.

Roach, D. R., Sjaarda, D., Castle, A. J., and Svircev, A. M. (2011). Bacteriophages as biopesticides: role of bacterial exopolysaccharides. Acta Horticulturae, *896*, 449-455.

Roach, D. R., Sjaarda, D. R., Castle, A. J., and Svircev, A. M. (2013). Host exopolysaccharide quantity and composition impact *Erwinia amylovora* bacteriophage pathogenesis. Applied and Environmental Microbiology, *79*, 3249-3256.

Roach, D. R., Sjaarda, D. R., Sjaarda, C. P., Ayala, C. J., Howcroft, B., Castle, A. J., and Svircev, A. M. (2015). Absence of lysogeny in wild populations of *Erwinia amylovora* and *Pantoea agglomerans*. Microbial Biotechnology, *8*, 510-518.

Roach, D. W. (2011). *Erwinia amylovora* bacteriophage resistance. Biological Sciences. PhD. Thesis, St. Catharines, ON Canada: Brock University.

Roberts, R., Hale, C., Van der Zwet, T., Miller, C., and Redlin, S. (1998). The potential for spread of *Erwinia amylovora* and fire blight via commercial apple fruit; a critical review and risk assessment. Crop Protection, *17*, 19-28.

Ruban, V. I., Tovkach, F. I., Seredyuk, R. M., and Kishko, Y. G. (1981). Properties of temperate phage 59 of *Erwinia carotovora*, its DNA and proteins. Mikrobiologicheskii Zhurnal, *43*, 88-92.

Russo, N. L., Burr, T. J., Breth, D. I., and Aldwinckle, H. S. (2008). Isolation of streptomycin-resistant isolates of *Erwinia amylovora* in New York. Plant Disease, *92*, 714-718.

Sapranauskas, R., Gasiunas, G., Fremaux, C., Barrangou, R., Horvath, P., and Siksnys, V. (2011). The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli*. Nucleic Acids Research, gkr606.

Sashital, D. G., Jinek, M., and Doudna, J. A. (2011). An RNA-induced conformational change required for CRISPR RNA cleavage by the endoribonuclease Cse3. Nature Structural and Molecular Biology, *18*, 680-687.

Sashital, D. G., Wiedenheft, B., and Doudna, J. A. (2012). Mechanism of foreign DNA selection in a bacterial adaptive immune system. Molecular Cell, *46*, 606-615.

Schattner, P., Brooks, A. N., and Lowe, T. M. (2005). The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. Nucleic Acids Research, *33*, W686-W689.

Schito, G., Rialdi, G., and Pesce, A. (1966). Biophysical properties of N 4 coliphage. Biochimica et Biophysica Acta (BBA)-Nucleic Acids and Protein Synthesis, *129*, 482-490.

Schnabel, E. L., Fernando, W. G. D., Meyer, M. P., Jones, A. L., and Jackson, L. E. (1999). Bacteriophage of *Erwinia amylovora* and their potential for biocontrol. Acta Horticulturae, 649-653.

Schnabel, E. L., and Jones, A. L. (1999). Distribution of tetracycline resistance genes and transposons among phylloplane bacteria in Michigan apple orchards. Applied and Environmental Microbiology, *65*, 4898-4907.

Schnabel, E. L., and Jones, A. L. (2001). Isolation and characterization of five *Erwinia amylovora* bacteriophages and assessment of phage resistance in strains of *Erwinia amylovora*. Applied and Environmental Microbiology, *67*, 59-64.

Scholl, D., Adhya, S., and Merril, C. (2005). *Escherichia coli* K1's capsule is a barrier to bacteriophage T7. Applied and Environmental Microbiology, *71*, 4872-4874.

Schroth, M. N., Thomson, S. V., and Moller, W. J. (1979). Streptomycin resistance in *Erwinia amylovora*. Phytopathology, *69*, 565-568.

Scott Jr, T. A., and Melvin, E. H. (1953). Determination of dextran with anthrone. Analytical Chemistry, *25*, 1656-1661.

Sebaihia, M., Bocsanczy, A. M., Biehl, B. S., Quail, M. A., Perna, N. T., Glasner, J. D., DeClerck, G. A., Cartinhour, S., Schneider, D. J., Bentley, S. D., Parkhill, J., and Beer, S. V. (2010). Complete genome sequence of the plant pathogen *Erwinia amylovora* strain ATCC 49946. Journal of Bacteriology, *192*, 2020-2021.

Seed, K. D., Lazinski, D. W., Calderwood, S. B., and Camilli, A. (2013). A bacteriophage encodes its own CRISPR/Cas adaptive response to evade host innate immunity. Nature, *494*, 489-491.

Semenova, E., Jore, M. M., Datsenko, K. A., Semenova, A., Westra, E. R., Wanner, B., van der Oost, J., Brouns, S. J., and Severinov, K. (2011). Interference by clustered regularly interspaced short palindromic repeat

(CRISPR) RNA is governed by a seed sequence. Proceedings of the National Academy of Sciences of the United States of America, *108*, 10098-10103.

Semenova, E., Nagornykh, M., Pyatnitskiy, M., Artamonova, I., and Severinov, K. (2009). Analysis of CRISPR system function in plant pathogen *Xanthomonas oryzae*. FEMS Microbiology Letters, *296*, 110-116.

Shaburova, O. V., Burkaltseva, M. V., Pleteneva, E. A., and Krylov, V. N. (2000). Temperate bacteriophages for rhizospheric pseudomonades *Pseudomonas putida*: Isolation and comparative study. Russian Journal of Genetics, *36*, 753-756.

Shah, S. A., Erdmann, S., Mojica, F. J., and Garrett, R. A. (2013). Protospacer recognition motifs: mixed identities and functional diversity. RNA Biology, *10*, 891-899.

Shah, S. A., Hansen, N. R., and Garrett, R. A. (2009). Distribution of CRISPR spacer matches in viruses and plasmids of crenarchaeal acidothermophiles and implications for their inhibitory mechanism. Biochemical Society Transactions, *37*, 23-28.

Sholberg, P. L., Bedford, K. E., Haag, P., and P., R. (2001). Survey of *Erwinia amylovora* isolates from British Columbia for resistance to bactericides and virulence on apple. Canadian Journal of Plant Pathology, *23*, 60-67.

Simpson, D. A., and Lamb, R. (1992). Alterations to influenza virus hemagglutinin cytoplasmic tail modulate virus infectivity. Journal of Virology, *66*, 790-803.

Sinkunas, T., Gasiunas, G., Fremaux, C., Barrangou, R., Horvath, P., and Siksnys, V. (2011). Cas3 is a single-stranded DNA nuclease and ATP-dependent helicase in the CRISPR/Cas immune system. The EMBO Journal, *30*, 1335-1342.

Sinkunas, T., Gasiunas, G., Waghmare, S. P., Dickman, M. J., Barrangou, R., Horvath, P., and Siksnys, V. (2013). *In vitro* reconstitution of Cascade-mediated CRISPR immunity in *Streptococcus thermophilu*s. The EMBO Journal, *32*, 385-394.

Sjaarda, D. (2012). Role of exopolysaccharides and monosaccharides in *Erwinia amylovora* resistance to bacteriophages. Biological Sciences. MSc. Thesis, St. Carharines, ON Canada: Brock University.

Sjulin, T. M., and Beer, S. V. (1978). Mechanism of wilt induction by amylovorin in cotoneaster shoots and its relation to wilting of shoots infected by *Erwinia amylovora*. Phytopathology*, 68*, 89-94.

Smith, H. S., Pizer, L. I., Pylkas, L., and Lederberg, S. (1969). Abortive infection of Shigella dysenteriae P2 by T2 bacteriophage. Journal of Virology*, 4*, 162-168.

Smits, T. H. M., Rezzonico, F., Kamber, T., Blom, J., Goesmann, A., Frey, J. E., and Duffy, B. (2010). Complete genome sequence of the fire blight pathogen *Erwinia amylovora* CFBP 1430 and comparison to other *Erwinia* spp. Molecular Plant-Microbe Interactions*, 23*, 384-393.

Sorek, R., Kunin, V., and Hugenholtz, P. (2008). CRISPR - a widespread system that provides acquired resistance against phages in bacteria and archaea. Nature Reviews Microbiology*, 6*, 181-186.

Sorek, R., Lawrence, C. M., and Wiedenheft, B. (2013). CRISPR-mediated adaptive immune systems in bacteria and archaea, Vol. 82, pp. 237-266.

Spilman, M., Cocozaki, A., Hale, C., Shao, Y., Ramia, N., Terns, R., Terns, M., Li, H., and Stagg, S. (2013). Structure of an RNA silencing complex of the CRISPR-Cas immune system. Molecular Cell*, 52*, 146-152.

St-Pierre, F., and Endy, D. (2008). Determination of cell fate selection during phage lambda infection. Proceedings of the National Academy of Sciences*, 105*, 20705-20710.

Stapp, C. (1961). Bacterial Plant Pathogens. Oxford University Press, London, 292.

Steinberger, E. M., and Beer, S. V. (1988). Creation and complementation of pathogenicity mutants of *Erwinia amylovora*. Molecular Plant-Microbe Interactions*, 1*, 135-144.

Steiner, P. W. (2000). Integrated orchard and nursery management for the control of fire blight, In J. L. Vanneste (ed.). Fire Blight: the disease and its causative agent, *Erwinia amylovora*, pp. 339-358. New York: CABI Pub.

Stewart, F. M., and Levin, B. R. (1984). The population biology of bacterial viruses: why be temperate. Theoretical Population Biology*, 26*, 93-117.

Stirm, S. (1968). *Escherichia coli* K bacteriophages, I. Isolation and introductory characterization of five *Escherichia coli* K bacteriophages. Journal of Virology*, 2*, 1107-1114.

Stockwell, V. O. (2014). Overview of concerns surrounding antibiotic use for control of fire blight. Acta Horticulturae Vol. 1056, pp. 39-42.

Stockwell, V. O., Johnson, K. B., Sugar, D., and Loper, J. E. (2011). Mechanistically compatible mixtures of bacterial antagonists improve biological control of fire blight of pear. Phytopathology*, 101*, 113-123.

Stummeyer, K., Schwarzer, D., Claus, H., Vogel, U., Gerardy-Schahn, R., and Muhlenhoff, M. (2006). Evolution of bacteriophages infecting encapsulated bacteria: lessons from *Escherichia coli* K1-specific phages. Molecular Microbiology*, 60*, 1123-1135.

Summers, W. C. (2001). Bacteriophage therapy. Annual Review of Microbiology*, 55*, 437-451.

Summers, W. C. (2005). Bacteriophage research: Early history, In E. E. Kutter and A. Sulakvelidze (eds.). Bacteriophages: biology and applications, pp. 5-27. Boca Raton, Fl.: CRC Press.

Sutherland, I. W. (1995). Polysaccharide lyases. FEMS Microbiology Reviews*, 16*, 323-347.

Sutherland, I. W., Hughes, K. A., Skillman, L. C., and Tait, K. (2004). The interaction of phage and biofilms. FEMS Microbiology Letters*, 232*, 1-6.

Svircev, A. M., Castle, A. J., and Lehman, S. M. (2010). Bacteriophages for control of phytopathogens in food production systems, In P. M. Sabour and M. W. Griffiths (eds.). Bacteriophages in the control of food- and waterborne pathogens, pp. 79-102. Washington: ASM Press.

Swarts, D. C., Mosterd, C., van Passel, M. W., and Brouns, S. J. (2012). CRISPR interference directs strand specific spacer acquisition. PLoS One*, 7*, e35888.

Szczepankowska, A. (2012). Role of CRISPR/cas System in the Development of Bacteriophage Resistance. Advances in Virus Research Vol. 82, pp. 289-338.

Tait, K., Skillman, L., and Sutherland, I. (2002). The efficacy of bacteriophage as a method of biofilm eradication. Biofouling, *18*, 305-311.

Tanaka, H., Negishi, H., and Maeda, H. (1990). Control of tobacco bacterial wilt by an avirulent strain of *Pseudomonas solanacearum* M4S and its bacteriophage. Annals of the Phytophological Society of Japan, *56*, 243-246.

Tancos, K. A., Villani, S., Kuehne, S., Borejsza-Wysocka, E., Breth, D., Carol, J., Aldwinckle, H. S., and Cox, K. D. (2016). Prevalence of streptomycin-resistant Erwinia amylovora in New York apple orchards. Plant Disease, *100*, 802-809.

Tang, T.-H., Bachellerie, J.-P., Rozhdestvensky, T., Bortolin, M.-L., Huber, H., Drungowski, M., Elge, T., Brosius, J., and Hüttenhofer, A. (2002). Identification of 86 candidates for small non-messenger RNAs from the archaeon *Archaeoglobus fulgidus*. Proceedings of the National Academy of Sciences, *99*, 7536-7541.

Tang, T. H., Polacek, N., Zywicki, M., Huber, H., Brugger, K., Garrett, R., Bachellerie, J. P., and Hüttenhofer, A. (2005). Identification of novel non-coding RNAs as potential antisense regulators in the archaeon *Sulfolobus solfataricus*. Molecular microbiology, *55*, 469-481.

Tharaud, M., Laurent, J., Faize, M., and Paulin, J. P. (1997). Fire blight protection with avirulent mutants of *Erwinia amylovora*. Microbiology, *143*, 625-632.

Thomas, J. A., Rolando, M. R., Carroll, C. A., Shen, P. S., Belnap, D. M., Weintraub, S. T., Serwer, P., and Hardies, S. C. (2008). Characterization of *Pseudomonas chlororaphis* myovirus 201φ2-1 via genomic sequencing, mass spectrometry, and electron microscopy. Virology, *376*, 330-338.

Thomas, R. (1935). A bacteriophage in relation to Stewart's disease of corn. Phytopathology, *25*, 371-372.

Thomson, S., Wagner, A., and Gouk, S. (1998). Rapid epiphytic colonization of apple flowers and the role of insects and rain. VIII International Workshop on Fire Blight 489 pp. 459-464.

Thomson, S. V. (2000). Epidemiology of fire blight, In J. L. Vanneste (ed.). Fire Blight: the disease and its causative agent, *Erwinia amylovora*, pp. 9-36. New York: CABI Pub.

Thomson, S. V., Gouk, S. C., Vanneste, J. L., Hale, C. N., and Clark, R. G. (1993). The presence of streptomycin resistant strains of *Erwinia amylovora* in New Zealand. Acta Horticulturae*, 338*, 223-230.

Thomson, S. V., Hansen, D. R., Flint, K. M., and Vandenberg, J. D. (1992). Dissemination of bacteria antagonistic to *Erwinia amylovora* by honey bees. Plant Disease*, 76*, 1052-1056.

Toth, I., Perombelton, M., Salmond, G. (1993). Bacteriophage φKP mediated generalised transduction in *Erwinia carotovora* subspecies *carotovora*. Journal of General Microbiology*, 139*, 2705-2709.

Touchon, M., Charpentier, S., Clermont, O., Rocha, E. P., Denamur, E., and Branger, C. (2011). CRISPR distribution within the *Escherichia coli* species is not suggestive of immunity-associated diversifying selection. Journal of Bacteriology*, 193*, 2460-2467.

Touchon, M., and Rocha, E. P. (2010). The small, slow and specialized CRISPR and anti-CRISPR of *Escherichia* and *Salmonella*. PLoS One*, 5*, e11126.

Tovkach, F. I. (2002a). Defective lysogeny in *Erwinia carotovora*. Microbiology*, 71*, 306-313.

Tovkach, F. I. (2002b). A study of *Erwinia carotovora* phage resistance with the use of temperate bacteriophage ZF40. Microbiology*, 71*, 72-77.

Tovkach, F. I. (2002c). Temperate bacteriophage ZF40 of *Erwinia carotovora*: Phage particle structure and DNA restriction analysis. Microbiology*, 71*, 65-71.

Triplett, L. R., Zhao, Y., and Sundin, G. W. (2006). Genetic differences between blight-causing *Erwinia* species with differing host specificities, identified by suppression subtractive hybridization. Applied and Environmental Microbiology*, 72*, 7359-7364.

Tsiantos, J., and Psallidas, P. (1993). Experiments on chemical control of fire blight (*Erwinia amylovora*) in Greece. Phytopathologia Mediterranea*, 32*, 201-205.

Tsiantos, J., Psallidas, P., and Chatzaki, A. (2003). Efficacy of alternatives to antibiotic chemicals for the control of fire blight of pears. Annals of Applied Biology*, 143*, 319-323.

Tyson, G. W., and Banfield, J. F. (2008). Rapidly evolving CRISPRs implicated in acquired resistance of microorganisms to viruses. Environmental microbiology, *10*, 200-207.

Uchiumi, T., Abe, M., and Higashi, S. (1998). Integration of the temperate phage φU into the putative tRNA gene on the chromosome of its host *Rhizobium leguminosarum* biovar *trifolii*. Journal of General and Applied Microbiology, *44*, 93-99.

Valyasevi, R., Sandine, W., and Geller, B. (1990). The bacteriophage kh receptor of *Lactococcus lactis* subsp. *cremoris* KH is the rhamnose of the extracellular wall polysaccharide. Applied and Environmental Microbiology, *56*, 1882-1889.

Van de Rijn, I. (1983). Streptococcal hyaluronic acid: proposed mechanisms of degradation and loss of synthesis during stationary phase. Journal of Bacteriology, *156*, 1059-1065.

van der Oost, J., Jore, M. M., Westra, E. R., Lundgren, M., and Brouns, S. J. J. (2009). CRISPR-based adaptive and heritable immunity in prokaryotes. Trends in Biochemical Sciences, *34*, 401-407.

van der Oost, J., Westra, E. R., Jackson, R. N., and Wiedenheft, B. (2014). Unravelling the structural and mechanistic basis of CRISPR-Cas systems. Nature Reviews Microbiology, *12*, 479-492.

van der Ploeg, J. R. (2009). Analysis of CRISPR in *Streptococcus mutans* suggests frequent occurrence of acquired immunity against infection by M102-like bacteriophages. Microbiology, *155*, 1966-1976.

van der Zwet, T., and Beer, S. V. (1995). Fire Blight: its nature, prevention and control : a practical guide to integrated disease management. Washington, DC USA: United States Department of Agriculture.

van der Zwet, T., and Keil, H. L. (1979). Fire Blight; a bacterial disease of rosaceous plants. Fruit Laboratory, Beltville Maryland: United States Depatment of Agriculture.

van der Zwet, T., and Sugar, D. (1994). Present distribution of fire blight and its mode of dissemination - a review. Acta Horticulturae, *367*, 391-401.

Van Laere, O., De Greef, M., and De Wael, L. (1980). Influence of the honeybee on fireblight transmission. II Symposium on Fire Blight 117 pp. 131-144.

Vandenbergh, P. A., Wright, A. M., and Vidaver, A. K. (1985). Partial purification and characterization of a polysaccharide depolymerase associated with phage-infected *Erwinia amylovora*. Applied and Environmental Microbiology*, 49*, 994-996.

Vanneste, J. L. (2000). Fire Blight: the disease and its causative agent, *Erwinia amylovora*. Wallingford, UK: CAB Pub.

Vanneste, J. L., and Voyle, M. D. (1998). Genetic basis of streptomycin resistance in pathogenic and epiphytic bacteria isolated from apple orchards in New Zealand. Acta Horticulturae*, 489*, 671-672.

Vantomme, R., Swings, J., Goor, M., Kersters, K., and de Ley, J. (1982). Phytopathological, serological, biochemical and protein electrophoretic characterization of *Erwinia amylovora* strains isolated in Belgium. Phytopathologische Zeitschrift*, 103*, 349-356.

Verhoyen, M. (1983). *Erwinia amylovora* (Burrill) Winslow et al.: its appearance in Belgium, hosts and losses caused. Parasitica*, 39*, 49-53.

Vidaver, A. K. (1976). Prospects for control of phytopathogenic bacteria by bacteriophages and bacteriocins. Annual Review of Phytopathology*, 14*, 451-465.

Vrancken, K., Holtappels, M., Schoofs, H., Deckers, T., and Valcke, R. (2013). Pathogenicity and infection strategies of the fire blight pathogen *Erwinia amylovora* in *Rosaceae*: state of the art. Microbiology (Reading, England)*, 159*, 823-832.

Wagner, P. L., Livny, J., Neely, M. N., Acheson, D. W. K., Friedman, D. I., and Waldor, M. K. (2002). Bacteriophage control of Shiga toxin 1 production and release by *Escherichia coli*. Molecular Microbiology*, 44*, 957-970.

Walsh, F., Pelludat, C., Duffy, B., Smith, D. P., Owens, S. M., Frey, J. E., and Holliger, E. (2014). Impact of streptomycin applications on antibiotic resistance in apple orchards. Agrarforschung Schweiz*, 5*, 300-305.

Wang, H., Jiang, X., Mu, H., Liang, X., and Guan, H. (2007). Structure and protective effect of exopolysaccharide from *P. agglomerans* strain KFS-9 against UV radiation. Microbiological Research*, 162*, 124-129.

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

Weinbauer, M. G., Agis, M., Bonilla-Findji, O., Maltis, A., and Winter, C. (2007). Bacteriophage in the environment. Nolfolk, UK: Caister Academic Press.

Westra, E. R., Pul, Ü., Heidrich, N., Jore, M. M., Lundgren, M., Stratmann, T., Wurm, R., Raine, A., Mescher, M., and Van Heereveld, L. (2010). H-NS-mediated repression of CRISPR-based immunity in *Escherichia coli* K12 can be relieved by the transcription activator LeuO. Molecular Microbiology*, 77*, 1380-1393.

Westra, E. R., Swarts, D. C., Staals, R. H., Jore, M. M., Brouns, S. J., and van der Oost, J. (2012a). The CRISPRs, they are a-changin': how prokaryotes generate adaptive immunity. Annual Review of Genetics*, 46*, 311-339.

Westra, E. R., van Erp, P. B., Künne, T., Wong, S. P., Staals, R. H., Seegers, C. L., Bollen, S., Jore, M. M., Semenova, E., and Severinov, K. (2012b). CRISPR immunity relies on the consecutive binding and degradation of negatively supercoiled invader DNA by Cascade and Cas3. Molecular Cell*, 46*, 595-605.

Wiedenheft, B., Lander, G. C., Zhou, K., Jore, M. M., Brouns, S. J., van der Oost, J., Doudna, J. A., and Nogales, E. (2011a). Structures of the RNA-guided surveillance complex from a bacterial immune system. Nature*, 477*, 486-489.

Wiedenheft, B., Sternberg, S. H., and Doudna, J. A. (2012). RNA-guided genetic silencing systems in bacteria and archaea. Nature*, 482*, 331-338.

Wiedenheft, B., van Duijn, E., Bultema, J. B., Waghmare, S. P., Zhou, K., Barendregt, A., Westphal, W., Heck, A. J., Boekema, E. J., and Dickman, M. J. (2011b). RNA-guided complex from a bacterial immune system enhances target recognition through seed sequence interactions. Proceedings of the National Academy of Sciences*, 108*, 10092-10097.

Wiedenheft, B., Zhou, K., Jinek, M., Coyle, S. M., Ma, W., and Doudna, J. A. (2009). Structural basis for DNase activity of a conserved protein implicated in CRISPR-mediated genome defense. Structure*, 17*, 904-912.

Wilkinson, B. J., and Holmes, K. M. (1979). *Staphylococcus aureus* cell surface: capsule as a barrier to bacteriophage adsorption. Infection and Immunity*, 23*, 549-552.

Willis, S., Kazmierczak, K., Carter, R., and Rothman-Denes, L. (2002). N4 RNA polymerase II, a heterodimeric RNA polymerase with homology to the

single-subunit family of RNA polymerases. Journal of Bacteriology*, 184*, 4952-4961.

Wilson, M., and Lindow, S. E. (1993). Interactions between the biological control agent *Pseudomonas fluorescens* A506 and *Erwinia amylovora* in pear blossoms. Phytopathology*, 83*, 117-123.

Wittmann, J., Dreiseikelmann, B., Rohde, M., Meier-Kolthoff, J. P., Bunk, B., and Rohde, C. (2014). First genome sequences of *Achromobacter* phages reveal new members of the N4 family. Virology Journal*, 11*, 1.

Wittmann, J., Klumpp, J., Switt, A. I. M., Yagubi, A., Ackermann, H.-W., Wiedmann, M., Svircev, A., Nash, J. H., and Kropinski, A. M. (2015). Taxonomic reassessment of N4-like viruses using comparative genomics and proteomics suggests a new subfamily-"*Enquartavirinae*". Archives of virology*, 160*, 1-10.

Yagubi, A. I., Castle, A. J., Kropinski, A. M., Banks, T. W., and Svircev, A. M. (2014). Complete Genome Sequence of *Erwinia amylovora* Bacteriophage vB_EamM_Ea35-70. Genome Announcements*, 2.*

Yamada, T., Kawasaki, T., Nagata, S., Fujiwara, A., Usami, S., and Fujie, M. (2007). New bacteriophages that infect the phytopathogen *ralstonia solanacearum*. Microbiology*, 153*, 2630-2639.

Yan, J., Mao, J., and Xie, J. (2014). Bacteriophage polysaccharide depolymerases and biomedical applications. BioDrugs*, 28*, 265-274.

Yosef, I., Goren, M. G., and Qimron, U. (2012). Proteins and DNA elements essential for the CRISPR adaptation process in *Escherichia coli*. Nucleic Acids Research*, 40*, 5569-5576.

Young, J. C., Dill, B. D., Pan, C., Hettich, R. L., Banfield, J. F., Shah, M., Fremaux, C., Horvath, P., Barrangou, R., and Verberkmoes, N. C. (2012). Phage-induced expression of CRISPR-associated proteins is revealed by shotgun proteomics in *Streptococcus thermophilus*. PLoS One*, 7*, e38077.

Zeng, L., Skinner, S. O., Zong, C., Sippy, J., Feiss, M., and Golding, I. (2010). Decision making at a subcellular level determines the outcome of bacteriophage infection. Cell*, 141*, 682-691.

Zhang, Y. X., and Geider, K. (1997). Differentiation of *Erwinia amylovora* strains by pulsed-field gel electrophoresis. Applied and Environmental Microbiology*, 63*, 4421-4426.

Zhao, Y., Sundin, G. W., and Wang, D. (2009). Construction and analysis of pathogenicity island deletion mutants of *Erwinia amylovora*. Canadian Journal of Microbiology*, 55*, 457-464.

Zong, C., So, L.-H., Sepúlveda, L. A., Skinner, S. O., and Golding, I. (2010). Lysogen stability is determined by the frequency of activity bursts from the fate-determining gene. Molecular Systems Biology*, 6*, 440.

Zurcher, T., Luo, G., and Palese, P. (1994). Mutations at palmitylation sites of the influenza virus hemagglutinin affect virus formation. Journal of Virology*, 68*, 5748-5754.