NOVEL VIRULENT PHAGES FOR *Xylella fastidiosa* AND OTHER MEMBERS OF THE *Xanthomonadaceae*

A Dissertation

by

STEPHEN JAMES AHERN

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Carlos F. Gonzalez
Committee Members,	Dennis Gross
	David Appel
	Ryland F. Young III
Head of Department,	Leland S. Pierson III

December 2013

Major Subject: Plant Pathology

ABSTRACT

The xylem-limited bacterium, *Xylella fastidiosa*, is the causal agent of several plant diseases, most notably Pierce's Disease of grapes and citrus variegated chlorosis. We report on the isolation and characterization of the first virulent phages for *X. fastidiosa*, Xfas103, Xfas106, Xfas303, Xfas304, with host ranges extending to *Xanthomonas*. Efficiency of plating on *X. fastidiosa* strain Temecula or *Xanthomonas* strain EC-12 ranged from 1.41×10^{-1} to 2.29×10^{-3}, depending on propagating and plating host. Phages propagated on homologous hosts had an observed adsorption rate constant of $(4.33 \pm 0.28) \times 10^{-12}$ ml cell^{-1} min^{-1} for strain Temecula, and ranged from $(4.58 \pm 0.30) \times 10^{-10}$ to $(7.26 \pm 0.42) \times 10^{-10}$ ml cell^{-1} min^{-1} for strain EC-12. Siphophages Xfas103 and Xfas106 exhibit over 80% nucleotide identity to each other and are syntenic to phage BcepNazgul. Here it is proposed that phage BcepNazgul is the founding member of a novel phage type, to which Xfas103 and Xfas106 belong. Podophages Xfas303 and Xfas304 show no significant DNA homology, both encode for a single subunit RNA polymerase at the right end of the class II gene cluster, and are new members of the phiKMV-like phage type. The four phages utilize Type IV pili as receptors to infect strains Temecula and EC-12. The phages may be useful as agents for an effective and environmentally friendly strategy for the control of diseases caused by *X. fastidiosa*.

DEDICATION

To my Mom, for all her love and support.

ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. Carlos Gonzalez, for his guidance and encouragement to always do the best science possible.

I would also like to thank my committee members, Dr. David Appel, Dr. Dennis Gross, and Dr. Ry Young, for their unique advice and support throughout the course of this research.

I would like to thank Guichun Yao for her continuous support, and Drs. Mayukh Das and Tushar Suvra Bhowmick for their guidance and advice. Without them, my work would not have been possible. I would like to thank all of the great undergraduate students whom I have had the opportunity to work and become friends with over the course of my time in the lab. You have always kept the lab fun.

TABLE OF CONTENTS

LIST OF FIGURES

Page

LIST OF TABLES

Page

CHAPTER I

INTRODUCTION AND LITERATURE REVIEW

The pathogen: *Xylella fastidiosa*

Classification

Xylella fastidiosa is a Gram-negative member of the Gamma-proteobacteria. It is a rod-shaped, non-flagellated bacterium, with distinctive rippled cell walls resembling *Rickettsia* (1-2). Two types of pili are located at the poles: type I pili (T1P), 0.4 to 1.0 μm in length, and type IV pili (T4P), 1 to 6 μm in length (1, 3). T1P are important for attachment, aggregation, and biofilm formation, whereas T4P are used for twitching motility and migration (4). *X. fastidiosa* is a fastidious aerobe that colonizes the nutritionally poor environmental niche of plant xylem vessels where it may or may not be the causal agent of disease, depending on the plant host (5).

Relatedness to Xanthomonas

The family *Xanthomonadaceae* is a wide-spread family of bacteria belonging to the Gram-negative Gamma-proteobacteria, including the two plant-associated genera *Xanthomonas* and *Xylella,* and the related genus *Stenotrophomonas* (6). The genus *Xanthomonas* currently comprises 27 species with validly published names that are important crop and horticultural pathogens (7). Genome analysis of different *X. fastidiosa* strains suggests that *X. fastidiosa* is most closely related to plant pathogenic members of the genus *Xanthomonas,* but with a genome reduced in size and with considerably lower GC content. This presumed evolutionary reduction may be related to

the fact that *X. fastidiosa* is xylem limited, in contrast to *Xanthomonas* (8). Individual species can comprise multiple pathovars. Pathogenic species and pathovars show a high degree of host plant specificity with many exhibiting tissue specificity, invading either the xylem elements of the vascular system or the intercellular spaces of the mesophyll parenchyma tissue of the host (9). *Xanthomonas* vascular pathogens such as *X. oryzae* pv. *oryzae* and *X. campestris* pv. *campestris* persist as epiphytes on plant surfaces before they enter the plant via natural openings such as hydathodes, stomata or wounds, and migrate to the lateral veins where they multiply. Inside the plant tissue, *Xanthomonas* can subsequently invade the surrounding mesophyll tissues resulting in disease symptoms (9). *X. fastidiosa* does not need plant invasion mechanisms, because cells are introduced into the xylem by insect vectors. In fact, genome analysis of *X. fastidiosa* has shown the presence of several genes previously only identified in animal pathogens, and it has been suggested that they may facilitate attachment and growth in insect vectors (10-11). Many genes have been implicated in the virulence in *Xanthomonas* species and many of these homologs are present in *X. fastidiosa*. These include T1P and afimbril adhesions for attachment and biofilm formation (12-13), xanthan-like exopolysaccharides (EPS) similar to *Xanthomonas* xanthan gum (12-14), a type II secretion system used to export exoenzymes to degrade the plant cell wall (12), and T4P used for twitching motility (12-13). Additionally, because of its nutrient-poor environment, *X. fastidiosa* has special mechanisms to concentrate and absorb nutrients. It is believed that it possesses EPS glycocayx-like fibers that may function in ion-

exchange, nutrient binding within cell aggregations, and conserving and concentrating digestive enzymes released by the bacterium (15).

Vectors

Several species of sharpshooter leafhoppers (Homoptera; *Cicadellidae*) and spittlebugs (Homoptera: *Cercopidae*) have been reported to be vectors of *X. fastidiosa* (16). For successful transmission of *X. fastidiosa* to plants, the insect must first acquire the bacterium from an infected plant. Second, the bacterium must attach to the insect's forgut cuticle. Lastly, the vector must feed upon and transmit the bacterium to a susceptible host. The ability of a vector to successfully acquire and transmit the bacterium depends on the combination of vector species and host plant (17). However, it appears that all species of sharpshooters can transmit all strains of *X. fastidiosa*, although with different degrees of efficiency (16). Sharpshooters are polyphagous insects (i.e., they feed on a wide variety of plants). Weeds found in riparian areas harbor *X. fastidiosa* and often serve as the primary source of inoculum for sharpshooters who then transmit the pathogen to crops. Infectivity is not lost in adults, so insect vectors are able to transmit the pathogen to several susceptible hosts for months after initial acquisition (16). *X. fastidiosa* utilizes T1P to successfully attach to the forgut of potential vectors (10). The attachment must be strong because of the considerable force asserted by the insect when feeding. Once attached, the bacterium is able to replicate within the insect's forgut. For an insect vector to acquire *X. fastidiosa* from an infected plant, a minimum concentration of 10^4 cells/g of plant tissue are required for most plants, and 10^6-

10^7cells/g of plant tissue are required for successful acquisition from grapes (18). However, as few as 100 cells are required for transmission from vector to plant (19).

Diseases caused by subspecies

Xylella fastidiosa causes a variety of diseases in over 150 plant species (5, 15, 20-29). Economically important diseases include Pierce's Disease (PD) of grape, citrus variegated chlorosis (CVC), phony peach disease (PPD), periwinkle wilt, almond leaf scorch (ALS), oleander leaf scorch (OLS) and coffee leaf scorch (CLS) (10). All strains of *Xylella* are currently classified as *X. fastidiosa*, with five recognized subspecies: (i) *X. fastidiosa* subsp. *fastidiosa* (most notable for PD); (ii) *X. fastidiosa* subsp. *multiplex* (most notable for PPD); (iii) *X. fastidiosa* subsp. *pauca* (most notable for CVC); (iv) *X. fastidiosa* subsp. *sandyi* (most notable for OLS); (v) *X. fastidiosa* subsp. *tashke* (associated with disease on the ornamental tree *Chitalpa tashkentensis*) (30-31). Except for reports of pear leaf scorch in Taiwan (32) and PD in Kosovo (33), diseases caused by *X. fastidiosa* are restricted to tropical and subtropical areas of the Americas, although some tree leaf scorch diseases occur in much colder climates (16). Crop diseases caused by *X. fastidiosa* appear to be "new encounter" diseases caused by endemic *X. fastidiosa.* Disease symptoms are often only observed in species exotic to the Americas, whereas *X. fastidiosa* is found in hundreds of symptomless indigenous plants (16).

Symptoms of disease are only observed when xylem vessels are highly colonized by the bacterium. The xylem transports water and soluble mineral nutrients from the roots throughout the plant, and used to replace water lost during transpiration and

photosynthesis. Xylem vessels are interconnected by bordered pits, which allow the passage of xylem sap, but block the passage of larger objects (such as bacteria) by a membrane (34). *X. fastidiosa* does not have a Type III secretion system, but Type II-secreted effectors including several β-1,4 endoglucanases, xylanases, xylosidases, and one polygalacturonase that are likely involved in degradation of different components of pit membranes and allow *X. fastidiosa* to spread between neighboring vessels (12-13, 35). Xylem vessels become occluded by dense colonization. When colonization becomes excessive, biofilms form along the xylem vessel and can block water flow. Limited water flow leads to water stress during high temperatures and can result in scorching symptoms.

Symptoms of PD in grapes begin with drying and necrosis of leaf margins, resulting in a "scorched" appearance. Scorched leaves usually drop from the distal end of the petiole, leaving bare petioles attached to canes, resulting in a "matchstick" appearance. Further defoliation, shoot dwarfing and cane stunting may occur. Vines may be reduced in growth, fruits may take on a dehydrated appearance, and eventual vine death may occur (36). PD precludes commercial production of European grape (*V. vinifera*) in much of the southeastern United States, since the pathogen and vector are endemic to the region. The first reports of PD in California date back to the 1880's, where the disease was usually confined to "hot spots" near water sources and riparian areas (5). The disease did not gain prominence until the introduction of the exotic glassy-winged sharpshooter (*Homalodisca vitripennis*, formerly known as *H. coagulata*), which is able to spread PD much more efficiently than the native blue-green sharpshooter (16).

Symptoms of CVC are observed on younger trees. Trees are generally not killed, but fruit production is severely reduced. Leaves develop small interveinal chlorotic spots, and wilting may occur. Fruits remain small, have high sugar content with hard rinds, and ripen early. Chronic symptoms include stunting and dieback of twigs (37). CVC was first reported in only a few orange trees in Argentina in 1987 (38). In less than a decade the disease had spread throughout South America (16). The rapid spread of this previously unknown disease suggests that the causal agent was an introduced strain of *X. fastidiosa* (16). Another theory, based on the close genetic homology of coffee strains of *X. fastidiosa* to citrus strains (39), is that the citrus strains may have recently evolved from endemic coffee strains. The rapid spread of CVC from Argentina to Brazil, and its ability to cause disease in grapes demonstrates the serious threat that CVC strains could pose to citrus production in North America and the potential for yet unknown diseases in additional crop and native plant species (16).

Symptoms of CLS on coffee plants first appear on young shoots as large necrotic-scorched areas on the tops or margins of mature leaves. New growth is stunted and pale. Fruit size and yield is also impaired. Premature fall of leaves and fruit is a more prominent symptom of disease the scorching. Water stress leads to more severe symptoms (37).

Peach trees infected with PPD generally decline in vigor and are more susceptible to other diseases and pests. Early symptoms include stunted young shoots that have earlier, more numerous and darker leaves than normal. Trees show early blooming and both leaves and flowers remain on the shoots longer than normal. Twigs

display increased lateral branching with shortened internodes. Lateral branches grow horizontally. Fruits are small and ripen early. Trees with PPD, when completely infected, appear uniform across the top, resembling a trimmed hedge (37).

Bacterial leaf scorch (BLS) affects many different shade tree species, including the American elm, red maple, sweet gum, sycamore and a several species of oak. Symptoms vary by tree species, but most include chloritic marginal leaf scorching. Symptoms appear in late summer to early fall. Leaves may curl and drop prematurely, branches die, and in some tree species the entire tree can die (37).

Current control measures

There are no methods proven to cure PD (16). Nearly all popular grape cultivars of European (*V. vinifiera*), American (*V. labrusca*) and French-American hybrid grapes are killed by PD (5). Species native to the southeastern United States (*V. muscadine*) show varying levels of resistance to PD, but are not desirable for wine production. The removal of infected grapevines only reduces secondary (vine to vine) spread of *X. fastidiosa* (40), and is generally ineffective in reducing the spread of disease within a plot because *X. fastidiosa* is primarily spread from sources outside vineyards, such as riparian areas, rather than vine to vine (16).

Infective overwintering adult vectors are responsible for establishing most chronic infections in early spring (16). Egg parasitoids such as *Gonatocerus* sp. have been used with limited success to control vector populations. However, their populations decrease strongly during winter when vector egg production is low, and therefore the first generation of vectors are only slightly parasitized (16). Significant efforts have been

made to control vector populations using systemic insecticides, especially neonicotinoids (imidaclorid). Area-wide insecticide control programs have been established, but have only been partially successful (40), and may be potentially harmful to the environment (41-42).

Given the current limited success in controlling *X. fastidiosa*, and the risk of its spread, there is a need for the development of an effective and environmentally friendly *X. fastidiosa* control strategy. Recently, there has been renewed interest in the application of bacteriophages (phage) for the control of bacterial plant disease (43). Phages have many attractions for both prophylactic and therapeutic control of bacterial plant diseases. Phages are self-replicating and self-limiting, specific to target bacteria and can be targeted against bacterial receptors that are essential for pathogenesis, without posing any harmful effects to plants, animals, humans or associated beneficial microflora (43).

Bacteriophages

History

Bacteriophages (or phages) were discovered independently by Frederick W. Twort in 1915 and by Felix d'Herelle in 1917. Phages are viruses of bacteria, and with an estimated viral population of greater than 10^{31}, are the most widely distributed biological entity in the biosphere (44). Phage can confer key phenotypes on their host and they play a key role in regulating bacterial populations the environment (45). The phage-bacterium relationship varies from the simple predator-prey model to a complex relationship that promotes the survival and evolutionary success of both (46).

Phages have played and continue to play a key role in bacterial genetics and molecular biology. In the 1930s and the years that followed, virologists such as Salvador Luria and Max Delbrück utilized phages as model systems to investigate many aspects of virology, including virus structure, genetics, and replication. These relatively simple agents have played a central role in some of the most significant discoveries in biological sciences, from the identification of DNA as the genetic material (47), to the deciphering of the genetic code (48), to the development of molecular biology (49). Research on phages has aided our understanding of the basic molecular mechanisms of gene action and biological structure. Phage genomics is revealing novel biochemical mechanisms for replication, maintenance and expression of the genetic material (49).

Classification

The dsDNA tailed phages, or Caudovirales, account for 96% of all the phages reported in the scientific literature (50). Tailed phages are further organized into families based on tail morphology; Myoviridae (long, contractile tails), Siphoviridae (long-flexuous, non-contractile tails) or Podoviridae (short-stubby, non-contractile tails).

Due to extensive mosaicism caused by frequent horizontal gene transfer events, the evolutionary relationships between phages cannot be easily established, and the concept of biological species or hierarchical lineages in the classical Linnaean sense are difficult to apply (51-52). DNA or amino acid identity relationships have been used to organize phages into groups (53-55), however the enormous diversity of phages has resulted in numerous phage groups with no obvious evolutionary linkage (56). A better

model proposed organizes phages into "phage types" based on shared genome organization and life style (57).

Phages can display one of two types of life cycle: virulent (lytic) or temperate (lysogenic). The phage life cycle consists of an extracellular search, attachment to susceptible bacteria, phage-genome uptake into bacteria, production of phage progeny, and subsequent release of these progeny into the extracellular environment (58). The extracellular search occurs via phage diffusion through an aqueous milieu. Phage contact with a host bacterium occurs via tail and tip proteins that have high affinity to specific bacterial surface molecules, including outer membrane proteins, pili, flagella, oligosaccharides or lipopolysaccharides (LPS) (59). Type IV pili (T4P) have been previously identified as phage receptors for several *P. aeruginosa* phages (60-66). For many phages, this binding step is reversible, and may require a second irreversible binding step. After successful attachment (or adsorption) to a specific receptor on the bacterial cell surface, the phage genome and associated proteins are released into the cell (the capsid proteins are usually stripped off and remain outside the cell). Phages express early genes that allow the phage to redirect bacterial protein expression systems and regulate the production of new virions. Expression of late phage genes allow production of additional copies of the phage genome and structural proteins, which are subsequently assembled into new phage virion particles. The lytic pathway results in the eventual lysis of the host cell (time determined by the phage) and the release of progeny phage that are free to bind and infect new host cells (67). When a temperate phage infects a bacterium, it can either replicate by means of the lytic life cycle, or it can incorporate its DNA

genome into the bacterium's genome and become a noninfectious prophage by means of the lysogenic life cycle. During the lysogenic cycle, phage expressed genes allow the phage genome to integrate into the host genome, and then suppress further expression of most phage genes, except those that are needed to maintain the lysogenic state. The prophage replicates with the bacterial host cell (or lysogen) until it is reactivated and resumes the lytic cycle (67). Phage induction is usually triggered by host stress caused by DNA damage. In the laboratory, prophages can be induced using mitomycin C or UV. Once induced, the phage goes through the typical lytic cycle, producing progeny phage and lyses the host cell.

The canonical lysis cassette of phages that infect Gram negative bacteria encodes a holin, endolysin, and a spanin complex (68). The endolysin, a peptidoglycan-degrading enzyme, accumulates in the cytosol at the end of the replication cycle. The holin, a small hydrophobic membrane spanning protein, is essential for the endolysin to gain access to the outer membrane. Holin genes have enormous diversity, with more than 50 unrelated gene families having been described (68). Holins form a membrane lesion in the cytoplasmic membrane at a genetically predetermined time, which permeabilizes the inner membrane for the endolysin to degrade the peptidoglycan layer (68). In the absence of an intact peptidoglycan layer, the cell lyses due to internal osmotic pressure. The presence of a dual start motif in some holin sequences results in the production of a holin and anti-holin, and further optimizes lysis time and phage fitness (68-69). A holin and an endolysin constitute the essential minimum for programmed lysis in most cases, however several phages encode Rz/Rz1 proteins (68). The Rz, an inner membrane

protein, and Rz1, an outer membrane lipoprotein, bind each other by C-terminal interactions, forming a complex that spans the entire periplasm. Once the endolysin destroys the cell wall, the spanin complexes oligomerize side-to-side and undergo a conformational change. The molecular basis of Rz and Rz1 function is currently unknown, but the complex is involved in the destruction of the outer membrane (70).

Phage therapy

Soon after their discovery, phages were utilized for the control of bacterial pathogens, also called phage therapy or phage biocontrol. Phage therapy was used in human and veterinary medicine to agricultural settings, and was initially used to successfully treat a variety of diseases ranging from typhoid and paratyphoid fevers, dysentery, cholera and pyogenic urinary tract infections (45, 71). With the introduction of broad-spectrum antibiotics in the 1940's, the utilization of phage therapy mostly ceased (71). In Eastern Europe and the former Soviet Union, research and the application of phage therapy (often as a compliment to antibiotic treatment) have continued (71). There is a current renewed interest in phage therapy in the West due to the emergence of drug-resistant pathogenic bacteria as well as concerns over traditional pathogen control. Phages offer a safe and novel alternative to control plant pathogenic bacteria (72). The potential of phage therapy to control plant pathogens was investigated early after the discovery of phage (73). Recent examples of phage therapy for plant disease control listed in TABLE 1. Phage therapy for plant pathogens has been utilized for controlling plant pathogens either in the rhizosphere or phyllosphere. Both of these environments pose challenges for successful phage therapy. In the soil, only a low number of viable

phage are available to attack target bacteria due to low rates of phage diffusion and high rates of phage inactivation (58). The phyllosphere is a harsh environment and phages applied to aerial tissues degrade rapidly due to UV and high temperature (74-77), can easily be washed away by rain, and are unable to infect in dry conditions (43, 58, 78). This transient survival of phages on plant leaf surfaces is a major limiting factor of phage treatment. The plant vascular system offers a unique opportunity to evaluate a closed “aqueous” system in which phage can be transported through the tissue and come in intimate contact with the causal pathogen. The vascular system is also a less harsh environment, in that phage are not exposed to UV.

TABLE 1 Application of phage therapy for control of plant pathogens

Host	Disease	Pathogen	References
Cabbage	Black rot	*Xanthomonas campestris pv. campestris*	(79)
Calla Lily	Bacterial soft rot	*Erwinia carotovora subsp. carotovora*	(80)
Citrus	Citrus canker	*Xanthomonas citri subsp. citri*	(81)
Citrus	Citrus bacterial spot	*Xanthomonas fuscans* subsp. *citrumelonis*	(81)
Geranium	Bacterial blight	*Xanthomonas campestris* pv. *pelargonii*	(82)
Mungbean	Bacterial leaf spot	*Xanthomonas axonopodis pv. vignaeradiatae*	(83)
Mushroom	Bacterial blotch	*Pseudomonas tolaasii*	(84-85)
Onion	Leaf blight	*Xanthomonas axonopodis pv. allii*	(86-88)
Pepper	Bacterial spot	*Xanthomonas campestris pv. vesicatoria*	(79)
Pomefruits	Fireblight	*Erwinia amylovora*	(89-94)
Potato	Potato scab	*Streptomyces scabies*	(95)
Rice	Bacterial leaf blight	*Xanthomonas oryzae*	(96)
Stonefruits	Bacterial spot	*Xanthomonas arboricola pv. pruni*	(76, 97-99)
Tobacco	Bacterial wilt	*Ralstonia solanacearum*	(100-101)
Tomato	Bacterial spot	*Xanthomonas campestris pv. vesicatoria*	(75, 102-104)
Tomato	Crown gall	*Agrobacterium tumefaciens*	(105)
Walnut	Walnut blight	*Xanthomonas campestris pv. juglandis*	(77)

Phages are ubiquitous in the environment and can easily be isolated from water, soil, plant cuttings and sewage (106-107). Samples can either be directly plated or enriched for phages in liquid culture with one or more bacterial hosts. Enrichment allows for the detection of phages that are initially found in low numbers. Liquid culture enrichment tends to select for rapidly growing virulent phages, which could bias selection for a single rapidly replicating phage, masking the appearance of slower-growing but perhaps broader host range phages. Additionally, growth conditions may alter host gene expression of potential phage receptor sites. Once environmental samples are collected and processed, they must be evaluated for the presence of phage able to lyse the host of interest. A panel of pathogenic host strains may be tested in order to increase the diversity of isolated phages. Additionally, surrogate host bacteria may be used as phage isolation hosts as long as phages isolated against the surrogate host are also able to efficiently adsorb too, infect and replicate in the target pathogenic strains. The use of surrogate hosts has many advantages, including ease of handling, ease of culture, and lower pathogenicity. Examples of previous surrogate host use include the production of *Mycobacterium tuberculosis* phage D29 using the easier to grow and less pathogenic *M. smegmatis* (108), and the production of *Listeria monocytogenes* phage P100 using the non-pathogenic *L. innocua* (109).

Generally, phages have narrow host ranges, often only infecting a particular strain or species of bacteria, whereas others may be able to infect numerous genera. Host range is determined by the ability of the phage to bind to the host surface, eject its DNA into the cell, and overcome potential bacterial CRISPR-Cas or restriction and

modification systems (110). In a given field or plant, there are different populations of target bacterium that may or may not be sensitive to a given phage, and the open nature of agricultural systems allow movement of new pathogen strains into and out of crop populations.

Phage therapy works by reducing the targeted pathogen (72, 111-113). Phage-resistant mutant bacterial strains may arise during the course of phage therapy, and initial decline may be followed by recovery of phage-resistant bacteria. However, bacterial resistance to phage attack often comes at a metabolic cost to bacteria and can affect virulence related properties. If a phage can be targeted against a bacterial receptor that is essential for virulence, resistant mutants would have attenuated virulence (43). Additionally, bacterial-resistance is less likely to develop with the use of phage cocktails, where bacteria are challenged with multiple phages that recognize different receptor sites, the equivalent of a multidrug treatment (114).

Temperate phage should not be used for phage therapy for a number of reasons. First, the genomes of temperate phage may contain genes which alter the phenotype of the host cell when a lysogen is established, a process called lysogen conversion. Temperate phages may encode genes that are able to transform the host cell they infect into a more effective disease causing agent. Second, lysogens are immune to superinfection by the same phage or related phages due to repression of transcription caused by the resident prophage. A third concern is that many temperate phages are capable of generalized transduction, in which the phage is able to package and transmit host DNA to other cells. Like lysogenic conversion, generalized transduction has the

potential to increase the disease causing potential of bacteria (114). For these reasons, temperate phages (or at least phages able to form stable lysogens) should not be used for phage therapy. Determining the temperate status of a novel phage is not always easy. Genomic characterization may identify genes commonly associated with lysogeny, such as repressor genes and integrases, although the diversity of phage genomes complicates their detection. The absence of plaque turbidity and the inability to form a stable lysogen under given culture conditions in a particular bacterial strain, does not necessarily mean that a phage is incapable of forming lysogens under different conditions or in other bacterial strains (114).

CHAPTER II

ISOLATION AND BIOLOGICAL CHARACTERIZATION OF PHAGES FOR *Xylella fastidiosa*

Introduction

The plant pathogenic bacterium, *Xylella fastidiosa*, is the cause of a number of economically important diseases, including Pierce's Disease (PD).There is currently no cure for disease caused by *X. fastidiosa*. Limited control efforts are directed toward preventing the spread of the disease by the use of tolerant varieties, cultural and hygienic measures, as well as chemical and biological insect vector control (16, 37). However, existing disease control methods are often only partially successful, and in the case of systemic insecticides used to control vector populations, may be potentially harmful to the environment (41-42).

Given the current limited success in controlling *X. fastidiosa*, and the risk of its spread, there is a need for the development of an effective and environmentally friendly *X. fastidiosa* control strategy. Recently, there has been renewed interest in the application of phage for the control of bacterial plant disease (43). Phages have many attractions for both prophylactic and therapeutic control of bacterial plant diseases. Phages are self-replicating and self-limiting, specific to target bacteria and can be targeted against bacterial receptors that are essential for pathogenesis, without posing any harmful effects to plants, animals, humans or associated beneficial microflora.

Recently, the first temperate (lysogenic) phage of *X. fastidiosa*, Xfas53, was isolated and characterized (115). However, temperate phage should not be used as biocontrol agents because of the potential risk of lysogenic conversion, superinfection immunity, and the increased risk of generalized transduction. Virulent (lytic) phages are needed for an effective and sustainable phage-based control strategy (114). Here it is described the isolation of phages for the plant pathogen *X. fastidiosa* and the further characterization of Xfas103, Xfas106, Xfas303, and Xfas304.

Materials and methods

Bacterial strains and culture conditions

Bacterial strains and plasmids used in this study are listed in TABLE 2. *X. fastidiosa* strains were cultured at 28°C in PW-M broth (PW-MB) (25) or PW-M agar plates [PW-MA; PW-MB amended with 20 g/liter plant cell culture tested agar (Sigma)] (115). PW-M soft agar (PW-MSA, PW-MB amended with 7.5 g/liter plant cell culture tested agar) was used for overlays. *Xanthomonas* strains were cultured at 28°C in tryptone nutrient broth (TNB) (116) or tryptone nutrient agar (TNA; TNB lacking KNO_3 and amended with 20 g/liter agar). TNA soft agar (TNSA, TNB lacking KNO_3 and amended with 7.5 g/liter agar) was used for overlays. Plant extracts were plated to PW-MA or TNA amended with cycloheximide (40 μg/ml), designated PW-MC or TNAC, for the detection of *X. fastidiosa* or enumeration and isolation of resident bacteria, respectively.

TABLE 2 Bacterial strains used in this study

Strain or plasmid	Genotype and relevant features [a]	Reference or source
X. fastidiosa		
Temecula1	PD strain (subsp. *fastidiosa*) (ATCC 700964)	(12)
Ann-1	Oleander isolate (subsp. *sandyi*) (ATCC 700598)	(117)
Dixon	Almond isolate (subsp. *multiplex*) (ATCC 700965)	(117)
XF53	Grape isolate (subsp. *fastidiosa*)	(115)
XF54	Grape isolate (subsp. *fastidiosa*)	(115)
XF95	Oleander isolate (subsp. *sandyi*)	(115)
XF134	Grape isolate (subsp. *fastidiosa*)	This study
XF136	Grape isolate (subsp. *fastidiosa*)	This study
XF140	Grape isolate (subsp. *fastidiosa*)	This study
XF141	Grape isolate (subsp. *fastidiosa*)	This study
Xanthomonas		
EC-12	*Xanthomonas* sp., rice isolate	This study
North 40	*X. axonopodis* pv. *citri*, sweet orange isolate	(N. Wang, U. Florida)
Ft. Basinger	*X. axonopodis* pv. *citri*, sweet orange isolate	(N. Wang, U. Florida)
Block 22	*X. axonopodis* pv. *citri*, sweet orange isolate	(N. Wang, U. Florida)
Jal-4	*X. euvesicatoria*, jalapeno isolate	Laboratory stock
Presidio-4	*Xanthomonas* sp., rice isolate	Laboratory stock

Bacteriophage isolation and purification

Plant and sewage samples were assayed for the presence of phage able to form plaques on *X. fastidiosa* strain Temecula. Plant extracts were prepared by macerating 10 g of plant tissue (*Oryzae sativa* or weeds obtained from rice fields in Jefferson County and Wharton County, Texas) in 50 ml of P-buffer (50 mM Tris-HCl pH 7.5, 100 mM NaCl, 8 mM MgSO4) using a mortar and pestle, strained through a sterile double layer of cheese cloth to remove large particles, and filter sterilized. Sewage influent collected from a municipal wastewater treatment plant in Brazos County, TX was centrifuged twice and filter sterilized. Plant and sewage filtrates were directly screened for phage activity using the spot-test and soft agar overlay method (115). Soft agar overlays containing 100 µl of strain Temecula (~10^8 CFU/ml) as an indicator host were allowed to solidify, spotted with 10 µl drops of serially diluted filtrates in P-buffer, and assayed after incubation at 28ºC for 5-7 days. Filtrates producing either plaques or cleared zones were serially diluted in P-buffer and assayed using the overlay method, where 100-µl of each dilution was directly mixed with the host suspension in tempered soft agar before overlaying. Individual plaques formed on overlays of strain Temecula were excised, suspended in P-buffer, filter sterilized and titered. This procedure was repeated three times to obtain single plaque isolates. Additionally, sewage filtrates were screened for phage activity using *Xanthomonas* strain EC-12 (isolated from plant sample, see below), as described above. Plaques were purified three times using strain EC-12, and purified phages were tested for activity on strain Temecula.

High-titer phage lysates (>1 × 10^{10} PFU/ml) were prepared by harvesting overlay plates of strain Temecula or EC-12 exhibiting confluent lysis with 5 ml of P-buffer, macerating the soft agar overlay, clarifying by centrifugation and filter sterilization. Lysates were stored at 4 ºC. High-titer phage lysates were further purified by CsCl buoyant density centrifugation as previously described (115).

Bacterial isolation, purification and identification

Non-filtered plant extracts and sewage samples were plated to PW-MC or TNAC, incubated at 28ºC and evaluated for growth (72 h on TNAC and 10 days on PW-MC). Representative single colonies were picked from plates and streak purified three times to obtain stocks on appropriate media. The identification of suspected *Xanthomonas* species was accomplished using the MIDI Sherlock® Microbial Identification System (Texas A&M University Plant Disease Clinic) and 16S-23S ITS region sequencing using primers Xan16-23F (5'-CTGAGCCAGGATCAAACTCT TC-3'), Xan16-23R (5'-CCTGACCTATCAACCACGTAGT-3') and Xan16SF1 (5'-AAGGAG GTGATCCAGCCGCA-3'). The genome of one *Xanthomonas* isolate, EC-12, was sequenced using the 454 pyrosequencing method. Additionally, samples were tested for the presence of *X. fastidiosa* using quantitative real time polymerase chain reaction (qRT-PCR) with *X. fastidiosa* specific primers (INF2 and INR1) as previously described (118).

Transmission electron microscopy

Electron microscopy of CsCl purified phage (~1 $\times 10^{11}$ PFU/ml) was performed by diluting stock with P-buffer. Phage were applied to thin 400-mesh carbon-coated Formvar grids, stained with 2% (wt/vol) uranyl acetate and air dried. Specimens were observed on a JEOL 1200EX transmission electron microscope operating at an acceleration voltage of 100 kV. Five virions of each phage were measured to calculate mean values and standard deviations for dimensions of capsid and tail, where appropriate.

Host range and efficiency of plating

Host ranges of purified phages (propagated on strain Temecula) were determined by the serial dilution spot-test method, as described above, using a panel of *X. fastidiosa* and *Xanthomonas* isolates as indicator hosts. Efficiency of plating (EOP) was determined by calculating the ratio of the phage plaque titer obtained with the heterologous (non-propagating) host, to that obtained on the homologous (propagating) host (in triplicate). Phage stocks were titered by the soft agar overlay method on either strain Temecula or EC-12, as described above.

One-step growth curve

Liquid cultures of logarithmically growing strain EC-12 were infected with individual phages at a multiplicity of infection (MOI) of ~3, incubated at room temperature for 5 min, and then diluted 1000-fold in TNB. Samples were kept at 28^{o}C with constant shaking (150 rpm). Two samples were taken at 3 min intervals. One

sample was filter sterilized and plated immediately without any treatment and the other was plated after treatment with 1% (vol/vol) chloroform (to release intracellular phages).

Bacteriophage adsorption

The kinetics of phage adsorption was determined as previously described (115). Liquid cultures of logarithmically growing cells (strain Temecula or EC-12) were infected with individual phages (propagated on homologous hosts) at a MOI of ~0.1. The mixture was incubated at room temperature with shaking (150 RPM). Samples were taken (2 h intervals for strain Temecula and 2 min intervals for strain EC-12), immediately filter sterilized, and titered. The rate of phage particle disappearance is defined as $dP/dt = -kBP$, where B is the concentration of bacteria, P is the concentration of free phage at any time (t), and k is the adsorption constant in ml cell^{-1} min^{-1} (119).

Lysogen formation assay

To assay for phage lysogens, Temecula and EC-12 spontaneous phage resistant mutants were isolated from soft agar overlays of bacteria mixed with Xfas phage (individually) at a high MOI (~3). Plates were monitored for cell growth (10-15 days for strain Temecula and 2-3 days for strain EC-12). Individual colonies that emerged were picked, purified (in triplicate) and tested for phage sensitivity by the soft agar overlay method. Primer pairs specific to Xfas103 primase (5'- AACCTGATCTGGTACGAC-3' and 5'- GGACATTTTTCAGTTCTCTC-3'), Xfas106 primase (5'-CAACCTCATCTGGTATGAC-3' and 5'-GTCTTGGGTAATTTCTTTCT-3'), Xfas303 helicase (5'- AACTACCTGACAGCGACT-3' and 5'-CGTACTAGCTTGGCTTCTA-3') or Xfas304 helicase (5'- AAGAAGCGTGGTTTGTTTGC-3' and 5'-

CTACCGGCTTCCCTAACTCC-3') were used to test for the presence of lysogens in phage resistant isolates. Wild type bacterial DNA was used as negative control and wild type bacterial DNA spiked with phage DNA served as positive controls.

Results and discussion

Phage isolation and characterization

Several phages were isolated from non-enriched plant and sewage filtrates using *X. fastidiosa* or *Xanthomonas* hosts and four were chosen for further characterization. Non-enriched plant extracts yielded phage titers of ~10 to 1×10^5 PFU/g of tissue when plated on strain Temecula. Phages Xfas103 and Xfas303 were isolated from non-enriched rice panicle tissue and Xfas304 was isolated from non-enriched weed tissue by direct spotting on strain Temecula. *X. fastidiosa* was not detected in either plant extracts or sewage (direct plating or qRT-PCR to limits of detection). However, plant extracts yielded several *Xanthomonas* isolates, including *Xanthomonas* strain EC-12. Based on draft genome sequence homology, strain EC-12 most closely resembles *X. sacchari* NCPPB 4393 (ZP_09852963). Direct plating of the non-enriched sewage filtrate showed no phage activity on strain Temecula, but had a titer of 1×10^3 PFU/ml when plated on strain EC-12. Phage Xfas106 was isolated from non-enriched sewage by direct plating to strain EC-12. Subsequent plating of purified Xfas106 formed plaques on strain Temecula at a reduced efficiency (see below) and may explain why the phage was not initially detected using strain Temecula. Phages Xfas103, Xfas303 and Xfas304, when propagated on strain Temecula, were also able to form plaques on strain EC-12. Phages Xfas103, Xfas106, Xfas303 and Xfas304, whether propagated on strain Temecula or

EC-12, were able to form plaques on three subspecies of *X. fastidiosa* and exhibited difference in their ability to form plaques on *X. fastidiosa* or *Xanthomonas* species (TABLE 3). Phages Xfas103 and Xfas106 produced small clear plaques, whereas phages Xfas303 and Xfas304 produced large clear plaques on lawns of strains Temecula and EC-12 (not shown). The shared ability for all four phages to form plaques on several of the same hosts but not others, suggests the presence of a shared primary receptor site (see below) recognized between the four phages, with yet unknown alternate secondary sites.

Phages Xfas103 and Xfas106 were found to have siphophage morphology, exhibiting capsids of 64 (±2.1) and 64 (±0.8) nm in diameter, with long non-contractile tails of 204 (±2.3) and 207 (±4.2) nm in length, respectively (FIG 1A,B respectively). Phages Xfas303 and Xfas304 were found to have podophage morphology, exhibiting capsids of 69 (±1.4) and 68 (±1.1) nm in diameter, respectively, with short non-contractile tails (FIG 1C,D respectively).

TABLE 3 Phage host range

Strain	Phage			
	Xfas103	Xfas106	Xfas303	Xfas304
X. fastidiosa				
Temecula1	+[a]	+	+	+
Ann-1	+	+	+	+
Dixon	+	+	+	+
XF53	+	+	+	+
XF54	+	+	+	+
XF95	+	+	+	+
XF134	-	+	+	+
XF136	-	+	+	-
XF140	+	+	+	-
XF141	-	+	+	-
Xanthomonas				
EC-12	+	+	+	+
Jal-4	-	-	+	+
Presidio-4	-	-	+	+
North 40	-	-	+	-
Ft. Basinger	-	-	+	-
Block 22	-	-	+	-

[a]Ability to form individual plaques.

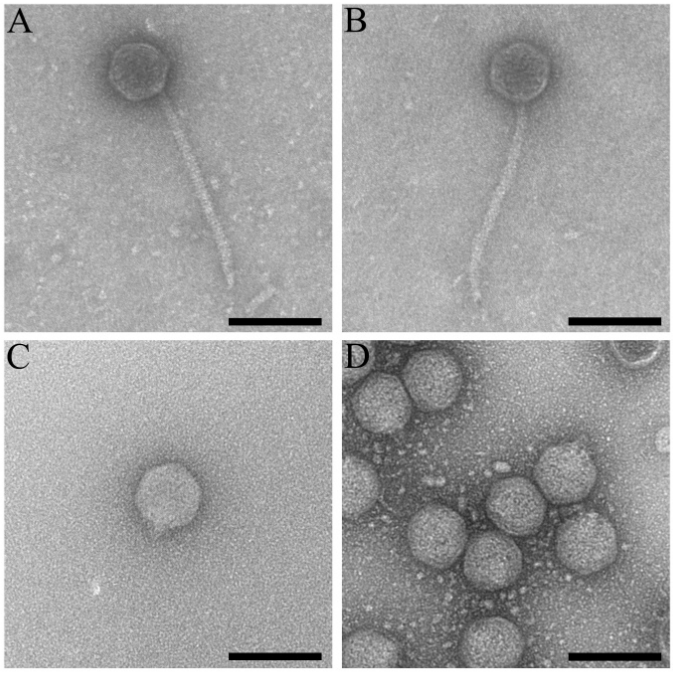

FIG 1 Morphology of phages Xfas103, Xfas106, Xfas303 and Xfas304. Transmission electron micrographs show (A) Xfas103, (B) Xfas106, (C) Xfas303 and (D) Xfas304 CsCl-purified phage particles. Samples were negatively stained with 2% (wt/vol) aqueous uranyl acetate. Bars, 100 nm.

Efficiency of plating

Phages Xfas106, Xfas303 and Xfas304 plated efficiently to strain Temecula and EC-12 (TABLE **4**). However, there was an approximate 2- to 3-log reduction in efficiency for phage Xfas103 when plated on Temecula or EC-12, respectively (TABLE **4**). This observation is suggestive of a restriction and modification system working in the two hosts. *X. fastidiosa* contains type I and type II restriction systems (120) and *Xanthomonas* spp. contain type II restriction systems (121), however no experimental information on the efficiency is available in the literature. These data suggest either the restriction and modification systems are relatively weak in both hosts, or the phages are subverting them, as has been reported in phage T7 (122)

Phage physiological parameters

Adsorption rate constants were determined from three replicate experiments using either strain Temecula or EC-12 as host. The observed rate constant for adsorption to

TABLE 4 Influence of bacterial host on efficiency of plating (EOP)

Production host	Indicator host	Phage isolate			
		Xfas103	Xfas106	Xfas303	Xfas304
Temecula	Temecula	1.0	1.0	1.0	1.0
Temecula	EC-12	$(1.2 \pm 0.2) \times 10^{-2}$	$(1.4 \pm 0.5) \times 10^{-1}$	$(2.3 \pm 0.2) \times 10^{-3}$	$(3.1 \pm 0.3) \times 10^{-1}$
EC-12	EC-12	1.0	1.0	1.0	1.0
EC-12	Temecula	$(1.4 \pm 0.7) \times 10^{-3}$	$(2.0 \pm 0.2) \times 10^{-1}$	$(9.5 \pm 0.5) \times 10^{-2}$	$(3.2 \pm 0.3) \times 10^{-1}$

[a] Data shown are the means of triplicate independent experiments ± SD.

Temecula for all phages propagated on Temecula was $(4.33 \pm 0.28) \times 10^{-12}$ ml cell^{-1} min^{-1}, similar to that determined for temperate phage Xfas53 (115). The determined rate constants on strain EC-12 for EC-12 propagated phage were $(5.48 \pm 0.23) \times 10^{-10}$, $(4.58 \pm 0.30) \times 10^{-10}$, $(7.26 \pm 0.42) \times 10^{-10}$ and $(5.98 \pm 0.37) \times 10^{-10}$ ml cell^{-1} min^{-1}, for Xfas103, Xfas106, Xfas303 and Xfas304, respectively. The observed difference in adsorption rates between hosts may be due to bacterial physiological characteristics that affect phage receptor site availability and function. An average burst size of 100, 113, 99, and 104 PFU/cell at ~52 min post infection was determined for phages Xfas103, Xfas106, Xfas303 and Xfas304, respectively, using strain EC-12 as host. The extremely slow phage adsorption rate to strain Temecula made it experimentally unfeasible to determine phage burst size in this host.

CHAPTER III

COMPLETE GENOME SEQUENCE OF PHAGES FOR *Xylella fastidiosa*

Introduction

The study of phage genomics is a fast growing science. Just 10 years ago, there were only 150 phage genomes deposited (123). Today there are over 1,000 phage genomes in the database. Comparative genome analysis reveals the evolutionary relatedness between phages, particularly those that may appear unrelated at first glance. Phages have evolved through horizontal gene transfer events between non homologous ancestral sequences, evident from the extreme genetic mosaicism found in phages (123). For responsible phage therapy, a phage genome sequence must be known. Genomic analysis provides unparalleled, high density information on a new phage. More than 96% of known phages are “tailed phages” with a linear, dsDNA genome packaged into an icosahedral or elongated icosahedral head. A "typical" genome is about 50 kb encoding ~ 60 proteins. Most of the proteins are involved in one of the following functional categories- head assembly, tail assembly, lysogenic/lytic decision and integration (if temperate), lysis, gene expression, DNA replication, DNA modification, host modification or DNA packaging. An integral approach in phage genomics is comparative analysis to known phages (123). However, often greater than 80% of predicted genes in a newly sequenced phage are unique to that phage or share homology only to a few hypothetical open reading frames (ORFs) in bacteria or other phages (56). Genomic analysis allows for rapid classification of some phages into known groups, and

allows for a more complete understanding of a phage's lifestyle and potential for general transduction (114). Here is described the genome sequence, characteristics and lifestyle of phages Xfas103, Xfas106, Xfas303 and Xfas304.

Materials and methods

DNA isolation and genome sequencing

DNA was isolated from CsCl purified phage suspensions as previously described (124). Phages Xfas103, Xfas106, Xfas303 and Xfas306 were sequenced as part of two pools of phage genomes using the 454 pyrosequencing method (Emory GRA Genomics Core). Phage genomic DNA was prepared from phage isolates as described above and mixed in equimolar amounts to a final concentration of ca. 100 ng/µl. The pooled DNA was sheared, ligated with a multiplex identifier (MID) tag specific for each of the four pools and sequenced by pyrosequencing using a full-plate reaction on a Roche FLX Titanium sequencer according to the manufacturer's protocols. The pooled phage DNA was present in two sequencing reactions. The trimmed FLX Titanium flowgram outputs corresponding to each of the four pools were assembled individually using the Newbler assembler version 2.5.3 (454 Life Sciences) by adjusting settings to include only reads containing a single MID identifier per assembly. The identity of individual contigs was determined by PCR using primers generated against contig sequences and individual phage genomic DNA preparations as template; the generation of the expected size product from a phage DNA template was used to match individual phages to their contigs. Phage Xfas103 was sequenced to 41-fold coverage, Xfas106 to 62-fold coverage, Xfas303 to 140-fold coverage and Xfas304 to 46–fold coverage. Sequencher

4.8 (Gene Codes Corporation) was used for sequence assembly and editing. Phage chromosomal end structures were determined as previously described (124).

ORF prediction and annotation

Protein-coding regions were initially predicted using GeneMark.hmm (125), refined by manual analysis in Artemis (126), and analyzed using BLAST (127). Proteins of particular interest were additionally analyzed by InterProScan (128), HHpred searches against the pdb70_29jun13 database (129), NCBI's Conserved Domain Database (130), LipoP (131), SignalP 4.1 (132), and TMHMM (133).

Lysogen formation assay

Xanthomonas strain EC-12 (see chapter II) was cultured at 28°C in medium TNB or TNA as described in Chapter 2. TNA soft agar (see Chapter II) was used for overlays. *Xylella* strain Temecula was cultured in medium PW-M or PW-MA as described in Chapter II. PW-M-soft agar (see Chapter II) was used for overlays. To assay for phage lysogens, Temecula and EC-12 spontaneous phage resistant mutants were isolated from soft agar overlays of bacteria mixed with Xfas phage (individually) at a high MOI (~3). Plates were monitored for cell growth (10-15 days for strain Temecula and 2-3 days for strain EC-12). Individual colonies that emerged were picked, purified (in triplicate) and tested for phage sensitivity by the soft agar overlay method. Primer pairs specific to Xfas103 primase (5'- AACCTGATCTGGTACGAC-3' and 5'-GGACATTTTTCAGTTCTCTC-3'), Xfas106 primase (5'-CAACCTCATCTGGTATGAC-3' and 5'-GTCTTGGGTAATTTCTTTCT-3'), Xfas303 helicase (5'- AACTACCTGACAGCGACT-3' and 5'-CGTACTAGCTTGGCTTCTA-

3') or Xfas304 helicase (5'- AAGAAGCGTGGTTTGTTTGC-3' and 5'-CTACCGGCTTCCCTAACTCC-3') were used to test for the presence of lysogens in phage resistant isolates. Wild type bacterial DNA was used as a negative control and wild type bacterial DNA spiked with phage DNA served as a positive control.

To examine possible phage repression under high-MOI conditions, we followed the procedure of Gill *et al.* (134). Overnight cultures of strain EC-12 were subcultured 1:50 into fresh TNB and cultured with aeration at 28°C to an OD600 of 0.3 to 0.5. Cells were pelleted by centrifugation and resuspended in 0.20 ml of phage lysate (harvested in TNB) or sterile TNB to a final OD600 of 1.0. The cell-phage mixtures were incubated for 30 min at 25°C and then centrifuged. Supernatants were removed and an aliquot (from the phage mixture) was serially diluted and plated to lawns of EC-12 to determine the phage titer remaining after initial incubation. To remove any reversibly bound phage, the cell pellets were resuspended in 0.20 ml of sterile TNB and were incubated for an additional 10 min at 25°C and then centrifuged. The supernatants were removed and an aliquot (from the phage mixture) was serially diluted and plated to lawns of EC-12 to determine the phage titer of unbound phage remaining post-incubation. The cell pellets were resuspended in 0.20 ml of sterile TNB, serially diluted and plated to enumerate the bacterial survivors remaining following phage exposure. The number of surviving CFU was compared to the CFU contained in the TNB-only control reaction mixtures to calculate the percentage of bacterial survivors. The post-incubation supernatant titers (initial and secondary) were compared to the titer of the phage stock to calculate the percentage of phage adsorbed to cells during the incubation period. The MOIactual (135)

was calculated as the ratio of the number of adsorbed phage to the number of CFU in the phage-free controls. The MOIactual values were used to calculate the proportions of cells adsorbing no phage, a single phage, and multiple phage based on a Poisson distribution (136). Three independent experiments were conducted using phages Xfas103 or Xfas303 with host strain EC-12.

Results and discussion

Phages Xfas103 and Xfas106 for X. fastidiosa

Nucleotide sequence and general sequence properties

The genomes of Xfas103 and Xfas106 were sequenced to completion by 454 pyrosequencing and genomic termini were determined experimentally. The general characteristics of the phage genomes are summarized in TABLE 5 and complete annotations with supporting evidence are provided in TABLE 6. The genomes of Xfas103 and Xfas106 were found to be approximately 56.1 and 55.6 kb, encoding 77 and 72 genes, respectively. It was experimentally determined that the genomes of Xfas103 and Xfas106 contain identical 12 bp (GTCAGCGCCCCC) 5'-protruding single-stranded cohesive (*cos*) ends. Head full packaging is not observed in phages with cohesive ends (e.g., phage lambda), indicating that is it unlikely Xfas103 and Xfas106 are capable of generalized transduction (137-138).

When analyzed by pairwise DNA MegaBLAST, the genomes of phages Xfas103 and Xfas106 can be aligned over 69.3% of their length, with an average 80.1% sequence

TABLE 5 General features of Xfas103, Xfas106, Xfas303 and Xfas304 genomes

	Phage			
Feature	Xfas103	Xfas106	Xfas303	Xfas304
Genome size (bp)	56,147	55,601	43,324	43,869
GC content (%)	62.3	63.0	65.3	60.2
Predicted No. of genes	77	72	52	51
Coding density (%)	95.4	96.9	96.0	93.9
Genomic termini[b]	5'-overhang	5'-overhang	DTR	DTR

TABLE 6 Comparative annotations of *Xylella* phages Xfas103 and Xfas106[a]

Gene(s)		% protein identity[b]	Strand[c]	Size (kDa)	Putative function	BcepNazgul homolog locus tag[d]	Evidence[e]
103gp*01,*	106gp*01*	52.2	R	10.2, 9.9	Hypothetical novel		
103gp*02,*	106gp*02*		R	7.1, 17.5	Hypothetical novel		
103gp*03,*	106gp*03*		R	7.2, 36.8	Hypothetical novel		
103gp*04,*	106gp*04*		R	13.1, 18.6	Hypothetical novel		
103gp*05,*	106gp*05*		R	5.0, 15.8	Hypothetical novel		
103gp*06*			R	7.3	Hypothetical novel		
103gp*07*			R	6.8	Hypothetical novel		
103gp*08*			R	20.3	Conserved protein		
103gp*09,*	106gp*06*	50.1	R	27.9, 22.2	Conserved protein		
	106gp*07*		R	5.4	Hypothetical novel		
103gp*10*			R	5.3	Hypothetical novel		
103gp*11,*	106gp*08*	21.4	R	7.3, 6.8	Hypothetical novel		
	106gp*09*		F	5.2	Hypothetical novel		
	106gp*10*		F	3.9	Hypothetical novel		
103gp*12*			R	19.6	Hypothetical novel		
103gp*13,*	106gp*11*	57.1	F	14.7, 15.1	Hypothetical novel		
103gp*14,*	106gp*12*	46.0	F	19.1, 20.0	Hypothetical novel		
103gp*15,*	106gp*13*	31.0	F	66.8, 61.9	Hypothetical novel		
	106gp*14*		F	9.5	Hypothetical novel		
103gp*16,*	106gp*15*	66.2	F	8.7, 9.2	Hypothetical novel		
103gp*17,*	106gp*16*	36.5	F	17.9, 19.2	Hypothetical novel		
103gp*18,*	106gp*17*	89.1	F	11.9, 12.0	Conserved protein		
103gp*19,*	106gp*18*	79.8	F	10.8, 10.8	Hypothetical novel		

TABLE 6 Continued,

Gene(s)		% protein identity[b]	Strand[c]	Size (kDa)	Putative function	BcepNazgul homolog locus tag[d]	Evidence[e]
103gp*20*			F	19.6	Hypothetical novel		
103gp*21*			F	6.5	Hypothetical novel		
103gp*22,*	106gp*19*	93.6	F	11.6, 11.7	Hypothetical novel		
103gp*23,*	106gp*20*	68.5	F	22.0, 23.1	Hypothetical novel		
103gp*24*			F	5.9	Hypothetical novel		
103gp*25,*	106gp*21*	56.5	F	11.8, 11.8	Hypothetical novel		
103gp*26,*	106gp*22*	70.3	F	8.2, 9.3	Hypothetical novel		
103gp*27,*	106gp*23*	34.1	F	12.4, 15.2	Hypothetical novel		
103gp*28,*	106gp*24*	68.4	F	26.3, 26.6	Conserved protein	NP_918968.1(E=1.24E-09, Dice=19.3)	
103gp*29,*	106gp*25*	52.1	F	16.0, 14.8	Hypothetical novel		
103gp*30,*	106gp*26*	49.4	F	19.5, 18.0	Hypothetical novel		
103gp*31,*	106gp*27*	57.6	F	24.8, 25.5	Hypothetical novel		
103gp*32,*	106gp*28*	12.8	F	10.2, 10.3	Hypothetical novel		
103gp*33,*	106gp*29*	62.5	F	21.9, 22.1	Hypothetical novel		
103gp*34,*	106gp*30*	76.9	F	17.0, 17.0	Conserved protein		IPR012816, PF08719, TIGR02464, SSF143990
	106gp*31*		F	12.7	Hypothetical novel		
103gp*35,*	106gp*32*	53.1	F	27.1, 27.7	Conserved protein		SSF52374, SSF52540, (PF13207, Xfas103 only) (IPR002891, PF01583 Xfas106 only)

TABLE 6 Continued,

Gene(s)		% protein identity[b]	Strand[c]	Size (kDa)	Putative function	BcepNazgul homolog locus tag[d]	Evidence[e]
103gp*36*			F	11.4	Hypothetical novel		
103gp*37,*	106gp*33*	75.9	F	34.1, 34.2	Conserved protein		(IPR002815, SSF56726, Xfas106 only)
103gp*38*			F	9.7	Hypothetical novel		
103gp*39,*	106gp*34*	30.1	R	11.2, 13.5	Rz1		(Lipoprotein adjacent to inner membrane spanin component, Xfas103 only) (lipoprotein overlapped with inner membrane spanin component, Xfas106 only)
103gp*40,*	106gp*35*	31.7	R	13.6, 15.4	Rz		N-terminal TMD
103gp*41,*	106gp*38*		R	17.7, 28.6	Putative antiholin		Predicted 1 TMD
103gp*42,*	106gp*37*		R	15.8, 16.3	Holin		(Predicted 4 TMDs, Xfas103 only) (predicted 2 TMDs, Xfas106 only)
103gp*43,*	106gp*36*		R	21.3, 31.5	Endolysin		(IPR024408, PF11860, HHpred hit 3gxr_A prob 97.7% E=4E-05, Xfas103 only) (LipoP, SignalP, HHpred hit 2anv_A prob 100% E=3E-37, Xfas106 only)

TABLE 6 Continued,

Gene(s)		% protein identity[b]	Strand[c]	Size (kDa)	Putative function	BcepNazgul homolog locus tag[d]	Evidence[e]
103gp*44,*	106gp*39*	46.0	R	37.4, 38.5	Tail fiber protein		(IPR001791, SM00282, Xfas103 only) IPR008985, SSF49899, PF13385, Positional evidence, see text
103gp*45,*	106gp*40*	72.8	R	34.8, 35.2	Conserved tail assembly protein		Positional evidence, see text
103gp*46,*	106gp*41*	85.7	R	6.3, 6.3	Conserved tail assembly protein		Positional evidence, see text
103gp*47,*	106gp*42*	83.9	R	27.2, 27.0	Conserved tail assembly protein		Positional evidence, see text
103gp*48,*	106gp*43*	75.2	R	90.7, 89.5	Conserved tail assembly protein	NP_918976.1 (E=1.45E-60, Dice=25.0)	PF13550, positional evidence, see text
103gp*49,*	106gp*44*	78.0	R	30.4, 30.3	Conserved tail assembly protein	NP_918979.1 (E=8.65E-18, Dice=26.0)	IPR018964, PF09356, IPR019228, PF09931, Positional evidence, see text
103gp*50,*	106gp*45*	75.1	R	64.0, 63.4	Conserved tail assembly protein	NP_918980.1(E=1.79E-13, Dice=11.5)	Positional evidence, see text
103gp*51,*	106gp*46*	72.5	R	33.2, 33.0	Conserved tail assembly protein		Positional evidence, see text
103gp*52,*	106gp*47*	80.8	R	36.4, 36.7	Conserved tail assembly protein		Positional evidence, see text
103gp*53,*	106gp*48*	77.1	R	156.6, 156.3	Tape measure protein		IPR013491, TIGR02675
103gp*54,*	106gp*49*	67.8	R	25.6, 25.7	Pre-tape measure frameshift protein GT	NP_918998.2(E=6.45E-21, Dice=27.0)	Positional evidence, see text

TABLE 6 Continued,

Gene(s)		% protein identity[b]	Strand[c]	Size (kDa)	Putative function	BcepNazgul homolog locus tag[d]	Evidence[e]
103gp*55,*	106gp*50*	63.2	R	18.6, 18.1	Pre-tape measure frameshift protein G	NP_918998.2(E=4.52E-04, Dice=16.1)	Positional evidence, see text, slippery sequence 5'-GGGAAAC-3'
103gp*56,*	106gp*51*	79.2	R	28.4, 28.7	Conserved, virion associated protein	NP_918986.1(E=8.77E-63, Dice=42.6)	
103gp*57,*	106gp*52*	78.9	R	19.2, 18.8	Conserved, virion associated protein	NP_918987.2(E=8.60E-22, Dice=30.6)	
103gp*58,*	106gp*53*	79.3	R	22.5, 22.8	Conserved protein	NP_918988.2(E=7.56E-32, Dice=38.9)	
103gp*59,*	106gp*54*	61.7	R	13.0, 12.6	Conserved protein	NP_918989.2(E=2.26E-20, Dice=32.3)	
103gp*60,*	106gp*55*	55.7	R	10.4, 9.8	Conserved protein		
103gp*61,*	106gp*56*	90.3	R	38.3, 38.5	Major capsid protein	NP_918991.1(E=1.81E-86, Dice=42.0)	IPR005564, PF03864
103gp*62,*	106gp*57*	75.4	R	14.2, 13.8	Decorator protein	NP_918992.1(E=1.72E-14, Dice= 28.9)	IPR004195, PF02924
103gp*63,*	106gp*58*	66.7	R	14.6, 15.0	Scaffold	NP_918993.2(E=4.01E-15, Dice=32.3)	
103gp*64,*	106gp*59*	75.2	R	47.8, 47.9	Prohead protease	NP_918994.2(E=1.36E-90, Dice=38.2)	IPR002142, PF01343, PTHR32497, SSF52096
103gp*65,*	106gp*60*	83.6	R	62.1, 61.6	Portal protein	NP_918995.1(E=3.37E-168, Dice=45.7)	IPR006429, PF05136, TIGR01539
103gp*66,*	106gp*61*	78.1	R	8.0, 8.1	Head-to-tail joining protein	NP_918996.1(E=2.70E-12, Dice=29.1)	IPR004174, PF02831, SSF64210
103gp*67,*	106gp*62*	83.6	R	80.4, 79.2	TerL	NP_918997.2(E=0, Dice=48.3)	IPR008866, PF05876

TABLE 6 Continued,

Gene(s)		% protein identity[b]	Strand[c]	Size (kDa)	Putative function	BcepNazgul homolog locus tag[d]	Evidence[e]
103gp*68,*	106gp*63*	82.0	R	22.0, 22.4	TerS	NP_918999.1(E=4.62E-43, Dice=30.4)	IPR009901, PF07278
103gp*69,*	106gp*64*	78.3	R	54.2, 53.9	Helicase	NP_919000.2(E=9.34E-66, Dice=29.3)	IPR000330, PF00176, IPR014001, SM00487, SSF52540
103gp*70,*	106gp*65*	77.2	R	11.7, 13.7	Holliday-junction resolvase	NP_919001.2(E=3.75E-18, Dice=30.8)	IPR014883, PF08774, SM00990
103gp*71,*	106gp*66*	78.6	R	79.1, 77.8	DNA polymerase I	NP_919002.1(E=6.40E-167, Dice=40.2)	IPR001098, PF00476, SM00482, SSF56672
103gp*72,*	106gp*67*	44.0	R	31.1, 30.3	Putative ligase	NP_919004.1(E=4.38E-25, Dice=25.4)	IPR012340, SSF50249, IPR022595, PF10991
103gp*73,*	106gp*68*	73.7	R	46.6, 44.9	Conserved protein	NP_919005.2(E=7.93E-59, Dice=27.9)	IPR021229, PF10926
103gp*74,*	106gp*69*	53.8	R	16.7, 16.5	Putative lipoprotein		LipoP
103gp*75,*	106gp*70*	86.0	F	10.3, 12.2	Hypothetical novel		

TABLE 6 Continued,

Gene(s)		% protein identity[b]	Strand[c]	Size (kDa)	Putative function	BcepNazgul homolog locus tag[d]	Evidence[e]
103gp*76,*	106gp*71*	80.5	F	95.5, 95.5	DNA primase	NP_919008.2(E=1.27E-158, Dice=33.0)	IPR014819, PF08707, IPR015330, PF09250, SM00943, SSF52540, SSF56747
103gp*77,*	106gp*72*	52.6	F	10.5, 11.6	Hypothetical novel		

[a] Putative assignable functions and protein molecular masses are shown.
[b] Protein identities between Xfas103 and Xfas106 homologs were calculated by the Dice method using BLASTp alignment identity scores.
[c] F/R, coding region on forward or reverse strand, respectively.
[d] E value and Dice identity reported for phage BcepNazgul homologs in Xfas103.
[e] Evidence includes detected conserved domains or motifs, HHpred hits, or other evidence as detailed in the text.

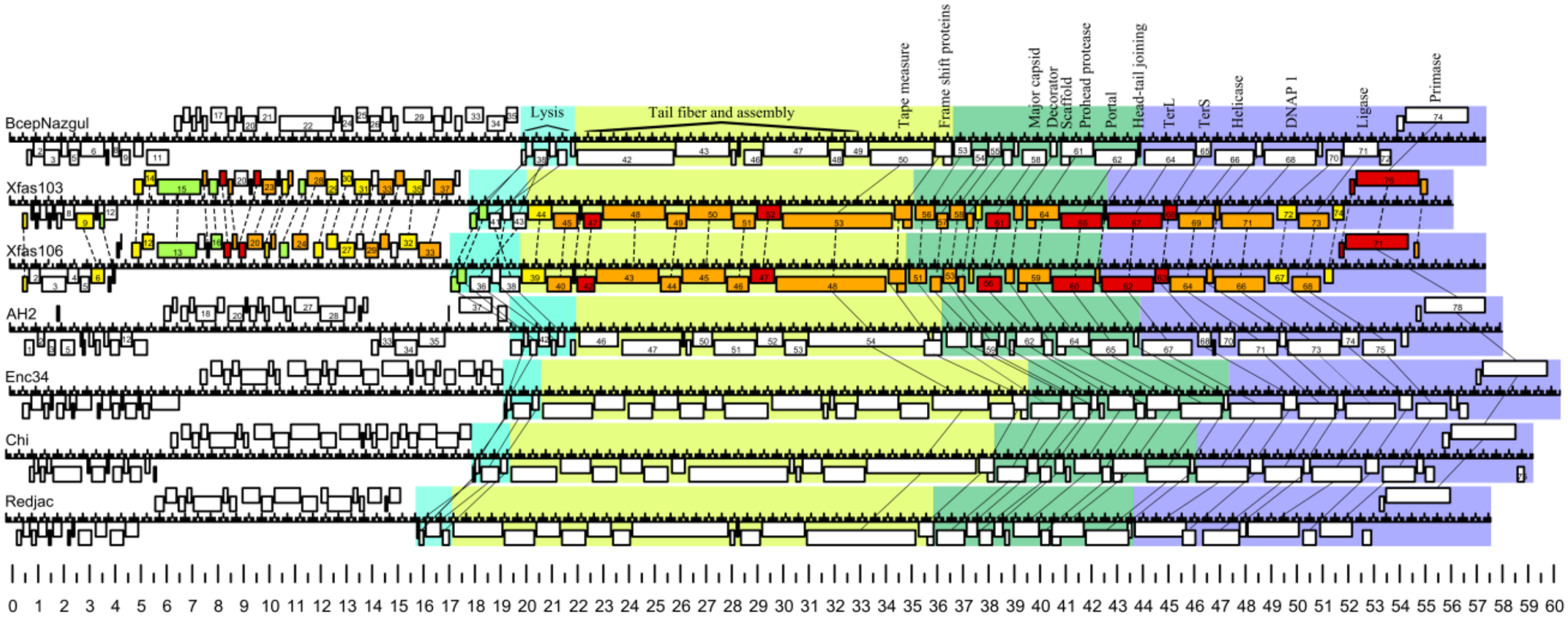

FIG 2 Synteny of phages Xfas103 and Xfas106 as compared to *Burkholderia* phage BcepNazgul, *Burkholderia* phage AH2, *Enterobacter* phage Enc34, *Escherichia* phage Chi and *Providencia* phage Redjac. connected by solid lines. Similarity at the amino acid level between homologous genes of Xfas303 and Xfas304 are indicated by color: white, <20%; green, 20-39.9%; yellow, 40-59.9%; orange, 60-79.9%; re green, head protein; turquoise green, tail protein; purple, DNA metabolism.

identity in aligned regions. Xfas103 and Xfas106 contain 58 homologous genes with amino acid sequence similarities ranging from 12.8% to 93.6%, and have retained genome synteny (TABLE 6, FIG 2). Although there is no direct evidence of transcriptional boundaries, there are few breaks in coding sequence large enough to contain promoters and there are no detected terminators, thus the simplest interpretation is that the genomes are organized into four divergent transcriptional units based on coding strand position. Genes found in transcriptional unit one (starting at the left end, reverse strand) and two (forward strand) primarily encode proteins of unknown function and show very limited similarity to other phage genes in the NCBI nr database. The lysis cassette, structural module and DNA metabolism module are all encoded for in transcriptional unit three (reverse strand), whereas transcriptional unit four (forward strand) encodes only a few genes, including a predicted primase.

When analyzed by BLASTp, Xfas103 and Xfas106 are most closely related to *Burkholderia* phage BcepNazgul (NC_005091); however, the relationship is distant. Of the 77 predicted Xfas103 genes, only 23 show homology to BcepNazgul, ranging from 11.0%-48.2% sequence identity at the amino acid level. These genes are primarily located in the structural and DNA metabolism regions (excluding tail fiber and assembly proteins). Despite the low observed homology, phages Xfas103 and Xfas106 are largely syntenic with BcepNazgul (FIG 2).

As previously noted by Gill *et al.*(56), the assignment of phage groups based on protein or DNA sequence relationships, without regard to overall genome architecture, results in numerous phage groups with no obvious evolutionary linkage to each other.

Casjens *et al.* (139) have proposed a "phage type" model that groups phages based on shared features of genome organization and relatedness of genes. BcepNazgul does not belong to any existing phage types because of its unique architecture and low homology to previously described phages. We therefore propose that BcepNazgul is the founder of a new phage type, as defined by Casjens, to which both Xfas103 and Xfas106 belong. Further analysis and reannotation of the genomes of *Burkholderia* phage AH2, *Enterobacter* phage Enc34, *Escherichia* phage Chi and *Providencia* phage Redjac indicate they are BcepNazgul-like phages (FIG 2). Another factor that must be taken into consideration when classifying phage type is lifestyle. Phage Enc34 has been reported to be virulent (140), whereas there is no report of the phage lifestyle of Chi and Redjac (141-142). Phage AH2 was not defined by the authors as temperate (139), but there is clearly evidence that it may be a temperate phage (addressed below).

Gene organization and homology

In the following, except where significant differences exist between Xfas103 and Xfas106, only the relationship between Xfas103 proteins and database homologs will be discussed in detail (FIG 2, TABLE 6).

Transcriptional units one and two

Most proteins encoded in transcriptional units one and two do not have phage homologs in the database. Despite this, 24 of the 38 proteins encoded in Xfas103 transcriptional units one and two have homologs in Xfas106, with 12.8% to 93.6% sequence identity at the amino acid level (FIG 2, TABLE 6). Only one protein (Xfas103 gp28) has a homolog in phage BcepNazgul (BcepNazgul gp34).

Lysis genes

Xfas106 and BcepNazgul contain lysis cassettes with the canonical gene order of holin-endolysin-Rz/Rz1 (genes Xfas106 *34-37* and genes BcepNazgul *36-39*, respectively), whereas in Xfas103 the holin gene (gp42) is found upstream of the endolysin (gp43) (FIG 2).

Spanins were identified by their unique primary and secondary features. Xfas103 Rz gp40 is a type II (N-in, C-out) cytoplasmic membrane protein (as predicted by TMHMM), and Xfas103 Rz1 gp39 is an OM protein (as predicted by LipoP). The Xfas103 spanin complex belongs to the separated spanin class, with overlapping *Rz* stop and *Rz1* start codons, as in phage T4. Despite protein homology between the spanin components of Xfas103 and Xfas106, Xfas106 *Rz* (gp35) and *Rz1* (gp34) are ordered in an overlapped arrangement, as in BcepNazgul (FIG 2).

Phage endolysins in Xfas103 (gp43) and Xfas106 (gp36) were identified by the presence of catalytic motifs and amino acid homology to phage endolysin proteins in the database (TABLE 6). The Xfas103 endolysin contains a DUF3380 domain (IPR024408), which has been defined as a peptidoglycan binding domain in phage PRD1 (143). Xfas106 gp36 is a lipoprotein (based on LipoP), containing a C-terminal lysozyme catalytic domain as detected by HHpred (hit 2anv_A, 100% probability), with a weak N-terminal signal sequence as predicted by SignalP, which suggests gp36 may be a SAR endolysin.

Phage holin genes are extremely diverse, but most have at least one TMD and are usually located within the lysis cassette adjacent to the endolysin and spanin genes.

Three experimentally verified holin topologies have been previously been identified: class I (3 TMDs with N-out, C-in); class II (2 TMDs with N-in, C-in); class III (one TMD with N-in, C-out) (67). Xfas103 gp42 is a putative holin, based on its location clustered within the lysis cassette and the presence of multiple TMDs. Amino acid charge analysis of Xfas103 gp42 suggests four strong TMDs with three net-positive soluble domains and a negatively charged C-terminal, which represents a novel holin membrane topology (FIG 3). Xfas106 *gp37* is predicted to encode a class II holin, but contains an unusually large total sequence and soluble domains (FIG 3). The BcepNazgul holin (gp39) has all the characteristics of a class I holin, including and net-positive N-terminal and basic C-terminal ends. However, TMD3 contains 13 amino acids bound by arginine residues, and is too small to normally be considered a TMD.

A

```
BcepNazgul
     -                        +   + +    -- +   +                  -   -
MNNLFELAAVVAAVALAALVAVFLYSRFFLKQRVQLVDDWRNLVKFSSTWFIAVGVVLQSWIEVAPDTLL
        10         20         30         40         50         60         70

       -- +         +              +  +  +  +   + -     -
HVWVLMPDDMKALIPPQYARFVPIAIIVLAVPFRIIKQARLARTQPRPETPNAPDHS
        80         90        100        110        120
Xfas103
 - + +   -           + +                 +      +
MEIKPRIGADLWHAITQGTTLKVRFILALQSLIFGASLLVRPMPAPTRWALFATLPVYAWSAMLWLSGSL
        10         20         30         40         50         60         70

   +    +                               -  +                 + -  + - -
MLWRVLSTRPVTWLAWASNVFALFTWFIICLSYAAFEGWRGLVGTHILALVMAVFCVLRTEATRSDLETA
        80         90        100        110        120        130        140
Xfas106
 -          - +   +++  -  ++                          +  - +   -   -
MDTNNSSPANPESKMPFRRRAWEFWRRNQIILFQLPVMLVFLIGSYIVLKSVDSRIGVEGFGDIFGYALN
        10         20         30         40         50         60         70

  +           ++     -  -+ - -  - ++   +-       -+ -             +
AVRITLIIFTTWWMKKWCWFDLHDRTELELFDLRKQGNRDAHWIVWTDRIEWMLALAFATFWYTR
        80         90        100        110        120        130
AH2
 + +  --  +  +                   -          - +-   -   +             +
MKLRLIDDWHKAWKLGSVQFAALFAFLFSIGPDLLHTWAFIPQDLKDALPEGTSRWIAVAALLLTLLGRV
        10         20         30         40         50         60         70

 + -++    --
FKFERKTPQGEEQ
        80
```

B

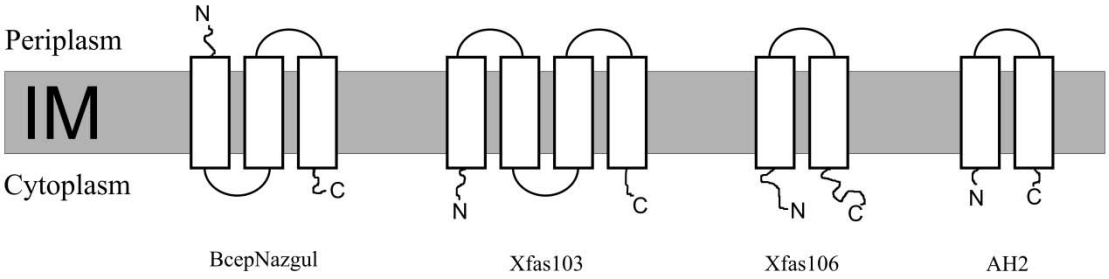

FIG 3 Primary and topological properties of holin proteins. (A) Primary structures of BcepNazgul gp39, Xfas103 gp42, Xfas106 gp37 and AH2 gp43. TMDs are shown in bold text, with charged residues indicated above. (B) Topological models of BcepNazgul, Xfas103, Xfas106 and AH2 predicted holins. IM: inner membrane.

The last TMD of the class II AH2 holin (gp43) also contains 13 amino acids bound by arginine residues (FIG 3). In both cases, the C-terminal TMD is too short and may have a kinetic defect in membrane localization. This can lead one to speculate that possibly the last TMD in these two otherwise unrelated proteins are kinetically challenged in terms of integration and are not inserted into the membrane until the membrane potential collapses. This would suggest that topological determinants can force C-terminal TMDs in the membrane despite "illegal" R to R spacing, possibly resulting in a new form of dynamic topological control of holin function.

Dual start and non-dual start phage antiholins are utilized by some phages to control phage lysis timing (68). Phages Xfas103, Xfas106 and BcepNazgul all have a hypothetical protein with 1 TMD was found adjacent (BcepNazgul gp40 and Xfas106 gp38) or within (Xfas103 gp41) the lysis cassette, and are thus predicted antiholins.

Structural genes

There is limited homology between the tail fiber and assembly genes of Xfas103 genes with BcepNazgul. However, Xfas103 gp*44-52* can be presumptively assigned the functions of tail fiber and assembly proteins based on synteny to similar genes in BcepNazgul (gp41-49). Xfas103 gp44 contains a laminin G domain (IPR001791) and a concanavalin A-like lectin/glucanase domain (IPR008985). Laminin G domains are associated with binding functions, and several con A-like domains are found in proteins involved in cell recognition and adhesion. BcepNazgul tail fiber protein (gp42) has a similar galactose-binding domain (IPR008979), which is associated with binding to specific ligands such as cell-surface-attached carbohydrate substrates. Xfas103 gp53 is a

predicted phage tape measure protein based on the detection of tape measure domain (IPR013491). Xfas103 gp54 and gp55 are predicted to encode frameshift proteins GT and G, respectively. A programmed translational frameshift is strongly conserved in tail assembly genes of dsDNA phages, and has been shown to be required for proper tail morphogenesis (144). A manual search of G proteins identified the slippery sequence 5'-GGGAAAC-3'for both Xfas103 and Xfas106.

Xfas103 head assembly and DNA packaging genes are syntenic and share homology to genes in BcepNazgul. Xfas103 encodes homologs of BcepNazgul gp53-56 (virion associated proteins), gp58 (major capsid), gp59 (decorator), gp60 (scaffold), gp61 (prohead protease), gp62 (portal) and gp63 (head-to-tail joining protein), with 11.0-42.6% sequence identity at the amino acid level. Therefore Xfas103 gp56-59, gp61, gp62, gp63, gp64, gp65 and gp66 are predicted to function as virion associated proteins, major capsid, decorator, scaffold, prohead protease, portal and head-tail joining protein, respectively. The protease and scaffold reading frames are fused, with the scaffold start codon downstream within the protease gene, similar to that observed in gpC and gpNu3 proteins of phage lambda (145). The predicted portal protein (gp65) and head-to-tail joining protein (gp66) are members of the phage lambda GpB portal (IPR006429) and head-to-tail joining protein W GpW family (IPR004174), respectively.

Xfas103 large subunit terminase (TerL) and small subunit terminase (TerS) proteins were identified by amino acid homology to BcepNazgul homologs gp64 and gp65, respectively. The TerL subunit contains all of the catalytic activities required for DNA maturation and packaging, whereas the TerS subunit is required for specific

recognition of viral DNA. As in phage lambda, the TerS of BcepNazgul-like phages recognize phage *cos* sites for initiation and termination of DNA packaging, removing the risk of generalized phage transduction.

DNA metabolism and replication

Xfas103 *gp69* is predicted to encode a phage helicase based on homology to the BcepNazgul helicase (gp66). Xfas103 gp70 is homologous to BcepNazgul gp67 (conserved phage protein), and is predicted to be a Holliday-junction resolvase. This function is supported by the presence of a VRR-NUC domain (IPR014883) that is associated with nuclease activity, and by HHpred predicted similarity to Holliday-junction resolvases (1hh1_A with 98.5% probability). Xfas103 gp71 is homologous to the BcepNazgul DNA polymerase 1 (gp68), and contains a DNA-directed DNA polymerase family A palm domain (IPR001098). Xfas103 gp72 is homologous to BcepNazgul gp70, a predicted conserved protein. Both Xfas gp72 and BcepNazgul gp70 contain a phage APSE-1 protein 50 family domain (IPR022595) and a nucleic acid-binding OB-fold domain (IPR012340). These domains are found in ATP and NAD-dependent DNA ligases, indicating that Xfas103 gp72 may be a putative DNA ligase. No phage RNA polymerase (RNAP) was identified in Xfas103, indicating that all transcription is accomplished by the host RNAP. Xfas103 gp76 is a predicted primase based on amino acid homology to BcepNazgul gp74 (primase) and other phage homologs. It contains a primase C-terminal 2 (IPR014819) and DNA primase/polymerase bifunctional N-terminal domain (IPR015330), and is separated from the DNA replication genes in transcriptional unit three.

Nucleotide sequence and general sequence properties

The genomes of Xfas303 and Xfas304 were sequenced to completion by 454 pyrosequencing, and genomic termini were determined experimentally. The general characteristics of the phage genomes are summarized in TABLE 5 and complete annotations with supporting evidence are provided in TABLE 7. The genomes of Xfas303 and Xfas304 were found to be approximately 43.3 and 43.8 kb, encoding 52 and 51 genes, respectively. It was experimentally determined that the genomes of Xfas103 and Xfas106 contain non-permutated short direct terminal repeats of 619 and 675 bp, respectively.

Analysis indicates that phages Xfas303 and Xfas304 share no significant DNA homology when compared by pairwise DNA megaBLAST. Despite extreme DNA sequence divergence, Xfas303 and Xfas304 contain 43 homologous genes with amino acid similarities ranging from 17.1% to 79.0%, and have retained canonical phiKMV-like gene order (FIG 4, TABLE 7). Similar to T7-like phages, phiKMV-like phage genomes consist of three functional gene clusters. Class I encodes for early genes with functions to overcome host restriction and to convert the metabolism of the host cell to

the production of phage proteins. Class II encodes for proteins involved in DNA metabolism. Class III encodes for structural and assembly proteins, DNA packaging proteins and lysis proteins. phiKMV-like phages are distinguished from T7-like phages in that they encode for a single subunit RNA polymerase (ssRNAP) at the end of the class II gene cluster, rather than in the early genomic region (146). Sequenced phiKMV-like phages have been reported for a wide range of hosts, including members of the Gamma-protebacteria and Beta-proteobacteria.

Xfas303 and Xfas304 are clearly phiKMV-like based on genome organization and protein homology with other phiKMV-like phages. Of the sequenced phiKMV-like phages, Xfas303 is most similar to *Caulobacter* phage Cd1 (GU393987), with 26 homologous genes ranging from 7.6-54.8% amino acid similarity.

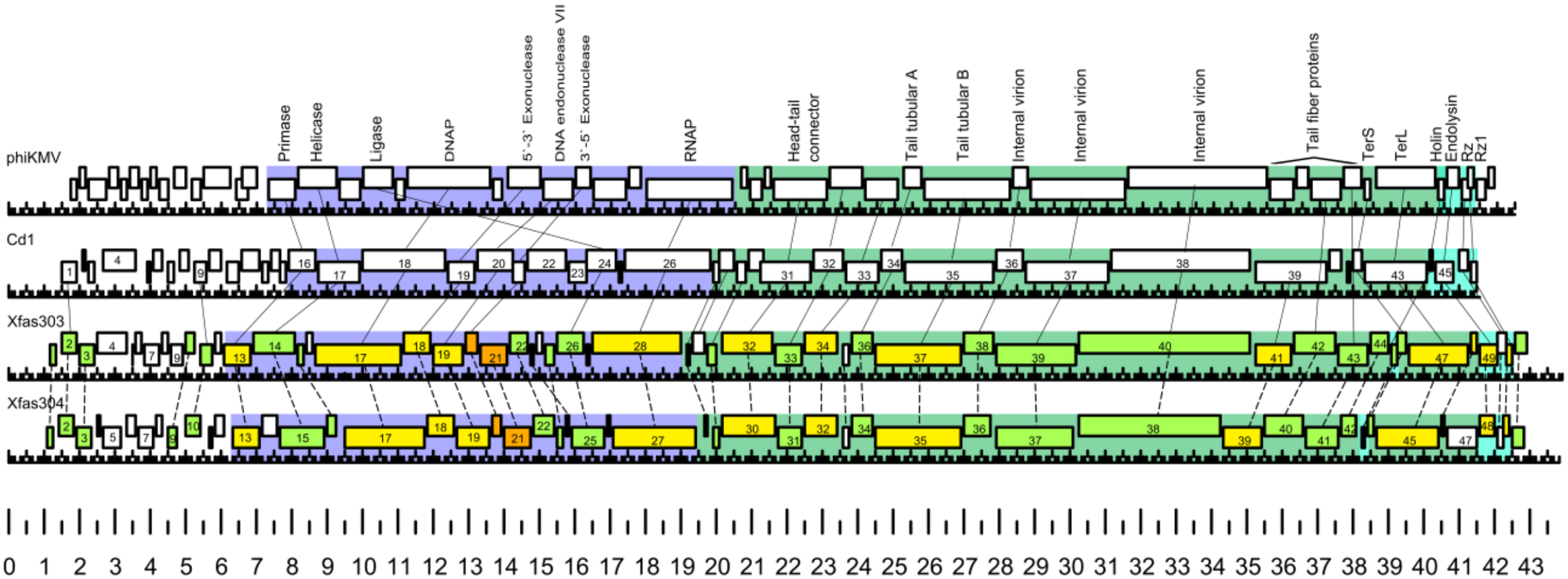

FIG 4 Synteny of phages Xfas303 and Xfas304 as compared to *Pseudomonas* phage phiKMV and *Caulobacter* phage Cd1. Predicted reading frames are indicated by boxes. Functionally equivalent genes are connected by lines. Similarity at the amino acid level of homologous genes between Xfas303 and Xfas304 is indicated by color: white, <20%; green, 20-39.9%; yellow, 40-59.9%; orange, 60-79.9%; red, ≥80%. Functional modules are represented by different colors: light blue, lysis; turquoise green, structural protein; purple, DNA metabolism.

TABLE 7 Comparative annotations of *Xylella* phages Xfas303 and Xfas304[a]

Gene(s)		% protein identity[b]	Strand[c]	Size (kDa)	Putative function	Phage Cd1 homolog locus tag[d]	Evidence[e]
303gp*01,*	304gp*01*	26.9	F	9.2, 10.3	Hypothetical novel		
303gp*02,*	304gp*02*	21.2	F	19.7, 19.6	Conserved protein	ADD21636.1(E=6.68E-07, Dice=26.2)	
303gp*03,*	304gp*03*	21.4	F	18.1, 16.8	Conserved protein		
303gp*04*			F	32.9	Conserved protein		
	304gp*04*		F	6.6	Hypothetical novel		
303gp*05,*	304gp*05*		F	5, 22.2	Hypothetical novel		
303gp*06,*	304gp*06*		F	8.2, 13.9	Hypothetical novel		
303gp*07,*	304gp*07*		F	19.5, 21.2	Hypothetical novel		
303gp*08,*	304gp*08*		F	11.3, 10.8	Hypothetical novel		
303gp*09*			F	13.5	Hypothetical novel		
303gp*10,*	304gp*09*	24.7	F	12.3, 15.8	Hypothetical novel		
303gp*11,*	304gp*10*	29.0	F	15, 15.8	Conserved protein	ADD21644.1(E=5.10E-08, Dice=22.1)	
	304gp*11*		F	6.9	Hypothetical novel		
303gp*12,*	304gp*12*		F	8.1, 11.8	Hypothetical novel		
303gp*13,*	304gp*13*	50.5	F	31.1, 31.8	DNA primase	ADD21651.1 (E= 9.66E-47, Dice=35.0)	cd01029
	304gp*14*		F	18.1	Hypothetical novel		
303gp*14,*	304gp*15*	38.4	F	48.4, 49.8	DNA helicase	ADD21652.1 (E=9.62E-109, Dice=40.9)	PF13481, (SSF52540 Xfas304 only)

TABLE 7 Continued,

Gene(s)		% protein identity[b]	Strand[c]	Size (kDa)	Putative function	Phage Cd1 homolog locus tag[d]	Evidence[e]
303gp*15,*	304gp*16*	25.3	F	10.6, 12.3	Hypothetical novel		
303gp*16*			F	7.5	Hypothetical novel		
303gp*17,*	304gp*17*	51.9	F	92.5, 88.9	DNA polymerase A	ADD21653.1(E=0, Dice=51.4)	IPR001098, PF00476, SM00482, IPR002298, PR00868, PTHR10133, SSF56672
303gp*18,*	304gp*18*	40.1	F	32.1, 31.7	Conserved protein	ADD21654.1(E=5.07E-54, Dice=35.3)	
303gp*19,*	304gp*19*	52.1	F	34.4, 35.1	5'-3' exonuclease	ADD21655.1(E=2.63E-72, Dice=38.5)	IPR020045,SSF47807
303gp*20,*	304gp*20*	70.2	F	14.1, 12.8	DNA endonuclease VII	ADD21656.1(E=6.39E-34, Dice=45.9)	IPR004211, PF02945, SSF54060
303gp*21,*	304gp*21*	79.2	F	31.8, 32.2	3' to 5' exonuclease		IPR012337, SSF53098
303gp*22,*	304gp*22*	30.3	F	22.1, 23.7	Conserved protein		
303gp*23,*	304gp*24*	75.2	F	8, 8.1	Conserved protein		
303gp*24*			F	8.2	Hypothetical novel		
303gp*25,*	304gp*23*	20.3	F	11.4, 9.3	Conserved protein		
303gp*26,*	304gp*25*	35.0	F	33.3, 37.6	ATP-dependent DNA ligase	ADD21659.1(E=2.41E-47, Dice=36.3)	(IPR012310, PF01068 Xfas303 only) IPR012340, SSF50249, SSF56091;
303gp*27,*	304gp*26*	43.9	F	6.7, 6.4	Hypothetical novel		

TABLE 7 Continued,

Gene(s)		% protein identity[b]	Strand[c]	Size (kDa)	Putative function	Phage Cd1 homolog locus tag[d]	Evidence[e]
303gp*28,*	304gp*27*	56.3	F	95.6, 91.7	DNA-dependent RNA polymerase	ADD21661.1(E=0, Dice=51.0)	IPR002092, PTHR10102, PF00940, SSF56672
303gp*29,*	304gp*28*	26.8	F	7.1, 6.5	Conserved protein	ADD21662.1(E=2.80E-06, Dice=23.6)	
303gp*30*			F	15.2	Conserved protein	ADD21663.1(E=2.42E-16, Dice=24.5)	
303gp*31,*	304gp*29*	30.1	F	12.1, 11.3	Conserved protein	ADD21664.1(E=3.00E-05, Dice=31.1)	
303gp*32,*	304gp*30*	47.2	F	56.5, 56.5	Head -tail connector	ADD21666.1(E=2.43E-154, Dice=44.5)	IPR020991, PF12236
303gp*33,*	304gp*31*	22.2	F	28.9, 26.3	Scaffold	ADD21667.1(E=4.29E-06, Dice=18.5)	(PHA01929 Xfas303 only)
303gp*34,*	304gp*32*	57.1	F	36.3, 36.3	Major capsid protein	ADD21668.1(E=2.91E-84, Dice=43.2)	cl18086
303gp*35,*	304gp*33*	19.4	F	7.4, 7.9	Hypothetical novel		
303gp*36,*	304gp*34*	33.6	F	23.9, 24.0	Tail tubular protein A	ADD21669.1(E=4.31E-35, Dice=29.3)	cl10205
303gp*37,*	304gp*35*	42.0	F	91.5, 93.8	Tail tubular protein B	ADD21670.1(E=0, Dice=36.8)	
303gp*38,*	304gp*36*	38.8	F	31.4, 30.2	Internal Virion protein	ADD21671.1(E=6.34E-21, Dice=21.7)	
303gp*39,*	304gp*37*	28.0	F	85.0, 85.1	Internal Virion protein	ADD21672.1(E=5.50E-29, Dice=12.0)	
303gp*40,*	304gp*38*	34.4	F	176.3, 147.1	Internal Virion protein with transclycosylase activity	ADD21673.1(E=2.85E-120, Dice=18.9)	IPR023346, SSF53955

TABLE 7 Continued,

Gene(s)		% protein identity[b]	Strand[c]	Size (kDa)	Putative function	Phage Cd1 homolog locus tag[d]	Evidence[e]
303gp*41,*	304gp*39*	46.0	F	39.8, 39.7	Tail fiber protein	ADD21674.1(E=1.57E-15, Dice=7.6)	IPR005604, PF03906
303gp*42,*	304gp*40*	32.8	F	44, 43.3	Putative tail fiber protein		See text
303gp*43,*	304gp*41*	24.5	F	33.6, 34.7	Putative tail fiber protein		See text
303gp*44,*	304gp*42*	36.2	F	21.2, 20.7	Putative tail fiber protein		See text
303gp*45,*	304gp*43*	38.2	F	7.3, 7.0	Holin		Predicted 2 TMDs, see text
303gp*46,*	304gp*44*	27.7	F	10.1, 10.2	TerS	ADD21677.1(E=1.80E-05, Dice=21.9)	
303gp*47,*	304gp*45*	59.8	F	65.8, 65.7	TerL	ADD21678.1(E=0, Dice=54.9)	
303gp*48,*	304gp*46*	41.9	F	8.5, 7.2	Hypothetical novel		
	304gp*47*		F	29.8	Conserved protein		
303gp*49,*	304gp*48*	49.3	F	19.2, 18.3	Endolysin	ADD21680.1(E=1.52E-25, Dice=32.0)	IPR002196, PF00959, IPR023346, SSF53955
303gp*50,*	304gp*49*	17.1	F	11.2, 10.8	Rz		N-terminal TMD

TABLE 7 Continued,

Gene(s)		% protein identity[b]	Strand[c]	Size (kDa)	Putative function	Phage Cd1 homolog locus tag[d]	Evidence[e]
303gp*51,*	304gp*50*	41.9	F	10.7, 10.2	Rz1	ADD21682.1(E=1.25E-05, Dice=20.1)	Lipoprotein overlapped with inner membrane spanin component
303gp*52,*	304gp*51*	39.7	F	15.1, 14.1	Conserved protein		

[a] Putative assignable functions and protein molecular masses are shown.
[b] Protein identities between Xfas303 and Xfas304 homologs were calculated by the Dice method using BLASTp alignment identity scores.
[c] F/R, coding region on forward or reverse strand, respectively.
[d] E value and Dice identity reported for phage Cd1 homologs in Xfas303.
[e] Evidence includes detected conserved domains or motifs, HHpred hits, or other evidence as detailed in the text.

Gene organization and homology

In the following, except where significant differences exist between Xfas303 and Xfas304, only the relationship between Xfas303 proteins and database homologs will be discussed in detail (FIG 4, TABLE 7).

Early genes

Early genes are involved in host conversion and have been reported to protect phage from bacterial defense mechanisms or alter host cell mechanisms (147). When comparing phages of the same phage type, phage proteins that interact most intimately with the bacterial host have been reported to be the most diverse (148). Five out of 12 Xfas303 early genes have homologs in Xfas304. Only two early genes, Xfas gp*02* and gp*11*, have homologs in other phiKMV-like phages, namely Cd1 gp*01* and gp*09*, respectively.

DNA metabolism

Phage Xfas303 DNA metabolism genes follow the canonical phiKMV-like gene order of DNA primase (gp13), DNA helicase (gp14), DNA polymerase A (gp17), 5'-3' exonuclease (gp19), DNA exonuclease VII (gp20), ATP-dependent DNA ligase (gp26), and an ssRNAP (gp28) (FIG 4). Xfas303 gp21 and gp22 do not have homologs in Cd1, but do have homologs in other phiKMV-like phages, such as *Ralstonia* phage RSB1 gp22 and gp24, respectively. The Xfas303 ssRNAP is located at the right end of the DNA metabolism gene cluster, a distinguishing characteristic of all reported phiKMV-like phages. However, it has been noted that the position of the DNA ligase gene can vary within phiKMV-like phages (149). In phiKMV, the DNA ligase is encoded at the

beginning of the DNA metabolism genes, whereas in Cd1 the DNA ligase is encoded directly before the RNA polymerase. Phage Xfas303 shares the gene order of phage Cd1 (FIG 4).

It is of note, that the DNA metabolism region of Xfas303 also shares high homology to several *Xanthomonas* phages, including CP1 (NC_019933), Xop411 (NC_009543), Xp10 (NC_004902), OP1 (NC_007709), phiL7 (NC_012742) and CP2 (NC_020205). These phages are non-phiKMV-like, yet have homology and preserved gene order characteristic of phiKMV-like DNA metabolism regions, including a ssRNAP located at the end of the region.

Structural and lysis genes

The Xfas303 structural region is in the canonical phiKMV-like gene order. It begins immediately downstream of the ssRNAP, with three small genes for conserved proteins of hypothetical function (gp29-31) that are homologs to Cd1 gp27-29. Next in sequence are genes encoding a head-tail connector (gp32), scaffold protein (gp33), major capsid protein (gp34), tail tube A (gp36), tail tube B (gp37), and three internal virion proteins (gp38-gp40). Xfas303 gp40 possesses a C-terminal lysozyme-like domain (IPR023346), identical to that found in Cd1 gp38. However, Xfas303 gp40 is ~300aa larger than any previously reported homolog in phiKMV-like phages. As observed in other phages, there is diversity within the tail fiber proteins of the phiKMV-like phages, both in terms of sequence homology and components (149). In phiKMV-like phages, tail structural proteins are found immediately after the internal virion proteins. Similarly, the Xfas303 tail fiber proteins (gp41-44) are encoded in the same location. Xfas303 gp41

contains a T7 tail fiber protein domain (IPR005604) and Xfas303 gp43 and gp44 have homologs to tail fiber proteins found in phiKMV-like phages. It appears, however that Xfas303 gp*41-44* may have been acquired through horizontal gene transfer with a non-phiKMV like phage. The four genes are found as a cluster and are homologs to genes gp*26-29* in *Xanthomonas* phage CP2, with 30.6-41.1% identity.

Xfas303 gp47 was assigned the function of TerL based on homology to other phiKMV-like phage TerL homologs in the database, and gp46 was designated as TerS based on its size and proximity to gp47. The lysis cassette of phage phiKMV has been studied in detail (150). The canonical phiKMV-like lysis cassette is organized with genes encoding for holin, endolysin, Rz and Rz1 found immediately downstream of TerS and TerL (150). The lysis cassette of Xfas303 is very similar to that of phiKMV. However, the predicted holin and endolysin are not found adjacent to each other. Instead they are separated by genes encoding for TerS (gp46), TerL (gp47) and a novel protein (gp48) (FIG 4).

Xfas303 *gp45* encodes a putative type II holin based on the presence of two TMDs (N-in, C-in). TMD2 (residues 31 to 48) is the most hydrophobic of the two. TMD1 (residues 13 to 29) is not predicted by TMHMM and has characteristics of a SAR domain, with a high percentage of weakly hydrophobic or polar residues. The class II topology and presence of a SAR domain followed by a typical TMD suggests that gp45 is a pinholin, similar to phiKMV gp45 (150). Xfas303 *gp49* encodes an endolysin, based on homology to other phiKMV-like phage endolysins in the database. Xfas303 endolysin gp49 exhibits an N-terminal hydrophobic domain rich in weakly hydrophobic

residues that is characteristics of a SAR endolysin (151). A glycoside hydrolase domain (IPR002196) is predicted by InterProScan and contains an E-8aa-D-8aa-T catalytic triad, similar to the E-8aa-D-5aa-T and E-9aa-D-6aa-T catalytic triads found in T4 protein E and phiKMV gp45, respectively (150, 152). However Xfas45 gp49 (173aa) is much smaller than the endolysin of phiKMV (898aa), and is more similar to the endolysin of Cd1 (185aa). SAR endolysins are embedded in the inner membrane where the enzyme is inactive. This prevents premature lysis and results in holin control of lysis timing (153). Xfas303 spanins were identified by characteristic gene orientation and proximity to the endolysin. Xfas303 Rz (gp50) and Rz1 (gp51) are subunits of a typical P2-like overlapped two component spanin complex, similar to that in phiKMV (150). Xfas303 gp52 is homologous with HslV family proteins of *Xanthomonas* phages OP1 (gp51), Xp10 (gp51) and Xop411 (gp51), with 38.4-42.1% identity, but does not contain the N-terminal hydrolase domain present in the other proteins.

Phage lifestyle

Based on genomic analysis, we can confidently predict that phages Xfas303 and Xfas304 are virulent, because they are phiKMV-like. The same confidence cannot be assigned to Xfas103 and Xfas106. Phages BcepNazgul, Xfas103 and Xfas106 all encode a small protein with a helix-turn-helix DNA-binding domain (gp73, gp75, and gp70, respectively). Similar proteins have been annotated as cro-like proteins in phages AH2, Enc34, Chi and Redjac (139-142). This would suggest that the BcepNazgul-like phages may have a temperate lifestyle. However, other than AH2, key genes required for lysogeny were not found in these phages. AH2 is predicted to encode three excisionase

proteins (gp30, gp31 and gp33) and an integrase (gp37). Homologs to these genes are not present in the other BcepNazgul-like phages. Furthermore, the AH2 excisionase and integrase genes are encoded in two adjacent transcriptional unites (based on coding strand position) not present in other BcepNazgul-like phages (FIG 2). This suggests that the BcepNazgul-like phages, excepting AH2, have diverged to become virulent, due to a possible ancient deletion of lysogeny-related genes. Phage AH2 is BcepNazgul-like, which brings into question whether phages with different lifestyles can truly be the same phage type.

To further examine the potential for lysogeny, 40 phage-insensitive isolates of strain Temecula and EC-12 each were recovered following a challenge by phage Xfas103, Xfas106, Xfas303, or Xfas304. PCR using phage specific primers did not detect the presence of phage lysogens in resistant isolates, indicating resistance was not due to lysogeny. However, the absence of stable lysogen formation is not a definitive indication for virulence (134). Additionally, we examined the effect of phage infection at a high MOI on cell survival, as previously described by Gill et al (134). As shown in TABLE 8, following phage exposure to Xfas103 or Xfas303, there was an observed 3.9-fold and 3.46-fold excess of survivors than would be predicted, respectively. The observed higher percent survival may be due to the high degree of reversible binding of Xfas103 and Xfas303 to strain EC-12 (55 and 46%, respectively). The observed excess percent survival of strain EC-12 following exposure to Xfas103 and Xfas303 (predicted virulent phage) was similar, and does not support the notion that Xfas103 is temperate.

TABLE 8 Actual MOI and predicted and actual bacterial survivors of *Xanthomonas* strain EC-12 following exposure to phage Xfas103 or Xfas303[a]

	Replicate No.	MOIactual	Predicted % surviving cells	Measured % surviving cells	Fold excess of survivors vs. prediction
Xfas103	1	2.64	7.14	26.1	3.65
	2	2.78	6.20	22.1	3.56
	3	2.88	5.61	25.7	4.58
Xfas303	1	2.81	6.02	13.9	2.30
	2	3.31	3.65	12.7	3.47
	3	3.67	2.55	11.8	4.62

[a] Predicted survivors were calculated from the Poisson distribution for the measured actual MOIs. Data shown are from three independent replicate experiments.

CHAPTER IV

RECEPTOR SITE IDENTIFICATION

Introduction

The susceptibility of a bacterium to phage infection is primarily dependent on whether or not the phage can adsorb to a specific attachment site(s) or receptor(s) on the cell surface (59). Since phages, like any other viruses, are obligate intracellular parasites, successful penetration into bacterial cell is an essential condition for continuation of their life cycle. A successful irreversible attachment may then be followed by the penetration of the phage nucleic acid into the cell, followed by intracellular replication of phage nucleic acid, ultimately leading in the release of new phage progeny (59).

The recognition of the receptor is a highly specific process and is part of the natural mechanism of host recognition (154). The cell surfaces of both Gram-negative and Gram-positive bacteria have been studied intensively from structural, biosynthetic, genetic, and functional viewpoints. As a result of these studies information of considerable value for investigations of phage receptor sites has emerged (59). Phages can use bacterial capsules, different components of lipopolysaccharide (LPS), flagella, fimbriae, outer membrane proteins or pili as receptors. Type IV pili (T4P) have been previously identified as phage receptors for several *P. aeruginosa* phages, including PO4, PP7, Pf, F116, D3112, DMS3, and B3 (60-66). Phages may also use enzymes to break down capsule-like materials on the bacterial surface to reach the cell wall of the bacterium (155). By mutating or losing a phage receptor, bacteria become resistant to

phages. Resistance may reduce the fitness of the bacteria or, if the receptor is a virulence determinant, the pathogen may have reduced virulence (106, 156). In bacterial populations of 10^6–10^8 bacteria, there is a high possibility of spontaneous phage-resistant mutants deficient in the receptor or with an altered receptor. It is therefore imperative the phage cocktails be composed of phages that recognize a diversity of receptors to avoid the development of phage resistance populations.

It has been reported that T4P are an essential virulence factor in *X. fastidiosa* and that insertion mutations in genes involved in T4P biogenesis such as *pilB* and *pilQ* result in loss of twitching motility, inhibition of basipetal movement in planta, and the ability to cause disease (10). It was on interest to determine if T4P acted as a receptor site for the Xfas phages. Here it is described the experimental determination of the primary receptor site for Xfas103, Xfas106, Xfas303 and Xfas304.

Materials and methods

Bacterial strains and culture conditions

Bacterial strains and plasmids used in this study are listed in TABLE 9. *X. fastidiosa* strains were cultured at 28°C in PW-M broth (PW-MB) (25) or PW-M agar plates [PW-MA; PW-MB amended with 20 g/liter plant cell culture tested agar (Sigma)] (115). PW-M soft agar (PW-MSA, PW-MB amended with 7.5 g/liter plant cell culture tested agar) was used for overlays. PD3 amended agar (20 g/liter) (157) was used for preparation of electrocompetent cells and transformations. For *X. fastidiosa* cultures carrying antibiotic resistance cassettes or plasmids, the medium was supplemented with chloramphenicol (Cm) (5 μg/ml) or kanamycin (Km) (5 μg/ml). *Xanthomonas* strains

were cultured at 28°C in tryptone nutrient broth (TNB) (116) or tryptone nutrient agar (TNA; TNB lacking KNO_3 and amended with 20 g/liter agar). TNA soft agar (TNSA, TNB lacking KNO_3 and amended with 7.5 g/liter agar) was used for overlays. For *Xanthomonas* cultures harboring plasmids, the medium was supplemented with Km (50 µg/ml). Yeast tryptone broth (YTB: 10 g/liter tryptone, 10 g/liter yeast extract) and yeast-tryptone agar amended with sucrose [YTSA; YTB amended with 20 g/liter, sucrose final concentration 15% (wt/vol)] were used for resolution of mutants. *Escherichia coli* strains were cultured at 37°C in LB broth or LB agar (158). For *E. coli* harboring plasmids, medium was supplemented with Cm (50 µg/ml), Km (50 µg /ml) or ampicillin (Amp) (100 µg /ml), as appropriate.

Construction of strain Temecula pilA deletion mutant and in trans *complementation*

Deletion of *pilA* in strain Temecula was performed as described by Feil (163) and modified by Matsumoto (161), with additional modifications. The *pilA* (NP_780105.1) open reading frame (ORF) from 2,256,755 to 2,257,201 (*X. fastidiosa* Temecula1, NC_004556.1) was removed and replaced with a Km resistance cassette by site-directed gene replacement. All plasmid constructs were generated using *E. coli* DH5α MCR as

TABLE 9 Bacterial strains and plasmids used in this study

Strain or plasmid	Genotype and relevant features [a]	Reference or source
X. fastidiosa		
Temecula1	PD strain (subsp. *fastidiosa*) (ATCC 700964)	(12)
XF15-1-1	XF15-1, NS1::Cm^r *pilA*	This study
TM1	Temecula1; *PD1693*::*Tn5*	(4)
*tonB*1	Temecula; *PD0843*::*Tn5*	(159)
Xanthomonas		
EC-12	*Xanthomonas* sp., rice isolate	This study
EC-12-1	EC-12, unmarked deletion of *pilA*	This study
EC-12-1-1	EC-121, pMo168::*pilA*	This study
E. coli		
DH5αMCR	F$^-$ *Φ80lacZΔM15* Δ(*lacZYA-argF*) *recA1 endA1gyrA96 thi-1hsdR17supE44 relAl deoR U169*	Life Tech., Inc. (Gaithersburg, Md.)
Plasmids		
pUC19	pMB1 replicon, cloning vector, Ap^r	(160)
pSJA101	Isogenic to pUC19; Temecula *pilA*-downstream flanking region cloned into XmaI and XbaI site	This study
pSJA102	Isogenic to pSJA101; Km^r cassette cloned into KpnI and XmaI site	This study
pSJA103	Isogenic to pSJA102; Temecula *pilA*-upstream flanking region cloned into SacI and KpnI site	This study
pAX1Cm	Cm^r cassette and multiple-cloning site on pAX1 backbone	(161)
pSJA104	Isogenic to pAX1Cm; Temecula *pilA* cloned into XbaI and XhoI site	This study
pMo130	Suicide vector for allelic exchange; ColE1 ori, RK2 *oriT*, *xylE, sacB,* Km^r	(162)
pSJA105	Isogenic to pMo130; EC-12 *pilA*-upstream flanking region cloned into NheI and BglII site	This study
pSJA106	pSJA105 with EC-12 *pilA*-downstream flanking region cloned into BglII and HindIII site	This study
pMo168	Replicative vector; ori pBBR1, mob+, *xylE*, Km^r	(162)
pSJA107	pMo168::EC-12 *pilA*	This study

[a] Ap, ampicillin; Cm, chloramphenicol; Km, kanamyacin.

host. Two 1-kb fragments flanking upstream and downstream of *pilA* were amplified from strain Temecula genomic DNA, using primers with added restriction sites. The 1-kb fragment downstream of *pilA* was amplified using 5'ATAACACCCGGGTGAATACACACAGCAACAC G-3' (the XmaI site is underlined) and 5'-GTAATATCTAGAGCGCAGGGAAACGATGGAAGGT-3'(the XbaI site is underlined). The 1-kb PCR product was digested with XmaI and XbaI, and cloned into pUC19, resulting in pSJA101. The Km resistance gene cassette from pXF004 was amplified using 5'-GCGGCGGGTACCTCGATGAATTGTGTCTCAAAA-3' (the KpnI site is underlined) and 5'-AGCAGCCCCGGGTTAGAAAAACTCATCGAGCAT-3' (the XmaI site is underlined). The PCR product was digested with KpnI and XmaI and cloned into pSJA101, resulting in pSJA102. The 1-kb fragment upstream of *pilA* was amplified using 5'-ATAATAGAGCTCGAGAGCAACCGCGACAATAGCA-3' (the SacI site is underlined) and 5'- GCGGCGGGTACCGTGTATACCTTCAATAAAAAG-3' (the Kpn1 site is underlined). The 1-kb PCR product was digested with SacI and KpnI, and cloned into pSJA102, resulting in pSJA103. pSJA103 was introduced into strain Temecula by electroporation as described (161) and transformants were selected on PW-MA containing Km. The deletion of *pilA* and Km replacement was confirmed by sequencing from upstream and downstream of *pilA*.

Complementation of the strain Temecula *pilA* deletion mutant, XF15-1, was accomplished by introducing a wild-type copy of *pilA* using a chromosome-based complementation system as described by Matsumoto (161). Briefly, *pilA* with its

regulatory region was amplified from genomic DNA using 5'-GCGGCGTCTAGATTCTGGAAGAATGGATCG-3' (the XbaI site is underlined) and 5'- AGGTGTCTCGAGTTATTTAGCGGCAGTGCA-3' (the XhoI site is underlined). The 0.66 kb PCR product was digested with XbaI and XhoI, and cloned into pAX1Cm, resulting in pSJA104, and transformed into strain XF15-1 by electroporation. Transformants were selected on PW-MA-Cm (5 μg/ml), and individual colonies were streak purified three times on selective medium. Insertion of *pilA* and Cm^r cassette into the neutral site (NS1) of putative complements was confirmed by PCR amplification using primers pair NS1-f and NS1-r (161), and sequenced for confirmation (Eton Bioscience).

Construction of strain EC-12 pilA deletion mutant and in trans *complementation*

Deletion and complementation of *pilA* in strain EC-12 was performed as described by Hamad et al. (162) with modifications. PCR primers were designed based on the annotated *pilA* sequence of strain EC-12 draft genome (Das et al., unpublished). All plasmid constructs were generated using *E. coli* DH5α MCR as host. A 1.0 kb fragment upstream of *pilA* was amplified using 5'-GCGGCGGCTAGCTTTGAAGACATCTTCGAATTG-3' (the NheI site is underlined) and 5'-ATGGCCAGATCTAATGTA CCCCTAGAGATGGTGT-3' (the BglII site is underlined). The PCR product was digested with NheI and BglII, and cloned into multiple cloning site (MCS) 1 of plasmid pMo130, to obtain plasmid pSJA105. A 1.0 kb fragment downstream of *pilA* was amplified using 5'-GCGCGCAGATCTGATTCGGTTGGATGGTTTTGAT-3' (the BglII site is underlined)

and 5'-ATAAATAAGCTTTTGACCGGCGCCTTGGCCACGC-3' (the HindIII site is underlined). The PCR product was digested with BglII and HindIII and cloned into pSJA105, to obtain pSJA106. Plasmid pSJA106 was introduced into electrocompetent EC-12 cells by electroporation as previously described (163). Cells were allowed to recover in TNB for 3 h at 28^{o}C before being dilution plated on TNA Km (50 µg/ml). After 72 h, resulting colonies were exposed to a mist of 0.45 M pyrocatechol to identify colonies in which single crossover events had occurred. Single yellow colonies were grown in YTB for 9 h and then plated to YTSA to select for resolved co-integrants. Colonies grown on YTSA were sprayed with pyrocatechol, and presumptive resolved co-integrants (exhibiting a white phenotype) were confirmed by sequencing across *pilA* upstream and downstream, to confirm *pilA* deletion.

To complement EC-12 *pilA* deletion mutant EC-12-1, wild-type *pilA* was introduced using a derivative of the replicative plasmid pMo168. *PilA,* with its regulatory region, was amplified from EC-12 genomic DNA using 5'-GCGGCGGCTAGCCTTATCCGTCAACACCAT-3' (the NheI site is underlined) and 5'- GTGAATAAGCTTTTACTGACAGGTCGCCGGGGC-3' (the HindIII site is underlined). The 0.47 kb PCR product was digested with NheI and HindIII and cloned into pMo168 MCS-1, resulting in pSJA107. Recombinant plasmid pSJA107 was transformed into EC-12-1 as described above, and plated on TNA Km 50 µg/ml for selection at 28°C. Transformants were sprayed with 0.45 M pyrocatechol to identify colonies containing pSJA107, and the presence of the wild-type *pilA* was confirmed by

PCR amplification using different combinations of primers. All PCR products were sequenced for confirmation.

Twitching motility assay

To assay twitching motility, late log phase liquid cultures of strain Temecula (5 day), EC-12 (18h) and their derivatives were centrifuged and resuspended to 200 µl in liquid medium as appropriate. Five microliter cell suspensions (1×10^9 CFU/ml) were spotted to appropriate solid medium modified with 1.2% agar, and incubated at 28˚C [Temecula and derivatives (5 day) and strain EC-12 and derivatives (18h)]. *Pseudomonas aeruginosa* strain PAO1 grown in a similar manner at 37˚C for 24h, was used as positive control. Following incubation, colony edge morphology was visualized using an SZ-PT Olympus Microscope equipped with QImaging Go-21 CMOS camera system.

Results and discussion

Our preliminary results demonstrated that two transposon insertion mutants of strain Temecula, TM1 (unpiliated) (159) and "*tonB1* mutant" (hyperpiliated) (4), were resistant to phages Xfas103, Xfas106, Xfas303, and Xfas304, indicating the four phages may be T4P dependent. To confirm these observations, we generated T4P mutants (in-frame deletions of *pilA*) in strains Temecula and EC-12, and tested for phage sensitivity with homologous propagated phage of each of the four phages, individually. PilA is the major subunit of T4P, and is required for T4P formation (164). Wild type Temecula and EC-12, both exhibited twitching motility (indicative of pilus retraction) (FIG 5 A,D respectively) and were sensitive to all four phages. Δ*pilA* derivatives, XF15-1 and EC-

12-1, were devoid in twitching motility (FIG 5B,E respectively) and were resistant to all four phages. Twitching motility and phage sensitivity were restored in XF15-1and EC-12-1 by complementation of *pilA* in *trans* (FIG 5C,F respectively), providing evidence that the four phages utilize T4P as primary receptors to infect strains Temecula and EC-12. However, it is not clear which component of the T4P is recognized by each phage. There is evidence that PilA is not the phage receptor in *P. aeruginosa.* Wild type *P. aeruginosa* strain PAK is sensitive to PO4 and resistant to phi6, whereas *P. syringae* strain DC3000 is sensitive to phi6 and resistant to PO4. The expression of *P. syringae pilA* in *P. aeruginosa* PAK has been shown to restore sensitivity to PO4 and confer sensitivity to phi6, indicating that a minor pilus-associated protein may indeed be the receptor site (165).

It is not clear if the described phages replicate on *Xanthomonas* and/or *X. fastidiosa* in the environment. The genomes of Xfas103, Xfas106, Xfas303 and Xfas304 have an average GC content of 63.5%, that is similar to that of sequenced *Xanthomonas* spp. (63.6-65.3%) and *Xanthomonas* phages (60%) (166); but higher than that of *X. fastidiosa* strains (49-55.6%) (167), suggesting the phages evolved in *Xanthomonas*. However, allopatric speciation would suggest if the phages had separated from one host for too long, they would lose the ability to adsorb and replicate in that host. Both genera are members of the *Xanthomonadaceae* and are associated with a broad range of plant hosts, where they do not necessarily cause disease (6, 168). The *Xanthomonas* and *X. fastidiosa* genomes are diverged, yet both have retained T4P, a required virulence factor in both *X. fastidiosa* (3) and *Xanthomonas* (169-172). Phylogenetic analysis suggests

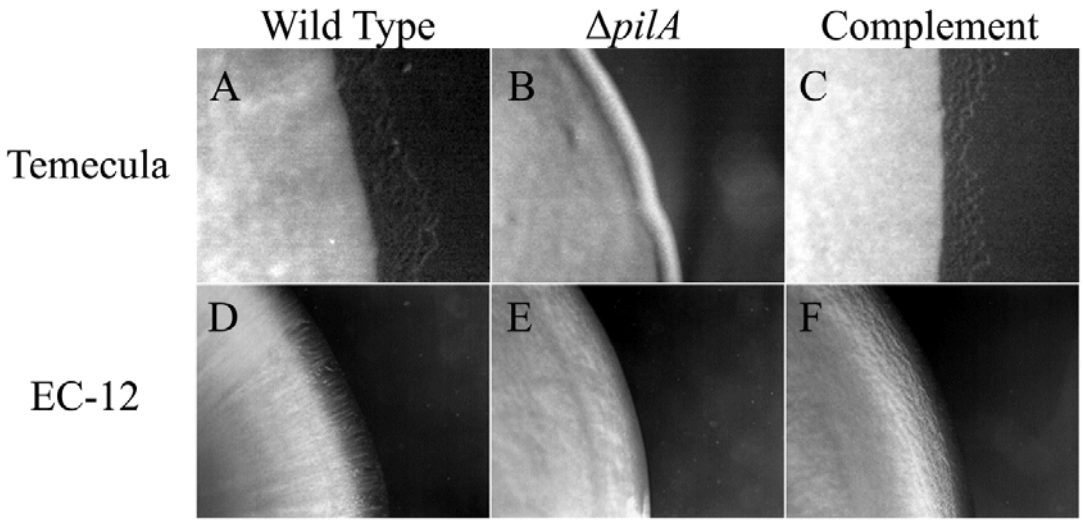

FIG 5 Bacterial colony edge morphology and twitching motility. Colony edge morphology and twitching motility for (A) wild-type strain Temecula, (B) strain Temecula Δp*ilA* mutant, (C) strain Temecula *ΔpilA* complement, (D) wild-type *Xanthomonas* strain EC-12, (E) strain EC-12 Δ*pilA* mutant and (F) strain EC-12 Δ*pilA* complement. Cultures of *X. fastidiosa* or *Xanthomonas* were spotted to appropriate solid medium modified with 1.2% agar, and incubated at 28˚C [*X. fastidiosa* (5 day) and *Xanthomonas* (18h)]. Wild type isolates exhibit twitching motility, Δp*ilA* mutants do not exhibit twitching motility, and complements are restored.

there is divergence in the T4P *pilE-fimT* cluster within the *Xanthomonadaceae* (6). *X. fastidiosa* contains two copies of the cluster, whereas *X. axonopodis pv. citri* has lost one copy and *X. campestris* pv. *campestris* has lost the other (6). We speculate that overlapping plant host ranges and shared environmental niches, along with the conservation of T4P in both *Xanthomonas* and *X. fastidiosa,* has led to the selection of phages able to adsorb and replicate in both genera.

CHAPTER V

CONCLUSIONS

In medicine and agriculture, the search for innovative solutions to control disease is an ongoing process. Currently, there are no economic or environmentally sustainable solutions for control of disease caused by *X. fastidiosa*. The wide host range of this pathogen has implications for economically important crops such as grape, citrus and coffee. Phage therapy offers a novel and environmentally acceptable treatment for the control of disease caused by *X. fastidiosa*.

In this study, I have developed a surrogate system using *Xanthomonas* that allows for the efficient isolation and propagation of phages from the environment for *X. fastidiosa*. Prior to this study, the only propagated phage was temperate phage Xfas53 isolated from supernatants of *X. fastidiosa* strain XF53. Phages Xfas53 which appears to have been formed by recombination between a widespread family of *X. fastidiosa* P2-related prophage elements and a podophage distantly related to phage P22 (115).

As noted in the prior text, temperate phage are not suitable for a phage based biocontrol system for control of disease caused by *X. fastidiosa*. During this study, I isolated and characterized the first virulent phages for *X. fastidiosa* , Xfas103, Xas106, Xfas303 and Xfas304, with a host range that extended to several species of *Xanthomonas*, including *X. axonopodis* pv. *citri* and *X. euvesicatoria.*

Adsorption rate constants determined for the four phages strain showed that as with phage Xfas53, the four phages exhibited unusually slow adsorption kinetics, possibly an adaptation to the confined niche of their slow-growing host.

I proposed that phage BcepNazgul is the founder of a novel phage type based on its unique genomic organization as compared to existing phage types, and that phages Xfas103 and Xfas106 belong to this new group. The genomes of this new group consisted of four divergent transcriptional units based on coding strand position, with genes found in transcriptional unit one and two primarily encoding proteins of unknown function, and genes found in transcriptional units three and four encoding the lysis cassette, structural module and DNA metabolism modules. I further proposed, based on genomic analysis that podophages Xfas303 and Xas304 are members of the virulent phiKMV-like group.

I determined the shared receptor site of all four phages by constructing deletion mutants in the *pilA* of *X. fastidiosa* strain Temecula and *Xanthomonas* strain EC-12. Mutants were impaired in twitching motility and were resistant to all four phages, whereas the complements were restored to wild-type phenotype and were phage sensitive. This confirmed that the Type IV pili are receptor for all four phages in both hosts.

The development of the surrogate system and the availability to isolate and propagate virulent phages for *X. fastidiosa* now provides the opportunity for the development of a phage based biocontrol system for disease caused by *X. fastidiosa*. Phage therapy offers compelling advantages for antibacterial therapy because of two

properties: specificity and exponential propagation. Phages specifically target the pathogen, without posing any harmful effects to humans, animals, and plants or associated beneficial microflora.

Future studies will include the testing of the efficacy of a phage based biocontrol system for control of Pierce's Disease in grape and determining the movement and persistence of phage in plants.

REFERENCES

1. **De La Fuente L, Montanes E, Meng Y, Li Y, Burr TJ, Hoch HC, Wu M.** 2007. Assessing adhesion forces of type I and type IV pili of *Xylella fastidiosa* bacteria by use of a microfluidic flow chamber. Appl Environ Microbiol **73:**2690-2696.

2. **Hopkins DL, Mollenha Hh.** 1973. *Rickettsia*-like bacterium associated with Pierces Disease of grapes. Science **179:**298-300.

3. **Mhedbi-Hajri N, Jacques MA, Koebnik R.** 2011. Adhesion mechanisms of plant-pathogenic *Xanthomonadaceae*. Adv Exp Med Biol **715:**71-89.

4. **Li Y, Hao G, Galvani CD, Meng Y, De La Fuente L, Hoch HC, Burr TJ.** 2007. Type I and type IV pili of *Xylella fastidiosa* affect twitching motility, biofilm formation and cell-cell aggregation. Microbiology **153:**719-726.

5. **Hopkins DL, Purcell AH.** 2002. *Xylella fastidiosa*: Cause of Pierce's Disease of grapevine and other emergent diseases. Plant Dis **86:**1056-1066.

6. **Moreira LM, De Souza RF, Digiampietri LA, Da Silva ACR, Setubal JC.** 2005. Comparative analyses of *Xanthomonas* and *Xylella* complete genomes. Omics **9:**43-76.

7. **Parkinson N, Aritua V, Heeney J, Cowie C, Bew J, Stead D.** 2007. Phylogenetic analysis of *Xanthomonas* species by comparison of partial gyrase B gene sequences. Int J Syst Evol Micr **57:**2881-2887.

8. **Beattie GA, Lindow SE.** 1995. The secret life of foliar bacterial pathogens on neaves. Annu Rev Phytopathol **33:**145-172.

9. **Ryan RP, Vorholter FJ, Potnis N, Jones JB, Van Sluys MA, Bogdanove AJ, Dow JM.** 2011. Pathogenomics of *Xanthomonas*: understanding bacterium-plant interactions. Nat Rev Microbiol **9:**344-355.

10. **Chatterjee S, Almeida RPP, Lindow S.** 2008. Living in two worlds: the plant and insect lifestyles of *Xylella fastidiosa*. Annu Rev Phytopathol **46:**243-271.

11. **Keen NT, Korsi-Dumenyo C, Yang C, Cooksey DA.** 2000. From rags to riches: insights from the first genomic sequence of a plant pathogenic bacterium. Genome Biol **1:**1019.

12. **Van Sluys MA, de Oliveira MC, Monteiro-Vitorello CB, Miyaki CY, Furlan LR, Camargo LE, da Silva AC, Moon DH, Takita MA, Lemos EG, Machado**

MA, Ferro MI, da Silva FR, Goldman MH, Goldman GH, Lemos MV, El-Dorry H, Tsai SM, Carrer H, Carraro DM, de Oliveira RC, Nunes LR, Siqueira WJ, Coutinho LL, Kimura ET, Ferro ES, Harakava R, Kuramae EE, Marino CL, Giglioti E, Abreu IL, Alves LM, do Amaral AM, Baia GS, Blanco SR, Brito MS, Cannavan FS, Celestino AV, da Cunha AF, Fenille RC, Ferro JA, Formighieri EF, Kishi LT, Leoni SG, Oliveira AR, Rosa VE, Sassaki FT, Sena JA, de Souza AA, Truffi D, Tsukumo F, Yanai GM, Zaros LG, Civerolo EL, Simpson AJ, Almeida NF, Setubal JC, Kitajima JP. 2003. Comparative analyses of the complete genome sequences of Pierce's Disease and citrus variegated chlorosis strains of *Xylella fastidiosa*. J Bacteriol **185:**1018-1026.

13. **Simpson AJG, Reinach FC, Arruda P, Abreu FA, Acencio M, Alvarenga R, Alves LMC, Araya JE, Baia GS, Baptista CS, Barros MH, Bonaccorsi ED, Bordin S, Bove JM, Briones MRS, Bueno MRP, Camargo AA, Camargo LEA, Carraro DM, Carrer H, Colauto NB, Colombo C, Costa FF, Costa MCR, Costa-Neto CM, Coutinho LL, Cristofani M, Dias-Neto E, Docena C, El-Dorry H, Facincani AP, Ferreira AJS, Ferreira VCA, Ferro JA, Fraga JS, Franca SC, Franco MC, Frohme M, Furlan LR, Garnier M, Goldman GH, Goldman MHS, Gomes SL, Gruber A, Ho PL, Hoheisel JD, Junqueira ML, Kemper EL, Kitajima JP, Krieger JE, Kuramae EE, Laigret F, Lambais MR, Leite LCC, Lemos EGM, Lemos MVF, Lopes SA, Lopes CR, Machado JA, Machado MA, Madeira AMBN, Madeira HMF, Marino CL, Marques MV, Martins EAL, Martins EMF, Matsukuma AY, Menck CFM, Miracca EC, Miyaki CY, Monteiro-Vitorello CB, Moon DH, Nagai MA, Nascimento ALTO, Netto LES, Nhani A, Nobrega FG, Nunes LR, Oliveira MA, de Oliveira MC, de Oliveira RC, Palmieri DA, Paris A, Peixoto BR, Pereira GAG, Pereira HA, Pesquero JB, Quaggio RB, Roberto PG, Rodrigues V, Rosa AJD, de Rosa VE, de Sa RG, Santelli RV, Sawasaki HE, da Silva ACR, da Silva AM, da Silva FR, Silva WA, da Silveira JF, Silvestri MLZ, Siqueira WJ, de Souza AA, de Souza AP, Terenzi MF, Truffi D, Tsai SM, Tsuhako MH, Vallada H, Van Sluys MA, Verjovski-Almeida S, Vettore AL, Zago MA, Zatz M, Meidanis J, Setubal JC, Org XfC.** 2000. The genome sequence of the plant pathogen *Xylella fastidiosa*. Nature **406:**151-157.

14. **da Silva FR, Vettore AL, Kemper EL, Leite A, Arruda P.** 2001. Fastidian gum: the *Xylella fastidiosa* exopolysaccharide possibly involved in bacterial pathogenicity. FEMS Microbiol Lett **203:**165-171.

15. **Purcell AH, Hopkins DL.** 1996. Fastidious xylem-limited bacterial plant pathogens. Annu Rev Phytopathol **34:**131-151.

16. **Redak RA, Purcell AH, Lopes JRS, Blua MJ, Mizell RF, Andersen PC.** 2004. The biology of xylem fluid-feeding insect vectors of *Xylella fastidiosa* and their relation to disease epidemiology. Annu Rev Entomol **49:**243-270.

17. **Purcell AH.** 1980. Almond Leaf Scorch - leafhopper (Homoptera, *Cicadellidae*) and spittlebug (Homoptera, *Cercopidae*) vectors. Journal of Economic Entomology **73:**834-838.

18. **Hill BL, Purcell AH.** 1997. Populations of *Xylella fastidiosa* in plants required for transmission by an efficient vector. Phytopathol **87:**1197-1201.

19. **Hill BL, Purcell AH.** 1995. Acquisition and retention of *Xylella fastidiosa* by an efficient vector, *Graphocephala atropunctata*. Phytopathol **85:**209-212.

20. **Barnard EL, Ash EC, Hopkins DL, McGovern RJ.** 1998. Distribution of *Xylella fastidiosa* in oaks in Florida and its association with growth decline in *Quercus laevis*. Plant Dis **82:**569-572.

21. **Costa HS, Raetz E, Pinckard TR, Gispert C, Hernandez-Martinez R, Dumenyo CK, Cooksey DA.** 2004. Plant hosts of *Xylella fastidiosa* in and near southern California vineyards. Plant Dis **88:**1255-1261.

22. **Henneberger TSM, Stevenson KL, Britton KO, Chang CJ.** 2004. Distribution of *Xylella fastidiosa* in sycamore associated with low temperature and host resistance. Plant Dis **88:**951-958

.

23. **Hernandez-Martinez R, de la Cerda KA, Costa HS, Cooksey DA, Wong FP.** 2007. Phylogenetic relationships of *Xylella fastidiosa* strains isolated from landscape ornamentals in southern California. Phytopathol **97:**857-864.

24. **Hernandez-Martinez R, Pinckard TR, Costa HS, Cooksey DA, Wong FP.** 2006. Discovery and characterization of *Xylella fastidiosa* strains in southern California causing mulberry leaf scorch. Plant Dis **90:**1143-1149.

25. **Hill BL, Purcell AH.** 1995. Multiplication and movement of *Xylella fastidiosa* within grapevine and 4 other plants. Phytopathol **85:**1368-1372.

26. **Hopkins DL.** 1989. *Xylella fastidiosa* - xylem-limited bacterial pathogen of plants. Annu Rev Phytopathol **27:**271-290.

27. **McGaha LA, Jackson B, Bextine B, McCullough D, Morano L.** 2007. Potential plant reservoirs for *Xylella fastidiosa* in south Texas. Am J Enol Viticult **58:**398-401.

28. **Montero-Astua M, Saborio-R G, Chacon-Diaz C, Garita L, Villalobos W, Hartung JS, Rivera C.** 2008. First report of *Xylella fastidiosa* in avocado in Costa Rica. Plant Dis **92:**175-175.

29. **Wistrom C, Purcell AH.** 2005. The fate of *Xylella fastidiosa* in vineyard weeds and other alternate hosts in California. Plant Dis **89:**994-999.

30. **Schaad NW, Postnikova E, Lacy G, Fatmi M, Chang CJ.** 2004. *Xylella fastidiosa* subspecies: *X. fastidiosa* subsp *piercei*, subsp. nov., *X. fastidiosa* subsp. *multiplex* subsp. nov., and *X. fastidiosa* subsp. *pauca* subsp. nov. Syst Appl Microbiol **27:**763-763.

31. **Randall JJ, Goldberg NP, Kemp JD, Radionenko M, French JM, Olsen MW, Hanson SF.** 2009. Genetic analysis of a novel *Xylella fastidiosa* subspecies found in the southwestern United States. Appl Environ Microb **75:**5631-5638.

32. **Leu LS, Su CC.** 1993. Isolation, cultivation, and pathogenicity of *Xylella fastidiosa*, the causal bacterium of Pear Leaf Scorch Disease in Taiwan. Plant Dis **77:**642-646.

33. **Berisha B, Chen YD, Zhang GY, Xu BY, Chen TA.** 1998. Isolation of Peirce's Disease bacteria from grapevines in Europe. Eur J Plant Pathol **104:**427-433.

34. **Tyree MT, Zimmermann MH.** 2002. Xylem structure and the ascent of sap, 2nd ed. Springer, New York, NY.

35. **Newman KL, Almeida RPP, Purcell AH, Lindow SE.** 2004. Cell-cell signaling controls *Xylella fastidiosa* interactions with both insects and plants. P Natl Acad Sci USA **101:**1737-1742.

36. **Krivanek AF, Stevenson JF, Walker MA.** 2005. Development and comparison of symptom indices for quantifying grapevine resistance to Pierce's Disease. Phytopathol **95:**36-43.

37. **Janse JD, Obradovic A.** 2010. *Xylella Fastidiosa*: its biology, diagnosis, control and risks. J Plant Pathol **92:**S35-S48.

38. **Rossetti V, Garnier M, Bove JM, Beretta MJG, Teixeira ARR, Quaggio JA, Denegri JD.** 1990. Occurrence of xylem-restricted bacteria in sweet orange trees affected by chlorotic variegation, a new citrus disease in Brazil. Cr Acad Sci Iii-Vie **310:**345-349.

39. **Qin XT, Miranda VS, Machado MA, Lemos EGM, Hartung JS.** 2001. An evaluation of the genetic diversity of *Xylella fastidiosa* isolated from diseased citrus and coffee in Sao Paulo, Brazil. Phytopathol **91:**599-605.

40. **Almeida RPP, Blua MJ, Lopes JRS, Purcell AH.** 2005. Vector transmission of *Xylella fastidiosa*: applying fundamental knowledge to generate disease management strategies. Ann Entomol Soc Am **98:**775-786.

41. **Stokstad E.** 2012. Agriculture field research on bees raises concern about low-dose pesticides. Science **335:**1555-1555.

42. **Whitehorn PR, O'Connor S, Wackers FL, Goulson D.** 2012. Neonicotinoid pesticide reduces bumble bee colony growth and queen production. Science **336:**351-352.

43. **Jones JB, Jackson LE, Balogh B, Obradovic A, Iriarte FB, Momol MT.** 2007. Bacteriophages for plant disease control. Annu Rev Phytopathol **45:**245-262.

44. **Rohwer F, Edwards R.** 2002. The phage proteomic tree: a genome-based taxonomy for phage. J Bacteriol **184:**4529-4535.

45. **Hagens S.** 2012. Bacteriophage applications: where are we now? J Vet Pharmacol Ther **35:**67-67.

46. **Abedon ST.** 2006. Phage ecology, p. 37-46. *In* Calendar R. (ed.), The bacteriophages, 2nd ed. Oxford University Press, New York, NY.

47. **Hershey AD, Chase M.** 1952. Independent functions of viral protein and nucleic acid in growth of bacteriophage. J Gen Physiol **36:**39-56.

48. **Crick FH, Barnett L, Brenner S, Watts-Tobin RJ.** 1961. General nature of the genetic code for proteins. Nature **192:**1227-1232.

49. **Summers WC.** 2006. Phage and the early development of molecular biology, p. 3-7. *In* Calendar R. (ed.), The bacteriophages, 2nd ed. Oxford University Press; New York, NY.

50. **Ackermann HW.** 2001. Frequency of morphological phage descriptions in the year 2000. Arch Virol **146:**843-857.

51. **Lawrence JG, Hatfull GF, Hendrix RW.** 2002. Imbroglios of viral taxonomy: genetic exchange and failings of phenetic approaches. J Bacteriol **184:**4891-4905.

52. **Hendrix RW, Smith MCM, Burns RN, Ford ME, Hatfull GF.** 1999. Evolutionary relationships among diverse bacteriophages and prophages: all the world's a phage. P Natl Acad Sci USA **96:**2192-2197.

53. **Hatfull GF, Jacobs-Sera D, Lawrence JG, Pope WH, Russell DA, Ko CC, Weber RJ, Patel MC, Germane KL, Edgar RH, Hoyte NN, Bowman CA, Tantoco AT, Paladin EC, Myers MS, Smith AL, Grace MS, Pham TT, O'Brien MB, Vogelsberger AM, Hryckowian AJ, Wynalek JL, Donis-Keller H, Bogel MW, Peebles CL, Cresawn SG, Hendrix RW.** 2010. Comparative genomic analysis of 60 mycobacteriophage genomes: genome clustering, gene acquisition, and gene size. Journal of Molecular Biology **397:**119-143.

54. **Lavigne R, Darius P, Summer EJ, Seto D, Mahadevan P, Nilsson AS, Ackermann HW, Kropinski AM.** 2009. Classification of myoviridae bacteriophages using protein sequence similarity. Bmc Microbiol **9:**224

55. **Lavigne R, Seto D, Mahadevan P, Ackermann HW, Kropinski AM.** 2008. Unifying classical and molecular taxonomic classification: analysis of the Podoviridae using BLASTP-based tools. Res Microbiol **159:**406-414.

56. **Gill JJ, Berry JD, Russell WK, Lessor L, Escobar-Garcia DA, Hernandez D, Kane A, Keene J, Maddox M, Martin R, Mohan S, Thorn AM, Russell DH, Young R.** 2012. The *Caulobacter crescentus* phage phiCbK: genomics of a canonical phage. Bmc Genomics **13:**542.

57. **Casjens SR.** 2008. Diversity among the tailed-bacteriophages that infect the *Enterobacteriaceae*. Res Microbiol **159:**340-348.

58. **Gill JJ, Abedon ST.** 2003. Bacteriophage ecology and plants. APSnet Feature http://www.apsnet.org/online/feature/phages/

59. **Lindberg AA.** 1973. Bacteriophage receptors. Annu Rev Microbiol **27:**205-241.

60. **Bradley DE.** 1973. Basic characterization of a *Pseudomonas aeruginosa* pilus-dependent bacteriophage with a long noncontractile tail. J Virol **12:**1139-1148.

61. **Bradley DE, Robertson D.** 1968. The structure and infective process of a contractile *Pseudomonas aeruginosa* bacteriophage. J Gen Virol **3:**247-254.

62. **Bradley DE.** 1973. The adsorption of the *Pseudomonas aeruginosa* filamentous bacteriophage Pf to its host. Can J Microbiol **19:**623-631.

63. **Byrne M, Kropinski AM.** 2005. The genome of the *Pseudomonas aeruginosa* generalized transducing bacteriophage F116. Gene **346:**187-194.

64. **Wang PW, Chu L, Guttman DS.** 2004. Complete sequence and evolutionary genomic analysis of the *Pseudomonas aeruginosa* transposable bacteriophage D3112. J Bacteriol **186:**400-410.

65. **Budzik JM, Rosche WA, Rietsch A, O'Toole GA.** 2004. Isolation and characterization of a generalized transducing phage for *Pseudomonas aeruginosa* strains PAO1 and PA14. J Bacteriol **186:**3270-3273.

66. **Roncero C, Darzins A, Casadaban MJ.** 1990. *Pseudomonas aeruginosa* transposable bacteriophages D3112 and B3 require pili and surface growth for adsorption. J Bacteriol **172:**1899-1904.

67. **Wang IN, Young R.** 2006. Phage lysis, p. 104-125. *In* Calendar R. (ed.), The bacteriophages, 2nd ed. Oxford University Press; New York, NY.

68. **Wang IN, Smith DL, Young R.** 2000. Holins: the protein clocks of bacteriophage infections. Annu Rev Microbiol **54:**799-825.

69. **Young I, Wang I, Roof WD.** 2000. Phages will out: strategies of host cell lysis. Trends Microbiol **8:**120-128.

70. **Berry J, Savva C, Holzenburg A, Young R.** 2010. The lambda spanin components Rz and Rz1 undergo tertiary and quaternary rearrangements upon complex formation. Protein Sci **19:**1967-1977.

71. **Summers WC.** 2001. Bacteriophage therapy. Annu Rev Microbiol **55:**437-451.

72. **Balogh B, Jones JB, Iriarte FB, Momol MT.** 2010. Phage therapy for plant disease control. Curr Pharm Biotechno **11:**48-57.

73. **Okabe N, Goto M.** 1963. Bacteriophages of plant pathogens. Ann. Rev. Phytopathol. **1:**397-418.

74. **Balogh B.** 2002. Strategies of improving the efficacy of bacteriophages for controlling bacterial spot of tomato. University of Florida, Gainesville, FL.

75. **Balogh B, Jones JB, Momol MT, Olson SM, Obradovic A, King P, Jackson LE.** 2003. Improved efficacy of newly formulated bacteriophages for management of bacterial spot on tomato. Plant Dis **87:**949-954.

76. **Civerolo EL, Keil HL.** 1969. Inhibition of bacterial ppot of peach foliage by *Xanthomonas pruni* bacteriophage. Phytopathol **59:**196-198.

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

78. **Iriarte FB, Balogh B, Momol MT, Smith LM, Wilson M, Jones JB.** 2007. Factors affecting survival of bacteriophage on tomato leaf surfaces. Appl Environ Microb **73:**1704-1711.

79. **Bergamin FA.** 1981. Studies on a bacteriophage isolated from *Xanthomonas campestris.* Summa Phytopathol **7:**35-43.

80. **Ravensdale M, Blom TJ, Gracia-Garza JA, Svircev AM, Smith RJ.** 2007. Bacteriophages and the control of *Erwinia carotovora* subsp *carotovora*. Can J Plant Pathol **29:**121-130.

81. **Balogh B, Canteros BI, Stall KE, Jones JB.** 2008. Control of citrus canker and citrus bacterial spot with bacteriophages. Plant Dis **92:**1048-1052.

82. **Flaherty JE, Harbaugh BK, Jones JB, Somodi GC, Jackson LE.** 2001. H-mutant bacteriophages as a potential biocontrol of bacterial blight of geranium. Hortscience **36:**98-100.

83. **Borah PK, Jindal JK, Verma JP.** 2000. Integrated management of bacterial leaf spot of mungbean with bacteriophages of Xav and chemicals. J. Mycol. Plant Pathol. **30:**19-21.

84. **Munsch P, Olivier JM.** 1995. Biocontrol of bacterial blotch of the cultivated mushroom with lytic phages: some practical considerations. Science and Cultivation of Edible Fungi. **2:**595-602.

85. **Munsch P, Olivier JM, Houdeau G.** 1991. Experimental control of bacterial blotch by bacteriophages. Science and Cultivation of Edible Fungi. **2:**389-396.

86. **Lang JM, Gent DH, Schwartz HF.** 2007. Management of *Xanthomonas* leaf blight of onion with bacteriophages and a plant activator. Plant Dis **91:**871-878.

87. **Gent DH, Schwartz HF.** 2005. Management of *Xanthomonas* leaf blight of onion with a plant activator, biological control agents, and copper bactericides. Plant Dis **89:**631-639.

88. **Lang JM, Schwartz HF, Gent DH.** 2004. Alternative strategies for onion *Xanthomonas* leaf blight management. Phytopathol **94:**S56-S56.

89. **Svircev AM, Lehman SM, Kim W, Barszcz E, Schneider KE, Castle AJ.** 2006. Proc. Int. Symp. Biol. Control Bact. Plant Dis, Seeheim/Darmstadt, Germany, 259. Berlin, Germany.

90. **Gill JJ, Svircev AM, Smith R, Castle AJ.** 2003. Bacteriophages of *Erwinia amylovora*. Appl Environ Microb **69:**2133-2138.

91. **Schnabel EL, Jones AL.** 2001. Isolation and characterization of five *Erwinia amylovora* bacteriophages and assessment of phage resistance in strains of *Erwinia amylovora*. Appl Environ Microb **67:**59-64.

92. **Schnabel EL, Fernando WGD, Meyer MP, Jones AL, Jackson LE.** 1999. Bacteriophage of *Erwinia amylovora* and their potential for biocontrol. Proceedings of the Eight International Workshop on Fire Blight **489:**649-653.

93. **Erskine JM.** 1973. Characteristics of *Erwinia amylovora* bacteriophage and its possible role in epidemiology of Fire Blight. Can J Microbiol **19:**837.

94. **Baldwin, CH, Goodman RN.** 1963. Prevalence of *Erwinia amylovora* in apple buds as detected by phage typing. Phytopathol **53:**1299–1303.

95. **McKenna F, El-Tarabily KA, Hardy GES, Dell B.** 2001. Novel in vivo use of a polyvalent *Streptomyces* phage to disinfest *Streptomyces* scabies-infected seed potatoes. Plant Pathol **50:**666-675.

96. **Kuo T, Cheng L, Yang C, Yang S.** Bacterial leaf blight of rice plant IV. Effect of bacteriophage on the infectivity of *Xanthomonas oryzae*. Bot. Bull. Acad. Sinica **12:**54.

97. **Zaccardelli M, Saccardi A, Gambin E, Minardi P, Mazzucchi U.** 1994. *Xanthomonas campestris* pv *pruni* bacteriophages on peach trees and their potential use for biological-control. Colloq Inra **66:**875-878.

98. **Civerolo EL.** 1973. Relationship of *Xanthomonas pruni* bacteriophages to bacterial spot disease in prunus. Phytopathol **63:**79-84.

99. **Randhawa PS, Civerolo EL.** 1986. Interaction of *Xanthomonas campestris* pv *pruni* with pruniphage and epiphytic bacteria on detached peach leaves. Phytopathol **76:**549-553.

100. **Wall GC, Sanchez JL.** 1993. A biocontrol agent for *Pseudomonas solanacearum*. Aciar Proc **45:**320-321.

101. **Tanaka H, Negishi H, Maeda H.** 1990. Control of tobacco bacterial wilt by an avirulent strain of *Pseudomonas solanacearunm* M4S and its bacteriophage. Ann Phytopathol. Soc. Jpn. **56:**243-246.

102. **Obradovic A, Jones JB, Momol MT, Balogh B, Olson SM.** 2004. Management of tomato bacterial spot in the field by foliar applications of bacteriophages and SAR inducers. Plant Dis **88:**736-740.

103. **Obradovic A, Jones JB, Momol MT, Olson SM, Jackson LE, Balogh B, Guven K, Iriarte FB.** 2005. Integration of biological control agents and systemic acquired resistance inducers against bacterial spot on tomato. Plant Dis **89:**712-716.

104. **Flaherty JE, Jones JB, Harbaugh BK, Somodi GC, Jackson LE.** 2000. Control of bacterial spot on tomato in the greenhouse and field with H-mutant bacteriophages. Hortscience **35:**882-884.

105. **Boyd RJ, Hildebra.Ac, Allen ON.** 1971. Retardation of Crown Gall Enlargement after Bacteriophage Treatment. Plant Dis Rep **55:**145.

106. **Kutter E, Sulakvelidze A.** 2005. Bacteriophages: biology and applications. CRC Press, Boca Raton, FL.

107. **Miedzybrodzki R, Fortuna W, Weber-Dabrowska B, Gorski A.** 2005. Bacterial viruses against viruses pathogenic for man? Virus Res **110:**1-8.

108. **David HL, Clavel S, Clement F.** 1980. Adsorption and growth of the bacteriophage-D29 in selected mycobacteria. Ann Inst Pasteur Vir **131:**167.

109. **Carlton RM, Noordman WH, Biswas B, de Meester ED, Loessner MJ.** 2005. Bacteriophage P100 for control of *Listeria monocytogenes* in foods: genome sequence, bioinformatic analyses, oral toxicity study, and application. Regul Toxicol Pharm **43:**301-312.

110. **Dupuis ME, Villion M, Magadan AH, Moineau S.** 2013. CRISPR-Cas and restriction-modification systems are compatible and increase phage resistance. Nat Commun **4:**2087.

111. **Kutter E, De Vos D, Gvasalia G, Alavidze Z, Gogokhia L, Kuhl S, Abedon ST.** 2010. Phage therapy in clinical practice: treatment of human infections. Curr Pharm Biotechno **11:**69-86.

112. **Hagens S, Loessner MJ.** 2010. Bacteriophage for biocontrol of foodborne pathogens: calculations and considerations. Curr Pharm Biotechno **11:**58-68.

113. **Abedon ST, Thomas-Abedon C.** 2010. Phage therapy pharmacology. Curr Pharm Biotechno **11:**28-47.

114. **Gill JJ, Hyman P.** 2010. Phage choice, isolation, and preparation for phage therapy. Curr Pharm Biotechno **11:**2-14.

115. **Summer EJ, Enderle CJ, Ahern SJ, Gill JJ, Torres CP, Appel DN, Black MC, Young R, Gonzalez CF.** 2010. Genomic and biological analysis of phage Xfas53 and related prophages of *Xylella fastidiosa*. J Bacteriol **192:**179-190.

116. **Hansen JB, Olsen RH.** 1978. Isolation of large bacterial plasmids and characterization of P2 I compatibility group plasmids Pmg1 and Pmg5. J Bacteriol **135:**227-238.

117. **Bhattacharyya A, Stilwagen S, Reznik G, Feil H, Feil WS, Anderson I, Bernal A, D'Souza M, Ivanova N, Kapatral V, Larsen N, Los T, Lykidis A, Selkov E, Walunas TL, Purcell A, Edwards RA, Hawkins T, Haselkorn R, Overbeek R, Kyrpides NC, Predki PF.** 2002. Draft sequencing and comparative genomics of *Xylella fastidiosa* strains reveal novel biological insights. Genome Res **12:**1556-1563.

118. **Bextine B, Child B.** 2007. *Xylella fastidiosa* genotype differentiation by SYBR Green-based QRT-PCR. FEMS Microbiol Lett **276:**48-54.

119. **Schwartz M.** 1975. Reversible interaction between coliphage lambda and its receptor protein. J Mol Biol **99:**185-201.

120. **Matsumoto A, Igo MM.** 2010. Species-specific type II restriction-modification system of *Xylella fastidiosa* temecula1. Appl Environ Microbiol **76:**4092-4095.

121. **Yu YJ, Yang MT.** 2007. A novel restriction-modification system from *Xanthomonas campestris* pv. *vesicatoria* encodes a m4C-methyltransferase and a nonfunctional restriction endonuclease. Fems Microbiology Letters **272:**83-90.

122. **Eskin B, Lautenbe.Ja, Linn S.** 1973. Host-controlled modification and restriction of bacteriophage T7 by *Escherichia coli*. B. J Virol **11:**1020-1023.

123. **Hendrix RW.** 2003. Bacteriophage genomics. Curr Opin Microbiol **6:**506-511.

124. **Summer EJ.** 2009. Preparation of a phage DNA fragment library for whole genome shotgun sequencing. Methods Mol Biol **502:**27-46.

125. **Lukashin AV, Borodovsky M.** 1998. GeneMark.hmm: new solutions for gene finding. Nucleic Acids Res **26:**1107-1115.

126. **Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, Barrell B.** 2000. Artemis: sequence visualization and annotation. Bioinformatics **16:**944-945.

127. **Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL.** 2009. BLAST plus : architecture and applications. Bmc Bioinformatics **10**.

128. **Hunter S, Jones P, Mitchell A, Apweiler R, Attwood TK, Bateman A, Bernard T, Binns D, Bork P, Burge S, de Castro E, Coggill P, Corbett M, Das U, Daugherty L, Duquenne L, Finn RD, Fraser M, Gough J, Haft D, Hulo N, Kahn D, Kelly E, Letunic I, Lonsdale D, Lopez R, Madera M, Maslen J, McAnulla C, McDowall J, McMenamin C, Mi HY, Mutowo-Muellenet P, Mulder N, Natale D, Orengo C, Pesseat S, Punta M, Quinn AF, Rivoire C, Sangrador-Vegas A, Selengut JD, Sigrist CJA, Scheremetjew M, Tate J, Thimmajanarthanan M, Thomas PD, Wu CH, Yeats C, Yong SY.** 2012. InterPro in 2011: new developments in the family and domain prediction database. Nucleic Acids Res **40:**D306-D312.

129. **Soding J, Biegert A, Lupas AN.** 2005. The HHpred interactive server for protein homology detection and structure prediction. Nucleic Acids Res **33:**W244-248.

130. **Marchler-Bauer A, Lu S, Anderson JB, Chitsaz F, Derbyshire MK, DeWeese-Scott C, Fong JH, Geer LY, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Jackson JD, Ke Z, Lanczycki CJ, Lu F, Marchler GH, Mullokandov M, Omelchenko MV, Robertson CL, Song JS, Thanki N, Yamashita RA, Zhang D, Zhang N, Zheng C, Bryant SH.** 2011. CDD: a conserved domain database for the functional annotation of proteins. Nucleic Acids Res **39:**D225-229.

131. **Juncker AS, Willenbrock H, Von Heijne G, Brunak S, Nielsen H, Krogh A.** 2003. Prediction of lipoprotein signal peptides in Gram-negative bacteria. Protein Science **12:**1652-1662.

132. **Petersen TN, Brunak S, von Heijne G, Nielsen H.** 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods **8:**785-786.

133. **Krogh A, Larsson B, von Heijne G, Sonnhammer ELL.** 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. Journal of Molecular Biology **305:**567-580.

134. **Gill JJ, Summer EJ, Russell WK, Cologna SM, Carlile TM, Fuller AC, Kitsopoulos K, Mebane LM, Parkinson BN, Sullivan D, Carmody LA, Gonzalez CF, LiPuma JJ, Young R.** 2011. Genomes and characterization of phages Bcep22 and BcepIL02, founders of a novel phage type in *Burkholderia cenocepacia*. J Bacteriol **193:**5300-5313.

135. **Kasman LM, Kasman A, Westwater C, Dolan J, Schmidt MG, Norris JS.** 2002. Overcoming the phage replication threshold: a mathematical model with implications for phage therapy. J Virol **76:**5557-5564.

136. **Adams MH.** 1959. Bacteriophages. Interscience Publishers, New York, NY.

137. **Gill JJ, Young R.** 2011. Therapeutic applications of phage biology: history, practice and recommendations. *In* Miller AA, Miller PF (ed.), Emerging trends in antibacterial discovery : answering the call to arms, Caister Academic Press, Norfolk, VA.

138. **Casjens SR, Gilcrease EB, Winn-Stapley DA, Schicklmaier P, Schmieger H, Pedulla ML, Ford ME, Houtz JM, Hatfull GF, Hendrix RW.** 2005. The generalized transducing *Salmonella* bacteriophage ES18: complete genome sequence and DNA packaging strategy. J Bacteriol **187:**1091-1104.

139. **Lynch KH, Stothard P, Dennis JJ.** 2012. Comparative analysis of two phenotypically-similar but genomically-distinct *Burkholderia cenocepacia*-specific bacteriophages. Bmc Genomics **13:**223.

140. **Kazaks A, Dislers A, Lipowsky G, Nikolajeva V, Tars K.** 2012. Complete genome sequence of the *Enterobacter cancerogenus* bacteriophage Enc34. J Virol **86:**11403-11404.

141. **Lee JH, Shin H, Choi Y, Ryu S.** 2013. Complete genome sequence analysis of bacterial-flagellum-targeting bacteriophage chi. Arch Virol. DOI:10.1007/s00705-013-1700-0

142. **Onmus-Leone F, Hang J, Clifford RJ, Yang Y, Riley MC, Kuschner RA, Waterman PE, Lesho EP.** 2013. Enhanced de novo assembly of high throughput pyrosequencing data using whole genome mapping. PLoS One **8:**61762.

143. **Rydman PS, Bamford DH.** 2002. The lytic enzyme of bacteriophage PRD1 is associated with the viral membrane. J Bacteriol **184:**1502-1502.

144. **Xu J, Hendrix RW, Duda RL.** 2004. Conserved translational frameshift in dsDNA bacteriophage tail assembly genes. Mol Cell **16:**11-21.

145. **Morgan GJ, Hatfull GF, Casjens S, Hendrix RW.** 2002. Bacteriophage Mu genome sequence: analysis and comparison with Mu-like prophages in *Haemophilus*, *Neisseria* and *Deinococcus*. J Mol Biol **317:**337-359.

146. **Ceyssens PJ, Lavigne R, Mattheus W, Chibeu A, Hertveldt K, Mast J, Robben J, Volckaert G.** 2006. Genomic analysis of *Pseudomonas aeruginosa* phages LKD16 and LKA1: establishment of the phi KMV subgroup within the T7 supergroup. J Bacteriol **188:**6924-6931.

147. **Roucourt B, Lavigne R.** 2009. The role of interactions between phage and bacterial proteins within the infected cell: a diverse and puzzling interactome. Environmental Microbiology **11:**2789-2805.

148. **Casjens SR, Thuman-Commike PA.** 2011. Evolution of mosaically related tailed bacteriophage genomes seen through the lens of phage P22 virion assembly. Virology **411:**393-415.

149. **Adriaenssens EM, Ceyssens PJ, Dunon V, Ackermann HW, Van Vaerenbergh J, Maes M, De Proft M, Lavigne R.** 2011. Bacteriophages LIMElight and LIMEzero of *Pantoea agglomerans*, belonging to the "phiKMV-like viruses". Appl Environ Microb **77:**3443-3450.

150. **Briers Y, Peeters LM, Volckaert G, Lavigne R.** 2011. phiKMV encodes a signal-arrest-release endolysin and a pinholin. Bacteriophage **1:**25-30.

151. **Xu M, Struck DK, Deaton J, Wang IN, Young R.** 2004. A signal-arrest-release sequence mediates export and control of the phage P1 endolysin. P Natl Acad Sci USA **101:**6415-6420.

152. **Kuty GF, Xu M, Struck DK, Summer EJ, Young R.** 2010. Regulation of a phage endolysin by disulfide caging. J Bacteriol **192:**5682-5687.

153. **Park T, Struck DK, Dankenbring CA, Young R.** 2007. The pinholin of lambdoid phage 21: control of lysis by membrane depolarization. J Bacteriol **189:**9135-9139.

154. **Rakhuba DV, Kolomiets EI, Dey ES, Novik GI.** 2010. Bacteriophage receptors, mechanisms of phage adsorption and penetration into host cell. Pol J Microbiol **59:**145-155.

155. **Pelkonen S, Aalto J, Finne J.** 1992. Differential activities of bacteriophage depolymerase on bacterial polysaccharide - binding is essential but degradation is inhibitory in phage infection of K1-defective *Escherichia coli*. J Bacteriol **174:**7757-7761.

156. **Levin BR, Bull JJ.** 2004. Population and evolutionary dynamics of phage therapy. Nat Rev Microbiol **2:**166-173.

157. **Davis MJ, French WJ, Schaad NW.** 1981. Axenic culture of the bacteria associated with phony disease of peach and plum leaf scald. Curr Microbiol **6:**309-314.

158. **Bertani G.** 1951. Studies on Lysogenesis: the mode of phage liberation by lysogenic *Escherichia coli*. J Bacteriol **62:**293-300.

159. **Cursino L, Li Y, Zaini PA, De La Fuente L, Hoch HC, Burr TJ.** 2009. Twitching motility and biofilm formation are associated with tonB1 in *Xylella fastidiosa*. FEMS Microbiol Lett **299:**193-199.

160. **Yanisch-Perron C, Vieira J, Messing J.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33:**103-119.

161. **Matsumoto A, Young GM, Igo MM.** 2009. Chromosome-based genetic complementation system for *Xylella fastidiosa*. Appl Environ Microb **75:**1679-1687.

162. **Hamad MA, Zajdowicz SL, Holmes RK, Voskuil MI.** 2009. An allelic exchange system for compliant genetic manipulation of the select agents *Burkholderia pseudomallei* and *Burkholderia mallei*. Gene **430:**123-131.

163. **White TJ, Gonzalez CF.** 1991. Application of electroporation for efficient transformation of *Xanthomonas campestris* pv *oryzae*. Phytopathol **81:**521-524.

164. **Craig L, Pique ME, Tainer JA.** 2004. Type IV pilus structure and bacterial pathogenicity. Nat Rev Microbiol **2:**363-378.

165. **Roine E, Raineri DM, Romantschuk M, Wilson M, Nunn DN.** 1998. Characterization of type IV pilus genes in *Pseudomonas syringae* pv tomato DC3000. Mol Plant Microbe In **11:**1048-1056.

166. **Damien F. Meyer AJB.** 2009. Genomics-driven Advances in *Xanthomonas* Biology. *In* Jackson RW (ed.), Plant pathogenic bacteria: genomics and molecular biology, Caister Academic, Norfolk, VA.

167. **Monteiro-Vitorello CB, De Oliveira MC, Zerillo MM, Varani AM, Civerolo E, Van Sluys MA.** 2005. *Xylella* and *Xanthomonas* Mobil'omics. Omics **9:**146-159.

168. **Vauterin L, Yang P, Alvarez A, Takikawa Y, Roth DA, Vidaver AK, Stall RE, Kersters K, Swings J.** 1996. Identification of non-pathogenic *Xanthomonas* strains associated with plants. Syst Appl Microbiol **19:**96-105.

169. **Wang L, Makino S, Subedee A, Bogdanove AJ.** 2007. Novel candidate virulence factors in rice pathogen *Xanthomonas oryzae* pv *oryzicola* as revealed by mutational amalysis. Appl Environ Microb **73:**8023-8027.

170. **Lim SH, So BH, Wang JC, Song ES, Park YJ, Lee BM, Kang HW.** 2008. Functional analysis of pilQ gene in *Xanthomanas oryzae* pv *oryzae*, bacterial blight pathogen of rice. J Microbiol **46:**214-220.

171. **Darsonval A, Darrasse A, Durand K, Bureau C, Cesbron S, Jacques MA.** 2009. Adhesion and fitness in the bean phyllosphere and transmission to seed of *Xanthomonas fuscans* subsp *fuscans*. Mol Plant Microbe In **22:**747-757.

172. **Qian W, Jia YT, Ren SX, He YQ, Feng JX, Lu LF, Sun QH, Ying G, Tang DJ, Tang H, Wu W, Hao P, Wang LF, Jiang BL, Zeng SY, Gu WY, Lu G, Rong L, Tian YC, Yao ZJ, Fu G, Chen BS, Fang RX, Qiang BQ, Chen Z, Zhao GP, Tang JL, He CZ.** 2005. Comparative and functional genomic analyses of the pathogenicity of phytopathogen *Xanthomonas campestris* pv *campestris*. Genome Res **15:**757-767.