Natural competence and intersubspecific recombination in *Xylella fastidiosa*; a causal agent of emerging plant diseases

by

Prem Prasad Kandel

A dissertation submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Auburn, Alabama
August 05, 2017

Keywords: *Xylella fastidiosa*, natural competence, twitching motility, intersubspecific recombination, disease emergence, genome

Approved by

Leonardo De La Fuente, Chair, Associate Professor of Plant Pathology
Joseph W. Kloepper, Professor of Plant Pathology
John F. Murphy, Professor of Plant Pathology
Paul A. Cobine, Associate Professor of Biological Sciences

Abstract

Xylella fastidiosa is a Gram-negative, vector-transmitted, plant-pathogenic bacterium causing incurable diseases to economically important crops such as grapevine, citrus, almond, and coffee. Recent epidemiological studies show that *X. fastidiosa* is expanding its plant-host range and is spreading to new geographical locations. Moreover, strains isolated from these new infections were genetically different and showed presence of homologous recombination (HR) in some cases. HR between subspecies of *X. fastidiosa* has been described as a major evolutionary force in creating genetic diversity and causing plant-host shift. As *X. fastidiosa* is naturally competent, this could be an explanation for intersubspecific recombination and disease emergence. However, limited information exists regarding natural competence of *X. fastidiosa* in natural habitats and the factors that influence it, the degree to which strains are naturally competent, whether intersubspecific recombination can be experimentally demonstrated, and if intersubspecific recombination can be detected in the genomes of *X. fastidiosa*. Therefore, experiments were performed by co-culturing combinations of recipient strains and donor DNA (plasmids, marker tagged live- and heat-killed cells) under various growth conditions; by selecting the recombinants based on the acquisition of marker, and testing their phenotypes *in vitro* and *in planta*; and by sequencing and comparing whole genomes of the parent and recombinant strains. Results demonstrated that natural competence readily occurs under *in vitro* conditions that mimic the natural habitats of the bacterium. Moreover, widespread natural competence ability was detected among strains isolated from diverse host plants, and intersubspecific recombination was validated experimentally. The rate of recombination was

significantly correlated with growth rate and twitching motility. Whole genome sequence analysis of *in vitro* generated recombinants and virulent strains isolated from infected fields demonstrated the presence of intersubspecific recombination in up to 10kb genomic regions. Results of this study are useful for policy makers in formulating strict quarantine measures to limit the introduction of new *X. fastidiosa* genotypes and prevent both intra- and intersubspecific recombination between the introduced and native strains, which will minimize the risk of new disease emergence and reduce economic losses due to the incurable diseases caused by *X. fastidiosa*.

Acknowledgments

I would like to express my sincere gratitude to my advisor Dr. Leonardo De La Fuente for his constant support and guidance throughout my Ph.D. His patience, motivation, and enthusiasm were always the motivating factors for me.

I would also like to sincerely thank the committee members Dr. Kloepper, Dr. Murphy, and Dr. Cobine for their guidance for my project; In addition, I want to thank outside university reader Dr. Peter Panizzi for his constructive suggestions.

I would like to thank Jennifer Parker, Luisa Cruz, Fernando Navarette, and Jonathan Oliver for teaching me the basics in the lab besides their suggestion with my project.

I thank lab members Sy, Hajeewaka, Edel, Hongyu, Eber, Laura, Qiang, and Marcus for their helpful suggestions.

I thank undergraduate interns Samantha, Lillian, and Courtney for helping me with the experiments.

I dedicate this work to my late father, Chet Prasad Kandel, for his devotion to my success and my mother, Dhanamaya Kandel, my brothers, Tilak Prasad Kandel and Hari Kandel, and all family members for their constant support and motivation.

Finally, I would like to thank my wife Bini Adhikaree who came into my life during the Ph.D. and constantly supports and encourages me along the way.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

BLS	Bacterial Leaf Scorch
BSA	Bovine Serum Albumin
CFU	Colony Forming Units
CTAB	Cetyltrimethylammonium Bromide
CVC	Citrus Variegated Chlorosis
DNA	Deoxyribose Nucleic Acid
DSF	Diffusible Signaling Factor
DUES	DNA Uptake Enhancing Sequences
HR	Homologous Recombination
IHR	Intersubspecific Homologous Recombination
MC	Microfluidic Chambers
MLSA	Multi-Locus Sequence Analysis
MLST	Multi-Locus Sequence Typing
PAMP	Pathogen Associated Molecular Patterns
PCR	Polymerase Chain Reaction
PD	Pierce's Disease
TEM	Transmission Electron Microscopy
XFM	*Xylella fastidiosa* Medium

Chapter 1

Overview of *Xylella fastidiosa* taxonomy, diversity, distribution, disease emergence, homologous recombination, and natural competence

1. *Xylella fastidiosa*

Xylella fastidiosa is a Gram-negative, fastidious, plant-pathogenic/endophytic bacterium causing serious diseases in crops of economic importance such as grapevine, citrus, almond, peach, plum, pecan, and coffee (Hopkins and Purcell, 2002). The bacterial cells are rod shaped with a length of 0.9-3.5µm and a radius of 0.25-0.35µm (Wells et al., 1987). The bacterium is slow-growing, strictly-aerobic, mesophilic (optimum growth occurs at 26-28°C), and neutrophilic (pH of 6.5-6.9) (Wells et al., 1987). Cells are aflagellar, but possess two types of pili, short type I pili (0.4 to 1.0 µm) and long type IV pili (1.0 to 5.8 µm), both positioned at one of the cell poles (Meng et al., 2005). Type I pili are involved in adhesion of cells to surfaces and auto-aggregation, while type IV pili are involved in twitching motility (Meng et al., 2005; De La Fuente et al., 2007): a characteristic type of bacterial motility that involves dragging of cells on surfaces facilitated by extension, tethering, and retraction of the type IV pili.

1.1. Disease mechanism, major diseases, and symptoms of *X. fastidiosa*

The disease mechanism, although not fully understood, involves attachment of *X. fastidiosa* cells and formation of biofilm-like aggregates in the xylem vessels which are conduits of water and mineral nutrients in the plant (Chatterjee et al., 2008). Biofilm, together with gum and tylose formation in response to *X. fastidiosa* infection (Sun et al., 2013), obstructs the flow of xylem sap through the vessels. This obstruction induces disease symptoms typical of water deficiency and mineral imbalance (Chatterjee et al., 2008; De La Fuente et al., 2013). Disease development

is associated with systemic colonization of the xylem vessels (Newman et al., 2003). Systemic colonization is facilitated by twitching motility (Meng et al., 2005), and ability of the bacterium to degrade pit membranes in the xylem vessels (Roper et al., 2007; Perez-Donoso et al., 2010). Recent studies have demonstrated the involvement of mineral nutrients of xylem sap, calcium in particular, in the disease process of surface attachment, biofilm formation, and twitching motility (Cruz et al., 2012; De La Fuente et al., 2013; Cruz et al., 2014; Parker et al., 2016).

The major diseases caused by *X. fastidiosa* are Pierce's Disease (PD) in grapes; citrus variegated chlorosis (CVC); bacterial leaf scorch (BLS) in almond, blueberry, oleander, oak; phony peach disease; plum leaf scald; and olive quick decline syndrome (EEPO, 2016). PD was the first disease described for *X. fastidiosa* and was reported in California around 1884 (Pierce, 1892). PD has now spread to several states in the southeastern U.S. including Florida, Georgia and Texas (Parker et al., 2012). *X. fastidiosa* has received significant research importance after CVC emerged in Brazil in 1987 and PD re-emerged in California in 1990s (Hopkins and Purcell, 2002). CVC was reported to affect approximately 40% of citrus trees in Brazil and cause an annual loss of around $120 million (Bove and Ayres, 2007). During the years from 1994-2000, PD destroyed more than 1,000 acres of grapevines in northern California causing damages of over $30 million (Brown et al., 2002). The host range of *X. fastidiosa* has been constantly expanding and now includes 359 plant species in 75 botanical families (EFSA, 2016) and also has a wide-spread geographical distribution (Almeida and Nunney, 2015).

Disease symptoms differ with the plant host and severity of infection. Symptoms associated with PD and BLS involve marginal leaf scorch that progresses towards the mid vein with a clear

yellow to purplish zone separating the scorched and green region, yellowing, wilting, cupping, and curling of leaves; twig dieback; leaf drop and match sticks; green islands on stems; dwarfing of plants; and shriveling (mummification) of berries (Hopkins and Purcell, 2002; Oliver et al., 2014; Oliver et al., 2015). Symptoms on citrus differ and do not involve leaf scorch, but leaves demonstrate interveinal chlorosis with gummy lesions on the underside of leaves, and reduced fruit size (Hopkins and Purcell, 2002).

1.2. Insect vectors of *X. fastidiosa*

X. fastidiosa is exclusively vector-transmitted by two groups of xylem sap feeding insects: sharpshooter leafhoppers (Cicadellidae) and spittlebugs (superfamily Cercopoidea, with five species of Aphrophoridae and two species of Clastopteridae) (Almeida and Purcell, 2003; Almeida et al., 2008). Vector colonization occurs in a persistent and non-circulative manner with bacterial cells forming biofilms at the cibarium and precibarium region of the foregut (Chatterjee et al., 2008). No transovarial or transstadial transmission has been demonstrated. A cell-cell communication signal (diffusible signaling factor; DSF) has been shown to be involved in biofilm formation in the vectors and transmission to the plant hosts (Newman et al., 2004). Two important aspects of vector transmission in *X. fastidiosa* are that there is no latent period for transmission and no specificity is involved between the vector species and *X. fastidiosa* isolates (Almeida and Nunney, 2015). The vectors are distributed worldwide and can feed on xylem sap from a diverse group of plant hosts. Two vector species implicated in the transmission and used in experimental studies are the blue green sharpshooter (*Graphocephala atropunctata*) and the glassy-winged sharpshooter (*Homalodisca vitripennis*) (Almeida and Nunney, 2015).

1.3. Taxonomy and genetic diversity of *X. fastidiosa*

Xylella fastidiosa is a gamma-proteobacterium in the family Xanthomonadaceae (Wells et al., 1987). Isolates were grouped within a single species of *X. fastidiosa*, but recently another species, *X. taiwanensis*, was described for the Taiwanese strains causing pear leaf scorch (Su et al., 2016). Initially, three subspecies were described within the species of *X. fastidiosa* by DNA-DNA hybridization: subsp. *fastidiosa*, subsp. *multiplex*, and subsp. *pauca* (Schaad et al., 2004). This subspecific designation was later confirmed by Multi Locus Sequence Typing (MLST) of seven housekeeping genes with two more subspecies described from isolates not included in the previous studies: subsp. *sandyi*, and subsp. *morus* (Scally et al., 2005; Yuan et al., 2010; Nunney et al., 2014b). Subspecies demonstrate a certain level of host specificity (Almeida et al., 2008; Oliver et al., 2014; Oliver et al., 2015), but this is rapidly changing due to extensive homologous recombination present in *X. fastidiosa* (Almeida et al., 2008; Nunney et al., 2012; Nunney et al., 2014a; Nunney et al., 2014b; Coletta-Filho et al., 2016). A recent study based on core genome analysis of 21 *X. fastidiosa* strains proposed only three clearly defined subspecies: *fastidiosa, multiplex* and *pauca,* with *morus* and *sandyi* strains grouped within subsp. *fastidiosa* (Marcelletti and Scortichini, 2016).

The most divergent strains (PD strain Temecula and CVC strain 9a5c) have 76% conserved sequences and share 98% of the genes as determined by comparative genomic analysis between the two strains, which explains that there is very limited genetic diversity in *X. fastidiosa* (Van Sluys et al., 2003). Despite this limited genetic variation, *X. fastidiosa* isolates differ in phenotypic characters associated with virulence such as plant host range, symptom type, and disease severity (Oliver et al., 2014; Oliver et al., 2015). With the use of MLST/MLSA and

MLSA of environmentally-mediated genes, genetic diversity has been detected among isolates between and within subspecies (Almeida et al., 2008; Yuan et al., 2010; Parker et al., 2012; Nunney et al., 2014a; Nunney et al., 2014b). More insights are likely to be generated given the ease with which whole genome sequences can be generated.

1.4. Geographical distribution of *X. fastidiosa*

X. fastidiosa was initially believed to be a virus restricted to North America (Hewitt, 1958). However, it was later described to be a fastidious bacterium inhabiting the xylem vessels by subsequent studies. *X. fastidiosa* is now present in the Americas, Taiwan, and Italy (Saponari et al., 2013; Retchless et al., 2014), and some yet to be confirmed cases in Iran (Amanifar et al., 2014) and Turkey (Güldür et al., 2005). More recently reports of *X. fastidiosa* detection from France and Spain are available (Strona et al., 2017).

MLST analysis of *X. fastidiosa* strains suggested evolution of different subspecies in geographic isolation with subsp. *fastidiosa* strains originated from Costa Rica in Central America (Nunney et al., 2010), subsp. *multiplex* strains native to temperate and subtropical North America (Nunney et al., 2012; Nunney et al., 2014a), and subsp. *pauca* strains native to South America (Nunney et al., 2012). Current reports show that subsp. *fastidiosa* is present in the U.S., Central America, and Taiwan; subsp. *multiplex* in the U.S. and Brazil; and subsp. *pauca* in South and Central America and Italy (Almeida and Nunney, 2015).

1.5. Homologous recombination and disease emergence in *X. fastidiosa*

Homologous recombination (HR) involves exchange of DNA fragments between two similar DNA molecules and is a hallmark of creating genetic diversity in almost every life form (Dorer et al., 2011). In bacteria, HR is a major mechanism involved in repair of double strand breaks in DNA and allows maintenance of genome integrity when the genome experiences interruption of the replication fork or DNA damages due to exposure to mutagens such as UV or reactive oxygen species (Dorer et al., 2011; Lenhart et al., 2012). HR in bacteria occurs in nature via three different modes of horizontal gene transfer: conjugation, transduction, and natural transformation. Conjugation and transduction involve plasmids and bacteriophages, respectively, to transfer DNA from one bacterial cell to the other, while natural transformation involves uptake of naked extra-cellular DNA fragments by the recipient cell and incorporation into the genome of the bacteria (Seitz and Blokesch, 2013). In *X. fastidiosa,* genetic studies mainly based on MLST have shown that genetic diversity results from horizontal gene transfer and genetic recombination. In fact, a MLST study showed that HR had a greater role than point mutation in creating allelic variation in this bacterium (Scally et al., 2005). A very widespread HR has been detected by studies utilizing *X. fastidiosa* strains collected from diverse geographic locations and plant hosts (Scally et al., 2005; Almeida et al., 2008; Nunney et al., 2010; Yuan et al., 2010; Nunney et al., 2012; Nunney et al., 2014a; Nunney et al., 2014b; Coletta-Filho et al., 2016). Moreover, some of these studies predicted HR between subspecies causing host shift in *X. fastidiosa* (Nunney et al., 2014a; Nunney et al., 2014b). Interestingly, *X. fastidiosa* was recently described to be naturally competent (Kung and Almeida, 2011), providing a justification for frequent HR events detected in *X. fastidiosa* strains.

1.6. Natural competence: definition, discovery, and distribution in the bacterial kingdom

Natural competence or natural DNA transformation is a phenomenon that involves uptake of naked DNA fragments from the environment and incorporation into the bacterial genome by HR (Lorenz and Wackernagel, 1994). Natural competence was initially proposed as a mechanism of nutrient acquisition as cells developed competence under starvation conditions (Seitz and Blokesch, 2013). Although, the exact role of natural competence is unclear, it is implicated in repair of damaged DNA (Dorer et al., 2010) and generation of genetic diversity that leads to adaptation (Baltrus et al., 2008) and increased virulence (Griffith, 1928; Coupat-Goutaland et al., 2011).

Natural competence was discovered in 1928 by Frederick Griffith (1928) and best described in 1999 by David Dubnau (Dubnau, 1999). Griffith worked with two strains of *Streptococcus pneumonae*, a type III-S (smooth) and type II-R (rough) strain. The smooth strain is lethal to the host as it covers itself with a polysaccharide capsule that enables it to evade host immune responses. The rough strain, however, lacks the protective capsule covering and succumbs to the host's immune response. By mixing heat-killed smooth strains with live rough strains, Griffith showed that the rough strains converted into lethal strains. He concluded that some 'transforming principle' moved from the dead to the live cells resulting in conversion of rough strains to lethal strains. Subsequent studies demonstrated that the transforming principle was DNA and that DNA serves as the genetic material in the cells (Lorenz and Wackernagel, 1994). Since the discovery of natural competence, more than 80 bacterial species have been described to be naturally competent with many aspects of competence development identified among diverse bacterial systems. For instance, although competence is constitutive at all growth phases of *Neisseria*

meningitides (Catlin, 1960), *N. gonorrhoeae* (Biswas et al., 1977), and *Helicobacter pylori* (Israel et al., 2000), *S. pneumoniae* is competent at mid-log phage (Håvarstein et al., 1995) whereas *Haemophilus influenzae* at stationary growth phase (Macfadyen et al., 2001). Another naturally competent bacterium, *Acinetobacter* sp., does not differentiate self or foreign DNA and is always competent, but the transformation frequency varies with the growth phase and is favored by a nutritional boost (Palmen et al., 1994). *H. influenzae* prefers species specific (self) DNA and the selection for self-DNA is mediated by DNA uptake sequences (DUS), short and conserved sequence motifs highly represented in the genome and recognized by hypothetical proteins on the cell surface (Smith et al., 1999). *Neisseria* spp. also use DUS or DUES (DNA uptake enhancing sequences) to facilitate uptake of DNA from closely related species (Elkins et al., 1991). In *H. pylori*, restriction modification systems have been found to limit the uptake of distantly related DNA (Humbert and Salama, 2008) although, it was believed that *H. pylori* can take up any DNA independent of its source.

1.7. Mechanisms of natural competence

Gram-negative and Gram-positive bacteria differ to some extent when it comes to the mechanisms of natural DNA transformation mainly because of the fact that Gram-negative bacteria possess an outer cell wall, which does not occur for Gram-positive bacteria. This would mean that the initial interaction between the DNA and the cell components is different in these two groups of bacteria. Research has shown that DNA translocation through the inner membrane is similar and carried out by orthologous proteins (Dubnau, 1999). The general scheme of natural competence that applies to both groups of bacteria can be presented as follows: 1. Binding of extra-cellular DNA by cell surface proteins, 2. Degradation of the 3' strand of the

DNA with an exonuclease, 3. Entry of the 5’ strand into the cytoplasm, 4. Binding of the single-stranded DNA (ssDNA) with SSB proteins such as DprA and RecA, and 5. Scanning for homology by RecA and homologous recombination with the recipient chromosome (Bakkali, 2013).

In Gram-positive bacteria, such as *Bacillus subtilis,* binding of DNA occurs to specific sites on the cell surface. No preference has been found for the nucleotide sequence (Dubnau, 1999). Competence development in Gram-positive bacteria is regulated by quorum sensing signals involving competence stimulating proteins (CSP) (Håvarstein et al., 1995). CSPs were shown to activate a two-component regulatory system formed by a histidine-kinase sensor ComD, and the response regulator ComE. Phosphorylation of ComE by ComDE activates *S*igX, an alternative sigma factor that activates the expression of competence genes (Håvarstein et al., 1995).

In Gram-negative bacteria, proteins involved in the biogenesis of type IV pili, type II secretion system, and twitching motility (PSTC proteins; pili, secretion, twitching, and competence) are involved in binding of extra-cellular DNA and its translocation to the periplasmic space (Hobbs and Mattick, 1993). In the periplasm, other competence-related proteins stabilize the DNA and direct it toward the inner membrane translocase. Concomitant degradation of the 3’ strand and entry of the 5’strand into the cytoplasm occurs in the inner membrane. In the cytoplasm, the ssDNA binds to a stabilizing protein, DprA, which recruits RecA. The RecA:ssDNA complex now scans for homology in the chromosomal DNA and recombination occurs between the incoming ssDNA and the chromosomal DNA (Seitz and Blokesch, 2013). The process of natural

competence that applies to majority of Gram-negative cells is shown below in figure 1-1.

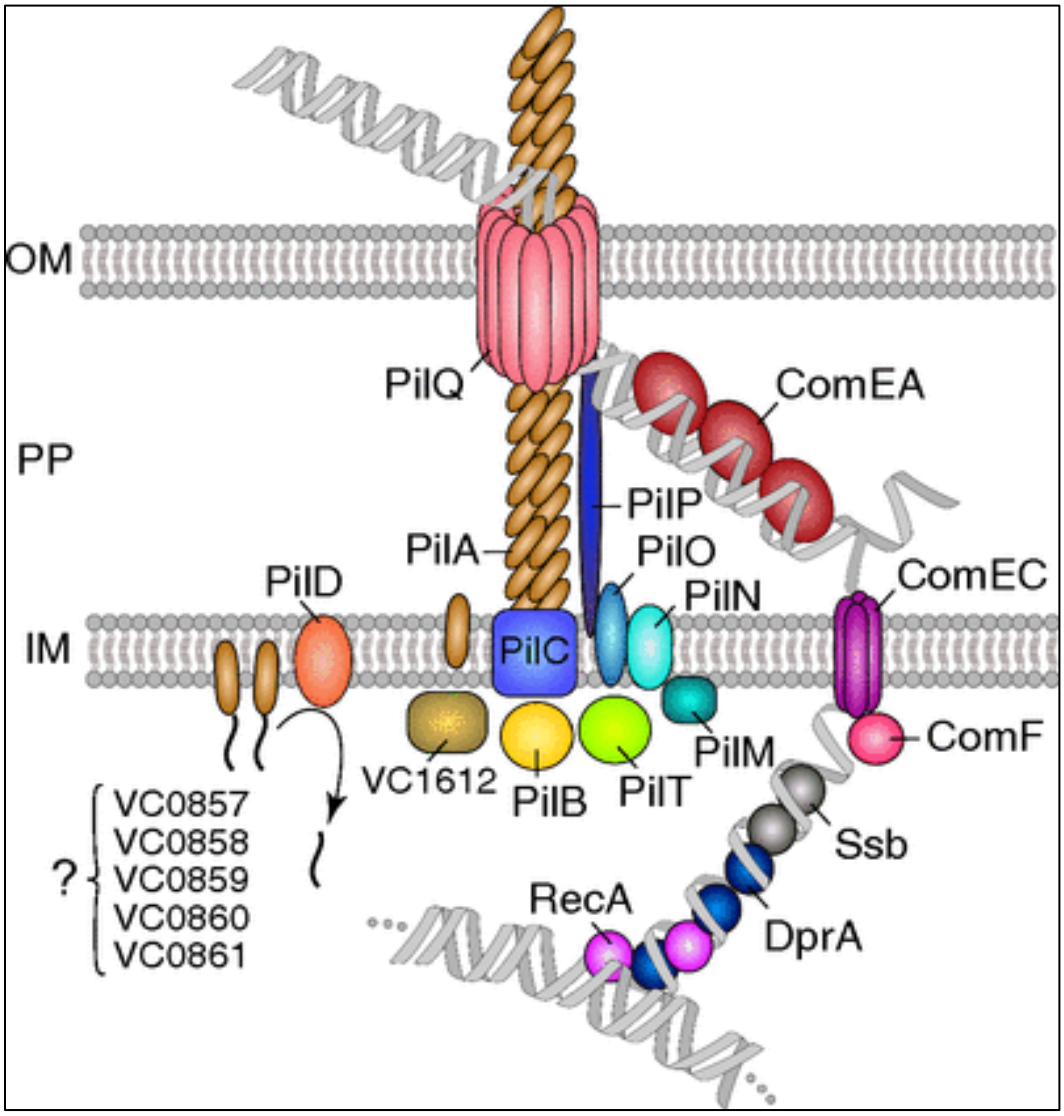


Fig. 1-1. Process of natural competence in Gram-negative bacteria (Seitz and Blokesch, 2013)

1.8. Regulation of competence induction

The reason for existence of natural competence in bacteria is debatable as different bacteria induce competence under different conditions. Three hypotheses have been presented to explain natural competence. First, DNA serves as a nutrient source under starvation conditions, and

induction of the competence state leads to uptake of DNA to serve as food for the bacterial cells. Second, incoming DNA serves as a repair template for damaged DNA caused by DNA damaging agents thereby providing a survival mechanism under stressful conditions. Third, natural competence drives evolution and adaptation by creating genetic diversity. The following section covers details of each of these hypotheses with specific examples from representative bacterial species.

1.8.1. Natural competence under starvation

As stated earlier, starvation signals induce natural competence in *H. influenzae*. Cells develop competence at the stationary growth phase and when transferred from a nutrient rich medium to a poor medium (Herriott et al., 1970). Moreover, supplementation of growth medium with cAMP, a secondary messenger that is accumulated in the cells under starvation condition, also induced competence in *H. influenzae* (Wise et al., 1973). The cAMP receptor protein (CRP) (Chandler, 1992), and adenylate cyclase (CyaA) (Dorocicz et al., 1993) were also required suggesting a signal transduction dependent induction. In addition, a cAMP dependent gene *sxy*, homologue of *tfoX* in other competent bacteria, was required for natural competence of *H. influenzae* (Redfield, 1991). The expression of *sxy* was up-regulated by addition of cAMP (Zulty and Barcak, 1995). Sxy was shown to influence the competence regulon consisting of 25 genes in 13 different transcriptional units identified using microarray expression profiling (Redfield et al., 2005). Each of these transcriptional units contains a competence regulatory element (CRE), a 22-bp element associated with the promoter (Karudapuram and Barcak, 1997). This element is renamed as CRP-s implying that it is a binding site for CRP, and is dependent on the Sxy protein (Cameron and Redfield, 2008). Based on these findings, and although a definite link between the

cAMP/CRP, CRP-s, and Sxy in regulating natural competence in *H. influenzae* has yet to be established, it is proposed that formation of a cAMP/CRP complex under starvation condition activates the expression of *sxy*, which then causes induction of natural competence state.

1.8.2. Natural competence for DNA damage repair

Unlike the type IV pili-like structures for DNA uptake in other Gram-negative bacteria, *H. pylori* uses type IV secretion system (T4SS), also referred to as *comB* system (Karnholz et al., 2006). *H. pylori* is constitutively competent (Israel et al., 2000), although the recombination frequencies differ according to growth stages (Baltrus and Guillemin, 2006) suggesting the involvement of some competence regulatory mechanisms. One signal that induces competence in *H. pylori* is DNA damage (Dorer et al., 2010; Dorer et al., 2011). DNA damage caused by ciprofloxacin, an antibiotic belonging to the second generation of fluoroquinolone, induced RecA dependent induction of genes such as genes for T4SS and a lysozyme-like protein (causes killing of neighboring cells and makes DNA available for uptake, recombination, and repair). As further support of this observation, mutants lacking DNA repair genes had increased expression of competence related genes (Dorer et al., 2010). This up-regulation of genes involved in DNA uptake and lysozyme dependent fratricide was also observed in the Gram-positive bacteria, *S. pneumoniae*, in response to DNA damage (Prudhomme et al., 2006). Moreover, natural competence in another Gram-negative pathogen, *Legionella pneumophilla*, is also induced by DNA damage caused by UV and antibiotics (Charpentier et al., 2011). Interestingly, these distantly related species that induce natural competence upon DNA damage lack an efficient SOS system for DNA damage repair (Seitz and Blokesch, 2013). This suggests that DNA

damage induces natural competence such that the incoming DNA could serve as a repair template.

2. Natural competence for evolution and adaptation

Natural competence has been proposed to bring genetic diversity and provide adaptive advantages to the recipient cells. However, there is a paucity of experimental studies demonstrating this phenomenon. The pioneer study of Griffith (1928) demonstrated the transfer of virulence from a non- pathogenic to pathogenic strain by natural competence. Moreover, an increased rate of adaptation following natural competence was reported in a human pathogen, *Helicobacter pylori* (Baltrus et al., 2008). A recent study in *Acinetobacter baylyi* reported a fitness benefit of DNA uptake independent of recombination and UV stress, while recombination reduced survival (Hülter et al., 2017).

Vibrio cholerae is an opportunistic, Gram-negative human pathogen that has extensive genetic diversity. Although no direct association of natural competence and adaptation has been established in this pathogen, the extensive genetic diversity suggests that natural competence plays a role probably in its evolution allowing it to evade the immune pressure. Natural competence in *V. cholerae* is induced when the bacteria grow on chitin surfaces (Meibom et al., 2005), an environment the bacterium experiences in its natural aquatic growth habitat. The regulatory mechanisms of natural competence have been extensively studied in this bacterium. Regulatory proteins named as TfoX (Scrudato and Blokesch, 2012) and HapR (Scrudato and Blokesch, 2013) control genes involved in DNA uptake, transportation, stability, and

homologous recombination of the incoming DNA with the chromosome. TfoX expression is dependent on growth of bacteria on chitin surfaces, while HapR expression is based on cell-cell communication signals. HapR has been described to be the main regulator of quorum sensing in *V. cholerae* (Scrudato and Blokesch, 2012).

Although the link between growth on chitin surfaces and expression of *tfoX* is not clear, a recent study (Dalia et al., 2014) identified a membrane bound transcriptional regulator, TfoS that is predicted to be a membrane spanning protein containing a periplasmic domain and a cytoplasmic helix-turn-helix DNA binding domain. The mutants of *tfoS* genes are not competent even when induced by culturing the cells with chitin (Dalia et al., 2014). It was further shown that the periplasmic domain is important for competence as a deletion mutant of this domain was non-transformable. Furthermore, overexpression of *tfoX* rescued the effect of the *tfoS* mutant indicating that TfoS acts upstream of TfoX and might have a regulatory effect on *tfoX* expression. Interestingly, it has been shown that initiation of translation of the *tfoX* mRNA is controlled by a Hfq based small RNA, TfoR. TfoR, by base pairing with the stem loop structure of *tfoX* mRNA, exposes its ribosome binding site facilitating translation. This induction of translation of *tfoX* activates the downstream competence related genes (Yamamoto et al., 2011).

HapR, the quorum sensing regulator which is produced only at high cell density, has been shown to repress the extracellular nuclease (Seper et al., 2011). This repression is important for natural competence as any surrounding DNA that can be transferred by natural competence is degraded in the presence of nucleases (Blokesch and Schoolnik, 2008). In addition, HapR also induces the expression of *comEA*, a gene encoding the periplasmic DNA binding protein (Meibom et al., 2005; Blokesch and Schoolnik, 2008) and a subset of other competence related genes such as

pilA and *comEC* (Scrudato and Blokesch, 2012). Hence, cell-cell communication signals promote competence in *V. cholera*e by suppressing the degradation of extracellular DNA and stabilizing the incoming DNA in the periplasmic space and in the cytoplasm (Scrudato and Blokesch, 2012).

2.1. Natural competence in plant pathogenic bacteria

Natural competence has been described in two plant pathogenic bacteria, *Ralstonia solanacearum* (Bertolla et al., 1997) and *X. fastidiosa* (Kung and Almeida, 2011). *R. solanacearum* showed the highest transformation frequency when cultured under minimal medium (Bertolla et al., 1997). Moreover, the cells were competent at the exponential growth phase with the transformation frequency reaching up to 4×10^{-7} per cell. The competent cells preferred DNA from closely related species and concentration of DNA in the transforming medium affected transformation frequency. These results suggest that *R. solanacearum* competence might be induced by starvation signals, and the preference for self-DNA might suggest the competence for DNA repair hypothesis. In another study, very large fragments of DNA (at least 39.4 - 78.9 kb) that included 20 new open reading frames and 11 new alleles were found to be transformed from the donor to the recipient cells (Coupat-Goutaland et al., 2011). One of the transformed strains was significantly more aggressive in causing diseases in tomato plants compared to the wild-type recipient. These results suggest that *R. solanacearum* could use natural competence to acquire new genes and alleles to become better adapted in its habitat. This supports the hypothesis of natural competence for evolution but more research will be required to verify this finding.

2.1.1. Natural competence in *Xylella fastidiosa*

Natural competence in *X. fastidiosa* was recently described (Kung and Almeida, 2011). Cells were competent at the exponential growth phase and a minimal medium was more favorable than a complex and rich medium. DNA methylated with *X. fastidiosa* methyl transferase was more efficiently recombined than non-methylated DNA suggesting a preference for self-DNA (Kung and Almeida, 2011). A recombination frequency of ~10^{-5} per recipient cell was described in the liquid medium which was later shown to increase by two orders of magnitude under solid media agar conditions (Kung and Almeida, 2011; Kung and Almeida, 2014). Cells required at least a 96bp flanking homology for natural competence to occur with maximum natural competence occurring with plasmids containing 1kb of flanking homology. Length of non-homologous insert was negatively correlated to the recombination frequency supporting a requirement of self-DNA for successful natural competence (Kung et al., 2013). Moreover, two strains were shown to recombine when co-cultured (Kung and Almeida, 2011), suggesting that natural competence could be a mechanism of horizontal gene transfer in *X. fastidiosa*.

Limited information exists, however, on the importance of natural competence in the epidemics of *X. fastidiosa*. More specifically, it is not known if natural habitats of the bacteria are conducive for natural competence, and if there are any environmental cues that govern natural competence in *X. fastidiosa*. Moreover, natural competence could provide an explanation for plant host shift and new disease emergence currently observed in *X. fastidiosa*, and proposed to be the result of homologous recombination between different subspecies of *X. fastidiosa*. Clearly, additional studies are required to understand the role of natural competence in *X. fastidiosa* genetic diversity and disease emergence.

Long-term goal

Define the role of natural competence and intersubspecific recombination in the diversity, virulence, and host adaptations of *X. fastidiosa*.

Specific objectives

1. Determine the environmental factors influencing natural competence and to test its occurrence under natural media flow condition mimicking the natural habitat
2. Test variability of natural competence and intersubspecific recombination *in vitro* in *X. fastidiosa*
3. Examine the role of intersubspecific recombination in *X. fastidiosa* virulence and disease emergence

Hypotheses

1. On co-culturing different strains of *X. fastidiosa*, recombination will occur in media flow condition.
2. Natural competence rates will be significantly different among *X. fastidiosa* strains.
3. Substantial numbers of recombinants will form on co-culturing strains of two *X. fastidiosa* subspecies.
4. Intersubspecific recombinants will demonstrate significantly different phenotype (twitching, settling, virulence) compared to the recipient parents.
5. Whole genome analysis will detect region(s) of homologous recombination in *X. fastidiosa*.

6. Linear homologous DNA fragments can be used to construct gene knock-outs in *X. fastidiosa*.

Experiments were designed to test natural competence in a system (microfluidic chambers) that mimics the natural habitat of the bacterium, to test natural competence and compare the recombination frequency among strains belonging to different subspecies and infecting different plant hosts, to test intersubspecific recombination *in vitro* between two most prevalent subspecies in the U.S., and to test if natural competence is the mechanism involved in intersubspecific recombination. Also, experiments aimed to test if randomly selected intersubspecific recombinants differ in the virulence phenotypes from their recipient parents, and to use whole genome sequencing and analysis of *X. fastidiosa* strains and recombinants produced *in vitro* to detect the presence of intersubspecific recombination.

The research performed in this dissertation has been divided into the following four chapters. Chapter 2 entitled "natural competence of *X. fastidiosa* occurs at a high frequency inside microfluidic chambers mimicking the bacterium's natural habitat" addresses the question of whether *X. fastidiosa* develops natural competence in the natural habitat of insect foregut and plant xylem vessels. This chapter was published in the American Society of Microbiology journal "Applied and Environmental Microbiology". Chapter 3 addresses the variability of natural competence among *X. fastidiosa* strains and if intersubspecific recombination occurs under experimental conditions. This chapter has been accepted for publication in the American Phytopathological Society journal "Molecular Plant-Microbe Interactions". Chapter 4 presents the results of genome sequencing and analysis of *X. fastidiosa* strains and recombinants. Results demonstrate that up to ~10kb of genomic regions can be transferred from donor to the recipient

strains, and regions intersubspecific homologous recombination derived from *X. fastidiosa* subsp. *multiplex* were detected frequently in the genomes of *X. fastidiosa* subsp. *fastidiosa* strains. Chapter 5 presents a possibility of overlap-extension PCR and natural transformation approach in generating gene-knockouts and genetic complementation in *X. fastidiosa* with a focus given to the major pilin genes of type IV pili. Results show that this method can be successfully used for genetic manipulation in *X. fastidiosa*, and one of the pilin subunits that is predicted to be under the control of σ-54 promoter is the functional pilin subunit of *X. fastidiosa*. Further research using both genomic and experimental approach will unravel the role of natural competence in *X. fastidiosa* disease emergence and host shift.

References

Almeida, R.P., and Purcell, A.H. 2003. Transmission of *Xylella fastidiosa* to grapevines by *Homalodisca coagulata* (Hemiptera: Cicadellidae). J Econ Entomol 96:264-271.

Almeida, R.P., and Nunney, L. 2015. How do plant diseases caused by *Xylella fastidiosa* emerge? Plant Dis 99:1457-1467.

Almeida, R.P., Nascimento, F.E., Chau, J., Prado, S.S., Tsai, C.W., Lopes, S.A., and Lopes, J.R. 2008. Genetic structure and biology of *Xylella fastidiosa* strains causing disease in citrus and coffee in Brazil. Appl Environ Microbiol 74:3690-3701.

Amanifar, N., Taghavi, M., Izadpanah, K., and Babaei, G. 2014. Isolation and pathogenicity of *Xylella fastidiosa* from grapevine and almond in Iran. Phytopathol Mediterr 53:318.

Bakkali, M. 2013. Could DNA uptake be a side effect of bacterial adhesion and twitching motility? Arch Microbiol 195:279-289.

Baltrus, D.A., and Guillemin, K. 2006. Multiple phases of competence occur during the *Helicobacter pylori* growth cycle. FEMS Microbiol Lett 255:148-155.

Baltrus, D.A., Guillemin, K., Phillips, P.C., and Travisano, M. 2008. Natural transformation increases the rate of adaptation in the human pathogen *Helicobacter pylori*. Evolution 62:39-49.

Bertolla, F., Van Gijsegem, F., Nesme, X., and Simonet, P. 1997. Conditions for natural transformation of *Ralstonia solanacearum*. Appl Environ Microbiol 63:4965-4968.

Biswas, G., Sox, T., Blackman, E., and Sparling, P. 1977. Factors affecting genetic transformation of *Neisseria gonorrhoeae*. J Bacteriol 129:983-992.

Blokesch, M., and Schoolnik, G.K. 2008. The extracellular nuclease Dns and its role in natural transformation of *Vibrio cholerae*. J Bacteriol 190:7232-7240.

Bove, J.M., and Ayres, A.J. 2007. Etiology of three recent diseases of citrus in Sao Paulo State: sudden death, variegated chlorosis and huanglongbing. IUBMB Life 59:346-354.

Brown, C., Lynch, L., and Zilberman, D. 2002. The economics of controlling insect-transmitted plant diseases. Am J Agr Econ 84:279–291.

Cameron, A.D., and Redfield, R.J. 2008. CRP binding and transcription activation at CRP-S sites. J Mol Biol 383:313-323.

Catlin, B.W. 1960. Transformation of *Neisseria meningitidis* by deoxyribonucleates from cells and from culture slime. J Bacteriol 79:579.

Chandler, M.S. 1992. The gene encoding cAMP receptor protein is required for competence development in *Haemophilus influenzae* Rd. Proc Natl Acad Sci USA 89:1626-1630.

Charpentier, X., Kay, E., Schneider, D., and Shuman, H.A. 2011. Antibiotics and UV radiation induce competence for natural transformation in *Legionella pneumophila*. J Bacteriol 193:1114-1121.

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

Coletta-Filho, H., Francisco, C., Lopes, J., Muller, C., and Almeida, R. 2016. Homologous recombination and *Xylella fastidiosa* host-pathogen associations in South America. Phytopathology.

Coupat-Goutaland, B., Bernillon, D., Guidot, A., Prior, P., Nesme, X., and Bertolla, F. 2011. *Ralstonia solanacearum* virulence increased following large interstrain gene transfers by natural transformation. Mol Plant Microbe Interact 24:497-505.

Cruz, L.F., Cobine, P.A., and De La Fuente, L. 2012. Calcium increases *Xylella fastidiosa* surface attachment, biofilm formation, and twitching motility. Appl Environ Microbiol 78:1321-1331.

Cruz, L.F., Parker, J.K., Cobine, P.A., and De La Fuente, L. 2014. Calcium-enhanced twitching motility in *Xylella fastidiosa* is linked to a single PilY1 homolog. Appl Environ Microbiol 80:7176-7185.

Dalia, A.B., Lazinski, D.W., and Camilli, A. 2014. Identification of a membrane-bound transcriptional regulator that links chitin and natural competence in *Vibrio cholerae*. MBio 5:e01028-01013.

De La Fuente, L., Burr, T.J., and Hoch, H.C. 2007. Mutations in type I and type IV pilus biosynthetic genes affect twitching motility rates in *Xylella fastidiosa*. J Bacteriol 189:7507-7510.

De La Fuente, L., Parker, J.K., Oliver, J.E., Granger, S., Brannen, P.M., van Santen, E., and Cobine, P.A. 2013. The bacterial pathogen *Xylella fastidiosa* affects the leaf ionome of plant hosts during infection. PLoS One 8:e62945.

Dorer, M.S., Fero, J., and Salama, N.R. 2010. DNA damage triggers genetic exchange in *Helicobacter pylori*. PLoS Pathog 6:e1001026.

Dorer, M.S., Sessler, T.H., and Salama, N.R. 2011. Recombination and DNA repair in *Helicobacter pylori*. Annu Rev Microbiol 65:329-348.

Dorocicz, I., Williams, P., and Redfield, R. 1993. The *Haemophilus influenzae* adenylate cyclase gene: cloning, sequence, and essential role in competence. J Bacteriol 175:7142-7149.

Dubnau, D. 1999. DNA uptake in bacteria. Annu Rev Microbiol 53:217-244.

EEPO. 2016. *Xylella fastidiosa*. EPPO Bulletin 46:407-418.

EFSA. 2016. Update of a database of host plants of *Xylella fastidiosa*: 20 November 2015. EFSA J 14:4378-n/a.

Elkins, C., Thomas, C., Seifert, H., and Sparling, P. 1991. Species-specific uptake of DNA by gonococci is mediated by a 10-base-pair sequence. J Bacteriol 173:3911-3913.

Griffith, F. 1928. The significance of pneumococcal types. J Hyg 27:113-159.

Güldür, M., Çağlar, B., Castellano, M., Ünlü, L., Güran, S., Yilmaz, M., and Martelli, G. 2005. First report of almond leaf scorch in Turkey. J Plant Pathol:246-246.

Håvarstein, L.S., Coomaraswamy, G., and Morrison, D.A. 1995. An unmodified heptadecapeptide pheromone induces competence for genetic transformation in *Streptococcus pneumoniae*. Proc Natl Acad Sci USA 92:11140-11144.

Herriott, R., Meyer, E., and Vogt, M. 1970. Defined nongrowth media for stage II development of competence in *Haemophilus influenzae*. J Bacteriol 101:517-524.

Hewitt, W.B. 1958. The Probable Home of Pierce's Disease Virus. Am J Enol Vitic 9:94-98.

Hobbs, M., and Mattick, J.S. 1993. Common components in the assembly of type 4 fimbriae, DNA transfer systems, filamentous phage and protein-secretion apparatus: a general system for the formation of surface-associated protein complexes. Mol Microbiol 10:233-243.

Hopkins, D.L., and Purcell, A.H. 2002. *Xylella fastidiosa*: Cause of Pierce's Disease of Grapevine and Other Emergent Diseases. Plant Dis 86:1056-1066.

Hülter, N., Sørum, V., Borch-Pedersen, K., Liljegren, M.M., Utnes, A.L., Primicerio, R., Harms, K., and Johnsen, P.J. 2017. Costs and benefits of natural transformation in *Acinetobacter baylyi*. BMC Microbiol 17:34.

Humbert, O., and Salama, N.R. 2008. The *Helicobacter pylori* HpyAXII restriction–modification system limits exogenous DNA uptake by targeting GTAC sites but shows asymmetric conservation of the DNA methyltransferase and restriction endonuclease components. Nucleic Acids Res 36:6893-6906.

Israel, D.A., Lou, A.S., and Blaser, M.J. 2000. Characteristics of *Helicobacter pylori* natural transformation. FEMS Microbiol Lett 186:275-280.

Karnholz, A., Hoefler, C., Odenbreit, S., Fischer, W., Hofreuter, D., and Haas, R. 2006. Functional and topological characterization of novel components of the comB DNA transformation competence system in *Helicobacter pylori*. J Bacteriol 188:882-893.

Karudapuram, S., and Barcak, G.J. 1997. The *Haemophilus influenzae* dprABC genes constitute a competence-inducible operon that requires the product of the tfoX (sxy) gene for transcriptional activation. J Bacteriol 179:4815-4820.

Kung, S.H., and Almeida, R.P. 2011. Natural competence and recombination in the plant pathogen *Xylella fastidiosa*. Appl Environ Microbiol 77:5278-5284.

Kung, S.H., and Almeida, R.P. 2014. Biological and genetic factors regulating natural competence in a bacterial plant pathogen. Microbiology 160:37-46.

Kung, S.H., Retchless, A.C., Kwan, J.Y., and Almeida, R.P. 2013. Effects of DNA size on transformation and recombination efficiencies in *Xylella fastidiosa*. Appl Environ Microbiol 79:1712-1717.

Lenhart, J.S., Schroeder, J.W., Walsh, B.W., and Simmons, L.A. 2012. DNA repair and genome maintenance in *Bacillus subtilis*. Microbiol Mol Biol Rev 76:530-564.

Lorenz, M.G., and Wackernagel, W. 1994. Bacterial gene transfer by natural genetic transformation in the environment. Microbiol Rev 58:563-602.

Macfadyen, L.P., Chen, D., Vo, H.C., Liao, D., Sinotte, R., and Redfield, R.J. 2001. Competence development by *Haemophilus influenzae* is regulated by the availability of nucleic acid precursors. Mol Microbiol 40:700-707.

Marcelletti, S., and Scortichini, M. 2016. Xylella fastidiosa CoDiRO strain associated with the olive quick decline syndrome in southern Italy belongs to a clonal complex of the subspecies pauca that evolved in Central America. Microbiology 162:2087-2098.

Meibom, K.L., Blokesch, M., Dolganov, N.A., Wu, C.-Y., and Schoolnik, G.K. 2005. Chitin induces natural competence in *Vibrio cholerae*. Science 310:1824-1827.

Meng, Y., Li, Y., Galvani, C.D., Hao, G., Turner, J.N., Burr, T.J., and Hoch, H.C. 2005. Upstream migration of *Xylella fastidiosa* via pilus-driven twitching motility. J Bacteriol 187:5560-5567.

Newman, K.L., Almeida, R.P., Purcell, A.H., and Lindow, S.E. 2003. Use of a green fluorescent strain for analysis of *Xylella fastidiosa* colonization of *Vitis vinifera*. Appl Environ Microbiol 69:7319-7327.

Newman, K.L., Almeida, R.P., Purcell, A.H., and Lindow, S.E. 2004. Cell-cell signaling controls *Xylella fastidiosa* interactions with both insects and plants. Proc Natl Acad Sci USA 101:1737-1742.

Nunney, L., Yuan, X., Bromley, R.E., and Stouthamer, R. 2012. Detecting genetic introgression: high levels of intersubspecific recombination found in *Xylella fastidiosa* in Brazil. Appl Environ Microbiol 78:4702-4714.

Nunney, L., Hopkins, D.L., Morano, L.D., Russell, S.E., and Stouthamer, R. 2014a. Intersubspecific recombination in *Xylella fastidiosa* strains native to the United States: infection of novel hosts associated with an unsuccessful invasion. Appl Environ Microbiol 80:1159-1169.

Nunney, L., Schuenzel, E.L., Scally, M., Bromley, R.E., and Stouthamer, R. 2014b. Large-scale intersubspecific recombination in the plant-pathogenic bacterium *Xylella fastidiosa* is associated with the host shift to mulberry. Appl Environ Microbiol 80:3025-3033.

Nunney, L., Yuan, X., Bromley, R., Hartung, J., Montero-Astua, M., Moreira, L., Ortiz, B., and Stouthamer, R. 2010. Population genomic analysis of a bacterial plant pathogen: novel insight into the origin of Pierce's disease of grapevine in the U.S. PLoS One 5:e15488.

Oliver, J.E., Cobine, P.A., and De La Fuente, L. 2015. *Xylella fastidiosa* isolates from both subsp. *multiplex* and *fastidiosa* cause disease on southern highbush blueberry (*Vaccinium sp.*) under greenhouse conditions. Phytopathology 105:855-862.

Oliver, J.E., Sefick, S.A., Parker, J.K., Arnold, T., Cobine, P.A., and De La Fuente, L. 2014. Ionome changes in *Xylella fastidiosa*-infected *Nicotiana tabacum* correlate with virulence and discriminate between subspecies of bacterial isolates. Mol Plant Microbe Interact 27:1048-1058.

Palmen, R., Buijsman, P., and Hellingwerf, K.J. 1994. Physiological regulation of competence induction for natural transformation in *Acinetobacter calcoaceticus*. Arch Microbiol 162:344-351.

Parker, J.K., Havird, J.C., and De La Fuente, L. 2012. Differentiation of *Xylella fastidiosa* strains via multilocus sequence analysis of environmentally mediated genes (MLSA-E). Appl Environ Microbiol 78:1385-1396.

Parker, J.K., Chen, H., McCarty, S.E., Liu, L.Y., and De La Fuente, L. 2016. Calcium transcriptionally regulates the biofilm machinery of *Xylella fastidiosa* to promote continued biofilm development in batch cultures. Environ Microbiol 18:1620-1634.

Perez-Donoso, A.G., Sun, Q., Roper, M.C., Greve, L.C., Kirkpatrick, B., and Labavitch, J.M. 2010. Cell wall-degrading enzymes enlarge the pore size of intervessel pit membranes in healthy and *Xylella fastidiosa*-infected grapevines. Plant Physiol 152:1748-1759.

Pierce, N. 1892. The California vine disease. US Dept Agric Div Veg Pathol Bull 2:22.

Prudhomme, M., Attaiech, L., Sanchez, G., Martin, B., and Claverys, J.-P. 2006. Antibiotic stress induces genetic transformability in the human pathogen *Streptococcus pneumoniae*. Science 313:89-92.

Redfield, R.J. 1991. sxy-1, a *Haemophilus influenzae* mutation causing greatly enhanced spontaneous competence. J Bacteriol 173:5612-5618.

Redfield, R.J., Cameron, A.D., Qian, Q., Hinds, J., Ali, T.R., Kroll, J.S., and Langford, P.R. 2005. A novel CRP-dependent regulon controls expression of competence genes in *Haemophilus influenzae*. J Mol Biol 347:735-747.

Retchless, A.C., Labroussaa, F., Shapiro, L., Stenger, D.C., Lindow, S.E., and Almeida, R.P.P. (2014). Genomic insights into *Xylella fastidiosa* interactions with plant and insect hosts. In Genomics of Plant-Associated Bacteria, L.-P.A. Gross DC, and Kole C, ed (Springer Berlin Heidelberg, Berlin, Heidelberg), pp. 177-202.

Roper, M.C., Greve, L.C., Warren, J.G., Labavitch, J.M., and Kirkpatrick, B.C. 2007. *Xylella fastidiosa* requires polygalacturonase for colonization and pathogenicity in *Vitis vinifera* grapevines. Mol Plant Microbe Interact 20:411-419.

Saponari, M., Boscia, D., Nigro, F., and Martelli, G.P. 2013. Identification of DNA sequences related to *Xylella fastidiosa i*n Oleander, Almond and Olive trees exhibiting leaf scorch symptoms in Apulia (Southern Italy). J Plant Pathol 95:668-668.

Scally, M., Schuenzel, E.L., Stouthamer, R., and Nunney, L. 2005. Multilocus sequence type system for the plant pathogen *Xylella fastidiosa* and relative contributions of recombination and point mutation to clonal diversity. Appl Environ Microbiol 71:8491-8499.

Schaad, N.W., Postnikova, E., Lacy, G., Fatmi, M., and Chang, C.J. 2004. *Xylella fastidiosa* subspecies: *X. fastidiosa* subsp. [correction] *fastidiosa* [correction] subsp. nov., X. *fastidiosa* subsp. *multiplex* subsp. nov., and *X. fastidiosa* subsp. *pauca* subsp. nov. Syst Appl Microbiol 27:290-300.

Scrudato, M.L., and Blokesch, M. 2012. The regulatory network of natural competence and transformation of *Vibrio cholerae*. PLoS Genet 8:e1002778.

Scrudato, M.L., and Blokesch, M. 2013. A transcriptional regulator linking quorum sensing and chitin induction to render *Vibrio cholerae* naturally transformable. Nucleic Acids Res:gkt041.

Seitz, P., and Blokesch, M. 2013. Cues and regulatory pathways involved in natural competence and transformation in pathogenic and environmental Gram-negative bacteria. FEMS Microbiol Rev 37:336-363.

Seper, A., Fengler, V.H., Roier, S., Wolinski, H., Kohlwein, S.D., Bishop, A.L., Camilli, A., Reidl, J., and Schild, S. 2011. Extracellular nucleases and extracellular DNA play important roles in *Vibrio cholerae* biofilm formation. Mol Microbiol 82:1015-1037.

Smith, H.O., Gwinn, M.L., and Salzberg, S.L. 1999. DNA uptake signal sequences in naturally transformable bacteria. Res Microbiol 150:603-616.

Strona, G., Carstens, C.J., and Beck, P.S. 2017. Network analysis reveals why *Xylella fastidiosa* will persist in Europe. Sci Rep 7:71.

Su, C.C., Deng, W.L., Jan, F.J., Chang, C.J., Huang, H., Shih, H.T., and Chen, J. 2016. *Xylella taiwanensis* sp. nov., causing pear leaf scorch disease. Int J Syst Evol Microbiol 66:4766-4771.

Sun, Q., Sun, Y., Walker, M.A., and Labavitch, J.M. 2013. Vascular occlusions in grapevines with Pierce's disease make disease symptom development worse. Plant Physiol 161:1529-1541.

Van Sluys, M.A., de Oliveira, M.C., Monteiro-Vitorello, C.B., Miyaki, C.Y., Furlan, L.R., Camargo, L.E., da Silva, A.C., Moon, D.H., Takita, M.A., Lemos, E.G., Machado, M.A., Ferro, M.I., da Silva, F.R., Goldman, M.H., Goldman, G.H., Lemos, M.V., El-Dorry, H., Tsai, S.M., Carrer, H., Carraro, D.M., de Oliveira, R.C., Nunes, L.R., Siqueira, W.J., Coutinho, L.L., Kimura, E.T., Ferro, E.S., Harakava, R., Kuramae, E.E., Marino, C.L., Giglioti, E., Abreu, I.L., Alves, L.M., do Amaral, A.M., Baia, G.S., Blanco, S.R., Brito, M.S., Cannavan, F.S., Celestino, A.V., da Cunha, A.F., Fenille, R.C., Ferro, J.A., Formighieri, E.F., Kishi, L.T., Leoni, S.G., Oliveira, A.R., Rosa, V.E., Jr., Sassaki, F.T., Sena, J.A., de Souza, A.A., Truffi, D., Tsukumo, F., Yanai, G.M., Zaros, L.G., Civerolo, E.L., Simpson, A.J., Almeida, N.F., Jr., Setubal, J.C., and Kitajima, J.P. 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.

Wells, J.M., Raju, B.C., Hung, H.Y., Weisburg, W.G., Mandelco-Paul, L., and Brenner, D.J. 1987. *Xylella fastidiosa* gen. nov., sp. nov.: Gram-negative, xylem-limited, fastidious plant bacteria related to *Xanthomonas* spp. Int J Syst Bacteriol 37:136-143.

Wise, E.M., Alexander, S.P., and Powers, M. 1973. Adenosine 3′: 5′-cyclic monophosphate as a regulator of bacterial transformation. Proc Natl Acad Sci USA 70:471-474.

Yamamoto, S., Izumiya, H., Mitobe, J., Morita, M., Arakawa, E., Ohnishi, M., and Watanabe, H. 2011. Identification of a chitin-induced small RNA that regulates translation of the tfoX gene, encoding a positive regulator of natural competence in *Vibrio cholerae*. J Bacteriol 193:1953-1965.

Yuan, X., Morano, L., Bromley, R., Spring-Pearson, S., Stouthamer, R., and Nunney, L. 2010. Multilocus sequence typing of *Xylella fastidiosa* causing Pierce's disease and oleander leaf scorch in the United States. Phytopathology 100:601-611.

Zulty, J.J., and Barcak, G.J. 1995. Identification of a DNA transformation gene required for com101A+ expression and supertransformer phenotype in *Haemophilus influenzae*. Proc Natl Acad Sci USA 92:3616-3620.

Chapter 2

Natural competence of *Xylella fastidiosa* occurs at a high frequency inside microfluidic chambers mimicking the bacterium's natural habitats

Abstract

Xylella fastidiosa is a xylem-limited bacterium that is the causal agent of emerging diseases in several economically important crops. Genetic diversity studies have demonstrated homologous recombination occurring among *X. fastidiosa* strains, which has been proposed to contribute to host plant shifts. Moreover, experimental evidence confirmed that *X. fastidiosa* is naturally competent for recombination *in vitro*. Here, as an approximation of natural habitats (plant xylem vessels and insect mouthparts), recombination was studied in microfluidic chambers (MC) filled with medium amended with grapevine xylem sap. First, different media were screened for recombination in solid agar plates using a pair of *X. fastidiosa* strains that were previously reported to recombine in co-culture. Highest frequency of recombination was obtained with PD3 medium, when compared to other two media (XFM, PW) used in previous studies. Dissection of media components lead to identification of bovine serum albumin as an inhibitor of recombination that was correlated to its previously known effect on inhibition of twitching motility. When recombination was performed in liquid culture, frequencies were significantly higher under flow (MC) as compared to batch conditions (test tubes). Recombination frequencies in MC and agar plate conditions were not significantly different from each other. Grapevine xylem sap from both susceptible and tolerant varieties allowed high recombination frequency in MC when mixed with PD3. These results suggest that *X. fastidiosa* can become naturally

competent in the natural growth environment of liquid flow, and this phenomenon could have implications in *X. fastidiosa* environmental adaptation.

Introduction

Natural competence is a phenomenon that allows bacteria to take up DNA segments from the environment and incorporate into the genome via homologous recombination (Lorenz and Wackernagel, 1994). Natural competence was first demonstrated in *Streptococcus pneumoniae* in 1928 by Frederick Griffith (1928). He showed that virulence genes were transferred from donor to recipient cells, converting the non-virulent recipients into virulent pathogens (Griffith, 1928). Since then, ~80 bacterial species in divergent phyla have been described as naturally competent (Johnston et al., 2014). Although the exact reasons for occurrence of natural competence in bacteria still remain unknown, studies showed that natural competence is induced under conditions of starvation (Herriott et al., 1970; Lorenz and Wackernagel, 1991; Bertolla et al., 1997a) and DNA damage (Dorer et al., 2010), and has been hypothesized that the incoming DNA serves as a food source and DNA repair material. Another proposition is that natural competence allows acquisition of new genes and alleles providing the recipient cells with adaptive advantages. In fact, a previous study showed increased rate of adaptation by natural competence in *Helicobacter pylori* (Baltrus et al., 2008). Interestingly, natural competence has been demonstrated in some of the most highly diverse and successful human pathogens such as *H. pylori* (Hofreuter et al., 1998; Humbert et al., 2011), *Neisseria meningitidis* and *N. gonorrhoeae* (Sparling, 1966; Hamilton and Dillard, 2006), and *Porphyromonas gingivalis* (Tribble et al., 2012), which require rapid adaptation to evade the immune response. Furthermore, natural competence also was described in two plant pathogens, *Ralstonia*

solanacearum (Bertolla et al., 1999) and *Xylella fastidiosa* (Kung and Almeida, 2011), both of which have very broad plant host ranges.

Xylella fastidiosa is a bacterial pathogen affecting many economically important crops such as grapevines, citrus, coffee, peach, and almond (Hopkins and Purcell, 2002). The disease process is not completely understood, but it is proposed that *X. fastidiosa* forms biofilm-like aggregates and blocks xylem vessels, the conduits for water and nutrient transport in the plants (Chatterjee et al., 2008). This blockage hinders xylem sap flow and starves the upper aerial parts of water and mineral nutrients, producing symptoms that resemble those of water and nutrient deficit. *X. fastidiosa* is transmitted by a number of xylem sap-feeding insects including sharpshooter leaf hoppers and spittlebugs in which *X. fastidiosa* forms biofilms in the foregut (Hill and Purcell, 1995; Redak et al., 2004). Taxonomically, *X. fastidiosa* is divided into five subspecies based on multi-locus sequence typing (MLST) (Scally et al., 2005; Nunney et al., 2014b). Even within subspecies, host range and genotype diversity has been described (Almeida et al., 2008; Parker et al., 2012), and recombination events among strains have been detected among field collected samples (Nunes et al., 2003; Almeida et al., 2008). In fact, homologous recombination was shown to have a greater effect in generating genetic diversity in *X. fastidiosa* than point mutation (Scally et al., 2005). Recent outbreaks of *X. fastidiosa* diseases in Europe (Saponari et al., 2013) and Asia (Leu and Su, 1993; Su et al., 2013) and also in new plant hosts such as olives (Saponari et al., 2013), blueberry (Chang et al., 2009), and pear (Leu and Su, 1993) suggest great adaptation potential of this pathogen.

In a number of plant species, *X. fastidiosa* is believed to live as a harmless endophyte without inducing disease symptoms (Newman et al., 2003; Chatterjee et al., 2008). Coexistence in the same xylem system of different strains for long time without killing the host, represents a fertile environment for exchange of DNA material. Several MLST-based studies detected inter-subspecific recombination among strains of *X. fastidiosa* and proposed recombination as the mechanism of new allele acquisition leading to plant host shift and disease emergence. Inter-subspecific recombination was described to generate strains that infect citrus and coffee (Nunney et al., 2012); mulberry (Nunney et al., 2014b); and blueberry and blackberry (Nunney et al., 2014a). A recent study also showed inter-subspecific recombination between coffee-infecting strains in plants intercepted in France (Jacques et al., 2015). Natural competence could be an explanation for the frequent recombination events detected in *X. fastidiosa*.

Natural competence in *X. fastidiosa* was recently described *in vitro* (Kung and Almeida, 2011), and the rate of homologous recombination was shown to be higher when the cells were growing exponentially in solid agar plates than in batch culture tubes, and minimal medium was more conducive than rich medium (Kung and Almeida, 2014). Using a plasmid as a donor DNA, a 96 bp of flanking homology was sufficient to initiate recombination (Kung et al., 2013). Moreover, some competence-related and type IV pili genes were shown to be involved in the process (Kung and Almeida, 2014). Although some of those studies were performed using plasmids as donor DNA, two strains were also shown to recombine in co-culture conditions (Kung and Almeida, 2011), although the capacity of these strains to act either as donor or recipient for DNA exchange was not determined in those studies.

The objective of this study was to test the hypothesis that natural competence in *X. fastidiosa* occurs under flow conditions (an approximation of the natural habitat of this bacterium). Associated with this objective was the aim to elucidate if previous observations of high frequency of *X. fastidiosa* natural competence *in vitro* (Kung and Almeida, 2011; Kung et al., 2013; Kung and Almeida, 2014) were dependent on batch culture conditions (both test tubes and agar plates), which allow cell-to-cell contact for longer time without replenishing of nutrients or removal of secreted molecules. Although natural competence and recombination is assumed to occur in natural habitats based on field surveys and DNA sequence data, experimental indication of its occurrence in the plant or insect host is not yet available for *X. fastidiosa*. Therefore, to circumvent the limitation of *X. fastidiosa* recombination tests in the natural hosts that are affected by uneven bacterial distribution and low populations (Newman et al., 2003; Das et al., 2015), we performed recombination experiments in a microfluidic chamber (MC) system that mimics the natural environment of xylem vessels and insect foreguts. MC allows continuous media flow conditions and formation of biofilms, and has been previously used to study the behavior of *X. fastidiosa* (De La Fuente et al., 2007; De La Fuente et al., 2008; Cruz et al., 2012; Navarrete and De La Fuente, 2014). Biofilm fraction of MC (referred hereafter as MC_in) and planktonic and detached cell fraction (referred hereafter as MC_out) can be collected separately and the behavior of cells in the two fractions determined. Two strains used in all the previous publications on this topic (Kung and Almeida, 2011; Kung and Almeida, 2014), were used in this current study to facilitate comparison with literature and to further our understanding of natural competence in *X. fastidiosa*. The results presented here show that growth under flow conditions support natural competence in *X. fastidiosa,* with recombination frequencies equivalent to that on solid media, previously described to be the most conducive environment for natural competence

in vitro (Kung and Almeida, 2014). These findings support the hypothesis that recombination occurs at high rates under flow conditions, representing natural habitats of *X. fastidiosa*.

Materials and Methods

Bacterial strains, media, and culture conditions

X. fastidiosa subspecies *fastidiosa* mutants NS1-CmR [a mutant of wild type strain Temecula1 in which a chloramphenicol resistant cassette was inserted in a non-coding region using the suicide plasmid pAX1-Cm (Matsumoto et al., 2009)], and *pglA*-KmR [a kanamycin resistant mutant of the Fetzer strain in which the gene encoding polygalacturonase is disrupted (Roper et al., 2007)], were used in this study. The mutants were cultured in periwinkle wilt (PW) agar medium (Davis et al., 1980) modified by omitting phenol red and adding 1.8 g liter^{-1} bovine serum albumin (BSA, Gibco Life Science Technology), and supplemented with the respective antibiotics. PD3 (Davis et al., 1981) and modified XFM (Kung and Almeida, 2011) were used when stated. Pectin was added to a final concentration of 0.01% as previously described (Killiny and Almeida, 2009). Kanamycin was used at 30 µg ml^{-1} and chloramphenicol at 10 µg ml^{-1}. Inocula were prepared by streaking cultures from the -80°C freezer stocks on PW agar plates and incubating the plates for 5-7 days at 28°C. Cultures were then re-streaked onto new plates and incubated for another 5-7 days before use.

Media selection for natural competence and growth in microfluidic chambers

To select a medium to test occurrence of natural competence in MC, three media (PW, modified XFM, and PD3) were first tested in solid agar plates. XFM and PW, used in previous studies (Kung and Almeida, 2011; Kung and Almeida, 2014), were selected as positive and negative

control media for recombination, respectively. Natural competence experiments were performed according to Kung et al (Kung and Almeida, 2011; Kung and Almeida, 2014) with some modifications. Briefly, cells (OD_{600} = 0.25) of NS1-CmR and *pglA*-KmR mutants were prepared in liquid media by scraping the cultures from PW antibiotics plates. Ten µl of each strain were spotted on top of each other on the agar plates of PW, XFM, and PD3 without antibiotics and the spots were allowed to dry for ~1 hour. The plates were then incubated at 28°C for ~3 days. Two spots from the same plate were then scraped off and suspended in 1 ml of PD3 to make one replication, and 3-4 replicates were included for each media type per experiment. The experiments were repeated independently twice for XFM and at least three times for PD3 and PW. Single mutant strains (donor and recipient) were included as controls. The suspensions were then serially diluted and 100 µl of appropriate dilutions were plated on PW agar plates in triplicate supplemented with both antibiotics (Km and Cm) to recover recombinants at the antibiotic resistant site, and with single antibiotics (Km or Cm) to check the growth of both parents in the mixture. Appropriate dilutions also were plated onto PW plates without antibiotics for enumeration of total viable cells. Plates were incubated at 28°C for at least 14 days before CFUs were enumerated. The recombination frequency at the antibiotic resistant site was calculated as the ratio of recombinant CFUs (CFUs from both antibiotics plates) to total CFUs (CFUs from appropriate no antibiotics plates) in equal volumes of suspension. After selecting the media that supported recombination in the agar plates, media (PD3 and XFM) were tested in the MC for cell attachment and biofilm formation.

Media components influencing natural competence and twitching motility

To test specific components that may influence natural competence, an initial screen was performed by removing or adding components to PW and PD3 media in solid agar plates as described above. The components tested were sodium citrate dehydrate, succinic acid, and starch (present only in PD3); and BSA and L-glutamine (present only in PW) and Pectin. The effect of BSA was further tested by supplementing PD3 and PD3+L-glutamine with BSA, and removing BSA from PW and XFM. Experiments were repeated three times independently with three replicates each time except for PD3+L-glutamine treatment that was performed once with three replicates. Twitching motility of both mutants was determined in media with and without BSA according to previous studies (Galvani et al., 2007; Cruz et al., 2014) with few modifications. Briefly, for PD3 and PW with and without BSA, media plates solidified with agar or Gelrite (rpi Research Products and International Corp, Mount Prospect, IL) were divided into two halves and 10-12 spots of each mutant strain were made using a sterile toothpick and plates were incubated at 28°C for 4-5 days. For XFM with and without BSA, plates solidified with agar were used and incubated for 10-12 days before measurements were recorded. Colony peripheral fringe was observed under 10x magnification using a Nikon Eclipse Ti Inverted Microscope (NIKON, Melville, NY), and fringe width was measured for six colonies per plate per strain, with at least seven measurements per colony using the Nikon DS-Q1 Digital Camera (Nikon, Melville, NY) connected to a Nikon Eclipse Ti Inverted Microscope (NIKON, Melville, NY), and controlled by NIS Elements imaging software version 3.0. Twitching experiments were performed at least three times independently for PD3 and PW with and without BSA and once for XFM with and without BSA.

Natural competence in different growth conditions

Three growth conditions were used; solid agar plates (plate, as described above), liquid culture tubes (tube), and continuous liquid flow (MC). PD3 without antibiotics was the medium used, and the initial inoculum of NS1-CmR and *pglA*-KmR mutants were prepared as described above.

I. Competence in tubes

Twenty-five ml glass test tubes containing three ml of PD3 were inoculated with 100 µl of each of the OD-adjusted strain suspensions as donor and recipient cells (final OD_{600} = 0.01). Tubes containing single strain inoculations were included as control treatments. Tubes were then incubated with shaking (180 rpm at 28˚C). After ~3 days the tubes were vortexed well to mix the biofilm formed on the air-liquid interface with the rest of the suspension, and serially diluted and plated as described above. Three independent experiments were performed and three replications were included in each experiment (n=9 in total).

II. Competence in MC

MCs were prepared as previously described (De La Fuente et al., 2007). Briefly, two parallel channels with separate inlets for bacterial cells and growing media were etched on a silicon wafer. The channels were then modeled into a polydimethyl siloxane (PDMS) and sandwiched between the PDMS layer and a glass cover slide. The inlets and outlets were then connected to tubings that were connected to syringes (Supplemental fig. 2-1). The syringes were connected to pumps which control media flow rate in the MC. The MC was mounted onto a Nikon Eclipse Ti Inverted Microscope (NIKON, Melville, NY) to observe cell attachment and biofilm formation using phase contrast and Nomarski Differential Interference Contrast optics. Time lapse video

was taken using a Nikon DS-Q1 Digital Camera (Nikon, Melville, NY) connected to the microscope and controlled by NIS Elements imaging software version 3.0.

For preparing the inocula for MCs, equal volumes of the strain pairs (OD_{600} = 0.25) were mixed and inoculated into the cell inlet syringes, and growing media (PD3) was injected to the media syringes. MCs were run for 5-7 days with a media flow rate of 0.25 µl min^{-1} until abundant growth of biofilm was observed. At the end of the experiment, the fraction of cells collected in the outlet syringe (MC_out) was harvested and the fraction formed inside the channels (MC_in) was detached and pushed to the outlet collection syringe by increasing the flow rate to 30-40 µl min^{-1}. Serial dilution, plating, CFU counts, and frequency of recombination calculations were done as described above. Four independent experiments were performed with seven replicates in total (n=7).

III. Competence in MC with grapevine sap

Grapevine (*Vitis vinifera*) sap was collected from a *X. fastidiosa* susceptible variety (Chardonnay) in Dahlonega, GA; and a tolerant variety (Blanc Du Bois) in Tallahassee, FL at the end of the dormant season (March/April). A new season cane was pruned and sap was collected in a 50-ml conical tube, which was stored in ice until it was brought back to the lab. Xylem sap was sterilized by filtering with 0.22 µm vacuum filter and stored at -80°C until used. Sap experiments were performed in the MC with both pure sap and 50% sap mixed in PD3 (v/v). Natural competence assays were same as with the MC experiment with PD3. Experiments were repeated at least three times for both sap types.

Natural competence with heat-killed donor cells and confirmation of homologous recombination

Confirmation of homologous recombination occurring via natural competence was performed by using heat-killed donor cells in the solid agar plates. Suspension of the donor cells (either NS1-CmR or *pglA*-KmR mutants) were incubated at 90°C for 15 min for heat killing. Complete killing was confirmed by plating an aliquot onto PW plates. The heat-killed donor and live recipients were then spotted on PD3 plates as described above. For confirmation of homologous recombination at the desired genomic region, randomly selected recombinant CFUs were re-streaked onto new double antibiotic PW plates, and colony PCR was performed using the primers targeting the flanking region of the construct used to generate the mutants according to Kung et al (Kung and Almeida, 2011). Sequences of the flanking regions of antibiotic cassettes insertion sites between the parent strains (NCBI Accession numbers KU873007-KU873014) were compared using the Muscle Pairwise Alignment algorithm within the Geneious 9.0.3 platform (Kearse et al., 2012).

Statistical Analysis

The number of recombinants, total CFUs, and recombination frequency data were analyzed in PROC GLIMMIX (SAS 9.3) which fits statistical models to data with non-normal distribution and non-constant variability. For the analysis of frequency, response distribution was used as the binomial distribution of number of recombinants/total cells. Least square differences of means among the treatments were separated by Tukey's HSD (honestly significance difference) test at the significance level of $P \leq 0.05$. For the repeated experiments, time factor was used as a

random variable. The fringe widths of bacterial colonies among different media also were compared using PROC GLIMMIX in SAS.

Results

Growth media influence natural competence in *X. fastidiosa*

Cell suspensions of the two strains were mixed together on agar plates and recombination was assessed by the acquisition of antibiotic resistance markers.

Results in solid agar plates of PW, XFM, and PD3 showed that PD3 is more conducive for recombination than other media (Fig. 2-1, Supplemental table 2-1) followed by XFM. Recombination frequency was $1.9\pm0.4 \times 10^{-5}$ (1.9 recombinants per ~10^5 cells) for PD3, $2.4\pm1.3 \times 10^{-7}$ for XFM, and $8.3\pm7.1 \times 10^{-9}$ for PW. Recombinants were readily recovered in PD3 (100-3000 recombinant CFUs $100\mu l^{-1}$). In XFM, the number varied from (0-12 recombinant CFUs $100\mu l^{-1}$). In PW only one recombinant was recovered of all the experiments performed and hence, the recombination frequency of 8.3×10^{-9} per recipient cell was set as the detection limit of the method. The total CFU counts were not statistically different among the three media (P = 0.22).

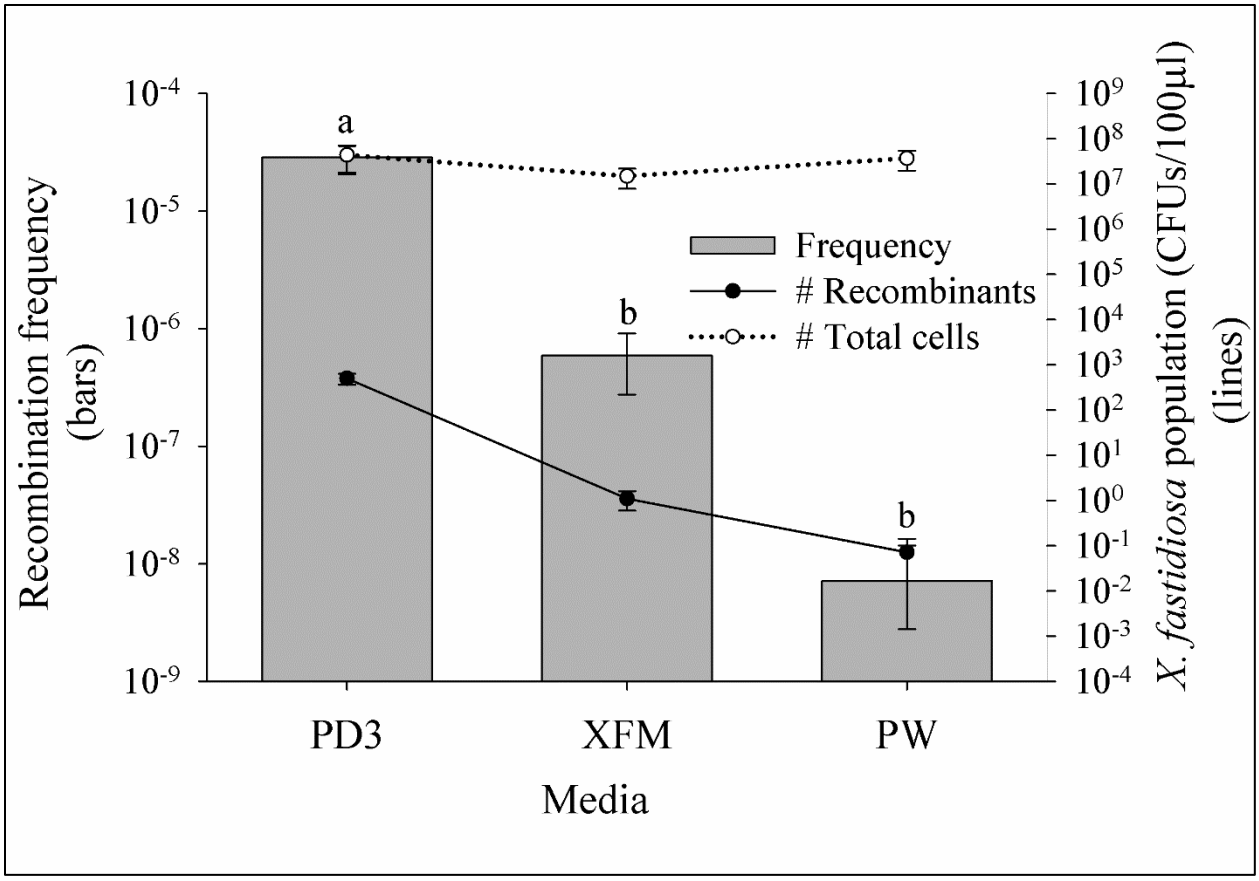


Fig. 2-1. Recombination in various growth media. Recombination frequencies, number of recombinants, and total CFU counts in different culture media in solid agar plates are shown. Cell suspensions of the two strains (NS1-CmR and pglA-KmR) were mixed together on agar plates and recombination was assessed by the acquisition of antibiotic resistance markers. Data represent means and standard errors from different experiments. Experiments were repeated independently twice for XFM (n=7) and at least three times for PD3 and PW (n=9). Different letters above the bars represent significant difference in the recombination frequency ($P < 0.05$) as analyzed with PROC GLIMMIX in SAS followed by least square mean comparisons by Tukey's HSD.

BSA impairs natural competence and twitching motility

Components that are only present in PD3 (sodium citrate dehydrate, succinic acid, and starch) and PW (BSA, L-glutamine) were initially screened for influence in natural competence as PD3 produced significantly higher number of recombinants than PW. Among the components tested, only BSA had a clear effect, as recombinants could not be recovered from treatments in both PW and PW-L-glutamine, both of which contain BSA, but could be recovered from both media when BSA was removed (Supplemental Table 2-1). Further test of the effect of BSA was performed in four media (PD3, PD3+L-glutamine, PW, and XFM) with and without BSA by conducting additional experiments. Results showed that BSA significantly reduced frequency of recombination when added to PD3 ($P < 0.001$) (Fig. 2-2) and PD3+L-glutamine ($P < 0.001$) (Supplemental table 2-1), and increased frequency when removed from PW ($P=0.0025$) and XFM ($P < 0.0001$) (Fig. 2-2, Supplemental table 2-1). Recombinants were readily recovered in PD3 (323±72 $100\mu l^{-1}$), and PD3+L-glutamine (252±30 $100\mu l^{-1}$), but were significantly reduced when supplemented with BSA (4 ±0.8 and 4 ±1.0 $100\mu l^{-1}$, respectively). Similarly, the number of recombinants increased in PW from 0 to 8.7±2.6 $100\mu l^{-1}$ and in XFM from 2.0±0.8 to 8.6±2.1 $100\mu l^{-1}$ on removing BSA (Supplemental table 2-1). Pectin supplementation to PD3 had no significant effect in the recombination frequency ($P=0.79$) (Supplemental table 2-1).

Twitching motility, as measured by colony fringe width, was higher in PD3 than in other media (Fig. 2-3). BSA significantly reduced twitching motility when supplemented to PD3 and significantly increased twitching motility when removed from both PW and XFM with the NS1-CmR mutant (Fig. 2-3). However, no fringe could be detected with the *pglA*-KmR mutant in any of the media tested (Fig. 2-3).

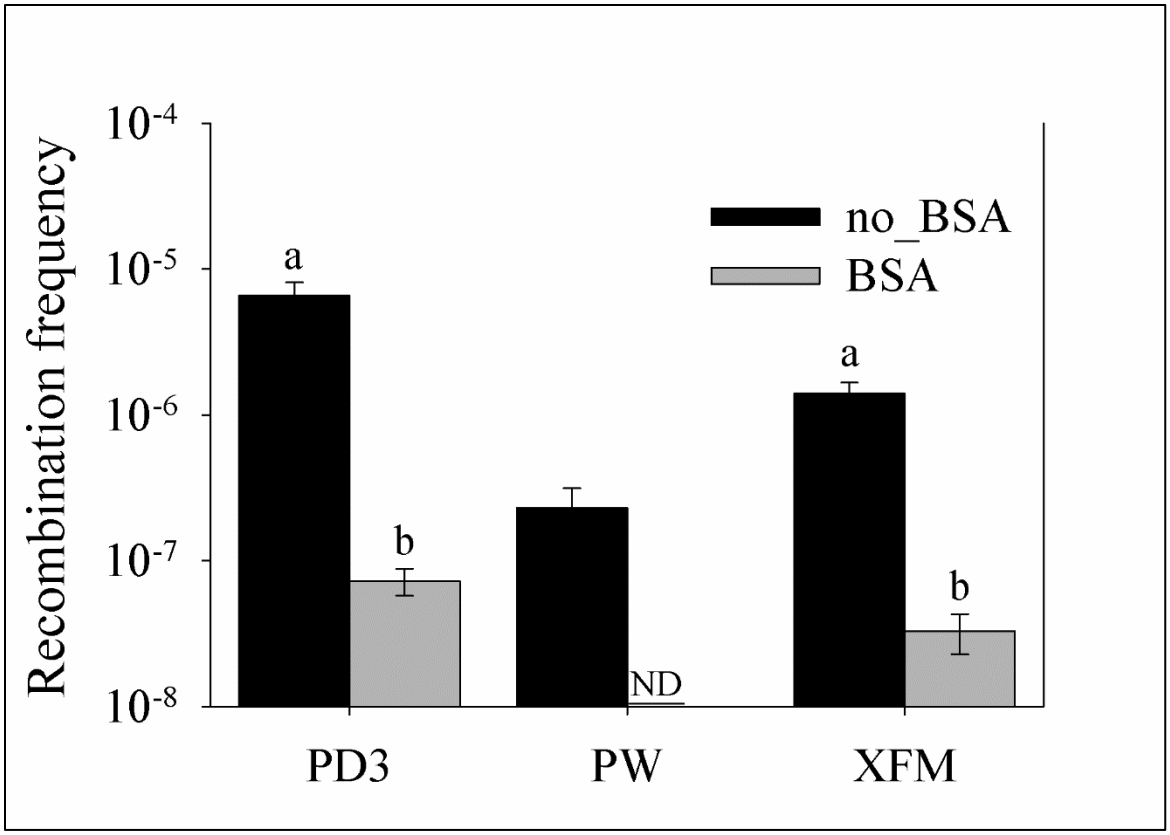


Fig. 2-2. Effect of BSA on recombination. Media used previously (Fig. 2-1) was modified by addition (PD3) or removal (PW, XFM) of BSA, and new set of experiments were conducted to measure recombination. Experiments were repeated three times with 2-4 replications (n≥8). Data represent means and standard errors from different experiments. BSA significantly reduced recombination frequency in each of the media. Different letters represent significant difference ($P < 0.05$) within the medium as analyzed with PROC GLIMMIX in SAS. ND: frequency below detection limit.

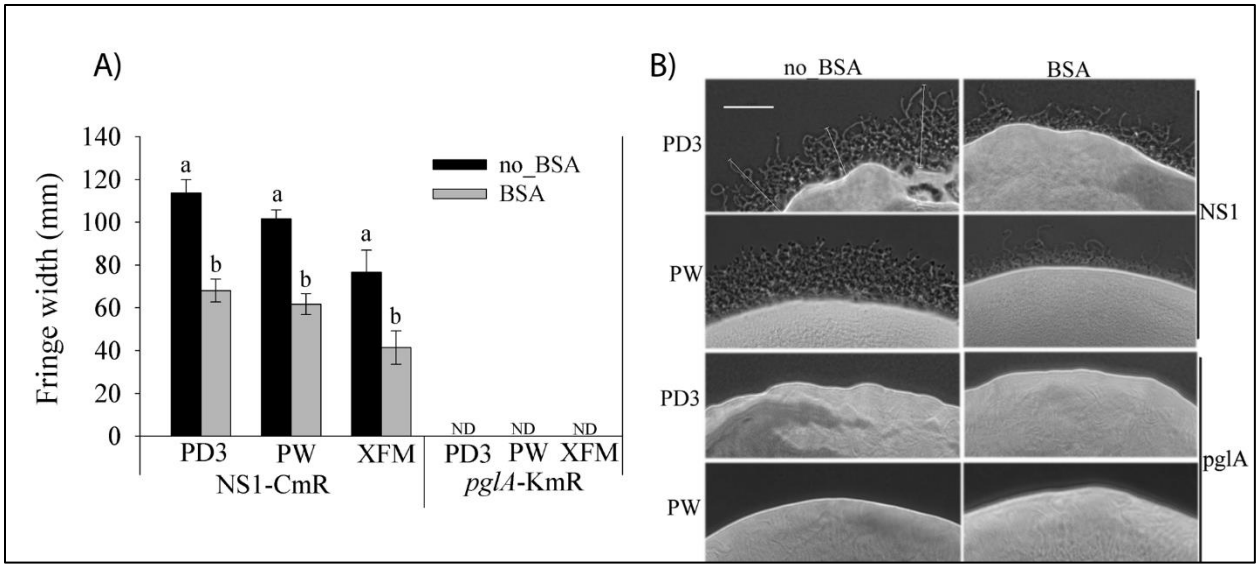


Fig. 2-3. Twitching motility in media with and without BSA. **A)** Fringe width was measured after spotting the bacteria in media plates and incubating at 28°C for 4-5 days for PD3 and PW and 10-12 days for XFM. Data represents means and standard errors from different experiments. Experiments were repeated independently three times for PD3 and PW (n=29-55) and once for XFM (n=4-9). Data were analyzed with PROC GLIMMIX in SAS. Different letters for each bar group indicate significant difference ($P < 0.05$). BSA significantly reduced twitching in all three media for NS1-CmR, while twitching motility was not detectable (ND) for *pglA*-KmR mutant in any media. **(B)** Colonies of NS1-CmR (fringe can be seen) and *pglA*-KmR mutant (no fringe) in media with and without BSA. Scale bar represents 100 µm.

Growth environment mimicking *X. fastidiosa* natural habitats supports natural competence

Surface attachment and biofilm formation of the cells was tested in the MC with PD3 and XFM, PW was not tested as it was not conducive for competence. Mixture of NS1-CmR and *pglA*-KmR mutant was inoculated into the channels at the beginning of the experiment. Both attachment and biofilm formation were pronounced with PD3, while XFM allowed very poor attachment of cells and formation of biofilms (Fig. 2-4). Hence, PD3 was selected as the medium to perform further natural competence experiments in MCs.

Within hours of inoculation, abundant *X. fastidiosa* cells were seen attached in the channels with PD3 medium (Fig. 2-4). Channels were filled with biofilm growth after ~ 5 days of injecting the cells. The frequency of recombination in the MC_in fraction that consist mainly of the biofilm formed under the flow conditions inside the microfluidic channels, was equivalent to that in the solid agar condition ($P = 0.52$) (Fig. 2-5). Noteworthy is that the frequency in the MC_in fraction was significantly higher than in the tubes ($P < 0.0001$) as well as in the MC_out fraction (consisting of cells collected downstream from the channels) ($P < 0.0001$) (Fig. 2-5). Recombinants were only occasionally detected in the MC_out fraction and the frequency was the lowest in this condition.

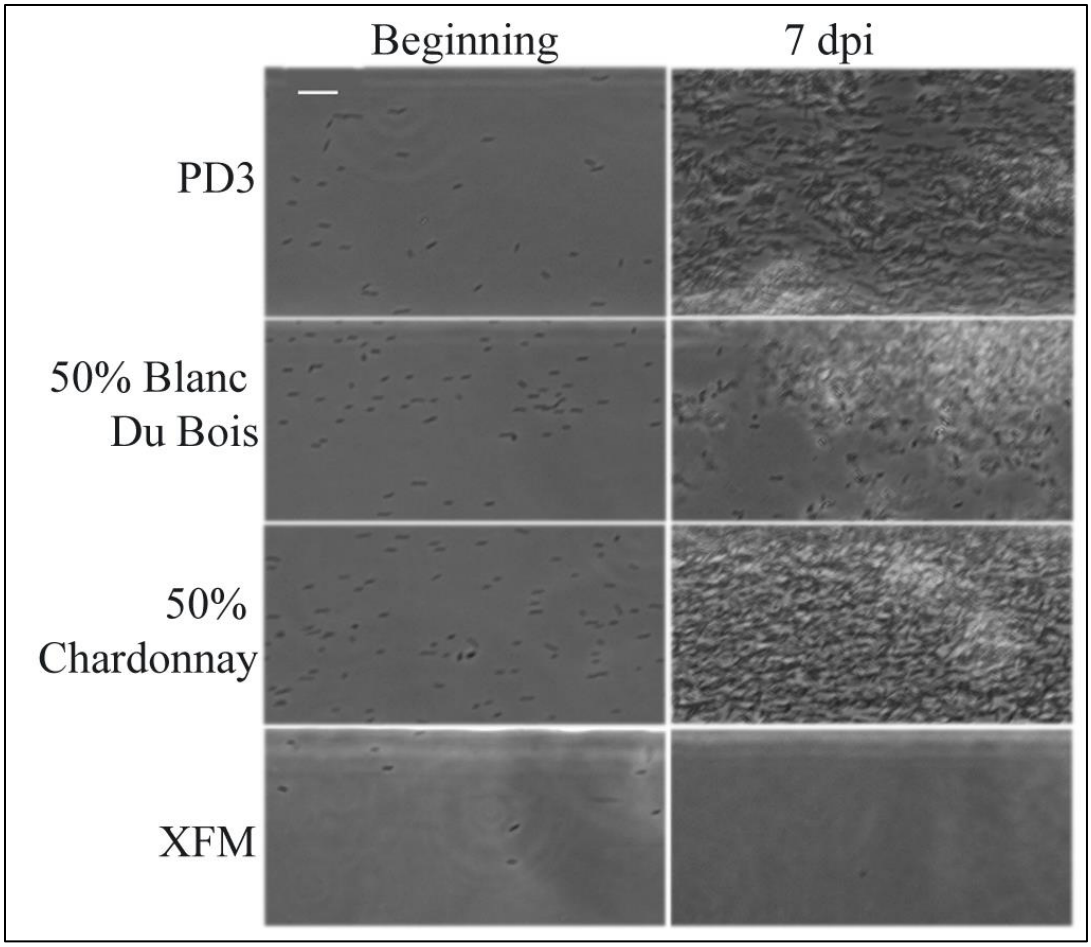


Fig. 2-4. Microfluidic channels showing cell attachment and biofilm formation in media and sap dilutions over time. Mixture of cells of NS1-CmR and *pglA*-KmR were injected into the channels. Cells attached well and copious biofilm growth was observed after seven days post inoculation (dpi) for both PD3 and 50% sap (susceptible cv. Chardonnay, tolerant Blanc Du Bois) but not for XFM. Scale bar on top left corner represents 10µm.

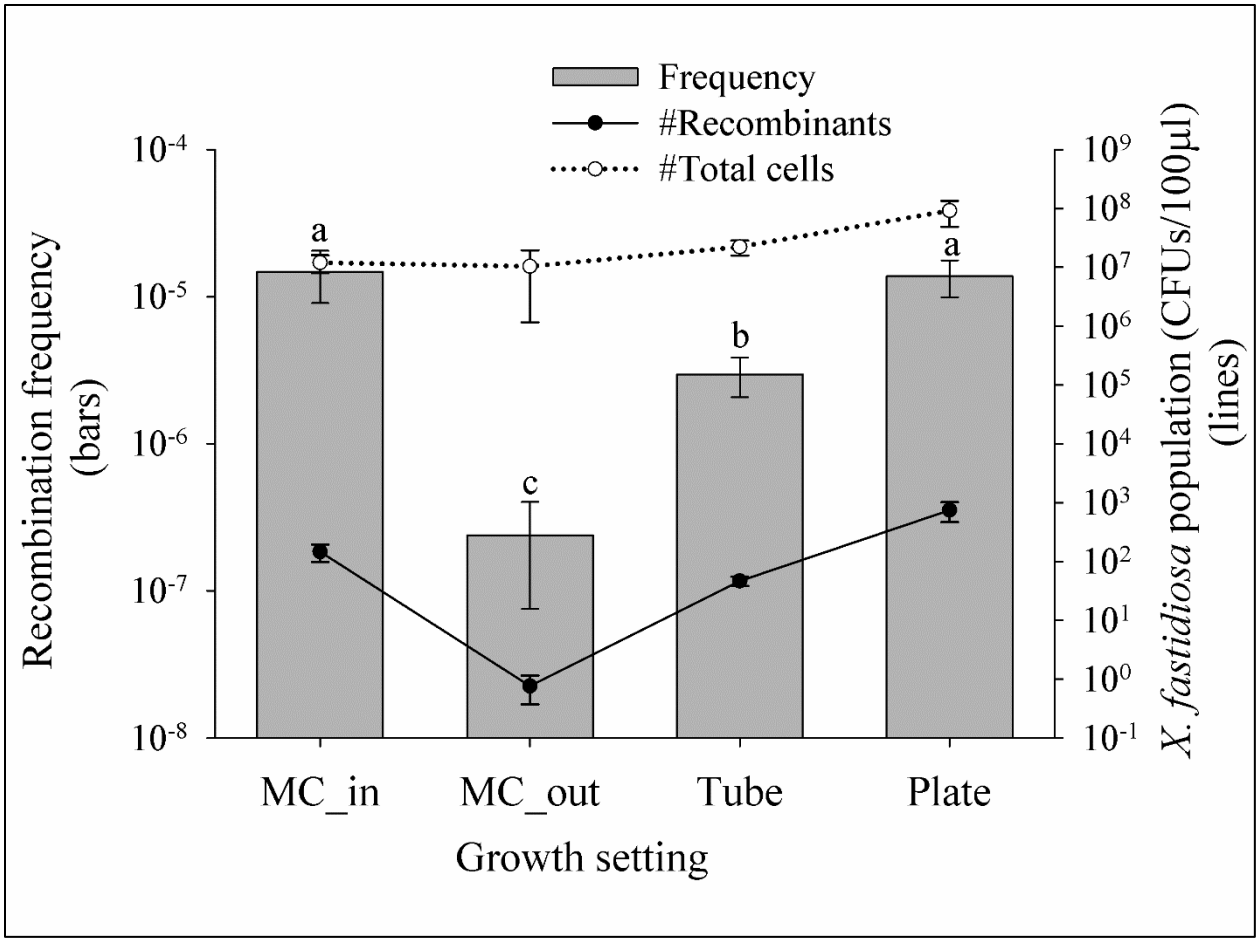


Fig. 2-5. Recombination frequencies, number of recombinants, and total CFUs in MC, tubes, and plates with PD3. Mixtures of the two strains were co-cultured in the three growth conditions and recombination at the antibiotics resistant marker region was assessed. For each growth condition, at least three independent experiments are included with n= 7, 7, 9 and 12 for MC_in, MC_out, tube and plate experiments, respectively. Data represents means and standard errors from different experiments. Different letters indicate significant differences for recombination frequency as analyzed by PROC GLIMMIX in SAS followed by Tukey's HSD test ($P < 0.05$).

Natural competence was maintained inside the microfluidic chambers with amendments of grapevine sap

Cell growth in pure xylem sap was slower than in PD3 and in 50% sap, nevertheless abundant cell aggregation was seen in the MCs after ~7 days. However, growth of only *pglA*-KmR mutant was observed in most of the experiments when pure sap was used, and generally no growth of NS1-CmR was observed, as revealed by plating the suspension in separate antibiotics plates at the end of the experiment (data not shown). During one experiment, both mutants were recovered and recombinants were formed in the MC_in fraction with pure sap (~16 ± 2 recombinant CFUs in 100 µl of the culture from six separate plates, and a frequency ~ 3.6×10^{-7}) (Fig. 2-6). Because of the inconsistent growth of NS1-CmR in pure sap, 50% sap (diluted with PD3) was used for further experiments. The frequency of recombination was highest in PD3, as compared to 50% Blanc Du Bois ($P = 0.001$) and 50% Chardonnay sap ($P < 0.0001$) (Fig. 2-6). Considering only grapevine sap, frequency of recombination was higher in 50% Blanc Du Bois than in 50% Chardonnay ($P = 0.0002$) (Fig. 2-6). Noteworthy is that with the sap experiments, recombinants could be recovered only from the MC_in fraction but not from the MC_out fraction (Supplemental table 2-1).

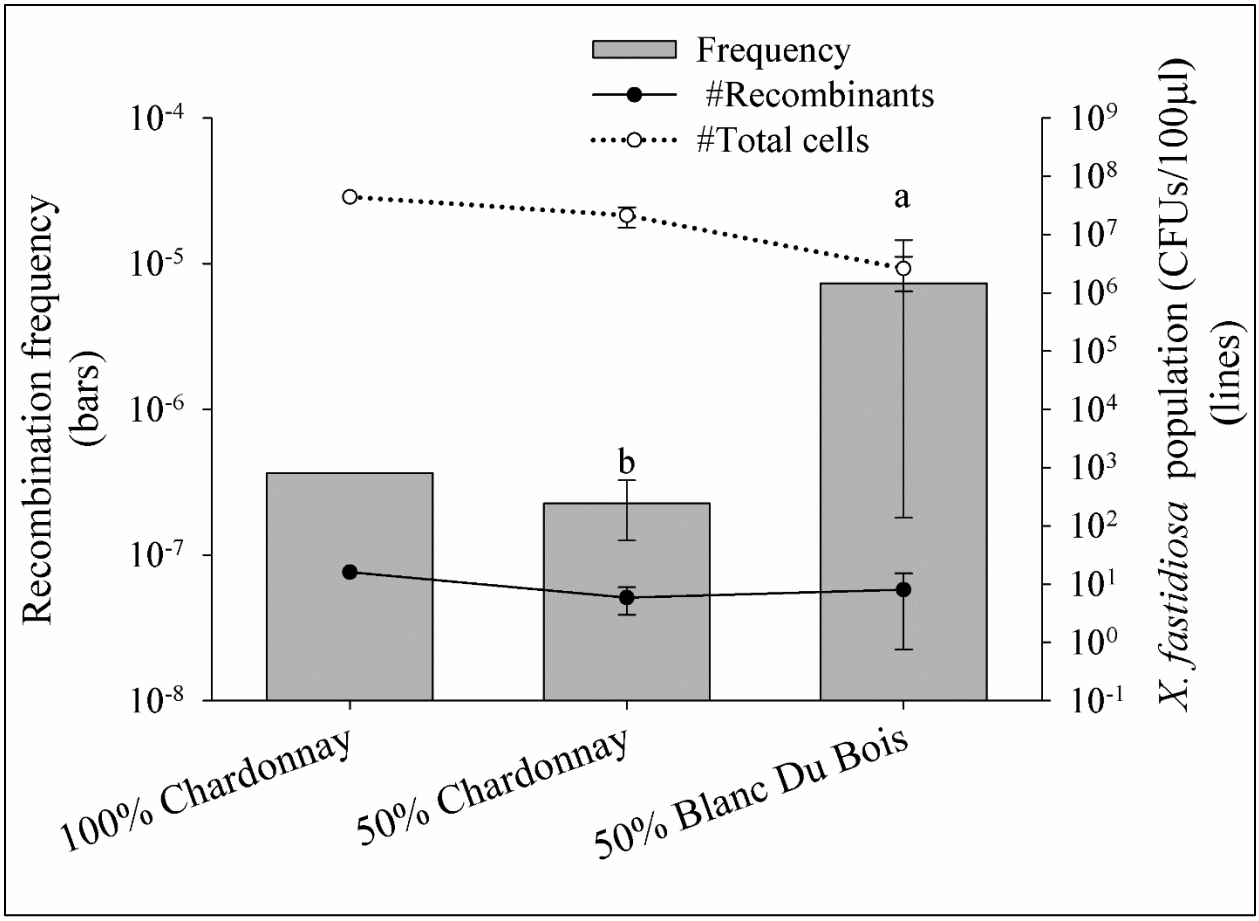


Fig. 2-6. Recombination frequency, number of recombinants, and total CFUs from the MC_in fraction with grapevine sap. Recombination frequency was significantly higher in 50% sap from tolerant Blanc Du Bois (n=2) than susceptible 50% Chardonnay sap (n=3), and pure sap (n=1), as analyzed in SAS with PROC GLIMMIX and Tukey HSD ($P < 0.05$). Data represent means and standard errors for different experiment.

Natural competence occurs with heat killed donor cells

Heat-killed cells of *pglA*-KmR mutant were used as donor with live cells of NS1-CmR as recipients, and vice versa, in PD3 agar plates. Double antibiotic resistant recombinant colonies (with a recombination frequency of $(3.6\pm1.4) \times 10^{-7}$, n=5) were recovered in both experiments

tested when NS1-CmR was used as the live recipient, confirming that natural competence is the process facilitating the homologous recombination. No recombinants, however, were obtained when *pglA*-KmR mutant was used as the live recipient. Experiments were performed also with pAX1.Cm plasmids as donor DNA and cells of *pglA*-KmR as recipients to confirm the non-competency of *pglA*-KmR, and no recombinants were detected (n=4, data not shown). Colony PCR of the randomly selected recombinant colonies confirmed that the antibiotic resistance marker gene is inserted in the targeted region by double recombination events (Supplemental fig. 2-2). Comparison of the flanking regions of the antibiotic resistance insertion sites between the parent strains (*pglA*-KmR and NS1-CmR) at the NS1 and *pglA* sites had 99.9% (one mismatch) and 100% identity, respectively (Supplemental fig. 2-3), ruling out the possibility that the non-competency of *pglA*-KmR is associated with the flanking homology of the recombining region.

Discussion

Several hypotheses have been proposed to explain the existence of natural competence in bacteria. One explanation is that starvation signals induce competence, and the incoming DNA serves as a nutrient source in poor nutrient conditions as demonstrated in *H. influenza* (Herriott et al., 1970), *P. stutzeri* (Lorenz and Wackernagel, 1991), and *R. solanacearum* (Bertolla et al., 1997b). Based on the results with a minimal medium (XFM), and a rich undefined medium (PW), a previous study (Kung and Almeida, 2011) speculated that growth in low nutrient medium favors natural competence in *X. fastidiosa.* However, the results of this study with these two media (XFM, PW) and PD3, another undefined rich medium, demonstrated that growth in PD3 significantly increases recombination frequency. This suggests that starvation is not necessary to induce competence in *X. fastidiosa*.

Two other components that differ between PD3 and PW are BSA and L-glutamine, which are only present in PW. Further investigation of the differences between PD3 and PW were performed by either removing or adding these components to/from one another. Initial screening with the components showed a pronounced effect of BSA in the number of recombinants recovered. Additional experiments confirmed that BSA significantly reduces the recombination frequency when present in PD3, PW, and XFM. Since both XFM and PW contain BSA, this may explain the lower recombination frequencies in these media. In a previous study, BSA had been found to reduce surface attachment and twitching motility of *X. fastidiosa* (Galvani et al., 2007). In fact, natural competence and twitching motility are dependent on the activity of type IV pili in *X. fastidiosa* (De La Fuente et al., 2007; Kung and Almeida, 2014). Therefore, in this study the correlation between twitching movement and natural competence in different media was investigated. Interestingly, PD3 allowed the highest fringe width and presence of BSA significantly reduced twitching motility in all three media. Twitching motility in XFM was lower than in both PD3 and PW as poor growth in XFM resulted in smaller colony sizes. Still the fringe width of colonies in XFM without BSA was bigger than in XFM with BSA. Most of the colonies spotted in XFM and XFM-BSA showed very little or no visible growth. This can be expected as XFM is a nutrient-limited minimal medium. Moreover, the *pglA*-KmR mutant that did not show twitching movement was not competent when tested with heat-killed NS1-CmR and plasmid DNA as the donor. These results of concomitant decrease in natural competence and twitching motility in BSA supplemented media and non-competency of twitch minus strain suggest that twitching motility is correlated with natural competence in *X. fastidiosa*. Natural competence in other gram negative bacteria is mediated by type IV pili like structures (Seitz and

Blokesch, 2013). Considering the effect of BSA on twitching, it remains to be determined if BSA only alters movement or biogenesis of type IV pili.

Our results with different growth settings showed that the recombination frequency is significantly higher in the MC_in fraction than in the MC_out fraction. The MC_in environment closely mimics xylem vessels and the insect foregut with respect to continuous liquid flow, adhesion of cells on channel walls in a fashion similar to adhesion of cells on xylem vessels and the insect foregut, and formation of biofilms. This environment is conducive for both biofilm formation and twitching motility as demonstrated in previous studies (De La Fuente et al., 2008; Cruz et al., 2012). Moreover, expression of some of the type IV pili genes were shown to be increased in the MC_in environment than in the other growth conditions (Cruz et al., 2014), implying that activity of type IV pili is increased in this system, which may explain higher rates of recombination in the MC_in fraction. The MC_out environment, on the other hand, consists mostly of planktonic cells and some detached biofilm fraction from MC_in, which is washed away with the liquid flow. Differences in recombination frequencies in these two environments suggest that the continuous media flow condition of the xylem vessels and growth in biofilm may increase the chances of recombination. Batch cultures in tubes also allowed recombination but at a lower rate than the continuous flow environment of MC_in, and surface attached condition of solid agar plates. A previous study also showed that growth in solid plates increase recombination compared to the growth in the liquid culture tubes (Kung and Almeida, 2014). Recombinants in the MC_out fraction were recovered when profuse biofilm growth was observed in the MC_in fraction with many recombinants formed. It is possible that the recombinants recovered in the MC_out fraction are due to detachment and washing away of

portions of biofilms from the MC_in fractions, supporting the proposition that biofilm formation induces competence. Biofilm formation and quorum sensing signals have been shown to induce natural competence in other naturally competent bacteria such as *V. cholerae* (Antonova and Hammer, 2011), *Acinetobacter* sp. (Hendrickx et al., 2003), and *Streptococcus mutans* (Li et al., 2001). Biofilms, in addition to having dense population of cells, contain elevated amounts of extracellular DNA (Steichen et al., 2011), which can be used for transformation by competent cells. Extracellular DNA has been shown to enhance biofilm formation in *X. fastidiosa* (Cheng et al., 2010). Moreover, Kung et al (Kung and Almeida, 2014) also showed that a knockout mutant on a biosynthetic gene for diffusible signaling factor (DSF), a cell-cell communication signal in *X. fastidiosa*, had reduced rate of recombination, implying that cell-cell communication signal also may be involved in regulating natural competence in *X. fastidiosa*.

MCs experiments with grapevine sap provide closer resemblance to the natural habitat than MCs with the artificial culture medium. Previously we have shown that biofilm structure in grapevine sap is more similar to the natural biofilm than are the aggregates observed in synthetic medium inside MC (Cogan et al., 2013). The experiments with amendments of sap in the MCs detected natural competence providing an indication that natural competence occurs in the xylem vessels of host plants and possibly in the insect vectors. Although results with pure sap experiments were not reproducible due to inconsistent growth of one of the strains used, recombinants were recovered once with pure Chardonnay sap as the medium. Recombinants were readily recovered with the 50% sap in PD3 for both tolerant and susceptible varieties. Maintenance of competence with addition of xylem sap, indicate that sap components support DNA acquisition and transformation. Natural competence occurring in environments resembling natural habitat also

have been demonstrated in other naturally competent bacteria such as *P. stutzeri* (Lorenz and Wackernagel, 1991) and *V. cholerae* (Meibom et al., 2005), in which artificial medium resembling natural soil extract, and natural growth substrate (chitin) induced competence, respectively. In *R. solanacearum*, another xylem colonizing plant pathogen, natural competence has been demonstrated *in planta* (Bertolla et al., 1999), and the recipient strains were shown to have increased virulence, acquiring DNA regions as long as 40 kb from donor strains.

Findings from competence experiments with grapevine sap and the MCs suggest that when two different strains are established together in the xylem vessels or in the vector foregut, recombination is possible. Noteworthy is the fact that in the experiments reported here, recombination was higher with sap from a tolerant grapevine variety, where infection by *X. fastidiosa* is symptomless. Co-infection by two genetically different isolates together in the same plant has been documented before (Chen et al., 2005), and there are reports of artificial mixed infection of a vector (Costa et al., 2006) as well as a single vector being able to transmit all four subspecies of *X. fastidiosa* (Almeida and Purcell, 2003). Moreover, it was shown that isolates from two different subspecies can cause disease in a single host (Oliver et al., 2015). Hence, possibility exists in nature that two different *X. fastidiosa* strains may encounter one another and exchange DNA as shown by MLST analyses. Donor DNA may be derived from dead cells or may be secreted by a type IV secretion system, as shown in *N. gonorrhoeae* (Dillard and Seifert, 2001). Moreover, the experiment with heat killed donor cells suggest that recombination is possible if homologous DNA fragments are present in the environment. Although the majority of recombination events will not be beneficial to the recipient cell, some may have adaptive advantage and increased virulence, among other phenotypes under selective pressure. For

example, the relative recent emergence of citrus variegated chlorosis and coffee leaf scorch (in 1980s) in South America is proposed to be due to inter-subspecific recombination between a subsp. *multiplex* donor and an unidentified native recipient based on MLST (Nunney et al., 2012). In addition, strains that are classified in the newly proposed subspecies '*morus'* that infect mulberry, have been suggested to be generated by recombination between a subsp. *fastidiosa* donor and a subsp. *multiplex* recipient (Nunney et al., 2014b). A similar mechanism may have resulted in strains that infect blueberry and blackberry (Nunney et al., 2014a).

Recombination events observed in this study are based on horizontal acquisition of antibiotic resistance markers (a single gene), which represent a small fraction of the genome of *X. fastidiosa*. Since the natural competence experiments were performed under conditions without any selective pressure, recombination events should be expected to have occurred at other regions of the genome as well, but were not detected due to the experimental approach used here. Under the simplistic assumption that gene exchange occurs randomly throughout the genome and with similar frequencies at all loci, the recombination frequencies reported in this study for one locus (~10^{-5} - 10^{-9} recombinants/total cells) could be as much as ~2.5 x 10^{3} higher, considering the size of the *X. fastidiosa* genome (~2.5Mb).

In summary, *X. fastidiosa* is naturally competent with a high rate of recombination when cultured under liquid flow conditions of MC, a system that mimics plant xylem vessels and the insect vector foregut. Natural competence in the MC was maintained even when the media was supplemented with grapevine xylem sap, suggesting that the natural habitat of *X. fastidiosa* supports natural competence. Moreover, habitats and media that favored increased biofilm

growth and increased twitching motility showed increased rate of recombination. This study advances the characterization of the phenomenon of natural competence in *X. fastidiosa* that needs to be further studied to understand the evolution and adaptation of this important plant pathogen.

Acknowledgements

This project was funded by HATCH AAES (Alabama Agricultural Experiment Station) program, California Agriculture Experiment Station, and Agriculture and Food Research Initiative competitive grant no. 2015-67014-23085 from the USDA National Institute of Food and Agriculture. S.M.L. was the recipient of an Undergraduate Research Fellowship from Auburn University. We would like to thank the Montaluce Winery, GA, and USDA station Tallahassee, FL for allowing us sap collection in their vineyard. We acknowledge the help of Fernando Navarrete, Luisa Cruz, Jennifer Parker, Hongyu Chen, Nick Cogan, Matt Donahue, and Mark Whitten during sample collection. We also thank Stephanie Kung for helpful suggestions.

References

Almeida, R.P., and Purcell, A.H. 2003. Transmission of *Xylella fastidiosa* to grapevines by *Homalodisca coagulata* (Hemiptera: Cicadellidae). J Econ Entomol 96:264-271.

Almeida, R.P., Nascimento, F.E., Chau, J., Prado, S.S., Tsai, C.W., Lopes, S.A., and Lopes, J.R. 2008. Genetic structure and biology of *Xylella fastidiosa* strains causing disease in citrus and coffee in Brazil. Appl Environ Microbiol 74:3690-3701.

Antonova, E.S., and Hammer, B.K. 2011. Quorum-sensing autoinducer molecules produced by members of a multispecies biofilm promote horizontal gene transfer to *Vibrio cholerae*. FEMS Microbiol Lett 322:68-76.

Baltrus, D.A., Guillemin, K., and Phillips, P.C. 2008. Natural transformation increases the rate of adaptation in the human pathogen *Helicobacter pylori*. Evolution 62:39-49.

Bertolla, F., VanGijsegem, F., Nesme, X., and Simonet, P. 1997a. Conditions for natural transformation of Ralstonia solanacearum. Appl Environ Microbiol 63:4965-4968.

Bertolla, F., Van Gijsegem, F., Nesme, X., and Simonet, P. 1997b. Conditions for natural transformation of *Ralstonia solanacearum*. Appl Environ Microbiol 63:4965-4968.

Bertolla, F., Frostegård, Å., Brito, B., Nesme, X., and Simonet, P. 1999. During infection of its host, the plant pathogen *Ralstonia solanacearum* naturally develops a state of competence and exchanges genetic material. Mol Plant Microbe Interact 12:467-472.

Chang, C.J., Donaldson, R., Brannen, P., Krewer, G., and Boland, R. 2009. Bacterial Leaf Scorch, a new blueberry disease Caused by *Xylella fastidiosa*. Hortscience 44:413-417.

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

Chen, J., Groves, R., Civerolo, E.L., Viveros, M., Freeman, M., and Zheng, Y. 2005. Two *Xylella fastidiosa* genotypes associated with Almond Leaf Scorch disease on the same location in California. Phytopathology 95:708-714.

Cheng, D.W., Lin, H., and Civerolo, E.L. 2010. Extracellular genomic DNA mediates enhancement of *Xylella fastidiosa* biofilm formation *in vitro*. J Plant Pathol 92:415-420.

Cogan, N.G., Donahue, M.R., Whidden, M., and De La Fuente, L. 2013. Pattern formation exhibited by biofilm formation within microfluidic chambers. Biophys J 104:1867-1874.

Costa, H.S., Guzman, A., Hernandez-Martinez, R., Gispert, C., and Cooksey, D.A. 2006. Detection and differentiation of *Xylella fastidiosa* strains acquired and retained by glassy-winged sharpshooters (Hemiptera: Cicadellidae) using a mixture of strain-specific primer sets. J Econ Entomol 99:1058-1064.

Cruz, L.F., Cobine, P.A., and De La Fuente, L. 2012. Calcium increases *Xylella fastidiosa* surface attachment, biofilm formation, and twitching motility. Appl Environ Microbiol 78:1321-1331.

Cruz, L.F., Parker, J.K., Cobine, P.A., and De La Fuente, L. 2014. Calcium-enhanced twitching motility in *Xylella fastidiosa* is linked to a single PilY1 homolog. Appl Environ Microbiol 80:7176-7185.

Das, M., Bhowmick, T.S., Ahern, S.J., Young, R., and Gonzalez, C.F. 2015. Control of Pierce's Disease by phage. PLoS One 10:e0128902.

Davis, M., French, W., and Schaad, N. 1981. Axenic culture of the bacteria associated with phony disease of peach and plum leaf scald. Curr Microbiol 6:309-314.

Davis, M.J., Purcell, A.H., and Thomson, S.V. 1980. Isolation media for the pierces disease bacterium. Phytopathology 70:425-429.

De La Fuente, L., Burr, T.J., and Hoch, H.C. 2008. Autoaggregation of *Xylella fastidiosa* cells is influenced by type I and type IV pili. Appl Environ Microbiol 74:5579-5582.

De La Fuente, L., Montanes, E., Meng, Y., Li, Y., Burr, T.J., Hoch, H.C., and 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.

Dillard, J.P., and Seifert, H.S. 2001. A variable genetic island specific for *Neisseria gonorrhoeae* is involved in providing DNA for natural transformation and is found more often in disseminated infection isolates. Mol Microbiol 41:263-277.

Dorer, M.S., Fero, J., and Salama, N.R. 2010. DNA damage triggers genetic exchange in *Helicobacter pylori*. PLoS Pathog 6:e1001026.

Galvani, C.D., Li, Y., Burr, T.J., and Hoch, H.C. 2007. Twitching motility among pathogenic *Xylella fastidiosa* isolates and the influence of bovine serum albumin on twitching-dependent colony fringe morphology. FEMS Microbiol Lett 268:202-208.

Griffith, F. 1928. The significance of pneumococcal types. J Hyg 27:113-159.

Hamilton, H.L., and Dillard, J.P. 2006. Natural transformation of *Neisseria gonorrhoeae*: from DNA donation to homologous recombination. Mol Microbiol 59:376-385.

Hendrickx, L., Hausner, M., and Wuertz, S. 2003. Natural genetic transformation in monoculture *Acinetobacter* sp. strain BD413 biofilms. Appl Environ Microbiol 69:1721-1727.

Herriott, R.M., Meyer, E.M., and Vogt, M. 1970. Defined nongrowth media for stage-ii development of competence in *haemophilus-influenzae*. J Bacteriol 101:517-524.

Hill, B.L., and Purcell, A.H. 1995. Acquisition and retention of *xylella fastidiosa* by an efficient vector, *Graphocephala-atropunctata*. Phytopathology 85:209-212.

Hofreuter, D., Odenbreit, S., Henke, G., and Haas, R. 1998. Natural competence for DNA transformation in *Helicobacter pylori*: identification and genetic characterization of the *comB* locus. Mol Microbiol 28:1027-1038.

Hopkins, D.L., and Purcell, A.H. 2002. *Xylella fastidiosa*: Cause of Pierce's Disease of Grapevine and Other Emergent Diseases. Plant Dis 86:1056-1066.

Humbert, O., Dorer, M.S., and Salama, N.R. 2011. Characterization of *Helicobacter pylori* factors that control transformation frequency and integration length during inter-strain DNA recombination. Mol Microbiol 79:387-401.

Jacques, M.A., Denance, N., Legendre, B., Morel, E., Briand, M., Mississipi, S., Durand, K., Olivier, V., Portier, P., Poliakoff, F., and Crouzillat, D. 2015. New coffee plant-infecting *Xylella fastidiosa* variants derived via homologous recombination. Appl Environ Microbiol 82:1556-1568.

Johnston, C., Martin, B., Fichant, G., Polard, P., and Claverys, J.P. 2014. Bacterial transformation: distribution, shared mechanisms and divergent control. Nat Rev Microbiol 12:181-196.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Mentjies, P., and Drummond, A. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28:1647-1649.

Killiny, N., and Almeida, R.P. 2009. Host structural carbohydrate induces vector transmission of a bacterial plant pathogen. Proc Natl Acad Sci USA 106:22416-22420.

Kung, S.H., and Almeida, R.P. 2011. Natural competence and recombination in the plant pathogen *Xylella fastidiosa*. Appl Environ Microbiol 77:5278-5284.

Kung, S.H., and Almeida, R.P. 2014. Biological and genetic factors regulating natural competence in a bacterial plant pathogen. Microbiology 160:37-46.

Kung, S.H., Retchless, A.C., Kwan, J.Y., and Almeida, R.P. 2013. Effects of DNA size on transformation and recombination efficiencies in *Xylella fastidiosa*. Appl Environ Microbiol 79:1712-1717.

Leu, L., and Su, C. 1993. Isolation, cultivation, and pathogenicity of *Xylella fastidiosa*, the causal bacterium of pear leaf scorch disease in Taiwan. Plant Dis 77:642-646.

Li, Y.H., Lau, P.C., Lee, J.H., Ellen, R.P., and Cvitkovitch, D.G. 2001. Natural genetic transformation of *Streptococcus mutans* growing in biofilms. J Bacteriol 183:897-908.

Lorenz, M.G., and Wackernagel, W. 1991. High frequency of natural genetic transformation of *Pseudomonas stutzeri* in soil extract supplemented with a carbon energy and phosphorus source. Appl Environ Microb 57:1246-1251.

Lorenz, M.G., and Wackernagel, W. 1994. Bacterial gene transfer by natural genetic transformation in the environment. Microbiol Rev 58:563-602.

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

Meibom, K.L., Blokesch, M., Dolganov, N.A., Wu, C.-Y., and Schoolnik, G.K. 2005. Chitin induces natural competence in *Vibrio cholerae*. Science 310:1824-1827.

Navarrete, F., and De La Fuente, L. 2014. Response of *Xylella fastidiosa* to zinc: decreased culturability, increased exopolysaccharide production, and formation of resilient biofilms under flow conditions. Appl Environ Microbiol 80:1097-1107.

Newman, K.L., Almeida, R.P., Purcell, A.H., and Lindow, S.E. 2003. Use of a green fluorescent strain for analysis of *Xylella fastidiosa* colonization of *Vitis vinifera*. Appl Environ Microbiol 69:7319-7327.

Nunes, L.R., Rosato, Y.B., Muto, N.H., Yanai, G.M., da Silva, V.S., Leite, D.B., Gonçalves, E.R., de Souza, A.A., Coletta-Filho Hé, D., Machado, M.A., Lopes, S.A., and de Oliveira, R.C. 2003. Microarray analyses of *Xylella fastidiosa* provide evidence of

coordinated transcription control of laterally transferred elements. Genome Res 13:570-578.

Nunney, L., Yuan, X., Bromley, R.E., and Stouthamer, R. 2012. Detecting genetic introgression: high levels of intersubspecific recombination found in *Xylella fastidiosa* in Brazil. Appl Environ Microbiol 78:4702-4714.

Nunney, L., Hopkins, D.L., Morano, L.D., Russell, S.E., and Stouthamer, R. 2014a. Intersubspecific recombination in *Xylella fastidiosa* strains native to the United States: infection of novel hosts associated with an unsuccessful invasion. Appl Environ Microbiol 80:1159-1169.

Nunney, L., Schuenzel, E.L., Scally, M., Bromley, R.E., and Stouthamer, R. 2014b. Large-scale intersubspecific recombination in the plant-pathogenic bacterium *Xylella fastidiosa* is associated with the host shift to mulberry. Appl Environ Microbiol 80:3025-3033.

Oliver, J.E., Cobine, P.A., and De La Fuente, L. 2015. *Xylella fastidiosa* isolates from both subsp. *multiplex* and *fastidiosa* cause disease on southern highbush blueberry (*Vaccinium sp.*) under greenhouse conditions. Phytopathology 105:855-862.

Parker, J.K., Havird, J.C., and De La Fuente, L. 2012. Differentiation of *Xylella fastidiosa* strains via multilocus sequence analysis of environmentally mediated genes (MLSA-E). Appl Environ Microbiol 78:1385-1396.

Redak, R.A., Purcell, A.H., Lopes, J.R.S., Blua, M.J., Mizell, R.F., and Andersen, P.C. 2004. The biology of xylem fluid-feeding insect vectors of *Xylella fastidiosa* and their relation to disease epidemiology. Annu Rev Entomol 49:243-270.

Roper, M.C., Greve, L.C., Warren, J.G., Labavitch, J.M., and Kirkpatrick, B.C. 2007. *Xylella fastidiosa* requires polygalacturonase for colonization and pathogenicity in *Vitis vinifera* grapevines. Mol Plant Microbe Interact 20:411-419.

Saponari, M., Boscia, D., Nigro, F., and Martelli, G.P. 2013. Identification of DNA sequences related to *Xylella fastidiosa i*n Oleander, Almond and Olive trees exhibiting leaf scorch symptoms in Apulia (Southern Italy). J Plant Pathol 95:668-668.

Scally, M., Schuenzel, E.L., Stouthamer, R., and Nunney, L. 2005. Multilocus sequence type system for the plant pathogen *Xylella fastidiosa* and relative contributions of recombination and point mutation to clonal diversity. Appl Environ Microbiol 71:8491-8499.

Seitz, P., and Blokesch, M. 2013. Cues and regulatory pathways involved in natural competence and transformation in pathogenic and environmental Gram-negative bacteria. FEMS Microbiol Rev 37:336-363.

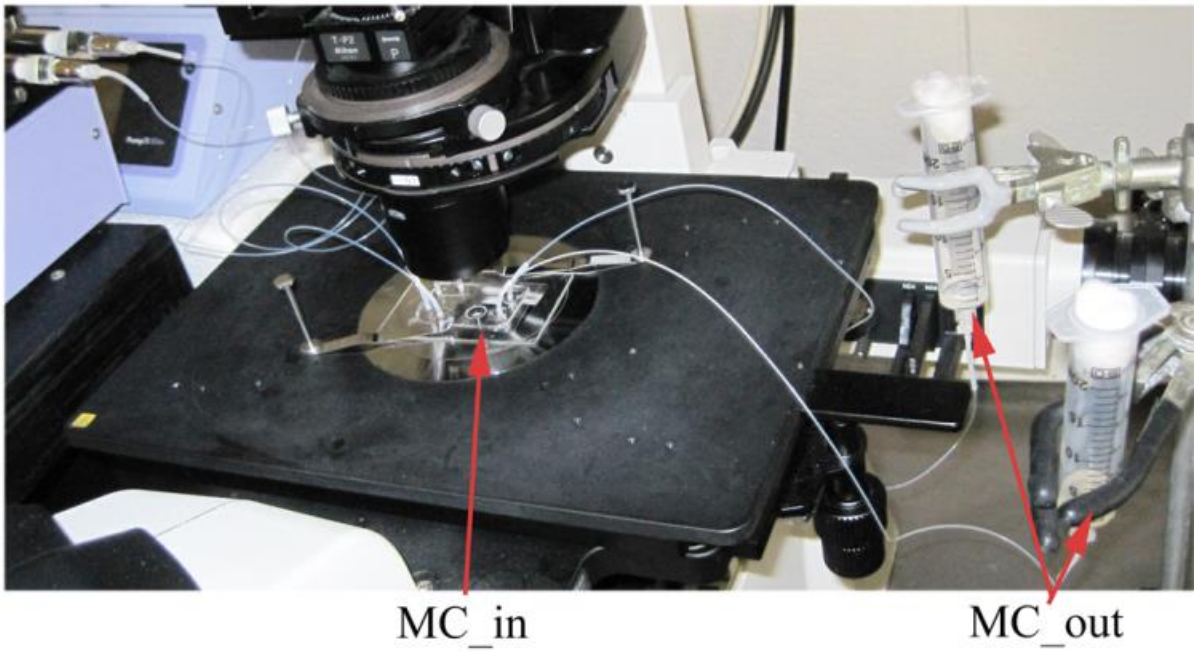
Sparling, P.F. 1966. Genetic transformation of *Neisseria gonorrhoeae* to streptomycin resistance. J Bacteriol 92:1364-1371.

Steichen, C.T., Cho, C., Shao, J.Q., and Apicella, M.A. 2011. The *Neisseria gonorrhoeae* biofilm matrix contains DNA, and an endogenous nuclease controls its incorporation. Infect Immun 79:1504-1511.

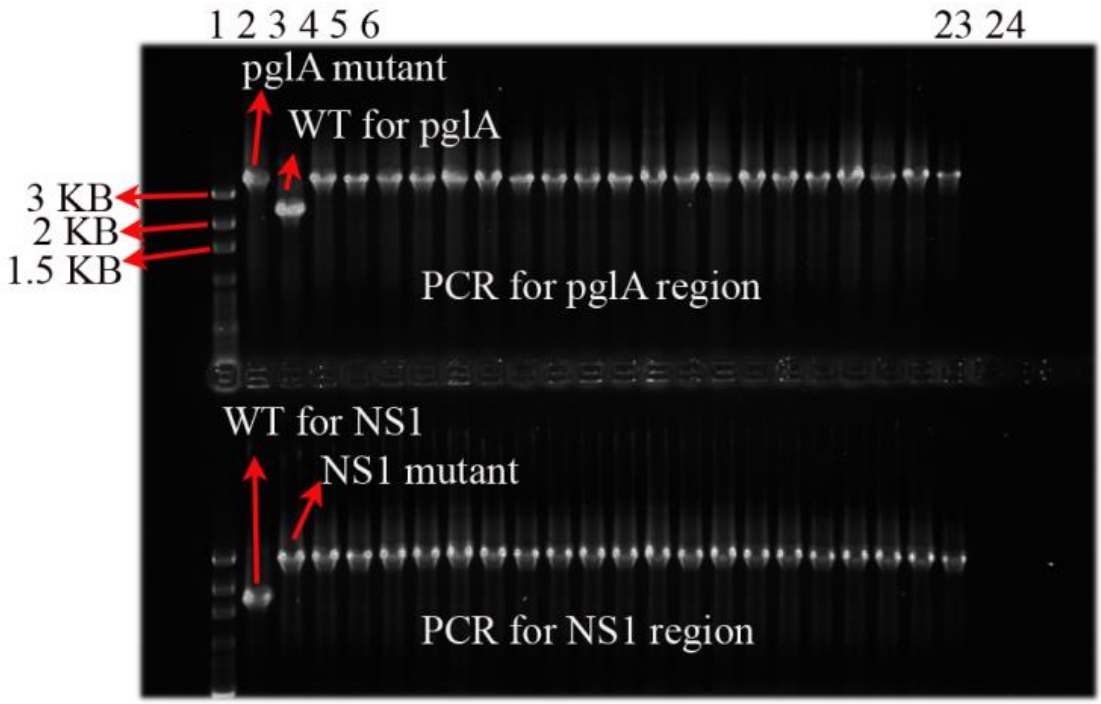
Su, C.C., Chang, C.J., Chang, C.M., Shih, H.T., Tzeng, K.C., Jan, F.J., Kao, C.W., and Deng, W.L. 2013. Pierce's disease of grapevines in Taiwan: isolation, cultivation and pathogenicity of *Xylella fastidiosa*. J Phytopathol 161:389-396.

Tribble, G.D., Rigney, T.W., Dao, D.H., Wong, C.T., Kerr, J.E., Taylor, B.E., Pacha, S., and Kaplan, H.B. 2012. Natural competence is a major mechanism for horizontal DNA transfer in the oral pathogen *Porphyromonas gingivalis*. MBio 3.

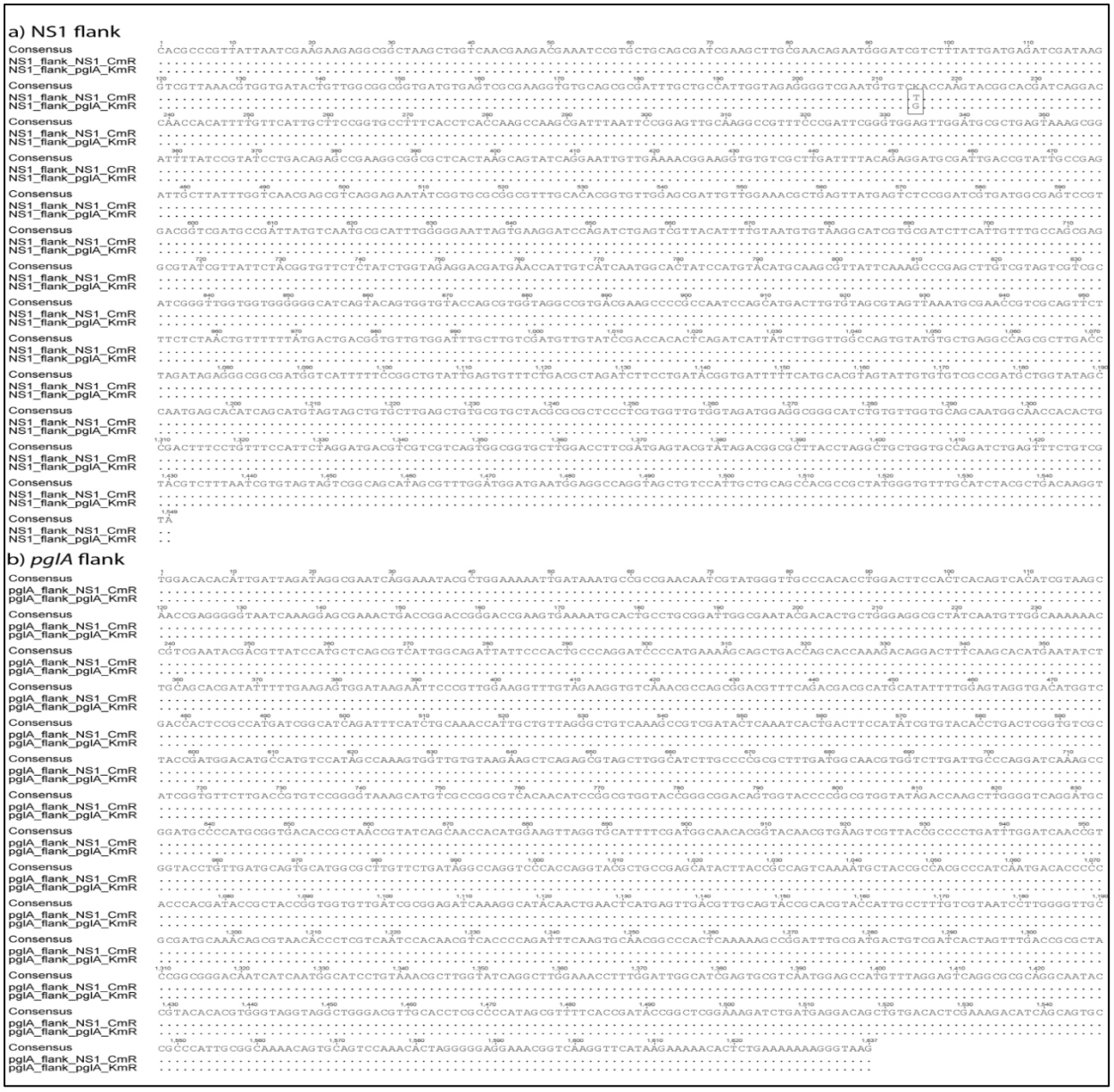
Supplemental material for chapter 2



Supplemental fig. 2-1. Microfluidic chamber system showing tubings connected to syringes filled with growth medium and bacterial suspensions on the left, and outlet collection syringes on the right. Cell attachment and biofilm formation occurs inside the channels (MC_in) (as shown in Fig. 2-4). The unattached cells, and detached biofilm fractions are collected in the outlet collection syringes (MC_out).



Supplemental fig. 2-2. Confirmation of homologous recombination between NS1-CmR and *pglA*-kmR mutant using colony PCR. The upper bands are PCR products with *pglA* region specific primers and the lower bands with NS1region specific primers. Lane 1; 100 bp ladder, lane 2- 3; wild type (WT) and mutant parents as indicated in the figure, lane 4-23 recombinants of NS1-CmR and *pglA*-KmR, lane 24 negative control. The wild type for pglA is 2326bp and the mutant with kanamycin cassette insertion is 3564 bp and for NS1 wild type is 1715 bp and mutant with chloramphenicol insertion is ~2900bp. The recombinants had both antibiotic cassette inserted in their genome.



Supplemental fig. 2-3. Sequence comparison of flanking regions between the parents (NS1-CmR and *pglA*-KmR) at antibiotics resistance insertion sites; **a)** NS1 site and **b)** *pglA* site. The flanking regions were 99.9% identical at the NS1 site and 100% identical at the *pglA* site between the parents. Square indicates mismatch at NS1 site between the two compared sequences.

Supplemental table 2-1: Number of recombinants, total CFUs, and recombination frequency under different media and settings

Media	Setting	#Recombinants[a]	#Total[a]	Frequency	N
PD3	Plate	322.8±72.3[b]	2.9±1.2 E+07	1.9±0.4 E-05	18
PD3+BSA[c]	Plate	3.8±0.8	5.2±0.3 E+07	7.3±1.5 E-08	9
PD3+BSA+L-glutamine	Plate	3.7±1.0	4.2±1.5 E+07	8.6±2.4 E-08	3
PD3+L-glutamine	Plate	252.0±30.2	3.1±0.2 E+07	8.2±1.0 E-06	3
PD3+Pectin	Plate	202.4±11.2	1.8±0.5 E+07	1.3±0.3 E-05	3
PD3-SA[c]	Plate	985.0	5.2 E+07	1.8 E-05	1
PD3-SCD[c]	Plate	575.5	5.2 E+07	1.1 E-05	1
PD3-SCD-SA	Plate	141.0	7.9 E+07	1.8 E-06	1
PD3-Starch	Plate	394.3	7.7 E+07	5.1 E-06	1
PW	Plate	0.1±01	3.7±1.7 E+07	8.3±7.1 E-09	12
PW-BSA	Plate	8.7±2.6	7.4±3.4 E+07	2.3±0.8 E-07	8
PW-L-glutamine	Plate	0	ND[d]	0	1
PW-BSA-L-glutamine	Plate	74.5	5.9 E+08	1.3 E-07	1
XFM	Plate	2.0±0.8	3.4±1.2 E+07	2.4±1.3 E-07	19
XFM-BSA	Plate	8.6±2.1	6.7±1.3 E+06	1.4±0.3 E-06	12
PD3	MC_in	146.3±48.1	1.2±0.4 E+07	1.5±0.5 E-05	7
PD3	MC_out	0.8±0.4	1.0±0.9 E+07	2.4±1.6 E-07	7
Chardonnay sap	MC_in	16.0	4.4 E+07	3.6 E-07	1
Chardonnay sap	MC_out	0	ND	0	1
50% Chardonnay sap	MC_in	5.9±2.9	2.1±0.8 E+07	2.3±1.0 E-07	3
50% Chardonnay sap	MC_out	0.0	2.3±2.0 E+06	0	3
50% Blanc Du Bois	MC_in	8.0±7.3	2.6±1.6 E+06	7.3±7.2 E-06	2
50% Blanc Du Bois	MC_out	0	3.0±0.3 E+05	0	2
PD3	Tube	46.7±8.4	2.2±0.6 E+07	3.0±0.9 E-06	9

[a]Values are expressed as CFU $100\mu l^{-1}$
[b]For the treatments where n >1, standard error of mean is included after the average number.
[c]Abbreviations. BSA: bovine serum albumin, SA: succinic acid, SCD: sodium citrate dehydrate.
[d]ND not determined. Total cells were not enumerated as no recombinants were detected for these treatments.

Chapter 3

Natural competence rates are variable among *Xylella fastidiosa* strains and homologous recombination occurs *in vitro* between subspecies *fastidiosa* and *multiplex*

Abstract

Xylella fastidiosa, an etiological agent of emerging crop diseases around the world, is naturally competent for the uptake of DNA from the environment that is incorporated into its genome by homologous recombination. Homologous recombination between subspecies of *X. fastidiosa* was inferred by in silico studies and was hypothesized to cause disease emergence. However, no experimental data are available on the degree to which *X. fastidiosa* strains are capable of competence and whether recombination can be experimentally demonstrated between subspecies. Here, using *X. fastidiosa* strains from different subspecies, natural competence was confirmed in eleven out of thirteen strains with plasmids containing antibiotic markers flanked by homologous regions; and in three out of five strains with dead bacterial cells used as source of donor DNA. Recombination frequency differed among strains and was correlated to growth rate and twitching motility. Moreover, intersubspecific recombination occurred readily between strains of subsp. *fastidiosa* and *multiplex*, as demonstrated by movement of antibiotic resistance and green fluorescent protein from donor to recipient cells, and confirmed by DNA sequencing of the flanking arms of recombinant strains. Results demonstrate that natural competence is widespread among *X. fastidiosa* strains and could have an impact in pathogen adaptation and disease development.

Introduction

Xylella fastidiosa is a xylem-limited, plant-pathogenic bacterium that causes destructive diseases in a number of economically important crops such as Pierce's disease in grapevines, citrus variegated chlorosis in citrus, and bacterial leaf scorch in coffee, plum, and almond (Hopkins and Purcell, 2002). In recent years, emergence of *X. fastidiosa* diseases has been reported in new host plants and geographic locations (Chang et al., 2009; Su et al., 2013; Martelli et al., 2016) with host range expanding to 359 plant species in 75 botanical families (European Food Safety Authority, 2016). Taxonomically, *X. fastidiosa* strains worldwide were classified under a single species, but recently another species was described for Taiwanese strains causing pear leaf scorch (Su et al., 2016). Studies based on Multi Locus Sequence Typing/Analysis (MLST/MLSA) have proposed five subspecies within *X. fastidiosa* (Scally et al., 2005; Nunney et al., 2014b): subsp. *fastidiosa*, *sandyi*, *morus*, *multiplex*, and *pauca*. The disease mechanism, although not fully understood, involves formation of biofilm in the xylem vessels of host plants, and obstruction of xylem sap transport leading to alteration of water and mineral nutrients in the aerial plant parts (Chatterjee et al., 2008; De La Fuente et al., 2013). *X. fastidiosa* strains are vector-transmitted by xylem sap-feeding insects such as sharpshooters and spittlebugs (Redak et al., 2004). There is no known specificity involved in recognition of insect vectors and *X. fastidiosa* subspecies, as a single insect vector was able to transmit multiple subspecies (Almeida and Nunney, 2015). However, host plant specificity has been described between and within subspecies (Almeida and Purcell, 2003; Almeida et al., 2008; Oliver et al., 2014; Harris and Balci, 2015; Oliver et al., 2015). Subspecies were believed to be geographically separated, but dissemination of plant materials and insect vectors has led to co-existence of two or more

subspecies in the same geographic location (Chen et al., 2005; Nunney et al., 2010; Parker et al., 2012; Nunney et al., 2014a).

Despite differences in plant host range, little genetic diversity has been described between *X. fastidiosa* strains that infect different host plants on a genome wide scale. For example, the grapevine strain Temecula and the citrus strain 9a5c, the most divergent genomes within *X. fastidiosa*, shared 98% of their genes and 96% of average amino acid identity when their whole genomes were compared (Van Sluys et al., 2003). However, with the application of MLST/MLSA techniques in *X. fastidiosa* phylogenetics, genetic differences have been described among and within subspecies (Scally et al., 2005; Almeida et al., 2008; Nunney et al., 2010; Yuan et al., 2010; Nunney et al., 2014b; Nunney et al., 2014c). For instance, strains collected from a single host and geographic location formed different haplotypes when they were analyzed by MLSA of environmentally-mediated genes (Parker et al., 2012). Interestingly, studies that detected genetic diversity among *X. fastidiosa* strains have also reported the presence of homologous recombination (HR) (Scally et al., 2005; Almeida et al., 2008; Nunney et al., 2010; Nunney et al., 2012; Nunney et al., 2013; Nunney et al., 2014b; Nunney et al., 2014c; Marcelletti and Scortichini, 2016; Coletta-Filho et al., 2017). In fact, HR was predicted to have greater effect in creating genetic diversity in *X. fastidiosa* than point mutation (Scally et al., 2005; Kung et al., 2013). Moreover, studies based on genetic data have shown occurrence of intersubspecific homologous recombination (IHR) between *X. fastidiosa* subspecies (Nunes et al., 2003), and hypothesized IHR to lead to plant host shift (Nunney et al., 2012; Nunney et al., 2014b; Nunney et al., 2014c). Co-existence of different *X. fastidiosa* subspecies has enabled the possibility of IHR to a greater extent. Interestingly, natural competence was described in *X. fastidiosa* (Kung

and Almeida, 2011) and provided an explanation for the frequent HR events detected by genetic studies.

Natural competence is a phenomenon that involves uptake of naked DNA fragments from the environment and their incorporation into the bacterial genome by HR (Lorenz and Wackernagel, 1994). Described in 1928 by Frederick Griffith (Griffith, 1928), natural competence was proposed as a mechanism of nutrient acquisition under starvation conditions (Seitz and Blokesch, 2013a), but has been also implicated in repair of damaged DNA (Dorer et al., 2010), and bacterial genome evolution providing adaptive advantage (Baltrus et al., 2008) and increased virulence (Coupat et al., 2011) to the recipient bacterium. In majority of naturally competent bacteria, competence development is regulated and induced by external environmental cues (Seitz and Blokesch, 2013a). Regulatory mechanisms involved in the induction of natural competence have been described in well studied bacterial systems such as *Vibrio cholerae* (Antonova and Hammer, 2015), and *Haemophilus influenzae* (Cameron et al., 2008).

However, very limited information exists regarding natural competence in *X. fastidiosa* as only few studies have been performed (Kung and Almeida, 2011; Kung et al., 2013; Kung and Almeida, 2014; Kandel et al., 2016). Studies so far have demonstrated natural competence of *X. fastidiosa in vitro* in batch cultures (Kung and Almeida, 2011; Kandel et al., 2016), solid media (Kung and Almeida, 2014; Kandel et al., 2016), and in microfluidic chambers (MC), a system designed to closely resemble the bacterium's natural habitat (Kandel et al., 2016). Natural competence occurred at a high frequency under flow conditions in MC including conditions with grapevine sap as growth medium (Kandel et al., 2016). Cells developed competence when

entering into the exponential growth phase (Kung and Almeida, 2011; Kung and Almeida, 2014), appeared to prefer DNA from self-sources over foreign-sources (Kung and Almeida, 2011), and required flanking region homology for recombination (Kung et al., 2013). One key knowledge gap is related to the variability of natural competence among strains as previous studies were based on few closely related strains belonging to the same subsp. *fastidiosa* (Kung and Almeida, 2011; Kung et al., 2013; Kung and Almeida, 2014; Kandel et al., 2016). Since both genotypic and phenotypic differences have been described in *X. fastidiosa* strains (Scally et al., 2005; Parker et al., 2012; Oliver et al., 2014; Antonova and Hammer, 2015; Oliver et al., 2015; Coletta-Filho et al., 2017), their natural competence abilities could differ as is the case reported in other naturally competent bacteria (Gromkova et al., 1998; Sikorski et al., 2002; Fujise et al., 2004; Coupat et al., 2008; Bosse et al., 2009; Maughan and Redfield, 2009; Evans and Rozen, 2013). Moreover, although IHR has been detected by MLST studies, and was hypothesized to lead to plant host shift (Nunney et al., 2012; Nunney et al., 2014b; Nunney et al., 2014c), there is still no experimental evidence demonstrating the generation of recombinants by mixing two different subspecies. Therefore, the objectives of this study were i) to determine the relative ability of natural competence in *X. fastidiosa* strains that belong to different subspecies ii) and to test experimentally if IHR occurs between the two subspecies prevalent in the US. Results showed that eleven of the thirteen strains were naturally competent, and recombination occurred at various regions of the genome. One highly virulent strain showed very high recombination with DNA from dead cells and plasmids tested as donors. Moreover, IHR was confirmed between subsp. *fastidiosa* and subsp. *multiplex* when combinations of donor and recipients from two subspecies were co-cultured. These findings demonstrate that *X. fastidiosa* isolates have different recombination potential, and co-existence of subspecies allows IHR that leads to

generation of novel genotypes, which based on the genomic regions of recombination, could show altered adaptation and/or virulence.

Materials and Methods

Bacterial strains, media, and culture conditions

X. fastidiosa subsp. *fastidiosa* strains WM1-1, Temecula1, Temecula1*, CCPM1, Fetzer, ConnCreek, and EB92-1; and subsp. *multiplex* strains AlmaEM3, BB08-1, BBI64, Birmingham Elm (BH Elm), and Georgia Plum were used in this study. Strain Chard1 was isolated from infected vines *Vitis vinifera* 'Chardonnay' in Dahlonega, Georgia in 2014 and is assumed to be a subsp. *fastidiosa* strain based on its plant host (Supplemental table 3-1). A Temecula1 mutant expressing green fluorescent protein (Temecula1-GFP) was used from a previous study (Newman et al., 2003). Mutant strains NS1::Km-WM1-1, NS1::Km-EB92-1, and NS1::Km-AlmaEM3 were generated by transforming the wild-type strains with pAX1.Km plasmids, and strain NS1::Cm-AlmaEM3 was generated by transforming with pAX1.Cm plasmids (Matsumoto et al., 2009), that insert the respective antibiotic resistant genes at neutral site (NS1) in the genome. Mutant strain *msrA*::Km-WM1-1 was generated by transforming pMSRA-Km (Supplemental table 3-1). All transformations were performed using the natural competence protocol described below.

All strains were cultured in PW (Davis et al., 1980) agar plates, modified by replacing phenol red with 1.8 g litre^{-1} Bovine Serum Albumin (BSA) (Gibco Life Sciences Technology), for one week at 28°C from -80°C glycerol stock, re-streaked onto new PW plates, and cultured for another week before use. Medium PD3 (Davis et al., 1981) was used to suspend the cells in liquid and

for natural competence experiments. Lauria Bertani (LB) medium was used to culture *Escherichia coli* cells. Kanamycin, chloramphenicol, and ampicillin were used at a concentration of 30, 10, and 100 µg ml^{-1}, respectively.

Test of natural competence of *X. fastidiosa* strains

Plasmids pAX1.Cm (Matsumoto et al., 2009), and pKLN61 (Newman et al., 2004) were used from previous studies. Plasmids pMSRA-Km, and pMOPB-Km were prepared as described earlier (Cruz et al., 2014) and are being characterized for another study in our laboratory (Chen et al., 2017 submitted). Briefly, ~800 bp long upstream and downstream fragments flanking open reading frames of methionine S-S-oxide reductase (*msrA*; PD0859), and outer membrane protein (*mopB*; PD1709), respectively were PCR amplified from the Temecula1 genomic DNA. The upstream and downstream fragments were digested using Asci restriction enzyme (Promega), ligated, and cloned into pJET1.2/blunt cloning vector and a kanamycin resistant cassette was inserted between the two fragments. All the plasmids were transformed into *E. coli* EAM1 competent cells that express *X. fastidiosa* DNA methylase (Matsumoto and Igo, 2010). Plasmids were prepared from the overnight cultures of EAM1 using an extraction kit (GeneJet Plasmid Miniprep Kit, Thermo Scientific), and concentration was adjusted to 100 ng/µl (NanoDrop 2000 spectrophotometer, Thermo Scientific). Aliquots were stored at -20˚C until use. Natural competence assays were performed in PD3 agar plates. The recipient strains were adjusted to optical density (OD_{600}) 0.25 (~10^8 cells/ml) in PD3 broth. Ten µl of this suspension were spotted onto PD3 agar plates, and one µg of plasmid in 10 µl volume was added to the spots. Following incubation at 28˚C for ~3 days (for non-competent strains this incubation time was extended up to 5 days), spots were suspended in 1 ml of PD3, and serial dilutions were plated in the

respective antibiotic (Km or Cm) PW plates in triplicates depending on which antibiotic cassette each plasmid carried, and PW plates without antibiotics. After 2-3 weeks of incubation at 28°C, CFUs were enumerated for recombinants (number of colonies formed on antibiotics supplemented PW plates) and total viable cells (number of colonies formed on appropriate dilutions of PW plates without antibiotics), followed by calculation of recombination frequency as the ratio of number of recombinants to total viable cells. For a given experiment, at least three repetitions were performed per strain, and the experiments had 2-6 biological replicates. For each strain, spots without the addition of plasmids were included as controls for every experiment. Genomic incorporation of the antibiotic resistant marker from the donor plasmid was confirmed by PCR as previously described (Kandel et al., 2016).

Comparison of the flanking region DNA sequences of donor and recipient strains

To compare flanking region DNA sequence homology among recipient *X. fastidiosa* strains with respect to each donor plasmid, up- and down-stream flanking regions of the antibiotic insertion sites of the plasmids pAX1.Cm, pKLN61, pMSRA-Km, and pMOPB-Km were obtained. Up-and downstream sequences homologous to each plasmids region were obtained from the genomes of the strains WM1-1, Temecula1, Temecula1*, BB08-1, AlmaEM3, BBI64 (see chapter 4), and EB92-1 (Zhang et al., 2011). The genomes (except EB92-1) were sequenced using an Illumina Miseq and PacBio sequencing systems, and resulting reads after quality trimming were mapped to the Temecula1 reference genome using the Geneious map to reference algorithm (Kearse et al., 2012). The two up-and down-stream sequences were concatenated, aligned using the Muscle Multiple Sequence Alignment tool in Geneious, and percent identity between each of the donor plasmids and recipient strains was determined.

Growth curve, biofilm, settling rate, and twitching motility of *X. fastidiosa* strains

Growth curves of strains were generated by culturing strains in PD3 medium in 96 well plates, and measuring OD_{600} value every day for eight days. At the beginning (day 0), wells of the 96 wells plate were inoculated with 190 µl of PD3 and 10 µl of cell suspension (OD_{600}=0.25) with eight wells (repetitions) used per strain. Eight wells were filled with 200 µl of PD3 to serve as controls. The plates were incubated at 27°C with shaking at 150 rpm/min. OD_{600} values were measured using Cytation 3 Multimode Imaging Reader (Biotek) and were adjusted by subtracting values from control wells. Growth rates were calculated as the slope of line obtained by natural log transformed growth values at the exponential growth phase (app. 1-4 days) using the formula [rate= {ln (OD day4) –ln (OD day1)}/time]. Biofilm was quantified using the crystal violet assay as previously described (Cruz et al., 2012) at the end of the growth curve experiment (eighth day). Settling rate, as a measure of cell to cell attachment, was determined by measuring OD_{600} of a cell suspension in a cuvette when cells appeared to settle exponentially {5-30 minutes post inoculation (mpi)}. Settling rate was calculated as in growth rate using the formula; [rate= {ln (OD 5mpi) –ln (OD 30mpi)}/time]. Twitching motility measurement was performed as previously described (Kandel et al., 2016). Experiments were repeated independently at least three times with at least three technical replicates per time.

Transmission Electron Microscopy (TEM)

X. fastidiosa strains WM1-1, AlmaEM3, and BBI64 were used with TEM imaging to observe the presence of type IV pilli (structure involved in natural competence and twitching motility). Strains were cultured in PD3 agar plates for three days from the second re-streak mentioned above. Cells from the edges of the colony were suspended in 100 µl of sterile water. Six µl of

this suspension was pipetted on a formvar-carbon coated TEM grid, and cells were allowed to settle for 10 min. The leftover liquid was blotted out with a filter paper. The grid was then negatively stained with six µl phopshotungistic acid (PTA) for two min, and after removing the excess PTA, grids were air dried and observed under Zeiss EM10 Transmission Electron Microscopy (Carl Zeiss, Jermany). Images were acquired with MaxIm DL software (Diffraction LTD, Canada).

Sequence comparision of genes of *X. fastidiosa* strains involved in natural competence

Sequence comparison of the genes involved in natural competence in *X. fastidiosa* (Kung and Almeida, 2014) strains that were positive and negative for natural competence was performed. Sequences of genes (*pilA, pilB, pilM, pilQ, pilO*, *recA*, *comA*, and *comF*) were extracted either from public databases or from reference mapping assembly (see above). Sequences were used in a multiple sequence alignment tool in Geneious. Mutations and insertion/deletions were analyzed. Insertion/deletions were further confirmed by Sanger sequencing of both forward and reverse fragments (Eurofin Genomics, USA) using specific primers (Supplemental table 3-3). Primers were designed with Primer3 Primer Design Tool in Geneious, and PCR was performed with standard protocol using iProof™ High-Fidelity PCR Kit (BioRad). PCR products were purified using PCR Extract Mini Kit (5 PRIME) before sequencing.

Test of intra- and intersubspecific homologous recombination

To assess IHR, *X. fastidiosa* strains belonging to subsp. *fastidiosa* and subsp. *multiplex* were used. In one set of experiments, subsp. *fastidiosa* strain WM1-1 was tagged with a kanamycin resistant marker at the *msrA* genomic region (*msrA*::Km-WM1-1) and subsp. *multiplex* strain

AlmaEM3 was tagged with chloramphenicol resistant marker at the NS1 site (NS1::Cm-AlmaEM3) using the natural competence protocol described above. These two mutant strains were co-cultured in MC for five days, and on solid agar plates for three days with PD3 medium as previously described (Kandel et al., 2016). Double antibiotic resistant recombinants were selected by plating the mixed culture in PW agar medium supplemented with both kanamycin and chloramphenicol. Experiments were performed once in the MC and three times with three replicates each time in the plate setting. Single strain spots were used as controls in the plates.

In another set of experiments, antibiotic/GFP marker tagged donor strains were heat-killed by incubating a suspension (OD_{600} ~0.8) at 90°C for 15 min. Heat-killed donor strains were spotted on PD3 plates on top of live recipient strains. NS1::Km-WM1-1, NS1::Km-EB92-1, and Temecula1-GFP from subsp. *fastidiosa*, and NS1::Km-AlmaEM3 from subsp. *multiplex* were used as donor strains while WM1-1, Temecula1, and EB92-1 (*fastidiosa*); and AlmaEM3, and BBI64 (*multiplex)* were used as recipient strains. The donor and recipients were mixed as in the plasmid recombination protocol described above. Experiments were repeated at least three times with three replications each time for WM1-1 and AlmaEM3 combinations. Recipient-only spots were included as controls. The complete killing of donor strains was confirmed by plating aliquots of 100 µl in PW plates.

Confirmation of intersubspecific recombination

To confirm that the recombinants acquired the antibiotic/GFP marker from the donor strain, PCR and sequencing approaches were performed. PCR confirmation was done as previously described (Kandel et al., 2016). Up- and down-stream flanking regions and antibiotic region of two selected recombinants of heat-killed NS1::Km-AlmaEM3 and WM1-1 (WM1-1 recombinant 1

and 2) and heat-killed NS1::Km-WM1-1 and AlmaEM3 (AlmaEM3 recombinant 1 and 2), were sequenced using Sanger sequencing platform as mentioned above using specific primers (Supplemental table 3-3). Primers were designed with Primer3 primer design tool in Geneious, and PCR was performed with standard protocol using iProof™ High-Fidelity PCR Kit (BioRad). PCR products were purified using PCR Extract Mini Kit (5 PRIME) before sequencing. Resulting sequences were concatenated, and aligned in a multiple sequence alignment together with the donor and recipient sequences to confirm the movement of antibiotic, up-stream, and down-stream region from the donor to the recipient. Selected recombinants of heat-killed Temecula-GFP donor and AlmaEM3 recipient were confirmed by both observing GFP fluorescence using Nikon Eclipse Ti Inverted Microscope (NIKON, Melville, NY), and precipitation test in Eppendorf tubes.

Virulence assessment *in vitro* and in planta

The recipient parent strains WM1-1, and AlmaEM3, and randomly selected recombinants of the two strains were tested *in vitro* for twitching motility, and settling as the parents differed in these two traits. Twitching motility and settling measurement was performed as described above. Virulence test of the parents and randomly selected recombinants (WM1-1 recombinant 2 and AlmaEM3 recombinant 2) was performed in the green house at Plant Science Research Centre (PSRC), Auburn University in a completely randomized design. The parent strains and their recombinants were inoculated in model plant tobacco (*Nicotiana tabaccum*) cultivar SR1 'Petite Havana', and host blueberry (*Vaccinium* 'Rebel' [PPA18, 138]). Disease incidence, and disease severity ratings on a scale of (0-7) for tobacco and blueberries, respectively were recorded as previously described (Oliver et al., 2014; Oliver et al., 2015). Briefly, for tobacco, #symptomatic

and #total leaves were counted, and % incidence was calculated for each plant as [(#symptomatic /#total leaves) × 100]. Symptom rating was performed weekly for 4-6 times after the appearance of first symptoms (~60 days post inoculation). For blueberry, severity scores of (0-7) were given to each of the two stems of each plant and average score of the two stems was calculated. AUDPC was calculated for each plant based on the weekly values of incidence rate for tobacco and average severity scores for blueberry. AUDPC values were calculated based on the midpoint rule method (Campbell and Madden, 1990); AUDPC = Σi^{n-1} $[(y_i+y_{i+1})/2]$ $(t_{i+1} - t_i)$ where n=number of times disease assessment was performed, y=score of incidence or severity for each plant, and t=time of assessment. Tobacco experiments were performed twice with ten plants per treatment and blueberry experiment was performed once with ten plants per treatment. Following disease scoring, selected plants were used to isolate the bacteria from petiole and midrib region of the leaves of plants inoculated with *X. fastidiosa* and buffer control as described in a previous study (Parker et al., 2012).

Statistical analysis

Recombination frequencies of the strains were log transformed and analyzed using PROC GLIMMIX in SAS 9.3 followed by post-hoc analysis using TUKEY-HSD (Honestly Significant Difference) at $P \leq 0.05$. Treatments or replicates that had zero frequency values were excluded from the analysis. Virulence traits (fringe width, settling rate, biofilm, AUDPC values) were also compared in SAS 9.3 with PROC GLIMMIX and TUKEY-HSD at $P \leq 0.05$. Correlation analysis among the variables was carried out in R 3.2.5 with cor.test and rcorr functions under Hmisc package (Ihaka and Gentleman, 1996). Nucleotide sequences of genes of natural

competence, flanking regions, and recombinant sequences were deposited in NCBI Genbank under accession numbers KY616728- KY616749.

Results

X. fastidiosa strains differed in their ability to undergo natural competence

Thirteen *X. fastidiosa* strains belonging to two subspecies were tested for natural competence using plasmids containing an antibiotic resistant marker flanked on either side by *X. fastidiosa* homologous genomic regions. Natural competence was confirmed in eleven of the thirteen strains tested, with recombination frequency ranging from below detection limit to 0.024 per recipient cell (Table 3-1). Two strains (BBI64 and Georgia Plum) did not recombine with any of the plasmids tested (Table 3-1). Generally, subspecies *fastidiosa* strains had higher recombination frequency compared to subspecies *multiplex* strains, although differences were not significant between the two groups. On comparing flanking DNA sequence homology of the recombination site, strains showed 94-100% identity with the homologous region of the donor plasmid (Supplemental fig. 3-1 and Supplemental table 3-2). Strains belonging to different subspecies had more divergence in the flanking region DNA sequences than the strains within the same subspecies (Supplemental table 3-2). Although no correlation was detected between recombination frequency and flanking region DNA similarity [pKLN61 (r=0.39, *P*=0.33); and pAX1.Cm (r=0.44, *P*=0.27)] due to strains within a subspecies having identical flanking DNA regions in most of the cases (Supplemental table 3-2), for a given strain (eg. strain Temecula1 with pMSRA-Km plasmid), recombination frequency was reduced when homology was low. However, strains that belonged to the same subspecies and had identical flanking region differed by orders of magnitude in recombination frequency (Table 3-1, Supplemental table 3-2). Strain

WM1-1, in particular, achieved highest recombination frequency with all the plasmids tested. Recombination occurred in ~2% of the cells when pKLN61 and pMSRA-Km plasmids were used as donor with WM1-1. Generally, recombination frequency was higher with pKLN61 plasmids than with other plasmids (Table 3-1).

Table 3-1: Recombination frequency of *X. fastidiosa* strains with various donor plasmids

Strains	subspecies/host plant	Recombination frequency [(mean ± SE) $\times 10^{-5}$][x]			
		Donor plasmid[y]			
		pAX1.Cm	pKLN61	pMOPB-Km	pMSRA-Km
WM1-1	*fastidiosa*/grape	[z]117±23.54[a]	2400±1200[a]	414.7±196.1[a]	1170±540.1[a]
Temecula1	*fastidiosa*/grape	14.83±6.83[b]	47.58±22.85[abcd]	0.60±0.19[b]	0.03±0.02[b]
Temecula1*	*fastidiosa*/grape	10.91±2.53[b]	21.54±8.67[bcd]	NA	NA
CCPM1	*fastidiosa*/grape	19.25±2.61[ab]	165±62.82[abc]	NA	NA
Chard1	*fastidiosa*/grape	18.66±2.73[ab]	50.94±16.80[abc]	NA	NA
Fetzer	*fastidiosa*/grape	0.83±0.15[bcd]	11.93±2.85[bcd]	NA	NA
ConnCreek	*fastidiosa*/grape	0.17±0.05[cd]	8.24±2.29[bcd]	NA	NA
EB92-1	*fastidiosa*/elderberry	0.017±0.01[d]	7.32±4.27[cd]	0.002±0.001[c]	0.21±0.21[b]
AlmaEM3	*multiplex*/blueberry	0.085±0.03[cd]	7.48±1.13[bcd]	0.01±0.006[bc]	0.07±0.04[b]
BB08-1	*multiplex*/blueberry	1.87±0.59[bc]	3.627±1.16[d]	NA	NA
BH Elm	*multiplex*/elm	0.27±0.18[bcd]	12.25±6.22[bcd]	NA	NA
Georgia Plum	*multiplex*/plum	0	0	NA	NA
BBI64	*multiplex*/blueberry	0	0	0	0

[x]Recombination frequency was calculated as the ratio of number of recombinants to total viable cells. SE indicates standard error for independent replicates. Experiments were repeated 2-6 times with at least three independent replicates per time with each type of plasmid, NA: data not available.

[y]Plasmids were constructed with homologous regions targeting different sections of the *X. fastidiosa* genome. For details see Materials and Methods.

[z]Log transformed frequency values for each donor plasmids were analyzed in SAS 9.3 with PROC GLIMMIX. Post hoc analysis was performed with Tukey HSD (Honestly Significant Difference) test at 5% significance level. Different letters in the same column represent significant difference in the recombination frequency.

Recombination frequency was associated to growth rate and twitching motility

Recombination frequencies of strains were analyzed for their association with biological parameters such as growth rate, twitching motility, biofilm formation, and cell-to-cell attachment (measured by settling rate). Growth curves (Fig. 3-1A) were generated by culturing the strains in 96 wells plates for eight days and measuring optical density at 600nm. By plotting natural log of OD_{600} values against time of growth, strains were estimated to attain exponential growth during 1-4 days post inoculation, from which growth rates were calculated (Fig. 3-1B). Growth curves and growth rate values showed that strains have different growth characteristics. Although, growth rates and recombination frequencies were significantly correlated for both plasmids tested with all strains [pKLN61 (r=0.61, *P*=0.026); and pAX1.Cm (r=0.65, *P*=0.015)], strains having lowest growth rates (AlmaEM3, BB08-1, BH Elm) were competent, while those having average growth rates (BBI64 and Georgia Plum) were not. Strains differed significantly in other parameters such as biofilm formation (Fig. 3-1C), settling rate (Fig. 3-1D), and twitching motility (Fig. 3-2A and 3-2B). Recombination frequency was correlated only with twitching motility [pKLN61 (r=0.59, *P*=0.034); and pAX1.Cm (r=0.71, *P*=0.006)], and not with biofilm [pKLN61 (r=0.00, *P*=0.99); and pAX1.Cm (r=0.05, *P*=0.87)], or settling rate [pKLN61 (r=-0.27, *P*=0.36); and pAX1.Cm (r=-0.36, *P*=0.22)]. Strains that lacked natural competence (BBI64, Georgia Plum) were also deficient in twitching motility (Fig. 3-2A and 3-2B). Transmission TEM imaging of selected strains differing in natural competence and twitching motility showed dense type IV pili (structure involved in twitching motility) for strain WM1-1, but no type IV pili was observed for strain BBI64 (Fig. 3-2C). Twitching motility of subsp. *fastidiosa* strains (except Fetzer and ConnCreek) was significantly higher than that of subsp. *multiplex* strains (Fig.

3-2A and 3-2B), while settling rate was higher with subsp. *multiplex* strains than with subsp. *fastidiosa* strains (Fig. 3-1D).

Insertion was detected in pili genes of non-competent and non-motile strains

Sequences of genes involved in natural competence and twitching motility of strains tested in this study were retrieved from whole genome sequences whenever available. All competence and pili genes screened (*pilA, pilB, pilM, pilN, pilO, pilQ, recA, comA,* and *comF*) were intact in all strains except BBI64, in which *pilQ* gene was found to be mutated. The BBI64 genomic reads, mapped to the reference Temecula1 genome, showed an insertion in *pilQ* coding region (PD1691) (Supplemental fig. 3-2), which was further confirmed by sequencing the *pilQ* fragment. BBI64 sequences, at other competence related regions, had identical sequence to at least one another naturally competent and motile strain of subspecies *multiplex*. This suggested that the insertion at *pilQ* rendered BBI64 non-competent and non-motile. Another strain (*pglA*$^{-}$-KmR) that lacked twitching and competence in a previous study (Kandel et al., 2016) showed an insertion in *pilM* (PD1695) coding region, which was also confirmed by further sequencing (data not shown). No frame shift mutation was observed in the competence-related genes in another non-competent and non-motile strain Georgia Plum (data not shown).

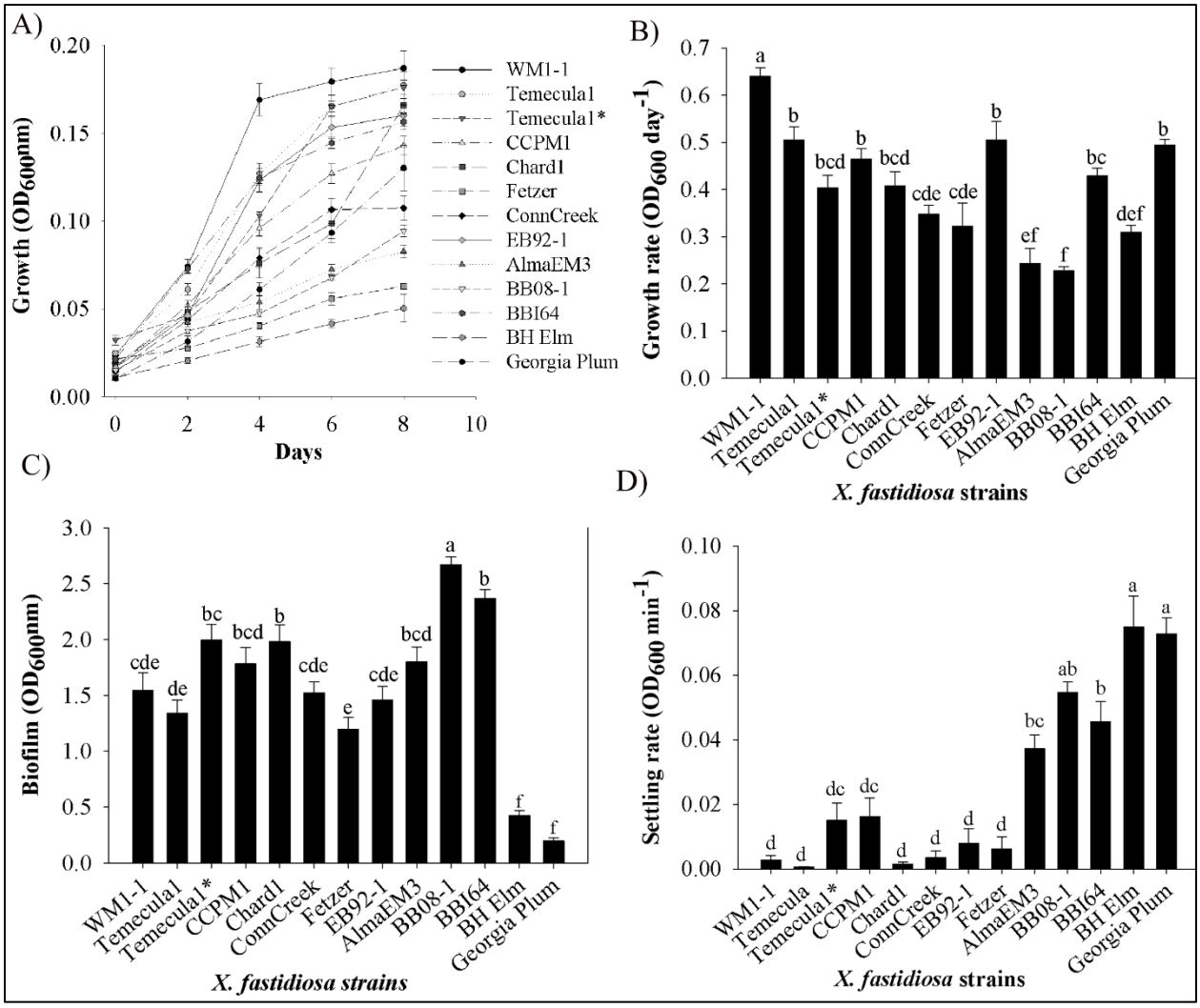


Fig. 3-1. Growth curves **(A)**, growth rates **(B)**, biofilm formation **(C)**, and settling rates **(D)** of *X. fastidiosa* strains used in this study. Growth curves were generated by culturing the bacteria in 96 wells plates and measuring OD_{600} values each day during eight days. Growth rate was calculated from the growth curve at the exponential growth phase (1-4 days). Biofilm was measured using crystal violet in the 96 wells plates at the end of growth curve experiment. Settling rate was measured by suspending the bacteria in a cuvette in 1ml of PD3 and measuring OD_{600} values at the interval of 5 min for 30 min. Each data point indicates mean value from at least three independent experiments. Error bars indicate standard errors. Different letters on top of column bars indicate significant difference as analyzed in SAS 9.3 with PROC GLIMMIX. Post hoc analysis was performed with Tukey HSD at 5% significance level.

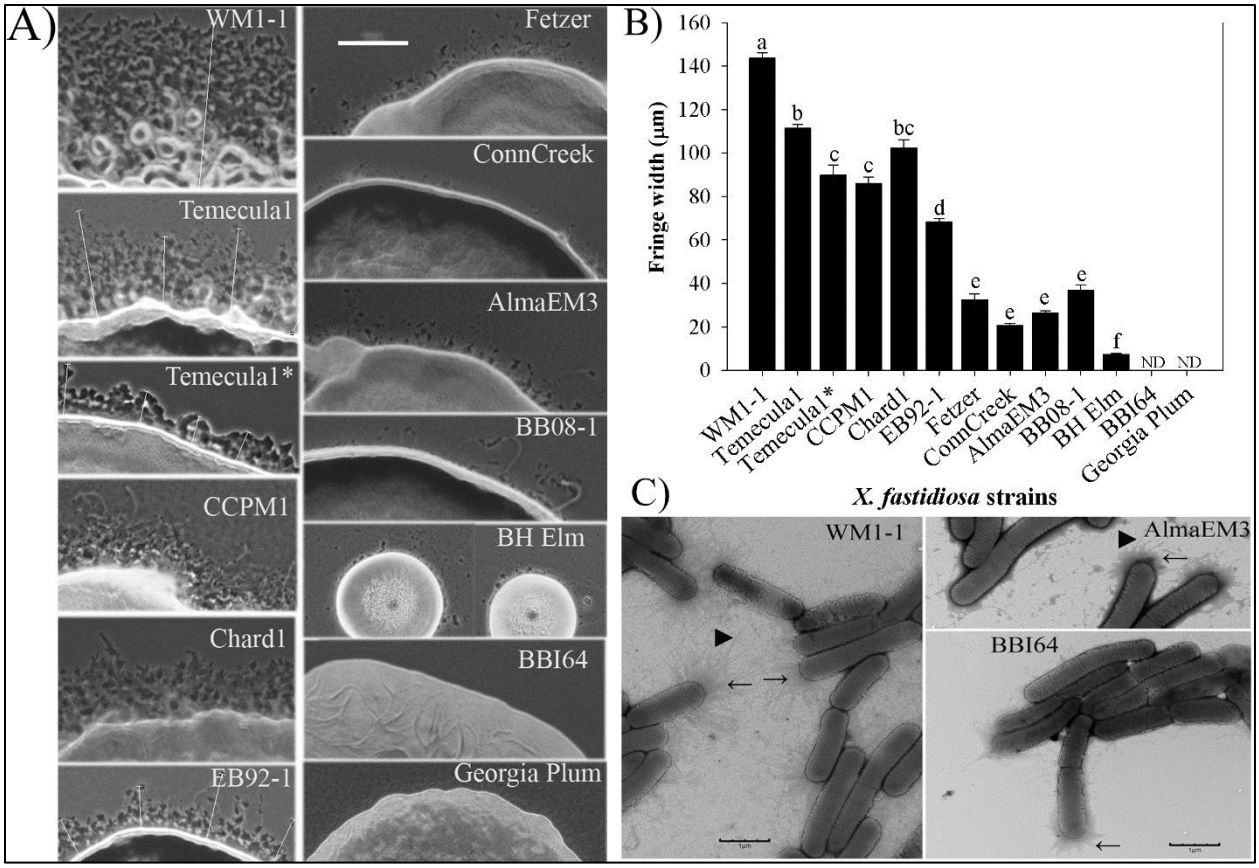


Fig. 3-2. Twitching motility and TEM of *X. fastidiosa* strains used in this study. Twitching motility was assessed by measuring the fringe width of the bacterial colonies spotted on PD3 plates after 3-5 days of incubation. Horizontal white bar on top right panel indicates 100µm **(A)**. Measurements included at least three biological replicates with four technical replicates per time except for BH Elm for which motility was measured only once with three technical replicates. Twitching motility differed among strains as analyzed in SAS 9.3 with PROC GLIMMIX. Post hoc analysis was performed with Tukey HSD at 5% significance level. Different letters on top of bar column represent significant difference in twitching motility. Error bars indicate standard error for independent replicates. ND: motility was not detected **(B)**. TEM images of selected strains. Strain WM1-1 and AlmaEM3 exhibited dense pili both long (type IV pili indicated by '►') and short (type I pili indicated by '←' and '→'), while no long pili were observed for BBI64. Horizontal black line indicates 1µm, magnification 12.5K **(C)**.

Intersubspecific homologous recombination occurred between *X. fastidiosa* subsp. *fastidiosa* and *X. fastidiosa* subsp. *multiplex*

IHR was confirmed by mixing different combinations of donor and recipient strains from two subspecies of *X. fastidiosa* under various growth conditions. First, different antibiotic resistant donor and recipient strains from subspecies *fastidiosa* and *multiplex* were co-cultured in MC and on agar plates with PD3 medium, and acquisition of double antibiotic resistance was assessed. Secondly, marker (kanamycin/GFP) tagged, heat-killed donor and wild-type live recipient strains from two subspecies were mixed on agar plates of PD3, and acquisition of the marker was assessed. IHR readily occurred between the two subspecies in MC and agar plates with both-live and live-dead combinations of donor and recipient (Fig. 3-3). From both-live experiments, on co-culturing *msrA*::Km-WM1-1 (subsp. *fastidiosa*) and NS1::Cm-AlmaEM3 (subsp. *multiplex*), 179 ±33 recombinant colonies per 100µl of culture with a frequency of 5.31×10^{-6} in MC, and 96±8 recombinant colonies per 100µl of culture with a frequency of 1.1×10^{-5} on agar plates were obtained (Fig. 3-3A). With live-dead combinations, strains WM1-1 and Temecula1 (subsp. *fastidiosa*) readily recombined DNA from heat killed NS1::Km-AlmaEM3 donor (subsp. *multiplex*), and strain AlmaEM3 (subsp. *multiplex*) recombined DNA from heat-killed NS1::Km-WM1-1, NS1::Km-EB92-1, and Temecula1-GFP (subsp. *fastidiosa*) (Fig. 3-3). The number of recombinants varied (1-412 colonies per 100µl of culture) depending on donor-recipient combinations. Strain EB92-1 did not recombine with DNA from heat-killed donor but could act as donor for strains of both subspecies (Fig. 3-3). Strain BBI64 was also unable to recombine with heat-killed donor cells. This strain was not used as heat-killed donor as it could not be tagged with the antibiotic-resistance gene due to lack of recombination. Recombination detected

with the heat-killed donor cells demonstrates that natural competence is the horizontal gene transfer mechanism involved.

The occurrence of IHR was further confirmed by PCR, sequencing, and phenotypic assessment. PCR confirmed insertion of antibiotic resistant cassette in the recombinant strain (data not shown). Sequence comparison of the flanking regions, in two recombinants each of heat-killed NS1::Km-AlmaEM3 donor and WM1-1 recipient (WM1-1 recombinants 1 and 2); and heat-killed NS1::Km-WM1-1 donor and AlmaEM3 recipient (AlmaEM3 recombinants 1 and 2) together with the donor and recipient strains, detected recombination of antibiotic cassette region as well as upstream and downstream flanking regions (Fig. 3-4, Supplemental table 3-4). The recombinant of AlmaEM3 with heat-killed Temecula1-GFP donor was kanamycin resistant and exhibited green fluorescence (characteristic of donor), and fast precipitation (characteristic of recipient) further confirming occurrence of IHR at kanamycin/GFP region (Fig. 3-5).

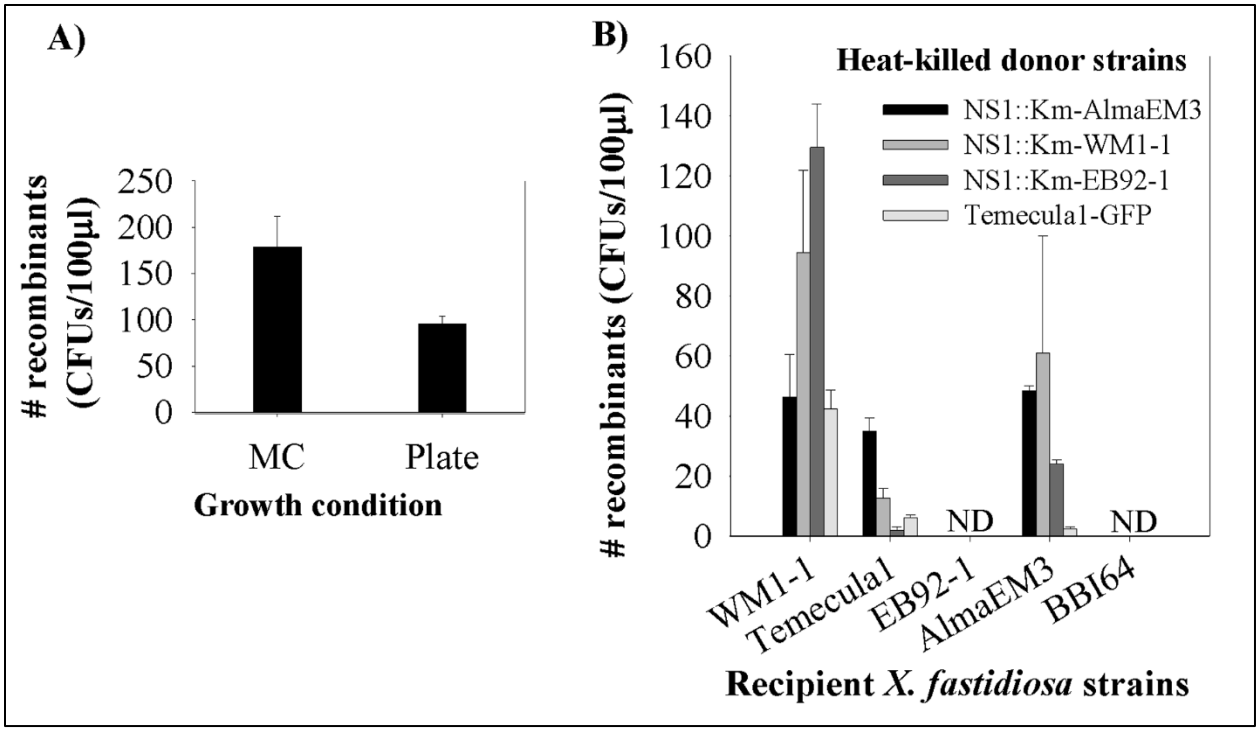


Fig. 3-3. Intra- and Intersubspecific recombination in *X. fastidiosa.* Recombination between antibiotic-resistance marker-tagged *X. fastidiosa* strains belonging to subspecies *fastidiosa (msrA*::Km-WM1-1) and subspecies *multiplex* (NS1::Cm -AlmaEM3) when cultured together in microfluidic chamber (MC) and agar plates with PD3**.** Double antibiotic resistant recombinant colonies were recovered from both settings **(A)**. Recombination of *X. fastidiosa* strains with heat-killed donor cells. Donor strains were tagged with kanamycin resistant marker at the neutral (NS1) site using pAX1.Km plasmids except for the Temecula1-GFP strain, and heat-killed by incubating at 90°C for 15 min. Heat-killed donor and recipient strains were mixed on spots of PD3 agar plates and incubated at 28°C for 3-4 days. Recombinants were selected by plating the mixed spots in PW-Km agar plates. Intersubspecific recombination was confirmed by mixing heat-killed strains of subspecies *fastidiosa* (NS1::Km -WM1-1, Temecula1-GFP, and NS1::Km-EB92-1) with a recipient of subspecies *multiplex* (AlmaEM3); and heat-killed NS1::Km-AlmaEM3 with recipients of subspecies *fastidiosa*. Bars represent mean number of recombinants obtained for each combination of recipient and donor. Experiments were repeated at least three times with three replicates each. ND: recombinants were not detected for these strains when used as recipients **(B)**.

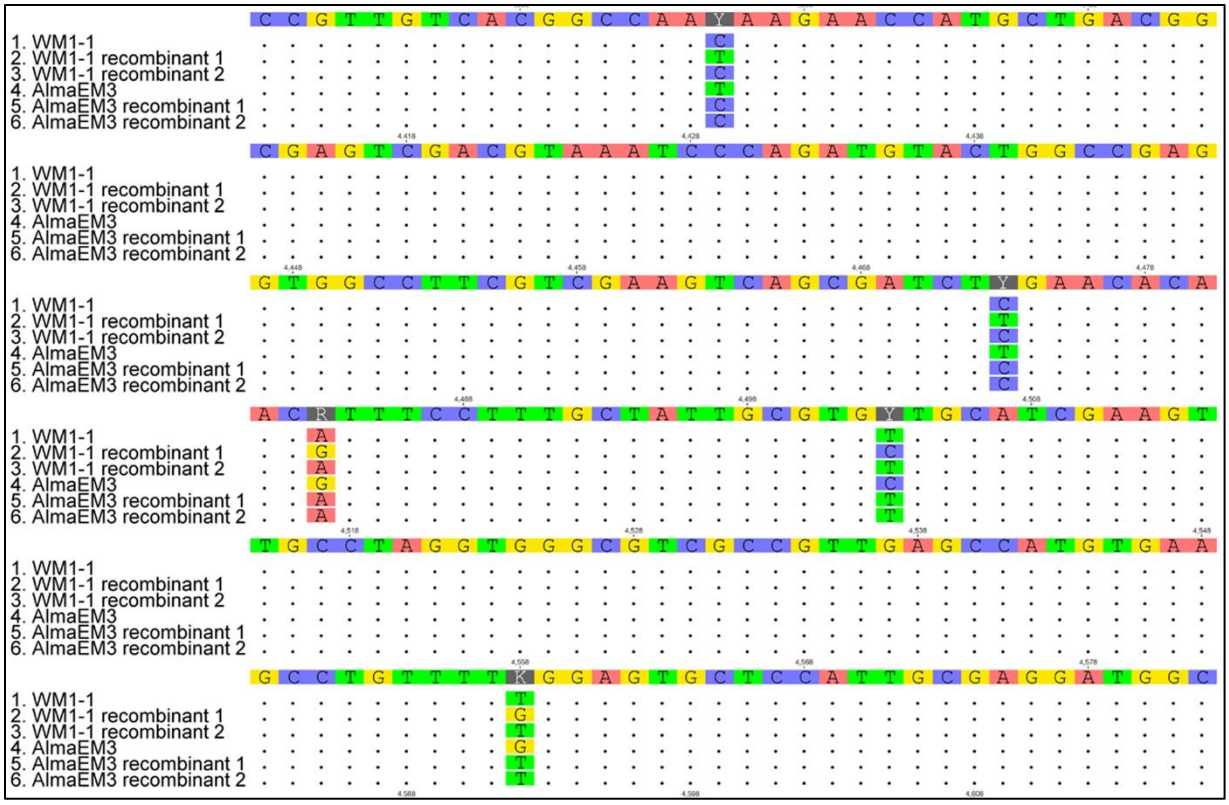


Fig. 3-4. Confirmation of intersubspecific recombination by sequencing**.** Multiple sequence alignment of sequences from donor, recipient, and recombinant strains in the upstream and downstream region of antibiotic insertion was performed. WM1-1 and AlmaEM3 are wild type strains belonging to subsp. *fastidiosa* and *multiplex*, respectively. WM1-1 recombinants were generated by mixing wild type WM1-1 with heat-killed NS1::Km-AlmaEM3, and AlmaEM3 recombinants were generated by mixing wild type AlmaEM3 with heat-killed NS1::Km-WM1-1. The shown region is 1037bp downstream of the kanamycin resistant gene insertion site and 247bp downstream of the homologous region used in the construct. Dots indicate no difference between donor and recipients. Recombination occurred in WM1-1 recombinant 1 and both AlmaEM3 recombinants but not in WM1-1 recombinant 2 at the region shown here (see supplemental table 3-4 for information on a bigger region).

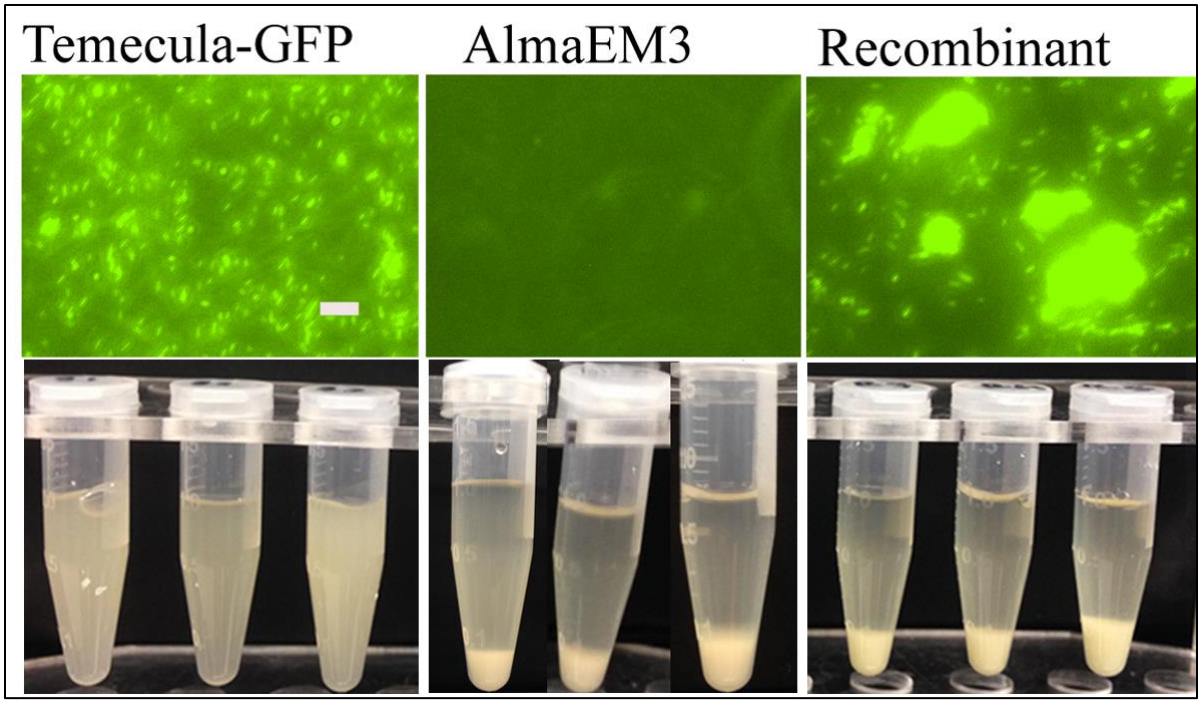


Fig. 3-5. Confirmation of inter-subspecific recombination in *X. fastidiosa* by movement of green fluorescence gene from Temecula-GFP donor (subsp. *fastidiosa*) to AlmaEM3 recipient (subsp. *multiplex*). Upper panel shows fluorescence characteristics of donor, recipient, and recombinant as observed under a fluorescence microscope, and lower panel shows the precipitation behavior of respective strains after suspending the cells in PD3 and allowing settling for ~1 hour. Recombinant showed fluorescence (characteristic of donor) as well as fast settling (characteristic of recipient), confirming that recombination occurred at the GFP locus. White line =10µm.

Randomly selected intersubspecific recombinants did not exhibit alterations in virulence phenotypes

Randomly selected recombinants generated by intersubspecific recombination on mixing whole genomic DNA of heat-killed NS1::Km-AlmaEM3 donor and WM1-1 recipient (WM1-1 recombinants), and heat-killed NS1::Km-WM1-1 donor and AlmaEM3 recipient (AlmaEM3 recombinants) were assessed for virulence phenotypes by measuring their settling rate and twitching motility *in vitro*, and plant virulence. The recombinants and parameters for their assessment were selected based on the differences observed in their parent strains (Fig. 3-1D, 3-2A and 3-2B; and 3-6A and 3-6B). On analyzing the results, the phenotype of recombinants was found to be like that of the recipient parent. WM1-1 recombinants had similar settling rate (0.0023±0.0007 OD min^{-1}, n=4), and twitching motility (177±20.31 µm, n=10), to that of WM1-1 (Fig. 3-1D, 3-2A, and 3-2B); and AlmaEM3 recombinants had similar settling rate (0.04±0.002 OD min^{-1}, n=5) and twitching motility (40±4.5 µm, n=10) to that of AlmaEm3 (Fig. 3-1D, 3-2A and 3-2B).Virulence scores from in planta disease assessment were also not significantly different between the recombinants and recipient parent. Both recipients (WM1-1 and AlmaEM3), and their recombinants colonized and produced characteristic symptoms of leaf scorch, yellowing, and wilting in tobacco (Fig. 3-6C). AUDPC values in tobacco derived from the incidence rate were higher for WM1-1 than for AlmaEM3 but not significantly different between the recipient strain and respective recombinants (Fig. 3-6A). In blueberry, AlmaEM3 and AlmaEM3 recombinant had significantly higher AUDPC values derived from the weekly disease severity ratings compared to WM1-1 and its recombinant (Fig. 3-6B). Symptoms such as leaf yellowing, leaf reddening, leaf scorch, match stick, and leaf drop were observed with the plants inoculated with AlmaEM3 and its recombinant only (Fig. 3-6C). Isolation results

confirmed infection by both parent strains and recombinants in tobacco, but only AlmaEM3 and its recombinant in blueberry. Buffer inoculated plants were free of *X. fastidiosa* in both tobacco and blueberry (data not shown).

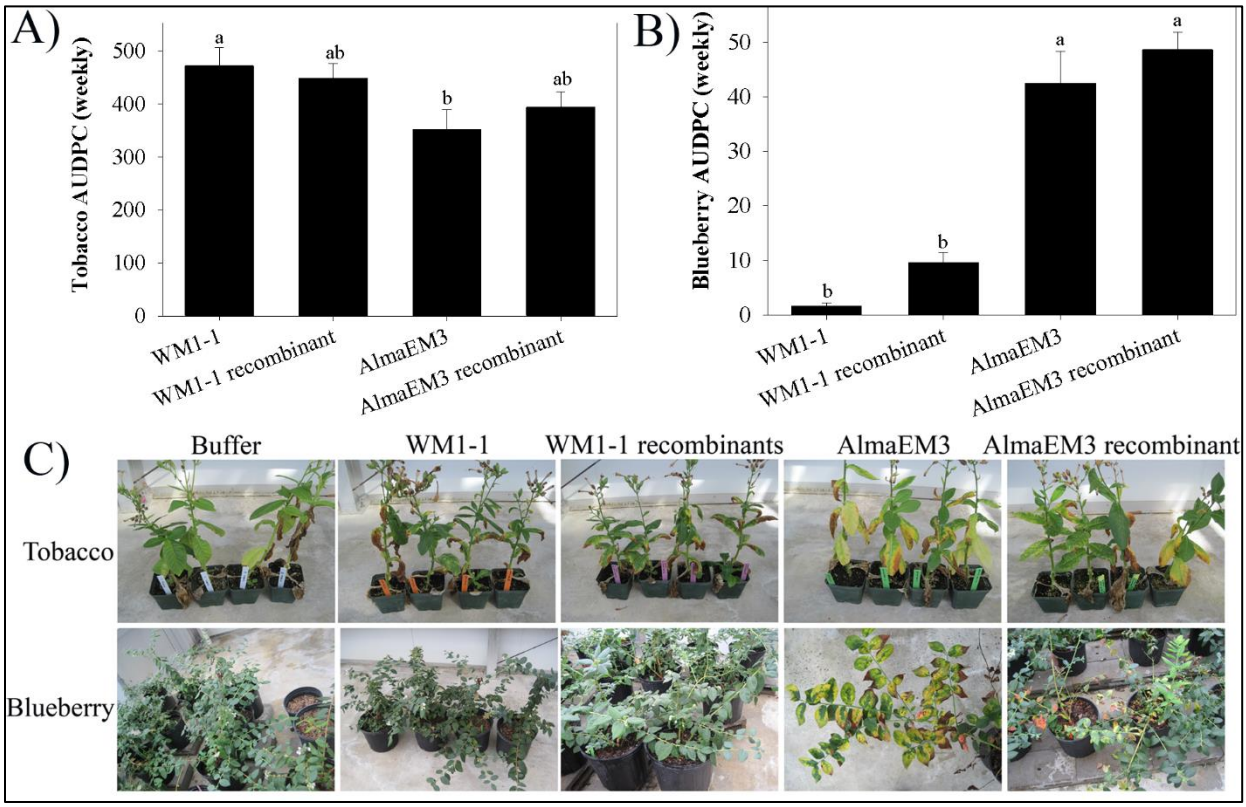


Fig. 3-6. Virulence test of donor, recipient, and recombinant strains in the greenhouse. Area Under Disease Progress Curve (AUDPC) values were different between parents but did not differ between the recipient and the recombinant strains suggesting that recombination did not change virulence behavior **(A and B)**. Tobacco experiments were performed twice with ten plants per treatment (data are shown from a representative experiment), and blueberry experiment was performed once with ten plants in each treatment. Both WM1-1 and AlmaEM3 and their recombinants produced symptoms in tobacco while only AlmaEM3 and their recombinants produced symptoms in blueberry **(C)**.

Intersubspecific recombination occurred in short and random fragments

Length of recombined region was estimated by further sequencing the flanking region of kanamycin gene insertion site in four of the WM1-1 and AlmaEM3 recombinants previously mentioned (including WM1-1 recombinant 2 and AlmaEM3 recombinant 2 that were used for in planta disease assessment). Length of recombined region in a single recombination event ranged from 739bp-3486bp with recombination extending randomly in the up- and downstream flanking regions (Supplemental table 3-4). Two independent recombination events separated by a ~1kb region were detected in WM1-1 recombinant 2 (Supplemental table 3-4). Possible recombination that could have occurred at other genomic positions in these recombinants was not assessed.

Discussion

Previous studies on natural competence in *X. fastidiosa* were based on a few strains from a single subspecies (*X. fastidiosa* subsp. *fastidiosa*), although recombination among strains of other subspecies has been described (Nunes et al., 2003; Nunney et al., 2012; Nunney et al., 2013; Nunney et al., 2014b; Nunney et al., 2014c; Coletta-Filho et al., 2017). On testing natural competence in 13 different strains, almost ubiquitous natural competence ability was detected. The frequency of recombination varied among strains even for a single genomic region (i.e., same donor plasmid) as in other naturally competent bacteria (Fujise et al., 2004; Coupat et al., 2008; Bosse et al., 2009; Evans and Rozen, 2013). Flanking region DNA similarity of the strains was not correlated with the recombination frequency, but most of the strains within a subspecies had identical flanking region. A clearer understanding on the rate of recombination and homology between recombining DNA could have been obtained by using donor DNA containing different level of similarity with the recipient strain at a given recombination region. However,

this was not performed in this study. Even if differences in recombination frequency, especially between strains of different subspecies, could be due to differences in homology between donor and recipient DNA sequences or differences in growth rates that showed positive correlation with recombination frequency, these parameters did not explain the non-competency of two strains that had average growth and similar sequence homology compared to the competent strains. Also, since growth of the strains was measured by optical density and strains appeared to differ in the rate of cell to cell attachment, growth values could have been biased, especially for the strains that showed high rate of precipitation (due to stronger cell to cell attachment). This was not further investigated, as it was beyond the scope of this study.

On further testing of other biological traits, twitching motility was significantly correlated with recombination frequency. Strain WM1-1 had the highest recombination frequency and showed highest motility among strains, while the two non-competent strains were non-motile. Positive correlation between recombination frequency and twitching motility was also suggested in our previous study using different media components (Kandel et al., 2016). Since components of type IV pili are involved in both natural competence and twitching motility in several naturally competent gram-negative bacteria (Seitz and Blokesch, 2013a), including *X. fastidiosa* (Kung and Almeida, 2014), the activity of type IV pili could govern both of these phenomena. On further analysis of competence and pili genes, defective pili genes were detected in the non-competent strains. One of the defective proteins detected was PilQ (strain BBI64), a member of secretin family that forms the secretin pore of the outer membrane (Collins et al., 2001), and is involved in type IV pili biogenesis and importing extra-cellular DNA into the periplasmic space (Seitz and Blokesch, 2013b). Previous studies in *X. fastidiosa* have demonstrated that *pilQ*-

mutants are non-motile (Meng et al., 2005) and non-competent (Kung and Almeida, 2014). Hence, we predict that the insertion in *pilQ* coding region is responsible for the lack of twitching and natural competence as BBI64 is unable to secrete the type IV pili. The lack of type IV pili was confirmed by TEM imaging. Motility has been described as a major virulence trait for *X. fastidiosa* (Meng et al., 2005; De La Fuente et al., 2007). BBI64 has no motility and WM1-1 has the highest motility in this study. Consistent with the critical role of twitching in virulence, BBI64 had reduced virulence while WM1-1 was highly virulent (Oliver et al., 2014; Oliver et al., 2015). A further observation supporting the correlation between twitching and natural competence was the fact that the Fetzer strain showed recombination in this study, while a mutant in the polygalacturonase gene *pglA* of this strain did not (Kandel et al., 2016). On closer examination Fetzer is motile (this study) while the *pglA* mutant is not (Kandel et al., 2016). Sequence data showed that *pglA* mutant had an insertion in *pilM*, a type IV pili biogenesis gene that was shown to be involved in twitching motility of *Acidovorax avanae* in a previous study (Bahar et al., 2009), most probably causing the lack of movement in this strain.

Additional factors could be involved in causing differences in natural competence of *X. fastidiosa* strains. *X. fastidiosa* genomes contain high levels of phage and phage-like regions (Nunes et al., 2003; Van Sluys et al., 2003; Varani et al., 2008), and natural competence could be a mechanism to help cells eliminate new integration of these regions by recombining the homologous DNA without phage sequences, as suggested by a recent study (Croucher et al., 2016). Other studies have reported restriction-modification (R-M) systems limiting transformation frequency (Niza et al., 2016). In this study, although all donor plasmids were extracted from an *E. coli* strain expressing *X. fastidiosa* DNA methyl transferase (Matsumoto

and Igo, 2010), it is possible that different strains especially from different genetic backgrounds, possess different forms of R-M systems, this could lead to differences in the amount of DNA available for recombination causing differences in recombination frequency. In this regard, a previous study has reported inability of a plasmid isolated from a citrus-infecting strain to transform a grape strain (Guilhabert and Kirkpatrick, 2003), suggesting existence of specific recognition mechanisms to differentiate DNA from self or foreign sources. Sequence analysis and annotation of the *X. fastidiosa* Temecula1 genome predicts at least four different types of R-M systems (Matsumoto and Igo, 2010). Future studies focused on these specific topics could explain the differences in recombination frequencies observed amongst *X. fastidiosa* strains.

Difference in recombination frequencies based on genomic positions was previously reported in *Ralstonia solanacearum* (Fall et al., 2007) with positions containing recombination hot-spots (*recA* and *mutS*) showing the highest frequency (Fall et al., 2007). In this study, higher recombination frequency was observed for pKLN61, plasmid that recombines in the region of *rpfF* gene, a diffusible signaling factor involved in cell-cell communication of *X. fastidiosa* (Newman et al., 2004); compared to pAX1.Cm that recombines at a neutral site (Matsumoto et al., 2009); and pMOPB-Km, and pMSRA-Km that recombine at regions whose functions are being characterized. Differences in the length of homologous flanking region and non-homologous insert have been found to influence recombination frequency in a previous study (Kung et al., 2013). However, the upstream and downstream flanking region length was higher in pAX1.Cm (759-790bp), pMOPB-Km (~800bp), and pMSRA-Km (~800bp) than for pKLN61 (~350bp at one end and ~750bp at the other end). The length of non-homologous insert between the homologous flanking regions was similar (~1,200bp) and the size of the plasmids is also

comparable (~5kb). Moreover, flanking region DNA sequence identity between the donor plasmids and recipient strains at these positions was also similar (except for EB92-1 at pAX1.Cm, and Temecula1 for pMSRA-Km). This suggests that the difference in recombination frequency at different genomic position is not associated with the characteristics of plasmid regions, and it remains to be determined if this difference holds any evolutionary significance.

Natural competence has been proposed to bring adaptive changes to the recipient bacteria such as repair of damaged DNA (Dorer et al., 2010) and generation of genetic diversity that can lead to adaptation (Baltrus et al., 2008). For the generation of adapted strains, recombining regions should come from a more successful and genetically distinct donor (Coupat et al., 2011). This could be possible when closely related but genetically different strains of a same species co-exist in a single habitat. Detection of IHR in *X. fastidiosa* by MLST/MLSA studies (Nunney et al., 2012; Nunney et al., 2014b; Nunney et al., 2014c) supported this possibility. In fact, these studies proposed IHR leading to plant host shift of *X. fastidiosa* in citrus (Nunney et al., 2012), mulberry (Nunney et al., 2014c), and blueberry (Nunney et al., 2014b). Moreover, mixed infection of the two subspecies have been suggested by previous studies. For example, almond leaf scorch strains isolated from the same orchard were found to be genetically different (Chen et al., 2005; Chen et al., 2010), and were grouped into two different subspecies; subsp. *fastidiosa* and *multiplex* (Yuan et al., 2010). Infection of a plum tree showing leaf scorch symptoms by subsp. *multiplex* and subsp. *pauca* strains was also reported by a recent study (Coletta-Filho et al., 2016). Results of this and previous studies (Oliver et al., 2014; Oliver et al., 2015) demonstrated that certain plants serve as hosts for strains from multiple subspecies. In addition, vectors of *X. fastidiosa* are found to be distributed worldwide in both temperate and tropical

climates and, unlike with plant hosts, exhibit no specificity for pathogen genotype (Almeida et al., 2005). In fact, a species of the vector (*Homalodisca vitripennis)* was able to transmit four subspecies of *X. fastidiosa* (Almeida and Nunney, 2015). All these observations suggest that strains belonging to different subspecies may co-exist within the same habitat (plant or insect), providing opportunities for recombination.

Although IHR was detected between subspecies when whole genomes of the donor (heat-killed) and recipient were mixed, recombinants did not differ significantly with the parent in virulence phenotypes suggesting that recombination did not bring phenotypic changes. On analyzing the flanking homologous region of recombination, 0.7-4kb regions were detected to have recombined, but the size could be higher as up to 80kb has been demonstrated to recombine by natural competence in *R. solanacearum* with the recombinant strain showing increased virulence (Coupat et al., 2011). Moreover, recombination could also occur at genomic regions other than the antibiotic insertion site as the whole genomic DNA was used as donor to generate these recombinants. However, due to the bias of the method used here that selected recombinants acquiring antibiotic resistance, recombination that occurred at other regions in the genome may not have been present in the recombinants selected for assessment in this study. In addition, only the minimum size of recombination events could be estimated, based on existing polymorphisms. Nonetheless, by targeting various genomic regions, it was confirmed that recombination occurred at multiple regions. Future studies by optimizing the selection procedures for recombinants in the context of pathogenicity to plants could reveal changes in virulence due to IHR in *X. fastidiosa.*

Interestingly, IHR was bidirectional meaning that both subspecies could act as both donor and recipient for one another. The evidence from field observation of *X. fastidiosa* disease emergence in new plant species and the detection of IHR in strains isolated from these infections by MLST/MLSA (Coletta-Filho et al., 2017), and from confirmation of natural competence in habitats mimicking natural environments (Kandel et al., 2016) to experimental validation of IHR (this study), suggests that IHR is occurring in nature and may have broader evolutionary implications in *X. fastidiosa* disease dynamics. In conclusion, *X. fastidiosa* strains showed extensive natural competence abilities, and the recombination potential differed among strains. Moreover, intersubspecific recombination occurred readily between *X. fastidiosa* subsp. *fastidiosa* and subsp. *multiplex* strains. These results emphasize the importance of quarantine measures to limit the introduction of novel genotypes of *X. fastidiosa* in areas with pre-existing infection. Moreover, measures to isolate host plants of different subspecies may be required to prevent mixed infections, minimizing the risk of generation of novel and virulent genotypes of *X. fastidiosa* by recombination.

Acknowledgements

This project was funded by HATCH AAES (Alabama Agricultural Experiment Station) program, and Agriculture and Food Research Initiative competitive grant no. 2015-67014-23085 from the USDA National Institute of Food and Agriculture. We thank Michael E. Miller, director of the AU Research and Instrumentation Facility, for his help with TEM; Steve Lindow (UC Berkeley) for providing a GFP-marked strain; Jennifer Parker for her help with Miseq library preparation and genome sequencing; Adam Retchless and Drake Stenger for sharing unpublished genome data; and Luisa Cruz for providing *mopB* and *msrA* mutant constructs.

References

Almeida, R.P., Blua, M.J., Lopes, J.R., and Purcell, A.H. 2005. Vector transmission of *Xylella fastidiosa*: applying fundamental knowledge to generate disease management strategies. Ann Entomol Soc Am 98:775-786.

Almeida, R.P.P., and Purcell, A.H. 2003. Transmission of *Xylella fastidiosa* to grapevines by *Homalodisca coagulata* (Hemiptera : Cicadellidae). J Econ Entomol 96:264-271.

Almeida, R.P.P., and Nunney, L. 2015. How do plant diseases caused by *Xylella fastidiosa* emerge? Plant Dis 99:1457-1467.

Almeida, R.P.P., Nascimento, F.E., Chau, J., Prado, S.S., Tsai, C.W., Lopes Sí, A., and Lopes, J.R.S. 2008. Genetic structure and biology of *Xylella fastidiosa* strains causing disease in citrus and coffee in Brazil. Appl Environ Microbiol 74:3690-3701.

Antonova, E.S., and Hammer, B.K. 2015. Genetics of natural competence in *Vibrio cholerae* and other vibrios. Microbiol spectr 3:VE-0010-2014.

Bahar, O., Goffer, T., and Burdman, S. 2009. Type IV pili are required for virulence, twitching motility, and biofilm formation of *Acidovorax avenae* subsp. *citrulli*. Mol Plant Microbe Interact 22:909-920.

Baltrus, D.A., Guillemin, K., and Phillips, P.C. 2008. Natural transformation increases the rate of adaptation in the human pathogen *Helicobacter pylori*. Evolution 62:39-49.

Bosse, J.T., Sinha, S., Schippers, T., Kroll, J.S., Redfield, R.J., and Langford, P.R. 2009. Natural competence in strains of *Actinobacillus pleuropneumoniae*. FEMS Microbiol Lett 298:124-130.

Cameron, A.D.S., Volar, M., Bannister, L.A., and Redfield, R.J. 2008. RNA secondary structure regulates the translation of sxy and competence development in *Haemophilus influenzae*. Nucleic Acids Res 36:10-20.

Campbell, C.L., and Madden, L.V. (1990). Introduction to Plant Disease Epidemiology (New York: John Wiley & Sons).

Chang, C.J., Donaldson, R., Brannen, P., Krewer, G., and Boland, R. 2009. Bacterial Leaf Scorch, a new blueberry disease Caused by *Xylella fastidiosa*. Hortscience 44:413-417.

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

Chen, J., Groves, R., Civerolo, E.L., Viveros, M., Freeman, M., and Zheng, Y. 2005. Two *Xylella fastidiosa* genotypes associated with Almond Leaf Scorch disease on the same location in California. Phytopathology 95:708-714.

Chen, J., Xie, G., Han, S., Chertkov, O., Sims, D., and Civerolo, E.L. 2010. Whole genome sequences of two *Xylella fastidiosa* strains (M12 and M23) causing almond leaf scorch disease in California. J Bacteriol 192:4534.

Coletta-Filho, H., Francisco, C., Lopes, J., Muller, C., and Almeida, R. 2017. Homologous recombination and *Xylella fastidiosa* host-pathogen associations in South America. Phytopathology 107:305-312.

Collins, R.F., Davidsen, L., Derrick, J.P., Ford, R.C., and Tonjum, T. 2001. Analysis of the PilQ secretin from *Neisseria meningitidis* by transmission electron microscopy reveals a dodecameric quaternary structure. J Bacteriol 183:3825-3832.

Coupat, B., Bernillon, D., Guidot, A., Prior, P., Nesme, X., and Bertolla, F. 2011. *Ralstonia solanacearum* virulence increased following large interstrain gene transfers by natural transformation. Mol Plant Microbe Interact 24:497-505.

Coupat, B., Chaumeille-Dole, F., Fall, S., Prior, P., Simonet, P., Nesme, X., and Bertolla, F. 2008. Natural transformation in the *Ralstonia solanacearum* species complex: number and size of DNA that can be transferred. FEMS Microbiol Ecol 66:14-24.

Croucher, N.J., Mostowy, R., Wymant, C., Turner, P., Bentley, S.D., and Fraser, C. 2016. Horizontal DNA transfer mechanisms of bacteria as weapons of intragenomic conflict. PLoS Biol 14:e1002394.

Cruz, L.F., Cobine, P.A., and De La Fuente, L. 2012. Calcium increases *Xylella fastidiosa* surface attachment, biofilm formation, and twitching motility. Appl Environ Microbiol 78:1321-1331.

Cruz, L.F., Parker, J.K., Cobine, P.A., and De La Fuente, L. 2014. Calcium-enhanced twitching motility in *Xylella fastidiosa* is linked to a single PilY1 homolog. Appl Environ Microbiol 80:7176-7185.

Davis, M., Purcell, A., and Thomson, S. 1980. Isolation media for the Pierce's disease bacterium. Phytopathology 70:425-429.

Davis, M., French, W., and Schaad, N. 1981. Axenic culture of the bacteria associated with phony disease of peach and plum leaf scald. Curr Microbiol 6:309-314.

De La Fuente, L., Burr, T.J., and Hoch, H.C. 2007. Mutations in type I and type IV pilus biosynthetic genes affect twitching motility rates in *Xylella fastidiosa*. J Bacteriol 189:7507-7510.

De La Fuente, L., Parker, J.K., Oliver, J.E., Granger, S., Brannen, P.M., van Santen, E., and Cobine, P.A. 2013. The bacterial pathogen *Xylella fastidiosa* affects the leaf ionome of plant hosts during infection. PLoS One 8:e62945.

Dorer, M.S., Fero, J., and Salama, N.R. 2010. DNA damage triggers genetic exchange in *Helicobacter pylori*. PLoS Pathog 6:e1001026.

European Food Safety Authority. (2016). Update of a database of host plants of *Xylella fastidiosa*: 20 November 2015. In EFSA Journal, pp. 4378-n/a.

Evans, B.A., and Rozen, D.E. 2013. Significant variation in transformation frequency in *Streptococcus pneumoniae*. ISME J 7:791-799.

Fall, S., Mercier, A., Bertolla, F., Calteau, A., Gueguen, L., Perriere, G., Vogel, T.M., and Simonet, P. 2007. Horizontal gene transfer regulation in bacteria as a "spandrel" of DNA repair mechanisms. PLoS One 2:e1055.

Fujise, O., Lakio, L., Wang, Y., Asikainen, S., and Chen, C. 2004. Clonal distribution of natural competence in *Actinobacillus actinomycetemcomitans*. Oral Microbiol Immunol 19:340-342.

Griffith, F. 1928. The significance of pneumococcal types. J Hyg 27:113-159.

Gromkova, R.C., Mottalini, T.C., and Dove, M.G. 1998. Genetic transformation in *Haemophilus parainfluenzae* clinical isolates. Curr Microbiol 37:123-126.

Guilhabert, M.R., and Kirkpatrick, B.C. 2003. Transformation of *Xylella fastidiosa* with broad host range RSF1010 derivative plasmids. Mol Plant Pathol 4:279-285.

Harris, J.L., and Balci, Y. 2015. Population structure of the bacterial pathogen *Xylella fastidiosa* among street trees in Washington D.C. PLoS One 10:e0121297.

Hopkins, D.L., and Purcell, A.H. 2002. *Xylella fastidiosa*: Cause of Pierce's Disease of Grapevine and Other Emergent Diseases. Plant Dis 86:1056-1066.

Ihaka, R., and Gentleman, R. 1996. R: a language for data analysis and graphics. J Comput Graph Stat 5:299-314.

Kandel, P.P., Lopez, S.M., Almeida, R.P., and De La Fuente, L. 2016. Natural competence of *Xylella fastidiosa* occurs at a high frequency inside microfluidic chambers mimicking the bacterium's natural habitats. Appl Environ Microbiol 82:5269-5277.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Mentjies, P., and Drummond, A. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28:1647-1649.

Kung, S.H., and Almeida, R.P. 2011. Natural competence and recombination in the plant pathogen *Xylella fastidiosa*. Appl Environ Microbiol 77:5278-5284.

Kung, S.H., and Almeida, R.P. 2014. Biological and genetic factors regulating natural competence in a bacterial plant pathogen. Microbiology 160:37-46.

Kung, S.H., Retchless, A.C., Kwan, J.Y., and Almeida, R.P. 2013. Effects of DNA size on transformation and recombination efficiencies in *Xylella fastidiosa*. Appl Environ Microbiol 79:1712-1717.

Lorenz, M.G., and Wackernagel, W. 1994. Bacterial gene transfer by natural genetic transformation in the environment. Microbiol Rev 58:563-602.

Marcelletti, S., and Scortichini, M. 2016. Genome-wide comparison and taxonomic relatedness of multiple *Xylella fastidiosa* strains reveal the occurrence of three subspecies and a new *Xylella* species. Arch Microbiol 198:803-812.

Martelli, G.P., Boscia, D., Porcelli, F., and Saponari, M. 2016. The olive quick decline syndrome in south-east Italy: a threatening phytosanitary emergency. Eur J Plant Pathol 144:235-243.

Matsumoto, A., and Igo, M.M. 2010. Species-specific type II restriction-modification system of *Xylella fastidiosa* Temecula1. Appl Environ Microbiol 76:4092-4095.

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

Maughan, H., and Redfield, R.J. 2009. Extensive variation in natural competence in *Haemophilus influenzae*. Evolution 63:1852-1866.

Meng, Y., Li, Y., Galvani, C.D., Hao, G., Turner, J.N., Burr, T.J., and Hoch, H.C. 2005. Upstream migration of *Xylella fastidiosa* via pilus-driven twitching motility. J Bacteriol 187:5560-5567.

Newman, K.L., Almeida, R.P.P., Purcell, A.H., and Lindow, S.E. 2003. Use of a green fluorescent strain for analysis of *Xylella fastidiosa* colonization of *Vitis vinifera*. Appl Environ Microbiol 69:7319-7327.

Newman, K.L., Almeida, R.P.P., Purcell, A.H., and Lindow, S.E. 2004. Cell-cell signaling controls *Xylella fastidiosa* interactions with both insects and plants. Proc Natl Acad Sci USA 101:1737-1742.

Niza, B., Merfa, M.V., Alencar, V.C., Menegidio, F.B., Nunes, L.R., Machado, M.A., Takita, M.A., and de Souza, A.A. 2016. Draft genome sequence of 11399, a transformable citrus-pathogenic strain of *Xylella fastidiosa*. Genome Announc 4:e01124-01116.

Nunes, L.R., Rosato, Y.B., Muto, N.H., Yanai, G.M., da Silva, V.S., Leite, D.B., Gonçalves, E.R., de Souza, A.A., Coletta-Filho Hé, D., Machado, M.A., Lopes, S.A., and de Oliveira, R.C. 2003. Microarray analyses of *Xylella fastidiosa* provide evidence of coordinated transcription control of laterally transferred elements. Genome Res 13:570-578.

Nunney, L., Yuan, X., Bromley, R.E., and Stouthamer, R. 2012. Detecting genetic introgression: high levels of intersubspecific recombination found in *Xylella fastidiosa* in Brazil. Appl Environ Microbiol 78:4702-4714.

Nunney, L., Ortiz, B., Russell, S.A., Ruiz Sanchez, R., and Stouthamer, R. 2014a. The complex biogeography of the plant pathogen *Xylella fastidiosa*: genetic evidence of introductions and subspecific introgression in Central America. PLoS One 9:e112463.

Nunney, L., Hopkins, D.L., Morano, L.D., Russell, S.E., and Stouthamer, R. 2014b. Intersubspecific recombination in *Xylella fastidiosa* strains native to the United States: infection of novel hosts associated with an unsuccessful invasion. Appl Environ Microbiol 80:1159-1169.

Nunney, L., Schuenzel, E.L., Scally, M., Bromley, R.E., and Stouthamer, R. 2014c. Large-scale intersubspecific recombination in the plant-pathogenic bacterium *Xylella fastidiosa* is associated with the host shift to mulberry. Appl Environ Microbiol 80:3025-3033.

Nunney, L., Vickerman, D.B., Bromley, R.E., Russell, S.A., Hartman, J.R., Morano, L.D., and Stouthamer, R. 2013. Recent evolutionary radiation and host plant specialization in the *Xylella fastidiosa* subspecies native to the United States. Appl Environ Microbiol 79:2189-2200.

Nunney, L., Yuan, X., Bromley, R., Hartung, J., Montero-Astua, M., Moreira, L., Ortiz, B., and Stouthamer, R. 2010. Population genomic analysis of a bacterial plant pathogen: novel insight into the origin of Pierce's disease of grapevine in the U.S. PLoS One 5:e15488.

Oliver, J.E., Cobine, P.A., and De La Fuente, L. 2015. *Xylella fastidiosa* isolates from both subsp. *multiplex* and *fastidiosa* cause disease on southern highbush blueberry (*Vaccinium sp.*) under greenhouse conditions. Phytopathology 105:855-862.

Oliver, J.E., Sefick, S.A., Parker, J.K., Arnold, T., Cobine, P.A., and De la Fuente, L. 2014. Ionome changes in *Xylella fastidiosa* infected *Nicotiana tabacum* correlate with virulence and discriminate between subspecies of bacterial isolates. Mol Plant Microbe Interact 27:1048-1058.

Parker, J.K., Havird, J.C., and De La Fuente, L. 2012. Differentiation of *Xylella fastidiosa* strains via multilocus sequence analysis of environmentally mediated genes (MLSA-E). Appl Environ Microbiol 78:1385-1396.

Redak, R.A., Purcell, A.H., Lopes, J.R.S., Blua, M.J., Mizell, R.F., and Andersen, P.C. 2004. The biology of xylem fluid-feeding insect vectors of *Xylella fastidiosa* and their relation to disease epidemiology. Annu Rev Entomol 49:243-270.

Scally, M., Schuenzel, E.L., Stouthamer, R., and Nunney, L. 2005. Multilocus sequence type system for the plant pathogen *Xylella fastidiosa* and relative contributions of recombination and point mutation to clonal diversity. Appl Environ Microbiol 71:8491-8499.

Seitz, P., and Blokesch, M. 2013a. Cues and regulatory pathways involved in natural competence and transformation in pathogenic and environmental Gram-negative bacteria. FEMS Microbiol Rev 37:336-363.

Seitz, P., and Blokesch, M. 2013b. DNA-uptake machinery of naturally competent *Vibrio cholerae*. Proc Natl Acad Sci USA 110:17987-17992.

Sikorski, J., Teschner, N., and Wackernagel, W. 2002. Highly different levels of natural transformation are associated with genomic subgroups within a local population of *Pseudomonas stutzeri* from soil. Appl Environ Microbiol 68:865-873.

Su, C.C., Deng, W.L., Jan, F.J., Chang, C.J., Huang, H., Shih, H.T., and Chen, J. 2016. *Xylella taiwanensis* sp. nov., causing pear leaf scorch disease. Int J Syst Evol Microbiol 66:4766-4771.

Su, C.C., Chang, C.J., Chang, C.M., Shih, H.T., Tzeng, K.C., Jan, F.J., Kao, C.W., and Deng, W.L. 2013. Pierce's disease of grapevines in Taiwan: isolation, cultivation and pathogenicity of *Xylella fastidiosa*. J Phytopathol 161:389-396.

Van Sluys, M.A., de Oliveira, M.C., Monteiro-Vitorello, C.B., Miyaki, C.Y., Furlan, L.R., Camargo, L.E., da Silva, A.C., Moon, D.H., Takita, M.A., Lemos, E.G., Machado, M.A., Ferro, M.I., da Silva, F.R., Goldman, M.H., Goldman, G.H., Lemos, M.V., El-Dorry, H., Tsai, S.M., Carrer, H., Carraro, D.M., de Oliveira, R.C., Nunes, L.R., Siqueira, W.J., Coutinho, L.L., Kimura, E.T., Ferro, E.S., Harakava, R., Kuramae, E.E., Marino, C.L., Giglioti, E., Abreu, I.L., Alves, L.M., do Amaral, A.M., Baia, G.S., Blanco, S.R., Brito, M.S., Cannavan, F.S., Celestino, A.V., da Cunha, A.F., Fenille, R.C., Ferro, J.A., Formighieri, E.F., Kishi, L.T., Leoni, S.G., Oliveira, A.R., Rosa, V.E., Jr., Sassaki, F.T., Sena, J.A., de Souza, A.A., Truffi, D., Tsukumo, F., Yanai, G.M., Zaros, L.G., Civerolo, E.L., Simpson, A.J., Almeida, N.F., Jr., Setubal, J.C., and Kitajima, J.P. 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.

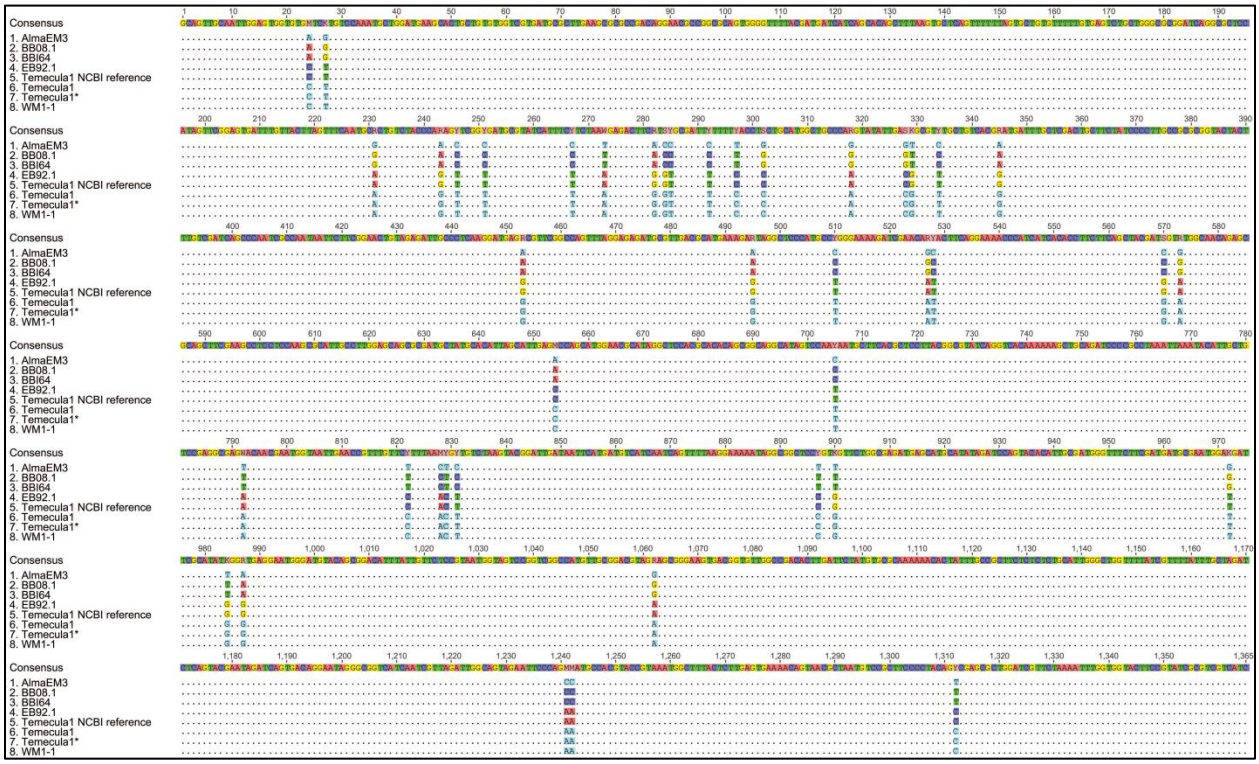
Varani, A.M., Souza, R.C., Nakaya, H.I., de Lima, W.C., Paula de Almeida, L.G., Kitajima, E.W., Chen, J., Civerolo, E., Vasconcelos, A.T., and Van Sluys, M.A. 2008. Origins of

the *Xylella fastidiosa* prophage-like regions and their impact in genome differentiation. PLoS One 3:e4059.

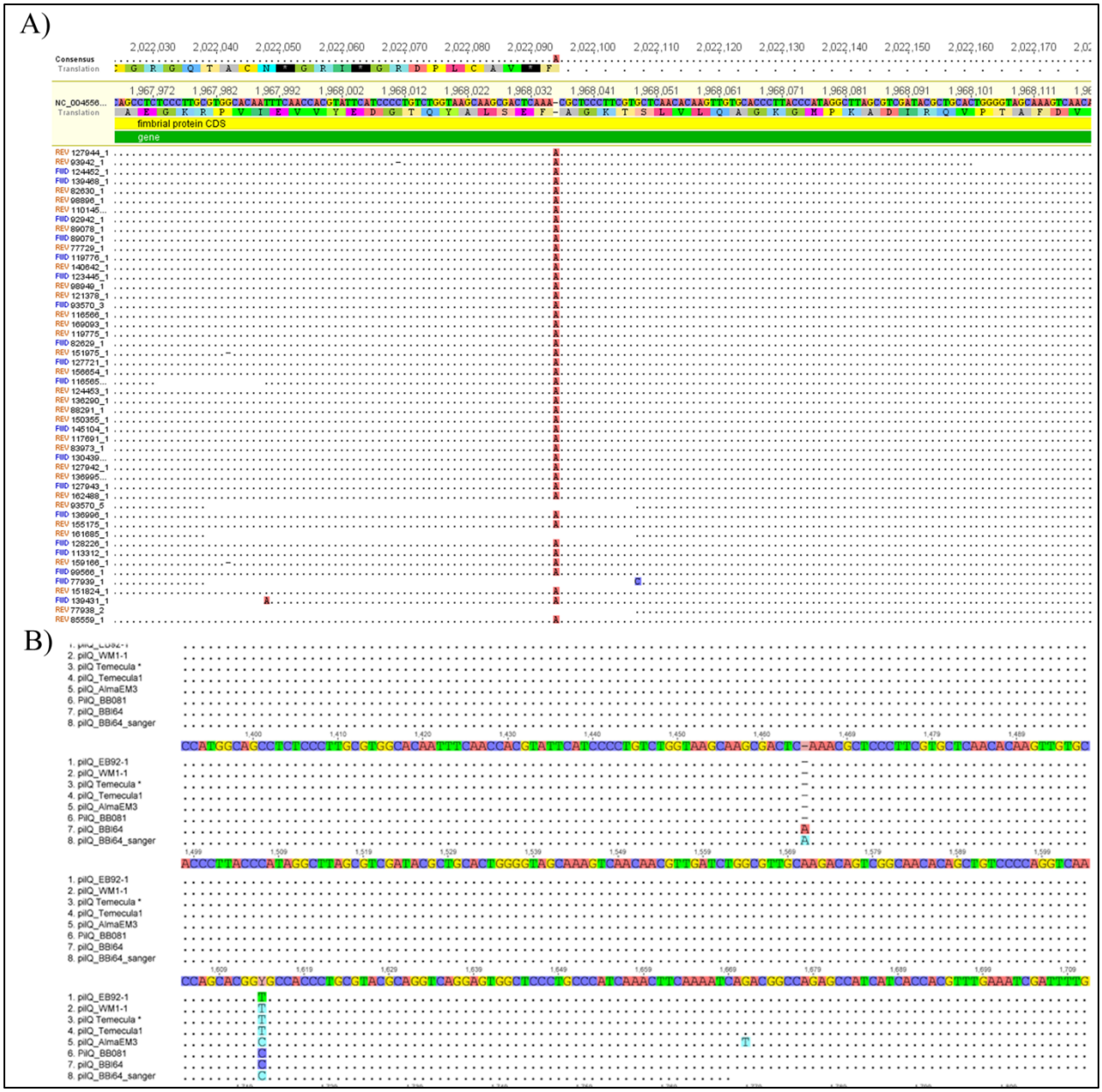
Yuan, X., Morano, L., Bromley, R., Spring-Pearson, S., Stouthamer, R., and Nunney, L. 2010. Multilocus sequence typing of *Xylella fastidiosa* causing Pierce's disease and oleander leaf scorch in the United States. Phytopathology 100:601-611.

Zhang, S., Flores-Cruz, Z., Kumar, D., Chakrabarty, P., Hopkins, D.L., and Gabriel, D.W. 2011. The *Xylella fastidiosa* biocontrol strain EB92-1 genome is very similar and syntenic to Pierce's disease strains. J Bacteriol 193:5576-5577.

Supplemental material for chapter 3



Supplemental fig. 3-1. Alignment of the flanking region DNA sequences of pKLN61 insertion site of *X. fastidiosa* strains. Homologous regions in the genome of strains relative to the donor plasmids were compared. The regions were identical within the subspecies but different between them.



Supplemental fig. 3-2. Read mapping of BBI64 with Temecula1 reference genome showing insertion of nucleotide 'A' in the coding region of *pilQ* gene **(A)**. Multiple sequence alignment of *pilQ* nucleotide sequence of various *X. fastidiosa* strains **(B)**. Insertion of A can be seen in BBI64 *pilQ* sequence which was further confirmed with Sanger sequencing of both forward and reverse segment.

Supplemental table 3-1: *X. fastidiosa* strains, mutants, and plasmids used in this study

Strain	Host plant	subspecies	Place of isolation	Reference/source
WM1-1	Grape	*fastidiosa*	Dahlonega/Lumpkin/GA	Parker et al., 2011
CCPM1	Grape (Petit Manseng)	*fastidiosa*	Dahlonega/Lumpkin/GA	Parker et al., 2011
Chard1	Grape	fastidiosa[a]	Dahlonega/Lumpkin/GA	This study
Temecula1	Grape	*fastidiosa*	Temecula/Riverside/CA	Van Sluyas et al., 2005
Temecula1*	Grape	*fastidiosa*	Temecula/Riverside/CA	Oliver et al., 2014
Fetzer	Grape	*fastidiosa*	Napa/CA	Parker et al., 2011
ConnCreek	Grape	*fastidiosa*	Rutherford/Napa/CA	Oliver et al., 2014
EB92-1	Elderberry	*fastidiosa*	Leesburg/Lake/FL	Hopkins, 2005
AlmaEM3	Blueberry (Emerald)	*multiplex*	Alma/Bacon/GA	Oliver et al., 2014
BB08-1	Blueberry (Windsor)	*multiplex*	Palatka/Putnam/FL	Oliver et al., 2014
BBI64	Blueberry (V1)	*multiplex*	Hoboken/Brantley/GA	Oliver et al., 2014
BH Elm	American Elm	*multiplex*	Birmingham/AL	Parker et al., 2011
Georgia Plum	Plum	*multiplex*	GA	Parker et al., 2011

Mutants	Wild type strain	Description	Reference/source
msrA::Km-WM1-1	WM1-1	*msrA* (PD0859) gene disrupted with kanamycin cassette	This study
NS1::Km-WM1-1	WM1-1	Km cassette inserted between PD0702 and PD0703, neutral site 1 (NS1)	This study

NS1::Km-EB92-1	EB92-1	Km cassette inserted in NS1	This study
NS1::Km-AlmaEM3	AlmaEM3	Km cassette inserted in NS1	This study
NS1::Cm-AlmaEM3	AlmaEM3	Chloramphenicol (Cm) cassette inserted in NS1	This study
Temecula1-GFP	Temecula1	Km cassette and GFP gene inserted in a different neutral site	Newman et al; 2003)
Plasmids	**Plasmid background**	**Description**	**Reference/source**
pAX1.Cm	pGEM-T	Plasmid containing regions of PD0702, Cm Cassette, multiple cloning site, and PD0703; inserts Cm cassette at NS1 when recombined with *X. fastidiosa*	Matsumoto et al; 2009
pAX1.Km	pGEM-T	Same as pAX1.Cm but contains Km cassette	Matsumoto et al; 2009
pKLN61	pGEM-5Zf (+)	Plasmid containing *X. fastidiosa oriC*, and TN903 Km cassette and *rpfF* gene (PD0407)	Newman et al; 2004
pMSRA-Km	pJET1.2	Plasmid Containining upstream and downstream region of *msrA* gene (PD0859) with Km cassette inserted in between	This study
pMOPB-Km	pJET1.2	Plasmid Containing upstream and downstream region of *mopB* gene (PD1709) with Km cassette inserted in between	This study

[a]Taxonomic classification not available but assumed to be a subsp. *fastidiosa* strain based on its plant host.

Supplemental table 3-2: Comparison of the flanking regions of the antibiotic insertion site between the donor plasmids and recipient *X. fastidiosa* strains.

	Donor plasmid			
Strains	**pAX1.cm[a]**	**pKLN61[a]**	**pMOPB[a]**	**pMSRA[a]**
WM1-1	99.9	100	100	100
Temecula1	99.9	100	100	95.7
Temecula1*	100	100	100	100
EB92-1	97.9	100	100	99.9
AlmaEM3	96.5	96.8	96.9	95.6
BBI64	94.5	96.8	96.9	95.6
BB08-1	96.5	96.8	96.9	95.6

[a]Numbers indicate the % identity between the homologous regions (upstream and downstream regions were concatenated for this analysis) in the donor plasmid and recipient strains

Supplemental table 3-3: Primers used in this study

Primer Name	Amplified region	Sequence (5'-3')	Description	Reference
pilQ_247_F	*pilQ* (PD1691)	CGTGGCGATAAGA AAGCATCAT	Amplify part of the *pilQ* gene	This study
pilQ_1052_R	*pilQ* (PD1691)	CTGGCCGTCTGATT TTGAAGTT	Same as above	This study
pilM_28_F	*pilM* (PD1695)	GACCAGCAGGCAT CGGTAT	Amplify part of the *pilM* gene	This study
pilM_736_R	*pilM* (PD1695)	ACAGGCCATCGTGG GAGTA	Same as above	This study
856,070_F	NS1 Up-stream region 1	TGCATAGGTAAAG ATGAGTATGCGG	Amplify region upstream of NS1 site (855158-857659 based on Temecula1 reference numbering)	This study
858,571_R	NS1 Up-stream region 1	CGGTAAAATGCAG CTTCTATCCAG	Same as above	This study
863,179_F	NS1 Down-stream region 1	TATGATTTCCACTC CTTCCAGGTG	Amplify region downstream of NS1 site (862230-864994 based on Temecula1 reference numbering)	This study
865,944_R	NS1 Down-stream region 1	AATGACTTGAATCA CCTCGCAATG	Same as above	This study
857461_F	NS1 Up-stream region 2	ATTTCAAGAGGGGT TTACCAAGGA	Amplify region upstream of NS1 site (857461-859413 based on Temecula1 reference numbering)	This study
859413_R	NS1 Up-stream region 2	GCAATGAACAAAA TGTGGTTGGTC	Same as above	This study
860517_F	NS1 Down-stream region 2	CTTGGACCTTCGAT GAGTACGTAT	Amplify region downstream of NS1 site (860517-862411) based on Temecula1 reference numbering)	This study
862411_R	NS1 Down-stream region 2	ACAATACCACCGGT CAAATTCCTA	Same as above	This study
NS1-f	NS1 flanking region	GTCAGCAGTTGCGT CAGATG	Confirm homologous recombination at NS1 site by PCR	Matsumoto et al., 2009
NS1-r	NS1 flanking region	AAAGCTGCCGACG CCAAATC	Same as above	Matsumoto et al., 2009

Supplemental table 3-4: Detection of recombination by sequencing

Position[a]	WM1-1	AlmaEM3	WM1-1 recombinant 1	WM1-1 recombinant 2	AlmaEM3 recombinant 1	AlmaEM3 recombinant 2
859009	C	A	C	A[b]	A	A
859142	G	A	G	A	A	A
859156	A	G	A	G	G	G
859216	T	C	T	C	C	C
859264	G	A	G	A	A	A
859624[c]	T	T	C	C	C	C
859913-859919	CTGTTA	CTGTTA	Kan	Kan	Kan	Kan
860858	A	G	G	A	A	A
860908	A	G	G	A	A	A
860955[d]	C	T	T	C	C	C
861033	C	T	T	C	C	C
861043	A	G	G	A	A	A
861064	T	C	C	T	T	T
861118[d]	T	G	G	T	T	T
861191	T	A	T	T	T	T
861194	C	T	C	C	C	C
861292	G	C	G	G	G	C
861330	A	G	A	A	A	G
861332	T	G	T	T	T	G
861350	C	T	C	C	C	T
861364	C	T	C	C	C	T
861382	T	C	T	T	T	C
861395	C	G	C	C	C	G
861397	G	T	G	G	G	T
861418-861498	81bp[e]	Deletion[f]	81bp	81bp	81bp	deletion
861558	deletion	TAGGTGTCGTCTTG	deletion	deletion	deletion	TAGGTGTCGTCTTG
861578	deletion	TGGGTTTGATG	deletion	deletion	deletion	TGGGTTTGATG
861660	G	A	G	G	G	A

861725	T	C	T	C	T	C
861899	G	T	G	T	G	T
862025	C	G	C	G	G	G
862037	C	A	C	A	A	A
862134	G	A	G	A	A	A
862160	G	A	G	A	A	A
862200	G	A	G	A	A	A
862202	A	G	A	G	G	G
862241	T	C	T	C	C	C
862269	C	T	C	T	T	T
862280	T	C	T	C	C	C
862364	T	G	T	G	G	G
862406	C	T	C	T	T	T
862463	C	T	C	T	T	T
Length[g]			2705bp	2910bp, 739bp	3486bp	2781bp

[a]Position is based on Temecula1 reference numbering. Only positions that were different between WM1-1 and AlmaEM3 at this region are reported. The regions between the reported positions were not different between the two strains (wild-type and respective recombinant).

[b]Yellow-highlighted bases were recombined from the donor strain to the recipient.

[c]Unresolved position, all recombinants had a different base at this position compared to the donor and recipient parents. This could be the break point of recombination when the donor strains were tagged with pAX1.Km plasmids.

[d]This region is shown in Figure 4.

[e]TCTGTACGATCATTCAGTCTATGCAATCATCCAGTGAATCGTGCATTGTGGTGCTAG GGTATTAACCGATTGCTCTGTTTG: This was the 81base pair sequence that is absent in AlmaEM3, also absent in M12 genome as evidenced by Blast (data not shown).

[f]This region is deleted in the strains relative to the other.

[g]Length of recombined region calculated as the length of flanking region that is recombined and the 1,211bp kanamycin cassette region

Chapter 4

Detection of intra-and inter-subspecific recombination in *Xylella fastidiosa* by whole genome sequence analysis

Abstract

Homologous recombination (HR) in *Xylella fastidiosa* has been described as a major cause of genetic variation and possible explanation for plant host shifts of this pathogen. Previous studies demonstrated that, although overall genetic diversity is limited among *X. fastidiosa* strains, phenotypic traits such as host range, symptom type, and disease severity differ. Multi Locus Sequence Typing (MLST), which is based on genetic analysis of specific loci or regions of the genome containing housekeeping genes, revealed genetic differences among strains at minute levels, and evidenced HR among strains. These observations, together with the results from recent experimental studies regarding natural competence and recombination in *X. fastidiosa,* support a greater role of HR in *X. fastidiosa* diversity and adaptation. Few studies focused on whole genome sequence (WGS) analysis of *X. fastidiosa* strains have generated fundamental information regarding the genetic relatedness and virulence mechanisms. With greater numbers of studies producing WGS data, these data can be useful to examine HR and its role in the evolution of *X. fastidiosa* in further details. Therefore, this study was aimed at determining the limits and abundance of HR events in the genomes of closely related *X. fastidiosa* subspecies *fastidiosa* strains. WGSs were obtained by both Illumina and PacBio sequencing of several *X. fastidiosa* wild-type (WT) strains, and recombinants experimentally produced *in vitro*. Regions of recombination in WT and recombinant subsp. *fastidiosa* strains that originate from subsp. *multiplex* strains were analyzed. Results show that intersubspecific HR regions ranging in size

from 5.4-8.7kb are frequently present and are randomly scattered across the genomes of WT subsp. *fastidiosa* strains. Moreover, a region of 10kb was detected to recombine in one event in the genome of one of the *in vitro* recombinants. Overall, this study demonstrates that genetic exchanges are extremely common among *X. fastidiosa* strains in nature, therefore intermixing of different host-plant genotypes should be avoided to prevent emergence of new and possibly more aggressive isolates.

Introduction

Horizontal gene transfer, a fundamental driving force of evolution in prokaryotes, occurs by conjugation, transduction, and transformation (Thomas and Nielsen, 2005). These processes facilitate acquisition of novel genetic elements including antibiotic resistance and virulence factors, resulting in the emergence of new pathotypes (Hacker and Carniel, 2001). Acquisition of new genes promotes expansion of the gene pool within a species. These genes can be transferred to closely related phyla of the same species by homologous recombination (HR). Bacteria can also exchange regions of their genome, which is loosely referred to as new allele acquisition (Didelot and Maiden, 2010) and is almost exclusively a result of RecA protein-mediated HR that is responsible for creating genetic variation in almost every life form (Dorer et al., 2011). In prokaryotes, repair of damaged DNA is proposed to be a key role of HR (Michod et al., 2008). DNA damage repair by HR not only promotes survival under stressful conditions, but may also produce new variants of the gene given that the homologous DNA originates from a closely related but genetically distinct source (Guttman, 1997). These new genetic variants can be important in promoting adaptive responses. For example, creating variability of surface proteins that include pathogen-associated molecular patterns (PAMPs) may support evasion of host defenses (Nelson and Selander, 1994; Feil and Spratt, 2001), transfer of virulence proteins can

cause disease emergence (Yan et al., 2008), and adaptation to new environments (Baltrus et al., 2008).

The relative role of HR and point mutation in generating genetic variation varies among bacterial species (Feil et al., 2001). Bacteria that have a high capacity for recombination may undergo rapid evolution, and contribute to disease emergence, as is often the case for emergence of pathogenic bacteria. For example, the emergence of *Escherichia coli* serotype O104:H4 that caused an outbreak of food-borne gastroenteritis in Germany in 2011, was suggested by whole genome sequencing of these strains to be a result of genetic exchange between a Shiga-toxin producing *E. coli* and an enteroaggregative *E. coli* (Rasko et al., 2011). Determining the genetic capabilities of mutations and recombination is crucial to understand the evolution of pathogen and its disease epidemics. These capabilities can be either experimentally determined or inferred from genetic data (Fraser et al., 2007; Kandel et al., 2017). For detection of HR from sequence data, the recombining DNA sequence must be different from the native fragment. Thus, all recombination events occurring in a population under a natural condition are impossible to be detected. This suggests that the rate of recombination in a species could be much higher than realized. In fact, the first observation of HR was achieved by detection of mosaic genes that contained fragments originating from different phyla of the same genus (Spratt et al., 1992). Therefore, co-existence of genetically diverse populations is crucial for detection of HR in a natural condition.

Use of Multi Locus Sequence Typing/Analysis (MLST/MLSA) in bacterial phylogenetics, a technique to study genetic diversity within a species based on sequence analysis of 5-11 house-keeping genes scattered across the genome (Maiden, 2006), has prompted or facilitated detection of HR in bacteria. Until recently, MLST had become a signature tool for bacterial phylogenetic

studies and for the detection of HR (Freel et al., 2013; Kong et al., 2013; Timilsina et al., 2015). However, use of few housekeeping genes may indicate a bias as HR that is present in other genomic regions cannot be represented by the MLST approach. As a result, studies based on whole genome sequence comparison have emerged in the recent years, and new tools for detecting HR in the bacterial genomes have been developed (Marttinen et al., 2012; Yahara et al., 2014; Croucher et al., 2015).

Xylella fastidiosa is a plant pathogenic bacterium that has a very broad plant host range (EFSA, 2016), and colonizes two distinct habitats: plant xylem vessels and foregut of insect vectors (sharpshooter leafhoppers and spittlebugs) (Hopkins and Purcell, 2002). Previously, *X. fastidiosa* was believed to be limited to the Americas, but it has now disseminated to other regions of the world including Europe (Saponari et al., 2013; Strona et al., 2017), and Asia (Su et al., 2016). It poses a serious threat in the production of wine grapes in the U.S., citrus and coffee in Brazil, and olives in Italy, and at present has become a worldwide threat due to its recent emergence around the world (Almeida and Nunney, 2015). Taxonomically, strains within *X. fastidiosa* are categorized into three subspecies (Marcelletti and Scortichini, 2016a), although additional subspecies have been suggested (Scally et al., 2005; Nunney et al., 2014b). Although, not clearly defined, subspecies are somewhat host-plant specific. For example, subsp. *fastidiosa* strains infect grapevines and some other plants but have not been reported to infect citrus plants, whereas subsp. *pauca* strains infect citrus, coffee, and olives but have not been reported to infect grapes (Nunney et al., 2012). Two of the subspecies *fastidiosa* and *multiplex* have existed in the U.S. for more than a century (Nunney et al., 2010).

Extensive HR has been described in *X. fastidiosa* by genetic diversity and phylogenetic studies based on MLST (Scally et al., 2005; Almeida et al., 2008; Nunney et al., 2010; Nunney et al., 2012; Coletta-Filho et al., 2016). These studies made very important predictions of disease emergence and epidemics in *X. fastidiosa* impacted by HR. For example, HR between *X. fastidiosa* isolates belonging to different subspecies has been proposed to be responsible for plant host shift and new disease emergences in isolates infecting citrus, mulberry, and blueberries (Nunney et al., 2010; Nunney et al., 2014a; Nunney et al., 2014b). Experimental studies have demonstrated natural competence and recombination in *X. fastidiosa* both *in vitro* (Kung and Almeida, 2011), and in habitats mimicking the natural growth environment of the bacterium (Kandel et al., 2016). Moreover, intersubspecific HR (IHR) between subsp. *fastidiosa* and subsp. *multiplex* has been recently demonstrated *in vitro* (Kandel et al., 2017). A previous study used two complete genomes of *X. fastidiosa* subsp. *fastidiosa,* Temecula1 and M23, and showed the presence of variable regions that were related to a subsp. *multiplex* isolate, M12 (Nunney et al., 2010). This approach could be more useful in further understanding the role of HR in plant host shift, virulence, and disease emergence than MLST as the former approach provides information on the whole genome. Therefore, the objective of this study was to utilize whole genome sequence data to determine IHR in *X. fastidiosa* wild-type strains isolated from field infections and in recombinant strains experimentally produced *in vitro* (Kandel et al., 2017). Results show that whole genome sequencing approach allows detection of IHR that was not detected by MLST. Genomic regions of up to 10kb were found to have recombined from a donor strain of subsp. *multiplex* to the recipient strain of subsp. *fastidiosa* in one of the recombinant produced *in vitro*. Moreover, regions of IHR derived from subsp. *multiplex* were frequently detected in multiple wild-type strains of subsp. *fastidiosa* and were randomly distributed across the genome.

A highly virulent subsp. *fastidiosa* strain WM1-1 showed two IHR regions that were more identical to subspecies *multiplex* strains than the *fastidiosa* strains. These data suggest that IHR is widespread in *X. fastidiosa* and could have broader implications in its adaptation and virulence.

Materials and methods

Bacterial strains, media, and culture conditions

X. fastidiosa subsp. *fastidiosa* wild-type strains WM1-1, Temecula1, Temecula1*, and CCPM1; and subsp. *multiplex* strain AlmaEM3 were used for this study. A variant strain that was isolated from infected grapevines in California and was initially presumed to be Temecula1 but differed based on whole genome sequence analysis in this study was also used and is referred to hereafter as TemeculaL. Mutant strains NS1-CmR, in which a chloramphenicol resistant gene is inserted in the neutral site 1 in the Temecula1 chromosome (Matsumoto et al., 2009); pglA-KmR, in which a kanamycin resistant gene is inserted to disrupt the gene encoding polygalacturonase in the Fetzer strain (Roper et al., 2007); and KLN59.3, in which green fluorescent protein (GFP) and kanamycin resistant gene are inserted in a neutral region of Temecula1 (Newman et al., 2003) were used. In addition, mutant strain NS1::Km-AlmaEM3 was generated by transforming the wild type AlmaEM3 with pAX1.Km plasmids (Matsumoto et al., 2009). The recombinant strains AlmaTemeculaL recombinants 1 and 2 were generated by mixing heat killed NS1::Km-AlmaEM3 with wild type TemeculaL. The NS1-pglA recombinant was generated by co-culturing strains NS1-CmR and pglA-KmR in PD3. The gfpWM1-1 recombinants 1 and 2 were generated by mixing heat-killed KLN59.3 with WM1-1 cells in PD3 agar medium as previously described (Kandel et al., 2017). Further information about strains, mutants, and recombinants used in this study is presented in Table 4-1.

Table 4-1: Strains, mutants, and recombinants used in this study

Strain	Host plant	subspecies	Place of isolation	Reference/source
WM1-1	Grape	*fastidiosa*	Dahlonega/GA	(Parker et al., 2012)
CCPM1	Grape	*fastidiosa*	Dahlonega/GA	(Parker et al., 2012)
Temecula1[a]	Grape	*fastidiosa*	Riverside/CA	(Van Sluys et al., 2003)
Temecula1*	Grape	*fastidiosa*	Riverside/CA	(Oliver et al., 2014)
TemeculaL[b]	Grape	*fastidiosa*	CA	This study
pglA (Fetzer)	Grape	*fastidiosa*	Napa/CA	(Roper et al., 2007)
M23	Grape	*fastidiosa*	Kern/CA	(Chen et al., 2010)
NS1-CmR	Mutant of Temecula1	*fastidiosa*	NA	(Matsumoto et al., 2009)
KLN58.3 (GFP)	Mutant of Temecula1	*fastidiosa*	NA	(Newman et al., 2003)
NS1pglA recombinant[c]	Recombinant of NS1-CmR and pglA mutant	*fastidiosa*	NA	(Kandel et al., 2016)
AlmaTemL recombinants 1 and 2[d]	Recombinants of TemeculaL with AlmaEM3 donor	*fastidiosa*	NA	This study
gfpWM1-1 recombinants 1 and 2[d]	Recombinants of WM1-1 with KLN59.3 donor	*fastidiosa*	NA	(Kandel et al., 2017)
AlmaEM3	Blueberry (Emerald)	*multiplex*	Alma/GA	(Oliver et al., 2014)
M12[a]	Almond	*multiplex*	Kern/CA	(Chen et al., 2010)
9A5C[a]	Sweet orange	*pauca*	São Paulo/ Brazil	(Simpson et al., 2000)

[a]Only genomes were used from NCBI database for these strains.
[b]Variant of Temecula1 that was confirmed by whole genome sequencing in this study.
[c]Recombinant was produced by co-culturing live cells of donor and recipient.
[d]Recombinant was produced by co-culturing heat-killed donor and cells and live recipient cells.

All strains were cultured in Periwinkle Wilt (PW) medium (Davis et al., 1981) agar plates, modified by replacing phenol red with 1.8 g litre^{-1} Bovine Serum Albumin (BSA) (Gibco Life Sciences Technology), for one week at 28°C. Each strain was initially taken from -80°C glycerol stock, re-streaked onto new PW plates, and cultured for another week before use. Kanamycin, chloramphenicol, and ampicillin were used at a concentration of 30, and 10 µg ml^{-1}, respectively.

DNA extraction, library preparation, and sequencing

X. fastidiosa strains WM1-1, TemeculaL, Temecula1*, AlmaEM3; mutants NS1-CmR, pglA-KmR, KLN59.3; and *in vitro* recombinant strains AlmaTemeculaL recombinants 1 and 2, NS1-pglA recombinant, and gfpWM1-1 recombinants 1 and 2 were sequenced using Illumina Miseq system at Auburn University. These strains were selected for whole genome sequencing based on their genetic relatedness, geographic location, and severity of virulence as reported in previous studies (Parker et al., 2012; Oliver et al., 2014). Strains WM1-1, CCPM1, AlmaEM3, BBI64, and BB08-1 genomes were sequenced using PacBio sequencing and assembled by Celera assembler by our collaborator's lab and regions of interest have been used in this study (Rodrigo Almeida, University of California Berkeley). Partial sequences of strains EB92-1 (Zhang et al., 2011), GB514 (Schreiber et al., 2010), Ann-1, MUL0034, U42D, Griffin-1, ATCC35871 (Marcelletti and Scortichini, 2016a) were retrieved from NCBI gene bank database whenever required.

For Illumina Miseq sequencing, cells cultured on PW agar plates were suspended in 400 µl of sterile MilliQ water and DNA was extracted using a modified CTAB protocol (Doyle, 1987). Extracted DNA was run through gel electrophoresis (1% agarose gel in TAE buffer) to detect RNA contamination followed by treatment with RNase A (ThermoFisher Scientific). Both DNA

and RNA concentrations were verified with a Qubit® 2.0 Fluorimeter (Life Technologies) using the Qubit dsDNA and Qubit RNA HS Assay Kit (Thermo Fisher Scientific), respectively. DNA samples of 20 ng were then adjusted to 50 µl (2.5 ng/µl) volume by diluting in sterile MilliQ water.

DNA libraries were prepared using Nextera® DNA Library Preparation Kit (Illumina) according to the manufacturer's protocol. Twelve separate genomic libraries were pooled together using different indices pairs used for each library. Clonal amplification and sequencing was performed on a MiSeq flow cell using MiSeq reagent kit V2. Paired-end sequencing was performed with 250 cycles for each end.

Quality filtering and trimming

Two FastQ files per genome (one each for forward and reverse reads) were generated as the final output files of the sequencing run. The FastQ files contain the actual nucleotide bases and the quality score for each base call of the fragments/reads. The FastQ files were run through FastQC (Babraham Bioinformatics), a quality control tool for high throughput sequence data, to assess the quality of sequencing runs.

Trimming was performed using Trimmomatic-0.35, a flexible trimmer for Illumina sequence data (Bolger et al., 2014). The command for Trimmomatic was designed such that it allowed removal of adapter contaminations, bases with Phred quality scores lower than 20 (1% error probability) from the beginning and end of the reads, and a sliding window trimming of 4:20. Reads with length shorter than 36bp after trimming were removed.

De Novo assembly

Trimming was followed by De Novo assembly of the trimmed reads using SPAdes 3.9.0 (Bankevich et al., 2012), a de Bruijn graph based assembly program. Assembly was performed using a command with careful mode switched on that allows mismatch correction. K-mer sizes of 21, 33, 55, 77, 99, and 127 were used. For two genomes WM1-1 and AlmaEM3 that had sequences available from both Miseq and PacBio sequencing, a hybrid assembly algorithm of SPAdes 3.9.0 was used. Resulting scaffolds or contigs of the SPAdes assembly were used for further analysis. Contigs of length shorter than 200bp were removed. Quality of assembly was assessed with QUAST, a quality assessment tool for genome assemblies (Gurevich et al., 2013), and the assembly statistics were recorded.

Reference mapping of raw reads and assembled contigs

Reference mapping was performed using Geneious map to reference algorithm (Kearse et al., 2012). The raw reads were quality trimmed using the Geneious trimming option such that reads containing bases that had error probabilities of >5% were removed. The reads were paired using 'set paired read' option and mapped to Temecula1 reference (Van Sluys et al., 2003) using default settings for medium sensitivity. The assembled contigs were aligned to Temecula1 genome using pairwise genome alignment program LASTZ plugin within Geneious (Harris, 2007).

Annotation

Annotation of assembled contigs was performed by a command line tool Prokka (Seemann, 2014). Selected genomes were also annotated using web-based RAST server (Aziz et al., 2008).

Detection of variable and recombinant regions in the genomes

A consensus contiguous-genome sequence was generated from the reference mapping for the genomes sequenced in this study. Regions with coverage less than 5x were designated as gaps and ambiguous bases were fixed manually with the help of reference assembly. Consensus sequences were used in a multiple sequence alignment using MAAFT 7.31 software (Katoh and Standley, 2013). The aligned genome sequences were analyzed by Gubbins 2.2.0-7 (Croucher et al., 2015) with minimum SNPs for recombination detection set to 100. After confirming the presence of regions of recombination in the genome, the reference assembly was manually visualized to detect presence of variable regions in the Geneious genome browser. Variable regions in AlmaTemeculaL recombinants 1 and 2 and subsp. *fastidiosa* strains WM1-1, CCPM1, TemeculaL, and pglA (Fetzer) were located based on the presence of elevated levels of locally-clustered SNPs. Variable regions detected in this way were used as query in a basic local alignment search tool (BLAST) against the NCBI gene bank database. Regions of the strains that yielded better matches to the sequences of different subspecies than their own (i.e. subsp. *fastidiosa)* were selected for further analysis.

Phylogenetic analysis of variable regions

Orthologous regions, in other strains, of the variable region detected above were identified by BLAST search in the gene bank database and a local BLAST database created by using the contigs generated by De Novo assembly. Sequences were aligned using Muscle Multiple Sequence Alignment program (Edgar, 2004). Phylogenetic trees were generated by a maximum likelihood (ML) analysis using RAxML version 7.2.8 with rapid bootstrapping and search for best-scoring ML tree algorithm using 1,000 replications. Sequence of the pear leaf scorch strain PLS229 (Su et al., 2016) was used as out-group. Variable regions that grouped with a different subspecies were inferred as regions of intersubspecific recombination. For comparison of the phylogeny, MLST sequences of seven house-keeping genes as described previously (Yuan et al., 2010) were retrieved for all strains used above (strain ATCC35871 was not used as two MLST genes *leuA*, and *petC* were missing in its draft genome, *petC* was missing in PLS229). The MLST sequences were concatenated, aligned, and phylogenetic tree was generated by RAxML using same parameters as with the variable regions. A phylogenetic tree of whole genomes was also generated by RAxML using 1,000 bootstrap replications and rapid bootstrap algorithm from the alignment file generated from MAFFT genome alignment with the subsp. *pauca* strain 9A5C used as outgroup. Nodes having less than 50% consensus support were collapsed to generate the consensus tree.

Results

Coverage, assembly statistics, and annotation of sequenced genomes

The Miseq sequencing platform generated 0.32-3.42 million reads per genome (Supplementary table 4-1) of which approximately 85% survived trimming. Average read length was 113.8-158.8bp. Average coverage ranged from of 20x-184x. The number of contigs generated by De Novo assembly ranged from 105-254 with a high N50 value ranging from 93,631bp -1,48,434bp except for AlmaTemL_rec2 for which the assembly contained 420 contigs with a N50 of 27,752bp. The total assembly size ranged from 2.45Mbp-2.74Mbp, which is within the range of *X. fastidiosa* genome size. A hybrid assembly approach using reads from Miseq and PacBio sequencing significantly improved the assembly as indicated by the reduction in the number of contigs and increase in N50 values for the genome assembled with this approach (Supplementary table 4-1). Quast results using Temecula1 as reference genome showed that the assemblies cover more than 95% of the genome (data not shown). Using repeat finder plugin from Geneious, it was detected that Temecula1 genome contains at least 235 repeat regions ranging from >100bp-7,465bp (data not shown) suggesting that a De Novo approach is highly unlikely to produce a complete single genome for *X. fastidiosa* even when both long and short reads are merged in assembling. Reference mapping with the Temecula1 showed that the sequenced genomes had >99% identity except for a subsp. *multiplex* strains AlmaEM3 that had 96.3% identity with the Temecula1 reference.

Annotation of strain WM1-1 genome with RAST server showed that the genome contains 2727 coding sequences with 55 RNA genes and a GC content of 51.6%. A total of 1082 (39.7%) of the coding regions had unknown functions and were assigned as hypothetical proteins. The genome

also contained 66 phage and phage-related proteins and 16 mobile element proteins. Number of genes involved in different metabolic pathways is shown (Fig. 4-1). AlmaEM3 annotation showed 2731 coding sequences with 56 RNA genes and a GC content of 51.7%. AlmaEM3 contained 1091, 72, and 12 hypothetical, phage, and mobile element proteins, respectively. Annotation with Prokka generated 2355-2616 gene features including 50-53 RNA genes (data not shown).

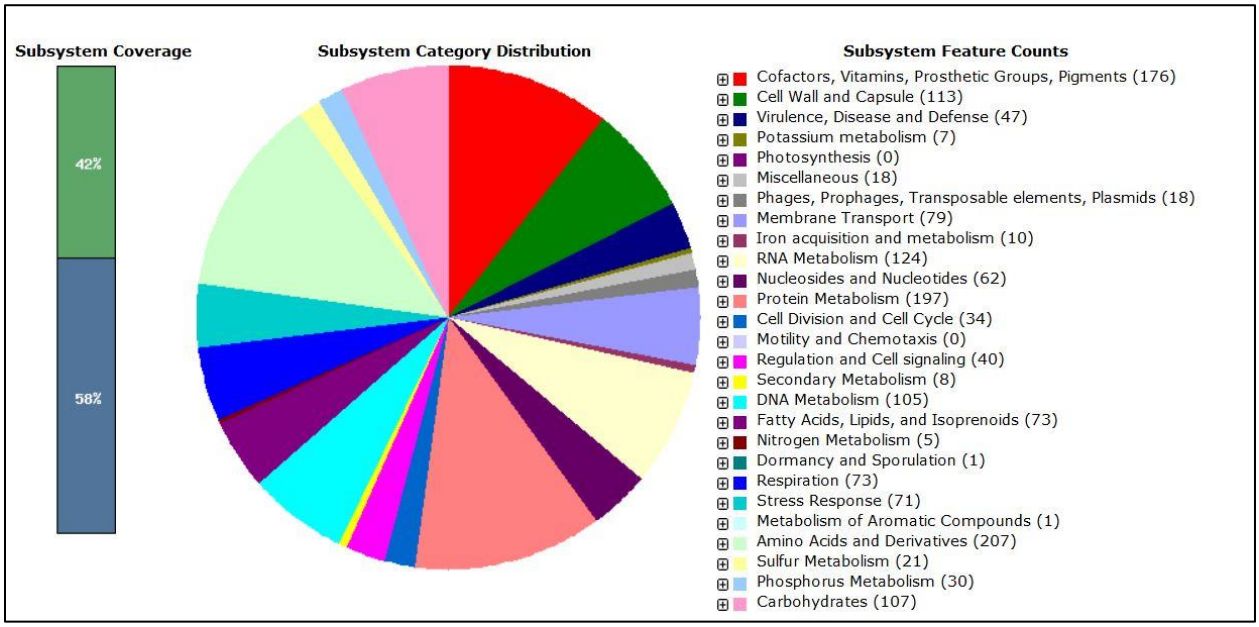


Fig. 4-1. Annotation of strain WM1-1 with RAST server showing number of genes involved in various functions.

Detection of homologous recombination in the genomes of *X. fastidiosa* strains by Gubbins

Whole genomes of *X. fastidiosa* obtained in this study and other reference genomes obtained from GenBank database (Table 4-1) were aligned with MAFFT software and the alignment file was used as input to detect regions of recombination in the genomes using Gubbins. Multiple recombination regions were detected in the genomes (Fig. 4-2). Some of these recombination

regions were shared by strains that belonged to multiple clades in the phylogenetic tree, where as other regions were unique to a specific clade. For example, the phylogenetic clade containing WM1-1 and its two recombinants (gfpWM1-1 rec1 and 2) contained five recombination regions that were unique to this clade and nine regions that were shared with strains belonging to other clades (Fig. 4-2).

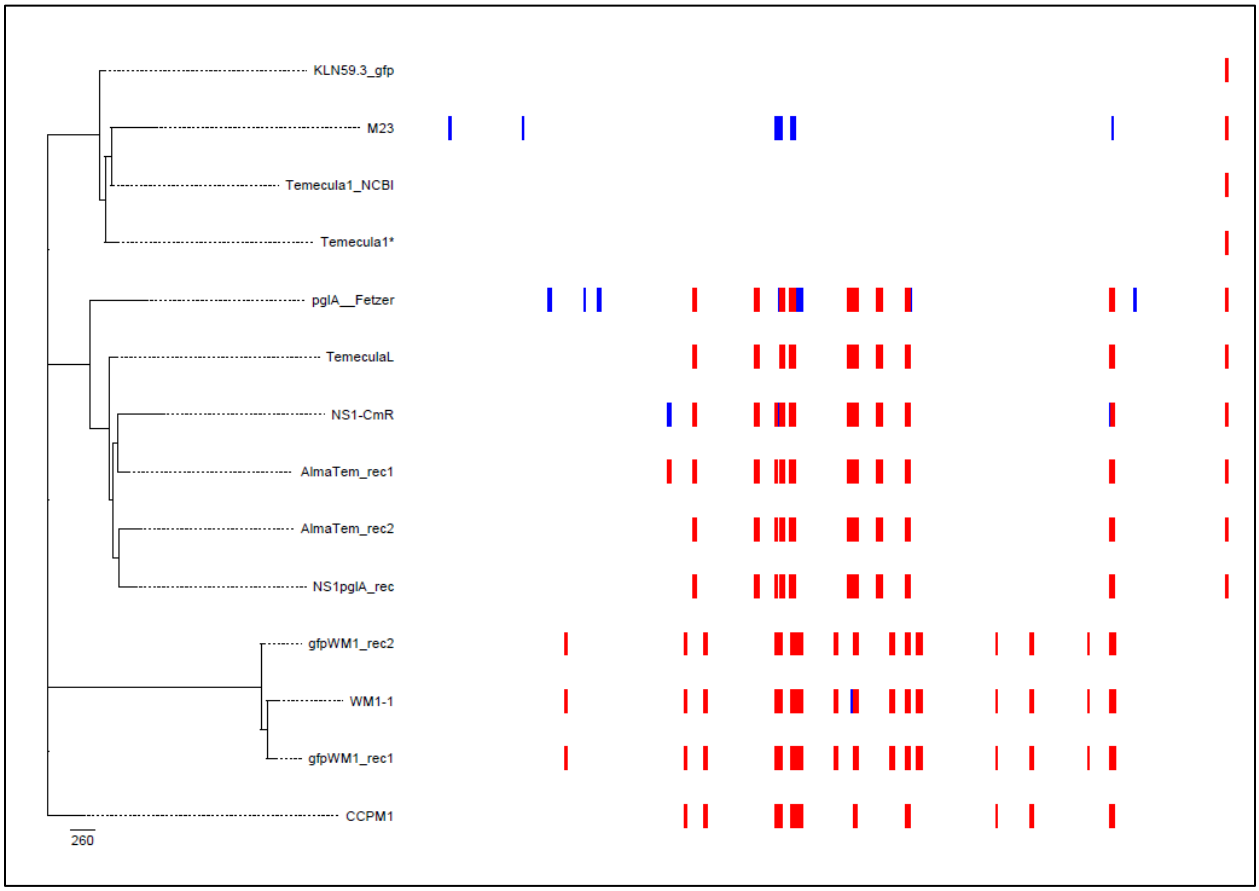


Fig. 4-2. Detection of recombination in the genomes of *Xylella fastidiosa* subsp. *fastidiosa* wild-type and *in vitro* recombinants using Gubbins. The panel on the left shows the resulting phylogeny based on the genome sequence. For each strain, colored blocks represent the regions of recombination. Red blocks indicate predicted recombination occurring on an internal branch, which are therefore shared by multiple isolates through common descent. Blue blocks represent recombination that occur on terminal branches, which are unique to individual isolates. Gubbin was run with default setting with minimum SNPs for recombination set to 100.

Detection of recombination in the genome of experimentally obtained recombinants

Miseq reads of donor, recipient, and recombinant strains were mapped with the reference Temecula1. Based on the reference mapping of two of the intersubspecific recombinants generated *in vitro,* recombination could be detected only in the flanking region of the antibiotic insertion site. In one of the recombinants (AlmaTemL_rec1), a ~10kb region of the genome in the flanking region was detected to have recombined (Fig. 4-3A). However, for the other recombinant (AlmaTemL_rec2), only regions of antibiotic resistance and short homologous flanking regions (regions used to generate the pAX1.Km plasmid) were detected to have recombined. On phylogenetic analysis based on this 10kb region, the strain (AlmaTemL_rec1) that showed recombination grouped with the donor strain AlmaEM3, while the one that lacked recombination (AlmaTemL_rec2) besides the antibiotic region, grouped with the recipient strain, TemeculaL (Fig. 4-3B). Intra-subspecific recombination could not be detected in the recombinants (NS1-pglA recombinant and gfpWM1-1 recombinants 1 and 2) as the flanking regions of antibiotic insertion did not have any sequence difference between the donor and recipient to allow detection of recombination. The two gfp recombinants recombined both the kanamycin region and the gfp region as indicated by phenotypic observation of kanamycin resistance and GFP-fluorescence as well as by the presence of these regions in the genomes of these recombinants (data not shown). Recombination could not be detected between the donor and recipient at other genomic regions. Recombinant regions detected by reference mapping were confirmed in the De Novo assembled contigs by local BLAST of the recombinant region as query against the contigs as database.

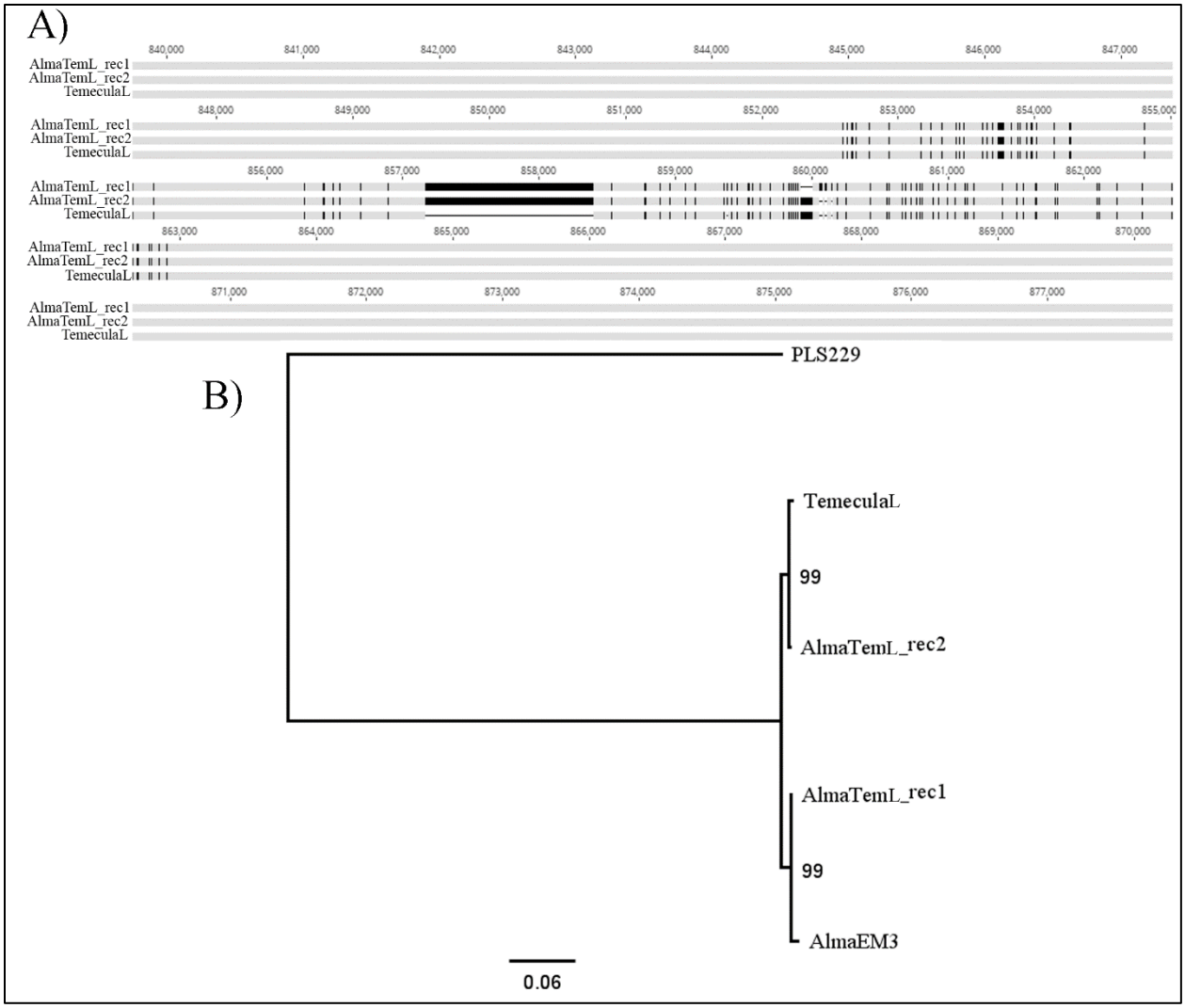


Fig. 4-3. Detection of intersubspecific recombination in the recombinants produced *in vitro*. Kanamycin marker-tagged, heat-killed, genomic DNA of *Xylella fastidiosa* subspecies *multiplex* strain AlmaEM3 was used as donor with the live recipient cells of subsp. *fastidiosa* strain TemeculaL. Recombinants were selected by acquisition of antibiotic resistance and whole genomes of two recombinants and the donor and recipient parents were sequenced. **A)** A region of ~10kb indicated by dark vertical strips is the recombined region in recombinant 1 that grouped with the donor strain at this region in the phylogenetic tree constructed by RAxML with 1000 bootstrap replications **(B);** Scale bar indicates number of substitutions per site and numbers at the tree nodes indicate % bootstrap support.

Detection of intersubspecific recombination in wild-type subsp. *fastidiosa* strains from subsp. *multiplex* donors

By using a similar approach as above, variable regions with respect to the Temecula1 reference genome as predicted from the Gubbins analysis, were detected in the genomes of subsp. *fastidiosa* strains WM1-1, TemeculaL, Fetzer, and CCPM1. Variable regions ranged in size from below 4kb up to 9kb (Table 4-2). These regions contained a greater number of SNPs with subsp. *fastidiosa* strains compared to that with subsp. *multiplex* strains (Table 4-2). To ensure that these variable regions are not due the bias of the reference mapping, De Novo assembled contigs were aligned to the Temecula1 genome, and the presence of these regions was confirmed in the contigs; an example with WM1-1 variable region 1 is presented in Fig. 4-4. BLAST search with the NCBI database showed that some variable regions identified in the genomes of subsp. *fastidiosa* strains were more identical to subsp. *multiplex* strains than with subsp. *fastidiosa* strains, and grouped perfectly with them on phylogenetic analysis based on these variable regions (Fig. 4-5A). This confirmed that the variable regions present in the genomes of subsp. *fastidiosa* strains were derived from subsp. *multiplex* strains by IHR. Phylogenetic analysis performed for the same strains based on MLST approach showed that strains containing IHR at specific genomic positions group with subspecies *fastidiosa* strains with MLST (Fig. 4-5B). This further reinforced that the variable regions detected in the genomes of subsp. *fastidiosa* strains originated from subsp. *multiplex.*

Table 4-2: Predicted genomic regions of intersubspecific recombination in *Xylella fastidiosa* subsp. *fastidiosa* strains from subsp. *multiplex* donors

Strain	**Subspecies**	**MLST Sequence type (ST)[a]**	**Length of variable region**	**#SNPs with *fastidiosa* strains[b]**	**#SNPs with *multiplex* strains[c]**
WM1-1	*fastidiosa*	2	8384bp (530784-539167)[d]	272-273	3-7
WM1-1	*fastidiosa*	2	5650bp (1893724-1899373)[d]	97- 98	11-17
TemeculaL	*fastidiosa*	1	8777bp (1066742-1075518)[d]	211-216	6-8
pglA (Fetzer)	*fastidiosa*	4	8777bp (1066742-1075518)[d]	211-216	6-8
CCPM1	*fastidiosa*	1	5432bp (172056-1725938)[d]	102-103	5-11

[a]Sequence type of strains based on MLST of seven house-keeping genes. All sequence types (1, 2, and 4) group within subsp. *fastidiosa* based on previous studies
[b]Number of single nucleotide polymorphism with subsp. *fastidiosa* strains at the variable region
[c]Number of single nucleotide polymorphism with subsp. *multiplex* strains at the variable region
[d]Position of variable region based on Temecula1 numbering of the genome

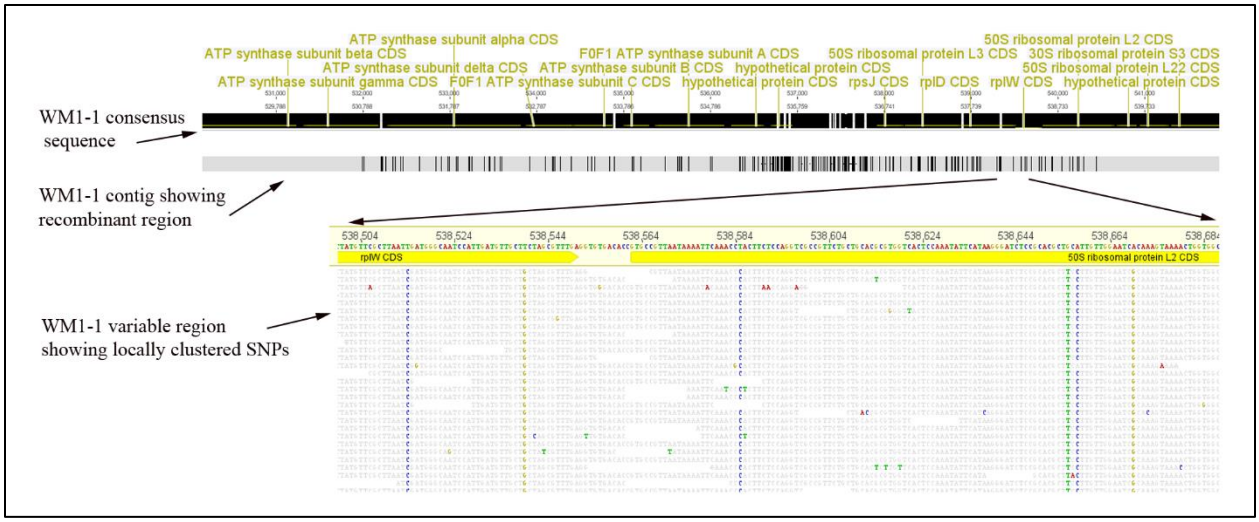


Fig. 4-4. Intersubspecific recombination in strain WM1-1. Variable region was observed as marked by dark stripes (middle panel) when contig generated by De Novo assembly was aligned to Temecula1 genome. A zoomed-in version of variable region observed in reference mapping of the WM1-1 reads showing six occurrences of SNPs within a span of 180bp is shown (lower panel). Region of WM1-1 changed by intersubspecific recombination. Genes involved in ATP synthase and 50S ribosomal subunits were altered by this intersubspecific recombination in WM1-1 (upper panel).

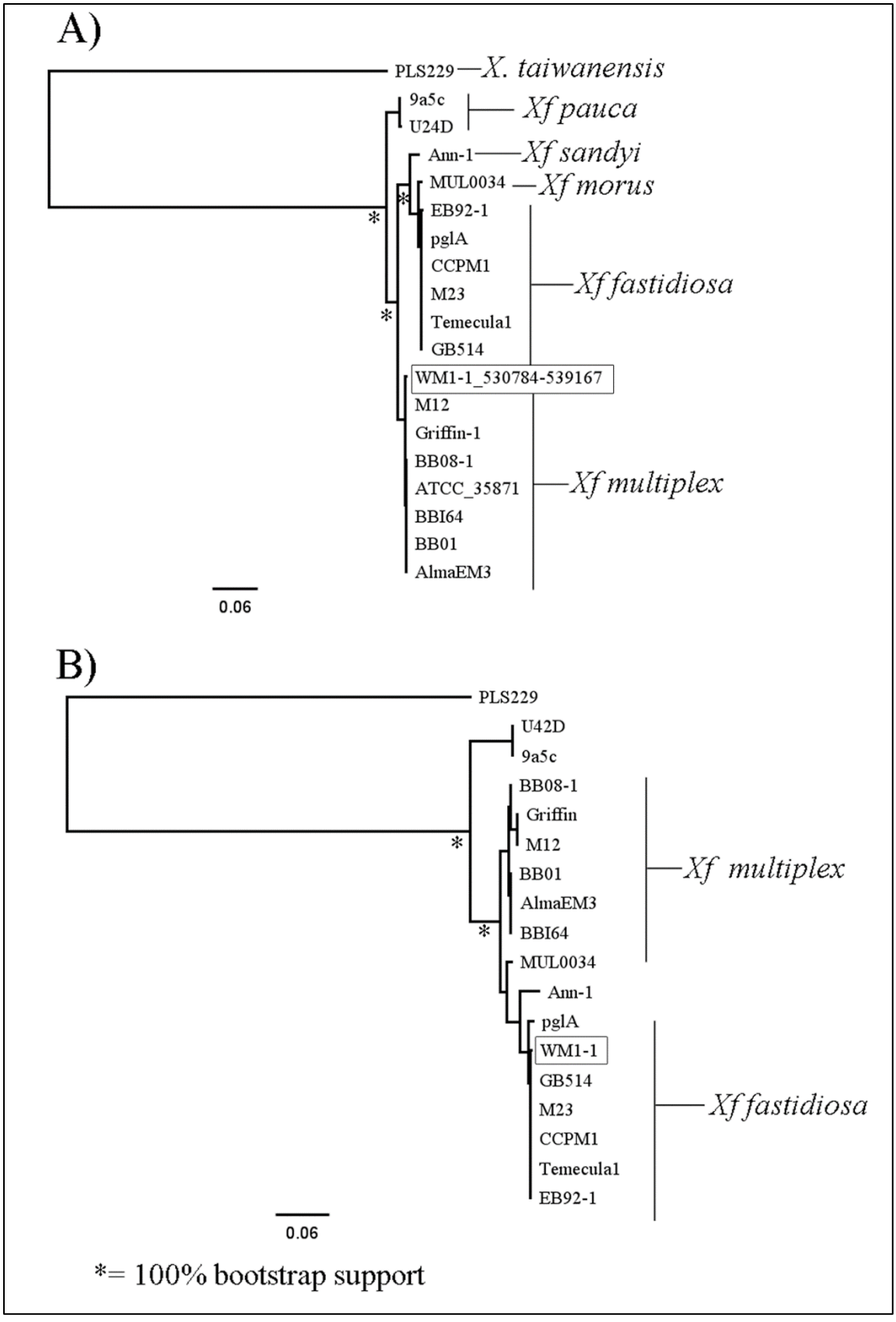

A)
PLS229
X. taiwanensis
9a5c
U24D
Xf pauca
Ann-1
Xf sandyi
MUL0034
Xf morus
EB92-1
pglA
CCPM1
M23
Temecula1
GB514
Xf fastidiosa
WM1-1_530784-539167
M12
Griffin-1
BB08-1
ATCC_35871
BBI64
BB01
AlmaEM3
Xf multiplex
0.06
B)
PLS229
U42D
9a5c
BB08-1
Griffin
M12
BB01
AlmaEM3
BBI64
Xf multiplex
MUL0034
Ann-1
pglA
WM1-1
GB514
M23
CCPM1
Temecula1
EB92-1
Xf fastidiosa
0.06
*= 100% bootstrap support

Fig. 4-5. (A) RAxML tree of *Xylella fastidiosa* strains from different subspecies based on WM1-1 recombinant sequence. WM1-1 is grouped with subspecies *multiplex* strains at the recombinant region (boxed). **(B)** RAxML tree of the same strains (except ATCC35871) based on concatenated sequences of seven MLST genes. Strain WM1-1 groups with subspecies *fastidiosa* strains based on MLSA. Scale bar represents number of substitutions per site. Trees were generated by RAxML version 7.2.8 with rapid bootstrapping and search for best-scoring ML tree algorithm using 1,000 replications. Scale bar indicates number of substitutions per site.

Phylogenetic tree of *X. fastidiosa* strains based on whole genome sequence alignment

A phylogenetic tree (Fig. 4-6) generated based on the genomes derived from the Temecula1 reference mapping and other reference genomes obtained from NCBI (Table 4-2) using the subsp. *pauca* strain 9A5C as outgroup showed that the wild-type *X. fastidiosa* subsp. *fastidiosa* strains group together forming separate clades within the tree. The *in vitro* recombinant strains group together within a same clade with their recipient parent confirming that recombination occurred only at a specific genomic location, and therefore this region is not sufficient to change the phylogenetic grouping. Moreover, the tree shows two separate groupings of the Temecula strains (Temecula1, Temecula1*, and gfp mutant grouping together, while TemeculaL, NS1-CmR, and the recombinants of TemeculaL and NS1pglA forming a separate clade). Strains WM1-1 isolated from Georgia formed a separate clade while the other Georgia strain CCPM1 was intermediate between Georgia and California strains. Subsp. *multiplex* strain AlmaEM3 isolated from blueberry plant in Georgia grouped together with another *multiplex* strain M12 isolated from almond plants in California (Fig. 4-6).

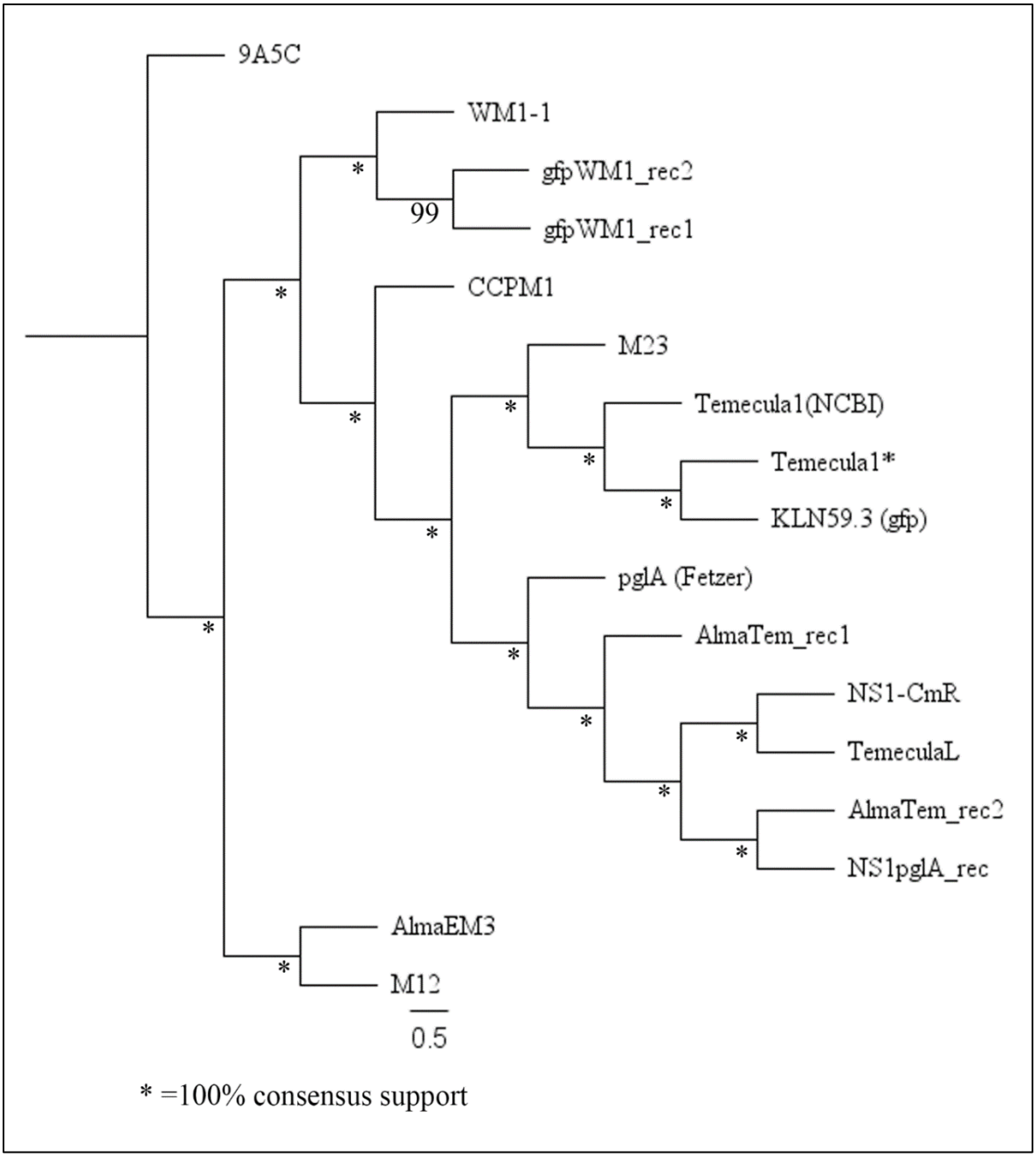


Fig. 4-6. Phylogenetic tree of *X. fastidiosa* wild-type strains and *in vitro* recombinants based on whole genome. All *in vitro* recombinants group together with the recipient parents while the wild-type strains form separate groupings. Genomes were aligned with MAFFT and tree was generated using RAxML with rapid bootstrapping algorithm using 1,000 replications strain 9A5C was used as outgroup. Consensus tree was generated by collapsing nodes with less than 50% support. Scale bar indicates number of substitutions per site.

Discussion

The ease with which sequences of whole genomes can now be obtained provides unprecedented opportunities to study various phenomena related to disease epidemics, pathogen emergence, and virulence of microbial pathogens associated with humans, animals, and plants. By comparing isolates that infect different hosts, virulence factors utilized by these pathogens in host colonization can be identified. Moreover, by determining the genetic relatedness of isolates infecting different crops or occurring at different geographic locations, the disease epidemics can be studied and decisions on management options made accordingly. More importantly, patterns of genetic exchanges that lead to genetic diversity and disease emergence can be studied in more detail with the availability of whole genome sequence data. However, use of next generation sequencing techniques in generating whole genome sequences is error-prone, mainly because genomes are sequenced in the form of very short (e.g. Illumina) to medium sized (e.g. PacBio) fragments (reads). Also there are inherent error probabilities associated with various sequencing platforms (Fox et al., 2014). Therefore, quality control is critical at every step of the sequencing plan from library preparation, quality filtering, assembly, to quality assessment of the genome. Availability of a closely identical reference genome is very useful in examining the quality of sequencing results. *X. fastidiosa* was the first plant pathogenic bacterium to have its full genome sequenced (Simpson et al., 2000). Subsequent complete genome sequencing of a grapevine strain, Temecula1, and comparative genomic analysis between the grapevine and citrus strain demonstrated that the two genomes are very similar (Van Sluys et al., 2003). More genomes representing other strains and subspecies have been sequenced almost to their entirety (Chen et al., 2010; Zhang et al., 2011). This provides a definite opportunity for draft genomes produced

by next-generation sequencing platforms to compare highly similar complete genomes of *X. fastidiosa* to assess the quality of draft genomes.

The initial sequencing run for this study had a high Phred score (>30 for more than 80% of the base calls), FASTQC analysis showed that most of the genomes passed the quality criteria, and finally a strict quality filtering was used to trim the data. Reference mapping of the reads showed that most of the reads were aligned to the reference genome resulting in a high reference sequence identity of 96-100%. All subsp. *fastidiosa* genomes had more than 99.3% identity with the Temecula1 reference while subsp. *multiplex* strain AlmaEM3 had 96.3% identity. Moreover, all the genomes achieved a high depth of coverage (at least 20x to 180x). De Novo assembly statistics indicated that a robust assembly of the genomes was achieved except for one genome (AlmaTemeculaL recombinant 2 that had relatively higher number of contigs and higher assembly size). The number of contigs was drastically reduced by a hybrid approach of assembling the Illumina and PacBio reads together. This can be expected as the longer read lengths of PacBio system can close repeat regions in the genome that cannot be covered by short reads of Illumina sequencing. However, even with the hybrid approach, a complete contiguous chromosome could not be obtained, possibly due to the presence of repeat regions that are >7kb according to Temecula1 genome analysis. Previous draft genome assemblies of *X. fastidiosa* have also reported similar results of genome fragmentation (Zhang et al., 2011; Chen et al., 2016). A BLAST search of the contigs generated showed that most of the contigs aligned perfectly with the genomes of *X. fastidiosa* in the NCBI database. Generally, the assembly size, GC content, and number of gene features obtained from annotation results were consistent with the genomes deposited in NCBI (Chen et al., 2016; Marcelletti and Scortichini, 2016b). Moreover, phylogenetic analysis based on MLST and whole genomes obtained from this study

were comparable to previously described *X. fastidiosa* taxonomy (Yuan et al., 2010; Marcelletti and Scortichini, 2016b) suggesting that the genome sequences obtained in this study are of high quality.

X. fastidiosa is a rapidly emerging plant pathogen around the world (Saponari et al., 2013; Almeida and Nunney, 2015). Genetic studies demonstrated extensive of HR within and between subspecies of *X. fastidiosa* (Scally et al., 2005; Almeida et al., 2008; Nunney et al., 2012; Nunney et al., 2014a; Nunney et al., 2014b; Coletta-Filho et al., 2016), predicting that HR between subspecies is involved in plant-host shift and disease emergence. Our recent study demonstrated experimentally that recombination potential differs among *X. fastidiosa* strains and IHR occurs *in vitro* between subsp. *multiplex* and subsp. *fastidiosa* (Kandel et al., 2017). Strains of subsp. *fastidiosa* showed higher recombination frequencies than those of subsp. *multiplex*. However, no evidence of phenotypic change was observed in the recombinants that were selected randomly (Kandel et al., 2017). This led to the determination of the minimum limits of recombination which were found to be 0.7-4kb per recombination event (Kandel et al., 2017). In that study, however, recombination was examined only at the flanking region of antibiotic insertion site due mainly to the limitation of the method used for sequencing (Sanger method). Therefore, to understand the limits and broader implications of HR, a whole genome analysis approach could be more useful.

A whole genome sequencing approach has been applied to detect recombination in recent studies (Marttinen et al., 2012; Yahara et al., 2014; Croucher et al., 2015). Approaches involved in detecting recombination look for regions of divergence that are not expected for the given clade under consideration in a multiple sequence alignment. These regions are hypothesized to be imported from a distinct donor (Croucher et al., 2015). Therefore, elevated levels of

polymorphisms in a nucleotide sequence alignment are an indication that this region could be a result of recombination. This approach may or may not be able to suggest the donor source depending upon whether the donor sequence is available for alignment. Although, this approach could be useful for alignment of short sequences such as MLSA, it becomes problematic when the sequence lengths become longer.

Other approaches such as ClonalFrame (Didelot and Falush, 2007) and its improvement for whole genome sequences, ClonalFrameML (Didelot and Wilson, 2015) take into account both SNPs at a specific region (e.g. the variable region) over the SNPs across the whole sequence and infer location of recombination based on the presence of elevated levels of SNPs. However, these approaches cannot model the origin or the donor source and cannot predict recombination if fewer than expected number of SNPs are present in recombination region (Didelot and Maiden, 2010). A similar approach referred to as Gubbins (Genealogies Unbiased By Recombination In Nucleotide Sequences) was developed for large datasets (Croucher et al., 2015) and can predict the regions of recombination in a genome, but still does not take into account the origin of the sequence. Moreover, this approach requires clonal populations (e.g. within a same sequence type). This method was used in our dataset and recombination regions were detected, but could not model the source of the recombinant region requiring a different approach. Another method, ClonalOrigin, can model recombination from a specific donor to a specific recipient (Didelot and Maiden, 2010), but requires clonal genealogy to be specified and is computationally intensive (Yahara et al., 2014).

In this study, a very straight-forward approach was used to detect recombination and infer its source in the genomes of *X. fastidiosa*. After detecting HR by Gubbins, genomes were compared to a common and very similar reference genome, and variable regions were identified. The

variable regions were then used in a BLAST search to identify their origin. Although, it was time consuming and required extra manual effort limiting its use for large datasets, the results are accurate as shown by the phylogenetic analysis. In a previous study using MLST (Nunney et al., 2012), a genetic introgression test was designed to detect IHR in short regions of subsp. *pauca* strains from Brazil. This method compared variation in native sequences to a probable donor, which had to be at least a different sequence than the native ones. Here, we compared native sequences to find variable regions, which were then inspected in the database to trace their origin.

With the method used here, IHR was detected both in the *in vitro* generated recombinant and wild-type *X. fastidiosa* strains collected from infected plants. Moreover, this approach provided the maximum limit of recombination to be up to 10kb per recombination event (based on our limitation to detect only regions that differ in sequence between the donor and the recipient strain). Recombination was randomly distributed across the genome such that some strains were recombinant at one genomic position, while others at other positions. For example, strains WM1-1 and CCPM1, that were isolated from different grapevine fields in Georgia, showed recombination at different genomic positions. Moreover, recombination present in TemeculaL and Fetzer was not present in Temecula1, although these strains originated from California. This indicates that recombination occurs randomly at different genomic locations whenever genetically different donor and recipient co-exist, producing several new variant strains containing genetic changes at different regions of the genome. Therefore, change in phenotype of recombinants can be only expected when recombination occurs at regions conferring virulence or fitness. In our test, the highly virulent strain WM1-1 (Oliver et al., 2014; Kandel et al., 2017) showed regions of recombination in the genes involved in ATP synthase and 50S ribosomal

subunit. The contribution of this recombinant region in the fitness and virulence of WM1-1 remains to be determined, but it can be expected that this genetic change can alter phenotype as it occurs in the genes important for fitness and metabolism. The presence of this variable region in strain WM1-1 was also confirmed by examining the genomes of two recombinants, gfpWM1-1 recombinants 1 and 2, that were generated by using WM1-1 as the recipient and therefore, possess the same region as WM1-1 at this position. Moreover, recombinant regions detected in TemeculaL were also detected in AlmaTemeculaL recombinants 1 and 2, in which TemeculaL was used as the recipient. It should be noted here that few other variable regions were also detected in the genomes but were not further analyzed as sequences of these regions were more identical to the strains of same subspecies i.e., subsp. *fastidiosa* than with a different subspecies. These regions could have originated by intra-subspecific recombination.

The high number of SNPs detected in the variable regions of strains used here are extremely unlikely to be generated by point mutation within a short fragment, as SNPs caused by point mutations should be randomly scattered, not locally-clustered. Moreover, it was shown that the polymorphic regions matched with subsp. *multiplex* genome sometimes with 100% identity. This confirms that the variable regions are generated by IHR between *X. fastidiosa* subspecies. MLSA of house-keeping genes of the strains that carried IHR demonstrated that they are true to type to their taxonomic grouping except at the IHR region, a result also supported by whole genome phylogeny.

IHR between subsp. *fastidiosa* and *multiplex* was also detected in North American strains in previous MLST studies (Nunney et al., 2014a; Nunney et al., 2014b). As the two subspecies have been reported to exist together in North America for more than 100 years (Nunney et al., 2010) recombination can be expected between them. Moreover, Nunney et al (2010) reported similar

variable regions when comparing two subsp. *fastidiosa* genomes: Temecula1 and M23. Strains of both subspecies served as both donors and recipients for one another showing a bi-directional pattern of HR in a previous study (Kandel et al., 2017). Therefore, data from both previous studies and this study confirm that intersubspecific recombination is extremely common, and strengthen the proposition of new disease emergence due to intersubspecific recombination in *X. fastidiosa*. Careful planning and implementation of policies to limit introduction of novel genotypes and to avoid co-existence of pre-existing genotypes is required to minimize this threat.

References

Almeida, R.P., and Nunney, L. 2015. How do plant diseases caused by *Xylella fastidiosa* emerge? Plant Dis 99:1457-1467.

Almeida, R.P., Nascimento, F.E., Chau, J., Prado, S.S., Tsai, C.W., Lopes, S.A., and Lopes, J.R. 2008. Genetic structure and biology of *Xylella fastidiosa* strains causing disease in citrus and coffee in Brazil. Appl Environ Microbiol 74:3690-3701.

Aziz, R.K., Bartels, D., Best, A.A., DeJongh, M., Disz, T., Edwards, R.A., Formsma, K., Gerdes, S., Glass, E.M., and Kubal, M. 2008. The RAST Server: rapid annotations using subsystems technology. BMC Genomics 9:75.

Baltrus, D.A., Guillemin, K., Phillips, P.C., and Travisano, M. 2008. Natural transformation increases the rate of adaptation in the human pathogen *Helicobacter pylori*. Evolution 62:39-49.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., Pyshkin, A.V., Sirotkin, A.V., Vyahhi, N., Tesler, G., Alekseyev, M.A., and Pevzner, P.A. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455-477.

Bolger, A.M., Lohse, M., and Usadel, B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114-2120.

Chen, J., Xie, G., Han, S., Chertkov, O., Sims, D., and Civerolo, E.L. 2010. Whole genome sequences of two *Xylella fastidiosa* strains (M12 and M23) causing almond leaf scorch disease in California. J Bacteriol 192:4534.

Chen, J., Wu, F., Zheng, Z., Deng, X., Burbank, L.P., and Stenger, D.C. 2016. Draft Genome Sequence of *Xylella fastidiosa* subsp. *fastidiosa* Strain Stag's Leap. Genome Announc 4:e00240-00216.

Coletta-Filho, H., Francisco, C., Lopes, J., Muller, C., and Almeida, R. 2016. Homologous recombination and *Xylella fastidiosa* host-pathogen associations in South America. Phytopathology.

Croucher, N.J., Page, A.J., Connor, T.R., Delaney, A.J., Keane, J.A., Bentley, S.D., Parkhill, J., and Harris, S.R. 2015. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. Nucleic Acids Res 43:e15.

Davis, M.J., French, W.J., and Schaad, N.W. 1981. Axenic culture of the bacteria associated with phony disease of peach and plum leaf scald. Curr Microbiol 6:309-314.

Didelot, X., and Falush, D. 2007. Inference of bacterial microevolution using multilocus sequence data. Genetics 175:1251-1266.

Didelot, X., and Maiden, M.C. 2010. Impact of recombination on bacterial evolution. Trends Microbiol 18:315-322.

Didelot, X., and Wilson, D.J. 2015. ClonalFrameML: efficient inference of recombination in whole bacterial genomes. PLoS Comput Biol 11:e1004041.

Dorer, M.S., Sessler, T.H., and Salama, N.R. 2011. Recombination and DNA repair in *Helicobacter pylori*. Annu Rev Microbiol 65:329-348.

Doyle, J.J. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem bull 19:11-15.

Edgar, R.C. 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics 5:113.

EFSA. 2016. Update of a database of host plants of *Xylella fastidiosa*: 20 November 2015. EFSA J 14:4378-n/a.

Feil, E.J., and Spratt, B.G. 2001. Recombination and the population structures of bacterial pathogens. Annu Rev Microbiol 55:561-590.

Feil, E.J., Holmes, E.C., Bessen, D.E., Chan, M.S., Day, N.P., Enright, M.C., Goldstein, R., Hood, D.W., Kalia, A., Moore, C.E., Zhou, J., and Spratt, B.G. 2001. Recombination within natural populations of pathogenic bacteria: short-term empirical estimates and long-term phylogenetic consequences. Proc Natl Acad Sci USA 98:182-187.

Fox, E.J., Reid-Bayliss, K.S., Emond, M.J., and Loeb, L.A. 2014. Accuracy of Next Generation Sequencing Platforms. Next Generat Sequenc & Applic 1:1000106.

Fraser, C., Hanage, W.P., and Spratt, B.G. 2007. Recombination and the Nature of Bacterial Speciation. Science 315:476-480.

Freel, K.C., Millán-Aguiñaga, N., and Jensen, P.R. 2013. Multilocus sequence typing reveals evidence of homologous recombination linked to antibiotic resistance in the genus Salinispora. Appl Environ Microbiol 79:5997-6005.

Gurevich, A., Saveliev, V., Vyahhi, N., and Tesler, G. 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics 29:1072-1075.

Guttman, D.S. 1997. Recombination and clonality in natural populations of *Escherichia coli*. Trends Ecol Evol 12:16-22.

Hacker, J., and Carniel, E. 2001. Ecological fitness, genomic islands and bacterial pathogenicity. A Darwinian view of the evolution of microbes. EMBO Rep 2:376-381.

Harris, R.S. 2007. Improved pairwise alignment of genomic DNA. The Pennsylvania State University.

Hopkins, D.L., and Purcell, A.H. 2002. *Xylella fastidiosa*: Cause of Pierce's Disease of Grapevine and Other Emergent Diseases. Plant Dis 86:1056-1066.

Kandel, P.P., Lopez, S.M., Almeida, R.P., and De La Fuente, L. 2016. Natural competence of *Xylella fastidiosa* occurs at a high frequency inside microfluidic chambers mimicking the bacterium's natural habitats. Appl Environ Microbiol 82:5269-5277.

Kandel, P.P., Almeida, R.P.P., Cobine, P.A., and De La Fuente, L. 2017. Natural competence rates are variable among *Xylella fastidiosa* strains and homologous recombination occurs in vitro between subspecies *fastidiosa* and *multiplex*. Mol Plant Microbe Interact 30:589-600.

Katoh, K., and Standley, D.M. 2013. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Mol Biol Evol 30:772-780.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., and

Drummond, A. 2012. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28:1647-1649.

Kong, Y., Ma, J.H., Warren, K., Tsang, R.S., Low, D.E., Jamieson, F.B., Alexander, D.C., and Hao, W. 2013. Homologous recombination drives both sequence diversity and gene content variation in *Neisseria meningitidis*. Genome Biol Evol 5:1611-1627.

Kung, S.H., and Almeida, R.P. 2011. Natural competence and recombination in the plant pathogen *Xylella fastidiosa*. Appl Environ Microbiol 77:5278-5284.

Maiden, M.C. 2006. Multilocus sequence typing of bacteria. Annu Rev Microbiol 60:561-588.

Marcelletti, S., and Scortichini, M. 2016a. Xylella fastidiosa CoDiRO strain associated with the olive quick decline syndrome in southern Italy belongs to a clonal complex of the subspecies pauca that evolved in Central America. Microbiology 162:2087-2098.

Marcelletti, S., and Scortichini, M. 2016b. Genome-wide comparison and taxonomic relatedness of multiple *Xylella fastidiosa* strains reveal the occurrence of three subspecies and a new *Xylella* species. Arch Microbiol 198:803-812.

Marttinen, P., Hanage, W.P., Croucher, N.J., Connor, T.R., Harris, S.R., Bentley, S.D., and Corander, J. 2012. Detection of recombination events in bacterial genomes from large population samples. Nucleic Acids Res 40:e6.

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

Michod, R.E., Bernstein, H., and Nedelcu, A.M. 2008. Adaptive value of sex in microbial pathogens. Infect Genet Evol 8:267-285.

Nelson, K., and Selander, R.K. 1994. Intergeneric transfer and recombination of the 6-phosphogluconate dehydrogenase gene (gnd) in enteric bacteria. Proc Natl Acad Sci USA 91:10227-10231.

Newman, K.L., Almeida, R.P., Purcell, A.H., and Lindow, S.E. 2003. Use of a green fluorescent strain for analysis of *Xylella fastidiosa* colonization of *Vitis vinifera*. Appl Environ Microbiol 69:7319-7327.

Nunney, L., Yuan, X., Bromley, R.E., and Stouthamer, R. 2012. Detecting genetic introgression: high levels of intersubspecific recombination found in *Xylella fastidiosa* in Brazil. Appl Environ Microbiol 78:4702-4714.

Nunney, L., Hopkins, D.L., Morano, L.D., Russell, S.E., and Stouthamer, R. 2014a. Intersubspecific recombination in *Xylella fastidiosa* strains native to the United States: infection of novel hosts associated with an unsuccessful invasion. Appl Environ Microbiol 80:1159-1169.

Nunney, L., Schuenzel, E.L., Scally, M., Bromley, R.E., and Stouthamer, R. 2014b. Large-scale intersubspecific recombination in the plant-pathogenic bacterium *Xylella fastidiosa* is associated with the host shift to mulberry. Appl Environ Microbiol 80:3025-3033.

Nunney, L., Yuan, X., Bromley, R., Hartung, J., Montero-Astua, M., Moreira, L., Ortiz, B., and Stouthamer, R. 2010. Population genomic analysis of a bacterial plant pathogen: novel insight into the origin of Pierce's disease of grapevine in the U.S. PLoS One 5:e15488.

Oliver, J.E., Sefick, S.A., Parker, J.K., Arnold, T., Cobine, P.A., and De La Fuente, L. 2014. Ionome changes in *Xylella fastidiosa*-infected *Nicotiana tabacum* correlate with virulence and discriminate between subspecies of bacterial isolates. Mol Plant Microbe Interact 27:1048-1058.

Parker, J.K., Havird, J.C., and De La Fuente, L. 2012. Differentiation of *Xylella fastidiosa* strains via multilocus sequence analysis of environmentally mediated genes (MLSA-E). Appl Environ Microbiol 78:1385-1396.

Rasko, D.A., Webster, D.R., Sahl, J.W., Bashir, A., Boisen, N., Scheutz, F., Paxinos, E.E., Sebra, R., Chin, C.-S., and Iliopoulos, D. 2011. Origins of the *E. coli* strain causing an outbreak of hemolytic–uremic syndrome in Germany. N Engl J Med 365:709-717.

Roper, M.C., Greve, L.C., Warren, J.G., Labavitch, J.M., and Kirkpatrick, B.C. 2007. *Xylella fastidiosa* requires polygalacturonase for colonization and pathogenicity in *Vitis vinifera* grapevines. Mol Plant Microbe Interact 20:411-419.

Saponari, M., Boscia, D., Nigro, F., and Martelli, G.P. 2013. Identification of DNA sequences related to *Xylella fastidiosa i*n Oleander, Almond and Olive trees exhibiting leaf scorch symptoms in Apulia (Southern Italy). J Plant Pathol 95:668-668.

Scally, M., Schuenzel, E.L., Stouthamer, R., and Nunney, L. 2005. Multilocus sequence type system for the plant pathogen *Xylella fastidiosa* and relative contributions of recombination and point mutation to clonal diversity. Appl Environ Microbiol 71:8491-8499.

Schreiber, H.L., Koirala, M., Lara, A., Ojeda, M., Dowd, S.E., Bextine, B., and Morano, L. 2010. Unraveling the first *Xylella fastidiosa* subsp. *fastidiosa* genome from Texas. Southwest Entomol 35:479-483.

Seemann, T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068-2069.

Simpson, A.J.G., Reinach, F., Arruda, P., Abreu, F., Acencio, M., Alvarenga, R., Alves, L.C., Araya, J., Baia, G., and Baptista, C. 2000. The genome sequence of the plant pathogen *Xylella fastidiosa*. Nature 406:151-157.

Spratt, B.G., Bowler, L.D., Zhang, Q.Y., Zhou, J., and Smith, J.M. 1992. Role of interspecies transfer of chromosomal genes in the evolution of penicillin resistance in pathogenic and commensal *Neisseria species*. J Mol Evol 34:115-125.

Strona, G., Carstens, C.J., and Beck, P.S. 2017. Network analysis reveals why *Xylella fastidiosa* will persist in Europe. Sci Rep 7:71.

Su, C.C., Deng, W.L., Jan, F.J., Chang, C.J., Huang, H., Shih, H.T., and Chen, J. 2016. *Xylella taiwanensis* sp. nov., causing pear leaf scorch disease. Int J Syst Evol Microbiol 66:4766-4771.

Thomas, C.M., and Nielsen, K.M. 2005. Mechanisms of, and barriers to, horizontal gene transfer between bacteria. Nat Rev Microbiol 3:711-721.

Timilsina, S., Jibrin, M.O., Potnis, N., Minsavage, G.V., Kebede, M., Schwartz, A., Bart, R., Staskawicz, B., Boyer, C., and Vallad, G.E. 2015. Multilocus sequence analysis of xanthomonads causing bacterial spot of tomato and pepper plants reveals strains generated by recombination among species and recent global spread of *Xanthomonas gardneri*. Appl Environ Microbiol 81:1520-1529.

Van Sluys, M.A., de Oliveira, M.C., Monteiro-Vitorello, C.B., Miyaki, C.Y., Furlan, L.R., Camargo, L.E., da Silva, A.C., Moon, D.H., Takita, M.A., Lemos, E.G., Machado, M.A., Ferro, M.I., da Silva, F.R., Goldman, M.H., Goldman, G.H., Lemos, M.V., El-Dorry, H., Tsai, S.M., Carrer, H., Carraro, D.M., de Oliveira, R.C., Nunes, L.R., Siqueira, W.J., Coutinho, L.L., Kimura, E.T., Ferro, E.S., Harakava, R., Kuramae, E.E., Marino, C.L., Giglioti, E., Abreu, I.L., Alves, L.M., do Amaral, A.M., Baia, G.S., Blanco, S.R., Brito, M.S., Cannavan, F.S., Celestino, A.V., da Cunha, A.F., Fenille, R.C., Ferro, J.A., Formighieri, E.F., Kishi, L.T., Leoni, S.G., Oliveira, A.R., Rosa, V.E., Jr., Sassaki, F.T., Sena, J.A., de Souza, A.A., Truffi, D., Tsukumo, F., Yanai, G.M., Zaros, L.G., Civerolo, E.L., Simpson, A.J., Almeida, N.F., Jr., Setubal, J.C., and Kitajima, J.P. 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.

Yahara, K., Didelot, X., Ansari, M.A., Sheppard, S.K., and Falush, D. 2014. Efficient inference of recombination hot regions in bacterial genomes. Mol Biol Evol 31:1593-1605.

Yan, S., Liu, H., Mohr, T.J., Jenrette, J., Chiodini, R., Zaccardelli, M., Setubal, J.C., and Vinatzer, B.A. 2008. Role of recombination in the evolution of the model plant pathogen *Pseudomonas syringae* pv. tomato DC3000, a very atypical tomato strain. Appl Environ Microbiol 74:3171-3181.

Yuan, X., Morano, L., Bromley, R., Spring-Pearson, S., Stouthamer, R., and Nunney, L. 2010. Multilocus sequence typing of *Xylella fastidiosa* causing Pierce's disease and oleander leaf scorch in the United States. Phytopathology 100:601-611.

Zhang, S., Flores-Cruz, Z., Kumar, D., Chakrabarty, P., Hopkins, D.L., and Gabriel, D.W. 2011. The *Xylella fastidiosa* biocontrol strain EB92-1 genome is very similar and syntenic to Pierce's disease strains. J Bacteriol 193:5576-5577.

Supplementary material for chapter 4

Supplementary table 4-1: Assembly statistics of sequenced genomes

Strain	**#Reads[a]**	**#Reads trim[b]**	**Read length**	**Covg[c]**	**#Ctgs[d]**	**Size (Mbp)[e]**	**N50[f]**	**L50[g]**	**Ref[h] (%)**
WM1-1	2942684	2587777	138.7	145.2	113 34[i]	2.45 2.47[i]	148434 238450[i]	5 4[i]	99.7
AlmaEM3	722996	654360	135.7	35.8	100 30[i]	2.46 2.48[i]	115714 271106[i]	7 4[i]	96.3
AlmaTemL_rec1	1268086	1143060	131.1	59.4	120	2.52	115978	7	99.3
AlmaTemL_rec2	1234660	640702	150.5	27.6	402	2.74	103360	6	99.9
TemeculaL	2140558	1334223	113.8	57.4	166	2.54	93631	8	99.8
Temecula1*	1960132	1751487	153.9	109.7	113	2.45	99386	6	100
pglA mutant (Fetzer)	818192	741373	127.4	37.9	146	2.49	97642	6	99.8
NS1-CmR	319068	316673	158.8	20.0	254	2.51	27752	29	99.7
NS1-pglA_rec	1400214	1285143	151.0	76.8	124	2.53	135060	5	99.9
KlN59.3 (gfp)	3415272	3090572	146.2	183.6	117	2.46	103359	6	100
GfpWM1-1_rec1	3133800	2842830	147.6	171.2	105	2.45	134853	6	99.6
GfpWM1-1_rec2	3048060	2730936	140.4	156.2	125	2.45	134871	6	99.6

[a]Number of reads produced by the sequencing run; [b]number of reads that survived trimming; [c]average depth of coverage; [e]number of scaffolds generated by SPAdes De Novo assembly; [f]size of the contig such that more than half of the assembly is contained in contigs longer than this value; [g]number of contigs that contain more than 50% of the total assembly; [h]percent identity with the reference Temecula1 genome based on reference mapping
[i]Assembly statistics of a hybrid assembly from Illumina and PacBio reads

Chapter 5

A rapid and efficient method for gene knockout and genetic complementation in *Xylella fastidiosa* using overlap-extension PCR and natural transformation confirms that the type IV pilin paralog predicted to be regulated by the alternative σ^{54} is the functional pilin

Abstract

Twitching motility is a major virulence factor of the plant-pathogenic bacterium *Xylella fastidiosa.* Cells of *X. fastidiosa* at one of the cell poles possess type IV pili, which are thread-like filamentous appendages involved in twitching motility. Genomic analysis of *X. fastidiosa* showed the presence of at least four paralogs of the *pilA* gene that encodes the major pilin subunit of type IV pili. However, whether all paralogs have a functional role in type IV pili biogenesis is not known. Here, using an efficient method of genome manipulation based on overlap-extension PCR and natural genetic transformation, deletion mutants of two of the *pilA* paralogs (*pilA1* PD1924 and *pilA2* PD1926) were generated in *X. fastidiosa* subspecies *fastidiosa* strains WM1-1 and TemeculaL. Twitching motility of the mutants was then assessed. Results show that deletion of *pilA2* that was predicted to be regulated by an alternative sigma factor (σ^{54}) caused loss of twitching motility, whereas deletion of *pilA1* did not influence twitching motility in either strains. Moreover, loss of twitching motility caused by *pilA2* deletion was restored when a wild-type copy of the *pilA2* gene was complemented at a neutral site in the genome of one of the mutants. This study demonstrates that homologous PCR templates generated by overlap-extension PCR can be used to rapidly generate gene knockouts and perform genetic complementation in *X. fastidiosa*, and that twitching motility in *X. fastidiosa* is controlled by regulating the transcription of the major pilin subunit that is predicted to be under the control of σ^{54} factor.

Introduction

Xylella fastidiosa is a Gram-negative, xylem-limited, insect-vectored, plant-pathogenic bacterium causing incurable diseases and substantial economic losses to the growers of grapevines, citrus, coffee, plum, and almond (Hopkins and Purcell, 2002; Chatterjee et al., 2008). Pierce's disease (PD) of grapevines in the United States and citrus variegated chlorosis (CVC) in Brazil, two most impactful diseases of *X. fastidiosa*, are major limiting factors in the production of these crops for those countries (Hopkins and Purcell, 2002). The xylem sap-feeding insect-vectors sharpshooter leafhoppers and spittlebugs transmit and inoculate the bacterial cells into the plant xylem vessels, where the cells attach to the surface, form biofilms, and move along and across the xylem vessels (Newman et al., 2003; Roper et al., 2007). The severity of disease symptoms depends on the extent of xylem vessel occlusions produced by bacterial aggregates (Newman et al., 2003). Therefore, the potential of cells for aggregation and systemic colonization of the xylem vessels are considered the major determinants of pathogenicity in *X. fastidiosa*.

X. fastidiosa cells are non-flagellated but possess two types of pili at one of their cell poles, short type I pili involved in cell attachment and biofilm formation, and long type IV pili involved in twitching motility (Meng et al., 2005; De La Fuente et al., 2007; Li et al., 2007). Twitching motility involves flagella-independent translocation of cells on a moist surface facilitated by extension, tethering, and retraction of type IV pili (Mattick, 2002). It was demonstrated in *X. fastidiosa* that a mutant lacking type I pili (*fimA* mutant) was biofilm-deficient but twitching-enhanced, whereas the mutants lacking type IV pili were biofilm-enhanced and twitching deficient (Meng et al., 2005; De La Fuente et al., 2007; Li et al., 2007). Moreover, mutants that

were twitching-enhanced traveled further upstream from the point of inoculation in grapevines (Meng et al., 2005). In addition, *X. fastidiosa* strains that lacked or exhibited reduced twitching motility (Kandel et al., 2017) had less severe disease symptoms in the model plant tobacco and natural host blueberry (Oliver et al., 2014; Oliver et al., 2015). Also, a polygalacturonase gene mutant that is compromised in lateral movement due to its inability to degrade xylem pit membranes was avirulent in grapevines (Roper et al., 2007) and was twitching deficient (Kandel et al., 2016). Furthermore, seven transposon mutants, including a mutant of hemagglutinin that mediates cell-cell aggregation, showed hypervirulent phenotypes and faster movement in grapevines than their wild-types (Guilhabert and Kirkpatrick, 2005). A previous study has also demonstrated the role of cell-cell communication system involving a diffusible signaling factor (DSF) in regulating host colonization, movement, and biofilm formation in *X. fastidiosa* (Newman et al., 2004). Interestingly, it was shown that mutant that was unable to produce DSF was hypervirulent to grapevines (Newman et al., 2004) and had increased expression of several type IV pili genes compared to the wild-type strain (Wang et al., 2012).

These observations suggest that motility of cells is one of the important virulence factors of *X. fastidiosa*. As noted before, type IV pili are the structures involved in motility of *X. fastidiosa* cells. Involvement of type IV pili and twitching motility in host colonization and virulence were also described in other Gram-negative bacteria including human pathogens such as *Pseudomonas aeruginosa* (Burrows, 2012), *Neisseria meningitidis* (Morand et al., 2009), *N. gonorrhoeae* (Wolfgang et al., 1998); and plant pathogens such as *Ralstonia solanacearum* (Liu et al., 2001), *Acidovorax citrulli* (Bahar et al., 2009), *Xanthomonas citri* (Dunger et al., 2014), and *Pseudomonas syringae* (Taguchi and Ichinose, 2011). Besides twitching motility, type IV pili

were also reported to contribute other functions like biofilm formation, electron transfer (Mattick, 2002; Craig et al., 2004), and natural DNA transformation (Seitz and Blokesch, 2013).

Structurally type IV pili are flexible, filamentous appendages composed of polymers of a single pilin protein-subunit that is encoded by *pilA* gene. However, extremely complex machineries are involved in the biogenesis and function of type IV pili. For example, in *P. aeruginosa*, where twitching motility and type IV pili have been characterized in detail, approximately 40 genes are required for the function of type IV pili including genes for pilin subunit, pilin assembly and retraction, and regulatory genes (Mattick, 2002). Genomes of *X. fastidiosa* (Simpson et al., 2000; Van Sluys et al., 2003) contain numerous genes that share signature motifs and properties of well described pili genes. Role of some of these genes such as *pilB, pilO, pilQ, pilR, pilY1* in type IV pili biogenesis and twitching motility was demonstrated in previous studies (Meng et al., 2005; De La Fuente et al., 2007; Li et al., 2007; Cursino et al., 2011; Cruz et al., 2014). Moreover, environmental signals such as calcium (Cruz et al., 2014), pectin (Killiny and Almeida, 2009), and chitin (Killiny et al., 2010) were shown to regulate one-to-few twitching motility genes. Other studies demonstrated the involvement of chemotaxis related genes (Pil-Chp) in regulating twitching motility and expression of type IV pili genes (Cursino et al., 2011; Hao et al., 2017). In addition, at least six open reading frames (ORFs) that encode characteristic pilin subunits were predicted in the genome of CVC strain 9A5C and one of the ORFs was predicted to be regulated by RpoN, the alternative sigma factor 54, σ^{54} (da Silva Neto et al., 2008). At least four homologs of the *pilA* ORFs are also present in the grapevine strain Temecula1 (Van Sluys et al., 2003). Protein products of two of these homologs, referred to as PilA1 (PD1924) and PilA2 (PD 1926) in previous studies (Wang et al., 2012; Parker et al., 2016), share complete set of functional

features of the type IV pilin subunit (Mattick, 2002). Expression analysis based on whole transcriptomics showed that both of these *pilA* paralogs are expressed in Temecula1, although the expression of *pilA2* was greater than that of *pilA1* (Parker et al., 2016). Other studies (de Souza et al., 2003; 2005; Caserta et al., 2010; Cruz et al., 2014) have considered *pilA1* paralog in their expression analysis to determine the influence of various environmental parameters on twitching motility. However, the exact role of these *pilA* paralogs remains poorly understood, and whether both or one of them has a major functional role in type IV pili biogenesis and twitching motility is not known, a clear understanding of which can help future gene expression studies to target the functional pilin paralog providing a better estimate of the regulation of type IV pili biogenesis and twitching motility in *X. fastidiosa*.

Therefore, the objective of this study was to characterize the function of the two pilin paralogs, *pilA1* (PD1924) and *pilA2* (PD1926), in two *X. fastidiosa* subspecies *fastidiosa* strains WM1-1 and TemeculaL. Presence of the *pilA* paralogs in the genomes of these strains was first confirmed, followed by screening of their promoter sequences to identify binding sites of regulatory elements. Separate deletion mutants of each paralogs were generated in both strains and double deletion mutant of both paralogs was generated in WM1-1. Moreover, a complemented mutant of *pilA2* deletion was generated in strain TemeculaL. Twitching motility of all deletion mutants and the complemented mutant was assessed. Noteworthy is that this study used a new gene deletion and complementation method that utilized overlap-extension PCR to generate a marker-tagged homologous template and natural genetic transformation of *X. fastidiosa*. Results show that this method can be used to rapidly and efficiently generate gene deletion mutants and complemented mutants of the deleted genes in *X. fastidiosa*. Furthermore,

the *pilA* paralog (previously referred to as *pilA2* in the Temecula1 strain) predicted to be regulated by σ^{54} (RpoN) based on the analysis of its promoter sequence was shown to be the major pilin subunit gene as deletion of this paralog caused loss of twitching motility that was restored by genetic complementation.

Materials and Methods

Bacterial strains, media, and culture conditions

X. fastidiosa subsp. *fastidiosa* strains WM1-1 and TemeculaL isolated from infected vineyards in Georgia and California, respectively were used in this study. Strains were cultured on PW (Davis et al., 1980) agar plates, modified by replacing phenol red with 1.8 g litre^{-1} of Bovine Serum Albumin (BSA, Gibco Life Sciences Technology). Each strain was initially cultured from -80°C glycerol stock for one week at 28°C and re-streaked onto new PW plates for another week before use. PD3 medium (Davis et al., 1981) was used for culturing and suspending cells in liquid. Antibiotics kanamycin and chloramphenicol were used at a concentration of 30 and 10 µg ml^{-1}, respectively.

DNA extraction and overlap-extension PCR for construction knockout template

For DNA extraction, cells cultured on PW agar plates were suspended in 400µl of sterile MilliQ water. DNA was extracted using Quick-DNA™ Fungal/Bacterial Miniprep Kit (Zymo Research) using manufacturers protocol. Bead beating for DNA extraction was performed for 5min using Mini-BeadBeater-96™ (BioSpec Products).

Knockout templates targeting four genomic regions according to Temecula1 reference sequence were constructed. The four regions were upstream non-coding region of *pilQ* (PD1691), downstream region of *pilM* (PD1695) that disrupted the *mrcA* gene encoding penicillin-binding protein 1A, *pilA1* coding region (PD1924), and *pilA2* coding region (PD1926). These templates were named as 1691 construct, 1695 construct, 1924 construct, and 1926 construct, respectively. Constructs 1691 and 1924 contained kanamycin resistant cassette, while the other two constructs 1695 and 1926 contained chloramphenicol resistant cassette. For creating each knockout construct, approximately 0.7-1.2kb upstream and downstream flanking regions of the targeted genomic region were PCR amplified using primer pairs UP_F/UP_R and Dn_F/Dn_R, respectively (Table 5-1). Antibiotic cassettes were amplified using Antb_F and Antb_R. DNA of strain WM1-1was used as template for upstream and downstream fragments, while plasmids pUC4K, and pAX1.Cm were used as template for kanamycin and chloramphenicol cassettes, respectively. Primers Up_R and Dn_F were 5'-extended with at least 21bp homologous regions of the antibiotic cassette to facilitate overlap-extension of these fragments. The three separately amplified fragments (upstream, antibiotic, and downstream) were purified from agarose gel and mixed in equal proportions in an overlap-extension PCR reaction using the end-primers (UP_F and Dn_R). The fusion product was purified from gel and stored at -20°C before use or used as template for amplification with the internal primer pair int_F/int_R. All the primers (Table 5-1) were designed with Primer3 Primer Design Tool in Geneious software (Biomatters). PCR was performed with standard protocol using iProof™ High-Fidelity PCR Kit (Bio-Rad) in S 1000^{TM} thermal cycler (Bio-Rad). PCR products were gel-purified using either Gel Extract Mini Kit (5 PRIME) or Freeze 'N Squeeze™ DNA Gel Extraction Spin Columns (Bio-Rad).

Natural transformation of *X. fastidiosa* strains with PCR template and confirmation of mutants

Transformation of *X. fastidiosa* cell was performed with the gel-purified or internal primer-amplified PCR template using natural transformation protocol as previously described (Kandel et al., 2016; Kandel et al., 2017). Briefly, recipient *X. fastidiosa* strain WM1-1 and TemeculaL cultures from PW plates were suspended in PD3 liquid medium and optical density (OD_{600}) was adjusted to 0.25 (~10^8 cells/ml). Ten µl of this suspension was spotted onto PD3 agar plates and 10µl PCR template was added on top of the spots. Spots were allowed to dry for about an hour and the plates were incubated at 28°C for three days. After three days, each spot was suspended in 0.5ml of PD3 and 100 µl aliquots were spread-plated onto PW plates containing respective antibiotics. After two weeks of incubation at 28°C, mutant CFUs were enumerated and re-streaked onto new antibiotic PW plates. Control spots of the strain cells without the addition of PCR template were also included for each experiment. Double mutants were generated by using *pilA1* mutants as recipients with the gel purified PCR template of *pilA2* deletion (1926 construct).

Confirmation of gene deletion was carried out by PCR and sequencing. PCR was performed with three sets of primer pairs, namely the out pairs (out_F/out_R) that target the genomic region outside the recombination region of the construct, antibiotic primers (antb_F/antb_R), and respective gene primers (pilA1_F/pilA1_R and PD1926_F/ PD1926_F). For each PCR, wild-type strains were included as controls. Sanger sequencing of the mutant at the deleted genomic region was performed using primers targeting the out-region, internal-region, and antibiotic-region. Information about all the primers is presented in Table 5-1.

Growth curve and biofilm formation

Growth curve and biofilm formation of the mutants were assessed together with the wild-types in 96-well plates for five days as previously described (Kandel et al., 2017). Experiments were repeated three times with three plates per time except for one experiment in which single plate was used.

Twitching motility

Twitching motility assessment of wild-type and mutant strains was performed in PD3 plates and PD3 plates supplemented with 50% Chardonnay sap as previously described (Kandel et al., 2017). Twitching motility was observed for at least ten transformant colonies for each mutant. Selected colonies were re-streaked and repeated observation of these colonies for twitching motility was performed.

Complementation of *pilA2* mutants

Complementation of *pilA2* mutants was performed with the PCR-amplified and gel-purified complementation template prepared from the fusion PCR of three fragments (upstream *pilA2* complementation template, *pilA2* template containing the *pilA2* ORF flanked by its promoter and terminator region, and downstream *pilA2* complementation template). The upstream *pilA2* complementation construct was amplified from NS1::Km WM1-1 (Kandel et al., 2017) using the primer pairs pilA2C_up_F and pilA2C_up_R (Table 5-1). These primers amplify the upstream flanking region of NS1 and the kanamycin cassette in a single fragment. The 950 bp *pilA2* fragment contained 237bp upstream promoter region, *pilA2* coding region, and 216bp downstream region. This fragment was amplified using primer pairs pilA2_F and pilA2_R using

the wild-type WM1-1 DNA as template. The downstream complementation product was also amplified from wild-type WM1-1 DNA using primers pilA2C_dn_F and pilA2C_dn_R and is same as the downstream region of NS1 site. Primers pilA2C_up_R, pilA2F, pilA2R, and pilA2C_Dn_F at their 5' end had a 27bp extension that was homologous to the adjoining fragment. The individual fragments were amplified, gel-purified, and fused with extension PCR as described above. The PCR fragments were then used in a natural transformation assay with the *pilA2* mutants of WM1-1 and TemeculaL.

RNA preparation and quantitative PCR

X. fastidiosa TemeculaL cells cultured in PW plates were suspended in PD3 and adjusted to OD_{600} of 0.25 as previously mentioned. Three 15 ml plastic tubes containing thee ml of one of the mediums (PD3, PD3 supplemented with three g/liter of BSA or 100% Chardonnay sap) were supplemented with 120 µl of the OD_{600} adjusted TemeculaL cell suspension. Tubes were incubated at 28°C with shaking at 150 rpm for three days. Cells were then precipitated by centrifugation of the tubes at 5000 rpm for 5 min at 4°C. The supernatant medium was discarded and the pellet formed at the bottom of the tube was suspended in 200 µl of DNA/RNA shield (Zymo Research) and RNA was extracted using Quick-RNA™ MiniPrep (Plus) kit (Zymo Research) using in-column DNase treatment following the manufacturers protocol. RNA concentration was measured by Nanodrop (NanoDrop 2000 spectrophotometer, Thermo Scientific). Complementary DNA (cDNA) was prepared using qScript cDNA SuperMix (Quantabio) with 170ng of normalized RNA in 20µl volume. Quantitative PCR was performed using 1 µl of cDNA in 20 µl volume containing 4 µl of 5x PerfeCTa MultiPlex qPCR ToughMix (Quantabio), 0.4 µM forward and reverse primers and 0.2 µM TaqMan probe (labeled with 5' 6-

carboxyfluorescein [FAM] and 3' Black Hole 588 Quencher-1 [BHQ-1]). Reactions were performed using CFX96^{TM} Real-Time System (Bio-Rad) with following cycling parameters: 95°C for 1 min, followed by 40 cycles of 95°C for 15 seconds and 60°C for 1 min. Fold change in expression of *pilA1*, and *pilA2* was calculated by $2^{-\Delta\Delta CT}$ method (Schmittgen and Livak, 2008) using expression of *nuoA* gene as the endogenous control. Primers and probes used for gene expression are also included in Table 5-1**.** Experiments were performed once with three replicates per media type.

Table 5-1: PCR and qRT-PCR primers and probes used in this study

Primer name	**Sequence (5'-3')**	**Description**	**Amplicon size (bp)**	**Reference**
1924_up_F	GCGGCACCACGTATATCAATAAAA	Amplify upstream region of *pilA1* (PD1924)	981	This study
1924_up_R	GCAACACCTTCTTCACGAGGCAGACAGAGAATTCATGTGTGGGGGTTTA[a]			
1924_dn_F	GAGATTTTGAGACACAACGTGGCTTATGAATACACACAGCAACACGATC	Amplify downstream region of *pilA1* (PD1924)	1190	This study
1924_dn_R	TTGGAGAAGAGGCGTGTTAAAAAC			
1924_int_F	ATATACAGGGTGCTTGCTGATTGA	Amplify internal region of *pilA1* construct	2952	This study
1924_int_R	GGATGGGTTTAGGGATGCTGATAA			
1924_out_F	AAACCACCACCGATAACAACAATC	Confirm *pilA1* mutants (primers target outside the pilA1 construct)	3509 2884[b]	This study
1924_out_R	AATAGCGTTGGTAAGAAATCCAGC			
1926_up_F	CATTTCACTTTGACTTCACCCGAA	Amplify upstream region of *pilA2* (PD1926)	1032	This study
1926_up_R	TGCCCCGTATTCAGTGTC			

	GCTGATTAATAGCGTTGGTAAGAAATCCAGC			
1926_dn_F	GCCTGGTGCTACGCCTGAATAAGTGAAACACGATTCATGGGTAAATGCTC	Amplify downstream region of *pilA2* (PD1926)	700	This study
1926_dn_R	TGAGCGTCAATTTTAGAGGATGGA			
1926_int_F	TCAGCAATACTCATACTGGCACTT	Amplify internal region of *pilA2* construct	2773	This study
1926_int_R	AACGTGTGCTTGAATCTTCGAATT			
1926_out_F	ACAAGAGTGAGCCGTTACAACTAT	Confirm pilA1 mutants (primers target outside the pilA1 construct)	3004 2352[b]	This study
1926_out_R	CTTTTCCAATGAGCAGTTATCGGG			
Kan_F	GTCTGCCTCGTGAAG	Amplify kanamycin cassette	1204	This study
Kan_R	AAGCCACGTTGTGT			
Cm_F	AATCAGCGACACTGAATACGG	Amplify chloramphenicol cassette	1119	This study
Cm_R	TCACTTATTCAGGCGTAGCAC			
pilA2C_up_F	CGCGCCCGTTATTAATCGAA	Amplify upstream fragment of *pilA2* complementation construct (contains upstream of NS1 region and kanamycin cassette)	1957	This study
pilA2C_up_R2[c]	ATATTGAAGGGTGCAATACAAAGCATCTAGTCTCAACCATCATCGATGAA			
pilA2C_dn_F2[c]	GTTTCAAGAGAGAGAGCGTTCAACACGATGCTGTTAACCATTGTCATC	Amplify downstream fragment of *pilA2* complementation construct (this is same as NS1 downstream fragment)	799	This study
pilA2C_dn_R	TAACCTTGTCAGCGTAGATG			

pilA2_F	ACAATTCATCGATGATG GTTGAGACTAGATGCTTT GTATTGCACCCTTCAAT	Amplify *pilA2* coding region, 237bp upstream, and 216bp downstream to include promoter and terminator regions	950	This study
pilA2_R	GCCATTGATGACAATGG TTAACAGCATCGTTTCAA GAGAGAGAGCGTTCAAC			
pilA_F	AAACACCGGACTTGCCA ACATCAC	Primer and probe to amplify fragment of *pilA1*	140	Cruz et al; 2014
pilA_R	TGTTGCATGTCCACTGAC CTCCAT			
pilA_P[d]	AAACCATCGCTTGGAAT CGTAGCGTCGA			
PD1926_F	CACTCCTAACGCTATTGG ACTAC	Primer and probe to amplify fragment of *pilA2*	117	Parker et al; 2016
PD1926_R	TTGACCTGACCATTACCA ATCA			
PD1926_P[d]	TGGTGGACATCACAACT ACTGGCG			
nuoA_F	AGACGCACGGATGAAGT TCGATGT	Reference gene for qRT-PCR		Cruz et al; 2016
nuoA_R	ATTCCAGCGCTCCCTTCT TCCATA			
nuoA_Pd[d]	TTCATCGTGCCTTGGACT CAGGTGTT			
1691_up F	AGGCAACCTGACAGCGA TAC	Amplify upstream region of 1691 construct	541	This study
1691_up_R2[c]	GCAACACCTTCTTCACGA GGCAGACCGTTGATGTTC GAGAAGTGCG			
1691_Dn_F2[c]	GAGATTTTGAGACACAA CGTGGCTTTGCATCAACC CCAAAGCTGA	Amplify downstream region of 1691 construct	504	This study
1691_Dn R	AAGCGATCCAATGAAGG GCT			
1691_int_F	ACGGCCCTTCTCTAAGAT TGC	Amplify internal region of 1691	2152 987[b]	This study

1691_int_R	TGTGTGGTGCTTGCATATTCTG	construct		
1695_up_F	CGATGGCCTGTTGATAGCGATA	Amplify upstream region of 1695 construct	1027	This study
1695_up_R	CCGTATTCAGTGTCGCTGATTGTCACCACGTTTGAGGAGTTTG			
1695_Dn_F	GTGCTACGCCTGAATAAGTGAAACGGGGTGAATGGACATTAG	Amplify downstream region of 1695 construct	1018	This study
1695_Dn_R	CAGATGGGGAGTGCTGCTTTA			
1695_int_F	AAAGACGAAATCCTGGAGCTGTAT	Amplify internal region of 1695 construct	2640 1546[b]	This study
1695_int_R	TTGATACCAATTGGAAGACAACGC			

[a]Underlined region indicates 5' extended region of the primer to facilitate overlap-extension.
[b]Amplicon size in the wild type strain without the insertion of antibiotic gene.
[c]These primers were re-designed as the previously designed primers for this region did not work.
[d]These were probes used for qRT-PCR and were labeled with FAM at the 5' end and BHQ1 at the 3'end.

Results

PCR template transformed *X. fastidiosa* cells generating targeted gene deletion mutants

Transformants were generated from all the PCR templates used for strains WM1-1 and TemeculaL (Table 5-2). Ten colonies per strain per construct were re-streaked on to new antibiotic plates. After confirming that the antibiotic gene is inserted in the desired genomic region, one colony per strain for each construct were selected for further analysis. The *pilA* mutants were selected for further analysis, while the other two mutants of upstream *pilQ* region and *mrcA* gene were not characterized and are not discussed further in this study. For *pilA* mutants, PCR confirmed that the construct replaced targeted *pilA* paralog from the desired

location of the genome (Fig. 5-1). Double mutants were also confirmed using the same method (data not shown). Sequencing results showed that the antibiotic cassette is inserted in the targeted genomic location without altering the flanking region sequence. The antibiotic resistant phenotype of the mutants was maintained even after re-streaking the mutant colonies for up to five generations onto new antibiotic PW plates and after culturing the mutant cells in non-antibiotic PW plates suggesting that the mutants generated by this method are stable.

Table 5-2: PCR transformants and twitching phenotype of mutants generated with various PCR templates

PCR Template	**Recipient strain/mutant**	**#transformants/0.1ml (mean±se)**	**Mutant phenotype**
pilA1 (1924)	WM1-1	53.4±15.9	Twitch plus
	TemeculaL	35.6±20.7	Twitch plus
pilA2 (1926)	WM1-1	27.6±4.5	Twitch minus
	TemeculaL	6.2±3.4	Twitch minus
	pilA1 WM1-1	19±5	Twitch minus
pilA2 complementation template	*pilA2* WM1-1	0	
	pilA2 TemeculaL	8.5±3.5	Twitch plus
Upstream *pilQ (*1691)[a]	WM1-1	51±24.8	Twitch plus
mrcA (1695)[a]	WM1-1	57±29.2	Twitch minus

[a]These regions were only used to confirm that PCR template can be used for transformation. These mutants are not further discussed in this study.

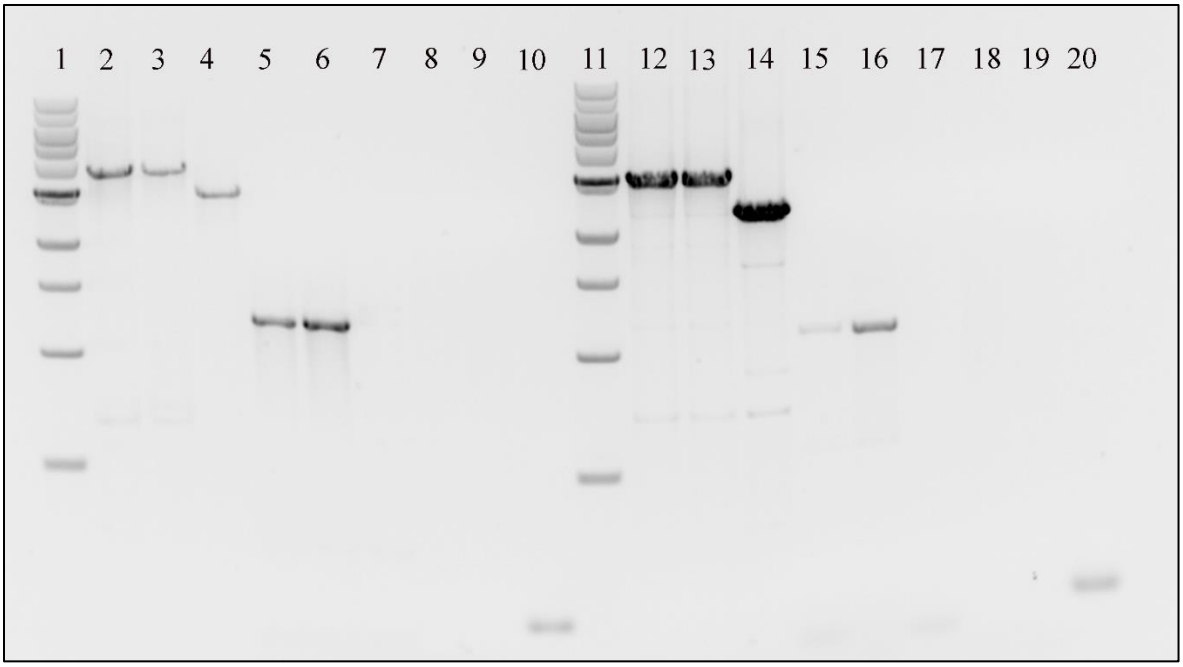


Fig. 5-1 Confirmation of *pilA1* and *pilA2* deletion from *Xylella fastidiosa* strains WM1-1 and TemeculaL. PCR confirmation was performed using three different PCR for each mutant. Out-PCR target genomic region outside the construct recombination site, kanamycin and chloramphenicol PCRs target the antibiotic region inserted in the mutants, and *pilA* PCRs target the respective *pilA* paralogs. Lanes 1 and 11; 1kb ladder, lanes 2 and 3; *pilA1* out PCR of *pilA1* mutants and lane 4 wild-type for *pilA1* out PCR, lanes 5 and 6; kanamycin PCR of *pilA1* mutants and lane 7 wild-type for kanamycin PCR , lanes 8 and 9; *pilA1* PCR of *pilA1* mutants and lane 10 wild-type for *pilA1* PCR, lanes 12 and 13; *pilA2* out PCR of *pilA2* mutants and lane 14 wild-type for *pilA2* out-PCR, lanes 15 and 16; chloramphenicol PCR of *pilA2* mutants and lane 17 wild-type for chloramphenicol PCR, ladder 18 and 19; *pilA2* PCR of *pilA2* mutants and ladder 20 wild-type for *pilA2* PCR. Longer PCR fragment from out-PCR in the mutants than in the wild-types suggest insertion of antibiotic cassettes, which is confirmed by the presence of band with antibiotic PCR in mutants but not in wild-types. Amplification with the target-gene specific PCR shows that the mutants lack the target genes and therefore no amplification, but the wild-type have intact gene therefore there is amplification (refer table 5-1 for amplicon sizes).

Growth curve and biofilm formation of *pilA* mutants

Growth curves of the mutants were generated by culturing the cells in 96-well plates and measuring the optical density every day for five days. Growth curves (Fig. 5-2) show that growth pattern is similar for the mutants and wild-type. Biofilm formation significantly increased in both *pilA* mutants and the double mutant compared to the wild-type in WM1-1, while only *pilA2* mutant formed significantly higher biofilm than the wild-type in TemeculaL backgrounds (Fig. 5-3).

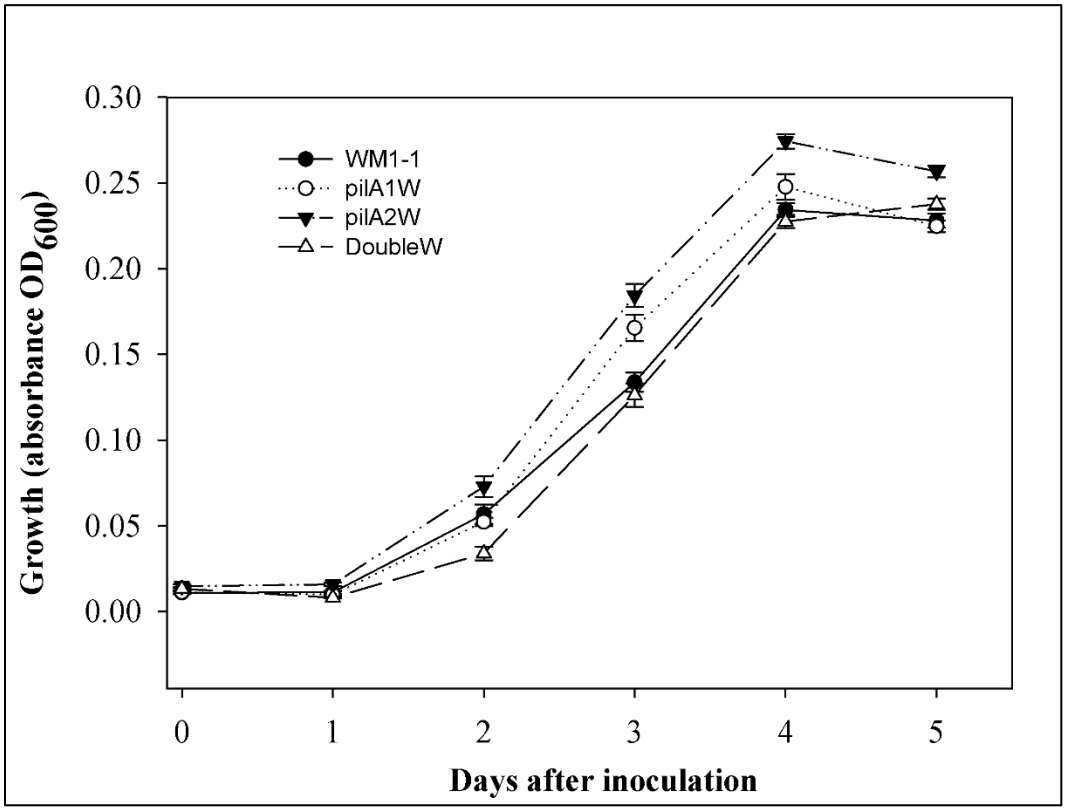


Fig. 5-2. Growth curves of *pilA1*, *pilA2*, and double mutants of strain WM1-1 assessed by measuring optical density (OD_{600}nm) every day for five days. Experiments were repeated three times with three plates per experiment for two experiments and on plate for one experiment. At least six wells per plate were used for each strain.

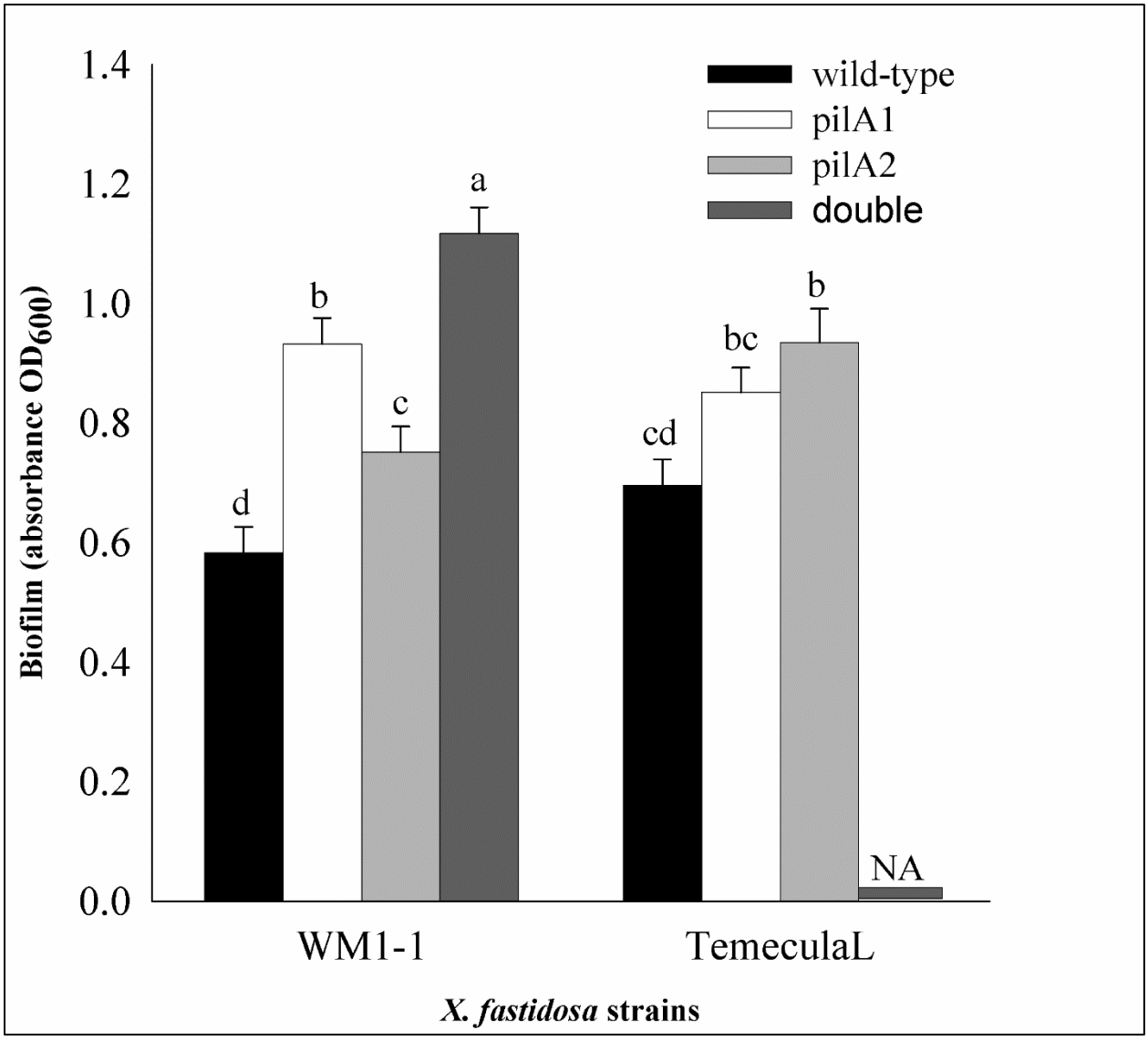


Fig. 5-3 Biofilm formation of *Xylella fastidiosa* wild-type strains and mutants as measured by 0.1% crystal violet staining. Staining of biofilm was performed at the end of the growth curve experiment. Data were analyzed in SAS 9.3 with PROC GLIMMIX. Error bars indicate standard error of mean. Different letters within each group indicate significant differences at 5% significance level. NA mutant not available for measurement.

Twitching motility of wild-type, mutant, and complemented strains

Formation of colony fringe was observed to assess twitching motility as shown in Fig. 5-4. Both wild-type and *pilA1* mutants of WM1-1 and TemeculaL displayed twitching motility, whereas the *pilA2* mutants were twitching-deficient for both strains. The double mutant of WM1-1 was also twitching deficient (Fig. 5-4). Similar pattern of twitching was observed in media supplemented with 50% chardonnay sap (data not shown).

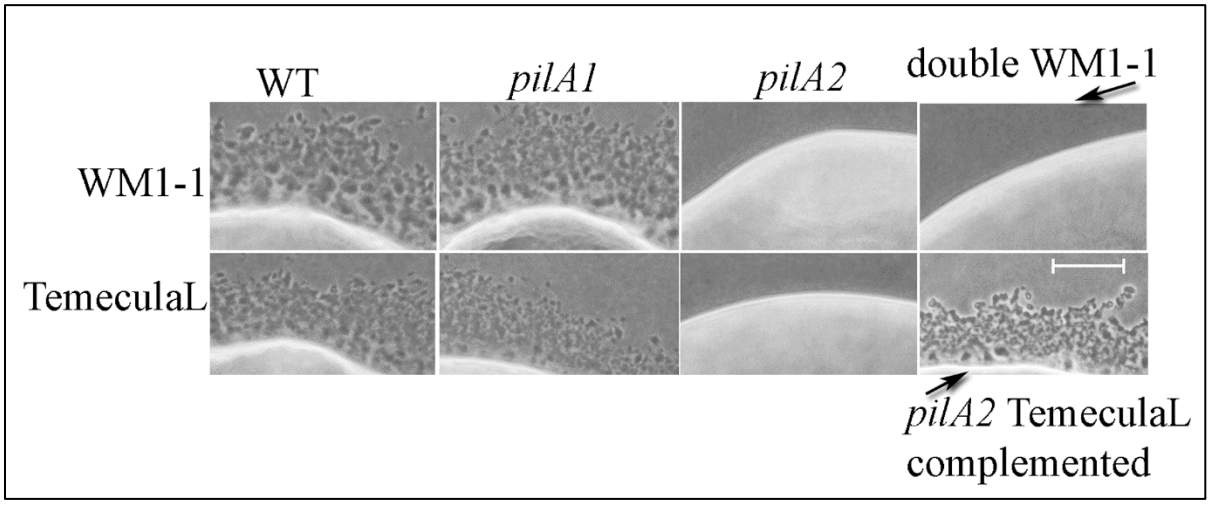


Fig. 5-4. Twitching motility of wild-type, *pilA* mutants, and complemented *pilA2* mutant colonies. Twitching motility was observed for at least ten independent transformants with equivalent results obtained in all cases. For selected transformants for each background, observation was made at different time points. Scale bar on lower right panel indicates 100µm.

Complementation of *pilA2* re-stored twitching motility

pilA2 mutants that were deficient in twitching motility were complemented using the wild-type copy of *pilA2* ORF and its native promoter and terminator at the neutral site 1 of *X. fastidiosa* TemeculaL genome. *pilA2* mutants of WM1-1 could not be transformed with the *pilA2*

complementation template. TemeculaL mutants could be successfully complemented with the fusion product of upstream, *pilA2*, and downstream PCR template. Seventeen transformants (Table 5-2) were generated in total, and five were re-streaked on new PW plates supplemented with both antibiotics (kanamycin and chloramphenicol). Two of these colonies were confirmed with PCR and sequencing for the insertion of *pilA2* fragment in the neutral site. Twitching motility of five transformant colonies selected for re-streaking showed that complementation re-stored twitching motility lost by *pilA2* deletion (Fig. 5-4).

Sequence analysis of *pilA* paralogs and their promoter

Based on gene expression analysis using a microarray experiment, a pilin paralog (XF2542) was predicted to be the major pilin of *X. fastidiosa* subsp. *pauca* strain 9A5C in a previous study (da Silva Neto et al., 2008). BLAST search using the 9A5C sequence of this paralog as query against Temecula1 genome showed PD1926 as the closest homologue with 98% coverage and 87% amino acid sequence identity, while the other paralog PD1924 had 97% coverage and 65% identity. On pairwise alignment, the two paralogs (PD1924 and PD1926) of Temecula1 were 68% identical by nucleotide sequence and 62% identical by amino acid sequence. However, both contained all the characteristic features of a pilin subunit including the leader peptide, cleavage site, N-terminal hydrophobic region, pilin domain, cytoplasmic, transmembrane, and extra-cellular domains etc. (data not shown). On comparing regulatory regions in the promoter based on the regions predicted by the previous study (da Silva Neto et al., 2008), putative binding sites of the alternative sigma factor σ^{54} were identified in the promoter of PD1926 in both Temecula1 and WM1-1 (Fig. 5-5). Therefore, results of deletion and complementation experiment and the promoter sequence analysis confirm that the *pilA* paralog encoded by PD1926 is the functional

pilin in *X. fastidiosa* subsp. *fastidiosa* strains WM1-1 and TemeculaL and the expression of this pili is predicted to be under σ^{54} regulation.

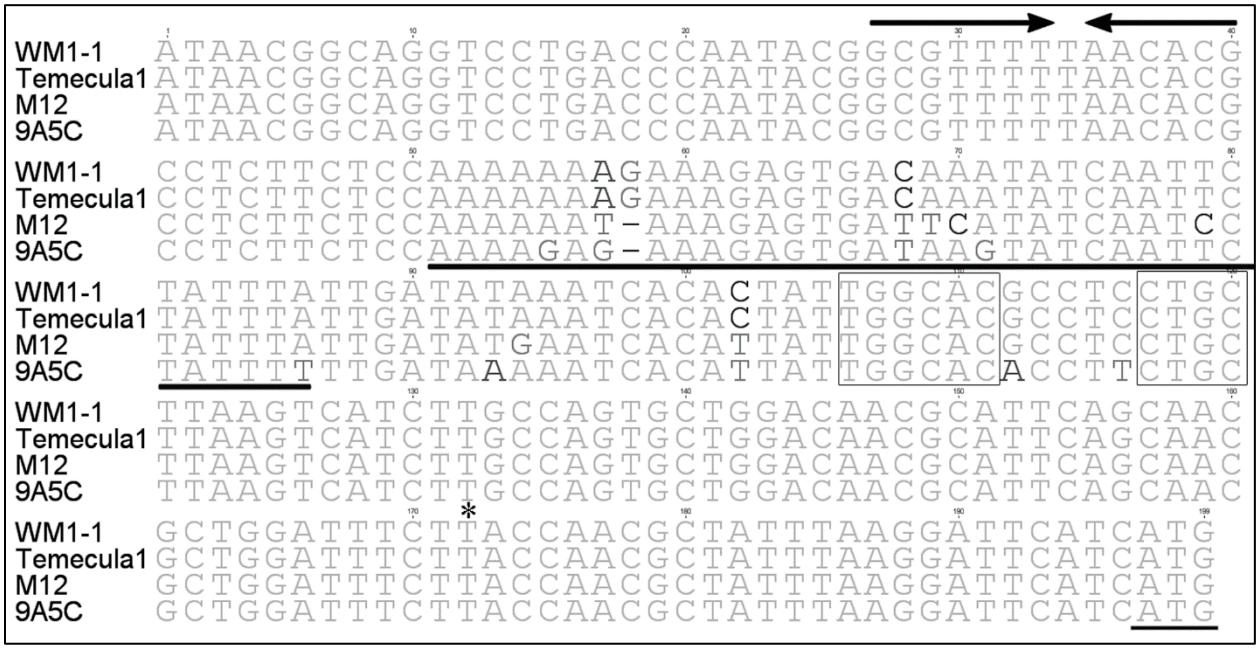


Fig. 5-5. Determination of characteristic features of a σ^{54} binding sites in the promoter of *PilA2* (PD1926) of *Xylella fastidiosa* subspecies *fastidiosa* strains WM1-1 and Temecula1. Subsp. *multiplex* strain M12 and subsp. *pauca* strain 9A5C are included for comparison. Translation start site (ATG), and integration host factor binding sites are underlined. Asterisk (*) indicates the conserved transcription start site that is 66bp upstream of the translation start site predicted previously for strain 9A5C. Boxed regions indicate conserved elements of -12 (CTGC) and -24 (TGGCAC) regions of a typical σ^{54} factor. Arrows indicate conserved inverted repeats, the possible enhancer binding sites of the σ^{54} promoter.

Expression of *pilA2* was higher than *pilA1* and supplementation of BSA in PD3 reduced *pilA2* expression

Expression of the two *pilA* paralogs was compared among three different media using qRT-PCR using the housekeeping gene *nuoA* as endogenous reference gene. As shown in Fig. 5-6A, expression of *pilA1* was not changed by growth in PD3 supplemented with BSA or grapevine sap. However, expression of *pilA2* was significantly reduced (P=0.006) by addition of BSA in PD3. Expression of *pilA2* was significantly higher than that of *pilA1* in all media tested (Fig. 5-6B).

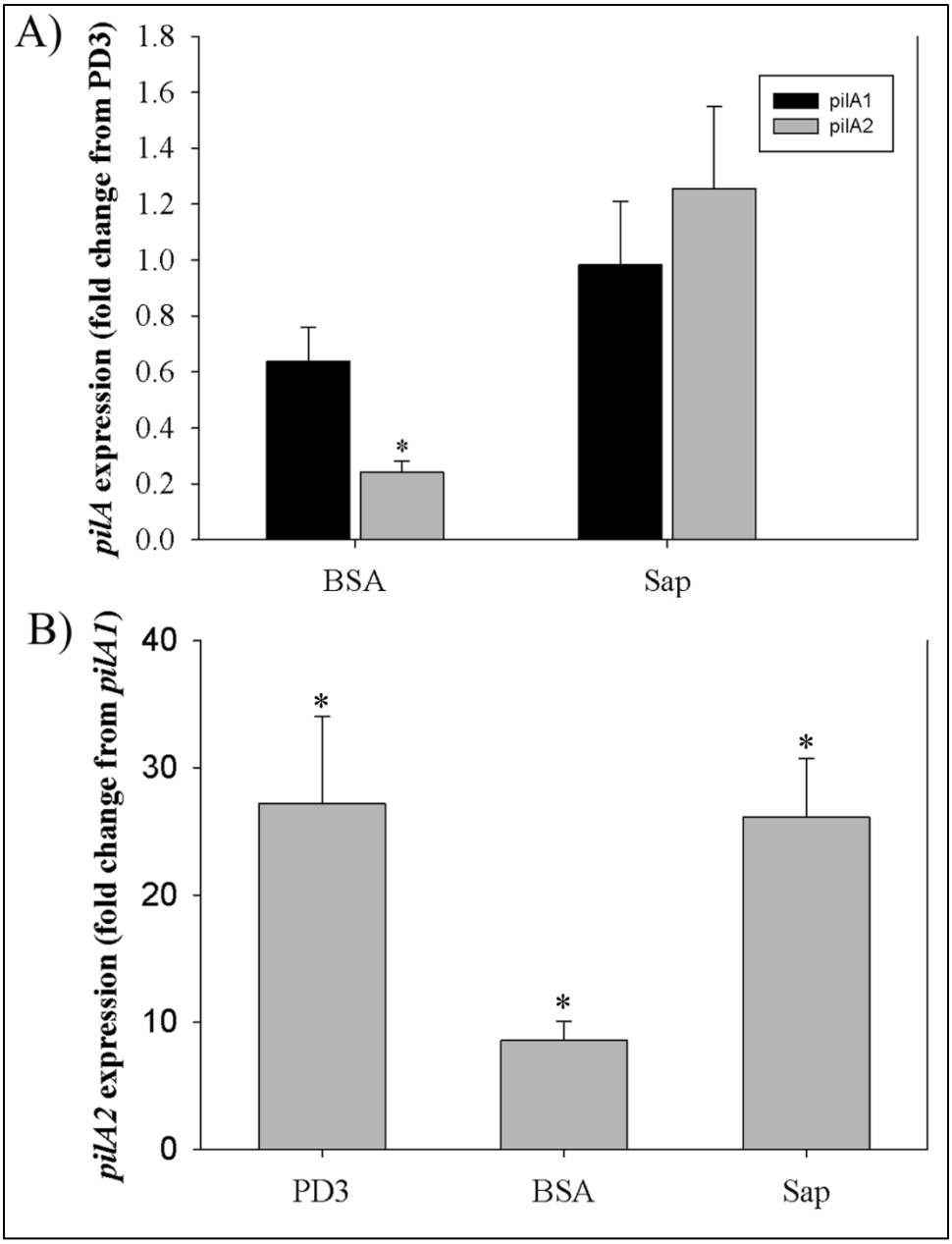


Fig. 5-6. Expression analysis of the two *pilA* paralogs of TemeculaL strain. **(A)** Results are expressed as fold change of *pilA* genes expression in cells cultured in PD3 medium supplemented with 3% of BSA (BSA) and 100% sap from grapevine variety Chardonnay (sap) relative to PD3. **(B)** Fold change in expression of *pilA2* relative to that of *pilA1* for same cultures under three different media. Gene expressions was normalized by using *nuoA* as endogenous control. Asterik indicates significant difference in fold change as compared by one-sample T-test (fold change $\neq 1$; $P < 0.01$). Error bars indicate standard errors of mean in fold change from three independent replicates of one experiment.

Discussion

Twitching motility has been described as the mechanism by which *X. fastidiosa* cells move in the xylem vessels of infected plants, even upstream from the point of inoculation where the flow of transpiration stream cannot transport the cells. Studies (Meng et al., 2005; De La Fuente et al., 2007) showed that *X. fastidiosa* cells at one cell pole possesses a long type IV pili that mediates twitching motility. Involvement of twitching motility in basipetal movement of cells in grapevines was demonstrated as the mutants lacking type IV pili were deficient in twitching motility and showed reduced up-stream migration in grapevines than the wild-types and twitching-enhanced mutants (Meng et al., 2005).

In the present study, role of one of the paralogs of type IV pilin in twitching motility was demonstrated. Despite *X. fastidiosa* genomes possessing several genes encoding pilin subunit (da Silva Neto et al., 2008), and at least two of the paralogs possessing similar structural features in the genomes of strains used in this study, the finding that only one of the *pilA* paralog (PilA2, PD1926) contributes to twitching motility was unexpected. Of all the mutants generated, *pilA2* mutants and double mutant lost twitching motility that was re-stored when a copy of *pilA2* gene was re-inserted at a different location in the genome. However, the *pilA1* mutants showed similar twitching motility to that of the wild-types, although biofilm formation was increased in these mutants in WM1-1 suggesting a different role of PilA1. Expression analysis in previous (Parker et al., 2016) and this current study showed that *pilA2* is expressed at a significantly higher level than *pilA1*. Moreover, the closest homologue of *pilA2* was demonstrated to be regulated differently by a σ^{54} factor in a subsp. *pauca* strain (da Silva Neto et al., 2008). Deletion of the *rpoN* gene encoding σ^{54} showed a pronounced inhibitory effect in the expression of this homolog

but not in the other *pilA* paralogs. Regulatory elements of σ^{54} in the *pilA2* promoter region of strains used in this study were confirmed by sequence comparison with the promoter of the subsp. *pauca* strain. Overall, loss of twitching in *pilA2* mutants, greater expression of *pilA2* than *pilA1*, and regulation of *pilA2* by a different sigma factor suggest a different route of the regulation of twitching motility in *X. fastidiosa*.

Previous studies in *P. aeruginosa* have demonstrated that mutants defective in σ^{54} protein production were non-motile and unable to produce type IV pili (Ishimoto and Lory, 1989) and flagella (Totten et al., 1990). Furthermore, expression of type IV pilin was demonstrated to be regulated by σ^{54} and PilR/PilS two-component regulatory system in a variety of other bacterial species (Parker et al., 2006; Kehl-Fie et al., 2009; Ray et al., 2015; Bretl et al., 2016). In *X. fastidiosa*, a mutant of *pilR* gene, the response regulator of the PilR/PilS system, was deficient in twitching (Li et al., 2007). Therefore, the findings from previous studies that one of the *pilA* paralog is deferentially expressed in a RpoN mutant (da Silva Neto et al., 2008), that the mutant of predicted σ^{54} activator is deficient in twitching (Li et al., 2007); and the results of the present study that the pilin paralog under σ^{54} regulation is twitching deficient, conclusively confirm that type IV pilin production in *X. fastidiosa* is regulated by σ^{54}.

In addition to its involvement in the regulation of genes involved in motility, the primary role of σ^{54} was described to be in nitrogen assimilation under nitrogen limited conditions (Zhao et al., 2009; Francke et al., 2011; Zhang and Buck, 2015). Whether nitrogen limitation co-regulates genes of nitrogen assimilation and motility responses through σ^{54}, or these two responses are regulated separately using different signal transduction mechanisms is not known. Interesting is

the fact that in previous studies (Galvani et al., 2007; Kandel et al., 2016) supplementation of a nitrogen rich compound bovine serum albumin (BSA) in the culture medium of *X. fastidiosa* significantly reduced twitching motility, and in the present study BSA was shown to decrease the expression of the major pilin *pilA2*. Although it is not known if *X. fastidiosa* cells use BSA as nitrogen source, it could act as an environmental signal that regulate the activation of the σ^{54}, which then activates genes related to nitrogen assimilation and motility.

Sigma factors are well-known to regulate expression of bacterial genes and operons by binding to the core RNA polymerase enzyme and directing it to recognize specific sequence motifs in the bacterial promoters to initiate transcription. With the first characterization of the house-keeping sigma factor σ^{70} in *Escherichia coli* (Burgess et al., 1969), at least seven other sigma factors regulating various functions have been identified and are broadly classified into two groups, referred to as σ^{70} family and σ^{54} family (Paget, 2015). The σ^{54} factors are unique in the sense that they differ in sequence from the σ^{70} class, recognize promoters with conserved GG and GC motifs located at -24 and -12 bases up-stream of the transcription start site (unlike -35 and -10 elements of σ^{70} class), and require ATP dependent activators known as enhancer binding proteins (EBPs) for their activation (Zhang and Buck, 2015). In *X. fastidiosa*, several σ^{54} binding promoters have been identified (da Silva Neto et al., 2010) and two of them (promoter of one of the *pilA* paralog and glutamine synthase gene) were experimentally verified (da Silva Neto et al., 2008; 2010). Moreover, nitrogen limitation induced these genes in σ^{54} dependent regulation system. In this study, expression of *pilA2* in grapevine sap was comparable to PD3, a medium that favors twitching motility (Kandel et al., 2016). This supports the previous finding that twitching motility of *X. fastidiosa* occurs in the plant xylem vessels (Meng et al., 2005), where

sap is the growth medium and nutrient source for *X. fastidiosa* cells. It is not known if twitching motility of *X. fastidiosa* occurs in the foregut of insect vectors, although cells are more likely to be involved in attachment and biofilm formation than in twitching motility in the insect foregut. The insect foregut system is composed mainly of chitin and embedded proteins and *X. fastidiosa* cells were able to degrade chitin and use it as a carbon source (Killiny et al., 2010; Labroussaa et al., 2017). With xylem sap being a nitrogen limited environment and insect foregut being a nitrogen rich environment due to imbedded proteins, *X. fastidiosa* cells may use σ^{54} gene regulation system to regulate nitrogen assimilation and twitching motility genes in these two distinct growth environments. However, further experiments will be required to substantiate these propositions.

Even if *X. fastidiosa* was the first plant pathogen to have its genome sequenced (Simpson et al., 2000), there is still limited understanding regarding the complete set of its virulence factors. This is because almost 40% of the proteins encoded by the genome do not have a known function and are designated as hypothetical proteins (Van Sluys et al., 2003). Also, *X. fastidiosa* has been a difficult pathogen for genetic manipulation as reported in previous studies (Monteiro et al., 2001; da Silva Neto et al., 2002). Although transposon mutagenesis (Meng et al., 2005; Roper et al., 2007) and homologous recombination protocols (Newman et al., 2004; Matsumoto et al., 2009; Cruz et al., 2014; Navarrete and De La Fuente, 2015) have been used, these protocols are time-consuming and inconsistent in efficiency. This is further complicated by the fastidious and slow-growing nature of the bacterium (Wells et al., 1987).

In this study, we developed a rapid and simplified method of genetic manipulation using overlap-extension PCR and natural transformation as *X. fastidiosa* was shown to be very efficient in natural genetic transformation (Kung and Almeida, 2011; Kandel et al., 2017). Using this method with a careful planning, gene knockouts can be generated within a month from the start of the bacterial culture from -80°C, and it does not require any plasmid vectors and steps such as restriction digestion, ligation, *E. coli* transformation, and plasmid preparation. Only few PCR steps are needed to generate an allelic exchange template that homologously recombines with the recipient bacterium, inserting a marker in exchange of the target gene. We believe that this easy, cheap, rapid, and efficient technique will augment genetic and functional studies aimed at understanding the virulence mechanisms of *X. fastidiosa* in more detail. This method was used to insert antibiotic marker genes at five different genomic locations. This included the *pilA* mutants and double mutants as well as the complemented mutant of *pilA2*. The location of genetic complementation was used as suggested and successfully utilized by a previous study (Matsumoto et al., 2009). Transformation of *X. fastidiosa* strains with linear DNA has been reported in a previous study (Kung and Almeida, 2011) and similar method of genetic transformation has been used in other bacterial systems that are capable of natural transformation (Aranda et al., 2010; Gomaa et al., 2017). In our experience three critical steps in the success of this method are, i) overlapping regions must be at least 21bp between the two fragments as regions shorter than this did not result in successful fusion of the PCR products, ii) growth stage of the cells used also caused some variation therefore, we suggest to use seven day old cultures from a second re-streak, although this may differ depending on the strain used for the study, and iii) medium used for natural transformation should be PD3 or PD2 (PW and XFM can be used without BSA supplementation). Moreover, in our experience, the concentration of PCR template

used could also influence the efficiency of transformation as gel purified PCR template (low DNA concentration) resulted in relatively lower number of transformant colonies than internal-primer amplified PCR template (high DNA concentration).

In our previous studies, we showed that twitching deficient strains or mutants lost the ability of natural transformation (Kandel et al., 2016; Kandel et al., 2017). In this study, however, we could naturally transform the twitching deficient *pilA2* mutant of TemeculaL. In a previous study (Kung and Almeida, 2014), twitching deficient mutants of *pilO* and *pilQ* were not transformable, whereas another twitching deficient mutant of *pilB* was transformable. Similarly, in our studies the twitch-minus and non-transformable strains contained mutations in *pilQ* and *pilM* genes. This may indicate that, although higher twitching motility increases the rate of natural transformation (Kandel et al., 2017), partial, rather than the complete, set of type IV pili structures may be involved in natural transformation of *X. fastidiosa*.

In summary, a rapid and efficient natural genetic transformation protocol was developed to generate gene deletion and complemented mutants in two strains of *X. fastidiosa* subsp. *fastidiosa*. This method revealed that the *pilA2* paralog of the ORF PD1926 encodes the functional pilin subunit of type IV pili that is involved in twitching motility of *X. fastidiosa*. Future studies could target this paralog to better understand the factors influencing twitching motility and virulence of *X. fastidiosa.*

References

Aranda, J., Poza, M., Pardo, B.G., Rumbo, S., Rumbo, C., Parreira, J.R., Rodríguez-Velo, P., and Bou, G. 2010. A rapid and simple method for constructing stable mutants of *Acinetobacter baumannii*. BMC Microbiol 10:279.

Bahar, O., Goffer, T., and Burdman, S. 2009. Type IV pili are required for virulence, twitching motility, and biofilm formation of *Acidovorax avenae* subsp. *citrulli*. Mol Plant Microbe Interact 22:909-920.

Bretl, D.J., Müller, S., Ladd, K.M., Atkinson, S.N., and Kirby, J.R. 2016. Type IV-pili dependent motility is co-regulated by PilSR and PilS2R2 two-component systems via distinct pathways in *Myxococcus xanthus*. Mol Microbiol 102:37-53.

Burgess, R.R., Travers, A.A., Dunn, J.J., and Bautz, E.K. 1969. Factor stimulating transcription by RNA polymerase. Nature 221:43-46.

Burrows, L.L. 2012. *Pseudomonas aeruginosa* twitching motility: type IV pili in action. Annu Rev Microbiol 66:493-520.

Caserta, R., Takita, M., Targon, M., Rosselli-Murai, L., De Souza, A., Peroni, L., Stach-Machado, D., Andrade, A., Labate, C., and Kitajima, E. 2010. Expression of *Xylella fastidiosa* fimbrial and afimbrial proteins during biofilm formation. Appl Environ Microbiol 76:4250-4259.

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

Craig, L., Pique, M.E., and Tainer, J.A. 2004. Type IV pilus structure and bacterial pathogenicity. Nat Rev Microbiol 2:363-378.

Cruz, L.F., Parker, J.K., Cobine, P.A., and De La Fuente, L. 2014. Calcium-enhanced twitching motility in *Xylella fastidiosa* is linked to a single PilY1 homolog. Appl Environ Microbiol 80:7176-7185.

Cursino, L., Galvani, C.D., Athinuwat, D., Zaini, P.A., Li, Y., De La Fuente, L., Hoch, H.C., Burr, T.J., and Mowery, P. 2011. Identification of an operon, Pil-Chp, that controls twitching motility and virulence in *Xylella fastidiosa*. Mol Plant Microbe Interact 24:1198-1206.

da Silva Neto, J.F., Koide, T., Gomes, S.L., and Marques, M.V. 2002. Site-directed gene disruption in *Xylella fastidiosa*. FEMS Microbiol Lett 210:105-110.

da Silva Neto, J.F., Koide, T., Gomes, S.L., and Marques, M.V. 2010. Global gene expression under nitrogen starvation in *Xylella fastidiosa*: contribution of the sigma54 regulon. BMC Microbiol 10:231.

da Silva Neto, J.F., Koide, T., Abe, C.M., Gomes, S.L., and Marques, M.V. 2008. Role of σ54 in the regulation of genes involved in type I and type IV pili biogenesis in *Xylella fastidiosa*. Arch Microbiol 189:249-261.

Davis, M., Purcell, A., and Thomson, S. 1980. Isolation media for the Pierce's disease bacterium. Phytopathology 70:425-429.

Davis, M.J., French, W.J., and Schaad, N.W. 1981. Axenic culture of the bacteria associated with phony disease of peach and plum leaf scald. Curr Microbiol 6:309-314.

De La Fuente, L., Burr, T.J., and Hoch, H.C. 2007. Mutations in type I and type IV pilus biosynthetic genes affect twitching motility rates in *Xylella fastidiosa*. J Bacteriol 189:7507-7510.

de Souza, A.A., Takita, M.A., Pereira, E.O., Coletta-Filho, H.D., and Machado, M.A. 2005. Expression of pathogenicity-related genes of *Xylella fastidiosa* in vitro and in planta. Curr Microbiol 50:223-228.

de Souza, A.A., Takita, M.A., Coletta-Filho, H.D., Caldana, C., Goldman, G.H., Yanai, G.M., Muto, N.H., de Oliveira, R.C., Nunes, L.R., and Machado, M.A. 2003. Analysis of gene expression in two growth states of *Xylella fastidiosa* and its relationship with pathogenicity. Mol Plant Microbe Interact 16:867-875.

Dunger, G., Guzzo, C.R., Andrade, M.O., Jones, J.B., and Farah, C.S. 2014. *Xanthomonas citri* subsp. *citri* type iv pilus is required for twitching motility, biofilm development, and adherence. Mol Plant Microbe Interact 27:1132-1147.

Francke, C., Groot Kormelink, T., Hagemeijer, Y., Overmars, L., Sluijter, V., Moezelaar, R., and Siezen, R.J. 2011. Comparative analyses imply that the enigmatic Sigma factor 54 is a central controller of the bacterial exterior. BMC Genomics 12:385.

Galvani, C.D., Li, Y., Burr, T.J., and Hoch, H.C. 2007. Twitching motility among pathogenic *Xylella fastidiosa* isolates and the influence of bovine serum albumin on twitching-dependent colony fringe morphology. FEMS Microbiol Lett 268:202-208.

Gomaa, A.E., Deng, Z., Yang, Z., Shang, L., Zhan, Y., Lu, W., Lin, M., and Yan, Y. 2017. High-frequency targeted mutagenesis in *Pseudomonas stutzeri* using a vector-free allele-exchange protocol. J Microbiol Biotechnol 27:335-341.

Guilhabert, M.R., and Kirkpatrick, B.C. 2005. Identification of *Xylella fastidiosa* antivirulence genes: hemagglutinin adhesins contribute to *X. fastidiosa* biofilm maturation and colonization and attenuate virulence. Mol Plant Microbe Interact 18:856-868.

Hao, L., Athinuwat, D., Johnson, K., Cursino, L., Burr, T., and Mowery, P. 2017. *Xylella fastidiosa* pil-chp operon is involved in regulating key structural genes of both type I and IV pili. VITIS 56:55-62.

Hopkins, D.L., and Purcell, A.H. 2002. *Xylella fastidiosa*: Cause of Pierce's Disease of Grapevine and Other Emergent Diseases. Plant Dis 86:1056-1066.

Ishimoto, K.S., and Lory, S. 1989. Formation of pilin in *Pseudomonas aeruginosa* requires the alternative sigma factor (RpoN) of RNA polymerase. Proc Natl Acad Sci USA 86:1954-1957.

Kandel, P.P., Lopez, S.M., Almeida, R.P., and De La Fuente, L. 2016. Natural competence of *Xylella fastidiosa* occurs at a high frequency inside microfluidic chambers mimicking the bacterium's natural habitats. Appl Environ Microbiol 82:5269-5277.

Kandel, P.P., Almeida, R.P.P., Cobine, P.A., and De La Fuente, L. 2017. Natural competence rates are variable among *Xylella fastidiosa* strains and homologous recombination occurs

in vitro between subspecies *fastidiosa* and *multiplex*. Mol Plant Microbe Interact 30:589-600.

Kehl-Fie, T.E., Porsch, E.A., Miller, S.E., and StGeme, J.W. 2009. Expression of *Kingella kingae* type IV pili is regulated by σ54, PilS, and PilR. J Bacteriol 191:4976-4986.

Killiny, N., and Almeida, R.P. 2009. *Xylella fastidiosa* afimbrial adhesins mediate cell transmission to plants by leafhopper vectors. Appl Environ Microbiol 75:521-528.

Killiny, N., Prado, S.S., and Almeida, R.P. 2010. Chitin utilization by the insect-transmitted bacterium *Xylella fastidiosa*. Appl Environ Microbiol 76:6134-6140.

Kung, S.H., and Almeida, R.P. 2011. Natural competence and recombination in the plant pathogen *Xylella fastidiosa*. Appl Environ Microbiol 77:5278-5284.

Kung, S.H., and Almeida, R.P. 2014. Biological and genetic factors regulating natural competence in a bacterial plant pathogen. Microbiology 160:37-46.

Labroussaa, F., Ionescu, M., Zeilinger, A.R., Lindow, S.E., and Almeida, R.P. 2017. A chitinase is required for *Xylella fastidiosa* colonization of its insect and plant hosts. Microbiology 163:502-509.

Li, Y., Hao, G., Galvani, C.D., Meng, Y., De La Fuente, L., Hoch, H., and Burr, T.J. 2007. Type I and type IV pili of *Xylella fastidiosa* affect twitching motility, biofilm formation and cell–cell aggregation. Microbiology 153:719-726.

Liu, H., Kang, Y., Genin, S., Schell, M.A., and Denny, T.P. 2001. Twitching motility of *Ralstonia solanacearum* requires a type IV pilus system. Microbiology 147:3215-3229.

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

Mattick, J.S. 2002. Type IV pili and twitching motility. Annu Rev Microbiol 56:289-314.

Meng, Y., Li, Y., Galvani, C.D., Hao, G., Turner, J.N., Burr, T.J., and Hoch, H.C. 2005. Upstream migration of *Xylella fastidiosa* via pilus-driven twitching motility. J Bacteriol 187:5560-5567.

Monteiro, P.c.B., Teixeira, D.C., Palma, R.R., Garnier, M., Bové, J.-M., and Renaudin, J. 2001. Stable transformation of the *Xylella fastidiosa* citrus variegated chlorosis strain with oriC plasmids. Appl Environ Microbiol 67:2263-2269.

Morand, P.C., Drab, M., Rajalingam, K., Nassif, X., and Meyer, T.F. 2009. *Neisseria meningitidis* differentially controls host cell motility through PilC1 and PilC2 components of type IV Pili. PLoS One 4:e6834.

Navarrete, F., and De La Fuente, L. 2015. Zinc detoxification is required for full virulence and modification of the host leaf ionome by *Xylella fastidiosa*. Mol Plant Microbe Interact 28:497-507.

Newman, K.L., Almeida, R.P., Purcell, A.H., and Lindow, S.E. 2003. Use of a green fluorescent strain for analysis of *Xylella fastidiosa* colonization of *Vitis vinifera*. Appl Environ Microbiol 69:7319-7327.

Newman, K.L., Almeida, R.P., Purcell, A.H., and Lindow, S.E. 2004. Cell-cell signaling controls *Xylella fastidiosa* interactions with both insects and plants. Proc Natl Acad Sci USA 101:1737-1742.

Oliver, J.E., Cobine, P.A., and De La Fuente, L. 2015. *Xylella fastidiosa* isolates from both subsp. *multiplex* and *fastidiosa* cause disease on southern highbush blueberry (*Vaccinium sp.*) under greenhouse conditions. Phytopathology 105:855-862.

Oliver, J.E., Sefick, S.A., Parker, J.K., Arnold, T., Cobine, P.A., and De La Fuente, L. 2014. Ionome changes in *Xylella fastidiosa*-infected *Nicotiana tabacum* correlate with virulence and discriminate between subspecies of bacterial isolates. Mol Plant Microbe Interact 27:1048-1058.

Paget, M.S. 2015. Bacterial sigma factors and anti-sigma factors: Structure, function and distribution. Biomolecules 5:1245-1265.

Parker, D., Kennan, R.M., Myers, G.S., Paulsen, I.T., Songer, J.G., and Rood, J.I. 2006. Regulation of type IV fimbrial biogenesis in *Dichelobacter nodosus*. J Bacteriol 188:4801-4811.

Parker, J.K., Chen, H., McCarty, S.E., Liu, L.Y., and De La Fuente, L. 2016. Calcium transcriptionally regulates the biofilm machinery of *Xylella fastidiosa* to promote continued biofilm development in batch cultures. Environ Microbiol 18:1620-1634.

Ray, S.K., Kumar, R., Peeters, N., Boucher, C., and Genin, S. 2015. *rpoN1*, but not *rpoN2*, is required for twitching motility, natural competence, growth on nitrate, and virulence of *Ralstonia solanacearum*. Front Microbiol 6:229.

Roper, M.C., Greve, L.C., Warren, J.G., Labavitch, J.M., and Kirkpatrick, B.C. 2007. *Xylella fastidiosa* requires polygalacturonase for colonization and pathogenicity in *Vitis vinifera* grapevines. Mol Plant Microbe Interact 20:411-419.

Schmittgen, T.D., and Livak, K.J. 2008. Analyzing real-time PCR data by the comparative CT method. Nat Protoc 3:1101-1108.

Seitz, P., and Blokesch, M. 2013. Cues and regulatory pathways involved in natural competence and transformation in pathogenic and environmental Gram-negative bacteria. FEMS Microbiol Rev 37:336-363.

Simpson, A.J.G., Reinach, F., Arruda, P., Abreu, F., Acencio, M., Alvarenga, R., Alves, L.C., Araya, J., Baia, G., and Baptista, C. 2000. The genome sequence of the plant pathogen *Xylella fastidiosa*. Nature 406:151-157.

Taguchi, F., and Ichinose, Y. 2011. Role of type IV pili in virulence of *Pseudomonas syringae* pv. tabaci 6605: correlation of motility, multidrug resistance, and HR-inducing activity on a nonhost plant. Mol Plant Microbe Interact 24:1001-1011.

Totten, P.A., Lara, J.C., and Lory, S. 1990. The *rpoN* gene product of *Pseudomonas aeruginosa* is required for expression of diverse genes, including the flagellin gene. J Bacteriol 172:389-396.

Van Sluys, M.A., de Oliveira, M.C., Monteiro-Vitorello, C.B., Miyaki, C.Y., Furlan, L.R., Camargo, L.E., da Silva, A.C., Moon, D.H., Takita, M.A., Lemos, E.G., Machado, M.A., Ferro, M.I., da Silva, F.R., Goldman, M.H., Goldman, G.H., Lemos, M.V., El-Dorry, H.,

Tsai, S.M., Carrer, H., Carraro, D.M., de Oliveira, R.C., Nunes, L.R., Siqueira, W.J., Coutinho, L.L., Kimura, E.T., Ferro, E.S., Harakava, R., Kuramae, E.E., Marino, C.L., Giglioti, E., Abreu, I.L., Alves, L.M., do Amaral, A.M., Baia, G.S., Blanco, S.R., Brito, M.S., Cannavan, F.S., Celestino, A.V., da Cunha, A.F., Fenille, R.C., Ferro, J.A., Formighieri, E.F., Kishi, L.T., Leoni, S.G., Oliveira, A.R., Rosa, V.E., Jr., Sassaki, F.T., Sena, J.A., de Souza, A.A., Truffi, D., Tsukumo, F., Yanai, G.M., Zaros, L.G., Civerolo, E.L., Simpson, A.J., Almeida, N.F., Jr., Setubal, J.C., and Kitajima, J.P. 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.

Wang, N., Li, J.-L., and Lindow, S.E. 2012. RpfF-dependent regulon of *Xylella fastidiosa*. Phytopathology 102:1045-1053.

Wells, J.M., Raju, B.C., Hung, H.Y., Weisburg, W.G., Mandelco-Paul, L., and Brenner, D.J. 1987. *Xylella fastidiosa* gen. nov., sp. nov.: Gram-negative, xylem-limited, fastidious plant bacteria related to *Xanthomonas* spp. Int J Syst Bacteriol 37:136-143.

Wolfgang, M., Park, H.-S., Hayes, S.F., Van Putten, J.P., and Koomey, M. 1998. Suppression of an absolute defect in type IV pilus biogenesis by loss-of-function mutations in pilT, a twitching motility gene in *Neisseria gonorrhoeae*. Proc Natl Acad Sci USA 95:14973-14978.

Zhang, N., and Buck, M. 2015. A perspective on the enhancer dependent bacterial RNA polymerase. Biomolecules 5:1012-1019.

Zhao, K., Liu, M., and Burgess, R.R. 2009. Promoter and regulon analysis of nitrogen assimilation factor, σ54, reveal alternative strategy for *E. coli* MG1655 flagellar biosynthesis. Nucleic Acids Res:gkp1123.