Natural competence distribution among *Xylella fastidiosa* strains and construction of a mutant library for RB-TnSeq

by

Ranlin Liu

A thesis submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Master of Science

Auburn, Alabama
December 11, 2021

Keywords: *Xylella fastidiosa*, natural competence, homologous recombination, genome-wide association study, RB-TnSeq

Copyright 2021 by Ranlin Liu

Approved by

Leonardo De La Fuente, Chair, Professor of Plant Pathology
Neha Potnis, Assistant Professor of Plant Pathology
Jeffrey J. Coleman, Assistant Professor of Plant Pathology

Abstract

The plant pathogen *Xylella fastidiosa* is a gram-negative, fastidious, xylem-limited, vector-transmitted bacterium, causing severe diseases on many economically important plants such as grapes, citrus, coffee, olive and almond among others. The outbreak of *X. fastidiosa* in new geographic regions and new hosts is a serious problem. Extensive evidence indicates that homologous recombination (HR) may be one of the major forces contributing to the appearance of new genotypes, host adaptation and host shifts in *X. fastidiosa.* Natural competence has been considered one of the most important means of HR in *X. fastidiosa*. To date, natural competence among *X. fastidiosa* has been described experimentally only for a few strains. However, little is known about the distribution of natural competence among diverse *X. fastidiosa* strains, to what extent natural competence correlates with HR detected in the core-genome in silico, and genetic mechanisms of natural competence in *X. fastidiosa*. Therefore, we characterized natural competence in 71 *X. fastidiosa* strains by measuring plasmid recombination efficiency; analyzed correlation between natural competence in vitro and HR in the genomes; and constructed a mutant library to screen genetic elements essential for natural competence using Random-Barcoded Transposon Sequencing (RB-TnSeq). Results showed that recombination rates differed extensively among strains with diverse origins specially between different subspecies. Surprisingly, results demonstrated that natural competence in vitro was slightly but significantly negatively correlated with HR in the genomes. Genome-wide association analysis showed that quorum sensing and phage-related genes may be associated with natural competence, although there was a high likelihood of false-positive associations probably due to the small and skewed sample size. On the other hand, construction of a saturated RB-TnSeq mutant library is still

ongoing based on the success of building a mini library despite its bias. Several steps of the library construction have been established, but additional troubleshooting is needed before an optimal library will be obtained for future research. Our findings suggest that natural competence may evolve differently among *X. fastidiosa* strains, and that other mechanisms in addition to natural competence may also be important for HR and host adaptation in this pathogen. Future efforts put on RB-TnSeq would help understand mechanisms of natural competence, and further clarify its role in HR and evolution in this pathogen.

Acknowledgments

I would like to express my sincere appreciation to my advisor Dr. Leonardo De La Fuente for his constant support and guidance for me throughout this project. His continuous suggestions, encouragement, patience, and enthusiasm motivated me to go further in this project and also my career and taught me the right attitudes I should have towards science and research. I would also like to sincerely thank the committee members Dr. Neha Potnis, Dr. Jeffrey Coleman for their guidance for my project. Special thanks to Dr. Neha Potnis for her constructive suggestions and helpful discussions in the bioinformatic analysis throughout the project.

Special thanks to Qing Ge for teaching me the basics in the lab and get me familiar with the all aspects in work and life during my initial time; to Marcus Merfa for teaching me important experiments for this project and helpful discussions about the experiments. I would like to thank all my lab members: Deepak Shantharaj, Eber Feliciano, Laura Gómez for helping me with experiments and also their helpful suggestions with my project and making me feel welcomed in a new environment. Finally, I would like to thank my family for always supporting me and encouraging me along the way.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

BSA	Bovine Serum Albumin
CA	California
CVC	Citrus Variegated Chlorosis
DAP	Diaminopimelic Acid
DSF	Diffusible Signaling Factor
FL	Florida
GA	Georgia
GWAS	Genome-Wide Association Study
HGT	Horizontal Gene Transfer
HR	Homologous Recombination
IHR	Inter-subspecific Homologous Recombination
IR	Inverted Repeat
Km	Kanamycin
ML	Maximum Likelihood
MLST	Multi Locus Sequence Typing
MTase	Methyl Transferases
OQDS	Olive Quick Decline Syndrome
PCR	Polymerase Chain Reaction
PD	Pierce's Disease
QQ	Quantile-Quantile
QS	Quorum Sensing

RB-TnSeq	Random-Barcoded Transposon Sequencing
REase	Restriction Endonucleases
R-M	Restriction-Modification
SNP	Single Nucleotide Polymorphism
ST	Sequence Type
TFP	Type IV pili
TIS	Transposon Insertion Sequencing
TX	Texas

Chapter 1

Literature Review

Xylella fastidiosa

The plant pathogen *Xylella fastidiosa* is a gram negative, fastidious bacterium, which inhabits the xylem vessel of plants and is only transmitted by insect vectors (Hopkins and Purcell, 2002; Chatterjee et al., 2008). The bacterium is rod-shaped, aerobic, and slow-growing with an optimum growth at 26-28 °C and pH of 6.5 - 6.9 (Wells et al., 1987). *X. fastidiosa* possesses two types of pili anchored at one of the cell pores but is non-flagellated (Meng et al., 2005). One type of the pili, type I pili, is shorter and is involved in biofilm formation and auto-aggregation of the cells, while the other pili, type IV pili (TFP), is longer and is involved in twitching motility, a type of movement on the surface, which helps *X. fastidiosa* move against the fluid current of xylem sap (Meng et al., 2005; De La Fuente et al., 2007; Li et al., 2007). *X. fastidiosa* has a broad host range, currently 638 plant species have been described as infected by this pathogen, although only a few plant species were symptomatic and susceptible to specific phylogenetic clades of *X. fastidiosa* (Almeida and Nunney, 2015; European Food Safety et al., 2021). However, among the symptomatic plants are many economically important crops such as grapes (Pierce's Disease, PD), citrus (Citrus Variegated Chlorosis, CVC), coffee (leaf scorch diseases), olive (Olive Quick Decline Syndrome, OQDS), and almond (leaf scorch diseases).

Virulence mechanisms of *X. fastidiosa*

The exact disease mechanism of *X. fastidiosa* infection is still unclear, yet, it has been found that attachment of cells and biofilm formation are important for disease development (Chatterjee et al.,

2008). Typical disease symptom caused by *X. fastidiosa* is leaf scorch associated with deficiency of water and disturbance of nutrients such as mineral in the xylem vessels where *X. fastidiosa* exclusively inhabits (Chatterjee et al., 2008). Such water deficiency is caused by a clog of xylem vessels by biofilm formation of *X. fastidiosa* cells, as well as gum and tyloses formed by plant cells (Sun et al., 2013). Once *X. fastidiosa* enters the xylem vessels at the point of infection, cells can move against the fluid current facilitated by the twitching motility, or move with the flow, and move to proximal vessels facilitated by degrading pit membranes, leading to systemic colonization (Chatterjee et al., 2008).

Insect vectors of *X. fastidiosa*

Infection of *X. fastidiosa* between plants exclusively relies on sap-feeding insect vectors mainly two groups, sharpshooter leafhoppers (Hemiptera, Cicadellidae) and spittlebugs (Hemiptera, Cercopidae). *X. fastidiosa* can be transmitted to another individual plant immediately once the insect acquires the cells when feeding on plant xylem due to no requirement of a latent period for infection (Chatterjee et al., 2008). Cells can also colonize and form biofilm in the foregut of the insects (Chatterjee et al., 2008). Remarkably, there is a lack of pathogen-vector specificity, and thus the bacterium cells are likely to be transmitted by any species from the two groups of insects and likewise the vectors can transmit any genotypes of *X. fastidiosa* (Almeida and Nunney, 2015). The insect vectors also feed on a broad range of plant species. However, sharpshooters prefer healthy plants while discriminating against infected plants (Daugherty et al., 2011). The transmission can be facilitated by cell-cell signaling which instead constrains virulence in plants in a population-dependent manner (Newman et al., 2004).

Taxonomy of *X. fastidiosa*

X. fastidiosa is in the family of Xanthomonadaceae, most closely related with another species *Xylella taiwanensis* and the genus Xanthomonas (Wells et al., 1987; Su et al., 2016). *X. taiwanensis*, causing leaf scorch on pears, was found in Taiwan and initially classified as a distant relative of *X. fastidiosa* (Su et al., 2014; Su et al., 2016). *X. fastidiosa* has been classified into three subspecies by DNA-DNA hybridization (Schaad et al., 2004), Multi Locus Sequence Typing (MLST) (Scally et al., 2005) and whole-genome analysis (Vanhove et al., 2019): subsp. *fastidiosa*, subsp. *multiplex*, and subsp. *pauca*, all of which are distinctly defined. The MLST analysis (Schuenzel et al., 2005) also suggested a fourth subspecies, subsp. *sandyi*, whose classification was also confirmed by whole genome phylogeny (Vanhove et al., 2019). Recently a fifth subspecies was proposed as subsp. *morus* to group strains infecting mulberry, although its classification is still in debate and subsp. *morus* was considered as a result of intersubspecific recombination between subsp. *fastidiosa* and subsp. *multiplex* (Nunney et al., 2014b; Potnis et al., 2019). Both subsp. *sandyi* and *morus* were also hypothesized to belong to subsp. *fastidosa* based on other phylogenetic analysis (Marcelletti and Scortichini, 2016). Genotypes or Sequence Type (ST) defined by MLST of *X. fastidiosa* strains to some extent indicate host specificity in terms of few hosts overlap between certain genotypes (Nunney et al., 2013; Coletta-Filho et al., 2017). It has been hypothesized that ST may be used for predicting potential host range of *X. fastidiosa* strains occurred in a new region (Sicard et al., 2018). However, this was challenged recently by whole-genome analysis of subsp. *multiplex* strains and experimental tests of host range of strains with different STs (Nunney et al., 2019; Landa et al., 2020).

Geographic distribution and dispersal

X. fastidiosa has allopatric origins limited to the Americas (Almeida and Nunney, 2015). Specifically, *X. fastidiosa* subsp. *fastidiosa* is native to Central America (Nunney et al., 2010; Castillo et al., 2020), whereas subsp. *pauca* to South America (Nunney et al., 2012) and subsp. *multiplex* to North America (Nunney et al., 2014a). However, they have now been dispersed to many other regions beyond the Americas especially in Europe (Saponari et al., 2018; Castillo et al., 2021b; Olmo et al., 2021) and Asia (Castillo et al., 2021b). Currently, all subspecies except for subsp. *morus* have been found in Europe as a result of multiple introductions from the Americas (Cella et al., 2018; Castillo et al., 2020; Landa et al., 2020; Castillo et al., 2021a). Recently, subsp. *fastidiosa* was found in Taiwan causing PD identified back in 2002, which has been considered as a result of introduction from the southeastern US (Su et al., 2013; Castillo et al., 2021b). In addition to Europe and Asia, *X. fastidiosa* has also been reported in Turkey (Çağlar et al., 2005), Iran (Amanifar et al., 2014) and Israel (Ofir et al., 2019), although this needs to be further confirmed.

Homologous recombination in *X. fastidiosa*

Homologous recombination (HR) is the genetic exchange including replacement and inversion of two homologous fragments of DNA. As the final process of Horizontal Gene Transfer (HGT) in bacteria, HR occurs through three mechanisms: transformation, conjugation, and transduction, all of which may end up integrating new genetic material into bacterial chromosomes (Sheppard et al., 2018). All three mechanisms have been described in *X. fastidiosa.* Transformation, also termed as natural transformation or natural competence, is an ability of cells to naturally uptake extracellular DNA and incorporate the DNA into the chromosome, which was observed in *X. fastidiosa* in vitro (Kung and Almeida, 2011). Conjugation is when the plasmid DNA is transferred

via direct contact of donor and recipient cells, which is also utilized as an approach of genetic manipulation for many bacteria (Schröder and Lanka, 2005). Transfer of mobile plasmids has been found between different subspecies of *X. fastidiosa* strains in vitro via conjugation dependent on the *tra* and *trb* operon that encode a conjugative Type IV secretion system, while the plasmid transfer was not observed between conjugative *Escherichia coli* donor and *X. fastidiosa* recipient cells (Burbank and Van Horn, 2017). Transduction is the process where a bacteriophage transfers DNA between the two bacteria during infections. Many genes were found to be phage-related in *X. fastidiosa* genomes and these genes seems to be associated with genome rearrangement, supporting the occurrence of HGT via transduction (Simpson et al., 2000; Van Sluys et al., 2003). Indeed, extensive HR, specifically inter-subspecific HR (IHR), has been detected throughout the genomes of *X. fastidiosa* strains (Potnis et al., 2019; Vanhove et al., 2019). In addition to recombination with closely related strains, HR in *X. fastidiosa* has been detected from distantly related bacterial species (Firrao et al., 2021). HR has been considered a primary force for generating genetic diversity in *X. fastidiosa* based on genomic data and hypothesized to contribute to the appearance of new genotypes, expansion of geographic regions, change of virulence and fitness, and host adaptation (Kung et al., 2013; Potnis et al., 2019; Vanhove et al., 2019). It has been demonstrated by MLST studies that host shifts in *X. fastidiosa* were closely associated with HR (Nunney et al., 2012; Nunney et al., 2014a; Nunney et al., 2014b). However, frequency of HR varied dramatically among different populations of *X. fastidiosa*. It was hypothesized that HR among *X. fastidiosa* strains was more extensive in native populations compared to introduced populations, since introduced populations were more likely to donate new alleles to native populations (Castillo et al., 2020; Landa et al., 2020; Castillo et al., 2021a). It has been found that many hotspots of recombined regions were involved in fundamental traits for *X. fastidiosa*

ecology, suggesting the importance of HR in this pathogen adaptation and evolution (Potnis et al., 2019).

Genetic mechanisms of natural competence

One of the mechanisms of HR is natural competence, which is widespread in bacteria (Johnston et al., 2014; Blokesch, 2016). In addition to being important in contributing to the genetic diversity in evolution (Griffith, 1928; Baltrus et al., 2008), the extracellular DNA obtained through natural competence can be used as a nutrient source under starvation conditions (Herriott et al., 1970; Lorenz and Wackernagel, 1991) and as sources for DNA repair (Dorer et al., 2010; Charpentier et al., 2011), although both of them may also be involved in the evolutionary process to some extent. TFP, which is also involved in adhesion, twitching motility and protein secretion, has been considered as DNA uptake machinery in natural competence in many Gram-negative naturally competent bacteria (Stone Barbara and Kwaik Yousef, 1999; Harding Christian et al., 2013; Ellison et al., 2018; Craig et al., 2019). *Vibrio cholerae* is a model organism for investigating DNA uptake by TFP in natural competence. In *V. cholerae*, successful natural competence starts by the uptake of foreign DNA via TFP. After retracting DNA into the periplasm, a DNA binding protein ComEA binds the DNA and delivers it into ComEC on the inner membrane where the DNA enters the cytoplasm. Finally, the DNA exists in the cytoplasm in a single-strand form would be combined by the DNA-binding proteins Ssb and DprA to prevent them from degradation and to be loaded on RecA, which recombines the DNA with the chromosome based on HR (Seitz and Blokesch, 2013). There are also many other factors mediating natural competence processes, which may vary in different species. For instance, mobile genetic elements have been found to inhibit natural competence in *Legionella* species, which may counteract defense of natural competence on selfish

mobile genes (Croucher et al., 2016; Durieux et al., 2019). Type VI secretion systems, a nanomachine injecting proteins into neighboring cells for survival and fitness, was found to be driven by competence regulators to kill prey cells for obtaining free DNA for natural competence in *V. cholerae* (Borgeaud et al., 2015).

Natural competence in *X. fastidiosa*

Natural competence in *X. fastidiosa* was first described by Kung and Almeida in 2011 (Kung and Almeida, 2011). *X. fastidiosa* showed natural competence under many experimental conditions such as batch culture, solid media, and microfluidic chambers, a system mimicking plant xylem vessels where *X. fastidiosa* inhabits, although occurrence of natural competence in natural environments including plant xylem vessels and the foregut of insect vectors remains to be confirmed (Kung and Almeida, 2011; Kandel et al., 2016). *X. fastidiosa* can uptake and recombine with different forms of donor DNA such as plasmid DNA, linear PCR-amplified fragments, and whole genomes released from dead cells of *X. fastidiosa* strains from the same or different subspecies (Kung and Almeida, 2011; Kandel et al., 2016). Its ability to recombine with other *X. fastidiosa* strains suggests potential of inter- and intra-subspecific HR, which was indeed observed in *X. fastidiosa* genomes (Kandel et al., 2017; Potnis et al., 2019; Castillo et al., 2021a). *X. fastidiosa* showed high recombination rates during exponential growth (Kung and Almeida, 2011; Kung et al., 2013; Kung and Almeida, 2014), and on solid media more than liquid media, preferably PD3 plates (Kung and Almeida, 2014; Kandel et al., 2016). At least 96 bp of homologous regions and no more than 6 kb of non-homologous insertion in the donor DNA was required for detecting natural competence in vitro. (Kung et al., 2013). Restriction-Modification (R-M) systems were also found to impact natural competence in *X. fastidiosa*. Recombination rates

with methylated donor DNA by using methylase of *X. fastidiosa* were significantly increased (Kung and Almeida, 2011). Like other Gram-negative naturally competent bacteria, TFP is involved in natural competence in *X. fastidiosa* (Kung and Almeida, 2014; Kandel et al., 2017; Kandel et al., 2018). A recent study exploring the association of TFP with natural competence in *X. fastidiosa* showed that TFP served as DNA uptake machinery and thus was crucial for natural competence in *X. fastidiosa* (Merfa et al., 2021). Since TFP is also involved in twitching motility, a positive correlation between twitching motility and natural competence was reported (Kandel et al., 2017; Merfa et al., 2021). While some genetic factors were demonstrated to affect the natural competence in *X. fastidiosa*, there may exist other genetic elements involving natural competence especially those involved in regulation and recombination.

Random-barcoded transposon sequencing

To study gene functions related to certain phenotypes or physiological traits, randomly inserted mutagenesis has been used with *X. fastidiosa* through delivering transposon plasmids, which randomly insert DNA fragments into the chromosome via transposases, delivered into cells via electroporation, generating a large number of mutants selected under certain conditions related to phenotypes of interest (Guilhabert and Kirkpatrick, 2005; Li et al., 2007). Although this approach succeeded in identifying various functions of different genes, it is still challenging to identify genetic factors associated with certain phenotypes on a genome-wide scale and to assess the phenotype with the whole mutant library (Newman et al., 2004; Guilhabert and Kirkpatrick, 2005; Li et al., 2007; Cruz et al., 2014; Kandel et al., 2018). Even with randomly inserted mutagenesis, generating a full mutant library containing mutants of almost every gene in *X. fastidiosa* has not been documented. With technical advancement, high-throughput sequencing allows us to screen a

complete mutant library at once, substantially increasing the efficiency of assessing phenotypes. One of the methods to rapidly find genes of interest is termed Transposon Insertion Sequencing (TIS), which is an approach identifying functional essential genes from pooled mutant libraries given certain conditions. Combining advantages of transposon mutagenesis with high-throughput sequencing, TIS is able to profile fitness reflected by relative abundance of mutants in competitive assays of genetic elements on a genome-wide level (Cain et al., 2020). One TIS method is termed Random-Barcoded Transposon Sequencing (RB-TnSeq), which was developed by incorporating barcodes into transposon sequencing. Adding barcodes in each mutant substantially reduces the laborious work of preparing a sequencing library because once the mutants are mapped, only DNA barcodes need to be sequenced following the fitness assays as each barcode represents a specific mutant, which is easier than re-sequencing the whole library. Therefore, RB-TnSeq facilitates multiple experiments applied for identifying essential genes responsible for different phenotypes. In fitness assays, the relative abundance of DNA barcodes in the library would be compared at different time points of competitive assays, which is used to indicate fitness in population (Wetmore et al., 2015).

Research goals

Understand the impact of natural competence on *X. fastidiosa* populations and expand the knowledge on its genetic mechanisms.

Specific objectives

1) Define relationships between natural competence ability in vitro of *X. fastidiosa* strains and their origin and classification, as well as the homologous recombination in their genome.

2) Identify novel genes involved in natural competence in *X. fastidiosa*

Hypotheses

1) Natural competence among *X. fastidiosa* strains is variable, correlates with geographical distribution, host range and homologous recombination in the genome.

2) *X. fastidiosa* possess novel molecular mechanisms that are involved in natural competence.

References

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

Amanifar, N., Taghavi, M., Izadpanah, K., and Babaei, G. 2014. Isolation and pathogenicity of *Xylella fastidiosa* from grapevine and almond in Iran. Phytopathol. Mediterr. 53:318-327.

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

Blokesch, M. 2016. Natural competence for transformation. Curr. Biol. 26:R1126-R1130.

Borgeaud, S., Metzger Lisa, C., Scrignari, T., and Blokesch, M. 2015. The type VI secretion system of *Vibrio cholerae* fosters horizontal gene transfer. Science 347:63-67.

Burbank, L.P., and Van Horn, C.R. 2017. Conjugative plasmid transfer in *Xylella fastidiosa* is dependent on *tra* and *trb* operon functions. J Bacteriol 199:e00388-00317.

Çağlar, B., Castellano, M., Ünlü, L., Güran, S., Yılmaz, M., and Martelli, G. 2005. First report of almond leaf scorch in Turkey. Plant Pathol. J. 87.

Cain, A.K., Barquist, L., Goodman, A.L., Paulsen, I.T., Parkhill, J., and van Opijnen, T. 2020. A decade of advances in transposon-insertion sequencing. Nat. Rev. Genet. 21:526-540.

Castillo, A.I., Chacon-Diaz, C., Rodriguez-Murillo, N., Coletta-Filho, H.D., and Almeida, R.P.P. 2020. Impacts of local population history and ecology on the evolution of a globally dispersed pathogen. BMC Genom. 21:369.

Castillo, A.I., Bojanini, I., Chen, H., Kandel, P.P., De La Fuente, L., and Almeida, R.P.P. 2021a. Allopatric plant pathogen population divergence following disease emergence. Appl. Environ. Microbiol. 87:e02095-02020.

Castillo, A.I., Tsai, C.-W., Su, C.-C., Weng, L.-W., Lin, Y.-C., Cho, S.-T., Almeida, R.P.P., and Kuo, C.-H. 2021b. Genetic differentiation of *Xylella fastidiosa* following the introduction into Taiwan. bioRxiv.

Cella, E., Angeletti, S., Fogolari, M., Bazzardi, R., De Gara, L., and Ciccozzi, M. 2018. Two different *Xylella fastidiosa* strains circulating in Italy: phylogenetic and evolutionary analyses. J. Plant Interact. 13:428-432.

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

Chatterjee, S., Almeida, R.P.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.D., Francisco, C.S., Lopes, J.R.S., Muller, C., and Almeida, R.P.P. 2017. Homologous recombination and *Xylella fastidiosa* host–pathogen associations in south America. Phytopathology 107:305-312.

Craig, L., Forest, K.T., and Maier, B. 2019. Type IV pili: dynamics, biophysics and functional consequences. Nat. Rev. Microbiol. 17:429-440.

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

Daugherty, M.P., Rashed, A., Almeida, R.P.P., and Perring, T.M. 2011. Vector preference for hosts differing in infection status: sharpshooter movement and *Xylella fastidiosa* transmission. Ecol. Entomol. 36:654-662.

De La Fuente, L., Montanes, E., Meng, Y., Li, Y., Burr Thomas, 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.

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

Durieux, I., Ginevra, C., Attaiech, L., Picq, K., Juan, P.-A., Jarraud, S., and Charpentier, X. 2019. Diverse conjugative elements silence natural transformation in *Legionella* species. Proc. Natl. Acad. Sci. U.S.A. 116:18613-18618.

Ellison, C.K., Dalia, T.N., Ceballos, A.V., Wang, J.C.-Y., Biais, N., Brun, Y.V., and Dalia, A.B. 2018. Retraction of DNA-bound type IV competence pili initiates DNA uptake during natural transformation in *Vibrio cholerae*. Nat. Microbiol. 3:773-780.

European Food Safety, A., Delbianco, A., Gibin, D., Pasinato, L., and Morelli, M. 2021. Update of the *Xylella* spp. host plant database - systematic literature search up to 31 December 2020. EFSA J 19:e06674.

Firrao, G., Scortichini, M., and Pagliari, L. 2021. Orthology-based estimate of the contribution of horizontal gene transfer from distantly related bacteria to the intraspecific diversity and differentiation of *Xylella fastidiosa*. Pathogens 10:46.

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

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. MPMI 18:856-868.

Harding Christian, M., Tracy Erin, N., Carruthers Michael, D., Rather Philip, N., Actis Luis, A., Munson Robert, S., and Taylor, R. 2013. *Acinetobacter baumannii* strain M2 produces type IV pili which play a role in natural transformation and twitching motility but not surface-associated motility. mBio 4:e00360-00313.

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.

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.

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.

Kandel, P.P., Chen, H., and De La Fuente, L. 2018. A short protocol for gene knockout and complementation in *Xylella fastidiosa* shows that one of the type IV pilin paralogs (PD1926) is needed for twitching while another (PD1924) affects pilus number and location. Appl. Environ. Microbiol. 84:e01167-01118.

Kandel, P.P., Lopez, S.M., Almeida, R.P.P., and Fuente, L.D.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*. MPMI 30:589-600.

Kung, S.H., and Almeida, R.P.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.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.P. 2013. Effects of DNA size on transformation and recombination efficiencies in *Xylella fastidiosa*. Appl. Environ. Microbiol. 79:1712-1717.
Landa, B.B., Castillo, A.I., Giampetruzzi, A., Kahn, A., Román-Écija, M., Velasco-Amo, M.P., Navas-Cortés, J.A., Marco-Noales, E., Barbé, S., Moralejo, E., Coletta-Filho, H.D., Saldarelli, P., Saponari, M., and Almeida, R.P.P. 2020. Emergence of a plant pathogen in Europe associated with multiple intercontinental introductions. Appl. Environ. Microbiol. 86:e01521-01519.
Li, Y., Hao, G., Galvani, C.D., Meng, Y., Fuente, L.D.L., Hoch, H.C., 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.
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. Microbiol. 57:1246-1251.
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.
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.
Merfa, M.V., Zhu, X., Shantharaj, D., Gomez, L.M., Naranjo, E., Potnis, N., Cobineb, P.A., and Fuente, L.D.L. (2021). Movement and evolution: complete functional analysis of type IV pilus of a re-emergent plant pathogen reveals a unique DNA receptor. In unpublished.
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. U.S.A. 101:1737-1742.
Nunney, L., Azad, H., and Stouthamer, R. 2019. An experimental test of the host-plant range of nonrecombinant strains of north American *Xylella fastidiosa* subsp. *multiplex*. Phytopathology 109:294-300.
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., 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-Astúa, 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.
Ofir, B., Orit, D., Lera, V., and Lior, B. (2019). A survey in Israel reveals the presence of *Xylella fastidiosa* in almond trees in the northern part of the country. In Proceedings of the 2nd European Conference on *Xylella Fastidiosa* 2019 (Zenodo).
Olmo, D., Nieto, A., Borràs, D., Montesinos, M., Adrover, F., Pascual, A., Gost, P.A., Quetglas, B., Urbano, A., García, J.d.D., Velasco-Amo, M.P., Olivares-García, C., Beidas, O., Juan, A., Marco-Noales, E., Gomila, M., Rita, J., Moralejo, E., and Landa, B.B. 2021. Landscape epidemiology of *Xylella fastidiosa* in the balearic islands. Agronomy 11:473.
Potnis, N., Kandel, P.P., Merfa, M.V., Retchless, A.C., Parker, J.K., Stenger, D.C., Almeida, R.P.P., Bergsma-Vlami, M., Westenberg, M., Cobine, P.A., and Fuente, L.D.L. 2019. Patterns of inter- and intrasubspecific homologous recombination inform eco-evolutionary dynamics of *Xylella fastidiosa*. ISME J 13:2319-2333.
Saponari, M., Giampetruzzi, A., Loconsole, G., Boscia, D., and Saldarelli, P. 2018. *Xylella fastidiosa* in olive in Apulia: where we stand. Phytopathology 109:175-186.
Scally, M., Schuenzel Erin, 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.B., and Chang, C.-J. 2004. *Xylella fastidiosa* subspecies: *X. fastidiosa* subsp *piercei*, subsp. nov., *X. fastidiosa* subsp. *multiplex* subsp. nov., and *X. fastidiosa* subsp. *pauca* subsp. nov. Syst. Appl. Microbiol. 27:290-300.
Schröder, G., and Lanka, E. 2005. The mating pair formation system of conjugative plasmids—a versatile secretion machinery for transfer of proteins and DNA. Plasmid 54:1-25.
Schuenzel, E.L., Scally, M., Stouthamer, R., and Nunney, L. 2005. A multigene phylogenetic study of clonal diversity and divergence in north American strains of the plant pathogen *Xylella fastidiosa*. Appl. Environ. Microbiol. 71:3832-3839.
Seitz, P., and Blokesch, M. 2013. DNA-uptake machinery of naturally competent *Vibrio cholerae*. Proc. Natl. Acad. Sci. U.S.A. 110:17987-17992.
Sheppard, S.K., Guttman, D.S., and Fitzgerald, J.R. 2018. Population genomics of bacterial host adaptation. Nat. Rev. Genet. 19:549-565.
Sicard, A., Zeilinger, A.R., Vanhove, M., Schartel, T.E., Beal, D.J., Daugherty, M.P., and Almeida, R.P.P. 2018. *Xylella fastidiosa*: insights into an emerging plant pathogen. Annu. Rev. Phytopathol. 56:181-202.
Simpson, A.J.G., Reinach, F.C., Arruda, P., Abreu, F.A., Acencio, M., Alvarenga, R., Alves, L.M.C., Araya, J.E., Baia, G.S., Baptista, C.S., Barros, M.H., Bonaccorsi, E.D., Bordin, S., Bové, J.M., Briones, M.R.S., Bueno, M.R.P., Camargo, A.A., Camargo, L.E.A., Carraro, D.M., Carrer, H., Colauto, N.B., Colombo, C., Costa, F.F., Costa, M.C.R., Costa-Neto, C.M., Coutinho, L.L., Cristofani, M., Dias-Neto, E., Docena, C., El-Dorry, H., Facincani, A.P., Ferreira, A.J.S., Ferreira, V.C.A., Ferro, J.A., Fraga, J.S., França, S.C., Franco, M.C., Frohme, M., Furlan, L.R., Garnier, M., Goldman, G.H., Goldman, M.H.S., Gomes, S.L., Gruber, A., Ho, P.L., Hoheisel, J.D., Junqueira, M.L., Kemper, E.L., Kitajima, J.P., Krieger, J.E., Kuramae, E.E., Laigret, F., Lambais, M.R., Leite,

L.C.C., Lemos, E.G.M., Lemos, M.V.F., Lopes, S.A., Lopes, C.R., Machado, J.A., Machado, M.A., Madeira, A.M.B.N., Madeira, H.M.F., Marino, C.L., Marques, M.V., Martins, E.a.L., Martins, E.M.F., Matsukuma, A.Y., Menck, C.F.M., Miracca, E.C., Miyaki, C.Y., Monteiro-Vitorello, C.B., Moon, D.H., Nagai, M.A., Nascimento, A.L.T.O., Netto, L.E.S., Nhani, A., Nobrega, F.G., Nunes, L.R., Oliveira, M.A., de Oliveira, M.C., de Oliveira, R.C., Palmieri, D.A., Paris, A., Peixoto, B.R., Pereira, G.a.G., Pereira, H.A., Pesquero, J.B., Quaggio, R.B., Roberto, P.G., Rodrigues, V., de M. Rosa, A.J., de Rosa, V.E., de Sá, R.G., Santelli, R.V., Sawasaki, H.E., da Silva, A.C.R., da Silva, A.M., da Silva, F.R., Silva, W.A., da Silveira, J.F., Silvestri, M.L.Z., Siqueira, W.J., de Souza, A.A., de Souza, A.P., Terenzi, M.F., Truffi, D., Tsai, S.M., Tsuhako, M.H., Vallada, H., Van Sluys, M.A., Verjovski-Almeida, S., Vettore, A.L., Zago, M.A., Zatz, M., Meidanis, J., and Setubal, J.C. 2000. The genome sequence of the plant pathogen *Xylella fastidiosa*. Nature 406:151-157.

Stone Barbara, J., and Kwaik Yousef, A. 1999. Natural competence for DNA transformation by *Legionella pneumophila* and its association with expression of type IV pili. J Bacteriol 181:1395-1402.

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. 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*. Phytopathology 161:389-396.

Su, C.C., Deng, W.L., Jan, F.J., Chang, C.J., Huang, H., and Chen, J. 2014. Draft genome sequence of *Xylella fastidiosa* pear leaf scorch strain in Taiwan. Genome Announc. 2:e00166-00114.

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.A., da Silva, A.C.R., Moon, D.H., Takita, M.A., Lemos, E.G.M., Machado, M.A., Ferro, M.I.T., da Silva, F.R., Goldman, M.H.S., Goldman, G.H., Lemos, M.V.F., 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.C., 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., Sassaki, F.T., Sena, J.A.D., de Souza, A.A., Truffi, D., Tsukumo, F., Yanai, G.M., Zaros, L.G., Civerolo, E.L., Simpson, A.J.G., Almeida, N.F., 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.

Vanhove, M., Retchless, A.C., Sicard, A., Rieux, A., Coletta-Filho, H.D., De La Fuente, L., Stenger, D.C., and Almeida, R.P.P. 2019. Genomic diversity and recombination among *Xylella fastidiosa* subspecies. Appl. Environ. Microbiol. 85:e02972-02918.

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. Evol. 37:136-143.

Wetmore, K.M., Price, M.N., Waters, R.J., Lamson, J.S., He, J., Hoover, C.A., Blow, M.J., Bristow, J., Butland, G., Arkin, A.P., and Deutschbauer, A. 2015. Rapid quantification of mutant fitness in diverse bacteria by sequencing randomly bar-coded transposons. mBio 6:e00306-00315.

Chapter 2

Natural competence rates differ among *Xylella fastidiosa* strains and are negatively correlated with recombination

Abstract

Horizontal gene transfer (HGT) contributes to genetic diversity and can influence adaptation of plant pathogens to the environment. Natural competence is one of the mechanisms of HGT that relies on uptake of extracellular DNA and recombination at homologous regions. To date, natural competence among *X. fastidiosa* has been described experimentally only in a few strains. The presence of natural competence in diverse *X. fastidiosa* strains and how it correlates with recombination detected on core-genome in silico remains unknown. Here we characterized natural competence in 71 *X. fastidiosa* strains from all subspecies - mainly subsp. *fastidiosa* (41 strains) and *multiplex* (26 strains) - by measuring plasmid pKLN61 recombination efficiency. Recombination rates above the detection limit were found in 45 out of 71 tested strains. Results showed that the majority of tested subsp. *fastidiosa* strains (96%) were naturally competent, while only 23% of tested subsp. *multiplex* strains showed natural competence in vitro. Interestingly, none of the six European strains we tested were naturally competent. Additionally, we analyzed homologous recombination (HR) in the core genome of 153 *X. fastidiosa* strains and correlated this data with natural competence in vitro. Surprisingly, results showed that natural competence in vitro was slightly but significantly negatively correlated with frequency of recombination events in the genomes. Genome-wide association analysis was performed on 48 strains for which phenotypic data in vitro and full genome sequences were available, to find genes associated with

natural competence. Of particular interest were two hypothetical proteins annotated as phage-related genes identified by a k-mer-based method, and two single nucleotide polymorphisms (SNPs) in another hypothetical protein and upstream of the gene *rpfG* identified by a SNP-based method. Unfortunately, probably due to the small and skewed sample size, there was a high likelihood of false-positive associations. Further investigation of GWAS using a larger dataset are needed to explore and confirm candidate genes associated with natural competence, which may provide an important insight on the genetic mechanisms of natural competence. Our study demonstrates variability of natural competence among *X. fastidiosa* strains and a negative correlation between natural competence and HR, suggesting that other mechanisms in addition to natural competence may also be important for HR in this pathogen.

Introduction

The plant pathogen *Xylella fastidiosa* is a gram negative bacterium that inhabits the xylem vessels of plants, and is only transmitted by insect vectors (Hopkins and Purcell, 2002; Chatterjee et al., 2008a). *X. fastidiosa* has a broad host range, currently 638 plant species have been described as infected by this pathogen, among which are many economically important crops including grapes (Pierce's disease), citrus (citrus variegated chlorosis), coffee (leaf scorch diseases), olive (olive quick decline syndrome) and almond (leaf scorch diseases) (European Food Safety et al., 2021). *X. fastidiosa* has been classified into five subspecies, three of which are very distinctly defined: subsp. *fastidiosa*, subsp. *multiplex*, and subsp. *pauca*. The other two subspecies, subsp. *sandyi* and subsp. *morus*, have been proposed to belong to subsp. *fastidiosa* based on phylogenetic analysis (Marcelletti and Scortichini, 2016; Denancé et al., 2019). *X. fastidiosa* subsp. *morus* is postulated to be the result of intersubspecific recombination between subsp. *fastidiosa* and subsp. *multiplex*

and its classification is still debatable (Nunney et al., 2014b; Jacques et al., 2016; Vanhove et al., 2019). The three main subspecies have allopatric origins limited to the Americas (Almeida and Nunney, 2015). Specifically, *X. fastidiosa* subsp. *fastidiosa* is native to Central America (Nunney et al., 2010; Castillo et al., 2020), whereas subsp. *pauca* to South America (Nunney et al., 2012) and subsp. *multiplex* to North America (Nunney et al., 2014a). However, they have now been dispersed to many other regions beyond the Americas especially Europe (Saponari et al., 2018; Olmo et al., 2021) and Asia (Castillo et al., 2021b). Currently, all subspecies except for subsp. *morus* have been found in Europe as a result of multiple introductions from the Americas (Cella et al., 2018; Castillo et al., 2020; Landa et al., 2020; Castillo et al., 2021a).

Despite its broad host range, host specificity of *X. fastidiosa* was shown to be somewhat correlated with the subspecies classification, since only a few hosts overlap among subspecies, but few hosts have been found to be infected by more than one subspecies (Nunney et al., 2013; Oliver et al., 2015; Coletta-Filho et al., 2017; Nunney et al., 2019). It is noticeable that the emergence of diseases caused by *X. fastidiosa* on new host plants, especially those that have great economic impact on agriculture, has become a big threat for agricultural systems. Such host shifts have been suggested to be associated with homologous recombination (HR) (Nunney et al., 2012; Nunney et al., 2014a; Nunney et al., 2014b). HR, as the final process of Horizontal Gene Transfer (HGT) in bacteria, occurs through three mechanisms: transformation, conjugation, and transduction, all of which end up integrating new genetic materials into bacterial chromosomes (Sheppard et al., 2018). HR has been further considered as a primary force for generating genetic diversity in *X. fastidiosa* based on genomic data and hypothesized to contribute to the appearance of new genotypes, expansion of geographic regions, change of virulence and fitness, and host adaptation (Potnis et al., 2019; Vanhove et al., 2019; Castillo et al., 2021a). Extensive HR especially Inter-

subspecific Homologous Recombination (IHR) has been detected throughout genomes of *X. fastidiosa* strains (Potnis et al., 2019; Vanhove et al., 2019; Castillo et al., 2020). In addition to recombination with closely related subspecies, *X. fastidiosa* has been shown to recombine with distantly related bacterial species (Firrao et al., 2021). However, the frequency of HR in strain's genomes varied dramatically among different populations. It was hypothesized that HR among *X. fastidiosa* strains was more extensive in native population compared to introduced populations, since introduced populations were more likely to donate new alleles to native population (Castillo et al., 2020; Landa et al., 2020; Castillo et al., 2021a). It has been found that many hotspots of recombined regions were involved in fundamental traits for *X. fastidiosa* ecology, suggesting the importance of HR in this pathogen adaptation and evolution (Potnis et al., 2019).

HR via natural competence has been experimentally confirmed in *X. fastidiosa* in vitro (Kung and Almeida, 2011). Natural competence for transformation or natural transformation is an ability to naturally uptake extracellular DNA and incorporate it into the chromosome (Blokesch, 2016). Such ability has been observed in some *X. fastidiosa* strains under several experimental conditions (Kandel et al., 2016). Many conditions and parameters have been described to affect natural competence in *X. fastidiosa* such as the flanking region homology, growth stage, media component, methylation of the DNA donor, and Type IV pili (TFP) (Kung and Almeida, 2011; Kung et al., 2013; Kung and Almeida, 2014; Kandel et al., 2016; Kandel et al., 2017). In the laboratory, the competence ability in *X. fastidiosa* dropped substantially when the size of the nonhomologous donor DNA increased (Kung et al., 2013). Another crucial factor, TFP, is involved not only in natural competence but also in twitching motility, adhesion and protein secretion (Kung and Almeida, 2014; Kandel et al., 2017; Kandel et al., 2018; Craig et al., 2019). Twitching motility, which mediates migration against flow, is also associated with virulence in *X. fastidiosa* (Li et al.,

2007). This movement appears to be positively correlated with the recombination rate of natural competence based on results from a dataset with a small number of strains (Kandel et al., 2017). Moreover, inter- and intra-subspecific HR have been documented in experimental conditions in *X. fastidiosa*, showing distinct recombination patterns (Kandel et al., 2017; Potnis et al., 2019). Although variation in recombination has been detected in many *X. fastidiosa* strains in silico, such differences were only experimentally reported in a few (13) strains. The presence of natural competence in diverse *X. fastidiosa* strains and how it correlates with recombination detected on the genomes remains poorly understood. Here we characterized natural competence in 71 *X. fastidiosa* strains mainly from subsp. *fastidiosa* and subsp. *multiplex*. Results showed that the majority of the tested subsp. *fastidiosa* strains were naturally competent, while only around a quarter of tested subsp. *multiplex* strains showed natural competence. Surprisingly, natural competence was found negatively correlated with HR as detected by fastGEAR based on core genome of 153 *X. fastidiosa* strains. The results of a Genome-Wide Association Study (GWAS) conducted using several different bioinformatic tools showed some interesting candidate genes that may be associated with natural competence, though there were no overlapping genes identified by all the tools, due to the limitation of sample size. These findings demonstrate distinct competent abilities between different subspecies and suggests that there may be other genetic mechanisms in addition to natural competence driving HR in different ecological environments.

Materials and methods

Bacterial strains, media, and culture conditions

Xylella fastidiosa strains used in this study are listed in Supplementary Table 2-1. Metadata of strains used for bioinformatic analysis in this study are listed in Supplementary material (available upon request). All strains were grown on PW (Davis et al., 1981) agar plates modified by omitting phenol red and using 1.8 g l^{-1} of bovine serum albumin (Gibco Life Sciences Technology) at 28°C for seven days, and re-streaked onto new PW plates for another week. Natural competence assays were performed using PD3 medium (Davis et al., 1981) used to suspend the cells in liquid and coculture the cells with donor DNA. LB (Luria Bertani) liquid medium was used to culture *Escherichia coli* cells for extracting plasmids. Kanamycin (Km) was used at a concentration of 30 µg ml^{-1}.

Test of natural competence in vitro in *X. fastidiosa* strains using donor plasmid DNA

Plasmid pKLN61 (Newman et al., 2004) was used as donor DNA, and it was purified from *E. coli* strain EAM1 that expresses *X. fastidiosa* DNA methylase PD1607 (Matsumoto and Igo, 2010). The plasmid was extracted from the cultures of EAM1 using a GeneJet Plasmid Miniprep kit (Thermo Scientific), and the concentration was adjusted to 100 ng/µl quantified using a Cytation 3 Image Reader spectrophotometer (BioTek Instruments Inc.). *X. fastidiosa* cells were collected from PW plates and suspended in 1 ml PD3 broth and further adjusted to a final OD_{600} of 0.25 (~10^8 cells/ml). Ten µl of the cell suspension were spotted onto PD3 plates, and 10 µl of 100 ng/µl pKLN61 plasmid were added on the top of the spots. Cultures were incubated at 28°C for 5~7 days depending on their growth rates (for strains that grew very slowly, the incubation time was increased to 10 days to collect sufficient cells). The spots were harvested in 1 ml of PD3 broth and diluted in 10-fold serial dilutions, which were plated on PW plates amended with or without Km. After 3 weeks of incubation at 28°C, colonies on PW plates amended with Km were counted as

recombinants and colonies on PW plates without Km were counted as total viable cells. Recombination frequency/rate was calculated as the ratio of number of recombinants to total viable cells. Five recombinants of each strain were selected to confirm recombination of Km marker from the donor plasmid using PCR as previously described (Kandel et al., 2016). At least two independent replicates were performed for each strain.

Confirmation of non-competence of *X. fastidiosa* strain De Donno

De Donno was selected as a representative strain to assess if the sequence of the homologous regions in the donor DNA targeted for recombination affected the non-competency of some strains. We set to test whether a non-competent strain show natural competence with donor DNA fragments composed of identical flanking regions to its own genome. Three different pili genes (*chpB*, *fimT2*, *pilV1*) were selected as the target genes, and their flanking regions were cloned from the genome of De Donno. Constructs of overlap-extension PCR products were obtained following a method previously described (Kandel et al., 2018). Briefly, upstream and downstream DNA fragments flanking the three genes of De Donno were amplified, purified and overlapped by PCR together with a Km resistance cassette fragment. Twenty µl of the constructed linear DNA fragments were used as donor DNA to conduct natural competence assays as described above.

Test of natural competence using heat-killed donor

To further confirm non-competency of strains that did not recombine with plasmid measured by the method described above, recombination with donor DNA from heat-killed cells of other *X. fastidiosa* strains was assessed. In one set of experiments, natural competence of *X. fastidiosa* subsp. *pauca* strain De Donno and XYL1961 was assessed with donor DNA of *X. fastidiosa* subsp.

fastidiosa mutants NS1::CmR, Temecula1-GFP, NS1::CmR/Km-WM1-1, NS1::CmR/Km-EB92-1 and *X. fastidiosa* subsp. *multiplex* mutant NS1::CmR/Km-AlmaEM3. In the other set of experiments, natural competence of *X. fastidiosa* subsp. *multiplex* strain BB08-1, XF3348, ESVL, XYL466, M12, and IVIA5901 were assessed with donor DNA of *X. fastidiosa* subsp. *multiplex* mutant NS1::Km-AlmaEM3. Donor cells (OD_{600} = 0.8 grown in PD3 broth) were heat-killed at 90°C for 15 min. Complete killing was confirmed by plating 100 µl of cell suspension onto PW plates. The heat-killed donor and live recipients (OD_{600} = 0.25 grown in PD3 broth) were spotted on PD3 plates and the following steps of plating were the same as the steps in natural competence assays as described above. Experiments were only carried out once.

Twitching motility of *X. fastidiosa* strains

Fresh (made at most 1 month before experiments) PW plates without Bovine Serum Albumin (BSA) were used to observe twitching motility (Galvani et al., 2007). Each plate was divided into 6 zones. Cells of each strain were scraped from PW plates using a sterile loop and then spotted on one zone using a toothpick, 10-12 spots of each strain were made in each zone. Two plates were used as technical replicates. After 4 days of incubation at 28°C, colony peripheral fringe of each spot was observed under 10X magnification using a Nikon Eclipse Ti Inverted Microscope (NIKON, Melville, NY), and photos of six colonies per plate per strain were taken for measurement of fringe width using ImageJ software (Schneider et al., 2012). Experiments were repeated independently at least three times.

Recombination detection and phylogenetic analysis

Recombination analysis was performed following the same pipeline used previously by our group to analyze HR of 55 *X. fastidiosa* strains (Potnis et al., 2019). All input genomes were re-annotated using prokka v1.13 (Seemann, 2014). Roary v3.13.0 (Page et al., 2015) was used to identify the core genomes of 48 (subset of strains tested in vitro) and 153 *X. fastidiosa* strains. The 0.72 Mb core-genome alignment of 153 strains was used to build a Maximum Likelihood (ML) tree using with RAxML v8.0.24 (Stamatakis, 2014) with the GTRGAMMA model (generalized time-reversible model) and 1,000 bootstrap replications. Recombinant regions in the core-genomes of 48 (subset of strains whose natural competence phenotype was tested in vitro) and 153 strains were identified by fastGEAR (Mostowy et al., 2017) with default parameters. fastGEAR first identified lineages in the dataset using hierBAPS followed by detecting ancestral recombination (recombination between two whole lineages) and recent recombination (recombination between a subset of strains between the two lineages) in the core genome alignment. To build recombination-free phylogeny of 153 strains, genomic sequences where recent recombination was detected were masked from the core genome alignment using the bng-recombination tool (https://github.com/tseemann/bng-tools/blob/master/bng-ma sk_recombination.pl). The masked alignment was used to build a phylogenetic tree without the impact of recombination. Phylogenetic trees were visualized in iToL (Letunic and Bork, 2019).

Genome-wide association study

To identify genetic factors associated with natural competence, the genomes of 48 strains characterized in vitro for natural competence were used. For the analysis, the natural competence phenotype was defined as a binary variable, where non-competent strains were assigned number "0" and competent strains were assigned number "1". To identify genes associated with HR,

recombination events detected by fastGEAR based on 153 strains were used as a continuous phenotype where recombination events were computed as a scale of 0-5. In the scale, strains with no recombination events were assigned "0" and strains with the most recombination events were assigned "5", and the percentage of number of recombination events of the strain compared to the most recombination events was given a corresponding number between 0 and 5. To increase the robustness of the study, we used two different tools, i.e., Pyseer (Lees et al., 2018) and DBGWAS v0.5.4 (Jaillard et al., 2018), with two methods, i.e., k-mer- and SNP-based methods. For SNP-based method, SNPs of the 48 strains were identified by ParSNP (Treangen et al., 2014) using Temecula1 as a reference genome. SNPs significantly associated with natural competence were identified by Pyseer using mixed effects model and annotated using snpEff (Cingolani et al., 2012). For k-mer-based method, we use both tools of Pyseer and DBGWAS, which used a mixed effects model and a fixed effects model respectively. Significant k-mers identified by Pyseer were mapped to three reference genomes of representative *X. fastidiosa* strain from three subspecies: Temecula1 (subsp. *fastidiosa*), AlmaEM3 (subsp. *multiplex*), and 9a5c (subsp. *pauca*). Significant unitigs identified by DBGWAS were mapped to the protein database of *X. fastidiosa* strain 9a5c downloaded from UniProt database (The UniProt, 2021). To help correct for population structure, we used the recombination-free phylogeny constructed as mentioned above. Manhattan plot was made using the ggplot2 package in R (Wickham, 2016). QQ plots were generated by scripts in the Pyseer program.

Statistical analysis

Natural competence rates > 0 were log-transformed to compute statistical differences compared with the recombination rate of the control strain TemeculaL using two-tailed Student's *t*-test.

Twitching motility data were also analyzed using two-tailed Student's *t*-test in R (Kassambara, 2020). To compare the difference of natural competence among different groups in subspecies, host, and geographic region, the recombination rates were transformed using formula of log(x+1) and analyzed using Kruskal-Wallis tests in R (Kassambara, 2020). We calculated the pairwise Pearson correlation coefficients between the transformed recombination rates of natural competence used in the Kruskal-Wallis test and the frequency of recombination events and plotted them using the ggplot2 package in R (Wickham, 2016). The frequency of recombination events was calculated as the ratio of number of recombination events of certain strain to the total number of recombination events based on 48 or 153 strains.

Results

Variation in natural competence among *X. fastidiosa* strains

Natural competence was measured by co-culturing *X. fastidiosa* cells with plasmid pKLN61 that contains a Km cassette inserted in the *rpfF* gene of Temecula1. Among 71 strains tested from 4 subspecies, 63% of strains were naturally competent in vitro. The majority of subsp. *fastidiosa* strains (96%) were naturally competent, while only 23% of subsp. *multiplex* strains were naturally competent in vitro (Fig. 1). Due to the small number of strains from subsp. *pauca* and subsp. *sandyi*, a conclusion on these subspecies could not be made, though none of them was competent. All six European strains tested were not naturally competent. To rule out a marker effect specific to *rpfF* gene, natural competence of the six European strains (De Donno, XYL1961, XF3348, ESVL, 466, IVIA5901) plus one USA strain (M12) was assessed using genomic DNA released

from heat-killed *X. fastidiosa* strains as donor DNA (NS1::CmR, Temecula1-GFP, NS1::CmR/Km-WM1-1, NS1::CmR/Km-EB92-1 and NS1::CmR/Km-AlmaEM3 for recipients De Donno and XYL1961; NS1::Km-AlmaEM3 for the rest of recipient strains). Results showed that they did not recombine with any of the DNA from heat-killed donors (data not shown). To assess if the sequence of the homologous regions in the donor DNA targeted for recombination affected the non-competency of some strains, we compared homologous region identity between donor DNA and recipient genomes of 48 strains (15 of them were non-competent) in the dataset whose genomes are available using BLAST. We found the identity of flanking regions between donor DNA and recipient genomes was 100% for all subsp. *fastidiosa* strains, 97.6% for subsp. *multiplex* and 96.2% for subsp. *pauca* on average (data not shown). Specially, subsp. *pauca* strain De Donno showed the least homologous region identity of 93.4%. Thus, to further assess the impact of homologous region identity on natural competence, De Donno was tested for natural competence in vitro using DNA fragments composed of flanking regions of three different pili genes cloned from its own genome. Nevertheless, no recombinants were obtained in these experiments (data not shown). Among naturally competent strains, recombination rates ranged from 5.8×10^{-7} (PD92-8, isolated from grape in FL) to 7.3×10^{-1} (Je17, isolated from grape in CA) (Fig. 1). The only two non-competent subsp. *fastidiosa* strains were isolated from grapes in Georgia (GA), but all the other 16 strains used in this study with the same origin were naturally competent.

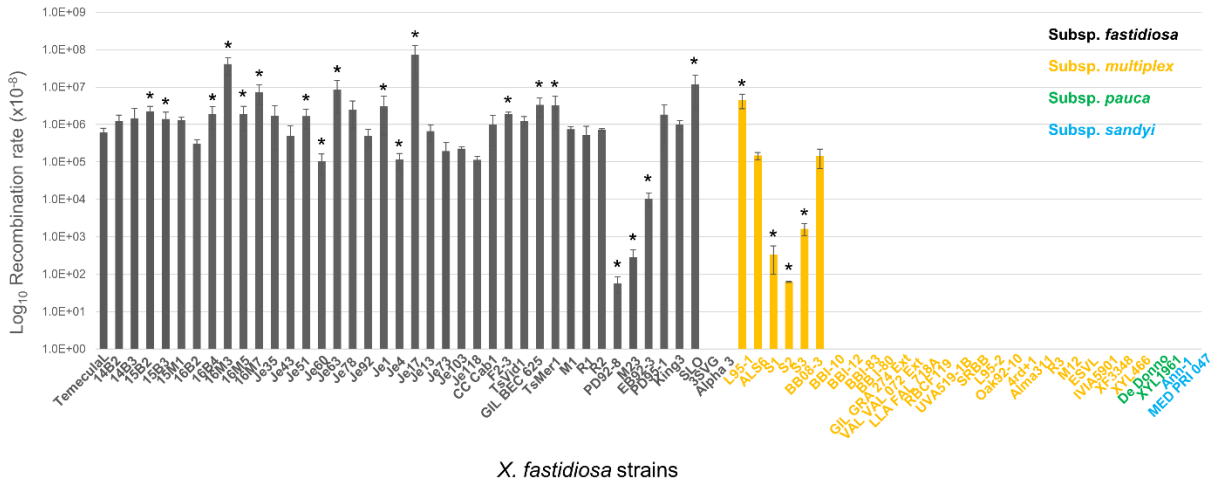


Fig. 1. Recombination frequency of 71 *X. fastidiosa* strains. The recombination frequency was $\log_{10}$-transformed. No bar indicates the recombination rate is "0" and the strain was non-competent. Different colors indicate different subspecies. The asterisk on the top of the bars indicates significant ($P < 0.05$) differences of recombination rates compared to control strain TemeculaL according to the two-tailed Student's *t*-test of log-transformed data. Experiments were repeated independently 2-3 times. Data represent means and standard errors.

Natural competence seems to be correlated with subspecies, and not with host or location

Next, we assessed the relationship between natural competence and subspecies classification, as well as ecological factors such as host and geographic origins, by comparing the average recombination rates among these three categories. Recombination rates of 13 strains previously measured by another study (Kandel et al., 2017) were also included in this analysis for a total of 84 strains. Results showed that subsp. *fastidiosa* strains had significantly higher recombination rates on average than subsp. *multiplex* strains, although the sample sizes of the two subspecies were not even (Fig. 2A). Recombination rates of subsp. *pauca* and *sandyi* were not compared due to their small sample size. Recombination rates varied among strains that infect different hosts and

come from different geographic locations. Due to limited access to strains with various origins, only three hosts and five geographic locations with > 6 strains in each host or location were compared. Only one strain isolated from pecan in Alabama was grouped with strains from Georgia (GA). The average recombination rate of strains isolated from grapes was significantly higher than that of those isolated from blueberry and almond (Fig. 2B), although this may be influenced by the high recombination rates of subsp. *fastidiosa,* as all strains isolated from grapes are classified as subsp. *fastidiosa*. None of the six European strains were naturally competent, but most (69%) of the US strains, mainly from CA, GA and Florida (FL), showed high recombination rates. CA strains showed highest average recombination rate among the groups, although it may be a result of the bias of subspecies classification as most CA strains were subsp. *fastidiosa* (Fig. 2C). Regardless of the variation in recombination rates among different groups, an obvious distribution pattern of natural competence in light of geographic locations and host was not observed, based on the limited data used here (Fig. 2D, E).

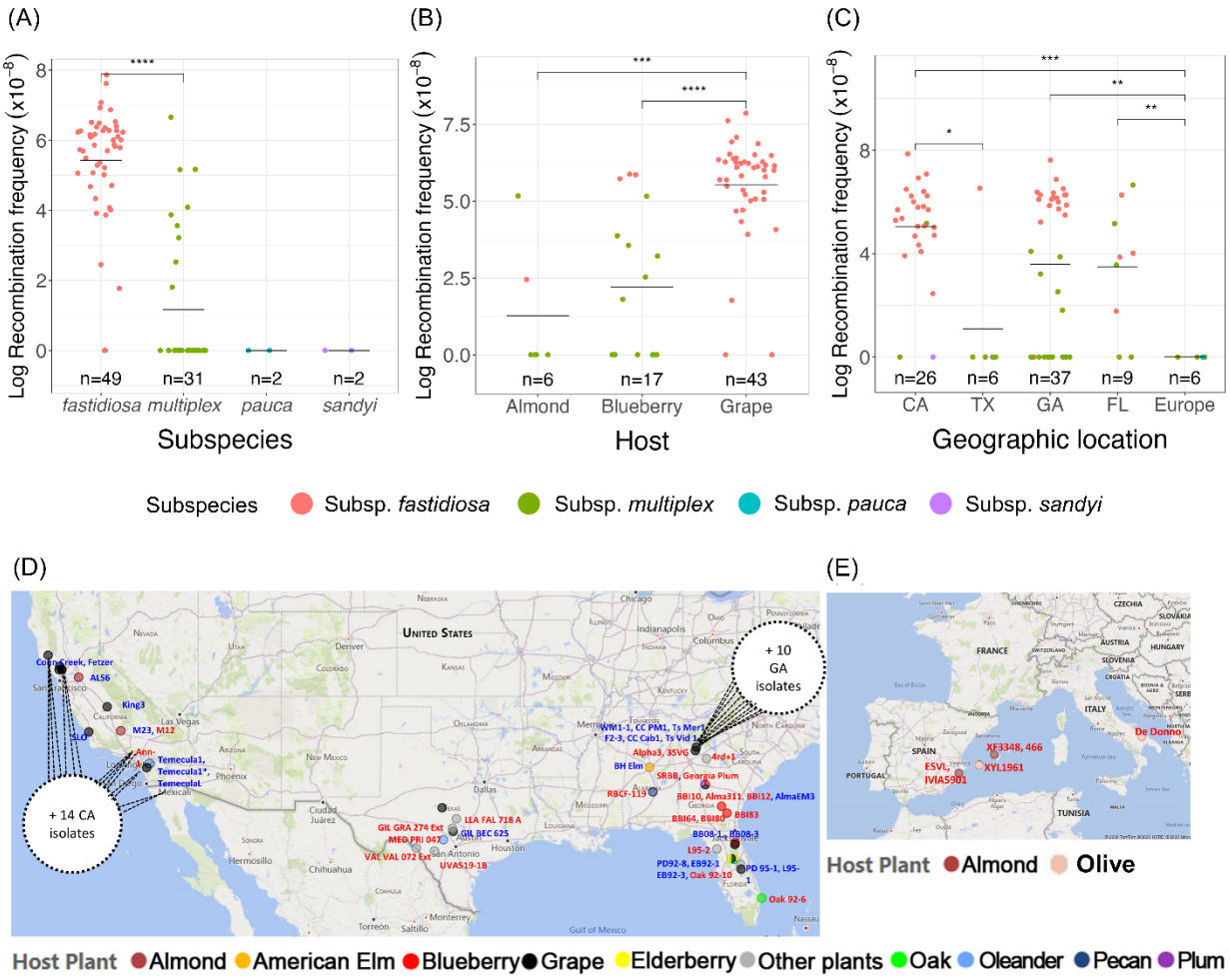


Fig. 2. Recombination rates in vitro grouped by subspecies (A), host (B), and geographic location (C) and distribution map of natural competence in the US (D) and Europe (E). For (A), (B), and (C), log transformed recombination rates were used for comparison of the mean values, represented as black horizontal lines, among different groups. Each subspecies was color-coded. Geographic locations use abbreviation for states in U.S.: California (CA), Texas (TX), Georgia (GA), Florida (FL). The asterisks above each subgroup indicate significant (* for $P < 0.05$, ** for $P < 0.01$, *** for $P < 0.001$, **** for $P < 0.0001$) differences between the two subgroups according to Kruskal-Wallis test. For (D) and (E), each circle represents a strain. The circle color signifies

host of the strain. The font color signifies whether the strain showed non-competency (red) or showed natural competence (blue).

Many non-competent strains were still motile via twitching

To assess whether twitching motility can explain the variation in natural competence, ten strains, seven of which were non-competent, were selected to measure fringe width of colonies representing twitching motility. These strains showed significantly different twitching motility compared to the control strain TemeculaL (Fig. 3). Surprisingly, all the non-competent strains except for strain 4rd+1 were capable of twitching, although over half of the strains showed limited twitching motility. Strain M23 that was naturally competent showed a deficiency in twitching motility. Twitching motility of SLO and L95-1 that had significantly higher recombination rates was significantly lower than that of TemeculaL. To investigate other factors that may be involved in natural competence, we conducted bioinformatic analyses.

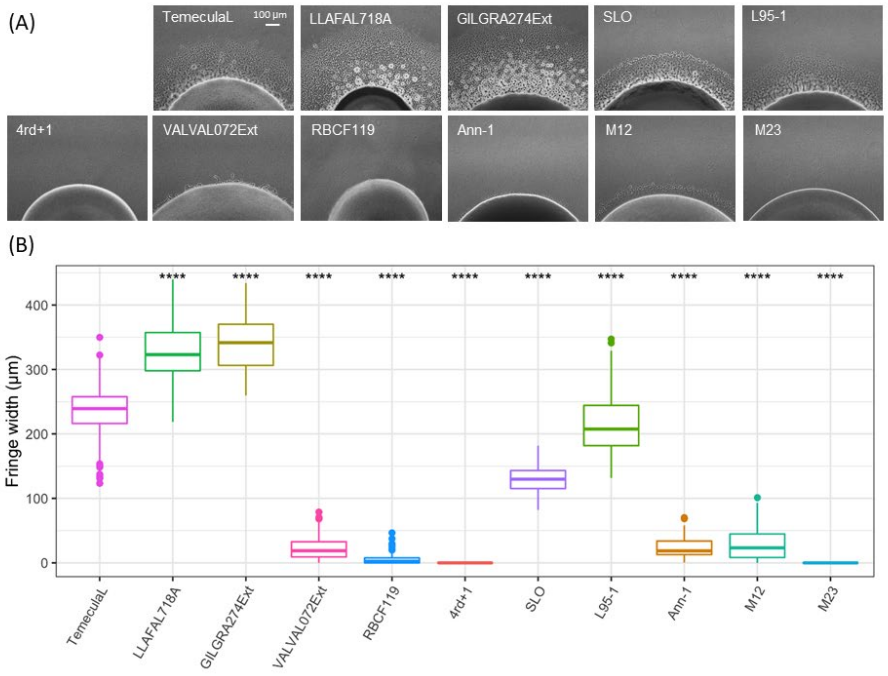


Fig. 3. Twitching motility of *X. fastidiosa* strains. Twitching motility was determined by measuring fringe width of colonies spotted on PW plates without BSA, after 4 days of growth. (A) Representative microscopic images of colony fringe of each strain. Scale bar on the first panel indicates 100 µm. (B) Box plot of colony fringe width of each strain. The four asterisks on the top of the boxes indicate significant ($P < 0.0001$) differences compared to TemeculaL according to the two-tailed Student's *t*-test. Measurements were repeated three times independently with at least 12 technical repeats each.

Recombination patterns among 153 *X. fastidiosa* strains

Previously IHR patterns among 55 *X. fastidiosa* strains were investigated (Potnis et al., 2019). Since more genomes of *X. fastidiosa* strains are currently available, we re-assessed IHR patterns with a larger dataset of 153 strains from five subspecies using fastGEAR. fastGEAR identified

four lineages including the three distinctly defined subspecies of subsp. *fastidiosa*, subsp. *multiplex*, subsp. *pauca*, and combined subsp. *sandyi* and *morus* into one lineage by using hierBAPS (Fig. 4). Due to large number of diverse *X. fastidiosa* strains, the core genome used to detect recombination events was only 0.72 Mb as defined by Roary. Ancestral (lineage-specific) recombination was detected between all lineages with a total number of 119 recombination events that varied from 15 bp to ~ 40 kb in length (Fig. 4A, Supplemental Table 2-2). The most frequent ancestral recombination occurred between subsp. *multiplex* and subsp. *fastidiosa* with 64 (54%) recombination events. The longest recombinant region between them was ~ 11 kb. Moreover, a total of 1757 recent (strain-specific) recombination events were detected with a size ranging from 10 bp to ~ 16 kb (Fig. 4B, Supplemental Table 2-3). *X. fastidiosa* subsp. *fastidiosa* was the least frequent recombinant group with only ~ 3 recombination events per strain on average, and over half of the strains only experienced 1 recombination event. The recombination events of subsp. *fastidiosa* were mainly from subsp. *multiplex* (73%), and none of them were from subsp. *pauca*. The length of recombinant regions varied from 16 bp to ~ 3.8 kb. For subsp. *multiplex*, the average number of recombination events was 14 per strain but showed clear polarization where strains in one group (IHR group) showed high number of recombination events, while strains in the other group (non-IHR group) showed only a few or no recombination events (Fig. 5A). ATCC35871, isolated from hybrid plum in GA, was previously assigned as an "intermediate" strain separating IHR and non-IHR groups as this strain showed moderate recombination compared to IHR and non-IHR strains (Nunney et al., 2014a; Landa et al., 2020). However, in this study, this strain also showed relatively high recombination events (59 events). Most (63%) of the recombination events were from subsp. *morus* and *sandyi* followed by subsp. *fastidiosa* (35%), while none of them were from subsp. *pauca*. The length of recombinant regions varied from 17 bp to ~ 4.5 kb. IHR also

affected phylogenetic relationships in terms of shortening the branch length that indicates genetic distances within subspecies while increasing the length between subspecies (Fig. 5B). The overall tree topology did not change much upon filtering recombination except for one subclade containing subsp. *multiplex* strains RBCF119, 4rd+1, Sy-VA and ATCC35871.

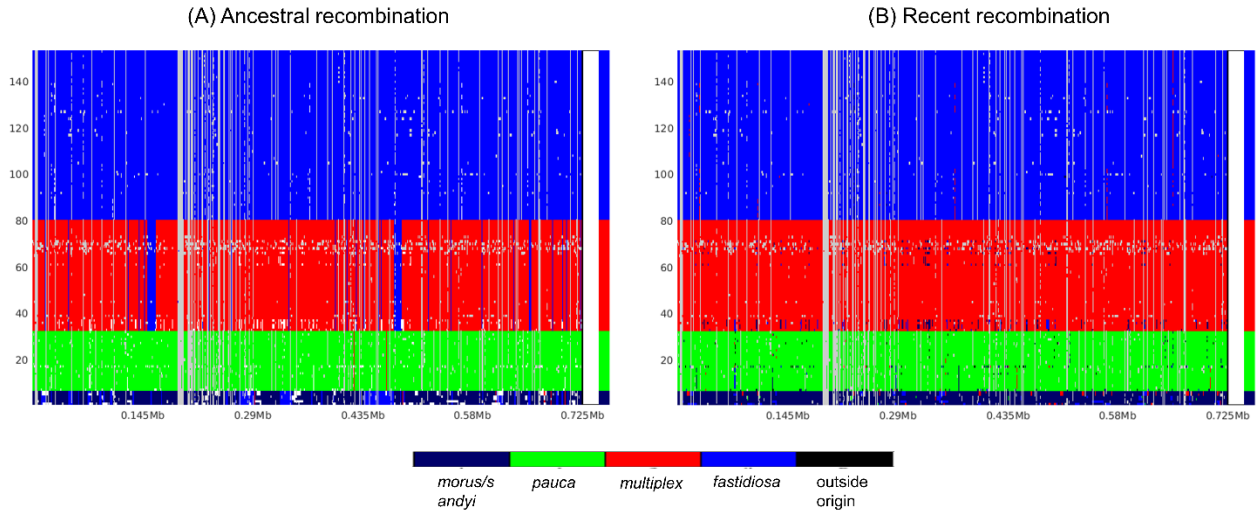


Fig. 4. (A) Ancestral (lineage-specific) and (B) recent (strain-specific) inter-lineage recombination across the core genome of 153 *X. fastidiosa* strains estimated by fastGEAR. Four lineages were identified based on hierBAPS analysis and were assigned with different colors: subsp. *fastidiosa* (blue), *pauca* (green), *multiplex* (red) and the combining group of subsp. *morus* and *sandyi* (dark blue). The black color signifies outside source. Stripes with different colors interfering the background color represent recombination events from the lineages the color infers in the genomic position of core genome.

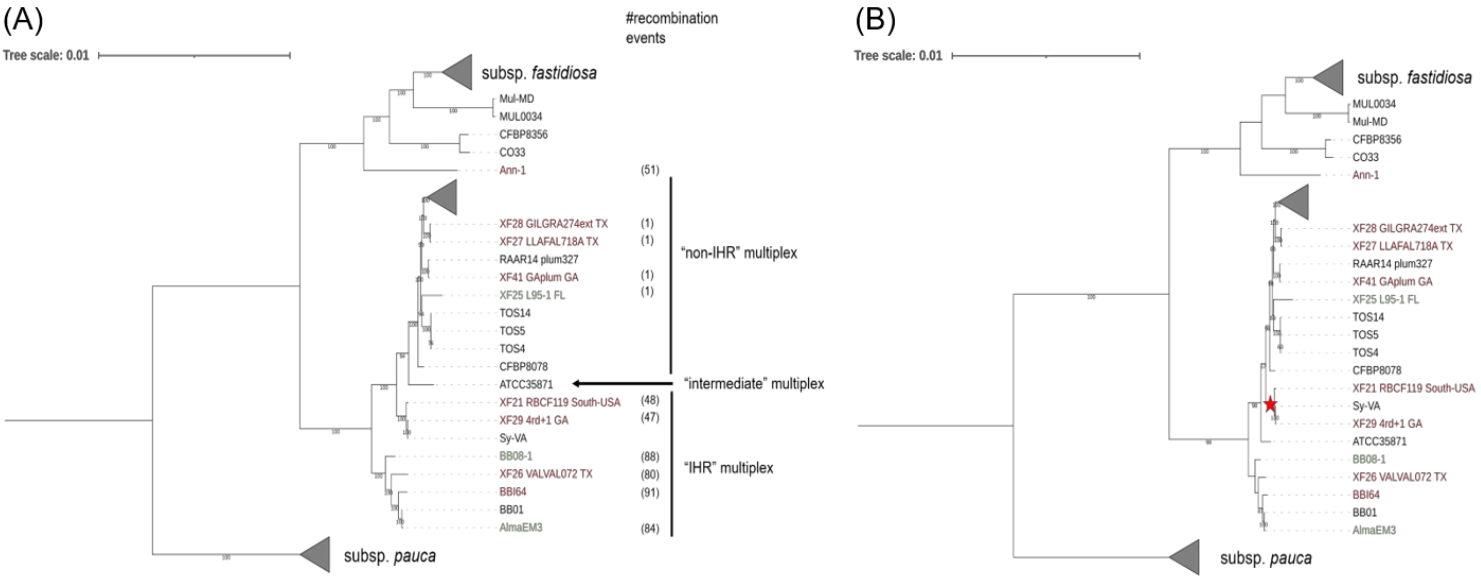


Fig. 5. Maximum Likelihood (ML) trees of 153 *X. fastidiosa* strains showed impact of recombination on phylogenetic relationships among strains. (A) ML tree of core genome alignment (0.72 Mb) of 153 strains from five subspecies. *X. fastidiosa* subsp. *fastidiosa*, subsp. *pauca*, and a subclade of subsp. *multiplex* were collapsed as the strain position did not change much upon filtering recombination. Strain names were colored based on their natural competence ability: competent (green) and non-competent (red), not tested (black). Numbers next to the strain names represent the number of recombination events detected by fastGEAR based on 153 strains. *X. fastidiosa* subsp. *multiplex* strains were divided into two groups, "IHR" (showing frequent inter-subspecific HR) and "non-IHR" (showing few inter-subspecific HR), by the "intermediate" recombinant strain ATCC35871 in the phylogenetic tree. (B) ML tree of core genome alignment masking recombinant regions detected by fastGEAR. The node with distinct position change was noted with a red star.

Natural competence is negatively correlated with recombination events detected by fastGEAR

To assess relationship between natural competence in vitro and HR in the genomes, we analyzed the correlation between natural competence and frequency of recombination events (calculated as a ratio of the number of recombination events against total recombination events) representing HR by calculating Pearson's correlation coefficient. Due to limited sample numbers, this analysis only included 48 strains with both recombination rates and number of recombination events available, including the 10 strains whose recombination rates were identified previously (Kandel et al., 2017). Since frequency of recombination events may differ when using larger datasets, this was measured for the 48 strains phenotyped in vitro using two fastGEAR results based on different datasets of 48 (only those used in vitro) and 153 (all genomes available, including the former 48 strains) strains. Surprisingly, natural competence in vitro was slightly but significantly negatively correlated with frequency of recombination events, no matter which fastGEAR results were used, with r values between -0.3 to -0.4 (Fig. 6). When using fastGEAR results of 153 strains, such linear association was strongly significant (P=0.0058). Such negative correlation was suspected within subsp. *fastidiosa*, most of which only experienced a few recombination events but were naturally competent, and within many subsp. *multiplex* strains in the IHR group, which experienced extensive recombination events but were non-competent when tested in vitro (Fig. 5A). However, low frequency of HR of many non-IHR subsp. *multiplex* strains corresponded to their non-competency in vitro. Moreover, subsp. *fastidiosa* showed obviously lower average frequency of recombination events than subsp. *multiplex*, which was opposite to the variation observed for natural competence in vitro.

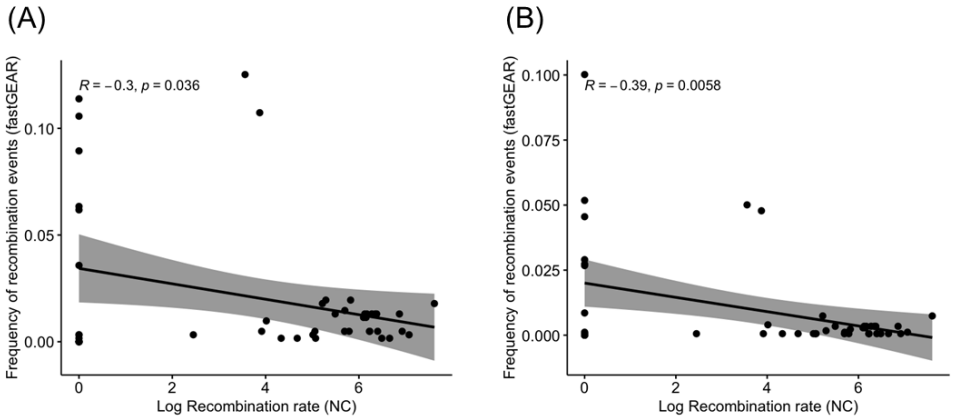


Fig. 6. Correlation results between recombination rate of natural competence in vitro (log-transformed) and frequency of recombination events based on 48 *X. fastidiosa* strains. Two different fastGEAR results from two datasets were analyzed: (A) the same 48 strains with natural competence phenotype available, (B) data of the 48 strains phenotyped in vitro from fastGEAR result of 153 strains. The Pearson correlation coefficient r and *P* values were noted in each figure.

Genome-wide association study

To understand genetic factors that may be associated with natural competence in vitro, whole-genome association study (GWAS) was performed to identify candidate genes or single nucleotide polymorphisms (SNPs) associated with this phenotype. To increase the robustness of the study, we used two methods, k-mer- and SNP-based, with two different tools, i.e., Pyseer (used both k-mer- and SNP-based methods) and DBGWAS (only used k-mer-based method). Due to limited genome samples, only the 48 strains that were used for correlation analysis were used due to their availability of both genomes and natural competence phenotype. A recombination-free phylogeny

of the 48 strains was built to help correct population structure for both tools and methods. The results showed that the top-hit genes identified by different methods were not overlapping, though results were consistent using the k-mer-based method no matter which tools were used. In the results of Pyseer based on both k-mer- and SNP-based methods which used a mixed effects model and a fixed effects model respectively, Quantile-Quantile (QQ) plots showed inflation, indicated by early deviation of the observed *P* value from the straight line (Supplementary Fig. 2-1.), suggesting a higher likelihood of residual false-positive associations. This may be due to the small sample size and biased dataset in which the natural competence phenotype was strongly associated with subspecies, and poor performance of the software used here on adequately correcting population structure which would be interfered by extensive recombination in *X. fastidiosa* genomes. Nonetheless, the identified significantly associated genes were somehow related to natural competence according to the literature. GWAS using Pyseer based on SNPs generated by Parsnp (using Temecula1 as reference genome) identified 428 significant SNPs (~ 5.8% of all the SNPs) with the *P* value of 6.8×10^{-7} as the threshold after Bonferroni correction (Top 10 gene hits were listed in Supplementary Table 2-4). The most significantly associated SNP ($P = 5.61 \times 10^{-37}$) was mapped to gene PD1306 (D-hexose-6-phosphate mutarotase) but annotated as a synonymous variant. The next most significantly associated SNP ($P = 7.81 \times 10^{-37}$) was mapped to upstream region of gene PD0405, which is known as the two component response regulator *rpfG* in Quorum Sensing (QS) in *X. fastidiosa*. However, no obvious peaks were observed in the Manhattan plot, indicating spurious associations of these two SNPs (Fig. 7A). The Pyseer results using k-mer-based method identified a total of 1379 significantly associated genes with a significance cutoff of 5.46×10^{-7} calculated by Pyseer program (Top 10 gene hits were listed in Supplementary Table 2-5). Such a high number of significantly associated genes (over 50% of total genes in *X. fastidiosa*)

indicates higher likelihood of spurious associations specially of those k-mers with only moderate significance. Such spurious associations were also supported by the QQ plot that showed an obvious inflation (Supplementary Fig. 2-2.). Interestingly, the most significant hits were mapped to PD1724 and PD1131 ($P = 1.58 \times 10^{-46}$) annotated as hypothetical phage-related proteins (Fig. 7B). Consistently, the same phage-related genes were identified as the most significantly associated genes by DBGWAS using k-mer-based method ($P = 1.41 \times 10^{-24}$) (Fig. 7C). In contrast to Pyseer results using the k-mer-based method, only 197 subgraphs were generated, indicating there were only 196 genetic events associated with natural competence, though a total of 799 significant unitigs were identified (Top 10 gene hits were listed in Supplementary Table 2-6). Moreover, as there was a weak negative correlation between natural competence in vitro and HR in the genomes, we tried to identify possible genetic factors associated with HR using DBGWAS to examine if similar genes would be identified, which would be associated with both natural competence and HR. The frequency of recombination events of 153 strains obtained before was computed as a scale of 0 to 5 as phenotype for GWAS. A total of 36 subgraphs were generated with only 139 significant unitigs. The top hit gene ($P = 1.53 \times 10^{-25}$) was XF0708 in strain 9a5c annotated as hypothetical phage holin family protein and is a homolog of PD1112 in Temecula1. Among the top 10 hit genes (Supplementary Table 2-7), two other hypothetical phage-related genes (XF1583 and XF1873) and gene XF0677 ($P = 1.5 \times 10^{-16}$) that is a homolog of PD1497 (*pilZ*) in Temecula1 were also identified. PilZ was annotated as a regulatory protein of TFP. The most significantly associated genes identified by GWAS did not overlap between competence-associated genes and HR-associated genes, although different hypothetical phage-related genes were found in both.

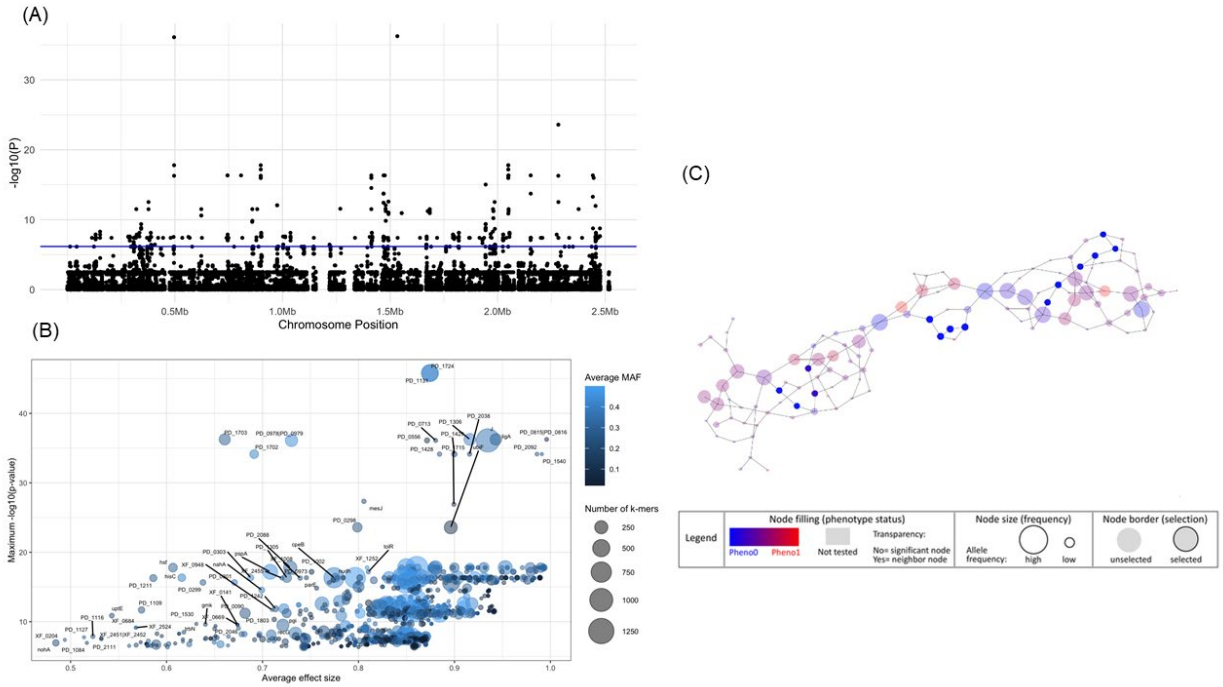


Fig. 7. GWAS results using different tools and methods for identifying competence-associated genes. (A). Manhattan plot of the GWAS mapping results using Pyseer and SNP-based method. Values of $-\log_{10}(P)$ in the y axis were plotted against base pair positions along the chromosome in the x axis. The line indicates the threshold for statistical significance of 6.8×10^{-7}. (B) Pyseer results using k-mers-based method. Values of maximum $-\log_{10}(P)$ in the y axis were plotted against average effect size in the x axis. (C) Top subgraph with minimum q value generated by DBGWAS.

Discussion

In the previous study assessing natural competence of 13 *X. fastidiosa* strains, a majority of strains were naturally competent (Kandel et al., 2017). However, our study showed almost half of the diverse 71 strains included in the dataset were not naturally competent, indicating natural competence may not be a common trait among *X. fastidiosa* strains. Such extensive variation on natural competence in vitro was also found in other bacteria (Coupat et al., 2008; Bossé et al., 2009; Maughan and Redfield, 2009; Evans and Rozen, 2013; Domingues et al., 2019; Durieux et al., 2019). It is noteworthy that the failure to detect natural competence may be simply due to an artifact of a lack of sensitivity of the experimental design, although the donor plasmid used here (pKLN61) induced higher recombination rates compared to other plasmids used in a previous study with the 13 *X. fastidiosa* strains (Kandel et al., 2017). For instance, growth rates of recipient strains and length of homologous flanking region and non-homologous insertion in donor plasmid DNA have been described to affect recombination rates in *X. fastidiosa* (Kung et al., 2013; Kandel et al., 2017). Therefore, it is likely that the media and donor plasmid DNA we used may not be conducive for high recombination rates for genetically diverse strains, especially for some subsp. *multiplex* strains that grow slower. However, we did observe in our dataset some strains that grew faster were non-competent, while some strains that grew slower were naturally competent. The relationship between growth rate and natural competence among different strains need further investigations. In terms of the homologous region sizes, the donor plasmid pKLN61 used here contains two homologous fragments of ~ 360 bp and ~ 750 bp. Only 96 bp flanking homologous regions were sufficient to allow *X. fastidiosa* initiate natural competence and usage of 1 kb flanking homologous regions would reach highest recombination rate, suggesting that the size of the flanking homologous regions in our experiments was adequate to induce relatively higher recombination rates (Kung et al., 2013). Sequence identity between donors and recipients was also

found to be correlated with recombination rate of natural competence in some other naturally competent bacteria (Zawadzki et al., 1995; Vulić et al., 1997; Majewski et al., 2000). Although such relationship in *X. fastidiosa* still remain unclear, it is highly likely that sequence identity would affect recombination of natural competence in *X. fastidiosa* due to the nature of homologous recombination in the process of natural competence. In our case, the flanking region identity between donor plasmid DNA and recipient genomes was at least above 97% even for 14 non-competent strains except for strain De Donno that showed lower but still high (93.4%) identity of homologous flanking regions. The non-competency of the subsp. *pauca* strain De Donno was further confirmed using three DNA fragments with flanking regions 100% identical to its own genome as donor DNA, but the results were in agreement with the failure of a recent attempt on transforming De Donno by natural competence (D'Attoma et al., 2020). These observations indicate the flanking region homology may not have a great impact on identifying non-competency under our experimental conditions. Investigating impact of identity of flanking region to recipient genomes on natural competence rates would help clarify this point. The non-competency of all six European strains tested using donor DNA released from heat-killed strains suggests marker effect in the donor plasmid may not affect identifying non-competency in our experimental setting. Nevertheless, measuring natural competence on different media or with different donor DNAs will help confirm our results. If the natural competence assessed in the laboratory truly reflected the natural competence ability of strains, such extensive variation in natural competence in vitro may suggest that natural competence has been gained or lost by different strains during evolution, probably due to a balance between benefits and costs of natural competence under various ecological environments. While natural competence may end up bringing new alleles that may be adaptive in certain environments, such new alleles may be deleterious or break down beneficial

allele combinations. There are also metabolic costs of expressing the machinery of DNA uptake and recombination. The trade-off of natural competence would largely depend on the fitness that natural competence would confer in a particular environments, which could varied among different strains with different origins (Vos, 2009).

When examining natural competence rates grouped by subspecies, host, and geographic location, subspecies classification seemed to be correlated to variation in natural competence in vitro. Due to the limited sample size and uneven sample structure, which was biased in subspecies and host, the conclusion may change when more strains with different host and geographic origins can be tested. Our result indi that natural competence may be correlated with subspecies classification as subsp. *fastidiosa* strains showed significantly higher recombination rate than subsp. *multiplex* strains. In the naturally competent bacterium *Pseudomonas stutzeri*, variation in natural competence rates was associated with genetic distances of the isolates, where genetically closely related isolates showed similar natural competence rates (Sikorski et al., 2002). Although the distribution pattern of natural competence rates in terms of genetic relationships among the strains was not studied, the fact that subsp. *fastidiosa* strains showed a higher average recombination rate than subsp. *multiplex* indicates natural competence may be associated with phylogenetic relationships among the strains. Additionally, in our dataset, all the six tested European strains were non-competent, corresponding with the hypothesis that introduced populations like European populations tend to have fewer recombination than native populations such as ones in Costa Rica and Brazil (Castillo et al., 2020; Landa et al., 2020). More European strains need to be tested for natural competence in the future.

Previous studies found twitching motility associated with natural competence in *X. fastidiosa* as both processes depend on TFP (Kung and Almeida, 2014; Kandel et al., 2017). However, in this study, many non-competent strains were still motile via twitching, and twitching motility did not seem to be related to natural competence rates, though this conclusion was based on a very small dataset of ten strains. More strains should be assessed for twitching motility in the future to find out the association between twitching motility and natural competence among different strains. Since TFP plays an important role in both twitching motility and DNA uptake in natural competence, variation in pili genes among *X. fastidiosa* strains may help explain the relationship between twitching motility and natural competence. Deletion of some TFP genes in TemeculaL causing non-competency or reduced natural competence rates did not completely abolish twitching motility (Merfa et al., 2021). Specially, deletion of the identified DNA uptake protein FimT3 in TemeculaL led to increased twitching motility (Merfa et al., 2021). This suggests those non-competent but motile strains in our small dataset may be varied in some specific TFP genes that play a more important role in DNA uptake compared to twitching motility. It is also possible that other factors in addition to TFP may be involved in natural competence. For this reason, we performed bioinformatic analyses.

Compared to our previous study investigating HR among 55 strains, the overall detected length of recombinant regions and frequency of recombination events in the core-genome in both ancestral and recent recombination were much lower (Potnis et al., 2019). This is probably due to the higher genomic diversity and as a result of this, the much smaller core-genome size of our dataset. Moreover, the relative number of recent recombination events of some strains and ancestral recombination between lineages also differed from the previous results, although most of them

were in accordance with the previous results (Potnis et al., 2019). For instance, Ann-1 was the most recombinant strain in subsp. *sandyi/morus* shown in this study, while it was previously the least recombinant in subsp. *sandyi/morus* (Potnis et al., 2019). The most frequent ancestral recombination was between subsp. *fastidiosa* and subsp. *multiplex* in this study, while previously it was between subsp. *fastidiosa* and subsp. *sandyi/morus*, suggesting that recombination between subsp. *fastidiosa* and subsp. *multiplex* tended to be more frequent in a more conserved core-genome (Potnis et al., 2019). Moreover, in contrast to previous findings that ATCC35871 showed moderate recombination and thus was classified as an "intermediate" strain between IHR and non-IHR groups (Nunney et al., 2014a; Landa et al., 2020), here it showed frequent recombination events compared to IHR strains, and was classified into IHR group. These results demonstrated that the addition of more genomes would change the detected recombination especially for certain strains, though the implications of these differences remain unclear. The strain CFBP8073, isolated from intercepted coffee plants imported from Mexico to France, was previously classified as subsp. *fastidiosa* by hierBAPS in fastGEAR based on 55 and 74 strains, by SNP-based analysis based on 27 strains, and by MLSA analysis based on 19 strains (Jacques et al., 2016; Giampetruzzi et al., 2017; Potnis et al., 2019; Vanhove et al., 2019). But it was classified as subsp. *sandyi* using a multiprimer PCR identification test and whole-genome alignment with 19 strains (Hernandez-Martinez et al., 2006; Jacques et al., 2016; Potnis et al., 2019; Vanhove et al., 2019). In this study, the fastGEAR results based on 153 strains classified CFBP8073 to belong to the lineage of subsp. *sandyi*/*morus* despite its comparatively lower number of recent recombination events compared to other subsp. *sandyi*/*morus* strains. Although our defined lineage was based on different methods and a relatively small core-genome size, this suggests CFBP8073 is more likely to be a subsp. *sandyi* strain and further confirmed the mosaic genomes of subsp. *sandyi*. In the previous analysis

of IHR in *X. fastidiosa*, CFBP8073 was considered as an outlier strain of subsp. *fastidiosa* as it showed extensive recombination events while other subsp. *fastidiosa* strains showed limited recombination events (Potnis et al., 2019; Vanhove et al., 2019). In our case, recombination of subsp. *fastidiosa* strains was more consistent showing limited recent recombination events compared to other lineages.

Surprisingly, a weak negative correlation between the recombination rate of natural competence and frequency of HR in the genomes was observed based on 48 strains. Larger datasets would be needed to test this correlation. It would be also interesting to assess natural competence of more subsp. *pauca*, *sandyi* and *morus* strains as many of them showed extensive recombination events in their genomes. It is worth noting that occurrence of natural competence in natural environments where *X. fastidiosa* inhabits, including plant xylem vessels and the foregut of insect vectors, remains to be confirmed, although some strains showed high natural competence rates in a microfluidic chamber, a system mimicking plant xylem vessels (Kandel et al., 2016). This negative correlation may suggest that natural competence in vitro may not correlate with natural competence occurring in natural environments. Furthermore, conjugation and transduction could also contribute to HR in *X. fastidiosa*. Transfer of mobile plasmids have been found between different subspecies of *X. fastidiosa* strains in vitro via conjugation dependent on *tra* and *trb* operon that encode conjugative Type IV secretion system (Burbank and Van Horn, 2017). In addition, conjugative and mobilizable plasmids were identified in silico in many strains' genomes (Burbank and Van Horn, 2017; Pierry et al., 2020). These findings suggest the potential of HR occurring via conjugation among *X. fastidiosa* populations. Conjugative or mobile elements have also been documented to interfere with essential genes for natural competence in other naturally competent

bacteria, indicating a possible negative interaction between conjugation and natural competence may exist in *X. fastidiosa* (Blokesch, 2017; Durieux et al., 2019). Moreover, transduction may also play a vital role in HR in *X. fastidiosa*, which is discussed in more detail in the next paragraph.

In GWAS analysis, the most significant shortcoming was the insufficient statistical power. In the future, GWAS results should be re-confirmed by using more strains with diverse origins as well as other bioinformatic tools that can better adjust population structure for highly recombinant species. Despite those drawbacks, the current results show some interesting candidate genes that were associated with natural competence and HR. A SNP in the upstream of *rpfG*, known as a response regulator in QS in *X. fastidiosa*, was the most significantly associated with natural competence. RpfG can be activated by RpfC, sensor of Diffusible Signal Factor (DSF) synthesized by RpfF, and was considered to regulate multiple phenotypes via the second messenger cyclic di-GMP (c-di-GMP) (Chatterjee et al., 2008b; Chatterjee et al., 2010; Beaulieu Ellen et al., 2013; de Souza Alessandra et al., 2013; Cursino et al., 2015; Baltenneck et al., 2021). In *X. fastidiosa*, cell-cell signaling was considered to be important for natural competence as a *rpfF* mutant of Temecula1 showed a significantly lower recombination rate with a donor plasmid DNA and even did not recombine with another mutant strain of Temecula1 in vitro, although natural competence did not seem to be affected by cell density (Kung and Almeida, 2011). Additionally, the *rpfF* mutant showed increased expression of 11 TFP genes including *fimT3*, whose product was identified as a DNA uptake protein, suggesting that RpfG may affect natural competence through regulating TFP via c-di-GMP in the pathway of QS (Wang et al., 2012; Merfa et al., 2021). Although the marker effect to *rpfF* may not affect identifying non-competency, it is still possible that the marker effect would influence measuring recombination rates of naturally competent

strains and thus may lead to the identification of the SNP in *rpfG*, which has close relationship with *rpfF*. Among the top genes associated with HR, one interesting candidate gene was *pilZ*, known as a regulatory gene in TFP. *pilZ* was shown to be an essential gene for natural competence as its deletion led to deficiency in natural competence in TemeculaL (Merfa et al., 2021). However, since it was not identified by GWAS as being significantly associated with natural competence, *pilZ* may play an important role in HR but not via natural competence. For instance, *pilZ* may also be associated with phage invasion which could contribute to HR. In other bacteria, mutation of *pilZ* resulted in failure of infection by phages, suggesting *pilZ* was a target of phages (Alm et al., 1996; Llontop et al., 2021). Furthermore, many hypothetical phage-related or prophage genes were identified as significantly associated with natural competence and HR, although different genes were detected. Prophage regions, which are viral DNA inserted into bacterial genomes but not infective until activation of lytic cycle, were commonly found in *X. fastidiosa* genomes of several strains including two highly recombinant strains CFBP8072 and CFBP8073 (Simpson et al., 2000; Van Sluys et al., 2003; Jacques et al., 2016). Prophages in *X. fastidiosa* have been associated with genome rearrangement affecting genome evolution (Van Sluys et al., 2003). These associations indicate that transduction may be related to natural competence and may contribute to HR in *X. fastidiosa*. However, all of these identified prophages are only hypothetical, and their integrities and functions remain to be confirmed. Studies showed that natural competence could be interfered by selfish mobile genetic elements including prophages (Blokesch, 2017). TFP have been described as a target of phages for *X. fastidiosa* and several other naturally competent bacteria (Waldor Matthew and Mekalanos John, 1996; Ahern et al., 2014; Chung et al., 2014; Harvey et al., 2018). In *Acinetobacter baylyi*, a prophage was observed to be reactivated and used a natural competence pilus as a receptor, inhibiting growth and natural competence during laboratory

evolution (Renda et al., 2016). It was hypothesized that *A. baylyi* may lose natural competence or limit expression of natural competence to avoid phage infections during evolution (Renda et al., 2016). Prophages were also shown to interfere with competence genes *comRS* which are also receptor-signal pair of QS in *Streptococcus pyogenes*, and as a result, the prophage inhibited natural competence of the bacterium (Rutbeek et al., 2021). These examples suggest possible ways that phage-related genes may affect natural competence in *X. fastidiosa*, although it requires more evidence.

Other genetic factors that were not identified here but may also be associated with natural competence in *X. fastidiosa*. Restriction-modification (R-M) systems were considered as a barrier of genetic transfer including process via natural competence to protect bacteria from invasion of foreign DNA which may result in deleterious mutations (Johnston et al., 2014). The systems, involved in DNA degradation, would hydrolyze the recombined DNA without methylation by Methyl Transferases (MTases) in certain recognition sites using Restriction Endonucleases (REases), killing recombinants and thus affecting recombination rates. In the model strain of *X. fastidiosa*, Temecula1, a type II R-M system was found to limit natural competence and electroporation efficiency, and a type I R-M system was also found to affect stable acquisition of foreigner DNA in electroporation (Guilhabert and Kirkpatrick, 2003; Matsumoto and Igo, 2010; Kung and Almeida, 2011). A recent study (O'Leary et al., 2020) showed that *X. fastidiosa* strains have diverse R-M systems with 3-4 type I R-M systems in their genomes. Methylation patterns of the type I R-M system varied at subspecies level to some extent, indicating that R-M systems may be one of the causes for variation in natural competence. Since different strains may have different R-M systems or methylation pattern, the donor DNA may contain some unknown motif recognized by unique MTases and REases the strain has and the REases may degrade the recombinants,

resulting in a phenotype of non-competency. Future work on understanding impact of R-M systems on natural competence in *X. fastidiosa* may help test this hypothesis.

To conclude, our results showed that natural competence varied extensively among *X. fastidiosa* strains especially among different subspecies, suggesting natural competence may evolve in response to different environments. The negative correlation between natural competence and HR and the GWAS results illustrates other genetic mechanisms such as conjugation and transduction may also contribute to HR in *X. fastidiosa*, indicating complicated evolving mechanisms. Future studies investigating the roles of environments as well as other mechanisms of HGT in *X. fastidiosa* evolution could provide insight into host shift and adaption.

References

Ahern, S.J., Das, M., Bhowmick, T.S., Young, R., and Gonzalez, C.F. 2014. Characterization of novel virulent broad-host-range phages of *Xylella fastidiosa* and *Xanthomonas*. J Bacteriol 196:459-471.

Alm, R.A., Bodero, A.J., Free, P.D., and Mattick, J.S. 1996. Identification of a novel gene, *pilZ*, essential for type 4 fimbrial biogenesis in *Pseudomonas aeruginosa*. J Bacteriol 178:46-53.

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

Baltenneck, J., Reverchon, S., and Hommais, F. 2021. Quorum sensing regulation in phytopathogenic bacteria. Microorganisms 9:239.

Beaulieu Ellen, D., Ionescu, M., Chatterjee, S., Yokota, K., Trauner, D., Lindow, S., and Ausubel Frederick, M. 2013. Characterization of a diffusible signaling factor from *Xylella fastidiosa*. mBio 4:e00539-00512.

Blokesch, M. 2016. Natural competence for transformation. Curr. Biol. 26:R1126-R1130.

Blokesch, M. 2017. In and out—contribution of natural transformation to the shuffling of large genomic regions. Curr. Opin. Microbiol. 38:22-29.

Bossé, 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.

Burbank, L.P., and Van Horn, C.R. 2017. Conjugative plasmid transfer in *Xylella fastidiosa* is dependent on *tra* and *trb* operon functions. J Bacteriol 199:e00388-00317.

Castillo, A.I., Chacon-Diaz, C., Rodriguez-Murillo, N., Coletta-Filho, H.D., and Almeida, R.P.P. 2020. Impacts of local population history and ecology on the evolution of a globally dispersed pathogen. BMC Genom. 21:369.

Castillo, A.I., Bojanini, I., Chen, H., Kandel, P.P., De La Fuente, L., and Almeida, R.P.P. 2021a. Allopatric plant pathogen population divergence following disease emergence. Appl. Environ. Microbiol. 87:e02095-02020.

Castillo, A.I., Tsai, C.-W., Su, C.-C., Weng, L.-W., Lin, Y.-C., Cho, S.-T., Almeida, R.P.P., and Kuo, C.-H. 2021b. Genetic differentiation of *Xylella fastidiosa* following the introduction into Taiwan. bioRxiv.

Cella, E., Angeletti, S., Fogolari, M., Bazzardi, R., De Gara, L., and Ciccozzi, M. 2018. Two different *Xylella fastidiosa* strains circulating in Italy: phylogenetic and evolutionary analyses. J. Plant Interact. 13:428-432.

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

Chatterjee, S., Wistrom, C., and Lindow, S.E. 2008b. A cell–cell signaling sensor is required for virulence and insect transmission of *Xylella fastidiosa*. Proc. Natl. Acad. Sci. U.S.A. 105:2670.

Chatterjee, S., Killiny, N., Almeida, R.P.P., and Lindow, S.E. 2010. Role of cyclic di-GMP in *Xylella fastidiosa* biofilm formation, plant virulence, and insect transmission. MPMI 23:1356-1363.

Chung, I.-Y., Jang, H.-J., Bae, H.-W., and Cho, Y.-H. 2014. A phage protein that inhibits the bacterial ATPase required for type IV pilus assembly. Proc. Natl. Acad. Sci. U.S.A. 111:11503.

Cingolani, P., Platts, A., Wang, L.L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X., and Ruden, D.M. 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. . Fly 6:80-92.

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

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.

Craig, L., Forest, K.T., and Maier, B. 2019. Type IV pili: dynamics, biophysics and functional consequences. Nat. Rev. Microbiol. 17:429-440.

Cursino, L., Athinuwat, D., Patel, K.R., Galvani, C.D., Zaini, P.A., Li, Y., De La Fuente, L., Hoch, H.C., Burr, T.J., and Mowery, P. 2015. Characterization of the *Xylella fastidiosa* PD1671 gene encoding degenerate c-di-GMP GGDEF/EAL domains, and its role in the development of Pierce's disease. PLOS ONE 10:e0121851.

D'Attoma, G., Morelli, M., De La Fuente, L., Cobine, P.A., Saponari, M., de Souza, A.A., De Stradis, A., and Saldarelli, P. 2020. Phenotypic characterization and transformation attempts reveal peculiar traits of *Xylella fastidiosa* subspecies *pauca* strain De Donno. Microorganisms 8:1832.

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 Souza Alessandra, A., Ionescu, M., Baccari, C., da Silva Aline, M., and Lindow Steven, E. 2013. Phenotype overlap in *Xylella fastidiosa* is controlled by the cyclic di-GMP phosphodiesterase Eal in response to antibiotic exposure and diffusible signal factor-mediated cell-cell signaling. Appl. Environ. Microbiol. 79:3444-3454.

Denancé, N., Briand, M., Gaborieau, R., Gaillard, S., and Jacques, M.-A. 2019. Identification of genetic relationships and subspecies signatures in *Xylella fastidiosa*. BMC Genom. 20:239.

Domingues, S., Rosário, N., Cândido, Â., Neto, D., Nielsen, K.M., and Da Silva, G.J. 2019. Competence for natural transformation is common among clinical strains of resistant *Acinetobacter* spp. Microorganisms 7:30.

Durieux, I., Ginevra, C., Attaiech, L., Picq, K., Juan, P.-A., Jarraud, S., and Charpentier, X. 2019. Diverse conjugative elements silence natural transformation in *Legionella* species. Proc. Natl. Acad. Sci. U.S.A. 116:18613-18618.

European Food Safety, A., Delbianco, A., Gibin, D., Pasinato, L., and Morelli, M. 2021. Update of the *Xylella* spp. host plant database - systematic literature search up to 31 December 2020. EFSA J 19:e06674.

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

Firrao, G., Scortichini, M., and Pagliari, L. 2021. Orthology-based estimate of the contribution of horizontal gene transfer from distantly related bacteria to the intraspecific diversity and differentiation of *Xylella fastidiosa*. Pathogens 10:46.

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.

Giampetruzzi, A., Saponari, M., Loconsole, G., Boscia, D., Savino, V.N., Almeida, R.P.P., Zicca, S., Landa, B.B., Chacón-Diaz, C., and Saldarelli, P. 2017. Genome-wide analysis provides evidence on the genetic relatedness of the emergent *Xylella fastidiosa* genotype in Italy to isolates from central America. Phytopathology 107:816-827.

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.

Harvey, H., Bondy-Denomy, J., Marquis, H., Sztanko, K.M., Davidson, A.R., and Burrows, L.L. 2018. *Pseudomonas aeruginosa* defends against phages through type IV pilus glycosylation. Nat. Microbiol. 3:47-52.

Hernandez-Martinez, R., Costa, H.S., Dumenyo, C.K., and Cooksey, D.A. 2006. Differentiation of strains of *Xylella fastidiosa* infecting grape, almonds, and oleander using a multiprimer PCR assay. PLANT DIS. 90:1382-1388.

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.

Jacques, M.-A., Denancé, N., Legendre, B., Morel, E., Briand, M., Mississipi, S., Durand, K., Olivier, V., Portier, P., Poliakoff, F., and Crouzillat, D. 2016. New coffee plant-infecting *Xylella fastidiosa* variants derived via homologous recombination. Appl. Environ. Microbiol. 82:1556-1568.

Jaillard, M., Lima, L., Tournoud, M., Mahé, P., van Belkum, A., Lacroix, V., and Jacob, L. 2018. A fast and agnostic method for bacterial genome-wide association studies: Bridging the gap between k-mers and genetic events. PLoS Genet. 14:e1007758.

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.

Kandel, P.P., Chen, H., and De La Fuente, L. 2018. A short protocol for gene knockout and complementation in *Xylella fastidiosa* shows that one of the type IV pilin paralogs (PD1926) is needed for twitching while another (PD1924) affects pilus number and location. Appl. Environ. Microbiol. 84:e01167-01118.

Kandel, P.P., Lopez, S.M., Almeida, R.P.P., and Fuente, L.D.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*. MPMI 30:589-600.

Kassambara, A. (2020). ggpubr: 'ggplot2' based publication ready plots.

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

Landa, B.B., Castillo, A.I., Giampetruzzi, A., Kahn, A., Román-Écija, M., Velasco-Amo, M.P., Navas-Cortés, J.A., Marco-Noales, E., Barbé, S., Moralejo, E., Coletta-Filho, H.D., Saldarelli, P., Saponari, M., and Almeida, R.P.P. 2020. Emergence of a plant pathogen in

Europe associated with multiple intercontinental introductions. Appl. Environ. Microbiol. 86:e01521-01519.
Lees, J.A., Galardini, M., Bentley, S.D., Weiser, J.N., and Corander, J. 2018. pyseer: a comprehensive tool for microbial pangenome-wide association studies. Bioinformatics 34:4310-4312.
Letunic, I., and Bork, P. 2019. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. Nucleic Acids Res. 47:W256-W259.
Li, Y., Hao, G., Galvani, C.D., Meng, Y., Fuente, L.D.L., Hoch, H.C., 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.
Llontop, E.E., Cenens, W., Favaro, D.C., Sgro, G.G., Salinas, R.K., Guzzo, C.R., and Farah, C.S. 2021. The PilB-PilZ-FimX regulatory complex of the type IV pilus from *Xanthomonas citri*. PLoS Pathog. 17:e1009808.
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.
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. Microbiol. 75:1679-1687.
Maughan, H., and Redfield, R.J. 2009. Extensive variation in natural competence in *Haemophilus Influenzae*. Evolution 63:1852-1866.
Merfa, M.V., Zhu, X., Shantharaj, D., Gomez, L.M., Naranjo, E., Potnis, N., Cobineb, P.A., and Fuente, L.D.L. (2021). Movement and evolution: complete functional analysis of type IV pilus of a re-emergent plant pathogen reveals a unique DNA receptor. In unpublished.
Mostowy, R., Croucher, N.J., Andam, C.P., Corander, J., Hanage, W.P., and Marttinen, P. 2017. Efficient inference of recent and ancestral recombination within bacterial populations. Mol. Biol. Evol. 34:1167-1182.
Newman Karyn, L., Almeida Rodrigo, P.P., Purcell Alexander, H., and Lindow Steven, 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. U.S.A. 101:1737-1742.
Nunney, L., Azad, H., and Stouthamer, R. 2019. An experimental test of the host-plant range of nonrecombinant strains of north American *Xylella fastidiosa* subsp. *multiplex*. Phytopathology 109:294-300.
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., 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-Astúa, 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.

O'Leary, M., Burbank, L., and Stenger, D.C. (2020). Distinct genetic lineages of *Xylella fastidiosa* carry conserved type I restriction-modification systems with diverse specificity subunits. In In Proceedings of the APS Annual Meeting "Plant Health 2020" (Online).

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.

Olmo, D., Nieto, A., Borràs, D., Montesinos, M., Adrover, F., Pascual, A., Gost, P.A., Quetglas, B., Urbano, A., García, J.d.D., Velasco-Amo, M.P., Olivares-García, C., Beidas, O., Juan, A., Marco-Noales, E., Gomila, M., Rita, J., Moralejo, E., and Landa, B.B. 2021. Landscape epidemiology of *Xylella fastidiosa* in the balearic islands. Agronomy 11:473.

Page, A.J., Cummins, C.A., Hunt, M., Wong, V.K., Reuter, S., Holden, M.T.G., Fookes, M., Falush, D., Keane, J.A., and Parkhill, J. 2015. Roary: rapid large-scale prokaryote pan genome analysis. Bioinformatics 31:3691-3693.

Pierry, P.M., Uceda-Campos, G., Feitosa-Junior, O.R., Martins-Junior, J., de Santana, W.O., Coletta-Filho, H.D., Zaini, P.A., and da- Silva, A.M. 2020. Genetic diversity of *Xylella fastidiosa* plasmids assessed by comparative genomics. Trop. plant pathol. 45:342-360.

Potnis, N., Kandel, P.P., Merfa, M.V., Retchless, A.C., Parker, J.K., Stenger, D.C., Almeida, R.P.P., Bergsma-Vlami, M., Westenberg, M., Cobine, P.A., and Fuente, L.D.L. 2019. Patterns of inter- and intrasubspecific homologous recombination inform eco-evolutionary dynamics of *Xylella fastidiosa*. ISME J 13:2319-2333.

Pratt, J.T., Tamayo, R., Tischler, A.D., and Camilli, A. 2007. PilZ domain proteins bind cyclic diguanylate and regulate diverse processes in *Vibrio cholerae*. J. Biol. Chem. 282:12860-12870.

Renda, B.A., Chan, C., Parent, K.N., and Barrick, J.E. 2016. Emergence of a competence-reducing filamentous phage from the genome of *Acinetobacter baylyi* ADP1. J Bacteriol 198:3209-3219.

Rutbeek, N.R., Rezasoltani, H., Patel, T.R., Khajehpour, M., and Prehna, G. 2021. Molecular mechanism of quorum sensing inhibition in *Streptococcus* by the phage protein paratox. J. Biol. Chem. 297:100992.

Ryjenkov, D.A., Simm, R., Römling, U., and Gomelsky, M. 2006. The PilZ domain is a receptor for the second messenger c-di-GMP: the PilZ domain protein YcgR controls motility in enterobacteria. J. Biol. Chem. 281:30310-30314.

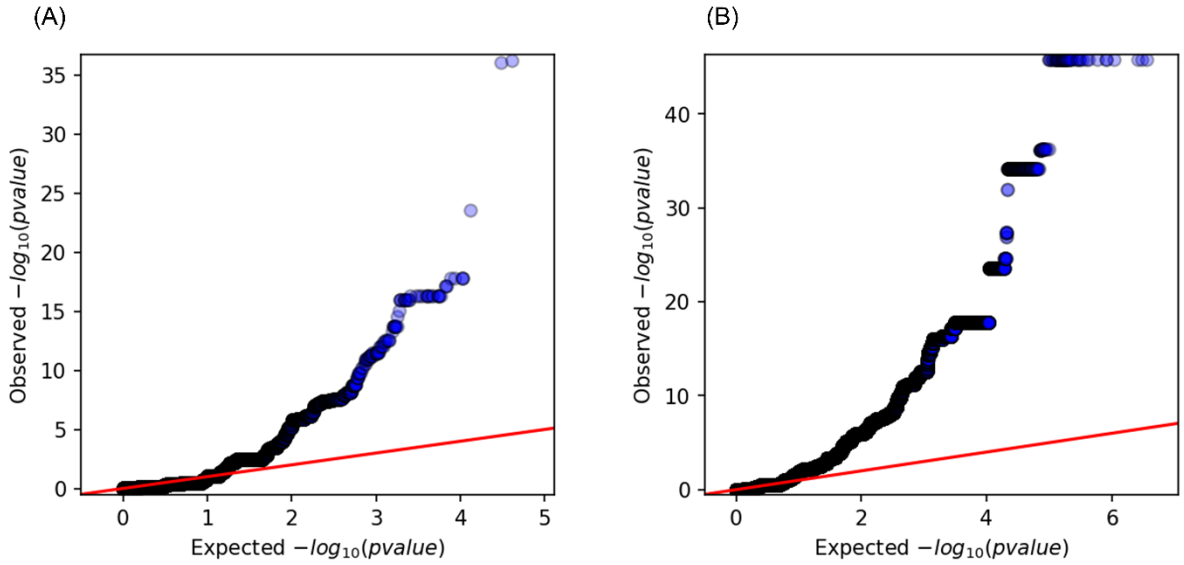
Saponari, M., Giampetruzzi, A., Loconsole, G., Boscia, D., and Saldarelli, P. 2018. *Xylella fastidiosa* in olive in Apulia: where we stand. Phytopathology 109:175-186.

Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. 2012. NIH Image to ImageJ: 25 years of image analysis. Nat. Methods 9:671-675.

Seemann, T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068-2069.
Sheppard, S.K., Guttman, D.S., and Fitzgerald, J.R. 2018. Population genomics of bacterial host adaptation. Nat. Rev. Genet. 19:549-565.
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.
Simpson, A.J.G., Reinach, F.C., Arruda, P., Abreu, F.A., Acencio, M., Alvarenga, R., Alves, L.M.C., Araya, J.E., Baia, G.S., Baptista, C.S., Barros, M.H., Bonaccorsi, E.D., Bordin, S., Bové, J.M., Briones, M.R.S., Bueno, M.R.P., Camargo, A.A., Camargo, L.E.A., Carraro, D.M., Carrer, H., Colauto, N.B., Colombo, C., Costa, F.F., Costa, M.C.R., Costa-Neto, C.M., Coutinho, L.L., Cristofani, M., Dias-Neto, E., Docena, C., El-Dorry, H., Facincani, A.P., Ferreira, A.J.S., Ferreira, V.C.A., Ferro, J.A., Fraga, J.S., França, S.C., Franco, M.C., Frohme, M., Furlan, L.R., Garnier, M., Goldman, G.H., Goldman, M.H.S., Gomes, S.L., Gruber, A., Ho, P.L., Hoheisel, J.D., Junqueira, M.L., Kemper, E.L., Kitajima, J.P., Krieger, J.E., Kuramae, E.E., Laigret, F., Lambais, M.R., Leite, L.C.C., Lemos, E.G.M., Lemos, M.V.F., Lopes, S.A., Lopes, C.R., Machado, J.A., Machado, M.A., Madeira, A.M.B.N., Madeira, H.M.F., Marino, C.L., Marques, M.V., Martins, E.a.L., Martins, E.M.F., Matsukuma, A.Y., Menck, C.F.M., Miracca, E.C., Miyaki, C.Y., Monteiro-Vitorello, C.B., Moon, D.H., Nagai, M.A., Nascimento, A.L.T.O., Netto, L.E.S., Nhani, A., Nobrega, F.G., Nunes, L.R., Oliveira, M.A., de Oliveira, M.C., de Oliveira, R.C., Palmieri, D.A., Paris, A., Peixoto, B.R., Pereira, G.a.G., Pereira, H.A., Pesquero, J.B., Quaggio, R.B., Roberto, P.G., Rodrigues, V., de M. Rosa, A.J., de Rosa, V.E., de Sá, R.G., Santelli, R.V., Sawasaki, H.E., da Silva, A.C.R., da Silva, A.M., da Silva, F.R., Silva, W.A., da Silveira, J.F., Silvestri, M.L.Z., Siqueira, W.J., de Souza, A.A., de Souza, A.P., Terenzi, M.F., Truffi, D., Tsai, S.M., Tsuhako, M.H., Vallada, H., Van Sluys, M.A., Verjovski-Almeida, S., Vettore, A.L., Zago, M.A., Zatz, M., Meidanis, J., and Setubal, J.C. 2000. The genome sequence of the plant pathogen *Xylella fastidiosa*. Nature 406:151-157.
Stamatakis, A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30:1312-1313.
The UniProt, C. 2021. UniProt: the universal protein knowledgebase in 2021. Nucleic Acids Res. 49:D480-D489.
Treangen, T.J., Ondov, B.D., Koren, S., and Phillippy, A.M. 2014. The Harvest suite for rapid core-genome alignment and visualization of thousands of intraspecific microbial genomes. Genome Biol. 15:524.
Van Sluys, M.A., de Oliveira, M.C., Monteiro-Vitorello, C.B., Miyaki, C.Y., Furlan, L.R., Camargo, L.E.A., da Silva, A.C.R., Moon, D.H., Takita, M.A., Lemos, E.G.M., Machado, M.A., Ferro, M.I.T., da Silva, F.R., Goldman, M.H.S., Goldman, G.H., Lemos, M.V.F., 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.C., 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., Sassaki, F.T., Sena, J.A.D., de Souza, A.A., Truffi, D., Tsukumo, F., Yanai, G.M., Zaros, L.G., Civerolo, E.L., Simpson, A.J.G., Almeida, N.F., 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.
Vanhove, M., Retchless, A.C., Sicard, A., Rieux, A., Coletta-Filho, H.D., De La Fuente, L., Stenger, D.C., and Almeida, R.P.P. 2019. Genomic diversity and recombination among *Xylella fastidiosa* subspecies. Appl. Environ. Microbiol. 85:e02972-02918.
Vos, M. 2009. Why do bacteria engage in homologous recombination? Trends Microbiol. 17:226-232.
Waldor Matthew, K., and Mekalanos John, J. 1996. Lysogenic conversion by a filamentous phage encoding cholera toxin. Science 272:1910-1914.
Wang, N., Li, J.-L., and Lindow, S.E. 2012. RpfF-dependent regulon of *Xylella fastidiosa*. Phytopathology 102:1045-1053.
Wickham, H. 2016. ggplot2: elegant graphics for data analysis. Springer-Verlag New York.

Supplementary materials for Chapter 2



Supplementary Fig. 2-1. Diagnostic Quantile-quantile (Q-Q) plots of expected *p*-values versus observed *p*-values of (A) SNPs and (B) k-mers identified using Pyseer.

Supplementary Table 2-1:

Strain	Subspecies	Host	Isolation year	Geographical Origin	Biosample or GenBank Accession No.
14B2	subsp. *fastidiosa*	*Vitis vinifera*	2014	USA, Georgia, Site 1	To be published
14B3	subsp. *fastidiosa*	*Vitis vinifera*	2014	USA, Georgia, Site 1	To be published
15B2	subsp. *fastidiosa*	*Vitis vinifera*	2015	USA, Georgia, Site 1	To be published
15B3	subsp. *fastidiosa*	*Vitis vinifera*	2015	USA, Georgia, Site 1	To be published
15M1	subsp. *fastidiosa*	*Vitis vinifera*	2015	USA, Georgia, Site 2	To be published
16B2	subsp. *fastidiosa*	*Vitis vinifera*	2016	USA, Georgia, Site 1	To be published
16B4	subsp. *fastidiosa*	*Vitis vinifera*	2016	USA, Georgia, Site 1	To be published
16M3	subsp. *fastidiosa*	*Vitis vinifera*	2016	USA, Georgia, Site 2	To be published
16M5	subsp. *fastidiosa*	*Vitis vinifera*	2016	USA, Georgia, Site 2	To be published
16M7	subsp. *fastidiosa*	*Vitis vinifera*	2016	USA, Georgia, Site 2	To be published
Je1	subsp. *fastidiosa*	*Vitis vinifera*	2016	USA, California, Santa Barbara, Lompoc	SAMN10358662
Je103	subsp. *fastidiosa*	*Vitis vinifera*	2016	USA, California, Temecula, Vista de Oro road	SAMN10358667

Je118	subsp. *fastidiosa*	*Vitis vinifera*	2016	USA, California, Temecula, Avenida Arriba	SAMN10358683
Je13	subsp. *fastidiosa*	*Vitis vinifera*	2016	USA, California, Sonoma, Bradford Mountain	SAMN10358689
Je17	subsp. *fastidiosa*	*Vitis vinifera*	2016	USA, California, Sonoma, Bradford Mountain	SAMN10358693
Je35	subsp. *fastidiosa*	*Vitis vinifera*	2016	USA, California, Napa valley, Conn Creek road	SAMN10358713
Je4	subsp. *fastidiosa*	*Vitis vinifera*	2016	USA, California, Santa Barbara, Cebada Canyon road	SAMN10358718
Je43	subsp. *fastidiosa*	*Vitis vinifera*	2016	USA, California, Napa valley, Rutherford	SAMN10358722
Je51	subsp. *fastidiosa*	*Vitis vinifera*	2016	USA, California, Napa valley, Veteran peak (nearby)	SAMN10358731
Je60	subsp. *fastidiosa*	*Vitis vinifera*	2016	USA, California, Bakersfiled, Tower Line road	SAMN10358741
Je63	subsp. *fastidiosa*	*Vitis vinifera*	2016	USA, California, Bakersfield, General Beale road	SAMN10358744

Je73	subsp. *fastidiosa*	*Vitis vinifera*	2016	USA, California, Bakersfield, Buena Vista boulevard	SAMN10358755
Je78	subsp. *fastidiosa*	*Vitis vinifera*	2016	USA, California, Bakersfield, General Beale road	SAMN10358760
Je92	subsp. *fastidiosa*	*Vitis vinifera*	2016	USA, California, Sonoma, Porter Creek	SAMN10358776
TemeculaL	subsp. *fastidiosa*	*Vitis vinifera*	-	USA, California, Temecula	PUJJ00000000.1
M23	subsp. *fastidiosa*	*Prunus dulcis*	2003	USA, California, San Joaquin Valley, Kern County	CP001011.1
SLO	subsp. *fastidiosa*	*Vitis vinifera*	1994	USA, California, San Luis Ovispo	SAMN09941414
3SVG	subsp. *fastidiosa*	*Vitis* spp.	-	USA, Georgia, North Georgia	-
Alpha 3	subsp. *fastidiosa*	*Vitis* spp.	-	USA, Georgia, North Georgia	-
CC Cab 1	subsp. *fastidiosa*	*Vitis vinifera*	2010	USA, Georgia	-
EB92-3	subsp. *fastidiosa*	*Sambucus* spp.	-	USA, Florida, Lake	-
F2-3	subsp. *fastidiosa*	*Vitis vinifera*	2009	USA, Georgia, Lumpking, Dahlonega	-
GIL BEC 625	subsp. *fastidiosa*	*Vitis vinifera*	-	USA, Texas, Stonewall, Gillespie	-

King 3	subsp. *fastidiosa*	*Vitis vinifera*	-	USA, California, Kingsburg, Fresno	-
M1	subsp. *fastidiosa*	*Vaccinium* spp.	2017	USA, Georgia, Pierce, Stevie Yong Farm	-
PD92-8	subsp. *fastidiosa*	*Vitis* sp.	-	USA, Florida, Lake, Leesburg	-
PD95-1	subsp. *fastidiosa*	*Vitis* sp.	-	USA, Florida, Osceola, St. Cloud	-
R1	subsp. *fastidiosa*	*Vaccinium* spp.	2017	USA, Georgia, Bacon, Blake Williams Farm	-
R2	subsp. *fastidiosa*	*Vaccinium* spp.	2017	USA, Georgia, Bacon, Blake Williams Farm	-
TsMer1	subsp. *fastidiosa*	*Vitis vinifera*	2010	USA, Georgia, Lumpking, Dahlonega	-
TSVid1	subsp. *fastidiosa*	*Vitis vinifera*	2010	USA, Georgia, Lumpking, Dahlonega	-
ESVL	subsp. *multiplex*	*Prunus dulcis*	2017	Spain, Alicante	NZ_QPQV01
IVIA5901	subsp. *multiplex*	*Prunus dulcis*	2017	Spain, Alicante	NZ_QPQW01
XF3348	subsp. *multiplex*	*Prunus dulcis*	2018	Spain, Mallorca	VDCL01
XYL466	subsp. *multiplex*	*Olea europaea*	-	Spain, Mallorca	To be published
L95-1	subsp. *multiplex*	*Lupinus aridonum*	-	USA, Florida, Orange	To be published
M12	subsp. *multiplex*	*Prunus dulcis*	2003	USA, California, Kern	CP000941.1

VAL VAL 072	subsp. *multiplex*	*Ambrosia trifida var. texana*	-	USA, Texas, Del Rio, Val Verde	To be published
LLA FAL 718A	subsp. *multiplex*	*Iva angustifolia*	-	USA, Texas, Llano, Tow	To be published
GIL GRA 274 ext	subsp. *multiplex*	*Helianthus annuus*	-	USA, Texas, Gillespie, Willow City	To be published
4rd+1	subsp. *multiplex*	*Platanus occidentalis*	-	USA, Georgia, Athens-Clarke	To be published
RBCF119	subsp. *multiplex*	*Carya illinoinensis*	2012	USA, Alabama, Lowndes, Lowndesboro	To be published
Alma311	subsp. *multiplex*	*Vaccinium* spp.	-	USA, Georgia, Alma, Bacon	-
ALS6	subsp. *multiplex*	*Prunus dulcis*	-	USA, California, San Joaquin	-
BB08-3	subsp. *multiplex*	*Vaccinium* spp.	-	USA, Florida, Palatka, Putnam	-
BBI-10	subsp. *multiplex*	*Vaccinium* spp.	-	USA, Georgia, Alma, Bacon	-
BBI-12	subsp. *multiplex*	*Vaccinium* spp.	-	USA, Georgia, Alma, Bacon	-
BBI-80	subsp. *multiplex*	*Vaccinium* spp.	-	USA, Georgia, Brantley, Hoboken	-
BBI-83	subsp. *multiplex*	*Vaccinium* spp.	-	USA, Georgia, Brantley, Hoboken	-
L95-2	subsp. *multiplex*	*Lupinus villosus*	-	USA, Florida, Levy	-
Oak92-10	subsp. *multiplex*	*Quercus nigra*	-	USA, Florida, Lake	-
R3	subsp. *multiplex*	*Vaccinium* spp.	2017	USA, Georgia, Bacon, Blake	-

S1	subsp. *multiplex*	*Vaccinium* spp.	2017	Williams Farm USA, Georgia, Bacon, Blake Williams Farm USA, Georgia, Bacon, Blake	-
S2	subsp. *multiplex*	*Vaccinium* spp.	2017	Williams Farm USA, Georgia, Bacon, Blake	-
S3	subsp. *multiplex*	*Vaccinium* spp.	2017	Williams Farm USA, Georgia,	-
SRBB	subsp. *multiplex*	*Vaccinium* spp.	-	South Georgia USA, Texas,	-
UVA519-1B	subsp. *multiplex*	*Cercis* spp.	-	Uvalde, Uvalde	-
De Donno	subsp. *pauca*	*Olea europaea*	2014	Italy, Apulia, Lecce, Galipolli - Li Sauli	CP020870.1
XYL1961	subsp. *pauca*	*Olea europaea*	-	Spain, Ibiza	To be published
Ann-1	subsp. *sandyi*	*Nerium oleander*	-	USA, California, Palm Springs	AAAM00000000.4
MED PRI 047	subsp. *sandyi*	*Nerium oleander*	-	USA, Texas	-

Mutant	**Wild type strain**	**Description**	**Reference/source**
NS1::CmR	Temecula1	Chloramphenicol (Cm) cassette inserted between PD0702 and PD0703, neutral site 1 (NS1)	(Matsumoto et al., 2009)
Temecula1-GFP	Temecula1	Km cassette and GFP gene inserted in a different neutral site	(Newman Karyn et al., 2003)
NS1::CmR/Km-WM1-1	WM1-1	Km/Cm cassette inserted between PD0702 and PD0703, neutral site 1 (NS1)	(Kandel et al., 2017)

NS1::CmR/Km -EB92-1	EB92-1	Km/Cm cassette inserted between PD0702 and PD0703, neutral site 1 (NS1)	(Kandel et al., 2017)
NS1::CmR/Km -AlmaEM3	AlmaEM3	Km/Cm cassette inserted between PD0702 and PD0703, neutral site 1 (NS1)	(Kandel et al., 2017)

Plasmid	**Plasmid Background**	**Description**	**Reference/source**
pKLN61	pGEM-5Zf (+)	Plasmid containing *X. fastidiosa oriC*, and TN903 Km cassette and *rpfF* gene (PD0407)	(Newman et al., 2004)

Supplementary Table 2-2: Ancestral recombination events in the core genome of 153 *X. fastidiosa* strains predicted by fastGEAR. Lineage 1 and 2 are donor or recipient lineages. The direction of ancestral recombination between donor and recipient is not distinguished. The length of recombination regions is the sum of all recombinant fragments.

Number of ancestral recombination events	Length of ancestral recombination region	Lineage1	Lineage2
64	54 kb	*fastidiosa*	*multiplex*
37	153 kb	*fastidiosa*	*morus/sandyi*
12	8.1 kb	*multiplex*	*morus/sandyi*
3	1.09 kb	*pauca*	*multiplex*
2	559 bp	*pauca*	*morus/sandyi*
1	141 bp	*pauca*	*fastidiosa*

Supplementary Table 2-3: Number and length of recent recombination events of 153 *X. fastidiosa* strains analyzed using fastGEAR

Strain	**Subspecies**	**Number of recombination events**	**Length of recombinant regions (bp)**
Ann-1	*morus/sandyi*	51	100109
MUL0034	*morus/sandyi*	36	84785
Mul-MD	*morus/sandyi*	36	82978
CO33	*morus/sandyi*	19	28326
CFBP8073	*morus/sandyi*	15	57636
CFBP8356	*morus/sandyi*	15	23344
XYL1961	*pauca*	176	82464
CFBP8072	*pauca*	136	58568
PD7211	*pauca*	54	18818
Hib4	*pauca*	35	11945
Fb7	*pauca*	34	14589

3124	*pauca*	21	6885
32	*pauca*	21	6924
COF0324	*pauca*	20	5482
Pr8x	*pauca*	20	5532
6c	*pauca*	19	5367
9a5c	*pauca*	17	6041
U24D	*pauca*	16	6067
COF0407	*pauca*	15	8753
CoDiRO	*pauca*	15	8763
DeDonno	*pauca*	15	8763
OLS0478	*pauca*	15	8803
OLS0479	*pauca*	15	8836
Salento-1	*pauca*	15	8764
Salento-2	*pauca*	15	8776
PD7202	*pauca*	14	7226
11399	*pauca*	3	2267
B111	*pauca*	3	2267
CVC0251	*pauca*	3	2267
CVC0256	*pauca*	3	2267
J1a12	*pauca*	3	2271
XRB	*pauca*	3	2261
BBI64	*multiplex*	91	73772
BB08-1	*multiplex*	88	78775
BB01	*multiplex*	87	74583
AlmaEM3	*multiplex*	84	73385
XF26_VALVAL072_TX	*multiplex*	80	72169
ATCC35871	*multiplex*	59	44449
Sy-VA	*multiplex*	49	40266
XF21_RBCF119_South-USA	*multiplex*	48	43813
XF29_4rd+1_GA	*multiplex*	47	39667
TOS14	*multiplex*	5	2973
TOS4	*multiplex*	5	2973
TOS5	*multiplex*	5	2909
CFBP8416	*multiplex*	3	144
Griffin-1	*multiplex*	2	136
XF27_LLAFAL718A_TX	*multiplex*	2	668
CFBP8417	*multiplex*	1	71
IVIA6902	*multiplex*	1	64
M12	*multiplex*	1	113
XF25_L95-1_FL	*multiplex*	1	17
XF28_GILGRA274ext_TX	*multiplex*	1	651

XF41_GAplum_GA	*multiplex*	1	51
CFBP8078	*multiplex*	0	0
CFBP8418	*multiplex*	0	0
Dixon	*multiplex*	0	0
ESVL	*multiplex*	0	0
Fillmore	*multiplex*	0	0
IAS-AXF-235T10	*multiplex*	0	0
IAS-AXF212H7	*multiplex*	0	0
IAS64H-11	*multiplex*	0	0
IAS64T-13	*multiplex*	0	0
IVIA5901	*multiplex*	0	0
IVIA6586-2	*multiplex*	0	0
IVIA6629	*multiplex*	0	0
IVIA6731	*multiplex*	0	0
IVIA6903	*multiplex*	0	0
LM10	*multiplex*	0	0
RAAR14_plum327	*multiplex*	0	0
RAAR6_Butte	*multiplex*	0	0
RH1	*multiplex*	0	0
XF-3960-18	*multiplex*	0	0
XF3348	*multiplex*	0	0
XYL1752	*multiplex*	0	0
XYL1966-18	*multiplex*	0	0
XYL1968-18	*multiplex*	0	0
XYL1981-18	*multiplex*	0	0
XYL1981	*multiplex*	0	0
XYL466	*multiplex*	0	0
XYL468	*multiplex*	0	0
16M3	*fastidiosa*	13	7849
CCPM1	*fastidiosa*	13	7849
GV156	*fastidiosa*	8	4812
EB92.1	*fastidiosa*	7	8022
14B3	*fastidiosa*	6	5146
15B2	*fastidiosa*	6	5146
15M1	*fastidiosa*	6	5146
16B4	*fastidiosa*	6	5146
16M5	*fastidiosa*	6	5186
16M7	*fastidiosa*	6	5229
Mus-1	*fastidiosa*	6	4779
WM1-1	*fastidiosa*	6	5322
14B2	*fastidiosa*	5	4587

15B3	*fastidiosa*	5	4587
ATCC35879	*fastidiosa*	5	4670
CFBP7969	*fastidiosa*	5	4670
CFBP7970	*fastidiosa*	5	4670
CFBP8082	*fastidiosa*	5	4670
DSM10026	*fastidiosa*	5	4670
GV230	*fastidiosa*	5	4763
NOB1	*fastidiosa*	5	4763
OK3	*fastidiosa*	5	4763
TPD3	*fastidiosa*	5	4763
TPD4	*fastidiosa*	5	4763
VB11	*fastidiosa*	5	4763
Je13_XF1_89_CA	*fastidiosa*	4	2669
Je65_XF2_45_CA	*fastidiosa*	4	2538
CFBP8351	*fastidiosa*	3	2930
Je17_XF1_93_CA	*fastidiosa*	3	2033
Je73_XF2_53_CA	*fastidiosa*	3	2033
Je9_XF1_85_CA	*fastidiosa*	3	2033
Je92_XF2_72_CA	*fastidiosa*	2	1512
RAAR4_SLO_CA	*fastidiosa*	2	1976
Bakersfield-1	*fastidiosa*	1	1131
CFBP8071	*fastidiosa*	1	1131
GB514	*fastidiosa*	1	1131
IVIA5235	*fastidiosa*	1	1131
IVIA5770	*fastidiosa*	1	1131
Je104_XF2_84_CA	*fastidiosa*	1	1131
Je114_XF2_94_CA	*fastidiosa*	1	1131
Je118_XF3_B1_CA	*fastidiosa*	1	1131
Je1_XF1_77_CA	*fastidiosa*	1	1131
Je35_XF2_15_CA	*fastidiosa*	1	1131
Je43_XF2_23_CA	*fastidiosa*	1	1131
Je4_XF1_80_CA	*fastidiosa*	1	1131
Je51_XF2_31_CA	*fastidiosa*	1	1131
Je60_XF2_40_CA	*fastidiosa*	1	1131
Je63_XF2_43_CA	*fastidiosa*	1	1131
Je72_XF2_52_CA	*fastidiosa*	1	1131
Je76_XF2_56_CA	*fastidiosa*	1	1131
Je78_XF2_58_CA	*fastidiosa*	1	1131
Je84_XF2_64_CA	*fastidiosa*	1	1131
Je97_XF2_77_CA	*fastidiosa*	1	1131
M23	*fastidiosa*	1	1131

RAAR7_Conn-creek_CA	*fastidiosa*	1	1131
StagsLeap	*fastidiosa*	1	1131
Temecula1	*fastidiosa*	1	1131
Temecula1Star	*fastidiosa*	1	1131
TemeculaL	*fastidiosa*	1	1131
XYL1732	*fastidiosa*	1	1131
XYL1978-18	*fastidiosa*	1	1131
XYL1980-18	*fastidiosa*	1	1131
XYL2014-18	*fastidiosa*	1	1131
XYL2017-18	*fastidiosa*	1	1131
XYL2055	*fastidiosa*	1	1131
XYL2093-18	*fastidiosa*	1	1131
XYL2106-18	*fastidiosa*	1	1131
XYL2107-18	*fastidiosa*	1	1131
XYL2153-18	*fastidiosa*	1	1131
XYL2177-18	*fastidiosa*	1	1131
XYL2400-18	*fastidiosa*	1	1131
XYL2508-18	*fastidiosa*	1	1131
XYL461	*fastidiosa*	1	1131

Supplementary Table 2-4: Top 10 genes significantly associated with natural competence from Pyseer results using SNP-based method on 48 strains

Gene	**POS**	**ref**	**alt**	**af**	**filter-pvalue**	**lrt-pvalue**	**beta**	**beta-std-err**	**variant_h2**	**notes**	**Variant_type**
PD1306	1533594	A	G	0.0625	0.228	5.61E-37	0.996	0.0254	0.985	bad-chisq	synonymous_variant
PD0405	495932	T	C	0.292	4.28E-11	7.81E-37	-0.992	0.0255	0.985	bad-chisq	upstream_gene_variant
PD1940	2281939	A	G	0.125	0.289	2.57E-24	0.947	0.0474	0.947	bad-chisq	synonymous_variant
PD0405	495933	A	G	0.333	3.95E-11	1.63E-18	-0.993	0.0696	0.903	bad-chisq	upstream_gene_variant
PD0731	898743	T	G	0.333	3.95E-11	1.63E-18	-0.993	0.0696	0.903	bad-chisq	synonymous_variant

PD1754	2048875	T	C	0.333	3.95E-11	1.63E-18	-0.993	0.0696	0.903	bad-chisq	upstream_gene_variant
PD1755	2049446	G	A	0.333	3.95E-11	1.63E-18	-0.993	0.0696	0.903	bad-chisq	upstream_gene_variant
PD1755	2049464	A	C	0.333	3.95E-11	1.63E-18	-0.993	0.0696	0.903	bad-chisq	upstream_gene_variant
PD0731	898707	G	A	0.292	3.44E-09	6.53E-18	-0.985	0.0716	0.897	bad-chisq	synonymous_variant
PD1755	2049150	T	G	0.292	3.44E-09	6.53E-18	-0.985	0.0716	0.897	bad-chisq	upstream_gene_variant
PD1755	2049296	G	A	0.292	3.44E-09	6.53E-18	-0.985	0.0716	0.897	bad-chisq	upstream_gene_variant

Supplementary Table 2-5: Top 10 genes significantly associated with natural competence from Pyseer results using k-mer-based method on 48 strains

gene	**hits**	**maxp**	**avg_af**	**avg_maf**	**avg_beta**
PD_1724	471	45.80134291	0.58173673	0.41826327	0.874390658
PD_1131	471	45.80134291	0.58173673	0.41826327	0.874390658
PD_0815/PD_0816	2	36.25103714	0.938	0.062	0.996
ligA	174	36.25103714	0.772706897	0.227293103	0.943068966
PD_1703	161	36.25103714	0.273242236	0.273242236	0.660583851
PD_1306	191	36.25103714	0.588068063	0.411931937	0.916005236
PD_0978/PD_0979	211	36.10734897	0.431483412	0.431483412	0.730189573
J (PD_0366)	1026	36.10734897	0.638974659	0.361025341	0.934892788
PD_0713	3	36.10734897	0.868	0.132	0.880333333
PD_0556	10	36.10734897	0.9002	0.0998	0.8714

Supplementary Table 2-6: Top 10 genes significantly associated with natural competence from DBGWAS results on 48 strains

Locus tag	Annotation
XF_2522/XF_0686	Hypothetical phage-related genes
XF_2724	Site-specific DNA-methyltransferase_(adenine-specific)
XF_0357	Uncharacterized protein
cysS	Cysteine—tRNA ligase
XF_2270	Putative_glucose-6-phosphate_1-epimerase
XF_1132	Transcriptional regulator (LysR family)
XF_0930/XF0931/XF0827	Uncharacterized protein
XF_0619	Periplasmic divalent cation tolerance protein
XF_a0012	Conjugal transfer protein
XF2137	Ferric enterobactin receptor

Supplementary Table 2-7: Top 10 genes significantly associated with homologous recombination from DBGWAS results on 153 strains

Locus tag	Annotation
XF_0708	Hypothetical phage holin family protein
XF_1233	Uncharacterized protein
XF_1583	Putative prophage protein
XF_0677	Type 4 fimbriae assembly protein
gloB	Hydroxyacylglutathione hydrolase
nagZ	Beta-hexosaminidase
XF_1873	Homolog of PD0917, PD0930, both of which are annotated as hypothetical phage-related protein
proS	Proline—tRNA ligase
tgt	Queuine tRNA-ribosyltransferase
XF_0292	Aconitate hydratase B

Chapter 3

Construction of a genome-wide mutant library in *Xylella fastidiosa* using RB-TnSeq

Abstract

Xylella fastidiosa is a gram-negative, fastidious, plant pathogenic bacterium. It is naturally competent, which is an ability to naturally uptake free DNA and recombine the DNA with the chromosome. Natural competence in *X. fastidiosa* has been shown as an important trait for its evolution. However, the mechanisms of natural competence in *X. fastidiosa* are largely unknown. To identify genetic elements essential for natural competence in *X. fastidiosa* in a genome-wide level, we employed a transposon insertion sequencing technique termed random-barcoded transposon sequencing (RB-TnSeq), which can efficiently screen for essential genes across the whole genome. Here we succeeded in constructing a mini-mutant library for quality check. Unfortunately, the mini-library turned out to be biased towards the *pilQ* gene, finding a solution to this issue is currently ongoing. Further efforts will be put on building a saturated mutant library and screening essential genetic features for natural competence.

Introduction

Xylella fastidiosa, a gram-negative fastidious bacterium, that inhabits in the xylem vessels of plants and can only be transmitted among hosts by sap-feeding insect vectors. It infects many economically essential crops causing many diseases such as Pierce's Disease (PD) on grapes, Citrus Variegated Chlorosis (CVC), Olive Quick Decline Syndrome (OQDS), leaf scorch diseases on coffee, almond, and many other plant hosts, which result in huge economic loss in agriculture. Host range of *X. fastidiosa* has expanded to 638 plant species from 87 families (European Food Safety et al., 2021). *X. fastidiosa* harbors a relatively small genome size with only ~ 2.5 Mb (Simpson et al., 2000). Although it is the first plant pathogen whose genome

was sequenced in 2000, genomic functions are still largely unknown (Simpson et al., 2000). The poor understanding on its genome is partly resulted from difficulties of genetic engineering in *X. fastidiosa* due to its slow-growing nature and laborious work (Raffini et al., 2020).

Site-directed mutagenesis and randomly inserted mutagenesis are common ways for characterizing gene functions or identifying phenotype-associated genes *in X. fastidiosa* respectively. Site-directed mutagenesis is mutation of a targeted genomic region through homologous recombination (HR) using homologous DNA fragments or plasmids that contain selective markers such as antibiotic markers or GFP expression cassettes. The DNA fragments or DNA vectors can be transformed into *X. fastidiosa* cells for mutagenesis via electroporation and a much easier method that was recently developed by our group, natural competence by which the cells will naturally uptake foreign DNA (da Silva Neto et al., 2002; Kandel et al., 2018). Randomly inserted mutagenesis has also been applied to *X. fastidiosa* through delivering transposon plasmids, which randomly insert DNA fragments into chromosome via transposases. Transposon constructs are delivered into cells via electroporation, generating a large number of mutants selected under certain conditions related to studied phenotypes (Guilhabert and Kirkpatrick, 2005; Li et al., 2007). Although these two approaches succeeded in identifying various functions of different genes, it is still challenging to identify genetic factors associated with a certain phenotype on a genome-wide scale, and to assess the phenotype with a whole mutant library (Newman et al., 2004; Guilhabert and Kirkpatrick, 2005; Li et al., 2007; Cruz et al., 2014; Kandel et al., 2018). Even with randomly inserted mutagenesis, generating a full mutant library containing mutants of almost every gene in *X. fastidiosa* has not been documented. With technical advances, high-throughput sequencing allows us to screen an unbiased mutant library at once, substantially increasing efficiency of assessing phenotypes of the mutant library. One of the methods to rapidly find genes of interest is termed Transposon Insertion Sequencing (TIS), which is an approach identifying functional

essential genes from pooled mutant libraries given certain conditions. Combining advantages of transposon mutagenesis with high-throughput sequencing, TIS is able to profile fitness reflected by relative abundance of mutants in competitive assays of genetic elements on the genome-wide level (Cain et al., 2020). One of the TIS methods is termed Random-Barcoded Transposon Sequencing (RB-TnSeq), which was developed by incorporating barcodes into transposon sequencing. Adding barcodes to each mutant substantially reduces laborious work of preparing sequencing libraries because once the mutants are mapped, only DNA barcodes need to be sequenced in following fitness assays as each barcode represents a specific mutant, which is easier than re-sequencing the whole library. Therefore, RB-TnSeq facilitates multiple experiments applied for identifying essential genes responsible for different phenotypes. In the fitness assays, relative abundance of DNA barcodes in the library would be compared in different time points of competitive assays, which is used to indicate fitness in the population (Wetmore et al., 2015).

RB-TnSeq has been applied to many microorganisms and discovered substantial genes essential for various biological processes (Liu et al., 2018). In particular, our research interest is genetic mechanisms of natural competence in *X. fastidiosa*. As mentioned above, natural competence is an ability of the cells to naturally uptake extracellular DNA and further incorporate the DNA into chromosome via HR, which possibly plays an important role in evolution of *X. fastidiosa* (Kandel et al., 2017a; Potnis et al., 2019). Natural competence is one of the mechanisms of Horizontal Gene Transfer (HGT) that also occurs through two other classic mechanisms: conjugation and transduction, both of which may result in genetic exchanges by HR (Sheppard et al., 2018; Sun, 2018; Straub et al., 2020). Conjugation is when plasmid DNA transfers via direct contact of donor and recipient cells, which is also utilized as an approach of genetic manipulation for many bacteria (Schröder and Lanka, 2005). Transfer of mobile plasmids has been found between different subspecies of *X. fastidiosa* strains in vitro

via conjugation dependent on the *tra* and *trb* operon that encode conjugative Type IV secretion system, while the plasmid transfer was not observed between conjugative *Escherichia coli* donor and *X. fastidiosa* recipient cells (Burbank and Van Horn, 2017).

Although there are substantial studies exploring genetic mechanisms of natural competence, most of them focused on bacteria infecting human or animals, while only a few investigations studied natural competence in plant pathogenic bacteria such as *X. fastidiosa* or *Ralstonia solanacearum*. RB-TnSeq has been used to identify essential genes for natural competence in cyanobacteria (Taton et al., 2020). So far, only a few genetic factors have been demonstrated to affect natural competence in *X. fastidiosa*, which can be used for confirmation of the RB-TnSeq results. One of these factors are Restriction-Modification (R-M) systems, which were also applied to genetic manipulation in *X. fastidiosa* by transforming methylated DNA into cells (Matsumoto and Igo Michele, 2010; Kung and Almeida, 2011; Kandel et al., 2018). When donor DNA was methylated by methylase of *X. fastidiosa,* the recombination rate of natural competence was significantly increased (Kung and Almeida, 2011). Like other Gram-negative naturally competent bacteria, Type IV pili (TFP) is involved in natural competence in *X. fastidiosa* (Kung and Almeida, 2014; Kandel et al., 2017b; Kandel et al., 2018). A recent study exploring the association of TFP with natural competence in *X. fastidiosa* showed TFP served as DNA uptake machinery and thus was crucial for natural competence in *X. fastidiosa* (Merfa et al., 2021). Despite these findings, there may exist other genetic elements involved in natural competence especially those involved in regulation and recombination.

Here, we aimed at using RB-TnSeq to screen essential genes for natural competence in *X. fastidiosa* across the whole genome. This is the first time that RB-TnSeq is applied to a fastidious prokaryote such as *X. fastidiosa*. Such application may further facilitate research about discovering other functional essential genes in *X. fastidiosa* and applying RB-TnSeq on other bacteria that are not model species. After much troubleshooting, a mini library was

constructed for testing quality and feasibility of the method. Unfortunately, the mini library was biased towards the *pilQ* gene, a finding a solution to this issue is currently ongoing. Further efforts will be put on building a saturated mutant library and screening essential genetic features for natural competence. This chapter will mainly focus on descripting different methods that we have attempted for constructing the mutant library.

Materials and methods

1. Bacterial strains, media, and culture conditions

Different subspecies of *X. fastidiosa* strains TemeculaL (subsp. *fastidiosa*), WM1-1 (subsp. *fastidiosa*), AlmaEM3 (subsp. *multiplex*) were used for creating the mutant library. All *X. fastidiosa* strains were grown on PW (Davis et al., 1980) agar plates modified by omitting phenol red and using 1.8 g l^{-1} of bovine serum albumin (Gibco Life Sciences Technology) at 28°C for seven days and re-streaked onto new PW plates for another week. PD3 medium (Davis et al., 1981) was used to culture *X. fastidiosa* cells in liquid and select mutants on agar plates. Two *E. coli* strains APA752 and APA766 that contained transposon plasmids were used as donor (Wetmore et al., 2015). LB (Luria Bertani) liquid medium was used to culture *E. coli* cells at 37°C.

2. Construction of a mutant library

We used two different randomly bar-coded transposon delivery vector plasmids that encode kanamycin (Km) resistance cassette and transposases: the pKMW3 mariner transposon vector library harbored by *E. coli* strain APA752, the pKMW7 Tn5 transposon vector library harbored by *E. coli* strain APA766. The two types of transposons differed in that the mariner-based transposons (pKMW3) target only TA dinucleotides, while Tn5-based transposons (pKMW7)

target random sites but preferentially insert in high GC regions. Mariner-based transposons have been shown to generate saturated mutant libraries more easily for many bacterial species, while Tn5-based transposons were also preferred when mariner-based transposons are not suitable for certain species (Chao et al., 2016).

2.1. Transfer of transposon plasmids via natural competence

Since transformation via natural competence is easy to perform and has been established in our lab, we first carried out natural competence assay with three strains: TemeculaL, WM1-1, and AlmaEM3. Different conditions tested were listed in Table 1 and 2.

Table 1. Experimental conditions tested using plasmids via natural competence

Assay description	**Recipient OD_{600}**	**Recipient volume (µl)**	**Plasmid quantity [a] (ng)**	**Kanamycin concentration (µg/ml)**	**Incubation days before selection**
Basic protocol	0.5, 0.25	10	200, 500, 1000	30	5
Methylation of plasmids [b]	0.3	20	1500-2000 ng Methylated pKMW3	30	5
Repeated transformation	0.3	20, 50	Addition of 200 ng or 1000 ng plasmids per 24h or 12h	30	7
Scale up	0.3	50	3500, 5000	10, 30	7

[a] If not mention, both pKWM3 and pKWM7 were tested in the experiments.

[b] Methylation was transforming plasmids into *E. coli* strain EAM1 that expresses methylase of *X. fastidiosa* and then extracting methylated plasmids from the transformants of EAM1.

Table 2. Experimental conditions tested using DNA from dead donor cells via natural competence

Recipient strain and OD_{600}	**Donor OD_{600}**	**Recipient: Donor volume ratio**	**Total cell volume (µl)**	**Incubation days**	**Kanamycin concentration (µg / mL)**
TemeculaL, 2.5	1.0	1:1, 2:5, 3:1	100	7	10, 30
TemeculaL, WM1-1, 5.0	2.5	1:1, 3:1, 1:3	100	7	10, 30

Procedure

1. Use plasmids as donor: extract plasmids from the cultures of APA752 and APA766 using an GeneJet Plasmid Miniprep kit (Thermo Scientific), adjust concentration to 100 ng/µl quantified using a Cytation 3 Image Reader spectrophotometer (BioTek Instruments Inc.).

   Use a large quantity of plasmids released from dead donor cells (APA752 and APA766) by either heat-kill at 90°C for 15 min or simply removing Diaminopimelic Acid (DAP) from LB medium by washing cells with PD3 broth twice (donor cells are DAP-auxotroph). Dead cells were confirmed by plating 100 µl of cell suspension onto LB plates.
2. Collect recipient cells from PW plates, suspend in 1 ml PD3 broth, adjust OD_{600}. Spot certain volume of cell suspension onto PD3 plates, add certain quantity of plasmids on the top of the cells. incubate cultures at 28°C for 5~7 days.
3. Harvest the cell spots to 1 ml of PD3 broth, dilute in 10-fold serial dilutions, plate them in PW plates amended with or without Km, incubate for one month at 28°C.

2.2. Transfer of transposon plasmids via conjugation

We try to use conjugation since the transposon plasmids have been successfully delivered to many other bacterial species via conjugation and conjugation is also easy to carry out. Additionally, *X. fastidiosa* strain Temecula1 can be transferred by other *X. fastidiosa* strains that carry conjugative plasmids via conjugation, though the attempt to conjugate *X. fastidiosa* using *E. coli* cells failed (Burbank and Van Horn, 2017). We tested both *E. coli* donor strains, APA752 and APA766, for different conditions listed in Table 3.

Table 3. Experimental conditions tested for delivering plasmids via conjugation

Recipient OD_{600}	**Donor OD_{600}**	**Recipient: Donor volume ratio**	**Total cell volume (µl)**	**Incubation Time (days)**	**Km concentration (µg / mL)**
TemeculaL, WM1-1, 5.0	2.5	1:1, 3:1, 1:3	100	5	10, 30
WM1-1, 2.5	1.0	1:1, 2:5, 3:1	100	1, 2, 5	10, 30

Procedure

1. Grow recipient *X. fastidiosa* strains for 5 days at 28°C on PD3 plates, harvest cells with PD3 broth, adjust OD_{600}. Grow donor strain APA752/APA766 in LB liquid medium amended with 50 µg/ml Km and 300 µM DAP until OD_{600} reaches 0.5, concentrate the cells and replace the LB medium with PD3 broth amended with DAP.
2. Pellet 1 ml donor culture for 1 min at 10,000 rpm, wash cells twice with 1 ml PD3 liquid medium to remove antibiotics from LB medium, finally dissolve each pellet in PD3 broth amended with DAP.
3. Combine donor and recipient cells at different ratios based on volume, mix by pipetting. Spot 100 µl mixture on 0.45-µm nitrocellulose filters (Millipore) overlaid on PD3 plates

supplemented with DAP, incubate for several days. Spot 100 µl donor or recipient cells separately on PD3 plates supplemented with DAP as control.

4. Harvest cells in PD3 broth, plate on PW plate with Km, incubate at 28 °C for one month.

2.3. Transfer of transposon plasmids via electroporation

Three strains: TemeculaL, WM1-1, and AlmaEM3 were tested under different conditions listed in Table 4.

Table 4. Experimental conditions tested for delivering plasmids via electroporation

Plasmid	**Plasmid quantity (ng)**	**Voltage (kV)**	**Competent cell quantity (µl)**
pKMW3	1000	2.5	100
	150, 500, 1000	3.0	
pKMW7	250, 500	3.0	50
	750, 1000, 1400	2.5	
pKMW7	1000, 2000	2.75	100
	150, 1000, 2000	3.0	

Procedure

1. Prepare *X. fastidiosa* electrocompetent cells: wash recipient cells 2-4 times with 10% iced glycerol and adjust OD_{600} to 2.5 (~10^9 cells ml-1).
2. Add plasmids into washed cells in 1.5 ml tubes, transfer the mixture to 2 mm cuvette, electroporate the plasmids into electrocompetent cells at 2.5 kV-3.0 kV, 200 Ω and 25 mF for 5 ms.
3. Add 1 ml PD3 broth to the transformed cells immediately, recover cells at 28°C with shaking at 150 rpm for 24h.

4. Concentrate the cell suspension for 5 min at 4000 rpm. Plate all cells on PW plates containing Km for selection and PW plates without Km as control. Incubate for one month at 28°C.

2.4. Transfer of a bar-coded EZ:Tn5 transpososome via electroporation

Another approach to generate a mutant library is to transfer mixture of transposases together with transposons, which is also termed as transpososome into recipient cells so that the recipient cells do not need to express transposases in vivo. We used TemeculaL as the recipient strain for the experiments.

2.4.1. Construction of a transpososome

1. To construct a bar-coded transposon DNA, clone the Km cassette from plasmid pKMW3 using primers (Table 5) that contain Tn5-specific Inverted Repeat (IR) and a random 20-bp DNA bar code flanked by the common BarSeq PCR priming sites U1 and U2 (Wetmore et al., 2015).
2. Digest the plasmid template with DpnI restriction enzyme for 1 h at 37°C (New England Biolabs). Methylation of transposons used MsssI enzyme that was treated for 1 h at 37°C (New England Biolabs).
3. Prepare the transpososome reaction mixture by adding transposons with EZ:Tn5 transposases (Epicentre) in the following order, mix by vortex, incubate for 30 min at room temperature, store at -20°C:

   2 µl 100 µg/ml EZ-Tn5 Transposon DNA in TE Buffer

   4 µl EZ-Tn5 Transposase

   2 µl 100% glycerol

Table 5. Primers used in this study

Name	Sequence (5'-3')	Purpose	Notes
RB-TN transposon-F	**CTGTCTCTTATACACA TCTcttgcagtgggcttacatg**	Amplify kanamycin cassette from transposon plasmid pKMW3	the sequence in bold is the Tn5 IR
RB-TN transposon-R	**CTGTCTCTTATACACA TCT***GTCGACCTGCAGCG TACG*NNNNNNNNNNNN NNNNNNNN*AGAGACCTC GTGGACATC*taatgctctgccagt gtcg		the sequence in bold is the Tn5 IR; the sequences in italics are the common priming sites for BarSeq

2.4.2. Electroporation with the transpososome

1. Electroporate a transpososome as described above using different conditions listed in Table 6.
2. Incubate 13-15 days till single colony shows up. Collect mutants by adding PD3 broth and then scraping cells from plates into the media. Dilute cell suspension with PD3 broth to OD_{600} of 0.2 and grow mutants with shaking at 150 rpm for 5 days.
3. Concentrate cell suspension to OD_{600} of 2.0-3.0, add 40% glycerol to make stocks.

Table 6. Experimental conditions tested for delivering a transpososome via electroporation

Assay description	Treatments	Km concentration (μg/ml)
Quantity of cells and transpososomes	CK [a]	5, 30, 50
	100 μl cell+1 μl [b]	5, 30, 50
	50 μl cells+1 μl [b]	5, 30, 50
	100 μl cells+2 μl [b]	5, 30, 50
	CK-2.5kV/2.75kV/3.0kV	5, 30, 50
Voltage [c]	2.5kV	5, 30, 50
	2.75kV	5, 30, 50
	3.0kV	5, 30, 50
Methylation and addition of inhibitor [d]	Transpososome	5
	Methylated transpososome	5
	Transpososome + 1 μl Inhibitor [e]	5
	Methylated transpososome + 1 μl Inhibitor [e]	5

[a] CK is the control that the competent cells were electroporated without adding a transpososome.

[b] The treatment is 100 µl competent cells plus 1 µl transpososome are electroporated under condition of 2500 V. The other two treatments are expressed in the same way.
[c] Different voltages were tested using 50 µl electrocompetent cells added with 1 µl transpososome.
[d] Methylation and addition of inhibitor were tested using 50 µl electrocompetent cells added with 1 µl transpososome
[e] Inhibitor is TypeOne™ Restriction Inhibitor (Lucigen), a phage protein that acts as a molecular decoy blocking the DNA binding site of type I R-M enzymes and inhibiting cleavage of unmodified DNA. Thus, the inhibitor can increase transformation efficiency by inhibiting activity of R-M systems.

2.4.3. Construction of transposons using modified PCR reaction and program

Since the mini-library constructed using the method of electroporating a transpososome described above was biased, we used a different PCR reaction and program shown in Table 7 and 8 with Q5 high-fidelity DNA polymerase (New England Biolabs) to generate a new EZ:Tn5 transpososome with the new amplified transposon DNA.

Component	Reaction	Final concentration
5X Q5 Reaction Buffer	10 µl	1X
5X Q5 High GC Enhancer	10 µl	1X
10 mM dNTPs	1 µl	200 µM
10 µM Forward Primer (RB-TN transposon-F)	2.5 µl	0.5 µM
10 µM Reverse Primer (RB-TN transposon-R)	2.5 µl	0.5 µM
Template DNA (pKMW3 purified plasmid)	10 ng	< 1,000 ng
Q5 High-Fidelity DNA Polymerase	0.5 µl	0.02 U/µl
DSMO	1.5 µl	
Nuclease-Free Water	to 50 µl	

Table 7. Modified PCR reaction system

Table 8. Modified PCR program

Step	Temp	Time
Initial Denaturation	98°C	30 seconds
6 x Cycles	98°C 58°C 72°C	5-10 seconds 10–30 seconds 90 seconds (20–30 seconds/kb)
Final Extension	72°C	2 minutes
Hold	4–10°C	

2.5. Genomic DNA extraction for TnSeq

1. Grow the mini mutant library in PD3 broth supplemented with 30 μg/ml Km at a start OD_{600} of 0.1 for a week to obtain sufficient cells for genomic DNA extraction.
2. Genomic DNA was extracted using a modified CTAB protocol (cetyltrimethylammonium bromide) (Doyle and Doyle) and further sent for whole-genome sequencing.

Results and Discussion

First, attempts of constructing a mutant library focused on delivering the whole transposon plasmids into *X. fastidiosa* cells via different methods of conjugation, natural competence, and electroporation. However, none of these attempts produced positive transformants. We hypothesized that *X. fastidiosa* may not be able to synthesize the transposase encoded in the transposon plasmid and thus mutants cannot be produced without the transposases. Then we used a transpososome to overcome the possible barrier of expressing transposases in vivo. We defined the optimal conditions for electroporation, which was using 50 μl electrocompetent cells added with 1 μl transpososome and 1 μl TypeOne™ restriction inhibitor under condition of 2.5 kV, 200 Ω and 25 mF. By using this condition, we obtained a mini library with around 10,000 colonies. Sequencing the library generated a total of 935967 reads, 95% of which contains barcodes. Estimated number of barcodes (may be inflated for sequencing error) was 10831, of which 6163 barcodes were distinct. We found 1454 transposon insertions in genome at 1251 different locations, of which 768 insertions (644 distinct locations) exist in central 10-90% of genes. Among those central insertions, 273 were in protein-coding genes where there are a total of 2078 coding genes. The percentage of insertions that the disrupted gene and

transposon were on same strand was 48.6%, indicative of no strand bias. However, the mini-library was biased to *pilQ* gene, whose amino acid product is a type IV pilus secretin channel, as around half of the reads were in mapped to this gene. It is likely the conditions that we used to select for the mutants led to strong positive selection for the loss of *pilQ*. However, we compared growth rate of *pilQ*-deleted mutant with wildtype in different liquid and solid media (PW, PD2, PD3) and found no obvious differences of the growth rate between different types of media (data not shown). Then we hypothesized this bias may be due to unusual barcodes with too many A/T nucleotides in the *pilQ* mutant. Therefore, we used a modified PCR reaction and program especially with much lower (6) amplification cycles and with GC enchancer, which would ideally produce a randomly bar-coded transposon with even distribution of A/T and G/C nucleotides. By using the new transposon, we created a new mutant mini-library with ~ 2,000 colonies. Currently, the new mini-library is in the process of sequencing.

Future work

1. If the new library is not biased to some gene (s), we will construct a saturated mutant library that ideally covers most genes and has at least 10 central insertions per gene for statistical power by using the modified transposon and repeated electroporation assays.
2. Once the saturated mutant library is obtained, we will carry out fitness assay for identifying genes essential for natural competence. Generally, one aliquot of the library would be thawed from -80°C and grow for 5-7 days. Genomic DNA of the library before natural competence assay would be sequenced as the population baseline. Then plasmid DNA that carry chloramphenicol expression cassette would be added to the library culture and incubate for 3-5 days. Next the culture would be plated on selective plates supplemented with both Km and chloramphenicol. Once the colonies appeared on plates, transformants would be collected, amplified and sequenced to compare with

the population baseline. Further analysis would be performed by our collaborator A. Deutschbauer (UC Berkeley).

References

Burbank, L.P., and Van Horn, C.R. 2017. Conjugative plasmid transfer in *Xylella fastidiosa* is dependent on *tra* and *trb* operon functions. J Bacteriol 199:e00388-00317.

Cain, A.K., Barquist, L., Goodman, A.L., Paulsen, I.T., Parkhill, J., and Opijnen, T.v. 2020. A decade of advances in transposon-insertion sequencing. Nature Reviews Genetics 21:526-540.

Chao, M.C., Abel, S., Davis, B.M., and Waldor, M.K. 2016. The design and analysis of transposon-insertion sequencing experiments. Nat. Rev. Microbiol. 14:119-128.

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

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.

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.

Doyle, J.J., and Doyle, J.L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue.

European Food Safety, A., Delbianco, A., Gibin, D., Pasinato, L., and Morelli, M. 2021. Update of the *Xylella* spp. host plant database - systematic literature search up to 31 December 2020. EFSA J 19:e06674.

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. MPMI 18:856-868.

Kandel, P.P., Chen, H., and De La Fuente, L. 2018. A short protocol for gene knockout and complementation in *Xylella fastidiosa* shows that one of the type IV pilin paralogs (PD1926) is needed for twitching while another (PD1924) affects pilus number and location. Appl. Environ. Microbiol. 84:e01167-01118.

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

Kandel, P.P., Almeida, R.P.P., Cobine, P.A., and De La Fuente, L. 2017b. Natural Competence Rates Are Variable Among Xylella fastidiosa Strains and Homologous Recombination Occurs In Vitro Between Subspecies fastidiosa and multiplex. MPMI 30:589-600.

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

Li, Y., Hao, G., Galvani, C.D., Meng, Y., Fuente, L.D.L., Hoch, H.C., 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., Price Morgan, N., Waters Robert, J., Ray, J., Carlson Hans, K., Lamson Jacob, S., Chakraborty, R., Arkin Adam, P., Deutschbauer Adam, M., and Typas, N. 2018. Magic pools: parallel assessment of transposon delivery vectors in bacteria. mSystems 3:e00143-00117.

Matsumoto, A., and Igo Michele, M. 2010. Species-Specific Type II Restriction-Modification System of Xylella fastidiosa Temecula1. Applied and Environmental Microbiology 76:4092-4095.

Merfa, M.V., Zhu, X., Shantharaj, D., Gomez, L.M., Naranjo, E., Potnis, N., Cobineb, P.A., and Fuente, L.D.L. (2021). Movement and evolution: complete functional analysis of type IV pilus of a re-emergent plant pathogen reveals a unique DNA receptor. In unpublished.

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. U.S.A. 101:1737-1742.

Potnis, N., Kandel, P.P., Merfa, M.V., Retchless, A.C., Parker, J.K., Stenger, D.C., Almeida, R.P.P., Bergsma-Vlami, M., Westenberg, M., Cobine, P.A., and Fuente, L.D.L. 2019. Patterns of inter- and intrasubspecific homologous recombination inform eco-evolutionary dynamics of *Xylella fastidiosa*. ISME J 13:2319-2333.

Raffini, F., Bertorelle, G., Biello, R., D'Urso, G., Russo, D., and Bosso, L. 2020. From nucleotides to satellite imagery: approaches to identify and manage the invasive pathogen *Xylella fastidiosa* and its insect vectors in Europe. Sustainability 12:4508.

Schröder, G., and Lanka, E. 2005. The mating pair formation system of conjugative plasmids—a versatile secretion machinery for transfer of proteins and DNA. Plasmid 54:1-25.

Sheppard, S.K., Guttman, D.S., and Fitzgerald, J.R. 2018. Population genomics of bacterial host adaptation. Nat. Rev. Genet. 19:549-565.

Simpson, A.J.G., Reinach, F.C., Arruda, P., Abreu, F.A., Acencio, M., Alvarenga, R., Alves, L.M.C., Araya, J.E., Baia, G.S., Baptista, C.S., Barros, M.H., Bonaccorsi, E.D., Bordin, S., Bové, J.M., Briones, M.R.S., Bueno, M.R.P., Camargo, A.A., Camargo, L.E.A., Carraro, D.M., Carrer, H., Colauto, N.B., Colombo, C., Costa, F.F., Costa, M.C.R., Costa-Neto, C.M., Coutinho, L.L., Cristofani, M., Dias-Neto, E., Docena, C., El-Dorry, H., Facincani, A.P., Ferreira, A.J.S., Ferreira, V.C.A., Ferro, J.A., Fraga, J.S., França, S.C., Franco, M.C., Frohme, M., Furlan, L.R., Garnier, M., Goldman, G.H., Goldman, M.H.S., Gomes, S.L., Gruber, A., Ho, P.L., Hoheisel, J.D., Junqueira, M.L., Kemper, E.L., Kitajima, J.P., Krieger, J.E., Kuramae, E.E., Laigret, F., Lambais, M.R., Leite, L.C.C., Lemos, E.G.M., Lemos, M.V.F., Lopes, S.A., Lopes, C.R., Machado, J.A., Machado, M.A., Madeira, A.M.B.N., Madeira, H.M.F., Marino, C.L., Marques, M.V., Martins, E.a.L., Martins, E.M.F., Matsukuma, A.Y., Menck, C.F.M., Miracca, E.C., Miyaki, C.Y., Monteiro-Vitorello, C.B., Moon, D.H., Nagai, M.A., Nascimento, A.L.T.O., Netto, L.E.S., Nhani, A., Nobrega, F.G., Nunes, L.R., Oliveira, M.A., de Oliveira, M.C., de Oliveira, R.C., Palmieri, D.A., Paris, A., Peixoto, B.R., Pereira, G.a.G., Pereira, H.A., Pesquero, J.B., Quaggio, R.B., Roberto, P.G., Rodrigues, V., de M. Rosa, A.J., de Rosa, V.E., de Sá, R.G., Santelli, R.V., Sawasaki, H.E., da Silva, A.C.R., da Silva, A.M., da Silva, F.R., Silva, W.A., da Silveira, J.F., Silvestri, M.L.Z., Siqueira, W.J., de Souza, A.A., de Souza, A.P., Terenzi, M.F., Truffi, D., Tsai, S.M., Tsuhako, M.H., Vallada, H., Van Sluys, M.A., Verjovski-Almeida, S., Vettore, A.L., Zago, M.A., Zatz, M., Meidanis, J., and Setubal, J.C. 2000. The genome sequence of the plant pathogen *Xylella fastidiosa*. Nature 406:151-157.

Straub, C., Colombi, E., and McCann, H.C. 2020. Population genomics of bacterial plant pathogens. Phytopathology 111:23-31.

Sun, D. 2018. Pull in and push out: mechanisms of horizontal gene transfer in bacteria. Front. Microbiol. 9:2154.

Taton, A., Erikson, C., Yang, Y., Rubin, B.E., Rifkin, S.A., Golden, J.W., and Golden, S.S. 2020. The circadian clock and darkness control natural competence in cyanobacteria. Nat. Commun. 11:1688.

Wetmore, K.M., Price, M.N., Waters, R.J., Lamson, J.S., He, J., Hoover, C.A., Blow, M.J., Bristow, J., Butland, G., Arkin, A.P., and Deutschbauer, A. 2015. Rapid quantification of mutant fitness in diverse bacteria by sequencing randomly bar-coded transposons. mBio 6:e00306-00315.