Project title:	*Xylella fastidiosa*: what are the factors that make this bacterium pathogenic and host-specific?
Project number:	3114 0052
Project leader:	Prof. Robert W. Jackson, University of Reading and University of Birmingham
Report:	Final report, May 2021
Previous report:	Annual report, October 2019
Key staff:	Dr Michelle Hulin, NIAB EMR Dr Emma Cascant López, NIAB EMR Dr Glyn Barrett, University of Reading Dr Richard Harrison, NIAB
Location of project:	NIAB EMR, East Malling ME19 6BJ
Industry Representative:	Mr Brian Fraser, Oakover Nurseries, Ashford TN26 1AR Mr Jamie Dewhurst, J&A Growers, Warwick CV35 8EB
Date project commenced:	01.10.2018

DISCLAIMER

While the Agriculture and Horticulture Development Board seeks to ensure that the information contained within this document is accurate at the time of printing, no warranty is given in respect thereof and, to the maximum extent permitted by law the Agriculture and Horticulture Development Board accepts no liability for loss, damage or injury howsoever caused (including that caused by negligence) or suffered directly or indirectly in relation to information and opinions contained in or omitted from this document.

AUTHENTICATION

We declare that this work was done under our supervision according to the procedures described herein and that the report represents a true and accurate record of the results obtained.

Michelle Hulin

Research Leader

NIAB EMR

Signature ..... Date .....9/8/21.....

Robert Jackson

Project leader

University of Reading and University of Birmingham

Signature ..... Date..21/7/21....

Report authorised by:

[Name]

[Position]

[Organisation]

Signature ..... Date .....

[Name]

[Position]

[Organisation]

Signature ..... Date .....

CONTENTS

GROWER SUMMARY

Headline

Fundamental knowledge of the bacterium Xylella fastidiosa – detected in over 350 different plant species in Europe alone and causing diseases in many crops – can lead to targeted treatment plans, instead of destroying acres of valuable crops.

Background

What is *Xylella fastidiosa*?

Xylella fastidiosa (*Xf*) bacterium that lives in the foregut of insects and the xylem of plants and causes diseases in several economically significant crops, including Pierce's disease (PD) of grapevine, phony peach disease (PPD), oak leaf scorch (OLS) and olive quick-decline syndrome (OQDS). *Xf* has been detected in over 350 different plant species in Europe alone, but detection of the bacterium in a plant does not necessarily lead to disease. However, these asymptomatic hosts may however act as a reservoir for insect vectors to further spread the bacteria to susceptible plants. Very little is known about the mechanisms behind what makes *Xf* cause symptoms in some plants but not in others. This research project investigates what makes *Xf* host-specific and pathogenic by means of molecular and computational biology. More specifically, the genes that encode effector proteins. Effector proteins are secreted by bacteria and interact with a host plant's immune system, the importance of which is explained below.

Currently, there is no treatment available for diseases caused by *Xf*. Management measures are restricted to vector control, pruning of infected plant tissue and destruction of the infected host. All surrounding potential plant hosts in a 100 m radius are destroyed and a demarcation order of a 5 km radius is set up banning the movement of any plant material outside this area.

Plant hosts of *X. fastidiosa*

There are over 350 different potential host plants in Europe alone, many of which are economically important crops and could devastate a country's economy if affected. The bacterium has not yet been detected in the UK, but the threat is very high as many of its host plants are cultivated here. These include, but are not limited to economically significant plants grown in forest nurseries, as protected crops and ornamental garden plants. Some of the most significant crops are:

> alfalfa, bay, blueberry, *Brassica*, *Cercis* (redbuds), *Chionanthus* (fringe tree), *Cytisus* (broom), elderberry, elm, fig, grapevine, *Hedera* (ivy), *Hypericum* (St. John's Wort), magnolia, maple, mulberry, *Nandina domestica* (sacred bamboo), lavender, oak, olive, pear, *Prunus* (e.g. apricot, cherry, plum), *Rubus* (e.g. raspberries, blackberries), *Rosa*, rosemary, strawberry, *Trifolium* (e.g. clover), walnut, willow.

Xf's large host range, its long incubation period, which may be up to six months in some plants, and rapid spread makes it a highly threatening pathogen. However, in order to come up with effective treatment plans for affected plants, it is important to understand the fundamental biology of the disease-causing bacteria.

How do plant bacteria cause disease?

Phytopathogens, or plant pathogens, have the ability to invade the host, evade host defence mechanisms and ensure disease progression by secreting virulence factors. Virulence factors are proteins, lipids and carbohydrates produced by the pathogen. One of the best-characterised virulence factors include effector proteins, which are secreted through secretion systems or channels in the bacterial cell. Those secreted through type 3 and type 4 secretion

systems (T3SS and T4SS, respectively) are among the most extensively studied in structure and function. The majority of bacterial phytopathogens have been found to secrete effectors through the T3SS, for example *Pseudomonas syringae*, *Erwinia* spp. and *Xanthomonas* spp. (Buttner, and He, 2009). However, *Xf* lacks the T3SS, meaning its strategy to cause disease may be quite different. Effectors may have several functions. In *Xf*, for example, a number of effector proteins are found to be involved in biofilm formation. Biofilm is an adhesive state of bacteria, where they aggregate in clusters. In the case of *Xf*, biofilm formation often leads to the blocking of the xylem, which stops the flow of water and minerals in the vessels and thus disease symptoms appear.

X. fastidiosa is an increasing threat in the European continent

Xf is believed to be native to the Americas and outbreaks of diseases caused by the bacterium within Europe have only been discovered in 2013. The first outbreak of *Xf* in Europe was detected in Italy, followed by France and Spain, and isolated cases in the Netherlands, Belgium, Switzerland, Germany and Portugal. *Xf* spread has been connected with human-mediated movement of infected plants, resulting in distribution of *Xf* across large geographical distances. In Europe, four *Xf* subspecies have been identified: *fastidiosa, multiplex, pauca and sandyi*. Subspecies *fastidiosa* originated in Central America, *multiplex* in North America and *pauca* in South America (Sicard, et al., 2018). The origin of *sandyi* is still under debate (Almeida, and Nunney, 2015). A fifth subspecies, *tashke*, has only been found in the Americas (Randall, et al., 2009; Janse, and Obradovic, 2010). And a sixth subspecies (*morus*) has been proposed but is still under review (Nunney, et al., 2014a).

The importance of studying *X. fastidiosa*

Xf is an increasing threat to British agriculture. Climate change makes the environment more suitable for *Xf* which is known to favour warmer regions (Henneberger, 2003; Feil, and Purcell,

2007; Meyer, and Kirkpatrick, 2008), but also international plant trade is growing every year, meaning the bacteria have many ways to enter the country.

This research will provide a better understanding of the evolutionary history of *Xf* and the molecules involved in disease progression, which can ultimately help with the generation of targeted treatments for plants infected by *Xf*. There is currently no treatment solution for plants infected by *Xf*, with the only option being destruction of the host, its surroundings and a quarantine order. This research may gain more insight into the complex host range of the bacterium, its yet unknown mode of action within the plant, and determine why the pathogen causes disease in some hosts but remains asymptomatic in others. Ultimately, the outcome of the study could pave the way to implementing further control measures and creating diagnostic tools for the prevention of an outbreak. Genomics can create diagnostics, and understanding how the pathogen causes disease could lead to a potential treatment rather than having to destroy bacteria-carrying hosts and face huge economic loss.

Summary

Currently, the only control measure of *Xf* is prevention and destruction of plant hosts. Coming up with a targeted treatment plan could mean saving potentially millions of pounds. There are several reasons why there is still very little known about *Xf*, some of which include its long incubation period in the plant, difficulty to cultivate *in vitro* and thus study in the laboratory, and the many asymptomatic host plants where the bacterium lives in without causing any disease. The ability of computational methods to investigate an organism's genome has become very powerful, allowing a better understanding of the organism. Investigating genes that are involved in disease development will help with the understanding of the bacterium's molecular biology. Understanding how the bacterium works in a molecular level could bring us a step closer to establish a targeted treatment plan for this devastating bacterium.

A number of interesting putative effector proteins are promising – but as this is an on-going research, more analyses need to be carried out. Understanding the molecules involved in pathogenesis of *Xf* could help in the development of an efficient treatment plan for plants infected by the bacteria. Furthermore, the first detection of *Xf* in Colombia is described here. Twelve *Coffea arabica* samples of five cultivars and one *C. arabica* of unknown cultivar have tested positive for *Xf*. The detection of *Xf* in the country opens more questions about this fascinating bacterium. Colombia does not appear to have an outbreak of *Xf*, unlike its neighbouring country Brazil, where coffee leaf scorch due to *Xf* is rampant. It would be interesting to perform comparative analyses between Brazilian and Colombian *Xf* strains to understand why *Xf* appears to be more pathogenic in one country/host than another.

Financial Benefits

The financial impact of *Xf* is difficult to accurately estimate due to its large host range. However, it is clear that the arrival of *Xf* in a country has an enormous impact in many sectors, as the detection of *Xf* would not only farms, but also nurseries, retailers, and importers/exporters. According to Lindow (2019), PD had caused an annual loss of US$ 104 million in California by 2014. Approximately US$ 50 million is also spent on preventative measures every year. In Brazil, 40% of citrus plants are affected by CVC, which had caused an annual loss of US$ 120 million by 2005 (Rapicavoli, et al., 2018a). In Europe, the first *Xf* epidemic was identified in Apulia in the South of Italy, where the bacterium was found to be the cause of OQDS. 40% of olive trees are grown in Apulia for the production of olive oil in Italy (Strona, Carstens, and Beck, 2017), and over 10 ha of olive trees have since been destroyed (Martelli, et al., 2016). Undeniably, the detection of *Xf* in a country has a significant economic impact. Prevention of the arrival of *Xf* in the UK is the only control measure currently available.

Action Points

As *Xf* has not yet been detected in the UK, there are no action points to be tackled for growers. However, it is advised to remain vigilant of symptoms and report any potential ones to the correct authority. If an outbreak is suspected, contact the APHA Plant Health and Seeds Inspector or PHSI Headquarters for England and Wales (planthealth.info@apha.gsi.gov.uk), the Scottish Government's Horticulture and Marketing Unit (hort.marketing@gov.scot), or the DAERA Plant Health Inspection Branch for Northern Ireland (planthealth@dardni.gov.uk). Be aware of *Xf* disease symptoms, these can vary between different plants. Visit the EPPO website (https://gd.eppo.int/taxon/XYLEFA/photos) for disease pictures and the European Commission website for an extensive list of susceptible *Xf* plant hosts (https://ec.europa.eu/food/plant/plant_health_biosecurity/legislation/emergency_measures/xylella-fastidiosa/susceptible_en). It is also advised to keep up-to-date with plant health news. And most importantly, to avoid importation of plants from areas affected by *Xf* (https://www.cabi.org/isc/datasheet/57195#todistributionDatabaseTable and https://www.cabi.org/isc/datasheet/57195#toDistributionMaps) and/or ensure the imported material holds appropriate plant passports and phytosanitary certificates.

According to DEFRA, import requirements for *Coffea* sp. plants, plant produce and products from third countries are regulated, meaning that imports are only allowed if products are reported prior to entering the UK and are accompanied by phytosanitary certificates. *Coffea* imported particularly from Costa Rica and Honduras are prohibited unless a scientific research license or derogation has been procured. Import from these two countries has been added to DEFRA's "prohibited" list most likely because coffee plants imported from these countries into the Netherlands have tested positive for *Xf* (Bergsma-Vlami, et al., 2015).

The detection of *Xf* in Colombia reveals that the prevalence of the bacteria in the Americas is more common than previously assumed. This means that *Xf* could also be more prevalent in the European continent than currently known, as the pathogen is mainly only reported in countries with existing outbreaks. In general, it is advised to avoid direct coffee plant imports from countries where *Xf* has been detected, particularly Argentina, Brazil, Colombia, Costa Rica, Honduras, Paraguay, Puerto Rico and Venezuela.

SCIENCE SECTION

Background: A description of the bacterium *Xylella fastidiosa*

Introduction

Xylella fastidiosa (*Xf*) is a bacterium that is most commonly known to cause olive quick-decline syndrome (OQDS) in the South of Italy, citrus variegated chlorosis (CVC) in Brazil, and Pierce's disease (PD) of grapevine in the USA. It is considered one of today's most devastating plant pathogens, disrupting international trade and causing huge economic loss for affected countries. As cited by White, et al. (2020), an estimate from Italy suggests that *Xf* currently affects olive trees that make up 10% of Italy's entire olive oil production, which amounts to a value of €390 million. This could even turn into a loss of up to €5.2 billion within 50 years if no further measures are put in place to stop the spread of the bacteria (Schneider, et al., 2020). Even though *Xf* was the first bacterial plant pathogen to have its complete genome sequenced (Simpson, et al., 2000), there are still many aspects to explore regarding its spread and molecular biology. This thesis aims to provide a description of *Xf* in detail (**Chapter I**). It includes the work that has been carried out to better understand the bacterium's biogeography, particularly the detection of *Xf* in Colombia (**Chapter II**). This part of the project emerged when an opportunity to join an expedition to Colombia in 2019 arose to search for *Xf* in the country. This was the first of such a survey to be conducted in Colombia, and the discovery of *Xf* in the country would allow for new genomes to be included in the analysis of *Xf* pathogenesis. This part of the thesis focusses on exploring virulence proteins that may be important in disease development (**Chapter III**).

A description of the bacteria

Biology

Xf is a rod-shaped gram-negative bacterium belonging to the Xanthomonadaceae family. Its diameter ranges from 0.25 to 0.5 µm, and its length from 0.9 to 4.0 µm. These sizes vary

between the different subspecies and conditions in which the cells are grown *in vitro* and *in vivo* (Chagas, Rossetti, and Beretta, 1992; Davis, Purcell, and Thomson, 1978; Wells, et al., 1987). Under an electron microscope, the characteristic rippled wall of *Xf* can be observed (Alves, et al., 2009; Newman, et al., 2003), and the cells have both short (type I pilus) and long fimbriae (type IV pilus). The long fimbriae, in particular, are responsible for the bacterium's twitching motility (Almeida, Coletta-Filho, and Lopes, 2014). *Xf*'s genome size is approximately 2.6 Mb, which is reduced compared to the 5.0 Mb genome of its phylogenetically closest taxon *Xanthomonas* spp. The other only known xylem-limited Xanthomonadaceae is *Xanthomonas albilineans*, which similarly to *Xf*, has a reduced genome size (3.8 Mb). It is believed that this reductive genome evolution resulted in the adaptation to a xylem-limited lifestyle (Pieretti, et al., 2009). *Xf* does not encode a type III secretion system (T3SS; Simpson, et al., 2000). This secretion system has been found to be essential for the virulence of other more widely studied phytopathogenic bacteria by facilitating host cell invasion and defence evasion (Galán, and Collmer, 1999).

Life cycle

Xf exists in both plants and insects. In plants, *Xf* resides in the xylem where it can exist in either a biofilm or planktonic state. The bacteria only proliferate as biofilm along the xylem wall (Chatterjee, Almeida, and Lindow, 2008). Nutrition is gained from sugars in the xylem sap and by degrading the cell wall using enzymes (Dow, and Daniels, 2000). This also allows the bacteria to move between xylem vessels and spread within the plant (Newman, et al., 2004). *Xf* relies on insect vectors for its dissemination. As *Xf* is xylem-limited, only xylem-sap feeding insects have the ability to acquire the bacteria from already affected plants. Once inside the insect, *Xf* attaches to the foregut of the insect and colonises the surface – though colonisation is not required for dissemination of the bacteria. These insects now have the ability to inoculate *Xf* into other plants during feeding, where

susceptible hosts may develop symptoms (Chatterjee, Almeida, and Lindow, 2008). See **Figure 1** for a diagram of the lifecycle. Leafhoppers, in particular, are able to re-transmit *Xf* to another plant host within one hour of acquiring the bacteria. *Xf* cells will colonise the xylem once again and infected plants may re-infect further insect vectors within one week of inoculation (Purcell, and Finlay, 1979).

Origin

The bacterium is believed to originate from the Americas, where it appears to be a generalist endophyte to native plant species (Hopkins, and Purcell, 2002; Chatterjee, Almeida, and Lindow, 2008). An endophyte is a microorganism that lives within a plant without deteriorating its fitness. *Xf* spread has been connected with human-mediated movement of infected plants, resulting in distribution of the bacteria across large geographical distances. The first disease associated with the bacterium was detected by Newton B. Pierce in 1892 in the USA. Previously confined to the Americas, the first outbreak of *Xf* in Europe was observed in 2013, largely affecting olive trees in the South of Italy (Saponari, et al., 2013). Today, *Xf* outbreaks have been rampant in Italy, France and Spain. In two countries – Netherlands (Bergsma-Vlami, et al., 2015) and Belgium (AVBS, 2018) – *Xf* was detected in imported plants that were intercepted at ports, and in another three countries – Switzerland (EPPO, 2015), Germany (EPPO, 2016a), and Portugal (EPPO, 2019) – the bacterium was detected in isolated cases only and is currently under eradication or has since been eradicated.

It has been suggested that *Xf* subsp. *pauca* diverged over 50,000 years ago. This was followed by *Xf* subsp. *multiplex*, then *fastidiosa* and *sandyi* approximately 20,000 to 40,000 years ago resulting from geographic isolation (Nunney, et al., 2013; Schuenzel, et al., 2005; Vanhove, et al., 2019).

Subspecies

To date, six *Xf* subspecies have been described. *Xf* subsp. *fastidiosa* evolved in Central America, *multiplex* in North America and *pauca* in South America (Sicard, et al., 2018). The origin of *Xf* subsp. *sandyi* and *morus* are still under debate (Schaad, et al., 2004; Scally, et al., 2005; Nunney, et al., 2014b; Marcelletti, and Scortichini, 2016a; Almeida, and Nunney, 2015), but it appears that *Xf* subsp. *morus* is a result of recombination between subsp. *fastidiosa* and *multiplex* (Nunney, et al., 2014b). Lastly, subspecies *tashke* has only been found in North America. In Europe, subspecies *fastidiosa*, *multiplex*, *pauca*, *sandyi*, and *morus* have been identified.

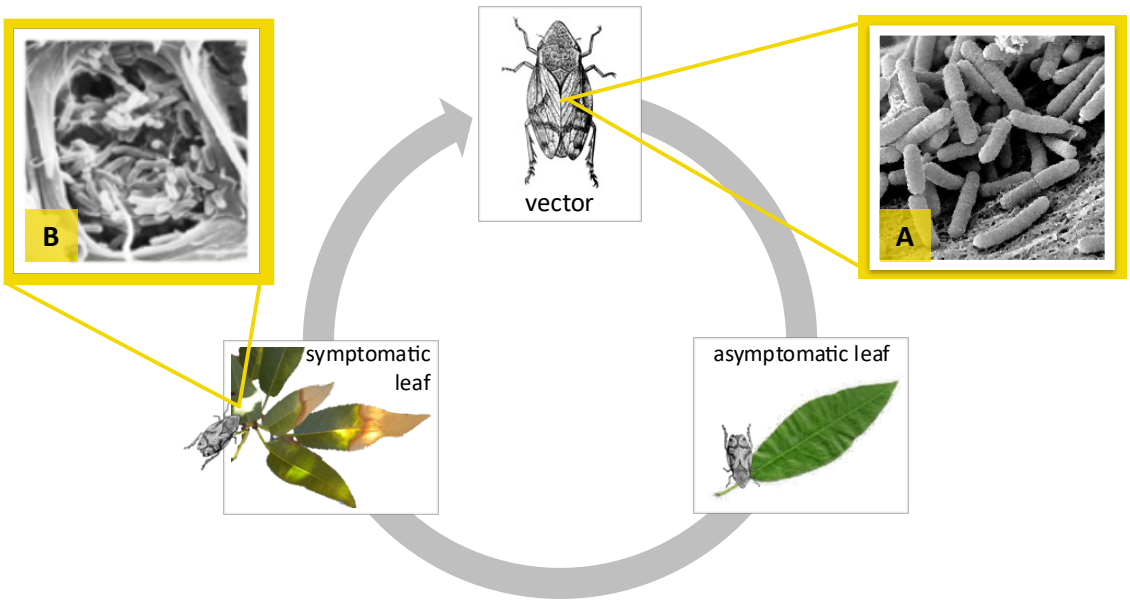


Figure 1: *Xylella fastidiosa* (*Xf*) lifecycle. *Xf* resides in the foregut of its xylem-sap feeding insect vectors (**A**: Electron Microscopy Laboratory, U.C. Berkeley). Insect vectors may inoculate new plants with *Xf* during feeding on the plant xylem, when the bacteria detach from the insect foregut and are disseminated to the new plant host. The bacteria proliferate as biofilm along the xylem wall (**B**: De Lima et al., 1998) may employ cell-wall degrading enzymes to move between xylem vessels and proliferate within the plant. More xylem-sap feeding insects may feed on the affected plant, acquire *Xf* and spread the bacteria to other plant hosts.

Hosts and vectors of the bacterium

Plant hosts

The plant host range of *Xf* includes both monocots and dicots. The bacterium has been detected in over 350 different botanical taxa, and dozens of crops are susceptible to *Xf*. The list of *Xf* host plants grows every year, particularly that of asymptomatic plant hosts. See **Appendix Table C** for an extensive list of *Xf* host plants. In the United Kingdom, the most significant crops include, but are not limited to: *Acer* (maple), *Brassica*, *Cercis* (redbuds), *Chionanthus* (fringe tree), *Cytisus* (broom), *Ficus carica* (fig), *Fragaria* (strawberry), *Hedera* (ivy), *Hypericum* (St. John's Wort), *Juglans* (walnut), *Lavandula* (lavender), *Magnolia*, *Medicago sativa* (alfalfa), *Morus alba* (mulberry), *Nandina domestica* (sacred bamboo), *Olea europaea* (olive), *Prunus* (stone fruit trees), *Pyrus* (pear), *Quercus* (oak), *Rosa* (rose), *Salix* (willow), *Salvia rosmarinus* (rosemary), *Sambucus* (elderberry), *Trifolium* (clover), *Ulmus* (elm), *Vaccinium* (blueberry), *Vitis* (grapevine). A plant does not remain a host for its entire lifetime when inoculated by a vector. Symptomatic plants can recover from diseases caused by *Xf* (Purcell, 1981) and bacteria have been found to die out in asymptomatic plant hosts (Purcell, and Saunders, 1999).

Insect vectors

Xf is transferred between plants through xylem-sap feeding insects, such as members of the Aphrophoridae (spittlebugs) and Cicadallinae (sharpshooters; Cavalieri and Porcelli, 2017). The fitness of insect vectors is not impacted by the bacteria. In Europe, *Philaenus spumarius* (meadow spittlebug) appears to be the main vector responsible for most of *Xf* spread (Cornara, et al., 2017; Rapicavoli, et al., 2018a). Vectors are capable of transmitting all *Xf* subspecies without specificity (Almeida, et al., 2005; Almeida, and Nunney, 2015; Freitag, 1951). Bacteria are transferred from a plant to an insect when the mouthpart of the vector penetrates an *Xf*-infected xylem vessel (Purcell, Finlay, and McLean, 1979). Inside the

insect, *Xf* resides in the foregut. Young insects lose their infectivity after the nymphal stages of vector moult, as the cuticular lining in the foregut is shed during each moult (Purcell, and Finlay, 1979). Adult insects do not moult and can therefore harbour *Xf* their entire life. A lack of transovarial transmission was also demonstrated by Freitag (1951). Insect vectors are able to transmit bacteria to plants even months after acquisition from an infected plant enabling the vectors' long-term infectivity.

X. fastidiosa is a global phytopathogen

Pathogenicity

The ability of a microorganism to cause damage to a host is known as pathogenicity (Casadevall, and Pirofski, 1999). To demonstrate pathogenicity, Koch's postulates must be fulfilled. For *Xf*, this was completed for PD in grapevine (Davis, Purcell, and Thomson, 1978), CVC (Chang, et al., 1993; Hartung, et al., 1994), OQDS (Saponari, et al., 2017) and a number of other plants (e.g. Hernandez-Martinez, Cooksey and Wong, 2009; Purcell et al., 1999). However, it must be noted that not all susceptible plant hosts of *Xf* have been verified by Koch's postulates, which is required to demonstrate that the different genotypes of a microorganism are indeed pathogenic to specific hosts (Almeida, and Nunney, 2015). Pathogenicity tests of *Xf* are difficult due to its fastidious growth *in vitro* to generate inoculum. *Xf* only causes disease in a small number of its very wide host list. Symptoms of susceptible plants resemble nutrient deficiencies, drought stress or infections caused by other pathogens. Over a dozen diseases have been associated with *Xf*-infection (see **Table 1**). However, as most plant hosts do not develop any disease symptoms and *Xf* exists as an endophyte, it may be unsuitable to designate the term "pathogen" to *Xf*. It appears that the most susceptible hosts to *Xf* disease are economically important crops. Though, this knowledge of the bacterium is biased as economically important crops are also the most studied plant hosts of

Xf. This approach of study often over-shadows that fact that the bacterium most probably evolved as an endophyte.

Virulence

It has been argued that *Xf* is primarily a commensal bacterium in plant hosts, but may adopt a virulent lifestyle due to molecular interactions with the host immune system (Baccari, and Lindow, 2011; Choi, et al., 2013; Roper, Castro, and Ingel, 2019). According to Casadevall and Pirofski (1999), virulence describes the quantitative degree of damage caused by a pathogen, or the degree of pathogenicity. Symptoms are a result of a pathogen's virulence. Symptoms vary among the different plant hosts infected by *Xf*, though the majority of host species do not develop any disease symptoms at all (Purcell, and Saunders, 1999). The pathogenesis, or mechanism by which a disease develops, is still unclear in *Xf*. It is suggested that *Xf*'s mechanism to self-restrict virulence and merely exist as an endophyte in a plant host is explained by cell-cell signalling within bacterial communities (Chatterjee, Almeida, and Lindow, 2008). It is not understood, however, why this mechanism breaks down in susceptible plant hosts, which results in severe symptoms. *Xf* symptoms typically result from impairing the flow of soluble nutrients and water through the xylem (Newman, et al., 2004). This impairment of water flow in the xylem results from the formation of biofilm, a matrix that consists of an aggregation of bacterial cells and molecules secreted by these bacteria. As part of the plant's immune response, vascular occlusion (VO) may also be induced in the xylem. VOs are structural modifications, for example formation of tyloses or thickening of the cell-wall by callose deposition and lignification (Choi, et al., 2013; Sun, et al., 2013; Zaini, et al., 2018). The formation of these structures allows the containment of the bacterium and therefore stop further spread within the xylem. However, as the mode of action of disease induction is believed to be the blockage of the xylem, the plant's intention to save itself only aids *Xf*'s pathogenicity. This was confirmed in PD-susceptible grapevine by Sun et al. (2013): VOs in

grapevine were found to suppress water conduction within the xylem and contribute to *Xf* symptoms even more (Alves, et al., 2009; De Benedictis, et al., 2017; Rapicavoli, et al., 2018a).

Xf exists in two states: planktonic and biofilm (Gouran, et al., 2016). In its biofilm state, *Xf* is embedded in extracellular polymeric substances (EPS) – which is made up of polysaccharides, DNA oligomers, proteins and peptides – and has an adhesive property. This allows *Xf* to colonise insects, where it resides in the foregut, but also inhibit the flow of water in plants by occlusion of the xylem. In its planktonic state, *Xf* exists as motile single cells and allows the dispersion of bacteria to different parts of the plant. Therefore, the *Xf* population is not homogenous in the entire plant and populations can accumulate at specific tissues (Hopkins, 1985). For instance, in PD of grapevine, *Xf* was found in larger populations in symptomatic leaf veins and petioles of the plant (Baccari, and Lindow, 2011; Krivanek, Sisterson, and Lin, 2005). A positive correlation between the severity of symptoms and the population of *Xf* within a plant host has also been shown in research conducted by Krivanek, Sisterson and Lin (2005).

The planktonic and biofilm states of *Xf* are controlled by an intercellular communication system known as quorum sensing (QS), which involves the diffusion of signalling molecules that accumulate as the bacterial population size increases (Camilli, and Bassler, 2006). When a certain threshold is reached, receptor proteins are activated, triggering signalling cascades that change gene expression (Von Bodman, Bauer, and Coplin, 2003). In *Xf*, QS is controlled by diffusible signal factors (DSFs). QS in *Xf* has been linked with promoting a non-adhesive phenotype at low cell densities (Roper, Castro, and Ingel, 2019). Interestingly, *Xf* in its planktonic state is hypervirulent. Research investigating defects in cell-cell aggregation, surface attachment and biofilm maturation have shown that the bacteria are fixed in dispersal

form and are able to systematically colonise the plant host very rapidly (Burbank, and Stenger, 2017; Gouran, et al., 2016; Guilhabert, and Kirkpatrick, 2005; Newman, et al., 2004). This indicates that *Xf* in its biofilm state is a self-limiting behaviour and a means to control its own movement within the host (Roper, Castro, and Ingel, 2019).

Table 1: Diseases caused by *Xylella fastidiosa* (*Xf*). Diseases caused by the bacterium result from the colonisation of the plant's xylem and blocking the flow of water and soluble nutrients. This leads to leaf scorch and stunt in many different plants. The following diseases have been associated with *Xf*.

Leaf scorch	Stunt
almond leaf scorch	alfalfa dwarf
coffee leaf scorch	citrus variegated chlorosis
elm leaf scorch	Lucerne dwarf
mulberry leaf scorch	periwinkle wilt
oak leaf scorch	phony peach disease
oleander leaf scorch	
olive quick-decline syndrome	
pear leaf scorch	
pecan leaf scorch	
Pierce's disease of grapevine	
plum leaf scald	

Economic impact

The financial impact of *Xf* is difficult to accurately estimate due to its large host range. However, it is clear that the arrival of *Xf* in a country has an enormous impact in many sectors, as the detection of *Xf* would not only affect farms, but also nurseries, retailers, and importers/exporters. According to Lindow (2019), PD had caused an annual loss of US$ 104 million in California by 2014. In addition, approximately US$ 50 million is spent on preventative measures every year. In Brazil, 40% of citrus plants are affected by CVC, which had caused an annual loss of US$ 120 million by

2005 (Rapicavoli, et al., 2018a). In Europe, the first *Xf* epidemic was identified in Apulia in the South of Italy, where the bacterium was found to be the cause of OQDS. 40% of olive trees are grown in Apulia for the production of olive oil in Italy (Strona, Carstens, and Beck, 2017), and over 10 ha of olive trees have since been destroyed due to *Xf* infection (Martelli, et al., 2016). Undeniably, the detection of *Xf* in a country has a significant economic impact. Prevention of the arrival of *Xf* in the UK is the only control measure currently available.

Detection

Culturing

Xf acquired its name for being a fastidious bacterium and requiring specific media for culturing (Wells, et al., 1987). The most commonly used media used to culture *Xf* include PD2 (Davis, Purcell, and Thomson, 1980), buffered charcoal-yeast extract (BCYE; Wells, et al., 1981), and periwinkle wilt Gelrite (PWG; Hill, and Purcell, 1995). Normally, *Xf* is grown in at least two different media for validation. Incubation takes place at 28°C and the first colonies are visible after 28 days. On all three media, the colonies appear circular, smooth-edged and slightly convex, and depending on the strain, colonies can be between 1.0 and 1.5 mm in diameter (EPPO, 2018). Due to its fastidious growth *in vitro*, serological assays and molecular methods are more commonly used to detect the presence of *Xf* in a tissue.

Serological assays

Serological techniques allow for the detection of antigens and antibodies of a microorganism of interest. A positive result is usually detected by a colorimetric change or change in fluorescence. A number of assays have been developed to detect the presence of *Xf* in tissue samples, and kits have become commercially available. For example, a double-antibody sandwich enzyme-linked immunosorbent assay (DAS-ELISA) developed by Sherald and Lei (1991) allows for rapid *Xf* detection. Briefly, tissue extracts suspected to harbour *Xf* (i.e. the

antigen) are applied to antibody-coated microwells. After a number of incubation and washing steps, a peroxidase-conjugated antibody is added to microwells, and the addition of a substrate to these wells will generate a colour change if the antigen of interest is present. Similarly, in an indirect immunofluorescence assay (IF) developed by Carbajal, Morano and Morano (2004) xylem exude is applied onto a previously prepared slide. This is then viewed under a microscope using epifluorescence illumination where *Xf*-positive samples are fluorescently visible. Both DAS-ELISA and IF are recommended to be used for plant tissue only due to its sensitivity. In contrast, direct tissue blot immunoassay (DTBIA) can be used for both plant and insect tissue and with minimal sample processing (Djelouah et al., 2014).

Molecular methods

In molecular detection methods, genetic sequences that are specific to *Xf* are targeted and amplified to determine the presence of the bacteria. One of the most commonly used methods to detect *Xf* in a sample is by polymerase chain reaction (PCR) due to its low cost and high sensitivity. Several protocols have been established using both conventional, and quantitative PCR (qPCR) techniques targeting different regions in the *Xf* genome (Francis, et al., 2006; Harper, Ward, and Clover, 2010; Li, et al., 2013; Minsavage, et al., 1994; Ouyang, et al., 2013). Another molecular detection method is by loop-mediated isothermal amplification (LAMP). A protocol advanced by Yaseen et al. (2015) and based on primers developed by Harper, Ward and Clover (2010) allows for on-site rapid detection of *Xf*. Molecular methods are suitable to test both plant and insect tissue as only an amount of as low as 10^2 cfu/mL of *Xf* is required to confirm its presence.

Current control measures

Treatments

There are no treatments for plants infected by *Xf*. In the European Union, when a host plant displays symptoms and is found to carry *Xf*, the host and all neighbouring potential hosts in a 100 m radius are destroyed (Commission Implement Decision (EU) 2015/789). A 5-10 km demarcation order is also implemented, preventing the movement of plant material outside this area and thus greatly affecting a region's economy. Currently, only preventative measures can be put to place to stop the spread of *Xf*.

Prevention

Some preventative measures, especially by targeting vectors, have been implemented to reduce the risk of an *Xf* outbreak (Dongiovanni, et al., 2018). However, this does not aid with those plants already affected by *Xf*. The lack of fundamental knowledge of the molecular biology of *Xf* makes it difficult to truly understand the mode of pathogenicity of the bacterium. Preventative measures include planting of certified nursery trees only, pruning affected branches and entire removal of severely symptomatic plants. Notably, *Xf* completely relies on insect vectors for spread. This limitation of the bacterium is a key to outbreak prevention by reducing vector population or preventing *Xf* from colonising a vector. Understanding *Xf* molecular biology could help with the development of a targeted treatment plan for infected plants, instead of relying on preventative measures only. With a better understanding of the molecules involved in pathogenesis, researchers will know what to target to stop disease progression.

The first detection of *X. fastidiosa* in Colombia

Introduction

Global occurrence of X. fastidiosa

Xf does not cause any symptoms in the majority of its plant hosts, however it is devastating to those hosts susceptible to the bacterium. *Xf* has been identified in Asia, Europe, and the Americas (see **Figure 2**) causing a variety of leaf scorch symptoms in crops. Before 2013, only sporadic cases of *Xf* have been described in Europe. In 2013, an outbreak of OQDS in Apulia in Italy was observed and determined to be caused by *Xf*. It is believed that the strain causing the outbreak in Italy evolved in Central America (Marcelletti, and Scortichini, 2016b) and was introduced through human-mediated movement. Researchers believe *Xf* evolved in the Americas (Hopkins, and Purcell, 2002; Chatterjee, Almeida, and Lindow, 2008) causing diseases in olive as these are particularly susceptible to *Xf* spread. Today, most diseases caused by *Xf* are reported in North and South America.

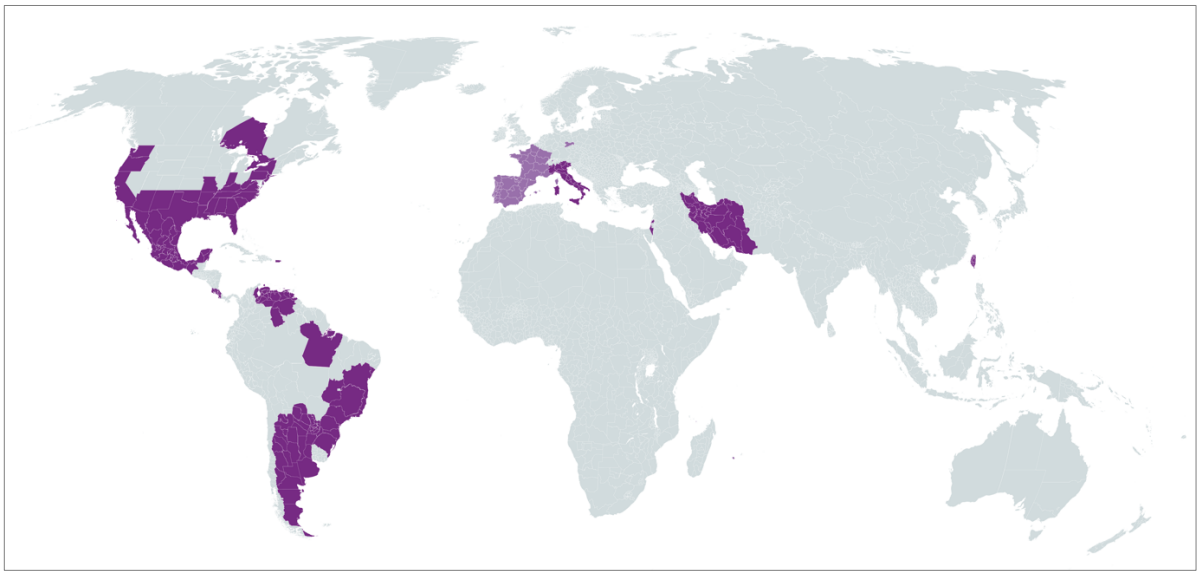


Figure 2: Global occurence of *Xylella* spp. This world map (https://mapchart.net/) depicts the global occurrence of *Xylella* spp. (highlighted in purple; lighter shaded areas indicate cases under control). *Xylella taiwanensis* has been identified in Taiwan only (Su et al., 2016). *Xylella fastidiosa* has been identified in the rest of the globe. The majority of cases have been reported in the Americas. The EPPO database provides an up-to-date distribution map

X. fastidiosa in Central and South America

Xf has been detected many times in different plants in Central America (citrus, coffee, avocado and oleander in Costa Rica; citrus and coffee in Puerto Rico), and South America (grapevine in Venezuela; citrus, coffee and plum in Brazil; plum in Paraguay; citrus and olive in Argentina). Population analyses by Nunney et al. (2010) suggest *Xf* subsp. *fastidiosa* to originate from the southern part of Central America. This subspecies is often associated with PD of grapevine and almond leaf scorch (ALS) in North America. *Xf* subsp. *pauca*, known to cause CVC and coffee leaf scorch (CLS) is hypothesised to be native in South America (Nunney, et al., 2012). *Xf* is associated with several disease outbreaks in Central and South America, particularly CVC and CLS. Those host species that harbour *Xf* have been introduced into the region, possibly suggesting that *Xf* is only pathogenic to plants with which it did not co-evolve. This explains why the introduction of *Xf* is very damaging to new geographical regions as the plant has not evolved the mechanisms to fend off the bacteria, in contrast to plant hosts that harbour *Xf* as an endophyte (Nunney, et al., 2014a).

X. fastidiosa is a quarantine pathogen

Xf is categorised as a quarantine pathogen in many countries. *Xf* has been associated with a variety of leaf scorch, stunt and dieback diseases in crops. In particular, the bacterium causes severe CVC and CLS largely caused by *Xf* subsp. *pauca* and restricted to agricultural and ornamental plants in Brazil (Nunney, et al., 2014a). In Costa Rica, where *Xf* subsp. *fastidiosa* is more common, a milder form of CLS – locally known as “crespera” – has been associated with the bacterium (Rodríguez, et al., 2001). However, unlike in European countries, *Xf* is endemic to the Americas. The EU has implemented strict regulations regarding the detection of bacteria in plant hosts to prevent the spread of *Xf* in the continent. As *Xf* is an invasive species in the European continent, plant hosts here are more likely to be susceptible to diseases caused by the bacterium.

Aim of research

This research describes the first known survey to be conducted in Colombia to determine the presence of *Xf*, which may provide a bigger picture of the spread of the bacteria. Several locations in the Antioquia department were chosen to collect leaves of different plant families where *Xf* had previously been detected in. Colombia is one of the most biodiverse countries in the world and is the connecting hub between Central and South America. No cases of *Xf* have been reported in Colombia – one of the largest coffee-producing countries – despite reported cases in many nearby countries. Furthermore, coffee is a plant host that has been found to harbour all *Xf* subspecies also found in Europe.

Methods

Sampling of plant leaves

Plant leaves were collected from seven different locations within the Antioquia department of Colombia (see **Figure 1**; Kahle, and Wickham, 2013). These included one coffee farm, one citrus farm, two research stations, one location within a rainforest, one university campus and one botanical garden in an urban area. Leaves of 15 different plant species from three families were collected: Malvaceae, Rubiaceae and Rutaceae. *Xf* had previously been detected in several species of these families. See **Appendix Table A** for full details of collected samples. Whenever possible, samples were taken from three plants of each plant species. Of each plant, at least three branches were selected and at least three leaves of each branch were collected (see **Appendix Figure A**) using scissors disinfected with 70% ethanol prior to each use. Sufficient leaves were collected per sample plant so that three batches of DNA extractions could be performed per sampled tree if needed. Each leaf was surface cleaned with 70% ethanol and air-dried before being placed into a clear polyethylene bag. This polyethylene bag was placed in an additional two polyethylene bags to prevent contamination and accidental *Xf* spread. All sample bags were stored until shipment to the United Kingdom

for processing.

DNA extraction

All leaf samples were processed in a laminar flow cabinet within a licensed pathogen laboratory at the National Institute of Agricultural Botany - East Malling Research (NIAB EMR) in Kent. When removed from the bags, samples were surface cleaned with 70% ethanol, followed by distilled water and subsequently left to air-dry. Once dried, leaves were cut as only the petiole, midrib and basal parts were required for DNA extraction. Cut leaf parts were placed in 2.0 mL Eppendorf tubes, frozen in liquid nitrogen and stored in a -20°C freezer until further processing. Total DNA was extracted from each sample using a cetyltrimethylammonium bromide (CTAB) method designed by EPPO (EPPO Bulletin, 2016) and modified in this research. The detailed protocol can be found in **Appendix Figure C**. All total DNA extracts were stored at -20°C. The remaining two batches of each sampled plant were stored at -80°C for future use.

Detection of X. fastidiosa

Three separate PCRs were prepared which are referred to as 16S, XF1 and XF2 hereafter. In the 16S PCR, primers 27F and 1492R targeted the 16S region of the genome to detect the presence of bacteria (Muyzer, De Waal, and Uitterlinden, 1993). The XF1 PCR was a primary method to determine the presence of *Xf*. In this PCR, *Xf*-specific primers RST31 and RST33 target the 3' end of *rpoD*, a gene encoding an RNA polymerase sigma-70 factor in the bacterium (Minsavage *et al.*, 1994). The XF2 PCR, using *Xf*-specific primers 16S-23F and 16S-23R, was a secondary method, and control to confirm the presence of *Xf* in a sample. This targets the 16S-23S intergenic spacer region of the bacterium (Martinati *et al.*, 2005). See **Appendix Table B** for complete sequences of each primer pair and PCR conditions for each reaction. 16S PCR was repeated thrice per sample to determine consistency of results.

Only samples that tested positive for 16S at least twice were tested for *Xf*. Sigma-Aldrich redTaq DNA polymerase was used for all PCRs. Finally, XF1 PCR was repeated on all positive Colombian samples using Invitrogen's high fidelity Platinum *Taq* DNA polymerase. *Xf* subsp. *fastidiosa* strain Temecula 1 DNA was acquired from the National Collection of Plant Pathogenic Bacteria (NCPPB) at Fera Science to use as a positive control.

Multi-locus sequence typing (MLST)

Multi-locus sequence typing (MLST) is a method to characterise bacteria by analysing sequence variations – known as the sequence type (ST) – in housekeeping genes, which are highly conserved sequences in the genome essential for the bacteria to survive (Maiden, *et al*., 1998; Maiden, 2006). This usually comprises of the amplification of species-specific housekeeping genes, subsequent sequencing, and comparison of the sequence variations using reference data. In *Xf*, seven housekeeping genes – *leuA, perC, malF, cysG, holC, nuoL* and *gltT* (see **Appendix Table B** for function and primer sequences of each target gene) – have been previously selected for MLST, which is important for the identification of the *Xf* subspecies of a strain (Scally, et al., 2005). The *Xf*-specific MLST protocol developed by Yuan et al. (2010) was followed and performed on four positive Colombian samples RUBCA03001, RUBCA03002, RUBCA03006 and RUBCA05001 (see **Appendix Table B** for PCR protocols) using Invitrogen's high fidelity Platinum *Taq* DNA polymerase. Sequencing was performed as described below and ST was allocated using the *Xf* PubMLST database (Jolley, Bray and Maiden, 2018; https://pubmlst.org/xfastidiosa/). For some *Xf* genomes available at NCBI, no subspecies information was provided. To identify the sequence type and thus the subspecies of those strains, the seven housekeeping genes of strains of interest were extracted from the genome using NCBI's BLAST, and the ST was determined using the PubMLST database (see **Appendix Table E** for subspecies information of each publicly available *Xf* genome).

An MLST phylogeny was created using concatenated sequences of four of the seven housekeeping genes targeted during MLST (*cysG*, *holC*, *nuoL*, *gltT*) of four Colombian strains on which MLST profiling was performed (RUBCA03001, RUBCA03002, RUBCA03006 and RUBCA05001) and 293 *Xf* strains obtained from the PubMLST database. Concatenated sequences of four target genes (*cysG, holC, nuoL, gltT*) from the MLST scheme refined by Yuan et al. (2010) were aligned with MUSCLE (Edgar, 2004) and the phylogeny was inferred by maximum-likelihood using IQ-TREE (Nguyen, et al., 2015). The subspecies, country of origin and host information of each strain are colour-coded.

Sequencing of positive samples

XF1 amplicons of samples RUBCA03001, RUBCA03002, RUBCA03003, RUBCA03005, RUBCA03006, RUBCA03007, RUBCA03008, RUBCA03010, RUBCA03011, RUBCA03013, RUBCA03015, RUBCA05001 and the positive control *Xf* subsp. *fastidiosa* strain Temecula-1 were selected for sequencing to refute contamination. PCR amplicons were purified using Biolabs Monarch PCR & DNA Cleanup Kit and Sanger sequenced using the Eurofins LightRun GATC service. Consensus sequences of sequencing data were acquired with Geneious 10.0.2 (https://www.geneious.com) and using the *rpoD* gene sequence of *Xf* subsp. *fastidiosa* strain 9a5c (obtained from NCBI) as the reference sequence. Multiple sequence alignment (MSA) by iterative progressive alignment was performed on consensus sequences using the program MUSCLE (Edgar, 2004). The alignment was visualised using JalView (Waterhouse, et al., 2009).

Figure 3: Map of collection sites in Colombia. GPS coordinates of all samples were collected during the survey. Using the R package *ggmap* (Kahle, and Wickham, 2013), the collection sites were mapped to the map of Colombia. Only samples in the Antioquia province of Colombia were collected. See **Appendix Table A** for full details of each collected sample.

Results

X. fastidiosa was detected in 13 Coffea arabica samples

Xf is believed to originate from the Americas (Hopkins, and Purcell, 2002; Chatterjee, Almeida, and Lindow, 2008). In Colombia, leaves of the plant families Malvaceae (e.g. hibiscus), Rubiaceae (e.g. coffee) and Rutaceae (citrus; see **Appendix Table A** for a full list of collected samples) were collected. The goal was to sample as many plants as possible from these families in both natural and cultivated environments and test these for *Xf*, the hypothesis being that *Xf* is an endophytic organism in South American endemic plants. EPPO has published standard *Xf* protocols for the extraction of total DNA from plant leaves for subsequent identification by molecular methods, particularly by PCR in this research (EPPO, 2016b). Here, the CTAB-based DNA extraction protocol was modified in order to get high concentrations of DNA from *C. arabica* samples. All leaves were surface sterilised with 70% ethanol, washed in distilled water and air-dried to avoid the DNA extraction of epiphytes, microorganisms living on the surfaces of plants. The leaves were freeze-dried and ground with a pestle and mortar instead of a mechanical homogeniser (EPPO, 2016b). Pre-heated CTAB buffer was added

after grinding. Due to the high concentrations of RNA in the first trials of the EPPO standard protocol, RNAse A was added after the CTAB step and incubated overnight to allow the RNAse to digest RNA in the sample. For the precipitation of DNA, room temperature 2-propanol instead of cold 2-propanol – as suggested in the standard protocol – was added to prevent excess salts of being precipitated with the DNA and thus get better concentrations. Lastly, each sample was washed three times with 70% ethanol to ensure all contaminants were removed from the sample.

A total of 51 plants were sampled in triplicates during the Colombian survey. In other words, for each sampled plant, enough leaves were collected to perform three batches of DNA extractions. Thirteen of these samples tested positive for *Xf* (see **Table 3**). Twelve of these were collected from a single coffee farm in Fredonia, and one from the EAFIT University campus in Medellín, both of which are located in the Antioquia department of the country. All positive samples originated from *C. arabica* plants. No *Xf* was identified using the XF1 PCR protocol in any of the Malvaceae and Rutaceae samples. From the coffee farm, plants of three positive samples did not display any *Xf*-specific symptoms. These were samples RUBCA03001 and RUBCA03002 (*C. arabica* cv. Geisha); and RUBCA03005 (*C. arabica* cv. Colombia). The positive sample collected from the University campus, RUBCA05001, did not display *Xf*-like symptoms, but were affected by coffee rust, a fungal disease caused by *Hemileia vastatrix*. The cultivar of this plant is unknown. The remaining nine samples that tested positive for *Xf* originated from the coffee farm and displayed mild leaf scorch symptoms similar to *Xf*-affected *C. arabica* plants found in Brazil and Costa Rica. These included samples RUBCA03003 (*C. arabica* cv. Geisha); RUBCA03006 (*C. arabica* cv. Colombia); RUBCA03007, RUBCA03008 (*C. arabica* cv. Caturra); RUBCA03010, RUBCA03011, RUBCA03012 (*C. arabica* cv. Pajarito); RUBCA03013 and RUBCA03015 (*C. arabica* cv. Castillo). All thirteen positive samples were tested by PCR using *Xf*-specific primers targeting

two different regions in the genome: XF1 PCR amplified the 3' end of *rpoD*, a gene encoding an RNA polymerase sigma-70 factor (see **Figure 5**; Minsavage *et al*., 1994), and XF2 PCR amplified the 16S-23S intergenic spacer region (Martinati *et al*., 2005).

XF1 PCR amplicons of the positive samples – RUBCA03001, RUBCA03002, RUBCA03003, RUBCA03005, RUBCA03006, RUBCA03007, RUBCA03008, RUBCA03010, RUBCA03011, RUBCA03012, RUBCA03013, RUBCA03015 and RUBCA05001 – and the positive control *Xf* subsp. *fastidiosa* strain Temecula-1 were sequenced. An MSA of the positive samples and the positive control reveals nucleotide differences in a number of sites of the sequences (see **Appendix Figure D**). This confirms that the positive Colombian samples were not contaminated with the positive *Xf* control.

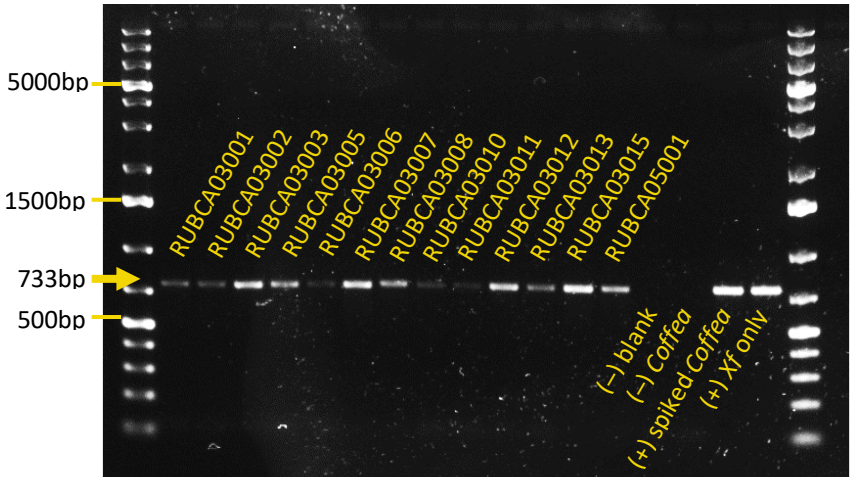


Figure 4: Gel image of positive Colombian *Xylella fastidiosa* (*Xf*) samples identified by XF1 PCR. XF1 PCR targets the *Xf*-specific *rpoD* gene was performed on all Colombian samples (Minsavage et al., 1994). This gel depicts all samples where *Xf* was identified, which comprise of several *Coffea arabica* cultivars from a single farm, and one *C. arabica* plant of unknown cultivar from a university campus. The resulting amplicon is 733bp long. Two negative controls (one blank and one *C. arabica* total DNA extract) and two positive controls (one *C. arabica* extract spiked with *Xf* DNA and one pure *Xf* DNA sample) were included.

X. fastidiosa subsp. pauca was confirmed in four Colombian strains

The subspecies of *Xf* is determined by determining a strain's MLST profile. This comprises the amplification of seven *Xf*-specific housekeeping genes (Yuan et al., 2010), subsequent sequencing, and ST profiling according to the PubMLST database (Jolley, Bray and Maiden, 2018; https://pubmlst.org/xfastidiosa/). Two samples originating from two different *C. arabica* cv. Geisha plants (RUBCA03001 and RUBCA03002) were determined to be of ST74 and ST66/ST74 (ST profile was unclear in this sample), and a third originating from a *C. arabica* cv. Colombia plant (RUBCA03006) of ST73. Loci *leuA*, *malF* and *petC* of the fourth sample originating from a C. arabica of unknown cultivar (RUBCA05001) were not included in the allelic profiling due to low quality sequencing data. However, allelic profiling of the remaining four loci was sufficient to determine the sequence type (ST74) and thus the subspecies (see **Table 2**). As summarised by EPPO (2018), all these ST profiles belong to *Xf* subsp. *pauca*, a subspecies believed to originate from South America (Sicard et al., 2018). A phylogeny was created using concatenated sequences of the seven housekeeping genes targeted during MLST (see **Figure 5**) including the four Colombian strains and 293 *Xf* strains obtained from the PubMLST database. The Colombian strains are highlighted with a red arrow on the phylogeny, and subspecies, country of origin and host information of each strain are colour coded.

Table 2: Multi-locus sequence type (MLST) allelic profiles of Colombian *Xylella fastidiosa* (*Xf*) strains. MLST was performed on four Colombian strains to determine the *Xf* subspecies. Briefly, seven loci (*cysG, gltT, holC, leuA, malF, nuoL* and *petC*) were amplified and sequenced to determine the allelic profile of each strain with help of the PubMLST database (Jolley, Bray and Maiden, 2018; https://pubmlst.org/xfastidiosa/). Sequence quality of RUBCA05001 *leuA*, *malF* and *petC* were not of sufficient quality to be profiled. This table also indicates whether the sequence matched the PubMLST database sequences only partially or exactly. Allelic profiles of strain RUBCA03006 differ in two loci (*cysG* and *holC*) when compared with the other three Colombian strains.

Locus	Strain	Cultivar	Phenotype	PubMLST match	Allele	ST
cysG	RUBCA03001	Geisha	A	partial	28	74
cysG	RUBCA03002	Geisha	A	partial	28	66 / 74
cysG	RUBCA03006	Colombia	S	partial	26	73
cysG	RUBCA05001	N/A	A	partial	28	74
gltT	RUBCA03001	Geisha	A	partial	8	74
gltT	RUBCA03002	Geisha	A	partial	8	66 / 74
gltT	RUBCA03006	Colombia	S	exact	8	73
gltT	RUBCA05001	N/A	A	exact	8	74
holC	RUBCA03001	Geisha	A	partial	11	74
holC	RUBCA03002	Geisha	A	partial	11	66 / 74
holC	RUBCA03006	Colombia	S	partial	10	73
holC	RUBCA05001	N/A	A	partial	11	74
leuA	RUBCA03001	Geisha	A	partial	7	74
leuA	RUBCA03002	Geisha	A	partial	7	66 / 74
leuA	RUBCA03006	Colombia	S	exact	7	73
malF	RUBCA03001	Geisha	A	partial	8	74
malF	RUBCA03002	Geisha	A	partial	8	66 / 74
malF	RUBCA03006	Colombia	S	partial	8	73
nuoL	RUBCA03001	Geisha	A	partial	16	74
nuoL	RUBCA03002	Geisha	A	partial	16	66 / 74
nuoL	RUBCA03006	Colombia	S	partial	16	73
nuoL	RUBCA05001	N/A	A	partial	16	74
petC	RUBCA03001	Geisha	A	partial	6	74
petC	RUBCA03002	Geisha	A	partial	8	66 / 74
petC	RUBCA03006	Colombia	S	partial	6	73

Table 3: A list of all Colombian samples that tested positive for *Xylella fastidiosa* (*Xf*). Different parameters were measured during the collection of leaf samples in Colombia, such as location information, metres above median sea level (MAMSL) and GPS coordinates in decimal degrees (dd). Below are details of the samples that tested positive for *Xf* by PCR. All samples underwent two *Xf*-specific PCRs amplifying different regions in the genome: XF1 PCR amplified the 3' end of rpoD, a gene encoding an RNA polymerase sigma-70 factor (Minsavage et al., 1994), and XF2 PCR amplified the 16S-23S intergenic spacer region (Martinati et al., 2005). Four samples that tested positive for *Xf* were collected from plants that did not display *Xf*-specific symptoms. However, one of these (RUBCA05001) was affected by coffee rust. The remaining nine samples that tested positive for *Xf* originated from plants that displayed leaf scorch, a symptom that has been observed in *Xf*-affected *Coffea arabica* plants in Brazil and Costa Rica.

ID	Date	Time	Family	Species	Cultivar	Symptoms	location	Location	MAMSL	GPS (dd)	Temp (°C)	Humidity (%)
RUBCA03001	20190627	15:25	Rubiaceae	*Coffea arabica*	Geisha	A	Coffee farm	Fredonia	1423	5.970375, -75.670041	24	59
RUBCA03002	20190627	15:30	Rubiaceae	*Coffea arabica*	Geisha	A	Coffee farm	Fredonia	1423	5.9703, -75.6701	24	59
RUBCA03003	20190627	15:45	Rubiaceae	*Coffea arabica*	Geisha	S	Coffee farm	Fredonia	1423	5.9704, -75.6704	24	59
RUBCA03005	20190627	16:07	Rubiaceae	*Coffea arabica*	Colombia	A	Coffee farm	Fredonia	1423	5.9730, -75.6700	24	59
RUBCA03006	20190627	16:12	Rubiaceae	*Coffea arabica*	Colombia	S	Coffee farm	Fredonia	1423	5.9730, -75.6701	24	59
RUBCA03007	20190627	16:42	Rubiaceae	*Coffea arabica*	Caturra	S	Coffee farm	Fredonia	1786	5.99748, -75.6644	24	59
RUBCA03008	20190627	16:46	Rubiaceae	*Coffea arabica*	Caturra	S	Coffee farm	Fredonia	1786	5.9749, -75.6643	24	59
RUBCA03010	20190627	16:54	Rubiaceae	*Coffea arabica*	Pajarito	S	Coffee farm	Fredonia	1786	5.9748, -75.6644	24	59
RUBCA03011	20190627	16:59	Rubiaceae	*Coffea arabica*	Pajarito	S	Coffee farm	Fredonia	1786	5.9747, -75.6644	24	59
RUBCA03012	20190627	17:07	Rubiaceae	*Coffea arabica*	Pajarito	S	Coffee farm	Fredonia	1786	5.9746, -75.6643	24	59
RUBCA03013	20190627	17:10	Rubiaceae	*Coffea arabica*	Castillo	S	Coffee farm	Fredonia	1786	5.9748, -75.6645	24	59
RUBCA03015	20190627	17:20	Rubiaceae	*Coffea arabica*	Castillo	S	Coffee farm	Fredonia	1786	5.9740, -75.6645	24	59
RUBCA05001	20190703	11:15	Rubiaceae	*Coffea arabica*	N/A	A	University campus	Medellín	1504m	6.2002, -75.5785	23	64

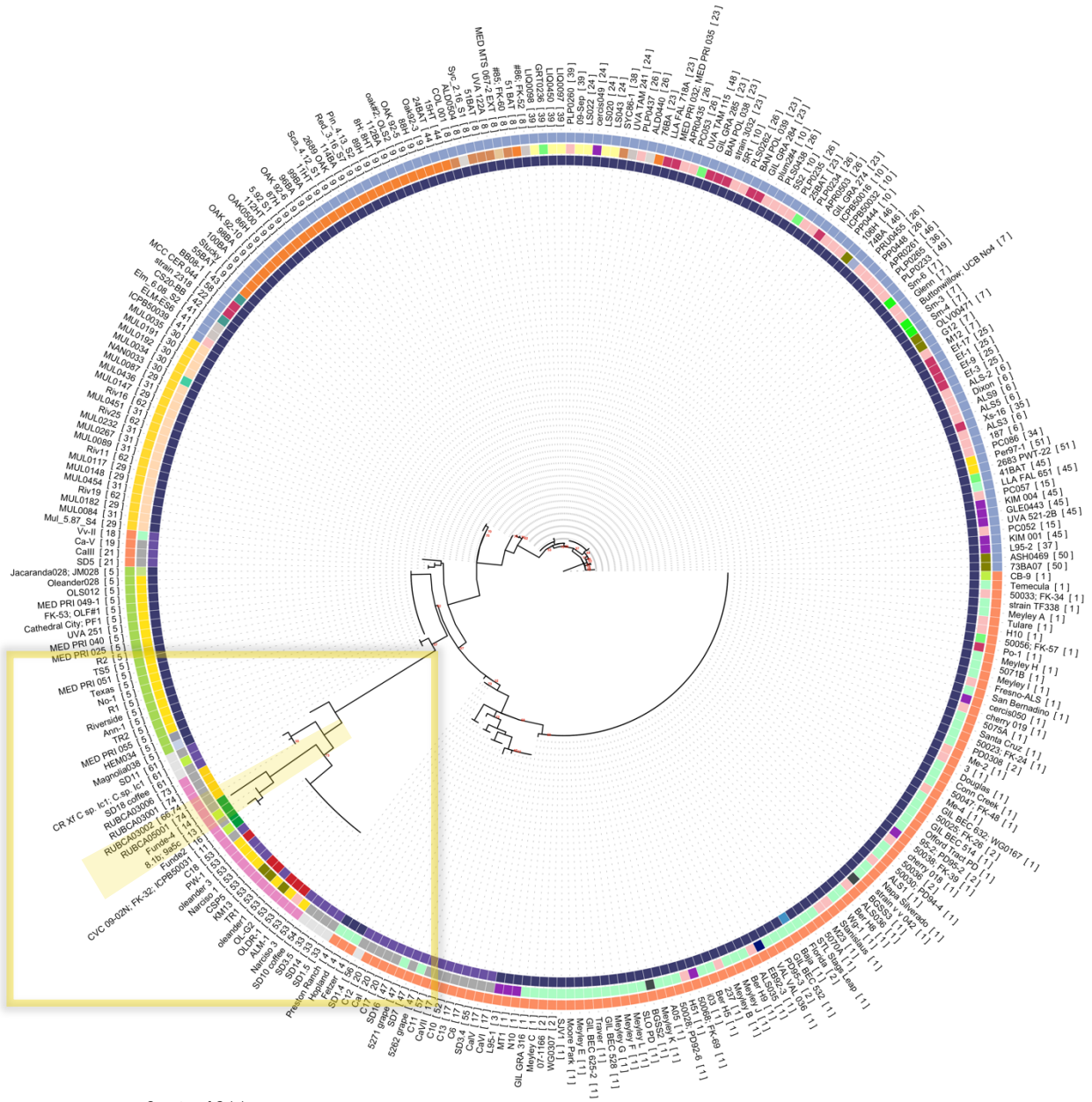


Figure 5: *Xylella fastidiosa* (*Xf*) multi-locus sequence type (MLST) phylogeny including Colombian strains. An *Xf* MLST phylogenetic tree showing Colombian strains RUBCA03001, RUBCA03002, RUBCA03006 and RUBCA05001 (position in phylogeny highlighted in **yellow**; see **Figure 6** for a detailed view) aligned with publicly available *Xf* MLST profiles obtained from the PubMLST database (Jolley, Bray and Maiden, 2018; https://pubmlst.org/xfastidiosa/). Concatenated sequences of four target genes (*cysG*, *holC*, *nuoL*, *gltT*) from the MLST scheme refined by Yuan et al. (2010) were aligned with MUSCLE (Edgar, 2004) and the phylogeny was inferred by maximum-likelihood using IQ-TREE (Nguyen et al., 2015). Bootstrap values below 75 are indicated in **red** at the respective nodes. The innermost coloured circle specifies the country of origin, the second circle the host family from which the strain was obtained, and the outermost circle the *Xf* subspecies (EPPO, 2018) of each strain. The sequence type (ST) of each strain is stated in brackets next to the strain name at the tip labels.

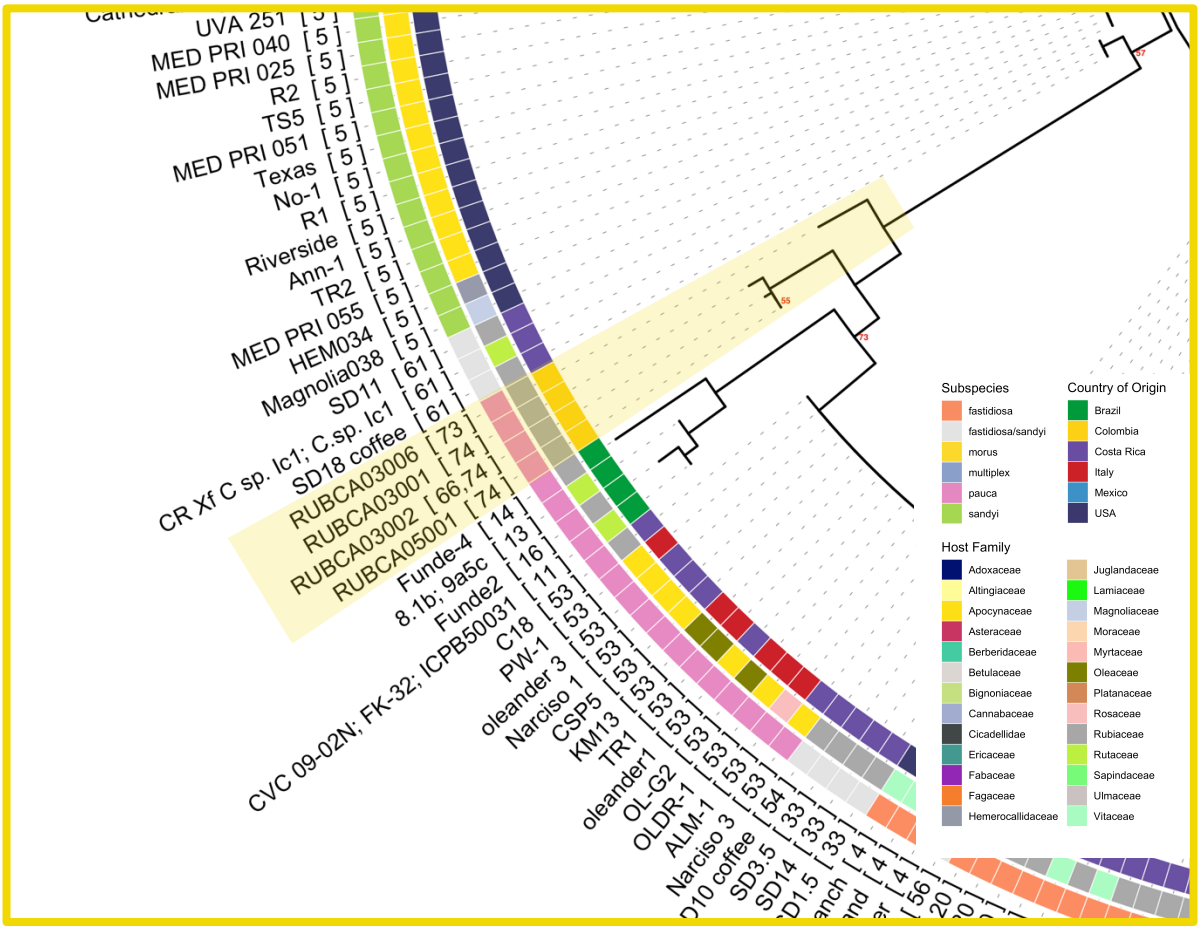


Figure 6: Spotlight on *Xylella fastidiosa* (*Xf*) multi-locus sequence type (MLST) phylogeny of Colombian strains. A detailed view of the MLST phylogeny from **Figure 5** including Colombian *Xf* strains (highlighted in **yellow**). Colombian strains cluster together with all other *pauca* strains (marked with **pink boxes** in the outermost coloured circle) originating from Brazil, Costa Rica, and Italy. The Colombian strains appear to have arisen from Cota Rican *pauca* strains.

Discussion

The presence of X. fastidiosa in Colombia

Xf is a familiar plant pathogen in the Americas. In Central and South America in particular, *Xf* is known to cause disease in citrus, coffee and *Prunus* spp. In Brazil, *Xf* is especially devastating as it is known to be the cause of CVC, a disease resulting in smaller and lower quality fruits, directly impacting the country's economy. In South America, *Xf* is also known to affect *C. arabica*, where it causes leaf scorch symptoms. *Coffea* spp. is a plant host known to harbour all four subspecies of interest – *fastidiosa, multiplex, pauca* and *sandyi* – which are

also found in Europe (EFSA, 2018). Interestingly, in many nearby countries, *Xf* has been detected in *C. arabica*, e.g. Venezuela, Brazil, Paraguay and Puerto Rico. However, no report of either the presence, nor absence of *Xf* in Colombia is available despite Colombia's leading market role in coffee. In this research, *Xf* was detected for the first time in *C. arabica* in Colombia. Samples of five different cultivars of *C. arabica* of a coffee farm in Fredonia, which lies in the Antioquia department, have been collected and *Xf* was detected in plants of all five cultivars. Samples from the families Malvaceae and Rutaceae have also been collected, but no *Xf* could be detected in those samples. However, this could be false negatives, as the detection by PCR may have been limited due to low *Xf* concentration in the sample and the PCR not being powerful enough to detect these concentrations.

The XF1 PCR was repeated on all thirteen positive Colombian samples and the positive control *Xf* subsp. *fastidiosa* strain Temecula-1 with Platinum *Taq* polymerase, a high-fidelity polymerase. High-fidelity polymerases provide better specificity during the replication process in PCR. All positive samples (RUBCA03001, RUBCA03002, RUBCA03003, RUBCA03005, RUBCA03006, RUBCA03007, RUBCA03008, RUBCA03010, RUBCA03011, RUBCA03012, RUBCA03013, RUBCA03015 and RUBCA05001) and the positive control were sent for Sanger sequencing. MSA of consensus sequences of all samples and the positive control show several differences between the sequences (see **Appendix Figure D**). This confirmed that the positive amplification of the Colombian samples was not in fact contamination from the positive control *Xf* subsp. *fastidiosa* strain Temecula-1, which was used throughout the detection process. An initial sequence comparison search using BLAST on NCBI of the XF1 amplicons suggested that the samples are of subspecies *pauca*, which may indicate a relation with *Xf* coffee strains in Brazil or Costa Rica.

Subspecies *pauca* was later confirmed in four Colombian strains (RUBCA03001,

RUBCA03002, RUBCA03006, RUBCA05001) by MLST. An MLST phylogeny was created with available *Xf* strains obtained from the PubMLST database (see **Figure 5**). The location of the Colombian strains is indicated with a red arrow, which cluster with other *Xf* subsp. *pauca* strains in the phylogeny. These other *pauca* strains originate from Brazil, Costa Rica and Italy. The Italian CoDiRo strain associated with OQDS in Apulia is believed to have evolved in Costa Rica and eventually introduced via a coffee plant to Italy (Marcelletti, and Scortichini, 2016b). No whole genome data could be found of other *Xf* strains with the same ST profiles as the Colombian *pauca* strains (ST66, 73 and 74). The ST profile of RUBCA03002 (ST66/74) was inconclusive which most likely resulted from sequencing errors due to low-quality DNA. However, both ST profiles belong to the *pauca* subspecies (EPPO, 2018).

C. arabica cv. Caturra was previously found to harbour X. fastidiosa in Costa Rica

Interestingly, a coffee plant of cultivar Caturra was also found to harbour *Xf* in Costa Rica (Rodríguez, et al., 2001). This cultivar was also sampled in Colombia and *Xf* was detected in two samples (RUBCA03007, RUBCA03008) with mild leaf scorch symptoms. Severe symptoms were described in the Costa Rican plants, although these differed from the typical leaf scorch symptoms classified as 'severe' in Brazil. Instead, the symptoms in the Costa Rican plants included leaf malformation and curling of edges, leaf chlorotic mosaic, and shortening of internodes. Unfortunately, the genome of the Costa Rican strains detected in *C. arabica* cv. Caturra are not available, and therefore it is unknown what *Xf* subspecies the plants harboured. *Xf* subsp. *fastidiosa* is suggested to be more common in Costa Rica (Montero-Astúa, et al., 2008; Rodríguez, et al., 2001), though infection by *Xf* subsp. *fastidiosa* in coffee in Costa Rica is less severe than infection by *Xf* subsp. *pauca* in Brazil (Li, et al., 2001). The Costa Rican Caturra strain is possible to be of subspecies *pauca* as this subspecies has also been detected in the country. If the Costa Rican strain was indeed *Xf* subsp. *pauca*, it would be interesting to compare it to the identified Colombian strains and

Brazilian strains causing severe CLS.

X. fastidiosa is more widely spread than currently known

The first detection of *Xf* in Colombia demonstrates that the bacterium is more widely spread than currently known. Four Colombian strains were confirmed to be of subspecies *pauca*, a subspecies primarily found in citrus and coffee in South America (Scally, et al., 2005; Schuenzel, et al., 2005). *Xf* subsp. *pauca* strain CoDiRo, which is associated with the Italian OQDS outbreak in Apulia, is suggested to have evolved in Costa Rica and eventually introduced via a coffee plant to Italy (Marcelletti, and Scortichini, 2016b). An invasive species in Italian olive groves, the local ecosystem was not prepared for *Xf* and therefore the bacterium was especially aggressive there. Currently, *Xf* research mainly focuses on economically significant plant hosts. Often, potential pathogens are only of interest when they cause outbreaks. However, pathogens that remain asymptomatic in plant hosts can also provide researchers with a lot of information. Asymptomatic hosts can help understand the pathogen's mode of action and why it remains a harmless endophyte in some hosts but causes devastating diseases in others. Undeniably, there could be many more countries similar to Colombia, where the bacterium resides in plant hosts as a harmless endophyte, particularly in the Americas, where *Xf* is endemic.

Conclusion

This first finding of *Xf* in *C. arabica* in Colombia is plausible as the bacterium is believed to be native to the Americas (Nunney et al., 2014) which is also confirmed by an MLST phylogeny with PubMLST *Xf* isolates (see **Figure 5**). *Xf* must therefore be more widely spread in the continent than currently known. Many more countries might be harbouring the bacterium without harm to endemic plants. Conducting *Xf* surveys in previously undetected countries will

increase the understanding of the genetic differences between the different strains and how these affect crops that are more susceptible to the bacterium.

Putative effectors of *X. fastidiosa*

Introduction

Bacterial effectors and their role in pathogenesis

Plants have evolved the ability to recognise molecular patterns of invading microorganisms, known as microbe- or pathogen-associated molecular patterns (MAMPs or PAMPs). PAMPs are recognised by pattern-recognition receptors (PRRs), molecular receptors of plant cells. Upon recognition, PAMP-triggered immunity (PTI) is activated (Pritchard and Birch, 2014). Some pathogens have the ability to overcome this basal host immunity by secreting effector proteins. These effectors are one of the best-characterised virulence factors. Virulence factors include proteins, lipids and carbohydrates produced by the pathogen and are involved in host invasion, disease progression and host defence evasion. Effectors, in particular, are also involved in enhancing disease susceptibility of the host by altering cellular responses and modulating transcription (Pelgrom, and Van den Ackerveken, 2016). Furthermore, effectors have a key role in the outcome of plant-pathogen interactions by functioning as both virulence and avirulence factors (Alfano, and Collmer, 2004; Khan, et al., 2018). Effectors may act as disease elicitor (virulence factor), or disease suppressor (avirulence factor) by altering the response of the plant immune system (Surico, 2013).

Bacterial secretion systems of gram-negative bacteria

Gram-negative bacteria secrete effectors into their surroundings or translocate them into a host cell through secretion systems (SS). These may be divided into Sec- and Tat-dependent, or Sec- and Tat-independent secretion pathways (Green, and Mecsas, 2016). Proteins are delivered to the periplasm through the Sec or Tat pathway. Proteins translocated through the Sec pathway are in their unfolded state, proteins translocated through the Tat pathway are in their folded state. More complex secretion pathways of effector proteins include Type 1 through Type 6 SS. Proteins secreted through T3SS and T4SS are among the most

extensively studied in structure and function. The majority of bacterial phytopathogens have been found to secrete effectors through the T3SS, for example *Pseudomonas syringae*, *Erwinia* spp. and *Xanthomonas* spp. (Buttner, and He, 2009). *Xf* lacks genes that make up the T3SS (Simpson, *et al*., 2000), one of the most extensively studied secretion systems (Dow, and Daniels, 2000), and instead encodes essential genes that make up components of T1SS, T2SS, T4SS and T5SS either within the bacterial chromosome (Simpson, et al., 2000; Sluys, et al., 2003) or on plasmids (Rogers and Stenger, 2012).

T1SS, T2SS and T5SS mainly aid in the secretion of effectors to be exported to the extracellular environment (Wandersman, 2013). T3SS and T4SS aid in effector translocation directly into host cells (Block, and Alfano, 2011; Zechner, Lang, and Schildbach, 2012; Russell, Peterson, and Mougous, 2014; Basler, 2015). Effectors secreted through T6SS mainly act in interactions with bacterial competitors, but have recently been found to also target eukaryotic cells *in vitro* (Boyer, et al., 2009; Bernal, Llamas, and Filloux, 2018; Lien, and Lai, 2017), such as Hcp and VgrG secreted by *Pectobacterium atrosepticum*, a bacterial phytopathogen that causes soft rot in potato. The two effectors were found to be expressed and target VasK during infection with potato. Gene knockout of *vasK* increased virulence in the host (Mattinen, et al., 2008; Ryu, 2015).

Similar to *Xf*, a number of bacterial phytopathogens have been described to lack a T3SS, such as the two Xanthomonadaceae species *X. albilineans* and *Xanthomonas sacchari* (5 Mb). Compared to other *Xanthomonas* spp., *X albilineans* also has a reduced genome size (3.8 Mb) and has the ability to invade the xylem, particularly of sugarcane (Jackson, et al., 2011; Pieretti, et al., 2009, 2015). *X. sacchari* also lacks genes that make up a T3SS and T6SS (Fang, et al., 2015). *Xylella taiwanensis*, a *Xylella* species that has been associated with pear

leaf scorch in Taiwan lacks a T3SS and T6SS, but unlike *Xf* also lacks a T4SS (Kanehisa and Goto, 2000; Kanehisa, 2019; Kanehisa et al., 2021; Su et al., 2016).

Aim of research

Effectors have a major role in pathogenesis and only few have been determined in *Xf.* The aim of this research was to identify effectors of *Xf* using the software PREFFECTOR developed by Dhroso, Eidson and Korkin (2018), which allows for the prediction of bacterial effectors secreted by T1SS through T6SS.

Methods

Host range of X. fastidiosa

An extensive list of documented *Xf* plant hosts was curated. Sources included the European Commission (EC, 2019), the European Food Safety Authority (EFSA, 2018) and the European and Mediterranean Plant Protection Organisation (EPPO, n.d.). The list includes details such as the *Xf* subspecies that was found in the plant, whether this was found in a natural or experimental setting, in which European country these were found, and diseases found in each plant host (see **Appendix Table C**). Wherever subspecies information was available for a host, a visualisation of the host range of each *Xf* subspecies was created using the R package *VennDiagram* (Chen, and Boutros, 2011).

Phylogeny inference

A total of 55 *Xf* complete and draft genomes and one *Xylella taiwanensis* complete genome was obtained from NCBI's GenBank database in November 2019 (see **Appendix Table D** for details of each genome). *X. taiwanensis* was used as an outgroup in phylogeny inference. Genomes were annotated with Prokka (Seemann, 2014), and filtered based on N50 statistics and contig number according to a paper published by Levy *et al*. in 2018, and CheckM to remove contaminated and incomplete genomes (Parks, et al., 2015). The core genome was

determined by identifying orthologous sequence groups, descendants of the same ancestral sequence that were separated due to speciation, between the genomes with OrthoFinder (Emms, and Kelly, 2015) using default parameters. Protein sequences were subsequently aligned with ClustalW (Thompson, Higgins, and Gibson, 1994), corrected with GBlocks (Castresana, 2000; Talavera, and Castresana, 2007), and protein models tested using ProtTest (Abascal, Zardoya, and Posada, 2005). Finally, a phylogeny was inferred by maximum-likelihood with IQ-Tree (Nguyen, et al., 2015) on concatenated protein sequence alignments of single-copy orthologous groups and visualised using the R package *ape* (Paradis, Claude, and Strimmer, 2004). The detailed pipeline can be found on GitHub (https://github.com/mirloupa/Xf/tree/master/Phylogeny).

Effector prediction and sequence orthology

Putative effectors were identified using the PREFFECTOR program released by Dhroso, Eidson and Korkin (2018). The program requires protein sequences of interest in FASTA format, which are uploaded to the PREFFECTOR web-server (http://korkinlab.org/preffector). Effector prediction was performed on all coding sequences for 55 *Xf* genomes (mean of 2,180 proteins per genome), one *X. taiwanensis* genome (2,143 proteins) and two *Xanthomonas* genomes (mean of 3,972 proteins per genome).

The probability of a sequence coding for a protein functioning as an effector is specified by a number between 0 and 1, where 0 indicates uncertainty of the probability and 1 indicates certainty. Only putative effector sequences with a probability of 1 have been included in further analyses (see **Appendix Table F** for the full list of PREFFECTOR results). 570 putative effectors were analysed for sequence homology by assembling these into orthologous groups. Sequences were annotated using the Prokka database (Seemann, 2014) and orthologues were identified and grouped using OrthoFinder (Emms, and Kelly, 2015). The presence and

absence of orthologous groups in each *Xf* species were translated into a matrix on *R* and mapped to an *Xf* phylogeny previously created.

Results

Host range of X. fastidiosa

From the curated list of *Xf* hosts, the presence of the subspecies most prevalent in different European countries was identified. Subspecies *fastidiosa* was found in Germany and Spain but is most prevalent in Spain. Subspecies *multiplex* was found in France, Portugal and Spain, and is most prevalent in France. Subspecies *pauca* was found in France, Italy and Spain, and is most prevalent in Italy. The subspecies *sandyi* has only been identified in France. A Venn diagram was then produced to visualise the number of hosts shared between the four subspecies of interest (Chen, and Boutros, 2011). Information of subspecies isolated from different hosts (see **Appendix Table C**) was used to create this Venn diagram of shared hosts (see **Figure 7**). *Xf* subsp. *multiplex* has the largest host range, with 88 hosts only affected by the subspecies. 48 hosts are affected by *Xf* subsp. *fastidiosa* only, 20 hosts solely by *pauca* and three by *sandyi*. Four plant hosts are shared among all four subspecies: *Coffea* spp. (coffee), *Nerium oleander* (oleander), *Polygala myrtifolia* (myrtle-leaf milkwort) and *Prunus dulcis* (almond). No shared plant hosts exist between the following subspecies:

- multiplex vs sandyi
- pauca vs sandyi
- fastidiosa vs multiplex vs sandyi
- fastidiosa vs pauca vs sandyi
- multiplex vs pauca vs sandyi

Finally, the curated list of *Xf* host range, subspecies and location information each was later used to determine any signatures in effector evolution (see **Figure 11**).

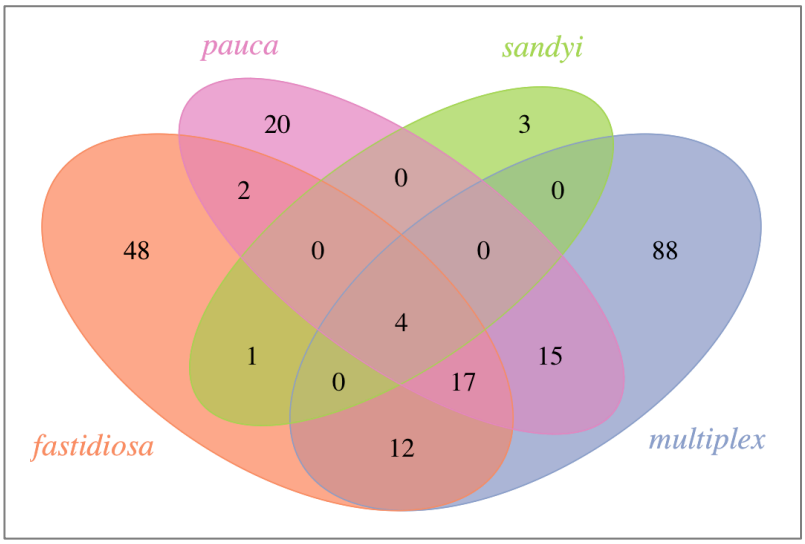


Figure 7: Venn diagram of shared host plants between *Xylella fastidiosa* (*Xf*) subspecies. In total, information of 206 plant hosts wherein *Xf* was detected was collected from EC (2018), EFSA (2018) and EPPO (n.d.). This diagram, created using the R package *VennDiagram* (Chen, and Boutros, 2011), depicts the number of hosts solely found in each subspecies and shared among other subspecies. A detailed list of the different subspecies detected in these plant hosts is found in **Appendix Table C**. Subspecies *multiplex* has the largest host range, whereas *sandyi* has the smallest. Four hosts are shared among all four subspecies: *Coffea* sp. (coffee), *Nerium oleander* (oleander), *Polygala myrtifolia* (myrtle-leaf milkwort) and *Prunus dulcis* (almond).

A phylogenetic tree of X. fastidiosa strains

A core-genome phylogeny of 55 *Xf* and one *X. taiwanensis* was generated (see **Figure 8**). The subspecies information of a number of *Xf* strains was available on NCBI GenBank, where the genomes were obtained from. Strains with known subspecies grouped together in the phylogenetic tree, thus allowing the inference of the subspecies of the other strains. The subspecies of these strains were later confirmed as described in the methods section above and included in the phylogeny. Origin information was present for most of the genomes which are coloured according to the country where the strain was found in. All South American strains appear to be of the subspecies *pauca*. Subspecies clearly group together in the phylogeny with the exception of *Xf* subsp. *fastidiosa* strain 6c grouping within the *Xf* subsp. *pauca* clade. Hosts are spread across the entire phylogeny with some convergence of different subspecies. For example, the plant family Vitaceae (e.g. grapevine) appears to only be infected by *Xf* subsp. *fastidiosa*, but the plant family Rosaceae is infected by *Xf* subsp. *fastidiosa*, *multiplex* and *pauca*. Bootstrap values are predominantly high with five exceptions where bootstrap values are below 70. This is most likely due to recombination. Recombination is the exchange of DNA between multiple organisms. In bacteria, this can occur through conjugation directly via cell-to-cell contact, transformation by the uptake of exogenous DNA from the surroundings of the bacterial cell, and transduction via virus-mediated DNA transfer. Recombination results in unreliable reconstruction of the tree topology (Hedge, and Wilson, 2014; Posada, and Crandall, 2002). However, González-Torres et al. (2019) have found that bootstrap support does not always provide sufficient information about the accuracy of a phylogeny. They have found that recombination may result in incorrect tress even with high bootstrap support.

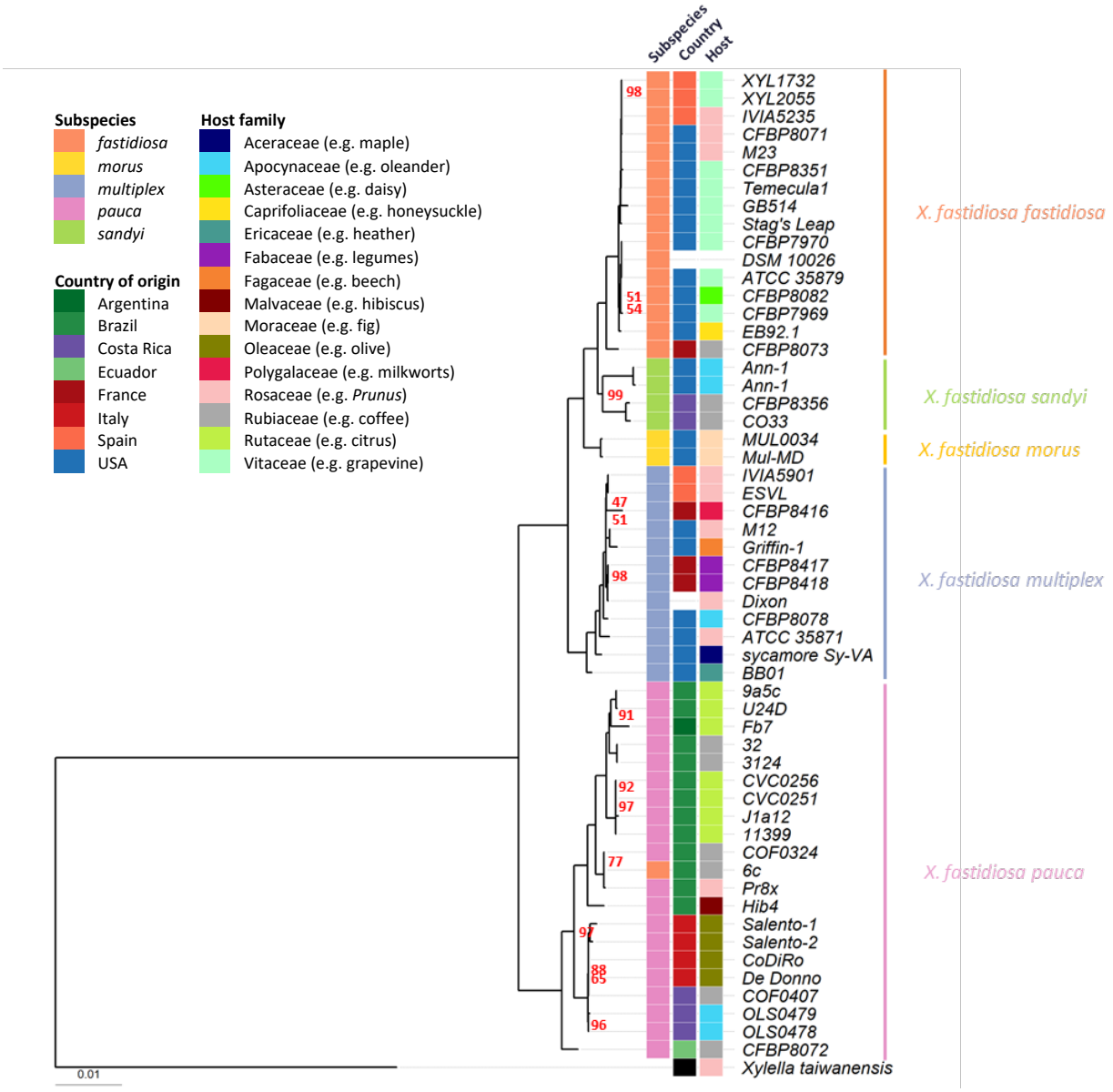


Figure 8: Phylogenetic tree of *Xylella fastidiosa* (*Xf*). A phylogenetic tree of 55 *Xf* and the *Xylella taiwanensis* genome (used outgroup) was created. This tree was generated using IQ-Tree's multiple sequence alignment by maximum-likelihood (Nguyen, et al., 2015). Bootstrap values below 100 are indicated in red. The tree was visualised using the *ape* package on R (Paradis, Claude and Strimmer, 2004). Location and host origin (where available), and subspecies information are highlighted in different colours.

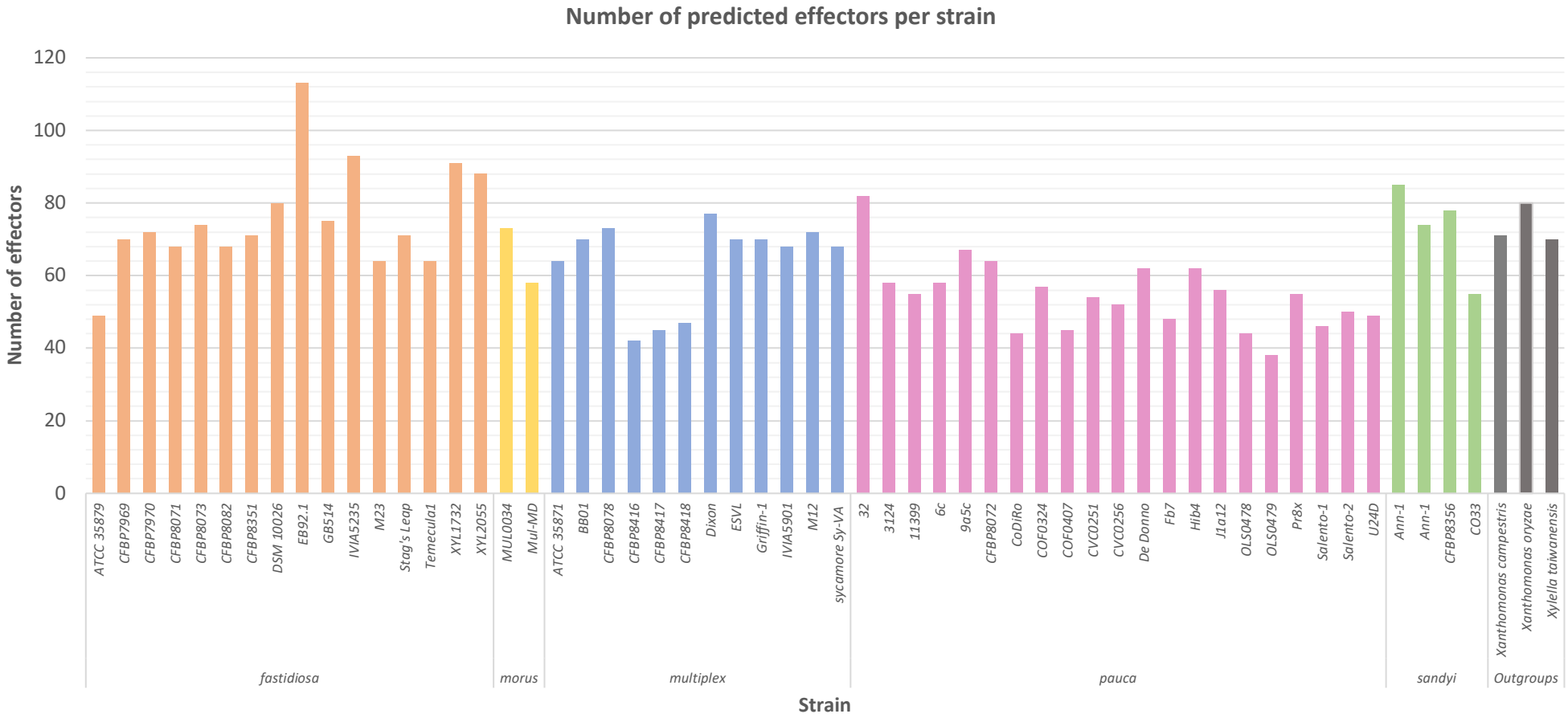


Figure 9: Number of predicted effectors per *Xylella fastidiosa* (*Xf*) strain. Effector prediction was performed using the PREFFECTOR software (Dhroso, Eidson and Korkin, 2018), which predicts effectors across all six secretion systems. Protein sequences of 55 *Xf* genomes, one *Xylella taiwanensis* genome and two *Xanthomonas* genomes were obtained from NCBI for analysis. A graphical interpretation of the number of effectors predicted per input genome. The colours indicate the different *Xf* subspecies. *Xf* subsp. *fastidiosa* strain EB92.1, which is associated with less pathogenic symptoms, interestingly has the highest number of predicted effectors.

Figure 10: Word cloud of predicted *Xylella fastidiosa* (*Xf*) effector sequences. This word cloud shows the resulting predicted effectors from the program PREFFECTOR (Dhroso, Eidson and Korkin, 2018) which predicts effectors across all six secretion systems. The resulting sequences were categorised and finally visually summarised using the R package wordcloud **(**Fellow, 2012). The majority of predicted effector sequences are hypothetical and uncharacterised proteins. Further analyses will be done to predict the functions of these by looking at the motifs, structure and similarity search. The font colours have no meaning and merely serve as visual aid.

Prediction of effector proteins in X. fastidiosa

Amino acid sequences of 55 *Xf* genomes, two *Xanthomonas* genomes and one *X. taiwanensis* genome were acquired from NCBI's GenBank database and uploaded to the PREFFECTOR webserver. As an output, a table was produced for each genome, listing the following information: a database ID generated by PREFFECTOR, a sequence ID identifying the sequence number within the original FASTA input file, the default minimum probability threshold of 0.9, the predicted probability calculated by PREFFECTOR, the effector categorisation, and the original FASTA sequence header of the predicted effector (see **Appendix Table F**). In total, 3,767 putative effectors were predicted by the software across the 58 genomes of interest, 3,546 of which were identified in *Xf* genomes. Interestingly, *Xf* strain EB92.1, a strain that appears to be less pathogenic than other *Xf* strains (Hopkins, 1951), has the largest number of predicted effectors (see **Figure 9**).

Proteins predicted by PREFFECTOR shows that the majority of sequences have not been characterised yet (see **Figure 10**; Fellows, 2012), which is not uncommon as the function of the majority of sequences in the genome is not yet known. Many predicted effectors of which the sequences have been previously described include various enzymes, transport proteins, membrane proteins, receptors, and haemagglutinins – which are described to be crucial in biofilm formation (De Souza, et al., 2003; Guilhabert, and Kirkpatrick, 2005).

A total of 570 coding sequences across 55 *Xf* strains, one *X. taiwanensis*, one *X. oryzae* and one *X. campestris* genome were included in the homologous sequence analysis. 453 of these sequences were grouped into 52 orthologous groups (OGs). The function of a number of proteins within seven OGs are known (see **Table 4**; see **Appendix Table F** for full list). Proteins within OG X1 have a cellulase function. X2 includes cold-shock proteins, and X15 quinones and alcohol dehydrogenases. Proteins within OG X29 are involved in with RTX toxins, X37

includes a haemagglutinin protein, X44 comprises of 30S ribosomal proteins only, and OG X50 includes proteins involved in surface adhesion. Most notably, OG X5 appears to only be present in *Xf* subsp. *pauca* (see **Figure 11**). No other OG pattern such as this is seen in the other subspecies.

Table 4: Function of *Xylella fastidiosa* (*Xf*) orthologous sequences. Orthologous groups, or orthogroups, of putative *Xf* effectors were determined using the PREFFECTOR program created by Dhroso, Eidson and Korkin (2018). Proteins with known function are listed below.

Orthogroup	Strain	Accession ID	Protein
X1	3124	ALQ97392.1	1,4-beta-cellobiosidase
X1	11399	OCA57933.1	1,4-beta-cellobiosidase
X1	6c	OJZ70903.1	1,4-beta-cellobiosidase
X1	9a5c	WP_010893338.1	endoglucanase
X1	9a5c	WP_010893773.1	1,4-beta-cellobiosidase
X1	Ann-1c	WP_024748856.1	1,4-beta-cellobiosidase partial
X1	ATCC_35879	KGM20724.1	1,4-beta-cellobiosidase partial
X1	CFBP7969	WP_128723174.1	1,4-beta-cellobiosidase
X1	CFBP7970	WP_128712456.1	1,4-beta-cellobiosidase
X1	CFBP8071	WP_128712519.1	1,4-beta-cellobiosidase
X1	CFBP8072	WP_058569679.1	1,4-beta-cellobiosidase
X1	CFBP8078	WP_128723671.1	4-beta-cellobiosidase
X1	CFBP8351	WP_128712519.1	1,4-beta-cellobiosidase
X1	CFBP8356	WP_128734966.1	1,4-beta-cellobiosidase
X1	CFBP8416	OMJ97057.1	1,4-beta-cellobiosidase
X1	CFBP8417	OMK00128.1	1,4-beta-cellobiosidase
X1	CFBP8418	OMJ99939.1	1,4-beta-cellobiosidase
X1	COF0324	KXB21420.1	1,4-beta-cellobiosidase
X1	CVC0251	KXB21968.1	1,4-beta-cellobiosidase
X1	CVC0256	KXB13296.1	1,4-beta-cellobiosidase
X1	Dixon	EAO14376.1	cellulase
X1	DSM_10026	SHG20270.1	cellulose binding domain-containing protein partial
X1	EB92.1	EGO81204.1	cellobiohydrolase A (1,4-beta-cellobiosidase A) partial
X1	EB92.1	EGO81385.1	endoglucanase BglC partial
X1	EB92.1	EGO82960.1	cellobiohydrolase A (1,4-beta-cellobiosidase A) partial
X1	ESVL	WP_128382978.1	1,4-beta-cellobiosidase
X1	Fb7	AWG45316.1	1,4-beta-cellobiosidase
X1	Griffin-1	ERI59813.1	1,4-beta-cellobiosidase
X1	Hib4	ALR07014.1	1,4-beta-cellobiosidase
X1	IVIA5235	RHW37904.1	1,4-beta-cellobiosidase
X1	IVIA5901	WP_128283863.1	hypothetical protein
X1	J1a12	ALR01763.1	1,4-beta-cellobiosidase
X1	M12	ACA11602.1	cellulase
X1	M23	ACB91997.1	cellulose 1,4-beta-cellobiosidase
X1	Mul-MD	EWG14499.1	cellulase
X1	MUL0034	AIC13557.1	hypothetical protein P303_02185
X1	Pr8x	ALR04597.1	1,4-beta-cellobiosidase
X1	Stags_Leap	WP_081095287.1	1,4-beta-cellobiosidase
X1	Temecula1	AAO28402.1	cellulose 1,4-beta-cellobiosidase

Orthogroup	Strain	Accession ID	Protein
X1	U24D	ALQ94365.1	endoglucanase
X1	U24D	ALQ94677.1	1,4-beta-cellobiosidase
X1	XYL1732	RUA39812.1	1,4-beta-cellobiosidase
X1	XYL2055	RUA38669.1	1,4-beta-cellobiosidase
X1	*Xylella taiwanensis*	WP_069636213.1	1,4-beta-cellobiosidase
X2	32	ETE34180.1	cold-shock protein
X2	3124	ALQ96546.1	cold-shock protein
X2	11399	OCA57322.1	cold-shock protein
X2	9a5c	WP_010894798.1	cold-shock protein
X2	Ann-1c	WP_004085832.1	cold-shock protein
X2	Ann-1f	AIC10508.1	cold-shock protein
X2	ATCC_35871	WP_004085832.1	cold-shock protein
X2	BB01	WP_004085832.1	cold-shock protein
X2	CFBP7969	WP_004085832.1	cold-shock protein
X2	CFBP7970	WP_004085832.1	cold-shock protein
X2	CFBP8071	WP_004085832.1	cold-shock protein
X2	CFBP8072	WP_010894798.1	cold-shock protein
X2	CFBP8073	WP_004085832.1	cold-shock protein
X2	CFBP8078	WP_004085832.1	cold-shock protein
X2	CFBP8082	WP_004085832.1	cold-shock protein
X2	CFBP8351	WP_004085832.1	cold-shock protein
X2	CFBP8356	WP_004085832.1	cold-shock protein
X2	CFBP8416	OMJ96975.1	cold-shock protein
X2	CoDiRo	KIA57572.1	cold-shock protein
X2	De_Donno	ARO68197.1	cold-shock protein
X2	Dixon	EAO12779.1	cold-shock protein DNA-binding
X2	DSM_10026	SHG79508.1	cold-shock DNA-binding protein family
X2	EB92.1	EGO81051.1	cold shock protein
X2	ESVL	WP_004085832.1	cold-shock protein
X2	GB514	ADN62224.1	cold shock protein
X2	Griffin-1	ERI60141.1	1,4-beta-cellobiosidase
X2	Hib4	ALR06040.1	cold-shock protein
X2	IVIA5235	RHW42932.1	cold-shock protein
X2	IVIA5901	WP_004085832.1	cold-shock protein
X2	J1a12	ALR01430.1	cold-shock protein
X2	M12	ACA12436.1	putative cold-shock DNA-binding domain protein
X2	M23	ACB92876.1	cold-shock DNA-binding domain protein
X2	MUL0034	AIC12651.1	cold-shock protein
X2	Pr8x	ALR03814.1	cold-shock protein
X2	Stags_Leap	WP_004085832.1	cold-shock protein
X2	Temecula1	AAO29227.1	cold shock protein
X2	U24D	ALQ95444.1	cold-shock protein
X2	XYL1732	RUA38378.1	cold-shock protein
X2	XYL2055	RUA37809.1	cold-shock protein
X2	*Xylella taiwanensis*	WP_038270170.1	cold-shock protein
X15	ATCC_35879	KGM20025.1	polyvinylalcohol dehydrogenase
X15	CFBP8073	WP_058564468.1	polyvinylalcohol dehydrogenase
X15	Dixon	EAO13916.1	quinoprotein
X15	DSM_10026	SHG72011.1	polyvinyl alcohol dehydrogenase (cytochrome)
X15	Griffin-1	ERI60323.1	hypothetical protein M233_04935
X15	M23	ACB92802.1	Pyrrolo-quinoline quinone
X15	Mul-MD	EWG15353.1	Pyrrolo-quinoline quinone
X15	MUL0034	AIC13805.1	polyvinylalcohol dehydrogenase
X15	sycamore_Sy-VA	KFA42156.1	hypothetical protein DF22_001285

Orthogroup	Strain	Accession ID	Protein
X15	*Xylella taiwanensis*	WP_081755433.1	polyvinylalcohol dehydrogenase
X29	CFBP8356	WP_057683294.1	hypothetical protein
X29	CO33	KQH73482.1	RTX toxin
X29	Mul-MD	EWG15232.1	hypothetical protein P910_001531
X29	MUL0034	AIC12677.1	RTX toxin Ca2+-binding protein
X37	EB92.1	EGO82688.1	hemagglutinin/hemolysin partial
X37	XYL1732	RUA34462.1	hypothetical protein DX878_11735 partial
X37	XYL2055	RUA34494.1	hypothetical protein DX877_11780 partial
X44	CFBP8078	WP_128723706.1	30S ribosomal protein THX
X44	*Xylella taiwanensis*	WP_081755402.1	30S ribosomal protein THX
X50	EB92.1	EGO81883.1	autotransporter adhesin partial
X50	IVIA5235	RHW48442.1	cell surface protein partial

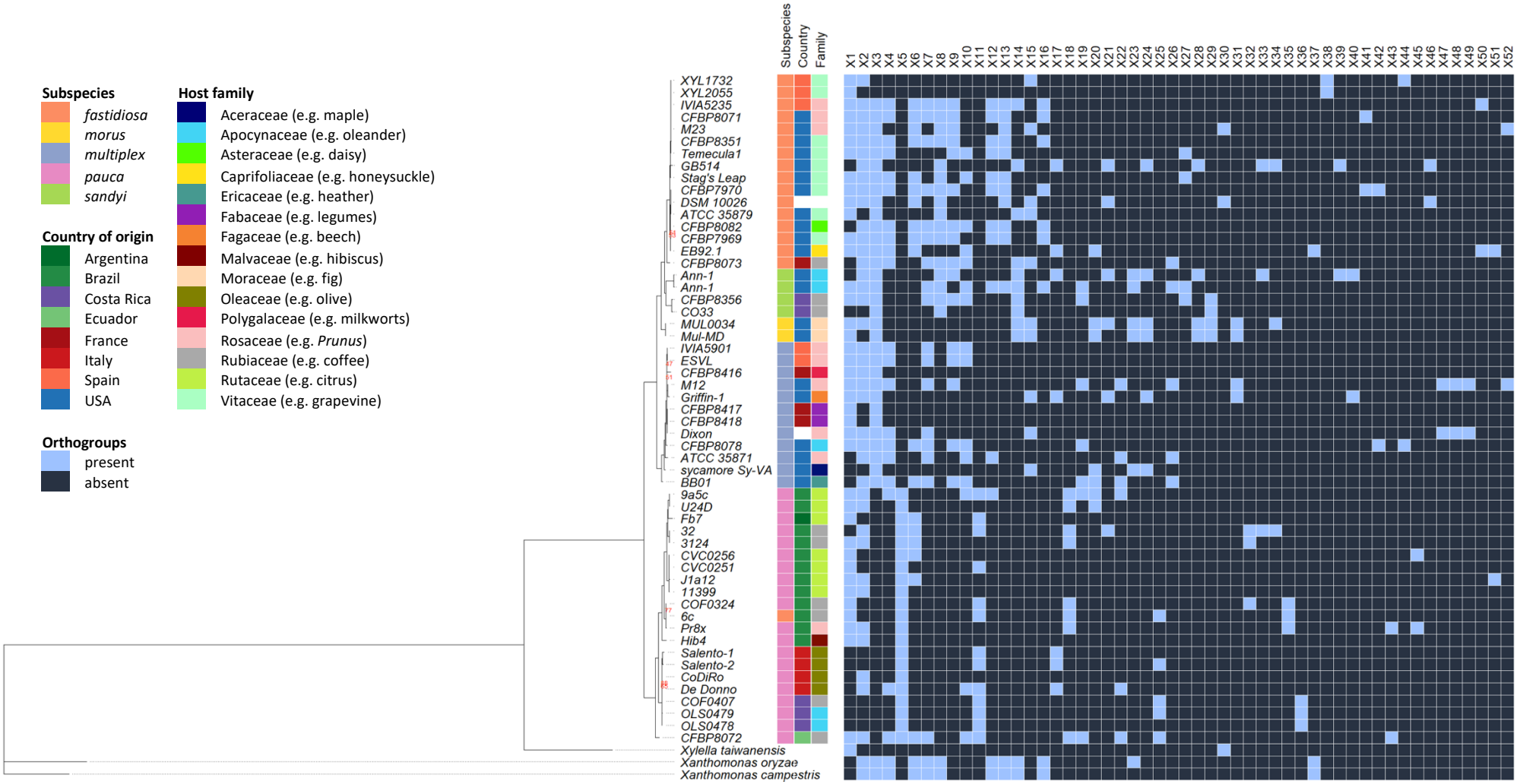


Figure 11: Presence-absence of *Xylella fastidiosa* (*Xf*) orthology groups mapped to a core-genome phylogeny. 570 putative effectors from 55 *Xf* strains, one *Xylella taiwanensis*, one *Xanthomonas oryzae* and one *Xanthomonas campestris* genome were discovered using the PREFFECTOR program (Dhroso, Eidson and Korkin, 2018) and analysed for sequence homology. Orthologous groups were identified and using OrthoFinder (Emms and Kelly, 2015). The presence and absence of orthologous groups in each *Xf* species were translated into a matrix on R and mapped to an *Xf* phylogeny previously created to determine evolutionary signatures.

Discussion

X. fastidiosa has a broad plant host range

A comprehensive literature search was conducted to assess the host-range of *Xf*. It was found that the different subspecies of interests share some hosts but can also be very host-specific. However, sampling bias – i.e. plants with *Xf*-symptoms are more likely to be tested for the bacterium – as well as a concentration too small for any molecular test to detect any bacteria could be limiting the knowledge of the true host range of *Xf*. As the subspecies *sandyi* was only established in 2005 by Schuenzel, *et al*., it might explain why only a limited number of hosts have been associated with the *sandyi* subspecies. A complete list of hosts affected by each subspecies is found in **Appendix Table C**. As the list of plants affected by *Xf* is incomplete, it is difficult to know whether multiplex really does have the largest host range, or whether this is just due to sampling bias. Also, the majority of plants where *Xf* has been isolated from are crops and ornamentals. Very limited research has been done on *Xf* in native plants, therefore it is very likely that *Xf* is present in such plants but has not been detected yet.

X. fastidiosa phylogeny

Phylogenies are a helpful way to understand the relationships between different strains of bacteria and how they might have diverged. However, the high instances of recombination between bacterial strains make it difficult to find a 'true' phylogenetic tree of a bacterial species. To create bacterial phylogenetic trees, one must look at the core genome instead, as these are usually more conserved between strains. This was done by implementing the OrthoFinder program on available *Xf* genomes (Emms, and Kelly, 2015). OrthoFinder finds orthologous genes – sequences that are descendants of the same ancestral sequence that were separated due to speciation – between the genomes of interest. Creating a phylogeny of all *Xf* genomes currently available and mapping the hosts where each strain was isolated

from allows the visualisation of any possible patterns of host specificity within and between subspecies.

According to the *Xf* phylogeny (see **Figure 8**), South American strains appear to be of the subspecies *pauca*. This supports the research conducted by (Marcelletti, and Scortichini, 2016b) that the *Xf* CoDiRo strain, associated with the first European outbreak of OQDS in Italy, belongs to the subspecies *pauca*, as the strain clusters in the same group. Unfortunately, genomes of only two strains of the subspecies *sandyi* were available. Subspecies *sandyi* has been described as the *Xf* subspecies with the most limited host range. Further research is needed to determine whether this limited host range is a result of the sporadic instances of *sandyi* in plants of interest, sampling bias, or whether *sandyi* is more prevalent in plants as a generalist endophyte and therefore only detected in pathogenic cases. As very limited research is available on generalist microorganisms, it might not be clear if the majority of subspecies *sandyi* strains are actually non-disease-causing. The heterogeneous location of origin (North America and Europe) of subspecies *fastidiosa* and *multiplex* strains implies that European strains were introduced from North America. Subspecies information of *Xf* strains MUL-MD and MUL0034 were not available. Nunney et al. (2014) have proposed the subspecies *morus* for strain MUL0034, however this novel subspecies is still under review. Further research is required to validate this or determine if the two strains belong to other subspecies, as they do not clearly group in any of the clades in this phylogeny.

From the curation of *Xf* hosts it is known that Rubiaceae is a plant family affected by all four *Xf* subspecies of interest, however the phylogeny does not show this. This is because there is no genome of *Xf* subsp. *multiplex* affecting Rubiaceae available. Unfortunately, in the sciences, it is often the case to be working with incomplete data. This uncovers many

questions: How many more plant hosts does *Xf* have? Is host-specificity between subspecies even more blurred than previously thought? In other words, how many more plant hosts are out there that can be affected by all four subspecies? A phylogeny will not be able to answer those questions, but it does provide a good indication of the signatures of host-specificity for *Xf*. This phylogeny will also be useful when analysing putative effectors of *Xf* to determine if these are host-specific or subspecies-specific.

Effector prediction

There are several methods by which bacterial effectors can be predicted. However, most of the available programs focus on effectors secreted by the T3SS. An issue with effector prediction is that the majority of proteins have not been characterised and the function of most proteins is unknown. Instead, one could look at protein motifs, the structure or detect similarity with proteins in other bacteria to find out the function of a similar protein in that bacterial strain. Furthermore, not all effectors are secreted through SS' and thus may not encode the signatures that effectors secreted through SS' would usually harbour. Small molecule effectors, such as phytotoxins or other effector molecules that may act as ligands to alter macromolecule activity, often do not rely on SS' but instead are determined by characterising non-ribosomal peptide and polyketide synthetases in the genome. Instead, these may be secreted via the Sec or Tat pathway only (Bender, Alarcón-Chaidez, and Gross, 1999; Collmer, Schneider, and Lindeberg, 2009).

Effectors and secretion systems of X. fastidiosa

A number of virulence factors have been previously described in *Xf*. Cell wall-degrading enzymes (CWDEs), such as endo-polygalacturonase (endo-PG) in combination with endoglucanase (EGase) which give *Xf* the ability to digest plant cell wall polymers (Zhang, et al., 2015). Lipopolysaccharides (LPS') are a structural component in Gram-negative bacterial

cell envelopes. LPS' are a type of pathogen-associated molecular pattern (PAMP), which allow the plant immune system to recognise the presence of a pathogen and induce an immune response. It has been found that *Xf* is able to modify the terminal O-antigen polysaccharide chain of its LPS', allowing the delay of recognition by the plant immune system (Rapicavoli, et al., 2018b). A putative CWDE, LipA – a lipase secreted through the T2SS – was also found to be abundantly secreted in PD symptomatic leaves (Nascimento, et al., 2016). A number of haemagglutinin and haemagglutinin-like proteins have also been shown to play a major role in biofilm formation, a key virulence factor in *Xf* pathogenesis (De Souza, et al., 2003; Guilhabert, and Kirkpatrick, 2005).

Effector proteins are a type of virulence factor and play a major role in pathogenicity. These effectors are secreted into the bacterial cell's surroundings or directly translocated into a host cell by secretion systems. *Xf* lacks genes that make up the T3SS (Simpson, et al., 2000) and instead encodes genes that make up components of T1SS, T2SS, T4SS and T5SS (Simpson, et al., 2000; Sluys, et al., 2003). Rogers, and Stenger (2012) found that *Xf* strain M23 and a strain of *Xf* subsp. *multiplex* both carry a plasmid that contains genes encoding homologues of a complete T4SS. Furthermore, effectors have been found to be present in pathogenic *Xf* strains but absent in the *Xf* strain EB92-1, which is often used as biocontrol as it only causes very weak symptoms in grapevine at ideal conditions. These included two enzymes secreted through T2SS (Zhang, et al., 2015).

Putative effectors of X. fastidiosa

In this research, putative effectors of *Xf* were determined using the PREFFECTOR program created by Dhroso, Eidson and Korkin (2018). Only putative effector sequences with a probability of 1 have been included in further analyses, as these were predicted to be effectors with the highest certainty. 570 putative effectors were analysed for sequence homology by

assembling these into orthologous groups. Orthologues arise from a speciation events from the last common ancestor of the species of interest. Evolutionary speaking, sequences of high similarity are more likely to be closely related than sequences with low similarity (Koonin, 2005). Generally, it is assumed that orthologues have a similar biological role in different species as they have been found to be under similar regulation and usually the same specificity in close organisms (Mirny, and Gelfand, 2002; Tatusov, Koonin, and Lipman, 1997). Here, orthologues were clustered based on a set of genes that descended from a single gene in the last common ancestor of *Xf*, *X. taiwanensis*, *X. oryzae* and *X. campestris* (Emms, and Kelly, 2015). These OGs were mapped to a core-genome phylogeny created for *Xf* to identify signatures of evolution between the different strains (see **Figure 11**).

Several proteins within these OGs have been previously annotated (see **Table 4**). The function of these proteins has been confirmed using the UniProt database (The UniProt Consortium, 2021). These include proteins with a cellulase or cellobiosidase – an enzyme hydrolysing O-glycosyl compounds involved in carbohydrate metabolic process – function in X1, a group of cold shock proteins in X2, dehydrogenase and pyrroloquinolone quinones, which oxidises oxidize polyvinyl alcohols in X15. Other proteins found to be grouped in OGs are pore-forming RTX toxins in X29, haemagglutinins in X37, the rRNA-binding 30S ribosomal protein THX – a protein involved in stabilising subunit structures – in X44, and an adhesin and cell surface protein of unknown function in X50. Unfortunately, the function of the majority of these orthologues is not known and have only been annotated as 'hypothetical protein' (HP). The function of such HPs can be determined by homology search. Sequences of the hypothetical protein may be compared to sequences of better studied species. Sequences may also be investigated for signatures that dictate protein structure. Interestingly, OG X3 is missing in all *Xf* subsp. *pauca.* No such pattern is observed in any of the other subspecies. NCBI BLAST search of one sequence within X3 suggests the protein is a homologue of a protein from the

septal ring lytic transglycosylase rare lipoprotein A (RlpA) family – a family involved in peptidoglycan binding – in *Luteimonas* spp., a genus within the Xanthomonadaceae family. RlpA was also studied by Jorgenson et al. (2014) in *Pseudomonas aeruginosa* and found to be involved in efficient daughter cell separation and preserving the bacterium's rod shape.

Conclusion

Understanding the role of effectors in *Xf* disease can elucidate why the bacterium remains harmless in certain plant hosts but become virulent to others. This study identified a group of genes to be present in *Xf* subsp. *pauca* only and absent in all other *Xf* subspecies, which may explain why this particular subspecies is so detrimental to Italy's olive groves. A functional analysis of these yet uncharacterised group of proteins, which are all putative effectors in the *pauca* subspecies, can shed light to their molecular function.

Final discussion and conclusion

How will the first detection of *X. fastidiosa* impact Colombia?

This research study describes the first report of *Xf* in Colombia. The subspecies of a number of Colombian strains was confirmed to be of *pauca*. *Xf* has been reported in several countries in Central and South America, and the bacterium is particularly known to cause CLS in Brazil, where *Xf* subsp. *pauca* is wide-spread, and a milder form in Costa Rica, where *Xf* subsp. *fastidiosa* is wide-spread. The first report of *Xf* in Colombia shows that the bacteria may likely be present in countries previously thought to be free of the bacteria. The owners of the farm where the positive coffee samples were collected from were notified, however, it is up to them to proceed regarding *Xf* eradication. As Colombia does not have the same strict *Xf* regulations as the European Union (i.e. destruction of potential hosts in a 100 m radius and demarcation of a 5 km radius), it is more likely that no extreme control measures will be implemented. It should be noted that no serious threat of the bacteria has been reported in Colombia, so it is very likely that *Xf* merely exists as an endophyte in the sampled plants. The affected farm has not described any concerns regarding *Xf*-like symptoms in their coffee plants. The detection of *Xf* in the country opens more questions about this fascinating bacterium. Colombia does not appear to have an outbreak of *Xf*, unlike its neighbouring country Brazil, where coffee leaf scorch due to *Xf* is rampant. International coffee plant trade of Colombia may be affected; however, the affected farm is centred around coffee bean production. Nevertheless, it would be interesting to perform comparative analyses between Brazilian and Colombian strains to see if they vary and understand why *Xf* appears to be more virulent in one country than another.

What can be done with the information of putative *X. fastidiosa* effectors?

This study investigated putative effectors of *Xf* and understand their role in the pathogenicity and host-range of the bacterium. More specifically, this study attempts to determine if certain

effectors are subspecies-specific, host-specific and/or only expressed in pathogenic strains. 3,546 putative effectors were identified in *Xf* genomes only. Those predicted with the highest certainty (n = 570) have been included in an orthology analysis, where 453 were grouped into 52 separate OGs. One OG in particular was found to only be present in *Xf* subsp. *pauca*. Proteins in this OG have not been characterised yet and homology search of sequences in this OG against the NCBI database does not give a definite answer. Therefore, the function of these proteins should be confirmed in *in vivo*, which will also give the opportunity to verify whether these proteins are indeed effectors and involved in pathogenicity. This is important, particularly because the largest *Xf* outbreak in Europe is in olives in Italy, which are caused by *Xf* subsp. *pauca*. Novel control measures could involve targeting effectors to stop symptoms from developing. It would also be interesting to investigate the expression of the identified effectors in the different *Xf* strains, and understand if certain host-subspecies combinations and effector expression are particularly detrimental to a plant host.

Future direction in *X. fastidiosa* research

It is important to study the various factors that make this bacterium pathogenic, and by investigating its molecular biology, genetics and community analyses of affected hosts, a better understanding of diseases caused by *Xf* can be gained. The aim of this research project was to understand the different factors that enable *Xf* to become pathogenic and host-specific. Several future questions can be considered: why is the bacterium pathogenic in some plants but remains asymptomatic in others? Do effectors play a role in symptomatic versus asymptomatic cases? Are there any effectors that are specific to symptomatic plants only? Does the microbiome play a role in *Xf* pathogenicity? Finding answers to these questions may give us a better understanding as to why *Xf* causes disease in some plants but not in others. In order to develop an effective control measure, or better yet a treatment plan for diseased hosts, research must be conducted to understand how the bacteria cause disease within a

plant. Understanding pathogenicity mechanisms and the diversity of different strains will aid in the design of novel targets to disrupt pathogen virulence.

Understanding the fundamental biology of this organism can help prevent diseases which result in with the enormous economic and even cultural loss that is caused by the bacterium around the world. Moreover, comparative genomic studies between *Xf* and *X. saccharis*, another xylem-invading Xanthomonadaceae lacking a T3SS, could elucidate what these two Xanthomonads have in common genetically speaking. Microbial community studies including *Xf* is also very limited, and more research is needed focusing on how community dynamics affect *Xf* pathogenicity and host susceptibility.

Knowledge and technology transfer

Table 5. List of attended knowledge and technology transfer events.

Date	Event	Activity
12/2020	Development and implementation of diagnostic tools for Xylella Virtual event	oral presentation
09/2020	Weekly Genetics, Genomics & Breeding department meeting Virtual event	oral presentation
02/2020	NIAB EMR Seminars East Malling, UK	seminar
02/2020	NIAB EMR PhD student meeting East Malling, UK	poster presentation
02/2020	The Linnean Society Student Conference 2020 London, UK	oral presentation
01/2020	AHDB Crop PhD Conference 2020 Nottingham, UK	poster presentation
11/2019	AHDB Soft Fruit Day 2019 NIAB EMR, East Malling, UK	poster presentation
10/2019	2nd European Conference on Xylella fastidiosa Ajaccio, France	poster presentation
10/2019	National Fruit Show 2019 Maidstone, UK	'Bacterial Diseases' co-exhibitor
10/2019	University of Nottingham Doctoral Training Programme student visit NIAB EMR, East Malling, UK	oral presentation
10/2019	The Worshipful Company of Gardeners' Association visit NIAB EMR, East Malling, UK	oral presentation
07/2019	Tropical Microbiology Course 2019 EAFIT University, Medellín, Colombia	seminar
06/2019	Soapbox Science 2019 Canterbury, UK	oral presentation
05/2019	Biosecurity and Xylella training RHS Garden Wisley UK	training
05/2019	AHDB industry visit and meeting with growers J&A Growers, Warwick, UK	industry visit
03/2019	Weekly Genetics, Genomics & Breeding department meeting NIAB EMR, East Malling, UK	oral presentation
03/2019	MBPP conference 2019 JIC Conference Centre, Norwich, UK	poster presentation
03/2019	NIAB Poster Day 2019 NIAB, Cambridge, UK	poster presentation
03/2019	Monthly PhD student meeting NIAB EMR, East Malling, UK	oral presentation
02/2019	Weekly Genetics, Genomics & Breeding department meeting NIAB EMR, East Malling, UK	oral presentation
11/2018	AHDB PhD Studentship Conference 2018 Solihul, UK	oral presentation
11/2018	Genetics, Genomics and Breeding Department Research Symposium 2018 Maidstone, UK	oral presentation

Glossary

AVBS	[Belgian Nurserymen and Growers' Federation]
BCYE	buffered charcoal-yeast extract
BLAST	basic local alignment search tool
bp	base pair(s)
CTAB	cetyltrimethylammonium (cetrimonium) bromide
cv.	cultivar
CVC	citrus variegated chlorosis
CWDE	cell-wall degrading enzyme
DTBIA	direct tissue blot immunoassay
EC	European Commission
EFSA	European Food Safety Authority
EGase	endoglucanase
ELISA	enzyme-limited immunosorbent assay
endo-PG	endo-polygalacturonase
EPPO	European and Mediterranean Plant Protection Organisation
EPS	extracellular polymeric substance(s)
HGT	horizontal gene transfer
HP	hypothetical protein
IF	immunofluorescence
LAMP	loop-mediated isothermal amplification
LPS	lipopolysaccharide
MAMP	microbe-associated molecular pattern
MAMSL	metres above median sea level
MSA	multiple sequence alignment
NCBI	National Center for Biotechnology Information
NCPPB	National Collection of Plant Pathogenic Bacteria
n.d.	no date
OG	orthologous group
OQDS	olive quick-decline syndrome
PAMP	pathogen-associated molecular pattern
PCR	polymerase chain reaction
PRR	pattern recognition receptor
PTI	pathogen-associated molecular pattern (PAMP)-triggered immunity
PD	Pierce's disease
PWG	periwinkle wilt Gelrite
qPCR	quantitative / real-time polymerase chain reaction
QS	quorum sensing
sp. / spp.	species (singular / plural)
subsp.	subspecies
T[1-6]SS	type [1-6] secretion system
VO	vascular occlusion
Xf	*Xylella fastidiosa*

References

Abascal, F, Zardoya, R, and Posada, D, 2005. ProtTest: Selection of best-fit models of protein evolution. Bioinformatics. 21(9), pp.2104–2105.

Alfano, JR, and Collmer, A, 2004. Type III secretion system effector proteins: Double agents in bacterial disease and plant defense. Annual Review of Phytopathology. 42, pp.385–414.

Almeida, RPP, Blua, MJ, Lopes, JRS, and Purcell, AH, 2005. Vector Transmission of Xylella fastidiosa: Applying Fundamental Knowledge to Generate Disease Management Strategies. Annals of the Entomological Society of America. [online] 98(6), pp.775–786. Available at: <https://academic.oup.com/aesa/article/98/6/775/96161>.

Almeida, RPP, Coletta-Filho, HD, and Lopes, JRS, 2014. Xylella fastidiosa. In: D. Liu, ed., *Manual of Security Sensitive Microbes and Toxins*. [online] CRC Press, pp.841–850. Available at: <https://nature.berkeley.edu/xylella/wp-content/uploads/2016/07/CVC_2014.pdf>.

Almeida, RPP, and Nunney, L, 2015. How do plant diseases caused by Xylella fastidiosa emerge? Plant Disease. [online] 99(11), pp.1457–1467. Available at: <http://dx.doi.org/10.1094/PDIS-02-15-0159-FE>.

Alves, E, Leite, B, Pascholati, SF, Ishida, ML, Andersen, PC, Andersen, C, and Andersen, PC, 2009. Citrus sinensis leaf petiole and blade colonization by Xylella fastidiosa: details of xylem vessel occlusion. Scientia Agricola. 66(2), pp.218–224.

AVBS, 2018. First discovery of Xylella fastidiosa in Belgium. [online] Available at: <https://www.avbs.be/actualiteit/eerste-detectie-van-xylella-fastidiosa-belgië>.

Baccari, C, and Lindow, SE, 2011. Assessment of the process of movement of Xylella fastidiosa within susceptible and resistant grape cultivars. Phytopathology. 101, pp.77–84.

Basler, M, 2015. Type VI secretion system: Secretion by a contractile nanomachine. Philosophical Transactions of the Royal Society B: Biological Sciences. [online] 370(1679). Available at: <https://royalsocietypublishing.org/doi/full/10.1098/rstb.2015.0021?url_ver=Z39.88-2003&rfr_id=ori:rid:crossref.org&rfr_dat=cr_pub%3Dpubmed>.

Bender, CL, Alarcón-Chaidez, F, and Gross, DC, 1999. Pseudomonas syringae Phytotoxins: Mode of Action, Regulation, and Biosynthesis by Peptide and Polyketide Synthetases. Microbiology and Molecular Biology Reviews. 63, pp.266–292.

De Benedictis, M, De Caroli, M, Baccelli, I, Marchi, G, Bleve, G, Gallo, A, Ranaldi, F, Falco, V, Pasquali, V, Piro, G, Mita, G, Di Sansebastiano, G Pietro, Benedictis, M De, Caroli, M De, Baccelli, I, Marchi, G, Bleve, G, Gallo, A, Ranaldi, F, Falco, V, Pasquali, V, Piro, G, Mita, G, Pietro, G, Sansebastiano, D, De Benedictis, M, De Caroli, M, Baccelli, I, Marchi, G, Bleve, G, Gallo, A, Ranaldi, F, Falco, V, Pasquali, V, Piro, G, Mita, G, Di Sansebastiano, G Pietro, Gallo, A, Pasquali, V, Di Sansebastiano, G Pietro, Ranaldi, F, De Benedictis, M, Baccelli, I, Bleve, G, Mita, G, De Caroli, M, Marchi, G, and Piro, G, 2017. Vessel occlusion in three cultivars of Olea europaea naturally exposed to Xylella fastidiosa in open field. Journal of Phytopathology. 165(9), pp.589–594.

Bergsma-Vlami, M, van de Bilt, JLJM, Tjou-Tam-Sin, NNA, van de Vossenberg, BTLH, and Westenberg, M, 2015. Imported From Costa Rica and Honduras in the Netherlands Xylella Fastidiosa in Coffea. Journal of Plant Pathology. 97(2), pp.391–403.

Bernal, P, Llamas, MA, and Filloux, A, 2018. Type VI secretion systems in plant-associated bacteria. Environmental Microbiology. [online] 20(1). Available at: <https://doi.org/10.1111/1462-2920.13956>.

Block, A, and Alfano, JR, 2011. Plant targets for Pseudomonas syringae type III effectors: Virulence targets or guarded decoys? Current Opinion in Microbiology. [online] 14(1), pp.39–46. Available at: <http://dx.doi.org/10.1016/j.mib.2010.12.011>.

Von Bodman, SB, Bauer, WD, and Coplin, DL, 2003. Quorum Sensing in Plant-Pathogenic Bacteria. Annual Review of Phytopathology. 41, pp.445–482.

Boyer, F, Fichant, G, Berthod, J, Vandenbrouck, Y, and Attree, I, 2009. Dissecting the bacterial type VI secretion system by a genome wide in silico analysis: What can be learned from available microbial genomic resources? BMC Genomics. [online] 10. Available at: <https://bmcgenomics.biomedcentral.com/articles/10.1186/1471-2164-10-104>.

Burbank, LP, and Stenger, DC, 2017. The DinJ/RelE toxin-antitoxin system suppresses bacterial proliferation and virulence of xylella fastidiosa in grapevine. Phytopathology. 107, pp.388–394.

Buttner, D, and He, SY, 2009. Type III Protein Secretion in Plant Pathogenic Bacteria. Plant Physiology. [online] 150(4), pp.1656–1664. Available at:

<http://www.plantphysiol.org/cgi/doi/10.1104/pp.109.139089>.
Camilli, A, and Bassler, BL, 2006. Bacterial small-molecule signaling pathways. Science. 311, pp.1113–1116.
Carbajal, D, Morano, KA, and Morano, LD, 2004. Indirect immunofluorescence microscopy for direct detection of Xylella fastidiosa in Xylem Sap. Current Microbiology. 96, pp.372–75.
Casadevall, A, and Pirofski, LA, 1999. Host-pathogen interactions: Redefining the basic concepts of virulence and pathogenicity. Infection and Immunity. 67(8), pp.3703–3713.
Castresana, J, 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Molecular Biology and Evolution. 17(4), pp.540–552.
Cavalieri, V, and Porcelli, F, 2017. Main insect vectors of Xylella fastidiosa worldwide & in Italy. In: A.M. D'Onghia, S. Brunel and F. Valentini, eds., *Xylella fastidiosa & the Olive Quick Decline Syndrome (OQDS). A serious worldwide challenge for the safeguard of olive trees.* Bari: CIHEAM, pp.31–32.
Chagas, CM, Rossetti, V, and Beretta, MJG, 1992. Electron microscopy studies of a xylem-limited bacterium in sweet orange affected with citrus variegated chlorosis disease in Brazil. Journal of Phytopathology. 134(306).
Chang, CJ, Garnier, M, Zreik, L, Rossetti, V, and Bové, JM, 1993. Culture and serological detection of the xylem-limited bacterium causing citrus variegated chlorosis and its identification as a strain of Xylella fastidiosa. Current Microbiology. 27(137).
Chatterjee, S, Almeida, RPP, and Lindow, S, 2008. Living in two Worlds: The Plant and Insect Lifestyles of Xylella fastidiosa. Annual Review of Phytopathology. 46, pp.243–271.
Chen, H, and Boutros, PC, 2011. VennDiagram: A package for the generation of highly-customizable Venn and Euler diagrams in R. BMC Bioinformatics. 12, p.35.
Choi, HK, Iandolino, A, Da Silva, FG, and Cook, DR, 2013. Water deficit modulates the response of vitis vinifera to the Pierce's disease pathogen xylella fastidiosa. Molecular Plant-Microbe Interactions. 26(6), pp.643–657.
Collmer, A, Schneider, DJ, and Lindeberg, M, 2009. Lifestyles of the Effector Rich: Genome-Enabled Characterization of Bacterial Plant Pathogens. Plant Physiology. [online] 150(August), pp.1623–1630. Available at: <http://www.plantphysiol.org/cgi/doi/10.1104/pp.109.140327>.
Cornara, D, Cavalieri, V, Dongiovanni, C, Altamura, G, Palmisano, F, Bosco, D, Porcelli, F, Almeida, RPP, and Saponari, M, 2017. Transmission of Xylella fastidiosa by naturally infected Philaenus spumarius (Hemiptera, Aphrophoridae) to different host plants. Journal of Applied Entomology. 141(1–2), pp.80–87.
Davis, MJ, Purcell, AH, and Thomson, S V, 1978. Pierce's disease of grapevines: Isolation of the causal bacterium. Science. 199(4324), pp.75–77.
Davis, MJ, Purcell, AH, and Thomson, S V, 1980. Isolation Media for the Pierce's Disease Bacterium. Phytopathology. 70, pp.425–429.
Dhroso, A, Eidson, S, and Korkin, D, 2018. Genome-wide prediction of bacterial effector candidates across six secretion system types using a feature-based statistical framework. Scientific Reports. [online] 8(1), p.17209. Available at: <http://dx.doi.org/10.1038/s41598-018-33874-1>.
Djelouah, K, Frasheri, D, Valentini, F, D'Onghia, AM, and Digiaro, M, 2014. Direct tissue blot immunoassay for detection of Xylella fastidiosa in olive trees. Phytopathologia Mediterranea. 53, pp.559–564.
Dongiovanni, C, Di Carolo, M, Fumarola, G, Tauro, D, Altamura, G, and Cavalieri, V, 2018. Evaluation of Insecticides for the Control of Juveniles of Philaenus spumarius L., 2015–2017. Arthropod Management Tests. 43(1), pp.1–2.
Dow, JM, and Daniels, MJ, 2000. Xylella Genomics and Bacterial Pathogenicity to Plants. Yeast. 1(4), pp.263–271.
EC, 2019. Commission database of host plants found to be susceptible to Xylella fastidiosa in the Union Territory – Update 12. *European Journal of Health Law*. [online] Available at: <http://booksandjournals.brillonline.com/content/10.1163/1571809042388581>.
Edgar, RC, 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. BMC Bioinformatics. [online] 5. Available at: <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-5-113>.
EFSA, 2018. Scientific report on the update of the Xylella spp. host plant database. EFSA Journal. [online] 16(9). Available at: <https://doi.org/10.2903/j.efsa.2018.5408>.
Emms, DM, and Kelly, S, 2015. OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. Genome Biology. [online] 16(157). Available at:

<https://genomebiology.biomedcentral.com/track/pdf/10.1186/s13059-015-0721-2>.
EPPO, 2015. Xylella fastidiosa detected in Coffea spp. plants imported into Switzerland. [online] Available at: <https://gd.eppo.int/reporting/article-5128>.
EPPO, 2016a. First report of Xylella fastidiosa subsp. fastidiosa on Nerium oleander in Germany. [online] Available at: <https://gd.eppo.int/reporting/article-5878>.
EPPO, 2016b. PM 7/24 (2) Xylella fastidiosa. In: *EPPO Bulletin*. [online] pp.463–500. Available at: <https://doi.org/10.1111/epp.12327>.
EPPO, 2018. PM 7/24 (3) Xylella fastidiosa. Bulletin OEPP/EPPO Bulletin.
EPPO, 2019. First report of Xylella fastidiosa subsp. multiplex in Portugal. [online] Available at: <https://gd.eppo.int/reporting/article-6447>.
EPPO, n.d. Xylella fastidiosa (XYLEFA): Hosts. [online] Available at: <https://gd.eppo.int/taxon/XYLEFA/hosts>.
Fang, Y, Lin, H, Wu, L, Ren, D, Ye, W, Dong, G, Zhu, L, and Guo, L, 2015. Genome sequence of Xanthomonas sacchari R1, a biocontrol bacterium isolated from the rice seed. *Journal of Biotechnology*.
Feil, H, and Purcell, AH, 2007. Temperature-Dependent Growth and Survival of Xylella fastidiosa in Vitro and in Potted Grapevines . Plant Disease. [online] 85(12), pp.1230–1234. Available at: <http://apsjournals.apsnet.org/doi/full/10.1094/PDIS.2001.85.12.1230?url_ver=Z39.88-2003&rfr_id=ori:rid:crossref.org&rfr_dat=cr_pub%3Dpubmed>.
Fellows, I, 2012. Wordcloud: Word clouds. *R package version*.
Francis, M, Lin, H, Rosa, JC, Doddapaneni, H, and Civerolo, EL, 2006. Genome-based PCR primers for specific and sensitive detection and quantification of Xylella fastidiosa. European Journal of Plant Pathology. 115, pp.203–213.
Freitag, JH, 1951. Host range of the Pierce's disease virus of Grapes as determined by insect transmission. Phytopathology. 41, pp.920–934.
Galán, JE, and Collmer, A, 1999. Type III secretion machines: Bacterial devices for protein delivery into host cells. Science. 284(5418), pp.1322–1328.
González-Torres, P, Rodríguez-Mateos, F, Antón, J, and Gabaldón, T, 2019. Impact of homologous recombination on the evolution of prokaryotic core genomes. mBio. 21: 829.
Gouran, H, Gillespie, H, Nascimento, R, Chakraborty, S, Zaini, PA, Jacobson, A, Phinney, BS, Dolan, D, Durbin-johnson, BP, Antonova, ES, Lindow, SE, Mellema, MS, Goulart, LR, and Dandekar, AM, 2016. The Secreted Protease PrtA Controls Cell Growth, Biofilm Formation and Pathogenicity in Xylella fastidiosa. Scientific Reports. [online] 6(July), pp.1–13. Available at: <http://dx.doi.org/10.1038/srep31098>.
Green, ER, and Mecsas, J, 2016. Bacterial Secretion Systems – An overview. Microbiology Spectrum. [online] 4(1). Available at: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4804464/pdf/nihms715968.pdf>.
Guilhabert, MR, and Kirkpatrick, BC, 2005. Identification of Xylella fastidiosa Antivirulence Genes: Hemagglutinin Adhesins Contribute to X. fastidiosa Biofilm Maturation and Colonization and Attenuate Virulence. Molecular Plant-Microbe Interactions. 181, pp.856–868.
Harper, SJ, Ward, LI, and Clover, GRGG, 2010. Development of LAMP and real-time PCR methods for the rapid detection of Xylella fastidiosa for quarantine and field applications. Phytopathology. 100, pp.1282–1288.
Hartung, JS, Beretta, J, Brlansky, RH, Spisso, J, and Lee, RF, 1994. Citrus variegated chlorosis bacterium: axenic culture, pathogenicity and serological relationships with otherstrains of Xylella fastidiosa. Phytopathology. 84(6), pp.591–597.
Hedge, J, and Wilson, DJ, 2014. Bacterial Phylogenetic Reconstruction from Whole Genomes Is Robust to Recombination but Demographic Inference Is Not. mBio. 5(6), pp.6–9.
Henneberger, TS, 2003. Effects of Low Temperature on Populations of Xylella fastidiosa in Sycamore. [online] University of Georgia. Available at: <https://getd.libs.uga.edu/pdfs/henneberger_tiffany_s_200308_ms.pdf>.
Hernandez-Martinez, R, Cooksey, DA, and Wong, FP, 2009. Leaf scorch of purple-leafed plum and sweetgum dieback: Two new diseases in southern california caused by xylella fastidiosa strains with different host ranges. Plant Disease. 93, pp.1131–1138.
Hill, BL, and Purcell, AH, 1995. Acquisition and retention of Xylella fastidiosa by an efficient vector, Graphocephala atropunctata. Phytopathology. 85, pp.209–212.
Hopkins, DL, 1951. Biological Control of Pierce' s Disease in the Vineyard with Strains of Xylella fastidiosa Benign to Grapevine. Plant Disease. [online] 89(12), pp.1348–1352. Available at:

<http://apsjournals.apsnet.org/doi/full/10.1094/PHYTO-07-18-0245-FI?url_ver=Z39.88-2003&rfr_id=ori:rid:crossref.org&rfr_dat=cr_pub%3Dpubmed>.
Hopkins, DL, 1985. Physiological and Pathological Characteristics of Virulent and Avirulent Strains of the Bacterium that Causes Pierce's Disease of Grapevine. Phytopathology. 75, pp.713–717.
Hopkins, DL, and Purcell, AH, 2002. Xylella fastidiosa : Cause of Pierce's Disease of Grapevine and Other Emergent Diseases. Plant Disease. [online] 86(10), pp.1056–1066. Available at: <https://apsjournals.apsnet.org/doi/10.1094/PDIS.2002.86.10.1056>.
Jackson, RW, Johnson, LJ, Clarke, SR, and Arnold, DL, 2011. Bacterial pathogen evolution: Breaking news. Trends in Genetics. 27(1), pp.32–40.
Janse, JD, and Obradovic, A, 2010. Xylella fastidiosa: Its biology, diagnosis, control and risks. Journal of Plant Pathology. [online] 92(1, Supplement), p.S1.35-S1.48. Available at: <http://www.sipav.org/main/jpp/index.php/jpp/article/view/2504>.
Jolley, KA, Bray, JE, and Maiden, MCJ, 2018. Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications [version 1; referees: 2 approved]. Wellcome Open Research. [online] 3(124). Available at: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6192448/pdf/wellcomeopenres-3-16155.pdf>.
Jorgenson, MA, Chen, Y, Yahashiri, A, Popham, DL, and Weiss, DS, 2014. The bacterial septal ring protein RlpA is a lytic transglycosylase that contributes to rod shape and daughter cell separation in Pseudomonas aeruginosa. Molecular Microbiology. 93(1), pp.113–128.
Kahle, D, and Wickham, H, 2013. ggmap: Spatial visualization with ggplot2. *R Journal*.
Kanehisa, M, 2019. Toward understanding the origin and evolution of cellular organisms. Protein Science. 28, pp.1947–1951.
Kanehisa, M, Furumichi, M, Sato, Y, Ishiguro-Watanabe, M, and Tanabe, M, 2021. KEGG: Integrating viruses and cellular organisms. Nucleic Acids Research. 49, pp.D545-551.
Kanehisa, M, and Goto, S, 2000. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Research. 28, pp.27–30.
Khan, M, Seto, D, Subramaniam, R, and Desveaux, D, 2018. Oh, the places they'll go! A survey of phytopathogen effectors and their host targets. Plant Journal. 93(4), pp.651–663.
Koonin, E V, 2005. Orthologs, paralogs, and evolutionary genomics. Annual Review of Genetics. 39(1), pp.309–338.
Krivanek, R, Sisterson, MS, and Lin, H, 2005. Vitis resistance to Pierce's disease is characterized by differential Xylella fastidiosa populations in stems and leaves. Phytopathology. 95, pp.44–52.
Levy, A, Salas Gonzalez, I, Mittelviefhaus, M, Clingenpeel, S, Herrera Paredes, S, Miao, J, Wang, K, Devescovi, G, Stillman, K, Monteiro, F, Rangel Alvarez, B, Lundberg, DS, Lu, TY, Lebeis, S, Jin, Z, McDonald, M, Klein, AP, Feltcher, ME, Rio, TG, Grant, SR, Doty, SL, Ley, RE, Zhao, B, Venturi, V, Pelletier, DA, Vorholt, JA, Tringe, SG, Woyke, T, Dangl, JL, Gonzalez, IS, Mittelviefhaus, M, Clingenpeel, S, Paredes, SH, Miao, J, Wang, K, Devescovi, G, Stillman, K, Monteiro, F, Alvarez, BR, Lundberg, DS, Lu, TY, Lebeis, S, Jin, Z, McDonald, M, Klein, AP, Feltcher, ME, Rio, TG, Grant, SR, Doty, SL, Ley, RE, Zhao, B, and Venturi, V, 2018. Genomic features of bacterial adaptation to plants. Nature Genetics. [online] 50(1), pp.138–150. Available at: <http://dx.doi.org/10.1038/s41588-017-0012-9>.
Li, W, Teixeira, DC, Hartung, JS, Huang, Q, Duan, Y, Zhou, L, Chen, J, Lin, H, Lopes, S, Ayres, AJ, and Levy, L, 2013. Development and systematic validation of qPCR assays for rapid and reliable differentiation of Xylella fastidiosa strains causing citrus variegated chlorosis. Journal of Microbiological Methods. 92, pp.79–89.
Li, WB, Pria, J, Teixeira, DC, Miranda, VS, Ayres, AJ, Franco, CF, Costa, MG, He, CX, Costa, PI, and Hartung, JS, 2001. Coffee leaf scorch caused by a strain of Xylella fastidiosa from citrus. Plant Disease. 85, pp.501–505.
Lien, Y-W, and Lai, E-M, 2017. Type VI Secretion Effectors: Methodologies and Biology. Frontiers in Cellular and Infection Microbiology. 7, p.254.
de Lima, JEO, Miranda, VS, Hartung, JS, Brlansky, RH, Coutinho, A, Roberto, SR, and Carlos, EF, 1998. Coffee Leaf Scorch Bacterium: Axenic Culture, Pathogenicity, and Comparison with Xylella fastidiosa of Citrus. Plant Disease. 82(1), pp.94–97.
Lindow, S, 2018. Money Matters: Fueling Rapid Recent Insight Into Xylella fastidiosa — An Important and Expanding Global Pathogen. Phytopathology. [online] 109(2), pp.210–212. Available at: <https://apsjournals.apsnet.org/doi/10.1094/PHYTO-09-18-0325-PER>.
Maiden, MCJ, 2006. Multilocus Sequence Typing of Bacteria. Annual Review of Microbiology. 60,

pp.561–588.
Maiden, MCJ, Bygraves, JA, Feil, E, Morelli, G, Russell, JE, Urwin, R, Zhang, Q, Zhou, J, Zurth, K, Caugant, DA, Feavers, IM, Achtman, M, and Spratt, BG, 1998. Multilocus sequence typing: A portable approach to the identification of clones within populations of pathogenic microorganisms. Proceedings of the National Academy of Sciences of the United States of America. 95(6), pp.3140–3145.
Marcelletti, S, and Scortichini, M, 2016a. Genome-wide comparison and taxonomic relatedness of multiple Xylella fastidiosa strains reveal the occurrence of three subspecies and a new Xylella species. Archives of Microbiology. [online] 198(8), pp.803–812. Available at: <https://dx.doi.org/10.1007/s00203-016-1245-1>.
Marcelletti, S, and Scortichini, M, 2016b. Xylella fastidiosa CoDIRO strain associated with the olive quick decline syndrome in southern Italy belongs to a clonal complex of the subspecies pauca that evolved in Central America. Microbiology (United Kingdom). [online] 162(12), pp.2087–2098. Available at: <http://mic.microbiologyresearch.org/pubmed/content/journal/micro/10.1099/mic.0.000388>.
Martelli, GP, Boscia, D, Porcelli, F, and Saponari, M, 2016. The olive quick decline syndrome in south-east Italy: a threatening phytosanitary emergency. European Journal of Plant Pathology. [online] 144(2), pp.235–243. Available at: <https://link.springer.com/article/10.1007/s10658-015-0784-7>.
Martinati, JC, Hansen Pacheco, FT, Oliveira De Miranda, VF, and Siu, MT, 2005. Phylogenetic relationships of Xylella fastidiosa strains based on 16s-23s rDNA sequences. Current Microbiology. 50(4), pp.190–195.
Mattinen, L, Somervuo, P, Nykyri, J, Nissinen, R, Kouvonen, P, Corthals, G, Auvinen, P, Aittamaa, M, Valkonen, JPT, and Pirhonen, M, 2008. Microarray profiling of host-extract-induced genes and characterization of the type VI secretion cluster in the potato pathogen Pectobacterium atrosepticum. Microbiology. 154(Pt 8), pp.2387–2396.
Meyer, MM, and Kirkpatrick, BC, 2008. Examining the Effects of Cold Therapy on Pierce ' s Disease-infected Grapevines and on the Viability of Xylella fastidiosa Cells in vitro. In: *Proceedings of the 2nd Annual National Viticulture Research Conference*.
Minsavage, G V, Thompson, CM, Hopkins, DL, Leite, RMVBC, and Stall, RE, 1994. Development of a Polymerase Chain Reaction Protocol for Detection of Xylella fastidiosa in Plant Tissue. Phytopathology. 84, pp.456–461.
Mirny, LA, and Gelfand, MS, 2002. Using orthologous and paralogous proteins to identify specificity determining residues. Genome biology. 3(3), p.preprint0002.1-0002.20.
Montero-Astúa, M, Chacón-Díaz, C, Aguilar, E, Rodríguez, CM, Garita, L, Villalobos, W, Moreira, L, Hartung, JS, and Rivera, C, 2008. Isolation and molecular characterization of Xylella fastidiosa from coffee plants in Costa Rica. Journal of Microbiology. 46, pp.482–490.
Muyzer, G, De Waal, EC, and Uitterlinden, AG, 1993. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. Applied and Environmental Microbiology. 59(3), pp.695–700.
Nascimento, R, Gouran, H, Chakraborty, S, Gillespie, HW, Almeida-Souza, HO, Tu, A, Rao, BJ, Feldstein, PA, Bruening, G, Goulart, LR, and Dandekar, AM, 2016. The Type II Secreted Lipase/Esterase LesA is a Key Virulence Factor Required for Xylella fastidiosa Pathogenesis in Grapevines. Scientific Reports. [online] 6(November 2015), pp.1–16. Available at: <http://dx.doi.org/10.1038/srep18598>.
Newman, KL, Almeida, RPP, Purcell, AH, and Lindow, SE, 2003. Use of a Green Fluorescent Strain for Analysis of Xylella fastidiosa Colonization of Vitis vinifera. Applied and Environmental Microbiology. 69(12), pp.7319–7327.
Newman, KL, Almeida, RPP, Purcell, AH, and Lindow, SE, 2004. Cell–cell signaling controls Xylella fastidiosa interactions with both insects and plants. Proceedings of the National Academy of Sciences. [online] 101(6), pp.1737–1742. Available at: <http://www.pnas.org/cgi/doi/10.1073/pnas.0308399100>.
Nguyen, LT, Schmidt, HA, Von Haeseler, A, and Minh, BQ, 2015. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Molecular Biology and Evolution. 32(1), pp.268–274.
Nunney, L, Ortiz, B, Russell, SA, Sánchez, RR, and Stouthamer, R, 2014a. The complex biogeography of the plant pathogen xylella fastidiosa: Genetic evidence of introductions and subspecific introgression in central America. PLoS ONE. [online] 9(11). Available at: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0112463>.
Nunney, L, Schuenzel, EL, Scally, M, Bromley, RE, and Stouthamerc, R, 2014b. Large-scale intersubspecific recombination in the plant-pathogenic bacterium xylella fastidiosa is associated with

the host shift to mulberry. Applied and Environmental Microbiology. [online] 80(10), pp.3025–3033. Available at: <http://aem.asm.org/cgi/pmidlookup?view=long&pmid=24610840>.
Nunney, L, Vickerman, DB, Bromley, RE, Russell, SA, Hartman, JR, Morano, LD, and Stouthamer, R, 2013. Recent evolutionary radiation and host plant specialization in the Xylella fastidiosa subspecies native to the United States. Applied and Environmental Microbiology. 79, pp.2189–2200.
Nunney, L, Yuan, X, Bromley, R, Hartung, J, Montero-Astúa, M, Moreira, L, Ortiz, B, Stouthamer, R, Montero-astu, M, 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. [online] 5(11). Available at: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0015488>.
Nunney, L, Yuan, X, Bromley, RE, and Stouthamer, R, 2012. Detecting genetic introgression: High levels of intersubspecific recombination found in Xylella fastidiosa in Brazil. Applied and Environmental Microbiology. 78(13), pp.4702–4714.
Ouyang, P, Arif, M, Fletcher, J, Melcher, U, and Ochoa Corona, FM, 2013. Enhanced reliability and accuracy for field deployable bioforensic detection and discrimination of Xylella fastidiosa subsp. pauca, causal agent of citrus variegated chlorosis using Razor Ex technology and taqman quantitative PCR. PLoS ONE. 8.
Paradis, E, Claude, J, and Strimmer, K, 2004. APE: Analyses of phylogenetics and evolution in R language. Bioinformatics. 20(2), pp.289–290.
Parks, DH, Imelfort, M, Skennerton, CT, Hugenholtz, P, and Tyson, GW, 2015. CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Research. 25(7), pp.1043–1055.
Pelgrom, AJ, and Van den Ackerveken, G, 2016. Microbial Pathogen Effectors in Plant Disease. eLS. [online] August(August), pp.1–10. Available at: <https://onlinelibrary.wiley.com/doi/full/10.1002/9780470015902.a0023724>.
Pieretti, I, Pesic, A, Petras, D, Royer, M, Süssmuth, RD, and Cociancich, S, 2015. What makes Xanthomonas albilineans unique amongst xanthomonads? Frontiers in Plant Science. [online] 6:289. Available at: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4408752/>.
Pieretti, I, Royer, M, Barbe, V, Carrere, S, Koebnik, R, Cociancich, S, Couloux, A, Darrasse, A, Gouzy, J, Jacques, MA, Lauber, E, Manceau, C, Mangenot, S, Poussier, S, Segurens, B, Szurek, B, Verdier, V, Arlat, M, and Rott, P, 2009. The complete genome sequence of Xanthomonas albilineans provides new insights into the reductive genome evolution of the xylem-limited Xanthomonadaceae. BMC Genomics. 10(616).
Posada, D, and Crandall, KA, 2002. The effect of recombination on the accuracy of phylogeny estimation. Journal of Molecular Evolution. 54(3), pp.396–402.
Pritchard, L, and Birch, PRJ, 2014. The zigzag model of plant-microbe interactions: Is it time to move on? Molecular Plant Pathology. 15(9), pp.865–870.
Purcell, AH, 1981. Vector Preference and Inoculation Efficiency as Components of Resistance to Pierce's Disease in European Grape Cultivars. Phytopathology. 20, pp.397–435.
Purcell, AH, and Finlay, AH, 1979. Evidence for Noncirculative Transmission of Pierce's Disease Bacterium by Sharpshooter Leafhoppers. Phytopathology. 69, pp.393–395.
Purcell, AH, Finlay, AH, and McLean, DL, 1979. Pierce's disease bacterium: Mechanism of transmission by leafhopper vectors. Science. 206(4420), pp.839–841.
Purcell, AH, and Saunders, SR, 1999. Fate of Pierce's disease strains of Xylella fastidiosa in common riparian plants in California. Plant Disease. 83(9), pp.393–395.
Purcell, AH, Saunders, SR, Hendson, M, Grebus, ME, and Henry, MJ, 1999. Causal Role of Xylella fastidiosa in Oleander Leaf Scorch Disease. Phytopathology. [online] 89(1), pp.53–58. Available at: <https://apsjournals.apsnet.org/doi/10.1094/PHYTO.1999.89.1.53>.
Randall, JJ, Goldberg, NP, Kemp, JD, Radionenko, M, French, JM, Olsen, MW, and Hanson, SF, 2009. Genetic analysis of a novel Xylella fastidiosa subspecies found in the Southwestern United States. Applied and Environmental Microbiology. 75(17), pp.5631–5638.
Rapicavoli, J, Ingel, B, Blanco-Ulate, B, Cantu, D, and Roper, C, 2018a. Xylella fastidiosa: an examination of a re-emerging plant pathogen. Molecular Plant Pathology. 19(4), pp.786–800.
Rapicavoli, JN, Blanco-Ulate, B, Muszyński, A, Figueroa-Balderas, R, Morales-Cruz, A, Azadi, P, Dobruchowska, JM, Castro, C, Cantu, D, and Roper, MC, 2018b. Lipopolysaccharide O-antigen delays plant innate immune recognition of Xylella fastidiosa. Nature Communications. [online] 9(1), p.390. Available at: <https://www.nature.com/articles/s41467-018-02861-5>.
Rodríguez, CM, Obando, JJ, Villalobos, W, Moreira, L, and Rivera, C, 2001. First Report of Xylella

fastidiosa Infecting Coffee in Costa Rica. Plant Disease. 85(9), p.1027.
Rogers, EE, and Stenger, DC, 2012. A Conjugative 38 kB Plasmid Is Present in Multiple Subspecies of Xylella fastidiosa. PLoS ONE. [online] 7(12). Available at: <http://dx.plos.org/10.1371/journal.pone.0052131>.
Roper, C, Castro, C, and Ingel, B, 2019. Xylella fastidiosa : bacterial parasitism with hallmarks of commensalism. Current Opinion in Plant Biology. [online] 50, pp.140–147. Available at: <https://doi.org/10.1016/j.pbi.2019.05.005>.
Russell, AB, Peterson, SB, and Mougous, JD, 2014. Type VI secretion system effectors: Poisons with a purpose. Nature Reviews Microbiology. 12(2), pp.137–148.
Ryu, CM, 2015. Against friend and foe: Type 6 effectors in plant-associated bacteria. Journal of Microbiology. 53(3), pp.201–208.
Saponari, M, Boscia, D, Altamura, G, Loconsole, G, Zicca, S, D'Attoma, G, Morelli, M, Palmisano, F, Saponari, A, Tavano, D, Savino, VN, Dongiovanni, C, and Martelli, GP, 2017. Isolation and pathogenicity of Xylella fastidiosa associated to the olive quick decline syndrome in southern Italy. Scientific Reports. 7(1), pp.1–13.
Saponari, M, Boscia, D, Nigro, F, and Martelli, GP, 2013. Identification of DNA sequences related to Xylella fastidiosa in oleander, almond and olive trees exhibiting leaf scorch symptoms in Apulia. Journal of Plant Pathology. 95(3), p.668.
Scally, M, Schuenzel, EL, 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. Applied and Environmental Microbiology. 71(12), pp.8491–8499.
Schaad, NW, Postnikova, E, Lacy, G, Fatmi, M, and Chang, CJ, 2004. Xylella fastidiosa subspecies: X. fastidiosa subsp piercei, subsp. nov., X. fastidiosa subsp. multiplex subsp. nov., and X. fastidiosa subsp. pauca subsp. nov. Systematic and Applied Microbiology. 27(3), pp.290–300.
Schneider, K, van der Werf, W, Cendoya, M, Mourits, M, Navas-Cortés, JA, Vicent, A, and Lansink, AO, 2020. Impact of Xylella fastidiosa subspecies pauca in European olives. Proceedings of the National Academy of Sciences of the United States of America. 117, pp.9250–9259.
Schuenzel, EL, 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. Applied and Environmental Microbiology. [online] 71(7), pp.3832–3839. Available at: <https://aem.asm.org/content/71/7/3832>.
Seemann, T, 2014. Prokka: Rapid prokaryotic genome annotation. Bioinformatics. 30(14), pp.2068–269.
Sherald, JL, and Lei, JD, 1991. Evaluation of a Rapid ELISA Test Kit for Detection of Xylella fastidiosa in Landscape Trees. Plant Disease. 75, pp.200–203.
Sicard, A, Zeilinger, AR, Vanhove, M, Schartel, TE, Beal, DJ, Daugherty, MP, and Almeida, RP, 2018. Xylella fastidiosa: Insights into an Emerging Plant Pathogen. Annual Review of Phytopathology. [online] 56(56), pp.181–202. Available at: <http://arjournals.annualreviews.org/doi/full/10.1146/annurev-phyto-080417-045849?url_ver=Z39.88-2003&rfr_id=ori:rid:crossref.org&rfr_dat=cr_pub%3Dpubmed>.
Simpson, AJG, Reinach, FC, Arruda, P, Abreu, FA, Acencio, M, Alvarenga, R, Alves, LMC, Araya, JE, Baia, GS, Baptista, CS, Barros, MH, Bonaccorsi, ED, Bordin, S, Bové, JM, Briones, MRS, Bueno, MRP, Camargo, AA, Camargo, LEA, Carraro, DM, Carrer, H, Colauto, NB, Colombo, C, Costa, FF, Costa, MCR, Costa-Neto, CM, Coutinho, LL, Cristofani, M, Dias-Neto, E, Docena, C, El-Dorry, H, Facincani, AP, Ferreira, AJS, Ferreira, VCA, Ferro, JA, Fraga, JS, França, SC, Franco, MC, Frohme, M, Furlan, LR, Garnier, M, Goldman, GHS, Goldman, MHS, Gomes, SL, Gruber, A, Ho, PL, Hoheisel, JD, Junqueira, ML, Kemper, EL, Kitajima, JP, Krieger, JE, Kuramae, EE, Laigret, F, Lambais, MR, Leite, LCC, Lemos, EGM, Lemos, MVF, Lopes, SA, Lopes, CR, Machado, JA, Machado, MA, Madeira, AMBN, Madeira, HMF, Marino, CL, Marques, M V, Martins, EAL, Martins, EMF, Matsukuma, AY, Menck, CFM, Miracca, EC, Miyaki, CY, Monteiro-Vitorello, CB, Moon, DH, Nagai, MA, Nascimento, ALTO, Netto, LES, Nhani, A, Nobrega, FG, Nunes, LR, Oliveira, MA, de Oliveira, MC, de Oliveira, RC, Palmieri, DA, Paris, A, Peixoto, BR, Pereira, GAG, Pereira, HA, Pesquero, JB, Quaggio, RB, Roberto, PG, Rodrigues, V, Rosa, AJM, de Rosa, VE, de Sá, RG, Santelli, R V, Sawasaki, HE, da Silva, ACR, da Silva, AM, da Silva, FR, Silva, WA, da Silveira, JF, Silvestri, MLZ, Siqueira, WJ, de Souza, AA, de Souza, AP, Terenzi, MF, Truffi, D, Tsai, SM, Tsuhako, MH, Vallada, H, Van Sluys, MA, Verjovski-Almeida, S, Vettore, AL, Zago, MA, Zatz, M, Meidanis, J, and Setubal, JC, 2000. The genome sequence of the plant pathogen Xylella fastidiosa: The Xylella fastidiosa consortium of the organization for nucleotide sequencing and analysis, Sao Paulo, Brazil. Nature. 406(6792), pp.151–157.

Sluys, MA Van, Oliveira, MC De, Miyaki, CY, Furlan, LR, Camargo, LEA, Silva, ACR, Moon, DH, Takita, MA, Lemos, EGM, Machado, MA, Ferro, MIT, Silva, FR, Goldman, MHS, Goldman, GH, Lemos, MVF, Tsai, SM, Carrer, H, Carraro, DM, Oliveira, RC De, Nunes, LR, Siqueira, WJ, Coutinho, LL, Kimura, ET, Ferro, ES, Harakava, R, Kuramae, EE, Marino, CL, Giglioti, E, Abreu, IL, Alves, LMC, Amaral, AM, Baia, GS, Blanco, SR, Brito, MS, Cannavan, FS, Celestino, A V, Cunha, AF, Fenille, RC, Ferro, JA, Formighieri, EF, Kishi, LT, Leoni, SG, Oliveira, AR, Jr, VER, Sassaki, FT, Sena, JAD, Souza, AA De, Truffi, D, Tsukumo, F, Yanai, GM, Zaros, LG, Civerolo, EL, Simpson, AJG, Jr, NFA, Setubal, JC, and Kitajima, JP, 2003. Comparative Analyses of the Complete Genome Sequences of Pierce ' s Disease and Citrus Variegated Chlorosis Strains of Xylella fastidiosa. Journal of Bacteriology. [online] 185(3), pp.1018–1026. Available at: <http://jb.asm.org/cgi/pmidlookup?view=long&pmid=12533478>.

De Souza, AA, Takita, MA, Coletta-Filho, HD, Caldana, C, Goldman, GH, Yanai, GM, Muto, NH, De Oliveira, RC, Nunes, LR, and Machado, MA, 2003. Analysis of gene expression in two growth states of Xylella fastidiosa and its relationship with pathogenicity. Molecular Plant-Microbe Interactions. 16(10), pp.867–875.

Strona, G, Carstens, CJ, and Beck, PSA, 2017. Network analysis reveals why Xylella fastidiosa will persist in Europe. Scientific Reports. [online] 7(1), pp.1–8. Available at: <http://dx.doi.org/10.1038/s41598-017-00077-z>.

Su, CC, Deng, WL, Jan, FJ, Chang, CJ, Huang, H, Shih, HT, and Chen, J, 2016. Xylella taiwanensis sp. nov., causing pear leaf scorch disease. International Journal of Systematic and Evolutionary Microbiology. 66, pp.4766–4771.

Sun, Q, Sun, Y, Walker, MA, and Labavitch, JM, 2013. Vascular Occlusions in Grapevines with Pierce's Disease Make Disease Symptom Development Worse. Plant Physiology. 161(3), pp.1529–1541.

Surico, G, 2013. The concepts of plant pathogenicity, virulence/avirulence and effector proteins by a teacher of plant pathology. Phytopathologia Mediterranea. 52(3), pp.399–417.

Talavera, G, and Castresana, J, 2007. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. Systematic Biology. 56(4), pp.564–577.

Tatusov, RL, Koonin, E V, and Lipman, DJ, 1997. A genomic perspective on protein families. Science. 6(3), p.e1000703.

The UniProt Consortium, 2021. UniProt: the universal protein knowledgebase in 2021. Nucleic Acids Research. 49:D1.

Thompson, JD, Higgins, DG, and Gibson, TJ, 1994. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Research. 22(22), pp.4673–4680.

Vanhove, M, Retchless, AC, Sicard, A, Rieux, A, Coletta-Filho, HD, De La Fuente, L, Stenger, DC, and Almeida, RPP, 2019. Genomic diversity and recombination among Xylella fastidiosa subspecies. Applied and Environmental Microbiology. 85(13).

Wandersman, C, 2013. Concluding remarks on the special issue dedicated to Bacterial secretion systems: Function and structural biology. Research in Microbiology. 164(6), pp.683–687.

Waterhouse, AM, Procter, JB, Martin, DMA, Clamp, M, and Barton, GJ, 2009. Jalview Version 2-A multiple sequence alignment editor and analysis workbench. Bioinformatics. 25(9), pp.1189–1191.

Wells, JM, Raju, BC, Hung, H-Y, Weisburg, WG, Mandelco-Paul, L, and Brenner, DJ, 1987. Xylella fastidiosa gen. nov., sp. nov: Gram-Negative, Xylem-Limited, Fastidious Plant Bacteria Related to Xanthomonas spp. International Journal of Systematic Bacteriology. 37(136).

Wells, JM, Raju, BC, Nyland, G, and Lowe, SK, 1981. Medium for Isolation and Growth of Bacteria Associated with Plum Leaf Scald and Phony Peach Diseases. Applied and Environmental Microbiology. 42, pp.357–363.

White, SM, Navas-Cortés, JA, Bullock, JM, Boscia, D, and Chapman, DS, 2020. Estimating the epidemiology of emerging Xylella fastidiosa outbreaks in olives. Plant Pathology. 69(8), pp.1403–1413.

Yaseen, T, Drago, S, Valentini, F, Elbeaino, T, Stampone, G, Digiaro, M, and D'Onghia, AM, 2015. On-site detection of Xylella fastidiosa in host plants and in 'spy insects' using the real-time loop-mediated isothermal amplification method. Phytopathologia Mediterranea. 54, pp.488–496.

Yuan, X, Morano, L, Bromley, R, Spring-pearson, S, Stouthamer, R, and Nunney, L, 2010. Multilocus sequence typing of Xylella fastidiosa causing Pierce's disease and oleander leaf scorch in the United States. Phytopathology. 100(6), pp.601–611.

Zaini, PAPAPA, Nascimento, R, Gouran, H, Cantu, D, Chakraborty, S, Phu, MMM, Goulart, LRLRLRLR, and Dandekar, AMAMAM, 2018. Molecular Profiling of Pierce's Disease Outlines the Response Circuitry of Vitis vinifera to Xylella fastidiosa Infection. Frontiers in Plant Science. [online] 9(June), pp.1–

16. Available at: <https://www.frontiersin.org/article/10.3389/fpls.2018.00771/full>.
Zechner, EL, Lang, S, and Schildbach, JF, 2012. Assembly and mechanisms of bacterial type IV secretion machines. Philosophical Transactions of the Royal Society B: Biological Sciences. 367(1592), pp.1073–1087.
Zhang, S, Chakrabarty, PK, Fleites, LA, Rayside, PA, Hopkins, DL, and Gabriel, DW, 2015. Three new pierce's disease pathogenicity effectors identified using xylella fastidiosa biocontrol strain EB92-1. PLoS ONE. [online] 10(7), pp.1–17. Available at: <http://dx.doi.org/10.1371/journal.pone.0133796>.

Appendix

Additional tables

Additional figures

Table E: Details of collected leaf samples from Colombia. This is a list of all 51 plants collected in Colombia. Samples were collected in triplicates for each plant sample (see **Error! Not a valid bookmark self-reference.** for a schematic). Several details were noted and measurements taken, including collection date and time, cultivar (Var) information if given, whether the plant had *Xylella*-like symptoms (**S**) or not (**A**), whether the plant was cultivated (**C**) or naturally occurring (**N**), location details, sea level in metres above median sea level (**MAMSL**), GPS coordinates in decimal degrees (**DD**; latitude, longitude), median aerial temperature in **°C**, humidity and notable observations.

ID	Date	Time	Family	Species	Var	Symptoms	Cultivation	Location description	Location	MAMSL	GPS (DD)	°C	Humidity	Notes
MALHR02001	20190625	15:00	Malvaceae	*Hibiscus rosa-sinensis*	N/A	A	C	Tulenapa research station	Urabá	30m	7.774001, -76.664901	29C	0.88	
MALHR02002	20190625	15:15	Malvaceae	*Hibiscus rosa-sinensis*	N/A	S	C	Tulenapa research station	Urabá	30m	7.774192, -76.664902	29C	0.88	
MALBCO4001	20190628	11:35	Malvaceae	*Theobroma cacao*	N/A	S	C	Farm	Sopetrán	521m	6.5377, -75.8318	23C	0.57	leafhopper on tree
MALBC04002	20190628	11:45	Malvaceae	*Theobroma cacao*	N/A	S	C	Farm	Sopetrán	521m	6.5374, -75.8318	23C	0.57	leafhopper on tree
UNKXX01001	20190625	09:51	N/A	*N/A*	N/A	A	N	Rainforest	Urabá	30m	7.7729, -76.6703	29C	0.88	
RUBAP02001	20190625	15:40	Rubiaceae	*Alibertia patinoi*	N/A	A	C	Tulenapa research station	Urabá	30m	7.775482, -76.665425	29C	0.88	
RUBAP02002	20190625	15:40	Rubiaceae	*Alibertia patinoi*	N/A	A	C	Tulenapa research station	Urabá	30m	7.775398, -76.665438	29C	0.88	
RUBAP02003	20190625	15:40	Rubiaceae	*Alibertia patinoi*	N/A	A	C	Tulenapa research station	Urabá	30m	7.775398, -76.665438	29C	0.88	
RUBAP02004	20190626	06:30	Rubiaceae	*Alibertia patinoi*	N/A	A	C	Tulenapa research station	Tulanepa	30m	7.773682, -76.654593	30C	0.74	
RUBAP02005	20190626	06:30	Rubiaceae	*Alibertia patinoi*	N/A	A	C	Tulenapa research station	Tulanepa	30m	7.775513, -76.665425	30C	0.74	
RUBAP02006	20190626	06:30	Rubiaceae	*Alibertia patinoi*	N/A	A	C	Tulenapa research station	Tulanepa	30m	7.773980, -76.656314	30C	0.74	
RUBAP02007	20190626	06:30	Rubiaceae	*Alibertia patinoi*	N/A	A	C	Tulenapa research station	Tulanepa	30m	7.773708, -76.654650	30C	0.74	
RUBCA03001	20190627	15:25	Rubiaceae	*Coffea arabica*	Geisha	A	C	Farm	Fredonia	1423m	5.970375, -75.670041	24C	0.59	
RUBCA03002	20190627	15:30	Rubiaceae	*Coffea arabica*	Geisha	A	C	Farm	Fredonia	1423m	5.9703, -75.6701	24C	0.59	
RUBCA03003	20190627	15:45	Rubiaceae	*Coffea arabica*	Geisha	S	C	Farm	Fredonia	1423m	5.9704,-75.6704	24C	0.59	
RUBCA03004	20190627	15:55	Rubiaceae	*Coffea arabica*	Colombia	S	C	Farm	Fredonia	1423m	5.9730, -75.6701	24C	0.59	
RUBCA03005	20190627	16:07	Rubiaceae	*Coffea arabica*	Colombia	A	C	Farm	Fredonia	1423m	5.9730, -75.6700	24C	0.59	

ID	Date	Time	Family	Species	Var	Symptoms	Cultivation	Location description	Location	MAMSL	GPS (DD)	°C	Humidity	Notes
RUBCA03006	20190627	16:12	Rubiaceae	*Coffea arabica*	Colombia	S	C	Farm	Fredonia	1423m	5.9730, -75.6701	24C	0.59	
RUBCA03007	20190627	16:42	Rubiaceae	*Coffea arabica*	Caturra	S	C	Farm	Fredonia	1786m	5.99748, -75.6644	24C	0.59	
RUBCA03008	20190627	16:46	Rubiaceae	*Coffea arabica*	Caturra	S	C	Farm	Fredonia	1786m	5.9749, -75.6643	24C	0.59	
RUBCA03009	20190627	16:49	Rubiaceae	*Coffea arabica*	Caturra	S	C	Farm	Fredonia	1786m	5.9748, -75.6642	24C	0.59	
RUBCA03010	20190627	16:54	Rubiaceae	*Coffea arabica*	Pajarito	S	C	Farm	Fredonia	1786m	5.9748, -75.6644	24C	0.59	
RUBCA03011	20190627	16:59	Rubiaceae	*Coffea arabica*	Pajarito	S	C	Farm	Fredonia	1786m	5.9747, -75.6644	24C	0.59	
RUBCA03012	20190627	17:07	Rubiaceae	*Coffea arabica*	Pajarito	S	C	Farm	Fredonia	1786m	5.9746, -75.6643	24C	0.59	
RUBCA03013	20190627	17:10	Rubiaceae	*Coffea arabica*	Castillo	S	C	Farm	Fredonia	1786m	5.9748, -75.6645	24C	0.59	
RUBCA03014	20190627	17:13	Rubiaceae	*Coffea arabica*	Castillo	S	C	Farm	Fredonia	1786m	5.9749, -75.6645	24C	0.59	
RUBCA03015	20190627	17:20	Rubiaceae	*Coffea arabica*	Castillo	S	C	Farm	Fredonia	1786m	5.9740, -75.6645	24C	0.59	
RUBCA05001	20190703	11:15	Rubiaceae	*Coffea arabica*	N/A	S	C	EAFIT Campus	Medellín	1504m	6.2002, -75.5785	23C	0.64	rust
RUBCA05002	20190703	11:25	Rubiaceae	*Coffea arabica*	N/A	S	C	EAFIT Campus	Medellín	1504m	6.2001, -75.5785	23C	0.64	rust
RUBTX06001	20190704	14:30	Rubiaceae	*Tocoyena*	N/A	S	C	Botanic gardens	Botanic Gardens, Medellín	1474m	6.2693, -75.5631	28C	0.51	rust
RUBGA06001	20190704	14:45	Rubiaceae	*Genipa americana*	N/A	S	C	Botanic gardens	Botanic Gardens, Medellín	1474m	6.2698, -75.5625	28C	0.51	
RUBPL06001	20190704	14:55	Rubiaceae	*Posoqueria latifolia*	N/A	S	C	Botanic gardens	Botanic Gardens, Medellín	1474m	6.2699, -75.5626	28C	0.51	
RUBPX06001	20190704	15:00	Rubiaceae	*Pogonopus*	N/A	S	C	Botanic gardens	Botanic Gardens, Medellín	1474m	6.2700, -75.5625	28C	0.51	
2251	20190704	15:15	Rubiaceae	*Cosmibuena grandiflora*	N/A	S	C	Botanic gardens	Botanic Gardens, Medellín	1474m	6.2713, -75.5626	28C	0.51	
RUBHP06001	20190704	15:20	Rubiaceae	*Hamelia patens*	N/A	A	C	Botanic gardens	Botanic Gardens, Medellín	1474m	6.2705, -75.5622	28C	0.51	
RUBHP06002	20190704	15:30	Rubiaceae	*Hamelia patens*	N/A	S	C	Botanic gardens	Botanic Gardens, Medellín	1474m	6.2706, -75.5623	28C	0.51	
RUBIJ06001	20190704	15:35	Rubiaceae	*Ixora javanica*	N/A	A	C	Botanic gardens	Botanic Gardens, Medellín	1474m	6.2708, -75.5623	28C	0.51	
RUBIH06001	20190704	15:55	Rubiaceae	*Isertia haenkeana*	N/A	S	C	Botanic gardens	Botanic Gardens, Medellín	1474m	6.2723, -75.5642	28C	0.51	
RUTCL02001	20190625	15:55	Rutaceae	*Citrus lemón*	N/A	S	C	Tulenapa research station	Urabá	30m	7.773901, -76.664054	29C	0.88	
RUTCH05001	20190703	10:30	Rutaceae	*Citrus hystrix*	N/A	S	C	EAFIT Campus	Medellín	1504m	6.2001, -75.5783	23C	0.64	

ID	Date	Time	Family	Species	Var	Symptoms	Cultivation	Location description	Location	MAMSL	GPS (DD)	°C	Humidity	Notes
RUTCH06001	20190704	15:40	Rutaceae	*Citrus hystrix*	N/A	S	C	Botanic gardens	Botanic Gardens, Medellín	1474m	6.2699, -75.5629	28C	0.51	
RUTCS07001	20190705	09:10	Rutaceae	*Citrus sinensis*	Valencia	S	C	Farm	La Pintada	729m	5.8284, -75.6082	24C	0.76	CVC
RUTCS07002	20190705	09:15	Rutaceae	*Citrus sinensis*	Valencia	S	C	Farm	La Pintada	729m	5.8284, -75.6082	24C	0.76	CVC
RUTCS07003	20190705	09:20	Rutaceae	*Citrus sinensis*	Valencia	S	C	Farm	La Pintada	729m	5.8283, -75.6082	24C	0.76	CVC
RUTCS07004	20190705	09:40	Rutaceae	*Citrus sinensis*	Salustiana	S	C	Farm	La Pintada	696m	5.8268, -75.6123	24C	0.76	CVC
RUTCS07005	20190705	09:45	Rutaceae	*Citrus sinensis*	Salustiana	S	C	Farm	La Pintada	696m	5.8269, -75.6124	24C	0.76	CVC
RUTCS07006	20190705	09:50	Rutaceae	*Citrus sinensis*	Salustiana	S	C	Farm	La Pintada	696m	5.8267, -75.6124	24C	0.76	CVC
RUTCL07001	20190705	10:40	Rutaceae	*Citrus lemón*	Tahiti	A	C	Farm	La Pintada	774m	5.8235, -75.6076	24C	0.76	
RUTCL07002	20190705	10:45	Rutaceae	*Citrus lemón*	Tahiti	S	C	Farm	La Pintada	774m	5.8235, -75.6075	24C	0.76	smaller fruits, lighter leaves
RUTCL07003	20190705	10:55	Rutaceae	*Citrus lemón*	Tahiti	A	C	Farm	La Pintada	774m	5.8235, -75.6072	24C	0.76	
RUTCL07004	20190706	11:55	Rutaceae	*Citrus lemón*	Tahiti	A	C	Farm	La Pintada	774m	5.8236, -75.6071	24C	0.76	

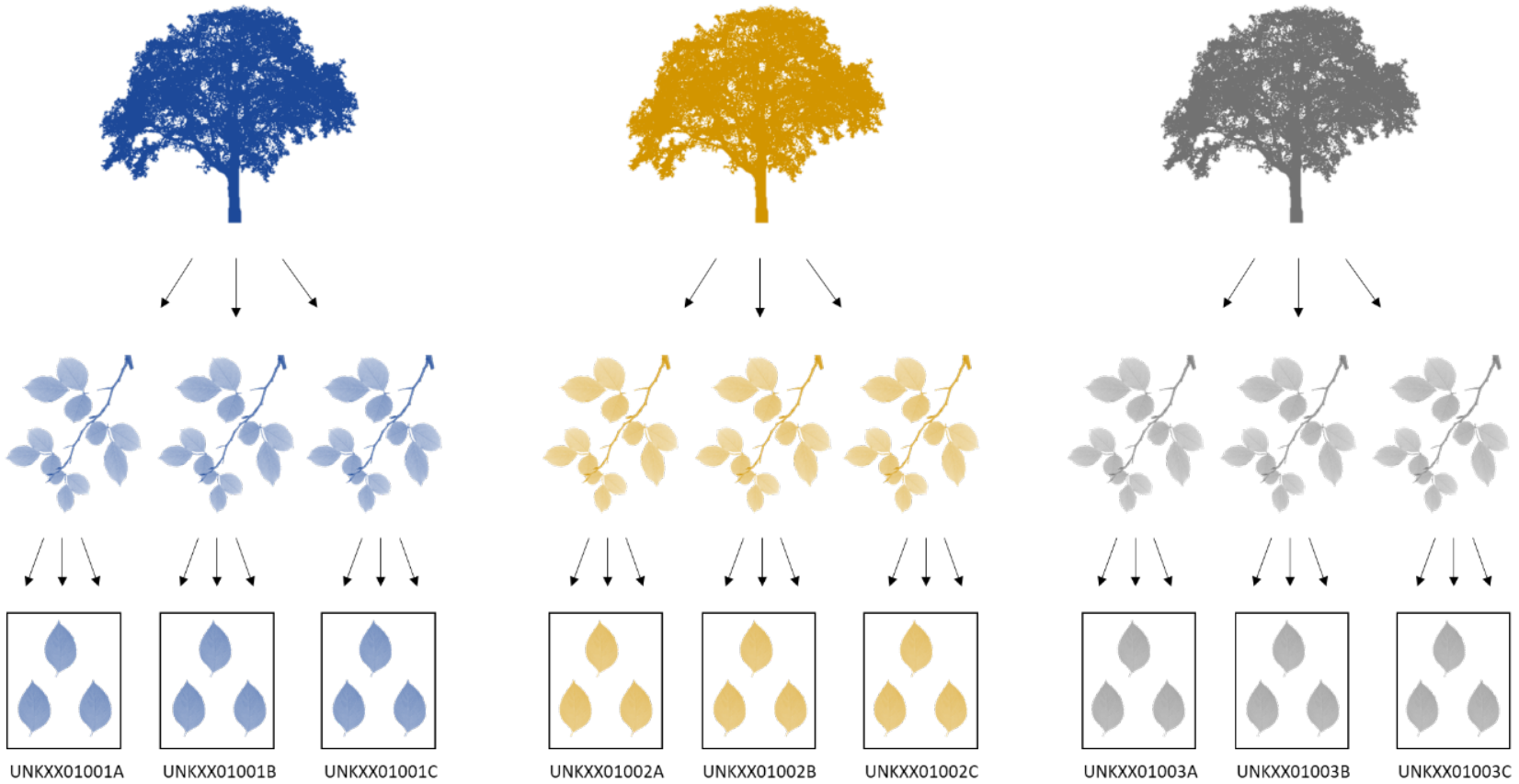


Figure L: A schematic of sampling leaves in Colombia. A total of 51samples were collected in the duration of a two-week survey in Colombia. Sampling was performed as follows: whenever possible, three plants of each species at each location were sampled. Per plant, three branches were selected, from which three leaves were removed using scissors disinfected in 70% ethanol and placed into a polyethylene bag. This would ultimately result in having triplicates of each plant originally sampled. Each sample was given a unique eleven-digit ID comprising of the first three letters of the plant family, the first letter of the genus, the first letter of the species, two digits indicating the location, three digits indicating the sample number, and a letter indicating the replicate (A, B or C).

RUBCA03001: *C. arabica* cv. Geisha

RUBCA03002: *C. arabica* cv. Geisha

RUBCA03003: *C. arabica* cv. Colombia

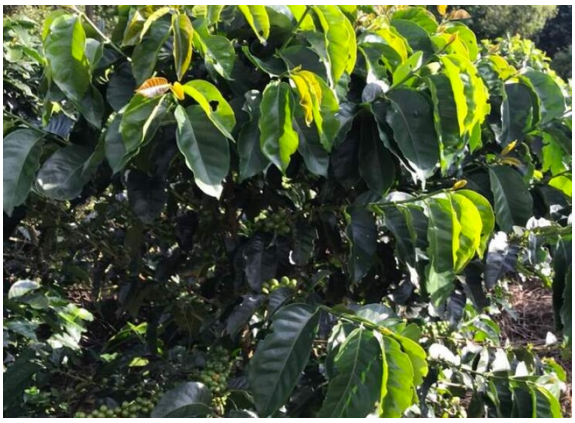
RUBCA03005: *C. arabica* cv. Colombia

RUBCA03006: *C. arabica* cv. Colombia

RUBCA03006: *C. arabica* cv. Colombia

RUBCA03007: *C. arabica* cv. Caturra

RUBCA03010: *C. arabica* cv. Pajarito



RUBCA03008: *C. arabica* cv. Caturra

RUBCA03008: *C. arabica* cv. Caturra

RUBCA03011: *C. arabica* cv. Pajarito

RUBCA03013: *C. arabica* cv. Castillo

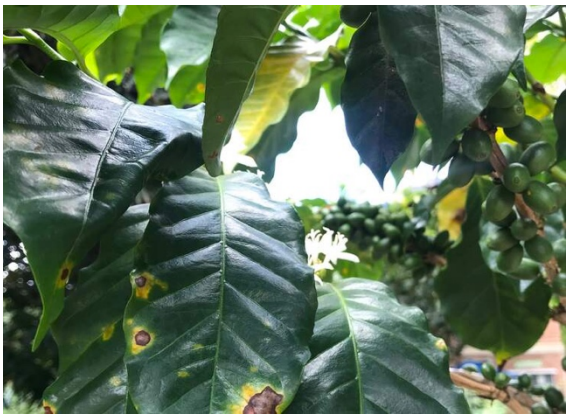
RUBCA03012: *C. arabica* cv. Pajarito

RUBCA03012: *C. arabica* cv. Pajarito

RUBCA03015: *C. arabica* cv. Castillo

RUBCA03015: *C. arabica* cv. Castillo

RUBCA05001: *C. arabica* cv. unknown



RUBCA05001: *C. arabica* cv. unknown

Figure M: Photographs of Colombian samples that tested positive for *Xylella fastidiosa* (*Xf*). *Xf* was identified by PCR in samples collected from the plants photographed above. See **Appendix Table E** for full details of each sample. Samples RUBCA03003, -006, -007, -008, -010, -011, -012, -013 and -015 had *Xf*-like symptoms, though difficult to see in some photographs.

1. Turn on water bath at 65°C
2. Clean surface of leaves to be used with ethanol
3. Place 0.5-1.0g of fresh small pieces of midribs, petioles, leaf basal part or twigs (1/4 of amount if lyophilised) into suitable tubes and immediately freeze dry in liquid nitrogen
4. Homogenise leaves using liquid nitrogen in pestle and mortar, or a tissue grinder
5. Add 5ml of CTAB buffer per 0.5-1.0g sample tube
6. Transfer sample and CTAB mix to 15ml falcon tubes and mix well
7. Heat at 65°C for 30min
8. Centrifuge at 16,000g, RT for 5min
9. Transfer 1ml aliquots of supernatant to fresh 2ml microcentrifuge tube (do not transfer any plant debris!)
10. Add 5µl of RNAse A (10mg/ml)
11. Incubate at 37°C overnight
12. Add 1ml of chloroform (isoamyl alcohol [24:1])
13. Mix well by shaking
14. Centrifuge at 16,000g for 10min
15. Transfer 700µl supernatant to new 1.5ml microcentrifuge tube
16. Add 490µl (or ~0.7 of available supernatant volume) of 2-propanol (room temperature)
17. Mix by inverting twice
18. Incubate at RT for 20min
19. Centrifuge at 16,000g, 4°C for 20min (recovery of pellet)
20. Remove supernatant
21. Wash pellet with 1ml of cold 70% ethanol
22. Centrifuge at 16,000g, 4°C for 10min
23. Remove supernatant and wash pellet again in 500µl of RT 70% ethanol
24. Mix by inversion
25. Centrifuge at 16,000g, 4°C for 10min
26. Remove supernatant and wash pellet again in 500µl of RT 70% ethanol
27. Mix by inversion
28. Centrifuge at 16,000g, 4°C for 10min
29. Remove supernatant and air-dry (~20min)
30. Re-suspend pellet in 100-150µl of TE buffer

Figure N: CTAB-based DNA extraction protocol. This is the modified protocol for the total DNA extraction from leaf samples. The original protocol was designed by EPPO (2016). The original protocol was modified as follows: homogenisation of leaf tissue was performed using pestle and mortar, and liquid nitrogen instead of a mechanical homogeniser; an overnight RNAse step was included to degrade unwanted RNA in the sample; room temperature 2-propanol was used for precipitation of DNA instead of cold 2-propanol to reduce the amount of salts being co-precipitated; finally, each sample was washed three times with 70% ethanol to ensure the removal of all contaminants.

Table F: Primer sequences used in this project. This is a list of all primer sequences used in this research, the target sequence and PCR conditions for each reaction.

Primer name	Amplicon size (bp)	Forward primer sequence	Reverse primer sequence	Target sequence	Reference	PCR conditions
27F / 1492R	~1,500	AGAGTTTGATCCTGGCTCAG	CTACGGCTACCTTGTTACGA	Bacteria-specific; 16S rRNA	Muyzer, De Waal and Uitterlinden, 1993	95°C 60s 95°C 30s, 51°C 30s, 72°C 120s (35x, redTaq polymerase) 72°C 5min
RST31 / RST33	733	GCGTTAATTTTCGAAGTGATTCGATTGC	CACCATTCGTATCCCGGTG	*Xylella*-specific; 3' end of the gene *rpoD*, coding for an RNA polymerase sigma-70 factor	Minsavage *et al.*, 1994	95°C 60s 95°C 30s, 57.9°C 30s, 72°C 45s (40x, Platinum Taq polymerase) 72°C 5min
16S-23SF / 16S-23SR	650	GATGACTGGGGTGAAGTCGT	GACACTTTTCGCAGGCTACC	*Xylella*-specific; 16S-23S intergenic spacer	Martinati *et al.*, 2005	95°C 60s 95°C 30s, 57°C 30s, 72°C 120s (40x, redTaq polymerase) 72°C 5min
***Xylella*-specific primers for multi-locus sequence typing (MLST)**						
leuA-F / leuA-R	708	GGTGCACGCCAAATCGAATG	GTATCGTTGTGGCGTACACTG	*leuA*, coding for 2-isopropylmalate synthase	Yuan *et al.*, 2010	95°C 60s 95°C 30s, 64°C 30s, 72°C 45s (35x, Platinum Taq polymerase) 72°C 5min
petC-F / petC-R	533	GCTGCCATTCGTTGAAGTACCT	GCACGTCCTCCCAATAAGCCT	*petC*, coding for ubiquinol cytochrome c oxidoreductase C1 subunit	Yuan *et al.*, 2010	95°C 60s 95°C 30s, 66°C 30s, 72°C 45s (35x, Platinum Taq polymerase) 72°C 5min
malF-F / malF-R	730	TTGCTGGTCCTGCGGTGTTG	GACAGCAGAAGCACGTCCCAGAT	*malfF*, coding for ABC transporter sugar permease	Yuan *et al.*, 2010	95°C 60s 95°C 30s, 68.5°C 30s, 72°C 45s (35x, Platinum Taq polymerase) 72°C 5min
cysG-F / cysG-R	600	GCCGAAGCAGTGCTGGAAG	GCCATTTTCGATCAGTGCAAAAG	*cysG*, coding for sirohaem synthase	Yuan *et al.*, 2010	95°C 60s 95°C 30s, 64.2°C 30s, 72°C 45s (35x, Platinum Taq polymerase) 72°C 5min
holC-F / holC-R	379	ATGGCACGCGCCGACTTCT	ATGTCGTGTTTGTTCATGTGCAGG	*holC*, coding for DNA polymerase III holoenzyme chi subunit	Yuan *et al.*, 2010	95°C 60s 95°C 30s, 67.1°C 30s, 72°C 45s (35x, Platinum Taq polymerase) 72°C 5min
nuoL-F / nuoL-R	557	TAGCGACTTACGGTTACTGGGC	ACCACCGATCCACAACGCAT	*nuoL*, coding for NADH ubiquinone oxidoreductase NQO12 subunit	Yuan *et al.*, 2010	95°C 60s 95°C 30s, 66.9°C 30s, 72°C 45s (35x, Platinum Taq polymerase) 72°C 5min
gltT-F / gltT-R	654	TCATGATCCAAATCACTCGCTT	ACTGGACGCTGCCTCGTAAACC	*gltT*, coding for glutamate symport protein	Yuan *et al.*, 2010	95°C 60s 95°C 30s, 63.2°C 30s, 72°C 45s (35x, Platinum Taq polymerase) 72°C 5min

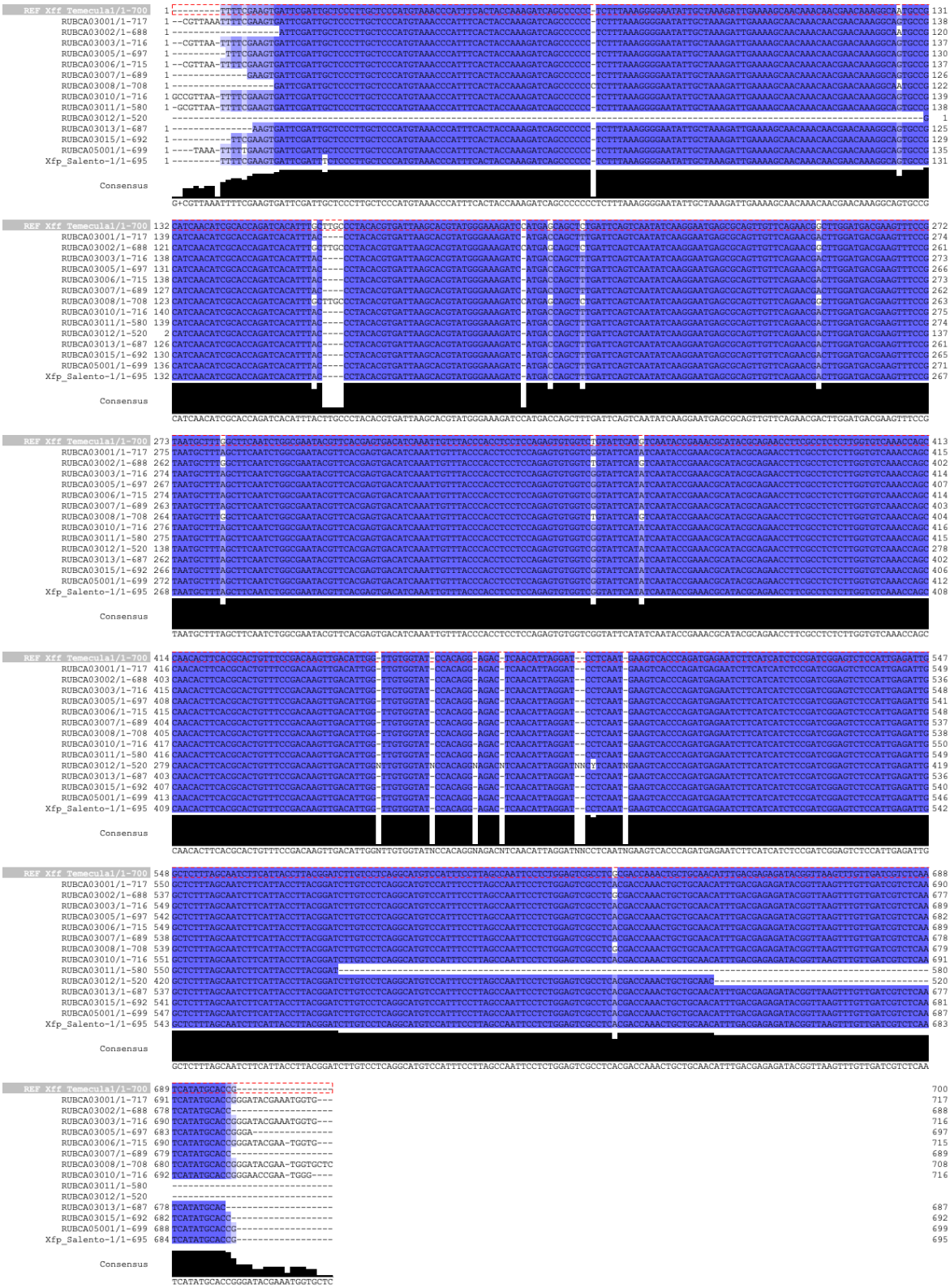


Figure O: Alignment of positive Colombian XF1 PCR amplicons. RUBCA03001, RUBCA03003, RUBCA03005, RUBCA03006, RUBCA03007, RUBCA03011, RUBCA03015 and RUBCA05001 were sent for Sanger sequencing using the Eurofins GATC LightRun service. Consensus sequences acquired using DNASTAR's Sanger Sequence Assembly and the *rpoD* gene sequence of *Xylella fastidiosa* (*Xf*) subsp. *fastidiosa* strain 9a5c, the *Xf* reference genome. Once consensus sequences were obtained, multiple sequence alignment by progressive strategy was performed using the program T-Coffee (Notredame, Higgins and Heringa, 2000). The alignment was finally visualised using JalView (Waterhouse, et al., 2009).

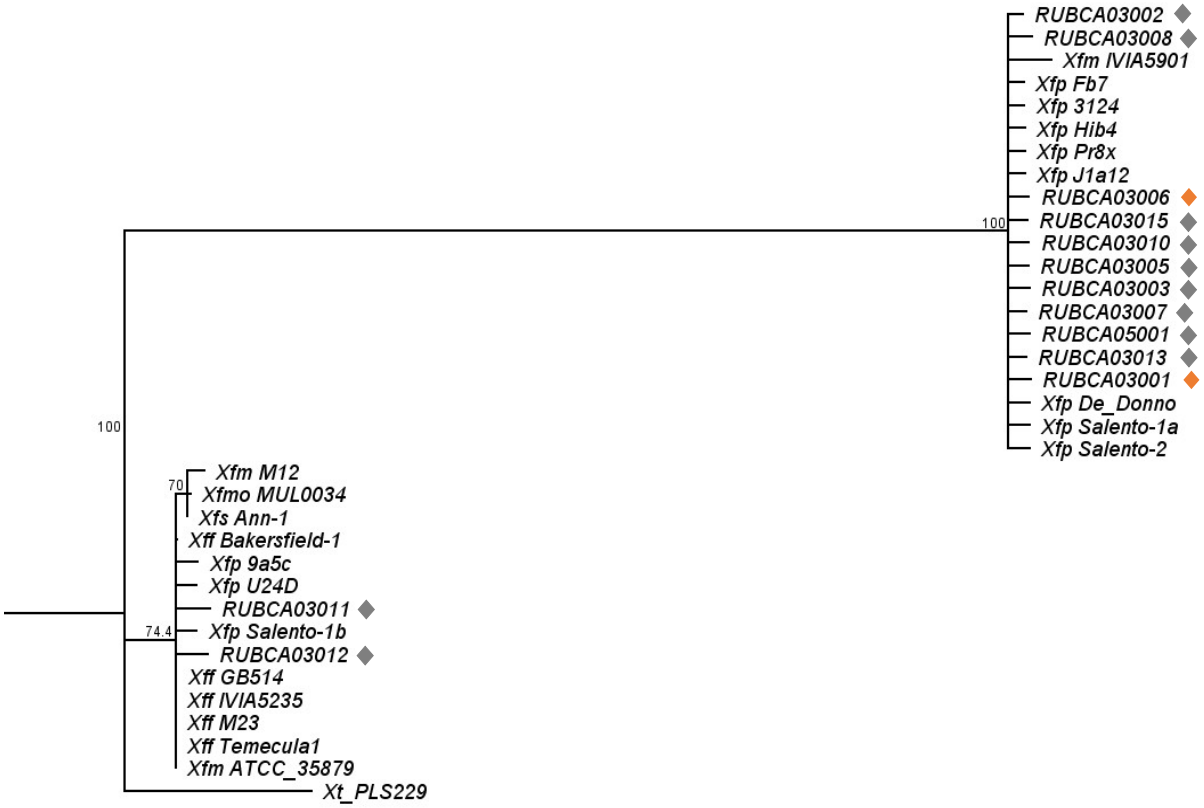


Figure P: Cladogram of Colombian *Xylella fastidiosa* (*Xf*) *rpoD* amplicons. Strains marked with a diamond are Colombian *Xf* strains. Those marked in **orange** orginated from a plant displaying *Xf*-like symptoms, and those marked in **grey** originated from asymptomatic plants.

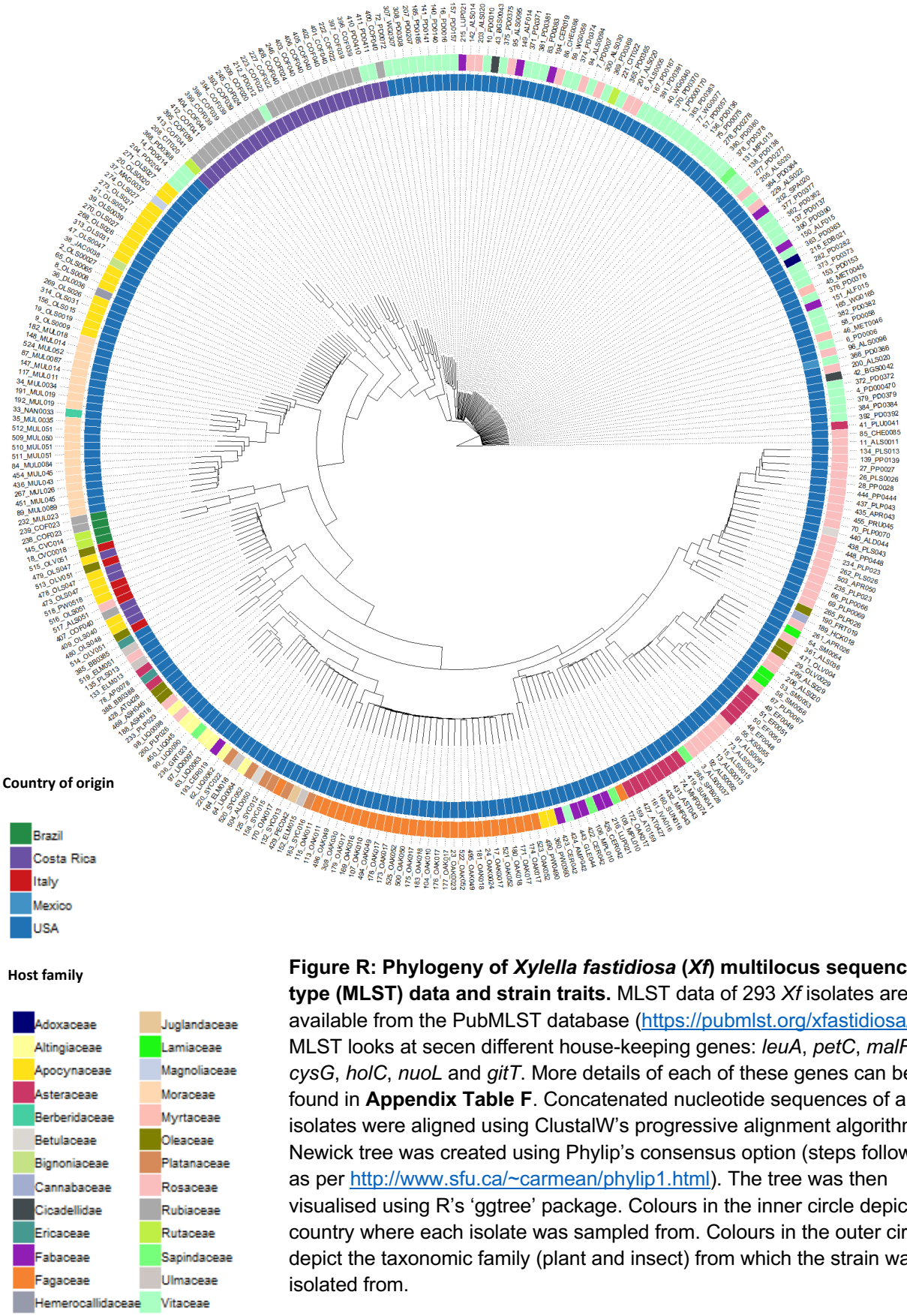


Figure R: Phylogeny of *Xylella fastidiosa* (*Xf*) multilocus sequencing type (MLST) data and strain traits. MLST data of 293 *Xf* isolates are available from the PubMLST database (https://pubmlst.org/xfastidiosa/). *Xf* MLST looks at secen different house-keeping genes: *leuA*, *petC*, *malF*, *cysG*, *holC*, *nuoL* and *gitT*. More details of each of these genes can be found in **Appendix Table F**. Concatenated nucleotide sequences of all 293 isolates were aligned using ClustalW's progressive alignment algorithm. A Newick tree was created using Phylip's consensus option (steps followed as per http://www.sfu.ca/~carmean/phylip1.html). The tree was then visualised using R's 'ggtree' package. Colours in the inner circle depict the country where each isolate was sampled from. Colours in the outer circle depict the taxonomic family (plant and insect) from which the strain was isolated from.

Table G: List of *Xylella fastidiosa* (*Xf*) host plants. A list of host plants wherein *Xf* was detected was compiled using data from EC (2018), EFSA (2018) and EPPO (n.d.). The list includes the *Xf* subspecies found in each host plant (if available). **N**, *Xf* detected in a natural setting; **E**, *Xf* detected in an experimental setting; **U**, no information available in which setting *Xf* was detected. The list also includes information whether *Xf* was found in the European countries France, Spain, Germany, Italy and Portugal. *Xf* was also detected in olive in Belgium, but no subspecies information has yet been published. No information could be found of the presence of host plants in Europe of rows highlighted in orange.

Hosts	Common names	*fastidiosa*	*multiplex*	*pauca*	*sandyi*	France	Spain	Germany	Italy	Belgium	Portugal	Disease	Reference
Acacia dealbata	silver wattle, blue wattle, mimosa		N			*multiplex*							EC (2018), EFSA (2018)
Acacia saligna	coojong, golden wreath wattle, orange wattle, blue-leafed wattle, Western Australian golden wattle		N	N		*multiplex*	*pauca*		*pauca*				EC (2018), EFSA (2018)
Acacia sp.			N	N			*multiplex, pauca*						EFSA (2018)
Acer griseum			N										EFSA (2018)
Acer platanoides			N										EFSA (2018)
Acer pseudoplatanus	sycamore		N			*multiplex*							EC (2018), EFSA (2018)
Acer rubrum			EN										EFSA (2018)
Acer sp.		N											EFSA (2018)
Alnus rhombifolia			N										EFSA (2018)
Amaranthus blitoides		E											EFSA (2018)
Ambrosia acanthicarpa		E											EFSA (2018)
Ambrosia psilostachya			N										EFSA (2018)
Ambrosia psilostachya var. texana			N										EFSA (2018)
Ambrosia trifida			N			*multiplex*							EFSA (2018)
Ampelopsis cordata			N										EFSA (2018)
Anthyllis hermanniae	Maltese yellow kindey vetch, Maltese shrubby kidney vetch		N			*multiplex*							EC (2018), EFSA (2018)
Artemisia arborescens	tree wormwood		N			*multiplex*							EC (2018), EFSA (2018)
Asparagus acutifolius	wild asparagus		N	N		*multiplex*			*pauca*				EC (2018), EFSA (2018)
Calicotome spinosa	thorn broom	N	U	U			*fastidiosa*						EC (2018), EFSA (2018)
Calicotome villosa	hairy thorny broom		N			*multiplex*							EC (2018), EFSA (2018)
Carya illinoinensis			EN										EFSA (2018)
Carya sp.			N										EFSA (2018)
Catharanthus roseus		E		EN					*pauca*				EFSA (2018)
Catharanthus sp.	periwinkles			U					*pauca*				EC (2018)
Celtis occidentalis			N										EFSA (2018)

Hosts	Common names	*fastidiosa*	*multiplex*	*pauca*	*sandyi*	France	Spain	Germany	Italy	Belgium	Portugal	Disease	Reference
Cercis canadensis			N										EFSA (2018)
Cercis occidentalis		N	N										EFSA (2018)
Cercis siliquastrum	Judas tree	N	N			*multiplex*							EC (2018), EFSA (2018)
Chenopdoium album	fat hen, lamb's quarters, melde, goosefoot (weed)			N					*pauca*				EC (2018), EFSA (2018)
Chenopodium quinoa		E											EFSA (2018)
Chionanthus sp.			N										EFSA (2018)
Cistus albidus	white leaved rock rose, grey-leaved cistus	U	U	U									EC (2018)
Cistus creticus	Cretan rock rose, pink rock-rose, hoary rock-rose		N	N		*multiplex*			*pauca*				EC (2018), EFSA (2018)
Cistus monspeliensis	Montpellier cistus	N	N			*multiplex*	*fastidiosa*						EC (2018), EFSA (2018)
Cistus salviifolius	sage-leaved rock-rose, salvia cistus, Gallipoli rose		U			*multiplex*							EC (2018), EFSA (2018)
Cistus sp.			N	N		*multiplex*							EFSA (2018)
Citroncirus sp.													EPPO (n.d.)
Citrus sp.				EN									EFSA (2018)
Citrus x sinensis	sweet orange	N	N	EN		*multiplex*						citrus-variegated chlorosis (CVC)	EFSA (2018)
Coffea arabica		N		N									EFSA (2018)
Coffea canephora		N			N								EFSA (2018)
Coffea sp.	coffee	U	U	N	N							coffee leaf scorch (CLS)	EC (2018), EFSA (2018)
Conium maculatum		E											EFSA (2018)
Convolvulus cneorum	shrubby bindweed, silverbush		U										EC (2018)
Convulus arvensis		E											EFSA (2018)
Coronilla glauca	scorpion vetch, shrubby scorpion-vetch		U										EC (2018)
Coronilla valentina	bastard senna, shrubby scorpion-vetch, scorpion vetch		N			*multiplex*							EC (2018), EFSA (2018)
Coronilla valentina ssp. glauca			N			*multiplex*							EFSA (2018)
Cyperaceae sp.													EPPO (n.d.)
Cyperus esculentus		E											EFSA (2018)
Cytisus racemosus			N										DEFRA (2016)
Cytisus scoparius	common broom, Scotch broom		N			*multiplex*							EC (2018), EFSA (2018)
Cytisus sp.			N			*multiplex*							EFSA (2018)
Cytisus villosus	hairy broom		N			*multiplex*							EC (2018), EFSA (2018)

Hosts	Common names	*fastidiosa*	*multiplex*	*pauca*	*sandyi*	France	Spain	Germany	Italy	Belgium	Portugal	Disease	Reference
Datura wrightii		E											EFSA (2018)
Dendranthema x grandiflorum		E											EFSA (2018)
Dodonaea viscosa	hopbush			N					*pauca*				EC (2018), EFSA (2018)
Echinochloa crus-galli		E											EFSA (2018)
Encelia farinosa			N										EFSA (2018)
Eremophila maculata	spotted fuchsia-bush, spotted emu bush			N					*pauca*				EC (2018), EFSA (2018)
Erigeron bonariensis	hairy fleabane, flax-leaf fleabane, wavy-leaf fleabane, Argentine fleabane (weed)			N					*pauca*				EC (2018), EFSA (2018)
Erigeron canadensis		E											EFSA (2018)
Erigeron sumatrensis	Guernsey fleabane (weed)			N					*pauca*				EC (2018), EFSA (2018)
Eriochloa graciis		E											EFSA (2018)
Erodium moschatum		E											EFSA (2018)
Erysimum hybrids		N											EFSA (2018)
Erysimum sp.	wallflower	U						*fastidiosa*					EC (2018)
Eucalyptus camaldulensis		E											EFSA (2018)
Eucalyptus globulus		E											EFSA (2018)
Euphorbia terracina	false caper, coastal spurge, Geraldton carnation weed			N					*pauca*				EC (2018), EFSA (2018)
Euryops chrysanthemoides	African bush daisy, bull's-eye		N			*multiplex*							EC (2018), EFSA (2018)
Fallopia japonica		N											EFSA (2018)
Ficus carica	common fig		N				*multiplex*						EC (2018), EFSA (2018)
Fortunella sp.													EPPO (n.d.)
Fraxinus americana			N										EFSA (2018)
Fraxinus angustifolia	narrow-leafed ash		N				*multiplex*						EC (2018), EFSA (2018)
Fraxinus sp.			N										EFSA (2018)
Genista corsica	broom		N			*multiplex*							EC (2018), EFSA (2018)
Genista ephedroides	broom		N			*multiplex*							EC (2018), EFSA (2018)
Genista lucida	broom	N	U	U			*fastidiosa*						EC (2018), EFSA (2018)
Genista sp.			N			*multiplex*							EFSA (2018)
Genista x spachiana (syn. Cytisus racemosus Broom)	sweet broom		N			*multiplex*							EC (2018), EFSA (2018)

Hosts	Common names	*fastidiosa*	*multiplex*	*pauca*	*sandyi*	France	Spain	Germany	Italy	Belgium	Portugal	Disease	Reference
Ginkgo biloba			N										EFSA (2018)
Gleditsia triacanthos			N										EFSA (2018)
Grevillea juniperina	juniper-leaf grevillea, juniper grevillea, prickly spider-flower		U	N					*pauca*				EC (2018), EFSA (2018)
Hebe sp.	shrubby veronica		N	N		*multiplex*			*pauca*				EC (2018)
Helianthus annuus		E	N										EFSA (2018)
Helianthus sp.			N										EFSA (2018)
Helichrysum italicum	curry plant, Italian strawflower, immortelle		N			*multiplex*							EC (2018), EFSA (2018)
Helicrysum stoechas	shrubby everlasting	U	U	U									EC (2018)
Heliotropium europaeum	common heliotrope, European heliotrope, European turn-sole			N					*pauca*				EC (2018), EFSA (2018)
Hemerocallis sp.					N								EFSA (2018)
Hibiscus rosa-sinensis			N										EFSA (2018)
Ipomoea purpurea		E											EFSA (2018)
Iva annua			N										EFSA (2018)
Jacaranda mimosifolia					N								EFSA (2018)
Juglans regia	common walnut, Persian walnut, English walnut, Circassian walnut	N					*fastidiosa*						EC (2018), EFSA (2018)
Koelreuteria bipinnata			N										EFSA (2018)
Lactuca serriola		E											EFSA (2018)
Lagerstroemia indica			N										EFSA (2018)
Lagerstroemia sp.			N										EFSA (2018)
Laurus nobilis	bay, bay laurel, sweet bay, true laurel, Grecian laurel, laurel tree, laurel		U	N					*pauca*				EC (2018), EFSA (2018)
Lavandula angustifolia	English lavender, lavender, true lavender		N	N		*multiplex*			*pauca*				EC (2018), EFSA (2018)
Lavandula dentata	French lavender, fringed lavender	U	N	N		*multiplex*	*multiplex, pauca*		*pauca*		*multiplex*		EC (2018), EFSA (2018)
Lavandula sp.			N			*multiplex*						asymptomatic	EFSA (2018)
Lavandula stoechas	French lavender, Spanish lavender, topped lavender		N	N		*multiplex*	unknown		*pauca*				EC (2018), EFSA (2018)
Lavandula x allardii (syn. Lavandula x heterophylla)	Allards lavender		U			*multiplex*							EC (2018)
Lavandula x chaytoriae	velvet lavender, Sawyers, lavender 'Sawyers'	U	U	U									EC (2018)
Lavandula x heterophylla			N			*multiplex*							EFSA (2018)
Lavandula x intermedia	fat lavender, hybrid lavender		N			*multiplex*							EC (2018), EFSA (2018)

Hosts	Common names	*fastidiosa*	*multiplex*	*pauca*	*sandyi*	France	Spain	Germany	Italy	Belgium	Portugal	Disease	Reference
Liquidambar styraciflua			EN										EFSA (2018)
Liriodendron tulipifera			N										EFSA (2018)
Lonicera japonica	Japanese honeysuckle, golden-and-silver honeysuckle		U										EC (2018)
Lupinus aridorum		N											EFSA (2018)
Lupinus villosus			N										EFSA (2018)
Magnolia grandiflora		N											EFSA (2018)
Malva parviflora		E											EFSA (2018)
Medicago sativa	alfalfa, lucerne	EN	N			*multiplex*						lucerne dwarf	EC (2018), EFSA (2018)
Metrosideros excelsa	pōhutukawa, New Zealand pohutukawa, New Zealand Christmas tree, New Zealand Christmas bush, iron tree		N			*multiplex*							EC (2018), EFSA (2018)
Metrosideros sp.		N											EFSA (2018)
Morus alba													EPPO (n.d.)
Morus rubra													EPPO (n.d.)
Myoporum insulare	blueberry tree, common boobialla, native juniper			N					*pauca*				EFSA (2018)
Myrtus communis	common myrtle		N	N		*multiplex*			*pauca*				EC (2018), EFSA (2018)
Nerium oleander	oleander	N	N	EN	EN		unknown	*fastidiosa*	*pauca*			oleander leaf scorch (OLS)	EC (2018), EFSA (2018)
Nicotiana clevelandii				E									EFSA (2018)
Nicotiana glauca		E											EFSA (2018)
Nicotiana tabacum		E	E	E									EFSA (2018)
Olea europaea	olive	E	EN	EN			*multiplex, pauca*		*pauca*			olive-quick-decline syndrome (OQDS)	EC (2018), EFSA (2018)
Olea europaea ssp. sylvestris	wild olive		N	N			*multiplex, pauca*					olive-quick-decline syndrome (OQDS)	EFSA (2018)
Olea sp.				N					*pauca*	NA			EFSA (2018)
Pelargonium graveolens	sweet scented geranium, rose geranium, old fashion rose geranium, rose-scent geranium		N			*multiplex*							EC (2018)
Pelargonium sp.			N			*multiplex*							EFSA (2018)
Pelargonium x fragrans	nutmeg pelargonium			N					*pauca*				EFSA (2018)
Persea americana													EPPO (n.d.)
Phagnalon saxatile			N			*multiplex*							EC (2018), EFSA (2018)
Phillyrea latifolia	green olive tree, mock privet			N					*pauca*				EC (2018), EFSA (2018)

Hosts	Common names	*fastidiosa*	*multiplex*	*pauca*	*sandyi*	France	Spain	Germany	Italy	Belgium	Portugal	Disease	Reference
Platanus occidentalis			EN										EFSA (2018)
Pluchea odorata		N											EFSA (2018)
Polygala moleracea		E											EFSA (2018)
Polygala myrtifolia	myrtle-leaf milkwort	N	EN	EN	N	*multiplex, pauca, sandyi*	*fastidiosa, multiplex, pauca*		*pauca*				EC (2018), EFSA (2018)
Polygala sp.			N			*multiplex*							EFSA (2018)
Polygala x dalmaisiana			N			*multiplex*							EFSA (2018)
Polygala x grandiflora nana			N			*multiplex*							EFSA (2018)
Portulaca oleracea		E											EFSA (2018)
Prunus angustifolia													EPPO (n.d.)
Prunus armeniaca			N										EFSA (2018)
Prunus avium	wild cherry, sweet cherry, gean	N	N	EN		*multiplex*	*fastidiosa*		*pauca*				EC (2018), EFSA (2018)
Prunus cerasifera	cherry plum, myrobalan plum		EN			*multiplex*							EC (2018), EFSA (2018)
Prunus cerasus	morello cherry, sour cherry, tart cherry, dwarf cherry		N										EC (2018), EFSA (2018)
Prunus domestica	common plum		N	EN			*multiplex*					plum leaf scald (PLS)	EFSA (2018)
Prunus dulcis	almond	EN	EN	EN	E	*multiplex, pauca*	*fastidiosa, multiplex, pauca*		*pauca*			almond leaf scorch (ALS)	EC (2018), EFSA (2018)
Prunus persica x P. Webbii		E	E										EFSA (2018)
*Prunus persica**	peach	N	N	EN		*pauca*						phony peach disease (PPD)	EFSA (2018)
Prunus salicina				E									EFSA (2018)
Prunus sp.		E	EN										EFSA (2018)
Prunus x amygdalo-persica				E									EFSA (2018)
Quercus coccinea			N										EFSA (2018)
Quercus falcata			N										EFSA (2018)
*Quercus ilex**	holm oak			EN		*pauca*							EFSA (2018)
Quercus laevis			N										EFSA (2018)
Quercus macrocarpa			N										EFSA (2018)
Quercus nigra			N										EFSA (2018)
Quercus palustris			N										EFSA (2018)
Quercus phellos			N										EFSA (2018)
Quercus pubescens				E									EFSA (2018)
Quercus robur			N										EFSA (2018)
Quercus rubra			N										EFSA (2018)
Quercus shumardii			N										EFSA (2018)

Hosts	Common names	*fastidiosa*	*multiplex*	*pauca*	*sandyi*	France	Spain	Germany	Italy	Belgium	Portugal	Disease	Reference
Quercus sp.			N										EFSA (2018)
Quercus suber	cork oak		N			*multiplex*							EC (2018), EFSA (2018)
Ratibida columnifera			N										EFSA (2018)
Rhamnus alaternus	Italian buckthorn, Mediterranean buckthorn	N	N	N			*fastidiosa, multiplex*		*pauca*				EC (2018), EFSA (2018)
Rosa floribunda	dog rose		N										DEFRA (2016)
Rosa canina			N			*multiplex*							EC (2018), EFSA (2018)
Rosa hybrids													EPPO (n.d.)
Rosa multiflora													EPPO (n.d.)
Rosa sp.			N										EFSA (2018)
Rosmarinus officinalis	rosemary	N	N	N		*multiplex*	*multiplex*	*fastidiosa*	*pauca*				EC (2018), EFSA (2018)
Rubus sp.			N										EFSA (2018)
Rubus ursinus		E	E										EFSA (2018)
Rumex crispus		E											EFSA (2018)
Salvia mellifera			N										EFSA (2018)
Sambucus canadensis		N											EFSA (2018)
Sambucus sp.		N	N										EFSA (2018)
Sapindus saponaria			N										EFSA (2018)
Simmondsia chinensis		E											EFSA (2018)
Solanum lycopersicum		E											EFSA (2018)
Solanum melongena		E											EFSA (2018)
Solidago virgaurea			N										EFSA (2018)
Sonchus oleraceus		E											EFSA (2018)
Sorghum halepense		E											EFSA (2018)
Spartium junceum	Spanish broom weaver's broom	N	N	N		*multiplex*			*pauca*				EC (2018), EFSA (2018)
Spartium sp.			N			*multiplex*							EFSA (2018)
Streptocarpus hybrids		N											EFSA (2018)
Streptocarpus sp.	Cape primrose	U						*fastidiosa*					EC (2018)
Teucrium capitatum	cat-thyme germander, felty germander	U	U	U									EC (2018)
Ulmus americana			N										EFSA (2018)
Ulmus crassifolia		N											EFSA (2018)
Vaccinium corymbosum		E	EN										EFSA (2018)
Vaccinium corymbosum x V. angustifolium hybrid			E										EFSA (2018)
Vaccinium sp.		E	EN										EFSA (2018)
Vaccinium virgatum													EPPO (n.d.)
Veronica elliptica	shore hebe, speedwell	U	U	U									EC (2018)

Hosts	Common names	*fastidiosa*	*multiplex*	*pauca*	*sandyi*	France	Spain	Germany	Italy	Belgium	Portugal	Disease	Reference
Vicia faba		E											EFSA (2018)
Vicia sativa		E											EFSA (2018)
Vinca major					E								EFSA (2018)
Vinca minor				N					*pauca*				EFSA (2018)
Vinca sp.	periwinkle		N	N					*pauca*				EC (2018), EFSA (2018)
Vitis aestivalis		N											EFSA (2018)
Vitis aestivalis hybrid		N											EFSA (2018)
Vitis candicans		N											EFSA (2018)
Vitis cinerea var. helleri x. V. vulpina		N											EFSA (2018)
Vitis girdiana		N											EFSA (2018)
Vitis labrusca													EPPO (n.d.)
Vitis rotundifolia		N											EFSA (2018)
Vitis sp.		N					*fastidiosa*						EFSA (2018)
Vitis vinifera	common grape vine	EN	E	E			*fastidiosa*					Pierce's disease (PD)	EC (2018), EFSA (2018)
Westringia fruticosa	coastal/Australian rosemary		N	N		*multiplex*			*pauca*				EC (2018), EFSA (2018)
Westringia glabra	violet westringia			N					*pauca*				EC (2018), EFSA (2018)
Xanthium strumarium		E	N										EFSA (2018)

Table H: Complete list of genomes used in the project. 55 Xylella fastidiosa genomes, one *Xylella taiwanensis* and two *Xanthomonas* genomes were obtained from NCBI's GenBank. Details on the genome size, sequencing information and origin are listed below. Information highlighted in **orange** were not described on the database and have been procured by sequence similarity search on NCBI's BLAST and the PubMLST database.

Strain	*Subspecies*	MLST	Accession	Size (Mb)	Assembly accession	Date added	Last updated	Submitted by	Assembly level	Assembly method	Coverage	Sequencing technology	Plasmids
32	*pauca*	16	AWYH01000001.1	2.60755	GCA_000506405.1	11/12/2013	02/04/2017	Universidade de Mogi das Cruzes	Contig	GS de novo Assembler v. 2.5.3	70x	454	NA
3124	*pauca*	16	CP009829.1	2.74859	GCA_001456195.1	03/12/2015	28/06/2017	Universidade de Sao Paulo	Complete Genome	Newbler v. 2.3; CROSSMATCH	267x	454 GS FLX Titanium	NA
11399	*pauca*	11	CM004499.1	2.73606	GCA_001684415.1	13/07/2016	11/04/2017	IAC - Centro de citricultura	Contig	CLC NGS Cell v. 6.0	70.0x	Illumina HiSeq	pXF51
6c	*pauca*	14	CM007617.1	2.60398	GCA_000506905.2	11/12/2013	06/04/2017	Universidade de Mogi das Cruzes	Contig	Bowtie2 v. 2.2.9	900x	Illumina MiSeq	pXF6c
9a5c	*pauca*	13	AE003849.1	2.73175	GCA_000006725.1	02/06/2000	29/03/2017	Sao Paulo state (Brazil) Consortium	Complete Genome	NA	NA	NA	pXF1.3, pXF51
Ann-1	*sandyi*	5	AAAM04000275.1	2.78091	GCA_000698805.1	06/06/2014	02/04/2017	University of California (LANL Genome Science Group)	Complete Genome	Velvet v. 1.0.13	22.3X	454; Illumina	unnamed1
Ann-1	*sandyi*	5	CP006696.1	2.51152	GCA_001886315.1	25/11/2016	05/04/2017	USDA-ARS	Scaffold	CLC Genomics Workbench v. 7.5	1271.0x	Illumina MiSeq	NA
ATCC 35871	*multiplex*	41	KE386775.1	2.41626	GCA_000428665.1	15/07/2013	01/04/2017	DOE Joint Genome Institute	Scaffold	NA	NA	Illumina HiSeq 2000	NA
ATCC 35879	*fastidiosa*	2	JQAP01000001.1	2.52233	GCA_000767565.1	21/10/2014	02/04/2017	Crop Diseases, Pests, Genetics Research Unit, San Joaquin Valley Agricultural Sciences Center, USDA	Contig	CLC Genomic Workbench v. 7.0.3	1380.0x	Illumina MiSeq	NA
BB01	*multiplex*	42	MPAZ01000045.1	2.72975	GCA_000166855.2	10/07/2002	11/04/2017	DOE Joint Genome Institute	Contig	ALLPATHS v. R37654	NA	Sanger	NA
CFBP7969	*fastidiosa*	2	PHFQ01000001.1	2.43128	GCA_004016275.1	14/01/2019	26/01/2019	INRA	Contig	Velvet v. 1.2.07; SOAPdenovo v. 2.04	900.0x	Illumina MiSeq	NA
CFBP7970	*fastidiosa*	2	PHFR01000001.1	2.48833	GCA_004016315.1	14/01/2019	26/01/2019	INRA	Contig	Velvet v. 1.2.07; SOAPdenovo v. 2.04	900.0x	Illumina MiSeq	NA
CFBP8071	*fastidiosa*	1	PHFP01000001.1	2.48429	GCA_004016295.1	14/01/2019	26/01/2019	INRA	Contig	Velvet v. 1.2.07; SOAPdenovo v. 2.05	900.0x	Illumina MiSeq	NA
CFBP8072	*pauca*	74	LKDK01000001.1	2.49666	GCA_001469345.1	18/12/2015	04/04/2017	INRA	Scaffold	Velvet v. 1.2.02	700.0x	Illumina HiSeq	NA
CFBP8073	*fastidiosa*	75	LKES01000001.1	2.58215	GCA_001469395.1	18/12/2015	04/04/2017	INRA	Scaffold	Velvet v. 1.2.02; SOAPdenovo v. 1.05	800.0x	Illumina HiSeq	NA
CFBP8078	*multiplex*	51	PHFS01000001.1	2.59655	GCA_004016365.1	14/01/2019	26/01/2019	INRA	Contig	Velvet v. 1.2.07; SOAPdenovo v. 2.04	1000.0x	Illumina MiSeq	NA
CFBP8082	*fastidiosa*	1	PHFT01000001.1	2.52668	GCA_004016375.1	14/01/2019	26/01/2019	INRA	Contig	Velvet v. 1.2.07; SOAPdenovo v. 2.04	900.0x	Illumina MiSeq	NA
CFBP8351	*fastidiosa*	1	PHFU01000001.1	2.47375	GCA_004016405.1	14/01/2019	26/01/2019	INRA	Contig	Velvet v. 1.2.07; SOAPdenovo v. 2.04	1000.0x	Illumina MiSeq	NA

CFBP8356	*sandyi*	76	PHFV01000001.1	2.53615	GCA_004016415.1	14/01/2019	26/01/2019	INRA	Scaffold	Velvet v. 1.2.07; SOAPdenovo v. 2.04	900.0x	Illumina MiSeq	NA
CFBP8416	*multiplex*	7	LUYC01000001.1	2.46675	GCA_001971475.1	25/01/2017	25/01/2017	INRA	Contig	Velvet v. 1.2.07; SOAPdenovo v. 2.04	125.0x	Illumina MiSeq	NA
CFBP8417	*multiplex*	6	LUYB01000001.1	2.50498	GCA_001971505.1	25/01/2017	06/04/2017	INRA	Contig	Velvet v. 1.2.07; SOAPdenovo v. 2.04	125.0x	Illumina MiSeq	NA
CFBP8418	*multiplex*	6	LUYA01000001.1	2.51397	GCA_001971465.1	25/01/2017	06/04/2017	INRA	Contig	Velvet v. 1.2.07; SOAPdenovo v. 2.04	125.0x	Illumina MiSeq	NA
CO33	*sandyi*	72	LJZW01000001.1	2.68193	GCA_001417925.1	28/10/2015	04/04/2017	National Research Council (C.N.R.), Institute for Sustainable Plant Protection	Contig	Velvet v. 1.2.8; SOAPdenovo v. 2.04; Edena v. 0.3; post-assembly SSPACE v. 1.0.7	310.0x	Illumina HiSeq	NA
CoDiRO	*pauca*	53	CM003178.1	2.54293	GCA_000811965.1	29/12/2014	03/04/2017	National Research Council (C.N.R.), Institute for Sustainable Plant Protection	Contig	Velvet v. 1.2.08; SOAPdenovo v. 2.04; EDENA v. 0.3; post-assembly SSPACE v. 1.0.7	345.0x	Illumina HiSeq	unnamed
COF0324	*pauca*	14	CM003758.1	2.77256	GCA_001549815.1	05/02/2016	04/04/2017	cBio Corp	Contig	Trimmomatic v. 0.32; SGA v. 0.10.13; iMetAMOS v. 1.5; samtools v. 1.1; FastQC v. 0.10.0; Spades v. 3.1.1; idba v. 1.1.1; Pilon v. 1.8; Quast v. 2.2; Prokka v. 1.7	736.432x	Illumina MiSeq	pXF-BHR-COF0324, pXF-P1.COF0324, pXF-PC_COF0324, pXF-RC.COF0324
COF0407	*pauca*	53	CM003762.1	2.53847	GCA_001549825.1	05/02/2016	04/04/2017	cBio Corp	Contig	Trimmomatic v. 0.32; SGA v. 0.10.13; iMetAMOS v. 1.5; samtools v. 1.1; FastQC v. 0.10.0; Spades v. 3.1.1; idba v. 1.1.1; Pilon v. 1.8; Quast v. 2.2; Prokka v. 1.7	612.211x	Illumina MiSeq	pXF-P1.OLS0479, pXF-P4.OLS0479, pXF-PS.OLS0479, pXF-RC.OLS0479
CVC0251	*pauca*	11	CM003754.1	2.74025	GCA_001549765.1	05/02/2016	04/04/2017	cBio Corp	Contig	Trimmomatic v. 0.32; SGA v. 0.10.13; iMetAMOS v. 1.5; samtools v. 1.1; FastQC v. 0.10.0; Spades v. 3.1.1; idba v. 1.1.1; Pilon v. 1.8; Quast v. 2.2; Prokka v. 1.7	944.475x	Illumina MiSeq	pXF-BHR.CVC0251, pXF-P1.CVC0251, pXF-P4.CVC0251, pXF-PS.CVC0251
CVC0256	*pauca*	11	CM003748.1	2.70214	GCA_001549745.1	05/02/2016	04/04/2017	cBio Corp	Contig	Trimmomatic v. 0.32; SGA v. 0.10.13; iMetAMOS v. 1.5; samtools v. 1.1; FastQC v. 0.10.0; Spades v. 3.1.1; idba v. 1.1.1; Pilon v. 1.8; Quast v. 2.2; Prokka v. 1.7	691.101x	Illumina MiSeq	pXF-BHR.CVC0256, pXF-P1.CVC0256, pXF-P4.CVC0256, pXF-PS.CVC0256

De Donno	*pauca*	53	CP020870.1	2.54374	GCA_002117875.1	04/05/2017	10/05/2017	POnTE (Pest Organisms Threatening Europe)	Complete Genome	SPAdes v. 3.9.0	636.0x	PacBio; Illumina HiSeq	pXF-De_Donno
Dixon	*multiplex*	6	AAAL02000032.1	2.62233	GCA_000166835.1	10/07/2002	30/03/2017	DOE Joint Genome Institute	Scaffold	NA	NA	NA	NA
DSM 10026	*fastidiosa*	2	FQWN01000063.1	2.43165	GCA_900129695.1	02/12/2016	06/04/2017	DOE Joint Genome Institute	Scaffold	NA	416x	NA	NA
EB92.1	*fastidiosa*	1	AFDJ01000168.1	2.47543	GCA_000219235.2	24/06/2011	22/11/2017	University of Florida	Contig	Newbler v. 2.3	194X	454 GS Titanium	NA
ESVL	*multiplex*	6	CM013391.1	2.5545	GCA_004023385.1	15/01/2019	18/01/2019	CNR	Contig	SPAdes v. 3.9	110.0x	Illumina HiSeq4000	pUCLA-ESVL, pXF64-Hb_ESVL
Fb7	*[pauca]*	[13]	CP010051.2	2.69932	GCA_001456335.3	03/12/2015	22/05/2018	Universidade de Sao Paulo	Complete Genome	NA	NA	NA	unnamed
GB514	*fastidiosa*	1	CP002165.1	2.51738	GCA_000148405.1	23/09/2010	11/04/2017	Research and Testing Laboratory	Complete Genome	NA	NA	NA	unnamed
Griffin-1	*multiplex*	7	AVGA01000001.1	2.38731	GCA_000466025.1	12/09/2013	11/04/2017	USDA	Contig	Newbler v. v2.6	30.0x	454	NA
Hib4	*pauca*	70	CP009885.1	2.87755	GCA_001456315.1	03/12/2015	28/06/2017	Universidade de Sao Paulo	Complete Genome	Newbler v. 2.3; CROSSMATCH	100x	454 GS FLX Titanium	pXF64-HB
IVIA5235	*fastidiosa*	1	CM010656.1	2.49157	GCA_003515915.1	10/09/2018	12/09/2018	Spanish National Research Council (CSIC), Institute for Sustainable Agriculture	Contig	SPAdes v. 3.9.0	450.0x	Illumina HiSeq 4000	pXFAS_5235
IVIA5901	*multiplex*	6	QPQW01000053.1	2.49356	GCA_004023395.1	15/01/2019	17/01/2019	Spanish National Research Council (CSIC), Institute for Sustainable Agriculture	Contig	SPAdes v. 3.9	309x	Ilumina HiSeq	NA
J1a12	*pauca*	11	CP009823.1	2.86724	GCA_001456235.1	03/12/2015	28/06/2017	Universidade de Sao Paulo	Complete Genome	Newbler v. 2.3; CROSSMATCH	65x	454 GS FLX Titanium	pXF27-J1, pXF51-J1
M12	*multiplex*	7	CP000941.1	2.47513	GCA_000019325.1	19/02/2008	30/03/2017	US DOE Joint Genome Institute	Complete Genome	NA	NA	NA	NA
M23	*fastidiosa*	1	CP001011.1	2.57399	GCA_000019765.1	11/04/2008	30/03/2017	US DOE Joint Genome Institute	Complete Genome	NA	NA	NA	pXFAS01
MUL0034	*morus*	30	CP006740.1	2.66658	GCA_000698825.1	06/06/2014	02/04/2017	University of California (LANL Genome Science Group)	Complete Genome	Newbler v. 2.3; VELVET v. 0.7.63	NA	454; Illumina	unnamed2
Mul-MD	*morus*	29	AXDP01000001.1	2.52055	GCA_000567985.1	10/02/2014	02/04/2017	FNPRU-USNA-ARS-USDA	Contig	Newbler v. 08-06-2012	5.0x	454	NA
OLS0478	*pauca*	53	CM003752.1	2.55541	GCA_001549755.1	05/02/2016	04/04/2017	cBio Corp	Contig	Trimmomatic v. 0.32; SGA v. 0.10.13; iMetAMOS v. 1.5; samtools v. 1.1; FastQC v. 0.10.0; Spades v. 3.1.1; idba v. 1.1.1; Pilon v. 1.8; Quast v. 2.2; Prokka v. 1.7	788.469x	Illumina MiSeq	pXF-P1.OLS0478, pXF-P4.OLS0478
OLS0479	*pauca*	53	CM003743.1	2.53996	GCA_001549735.1	05/02/2016	04/04/2017	cBio Corp	Contig	Trimmomatic v. 0.32; SGA v. 0.10.13; iMetAMOS v. 1.5; samtools v. 1.1; FastQC v.	844.258x	Illumina MiSeq	pXF-P1.COF0407, pXF-

										0.10.0; Spades v. 3.1.1; idba v. 1.1.1; Pilon v. 1.8; Quast v. 2.2; Prokka v. 1.7			P4.COF0407, pXF-PS.COF0407, pXF-RC.COF0407
Pr8x	*pauca*	14	CP009826.1	2.70582	GCA_001456295.1	03/12/2015	28/06/2017	Universidade de Sao Paulo	Complete Genome	Newbler v. 2.3; CROSSMATCH	63x	454 GS Titanium	pXF39
Salento-1	*[pauca]*	[53]	CP016608.1	2.54337	GCA_002954185.1	27/02/2018	04/03/2018	CNR	Complete Genome	HGAP v.2 + Circlator v. 1.2.1	402.7x	PacBio	pSal1
Salento-2	*[pauca]*	[53]	CP016610.1	2.54357	GCA_002954205.1	27/02/2018	04/03/2018	CNR	Complete Genome	HGAP v.2 + Circlator v. 1.2.1	349.25x	PacBio	pSal2
Stag's Leap	*fastidiosa*	1	LSMJ01000001.1	2.5108	GCA_001572105.1	24/02/2016	04/04/2017	USDA-ARS	Contig	Bowtie 2 v. 2.2.6	750.0x	Illumina MiSeq	NA
sycamore Sy-VA	*multiplex*	8	JMHP01000001.1	2.47588	GCA_000732705.1	22/07/2014	02/04/2017	Beltsville Agricultural Research Center	Contig	Newbler v. 2.7	70.0x	454	NA
Temecula1	*fastidiosa*	1	AE009442.1	2.52115	GCA_000007245.1	29/01/2003	29/03/2017	Sao Paulo state (Brazil) Consortium	Complete Genome	NA	NA	NA	pXFPD1.3
U24D	*pauca*	13	CP009790.1	2.73249	GCA_001456275.1	03/12/2015	28/06/2017	Universidade de Sao Paulo	Complete Genome	Newbler v. 2.3; CROSSMATCH	81x	454 GS FLX Titanium	pXF51ud
XYL1732/17	*fastidiosa*	1	QTJT01000001.1	2.444109	GCA_003973705.1	27/12/2018	04/01/2018	University of Balearic Islands	Contig	Newbler v. 2.9	102.0x	Illumina MiSeq	pXFAS01, pXFAS_5235
XYL2055/17	*fastidiosa*	1	QTJS01000001.1	2.45678	GCA_003973695.1	27/12/2018	04/01/2018	University of Balearic Islands	Contig	Newbler v. 2.9	151.0x	Ilumina HiSeq	pXFAS01, pXFAS_5235
Xylella taiwanensis strain PLS229	NA	NA	JDSQ01000001.1	2.82488	GCA_013177435.1	NA	NA	NA	Complete	NA	NA	NA	NA
Xanthomonas campestris pv. *campestris* strain ATCC 33913	NA	NA	NC_003902.1	5.08	GCA_000007145.1	28/11/2001	NA	NA	Complete	NA	NA	NA	NA
Xanthomonas oryzae pv. *oryzae* PXO99A	NA	NA	NC_010717.2	5.24	GCA_000019585.2	16/05/2008	NA	NA	Complete	NA	NA	NA	NA

Table I: Details of each strain displayed in the phylogenetic tree. This table lists details of the genomes from which a phylogenetic tree was created. Information includes GenBank accession number of each strain and location and host from which the isolate was obtained from.

Accession Number	Strain	Subspecies	Continent	Country	Detailed location	Host species	Common host name
AWYH01000001.1	32	*pauca*	South America	Brazil	Sao Paulo	*Coffea*	coffee
CP009829.1	3124	*pauca*	South America	Brazil	Matao, Sao Paulo	*Coffea*	coffee
CM004499.1	11399	*pauca*	South America	Brazil	NA	*Citrus x sinensis*	sweet orange
CM007617.1	6c	*fastidiosa*	South America	Brazil	Sao Paulo	*Coffea*	coffee plant
AE003849.1	9a5c	*pauca*	South America	Brazil	Macaubal, Sao Paulo	*Citrus × sinensis pummelo x mandarin orange*	Valencia sweet orange
AAAM04000275.1	Ann-1	*sandyi*	North America	USA	Palm Springs, California	*Nerium oleander*	oleander
CP006696.1	Ann-1	*sandyi*	North America	USA	NA	*Nerium oleander*	oleander
KE386775.1	ATCC 35871	*multiplex*	North America	USA	Georgia	*Prunus*	hybrid plum
JQAP01000001.1	ATCC 35879	*fastidiosa*	North America	USA	Florida	*Vitis vinifera*	grapevine
MPAZ01000045.1	BB01	*multiplex*	North America	USA	Georgia	*Vaccinium corymbosum*	blueberry
PHFP01000001.1	CFBP7969	*fastidiosa*	North America	USA	North Carolina	*Vitis rotundifolia* cv Carlos	grapevine
PHFR01000001.1	CFBP7970	*fastidiosa*	North America	USA	Florida	Vitis vinifera	grapevine
PHFP01000001.1	CFBP8071	*fastidiosa*	North America	USA	California	*Prunus dulcis*	almond
LKDK01000001.1	CFBP8072	*pauca*	South America	Ecuador	imported from Ecuador to France	*Coffea arabica*	Arabica coffee
LKES01000001.1	CFBP8073	*fastidiosa*	Europe	France	NA	*Coffea canephora*	Robusta coffee
PHFS01000001.1	CFBP8078	*multiplex*	North America	USA	Florida	*Vinca* sp.	periwinkle
PHFT00000000.1	CFBP8082	*fastidiosa*	North America	USA	Florida	*Ambrosia artenisiifolia*	common ragweed
PHFU01000001.1	CFBP8351	*fastidiosa*	North America	USA	California	*Vitis vinifera L*	grapevine
PHFV01000001.1	CFBP8356	*sandyi*	North America	Costa Rica	(intercepted in France)	*Coffea arabica*	coffee
LUYC01000001.1	CFBP8416	*multiplex*	Europe	France	Propriano, Corse	*Polygala myrtifolia*	myrtle-leaf milkwort
LUYB01000001.1	CFBP8417	*multiplex*	Europe	France	Alata, Corse	*Spartium junceum*	Spanish broom
LUYA01000001.1	CFBP8418	*multiplex*	Europe	France	Alata, Corse	*Spartium junceum*	Spanish broom
LJZW01000001.1	CO33	*sandyi*	North America	Costa Rica	imported from Costa Rica through Netherlands and to northern Italy	*Coffea*	coffee plant
CM003178.1	CoDiRo	*pauca*	Europe	Italy	Apulia	*Olea europaea*	common olive
CM003758.1	COF0324	*pauca*	South America	Brazil	Varginha, Minas Gerais	*Coffea*	coffee
CM003762.1	COF0407	*pauca*	North America	Costa Rica	Curridabat, San Jose	*Coffea*	coffee
CM003754.1	CVC0251	*pauca*	South America	Brazil	Bebedouro, Sao Paulo	*Citrus x sinensis*	sweet orange
CM003748.1	CVC0256	*pauca*	South America	Brazil	Colina, Sao Paulo	*Citrus x sinensis*	sweet orange
CP020870.1	De Donno	*pauca*	Europe	Italy	Apulia	*Olea europaea*	common olive
AAAL02000032.1	Dixon	*multiplex*	NA	NA	NA	*Prunus dulcis*	almond
FQWN01000063.1	DSM 10026	*fastidiosa*	NA	NA	NA	*NA*	NA
AFDJ01000168.1	EB92.1	*fastidiosa*	North America	USA	Leesburg	*Sambucus canadensis*	common elderberry
CM013391.1	ESVL	*multiplex*	Europe	Spain	Benimantell, Alicante	*Prunus dulcis*	almond
CP010051.2	Fb7	*pauca*	South America	Argentina	Corrientes	*Citrus*	citrus
CP002165.1	GB514	*fastidiosa*	North America	USA	Texas	*Vitis vinifera*	grapevine
AVGA01000001.1	Griffin-1	*multiplex*	North America	USA	Griffin, Georgia	*Quercus rubra*	red oak tree
CP009885.1	Hib4	*pauca*	South America	Brazil	Jarinu, Sao Paulo	*Hibiscus*	hibiscus
CM010656.1	IVIA5235	*fastidiosa*	Europe	Spain	Mallorca Island	*Prunus avium*	sweet cherry
QPQW01000053.1	IVIA5901	*multiplex*	Europe	Spain	Bolulla, Alicante	*Prunus dulcis*	almond
CP009823.1	J1a12	*pauca*	South America	Brazil	Jales, Sao Paulo	*Citrus*	citrus

Accession Number	Strain	Subspecies	Continent	Country	Detailed location	Host species	Common host name
CP000941.1	M12	*multiplex*	North America	USA	San Joaquin Valley, California	*Prunus dulcis*	almond
CP001011.1	M23	*fastidiosa*	North America	USA	San Joaquin Valley, California	*Prunus dulcis*	almond
CP006740.1	MUL0034	*morus*	North America	USA	NA	*Morus*	mulberry
AXDP01000001.1	Mul-MD	*morus*	North America	USA	Beltsville, Maryland	*Morus*	mulberry
CM003752.1	OLS0478	*pauca*	North America	Costa Rica	Sabanilla, San Jose Province	*Nerium oleander*	oleander
CM003743.1	OLS0479	*pauca*	North America	Costa Rica	Sabanilla, San Jose Province	*Nerium oleander*	oleander
CP009826.1	Pr8x	*pauca*	South America	Brazil	Jarinu, Sao Paulo	*Prunus*	plum
CP016608.1	Salento-1	*pauca*	Europe	Italy	Taviano, Lecce, Apulia	*Olea europaea*	common olive
CP016610.1	Salento-2	*pauca*	Europe	Italy	Ugento, Lecce, Apulia	*Olea europaea*	common olive
LSMJ01000001.1	Stag's Leap	*fastidiosa*	North America	USA	Napa Valley, California	*Vitis vinifera*	grapevine
JMHP01000001.1	sycamore Sy-VA	*multiplex*	North America	USA	Virginia	*Acer pseudoplatanus*	sycamore tree
AE009442.1	Temecula1	*fastidiosa*	North America	USA	Temecula, California	*Vitis vinifera*	grapevine
CP009790.1	U24D	*pauca*	South America	Brazil	Ubarana, Sao Paulo	*Citrus x sinensis*	sweet orange
QTJT01000001.1	XYL1732	*fastidiosa*	Europe	Spain	Manacor, Mallorca	*Vitis vinifera*	grapevine (white grape cultivar Paradella)
QTJS01000001.1	XYL2055	*fastidiosa*	Europe	Spain	Manacor, Mallorca	*Vitis vinifera*	grapevine
JDSQ01000001.1	*Xylella taiwanensis*	NA	Asia	Taiwan	NA	*Pyrus* L.	pear
NC_003902.1	*Xanthomonas campestris*	NA	NA	NA	NA	NA	NA
NC_010717.2	*Xanthomonas oryzae*	NA	NA	NA	NA	NA	NA

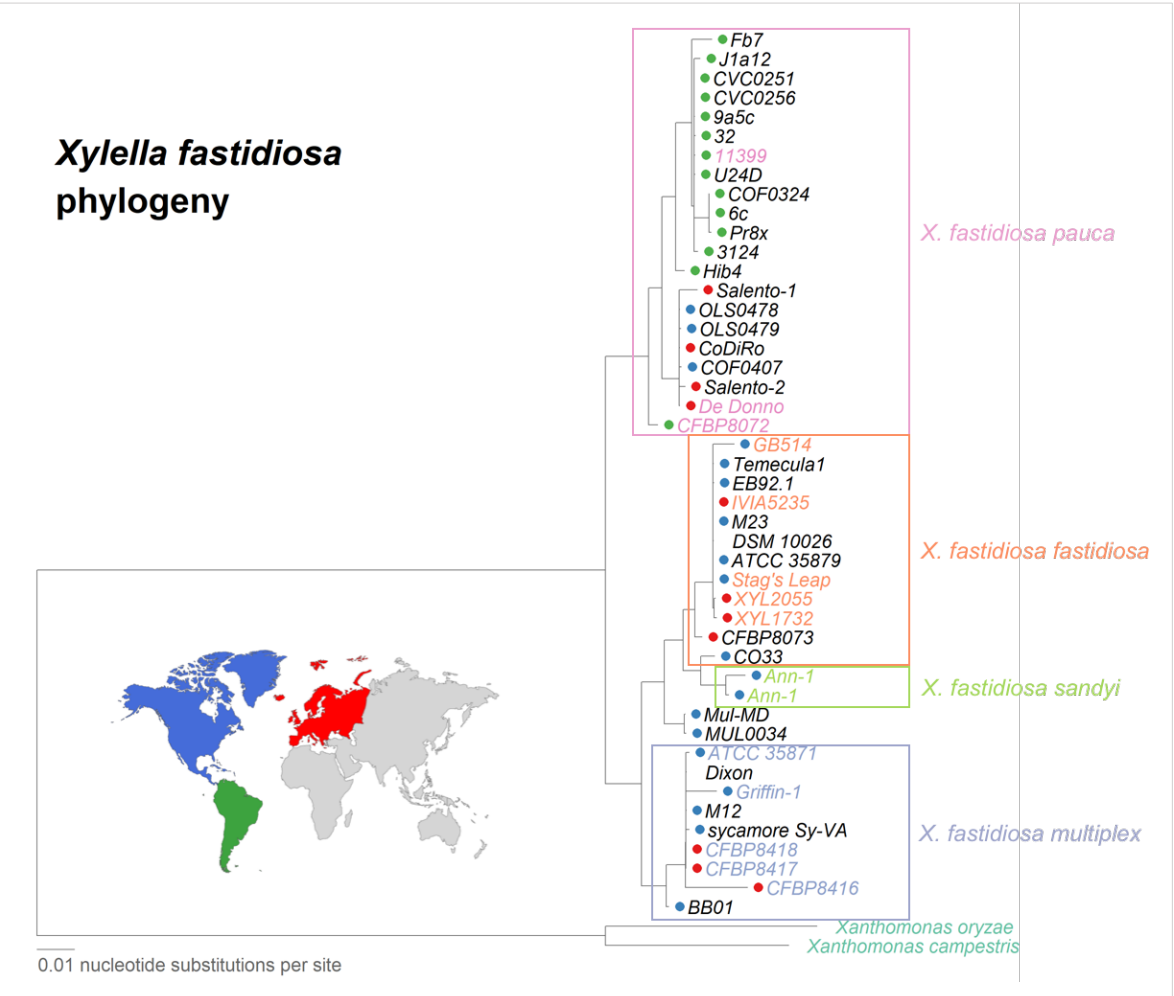


Figure S: First draft of a *Xylella fastidiosa* (*Xf*) phylogeny. A phylogenetic tree of 46 *Xf* and two *Xanthomonas* genomes (outgroups) was created. This tree was generated using FastTree's multiple sequence alignment by maximum-likelihood. The tree was visualised using the 'ape' package on R. Location of origin is highlighted by coloured circles corresponding to continents in the world map at the bottom left. The subspecies of strains with coloured fonts were confirmed by previous research. The subspecies of *Xf* strains Mul-MD and MUL0034 are less clear.

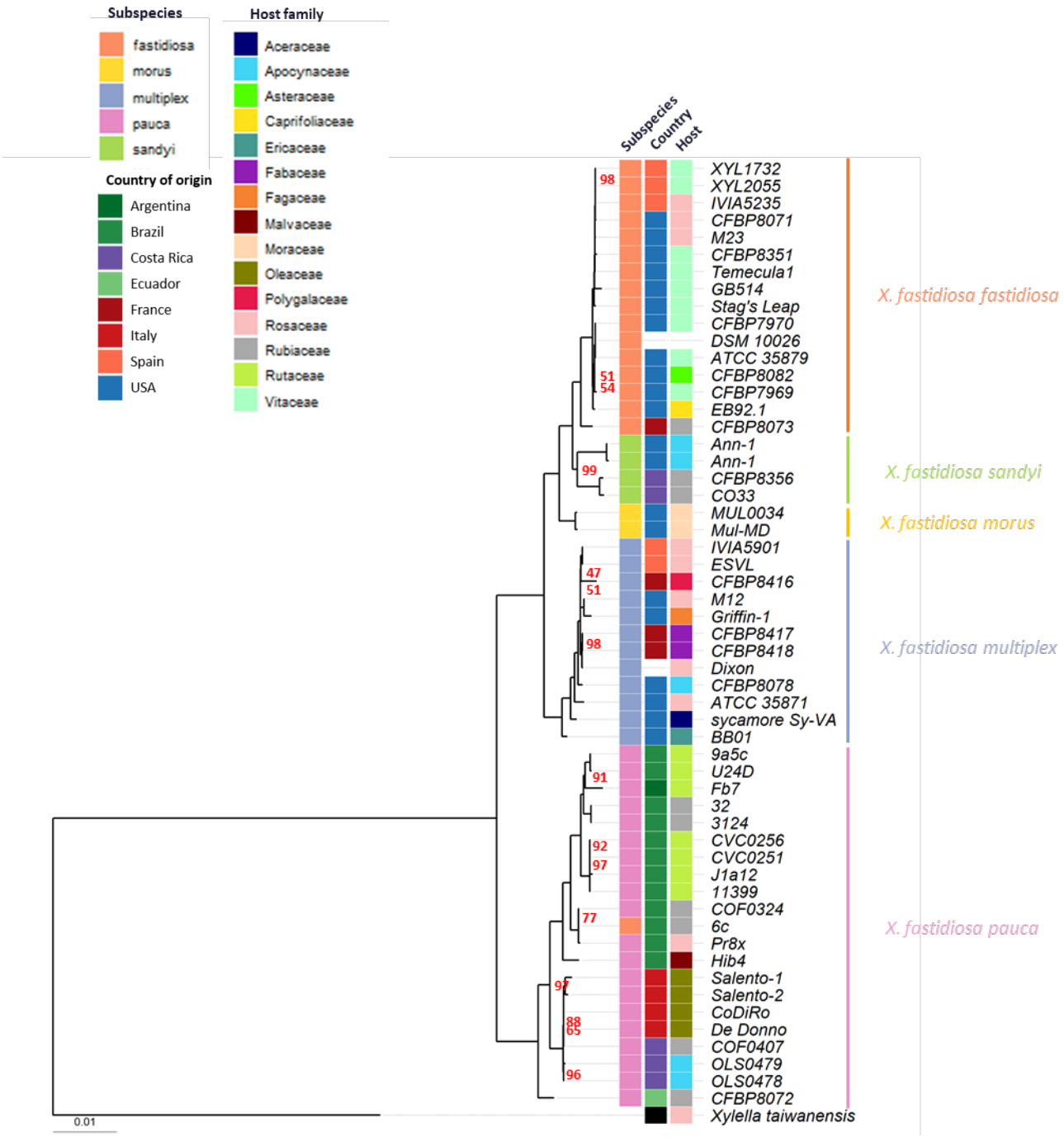


Figure U: Phylogenetic tree of *Xylella fastidiosa* (*Xf*). A phylogenetic tree of 55 *Xf* and the *Xylella taiwanensis* genome (outgroup) was created. This tree was generated using IQ-Tree's multiple sequence alignment by maximum-likelihood (Nguyen, et al., 2015). Bootstrap values below 100 are indicated in **red**. The tree was visualised using the *ape* package on R (Paradis, Claude and Strimmer, 2004). Location and host origin (where available), and subspecies information are highlighted in different colours.

Table J: Full list of PREFFECTOR results. This table includes all results of the PREFFECTOR program which was used to determine putative *Xylella fastidiosa* effectors across 55 genomes.

Please visit

https://github.com/mirloupa/Xf/blob/5371dd15716d862b83b93478371f7462d3da017a/Appendix/preffector_results_all.tsv to download the table.

Table K: Function of *Xylella fastidiosa* (*Xf*) orthologous sequences. Orthologous groups, or orthogroups, of putative *Xf* effectors were determined using the PREFFECTOR program created by Dhroso, Eidson and Korkin (2018). The function of each putative effector is listed below.

Orthogroup	Strain	Accession ID	Protein
X1	3124	ALQ97392.1	1,4-beta-cellobiosidase
X1	11399	OCA57933.1	1,4-beta-cellobiosidase
X1	6c	OJZ70903.1	1,4-beta-cellobiosidase
X1	9a5c	WP_010893338.1	endoglucanase
X1	9a5c	WP_010893773.1	1,4-beta-cellobiosidase
X1	Ann-1c	WP_024748856.1	1,4-beta-cellobiosidase partial
X1	ATCC_35879	KGM20724.1	1,4-beta-cellobiosidase partial
X1	CFBP7969	WP_128723174.1	1,4-beta-cellobiosidase
X1	CFBP7970	WP_128712456.1	1,4-beta-cellobiosidase
X1	CFBP8071	WP_128712519.1	1,4-beta-cellobiosidase
X1	CFBP8072	WP_058569679.1	1,4-beta-cellobiosidase
X1	CFBP8078	WP_128723671.1	4-beta-cellobiosidase
X1	CFBP8351	WP_128712519.1	1,4-beta-cellobiosidase
X1	CFBP8356	WP_128734966.1	1,4-beta-cellobiosidase
X1	CFBP8416	OMJ97057.1	1,4-beta-cellobiosidase
X1	CFBP8417	OMK00128.1	1,4-beta-cellobiosidase
X1	CFBP8418	OMJ99939.1	1,4-beta-cellobiosidase
X1	COF0324	KXB21420.1	1,4-beta-cellobiosidase
X1	CVC0251	KXB21968.1	1,4-beta-cellobiosidase
X1	CVC0256	KXB13296.1	1,4-beta-cellobiosidase
X1	Dixon	EAO14376.1	Cellulase
X1	DSM_10026	SHG20270.1	Cellulose binding domain-containing protein partial
X1	EB92.1	EGO81204.1	Cellobiohydrolase A (1,4-beta-cellobiosidase A) partial
X1	EB92.1	EGO81385.1	Endoglucanase BglC partial
X1	EB92.1	EGO82960.1	Cellobiohydrolase A (1,4-beta-cellobiosidase A) partial
X1	ESVL	WP_128382978.1	1,4-beta-cellobiosidase
X1	Fb7	AWG45316.1	1,4-beta-cellobiosidase
X1	Griffin-1	ERI59813.1	1,4-beta-cellobiosidase
X1	Hib4	ALR07014.1	1,4-beta-cellobiosidase
X1	IVIA5235	RHW37904.1	1,4-beta-cellobiosidase
X1	IVIA5901	WP_128283863.1	hypothetical protein
X1	J1a12	ALR01763.1	1,4-beta-cellobiosidase
X1	M12	ACA11602.1	Cellulase
X1	M23	ACB91997.1	cellulose 1,4-beta-cellobiosidase
X1	Mul-MD	EWG14499.1	cellulase
X1	MUL0034	AIC13557.1	hypothetical protein P303_02185
X1	Pr8x	ALR04597.1	1,4-beta-cellobiosidase
X1	Stags_Leap	WP_081095287.1	1,4-beta-cellobiosidase
X1	Temecula1	AAO28402.1	cellulose 1,4-beta-cellobiosidase
X1	U24D	ALQ94365.1	endoglucanase
X1	U24D	ALQ94677.1	1,4-beta-cellobiosidase
X1	XYL1732	RUA39812.1	1,4-beta-cellobiosidase
X1	XYL2055	RUA38669.1	1,4-beta-cellobiosidase
X1	*Xylella taiwanensis*	WP_069636213.1	1,4-beta-cellobiosidase
X2	32	ETE34180.1	cold-shock protein
X2	3124	ALQ96546.1	cold-shock protein
X2	11399	OCA57322.1	cold-shock protein

Orthogroup	Strain	Accession ID	Protein
X2	9a5c	WP_010894798.1	cold-shock protein
X2	Ann-1c	WP_004085832.1	cold-shock protein
X2	Ann-1f	AIC10508.1	cold-shock protein
X2	ATCC_35871	WP_004085832.1	cold-shock protein
X2	BB01	WP_004085832.1	cold-shock protein
X2	CFBP7969	WP_004085832.1	cold-shock protein
X2	CFBP7970	WP_004085832.1	cold-shock protein
X2	CFBP8071	WP_004085832.1	cold-shock protein
X2	CFBP8072	WP_010894798.1	cold-shock protein
X2	CFBP8073	WP_004085832.1	cold-shock protein
X2	CFBP8078	WP_004085832.1	cold-shock protein
X2	CFBP8082	WP_004085832.1	cold-shock protein
X2	CFBP8351	WP_004085832.1	cold-shock protein
X2	CFBP8356	WP_004085832.1	cold-shock protein
X2	CFBP8416	OMJ96975.1	cold-shock protein
X2	CoDiRo	KIA57572.1	cold-shock protein
X2	De_Donno	ARO68197.1	cold-shock protein
X2	Dixon	EAO12779.1	Cold-shock protein DNA-binding
X2	DSM_10026	SHG79508.1	cold-shock DNA-binding protein family
X2	EB92.1	EGO81051.1	Cold shock protein
X2	ESVL	WP_004085832.1	cold-shock protein
X2	GB514	ADN62224.1	cold shock protein
X2	Griffin-1	ERI60141.1	1,4-beta-cellobiosidase
X2	Hib4	ALR06040.1	cold-shock protein
X2	IVIA5235	RHW42932.1	cold-shock protein
X2	IVIA5901	WP_004085832.1	cold-shock protein
X2	J1a12	ALR01430.1	cold-shock protein
X2	M12	ACA12436.1	putative cold-shock DNA-binding domain protein
X2	M23	ACB92876.1	cold-shock DNA-binding domain protein
X2	MUL0034	AIC12651.1	cold-shock protein
X2	Pr8x	ALR03814.1	cold-shock protein
X2	Stags_Leap	WP_004085832.1	cold-shock protein
X2	Temecula1	AAO29227.1	cold shock protein
X2	U24D	ALQ95444.1	cold-shock protein
X2	XYL1732	RUA38378.1	cold-shock protein
X2	XYL2055	RUA37809.1	cold-shock protein
X2	*Xylella taiwanensis*	WP_038270170.1	cold-shock protein
X3	Ann-1c	WP_024748838.1	hypothetical protein
X3	Ann-1f	AIC09613.1	hypothetical protein D934_03670
X3	ATCC_35871	WP_027700566.1	hypothetical protein
X3	ATCC_35879	KGM21287.1	hypothetical protein JT24_02740
X3	BB01	WP_071869870.1	hypothetical protein
X3	CFBP7969	WP_004087222.1	hypothetical protein
X3	CFBP7970	WP_004087222.1	hypothetical protein
X3	CFBP8071	WP_004087222.1	hypothetical protein
X3	CFBP8073	WP_004087222.1	hypothetical protein
X3	CFBP8078	WP_027700566.1	hypothetical protein
X3	CFBP8082	WP_004087222.1	hypothetical protein
X3	CFBP8351	WP_004087222.1	hypothetical protein
X3	CFBP8356	WP_057683677.1	hypothetical protein
X3	CFBP8416	OMJ96558.1	hypothetical protein XYFPCFBP8416_09735

Orthogroup	Strain	Accession ID	Protein
X3	CFBP8417	OMJ98647.1	hypothetical protein XYFPCFBP8417_09565
X3	CFBP8418	OMJ98807.1	hypothetical protein XYFPCFBP8418_09615
X3	CO33	KQH72984.1	hypothetical protein AOT81_10970
X3	Dixon	EAO14407.1	conserved hypothetical protein
X3	DSM_10026	SHG25783.1	hypothetical protein SAMN05660380_00228
X3	EB92.1	EGO83061.1	hypothetical protein XFEB_00064
X3	ESVL	WP_012337698.1	hypothetical protein
X3	GB514	ADN63494.1	hypothetical protein XFLM_07930
X3	Griffin-1	ERI59435.1	hypothetical protein M233_09605
X3	IVIA5235	RHW44452.1	hypothetical protein D1605_01650
X3	IVIA5901	WP_012337698.1	hypothetical protein
X3	M12	ACA11567.1	conserved hypothetical protein
X3	M23	ACB91945.1	conserved hypothetical protein
X3	Mul-MD	EWG14466.1	hypothetical protein P910_002160
X3	MUL0034	AIC12056.1	hypothetical protein P303_02010
X3	Stags_Leap	WP_004087222.1	hypothetical protein
X3	sycamore_Sy-VA	KFA42019.1	hypothetical protein DF22_001,462
X3	Temecula1	AAO28378.1	conserved hypothetical protein
X3	XYL1732	RUA35871.1	hypothetical protein DX878_09475
X3	XYL2055	RUA35841.1	hypothetical protein DX877_09690
X4	9a5c	WP_075584605.1	hypothetical protein
X4	ATCC_35871	WP_012337575.1	hypothetical protein
X4	BB01	WP_012337575.1	hypothetical protein
X4	CFBP7969	WP_004572979.1	hypothetical protein
X4	CFBP7970	WP_004572979.1	hypothetical protein
X4	CFBP8071	WP_004572979.1	hypothetical protein
X4	CFBP8072	WP_081044378.1	hypothetical protein
X4	CFBP8073	WP_081046799.1	hypothetical protein
X4	CFBP8078	WP_012337575.1	hypothetical protein
X4	CFBP8082	WP_004572979.1	hypothetical protein
X4	CFBP8351	WP_004572979.1	hypothetical protein
X4	De_Donno	ARO67822.1	hypothetical protein B9J09_00915
X4	Dixon	EAO12591.1	conserved hypothetical protein
X4	EB92.1	EGO80923.1	hypothetical protein XFEB_02287
X4	ESVL	WP_128283928.1	hypothetical protein
X4	IVIA5235	RHW41807.1	hypothetical protein D1605_05685
X4	IVIA5901	WP_128283928.1	hypothetical protein
X4	M12	ACA11214.1	conserved hypothetical protein
X4	Stags_Leap	WP_004572979.1	hypothetical protein
X4	Temecula1	AAO28063.1	conserved hypothetical protein
X4	XYL1732	RUA38802.1	hypothetical protein DX878_03020
X4	XYL2055	RUA38484.1	hypothetical protein DX877_03160
X5	32	ETE31348.1	hypothetical protein B398_08190
X5	3124	ALQ97347.1	hypothetical protein XFC3_08185
X5	11399	OCA57974.1	hypothetical protein AA93_05575
X5	6c	OJZ70856.1	hypothetical protein B375_0207520
X5	9a5c	WP_010893841.1	hypothetical protein
X5	CFBP8072	WP_046419372.1	hypothetical protein
X5	CoDiRo	KIA58106.1	hypothetical protein RA12_06135
X5	COF0324	KXB21468.1	hypothetical protein ADT30_03735
X5	COF0407	KXB13848.1	hypothetical protein ADT33_08040

Orthogroup	Strain	Accession ID	Protein
X5	CVC0251	KXB22011.1	hypothetical protein ADT28_03850
X5	CVC0256	KXB13340.1	hypothetical protein ADT29_08520
X5	De_Donno	ARO68989.1	hypothetical protein B9J09_08050
X5	Fb7	ALR09115.2	hypothetical protein XFFB_07815
X5	Hib4	ALR06968.1	hypothetical protein XFHB_09060
X5	J1a12	ALR01722.1	hypothetical protein OY18_05120
X5	OLS0478	KXB10068.1	hypothetical protein ADT32_10160
X5	OLS0479	KXB17166.1	hypothetical protein ADT31_05050
X5	Pr8x	ALR04553.1	hypothetical protein XFPR_07925
X5	Salento-1	AVI21046.1	hypothetical protein BCV75_07530
X5	Salento-2	AVI23070.1	hypothetical protein BC375_07590
X5	U24D	ALQ94720.1	hypothetical protein XFUD_05630
X6	32	ETE31353.1	hypothetical protein B398_08340
X6	3124	ALQ97369.1	hypothetical protein XFC3_08325
X6	BB01	WP_071869871.1	hypothetical protein
X6	CFBP7969	WP_004090581.1	hypothetical protein
X6	CFBP7970	WP_004090581.1	hypothetical protein
X6	CFBP8071	WP_004090581.1	hypothetical protein
X6	CFBP8072	WP_081044462.1	hypothetical protein
X6	CFBP8078	WP_071869871.1	hypothetical protein
X6	CFBP8082	WP_004090581.1	hypothetical protein
X6	CFBP8351	WP_004090581.1	hypothetical protein
X6	CVC0256	KXB13319.1	hypothetical protein ADT29_08385
X6	DSM_10026	SHG27538.1	hypothetical protein SAMN05660380_00291
X6	EB92.1	EGO81235.1	hypothetical protein XFEB_01923
X6	Fb7	ALR09137.2	hypothetical protein XFFB_07955
X6	IVIA5235	RHW37878.1	hypothetical protein D1605_09870
X6	J1a12	ALR01741.1	hypothetical protein OY18_05260
X6	M23	ACB92027.1	hypothetical protein XfasM23_0583
X6	Stags_Leap	WP_004090581.1	hypothetical protein
X6	XYL1732	RUA39787.1	hypothetical protein DX878_01000
X6	XYL2055	RUA38644.1	hypothetical protein DX877_02555
X7	Ann-1c	WP_080702522.1	hypothetical protein
X7	ATCC_35871	WP_080654466.1	hypothetical protein
X7	BB01	WP_071869592.1	hypothetical protein
X7	CFBP7969	WP_100206152.1	hypothetical protein
X7	CFBP7970	WP_100206152.1	hypothetical protein
X7	CFBP8071	WP_100206152.1	hypothetical protein
X7	CFBP8072	WP_081044396.1	hypothetical protein
X7	CFBP8073	WP_081033415.1	hypothetical protein
X7	CFBP8078	WP_128723656.1	hypothetical protein
X7	CFBP8082	WP_100206152.1	hypothetical protein
X7	CFBP8351	WP_100206152.1	hypothetical protein
X7	CFBP8356	WP_081033415.1	hypothetical protein
X7	Dixon	EAO13830.1	conserved hypothetical protein
X7	ESVL	WP_076613215.1	hypothetical protein
X7	IVIA5235	RHW37159.1	hypothetical protein D1605_10345
X7	IVIA5901	WP_128283835.1	hypothetical protein
X7	M12	ACA12860.1	conserved hypothetical protein
X7	XYL1732	RUA36100.1	hypothetical protein DX878_09055
X7	XYL2055	RUA36072.1	hypothetical protein DX877_09165

Orthogroup	Strain	Accession ID	Protein
X8	Ann-1c	WP_042836348.1	hypothetical protein
X8	Ann-1f	AIC10957.1	hypothetical protein D934_01015
X8	ATCC_35879	KGM21314.1	hypothetical protein JT24_00110
X8	CFBP7969	WP_012382406.1	hypothetical protein
X8	CFBP7970	WP_012382406.1	hypothetical protein
X8	CFBP8071	WP_012382406.1	hypothetical protein
X8	CFBP8073	WP_012382406.1	hypothetical protein
X8	CFBP8082	WP_012382406.1	hypothetical protein
X8	CFBP8351	WP_012382406.1	hypothetical protein
X8	CFBP8356	WP_012382406.1	hypothetical protein
X8	CO33	KQH73083.1	hypothetical protein AOT81_10450
X8	DSM_10026	SHH02979.1	hypothetical protein SAMN05660380_01960
X8	GB514	ADN63013.1	hypothetical protein XFLM_05325
X8	IVIA5235	RHW43970.1	hypothetical protein D1605_02370
X8	M23	ACB91476.1	hypothetical protein XfasM23_0018
X8	Stags_Leap	WP_012382406.1	hypothetical protein
X8	XYL1732	RUA36537.1	hypothetical protein DX878_08170
X8	XYL2055	RUA36346.1	hypothetical protein DX877_08415
X9	Ann-1f	AIC11168.1	hypothetical protein D934_05670
X9	BB01	WP_071869633.1	hypothetical protein
X9	CFBP7969	WP_128723188.1	hypothetical protein
X9	CFBP7970	WP_128712471.1	hypothetical protein
X9	CFBP8071	WP_012382736.1	hypothetical protein
X9	CFBP8073	WP_081046836.1	hypothetical protein
X9	CFBP8078	WP_128723675.1	hypothetical protein
X9	CFBP8082	WP_128712471.1	hypothetical protein
X9	CFBP8351	WP_012382736.1	hypothetical protein
X9	CFBP8356	WP_128735160.1	hypothetical protein
X9	ESVL	WP_128283828.1	hypothetical protein
X9	IVIA5235	RHW43071.1	hypothetical protein D1605_05685
X9	IVIA5901	WP_128283828.1	hypothetical protein
X9	M12	ACA12828.1	conserved hypothetical protein
X9	M23	ACB93290.1	conserved hypothetical protein
X9	Temecula1	AAO29621.1	conserved hypothetical protein
X10	9a5c	WP_010895081.1	hypothetical protein
X10	Ann-1c	WP_071869525.1	hypothetical protein
X10	ATCC_35871	WP_076613198.1	hypothetical protein
X10	BB01	WP_071869525.1	hypothetical protein
X10	CFBP8072	WP_080939625.1	hypothetical protein
X10	CFBP8073	WP_081046794.1	hypothetical protein
X10	CFBP8078	WP_076613198.1	hypothetical protein
X10	CFBP8082	WP_011098336.1	hypothetical protein
X10	CFBP8356	WP_071869525.1	hypothetical protein
X10	De_Donno	ARO69520.1	hypothetical protein B9J09_11,405
X10	EB92.1	EGO82941.1	hypothetical protein XFEB_00186
X10	ESVL	WP_076613198.1	hypothetical protein
X10	IVIA5901	WP_076613198.1	hypothetical protein
X10	Stags_Leap	WP_011098336.1	hypothetical protein
X10	Temecula1	AAO29857.1	conserved hypothetical protein
X11	32	ETE33528.1	hypothetical protein B398_04735

Orthogroup	Strain	Accession ID	Protein
X11	6c	OJZ71494.1	hypothetical protein B375_0204435
X11	9a5c	WP_023906433.1	hypothetical protein
X11	CFBP8072	WP_023906433.1	hypothetical protein
X11	COF0324	KXB20454.1	hypothetical protein ADT30_07460
X11	COF0407	KXB10569.1	hypothetical protein ADT33_11170
X11	CVC0251	KXB18692.1	hypothetical protein ADT28_12705
X11	De_Donno	ARO68399.1	hypothetical protein B9J09_04535
X11	Fb7	ALR08862.2	hypothetical protein XFFB_06160
X11	OLS0478	KXB13148.1	hypothetical protein ADT32_01130
X11	OLS0479	KXB15169.1	hypothetical protein ADT31_08620
X11	Salento-1	AVI20533.1	hypothetical protein BCV75_04240
X11	Salento-2	AVI22547.1	hypothetical protein BC375_04285
X12	9a5c	WP_010893609.1	hypothetical protein
X12	Ann-1c	WP_080702450.1	hypothetical protein
X12	ATCC_35871	WP_080654439.1	hypothetical protein
X12	CFBP7969	WP_011097630.1	hypothetical protein
X12	CFBP7970	WP_011097630.1	hypothetical protein
X12	CFBP8071	WP_011097630.1	hypothetical protein
X12	CFBP8082	WP_011097630.1	hypothetical protein
X12	CFBP8351	WP_011097630.1	hypothetical protein
X12	IVIA5235	RHW40557.1	hypothetical protein D1605_07495
X12	Stags_Leap	WP_011097630.1	hypothetical protein
X12	Temecula1	AAO28270.1	conserved hypothetical protein
X12	XYL1732	RUA39309.1	hypothetical protein DX878_01355
X12	XYL2055	RUA38957.1	hypothetical protein DX877_02350
X13	Ann-1c	WP_080702445.1	hypothetical protein
X13	CFBP7969	WP_011098370.1	hypothetical protein
X13	CFBP7970	WP_011098370.1	hypothetical protein
X13	CFBP8071	WP_011098370.1	hypothetical protein
X13	CFBP8082	WP_011098370.1	hypothetical protein
X13	CFBP8351	WP_011098370.1	hypothetical protein
X13	DSM_10026	SHG49753.1	hypothetical protein SAMN05660380_00832
X13	IVIA5235	RHW37147.1	hypothetical protein D1605_11880
X13	M23	ACB93594.1	hypothetical protein XfasM23_2199
X13	Stags_Leap	WP_011098370.1	hypothetical protein
X13	Temecula1	AAO29915.1	conserved hypothetical protein
X13	XYL1732	RUA35308.1	hypothetical protein DX878_10580
X13	XYL2055	RUA35511.1	hypothetical protein DX877_10330
X14	Ann-1c	WP_020851938.1	hypothetical protein
X14	Ann-1f	AIC10956.1	hypothetical protein D934_01000
X14	ATCC_35879	KGM21311.1	hypothetical protein JT24_00090
X14	CFBP8073	WP_058564914.1	hypothetical protein
X14	CFBP8356	WP_126709104.1	hypothetical protein
X14	CO33	KQH73081.1	hypothetical protein AOT81_10435
X14	GB514	ADN63011.1	hypothetical protein XFLM_05315
X14	IVIA5235	RHW43967.1	hypothetical protein D1605_02345
X14	Mul-MD	EWG13682.1	hypothetical protein P910_003083
X14	MUL0034	AIC14122.1	hypothetical protein P303_11870
X14	XYL1732	RUA36534.1	hypothetical protein DX878_081,45
X14	XYL2055	RUA36348.1	hypothetical protein DX877_08440

Orthogroup	Strain	Accession ID	Protein
X15	ATCC_35879	KGM20025.1	polyvinylalcohol dehydrogenase
X15	CFBP8073	WP_058564468.1	polyvinylalcohol dehydrogenase
X15	Dixon	EAO13916.1	quinoprotein
X15	DSM_10026	SHG72011.1	polyvinyl alcohol dehydrogenase (cytochrome)
X15	Griffin-1	ERI60323.1	hypothetical protein M233_04935
X15	M23	ACB92802.1	Pyrrolo-quinoline quinone
X15	Mul-MD	EWG15353.1	Pyrrolo-quinoline quinone
X15	MUL0034	AIC13805.1	polyvinylalcohol dehydrogenase
X15	sycamore_Sy-VA	KFA42156.1	hypothetical protein DF22_001285
X15	*Xylella_taiwanensis*	WP_081755433.1	polyvinylalcohol dehydrogenase
X16	Ann-1c	WP_080702461.1	hypothetical protein
X16	CFBP7969	WP_100206166.1	hypothetical protein
X16	CFBP7970	WP_100206166.1	hypothetical protein
X16	CFBP8071	WP_100206166.1	hypothetical protein
X16	CFBP8082	WP_100206166.1	hypothetical protein
X16	CFBP8351	WP_100206166.1	hypothetical protein
X16	IVIA5235	RHW37853.1	hypothetical protein D1605_09725
X16	XYL1732	RUA39761.1	hypothetical protein DX878_00855
X16	XYL2055	RUA38619.1	hypothetical protein DX877_02410
X17	Ann-1f	AIC11426.1	hypothetical protein D934_09155
X17	Ann-1f	AIC11450.1	hypothetical protein D934_09355
X17	De_Donno	ARO69062.1	hypothetical protein B9J09_08535
X17	EB92.1	EGO81085.1	hypothetical protein XFEB_02081 partial
X17	EB92.1	EGO81107.1	hypothetical protein XFEB_02058 partial
X17	GB514	ADN63933.1	hypothetical protein XFLM_10280
X17	Griffin-1	ERI59206.1	hypothetical protein M233_10900
X17	Salento-1	AVI21112.1	hypothetical protein BCV75_07945
X17	Salento-2	AVI23134.1	hypothetical protein BC375_08005
X18	32	ETE31344.1	hypothetical protein B398_07770
X18	3124	ALQ97281.1	hypothetical protein XFC3_07785
X18	6c	OJZ70788.1	hypothetical protein B375_0207125
X18	9a5c	WP_023906732.1	hypothetical protein
X18	CFBP8072	WP_058569180.1	hypothetical protein
X18	COF0324	KXB21537.1	hypothetical protein ADT30_04125
X18	Pr8x	ALR04485.1	hypothetical protein XFPR_07525
X18	U24D	ALQ94784.1	hypothetical protein XFUD_06025
X19	9a5c	WP_010895066.1	hypothetical protein
X19	Ann-1c	WP_012338119.1	hypothetical protein
X19	BB01	WP_012338119.1	hypothetical protein
X19	CFBP8072	WP_081044425.1	hypothetical protein
X19	CFBP8078	WP_012338119.1	hypothetical protein
X19	CFBP8356	WP_081033408.1	hypothetical protein
X19	M12	ACA13068.1	conserved hypothetical protein
X20	9a5c	WP_042462811.1	hypothetical protein
X20	BB01	WP_038230917.1	hypothetical protein
X20	EB92.1	EGO82211.1	hypothetical protein XFEB_00954
X20	Mul-MD	EWG15085.1	hypothetical protein P910_001384
X20	MUL0034	AIC13874.1	hypothetical protein P303_07620
X20	sycamore_Sy-VA	KFA41407.1	hypothetical protein DF22_001986
X20	U24D	ALQ94194.1	hypothetical protein XFUD_02410

Orthogroup	Strain	Accession ID	Protein
X21	32	ETE35800.1	hypothetical protein B398_01270
X21	Ann-1f	AIC11000.1	hypothetical protein D934_02200
X21	DSM_10026	SHG52207.1	hypothetical protein SAMN05660380_00865
X21	GB514	ADN63229.1	hypothetical protein XFLM_06490
X21	Griffin-1	ERI60635.1	hypothetical protein M233_03245
X21	MUL0034	AIC14072.1	hypothetical protein P303_10650
X22	9a5c	WP_010894783.1	hypothetical protein
X22	ATCC_35871	WP_012337970.1	hypothetical protein
X22	BB01	WP_012337970.1	hypothetical protein
X22	CFBP8072	WP_081044418.1	hypothetical protein
X22	De_Donno	ARO68210.1	hypothetical protein B9J09_03320
X22	M12	ACA12424.1	conserved hypothetical protein
X23	Ann-1c	WP_024748928.1	hypothetical protein
X23	Ann-1f	AIC10997.1	hypothetical protein D934_02110
X23	Mul-MD	EWG13319.1	hypothetical protein P910_003433 partial
X23	MUL0034	AIC14075.1	hypothetical protein P303_10740
X23	sycamore_Sy-VA	KFA40036.1	hypothetical protein DF22_003374 partial
X23	XYL2055	RUA34445.1	hypothetical protein DX877_11825 partial
X24	Ann-1f	AIC11124.1	hypothetical protein D934_04800
X24	GB514	ADN63662.1	hypothetical protein XFLM_08850
X24	Griffin-1	ERI61041.1	hypothetical protein M233_01135
X24	Mul-MD	EWG13824.1	hypothetical protein P910_002916
X24	MUL0034	AIC13605.1	hypothetical protein P303_03195
X24	sycamore_Sy-VA	KFA41124.1	hypothetical protein DF22_002288
X25	6c	OJZ70495.1	hypothetical protein B375_0205340
X25	CFBP8072	WP_058569749.1	hypothetical protein
X25	COF0407	KXB17099.1	hypothetical protein ADT33_01285
X25	OLS0479	KXB16511.1	hypothetical protein ADT31_06310
X25	Salento-2	AVI22762.1	hypothetical protein BC375_05650
X26	Ann-1c	WP_020851770.1	hypothetical protein
X26	ATCC_35871	WP_012337738.1	hypothetical protein
X26	BB01	WP_071869659.1	hypothetical protein
X26	CFBP8073	WP_020851770.1	hypothetical protein
X26	M12	ACA11674.1	hypothetical protein Xfasm12_0676
X27	Ann-1c	WP_080702480.1	hypothetical protein
X27	CFBP8356	WP_128735178.1	hypothetical protein
X27	Stags_Leap	WP_011098217.1	hypothetical protein
X27	Temecula1	AAO29538.1	conserved hypothetical protein
X28	Ann-1f	AIC10973.1	hypothetical protein D934_01,405
X28	GB514	ADN63081.1	hypothetical protein XFLM_05710
X28	Mul-MD	EWG13920.1	hypothetical protein P910_002844
X28	MUL0034	AIC14107.1	hypothetical protein P303_11,480
X29	CFBP8356	WP_057683294.1	hypothetical protein
X29	CO33	KQH73482.1	RTX toxin
X29	Mul-MD	EWG15232.1	hypothetical protein P910_001531
X29	MUL0034	AIC12677.1	RTX toxin Ca2+-binding protein
X30	DSM_10026	SHG60814.1	hypothetical protein SAMN05660380_01057
X30	M23	ACB93005.1	hypothetical protein XfasM23_1597
X30	*Xylella taiwanensis*	WP_038271518.1	hypothetical protein

Orthogroup	Strain	Accession ID	Protein
X30	*Xylella taiwanensis*	WP_038271520.1	hypothetical protein
X31	Griffin-1	ERI60788.1	hypothetical protein M233_02020
X31	M12	ACA13072.1	hypothetical protein Xfasm12_2221
X31	Mul-MD	EWG15019.1	hypothetical protein P910_001777
X31	MUL0034	AIC14162.1	hypothetical protein P303_12405
X32	32	ETE33419.1	hypothetical protein B398_04190
X32	3124	ALQ96690.1	hypothetical protein XFC3_04045
X32	COF0324	KXB19940.1	hypothetical protein ADT30_08475
X33	32	ETE34976.1	hypothetical protein B398_02560
X33	Ann-1f	AIC11594.1	hypothetical protein D934_11720
X33	GB514	ADN62368.1	hypothetical protein XFLM_01825
X34	32	ETE35666.1	hypothetical protein B398_01790
X34	GB514	ADN62522.1	hypothetical protein XFLM_02615
X34	MUL0034	AIC13939.1	hypothetical protein P303_08415
X35	6c	OJZ70003.1	hypothetical protein B375_0209695
X35	COF0324	KXB21939.1	hypothetical protein ADT30_02250
X35	Pr8x	ALR04924.1	hypothetical protein XFPR_10115
X36	COF0407	KXB10262.1	hypothetical protein ADT33_11255 (plasmid)
X36	OLS0478	KXB09756.1	hypothetical protein ADT32_00275 (plasmid)
X36	OLS0479	KXB13419.1	hypothetical protein ADT31_00315 (plasmid)
X37	EB92.1	EGO82688.1	hemagglutinin/hemolysin partial
X37	XYL1732	RUA34462.1	hypothetical protein DX878_11735 partial
X37	XYL2055	RUA34494.1	hypothetical protein DX877_11780 partial
X38	XYL1732	RUA34481.1	hypothetical protein DX878_11720 partial
X38	XYL2055	RUA34461.1	hypothetical protein DX877_11810 partial
X38	XYL2055	RUA34480.1	hypothetical protein DX877_11795 partial
X39	Ann-1f	AIC10963.1	hypothetical protein D934_01150
X39	Ann-1f	AIC11032.1	hypothetical protein D934_03075
X40	GB514	ADN63039.1	hypothetical protein XFLM_05465
X40	Griffin-1	ERI60927.1	hypothetical protein M233_01880
X41	CFBP7970	WP_126715040.1	hypothetical protein
X41	CFBP8071	WP_126715040.1	hypothetical protein
X42	CFBP7970	WP_128712485.1	hypothetical protein
X42	CFBP8078	WP_128723916.1	hypothetical protein
X43	CFBP8072	WP_058569714.1	hypothetical protein
X43	Pr8x	ALR03934.1	hypothetical protein XFPR_04125
X44	CFBP8078	WP_128723706.1	30S ribosomal protein THX
X44	*Xylella taiwanensis*	WP_081755402.1	30S ribosomal protein THX
X45	CVC0256	KXB15101.1	hypothetical protein ADT29_04315
X45	Pr8x	ALR04893.1	hypothetical protein XFPR_09955
X46	DSM_10026	SHG83966.1	hypothetical protein SAMN05660380_01558
X46	GB514	ADN63837.1	hypothetical protein XFLM_09770
X47	Dixon	EAO12308.1	conserved hypothetical protein
X47	M12	ACA11898.1	conserved hypothetical protein
X48	Dixon	EAO12662.1	hypothetical protein XfasaDRAFT_0679
X48	M12	ACA12296.1	hypothetical protein Xfasm12_1366

Orthogroup	Strain	Accession ID	Protein
X49	Dixon	EAO13347.1	conserved hypothetical protein
X49	M12	ACA11357.1	conserved hypothetical protein
X50	EB92.1	EGO81883.1	Autotransporter adhesin partial
X50	IVIA5235	RHW48442.1	cell surface protein partial
X51	EB92.1	EGO82873.1	hypothetical protein XFEB_00205
X51	J1a12	ALR02911.1	hypothetical protein OY18_12675
X52	M12	ACA11352.1	hypothetical protein Xfasm12_0333
X52	M23	ACB91752.1	hypothetical protein XfasM23_0303