Dipartimento di Scienze Agrarie e Forestali

Corso di Dottorato di Ricerca in

Scienze delle Produzioni Vegetali e Animali - XXXI Ciclo

Approaching the micro-evolution of bacterial pathogens of plants using MLVA: invasive outbreaks vs endemic diseases

AGR/12 - Patologia Vegetale

PhD student
Dr. Yaseen Jundi RAHI

PhD Coordinator
Prof. Roberta BERNINI

Supervisors
Prof. Giorgio Mariano BALESTRA

Prof. Stefania MASCI

Prof. Angelo MAZZAGLIA

Co- advisor

Dr. FRANCO VALENTINI

AA 2019/2020

Thesis abstract

Approaching micro-evolution of bacterial pathogens of plants using MLVA: invasive outbreaks *vs* endemic diseases

Introduction

The movement of goods around the world and broad borders have facilitated the diffusion of plant diseases through the movement of pathogens into a new area, and climate change has helped the emergence or re-emergence of new pathogens that negatively affect food production. In this thesis, we have dealt with three important pathogens *Pseudomonas syringae* pv. *actinidiae, Xylella fastidiosa* , as invasive outbreaks pathogens and *Pseudomonas savastanoi* an endemic pathogen. These pathogens have a very important impact on food production not only in Italy but also worldwide. In order to understand the genetic characteristics, various methods have been developed for analysing plant pathogenic bacteria, these methods are based on either restriction enzymes or DNA segments amplified by PCR. These methods have differed in their discriminative power, reproducibility, and ease of results interpretation. Hence, we have developed and applied MLVA (Multiple Loci Variable Number of Tandem Repeats Analysis) for genotyping the three pathogens.

The aims

To develop a robust MLVA (Multiple Loci Variable Number of Tandem Repeats Analysis) scheme suitable for global and local studies of three main bacterial diseases (endemic and invasive) in Italy: Knot disease on olive, oleander, and ash which is caused by *Pseudomonas savastanoi* pathovars, as an endemic pathogen, then canker of kiwifruit which is caused by *P. syringae* pv. *actinidiae* (Psa), and the new emerge olive quick decline syndrome which is caused by *Xylella fastidiosa* subspecies *pauca* st53 as an invasive outbreak. Then, we used this

approach to understand differences in populations of each pathogen, elucidate the diffusion and movements of these populations.

The results

The results obtained in these studies point out that MLVA represents a very promising first-line assay for large-scale routine genotyping prior to whole-genome sequencing of only the most relevant samples in case of outbreaks. In addition, its ability to differentiate the pathovars, biovars, and subspecies, `the method went beyond that when it was able to differentiate the sequence types of the same subspecies. MLVA assay has great potential as an easy and effective tool not only to recognize and schedule the presence of the above-mentioned types of bacterial species all over the world but above all to trace their movements on local and to international scale, in addition to its ability supporting the simple detection of contaminated materials with key information concerning specific haplotypes populations.

Thesis committee

- Prof.

- Prof.

-Prof.

-Prof.

Acknowledgments:

*(**Alhamdu**[1] **LILAH**) before and after and forever, Hamdan as he deserves, and as it supposed to be for his generosity.*

First of all , I would like to thank the Italian Ministry of foreign affairs and the Italian embassy in Baghdad, Iraq, for their great cooperation in the agriculture field with the ministry of agriculture Iraq.

I want to thank the Mediterranean Agronomic Institute of Bari (CIHEAM-MAIB) that accepted me as a member and opened the door for a better future and a better life. My personal gratitude goes to the Director of the institute Dr. M. Raeli, Then DR. Biagio and Dr. Olympia Antonelli. All the appreciations to the (CIHEAM-MAIB) staff.

My sincere thanks and appreciations go to the coordinator of the "Integrated Pest Management" department in CIHEAM-MAIB, Dr. Anna Maria D'Onghia, DR. Khaled dejaluh, Dr. Franco Valentini, DR. Thaer Yaseen and all the people in Integrated Pest Management" department in CIHEAM-MAIB.

I would like to thank the University of Tuscia for accepting me as one of her students, and the support I get all these years. My thanks to the D.A.F.N.E. department.

A special thanks from the heart to Prof. Leonardo Varvaro.

I would like to express my sincere gratitude and gratefulness to my supervisor Giorgio Mariano BALESTRA for his continuous support, guidance, and great help throughout my studies. Prof, Thank you so much again.

There are no enough words to show my appreciation to you Angelo Mazzaglia, but at least I can thank you from my deep heart truly, truly grateful for your great support, encouragement, and guidance throughout the time of my study during the years without you this work could not be done.

Dear Alfredo Fabi, thank you from my deep heart. You are the brother, I have obtained in this city, you were always standing with me looking after me and my

[1] thanks

family. As always you say may Allah keep our families in his hands. Thank you and your family all of them you were my second family.

Also, my sincere thanks and gratitude to, Dr. Maria Claudia Taratufolo, Dr. Maria pia, Dr. Silvia Turco, my sincere thanks to Phydia staff Dr. Vincenzo Tagliavento, and Dr. Serena Ciarroni.

To the best friends in the university Sara Fransconi and Elena Chaboteaux. Sara you know, I will always remember your countless help not only in the lab but in life too. You are a great friend. I will always remember you (Baad eyounee).

Elena the girl who took my hands for the first steps in my work. Elena, I will never your help, kindness, time, our friendship you did the best for me.

To all my Colleagues in bacteriology and Mycology lab for their kind support and collaboration.

Thanks to all my friends from Iraq specially Mahdi Bdair, Mukhtar Al -obaidi, who support me along with my studies.

A special thanks to Dr. Hussein Ali saoud AL- Jumaily, Dr. Faisel Rasheed, Dr. Mahdi Al-Qaisi, and Dr. Qutaiba Al-Jubouri.

I would like to express my sincere appreciation to my wife and love for her patience, understanding, care and moral support. Thank you so much, you are the real women in my back.

To my children; Mohammed (my friend) you were so patient with me, you will understand more when you grew up. Murtda (Viterpicolo) you add happiness to our life thank you.

Thank you, my mother, for your prayers my dear brothers and father may Allah bless your souls, Thank you, my brothers and sisters.

List of Figures

List of Tables

List of abbreviations

MLOs	Mycoplasma-like organisms
OQDS	Olive Quick Decline Syndrome
HLB	Huanglongbing
IPM	Integrated pest management
FAO	Food and Agriculture Organization
SAR	Systemic acquired resistance
EBDC	Ethylenebisdithiocarbamate
WGS	Whole-genome sequences
DNA	Deoxyribonucleic acid
kb	kilobases
TRs	Tandem repeats
HGT	Horizontal gene transfer
PFGE	pulsed-field gel electrophoresis
Rep-PCR	Repetitive sequence-based PCR
RAPD	Random amplification of polymorphic DNA
RFLP	Restriction fragment length polymorphism PCR
AFLP	Amplified Fragment Length Polymorphism
MLST	Multilocus sequence typing
MLVA	multilocus variable-number tandem repeat analysis
Co-PCR	Co-operational polymerase chain reaction
SSRS	Simple sequence repeats
SNPs	Single Nucleotide Polymorphisms
CNVs	Copy Number Variations
VNTRs	Variable-number tandem repeats
Psa	P. syringae pv. *actinidiae*
IAA	Indole acetic acid
iaaH	Indole acetamide hydrolase
QS	Quorum sensing
TTSS	Type III secretion system
c-di GMP	Cyclic diguanylate
Psv	*Pseudomonas savastanoi*
WHOP	Woody host and Pseudomonas
CVC	Citrus Variegated Chlorosis
X. fastidiosa	*Xylella fastidiosa*
CoDiRO	Complesso del disseccamento Rapido dell'olivo
%	Percent
°C	Degree Celsius
bp	Base pairs
LOPAT	Levan, Oxidase, Potato, Arginine and Tobacco tests
Sp.	Species

PCR	Polymerase Chain Reaction
REP	Repetitive Extragenic Palindromic sequence
rpm	Revolutions per minute
SDW	Sterile distilled water
sec	Second
sp.	Species
TAE	Tris Acetate-EDTA
Taq	Thermophilus aquaticus
µl	Microliter
µM	Micromole
pv.	Pathovar
MLGs	Multilocus genotypes
Subsp.	subspecies

TABLE OF CONTENTS

Chapter 1 **: Introduction**

1.1 **Plant disease**

Plant diseases remain an endless threat to our food, crops, cash crops, social life, and landscape. Disease in plants can be defined as follows: the series of visible and invisible reactions of plant cells and tissues to biotic (pathogenic organisms) and abiotic (environmental factors) or their combinations, which result in adverse changes in the form, function, or integrity of the plant and may lead to partial damage of plant parts and or in severe cases death of the entire plant or fields plants. The inherited defects (e.g. genetic disease) of the plant can be added to the definition, due to their important roles. Human involvement and activities also play an important role in the development of plant disease (1).

Abiotic factors, i.e. environmental or physiological conditions, such as temperature, moisture, mineral nutrients, pollutants, malnutrition, toxic agents, *etc*., occurring above or below certain levels tolerated by plants can disturb plant functions (2). In simple words, the disease is any kind of distribution in plant function due to biotic and abiotic or human activities that make the plant unable to carry out its physiological functions correctly. Therefore, plant diseases may be the result of living and/or non-living causes.

Biotic diseases are caused by living organisms, known as pathogens, which include microorganisms (e.g., viruses, bacteria, oomycetes, fungi, nematodes, and parasitic plants) (3). A plant pathogen is an organism that can inhabit the rhizosphere (zone of interaction of root and soil) and/or phyllosphere (aerial part of the plant) and the endosphere (internal parts of the plants) and can disturb the growth, production, and survival of the plant by causing changes in the morphology and behavior of the plant or its tissues (4). For example, when bacteria attack plants, they usually disturb plant cell metabolism by secreting enzymes, toxins, growth regulators, and other substances, and by absorbing nutrients from the host cells for their own use (5). Some pathogens may also cause disease by growing and multiplying in the xylem or phloem vessels of plants, thereby blocking the upward transportation of water and nutrients, as in the case of *Xylella fastidiosa* , the causal agent of Pierce disease and olive quick decline syndrome (OQDS) in Italy, and of *Candidatus Liberibacter asiaticus*, the causal agent of Huanglongbing (HLB) disease (1).

Figure 1: The effects of the disease in fields of (A) Horse chestnut plants[2]; (B) Olive plants[3]; C) Pears.[4]

Plant diseases vary in their effects, ranging from insignificant symptoms in the garden to disasters that destroy main food crops planted over large areas (6). Food losses due to crop infections from pathogens are a persistent problem in agriculture since these organisms reduce quality, yields, shelf-life, and consumer satisfaction, in both pre- and post-harvest (7). Major disease outbreaks may cause starvation, hunger, and migration, such as potato blight disease caused by *Phytophthora infestans*, the Great Bengal Famine of 1943, black stem rust of wheat in the USA, and many others (8). It is difficult to control plant pathogens due to their population differences in time, space, and genotype. Another important aspect is that a pathogen may evolve to overcome the resistance of the plant (9). In order to meet the increasing demands of feeding the growing population, the improvement of crop protection strategies to prevent losses due to pathogens plays an important role in securing both food quality and quantity (7).

In conclusion, it can be said that not only do biotic and abiotic factors cause plant disease but so does the interaction between them. In addition, human activities consisting of agricultural

[2] https://www.dailymail.co.uk/news/article-3223818/
[3] https://www.cbsnews.com/news/pathogen-destroying-olive-trees-in-italy-*xylella-fastidiosa*/
[4] https://wiki.bugwood.org/ *Erwinia amylovora*

practices such as cropping systems and extensive pesticide applications could also lead to the development of this complex/disease (10). We will discuss biotic factors which include three bacterial pathogens; hence, other aspects will not be discussed here.

1.2 Bacterial plant pathogens

Bacteria are single-celled prokaryotic microorganisms, classified as eubacteria and archaebacteria according to differences in the composition of their cell wall and cytoplasmic membrane. Eubacterial plant pathogens are sub-divided into gram-negative bacteria, gram-positive bacteria, and Mycoplasma-like organisms (MLOs). The difference between them and Mollicutes (1) is the presence of cell wall characteristics. MLOs lack a cell wall; their outer boundary is instead a cytoplasmic membrane, which gives them some unusual properties not found in most eubacteria.

Common plant pathogenic bacteria are all small, single, rod-shaped cells, apart from the filamentous Streptomyces bacteria. Bacteria range from 0.6 to 3.5 μm in diameter (11). The cell wall of most bacterial species is surrounded by a viscous material known as a capsule, which provides protection against stresses, immune responses, antibacterial agents, and antibiotics. Gram-positive bacteria, on the other hand, produce endospores to help them survive unfavorable conditions (12).

Many bacterial species are motile by means of their flagella, which are filamentous structures consisting of proteins. The number and distribution of flagella vary among species. Some species have only one flagellum present at one end (polar or monotrichous flagellation). In some species have flagella located at both ends of the cell (lophotrichous bacteria), while other species have flagella at various places on the cell (peritrichous bacteria) (1).

There are over 100 bacterial species responsible for plant diseases affecting productivity. Difficulties in controlling bacterial diseases (both chemically and biologically) are due to different factors such as limits on using antibiotics in Europe, the development of resistance, wide host range, population size, and evolution. Bacterial diseases of fruit trees, for example, may cause losses that reach 50% -100% and the whole crop could be destroyed as a result of the bacterial attack (13).

Bacteria multiply asexually by binary division. The cell grows to twice its size and splits into two cells; this division must occur at the correct time and correct location in the cell.

Progeny cells receive a complete set of genes of their parent cell. On the other hand, the gram-positive bacteria Streptomyces species were found to multiply by forming a spore (14)(15). Bacteria replicate at a high rate. A single bacterial cell may produce one million progeny bacteria in less than a day if there are optimal environmental factors. However, this high rate will gradually slow and finally stop at a certain point (16). Plant-associated bacteria may be beneficial or harmful; they can live in the rhizosphere (zone of root and soil interaction) and/or phyllosphere (aerial part of the plant), and the endosphere (internal parts of the plants). Plant pathogenic bacteria are found as saprophytes in plant debris or in the soil, but mainly develop in the host plant as parasites, or on the plant surface as epiphytes (1).

Pathogenic bacterial species that cause serious diseases of plants throughout the world have increased from several genera (five) to almost 40 genera, belonging to three gram-negative bacterial families (*Xanthomonadaceae, Pseudomonadaceae,* and *Enterobacteriaceae*), and one gram-positive family (*Corynebacteriaceae*) (17,18). There are more than 150 bacterial species, including over 10 species considered to be the most important bacterial plant pathogens because they cause high economic losses: *Pseudomonas syringae* pathovars; *Ralstonia solanacearum, Agrobacterium tumefaciens*; *Xanthomonas oryzae* pv. oryzae; *Xanthomonas campestris* pathovars, *Xanthomonas axonopodis* pathovars; *Erwinia amylovora*; *Xylella fastidiosa ,* Dickeya (*dadantii* and *solani)*; *Pectobacterium carotovorum,* and *Pectobacterium atrosepticum* (19). These species are the most scientifically/ economically important bacterial pathogens (20).

1.3 Economic importance of plant diseases

The emergence and introduction of new plant pathogens have been aided by different factors, such as open borders, food and goods trade, climate changes, multi-cropping, monoculture production systems, changing from traditional to high-intensity agriculture systems, excessive use of pesticides, and many others (15). The interaction of these factors has affected the natural plant-pathogen interaction, in turn disturbing the coevolutionary system by creating a unique condition that makes the bacterial population evolve and causes epidemics (21). Bacterial plant pathogens are a major problem worldwide for agriculture because they are difficult to control, and their impact on agriculture is huge. Meanwhile, the results of any bacterial attack on a crop will be more significant at the global level, but this is related to different factors, for example, local climate, cropping practices, the choice of crop and plant cultivar, plant species, even more, the characteristics of pathogens (pathovars, biovars,

sequence typing) (22). Many plant pathogenic bacteria are asymptomatic, and any simple change in one or some of the previous factors may lead to favorable conditions that help the emergence or re-emergence of plant pathogens (23). In 2002, the bleeding canker of the European horse chestnut disease caused by *Pseudomonas syringae* pv. *aesculi* was reported in England, Wales, Scotland, Netherlands, Belgium, France, and Germany. It affected hundreds of thousands of trees in these countries, resulting in severe damage to rural and urban landscape features (15). *Pseudomonas syringae* pathovars cause diseases of a wide host range, e.g. monocots, herbaceous dicots, and woody dicots, worldwide. For example, *Pseudomonas syringae* pv. *actinidiae* was present in Japan since 1984, but with a limited impact on the local production in Japan and China (24–27). However, in 2008 a new biovar spread worldwide starting in Italy, New Zealand, Chile, China and other countries (28,29) and caused severe losses in all major areas of kiwifruit cultivation (30). The emergence of the bacterium *Xylella fastidiosa subsp. pauca* in Italy (Apulia region) in 2013 poses a great threat to Europe and the Mediterranean basin because of its host's number over 350 plant species, and vectors. The bacterium has also been detected in France and Spain (31). The disease was not reliant only on the bacterium but also on the interaction of plant, pathogen, and vectors and the effects of abiotic factors (32). The impact of *Xylella fastidiosa* in Italy has caused not only economic losses, but also agricultural, environmental, political, and cultural damage (32). There is no real estimate of the financial cost, but it is estimated that there are approximately 25 million olive trees in the demarcated area, so that the loss to olive agriculture is substantial (Bosca personal communication). Another important species is *Agrobacterium* spp., which can attack more than 500 plant species causing significant economic losses in the production of many plants worldwide. Most of these species affect fruit species: stone and pome fruits, nut trees and grapevine, as well as some perennial ornamentals (33). Precise estimation of losses in crops attacked by bacteria is difficult, because comparison of crop sizes in protected and non-protected plantations is impossible.

1.4 Epidemiology of plant bacterial diseases

A plant disease to become established in a new area, it needs three factors: a susceptible host, pathogen virulence, and suitable environmental conditions (humidity and temperature). Human activities can cause wounds or damage to the plant tissue, in addition, frost, hail, wind, and rain, all these factors can facilitate bacterial entrance (34). Epidemiology studies plant disease occurrence over space and time, with information related to the pathogen’s ability to survive (11). The bacterial disease process includes two important steps: invasion of plant tissue and

multiplication. Plant pathogenic bacteria may live as epiphytes and some as saprotrophs on a host plant, but they are unable to enter plant tissue directly. The process of infection starts with the attachment of bacteria to the plant surface followed by bacteria penetrating through the wounds or natural openings, such as hydathodes, stomata, and nectaroides then is the multiplication of the bacteria in the new environment; they multiply in the intercellular spaces from where they can spread into different plant tissue (35). The environmental conditions, especially humidity and temperature, can further develop or limit the pathogen, disease, and development of symptoms. Bacterial numbers can decrease rapidly when the infected tissue dies. Moreover, many bacteria are unable to survive on dead plant debris, which in the end reduces the source of inoculum (36).

Bacteria may spread by one or multiple methods, which are based on bacterial location. For example, bacteria may be transferred by infected plant materials or contaminated seed, others may be transferred by animals, insects, birds, mites or by human equipment (machinery and tools), and activities such as cultivation, grafting or pruning. They can also be spread by water, (rain, overhead irrigation) and wind (11,17). The most important method of bacterial dissemination involves vectors, as for OQDS caused by *Xylella fastidiosa* and HLB caused by *Candidatus Liberibacter asiaticus*, both of which are transferred by insects. Bacteria may survive in insects, knots, plant debris, soil, water, or seeds.

1.5 Control measures

One of the main goals of growers is to earn an income from their crops by supplying the local or international markets to feed humans and animals. On the other hand, consumers seek safe and tasty products with an attractive appearance. In addition, they are also concerned about the conservation of the ecosystem and natural resources (1). These targets face the problem of plant disease, which disease control helps to solve and to achieve these aspects (37). The main methods of disease control are based on avoidance, exclusion, eradication, protection, developing plant resistance, and finally cure (38).

The best method of disease control is what is now known as Integrated Pest Management (IPM) and consists of various integrated activities. According to the Food and Agriculture Organization (FAO), IPM is defined as ‘the careful consideration of all available pest control techniques and subsequent integration of appropriate measures that discourage the development of pest populations and keep pesticides and other interventions to levels that are economically justified and reduce or minimize risks to human health and the environment (39).

Cultural, physical, chemical and biological control methods, resistant varieties, improved resistance, quarantine and inspection and the recent trend methods of integrated pest management should be used alone or together to support disease management and to reduce the pathogen population below an acceptable threshold (40). Combinations of control techniques aim to reduce the pathogen population to a harmless level and to reduce the use of pesticides, thereby reducing the cost of control.

All these methods have different degrees of success and their own limitations. Factors that interact with each other and on disease control include pathogen-host interaction, pathogen epidemiology, environmental factors, and the seasonal or physiological stages of both host and pathogen (1). The best approach is to use appropriate integrated methods to help control the disease. The next sections will discuss these methods in a general way.

1.5.1. **Chemical methods**

A plant disease caused by phytopathogenic bacteria can be difficult to control, especially if the bacterial pathogen has become established and diffused. The use of chemical compounds will be more effective as a preventive measure, or at an early stage in the bacterial infection, which in general can be used for soil treatment, control of insect vectors, and fumigation (41). Agriculture has long made use of copper-based compounds, such as Bordeaux mixture, cupric hydroxide, copper sulfate, ammoniacal copper and copper salts of fatty acids against bacteria as a foliar treatment to protect plants from bacterial diseases (42).

Due to their effectiveness, copper-based compounds are used widely to control or reduce pathogens and have been used successfully against many bacterial plant diseases, including olive knot, Tomato bacterial speck, the bacterial canker of apricot and cherry, canker of kiwi fruit (43). Copper mixed with fertilizer to control the bacterial disease has given good results (44–48). The over-use of copper compounds has, however, seriously damaged agricultural and natural ecosystems. In addition to the limitations of their use, the emergence of resistant strains has been detected in different bacterial species, and this could prevent control of the disease (49).

Phytotoxin accumulation in the soil due to repeated spraying has also been observed; in a study conducted in Italy, it was found that copper concentrations were high enough to represent a risk of toxicity (50,51).

Although of its importance, but the accumulation of copper in soil plays a toxic role in the plant and can contaminate the food chain, although this varies from plant to plant because some plant species can accumulate heavy metals in their tissue (52,53). Not only, copper is a soil toxin, but many bacteria have also developed a resistance to copper (52,54).

In order to overcome copper resistance, farmers started to mix it with other Ethylenebisdithiocarbamate (EBDC) compounds such as maneb or mancozeb, but these compounds still have limited efficacy (55). In the treatment of endophyte pathogenic bacteria, copper is considered as toxic (51).

In the mid-ninetieth century, antibiotics were used extensively against bacterial plant pathogens. Four antibiotics are used in plant disease control (Streptomycin, Oxytetracycline, Gentamicin, and Oxolinic acid) to prevent different plant pathogens, such as *Erwinia* spp *Pseudomonas spp.*, *Xanthomonas campestris*, *Agrobacterium tumefaciens*, mainly *phytoplasmas*. The main areas using these antibiotics are the USA, Mexico, New Zealand, and the Middle East, while some European countries, like Germany, Austria, and Switzerland, have allowed the use of antibiotics on an emergency basis, and under strictly limited conditions (56).

The main use of antibiotics and the widest application was against *Erwinia amylovora*, causal agents of fire blight (57). Extensive use of antibiotics has revealed the development of resistance in bacterial plant pathogens (58). The antibiotics remain one week on plant surfaces and their activity can quickly diminish. One problem for scientists is the possibility of transferring resistance genes from plant pathogenic bacteria into human and animal pathogenic bacteria, but most studies focus on transfer through the soil and deactivation of bacteria in the soil or on plant surfaces (59–62). The possibility still presents if the other bacterial transmission methods are considered and if the plant carries out some of this bacterial resistance to antibiotics.

1.5.2. **Biological control methods**

These methods appeared in reaction to copper and antibiotic resistance, and to the pollution due to the excessive use of pesticides, whose accumulation has negative impacts on ecosystems. Biological control methods were developed for entomology and then plant pathology. For plant pathogenic bacteria, the process is based on using microorganisms as antagonists or natural substances from plants or any other sources to suppress the pathogen or diseases (63). Biocontrol agents work through multiple beneficial characteristics such as

rhizosphere competence, antagonistic potential, and ability to produce antibiotics, lytic enzymes, and toxins (64). These antagonists are either direct or indirect; the mechanism of a direct antagonist is based on direct contact or a high degree of selectivity to the pathogen, whereas the activity of an indirect antagonist does not involve sensing or targeting a pathogen (8).

Bacteria are the main group of beneficial microorganisms. These beneficial bacteria produce different kinds of metabolites, e.g. secretion of siderophores and enzymes, production of growth regulators and antibiotics, and induction of systemic resistance (65).

Bacterial species used to control and suppress plant disease include *Bacillus spp*. *Pseudomonas spp*., *Agrobacterium spp*., and some fungus have also been used, such as *Trichoderma spp*. (66).

Many studies have reported the use of bacterial species or their products against bacterial plant pathogens directly like competition (67,68). Inducing systemic acquired resistance (SAR) in the plant (69–75). Several applications have been approved for controlling bacterial plant pathogens of tomato, pepper, onion and fire blight (76).

Bacteriophages can also be used for biological control in bacterial diseases, in combination with other control methods under the umbrella of IPM strategies. Phages have been used in wide applications against different human, animal and plant pathogens. Many studies have reported the use of phages against plant disease (77–81).

Breeding programs have also been widely used in agriculture, involving the enhancement of plant resistance to disease. The mechanisms used include reengineering plants to produce antibacterial proteins of non-plant origin, inhibiting bacterial pathogenicity or virulence factors, enhancing natural plant defenses, and artificially inducing programmed cell death at the site of infection (82–87).

Different applications on various bacterial plant diseases have been applied to breeding programs(82). The use of certified plant material is an important method to control plant disease and aims to provide planting material free of phytopathogenic bacteria (39,83–88). Different programs have been used for certified plant material to control plant disease (89).

Physical methods are also widely used against plant pathogens, aiming to deactivate or eradicate the pathogen from seeds. These treatments include the use of hot water, hot air, steam, aerated steam, moist hot air, and solar heat, while other methods for the treatment of soil-borne pathogens are less used on bacterial pathogens, such as soil solarization, and hot air sterilization (41,90).

Cultural methods such as host eradication, crop rotation, sanitation by creating unfavorable conditions, plastic traps and mulches have been applied for the control of plant bacterial diseases, with the aim of eradicating and/or reducing pathogen inoculum (91). Most of the above-mentioned methods have their positive and negative impacts cooper-antibiotic creates resistance in bacteria (92). The use of SAR also has a negative impact on the plant (50,93–95).

Development of strains that overcome the resistance genes were also noticed in some breeding programs (96). One of the most important methods in controlling the disease is monitoring and early detection of pathogens, which can help greatly in the reduction of disease spread and the preparation of a rapid management plan (97,98). The concept of IPM appears a necessity in order to reduce the use of chemicals and to control plant pathogens effectively (99). In addition to economic aspects (100).

Detection and identification are critical steps for the suitable application of phytosanitary measures (101). The precise estimation of disease occurrence and severity, or disease effects, in addition, to forecasting temporal and spatial disease spread in specific growing regions are important to set up the control measures. Different methods are used to diagnose and detect disease, including visual plant disease estimation and microscopic evaluation of morphological features to identify pathogens, as well as molecular, serological, and microbiological diagnostic techniques (102). Detection methods will be discussed below.

1.6 Population genetic analysis

The demand for food has been increased over the last century and this will continue due to the rising of human population. Indicators suggest that an additional 70% of food production is required for the next fifty years. The food supply chain or the decrease in agricultural productivity is facing many challenges, an important one of which is plant disease (103). The movement of goods around the world and open borders have facilitated the diffusion of plant diseases through the movement of pathogens into new areas. Furthermore, climate change has helped the emergence or re-emergence of new pathogens which negatively affect food

production (104). Many control methods have been used in order to set up disease management plans and face food shortages, including plant breeding (resistance variety) programs. However, these resistant varieties have faced failure due to the emergence of new variant strains that overcome these resistance genes. These strains have developed resistance to pesticides or antibiotics (105). It is very important to understand how a pathogen emerges and adapts to the environment. Population genetic analysis could be the best tool to understand the previous two concepts (106). Knowledge of population genetics has been increased by coalescent theory, computational methods, and molecular biology, which have opened up the genomic era (107).

Several advantages are obtained from the population genetic analysis. for example, it can helps to identify the source of infection of the plant pathogen and to determine if the pathogen has been introduced or has emerged locally (108). It also reveals the genetic patterns and evolutionary process (109), and another advantage, it helps in understanding the biologically relevant genetic variation within individuals and the effect of evolutionary forces and their contributions within and among the population through space and time (110). Population genetics aims to explain how genetic variation is maintained within and among the population over space and time (111). The information retrieved from these studies will help the creation of plant disease management strategies, through the understanding of disease epidemiology, ecology, biology and evolution (106,112).

The first step in the analysis of population genetics is to understand the population's genetic structure: how they form, how they vary and differences among the strains or populations using molecular markers (113). The population consists of a group of individuals of the same species living in a given area, interacting with each other and impacted by natural selection and evolution (113). The population is considered as the smallest unit of evolution (114). The discussion here will be limited to bacteria. Bacteria live in community as a population, they vary in their phenotypic and genotypic traits, and the study of population genetics of this organism can provide insight about the origin and dispersal of the bacterial pathogens (115). Population structure studies aim to understand how the pathogen evolved and the forces driving evolution, and apply this knowledge to achieve sustainable plant disease management (116).

1.7 Genetic variation

New developments in genomics allow discovery of the genome components. Whole-genome sequences (WGS) and comparative analysis of bacterial genomes have revealed two

types of genes within the genome: the core and flexible genomes (117,118). These findings have revealed more about genetic variation among bacterial species that seem homogeneous but are heterogeneous at the cell and genetic levels. In addition, these findings have helped plant breeders to develop resistant cultivars (119,120).

The natural changes in the genetic information of a population or species are known as genetic variation. These variations help individuals whose genes give rise to new characteristics that are best adapted to the environment and will be the fittest to survive, reproduce and transfer these variations to the next generations, in a process known as evolution (121). Genetic variation occurs due to the interaction of different evolutionary forces, including mutations, horizontal gene transfer or gene flow, random genetic drift , natural selection and other environmental and human activities (122).

When bacteria reproduce by binary fission, all individuals within a species are identical "clones" because they originate from a single parent (wild-type) (123). Bacteria must ensure the transfer and maintenance of the genetic information with the correct sequence to their offspring. Different mechanisms are involved in the organization of this process (2). One mechanism that leads to the emergence of new descents through vertical gene transfer is de novo mutation and the accumulation of mutations over generations (124).

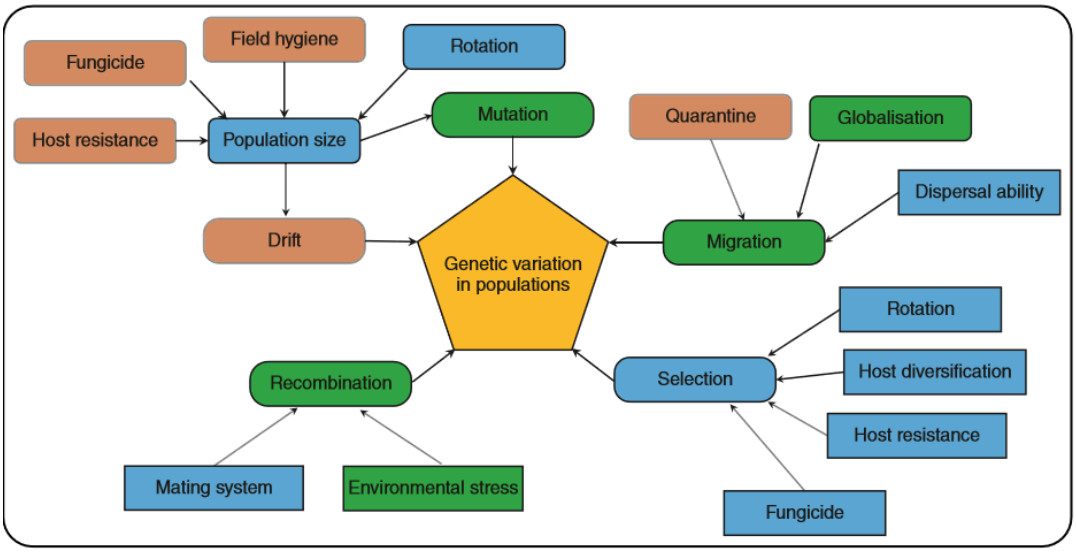


FIGURE 2: Circumstances involved in the generation and maintenance of the genetic variation of the pathogen population. green, brown and blue indicate a positive, negative and variable (positive or negative) impact on the evolution of genetic variation, retrieved from (116).

A bacterial species with greater genetic variation has greater adaptability to changing environments (38). A great variety is observed with bacterial genomes, and these variations are due to the acquisition and loss of functional accessory genes by the interaction of the evolutionary forces (125). Lower genetic variability among the population decreases the fitness and adaptability of the population, while greater genetic variability helps the population to be more adaptable, and can help pathogens to evolve more rapidly to challenge different circumstances, including control measures (126). For example, when a new variant of the same pathogen appears and overcomes a resistant cultivar in one region, it then starts to spread to new regions or areas. The development of virulence genes that overcome resistance via mutation or other mechanisms in the local bacterial population is evolution and causes genetic variation (38). The new virulence strain can move into other local areas or countries as a result of natural and/or human causes that spread disease. This concept also applies to the use of chemicals or antibiotics, to which pathogens can develop resistance due to evolution (127,128).

1.8 Evolution

Population genetics answers the questions about what drives genetic variety and how it is maintained over time and explains the reasons for heterogeneity in bacteria (113). This is because population genetics studies the genetic structure (frequencies of alleles and genotypes) within and among the population to understand how evolution forces occur and act on populations (114). There is no clear picture of the essence of evolution because of the ongoing dispute regarding as to whether the concept of evolution is microevolution or macroevolution,

and as to which one leads to the evolution of species (129). As a definition, evolution means changes in the inherited characteristics of a population over time, as over a long time (thousands of years) which leads to the development of new species, and over a short time (days weeks, months, years) which leads to the development of new variants through interactions of different processes, in order to adapt and to be more suited to the environment (130).

The processes of evolution are mutations, recombination and horizontal gene transfer, genetic drift and natural selection; all these processes have interacted alone or with each other and led to evolution (131).

Macroevolution is the production of a new species from previously existing species and involves morphological innovations and ecological transitions (132). The term refers to the development of new species, genera, family or clades or their extinction through the long-term process under wide environmental changes (133). It focuses on phenotypic evolutionary trends over geological time. The mechanisms that lead to macroevolutionary transitions are an unsolved issue (134).

There are two schools that deal with these concepts. The first Raia *et al*., 2015 holds that micro-macro evolution has the same process to develop above and within-species level (135). The other school (136) basically explains that the process which produces macroevolution is different from the process of microevolution and that microevolution is not sufficient to produce macroevolution. Richard, 1940 has set three mechanisms involved in the development of new species; generation and sorting of variation in addition to natural selection. For bacteria the same ongoing argument regards the mechanisms responsible for microevolution or macroevolution (137); some authors (138) defend the idea that microevolution is responsible for macroevolution via a long-time process, while others (139,140) disagree, stating that macroevolution cannot be based on the processes of microevolution. (141,142) reported that qualitative change resulted in macroevolution through a long-term process. It is suggested that horizontal gene transfer among taxa higher than the species level could occur and lead to macroevolution (133). To sum up, there is a gap in understanding of the mechanisms of macroevolution, and it seems that the difference concerns only the time periods involved in the two processes (142).

It is not my intention to discuss the mechanisms involved in the process of macroevolution for two reasons; firstly, my topic focuses on microevolution, and secondly, the

topic has been discussed by other authors in the past (143). Three approaches have described the evolution of bacteria: age Inferred by association with ecological events, age inferred from eukaryotic molecular clocks and age inferred from the host fossil record (136,137,151–154,139,144–150). The long-time period required by macroevolution makes it impossible to study its mechanisms in the laboratory, which leaves a big gap in understanding of this system (155).

Microevolution is the process involving changes in the gene pool that occur during a short period of time, i.e. days, weeks or months even years. These changes, which occur in the accessory genes, can lead to fast phenotypic changes and new bacterial variants; the final consequences of these processes are the development of epidemic diseases (133). Microevolution in bacteria is responsible for the rapid emergence of variants (132,156,157). Whole-genome sequences (WGS) and comparative analysis of bacterial genomes have revealed two types of genes within the genome: the core and flexible genomes (140,158–161). The core genes are essential for the survival of the organism, whereas flexible genes consist of genes responsible for adaptation to specific niches, hosts, or environments. Core genes are evolved in mutation and natural selection but rarely in horizontal gene transfer and are considered as the raw material for evolution, while the flexible genome evolves largely through horizontal genetic exchange. Flexible genomes have an important role from the biological point of view; they are involved in pathogenicity, virulence features such as adhesins, capsules or toxins, and developing resistance to chemicals or antibiotics (118,162). Genes responsible for these factors normally lie within the flexible gene pool or plasmids of the bacterial genome. This can offer advantages under particular conditions due to human activity and environmental conditions, whereas the evolutionary forces take advantage and play their role mainly on these groups of genes (163–165).

The major driving forces for novel genetic variants in bacteria are mutations, horizontal gene transfer and recombination, genetic drift and natural selection. These mechanisms continually help bacteria to rapidly generate a new variants (118,166). Horizontal gene transfer of genes lost by mutation is thought to be the primary mechanisms of prokaryotic adaptation leading to speciation. Since a very great number of DNA sequences between 10 and 100 kilobases (kb) in length are present in the genome termed genomic islands, a huge number of genes in the genomic islands are present too. These genes may be transferred and recombined all together in a new bacterial genome, which may lead to changes in the genome functions and

evolution in important leaps (140). In the study conducted by (118,167) on *Salmonella enterica*, the result showed that a significant amount of the DNA segment is from distantly related species which indicates lateral gene transfer involvement in evolution. All of the evolutionary forces will be discussed below.

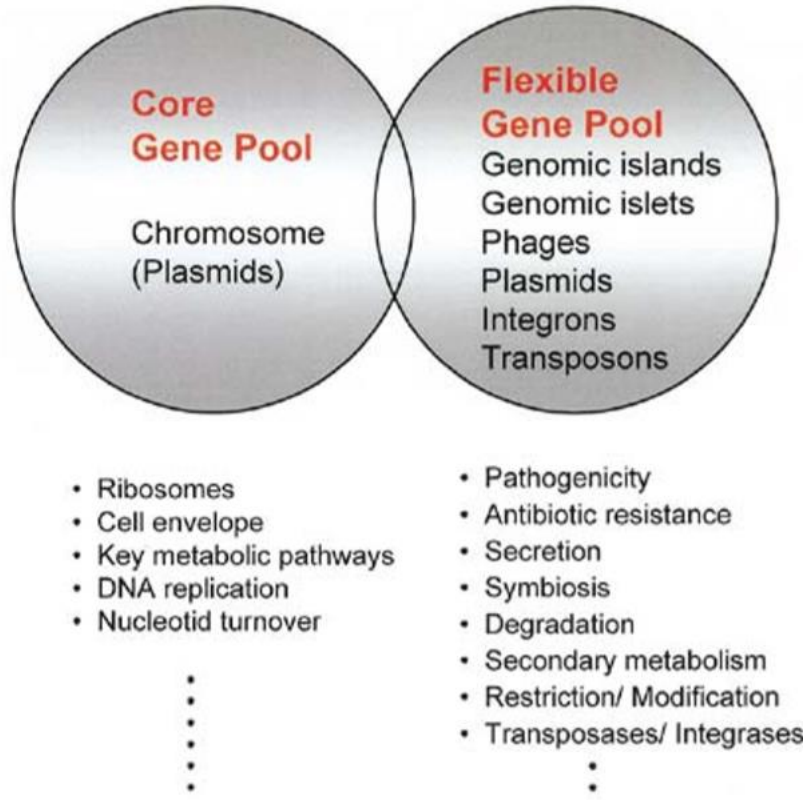


FIGURE 3: Mode of the DNA pools in the genomes of prokaryotes, retrieved from (168).

1.9 Evolutionary forces

1.9.1. Mutation

This is the primary source for genetic variation and raw material of evolution (118). A mutation is a change occurring in DNA nucleotide sequences structures that are transmitted from parent to offspring (169,170). There are three types of mutation based on their effect on genomes: the first type of mutation can increase the fitness of the host, the second can decrease its fitness (harmful), and the third has no effect on its fitness. Mutations can result from substitution, deletion, insertion, inversion, reciprocal translocation and chromosomal

rearrangements of DNA sequences (171). A small portion of mutation has a beneficial advantageous effect, while the majority has a deleterious effect (172). Point mutation and DNA rearrangements (inversions, duplications, insertions, deletions, or transposition of large sequences of DNA from one location of a bacterial chromosome or plasmid to another can occur during DNA replications, and could on an evolutionary time-scale lead to a modification in the genetic map (173). Mutation rates vary among bacterial genomes between species, or within the strains of the same species, and haploid pathogens differ from diploid pathogens in the change of the mutation numbers (171,174). Mutation of a single nucleotide could cause the protein to lose or change its function to a new function (175).

These mutations can lead to the loss of genetic functions or entire genes, or add a new function in the case of duplication, for example (176). Mutation alone cannot have a great effect on the evolutionary process unless it is combined with other mechanisms. When a mutation occurs, natural selection decides whether to keep or eliminate this mutation from the population (170,177).

Most mutations are harmful to the bacteria, while others can provide advantages under the right conditions and based on natural selection (178). Mutation rate variability effects evolutionary adaptation. In a study Lopreato, 2001 on *E. coli*, it was found that evolutionary adaptation accelerated when the mutation rate increased (179).

A point mutation can be generated by slipped strands mispairing, which positively or negatively affects gene expression; one of its advantages is a contribution to the development of bacterial resistance to antibiotics (180). The modification of structural and regulatory genes in bacterial species that are opportunistic or non-primary pathogens have the ability to modify or knock out the encoded proteins and influence their function (pathoadaptive mutagenesis) during pathogenesis (140). A pathoadaptive mutation is another microevolutionary concept, which enables some bacteria to be pathogenic without modification or loss of pre-existing genes that had been adapted for life as a commensal organism (181). Deletion is another microevolutionary process within mutation, in which bacteria can adapt to the environment through the manipulation of large DNA sequences (181). The role of mutation in plant pathogenic bacteria has been clarified and confirmed in different studies (162,182), *etc*. Within the concept of mutation, a very important DNA motif called tandem repeats (TRs) plays an important role in the evolution of bacteria. These motifs developed due to DNA replication

slippage and their mutation rates are higher compared to point mutations (183–187). This aspect will be discussed later on this chapter.

1.9.2. **Horizontal gene transfer (HGT), gene flow**

HGT is the movement and incorporation of genetic material between closely related organisms or distantly related organisms (188). This process can either moderate or deepen genetic variability among subpopulations or strains in the case of bacteria inhabiting the same environment (189–191). This mechanism is considered the most important mechanism in bacterial evolution (187). Horizontal gene transfer among and within bacterial populations occurs via the release of naked DNA from the community (eukaryotic, prokaryotic) followed by uptake and recombination by the bacteria (192). Three different mechanisms are responsible for HGT: natural transformation, transduction, and conjugation. In order to accomplish the transfer of DNA sequences, mobile genetic elements such as plasmids, bacteriophages, transposons, and integrons are responsible for the transfer of DNA among bacterial populations(193). Many factors could limit HGT such as the establishment, expression, and temporospatial functions of the agents affecting DNA movement: plasmids, bacteriophages, and transposons (194,195).

HGT plays a major role in the evolution of most prokaryotes, alongside other mechanisms. It is responsible for resistance genes, gene clusters encoding biodegradative pathways and pathogenicity determinants, and is responsible for speciation and sub-speciation in bacteria (194). In addition to genome size, other factors like growth temperature, oxygen utilization, and pathogenicity may affect HGT (196–199). One study revealed that HGT is the driving force responsible for much of the genetic variation among prokaryotic microbes (191). HGT and recombination have formed the diversity of the bacterial population and the variation of their lineages. While HGT has introduced novel DNA distinct bacterial genome, the latter mediates the exchange of DNA among the individuals of the same species (200). The HGT mechanism depends on the recombination and incorporation of genes and genetic elements from strains of bacterial species and eukaryotic species (189,201). In HGT recombination, the bacteria achieved recombination through the same processes; in transformation, bacteria can acquire DNA from the neighbouring environment, which causes bacterial DNA modification; in conjunction, bacteria transfer DNA sequences by direct contact, and in transduction, the involvement of a bacteriophage is required for the bacteria to be able to exchange DNA. These three processes help to achieve HGT (189,193). A study conducted recently showed the

involvement of HGT in a wide range of antibiotic resistance in bacteria (Zhan, 2016). It was found that different numbers of genes were shared between archaea and the bacteria that live in similar ecological conditions, due to HGT (196,197,199). Different There are several advantages derived from recombination: it can achieve purging deleterious mutations, the new combination (multilocus genotypes) helps the population to adapt, and it reduces or eliminates linkage disequilibrium (191). Additionally, it increases the diversity of bacterial clonality, and/because recombination work to homogenize the species gene pool(202). Recombination could disturb HGT, especially on the gene on plasmids. It can substitute for the acquisition of adaptations from both close and distant species (203). Different plant pathogenic bacteria have witnessed HGT and recombinant (198).

Gene flow or gene migration refers to the transfer of genetic variation from one geographic population to another population, which could happen between two populations of the same species or between two different species through horizontal gene transfer (204–207). Gene flow plays two important roles in evolution: the first is to increase the range of a species, and the second is to limit the genetic divergence of populations that would occur by random genetic (208).

1.9.3. **Genetic drift**

This is the process in which there is a probability that newly generated mutants can survive and exist in the population. In other words, the mutants occurring in a large population, which are constant and usually quite low, are able to remain and persist, which then leads to more genetic variability than in a small population (209).

The random change in alleles frequency in a population occurs by mutants/mutations (genetic sampling error) over a short period of time. It is more effective in small populations (210). The influence of genetic drift is limited by the number of individuals that survive to be able to transmit the modified genes to their offspring, and also by the size of the population: drift occurs faster in small populations. At the same time, the effect of genetic drift will decline over time with an increasing population (211). Genetic drift may lead to a reduction in genetic variation (due to the disappearance of the gene variants), and to genetic differentiation among populations (209,212). Immigration, emigration, founder effects, and population bottlenecks are responsible for genetic drift. A population bottleneck occurs when the size of the population decreases sharply due to external factors, and the alleles fluctuation is varied. The founder effect is a type of population bottleneck and occurs when a few individuals migrate out of the

population to establish a new subpopulation (126). The replacement rate of a gene is constant between mutation and genetic drift. When a new allelic variation is created by a mutation in DNA, the genetic drift works slowly to remove it in order to create a stable state (208). Allele frequencies can be changed due to mutation, recombination and genetic drift (213). Genetic drift plays a crucial role in forming the genetic structure of a pathogen in a local population due to the environmental condition surrounding the pathogen. Control measures can reduce the rate of evolution by decreasing the effective population size (212).

Genetic drifts affect bacterial genomes through the fixation of harmful mutations (113). It promotes genome reduction in bacteria. and education of bacterial genome size is a consequence of genetic drift coupled with a mutational bias toward deletions (126). Genetic drift segregates a bacterial gene pool into well-defined regions, and reduce the genetic diversity in bacteria (214). It has played an important role in the diversification of bacteria such as *Pyrococcus, Mycobacterium tuberculosis* (215).

1.9.4. **Natural selection**

Natural selection is a model of how evolution works or what makes the organisms adapt to their environment. It plays an important role in genetic variations, and it also shapes the phenotypic, biological, and behavioral adaptations of bacteria across generations (216,217).

This concept was described by Darwin, an English naturalist, with the famous phrase "descents with modifications" or survival of the fittest (fitness means the organisms possessing the best characteristics for that particular environment will survive (218). This scientific theory explains the process that causes a change in the characteristics of organisms over time to be more suited to their environment reproduce and disperse, which could lead in turn after a long time period to the evolution of a new species (209,219,220). The interaction between natural selection and evolutionary forces can result in evolution (221).

Natural selection results from several interacting steps; under certain pressure conditions the presence of variation among descendants appears due to evolutionary processes, such as mutation or HGT, the number of individuals who carry the variations traits should be pass to the offspring through vertical gene transfer, the reproduction of the variants varies among strains, the successful variations accumulate during time and disperse to the survival progeny that can continue and reproduce under that certain conditions the parents escape through which in turn give the best fit for the environment (122). Natural selection explains the predictability

of one genotype to be fit for the pressure of the environment; it sometimes leads to the origin of new species, besides the evolution of the species by adapting to their environment (209). When newly-generated genotypes have harmful fitness effects, natural selection will work to eliminate them from the population, but if the generated genotype is neutral, there will be no effect on the fate of these variants (222). These mechanisms have an impact on the population which can lead a single population to be divided into two subpopulations, different from each other intolerance to the environmental conditions, pathogenicity, resistance to chemicals, invasion, *etc.*, but this new subpopulation can share a common core genome (223). An example of natural selection on bacteria is resistant to antibiotics and pesticides. When the harmful bacteria are exposed to antibiotics a resistant bacterium can survive for multiple reasons like incomplete dosage or mutations. The surviving bacteria can reproduce and increase over time by natural selection, which is the strongest evidence of Darwin (224). According to natural selection, bacteria can suffer from resistance cost; when bacteria developed resistance in the case of extreme use of antibiotics, this resistance will be eliminated when the use of antibiotic stops (222,225). During the last decade, many studies have shown that bacterial resistance to antibiotics does not belong to the mutation & natural selection mechanisms, but it belongs to different means; for example, it comes from the selection of an existing genetic trait, from gene transfer, or genetic loss due to mutation (226). Natural selection is involved in the limitation of neutral genetic diversity across many species (227–229). According to (218), molecular variability among species is caused by random drift of mutant genes, and not by Darwinian selection. A new phenotype will be subject to natural selection forces, which will decide if the new phenotype stays or is discarded from the genome (122). The natural selection theory cannot completely explain the fact of evolution (132).

In conclusion, the above-mentioned processes play an important role in evolution. Mutation, recombination and horizontal gene transfer can increase genetic variability by adding or losing new DNA sequences. While genetic drift decreases genetic variation, natural selection plays an important role by increasing or decreasing the genetic variation based on the generated variation (230). The emergence of plant disease especially bacteria results from DNA modifications due to the evolutionary process. In any case, these processes are vulnerable to environmental conditions, which can shape the rearrangement of the earlier pathogenicity determinant or the new ones (113). Understanding how genetic variation is maintained within and among the population over space and time will help the establishment of sustainable plant disease control (187). Molecular markers can solve the problem. Molecular markers need to

be chosen appropriately to be cheap, fast, reasonably polymorphic, reproducible, and provide insights at the right evolutionary scale (106). The following pages will discuss this concept.

1.10 Genotyping and Molecular methods

Plant disease management strategies are based on crucial steps. One of these steps is understanding the ecology and biology of plant pathogens (105), the other essential step in setting up active control measures is surveillance (monitoring and early detection) of the disease and its causal agents (99). The ability to quickly and reliably differentiate among related bacterial isolates is essential for epidemiological surveillance (Sankaran et al., 2010). These methods and the estimations of their strength are based on a knowledge of the genetic and phenotypic diversity of plant pathogens (231).

Bacterial pathogens have evolved from a small number of species that have differences in their phenotypic and genotypic traits. These species are relatively close and able to cause disease in different host plants (232). These species have caused disease due to the emergence of new variants as a result of evolution or are newly discovered pathogens (233). Evolutionary forces have helped these species to adapt and resist the environmental conditions, which in turn allowed the pathogen to evolve and re-emerge, causing severe diseases (234). These species are comprised of strains from one lineage with typical features, and they are similar to each other, or differ in some of their genotypic and/or phenotypic traits (235). Moreover, a group of strains based on special features can form one of the following groups: biovars, serovars, phagovars, and pathovars (236). The results of the driving evolutionary forces are genetic diversity among bacterial strains or populations (237).

Genotyping methods can help to understand the genetic diversity in the population structure of a pathogen. Molecular markers need to be chosen appropriately to be cheap, fast, reasonably polymorphic, reproducible, and provide insights at the right evolutionary scale (233). Understanding the recent divergence in bacterial pathogen requires markers with high mutation rates, such as tandem repeats (105). The other loci in mitochondrial, nuclear, that have slow mutation rates can provide a more distant evolutionary history (238–240).

Different methods have been used in the past: morphological, biochemical, nutritional, and immune tests were used to differentiate among strains, and this has led to the species concept (110,241,242). There are a number of practical limitations that make the previous methods unsuitable for bacterial population structure studies and dynamics also for scientific

purpose is less determined, but very critical (243). These methods have several drawbacks: they are usually labor-intensive and time-consuming and provide false-positive and false-negative results; they are prone to errors, and unable to distinguish among closely related strains or track their origin; it is difficult to compare the behavior of different pathogens, and they are highly variable or unstable (244). For these reasons, researchers have worked hard to find easier alternative methods. Methods can easily apply, require less time and labor and are even less expensive. The golden era of detection and identification or differentiation among strains starts with a nucleic acid-based technique. The first DNA based method used among *Ps. syringae* strains was a DNA-DNA hybridization method (245–247). Subsequently, it became necessary to identify a dependable method that can accurately discriminate between strains (248). Molecular typing methods are fast and powerful tools that enable us to differentiate closely related strains (249).

Genotyping or discrimination among closely related strains based on their genetic bases is very important for plant pathogen detection, the study of population genetic structure (250), evolutionary history and host specificity (251), epidemiology studies (outbreak investigation, causal agents, transmission, and surveillance) (250), and taxonomy structure (223). This information will aid substantially in setting up plant disease management (233). Genotyping started with random PCR amplification-based methods. The identified genetic profile of each strain is different, based on the method which gives a specific result as a fingerprint (252).

Many methods have been applied in the past twenty-five years for genotyping of human, animal, and plant pathogenic bacteria. These methods are applied in order to quickly and accurately discriminate the related strains and/or understand the source of infection and how do they spread (253). Van Belkum, 1999 has identified three classes of bacterial genotyping methods: DNA banding pattern-based genotyping methods show that strains differ in their yield size of DNA bands by amplification of genomic DNA using restriction enzymes (REs)(254). DNA sequencing-based genotyping methods are based on the differences between the obtained DNA sequences after PCR amplification. Finally, DNA hybridization-based methods discriminate bacterial strains by analyzing the hybridization of their DNA to probes of known sequences. In other words, these methods are either based on restriction enzymes analysis of the total genome or depend on DNA segments amplified by PCR (253).

The most common and widely applied methods in plant pathogenic bacteria are pulsed-field gel electrophoresis (PFGE) and ribotyping, Repetitive sequence-based PCR (Rep-PCR)

(255). Random amplification of polymorphic DNA (RAPD) (256), restriction fragment length polymorphism PCR-(RFLP) (257), Amplified Fragment Length Polymorphism (AFLP) (258), Multilocus sequence typing (MLST) (259), SNPs-based genotyping (260), and Whole Genome Sequence (WGS) (261). Due to our studies of the three diseases (Olive knot, Olive Quick Decline Syndrome, and Canker of Kiwi) the disease caused by *Pseudomonas savastanoi*, *Xylella fastidiosa* and *Pseudomonas syringae* pv. *actindiae* the methods used with the three diseases will be discussed below.

Methods applied to genotyping and population studies of *Pseudomonas savastanoi* include SDS-PAGE (262,263), restriction fragment length polymorphism (RFLP) (264), (ERIC, BOX, and REP) (265), fluorescent amplified fragment length polymorphism (f-AFLP) analysis (266,267), Random amplification of polymorphic DNA (RAPD) (268), High-*Resolution* Melting Analysis (HRMA) (269), rep-PCR and MLST (270).

While the methods used for the differentiation and genotyping of *Pseudomonas syringae* pv. *actindiae* are rep-PCR, IS50-PCR and RAPD, from a wide collection of Italian, Korean and Japanese strains (271), Genome sequencing and SNP analysis (272), Comparative genomic analyses (273), Multilocus sequence analysis (MLSA) of housekeeping, type III effector and phytotoxin genes were also used for differentiation (274,275).

Lastly, methods used to genotype and differentiate *X. fastidiosa* are (RFLP)(276), Random amplification of polymorphic DNA (RAPD) , combination of different methods such (RFLP), (RAPD), ERIC and REP(277), PCRs were used in later years (278), Short Sequence repeats (SSR) applying a variable number of tandem repeats analysis (VNTR) (279,280), MLST/MLSA Multilocus sequence typing analysis (251,281–283), Combination of SNPs and SSR marker methods (284–286), and whole-genome analysis (287).

These methods have been described intensively and applied to different human, animal, environmental bacterial pathogens in addition to plant pathogenic bacteria. Each method has its own positive and negative aspects. For example, methods based on restriction enzymes, such as PCR-(RFLP), and (PFGE) are time-consuming and labor-intensive, show lower resolution than DNA- DAN hybridization among related strains, and require a large amount of genomic DNA. The results of these methods cannot be exchanged among laboratories (288–290). AFLP also requires a large amount of high-quality genomic DNA and it shows limited genetic diversity (258,291). RAPD methods have shown low reproducibility and sensitivity to

other factors involved in PCR procedures (292). MLST is the most widely applied method; despite its accuracy, reproducibility, and portability, MLST has faced some challenges in its results for monomorphic bacteria, where it shows low resolution for this type of pathogen due to the low sequence diversity in these genes. Therefore, it fails to resolve the evolutionary pattern of plant pathogenic bacteria, since many bacterial pathogens are monomorphic (293). Whole-genome sequencing (WGS) and SNPs methods have provided more details about genome properties but are expensive and still require advances in next-generation sequencing technologies, or knowledge in computational biology and facing infrastructural constraints (294). As it appears from this literature, the methods vary in their reproducibility, accuracy, applicability, and technical problems. They can be expensive, time-consuming and laborious, and not comparable across the laboratories which are different from what it is needed.

In order to provide and compare three different diseases with three different times of emergence, one as endemic (Olive knot) and two as epidemic (Canker of Kiwifruit and olive quick decline syndrome), the molecular marker should be carefully chosen, and the method should be independent, reasonably polymorphic, reproducible, less expensive, less laborious, portable and easy to compare across laboratories. It should also provide an accurate evolutionary scale (295,296). Short tandem repeats (DNA motifs) that have high mutation rates (103-109) can provide understandings into recent divergence (105,113). Within bacterial populations, a promising tool is Multiple Locus Variable Number of Tandem Repeats (MLVA) which benefits from sequencing technology and computer analysis. The method was developed for *Haemophilus influenzae* strains and bacterial human pathogens (238–240) and will be discussed next.

In the past, different methods have been used to detect and identify the plant pathogens, such as cultural, biochemical, physiological and immunological assays. While the identification of the pathogens was based on specific morphological features, biochemical and immune tests led to the identification and assignment of the pathogens to the species level (297–299). The polymerase chain reaction (PCR) has replaced a wide range of traditional methods and become the cornerstone for bacterial detection and identification. In addition to the normal PCR, advanced PCR methods have been invented and have been used widely for detecting plant pathogens. These methods include reverse-transcription PCR (RT-PCR), Multiplex PCR, Real-time PCR, Co-operational polymerase chain reaction (Co-PCR), and nested-PCR. Many

reviews have been carried out for these methods and other Nucleic acid-based methods (300–302).

1.11 Multiple Locus Variable-number Tandem Repeat Analysis (MLVA)

The whole-genome sequences analysis era and the development of computer science and software have made it possible to discover repetitive DNA sequences (104,300,303–307). These DNA repetitive sequences were found distributed throughout the eukaryotic and prokaryotic genomes (308). TRs have different locations within the genome, in intergenic regions, coding regions or pseudogenes (309).

These repeats were first discovered in eukaryotes in 1960 (310). In addition to their presence in multiple copies, they have differences in other features such as size, location, complexity, and repeat style (311). Different criteria were used to classify these repetitive DNA sequences, such as their functional role, degree of repetitiveness, and how they group together (312). TRs are classified based on their organization on the genome into two types. There are interspersed repeats, in which these repeat units are scattered in the genome in a random way (313). Interspersed repeats are transposable elements which include the transposable elements, short interspersed elements retrogenes, and retropseudogenes, and the differences in genome size are generated by this type of repeats (314). Then there are Tandem Repeats (TRs), where the repetitive sequences are aligned in an array one after one, and these include gene families, microsatellite, minisatellite, satellite DNAs and mixes of both repeats, such as segmental duplications (SDs) and duplicated DNA fragments (315). Tandem repeats (TRs) or satellite DNA are DNA sequences repeated in sequences from head to tell manner and with a different number of repeats in each locus of each individual (2,313). DNA sequences (short nucleotides) that form a unit, and these units are repeated in array several times in the locus within the genome. These sequences are divided into three types based on the size of the repeated unit are divided into three types: microsatellites, mini-satellites, and macro-satellites or mega-satellites. It is not clear whether microsatellites are from 1-6 Base pairs (bp) or from 1-9bp, as different studies have identified microsatellites from the size of the repeated unit from 1bp to 6bp nucleotide long, and also called them Short TRs, microsatellites, simple sequence repeats (SSRs) (316). Other studies have identified the microsatellite from 1-9bp for each unit size. For minisatellites, the size of the unit is from ten nucleotides and until to 100bp (317,318). The macro-satellite or mega-satellite contains more than 100bp per unit (312,315,319,320). Furthermore, TRs can be placed in two categories degenerated and identical TRs. Degenerated

TRs only in which one nucleotide difference in the unit that arising from point mutation. On the other hand, in identical TRs the nucleotides match each other in all of the units(321,322).

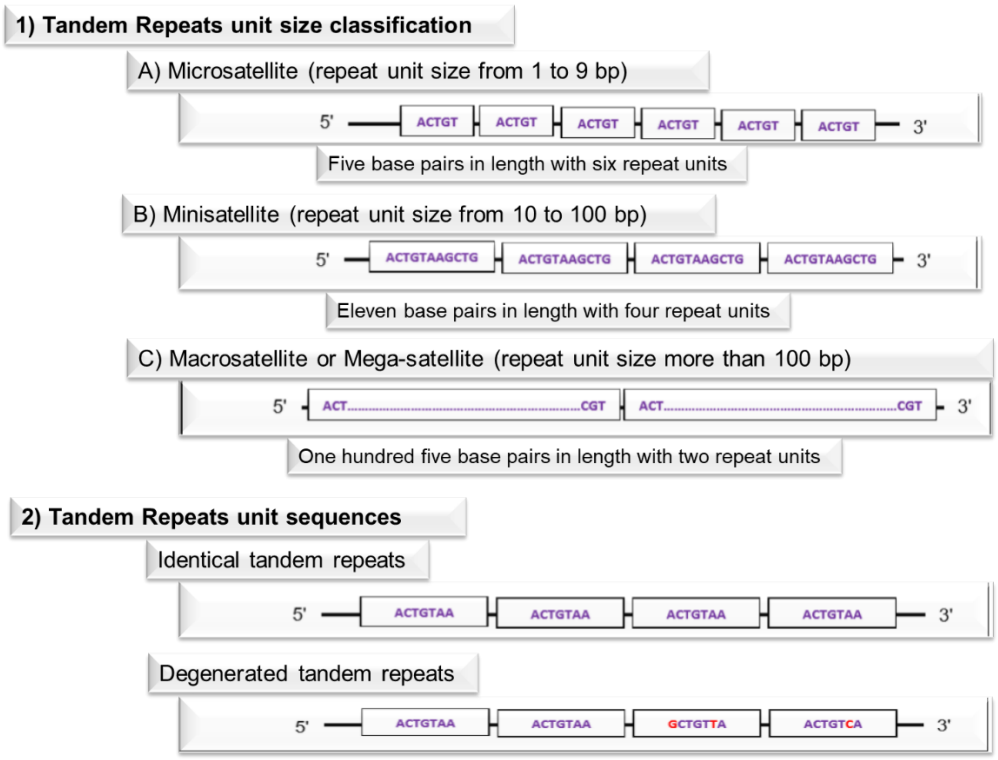


FIGURE 4: Tandem repeats classifications and types

Among these microsatellites, or simple sequence repeats (STRs) are considered of most important for their abundance comparing to minisatellite, in addition, the numbers of repeats within specific STRs tend to be highly variable, their easiness to be analyzed by PCR (308,322). Weber and Wong have found out the variability among these TRs in human genomes and identify the possibility of using these TRs as a genetic marker (323). Because of their high variability, they are also referred to as a variable number of tandem repeats (VNTRs) (324). One of the main characteristics of these tandem repeats are the instability, they mutate in very high-speed 10^{-2}- 10^{-7} higher than regular mutation rates (point mutation) (297,319). The mutations are happened mainly due to (deletion, addition) in the repeat units, while the mutation within the repeat units is rare (311). In general, there are three mechanisms that explained TRs mutation in eukaryotes: recombination, Retro-transposition mechanism, Strand-slippage replication (315), other studies have focused on only two mechanisms: recombination and DNA polymerase slippage (311). In prokaryotes, the mechanisms for the TRs mutation are

the same which includes DNA strand slippage and their effect of contraction and expansion, recombination, replication, and a variety of DNA repair pathways (325–327).

1.11.1. **Identification of tandem repeats**

The development of genome sequencing, bioinformatics, and computer science has helped the identification and analysis of repeats. Different algorithms have been used to identify these TRs. These algorithms are varying in their ability to detect different types of TRs (319,325,328–330). The problem is the ability of these algorithms to detect the different types of TRs, therefore the use of different types of algorithms is recommended (331). Furthermore, in order to obtain the right number and consensus TRs, there are three important parameters should be applied while using these algorithms: alignment weights, type of repeats, and threshold scores. These parameters strongly affect the results of TRs (308). Various methods and algorithms have been applied to detect TRs, these methods have been reviewed in (332). There are five main and Widely applied algorithms TRF, Mreps, Sputnik, STAR, and Repeat Masker (332–334). We have used Tandem Repeat Finder which is widely used (TRF the most applied algorithms) (335). TRF program shows more flexibility with the easiness of handling the parameters to obtain the exact results (336).

1.11.2. **The role of tandem repeats**

TRs were thought to be useless DNA, due to their high variability and their sequence simplicity (332). Studies in the last 25 years have proved the importance of these repeats. The role that tandem repeats come from their location in the genome. Analysis of the genomes has clarified, when they fall with the coding region, regulatory sequences they affect gene expression (337). TRs are found within promoters of stress-induced genes and within the coding regions of genes encoding cell-surface and regulatory proteins. The changes in these repeats are accompanied by phenotypic changes (319). A different role has been identified for TRs, In addition to their role in genetic mapping, genotyping, and forensic studies (315). In any case, the abundance of repetitive sequences in eukaryotes is not meaningless. Since the 1990s, different studies have shown the role of TRs in human disease. More than 20 diseases have been found in humans related to TRs, hereditary neurological diseases including fragile X syndrome, Huntington's disease, myotonic dystrophy, *etc*. (338). Furthermore, it was found that TRs interfere with cellular function (329,339). In animals, an association was found between tandem repeats and canine epilepsy disease (340). In other eukaryotes, such as yeast, fluctuation in gene size which leads to quantitative alterations in phenotype are associated with

TRs (341). In the plant Arabidopsis thaliana, STRs can influence gene and protein function, which leads to phenotypic variability (310). Also, a correlation was found between STR diversity, and ecological factors in wheat and barley (342,343). TRs were found to play an important role in animal disease (344,345). More details about Tandem repeats and their roles have been reviewed (341,346). The first study on bacteria was on the *H. influenzae* phase variation gene in 1989 (319). Phase variation is one of the bacterial adapting strategies, thus the bacteria will be able to face the changing environment by modulating the function of specific genes (347). The repeats are involved in different cell roles such as disease development, cellular differentiation, genome stability, transcriptional control, and evolution (308,348). TRs abundancy could speed the loss of gene order, through chromosome rearrangement (320). TRs play a role in bacterial pathogenicity and genomic variability (Rocha, 2003). In the study conducted by (340,349). TRs were found to play a role in genetic variation within a population, and they facilitate evolutionary changes. Many other studies have been conducted to reveal the importance of TRs, especially microsatellites and their role in bacteria, in the last two decades (343). An important role of TRs is that they can reveal the recent divergence in the microevolutionary level due to the high mutation rates (318, 331, 350, 351). Three genetic variations have been used for comparative genomics studies: Single Nucleotide Polymorphisms (SNPs), Copy Number Variations (CNVs) and Tandem Repeats (TRs) (238,240). As the number of these tandem repeats in a particular locus varies from strain to strain, they are known as variable-number tandem repeats or VNTRs (320). The particularity of the TRs used in this process is falling under the umbrella of variability in TRs length, sequence, and position of each strain (352,353).

1.11.3. **The MLVA method**

MLVA or Multilocus Variable-Number Tandem-Repeat Analysis is a molecular method based upon the calculation of Variable copy Numbers of Tandem Repeats (VNTR). These regions are the most variable regions in bacterial genomes and therefore have the potential to resolve the genetic diversity of monomorphic bacterial pathogens (309,354). The first step in this method is the identification of the TRs from *in silico* programs, according to certain criteria. Secondly, in order to calculate these TRs, PCR primers must be designed from the conserved region around these TRs, the left and right flanking regions. This is followed by sample preparation and DNA extraction for each sample, then PCR amplification. Finally, the differences among the strains should be measured by gel electrophoresis and capillary electrophoresis, in which the results translated numerically, after the calculation of these loci by

detracting the right and the left flanking regions from the amplicons from PCR and dividing the remaining length by the size of the corresponding repeat unit in that locus. The results obtained from both approaches are imported into sophisticated software for the assignment of allele numbers. A string of alleles is created from the number of repeat units for each locus forming the MLVA profile and is eventually used to assign an MLVA type. The generated numeric data can be exchanged among laboratories around the world via the public MLVA database (223,312).

Its great ability to discriminate the strains based on differences in the number of the repeats in each locus generated by amplification of the genomic DNA of several loci results in a high level of polymorphism (322,355). In addition, MLVA is a rapid procedure for monitoring short-term, local outbreaks of bacterial pathogens, revealing insights about the relationships at a microevolutionary level (253).

MLVA methodology was primarily developed for *Haemophilus influenzae* strains and bacterial human pathogens (322), and then for bacterial animal pathogens (298,299,352). MLVA was applied for the first time in subtyping the plant pathogenic bacteria *Xylella fastidiosa* (356), and since then has been used to subtype a wide range of plant pathogenic bacteria, as shown in the box below.

TABLE I: USE OF MLVA METHODS FOR SUBTYPING BACTERIAL PLANT PATHOGENS

Authors	*Year*	*N of VNTR Loci*	*Pathogen species*	*Host plant species*
Della Coletta-Filho et al.	2001	8	*Xylella fastidiosa*	*Citrus sinensis; Coffea arabica*
Ngoc et al.	2009	14	*Xanthomonas citri* pv. *citri*	Citrus spp.
Bergsma-Vlami et al.	2012	6	*Xanthomonas arboricola* pv. *pruni*	*Prunus laurocerasus*
Gironde and Manceau	2012	8	*Pseudomonas syringae* pv. *maculicola; Pseudomonas syringae* pv. *tomato*	*Brassicaceae fam.; Lycopersicon esculentum*
Zhao et al.	2012	25	*Xanthomonas oryzae* pv. *oryzicola*	*Oryza sativa*
N'Guessan et al.	2013	26	*Ralstonia solanacearum*	*Solanaceae fam*
Zaluga et al.	2013	8	*Clavibacter michiganensis subsp. michiganensis*	*Lycopersicon esculentum*
Pruvost et al.	*2014*	*31*	*Xanthomonas citri* pv. *citri*	*Citrus spp.*
Vernière et al.	*2014*	*14*	*Xanthomonas citri* pv. *citri*	*Citrus spp.*
Bühlmann et al	*2014*	*6*	*Erwinia amylovora*	*Pomaceae fam*
Ciarroni et al.	2015	13	*Pseudomonas syringae* pv. *actinidiae*	*Actinidia deliciosa*
Guinard et al.	2017	7	*Ralstonia solanacearum*	Mix populations

1.11.4. **Use of MLVA**

The method has great potential as an easy and effective tool, not only in the recognition and schedule/detection of the presence of different bacterial types worldwide but above all to trace their movements on a local to the international scale, supporting the simple detection of contaminated materials with key information concerning specific haplotypes (251). The first MLVA scheme for a plant pathogenic bacterium was developed for *Xylella fastidiosa* in the early 2000s. The method provided a high-resolution tool for epidemiological, genetic, and ecological analysis of citrus-specific *Xylella fastidiosa* strains; this (357) made it possible to obtain high resolution and robust genotyping of *Xanthomonas citri* pv. citri (251). (358) reported the usefulness of the method in population structures and epidemiological monitoring, in addition to fast, reliable, and cost-effective molecular typing. MLVA was reported to be a very promising first-line assay for large-scale routine genotyping prior to whole-genome sequencing (Zhao et al., 2012a). It was reported as a very useful tool for gaining insight into geographic diversity, and for understanding the dynamic evolution of the pathogen (357). MLVA is a promising typing technique for local surveillance and outbreak investigation in epidemiological studies; it can unravel the intra-pathovar structure.

1.12 References

1. Agrios GN. Introduction to plant pathology. Elsevier Acad Press Publ. 2005;

2. Atwood T, Campbell P, Parish H, Smith T, Stirling J, Vella F, et al. Oxford Dictionary of Biochemistry and Molecular Biology. Oxford Univ. Press, Oxford; 2006.

3. Cramer GR, Urano K, Delrot S, Pezzotti M, Shinozaki K. Effects of abiotic stress on plants: a systems biology perspective. BMC Plant Biol. 2011;11(1):163.

4. Lucas GB, Campbell L. Introduction to plant diseases: identification and management. Springer Science & Business Media; 2012.

5. Mikkelsen L, Elphinstone J, Jensen DF. Literature review on detection and eradication of plant pathogens in sludge, soils and treated biowaste. Desk study bulk density Bruxelles Eur Comm DG RTD under Framew. 2006;6.

6. Janse JD, Obradovic A. *Xylella fastidiosa* : its biology, diagnosis, control and risks. J Plant Pathol. 2010;S35–48.

7. Strange RN, Scott PR. Plant disease: a threat to global food security. Annu Rev Phytopathol [Internet]. 2005;43:83–116. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16078878

8. Singh HB. Management of Plant Pathogens with Microorganisms. Proc Indian Natl Sci Acad. 2014;80(2):443–54.

9. Schumann GL, D'Arcy CJ. Hungry planet: stories of plant diseases. APS Press; 2012.

10. Lucas JA. Advances in plant disease and pest management. J Agric Sci. 2010;149(S1):91–114.

11. Janse JD. Phytobacteriology: principles and practice. Cabi; 2005.

12. Wen Z, Zhang J-R. Bacterial Capsules. 2015;33–53.

13. Moens S, Vanderleyden J. Functions of bacterial flagella. Crit Rev Microbiol. 1996;22(2):67–100.

14. Jackson RW. Plant pathogenic bacteria: genomics and molecular biology. Horizon Scientific Press; 2009.

15. janse jd. emerging bacterial diseases of fruit trees and some other crops that are or may become a threat for southern europe: notes on epidemiology, risks, prevention and management on first occurrence. pred ref. 2011;9.

16. Angert ER. Alternatives to binary fission in bacteria. Nat Rev Microbiol [Internet]. 2005;3(3):214. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15738949

17. Sobiczewski P. Bacterial diseases of plants: Epidemiology, diagnostics and control. Zemdirbyste. 2008;95:151–7.

18. Kumar J, Singh D, Ghosh P, Kumar A. Endophytic and Epiphytic Modes of Microbial Interactions and Benefits. Plant-Microbe Interact Agro-Ecological Perspect. 2017;227–53.

19. Feng J. Recent advances in taxonomy of plant pathogenic bacteria. Sci Agric Sin. 2017;50(12):2305–14.

20. Mansfield J, Genin S, Magori S, Citovsky V, Sriariyanum M, Ronald P, et al. Top 10 plant pathogenic bacteria in molecular plant pathology. Mol Plant Pathol [Internet]. 2012;13(6):614–29. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22672649

21. Pautasso M, Petter F, Rortais A, Roy A-S. Emerging risks to plant health: a European perspective. Health (Irvine Calif). 2015;50:52.

22. Zhan J, Thrall PH, Burdon JJ. Achieving sustainable plant disease management through evolutionary principles. Trends Plant Sci. 2014;19(9):570–5.

23. Álvarez B, Biosca EG, López MM. On the life of *Ralstonia solanacearum*, a destructive bacterial plant pathogen. Curr Res Technol Educ Top Appl Microbiol Microb Biotechnol. 2010;1:267–79.

24. Commission F. Report on the national survey to assess the presence of bleeding canker of horse chestnut trees in Great Britain. Edinburgh, UK: Forestry Commission. 2008.

25. Webber JF, Parkinson NM, Rose J, Stanford H, Cook RTA, Elphinstone JG. Isolation and identification of *Pseudomonas syringae* pv. *aesculi* causing bleeding canker of horse chestnut in the UK. Plant Pathol. 2008;57(2):368.

26. Rodriguez-Moreno L, Jimenez AJ, Ramos C. Endopathogenic lifestyle of *Pseudomonas savastanoi* pv. *savastanoi* in olive knots. Microb Biotechnol [Internet]. 2009;2(4):476–88. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21255279

27. Ferguson AR. Kiwifruit: evolution of a crop. In: VII International Symposium on Kiwifruit 913. 2010. p. 31–42.

28. Takikawa Y, SERIZAWA S, ICHIKAWA T, TSUYUMU S, GOTO M. *Pseudomonas syringae* pv. *actinidiae* pv. nov. Japanese J Phytopathol. 1989;55(4):437–44.

29. Fang Y, Zhu X, Wang Y. Preliminary studies on kiwifruit diseases in Hunan Province. Sichuan Fruit Sci Technol. 1990;18:28–9.

30. Marcelletti S, Scortichini M. Definition of Plant-Pathogenic Pseudomonas Genomospecies of the *Pseudomonas syringae* Complex Through Multiple Comparative Approaches. Phytopathology [Internet]. 2014;104(12):1274–82. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24875383

31. Ferrante P, Scortichini M. Redefining the global populations of *Pseudomonas syringae* pv. *actinidiae* based on pathogenic, molecular and phenotypic characteristics. Plant Pathol. 2015;64(1):51–62.

32. Sicard A, Zeilinger AR, Vanhove M, Schartel TE, Beal DJ, Daugherty MP, et al. *Xylella fastidiosa* : Insights into an Emerging Plant Pathogen. Annu Rev Phytopathol [Internet]. 2018;56(1):181–202. Available from: https://www.annualreviews.org/doi/abs/10.1146/annurev-phyto-080417-045849

33. Almeida RPP. ECOLOGY. Can Apulia's olive trees be saved? Science (80-) [Internet]. 2016;353(6297):346–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27463657

34. Puławska J, Gétaz M, Kałuzna M, Kuzmanović N, Obradović A, Pothier JF, et al. 15 Bacterial Diseases. Cherries Bot Prod Uses. 2017;365.

35. Scortichini M. Epidemiology and predisposing factors of some major bacterial diseases of stone and nut fruit trees species. J Plant Pathol. 2010;S73–8.

36. Kannan VR, Bastas KK. Sustainable Approaches to Controlling Plant Pathogenic Bacteria. CRC Press; 2015.

37. Silva Dias J. World importance, marketing and trading of vegetables. In: XXVIII International Horticultural Congress on Science and Horticulture for People (IHC2010): International Symposium on 921. 2010. p. 153–69.

38. He D, Zhan J, Xie L. Problems, challenges and future of plant disease management: from an ecological point of view. J Integr Agric. 2016;15(4):705–15.

39. Maloy OC. Plant disease management. plant Heal Instr. 2005;10:1–11.

40. FAO. International Code of Conduct on the Distribution and Use of Pesticides. Food Agric Organ United Nations, Rome, Italy. 2002;

41. Grondeau C, Samson R, Sands DC. A Review of Thermotherapy to Free Plant Materials from Pathogens, Especially Seeds from Bacteria. CRC Crit Rev Plant Sci. 2011;13(1):57–75.

42. Seguí EM, González MML. Métodos de control de las bacteriosis. In: Patología vegetal. Phytoma-España; 1996. p. 653–78.

43. Dollwet HH. Historic uses of copper compounds in medicine. Trace Elem Med. 1985;2(pk):80–7.

44. Yunis H, Bashan Y, Okon Y, Henis Y. Weather dependence, yield losses and control of bacterial speck of tomato caused by Pseudomonas tomato. Plant Dis. 1980;64:937–9.

45. Wimalajeewa DLS, Cahill R, Hepworth G, Schneider HG, Washbourne JW. Chemical control of bacterial canker (*Pseudomonas syringae* pv. *syringae)* of apricot and cherry in Victoria. Aust J Exp Agric. 1991;31(5):705–8.

46. Aldwinckle HS, Bhaskara Reddy M V, Norelli JL. Evaluation of control of fire blight infection of apple blossoms and shoots with SAR inducers, biological agents, a growth regulator, copper compounds, and other materials. In: IX International Workshop on Fire Blight 590. 2001. p. 325–31.

47. Balestra GM, Bovo M. Effectiveness of copper compounds in the control of bacterial diseases of kiwifruit plants. Acta Hortic. 2003;

48. Teviotdale BL, Krueger WH. Effects of timing of copper sprays, defoliation, rainfall, and inoculum concentration on incidence of olive knot disease. Plant Dis. 2004;88(2):131–5.

49. Sayler RJ, Kirkpatrick BC. The effect of copper sprays and fertilization on bacterial canker in French prune. Can J Plant Pathol. 2003;25(4):406–10.

50. Cooksey DA. Plasmid-determined copper resistance in *Pseudomonas syringae* from impatiens. Appl Environ Microbiol. 1990;56(1):13–6.

51. Vallad GE, Pernezny KL, Balogh B, Wen A, Figueiredo JFL, Jones JB, et al. Comparison of kasugamycin to traditional bactericides for the management of bacterial spot on tomato. HortScience. 2010;45(12):1834–40.

52. Mantovi P, Bonazzi G, Maestri E, Marmiroli N. Accumulation of copper and zinc from liquid manure in agricultural soils and crop plants. Plant Soil. 2003;250(2):249–57.

53. Pietrzak U, McPhail DC. Copper accumulation, distribution and fractionation in vineyard soils of Victoria, Australia. Geoderma. 2004;122(2–4):151–66.

54. Nagajyoti PC, Lee KD, Sreekanth TVM. Heavy metals, occurrence and toxicity for plants: a review. Environ Chem Lett. 2010;8(3):199–216.

55. Cervantes C, Gutierrez-Corona F. Copper resistance mechanisms in bacteria and fungi. FEMS Microbiol Rev. 1994;14(2):121–37.

56. Van Der Zwet T. Fire blight: its nature, prevention, and control: A practical guide to integrated disease management. ok; 1991.

57. McManus PS, Stockwell VO, Sundin GW, Jones AL. ANTIBIOTIC USE IN PLANT AGRICULTURE. Annu Rev Phytopathol [Internet]. 2002;40:443–65. Available from: http://www.ncbi.nlm.nih.gov/pubmed/12147767

58. Goodman RN. Fire blight control with sprays of agri-mycin, a streptomycin–terramycin combination. Plant Dis Rep. 1954;38(citaion):874–8.

59. Burr TJ, Norelli JL, Katz B, Wilcox WF, Hoying SA. Streptomycin resistance of *Pseudomonas syringae* pv. *papulans* in apple orchards and its association with a conjugative plasmid. Phytopathology. 1988;78(4):410–3.

60. Debener T, Lehnackers H, Arnold M, Dangl JL. Identification and molecular mapping of a single Arabidopsis thaliana locus determining resistance to a phytopathogenic *Pseudomonas syringae* isolate. Plant J. 1991;1(3):289–302.

61. Ritchie DF, Dittapongpitch V. Copper-and streptomycin-resistant strains and host differentiated races of *Xanthomonas campestris* pv. *vesicatoria* in North Carolina. Plant Dis. 1991;75:733–6.

62. McManus PS, Jones AL. Epidemiology and genetic analysis of streptomycin-resistant *Erwinia amylovora* from Michigan and evaluation of oxytetracycline for control. Phytopathol. 1994;(ok).

63. Stockwell VO, Duffy B. Use of antibiotics in plant agriculture. 2012;

64. Pal KK, Gardener BM, McSpadden Gardener B. Biological Control of Plant Pathogens. Plant Heal Instr. 2006;

65. Aboutorabi M. A Review on the Biological Control of Plant Diseases using Various Microorganisms. J Res Med Dent Sci. 2018;6(4):30–5.

66. Singh RP, Jha PN. Plant growth promoting potential of ACC deaminase rhizospheric bacteria isolated from Aerva javanica: a plant adapted to saline environments. Int J Curr Microbiol App Sci. 2015;4(7):142–52.

67. Amusa NA, Odunbaku OA. Biological control of bacterial diseases of plants in Nigeria: problems and prospects. Res J Agric Biol Sci. 2007;3:979–82.

68. Mota MS, Gomes CB, Souza Júnior IT, Moura AB. Bacterial selection for biological control of plant disease: criterion determination and validation. brazilian J Microbiol. 2017;48(1):62–70.

69. Quimio AJ, Ayo AL. Biological Control of Tobacco Bacterial Wilt with Avirulent Bacteriocin Producer Strain of *Pseudomonas solanacearum*. 7th Int. In: Conf Plant Path Bact Budapest. 1989.

70. Krishnamurthy K, Gnanamanickam SS. Biocontrol of rice sheath blight with formulated *Pseudomonas putida*. Indian Phytopathol. 1998;51(3):233–6.

71. Köhl J, Postma J, Nicot P, Ruocco M, Blum B. Stepwise screening of microorganisms for commercial use in biological control of plant-pathogenic fungi and bacteria. Biol Control. 2011;57(1):1–12.

72. Nguyen MT, Ranamukhaarachchi SL, Hannaway DB. Efficacy of antagonist strains of *Bacillus megaterium*, *Enterobacter cloacae*, *Pichia guilliermondii* and *Candida ethanolica* against bacterial wilt disease of tomato. J Phytol. 2011;

73. Roselló G, Bonaterra A, Francés J, Montesinos L, Badosa E, Montesinos E. Biological control of fire blight of apple and pear with antagonistic *Lactobacillus plantarum*. Eur J plant Pathol. 2013;137(3):621–33.

74. Sundin GW, Castiblanco LF, Yuan X, Zeng Q, Yang CH. Bacterial disease management: challenges, experience, innovation and future prospects: Challenges in bacterial molecular plant pathology. Mol Plant Pathol [Internet]. 2016;17(9):1506–18. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27238249

75. Daranas N, Roselló G, Cabrefiga J, Donati I, Francés J, Badosa E, et al. Biological control of bacterial plant diseases with *Lactobacillus plantarum* strains selected for their broad-spectrum activity. Ann Appl Biol. 2019;174(1):92–105.

76. Mauch-Mani B, Métraux J-P. Salicylic acid and systemic acquired resistance to pathogen attack. Ann Bot. 1998;82(5):535–40.

77. Romero AM, Kousik CS, Ritchie DF. Resistance to bacterial spot in bell pepper induced by acibenzolar-S-methyl. Plant Dis. 2001;85(2):189–94.

78. Maxson-Stein K, He S-Y, Hammerschmidt R, Jones AL. Effect of treating apple trees with acibenzolar-S-methyl on fire blight and expression of pathogenesis-related protein genes. Plant Dis. 2002;86(7):785–90.

79. Obradovic A, Jones JB, Momol MT, Balogh B, Olson SM. Management of tomato bacterial spot in the field by foliar applications of bacteriophages and SAR inducers. Plant Dis. 2004;88(7):736–40.

80. Van Loon LC, Bakker P. Induced systemic resistance as a mechanism of disease suppression by rhizobacteria. In: PGPR: Biocontrol and biofertilization. Springer; 2005. p. 39–66.

81. Ferraz HGM, Resende RS, Moreira PC, Silveira PR, Milagres EA, Oliveira JR, et al. Antagonistic rhizobacteria and jasmonic acid induce resistance against tomato bacterial spot. Bragantia. 2015;(AHEAD):0.

82. Mourgues F, Brisset M-N, Chevreau E. Strategies to improve plant resistance to bacterial diseases through genetic engineering. Trends Biotechnol. 1998;16(5):203–10.

83. Bariana HS. Diseases| breeding for disease resistance. 2003;

84. Pink DAC, Hand P. Plant resistance and strategies for breeding resistant varieties. PLANT Prot Sci. 2003;38:9–14.

85. Medeiros JGS, Citadin I, Santos I dos, Assmann AP. Reaction of peach tree genotypes to bacterial leaf spot caused by *Xanthomonas arboricola* pv. *prunis*. Sci Agric. 2011;68(1):57–61.

86. van Bueren ETL, Jones SS, Tamm L, Murphy KM, Myers JR, Leifert C, et al. The need to breed crop varieties suitable for organic farming, using wheat, tomato and broccoli as examples: a review. NJAS-Wageningen J Life Sci. 2011;58(3–4):193–205.

87. Horvath DM, Stall RE, Jones JB, Pauly MH, Vallad GE, Dahlbeck D, et al. Transgenic resistance confers effective field level control of bacterial spot disease in tomato. PLoS One. 2012;7(8):e42036.

88. Huet G. Breeding for resistances to *Ralstonia solanacearum*. Front Plant Sci. 2014;5:715.

89. Janse JD, Wenneker M. Possibilities of avoidance and control of bacterial plant diseases when using pathogen-tested (certified) or-treated planting material. Plant Pathol. 2002;51(5):523–36.

90. Janse JD. Control of Bacterial Diseases Through Pathogen Freedom of Planting Material. In: Plant Pathogenic Bacteria. Springer; 2001. p. 6–18.

91. Singh VK, Pandey P. Physical methods in management of Plant diseases. B Ecofriendly Innov approaches Plant Dis Manag VK Singh, Y Singh A Singh, Int B Distrib Publ New Delhi. 2012;21–9.

92. Ogle HJ, Dale M. Disease management: cultural practices. Plant Pathog plant Dis. 1997;390–404.

93. Stall RE, Loschke DC, Jones JB. Linkage of copper resistance and avirulence loci on a self-transmissible plasmid in *Xanthomonas campestris* pv. *vesicatoria*. Phytopathology. 1986;76(2):240–3.

94. Bender CL, Cooksey DA. Molecular cloning of copper resistance genes from *Pseudomonas syringae* pv. *tomato*. J Bacteriol. 1987;169(2):470–4.

95. Manulis S, Zutra D, Kleitman F, Dror O, David I, Zilberstaine M, et al. Distribution of streptomycin-resistant strainsof *Erwinia amylovora* in Israel and occurrence of blossom blight in the autumn. Phytoparasitica. 1998;26(3):223–30.

96. Graham JH, Leite Jr RP. Lack of control of citrus canker by induced systemic resistance compounds. Plant Dis. 2004;88(7):745–50.

97. Kousik CS, Ritchie DF. Disease potential of pepper bacterial spot pathogen races that overcome the Bs2 gene for resistance. Phytopathology. 1996;

98. Gassmann W, Dahlbeck D, Chesnokova O, Minsavage G V, Jones JB, Staskawicz BJ. Molecular evolution of virulence in natural field strains of *Xanthomonas campestris* pv. *vesicatoria*. J Bacteriol. 2000;182(24):7053–9.

99. Martinelli F, Scalenghe R, Davino S, Panno S, Scuderi G, Ruisi P, et al. Advanced methods of plant disease detection. A review. Agron Sustain Dev. 2014;35(1):1–25.

100. Ciancio A, Mukerji KG. Integrated management of plant pests and diseases. Springer; 2007.

101. Gadoury DM, MacHardy WE, Rosenberger DA. Integration of pesticide application schedules for disease and insect control in apple orchards of the northeastern United States. Plant Dis. 1989;

102. Scala V, Pucci N, Loreti S. The diagnosis of plant pathogenic bacteria: a state of art. Front Biosci (Elite Ed). 2018;10:449–60.

103. Bock CH, Poole GH, Parker PE, Gottwald TR. Plant disease severity estimated visually, by digital photography and image analysis, and by hyperspectral imaging. CRC Crit Rev Plant Sci. 2010;29(2):59–107.

104. Fang Y, Ramasamy RP. Current and Prospective Methods for Plant Disease Detection. Biosens [Internet]. 2015;5(3):537–61. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26287253

105. Grünwald NJ, Goss EM. Evolution and population genetics of exotic and re-emerging pathogens: novel tools and approaches. Annu Rev Phytopathol. 2011;49:249–67.

106. Gillespie JH. Population genetics: a concise guide. JHU Press; 2004.

107. Grunwald NJ, Everhart SE, Knaus BJ, Kamvar ZN. Best Practices for Population Genetic Analyses. Phytopathology [Internet]. 2017;107(9):1000–10. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28513284

108. Hartl DL, Clark AG. Molecular population genetics. Princ Popul Genet. 2007;317(ok):384.

109. Grünwald NJ, Flier WG. The biology of *Phytophthora infestans* at its center of origin. Annu Rev Phytopathol. 2005;43.

110. Goss EM, Larsen M, Chastagner GA, Givens DR, Grünwald NJ. Population genetic analysis infers migration pathways of *Phytophthora ramorum* in US nurseries. PLoS Pathog. 2009;5(9):e1000583.

111. Milgroom MG, del Mar Jiménez-Gasco M, García CO, Drott MT, Jiménez-Díaz RM. Recombination between clonal lineages of the asexual fungus *Verticillium dahliae* detected by genotyping by sequencing. PLoS One. 2014;9(9):e106740.

112. Amorium A, Maloy S, Hughes K. Population Gentics . 2nd ed. Vol. 5, Brenner's Encyclopedia of Genetics. Elsevier : Academic Press; 2013.

113. Zhan J. Population Genetics of Plant Pathogens. ELS John Wiley Sons Ltd, Chichester http//www els net [doi 101002/9780470015902 a0021269]. 2016;1–7.

114. Hastings A. Population biology: concepts and models. Springer Science & Business Media; 2013.

115. Jonckers LHM. The concept of population in biology. Acta Biotheor. 1973;22(2):78–108.

116. Spratt BG, Maiden MCJ. Bacterial population genetics, evolution and epidemiology. Philos Trans R Soc London B Biol Sci. 1999;354(1384):701–10.

117. Zhan J, Thrall PH, Papaix J, Xie L, Burdon JJ. Playing on a pathogen's weakness: using evolution to guide sustainable plant disease control strategies. Annu Rev Phytopathol. 2015;53:19–43.

118. Hacker J, Carniel E. Ecological fitness, genomic islands and bacterial pathogenicity: a Darwinian view of the evolution of microbes. EMBO Rep. 2001;2(5):376–81.

119. Schuster ML, Coyne DP. Genetic variation in bean bacterial pathogens. Euphytica. 1975;24(1):143–7.

120. Mira A, Martín-Cuadrado AB, D'Auria G, Rodríguez-Valera F, Martin-Cuadrado AB, D'Auria G, et al. The bacterial pan-genome:a new paradigm in microbiology. Int Microbiol [Internet]. 2010;13(2):45–57. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20890839

121. Neik TX, Barbetti MJ, Batley J. Current status and challenges in identifying disease resistance genes in Brassica napus. Front Plant Sci. 2017;8:1788.

122. Kimura M. The neutral theory of molecular evolution. Sci Am. 1979;241(5):98–129.

123. Zhan J, McDonald BA. The interaction among evolutionary forces in the pathogenic fungus Mycosphaerella graminicola. Fungal Genet Biol. 2004;41(6):590–9.

124. Sniegowski P. Mismatch repair: Origin of species? Curr Biol. 1998;8(2):R59–61.

125. Darmon E, Leach DRF. Bacterial genome instability. Microbiol Mol Biol Rev. 2014;78(1):1–39.

126. Kuo C-HH, Moran NA, Ochman H. The consequences of genetic drift for bacterial genome complexity. Genome Res [Internet]. 2009;19(8):1450–4. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19502381

127. Burdon JJ. Genetic variation in pathogen populations and its implications for adaptation to host resistance. In: Durability of disease resistance. Springer; 1993. p. 41–56.

128. Bahri B, Kaltz O, Leconte M, de Vallavieille-Pope C, Enjalbert J. Tracking costs of virulence in natural populations of the wheat pathogen, Puccinia striiformis f. sp. tritici. BMC Evol Biol. 2009;9(1):26.

129. Rimoin DL, Pyeritz RE, Korf B. Emery and Rimoin's essential medical genetics. Elsevier; 2013.

130. Kokić T. Is there a unique process which governs macroevolution? Nov Prisut časopis za Intelekt i Duh pitanja. 2015;13(3):319–33.

131. Hall B, Strickberger MW. Strickberger's evolution. Jones & Bartlett Learning; 2008.

132. Morschhäuser J, Köhler G, Ziebuhr W, Blum-Oehler G, Dobrindt U, HACKER J. Evolution of microbial pathogens. In: The Activities Of Bacterial Pathogens In Vivo: Based on Contributions to a Royal Society Discussion Meeting. World Scientific; 2000. p. 349–73.

133. Bell G. Experimental macroevolution. Proc R Soc B Biol Sci [Internet]. 2016;283(1822):20152547. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26763705

134. Huneman P. Macroevolution and Microevolution: Issues of Time Scale in Evolutionary Biology. Time Nat Nat Time. 2017;326:315–57.

135. Raia P, Passaro F, Carotenuto F. Cope's rule and the universal scaling law of ornament complexity. Am Nat. 2015;186(2):165–75.

136. Dobzhansky T. Genetics and the origin of species. NY: Columbia Univ. Press. 364 p. 1937;

137. Richard G. The material basis of evolution. New Haven, Yale University Press; 1940.

138. Jablonski D. Approaches to Macroevolution: 1. General Concepts and Origin of Variation. Evol Biol [Internet]. 2017;44(4):427–50. Available from: http://www.ncbi.nlm.nih.gov/pubmed/29142333

139. Vislobokova IA. The concept of macroevolution in view of modern data. Paleontol J. 2017;51(8):799–898.

140. Ziebuhr W, Ohlsen K, Karch H, Korhonen T, Hacker J. Evolution of bacterial pathogenesis. Cell Mol Life Sci C. 1999;56(9–10):719–28.

141. Spetner LM. Not by Chance!: Shattering the Modern Theory of Evolution. ok: Judaica Press; 1998.

142. Hartman H. Macroevolution, catastrophe and horizontal transfer. In: Horizontal Gene Transfer (Second Edition). Elsevier; 2002. p. 421–5.

143. Provorov NA, Vorobyov NI. Macroevolution of symbiosis as self-organization of superspecies system controlled by natural selection. Biol Bull Rev. 2013;3(4):274–85.

144. Wright S. The material basis of evolution. Sci Mon. 1941;53(2):165–70.

145. Gould SJ, Eldredge N. Punctuated equilibria: the tempo and mode of evolution reconsidered. Paleobiology. 1977;3(2):115–51.

146. Jablonski D. Micro- and macroevolution: Scale and hierarchy in evolutionary biology and paleobiology. Paleobiology. 2000;26:15–52.

147. Gould SJ. The structure of evolutionary theory. Cambridge, MA: Harvard Univ. Press; 2002.

148. Grantham T. Is macroevolution more than successive rounds of microevolution? Palaeontology. 2007;50(1):75–85.

149. Jablonski D. Species selection: Theory and data. Annu Rev Ecol Evol Syst. 2008;39:501–24.

150. Plutynski A. Speciation and macroevolution. 2008;

151. Bell MA. Evolution since Darwin: the first 150 years. ok; 2010.

152. Raia P, Carotenuto F, Passaro F, Fulgione D, Fortelius M. Ecological specialization in fossil mammals explains Cope's rule. Am Nat. 2012;179:328–37.

153. Serrelli E, Gontier N, Futuyma DJ. Macroevolution: Explanation, Interpretation and Evidence. 2015;2.

154. Gontier N. Uniting Micro- with Macroevolution into an Extended Synthesis: Reintegrating Life's Natural History into Evolution Studies. In: Macroevolution. Springer; 2015. p. 227–75.

155. Ochman H, Elwyn S, Moran NA. Calibrating bacterial evolution. Proc Natl Acad Sci. 1999;96(22):12638–43.

156. Achtman M. Microevolution and epidemic spread of serogroup A Neisseria meningitidis–a review. Gene. 1997;192(1):135–40.

157. Jordan IK, Rogozin IB, Wolf YI, Koonin E V. Microevolutionary genomics of bacteria. Theor Popul Biol. 2002;61(4):435–47.

158. Coupat B, Chaumeille-Dole F, Fall S, Prior P, Simonet P, Nesme X, et al. Natural transformation in the *Ralstonia solanacearum* species complex: number and size of DNA that can be transferred. FEMS Microbiol Ecol. 2008;66(1):14–24.

159. Lovell HC, Mansfield JW, Godfrey SAC, Jackson RW, Hancock JT, Arnold DL. Bacterial evolution by genomic island transfer occurs via DNA transformation in planta. Curr Biol. 2009;19(18):1586–90.

160. Green S, Studholme DJ, Laue BE, Dorati F, Lovell H, Arnold D, et al. Comparative genome analysis provides insights into the evolution and adaptation of *Pseudomonas syringae* pv. *aesculi* on Aesculus hippocastanum. PLoS One. 2010;5(4):e10224.

161. He M, Sebaihia M, Lawley TD, Stabler RA, Dawson LF, Martin MJ, et al. Evolutionary dynamics of *Clostridium difficile* over short and long time scales. Proc Natl Acad Sci. 2010;107(16):7527–32.

162. Hacker J, Blum-Oehler G, Mühldorfer I, Tschäpe H. Pathogenicity islands of virulent bacteria: structure, function and impact on microbial evolution. Mol Microbiol. 1997;23(6):1089–97.

163. Sarkar SF, Guttman DS. Evolution of the core genome of *Pseudomonas syringae*, a highly clonal, endemic plant pathogen. Appl Env Microbiol [Internet]. 2004;70(4):1999–2012. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15066790

164. McBride SM, Coburn PS, Baghdayan AS, Willems RJL, Grande MJ, Shankar N, et al. Genetic variation and evolution of the pathogenicity island of *Enterococcus faecalis*. J Bacteriol [Internet]. 2009;191(10):3392–402. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19270086

165. Silva C, Wiesner M, Calva E, Calv E. The Importance of Mobile Genetic Elements in the Evolution of Salmonella: Pathogenesis, Antibiotic Resistance and Host Adaptation. In: Salmonella-A Diversified Superbug. InTech; 2012.

166. Zhou Z, McCann A, Litrup E, Murphy R, Cormican M, Fanning S, et al. Neutral genomic microevolution of a recently emerged pathogen, *Salmonella enterica* serovar Agona. PLoS Genet. 2013;9(4):e1003471.

167. Pitman AR, Jackson RW, Mansfield JW, Kaitell V, Thwaites R, Arnold DL. Exposure to host resistance mechanisms drives evolution of bacterial virulence in plants. Curr Biol. 2005;15(24):2230–5.

168. Lan R, Reeves PR. Gene transfer is a major factor in bacterial evolution. Mol Biol Evol. 1996;13(1):47–55.

169. Bergman J. Mutations and Evolution. CEN Tech J. 1995;9(2):146–54.

170. Opal SM, Pop-Vicas A. Molecular mechanisms of antibiotic resistance in bacteria. In: Mandell, Douglas, and Bennett's Principles and Practice of Infectious Diseases (Eighth Edition). Elsevier; 2015. p. 235-251. e3.

171. Saitou N. Introduction to evolutionary genomics. Springer; 2013.

172. Bartlett JL. Towards a creationary classification of mutations. Answers Res J. 2009;2:169–74.

173. Eyre-Walker A, Keightley PD. The distribution of fitness effects of new mutations. Nat Rev Genet [Internet]. 2007;8(8):610–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17637733

174. Roth JR, Benson N, Galitski T, Haack K, Lawrence JG, Miesel L. Rearrangements of the bacterial chromosome: formation and applications. *Escherichia coli,* Salmonella Cell Mol Biol. 1996;2:2256–76.

175. Baer CF, Miyamoto MM, Denver DR. Mutation rate variation in multicellular eukaryotes: causes and consequences. Nat Rev Genet. 2007;8(8):619.

176. Clancy S. Genetic mutation. Nat Educ. 2008;1(1):187.

177. Wahab T. PCR and sequencing based methods for detection and typing of pathogenic microorganisms. 2011;

178. Hershberg R. Mutation--The Engine of Evolution: Studying Mutation and Its Role in the Evolution of Bacteria. Cold Spring Harb Perspect Biol [Internet]. 2015;7(9):a018077. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26330518

179. Lopreato J. Biology's Influence on Sociology: Human Sociobiology. 2001;

180. Sprouffske K, Aguílar-Rodríguez J, Sniegowski P, Wagner A. High mutation rates limit evolutionary adaptation in *Escherichia coli*. PLoS Genet. 2018;14(4):e1007324.

181. Sokurenko E V, Hasty DL, Dykhuizen DE. Pathoadaptive mutations: gene loss and variation in bacterial pathogens. Trends Microbiol. 1999;7(5):191–5.

182. Birch A, Häusler A, Rüttener C, Hütter R. Chromosomal deletion and rearrangement in Streptomyces glaucescens. J Bacteriol. 1991;173(11):3531–8.

183. Wichmann G, Ritchie D, Kousik CS, Bergelson J. Reduced genetic variation occurs among genes of the highly clonal plant pathogen *Xanthomonas axonopodis* pv. *vesicatoria*, including the effector gene avrBs2. Appl Environ Microbiol. 2005;71(5):2418–32.

184. Ma W, Dong FFT, Stavrinides J, Guttman DS. Type III effector diversification via both pathoadaptation and horizontal transfer in response to a coevolutionary arms race. PLoS Genet. 2006;2(12):e209.

185. Trivedi P, Wang N. Host immune responses accelerate pathogen evolution. ISME J. 2014;8(3):727.

186. Bartoli C, Lamichhane JR, Berge O, Varvaro L, Morris CE. Mutability in P seudomonas viridiflava as a programmed balance between antibiotic resistance and pathogenicity. Mol Plant Pathol. 2015;16(8):860–9.

187. Bartoli C, Roux F, Lamichhane JR. Molecular mechanisms underlying the emergence of bacterial pathogens: an ecological perspective. Mol Plant Pathol. 2016;17(2):303–10.

188. Mallet J. Gene flow. In: SYMPOSIUM-ROYAL ENTOMOLOGICAL SOCIETY OF LONDON. 1999. p. 337–60.

189. Ochman H, Lawrence JG, Groisman EA. Lateral gene transfer and the nature of bacterial innovation. Nature. 2000;405(6784):299.

190. Polz MF, Alm EJ, Hanage WP. Horizontal gene transfer and the evolution of bacterial and archaeal population structure. Trends Genet. 2013;29(3):170–5.

191. Fuchsman CA, Collins RE, Rocap G, Brazelton WJ. Effect of the environment on horizontal gene transfer between bacteria and archaea. PeerJ [Internet]. 2017;5:e3865. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28975058

192. Boucher Y, Douady CJ, Papke RT, Walsh DA, Boudreau MER, Nesbø CL, et al. Lateral gene transfer and the origins of prokaryotic groups. Annu Rev Genet. 2003;37(1):283–328.

193. Thomas CM, Nielsen KM. Mechanisms of, and barriers to, horizontal gene transfer between bacteria. Nat Rev Microbiol [Internet]. 2005;3(9):711. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16138099

194. Frost LS, Leplae R, Summers AO, Toussaint A. Mobile genetic elements: the agents of open source evolution. Nat Rev Microbiol [Internet]. 2005;3(9):722. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16138100

195. Olendzenski L, Gogarten JP. Gene transfer: who benefits? In: HorizontalGeneTransfer: GenomesinFlux book. Springer; 2009. p. 3–9.

196. De la Cruz F, Davies J. Horizontal gene transfer and the origin of species: lessons from bacteria. Trends Microbiol. 2000;8(3):128–33.

197. Kay E, Vogel TM, Bertolla F, Nalin R, Simonet P. In situ transfer of antibiotic resistance genes from transgenic (transplastomic) tobacco plants to bacteria. Appl Environ Microbiol. 2002;68(7):3345–51.

198. Wiedenbeck J, Cohan FM. Origins of bacterial diversity through horizontal genetic transfer and adaptation to new ecological niches. FEMS Microbiol Rev [Internet]. 2011;35(5):957–76. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21711367

199. Gyles C, Boerlin P. Horizontally transferred genetic elements and their role in pathogenesis of bacterial disease. Vet Pathol. 2014;51(2):328–40.

200. Gogarten JP, Doolittle WF, Lawrence JG. Prokaryotic evolution in light of gene transfer. Mol Biol Evol. 2002;19(12):2226–38.

201. Lawrence JG. Gene transfer, speciation, and the evolution of bacterial genomes. Curr Opin Microbiol. 1999;2(5):519–23.

202. Witte H. Introduction to Population Biology and Evolution. JSTor; 1982.

203. Guttman DS, Dykhuizen DE. Clonal divergence in Escherichia coli as a result of recombination, not mutation. Science (80-). 1994;266(5189):1380–3.

204. Gabriel DW. Why do pathogens carry avirulence genes? Physiol Mol Plant Pathol. 1999;55:205–14.

205. Didelot X, Maiden MCJ. Impact of recombination on bacterial evolution. Trends Microbiol. 2010;18(7):315–22.

206. Mhedbi-Hajri N, Hajri A, Boureau T, Darrasse A, Durand K, Brin C, et al. Evolutionary history of the plant pathogenic bacterium *Xanthomonas axonopodis*. PLoS One. 2013;8(3):e58474.

207. Wicker E, Lefeuvre P, De Cambiaire J-C, Lemaire C, Poussier S, Prior P. Contrasting recombination patterns and demographic histories of the plant pathogen *Ralstonia solanacearum* inferred from MLSA. ISME J. 2012;6(5):961.

208. Choudhuri S. Bioinformatics for beginners: genes, genomes, molecular evolution, databases and analytical tools. Elsevier; 2014.

209. Milgroom MG. Population biology of plant pathogens: genetics, ecology, and evolution. APS Press; 2015.

210. McDonald BA, Linde C. The population genetics of plant pathogens and breeding strategies for durable resistance. Euphytica. 2002;124(2):163–80.

211. McDonald BA, McDermott JM. Population genetics of plant pathogenic fungi. Bioscience. 1993;43(5):311–9.

212. Masel J. Genetic drift. Curr Biol [Internet]. 2011;21(20):R837–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22032182

213. Honnay O. Genetic Drift. Brenner's Encycl Genet . 2013;251–3.

214. Bobay L-MM, Ochman H. The Evolution of Bacterial Genome Architecture. Front Genet [Internet]. 2017;8:72. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28611826

215. Hallatschek O, Hersen P, Ramanathan S, Nelson DR. Genetic drift at expanding frontiers promotes gene segregation. Proc Natl Acad Sci U S A [Internet]. 2007;104(50):19926–30. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18056799

216. Escobar-Paramo P, Ghosh S, DiRuggiero J, Escobar-Páramo P, Ghosh S, DiRuggiero J, et al. Evidence for genetic drift in the diversification of a geographically isolated population of the hyperthermophilic archaeon Pyrococcus. Mol Biol Evol [Internet]. 2005;22(11):2297–303. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16079249

217. Pepperell CS, Casto AM, Kitchen A, Granka JM, Cornejo OE, Holmes EC, et al. The role of selection in shaping diversity of natural M. tuberculosis populations. PLoS Pathog. 2013;9(8):e1003543.

218. Corbett-Detig RB, Hartl DL, Sackton TB. Natural selection constrains neutral diversity across a wide range of species. PLoS Biol [Internet]. 2015;13(4):e1002112. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25859758

219. Charles D. On the origin of species by means of natural selection. Murray, London. 1859;(ok).

220. Darwin C. On the origin of species, 1859. species the origin of, editor. Routledge; 2004.

221. König B. Natural selection. 2001;ok.

222. Futuyma D. Natural selection: How evolution works. Retrieved Sept. 2004;26(http://www.actionbioscience.org/evolution/futuyma.html/lessons/contactus.html):2015.

223. van Belkum A, Struelens M, de Visser A, Verbrugh H, Tibayrenc M. Role of genomic typing in taxonomy, evolutionary genetics, and microbial epidemiology. Clin Microbiol Rev [Internet]. 2001;14(3):547–60. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11432813

224. Doolittle WF, Papke RT. Genomics and the bacterial species problem. Genome Biol [Internet]. 2006;7(9):116. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17020593

225. Purdom G. Antibiotic resistance of bacteria an example of evolution in action. Answers. 2007;2(3):74–6.

226. Lenski RE. Bacterial evolution and the cost of antibiotic resistance. Int Microbiol. 1998;1(4):265–70.

227. Blazquez J, Oliver A, Gómez-Gómez J-M. Mutation and evolution of antibiotic resistance: antibiotics as promoters of antibiotic resistance? Curr Drug Targets. 2002;3(4):345–9.

228. Anderson KL. Is bacterial resistance to antibiotics an appropriate example of evolutionary change. Creat Res Soc Q. 2005;41(4):318–26.

229. Quesada JM, Perez-Martinez I, Ramos C, Lopez MM, Penyalver R, Pérez-Martínez I, et al. IS53: an insertion element for molecular typing of *Pseudomonas savastanoi* pv. *savastanoi* . Res Microbiol [Internet]. 2008;159(3):207–15. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18359611

230. Evan T. Picking Holes in the Concept of Natural Selection. Oxford University Press; 2014.

231. Van Belkum A, Tassios PT, Dijkshoorn L, Haeggman S, Cookson B, Fry NK, et al. Guidelines for the validation and application of typing methods for use in bacterial epidemiology. Clin Microbiol Infect. 2007;13:1–46.

232. Morris CE, Barny MA, Berge O, Kinkel LL, Lacroix C. Frontiers for research on the ecology of plant-pathogenic bacteria: fundamentals for sustainability: Challenges in Bacterial Molecular Plant Pathology. Mol Plant Pathol [Internet]. 2017;18(2):308–19. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27862839

233. Fegan M, Hayward C. Genetic diversity of bacterial plant pathogens. In: Plant Microbiology. Taylor & Francis; 2004. p. 206–31.

234. Anderson PK, Cunningham AA, Patel NG, Morales FJ, Epstein PR, Daszak P. Emerging infectious diseases of plants: pathogen pollution, climate change and agrotechnology

drivers. Trends Ecol Evol [Internet]. 2004;19(10):535–44. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16701319

235. Altizer S, Harvell D, Friedle E. Rapid evolutionary dynamics and disease threats to biodiversity. Trends Ecol Evol. 2003;18(11):589–96.

236. Dijkshoorn L, Ursing BM, Ursing JB. Strain, clone and species: comments on three basic concepts of bacteriology. J Med Microbiol. 2000;49(5):397–401.

237. Brenner DJ, Staley JT, Krieg NR. Classification of procaryotic organisms and the concept of bacterial speciation. In: Bergey's Manual «of Systematic Bacteriology. Springer; 2005. p. 27–32.

238. Atallah ZK, Maruthachalam K, Du Toit L, Koike ST, Davis RM, Klosterman SJ, et al. Population analyses of the vascular plant pathogen *Verticillium dahliae* detect recombination and transcontinental gene flow. Fungal Genet Biol. 2010;47(5):416–22.

239. Berbegal M, Pérez-Sierra A, Armengol J, Grünwald NJ. Evidence for multiple introductions and clonality in Spanish populations of *Fusarium circinatum*. Phytopathology. 2013;103(8):851–61.

240. Dunn AR, Bruening SR, Grünwald NJ, Smart CD. Evolution of an experimental population of *Phytophthora capsici* in the field. Phytopathology. 2014;104(10):1107–17.

241. Stukenbrock EH, Banke S, Javan-Nikkhah M, McDonald BA. Origin and domestication of the fungal wheat pathogen Mycosphaerella graminicola via sympatric speciation. Mol Biol Evol. 2006;24(2):398–411.

242. Schoebel CN, Stewart J, Gruenwald NJ, Rigling D, Prospero S. Population history and pathways of spread of the plant pathogen Phytophthora plurivora. PLoS One. 2014;9(1):e85368.

243. Young JM. Taxonomy of *Pseudomonas syringae*. J Plant Pathol. 2010;S5–14.

244. Alvarez F, De Los Ríos JG, Jimenez P, Rojas A, Reche P, Troya MT. Phenotypic variability in different strains of *Pseudomonas syringae* subsp. *savastanoi* isolated from different hosts. Eur J plant Pathol. 1998;104(6):603–9.

245. O'Connor L. Diagnostic bacteriology protocols. Vol. 345. Springer Science & Business Media; 2006.

246. Bentley SD, Parkhill J. Genomic perspectives on the evolution and spread of bacterial pathogens. Proc R Soc B Biol Sci [Internet]. 2015;282(1821):20150488. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26702036

247. Braga PAC, Tata A, Gonçalves dos Santos V, Barreiro JR, Schwab NV, Veiga dos Santos M, et al. Bacterial identification: from the agar plate to the mass spectrometer. Rsc Adv. 2013;3(4):994–1008.

248. Gardan L, Bollet C, Ghorrah MA, Grimont F, Grimont PAD. DNA relatedness among the pathovar strains of *Pseudomonas syringae* subsp. *savastanoi* Janse (1982) and proposal of *Pseudomonas savastanoi* sp. nov. Int J Syst Evol Microbiol. 1992;42(4):606–12.

249. Cesbron S, Pothier J, Gironde S, Jacques M-A, Manceau C. Development of multilocus variable-number tandem repeat analysis (MLVA) for *Xanthomonas arboricola* pathovars. J Microbiol Methods. 2014;100:84–90.

250. Vinatzer BA, Monteil CL, Clarke CR. Harnessing population genomics to understand how bacterial pathogens emerge, adapt to crop hosts, and disseminate. Annu Rev Phytopathol [Internet]. 2014;52:19–43. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24820995

251. Della Coletta-Filho H, Takita MA, de Souza AA, Aguilar-Vildoso CI, Machado MA, Coletta-Filho HD, et al. Differentiation of strains of *Xylella fastidiosa* by a variable number of tandem repeat analysis. Appl Env Microbiol [Internet]. 2001;67(9):4091–5. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11526010

252. Liang Y, Zhang X, Tian C, Gao A, Wang P. Pathogenic indentification of kiwifruit bacterial canker in Shaanxi. J Northwest For Coll. 2000;15(1):37–9.

253. Li W, Raoult D, Fournier P-EE. Bacterial strain typing in the genomic era. FEMS Microbiol Rev [Internet]. 2009;33(5):892–916. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19453749

254. van Belkum A. The role of short sequence repeats in epidemiologic typing. Curr Opin Microbiol. 1999;2(3):306–11.

255. Jannes G, De Vos D. A review of current and future molecular diagnostic tests for use in the microbiology laboratory. In: Diagnostic Bacteriology Protocols. Springer; 2006. p. 1–21.

256. Smith JJ, Palmano S, Saddler GS. Characterisation of *Curtobacterium flaccumfaciens* pathovars by AFLP, rep-PCR and pulsed-field gel electrophoresis. Eur J plant Pathol. 2003;109(8):817–25.

257. Jaunet TX, Wang J-F. Variation in genotype and aggressiveness of Ralstonia solanacearum race 1 isolated from tomato in Taiwan. Phytopathology. 1999;89(4):320–7.

258. Manceau C, Horvais A. Assessment of genetic diversity among strains of *Pseudomonas syringae* by PCR-restriction fragment length polymorphism analysis of rRNA operons with special emphasis on P. syringae pv. tomato. Appl Environ Microbiol. 1997;63(2):498–505.

259. Savelkoul PHM, Aarts HJM, De Haas J, Dijkshoorn L, Duim B, Otsen M, et al. Amplified-fragment length polymorphism analysis: the state of an art. J Clin Microbiol. 1999;37(10):3083–91.

260. Maiden MCJ, Bygraves JA, Feil E, Morelli G, Russell JE, Urwin R, et al. Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. Proc Natl Acad Sci. 1998;95(6):3140–5.

261. Qian W, Jia Y, Ren S-X, He Y-Q, Feng J-X, Lu L-F, et al. Comparative and functional genomic analyses of the pathogenicity of phytopathogen *Xanthomonas campestris* pv. *campestris*. Genome Res. 2005;15(6):757–67.

262. Bhattacharyya A, Stilwagen S, Reznik G, Feil H, Feil WS, Anderson I, et al. Draft sequencing and comparative genomics of *Xylella fastidiosa* strains reveal novel biological insights. Genome Res. 2002;12(10):1556–63.

263. Mondal KK, Shanmugam V. Advancements in the diagnosis of bacterial plant pathogens: An overview. Biotechnol Mol Biol Rev. 2013;8(1):1–11.

264. Mugnai L, Giovannetti L, Ventura S, Surico G. The grouping of strains of *Pseudomonas syringae* subsp. *savastanoi* by DNA restriction fingerprinting. J Phytopathol. 1994;142(4):209–18.

265. Sisto A, Morea M, Baruzzi F, Palumbo G. Differentiation of *Pseudomonas syringae* subsp. *savastanoi* strains isolated from various host plants by restriction fragment length polymorphism. Phytopathol Mediterr. 2002;41(1):63–71.

266. Scortichini M, Rossi MP, Salerno M. Relationship of genetic structure of *Pseudomonas savastanoi* pv. *savastanoi* populations from Italian olive trees and patterns of host genetic diversity. Plant Pathol. 2004;53(4):491–7.

267. Nguyen KA, Förster H, Adaskaveg JE. Genetic Diversity of *Pseudomonas savastanoi* pv. *savastanoi* in California and Characterization of Epidemiological Factors for Olive Knot Development. Plant Dis. 2018;102(9):1718–24.

268. Sisto A, Cipriani MG, Tegli S, Cerboneschi M, Stea G, Santilli E. Genetic characterization by fluorescent AFLP of *Pseudomonas savastanoi* pv. *savastanoi* strains isolated from different host species. Plant Pathol. 2007;56(3):366–72.

269. Krid S, Rhouma A, Quesada JM, Penyalver R, Gargouri A. Delineation of *Pseudomonas savastanoi* pv. *savastanoi* strains isolated in Tunisia by random-amplified polymorphic DNA analysis. J Appl Microbiol [Internet]. 2009;106(3):886–94. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19191961

270. Gori A, Cerboneschi M, Tegli S. High-resolution melting analysis as a powerful tool to discriminate and genotype *Pseudomonas savastanoi* pathovars and strains. PLoS One. 2012;7(1):e30199.

271. Moretti C, Vinatzer BA, Onofri A, Valentini F, Buonaurio R. Genetic and phenotypic diversity of Mediterranean populations of the olive knot pathogen, *Pseudomonas savastanoi* pv. *savastanoi*. Plant Pathol. 2017;66(4):595–605.

272. Mazzaglia A, Renzi M, Balestra GM. Comparison and utilization of different PCR-based approaches for molecular typing of *Pseudomonas syringae* pv. *actinidiae* strains from Italy. Can J Plant Pathol. 2011;33(1):8–18.

273. Mazzaglia A, Studholme DJ, Taratufolo MC, Cai R, Almeida NF, Goodman T, et al. *Pseudomonas syringae* pv. *actinidiae* (PSA) isolates from recent bacterial canker of kiwifruit outbreaks belong to the same genetic lineage. PLoS One. 2012;7(5):e36518.

274. Marcelletti S, Ferrante P, Petriccione M, Firrao G, Scortichini M. *Pseudomonas syringae* pv. *actinidiae* draft genomes comparison reveal strain-specific features involved in adaptation and virulence to *Actinidia* species. PLoS One. 2011;6(11):e27297.

275. Butler MI, Stockwell PA, Black MA, Day RC, Lamont IL, Poulter RTM. *Pseudomonas syringae* pv. *actinidiae* from recent outbreaks of kiwifruit bacterial canker belong to different clones that originated in China. PLoS One. 2013;8(2):e57464.

276. Chapman JR, Taylor RK, Weir BS, Romberg MK, Vanneste JL, Luck J, et al. Phylogenetic relationships among global populations of *Pseudomonas syringae* pv. *actinidiae*. Phytopathology [Internet]. 2012;102(11):1034–44. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22877312

277. Chen J, Chang CJ, Jarret RL, Gawel N. Genetic variation among *Xylella fastidiosa* strains. Phytopathology. 1992;82(9):973–7.

278. Chen J, Lamikanra O, Chang CJ, Hopkins DL. Randomly amplified polymorphic DNA analysis of *Xylella fastidiosa* Pierce's disease and oak leaf scorch pathotypes. Appl Environ Microbiol. 1995;61(5):1688–90.

279. Mehta A, Leite RP, Rosato YB. Assessment of the genetic diversity of *Xylella fastidiosa* isolated from citrus in Brazil by PCR-RFLP of the 16S rDNA and 16S-23S intergenic spacer and rep-PCR fingerprinting. Antonie Van Leeuwenhoek. 2001;79(1):53–9.

280. Qin X, Miranda VS, Machado MA, Lemos EGM, Hartung JS. An evaluation of the genetic diversity of *Xylella fastidiosa* isolated from diseased citrus and coffee in Sao Paulo, Brazil. Phytopathology. 2001;91(6):599–605.

281. Lin H, Civerolo EL, Hu R, Barros S, Francis M, Walker MA. Multilocus simple sequence repeat markers for differentiating strains and evaluating genetic diversity of *Xylella fastidiosa* . Appl Environ Microbiol. 2005;71(8):4888–92.

282. Lin H, Islam MS, Cabrera-La Rosa JC, Civerolo EL, Groves RL. Population structure of *Xylella fastidiosa* associated with almond leaf scorch disease in the San Joaquin Valley of California. Phytopathology. 2015;105(6):825–32.

283. Francisco CS, Ceresini PC, Almeida RPP, Coletta-Filho HD. Spatial genetic structure of coffee-associated *Xylella fastidiosa* populations indicates that cross infection does not occur with sympatric citrus orchards. Phytopathology. 2017;107(4):395–402.

284. Scally M, Schuenzel EL, Stouthamer R, Nunney L. Multilocus sequence type system for the plant pathogen *Xylella fastidiosa* and relative contributions of recombination and point mutation to clonal diversity. Appl Env Microbiol [Internet]. 2005;71(12):8491–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16332839

285. Elbeaino T, Valentini F, Kubaa RA, Moubarak P, Yaseen T, Digiaro M. Multilocus sequence typing of *Xylella fastidiosa* isolated from olive affected by "olive quick decline syndrome" in Italy. Phytopathol Mediterr. 2014;53(3):533–42.

286. Denancé N, Legendre B, Briand M, Olivier V, De Boisseson C, Poliakoff F, et al. Several subspecies and sequence types are associated with the emergence of *Xylella fastidiosa* in natural settings in France. Plant Pathol. 2017;66(7):1054–64.

287. Montes Borrego M, Lopes JRS, Jiménez-Díaz RM, Landa BB. Combined use of a new SNP-based assay and multilocus SSR markers to assess genetic diversity of *Xylella fastidiosa* subsp. pauca infecting citrus and coffee plants. 2015;

288. Van Sluys MA, De Oliveira MC, Monteiro-Vitorello CB, Miyaki CY, Furlan LR, Camargo LEA, et al. Comparative analyses of the complete genome sequences of Pierce's disease and citrus variegated chlorosis strains of *Xylella fastidiosa* . J Bacteriol. 2003;185(3):1018–26.

289. Marcelletti S, Scortichini M. Genome-wide comparison and taxonomic relatedness of multiple *Xylella fastidiosa* strains reveal the occurrence of three subspecies and a new Xylella species. Arch Microbiol. 2016;198(8):803–12.

290. Giampetruzzi A, Saponari M, Loconsole G, Boscia D, Savino VN, Almeida RPP, et al. Genome-wide analysis provides evidence on the genetic relatedness of the emergent *Xylella fastidiosa* genotype in Italy to isolates from Central America. Phytopathology. 2017;107(7):816–27.

291. Lin J-J, Kuo J, Ma J. A PCR-based DNA fingerprinting technique: AFLP for molecular typing of bacteria. Nucleic Acids Res. 1996;24(18):3650.

292. Foster JT, Bull RL, Keim PS. Ricin Forensics: Comparisons to Microbial Forensics. In: Microbial Forensics. Elsevier; 2011. p. 315–26.

293. Lin J-J, Kuo J, Ma J, Saunders JA, Beard HS, MacDonald MH, et al. Identification of molecular markers in soybean comparing RFLP, RAPD and AFLP DNA mapping techniques. Plant Mol Biol Report. 1996;14(2):156–69.

294. Achtman M. Evolution, population structure, and phylogeography of genetically monomorphic bacterial pathogens. Annu Rev Microbiol [Internet]. 2008;62:53–70. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18785837

295. Davis Ii EW, Weisberg AJ, Tabima JF, Grunwald NJ, Chang JH. Gall-ID: tools for genotyping gall-causing phytopathogenic bacteria. PeerJ [Internet]. 2016;4:e2222. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27547538

296. Poulin L, Grygiel P, Magne M, Gagnevin L, Rodriguez RL, Forero Serna N, et al. New multilocus variable-number tandem-repeat analysis tool for surveillance and local epidemiology of bacterial leaf blight and bacterial leaf streak of rice caused by *Xanthomonas oryzae*. Appl Env Microbiol [Internet]. 2015;81(2):688–98. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25398857

297. Van Belkum A, Scherer S, Van Leeuwen W, Willemse D, Van Alphen L, Verbrugh H. Variable number of tandem repeats in clinical strains of *Haemophilus influenzae*. Infect Immun. 1997;65(12):5017–27.

298. Noller AC, McEllistrem MC, Pacheco AGF, Boxrud DJ, Harrison LH. Multilocus variable-number tandem repeat analysis distinguishes outbreak and sporadic *Escherichia coli* O157: H7 isolates. J Clin Microbiol [Internet]. 2003;41(12):5389–97. Available from: http://www.ncbi.nlm.nih.gov/pubmed/14662916

299. Marsh JW, O'Leary MM, Shutt KA, Pasculle AW, Johnson S, Gerding DN, et al. Multilocus variable-number tandem-repeat analysis for investigation of *Clostridium difficile* transmission in hospitals. J Clin Microbiol [Internet]. 2006;44(7):2558–66. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16825380

300. Alvarez AM. Integrated approaches for detection of plant pathogenic bacteria and diagnosis of bacterial diseases. Annu Rev Phytopathol [Internet]. 2004;42:339–66. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15283670

301. Priyanka B, Patil RK, Dwarakanath S. A review on detection methods used for foodborne pathogens. Indian J Med Res. 2016;144(3):327.

302. Váradi L, Luo JL, Hibbs DE, Perry JD, Anderson RJ, Orenga S, et al. Methods for the detection and identification of pathogenic bacteria: past, present, and future. Chem Soc Rev [Internet]. 2017;46(16):4818–32. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28644499

303. Almeida RPP, Purcell AH. Transmission of *Xylella fastidiosa* to grapevines by Homalodisca coagulata (Hemiptera: Cicadellidae). J Econ Entomol. 2003;96(2):264–71.

304. Vincelli P, Tisserat N. Nucleic Acid-Based Pathogen Detection in Applied Plant Pathology. Plant Dis [Internet]. 2008;92(5):660–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/30769590

305. Palacio-Bielsa A, Cambra MA, López MM. PCR detection and identification of plant-pathogenic bacteria: updated review of protocols (1989-2007). J Plant Pathol. 2009;249–97.

306. Prasad D, Vidyarthi AS. DNA based methods used for characterization and detection of food borne bacterial pathogens with special consideration to recent rapid methods. African J Biotechnol. 2009;8(9).

307. Lau HY, Botella JR. Advanced DNA-based point-of-care diagnostic methods for plant diseases detection. Front Plant Sci [Internet]. 2017;8:2016. Available from: http://www.ncbi.nlm.nih.gov/pubmed/29375588

308. Zhou K, Aertsen A, Michiels CW. The role of variable DNA tandem repeats in bacterial adaptation. FEMS Microbiol Rev [Internet]. 2014;38(1):119–41. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23927439

309. Van Belkum A, Scherer S, Van Alphen L, Verbrugh H. Short-sequence DNA repeats in prokaryotic genomes. Microbiol Mol Biol Rev. 1998;62(2):275–93.

310. Verstrepen KJ, Jansen A, Lewitter F, Fink GR. Intragenic tandem repeats generate functional variability. Nat Genet [Internet]. 2005;37(9):986. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16086015

311. Fan H, Chu J-Y. A Brief Review of Short Tandem Repeat Mutation. Genomics Proteomics Bioinformatics. 2007;5(1):7–14.

312. Lindstedt BA. Multiple-locus variable number tandem repeats analysis for genetic fingerprinting of pathogenic bacteria. Electrophoresis [Internet]. 2005;26(13):2567–82. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15937984

313. López-Flores I, Garrido-Ramos MA. The repetitive DNA content of eukaryotic genomes. In: Repetitive DNA. Karger Publishers; 2012. p. 1–28.

314. Zhou H, Du L, Yan H. Detection of tandem repeats in DNA sequences based on parametric spectral estimation. IEEE Trans Inf Technol Biomed. 2009;13(5):747–55.

315. Jansen A, Gemayel R, Verstrepen KJ. Unstable microsatellite repeats facilitate rapid evolution of coding and regulatory sequences. In: Repetitive DNA. Karger Publishers; 2012. p. 108–25.

316. Sawaya S, Bagshaw A, Buschiazzo E, Kumar P, Chowdhury S, Black MA, et al. Microsatellite tandem repeats are abundant in human promoters and are associated with regulatory elements. PLoS One. 2013;8(2):e54710.

317. Schlotterer C, Imhof M, Wang H, Nolte V, Harr B, Schlötterer C, et al. Low abundance of *Escherichia coli* microsatellites is associated with an extremely low mutation rate. J Evol Biol [Internet]. 2006;19(5):1671–6. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16910996

318. Quilez J, Guilmatre A, Garg P, Highnam G, Gymrek M, Erlich Y, et al. Polymorphic tandem repeats within gene promoters act as modifiers of gene expression and DNA methylation in humans. Nucleic Acids Res [Internet]. 2016;44(8):3750–62. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27060133

319. Gemayel R, Vinces MD, Legendre M, Verstrepen KJ. Variable tandem repeats accelerate evolution of coding and regulatory sequences. Annu Rev Genet [Internet]. 2010;44:445–77. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20809801

320. Gemayel R, Cho J, Boeynaems S, Verstrepen KJ. Beyond junk-variable tandem repeats as facilitators of rapid evolution of regulatory and coding sequences. Genes (Basel) [Internet]. 2012;3(3):461–80. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24704980

321. Thierry A, Bouchier C, Dujon B, Richard G-FF. Megasatellites: a peculiar class of giant minisatellites in genes involved in cell adhesion and pathogenicity in *Candida glabrata*. Nucleic Acids Res [Internet]. 2008;36(18):5970–82. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18812401

322. Ee CS, Ismail A, Singh KKB. Multiple-locus variable-number tandem repeat analysis (MLVA) for bacterial typing. 2016;

323. Weber JL, Wong C. Mutation of human short tandem repeats. Hum Mol Genet. 1993;2(8):1123–8.

324. Jeffreys AJ, Wilson V, Thein SL. Individual-specific 'fingerprints' of human DNA. Nature. 1985;316(6023):76.

325. Tachida H, Iizuka M. Persistence of repeated sequences that evolve by replication slippage. Genetics. 1992;131(2):471–8.

326. Sia EA, Jinks-Robertson S, Petes TD. Genetic control of microsatellite stability. Mutat Res Repair. 1997;383(1):61–70.

327. Richard G, Pâques F. Mini-and microsatellite expansions: the recombination connection. EMBO Rep. 2000;1(2):122–6.

328. Carraway M, Marinus MG. Repair of heteroduplex DNA molecules with multibase loops in Escherichia coli. J Bacteriol. 1993;175(13):3972–80.

329. Wells RD, Parniewski P, Pluciennik A, Bacolla A, Gellibolian R, Jaworski A. Small slipped register genetic instabilities in *Escherichia coli* in triplet repeat sequences associated with hereditary neurological diseases. J Biol Chem. 1998;273(31):19532–41.

330. Bichara M, Wagner J, Lambert IB. Mechanisms of tandem repeat instability in bacteria. Mutat Res [Internet]. 2006;598(1–2):144–63. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16519906

331. Treangen TJ, Abraham A-LL, Touchon M, Rocha EPC. Genesis, effects and fates of repeats in prokaryotic genomes. FEMS Microbiol Rev [Internet]. 2009;33(3):539–71. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19396957

332. Lim KG, Kwoh CK, Hsu LY, Wirawan A. Review of tandem repeat search tools: a systematic approach to evaluating algorithmic performance. Brief Bioinform. 2012;14(1):67–81.

333. Merkel A, Gemmell N. Detecting short tandem repeats from genome data: opening the software black box. Br Bioinform [Internet]. 2008;9(5):355–66. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18621747

334. Sharma PC, Grover A, Kahl G. Mining microsatellites in eukaryotic genomes. Trends Biotechnol [Internet]. 2007;25(11):490–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17945369

335. Leclercq S, Rivals E, Jarne P. Detecting microsatellites within genomes: significant variation among algorithms. BMC Bioinformatics [Internet]. 2007;8(1):125. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17442102

336. Benson G. Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res. 1999;27(2):573–80.

337. Orgel LE, Crick FHC. Selfish DNA: the ultimate parasite. Nature. 1980;284(5757):604.

338. Legendre M, Pochet N, Pak T, Verstrepen KJ. Sequence-based estimation of minisatellite and microsatellite repeat variability. Genome Res [Internet]. 2007;17(12):1787–96. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17978285

339. Cummings CJ, Zoghbi HY. Fourteen and counting: unraveling trinucleotide repeat diseases. Hum Mol Genet. 2000;9(6):909–16.

340. Yang J, Li F. Are all repeats created equal? Understanding DNA repeats at an individual level. Curr Genet [Internet]. 2017;63(1):57–63. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27260214

341. Lohi H, Young EJ, Fitzmaurice SN, Rusbridge C, Chan EM, Vervoort M, et al. Expanded repeat in canine epilepsy. Science (80-). 2005;307(5706):81.

342. Sureshkumar S, Todesco M, Schneeberger K, Harilal R, Balasubramanian S, Weigel D. A genetic defect caused by a triplet repeat expansion in Arabidopsis thaliana. Science (80-). 2009;323(5917):1060–3.

343. Reinar WB, Knutsen JB, Jentoft S, Tørresen OK, Butenko MA, Jakobsen KS. Evolutionary adaptability linked to length variation in short genic tandem repeats . 2018;

344. Li Y-C, Röder MS, Fahima T, Kirzhner VM, Beiles A, Korol AB, et al. Natural selection causing microsatellite divergence in wild emmer wheat at the ecologically variable microsite at Ammiad, Israel. Theor Appl Genet. 2000;100(7):985–99.

345. Turpeinen T, Tenhola T, Manninen O, Nevo E, Nissilä E. Microsatellite diversity associated with ecological factors in *Hordeum spontaneum* populations in Israel. Mol Ecol. 2001;10(6):1577–91.

346. Drögemüller C, Karlsson EK, Hytönen MK, Perloski M, Dolf G, Sainio K, et al. A mutation in hairless dogs implicates FOXI3 in ectodermal development. Sci (New York, NY). 2008;321(5895):1462.

347. Weiser JN, Love JM, Moxon ER. The molecular mechanism of phase variation of *H. influenzae* lipopolysaccharide. Cell. 1989;59(4):657–65.

348. Saunders NJ, Jeffries AC, Peden JF, Hood DW, Tettelin H, Rappuoli R, et al. Repeat-associated phase variable genes in the complete genome sequence of *Neisseria meningitidis* strain MC58. Mol Microbiol. 2000;37(1):207–15.

349. Oliveira EJ, Pádua JG, Zucchi MI, Vencovsky R, Vieira MLC. Origin, evolution and genome distribution of microsatellites. Genet Mol Biol. 2006;29(2):294–307.

350. Rocha EPC, Danchin A, Viari A. Functional and evolutionary roles of long repeats in prokaryotes. Res Microbiol. 1999;150(9–10):725–33.

351. Vieira MLC, Santini L, Diniz AL, Munhoz Cde F, Munhoz C de F. Microsatellite markers: what they mean and why they are so useful. Genet Mol Biol [Internet]. 2016;39(3):312–28. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27561112

352. Van Belkum A, Eriksen NHR, Sijmons M, Van Leeuwen W, Van Den Bergh M, Kluytmans J, et al. Coagulase and protein A polymorphisms do not contribute to persistence of nasal colonisation by Staphylococcus aureus. J Med Microbiol. 1997;46(3):222–32.

353. van Belkum A, Tassios PT, Dijkshoorn L, Haeggman S, Cookson B, Fry NK, et al. Guidelines for the validation and application of typing methods for use in bacterial epidemiology. Clin Microbiol Infect [Internet]. 2007;13 Suppl 3:1–46. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17716294

354. Jansen R, Embden JDA van, Gaastra W, Schouls LM. Identification of genes that are associated with DNA repeats in prokaryotes. Mol Microbiol. 2002;43(6):1565–75.

355. Grissa I, Bouchon P, Pourcel C, Vergnaud G. On-line resources for bacterial micro-evolution studies using MLVA or CRISPR typing. Biochimie [Internet]. 2008;90(4):660–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17822824

356. Le Flèche P, Jacques I, Grayon M, Al Dahouk S, Bouchon P, Denoeud F, et al. Evaluation and selection of tandem repeat loci for a Brucella MLVA typing assay. BMC Microbiol [Internet]. 2006;6(1):9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16469109

357. Ciarroni S, Gallipoli L, Taratufolo MC, Butler MI, Poulter RTM, Pourcel C, et al. Development of a Multiple Loci Variable Number of Tandem Repeats Analysis (MLVA) to Unravel the Intra-Pathovar Structure of *Pseudomonas syringae* pv. *actinidiae* Populations Worldwide. PLoS One [Internet]. 2015;10(8):e0135310. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26262683

358. Pruvost O, Magne M, Boyer K, Leduc A, Tourterel C, Drevet C, et al. A MLVA genotyping scheme for global surveillance of the citrus pathogen *Xanthomonas citri* pv. *citri* suggests a worldwide geographical expansion of a single genetic lineage. PLoS One [Internet]. 2014;9(6):e98129. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24897119

Chapter 2 **: Canker of kiwifruit disease**

2.1 **Kiwi fruit (Actinidia spp.) kiwi, Chinese gooseberry, Yang-tao, Fuzzy-Skinned**

The genus *Actinidia* (*A.*) belongs to the family *Actinidiaceae*, which includes 54 species and 21 varieties. The famous commercial varieties distributed around the world are *A. chinensis*, *A deliciosa* and *A. chevalier* (1). The genus *Actinidia* is widely distributed in eastern Asia, and most species are present in China endemically. Their distribution lies in mountains and hills of the southwest to the northeast, and for this reason, China is considered the origin of the genus (2). Although the wild taxa of *actinidiae* are present in China, it was not domesticated until the last century, when it was domesticated for commercial purposes for the good-flavoured, and large-fruited species. Two of the main species distributed around the world are *A. chinensis* and *A. deliciosa* and they have been domesticated over a wide area of China, after the success of domestication in New Zealand (3). In New Zealand, the first domestication and commercial cultivation started approximately in 1904 when a school teacher brought seeds from China and passed them to the botanist Alexander Allison, and commercial cultivation started (4). The fruit was actually planted in Europe between the 17th and 18th centuries for ornamental purposes (5). The genus is a dioecious plant and it has different advantages for human health, where it considered to be beneficial for gut health and digestion (6). The fruit helps to improve digestive, immune and metabolic health. It is rich in vitamin C, dietary fiber, potassium, vitamin E and folate, besides many bioactive components which include antioxidants, phytonutrients, and enzymes (7).

Kiwifruit production has increased by more than 50% in the last twenty years, and the production of this fruit has reached 0.22% of global food production, after apple, orange, and banana (6). Five countries are the world's main producers of kiwi: China, Italy, New Zealand, Chile, and Greece. Of these, Italy, New Zealand, Chile, and Greece account for 80% of total production, whereas China exports only 1% of their production while the other amount is used for national consumption (8). The kiwifruit ranks after citrus, apples, table grapes, peaches / nectarines, and pears in terms of value. In Italy, kiwifruit accounts for about 3.5% of the total area under fruit crops and about 4% of total fruit production by weight (9).

Plant pathogens are the main constraints for plant production. The Kiwi plant is attacked by different pathogens, such as fungi, viruses, and bacteria. Different plant pathogens have

been reported such as bacterial canker, blossom blight, fruit rot, stem canker, gray mold, anthracnose, Nectria canker, leaf blight, brown leaf spot, brown felt, and root-knot nematode (10). The main fungal diseases of kiwifruit are Field rot, caused by *Sclerotinia sclerotiorum*, Storage rot, caused by *Botrytis cinerea*, and Ripe rot, caused by *Botryosphaeria dothidea* (11). It is also attacked by other fungi species, such as *Phaeoacremonium* spp., *Phaemoniella* spp. *and Fomitiporia punctate*, which cause wood decay and vine decay (12). Other fungi cause nursery damage on kiwifruit on the young plant (13). Viruses, also, attack kiwi plants. Three different groups of viruses have been identified, and of these two viruses have been identified as causing severe damage to the commercial orchards of kiwifruit: Cherry leaf roll virus (CLRV) and Pelargonium zonate spot virus (PZSV) (10). Bacterial diseases also affect kiwifruit. Two species *Pseudomonas viridiflava* and *Pseudomonas syringae* pv. *syringae* causes Blossom blight and necrotic leaf spotting (14). In 2008 a new virulent strain of *Pseudomonas syringae* pv. *actinidiae* has attacked the kiwi fruit in Italy causes a severe damages to the Kiwi production sector.

2.2 Importance and distribution of the disease

The disease is caused by virulent strains of the gram-negative bacterium *Pseudomonas syringae* pv. *actinidiae* (Psa) (15,16). The disease has caused great damage to kiwifruit productions around the world (17), It was first described in 1984 in Japan on *Actinidia deliciosa* as a new pathovar belonging to *Pseudomonas syringae* complex (18). The disease was described in China in 1994 as attacking kiwi fruit (17). Since the earliest identification of the disease in Japan, the disease has been officially reported in different countries in China (19), Korea since the 1990s (20–23). In addition, Italy (24,25) reported the disease in the early 1990s. In 2008 a new and very aggressive strain has emerged in Italy, causing severe damage and spreading quickly all over the country, attacking *Actinidia chinensis*, and *Actinidia deliciosa* (26). Globally, after the detection of the Italian outbreak, the disease has been reported from different countries on *Actinidia chinensis*, and *Actinidia deliciosa* in Portugal, New Zeeland, France, Switzerland, Spain, Chile, Turkey, Germany, Slovenia, Greece, Georgia, and finally Argentina (16,27).

FIGURE 5**:** Distribution of bacterial canker of kiwi fruit around the world.

Under favorable conditions, the disease can cause great economic losses by destroying the orchard in one or just a few seasons (28–33). The disease has caused severe damage in orchards and extensive economic losses worldwide (34,35). It is already responsible for worldwide economic losses amounting to hundreds of millions of euros and poses a very big threat to the entire kiwifruit industry in New Zealand (36). In New Zealand the problem has had a bigger impact; in the first five years of the outbreak losses of around NZ$400 million were expected, but about NZ$900 million in just one year (35). The situation in New Zealand has now been recovered by the replacement of the sensitive variety 'Hort16A' with *Actinidia chinensis* var. *chinensis* 'Zesy002' (37). In Italy for example, losses reached about €2 million euros in 2009 (38). The situation has increased according to the infected area, which reached 2000 ha between 2010-2012, and each lost hectare cost about €85.000 (39).

Different measures have been taken in order to control and manage the disease in Italy. These include spray treatments, cultural practices, such as pruning, tool disinfections, and covering pruning cuts, managing nitrogen fertilization and water stress in a correct manner, and many others (40). The recovery of kiwi plantation in Italy is proceeding slowly, despite the cultivation of more than 25,000 ha around the country (41).

2.3 The disease processes

The development of the disease depends on three major environmental factors: moisture, temperature, and light (42). Multiplication of the bacteria depends on favourable conditions; optimum temperatures (which range from 12 to18 °C), and humid conditions can greatly favour the multiplication of the bacterium. On the other hand, high temperatures negatively affect the multiplication and dispersal of the bacterium (Young, 2012). High temperatures reduce the amount of ooze, while high winds increase the severity of the disease (43,44).

The bacteria can survive very well in its host plant (Kiwi fruits), and in water, where it can survive for a long period of time especially in raindrops. In soils and compost, the bacteria may survive but not colonize. The spread of the disease is mainly based on the means of bacterial survival. Several methods have been reported for the dispersal of the pathogen *Pseudomonas syringae* pv. *actinidiae*, whether over long or short distances. Agronomical techniques, as well as frost, wind, rain and hail storms, can help disperse the disease (45), and insects such as honey bees were found to be able to carry the bacterium (43). In addition to the environmental conditions, control management can limit disease dispersal and severity. Disease severity differs from region to region or season to season (44). The development of symptoms depends on the level of inoculum of the previous fall/autumn and early winter. The pathogen can overwinter in diseased plants.

Primary infection occurs in late winter when the bacterial exudates ooze out and disperse the pathogen to the field. The pathogen can invade twigs through both wounds and natural openings, such as harvest wounds, leaf fall and high activity of stomata and lenticels, in addition to pruning, mainly in winter (46). These points of entry are not the only ones, but environmental conditions, such as frosts, hailstorms, and heavy storms can also create wounds, and their role also disperses the pathogen and creates a secondary infection (29).

In association with the entrance of the bacteria where it colonizes the branches through wounds, movement of the bacteria then occurs systemically, and it colonizes the one-year-old twigs' vascular system. This process may happen many times in winter until the migration of the bacteria reaches the main trunk in the following season (47). The pathogen overwinters in the diseased plants. The pathogen reaches and colonizes the root of the kiwi plant, and bacteria are also able to colonize and migrate from young leaves to twigs systemically (48). Bacteria migrate from leaf veins to petioles and twigs where the endophytic phase could happen can

occur (34). The development of symptoms depends on the level of inoculum, in the previous fall/autumn and early winter.

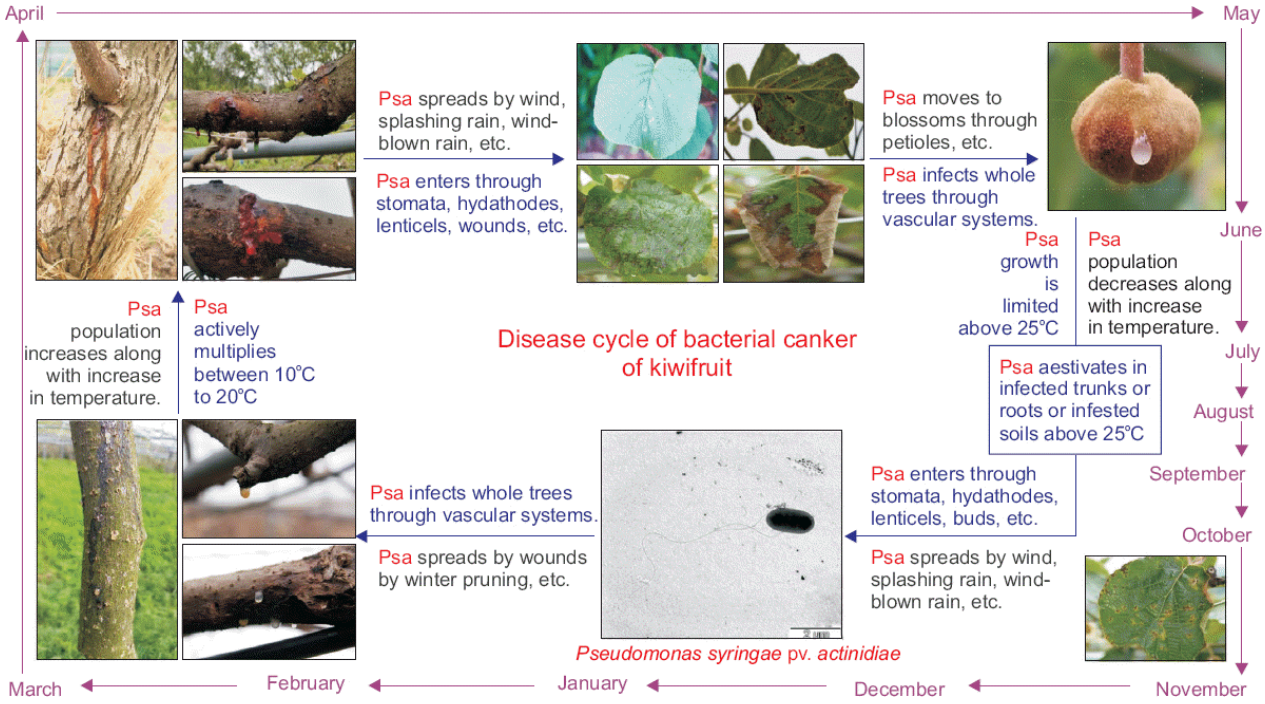


FIGURE 6: Proposed disease cycle of bacterial canker on green-fleshed kiwifruit (*Actinidia deliciosa*) caused by *Pseudomonas syringae* pv. *actinidiae* (Psa).

2.4 Symptoms of the disease

The symptoms on kiwi plants are easy to recognize and depend on affected plant parts and the season. Symptoms mainly start with a water-soaked lesion on a leaf, turning into brown /spots which are usually surrounded by a yellow halo. Sometimes it depends on the season; the desiccation of flowery parts of the plant can be seen clearly and fruit collapse is clear from symptoms on vegetative tissue. Dieback on woody parts of the kiwi plants, such as twigs, leaders, and canes, is a sign of disease, and development of a typical canker formation on the vines and trunks is another sign, with abundant white shiny milky ooze or dark red exudates excreted from the bark (49). In a winter infection, for example, plants start to release the exudate, which is a clear sign of the bacterial infection. Meanwhile, canker develops alongside the withering of main twigs and trunks.

FIGURE 7: Symptoms of Canker disease of Kiwi fruit; (A) Brown spots surrounded by yellow haloes are visible in spring on kiwifruit leaves; (B) Trunk affected by *Pseudomonas syringae* pv. *actinidiae* shows cankers which usually ooze red exudates; (C) Foliar wilting at the beginning of the season; (D)Bacterial ooze due to infection on kiwi plant; (E) Bacterial exudate oozing from infected Actinidia twig.

2.5 Characteristic of the bacterial canker pathogen

The taxonomic tree is as follows:

Domain: Bacteria

Phylum: Proteobacteria

Class: Gammaproteobacteria

Order: *Pseudomonadales*

Family: *Pseudomonadaceae*

Genus: *Pseudomonas*

Species: *Pseudomonas syringae* pv. *actinidiae*

The causal agent of bacterial canker of kiwi fruit as described by (39,47), *P. syringae* pv. *actinidiae* (Psa), belongs to the wide broad complex species *P. syringae* with another 60 pathovars recognized based on their host preference, few of them are pathogenic, while others are environmental bacteria (17). The bacterium is Gram-negative and rod-shaped, round, convex, glistening, translucent and white on a nutrient agar plate. It has 1-3 polar flagella for movement. It is positive to catalase, levan, and tobacco hypersensitive reaction, but negative for the pigment on King B and oxidase (50). According to (17) they belong to LOPAT group I of *Pseudomonas syringae.* The pathovar was not included in the work by (51) for the classification of *Pseudomonas syringae* pathovars based on DNA-DNA hybridization. It was placed in genomospecies 8 based on ANI analyses (52,53).

The different populations of *P. syringae* pv. *actinidiae* (Psa) have been identified based on whole-genome analysis, MLST, 16SrDNA, phytotoxin production, and finally TT3E genes (54). The analysis revealed four biovars until 2014 (44,55–58) when analysis based on MLST transfer biovar four into a new pathovar *Pseudomonas syringae* pv. *actinidifoliorum*. From the first discovery and description of *Pseudomonas syringae* pv. *actinidiae* in Japan in 1984, then in China (56) and until 2014, four biovars were described, which were first based on biochemical, genetic and pathogenicity characteristics (17), in addition to phytotoxin production (44) and phylogenetic analysis (58).

The four biovars were described as follows:

Biovar 1 includes the strains isolated 1984-1992 from Japan, Italy.
Biovar 2 includes strains isolated in South Korea in 1990 (55).
Biovar 3 is the widely distributed and pandemic biovar detected in Italy in 2008 (24) and its detection around the world followed until 2018.
Biovar 4 was described in New Zealand (59) but then transferred as a new pathovar with the name of *Pseudomonas syringae* pv. *actinidifoliorum* (60).

In Japan, two new biovars were described and classified as biovars 5 and 6 based on the phenotypic and genomic analysis, in addition to the symptoms they produced. The former was discovered in 2014 in local areas and the latter was found in 2016 in Nagano Prefecture, Japan (56).

Plant bacterial pathogens such as *Pseudomonas syringe* pathovars have different kinds of virulence factors that help the bacteria to invade the plant immunity system (Type III effectors, phytotoxins, plant hormones and determinants, cell wall degrading enzymes), that are of

importance for epiphytic fitness (61,62). Two main virulence factors were found to play an important role in the pathogenicity/virulence of *Pseudomonas syringae* pv. *actinidiae*: phytotoxins and Type III effectors (63). In bacteria, phytotoxins are small molecules with different structures that are toxic to plants and affect their behavior even with small concentrations. Plants react with different kinds of symptoms, such as water soaking, leaf spots, chlorosis wilt, blights, and necrosis within the host plant. (57) Phytotoxins may also be produced by the plant as a reaction to avoid plant pathogens (64,65). Two main phytotoxins produced by *Pseudomonas syringae* pv. *actinidiae* are phaseolotoxin and coronatine, according to the biovars. One of the main differences among these biovars is the production of phytotoxins. It was found that biovar 1 can produce phaseolotoxin, and biovar 2 can produce coronatine, but biovar 3 and 5 do not produce phytotoxins, while biovar 6 produces both phaseolotoxin and coronatine (66).

Another virulence factor is the presence of Type III effector genes. *Pseudomonas syringae* pv. *actinidiae* uses the Type III secretion system in order to change and control plant cell performance to its advantage by injecting and delivering bacterial proteins (effectors) into the intracellular fluid of plant host. These effectors can subvert the plant's immune system and help the bacteria to adapt and start the disease process (58). The main aim of effector genes is to suppress the plant's immunity system, in order to establish the pathogen and start the disease, in addition to performing different biochemical roles (67). Different groups of Type III effector genes were observed among the Psa groups (1,2,3) isolated in a different period between 1984 and 2015 (68). Differences in the composition of effectors were observed in biovar 5, and biovar 6, in spite of the real number of effector genes in biovar 5 (69). Another study (58) found that effector genes in biovar 6 were fewer than in other biovars, but with two unique effectors

2.6 References

1. Alnaasan Y, Valentini F, Balestra GM, Mazzaglia A, D'Onghia AM, Varvaro L. Modification of a multiple-locus variable number tandem repeat analysis (MLVA) for typing isolates of *Erwinia amylovora*. Plant Pathol. 2017;

2. Li X, Li J, Soejarto DD. New synonyms in Actinidiaceae from China. 2007;

3. Huang H. Kiwifruit: The Genus Actinidia. Academic Press; 2016.

4. Ferguson AR. Kiwifruit: evolution of a crop. In: VII International Symposium on Kiwifruit 913. 2010. p. 31–42.

5. Ferguson AR. 1904—the year that kiwifruit (*Actinidia deliciosa*) came to New Zealand. New Zeal J Crop Hortic Sci. 2004;32(1):3–27.

6. Richardson DP, Ansell J, Drummond LN. The nutritional and health attributes of kiwifruit: A review. Eur J Nutr. 2018;57(8):2659–76.

7. Moughan PJ, Rutherfurd SM, Balan P. Kiwifruit, mucins, and the gut barrier. In: Advances in food and nutrition research. Elsevier; 2013. p. 169–85.

8. Ward C, Courtney D. Kiwifruit: taking its place in the global fruit bowl. In: Advances in food and nutrition research. Elsevier; 2013. p. 1–14.

9. FAOSTAT. No Title. 2016; Available from: http://www.fao.org

10. Guroo I, Wani SMA, Wani SMA, Ahmad M, Mir SA, Masoodi FA. A review of production and processing of kiwifruit. J Food Process Technol. 2017;8(10).

11. Koh YJ, Jung JS, Hur JS. Current status of occurrence of major diseases on kiwifruits and their control in Korea. In: V International Symposium on Kiwifruit 610. 2002. p. 437–43.

12. Pennycook SR. Fungal fruit rots of *Actinidia deliciosa* (kiwifruit). New Zeal J Exp Agric. 1985;13(4):289–99.

13. Riccioni L, Manning M, Valvassori M, Haegi A, Casonato S, Spinelli R. A new disease: leader die-back in *Actinidia chinensis*'hort16a'in Italy. In: VI International Symposium on Kiwifruit 753. 2006. p. 669–76.

14. Blouin AG, Pearson MN, Chavan RR, Woo ENY, Lebas BSM, Veerakone S, et al. Viruses of kiwifruit (Actinidia species). J Plant Pathol. 2013;221–35.

15. Balestra GM, Varvaro L. *Pseudomonas syringae* pv. *syringae* causal agent of disease on floral buds of Actinidia deliciosa (A. Chev) Liang et Ferguson in Italy. J Phytopathol. 1997;145(8-9):375–8.

16. Balestra GM, Mazzaglia A, Spinelli R, Graziani S, Quattrucci A, Rossetti A. Cancro batterico su *Actinidia chinensis*. L'Informatore Agrar. 2008;38:75–6.

17. Takikawa Y, Serizawa S, Ichikawa T, Tsuyumu S, Goto M. *Pseudomonas syringae* pv. *actinidiae* pv. nov. Japanese J Phytopathol. 1989;55(4):437–44.

18. Costa G, Ferguson AR. Bacterial canker of kiwifruit: response to a threat. In: I International Symposium on Bacterial Canker of Kiwifruit 1095. 2013. p. 27–40.

19. Fang Y, Zhu X, Wang Y. Preliminary studies on kiwifruit diseases in Hunan Province. Sichuan Fruit Sci Technol. 1990;18:28–9.

20. Zhu XX, Fang YZ, Liao XG. Study on the pathogen of bacterial canker of kiwifruit. Agric Sci Technol. 1993;6:31–3.

21. Shen Z, Huang LL, Kang ZS. The investigation of kiwifruit bacterial canker in Guanzhong zone of Shaanxi Province. Acta Agric Boreali-occidentalis Sin. 2009;18:191–3.

22. Tu X, Shen Z, Gao Z, Kang Z, Huang L. Investigation of kiwifruit bacterial canker disease (*Pseudomonas syringae* pv. *actinidiae*) in Guanzhong area of Shanxi Province and its biological control. Plant Dis Pests. 2011;2(1):17–20.

23. Zhao ZB, Gao XN, Huang QL, Huang LL, Qin HQ, Kang ZS. Identification and characterization of the causal agent of bacterial canker of kiwifruit in the Shaanxi province of China. J Plant Pathol. 2013;155–62.

24. Koh YJ. Outbreak and spread of bacterial canker in kiwifruit. Korean. J Plant Pathol. 1994;10:68–72.

25. Koh YJ, Kim GH, Jung JS, Lee YS, Hur JS. Outbreak of bacterial canker on Hort16A (*Actinidia chinensis* Planchon) caused by *Pseudomonas syringae* pv. *actinidiae* in Korea. New Zeal J Crop Hortic Sci. 2010;38(4):275–82.

26. Scortichini M. Occurrence of *Pseudomonas syringae* pv. *actinidiae* on kiwifruit in Italy. Plant Pathol. 1994;43(6):1035–8.

27. Rossetti A, Fratarcangeli L, Mazzaglia A, Quattrucci A, Renzi M, Ricci L, et al. Caratteristiche e diffusione dei batteri fitopatogeni su Actinidia spp. Italus Hortus. 2009;16(5):32–4.

28. Abelleira A, López MM, Peñalver J, Aguín O, Mansilla JP, Picoaga A, et al. First report of bacterial canker of kiwifruit caused by *Pseudomonas syringae* pv. *actinidiae* in Spain. Plant Dis. 2011;95(12):1583.

29. Balestra GM, Renzi M, Mazzaglia A. First report of bacterial canker of Actinidia deliciosa caused by *Pseudomonas syringae* pv. *actinidiae* in Portugal. New Dis Reports. 2010;22(10):2044-0588.2010.

30. EPPO. First report of *Pseudomonas syringae* pv. *actinidiae* in Germany. EPPO Glob Database. 2011;(Num. article: 2011/055).

31. Bastas KK, Karakaya A. First report of bacterial canker of kiwifruit caused by *Pseudomonas syringae* pv. *actinidiae* in Turkey. Plant Dis. 2012;96(3):452.

32. Meparishvili G, Gorgiladze L, Sikharulidze Z, Muradashvili M, Koiava L, Dumbadze R, et al. First report of bacterial canker of kiwifruit caused by *Pseudomonas syringae* pv. *actinidiae* in Georgia. Plant Dis. 2016;100(2):517.

33. Balestra GM, Buriani G, Cellini A, Donati I, Mazzaglia A, Spinelli F. First Report of *Pseudomonas syringae* pv. *actinidiae* on Kiwifruit Pollen from Argentina. Plant Dis. 2018;102(1):237.

34. Scortichini M, Marcelletti S, Ferrante P, Petriccione M, Firrao G. *Pseudomonas syringae* pv. *actinidiae*: a re-emerging, multi-faceted, pandemic pathogen. Mol Plant Pathol. 2012;13(7):631–40.

35. Vanneste JL. *Pseudomonas syringae* pv. *actinidiae* (Psa): a threat to the New Zealand and global kiwifruit industry. New Zeal J Crop Hortic Sci. 2012;40(4):265–7.

36. Do KS, Chung BN, Joa JH. D-PSA-K: a model for estimating the accumulated potential damage on kiwifruit canes caused by bacterial canker during the growing and overwintering seasons. plant Pathol J. 2016;32(6):537.

37. Birnie D, Livesey A. Lessons Learned from the Response to Psa-V. . Wellington, NZ Sapere Res Gr [Internet]. 2014; Available from: http://www.kvh.org.nz/vdb/document/100538

38. Vanneste JL. The scientific, economic, and social impacts of the New Zealand outbreak of bacterial canker of kiwifruit (*Pseudomonas syringae* pv. *actinidiae*). Annu Rev Phytopathol. 2017;55:377–99.

39. Balestra GM, Mazzaglia A, Quattrucci A, Renzi M, Rossetti A. Current status of bacterial canker spread on kiwifruit in Italy. Australas Plant Dis Notes. 2009;4(1):34–6.

40. Donati I, Buriani G, Cellini A, Mauri S, Costa G, Spinelli F. New insights on the bacterial canker of kiwifruit (*Pseudomonas syringae* pv. *actinidiae*). J Berry Res. 2014;4(2):53–67.

41. Marcelletti S, Scortichini M. Genome-wide comparison and taxonomic relatedness of multiple *Xylella fastidiosa* strains reveal the occurrence of three subspecies and a new *Xylella* species. Arch Microbiol. 2016;198(8):803–12.

42. Baudino C, Giuggioli NR, Briano R, Massaglia S, Peano C. Integrated methodologies (SWOT, TOWS, LCA) for improving production chains and environmental sustainability of kiwifruit and baby kiwi in Italy. Sustainability. 2017;9(9):1621.

43. Ferrante P, Fiorillo E, Marcelletti S, Marocchi F, Mastroleo M, Simeoni S, et al. The importance of the main colonization and penetration sites of *Pseudomonas syringae* pv. *actinidiae* and prevailing weather conditions in the development of epidemics in yellow kiwifruit, recently observed in central Italy. J Plant Pathol. 2012;94(2):455–61.

44. Vanneste JL, Yu J, Cornish DA, Tanner DJ, Windner R, Chapman JR, et al. Identification, virulence, and distribution of two biovars of *Pseudomonas syringae* pv. *actinidiae* in New Zealand. Plant Dis. 2013;97(6):708–19.

45. Serizawa S, Ichikawa T. Epidemiology of bacterial canker of kiwifruit, 3: The seasonal changes of bacterial population in lesions and of its exudation from lesion. Ann Phytopathol Soc Japan. 1993;

46. Young JM. *Pseudomonas syringae* pv. *actinidiae* in New Zealand. J Plant Pathol. 2012;94(1sup):1-5-S1. 10.

47. Renzi M, Copini P, Taddei AR, Rossetti A, Gallipoli L, Mazzaglia A, et al. Bacterial canker on kiwifruit in Italy: anatomical changes in the wood and in the primary infection sites. Phytopathology. 2012;102(9):827–40.

48. Marcelletti S, Scortichini M. Definition of Plant-Pathogenic Pseudomonas Genomospecies of the *Pseudomonas syringae* Complex Through Multiple Comparative Approaches. Phytopathology [Internet]. 2014;104(12):1274–82. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24875383

49. Petriccione M, Di Cecco I, Arena S, Scaloni A, Scortichini M. Proteomic changes in *Actinidia chinensis* shoot during systemic infection with a pandemic *Pseudomonas syringae* pv. *actinidiae* strain. J Proteomics. 2013;78:461–76.

50. Young JM. Taxonomy of *Pseudomonas syringae*. J Plant Pathol. 2010;S5–14.

51. Lelliott RA, Billing E, Hayward AC. A determinative scheme for the fluorescent plant pathogenic pseudomonads. J Appl Bacteriol. 1966;29(3):470–89.

52. Gardan L, Bollet C, Ghorrah MA, Grimont F, Grimont PAD. DNA relatedness among the pathovar strains of *Pseudomonas syringae* subsp. *savastanoi* Janse (1982) and proposal of *Pseudomonas savastanoi* sp. nov. Int J Syst Evol Microbiol. 1992;42(4):606–12.

53. Gardan L, Shafik H, Belouin S, Broch R, Grimont F, Grimont PAD. DNA relatedness among the pathovars of *Pseudomonas syringae* and description of *Pseudomonas tremae* sp. nov. and *Pseudomonas cannabina* sp. nov.(ex Sutic and Dowson 1959). Int J Syst Evol Microbiol. 1999;49(2):469–78.

54. Scortichini M, Ferrante P, Marcelletti S, Petriccione M. Omics, epidemiology and integrated approach for the coexistence with bacterial canker of kiwifruit, caused by *Pseudomonas syringae* pv. *actinidiae*. Ital J Agron. 2014;163–5.

55. Chapman JR, Taylor RK, Weir BS, Romberg MK, Vanneste JL, Luck J, et al. Phylogenetic relationships among global populations of *Pseudomonas syringae* pv. *actinidiae*. Phytopathology [Internet]. 2012;102(11):1034–44. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22877312

56. Cunty A, Poliakoff F, Rivoal C, Cesbron S, Fischer-Le Saux M, Lemaire C, et al. Characterization of *Pseudomonas syringae* pv. *actinidiae* (Psa) isolated from France and assignment of Psa biovar 4 to a de novo pathovar. Plant Pathol. 2014;

57. Sawada H, Fujikawa T. Genetic diversity of *Pseudomonas syringae* pv. *actinidiae*, pathogen of kiwifruit bacterial canker. Plant Pathol. 2019;68(7):1235–48.

58. Fujikawa T, Sawada H. Genome analysis of the kiwifruit canker pathogen *Pseudomonas syringae* pv. *actinidiae* biovar 5. Sci Rep. 2016;6:21399.

59. Balestra GM, Mazzaglia A, Rossetti A. Outbreak of bacterial blossom blight caused by *Pseudomonas viridiflava* on *Actinidia chinensis* kiwifruit plants in Italy. Plant Dis. 2008;92(12):1707.

60. Vanneste JL, Yu J, Cornish DA, Oldham JM, Spinelli F, Pattemore DE, et al. Survival of *Pseudomonas syringae* pv. *actinidiae* in the environment. In: I International Symposium on Bacterial Canker of Kiwifruit 1095. 2013. p. 105–10.

61. Sawada H. Diversity of pathogens causing bacterial canker of kiwifruit in Japan . Trends Sci. 2016;21(8): 8-2:103-110.

62. Sawada H, Miyoshi T, Ide Y. Novel MLSA group (Psa5) of *Pseudomonas syringae* pv. *actinidiae* causing bacterial canker of kiwifruit (*Actinidia chinensis*) in Japan. Japanese J Phytopathol. 2014;80(3):171–84.

63. Sundin GW. Genomic insights into the contribution of phytopathogenic bacterial plasmids to the evolutionary history of their hosts. Annu Rev Phytopathol. 2007;45:129–51.

64. Tuerkkan M, Dolar FS. Role Of Phytotoxins In Plant Diseases. TARIM Bilim DERGISI-JOURNAL Agric Sci. 2008;14(1):87–94.

65. Graniti A. Phytotoxins and their involvement in plant diseases. Introduction. Experientia. 1991;47(8):751–5.

66. Bender CL, Alarcón-Chaidez F, Gross DC. *Pseudomonas syringae* phytotoxins: mode of action, regulation, and biosynthesis by peptide and polyketide synthetases. Microbiol Mol Biol Rev. 1999;63(2):266–92.

67. Lindeberg M, Cunnac S, Collmer A. *Pseudomonas syringae* type III effector repertoires: last words in endless arguments. Trends Microbiol. 2012;20(4):199–208.

68. Choi S, Jayaraman J, Segonzac C, Park H-JH, Park H-JH, Han S-W, et al. *Pseudomonas syringae* pv. *actinidiae* type III effectors localized at multiple cellular compartments activate or suppress innate immune responses in *Nicotiana benthamiana*. Front Plant Sci. 2017;8:2157.

69. Ferrante P, Takikawa Y, Scortichini M. *Pseudomonas syringae* pv. *actinidiae* strains isolated from past and current epidemics to Actinidia spp. reveal a diverse population structure of the pathogen. Eur J plant Pathol. 2015;142(4):677–89.

2.7 Article 1:

Improved MLVA typing reveals a highly articulated structure in *Pseudomonas syringae* pv. *actinidiae* populations

Angelo Mazzaglia[1], Silvia Turco[1], Maria Claudia Taratufolo[1], Yaseen Jundi Rahi[1], Lorenzo Gallipoli[1], Giorgio Mariano Balestra[1]

[1]*Università degli Studi della Tuscia, Department of Agricultural and Forest Sciences (DAFNE), via San Camillo de Lellis snc, 01100 Viterbo, Italy*

[*]**Corresponding author**: angmazza@unitus.it

Summary

Pseudomonas syringae pv. *actinidiae* (Psa) is the causal agents of the bacterial canker of kiwifruit, probably the most limiting disease of this cultivation worldwide. Since its first isolation and description in 1984, the pathogen has been reported scattered in Southeastern Asia until 2008, when a new pandemic occurred affecting all the most important areas of kiwifruit cultivation in Europe, New Zealand and South America within few years. The consequent research boost elucidated different facets of this disease, including a thoughtful genetic characterization of the pathogen that has led, to date, to an intrapathovar distinction in five biovars, mainly circumscribed to Asiatic countries. Despite this, it is conceivable that additional genetic diversity still remains undiscovered most likely in those areas where the pathogen and the host plant species have coevolved. Here an enhanced Multi Locus VNTR Analysis (MLVA) assay based on 19 loci was applied to a comprehensive collection of 152 strains representative of all the main infected areas in the world and the results suggest that a high number of intrapathovar types of Psa possibly exists.

Keywords: bacterial canker, kiwifruit, *Actinidia* spp., biovars, molecular markers, tandem repeats, VNTR

Introduction

The canker of kiwifruit is an emerging infectious plant disease caused by the Gram-negative bacterium *Pseudomonas syringae* pv. *actinidiae (Psa)*, affecting the cultivations worldwide. A series of typical symptoms are prompted by Psa in kiwifruit plants, such as brown leaf spots with chlorotic haloes, whitish to reddish exudates on trunks and twigs, sudden death of young vines, collapse of fruits and brown discoloration of buds (Serizawa *et al.*, 1989). It was primarily isolated, recognized as a causal agent of the bacterial canker of kiwifruit, and described as a new pathovar in Japan in 1984 (Takikawa *et al.*, 1989). Almost at the same time, the disease was reported in Hunan province, China, in 1984 (Fang et al., 1990) and shortly after in Korea (Koh *et al.*, 1994). After these reports, many others have come during time testifying the progressive spread of the disease throughout these countries. However, ten years ago, the pathogen was confined to South-Eastern Asia, with a single occasional exception in Italy (Scortichini, 1994). In 2008 a new, very aggressive, Psa variant suddenly emerged in Italy, causing severe damages (Balestra *et al.*, 2009) all over the Country on both *Actinidia chinensis* var. *chinensis* and var. *deliciosa*. Since then an impressive progression of reports occurred all over the kiwifruit cultivated areas worldwide, up to the most recent in Argentina (Balestra *et al.*, 2017).

As a consequence of this hypervirulent pandemic, research efforts have been multiplied and several molecular approaches have been used to identify and genetically characterize the pathogen. In 2012, an intra-pathovar taxonomic division in 4 biovars has been proposed (Chapman *et al.*, 2012): the biovar Psa1, mainly referable to strains isolated during 80's in Japan, plus the Italian occurrence in 1992; the biovar Psa2, associated to Korean strains isolated in late 90's; the biovar Psa3, referring to the hypervirulent strains isolated in European

countries, New Zealand, China, South America, and later on in Korea (Kim *et al.*, 2016) and Japan (Sawada *et al.*, 2015); the biovar Psa4, indicating strains of paltry virulence on kiwifruit, later redefined as new pathovar *actinidifoliorium* (Cunty *et al.*, 2015b). An additional biovar, named Psa5, was described in 2012 (Sawada *et al.*, 2014), circumscribed to Saga Prefecture, in South Western Japan, and considered an endemic type of Psa. The same assumption was made for the biovar Psa6, described in 2015 in Nagano Prefecture, Japan, differing for the production of both phaseolotoxin and coronatine (Sawada *et al.*, 2016).

Regarding the origin of the pandemic outbreaks referable to Psa3, different hypotheses have been proposed. A first scenario advises for a Chinese origin of the epidemics, likely by independent events (Mazzaglia *et al.*, 2012; Butler *et al.*, 2013; McCann *et al.*, 2013). Recently, another prospect, relying on the similarity between Korean and Japanese strains belonging to Psa biovar 3 and the strains of the New Zealand outbreak, and on the higher number of biovars occurring in Japan and Korea respect to China, leans toward alternative origin for the Psa3 pandemic (McCann *et al.*, 2017).

In this complex and highly dynamic scenario, a fine-tuned depiction of how this pathogen is structured in diverging populations is essential to outline new occurrences promptly and track the pathogen's variants. The ever-increasing availability of bacterial genomes allows identifying genetic traits to be used as a marker of variability among populations belonging to the same taxon. Bacterial genomes were demonstrated to be crowded with DNA stretches of different lengths, repeated in a head-to-tail mode (Tandem Repeats - TR) and having high evolutionary speed. These repeats are particularly well-suited for epidemiological studies and the loci containing them, named Variable Number of Tandem Repeats (VNTRs), often organized in multiple contemporaneous assays (MLVA - Multiple Locus VNTR Analysis), have repeatedly proved successful for genotyping purposes (van Belkum *et al.*, 1998), including bacterial plant pathogens.

This molecular approach has been already implemented on Psa in the past. Ciarroni and colleagues (Ciarroni *et al.*, 2015) depicted an MLVA assay based on the analysis of thirteen VNTR loci allowing the discrimination of biovars Psa1, Psa2, Psa3 and Psa4 with additional information about relationships among individuals. In the same year, another MLVA assay based on eleven VNTR loci was independently developed to characterize Psa3 and Psa4 in France (Cunty *et al.*, 2015a).

Here, we merged the two assays in one and combined it with clustering methods and bioinformatics algorithms in order to establish an even more powerful tool for the assessment of the population's structure of the pathogen. The method was thereafter applied on a large selection of strains representative of all the main Psa reports, including all biovars as well as past and current outbreaks. The analysis has been approached progressively, starting from a selected set of strains whose biovar assessment has been thoroughly accomplished in previous literature; then, the results from their multiple clustering and multivariate approaches analysis have been used as guiding thread to investigate all the remaining isolates, avoiding any prior grouping assumption. By this approach, we aimed to draw up a clear and better-fine-tuned picture of the genetic variability of this dangerous pathogen.

.

Materials and methods

Pseudomonas syringae pv. *actinidiae* strains and DNA extraction

The list of the 152 Psa strains under examination, with details concerning geographic origin and time of isolation, is reported in Table S1. It includes a selection of 96 strains, 39 from 10 regions of China, 27 from 7 regions of Japan, and 30 strains from 13 different areas of Korea, isolated in a time range from 1984 to 2015, to represent the most of the temporal and spatial variability in Asiatic area. Then, 56 strains representatives of the outbreaks all over the world, isolated from 2008 up to 2017, were added. Among these, 34 individuals (marked with § in

Table S2) conclusively attributable to biovar Psa1 (4 strains), Psa2 (3 strains), Psa3 (19 strains), Psa5 (4 strains), and Psa6 (4 strains) were chosen as "control dataset" to check the assay affordability and to draw up a consistent bioinformatic pipeline for the analysis. *Pseudomonas syringae* pv. *actinidifoliorium*, ex-biovar 4, was not included in this analysis because of its genetic divergence from the other Psa, as already demonstrated (Ciarroni *et al.*, 2015; Cunty *et al.*, 2015a).

Each strain has been freshly grown on King's B medium for 48-72 hrs at 26 °C and about 200 mg of bacterial cells were collected for DNA extraction with a commercial kit (Macherey Nagel) following manufacturer's instructions. The DNA concentration was evaluated by a fluorometer (Qubit, Qiagen) in each sample and adjusted to the final concentration of 20 ng/µl for the following analytical steps.

PCR amplification of VNTR loci (MLVA) and capillary electrophoresis

The experimental plan provided for the amplification of 24 VNTR loci, 13 published in Ciarroni *et al.* (2015) and 11 in Cunty *et al.* (2015a), per each strain, using the primer pairs reported in Table 1.

During a preliminary *in silico* check emerged that four of these loci, as independently reported in the two papers above, were exactly the same. This concurrence was detected for VNTR loci Psa1=TR19II, Psa10=TR39II, GM1553=TR64II, and GM1834=TR10I. Hence, only one pair of primers per each duplicate was further considered (primers marked with § in Table 1), reducing the total number of loci under investigation to 20.

Each PCR reaction was composed of 12.5 µl of GoTaq Colorless Master Mix 2X (Promega Corporation, Madison, WI, USA), 1 µl of template DNA (20 ng), 1 µl of forward and 1 µl of reverse primer corresponding to 10 µM concentration each, 9.5 µl of nuclease-free water to the final volume of 25 µl. The PCR amplifications were performed on a C1000 thermal cycler (Biorad Laboratories Inc., Ca., USA) using the thermic profiles reported in Table S2.

The PCR protocols were primarily tested on a pool of 10 randomly selected strains. Four primer couples (marked with * in Table 1) from the paper of Cunty *et al.*, 2015a failed to amplify any sample; a further *in silico* check explained that the sequences of reverse primers in each couple, as reported in the paper, were directionally inverted. After correction, the respective VNTR loci were perfectly amplified. Once the efficiency of the protocols has been verified, PCR amplification of all the loci on all the strains was performed and repeated twice to test the reproducibility of the assay.

Table 1: Primer pairs used for the amplification of VNTR loci.

Name	Forward primer (5'→3')	Tm(C°)	Reverse primer (5'→3')	Tm(C°)	Ref.
Psa01	CAAGCAGGAGATGGAAGAGC	60,5	CATGCGGGCAATCTGATAGT	58,4	Ciarroni *et al.*, 2015
Psa03	TTATCGGCGGGATGTGTATT	56,4	ACTGCGTCTGGTCGATAACC	60,5	
Psa04	ACGAGTCCGCTCCTACAAAA	58,4	TACAACCAAGGTGGCCTGTT	58,4	
Psa05	GTAGGCCGCGCTTTAAT	56,1	TGCACTTCTTTTTCGCCTCT	56,4	
Psa06	TTAACGCAAGCAATCCTAACC	58,4	TGTGCAATAAATGCGGGTTA	60,5	
Psa07	GCCTACCTTTTACGCCATGA	55,9	CCGCCTCCAGTCAGGTTAAT	53,2	
Psa08	GTCATTGGCGAACTGATCCT	58,4	CTTTTCATGCTGAAAGTCATGC	58,4	
Psa09	CGCTGTCTGGCTTTGAAAAT	56,4	TAGGACGGCCGAAGGTTTAT	58,4	
Psa10	AAGCCTGAGTAAGCGGTTCA	58,4	GCCCCAGTCCCAGTTGTAAT	60,5	
GM254	CGTGTCACTGAAAGTCACCAT	58,4	TATTTACCCGGTGTTGAGGC	58,4	
GM1553	CTGGCACGAGACGAGTCC	60,7	GCTGAGCTTGAAGGAGACG	59,5	
GM1834	CGAGTTCTATTTGCGTCAGG	58,4	TGTCCAGCGTAATCTTGCTC	58,4	
GM4076	TGGGTGGAATACAGCCGCCA	62,5	CTTGTTCGGGAGCGGCAAGCT	65,3	
TR10I*§	AGTCTCTGCGCCTCAGGAT	58,8	GGCACTGGATTTTTCCAGAC	57,3	Cunty *et al.*, 2015
TR14I*	CTGGAAAACGTCCTGAGCAT	57,3	GTGAGTCAGGCAAACCGAGT	59,4	
TR15I*	TCGAGAGGAACACCAATGTG	57,3	TGGAAACATCGTCTGCAAAA	53,2	
TR30I*	ACGTTACTTCGAGCGGAGTC	59,4	CCGACCTACCCGAATATGTG	59,4	
TR1II	TGCCTGAGTACCTTTACCGG	59,4	CACCCAGCTCGACAATCAAG	59,4	
TR2II	GTCATAACGGGTGAGAGTGC	59,4	ACGGCCCTTGAAAGTGACTA	57,3	
TR3II	CGTGAGGCTCTGACTTTCTG	59,4	AAATCCGGGCTGTTTATCGC	57,3	
TR39II§	CGGTGGACTTGAAGAACACG	59,4	CACCCTGAACTGATTGCACC	59,4	
TR11II	GATTGGTGACGTTGCGATGA	57,3	TTGTTGCCCTACACGCTCTA	57,3	

TR19 II§	CCCAGAAAGAATGCGGACTG	59,4	AGCAGGAGATGGAAGAGCTG	59,4	
TR64 II§	GTTGGCGGGTATGTGTCTG	58,8	CACCACGCTTCTTCTTGCAG	59,4	
* for these loci, the reverse primer was corrected respect to what reported in Cunty *et al.*, 2015; § these primer couples were discarded because amplifying the same TR of primer couples GM1834, Psa10, Psa1 and GM1554 respectively.					

The PCR amplicons have been visualized and their size assessed by capillary electrophoresis using QiaXcel system (QIAGEN, Milan, Italy). Results were interpreted by means of Screengel software (QIAGEN). The tandem repeat number was calculated subtracting the flanking region size from the amplicon size and then dividing the remaining by the repeat unit length. Per each locus, a random selection of strains was Sanger sequenced to confirm the exactness of previous calculations.

Data analysis

Data elaboration has been postulated for three independent approaches: hierarchical clustering, STRUCTURE, and Discriminant Analysis of Principal Components (DAPC). Firstly, it was entirely performed on the "control dataset" to test the efficiency of the assay and the robustness of the analytical pipeline. Then, the same whole analytical approach was repeated including all the remaining strains, for which the lack of any prior knowledge of group assignment was assumed.

The two data matrices for the "control dataset" and the "complete dataset", were imported into R version 3.4.4 and transformed into *genind* objects using the R package *adegenet* 2.1.1 (Jombart, 2008). All the analyses of population structure and genetic diversity were performed using R packages *poppr* version 2.8.1 (Kamvar *et al.*, 2014, 2015), *adegenet* 2.1.1 (Jombart, 2008) and *ade4* version 1.7-13 (Dray & Dufour, 2007).

The genotype accumulation curve was obtained by *informloci()* and *mlg.table()* functions of the *poppr* package to assess if the number of loci is sufficient to explain all the observed variability (MLGs). Standard statistics were calculated by *poppr()* function, whilst Genotypic richness and Evenness were evaluated using the R package *vegan* 2.5-6 (Oksanen *et al.*, 2019).

Hierarchical clustering

The genetic inter-individual distance computation obtained without any prior information about the population structure for the individuals under examination constituted the basis for clustering. For molecular markers as tandem repeats, Bruvo's distance is specifically appropriate for considering the repeat length and for being insensitive to different ploidy levels (Bruvo *et al.*, 2004). Hierarchical clustering was performed with *hclust()* function of the R package *stats* using Bruvo's distance to create genetic distance matrices and UPGMA as algorithm for agglomerative clustering. Bruvo's distances were bootstrapped using the *poppr bruvo.boot()* function and a cut-off threshold of 80 was set. Dendrograms were visualized with the R package *factoextra* version 1.0.5 (Kassambara & Mundt, 2016).

Starting with the "control dataset", the genetic distance threshold able to distinguish the biovar association of strains was assessed by calculating the limits at which the multilocus lineages (MLLs) can be progressively resolved. The *poppr mlg.filter()* and *cutoff_predictor()* functions determined the largest difference between thresholds and the average value between these values was the selected cut-off threshold separating clonal lineages that were subsequently imposed in the hierarchical clustering of all the individuals ("complete dataset") to discriminate populations.

STRUCTURE

The same dataset was independently tested by STRUCTURE software, version 2.3.4 (Falush *et al.*, 2007), a well-known Bayesian model-based algorithm for clustering many types of genetic data. If the number of groups (K) is not user-defined, it has to be heuristically chosen to compare the results of penalized log-likelihoods in a defined range of possible clusters.

Here, the optimal K was estimated with Evanno method (Evanno *et al.*, 2005) on 10 iterations for each K arranged in a range of values using the *evannoMethodStructure()* function in the R package *pophelper* (Francis, 2017). Once K was set, the differences in genetic variants and the

proportional assignment to multiple populations ("admixture model") were evaluated by Markov Chain Monte Carlo (MCMC) estimation. The analysis parameters were set to 1.000 (burnin period) and 10.000 (MCMC reiterations) for the whole dataset. Finally, ten reiterated outputs for the chosen K were combined and merged by *clumppExport()* function of CLUMPP and visualized using the function *plotQ()* of the R package *pophelper*.

Discriminant Analysis of Principal Components (DAPC)

Data were also explored by DAPC, a multivariate method to infer the population structure of clonal organisms by determining the number of clusters observed without any prior knowledge (Jombart *et al.*, 2010; Grünwald & Goss, 2011). All the DAPC analysis was carried out with the R package *adegenet* (Jombart, 2008). Then, the posterior assignment probability was calculated and visualized by *compoplot()* function. Cross-Validation by the function *xval.Dapc()* checked if the correct number of principal components was retained.

Results

VNTR loci amplification

Twenty VNTR loci were amplified on 152 Psa strains, for a total of 3040 reactions. PCR amplifications yielded correctly a single amplicon in all the loci confirming the specificity of the primers used, and the exact dimension of each amplicon was assessed by capillary electrophoresis. Among them, the locus TR-15I yielded a unique invariable amplicon of 164 bp on all the strains in agreement with results from Cunty et al. (2015a), where this locus shows variability only toward *Pfm.* Being uninformative, it was discarded, and 19 loci were kept for downstream analyses. Amplification failures were observed at locus Psa-04 for six Japanese strains isolated from 1982 to 1988, and for the single Italian strain isolated in 1992, and at locus Psa-09 for seventeen Korean strains. Both instances were already reported in Ciarroni *et al.*, 2015. In addition, Locus TR30. I failed to amplify all the strains belonging to the biovar Psa5. The amplification failure was coded as "0". The Sanger sequencing of randomly selected

amplicons confirmed the exactness of repeat numbers as assessed by capillary electrophoresis. The features of each locus are reported in Table 2.

Table 2: Features of the VNTR loci.

VNTR locus	TR length	Flanking left	Flanking right	Flanking size
PSA01	7	63	97	160
PSA03	7	86	66	152
PSA04	33	66	58	124
PSA05	7	86	68	154
PSA06	8	68	55	123
PSA07	8	39	25	64
PSA08	9	72	72	144
PSA09	99-112	26	63	89
PSA10	7	79	67	146
GM254	8	109	226	335
GM1553	7	139	32	171
GM1834	6	118	58	176
GM4076	7	28	156	184
TR14I	6	53	52	105
TR15I	13	76	62	138
TR30I	7	44	50	94
TR1II	8	128	102	230
TR2II	9	105	126*	231
TR3II	9	58	69	127
TR11II	11	94	42	136

Results were coded in a matrix of 152 lines (strains) reporting the ordered number of repetitions per each of the 19 loci and stored as a "complete dataset" (Table S3). From this dataset, 34 individuals conclusively attributable to the 5 known biovars (Table S1 - marked with [§]) were selected as "control dataset" to check the assay's affordability and to draw up a consistent bioinformatics pipeline for the analysis.

Among the 152 individuals analysed in this study, a total of 129 alleles were identified across the 19 loci. Each individual is represented by a single independent multilocus genotype (MLG); however, checking for the presence of clones, an inspection with *mlg()* function showed that the 152 individuals in "complete dataset" shared only 85 different MLGs, whilst the 34 individuals in "control dataset" shared only 22; any clone replication was removed with *clonecorrect()* function.

The genotype accumulation curve shown in Figure S1 confirmed that all the loci are polymorphic and beneficial to describe the diversity among the 85 multilocus genotypes (MLGs). The curve shows that the plateau has been reached, and the addition of more loci would not increase the discriminatory power of the assay (Kamvar *et al*., 2015).

Testing the genetic structure of *Pseudomonas syringae* pv. *actinidiae* on the "control dataset"

The hierarchical clustering on the "control dataset" constructed on Bruvo's distance and UPGMA unequivocally led to a dendrogram constituted by five clear clusters, all with 100% bootstrap support and all the individuals were distributed in five groups perfectly fitting to the five known biovars (Figure 1). Proven the consistency of the assay to correctly assign strains to biovars in "control dataset", it was possible to determine the most affordable value of genetic distance below which two individuals are considered belonging to the same group.

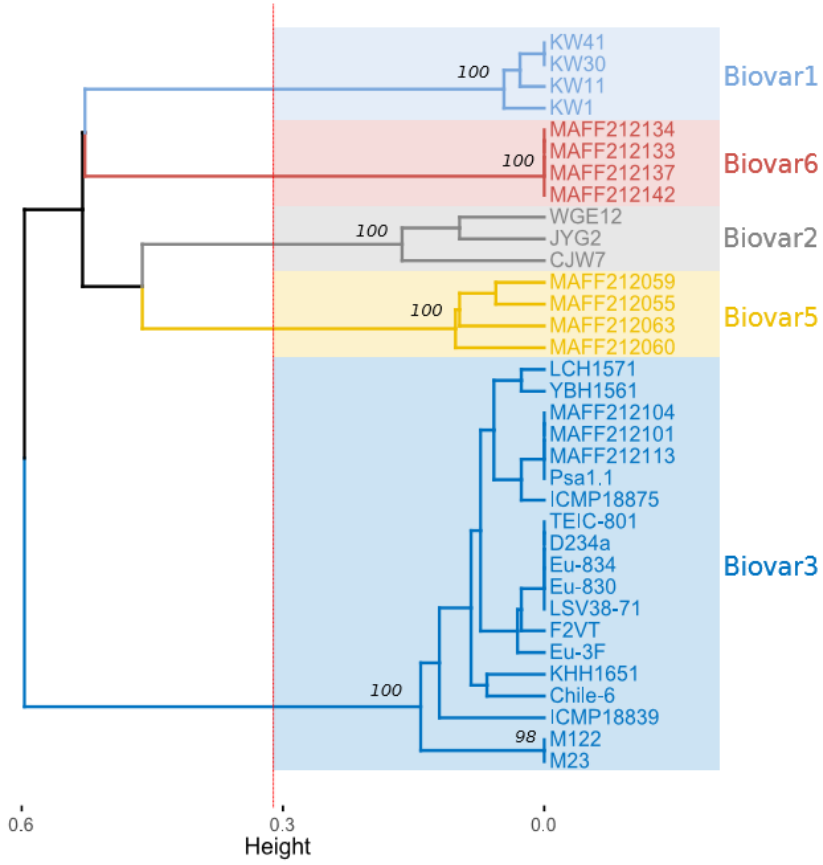


Figure 8: UPGMA dendrogram of Bruvo's genetic distance between a selection of 34 individuals of *P. syringae* pv. *actinidiae*, created using 1,000 bootstrap replications. The red dot line indicates the cut-off threshold calculated by mlg.filter().

All of the distance thresholds progressively resolving the 22 unique multilocus lineages (MLLs) were elaborated by *cutoff_predictor()* function, and based on Bruvo's distance and UPGMA as clustering algorithm, the largest gap between distances was the one collapsing the 22 MLGs into 5 MLLs corresponding to the 5 biovars, and the related average cut-off threshold was 0.3124 (Figure 2).

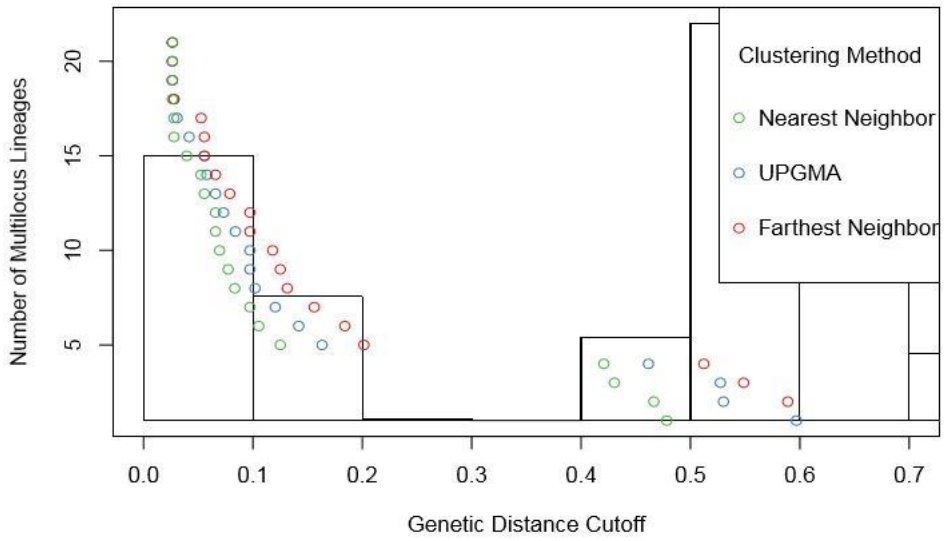


Figure 9: Graphical representation of clustering algorithms (UPGMA was selected for this study) collapsing MLGs of the "control dataset" in MLLs. The x-axis represents Bruvo's distance cut-off values, the y-axis the number of MLLs observed at different values of genetic distance.

When the "control dataset" was analysed by STRUCTURE, 10-fold iterated runs of K from 2 to 8 were performed. The ΔK was plotted to establish the number of clusters that best describe the data according to Evanno's criterion and the most affordable value was K=5 (Figure S2a). Consequently, the STRUCTURE results of 10-iterated output for K=5, combined and merged by CLUMPP, produced the barplot graph describing a population structure perfectly fitting with the previous clustering (Figure S2b).

The Bayesian information criterion (BIC) as a function of the number of clusters in the Discriminant Analysis of Principal Components (DAPC) also indicated K=5 (Figure S3a). Four principal components and two discriminant eigenvalues were retained in agreement with the cross-validation analysis. The DAPC scatterplot graph (Figure S3b) shows five plainly separated groups, each constituted by tightly linked individuals. The barplot of posterior assignment (Figure S3c) reporting membership coefficient of individuals excluded any admixture, and each group as its respective individuals matched exactly the five biovars.

New individuals: population assignment without a priori constraints

Since the main aim of the paper was to test the MLVA assay for positioning the strains in population without any prior constraint, 118 Psa individuals with a large span of geographic and time origins, were added to the "control dataset" of 34, using their MLGs as one and only information, to restore the "complete dataset" of 152 individuals. This dataset underwent to the same analytical pipeline described above.

Eighty-five different MLGs on 152 individuals were detected confirming the high clonality of the pathogen in general, and of Psa3 in particular. In fact, in most cases, identical MLGs included individuals isolated in the same country or in the same geographic region (Table S4). Interestingly, the MLG71, the most abundant of Psa3, is shared by 13 isolates from very different countries, i.e. Japan, New Zealand, and Argentina.

The dendrogram obtained by hierarchical clustering using Bruvo's distance and UPGMA running 1000 bootstrap replications is shown in Figure 3. It is worth noting how the 34 strains of the "control dataset" preserve their clustering structure being allocated in five different clusters. When the cut-off threshold of 0.3124 obtained with a "control dataset" was applied in this hierarchical clustering, a new articulated population structure appeared. The contingency table showing the individual's distribution among the inferred groups using this threshold assigns most of the unknown individuals to biovar Psa1, biovar Psa2 and, particularly, biovar

Psa3, but not to the biovars Psa5 and Psa6. According to this threshold, the number of populations required to allocate all the individuals increased to fourteen (Figure S4).

According to the clustering, one group include all the biovar Psa1 strains isolated in Japan in 1984 together with the Italian strain isolated in 1992, three additional strains from Japan, two isolated in 1988 (Pa429 and PA459) and one isolated in 2011 (Jap2820), plus two strains from Korea isolated in 1989; this group was assumed as biovar Psa1.

The cluster including the three strains documented as biovar Psa2, contains fourteen additional strains, exclusively from Korea, isolated from 1999 to 2013 in Jeju-do and Jeollanam-do provinces; this group represents in this study the biovar Psa2.

The most populous cluster contains 79 strains including all the strains from European, New Zealand and South American outbreaks without any exception, 9 from Japan isolated in 2014 and reported as belonging to biovar Psa3 (McCann *et al.*, 2017), 4 from Korea isolated in 2014-2016 also referable to biovar Psa3 (Kim *et al.*, 2016), and 11 from China in 2010 (Shaanxi), in 2012 (Sichuan), in 2013 (Anqing) and in 2017 (Kunmin); all the 19 Psa3 strains from "control dataset" are coherently included among them. Thus, we attributed this entire group to biovar Psa3. Remarkably, even if with modest statistical support, a further division of this cluster in three subgroups is conceivable, one constituted by European strains, another with Chilean strains, and another including strains from Shaanxi (China), New Zealand, Argentina, Japan, and Korea. Then, some strains from different Chinese regions form separate subclusters within biovar3.

Biovars 5 and 6 form two independent clusters composed exclusively by their representative strains, both fully supported by statistics (100% bootstrap).

Besides these groups coherent with current biovars, nine clusters remained independent with a strong statistical support and thus were named as populations from 1 to 9 (Pop1-Pop9), constituted only by strains from South-East Asian countries. The population Pop9 includes 12

strains from Shaanxi, Sichuan, Anhui and Shanghai regions of China; all the others are singletons Pop3 and Pop8 (Korea) or composed by few representatives, with the same geographic and time origin, such as Pop1 (Japanese), Pop4 and Pop5, (Chinese), Pop6 (Korean) or even admixed as Pop2 (Jap2726 from Japan; KN2 from Korea) and Pop7 (M218, M228, and JZHY7 from China; KTS1471 from Korea).

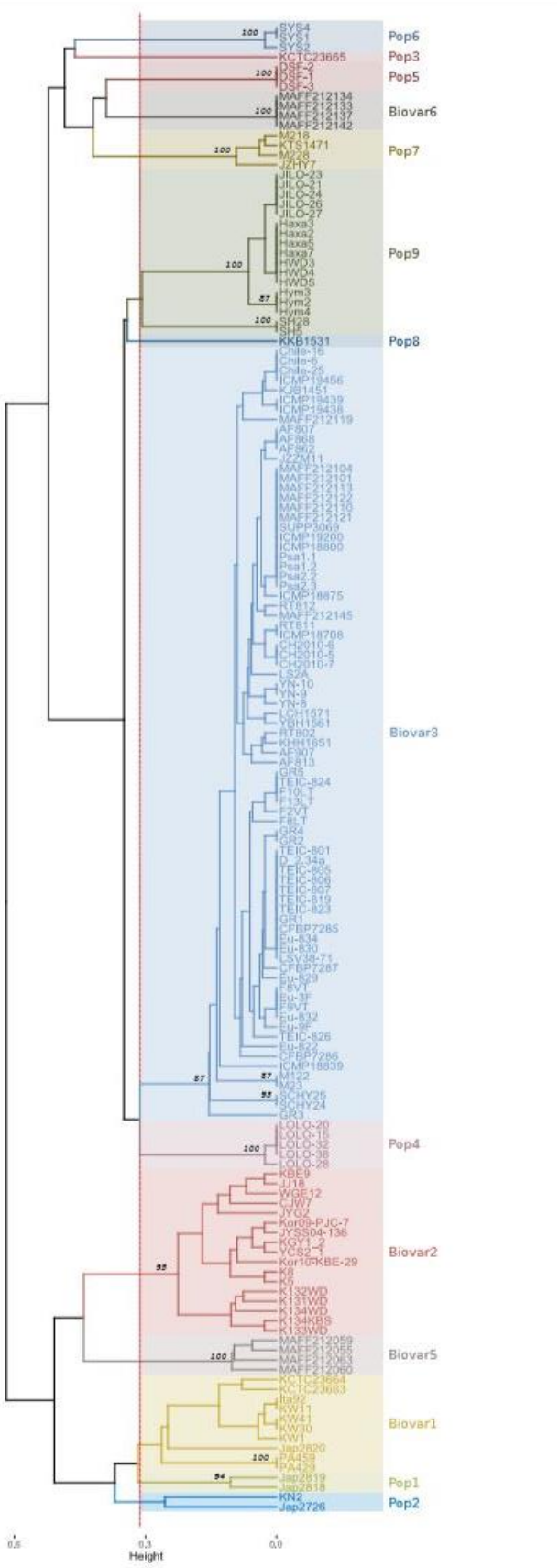


Figure 10: UPGMA dendrogram of Bruvo's genetic distance between 152 individuals of *P. syringae* pv. *actinidiae*, created using 1,000 bootstrap replications. The red dot line indicates the cut-off threshold calculated by mlg.filter().

When the same dataset was analysed by STRUCTURE, Evanno's method didn't provide a clear preferable number of clusters to retain. However, because K=14 returned the highest ΔK value ranging K from 5 to 15, in agreement with hierarchical clustering, we analysed this population arrangement. The 10 iterated output results for K=14, combined and merged by CLUMPP, showed a distribution of individuals in the populations that almost perfectly fits with the results of hierarchical clustering (Figure 4). In detail, exact matching was obtained for individuals of biovars Psa2, Psa3, Psa5 and Psa6, and for populations Pop3, Pop4, Pop5, Pop6, Pop7, and Pop8. Only in biovar Psa1 two strains, PA429 and PA459, form an independent group, whilst strains Jap2819 and Jap2726, previously parts of Pop1 and Pop2 respectively, are newly included in it. Also, 2 strains from these populations, Jap2818 and KN2, cluster in group 3.

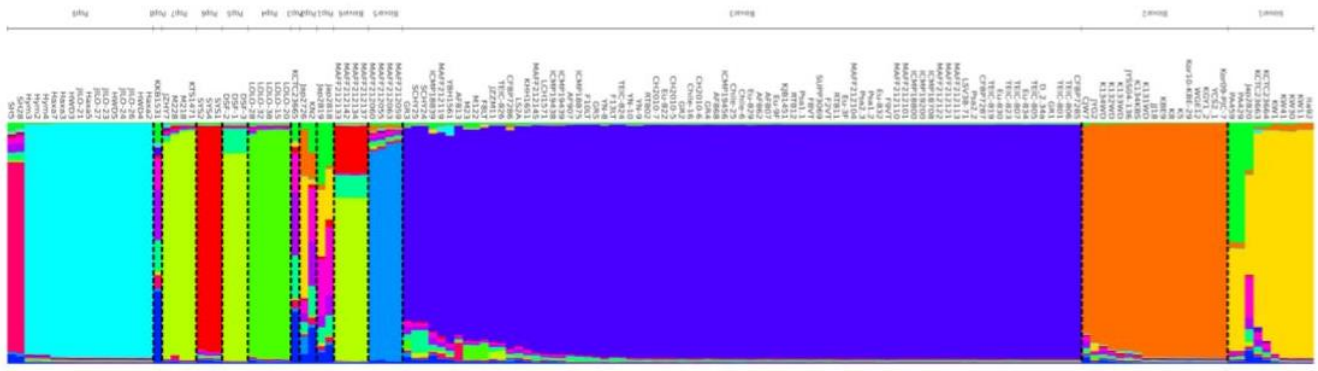


Figure 11: Population genetic structure of *P. syringae* pv. *actinidiae* obtained with STRUCTURE and CLUMPP when K = 14, burnin period = 1000 and MCMC reiterations 10000.

It is not accidental that all these strains show a highly admixed posterior assignment, as disclosed by their multiple colours bars in Figure 5. More evident in a posterior assignment is the exclusion of strains SH5 and SH8 from Pop9, leading to the newly independent group 11.

DAPC came out with comparable results too. The BIC, as for ΔK above, didn't suggest a distinctive number of groups, but 14 remained among the most conceivable solutions for this analysis (Figure 5a). Cross validating this value, 18 principal components, sufficient to explain 77,8% of the whole variance, and 4 Linear Discriminants were retained. The resulting scatterplot (Figure 5b) shows the distribution of the groups on LD axes, while the posterior assignment is reported as a barplot in Figure 5c. Group 10 includes all Psa1 strains except Jap2820; groups 9 and 14 includes Psa2 individuals, however strongly intermixed. Individuals referable to Psa3 are split in groups 4 and 5, discriminating European strains and strains from any other origin, in partial agreement with results from hierarchical clustering; also, DAPC places the strains SCHY24 and SCHY25 in a separate group (group 7). Then, group 12 matches exactly biovar Psa5 and group 13 biovar Psa6.

Concerning the other populations, DAPC joins together in group 1 strains from Japan and Korea, defined as Pop1, Pop2, and Pop3 in the hierarchical clustering, and in group 2 strains from China and Korea, which include Pop7 and part of Pop9. The remaining groups 3, 8, and 11, perfectly match those obtained in previous analyses, Pop5, Pop8 and Pop4 respectively.

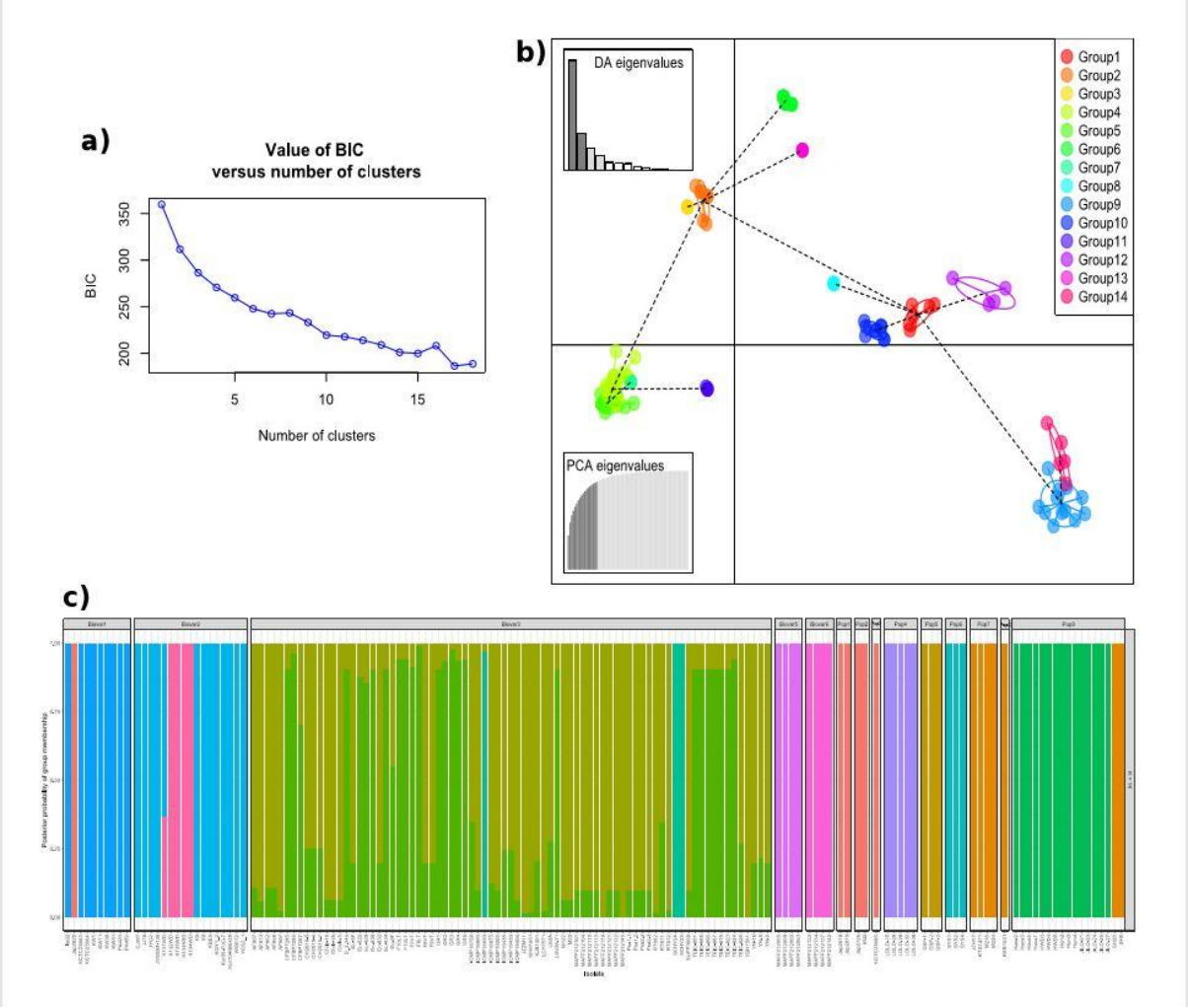


Figure 12: Discriminant analysis of principal component (DAPC) for 152 *Pseudomonas syringae* pv. *actinidiae* individuals. a) Bayesian information criterion (BIC) as a function of the number of clusters (K). b) Scatterplot of individuals arranged in 14 different clusters. c) Posterior assignment probability of individuals when K = 14.

The assignment of each of the strains to populations according to the three analytical methods above described is summarized in Table S5.

Genetic diversity is depicted by the *poppr()* function as genotypic richness, genotypic evenness, Nei's (Hexp) and Simpon's (λ) indices, as reported in Table 3. The highest of Simpson's diversity (λ) was observed in Psa2, followed by Psa3. On the contrary, Psa6, represented by four clones, showed a λ equal to zero. Genotypic evenness E.5 is a function of the MLGs frequencies, equal to 1 when all genotypes are present at equal frequencies and to 0

when there is only one MLG. Indeed Psa3, with the highest number of MLG, showed an Evenness close to 0.6 meaning that the 37 MLGs present don't have the same frequency. Nei's diversity index resulted to be highest in Psa2 and zero for Psa6 and Pop4.

Discussion

Bacterial diseases always represent a threat to many plant species, but in the last decades numerous cases of sudden and pernicious epidemics occurred worldwide, thus falling in the so-called "emerging infectious diseases". Typically, a brand-new interaction between a pathogen and a susceptible host is at the origin of these emergencies, but it is in turn related to the movement of pathogens in new territories or the introduction of plants out of their natural range, to the erosion of genetic diversity for breeding, or to environmental changes that modify the distribution areas of plants and pathogens.

The bacterial canker of kiwifruit by *Pseudomonas syringae* pv. *actinidiae* is an example of these. China is the native environment of almost all the *Actinidia* species, but kiwifruit wasn't considered for farming there for a very long time. Instead, it was New Zealand the first country where the economic potential of kiwifruit cultivation and commercialization took place since the middle '60s, particularly with *deliciosa* variety derived from a single seed introduction in 1904. Then, all the other countries worldwide started to cultivate kiwifruit extensively using plant material with an extremely homogeneous genetic basis (Ferguson & Bollard, 1990). In this context, the bacterial canker and its causal agent were reported and described since the middle '80s in South-eastern Asia, and then all over the world after 2008.

The ensuing investigations described five biovars (plus the former biovar 4, now *Pfm*) to explain the diversity within the pathogen. However, *Pseudomonas syringae* referable bacteria possess dramatic plasticity in their genetic background allowing themselves to adapt to virtually all the existent environmental niches (Baltrus *et al.*, 2017). Thus, also considering the

vastness and the variability of the geographic area where *Actinidia* species reside, a more articulated intrapathovar structure for Psa is conceivable.

Here, a molecular tool based on selectively neutral, locus-specific, highly polymorphic and easily repeatable markers, specifically suited for bacterial population studies was arranged. In our MLVA assay, repeats of different lengths were included, assuring variable mutation rates, but rarely exceeding 20 nucleotides, to keep homoplasy risk to a minimum. Outcoming data were then explored by three independent analytical methods, which led to groupings having significant reciprocal consistencies.

Strains belonging to biovar Psa1, besides MLST categorization, have a specific phytotoxin pattern and a genomic island containing the gene cluster for their biosynthesis (Sawada & Fujikawa, 2019). This biovar was found infecting *Actinidia arguta*, one of the 5 wild species of *Actinidia* naturally occurring in mountain regions of Japan (Ushiyama *et al.*, 1992), thus likely being a Japanese endemism, even if related strains were reported during time also in Italy (Scortichini, 2014) and Korea (Mazzaglia *et al.*, 2012; McCann *et al.*, 2017). Our MLVA assay correctly exemplifies this situation with a cluster formed by a core of strains from the first Japanese report of PSA and the Italian strain, together with other old Japanese and Korean strains. Two groups constituted by Japanese strains isolated in Chubu (2011) and one Korean strain isolated in 1997, here indicated as Pop1 and Pop2, clustered just beyond the genetic threshold imposed by the clustering algorithm, suggesting a clear connection to biovar Psa1.

The biovar Psa2 is instead plausibly native of Korea, where it was reported first in 1988 in Jeju island (Koh *et al.*, 1994) and afterward in the other Korean kiwifruit cultivation areas (Koh *et al.*, 2010). To date, Psa2 has never been reported elsewhere. It is characterized by the presence of a cluster of genes for coronatine production (Han *et al.*, 2003). The strains belonging to Psa2 group together in our analysis; however, strains isolated from 1999 to 2010 in different location of Jeollanam and Jeju provinces are separated from strains isolated in Wando, Jeonnam

province, during 2013, in both hierarchical clustering and DAPC with proper statistical support, possibly representing two somewhat different Psa2 groups.

The group corresponding to biovar Psa3 has been steadily recognized with all the grouping methods. First, it includes all the strains isolated from any outbreaks outside of Southeastern Asia without any exception. Then, it includes all the strains representatives of documented Psa3 outbreaks in the last years in both Korea and Japan, i.e. the strains from the Korean outbreak occurred since 2014, first in Jeju Island and in Jeollanam and Gyeongnam provinces soon after, as result of a possible introduction of contaminated pollen from China (Kim *et al.*, 2016), and the eight strains representing the first Psa3 outbreak in Japan in 2014 (Sawada et. al., 2015). The Psa3 cluster also includes strains from several areas of China, i.e. Shaanxi, Sichuan, Anqing, and Kunmin. Some of them, isolated in Shaanxi in 2010, has been correlated to the New Zealand outbreak (Butler *et al.*, 2013; Ciarroni *et al.*, 2015), but the remaining has never been described before. A significant step forward to understand the main routes of transmission of Psa3 has been made through the comparison of Integrative and Conjugative Elements (ICEs) providing evidences for three independent contamination events that led to epidemics in New Zealand (Pac_ICE1), Europe (Pac_ICE2) and Chile (Pac_ICE3) (Butler *et al.*, 2013; Butler & Poulter, 2015). Our hierarchical clustering not only distinguished three sub-clusters within Psa3, fitting perfectly with this subdivision but also portrayed additional independent groups epitomizing the existence of further genetic variants in China, as also recently reported (He et al., 2019). As for biovar Psa1, two groups and 1 singleton (Pop4, Pop9 and Pop8, respectively in Figure 3) cluster autonomously just above the genetic threshold that outlines the border of biovar Psa3 group, advising for some relationship with this. One of them, Pop4, is constituted by strains isolated in 2012 in Guizhou, a mountain region in the southwest of China characterized by a multifaceted topography and different microclimates that gave rise to a great diversity in the taxa of the genus *Actinidia* (Ferguson & Huang, 2007); for the same reason, it

seems plausible that a distinct variant of Psa, or rather of Psa3, has evolved there. The second cluster, Pop9, is conversely composed by strains isolated in three regions of Central China, all characterized by temperate climate: Shaanxi (Xi'an), Sichuan (Deyang), and Anhui (Anqing, Yuexi). This group deserves particular attention for being distributed in the most important kiwifruit cultivation areas of China, probably since a long time ago, and for being consequently at high risk of diffusion abroad. Within Pop9 fall also two strains from the Jiading district, a suburban area of Shanghai, where *Actinidia* cultivation is factually absent and thus the infected plants were probably introduced from elsewhere. The singleton KKB1531, here Pop8, is a unique strain isolated in 2015 from red-fleshed kiwifruit in Bongkae, Korea, whose origin cannot be further ascertained.

Since 2010 a new biovar named Psa5, with distinct physiological and genetic features, was detected on yellow-fleshed kiwifruit plants cv. Hort16A in the Northwest part of the Kyushu Island (Saga prefecture) with symptoms less severe than usual; genomic studies demonstrated a relevant genetic affiliation with biovar 2 (Fujikawa & Sawada, 2016). However, despite the intense monitoring of Psa carried out in all Japan, Psa5 was never found elsewhere (Sawada & Fujikawa, 2019). Our data fit again with all these observations grouping these strains independently with 100% bootstrap support but being at the same time rather correlated to biovar 2.

A similar history took place for Psa6 when diseased plants were noticed in a central area of Honshu Island on green-fleshed *Actinidia deliciosa* cv Hayward in 2015. The associated strains again showed peculiar phenotypic and genetic features justifying its acknowledgment of new biovar Psa6 (Sawada *et al.*, 2016). Here the four Psa6 isolates shared an identical haplotype in this analysis and coherently clustered apart from the others with full statistical support in all the attempted approaches.

Finally, three additional groups and one singleton come distinctly out from any clustering referable to known biovars as above described.

The group Pop6 is constituted by three Korean strains isolated in Goheung, Jeollanam province, from diseased plants of Yellow-king and Hongyang cultivars and reported as first occurrence of Psa3 in Korea in 2011 (Koh *et al.*, 2012) as consequence of the introduction of infected plant material from China (Kim *et al.*, 2016). The apparent incongruency of clustering outside Psa3 group is instead a further proof of the reliability of our analytical approach; a genomic study indeed demonstrated that they are distinguishable from other Psa3 strains for the presence of a very large plasmid and a different Pac_ICE, named Pac_ICE7_kr (Butler *et al*., 2015), thus belonging to a virulent Psa alternative to Psa3.

Another distinct group, Pop5, includes few strains from Shimen county, Hunan province, in the south side of the Yangtze River, where only a few scattered *Actinidia* orchards are present (Ferguson & Huang, 2007). Again, it is conceivable that geographic isolation has led to an endemic form of Psa.

The group Pop7 encompasses four strains, three from China (Xianyan and Baoji, Shaanxi, 2010 and Anqing, Anhui, 2013) and one from Korea (Jocheon, Jeju Island, 2014), proving evidence of another lineage shared among different countries, which could have been transmitted in recent times.

Finally, the last singleton (Pop3) appears in the analysis, the strain KCTC23665 isolated in Jeollanam, Korea, in 1989, which unexpectedly is separated from strains KCTC23663 and KCTC23664 of the same geographic and time origin. This discrepancy, hardly explainable by consistent biological means, might be related to inaccuracies during routine renewals in long-time storage.

Considering the above, the MLVA enlarged assay, coupled with the approaches to data analyses here described, proved a strong consistency with most of the Psa population structure

as currently depicted by many different studies and regardless of the used methods. But it also is the reason why, in our opinion, the new suggestions arising from this research should have to be duly taken into account. Indeed, this study provides evidence that the genetic variability of Psa is still far from being exhaustively described.

Whilst it is confirmed that only Psa3 has outcrossed the boundaries of the south-eastern Asiatic area becoming responsible for the worldwide pandemic, nonetheless, several other types of Psa seem to be scattered throughout Asia. Our present results seemingly advise for a substantial equilibrium in genetic richness in this area, being the Chinese strains scattered in five, Japanese in six and Korean in eight of the 14 groups depicted. However, we still deem the highest variability resides in China for various reasons. Firstly, at least some of the outbreaks in Korea and Japan, i.e. those referable to Psa3, are reported as consequences of the introduction of contaminated material from China (Koh et al., 2012; Kim et al., 2016; Sawada & Fujikawa 2019). The second descends as ecological deduction, considering both the enormous diversity of natural environments where kiwifruit grows in China, in term of land vastity, richness of diverse biological niches, and number of domesticated and wild *Actinidia* species, and that Psa is likely capable of infecting all of them as well as other wild plant species there (Liu *et al*., 2016). This also reflects on the hypothesis that all the research efforts aimed to monitor the disease and to estimate Psa diversity in both Japan and Korea could have almost completely elucidated their respective contexts, while we could have just seen the tip of the iceberg about the Chinese situation. Feasibly, only in-depth investigation on wild *Actinidia* species in natural uncontaminated environments scattered through south eastern Asia could shed light on this very intricate scenario. In the light of these considerations, the improvement of specific and highly sensitive molecular typing tools, like the assay reported here, could make easier and more effective not only to investigate about the presence of Psa but even to discriminate

thoroughly its variants in order to track the diffusion pathways of known populations as well to promptly recognize the advent of new conceivable Psa types

Table II**:** Poppr population statistics.

Pop	N	MLG	lambda	E.5	Hexp
Biovar1	10	7	0.84	0.938	0.264
Pop2	2	2	0.5	1	0.421
Pop1	2	2	0.5	1	0.211
Biovar3	79	37	0.93	0.597	0.128
Biovar5	4	4	0.75	1	0.158
Biovar6	4	1	0	NA	0
Pop3	1	1	0	NA	NA
Biovar2	17	17	0.94	1	0.271
Pop6	3	2	0.44	0.899	0.035
Pop7	4	4	0.75	1	0.123
Pop8	1	1	0	NA	NA
Pop4	5	2	0.32	0.725	0.021
Pop9	17	4	0.7	0.89	0.145
Pop5	3	1	0	NA	0
Total	152	85	0.97	0.626	0.537

Acknowledgments

We are grateful to Dr. Takashi Fujikawa, Institute of Fruit Tree and Tea Science, National Agriculture and Food Research Organization (NARO), Japan, for providing DNAs of strains belonging to biovar 6. We thank Dr. Alessandro Grottoli for technical support. The research was carried out in the frame of the MIUR (Ministry for Education, University and Research) initiative “Department of Excellence” (Law 232/2016). The authors declare that no competing interests exist.

2.8 References in article 1

Balestra GM, Buriani G, Cellini A, Donati I, Mazzaglia A, Spinelli F, 2017. First Report of *Pseudomonas syringae* pv. *actinidiae* on kiwifruit pollen from Argentina. *Plant Disease* **102**, 237–237.

Balestra GM, Mazzaglia A, Quattrucci A, Renzi M, Rossetti A, 2009. Current status of bacterial canker spread on kiwifruit in Italy. *Australasian Plant Disease Notes* **4**, 34–36.

Baltrus DA, McCann HC, Guttman DS, 2017. Evolution, genomics, and epidemiology of *Pseudomonas syringae*. *Molecular Plant Pathology* **18**, 152–168.

Belkum A van, Scherer S, Alphen L van, Verbrugh H, 1998. Short-sequence DNA repeats in prokaryotic genomes. *Microbiology and Molecular Biology Reviews* **62**, 275–93.

Bruvo R, Michiels NK, D'Souza TG, Schulenburg H, 2004. A simple method for the calculation of microsatellite genotype distances irrespective of ploidy level. *Molecular Ecology* **13**, 2101–2106.

Butler MI, Poulter RTM, 2015. The accessory genome of PSA strains from China, New Zealand, Chile, Italy, Japan, and Korea. In: *2nd International Symposium on Bacterial Canker of Kiwifruit Abstract and Programme Book*, 2015. Bologna, Italy: ISHS, 94

Butler MI, Stockwell PA, Black MA, Day RC, Lamont IL, Poulter RTM, 2013. *Pseudomonas syringae* pv. *actinidiae* from recent outbreaks of kiwifruit bacterial canker belong to different clones that originated in China. *PLoS One* **8**, e57464.

Butler MI, Jung JS, Kim GH, Lamont IL, Stockwell PA, Koh YJ, Poulter R, 2015. Genome features of *Pseudomonas syringae* pv. *actinidiae* recently isolated in Korea. *Acta Horticulturae* **1095**, 75-80

Chapman JR, Taylor RK, Weir BS, Romberg MK, Vanneste JL, Luck J, Alexander BJR, 2012. Phylogenetic relationships among global populations of *Pseudomonas syringae* pv. *actinidiae*. *Phytopathology* **102**, 1034–1044.

Ciarroni S, Gallipoli L, Taratufolo MC, *et al.*, 2015. Development of a multiple loci variable number of tandem repeats analysis (MLVA) to unravel the intra-pathovar structure of *Pseudomonas syringae* pv. *actinidiae* populations worldwide. *PLoS One* 10:e0135310.

Cunty A, Cesbron S, Poliakoff F, Jacques MA, Manceau C, 2015a. Origin of the outbreak in France of *Pseudomonas syringae* pv. *actinidiae* biovar 3, the causal agent of bacterial canker of kiwifruit, revealed by a multilocus variable-number tandem-repeat analysis *Applied Environmental Microbiology* **81**, 6773–6789.

Cunty A, Poliakoff F, Rivoal C, *et al*., 2015b. Characterization of *Pseudomonas syringae* pv. *actinidiae* (Psa) isolated from France and assignment of Psa biovar 4 to a de novo pathovar: *Pseudomonas syringae* pv. *actinidifoliorum* pv. nov. *Plant Pathology* **64**, 582–596.

Dray S, Dufour AB, 2007. The ade4 package: Implementing the duality diagram for ecologists. *Journal of statistical software* **22**, 1–20.

Evanno G, Regnaut S, Goudet J, 2005. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology* 14, 2611–2620.

Falush D, Stephens M, Pritchard JK, 2007. Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology Notes* **7**, 574–578.

Fang Y, Xiaoxiang Z, Tao WY, 1990. Preliminary studies on kiwifruit disease in Hunan province. *Sichuan Fruit Science and Technology* **18**, 28–29.

Ferguson AR, Bollard EG, 1990. Domestication of the kiwifruit. In: I.J. Warrington & G.C. Weston eds. *Kiwifruit: science and management*. Auckland, New Zealand: Ray Richards in Association with the New Zealand Society of Horticultural Science, 165-246.

Ferguson AR, Huang H, 2007. Genetic Resources of Kiwifruit: Domestication and Breeding. *Horticultural Reviews*, 1–121.

Francis RM, 2017. POPHELPER: An R package and web app to analyse and visualize population structure. *Molecular Ecology Resources* **17**, 27–32.

Fujikawa T, Sawada H, 2016. Genome analysis of the kiwifruit canker pathogen *Pseudomonas syringae* pv. *actinidiae* biovar 5. *Scientific Reports* **6**, 21399–21311.

Grünwald NJ, Goss EM, 2011. Evolution and population genetics of exotic and re-emerging pathogens: novel tools and approaches. *Annual Review of Phytopathology* **49**, 249–267.

Han HS, Koh YJ, Hur JS, Jung JS, 2003. Identification and characterization of coronatine-producing *Pseudomonas syringae* pv. *actinidiae*. *Journal of Microbiology and Biotechnology* **13**, 110–118.

He R, Liu P, Jia B, *et al*. 2019. Genetic diversity of *Pseudomonas syringae* pv. *actinidiae* strains from different geographic regions in China. *Phytopathology* **109**, 347–357.

Jombart T, 2008. Adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics* **24**, 1403–1405.

Jombart T, Devillard S, Balloux F, 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics* 11, 94

Kamvar ZN, Brooks JC, Grünwald NJ, 2015. Novel R tools for analysis of genome-wide population genetic data with emphasis on clonality. *Frontiers in Genetics* **6**, 1–10.

Kamvar ZN, Tabima JF, Grunwald NJ, 2014. Poppr: An R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ* **2**, e281.

Kassambara A, Mundt F, 2016. R factoextra. [http://www.sthda.com/english/rpkgs/factoextra]

Kim GH, Kim KH, Son KI, et al., 2016. Outbreak and spread of bacterial canker of kiwifruit caused by *Pseudomonas syringae* pv. *actinidiae* biovar 3 in Korea. *The Plant Pathology Journal* **32**, 545–551.

Koh YJ, Chung HJ, Cha BJ, Lee DH, 1994. Outbreak and spread of bacterial canker in kiwifruit. *Korean Journal of Plant Pathology* 10, 68–72.

Koh YJ, Kim GH, Koh HS, Lee YS, Kim SC, Jung JS, 2012. Occurrence of a new type of *Pseudomonas syringae* pv. *actinidiae* strain of bacterial canker on kiwifruit in Korea. *The Plant Pathology Journal* **28**, 423–427.

Liu P, Xue S, He R, *et al.*, 2016. *Pseudomonas syringae* pv. *actinidiae* isolated from non-kiwifruit plant species in China. *European Journal of Plant Pathology* **145**, 743–754.

Mazzaglia A, Studholme DJ, Taratufolo MC, *et al.*, 2012. *Pseudomonas syringae* pv. *actinidiae* (PSA) isolates from recent bacterial canker of kiwifruit outbreaks belong to the same genetic lineage *PLoS One* **7**, e36518.

McCann HC, Li L, Liu Y, *et al.*, 2017. Origin and evolution of the kiwifruit canker pandemic. *Genome Biology and Evolution* **9**, 932–944.

McCann HC, Rikkerink EHA, Bertels F, *et al.*, 2013. Genomic analysis of the kiwifruit pathogen *Pseudomonas syringae* pv. *actinidiae* provides insight into the origins of an emergent plant disease. *PLoS Pathogens* **9**, e1003503.

Oksanen J, Blanchet FG, Friendly M, *et al*., 2019. vegan: Community Ecology Package v2.5-6. *https//cran.r-project.org, https//github.com/vegandevs/vegan*.

Sawada H, Fujikawa T, 2019. Genetic diversity of *Pseudomonas syringae* pv. *actinidiae*, pathogen of kiwifruit bacterial canker. *Plant Pathology* **68**, 1235–1248.

Sawada H, Kondo K, Nakaune R, Nakaline R, 2016. Novel biovar (biovar 6) of *Pseudomonas syringae* pv. *actinidiae* causing bacterial canker of kiwifruit (*Actinidia deliciosa*) in Japan. *Japanese Journal of Phytopathology.* **82**, 101-115.

Sawada H, Miyoshi T, Ide Y, 2014. Novel MLSA group (Psa5) of *Pseudomonas syringae* pv. *actinidiae* causing bacterial canker of kiwifruit (*Actinidia chinensis*) in Japan. *Japanese Journal of Phytopathology* **80**, 171–184.

Sawada H, Shimizu S, Miyoshi T, *et al*., 2015. Characterization of biovar 3 strains of *Pseudomonas syringae* pv. *actinidiae* isolated in Japan. *Japanese Journal of Phytopathology* **81**, 111-126.

Scortichini M, 1994. Occurrence of *Pseudomonas syringae* pv. *actinidiae* on kiwifruit in Italy. *Plant Pathology* **43**, 1035–1038.

Serizawa S, Ichikawa T, Takikawa Y, Tsuyumu S, Goto M, 1989. Occurrence of bacterial canker of kiwifruit in Japan: description of symptoms, isolation of the pathogen and screening of bactericides. *Japanese Journal of Phytopathology* **55**, 427–436.

Takikawa Y, Serizawa S, Ichikawa T, Tsuyumu S, Goto, M, 1989. *Pseudomonas syringae* pv. *actinidiae* pv. nov.: The causal bacterium of canker of kiwifruit in Japan. *Japanese Journal of Phytopathology.* 55, 437–444.

Ushiyama K, Kita N, Suyama K, Aono N, Ogawa J, Fujii H, 1992. Bacterial canker disease of wild *Actinidia* plants as the infection source of outbreak of bacterial canker of kiwifruit caused by *Pseudomonas syringae* pv. *actinidiae*. *Japanese Journal of Phytopathology* 58, 426–430.

2.9 Supplementary materials of article 1

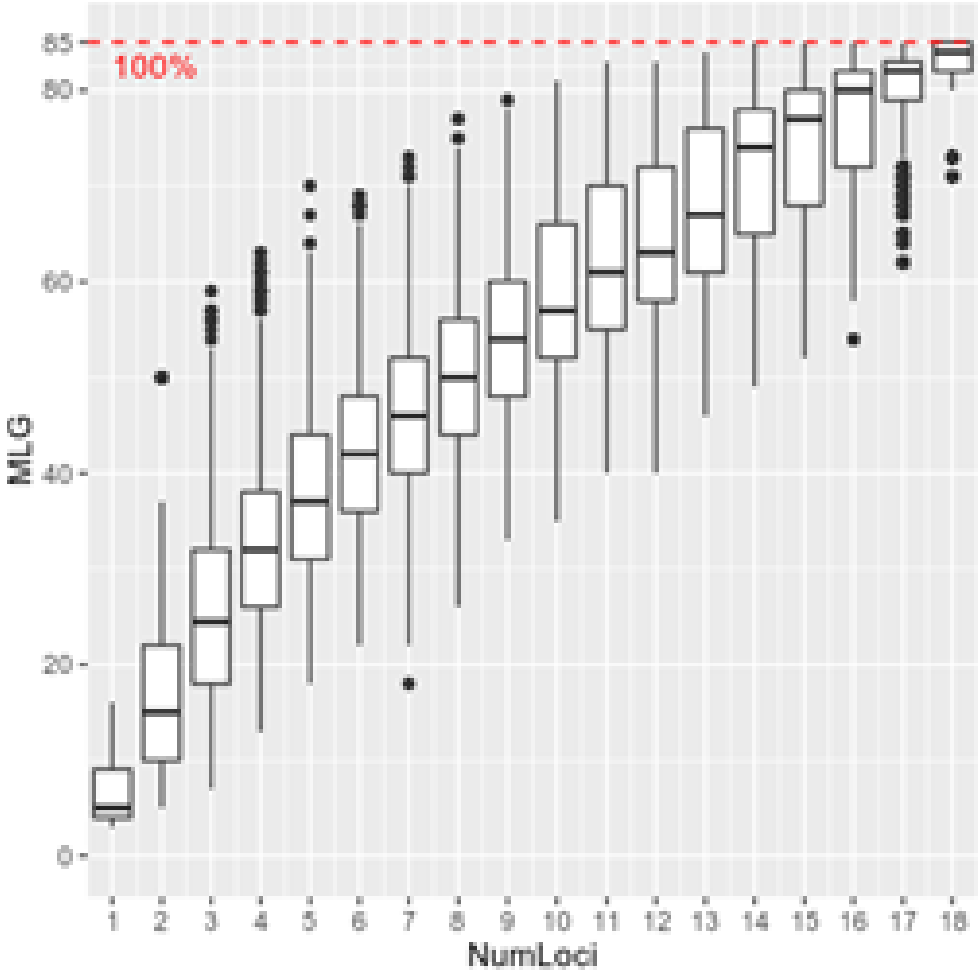


Figure S1: Genotype accumulation curve for 152 isolates of *P. syringae* pv. *actinidiae*, the x-axis represents the number of loci that were randomly sampled 1000 times up to (n 1) loci while the y-axis represents the number of unique MLGs observed in the dataset. The dashed lines at the plateau indicate that enough number of loci was used to discriminate between individuals.

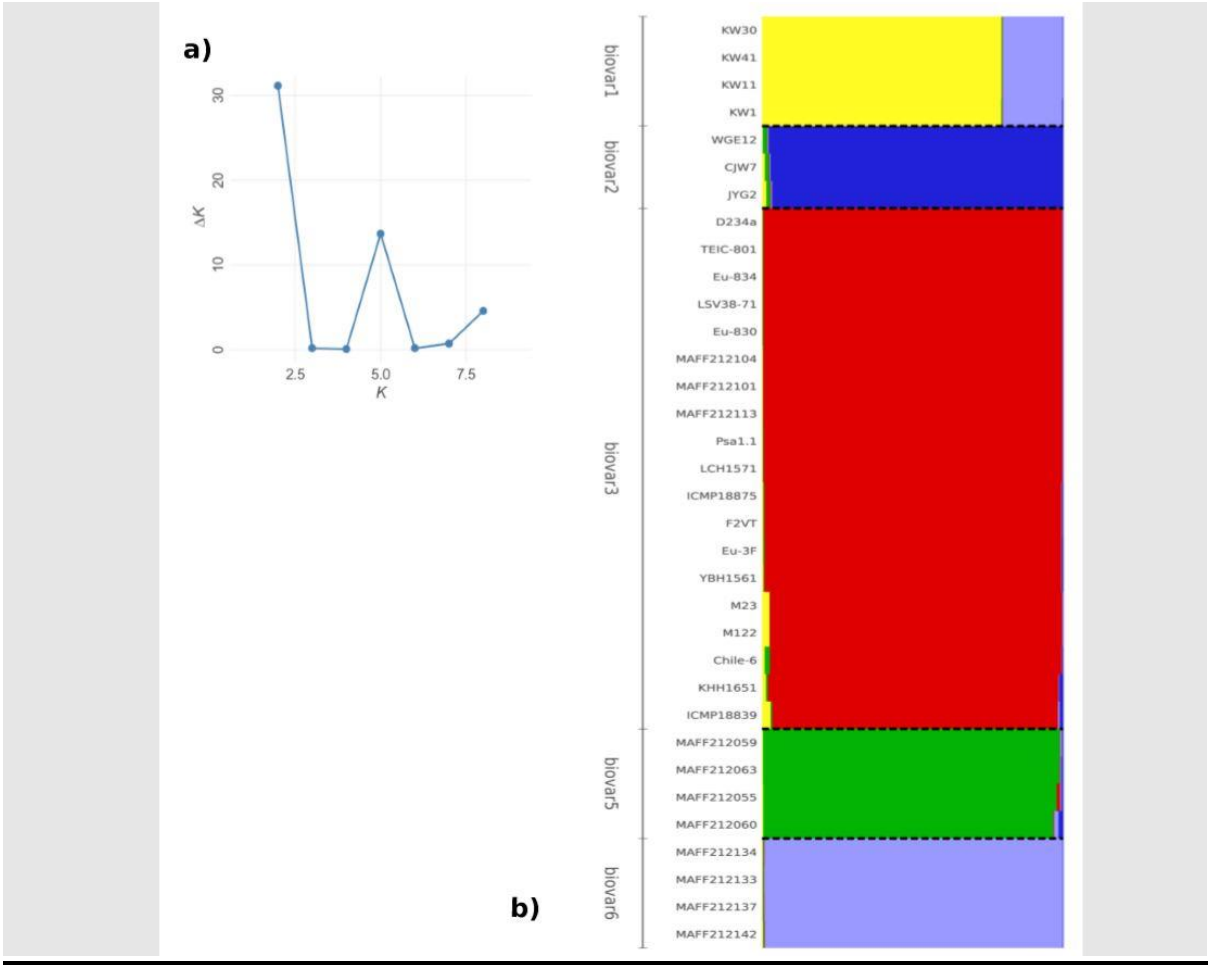


Figure S2: STRUCTURE and CLUMPP analysis run with K = 5, burnin period = 1000 and MCMC reiterations 10000. **a)** Delta K values calculated by Evanno's method to infer the number of groups that best suits the data. b) Population genetic structure of a selection of 34 individuals of *P. syringae* pv. *actinidiae* used as a "control dataset".

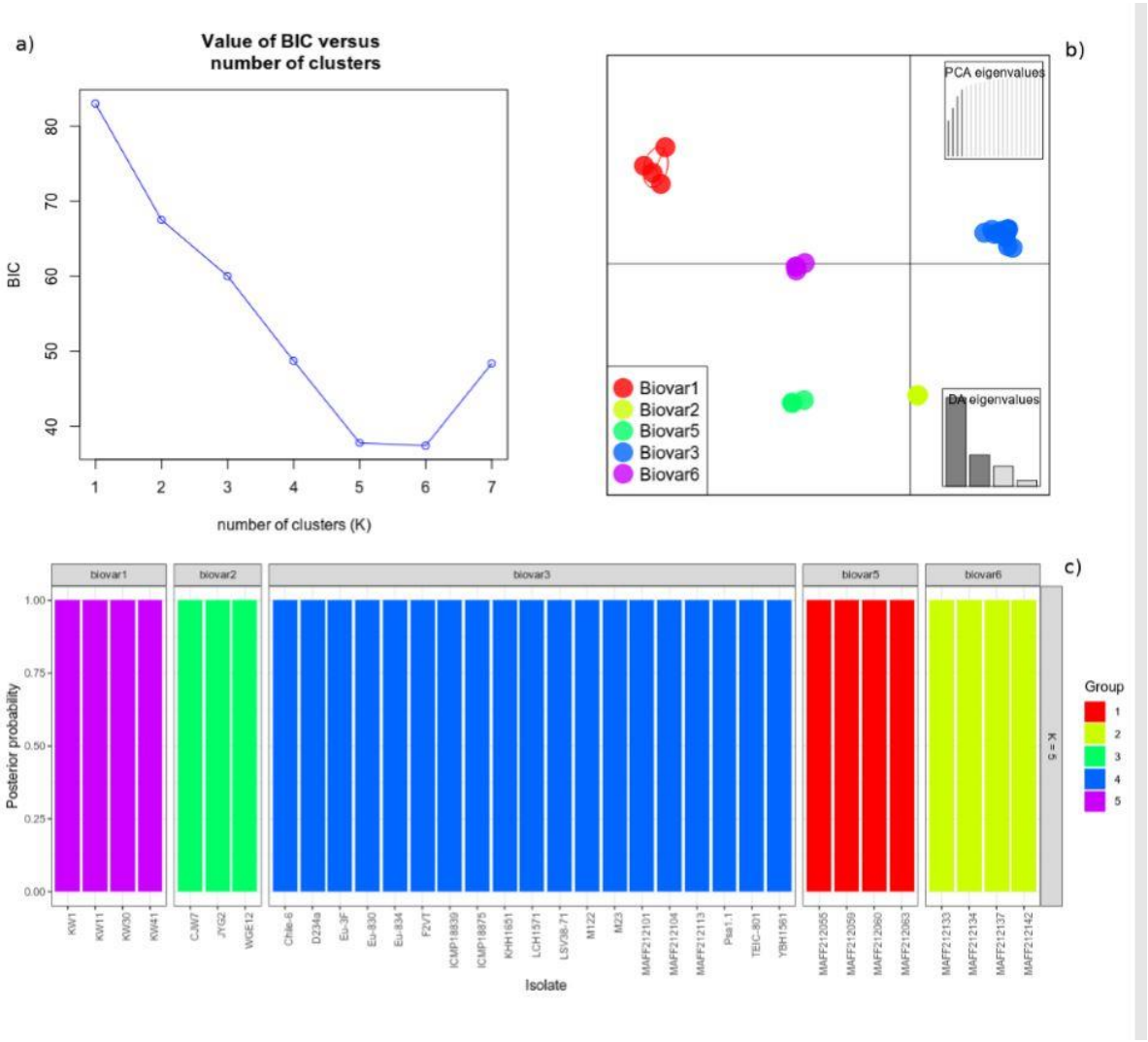


Figure S3: Discriminant analysis of principal component (DAPC) of the "control dataset". a) Bayesian information criterion (BIC) as a function of the number of clusters (K). b) Scatterplot of individuals arranged by groups on the two principal components of DAPC. c) Posterior assignment of individuals when K=5.

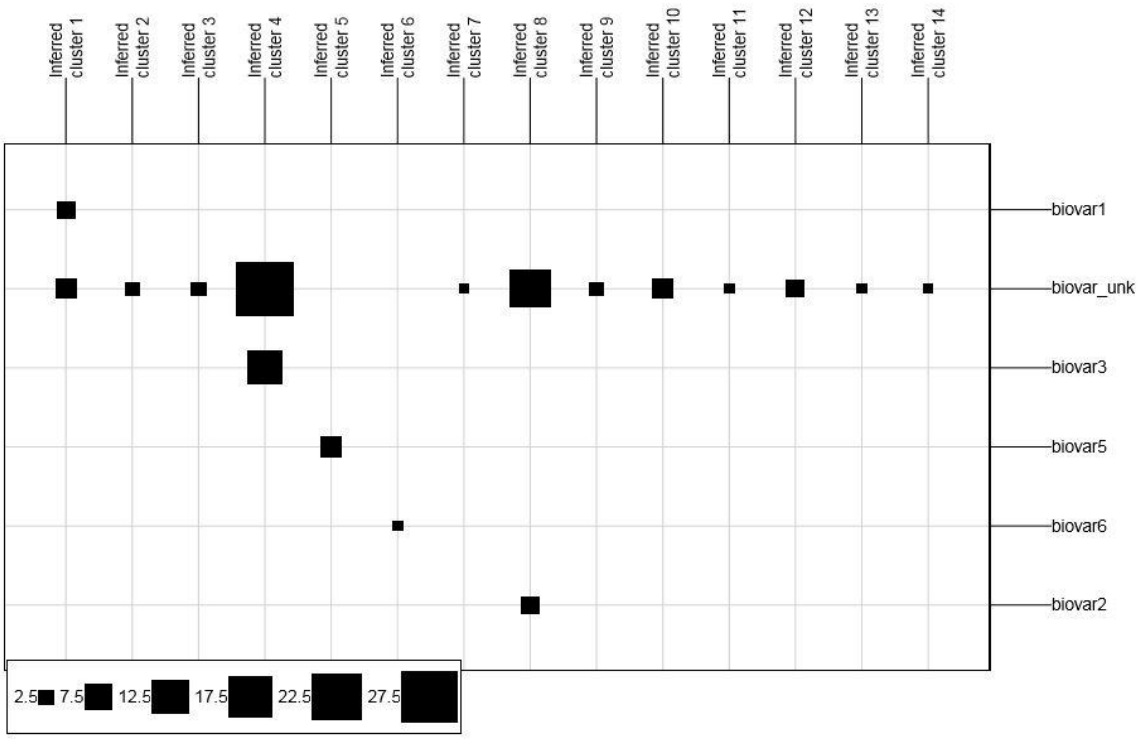


Figure S4: Individuals distribution between original and inferred populations of *P. syringae* pv. *actinidiae*.

Table S1: Geographic origin and time of isolation of the 152 Psa individuals.

N°	Strain	Country	Region	Year
1§	KW30	Japan	Chubu	1984
2§	KW41	Japan	Chubu	1984
3§	KW1	Japan	Chubu	1984
4§	KW11	Japan	Chubu	1984
5	PA429	Japan	unknown_J	1988
6	PA459	Japan	unknown_J	1988
7	Jap2726	Japan	Chubu	2009
8	Jap2820	Japan	Shikoku	2011
9	Jap2818	Japan	Chubu	2011
10	Jap2819	Japan	Chubu	2011
11§	MAFF212101	Japan	Kyushu	2014
12§	MAFF212104	Japan	Shikoku	2014
13§	MAFF212113	Japan	Fukuoka	2014
14	MAFF212119	Japan	Fukuoka	2014
15	MAFF212122	Japan	Kyushu	2014
16	MAFF212110	Japan	Kansai	2014
17	MAFF212121	Japan	Kyushu	2014
18	MAFF212145	Japan	Kyushu	-
19	SUPP3069	Japan	Chubu	2014
20§	MAFF212055	Japan	Kyushu	2012
21§	MAFF212059	Japan	Kyushu	2012
22§	MAFF212063	Japan	Kyushu	2012
23§	MAFF212060	Japan	Kyushu	2012
24§	MAFF212133	Japan	Chubu	2015
25§	MAFF212134	Japan	Chubu	2015
26§	MAFF212137	Japan	Chubu	2015
27§	MAFF212142	Japan	Chubu	2015
28	KCTC23663	Korea	Jeollanam-do	1989
29	KCTC23664	Korea	Jeollanam-do	1989
30	KCTC23665	Korea	Jeollanam-do	1989
31	KN2	Korea	Unknown_K	1997
32§	CJW7	Korea	Jeju-do	1999
33§	JYG2	Korea	Jeollanam-do	1999
34§	WGE12	Korea	Jeollanam-do	1999
35	K5	Korea	Jeju-do	2008
36	YCS2_1	Korea	Jeju-do	2008
37	JJ18	Korea	Jeju-do	2008
38	K8	Korea	Jeju-do	2008

39	KGY1_2	Korea	Jeju-do	2008
40	KBE9	Korea	Jeju-do	2008
41	JYSS04-136	Korea	Jeollanam-do	2008
42	Kor09-PJC-7	Korea	Jeju-do	2009
43	Kor10-KBE-29	Korea	Jeju-do	2010
44	SYS1	Korea	Jeollanam-do	2011
45	SYS2	Korea	Jeollanam-do	2011
46	SYS4	Korea	Jeollanam-do	2011
47	K133WD	Korea	Jeollanam-do	2013
48	K134WD	Korea	Jeollanam-do	2013
49	K131WD	Korea	Jeollanam-do	2013
50	K132WD	Korea	Jeollanam-do	2013
51	K134KBS	Korea	Jeollanam-do	2013
52	KTS1471	Korea	Jeju-do	2014
53	KJB1451	Korea	Jeollabuk-do	2014
54§	YBH1561	Korea	Jeollanam-do	2015
55§	LCH1571	Korea	Jeollanam-do	2015
56	KKB1531	Korea	Jeju-do	2015
57§	KHH1651	Korea	Gyeongsangnam-do	2016
58§	M23	China	Shaanxi	2010
59§	M122	China	Shaanxi	2010
60	M218	China	Shaanxi	2010
61	M228	China	Shaanxi	2010
62	CH2010-5	China	Shaanxi	2010
63	CH2010-6	China	Shaanxi	2010
64	CH2010-7	China	Shaanxi	2010
65	LOLO-15	China	Guizhou	2012
66	LOLO-20	China	Guizhou	2012
67	LOLO-28	China	Guizhou	2012
68	LOLO-32	China	Guizhou	2012
69	LOLO-38	China	Guizhou	2012
70	Haxa2	China	Shaanxi	2012
71	Haxa3	China	Shaanxi	2012
72	Haxa5	China	Shaanxi	2012
73	Haxa7	China	Shaanxi	2012
74	Hym2	China	Sichuan	2012
75	Hym3	China	Sichuan	2012
76	Hym4	China	Sichuan	2012
77	JILO-21	China	Anhui	2012
78	JILO-23	China	Anhui	2012
79	JILO-24	China	Anhui	2012
80	JILO-26	China	Anhui	2012

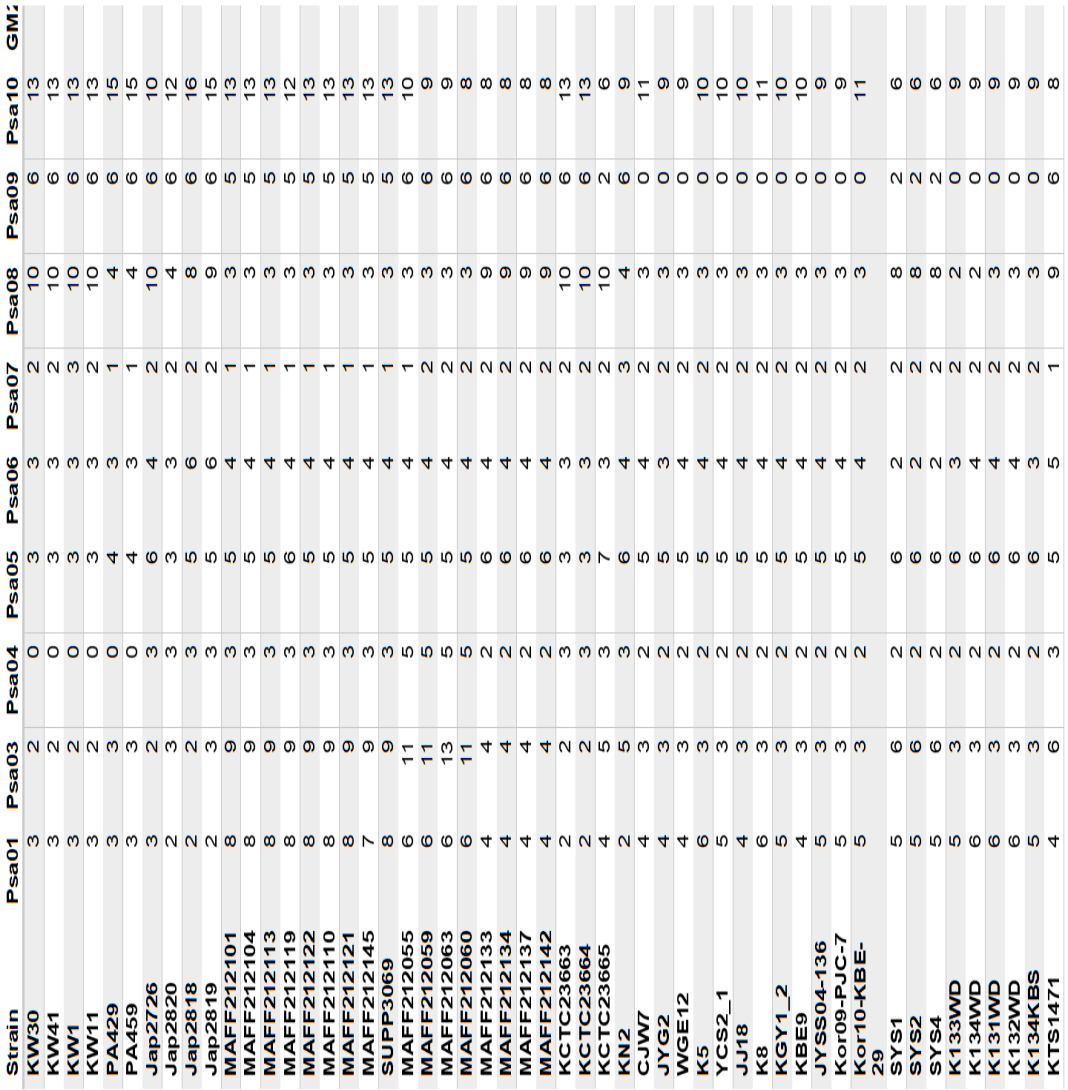
81	JILO-27	China	Anhui	2012
82	JZHY7	China	Anhui	2013
83	JZZM11	China	Anhui	2013
84	SH5	China	Shangai	2013
85	SH28	China	Shangai	2013
86	SCHY24	China	Sichuan	2012
87	SCHY25	China	Sichuan	2012
88	HWD3	China	Shaanxi	2012
89	HWD4	China	Shaanxi	2012
90	HWD5	China	Shaanxi	2012
91	DSF-1	China	Shimen	2015
92	DSF-2	China	Shimen	2015
93	DSF-3	China	Shimen	2015
94	YN-8	China	Kunmin	2017
95	YN-9	China	Kunmin	2017
96	YN-10	China	Kunmin	2017
97§	Eu-3F	France	Rhone Alpes	2010
98	Eu-9F	France	Rhone Alpes	2010
99§	D_2.34a	Germany	Bavaria	2013
100§	TEIC-801	Greece	Central Macedonia	2014
101	TEIC-805	Greece	Central Macedonia	2014
102	TEIC-806	Greece	Central Macedonia	2014
103	TEIC-807	Greece	Central Macedonia	2014
104	TEIC-819	Greece	Central Macedonia	2014
105	TEIC-823	Greece	Central Macedonia	2014
106	TEIC-824	Greece	Central Macedonia	2014
107	TEIC-826	Greece	Central Macedonia	2014
108	GR1	Greece	Central Macedonia	2014
109	GR2	Greece	Central Macedonia	2014
110	GR3	Greece	Central Macedonia	2014
111	GR4	Greece	Central Macedonia	2014
112	GR5	Greece	Central Macedonia	2014
113	Ita92	Italy	Latium	1992
114	CFBP7285	Italy	Veneto	2008
115	CFBP7286	Italy	Latium	2008
116	CFBP7287	Italy	Latium	2008
117§	F2VT	Italy	Latium	2017
118	F8VT	Italy	Latium	2017
119	F9VT	Italy	Latium	2017
120	F8LT	Italy	Latium	2017
121	F10LT	Italy	Latium	2017
122	F13LT	Italy	Latium	2017

123[§]	Eu-834	Portugal	Norte	2011
124	Eu-822	Portugal	Norte	2011
125	Eu-832	Portugal	Norte	2011
126[§]	Eu-830	Spain	Galicia	2011
127	Eu-829	Spain	Galicia	2011
128[§]	LSV38-71	Switzerland	Canton of Geneva	2011
129[§]	ICMP18839	New Zealand	Bay of Plenty	2011
130[§]	ICMP18875	New Zealand	Bay of Plenty	2011
131	ICMP19200	New Zealand	Auckland	2011
132	ICMP18708	New Zealand	Bay of Plenty	2010
133	ICMP18800	New Zealand	Bay of Plenty	2010
134	RT802	New Zealand	Bay of Plenty	2015
135	RT811	New Zealand	Bay of Plenty	2015
136	RT812	New Zealand	Bay of Plenty	2015
137	LS2A	New Zealand	Bay of Plenty	2016
138	AF868	New Zealand	Auckland	2014
139	AF807	New Zealand	Far North	2014
140	AF907	New Zealand	Waikato	2014
141	AF862	New Zealand	Auckland	2014
142	AF813	New Zealand	Auckland	2014
143[§]	Chile-6	Chile	VII Reg Maule	2013
144	Chile-16	Chile	VII Reg Maule	2013
145	Chile-25	Chile	VIII Reg Bio Bio	2013
146	ICMP19438	Chile	VII Reg Maule	2011
147	ICMP19439	Chile	VII Reg Maule	2011
148	ICMP19456	Chile	VII Reg Maule	2010
149[§]	Psa1.1	Argentina	Mar del Plata	2014
150	Psa1.2	Argentina	Mar del Plata	2014
151	Psa2.2	Argentina	Mar del Plata	2014
152	Psa2.3	Argentina	Mar del Plata	2014

Table S2: PCRs amplification protocol for the detection of the 19 VNTR loci selected in this study.

Locus	Initial denaturation		Cycle						Number of cycles	Final extension	
			Denaturation		Annealing		Extension				
	C°	s	C°	s	C°	s	C°	s	n	C°	s
PSA01	95	300	95	30	57	30	72	30	35	72	600
PSA03	95	300	95	30	59	30	72	30	35	72	600
PSA04	95	300	95	30	60	30	72	30	35	72	600
PSA05	95	300	95	30	59	30	72	30	35	72	600
PSA06	95	300	95	30	56	30	72	30	35	72	600
PSA07	95	300	95	30	57	30	72	30	35	72	600
PSA08	95	300	95	30	59	30	72	30	35	72	600
PSA09	95	300	95	30	62	30	72	48	35	72	600
PSA10	95	300	95	30	59	30	72	30	35	72	600
GM254	95	300	95	30	55	30	72	30	35	72	600
GM1553	95	300	95	30	52	30	72	30	35	72	600
GM1834	95	300	95	30	55	30	72	30	35	72	600
GM4076	95	300	95	30	66	45	72	30	35	72	600
TR14I	95	300	95	30	53	30	72	30	35	72	600
TR30I	95	300	95	30	53	30	72	30	35	72	600
TR1II	95	300	95	30	59	30	72	30	35	72	600
TR2II	95	300	95	30	59	30	72	30	35	72	600
TR3II	95	300	95	30	59	30	72	30	35	72	600
TR11II	95	300	95	30	60	30	72	30	35	72	600

Table S3: Number of repetitions for each of the 19 VNTR loci detected through MLVA analysis.



Strain	Psa01	Psa03	Psa04	Psa05	Psa06	Psa07	Psa08	Psa09	Psa10	GM2
KW30	3	2	0	3	3	2	10	6	13	
KW41	3	2	0	3	3	2	10	6	13	
KW1	3	2	0	3	3	3	10	6	13	
KW11	3	2	0	3	3	2	10	6	13	
PA429	3	3	0	4	3	1	4	6	15	
PA459	3	3	0	4	3	1	4	6	15	
Jap2726	3	2	3	6	4	2	10	6	10	
Jap2820	2	3	3	3	3	2	4	6	12	
Jap2818	2	2	3	5	6	2	8	6	16	
Jap2819	2	3	3	5	6	2	9	6	15	
MAFF212101	8	9	3	5	4	1	3	5	13	
MAFF212104	8	9	3	5	4	1	3	5	13	
MAFF212113	8	9	3	5	4	1	3	5	13	
MAFF212119	8	9	3	6	4	1	3	5	12	
MAFF212122	8	9	3	5	4	1	3	5	13	
MAFF212110	8	9	3	5	4	1	3	5	13	
MAFF212121	8	9	3	5	4	1	3	5	13	
MAFF212145	7	9	3	5	4	1	3	5	13	
SUPP3069	8	9	3	5	4	1	3	5	13	
MAFF212055	6	11	5	5	4	1	3	6	10	
MAFF212059	6	11	5	5	4	2	3	6	9	
MAFF212063	6	13	5	5	4	2	3	6	9	
MAFF212060	6	11	5	5	4	2	3	6	8	
MAFF212133	4	4	2	6	4	2	9	6	8	
MAFF212134	4	4	2	6	4	2	9	6	8	
MAFF212137	4	4	2	6	4	2	9	6	8	
MAFF212142	4	4	2	6	4	2	9	6	8	
KCTC23663	2	2	3	3	3	2	10	6	13	
KCTC23664	2	2	3	3	3	2	10	6	13	
KCTC23665	4	5	3	7	3	2	10	2	6	
KN2	2	5	3	6	4	3	4	6	9	
CJW7	4	3	2	5	4	2	3	0	11	
JYG2	4	3	2	5	3	2	3	0	9	
WGE12	4	3	2	5	4	2	3	0	9	
K5	6	3	2	5	4	2	3	0	10	
YCS2_1	5	3	2	5	4	2	3	0	10	
JJ18	4	3	2	5	4	2	3	0	10	
K8	6	3	2	5	4	2	3	0	11	
KGY1_2	5	3	2	5	4	2	3	0	10	
KBE9	4	3	2	5	4	2	3	0	10	
JYSS04-136	5	3	2	5	4	2	3	0	9	
Kor09-PJC-7	5	3	2	5	4	2	3	0	9	
Kor10-KBE-29	5	3	2	5	4	2	3	0	11	
SYS1	5	6	2	6	2	2	8	2	6	
SYS2	5	6	2	6	2	2	8	2	6	
SYS4	5	6	2	6	2	2	8	2	6	
K133WD	5	3	2	6	3	2	2	0	9	
K134WD	6	3	2	6	4	2	2	0	9	
K131WD	6	3	2	6	4	2	3	0	9	
K132WD	6	3	2	6	4	2	3	0	9	
K134KBS	5	3	2	6	3	2	3	0	9	
KTS1471	4	6	3	5	5	1	9	6	8	

Table S4: *Pseudomonas syringae* pv. *actinidiae* strains grouped by their relative MultiLocus Genotypes (MLGs) matching the country of origin.

MLG	Country	Strains
MLG1	China	DSF-1, DSF-2, DSF-3
MLG2	China	JZZM11
MLG3	Korea	SYS2
MLG4	Korea	SYS1, SYS4
MLG5	China	LOLO-28
MLG6	China	LOLO-15, LOLO-20, LOLO-32, LOLO-38
MLG7	Korea	KKB1531
MLG8	Korea	Kor10-KBE-29

MLG9	Korea	JYSS04-136
MLG10	Korea	Kor09-PJC-7
MLG11	Korea	KGY1_2
MLG12	Korea	YCS2_1
MLG13	Korea	K133WD
MLG14	Korea	K134KBS
MLG15	China	SH5, SH28
MLG16	China	JZHY7
MLG17	Korea	KTS1471
MLG18	China	M228
MLG19	China	M218
MLG20	Korea	KCTC23665
MLG21	Japan	MAFF212133, MAFF212134, MAFF212137,MAFF212142
MLG22	China	Hym2, Hym3, Hym4
MLG23	China	JILO-21, JILO-23, JILO-24, JILO-26, JILO-27
MLG24	China	Haxa2, Haxa3, Haxa5, Haxa7, HWD3, HWD4, HWD5
MLG25	Korea	CJW7
MLG26	Korea	WGE12
MLG27	Korea	JJ18
MLG28	Korea	KBE9
MLG29	Korea	JYG2
MLG30	Japan	MAFF212063
MLG31	Japan	MAFF212055
MLG32	Japan	MAFF212060
MLG33	Japan	MAFF212059
MLG34	Korea	K8
MLG35	Korea	K5
MLG36	Korea	K134WD
MLG37	Korea	K132WD
MLG38	Korea	K131WD
MLG39	China	SCHY24
MLG39	China	SCHY25
MLG40	Japan	MAFF212145
MLG41	Chile	ICMP19438, ICMP19439
MLG42	New Zealand	LS2A
MLG43	New Zealand	AF813
MLG44	Portugal	Eu-822
MLG45	France	Eu-9F
MLG46	New Zealand	ICMP18708, RT811
MLG47	China	CH2010-5, CH2010-6, CH2010-7
MLG48	China	M23, M122
MLG49	Greece	GR3
MLG50	Greece	TEIC-824, GR5
	Italy	F10LT, F13LT
MLG51	Greece	GR2, GR4

MLG52	Italy	F8LT
MLG53	Italy	F2VT
MLG54	Italy	CFBP7287
MLG55	Italy	CFBP7286
MLG56	France	Eu-3F
	Italy	F8VT, F9VT
	Portugal	Eu-832
MLG57	Germany	D_2.34a
	Greece	TEIC-801, TEIC-805, TEIC-806, TEIC-807, TEIC-819, TEIC-823, GR1
	Italy	CFBP7285
	Portugal	Eu-834
	Spain	Eu-830
	Switzerland	LSV38-71
MLG58	New Zealand	AF907
MLG59	New Zealand	ICMP18875
MLG60	Greece	TEIC-826
MLG61	Korea	KHH1651
MLG62	Chile	Chile-6, Chile-16, Chile-25, ICMP19456
MLG63	Korea	KJB1451
MLG64	China	YN-8
MLG65	China	YN-9, YN-10
MLG66	New Zealand	AF868, AF807, AF862
MLG67	Korea	YBH1561
MLG68	Korea	LCH1571
MLG69	New Zealand	RT802
MLG70	New Zealand	RT812
MLG71	Japan	MAFF212101, MAFF212104, MAFF212113, MAFF212122, MAFF212110, MAFF212121, SUPP3069
	New Zealand	ICMP19200, ICMP18800
	Argentina	Psa1.1, Psa1.2, Psa2.2, Psa2.3
MLG72	Japan	MAFF212119
MLG73	Spain	Eu-829
MLG74	Korea	KN2
MLG75	Japan	Jap2819
MLG76	Japan	Jap2820
MLG77	Japan	Jap2818
MLG78	Korea	KCTC23663
MLG79	Korea	KCTC23664
MLG80	New Zealand	ICMP18839
MLG81	Japan	PA429, PA459
MLG82	Japan	Jap2726
MLG83	Japan	KW1
MLG84	Japan	KW11, Ita92
MLG85	Japan	KW30, KW41

Table S5: Assignment of the 152 *Pseudomonas syringae* pv. *actinidiae* strains to different groupings according to the three analytical methods used. For hierarchical clustering, the numbering corresponds to clusters described in Figure 3, whilst for STRUCTURE and for DAPC the numbering corresponds to the highest percentage of posterior assignment.

Strain	Hierarchical clustering	STRUCTURE	DAPC
KW30	1	6	10
KW41	1	6	10
KW1	1	6	10
KW11	1	6	10
PA429	1	9	10
PA459	1	9	10
Jap2820	1	6	1
KCTC23663	1	6	10
KCTC23664	1	6	10
Ita92	1	6	10
MAFF212055	5	10	12
MAFF212059	5	10	12
MAFF212063	5	10	12
MAFF212060	5	10	12
MAFF212133	6	12	13
MAFF212134	6	12	13
MAFF212137	6	12	13
MAFF212142	6	12	13
MAFF212101	4	5	4
MAFF212104	4	5	4
MAFF212113	4	5	4
MAFF212119	4	5	4
MAFF212122	4	5	4
MAFF212110	4	5	4
MAFF212121	4	5	4
MAFF212145	4	5	4
SUPP3069	4	5	4
KJB1451	4	5	4
YBH1561	4	5	4
LCH1571	4	5	4
KHH1651	4	5	4
M23	4	5	4
M122	4	5	4

CH2010-5	4	5	4
CH2010-6	4	5	4
CH2010-7	4	5	4
JZZM11	4	5	4
SCHY24	4	5	7
SCHY25	4	5	7
YN-8	4	5	4
YN-9	4	5	4
YN-10	4	5	4
Eu-3F	4	5	5
Eu-9F	4	5	5
D_2.34a	4	5	5
TEIC-801	4	5	5
TEIC-805	4	5	5
TEIC-806	4	5	5
TEIC-807	4	5	5
TEIC-819	4	5	5
TEIC-823	4	5	5
TEIC-824	4	5	5
TEIC-826	4	5	5
GR1	4	5	5
GR2	4	5	5
GR3	4	5	5
GR4	4	5	5
GR5	4	5	5
CFBP7285	4	5	5
CFBP7286	4	5	5
CFBP7287	4	5	5
F2VT	4	5	5
F8VT	4	5	5
F9VT	4	5	5
F8LT	4	5	5
F10LT	4	5	5
F13LT	4	5	5
Eu-834	4	5	5
Eu-822	4	5	5
Eu-832	4	5	5
Eu-830	4	5	5
Eu-829	4	5	5
LSV38-71	4	5	5
ICMP18839	4	5	4
ICMP18875	4	5	4
ICMP19200	4	5	4

ICMP18708	4	5	4
ICMP18800	4	5	4
RT802	4	5	4
RT811	4	5	4
RT812	4	5	4
LS2A	4	5	4
AF868	4	5	4
AF807	4	5	4
AF907	4	5	4
AF862	4	5	4
AF813	4	5	4
Chile-6	4	5	4
Chile-16	4	5	4
Chile-25	4	5	4
ICMP19438	4	5	4
ICMP19439	4	5	4
ICMP19456	4	5	4
Psa1.1	4	5	4
Psa1.2	4	5	4
Psa2.2	4	5	4
Psa2.3	4	5	4
CJW7	8	8	9
JYG2	8	8	9
WGE12	8	8	9
K5	8	8	9
YCS2_1	8	8	9
JJ18	8	8	9
K8	8	8	9
KGY1_2	8	8	9
KBE9	8	8	9
JYSS04-136	8	8	9
Kor09-PJC-7	8	8	9
Kor10-KBE-29	8	8	9
K133WD	8	8	14
K134WD	8	8	14
K131WD	8	8	14
K132WD	8	8	14
K134KBS	8	8	14
Jap2818	3	3	1
Jap2819	3	6	1
Jap2726	2	6	1
KN2	2	3	1
KCTC23665	7	2	1

LOLO-15	12	13	11
LOLO-20	12	13	11
LOLO-28	12	13	11
LOLO-32	12	13	11
LOLO-38	12	13	11
DSF-1	14	12	3
DSF-2	14	12	3
DSF-3	14	12	3
SYS1	9	4	8
SYS2	9	4	8
SYS4	9	4	8
KTS1471	10	12	2
M218	10	12	2
M228	10	12	2
JZHY7	10	12	2
KKB1531	11	1	2
SH5	13	11	2
SH28	13	11	2
Haxa2	13	7	6
Haxa3	13	7	6
Haxa5	13	7	6
Haxa7	13	7	6
Hym2	13	7	6
Hym3	13	7	6
Hym4	13	7	6
JILO-21	13	7	6
JILO-23	13	7	6
JILO-24	13	7	6
JILO-26	13	7	6
JILO-27	13	7	6
HWD3	13	7	6
HWD4	13	7	6
HWD5	13	7	6

Chapter 3 : Olive Diseases

3.1 Olive (*Olea europaea* L.)

Olive is one of the most ancient cultivated fruit crops in the Mediterranean basin. Its cultivation started in the eastern Mediterranean area and later spread worldwide. The Phoenicians, Greeks, and Romans contributed to the expansion of olive cultivation (1). Olive (*Olea europaea* L.) belongs to the Oleaceae family, which contains 600 species within 30 genera. The genus *Olea* includes 30 species, of which only *Olea europaea* is cultivated and has spread to Europe, Asia, Oceania and Africa (2). Olive is considered, as a Mediterranean climate tree.

The olive tree is vulnerable to pest and disease attack. Causal agents are either present in the area, or recently introduced, or have evolved. In addition, some causal agents are well known to have been endemic for a long time, while others are epidemic, in the sense that they have emerged or migrated from other hosts or places (3). Insect species (4), viruses and virus-like pathogens (5), nematodes (6), *Phytoplasma* (7), fungi (8,9) and bacterial species vary in their effect on olive (10–12).

Bacterial diseases are some of the most important diseases due to the difficulties of control. Of two bacterial diseases that attack olive, the olive knot is endemic and has long been known; it is caused by *Pseudomonas savastanoi* pathovars. The other disease is an epidemic that appeared recently in southern Italy, which causes severe death to trees and is known as olive quick decline syndrome (OQDS) (13).

The Mediterranean basin is a very important area of olive production, and olive trees are a typical landscape feature. Nowadays there are approximately 805 million olive trees worldwide, 98% of which are in the Mediterranean countries (14). Mediterranean countries produce more than 90% of the world's olives. The olive sector in the Mediterranean basin leads world olive production, as 97% of olive production comes from this area (12). Currently, Italian olive production ranks second after Spain, which produces approximately 7,875,800 tonnes, while Italy produces 2,940,545, Greece 2,000,000, Turkey 1,676,000, and Tunisia 1,100,000 (FAO STAT,2013). Italy produces about 27% of the total EU olive production (FAO STAT,

2013). Olive quick decline syndrome in southern Italy has emerged as a major threat, not only within the EU or Mediterranean basin but at the global level (13).

3.2 Knot diseases

Overgrowth and knots on olive trees have been noted since they were observed around 300 BC by ancient Greek botanist Theophrastus. The bacteria were isolated by Arkangeli and savastanoi, who experimentally fulfilled Koch's postulate (15). The disease is caused by the gram-negative bacteria *Pseudomonas savastanoi* pathovars, which belong to the *Pseudomonas syringae* pathovars (16).

Knot disease causes rapid cell development in infected tissue, and excessive growth leads to abnormal growths called tumours, galls, or knots. These formations can be seen clearly on the aerial parts of the infected plant, mostly on stems and twigs, but not often on leaves and fruits (17). The presence of knots has been detected in different members of plant families including *Oleaceae* (Olive, Jasminum, Forsythia, Phillyrea, and Ligustrum), *Fabaceae* (Retama), *Rhamnaceae* (Rhamnus), *Myrtaceae* (Myrtus), or *Apocinaceae* (Nerium, Mandevilla) and *Lythraceae pomegranate* (18–22). The bacterium causes different types of symptoms on ash (*Fraxinus excelsior*), with the development of vertical and lateral cracks rather than tumours or knots (23–25). The disease is different from the black knot caused by the fungal agent *Apiosporina morbosa*, which attacks plum, prune, and cherry (26).

3.3 Importance and distribution of the disease

The gram-negative bacteria *Pseudomonas savastanoi* pathovars are the causal agents of knot and canker disease. The disease is considered to be one of the extreme diseases attacking olive plants under favourable conditions (temperature and relative humidity), causing crop losses and severe damage in olive groves (27). Bacterial attack reduces tree vigour over time, as a result of heavy tumours, tree defoliation, and branch dieback, which in turn reduce the quality and quantity of olive production (28,29). Olive knot is a common disease that occurs in almost all regions of the world where olive is grown. The disease is present in Europe, Asia, America, Africa, Australia, and New Zealand (30), but there is no exact estimate of losses due to this disease. As regards oleander, there is usually no serious threat to overall plant health (31). The spread of the disease is associated with bacterial existence.

.

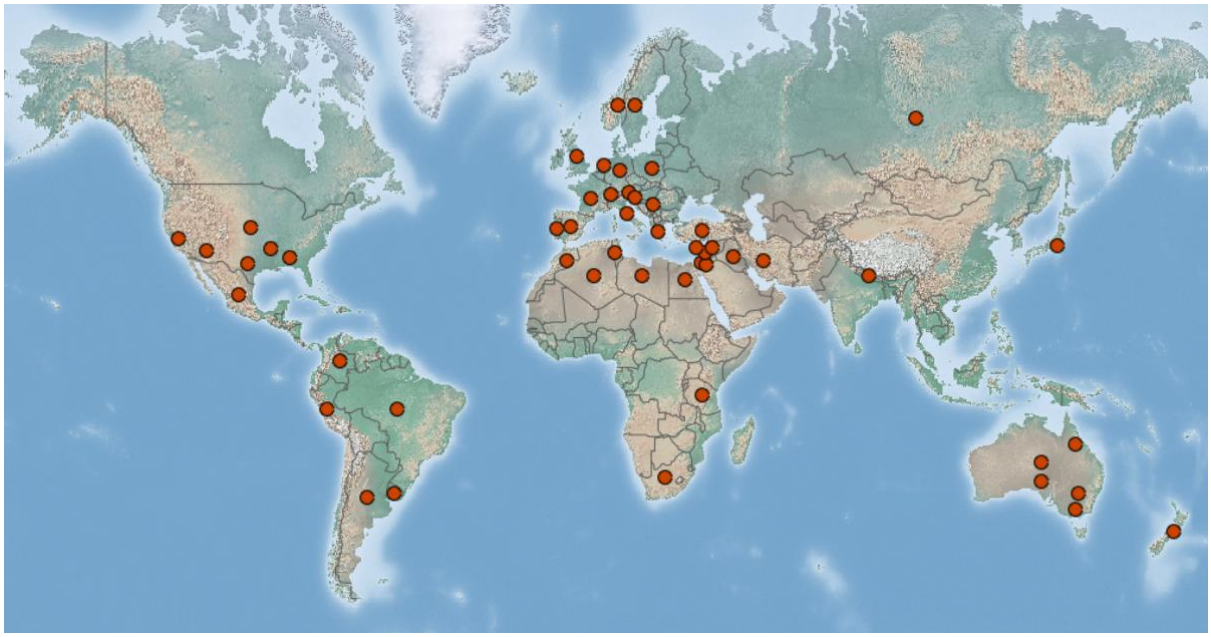


FIGURE 13: World distribution of olive knot disease

3.4 Epidemiology

Pseudomonas savastanoi exists as an epiphyte on olive trees, where it can survive and multiply, and the bacterium also exists as an endophyte (32). *Pseudomonas savastanoi* will quickly deteriorate when it reaches the soil (33,34). Different methods have been reported for the spread of the bacteria in the field. For example, it spreads over long distances on infected plant material, and can thus enter new areas besides spreading locally. Its spread in local areas or over short distances by rain and wind has been proved experimentally (35). It can also be spread by agricultural practices like pruning, grafting, irrigation and fruit harvesting (36). The involvement of the fruit fly in the bacterial spread has not been observed (29,31), except for a recent study (16,37) the reported olive fly as a vector over a short distance. The population size of the bacteria when present in leaves will be affected by the environmental conditions; while the bacteria can decline in dry cold weather, it can be seen to reach a higher level in warm and wet weather conditions. The increase in population size was reported to be high in spring and fall/autumn (38). Different bacterial species were detected on olive leaf surfaces, including three *Pantoea agglomerans, Xanthomonas campestris* and *Pseudomonas savastanoi* (28). Other species such as *Erwinia toletana,* and other bacteria from the genera *Burkholderia, Hafnia, Pseudomonas* and *Stenotrophomonas* were also reported in different studies (39).

3.5 Taxonomy

Domain: Bacteria

Phylum: Proteobacteria

Class: Gammaproteobacteria

Order: Pseudomonadales

Family: Pseudomonadaceae

Genus: Pseudomonas syringe complex

Species: *Pseudomonas savastanoi*

The pathogen belongs to the so-called *Pseudomonas syringae* complex. *Pseudomonas syringae* species complex is subdivided into more than 60 pathovars based on pathogenicity and host range, but it is divided into nine genomospecies based on DNA–DNA hybridization, and finally into 13 phylogenetic groups based on multilocus sequence analysis (19,40). *Pseudomonas savastanoi* is placed in genomospecies 2 Based on DNA-DNA relatedness (17,41,42) and phylogroup 3 based on MLST analysis of housekeeping genes (43). *Pseudomonas savastanoi* contains six pathovars, which have been classified on the basis of their host range and genetic information. These were initially pv. *savastanoi*, pv. *phaseolicola*, and *glycinea* (44), and then pv. *fraxini* and *nerii* (17), *and retacarpa* were added later (45).

3.6 Characterization of the pathogen

Pseudomonas savastanoi (17,42) is a gram-negative, rod-shaped, non-spore forming, an anaerobic bacterium with one to five polar flagella for motility. It is generally 0.4 -0.8-1.0–3.0 mm diameter in length, and optimum temperatures for growth are between 25–30 °C. Regarding its biochemical properties, it is Levan negative, oxidase negative and catalase-positive, and on King's B medium the bacteria produce a weak, blue-green fluorescent pigment (Smith, 1908). According to (20), *Pseudomonas savastanoi* strains were placed in the Ib group according to the LOPAT identification scheme. The bacterial biochemical tests are oxidase, arginine dehydration, and potato rot negative; it shows catalase-positive, and hyper-sensitivity in tobacco leaves is positive), also with the presence of some Levan positive (46). The six pathovars of *Pseudomonas savastanoi* can infect a wide range of herbaceous and woody plants, and develop different types of symptoms, which vary from leaf spots and blights to soft rots of fruits, wilts, scabs, overgrowths, and cankers based on the host (47). Overgrowths and cankers

represented our 4 pathovars. (48,49) has reported the particularity of olive and fraxini strains to be specific to olive and *fraxini* plants, while *nerii* is able to infect oleander and susceptible olive cultivars. Further investigation over the years has revealed the strain's pathogenicity and ability to attack different hosts. For example, the strains found on olive, jasmine, privet, pomegranate, *Forsythia spp.*, and *Phillyrea* spp. belong to pathovar *savastanoi*, while pathovar nerii was found on oleander, pathovar fraxini was detected and isolated from ash, and then pathovar retacarpa was found on *Retama sphaerocarpa*. All these pathovars are responsible for the formation of knots or canker on the above-mentioned plants (31). There is a wide range of diversity among *Pseudomonas savastanoi* pv. *savastanoi* strains, for example, they vary in virulence, and their effects vary from high to low according to olive cultivar. Other differences were also noticed in their motility, colony size, and morphology, and the size of tumours they produce in olive explants (25). Concerning the phenotypic differences, Levan positive was identified, and non-pigmented producing strains were also found (49,50).

In addition to these pathovars, *Pseudomonas savastanoi* includes two other pathovars: glycinea, which is responsible for bacterial blight of soybean, and pathovar *phaseolicola*, which is responsible for Halo blight of bean (*Phaseolus vulgaris* L.) (47).

3.7 The disease processes

Pseudomonas savastanoi pathovars colonized the lower surface of the leaves (17), in stems and leaves where the bacteria can multiply (51). The bacterium has been detected in the natural openings of the leaves, such as the stomata, and in symptomless plants (52). The population size of the bacteria when present in leaves will be affected by the environmental conditions. While the bacteria can decline in dry cold weather, it can reach a higher level in warm and wet weather conditions. Population size was reported to increase more in spring and fall/autumn (53). A crucial step for many plant pathogenic bacteria in their disease process is to be in contact with the hosts. The presence of epiphytic or endophytic phases of *Pseudomonas savastanoi* does not initiate the disease but these are considered the major source of available inoculum for its development (28). The bacteria may attack any type of wounds in the plant organs due to pruning, shoot emergence, and frost or hail injuries, insects, winds and birds or any mechanical damage (52). The bacteria can start the infection with low to medium temperatures of 5-27 °C but the development of symptoms depends on favourable conditions, such as temperature and humidity, mainly in fall and spring (36). The saprophytic phase of the

P. savastanoi could become pathogenic at any time, according to the weather conditions, inoculum, and availability of wounds (21). The bacteria can survive inside knots, which play a very important role as a source of inoculum, and spread the pathogen under favourable conditions, such as wind, rain and frost or other conditions. On all hosts, the pathogen initially colonizes the tissues around the wounds and disrupts the integrity of the host cells by degrading the cell wall with enzymes and indole acetic acid (IAA) production (55), which results in bores filled with bacteria. The bacteria produce IAA, which causes the plant tissue to increase in size (hypertrophy), and then the cells start divide abnormally (hyperplasia). Finally, there is a differentiation of certain cells of the hyperplastic area, elements of xylem and phloem (56).

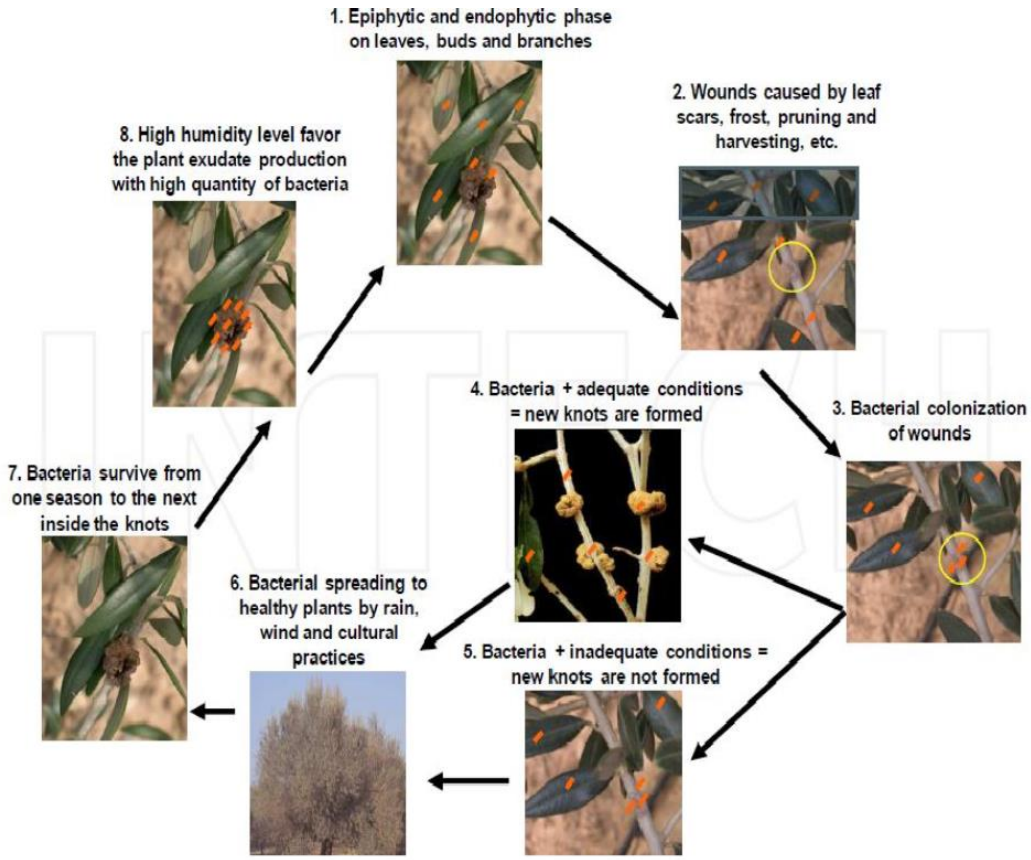


Figure 14: The disease process of olive knot caused by *Pseudomonas savastanoi* pv. *savastanoi*

3.8 Symptoms

It is easy to recognize the symptoms of the olive knot in the field. The most noticeable symptoms on olive tree or oleander are the hyperplastic growths of tumorous galls or knots, and these overgrowth tissues can easily be seen on the stems and branches of the host plant (21). Their presence on the leaves and fruits of olive is unusual but has been observed (19).

A cross-section of a knot on olive shows a dark core surrounded by healthy tissue. The symptoms on olive of heavy infection can cover the entire branch of the infected tree. In oleander, besides the symptoms mentioned, the bacterium also deforms the flowers and seed pods, and/or decreases blooming, and causes the death of pistils (18). The symptoms are a little different on ash since there are no tumours, but cankers or wart-like excrescences are present (57). Different virulence factors contribute to the formation of the olive knot, as will be discussed next.

Figure 15: Knot disease symptoms : (A) Knots on Olive twigs; (B) Horizontal cut on knot; (C) Oleander knots;(D) Canker on Ash ; (E) Infected olive tree.

3.9 Pathogenicity and virulence factors

Different systems are used by bacterial plant pathogens to invade the host plant. The role of these systems in the invasion is to damage the host cells and/or subvert its immune system so that the pathogen can become established and multiply in the new environment (24,58). *Pseudomonas syringae* pathovars have a wide range of important factors that play diverse roles in pathogenicity and virulence. These include regulation phytotoxins, cell wall-degrading

enzymes, extracellular polysaccharides, iron uptake systems, resistance to plant-derived antimicrobials, adhesion, and the general processes of motility and chemotaxis (59). In the olive knot pathogen, many virulence factors have been identified as playing a role in the development of knot disease, but the most important ones are phytohormones, Type III secretion system (T3SS) and its effectors, and quorum sensing (QS)(19).

The pathogen *Pseudomonas savastanoi* pv. *savastanoi* uses several virulence factors that help it to invade the plant and form the knots. Phytohormones used are mainly indole-3acetic acid (IAA) and cytokinin. Studies have proved that the development of overgrowth is due to the production of indole-3acetic acid (IAA) and cytokinin by the bacteria (60). Indole-3acetic acid (IAA) is involved in the modification of plant cell size and rapid cell division using a pathway different from the one used in the plant to grow (61–63). In this pathway, the iaaM gene converts the tryptophan into indole acetamide iaaM, then the iaaH gene catalyzes the conversion of IAM into IAA. The activities of the iaaM, and iaaH genes which convert tryptophan 2-monooxygenase and indole acetamide hydrolase and the two enzymes produced by these two genes catalyze the synthesis of IAA from L-tryptophan (21,64).

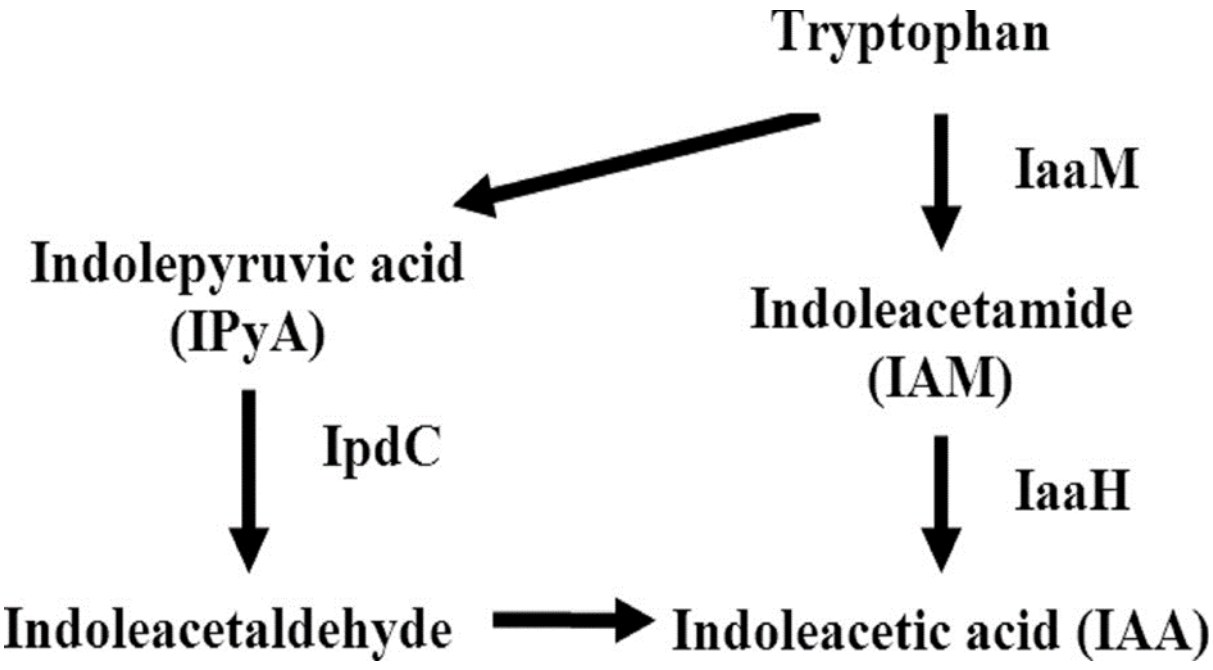


FIGURE 16: INDOLE ACETIC ACID PATHWAYS (61).

Cytokinin plays a crucial role in plant development. There is a positive correlation between *Pseudomonas savastanoi* pv. *savastanoi* and pv. *neri* strains in the production of large knots

(65). Cytokinin is involved in knot formation as a contributor to tumour formation but is not essential. In other words, it helps to increase the size of the knots, but not to form them (61). Another possible factor is the phytohormone ethylene, which is produced by the plant due to the presence of the pathogen *Pseudomonas savastanoi*. the role of this compound in addition to its involvement in abscission and senescence, ethylene is also responsible for chlorosis of the plant tissues (60,64).

The Type III secretion system (TTSS) determined by the hrp and/or hrc gene clusters is an important factor used by *P. savastanoi* pathovars and is also a character among different plants and animals gram-negative bacterial pathogen. Through the Type III secretion system, the pathogen delivers multiple molecules (proteins) into plant cells to start the infection, subvert the immune system and damage the plant cells (66). There are more than 40 effecters present in the *P. savastanoi* pv. *savastanoi* genome and each one or group attacks a different part of the host cell; some effectors attack plant immunity, and others attack plasma membrane components

The involvement of the Type III secretion system (TTSS) determined by the hrp and/or hrc gene clusters was demonstrated in the knot formation, whereas it plays a key role in secreted phytohormones or virulence factors and other metabolic pathways (67).

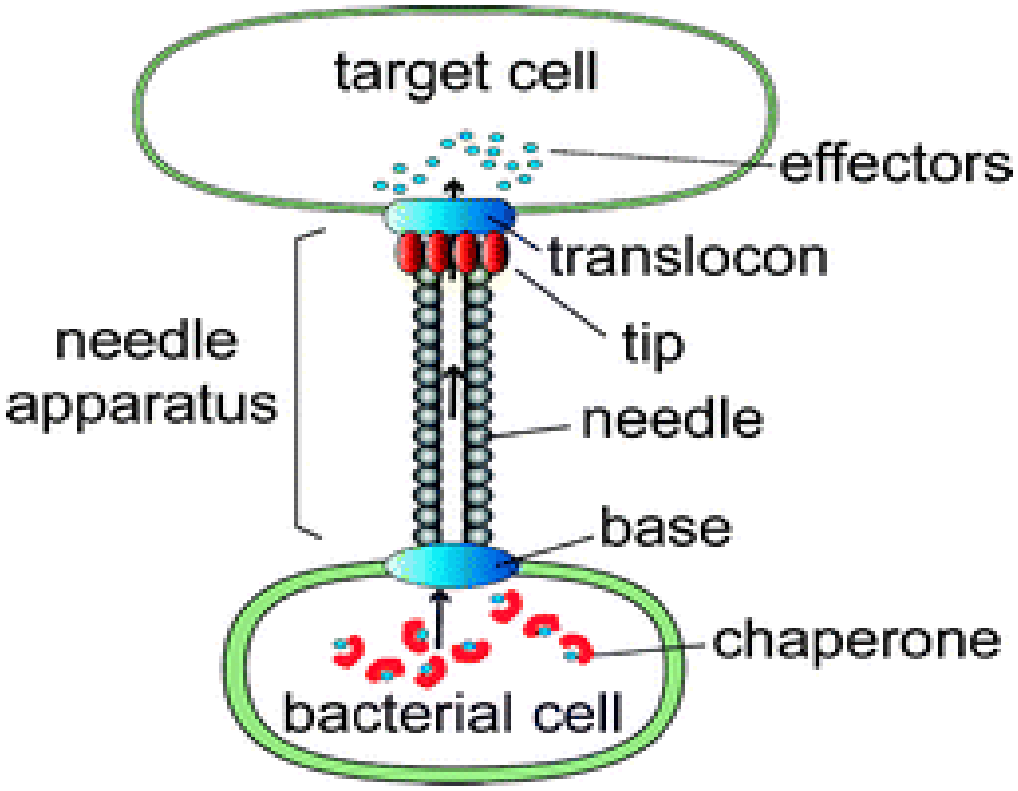


Figure 17:Type III secretion system (TTSS) and effectors delivery (68,69).

This idea was supported by a study (70) that proved that the necrosis of the hyperplastic tissue in *Pseudomonas savastanoi* pv. *savastanoi* is based on T3SS. An additional aspect that has an effect on tumour size is cyclic diguanylate (c-di-GMP), a secondary messenger in bacteria that controls exchange between mobile and fixed lifestyles (29). In agreement with what has been mentioned above different factors have been identified as involved in knot formation, such as Type II and IV secretion system genes, tolerance and detoxification of reactive oxygen species genes, a methyl-accepting chemotaxis protein gene, the production of complex molecules of the cell wall (71). The woody host and *Pseudomonas* (WHOP) region which is comprised of 15 kb gene clusters used to metabolize the phenolic compounds, was identified and found to be involved in knot formation in *Pseudomonas savastanoi* pv. *savastanoi* (68). Plasmids and insertion sequences pathogens can acquire new genes that can be helpful to the pathogen for adaptation or colonization of a new environment have also been found to be involved in pathogenicity and virulence activities (60,72).

In addition to growth regulator factors, (73) have pointed out the involvement of other virulence factors in knot formation. A sort of cooperation and communication (Quorum sensing

molecules) between *Pantoea agglomernas* and *Erwinia toletana* strains and *Pseudomonas savastanoi* was found to be involved in increase knot formation (74).

3.10 References

1. Sawada H, Fujikawa T. Genetic diversity of *Pseudomonas syringae* pv. *actinidiae*, pathogen of kiwifruit bacterial canker. Plant Pathol. 2019;68(7):1235–48.

2. Zohary D, Spiegel-Roy P. Beginnings of fruit growing in the Old World. Science (80-). 1975;187(4174):319–27.

3. Bracci T, Busconi M, Fogher C, Sebastiani L. Molecular studies in olive (Olea europaea L.): overview on DNA markers applications and recent advances in genome analysis. Plant Cell Rep. 2011;30(4):449–62.

4. Hall BH. New challenges for pest and disease management in olive orchards and nurseries. In: XXVIII International Horticultural Congress on Science and Horticulture for People (IHC2010): Olive Trends Symposium-From the 924. 2010. p. 127–35.

5. Delrio G. Integrated control in olive groves. In: Biological Control and Integrated Crop Protection: Towards Environmentally Dafer Agriculture, Proc Int Conf IOBC/WPRS, Veldhohen, Netherlands (8-13 September, 1991). 1992. p. 67–76.

6. Barba M. Viruses and virus-like diseases of olive 1. EPPO Bull. 1993;23(3):493–7.

7. Castillo P, Nico AI, Navas-Cortes JA, Landa BB, Jimenez-Diaz RM, Vovlas N, et al. Plant-parasitic nematodes attacking olive trees and their management. Plant Dis [Internet]. 2010;94(2):148–62. Available from: http://www.ncbi.nlm.nih.gov/pubmed/30754270

8. Çağlayan K, Faggioli F, Barba M. Viruses, phytoplasmas, and diseases of unknown etiology of olive trees. Virus Virus-Like Dis Pome Stone Fruits. 2011;289–97.

9. Martelli GP. A brief outline of infectious diseases of olive. Palest Tech Univ Res J. 2013;1(1):1–9.

10. Goud JC, Hiemstra JA. Verticillium wilt of major tree hosts: other tree species. 1998;ok.

11. Sergeeva V, Spooner-Hart R. Olive diseases and disorders in Australia. Olive Dis Disord Aust. 2010;59:29–32.

12. Saponari M, Boscia D, Nigro F, Martelli GP. Identification of DNA sequences related to *Xylella fastidiosa* in oleander, almond and olive trees exhibiting leaf scorch symptoms in Apulia (Southern Italy). J Plant Pathol. 2013;95(3).

13. Therios IN. Olives. Vol. 18. CABI; 2009.

14. Tsitsipis J, Varikou K, Kalaitzaki A, … VA-4th EM of the, 2009 undefined. Chemical control of olive pests: Blessing or curse.

15. Almeida RPP, De La Fuente L, Koebnik R, Lopes JRS, Parnell S, Scherm H. Addressing the new global threat of *Xylella fastidiosa* . Phytopathology. 2019;109(2):172–4.

16. Panagopoulos CG. Olive knot disease in Greece 1. EPPO Bull. 1993;23(3):417–22.

17. Gardan L, Bollet C, Ghorrah MA, Grimont F, Grimont PAD. DNA relatedness among the pathovar strains of *Pseudomonas syringae* subsp. *savastanoi* Janse (1982) and proposal of *Pseudomonas savastanoi* sp. nov. Int J Syst Evol Microbiol. 1992;42(4):606–12.

18. Iacobellis NS. Olive knot. Encycl Plant Pathol. 2001;2:713–5.

19. Ramos C, Matas IM, Bardaji L, Aragón IM, Murillo J, Aragon IM, et al. *Pseudomonas savastanoi* pv. *savastanoi* : some like it knot. Mol Plant Pathol [Internet]. 2012;13(9):998–1009. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22805238

20. Janse JD. *Pseudomonas syringae* subsp. *savastanoi* (ex Smith) subsp. nov., nom. rev., the bacterium causing excrescences on Oleaceae and *Nerium oleander* L. Int J Syst Evol Microbiol. 1982;32(2):166–9.

21. Quesada JM, Penyalver R, López MM. Epidemiology and control of plant diseases caused by phytopathogenic bacteria: the case of olive knot disease caused by *Pseudomonas savastanoi* pv. *savastanoi* . Plant Pathol. 2012;299–326.

22. Buonaurio R, Moretti C, da Silva DP, Cortese C, Ramos C, Venturi V. The olive knot disease as a model to study the role of interspecies bacterial communities in plant disease. Front Plant Sci [Internet]. 2015;6:434. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26113855

23. Bradbury JF. Guide to plant pathogenic bacteria. CAB international; 1986.

24. Iacobellis NS, Caponero A, Evidente A. Characterization of *Pseudomonas syringae* sp. *savastanoi* strains isolated from ash. Plant Pathol. 1998;47(1):73–83.

25. Moretti C, Vinatzer BA, Onofri A, Valentini F, Buonaurio R. Genetic and phenotypic diversity of Mediterranean populations of the olive knot pathogen, *Pseudomonas savastanoi* pv. *savastanoi* . Plant Pathol. 2017;66(4):595–605.

26. Janse JD. The bacterial disease of ash (Fraxinus excelsior), caused by *Pseudomonas syringae* subsp. *savastanoi* pv. Fraxini II. Etiology and taxonomic considerations. For Pathol. 1981;11(7):425–38.

27. Snover KL, Arneson PA. Symptoms and signs. 2002;

28. Lamichhane JR. Phytobacteriological studies related to the native (*Olea cuspidata* Wall.) and introduced (*Olea europaea* L.) olive species in Nepal. 2011;

29. Rodriguez-Palenzuela P, Matas IM, Murillo J, Lopez-Solanilla E, Bardaji L, Perez-Martinez I, et al. Annotation and overview of the *Pseudomonas savastanoi* pv. *savastanoi* NCPPB 3335 draft genome reveals the virulence gene complement of a tumour-inducing pathogen of woody hosts. Env Microbiol [Internet]. 2010;12(6):1604–20. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20370821

30. Schroth MN, Osgood JW, Miller TD. Quantitative assessment of the effect of the olive knot disease on olive yield and quality. Phytopathology. 1973;63(8):1064–5.

31. Young JM. Olive knot and its pathogens. Australas Plant Pathol. 2004;33(1):33–9.

32. CABI. *Pseudomonas savastanoi* pv. *savastanoi* (oleander knot). 2018; Available from: http://www.cabi.org/isc.

33. Krid S, Rhouma A, Mogou I, Quesada JM, Nesme X, Gargouri A. *Pseudomonas savastanoi* endophytic bacteria in olive tree knots and antagonistic potential of strains of *Pseudomonas fluorescens* and *Bacillus subtilis*. J Plant Pathol. 2010;335–41.

34. Nguyen K. Epidemiology and Management of Olive Knot Caused by *Pseudomonas savastanoi* pv. *savastanoi* in California Olive Production. UC Riverside; 2017.

35. Surico G. Histological observations on tumours of olive knot. Phytopathol Mediterr. 1977;16(2/3):109–25.

36. Wilson EE. The Olive knot disease: its inception, development, and control. Hilgardia. 1935;9(4).

37. Lüthy P, Studer D, Yamvrias C. The bacterial symbiote of the olive fruit fly (Dacus oleae). Experientia. 1983;39(12).

38. Quesada JM, Penyalver R, Pérez-Panadés J, Salcedo CI, Carbonell EA, López MM. Dissemination of *Pseudomonas savastanoi* pv. *savastanoi* populations and subsequent appearance of olive knot disease. Plant Pathol. 2010;59(2):262–9.

39. Ercolani GL. Distribution of epiphytic bacteria on olive leaves and the influence of leaf age and sampling time. Microb Ecol. 1991 Dec;21(1):35–48.

40. Hosni T, Moretti C, Devescovi G, Suarez-Moreno ZR, Fatmi MB, Guarnaccia C, et al. Sharing of quorum-sensing signals and role of interspecies communities in a bacterial plant disease. ISME J [Internet]. 2011;5(12):1857. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21677694

41. Baltrus DA, Nishimura MT, Romanchuk A, Chang JH, Mukhtar MS, Cherkis K, et al. Dynamic evolution of pathogenicity revealed by sequencing and comparative genomics of 19 *Pseudomonas syringae* isolates. PLoS Pathog [Internet]. 2011;7(7):e1002132. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21799664

42. Young JM. Taxonomy of *Pseudomonas syringae*. J Plant Pathol. 2010;S5–14.

43. Gardan L, Shafik H, Belouin S, Broch R, Grimont F, Grimont PAD. DNA relatedness among the pathovars of *Pseudomonas syringae* and description of *Pseudomonas tremae* sp. nov. and *Pseudomonas cannabina* sp. nov.(ex Sutic and Dowson 1959). Int J Syst Evol Microbiol. 1999;49(2):469–78.

44. Parkinson N, Bryant R, Bew J, Elphinstone J. Rapid phylogenetic identification of members of the *Pseudomonas syringae* species complex using the rpoD locus. Plant Pathol. 2011;60(2):338–44.

45. De Los Rios JEG. Retama sphaerocarpa (L.) Boiss., a new host of *Pseudomonas savastanoi*. Phytopathol Mediterr. 1999;54–60.

46. Lelliott RA, Stead DE. Methods for the diagnosis of bacterial diseases of plants. Blackwell Scientific Publications; 1987.

47. Marchi G, Viti C, Giovannetti L, Surico G. Spread of levan-positive populations of *Pseudomonas savastanoi* pv. *savastanoi* , the causal agent of olive knot, in central Italy. Eur J plant Pathol. 2005;112(2):101–12.

48. Kennelly MM, Cazorla FM, de Vicente A, Ramos C, Sundin GW. *Pseudomonas syringae* diseases of fruit trees: progress toward understanding and control. Plant Dis [Internet]. 2007;91(1):4–17. Available from: http://www.ncbi.nlm.nih.gov/pubmed/30781059

49. Perez-Martinez I, Zhao Y, Murillo J, Sundin GW, Ramos C, Pérez-Martínez I, et al. Global genomic analysis of *Pseudomonas savastanoi* pv. *savastanoi* plasmids. J Bacteriol [Internet]. 2008;190(2):625–35. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17993520

50. Penyalver R, Garcia A, Ferrer A, Bertolini E, Quesada JM, Salcedo CI, et al. Factors Affecting *Pseudomonas savastanoi* pv. *savastanoi* Plant Inoculations and Their Use for Evaluation of Olive Cultivar Susceptibility. Phytopathology [Internet]. 2006;96(3):313–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18944447

51. Surico G. Scanning electron microscopy of olive and oleander leaves colonized by *Pseudomonas syringae* subsp. *savastanoi* . J Phytopathol. 1993;138(1):31–40.

52. Quesada JM, Penyalver R, López MM. Epidemiological Basis for an Efficient Control of *Pseudomonas savastanoi* pv. *savastanoi* on Olive Trees. In: *Pseudomonas syringae* Pathovars and Related Pathogens–Identification, Epidemiology and Genomics. Springer; 2008. p. 57–64.

53. Teviotdale BE. Diseases of olive. Olive Prod Man. 1994;3353:107–9.

54. Quesada JM, García A, Bertolini E, López MM, Penyalver R. Recovery of *Pseudomonas savastanoi* pv. *savastanoi* from symptomless shoots of naturally infected olive trees. Int Microbiol. 2007;10(2):77.

55. Marchi G, Mori B, Pollacci P, Mencuccini M, Surico G. Systemic spread of *Pseudomonas savastanoi* pv. *savastanoi* in olive explants. Plant Pathol. 2009;58(1):152–8.

56. Magie AR. Physiological factors involved in tumor production by the oleander knot pathogen, *Pseudomonas savastanoi* . University of California, Davis.; 1963.

57. Smith CO. Oleander bacteriosis in California. Phytopathology. 1928;18(6).

58. Cinelli T, Marchi G, Cimmino A, Marongiu R, Evidente A, Fiori M. Heterogeneity of *Pseudomonas savastanoi* populations infecting M yrtus communis in S ardinia (I taly). Plant Pathol. 2014;63(2):277–89.

59. Green ER, Mecsas J. Bacterial Secretion Systems: An Overview. Microbiol Spectr [Internet]. 2016;4(1). Available from: http://www.ncbi.nlm.nih.gov/pubmed/26999395

60. Caballo-Ponce E, Murillo J, Martinez-Gil M, Moreno-Perez A, Pintado A, Ramos C, et al. Knots Untie: Molecular Determinants Involved in Knot Formation Induced by *Pseudomonas savastanoi* in Woody Hosts. Front Plant Sci [Internet]. 2017;8:1089. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28680437

61. Surico G, Iacobellis NS, Sisto A. Studies on the role of indole-3-acetic acid and cytokinins in the formation of knots on olive and oleander plants by *Pseudomonas syringae* pv. *savastanoi* . Physiol Plant Pathol. 1985;26(3):309–20.

62. Akiyoshi DE, Regier DA, Gordon MP. Cytokinin production by *Agrobacterium* and *Pseudomonas* spp. J Bacteriol. 1987;169(9):4242–8.

63. Cerboneschi M, Decorosi F, Biancalani C, Ortenzi M V, Macconi S, Giovannetti L, et al. Indole-3-acetic acid in plant-pathogen interactions: a key molecule for in planta bacterial virulence and fitness. Res Microbiol [Internet]. 2016;167(9–10):774–87. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27637152

64. Rodríguez-Moreno L, Barceló-Muñoz A, Ramos C, Rodriguez-Moreno L, Barcelo-Munoz A, Ramos C, et al. In vitro analysis of the interaction of *Pseudomonas savastanoi* pvs. *savastanoi* and nerii with micropropagated olive plants. Phytopathology [Internet]. 2008;98(7):815–22. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18943258

65. Yang S, Zhang Q, Guo J, Charkowski AO, Glick BR, Ibekwe AM, et al. Global effect of indole-3-acetic acid biosynthesis on multiple virulence factors of *Erwinia chrysanthemi* 3937. Appl Environ Microbiol. 2007;73(4):1079–88.

66. Mattoo AK. The Plant Hormone Ethylene: 0. CRC press; 2017.

67. Cunnac S, Lindeberg M, Collmer A. *Pseudomonas syringae* type III secretion system effectors: repertoires in search of functions. Curr Opin Microbiol [Internet]. 2009;12(1):53–60. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19168384

68. Matas IM, Lambertsen L, Rodríguez-Moreno L, Ramos C, Rodriguez-Moreno L, Ramos C. Identification of novel virulence genes and metabolic pathways required for full fitness of *Pseudomonas savastanoi* pv. *savastanoi* in olive (*Olea europaea)* knots. New Phytol [Internet]. 2012;196(4):1182–96. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23088618

69. Matas IM, Castaneda-Ojeda MP, Aragon IM, Antunez-Lamas M, Murillo J, Rodriguez-Palenzuela P, et al. Translocation and functional analysis of *Pseudomonas savastanoi* pv. *savastanoi* NCPPB 3335 type III secretion system effectors reveals two novel effector families of the *Pseudomonas syringae* complex. Mol Plant Microbe Interact [Internet]. 2014;27(5):424–36. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24329173

70. Wang Y, Zhang L, Picking WLWD, Picking WLWD, De Guzman RN. Structural dissection of the extracellular moieties of the type III secretion apparatus. Mol Biosyst. 2008;4(12):1176–80.

71. Aragón IM, Pérez-Mendoza D, Gallegos M, Ramos C. The c-di-GMP phosphodiesterase BifA is involved in the virulence of bacteria from the *Pseudomonas syringae* complex. Mol Plant Pathol. 2015;16(6):604–15.

72. Caballo-Ponce E, van Dillewijn P, Wittich RM, Ramos C. WHOP, a Genomic Region Associated With Woody Hosts in the *Pseudomonas syringae* Complex Contributes to the Virulence and Fitness of *Pseudomonas savastanoi* pv. *savastanoi* in Olive Plants. Mol Plant Microbe Interact [Internet]. 2017;30(2):113–26. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28027024

73. Marchi G, Cinelli T, Surico G. 6. A review on *Pseudomonas savastanoi* genetic traits involved in disease development and in symptom induction. 2014;

74. Iacobellis NS, Sisto A, Surico G, Evidente A, DiMaio E. Pathogenicity of *Pseudomonas syringae* subsp. *savastanoi* mutants defective in phytohormone production. J Phytopathol. 1994;140(3):238–48.

3.11 **Article 2**

Genetic diversity and population structure of *Pseudomonas savastanoi,* an endemic pathogen of the Mediterranean area, as revealed by a novel MLVA assay

Yaseen Jundi Rahi[1,3], Maria Claudia Taratufolo[1], Silvia Turco[1], Matteo Cerboneschi[2], Stefania Tegli[2], Franco Valentini[3], Annamaria D'Onghia[3], Nicola Sante Iacobellis[4], Giorgio Mariano Balestra[1], Angelo Mazzaglia[1]

[1]University of Tuscia, Department of Agricultural and Forestry Sciences, Via S. Camillo de Lellis snc, 01100 Viterbo, Italy;
[2]University of Florence, Agrifood Production and Environmental Sciences (DISPAA), Piazzale delle Cascine 18, 50144 Florence, Italy
[3]CIHEAM-Mediterranean Agronomic Institute of Bari, Via Ceglie 9, 70010 Valenzano (Ba), Italy,
[4]School of Agricultural, Forestry, Food, and Environmental Sciences, University of Basilicata, Viale dell'Ateneo Lucano, 85100 Potenza, Italy

Corresponding author: *angmazza@unitus.it*

Abstract

Pseudomonas savastanoi is a bacterial species included in the *Pseudomonas syringae* complex, that is further subtyped in pathovars causing a group of diseases of woody plants, such as "knot disease" on olive and oleander and bacterial canker on ash. All these diseases are present in the Mediterranean area for a long time so that this species can be considered an endemic pathogen. Here, an MLVA approach was developed to assess genetic relationships among and within these pathovars, with a specific focus on *P. savastanoi* pv. *savastanoi*. Analysing the genome of the *P. savastanoi* pv. *savastanoi* strain NCPPB 3335 (accession n°CP008742), 14 Tandem Repeat (TR) loci were identified and corresponding primers were designed and used for the amplification of genomic DNAs from 84 strains belonging to *Pseudomonas savastanoi* pathovars. The data were analysed using different approaches, hierarchical clustering, STRUCTURE, and DAPC, to evaluate the effectiveness of the assay in depicting pathovars and population structure of the pathogen. Results reveal a very complex articulation of genetic relationship as expected for a long-time evolving pathogen, being however able to discriminate the pathovars each other. At intrapathovar level, the MLVA assay cluster together isolates mainly according to their hosts and geographic origin and resulted particularly useful in identification and tracking *P. savastanoi* populations at local level.

Keywords*: knot disease, VNTR, tandem repeats, MLVA, DAPC, population structure, endemic disease.*

INTRODUCTION

The term “Knot disease” in plant pathology describes the excessive growth of plant tissues or organs that leads to hyperplastic and/or hypertrophic masses in the shape of tumours or knots. These formations can be clearly observed on the aerial parts of diseased plants mostly on the woody parts of the plant, such as stem, branches or twigs, more rarely on leaves and fruits. The disease is in general caused by the Gram-negative bacterium *Pseudomonas savastanoi* (1–5). It is a monomorphic pathogenic gamma-proteobacterium species belong to genomospecies 2 of the *Pseudomonas syringae* complex (6). At the moment, six pathovars have been described within this species: pv. *savastanoi* (7,8), pv. *fraxini* and pv. *nerii* (1), and pv. *retacarpa* (9). They attack different members of plant families including Oleaceae (Olea europaea, *Jasminum officinalis*, *Forsythia spp*, *Phillyrea spp*, and *Ligustrum japonicum*), *Fabaceae* (Retama), *Rhamnaceae Rhamnaceae* spp. *Myrtaceae* (*Myrtus communis*), *Apocinaceae* (*Nerium oleander*, *Mandevilla sanderi*) and *Lythraceae* (*Punica granatum*) (10–13). In ash, the symptoms are slightly different, being tumors or knots rather substituted by vertical and lateral cracks and canker along with the plant (14). Then, two pathovars are currently classified under this species: pv. *phaseolicola* (8,15) is responsible for the halo blight of beans (*Phaseolus vulgaris* L.) (16), an exception was reported that some strains of pv. *phaseolicola* attack mulberry (17), while pv. *glycinea* (18), the causal agent of bacterial halo blight of soybean (Glycine max) and restricted to this host. *Pseudomonas savastanoi* pathovars are classified based on host range and genetic information (19). Early studies have focused on the differentiation and clarification of different pathovars using different morphological, serological, physiological, nutritional, biochemical, numerical taxonomy and DNA-DNA hybridization (1,8). These methods have shown during the time some limitations, especially when differentiation or relatedness at strain level are required

(20). The movement of goods around the world and open borders have facilitated the diffusion of plant diseases through the movement of pathogens into a new area, and climate change has helped the emergence or re-emergence of new pathogens which negatively effect on food production (21). It is very important to understand how a pathogen emerges and adapts to the environment. Population genetic analysis could be the best tool to understand the previous two concepts (22). Knowledge of population genetics has been increased by coalescent theory, computational methods, and molecular biology, which have opened up the genomic era (23).

Only a few tools have been developed to investigate in detail and understand the genetic diversity among the strains of *Pseudomonas savastanoi* pathovars. The need for a reliable method that can accurately discriminate between strains for crop surveillance, outbreak investigation, and study the evolutionary of plant pathogens to establishing disease control strategies is a must (24). Molecular typing methods are fast and powerful tools that enable us to differentiate closely related strains, as a fast tool in an epidemiological survey, to determine the relatedness among the strains and to track their origin and pathways of spread. Also, molecular approaches can well define the evolutionary, host adaption and genetic diversity (25–27). Furthermore, it could reveal the biological features of the pathogen that can affect virulence, pathogenicity, and host specificity, helping us to set up sustainable control measures (28,29).

The first approach applied to study *Pseudomonas savastanoi* was (RFLP) Restriction Fragment Length Polymorphism on a group of Italian strains from olive, oleander, and ash for evaluating differences among these strains (30). In other studies, the genetic diversity of *Pseudomonas savastanoi* pathovars strains was calculated using repetitive PCR and Random amplification of polymorphic DNA (RAPD) methods to understand the relationships between the strains and their geographical distribution (26,31)

Further, fluorescent amplified fragment length polymorphism (f-AFLP) analysis was also used to understand the genetic variability among the pool of representative strains from different olive regions around the world (32). In Spain, Quesada and others have investigated the genetic diversity of *Pseudomonas savastanoi* pv. *savastanoi* using the Insertion element (IS53) for typing (33). More recently, (13) have studied the genotypic diversity of Mediterranean populations using repetitive element palindromic polymerase chain reaction rep-PCR and multilocus sequence typing (MLST) of four different genes, gap, gltA, gyrB, and rpoD. Despite some good results obtained by these methods, still, they are facing struggles such as being laborious and time-consuming, and often being poorly reproducible, which is the hardest limitation to compare results across laboratories (34,35). MLST is one of the best and widest applied methods for subtyping bacterial plant pathogens, but it also has limits. However, the results obtained by MLST are similar to that one obtained by DNA-DNA hybridization in the *P. syringae* complex (36). Indeed, the housekeeping genes sequences compared in this method can have very small amount of sequence diversity between individuals of the same species, which results in failures when the aim is to resolve the evolutionary patterns of the bacterial populations (37), or to understand the populations pattern of highly homogeneous bacterial pathogens (38). In recent years whole-genome sequencing (WGS) and (SNPs) have been widely applied to different bacteria, including *Pseudomonas savastanoi* pathovars (39–42). These methods provide all the details about the entire genome sequences, but their cost is high, require advanced next-generation sequencing technologies, and most of all, they require high knowledge and computational biology skills, as well as devoted infrastructural facilities (43,44).

MLVA (Multiple Loci Variable number of tandem repeat Analysis) is a very promising typing technique. This molecular method is based upon the calculation of Variable copy Numbers of Tandem Repeats (VNTR). Tandemly repeated sequences (TR loci) have been

found scattered throughout the prokaryotic and eukaryotic genomes (45,46) and are among the most variable regions in bacterial genomes. Therefore, MLVA has the potential to resolve the genetic diversity of monomorphic pathogens. TRs occur in the genome due to a mismatch in DNA strands during replication (slipped-strand mispairing) or DNA recombination error (47,48). TRs have been found to play an important role in bacterial adaption (49), phase variation (47), speed of loss gene disorder (50) they facilitate evolutionary change (51,52) genome plasticity and variation (53). In order to calculate these TRs, PCR primers must be designed from the conserved region around these TRs, while the differences among the strains could be measured by the capillary electrophoreses and translated numerically after the calculation of these loci by detracting the right and the left flanking regions from the amplicons from PCR and dividing the remaining length by the size of the corresponding repeat unit in that locus. The generated numeric data can be exchanged among laboratories around the world via the public MLVA database (38).

Indeed, its great discriminatory power for the strains based on differences in the number of the repeats in each locus generated by amplification of the genomic DNA of several loci results in a high level of polymorphism (54). In addition, MLVA is a fast procedure for monitoring short-term, local outbreaks of bacterial pathogens revealing insights about the relationships at a microevolutionary level (55).

MLVA methodology primarily was developed for *Haemophilus influenza* strains and bacterial human pathogens (56–58), and animal pathogens (59). MLVA was applied for the first time for the plant pathogenic bacteria *Xylella fastidiosa* subtyping (60), and after that, it has been used in a wide range for subtyping plant pathogenic bacteria such as *Xanthomonas citri* pv. *citri* (61,62), *Xanthomonas arboricola* pv. *pruni* (63), *Pseudomonas syringae* pv. *maculicola*; and *Pseudomonas syringae* pv. *tomato* (64), *Xanthomonas oryzae* pv. *oryzicola* (35), *Ralstonia solanacearum* (65), *Clavibacter michiganensis subsp. michiganensis* (66),

Erwinia amylovora (67) (Bühlmann et al., 2014), *Pseudomonas syringae* pv. *actinidiae* (68), and *Ralstonia solanacearum* (69).

In this study, we identified for the first time VNTRs loci, and designed an MLVA assay, to recognize *Pseudomonas savastanoi* pathovars, and understand genetic relationships between populations within *Pseudomonas savastanoi* pv. *savastanoi* in a wide collection of strains including worldwide isolates from an olive orchard from one region and a single tree of the same field.

MATERIAL AND METHODS

Bacterial strains and growth conditions

The study was conducted on a group of 84 strains of *Pseudomonas savastanoi* (Psv) strains representative of the pathovars *savastanoi, neri, fraxini, retacarpa,* and *phaseolicola.*

Sixty-two strains were obtained from the DAFNE collection, the University of Tuscia in Viterbo, and from CIHEAM collection in Bari. Twenty-four additional samples were isolated from an olive orchard in Viterbo, Italy, 14 of them were collected randomly in the orchard, whilst 10 were obtained from a single tree. The bacteria were grown on King's B medium (KB) at 28°C for 48hrs before DNA extraction.

Genomic DNA extraction

Genomic DNA was extracted from freshly grown colonies using the QIAGEN kit (GmbH, Germany) following the manufacturer's instruction. DNA quantification was obtained by a Qubit Fluorometer (Invitrogen, Life Technologies Italia, Monza, Italy), then its concentration was adjusted to 40 ng/µl with TE (10mMTris-HCl, 1mM EDTA) buffer at pH 8.0, before storage at -20 C° until use. The identity of all the strains as *Pseudomonas savastanoi* was confirmed through iaal gene amplification (70).

Tandem repeats identification and design of VNTR primers

The complete genome sequence of the *Pseudomonas savastanoi* pv. *savastanoi* strain NCPPB 3335 (Gene Bank accession number CP008742) was analyzed for the presence of candidate VNTR loci using Tandem Repeats Finder program (71) with the following parameters: tandem repeat ranges from 50 to 1000 bp, repeat unit length from 5 to 300 bp and similarity higher than 80% within the copies of the tandem repeat array.

The VNTRs matching the predicted polymorphism size among *savastanoi, neri,* and *fraxini* pathovars were selected. In order to confirm the presence of these regions in all the available WGS sequences, both the right and left flanking regions of about 100 bp each was confirmed by BLAST (72). Primer pairs were designed by Primer3plus software (http://www.bioinformatics.nl/cgi-qbin/primer3plus/primer3plus) in the flanking regions of each selected tandem repeat.

PCR amplification, agarose gel electrophoresis.

The selected VNTR loci were amplified using a C1000™ Thermal Cycler (Bio-Rad, USA). Each PCR reaction mix contained 12.5 µL 2X GoTaq® master mix (Promega, Madison, USA), 9.5 µL nuclease-free water and 1 µL (40 ng) of template DNA, and 1µL of forward primer and 1µl reverse primer corresponding to 10 µM concentration, primers reported in Table (I) to a final volume of 25 µL. All the PCR reactions were run with an initial denaturation step of 5 min at 94°C followed by 35 cycles at 94°C for 30 seconds, 50 °C to 67°C (depending on the primer pair) for 30 seconds and 72°C for 2 minutes with a final extension step of 10 min at 72°C. PCR products were separated on 1.5% agarose gels and visualized under UV upon Gel Red® Nucleic Acid to confirm positive amplification. A random selection of samples underwent Sanger sequencing to check the number of tandem repeats.

VNTR analysis by capillary electrophoresis

The amplicons were analyzed using a QIAxcel multi-capillary electrophoresis system (QIAGEN, Milan, Italy). A DNA High-Resolution gel cartridge and the OM800 method were used to estimate amplicons sizes below 600 bp, whereas the OM500 method was used for amplicons larger than 600 bp, with the following run parameters: 10s of sample injection time; 5kV of sample injection voltage; 3kVof separation voltage for 700s of separation time while for the OM500, that is characterized by higher separation voltage (5kV) and shorter separation time (500 s). The final results were analyzed and interpreted by means of the Screen gel software (QIAGEN), which gives estimates of both size and concentration of amplicons.

Data analysis

The data matrix containing the tandem repeats number for 15 different loci, among 84 isolates, was imported into R version 3.4.4 (73) and transformed into genind object using the R package *adegenet* 2.1.1 (74). All the analyses of population structure and genetic diversity were performed using R packages *poppr*, version 2.8.1 (75,76), *adegene*t 2.1.1 (74), and ade4

version 1.7-13 (77). R package vegan, version 2.5-4, was used to calculate the genotype richness and evenness (78).

In order to reveal the population structure, a hierarchical clustering analysis was performed with *hclust*() function of the package stats (73), using Bruvo's distance and UPGMA algorithm. The result was visualized as a dendrogram with the R package factoextra version 1.0.5 (79). Furthermore, poppr *bruvo.boot*() function was used to calculate Bruvo's distance with bootstraps and a cut-off threshold of 80 was set.

The population structure was also evaluated using STRUCTURE software version 2.3.4 with K number set from 2 to 10 and 10 iteration runs per each K (80). Plus, the Evanno method was applied to obtain the optimal ΔK estimation (81). STRUCTURE results were combined using the CLUMPP algorithm through *clumppExport*() function of the same package and visualized with the function *plotQ*().

A Discriminant Analysis of Principal Component (DAPC) was carried out using the R package *adegenet* (82). The Bayesian Information Criterion (BIC) and the function *xval.Dapc*() was used to evaluate the number of clusters to be selected and the correct number of principal components to retain.

RESULTS

The investigation of the *Pseudomonas savastanoi* pv. *savastanoi* genome sequence NCPPB 3335 by TRF program resulted in about 110 candidate-VNTR loci. The comparison of the candidate loci against the WGS of *P. savastanoi* strains available in NCBI showed that less than 30% of these loci varied in size. A total of 20 candidate-VNTR loci having a match percentage higher than 90% were selected and respective primer pairs were designed consequently. The TRs characteristics, sequence of the primers, the position of the amplicon along the reference genome of NCPPB 3335, and, where relevant, the putative gene functions are indicated in Table I. The 20 VNTRs loci were amplified on the 84 isolates for a total of 1680 reactions. The amplicons were separated via the QIAxcel multi-capillary electrophoresis system. Five of the 20 loci gave multiple products or a lack of reproducibility, which prompted their exclusion from further analysis.

TABLE III: SHOWS TRS CHARACTERISTICS, PUTATIVE GENE FUNCTION, PRIMERS, AND ANNEALING TEMPERATURE.

Name of TRs	TR unit length (bp)	TRs length	start-end	Gene function	Primers
TRsav1	9	158	995987 996340	*Hypothetical protein*	*FR:ATTTCCTGAGCGTCCTGTGT* *RV:ATTAAAGTGTTGATTCTTTC*
TRsav2	6	152	1659360 1659565	*Tellurium resistance protein Tera*	*FR: CTTGAACCGCTGGCAAAA* *RV:CCGAAACCGGCGCTGGATTT*
TRsav3	7	166	1988604 1990085	*Glycosyl transferase*	*FR: ATCTGGTGGGTTTCATGACC* *RV: CTCTGCATAATCGTATCCCT*
TRsav6	7	193	935627 938149	*Flavodoxin*	*FR:CTGGTGGATAACCGTCAGGT* *RV:AGCTGATCGAGCAAGGACGT*
TRsav7	8	179	2469543 2469753	*Transposase*	*FR: CTTGCCCATCTTGTCGACTT* *RV:GGCAACGCGCAGGCTCTGGA*
TRsav8	6	153	5324874 5325044	*DNA topoisomerase IV subunit A*	*FR: CGACCGTGAACAGAACTG* *RV: ACCGCCAGATCGGTCACATA*
TRsav10	8	169	2528494 2528694	*4-hydroxy-tetrahydrodipicolinate synthase*	*FR: GATGCTGGCTGAGGGTTG* *RV: GCGCAGATGCCCTGTCGATC*
TRsav11	6	175	2659925 2660123	*heme ABC transporter ATP-binding protein CcmA*	*FR: GCTCAATCTGTTGTGGGTTG* *RV:CTGCTGCCCGGCCGACAAGG*
TRsav13	24	130	1219712 1219889	*Hypothetical protein*	*FR: AAGATTTGGTACGCCAGCAG* *RV: GGTTTTACAGGTGGCCTCAC*
TRsav15	8	194	134818 135075	*unknown*	*FR: TTTGATGATCAGCCTTCGTG* *RV: GGCGCGATGATGGAGCGG*
TRsav16	6	133	259976 260396	*Hypothetical protein*	*FR: GTCAGATGCTTTTGGCTTGA* *RV:TGGAGATCCCTTTATTAATGAC*
TRsav17	6	187	495079 495294	*Phosphodiesterase*	*FR: ACCTATGGCGTGGTCGATAC* *RV: TTGTGTCGATCGTCATGATT*
TRsav18	9	157	996131 996323	*Hypothetical protein*	*FR: ATAACCAGTCCGCGAGCTAA* *RV: TGTTTCATGAGCTAGAGAAA*
TRsav19	7	173	1990097 1990297	*GDP-mannose dehydrogenase*	*FR: TAAGTCAGTTGCGAGCCTCA* *RV: GACTCCCGAAGGCAAGCGCG*
TRsav20	8	149	3346599 3346771	*unknown*	*FR: GGTTCACTGCATCAAACCAG* *RV: ATGGGCGAGGGTTGCTGTTC*

The total number of alleles in the final 14 VNTR loci resulted to be 129, ranging from only 2 alleles for TRsav 8 and TR sav17 loci, to a maximum of 30 alleles in the locus TRsav16 (data not shown).

The final data matrix containing the number of repeats was imported on R version 3.4.4 (73) and transformed into *genind* object using the R package *adegenet* 2.1.1 (74). In order to discard duplicated genotypes and to remove biases they could induce, the *clonecorrect* () function of the R package *poppr* was used to remove the clones and collapse the 84 individuals to 78 unique multilocus genotypes (MLGs). To evaluate if the number of loci was enough to describe the diversity between individuals, the *genotype_curve*() function of the poppr package was used with 1000 randomly sampling, to create the curve represented in (Fig. 1), where the number of n-1 loci show that the number of loci used can totally discriminate the individuals (76).

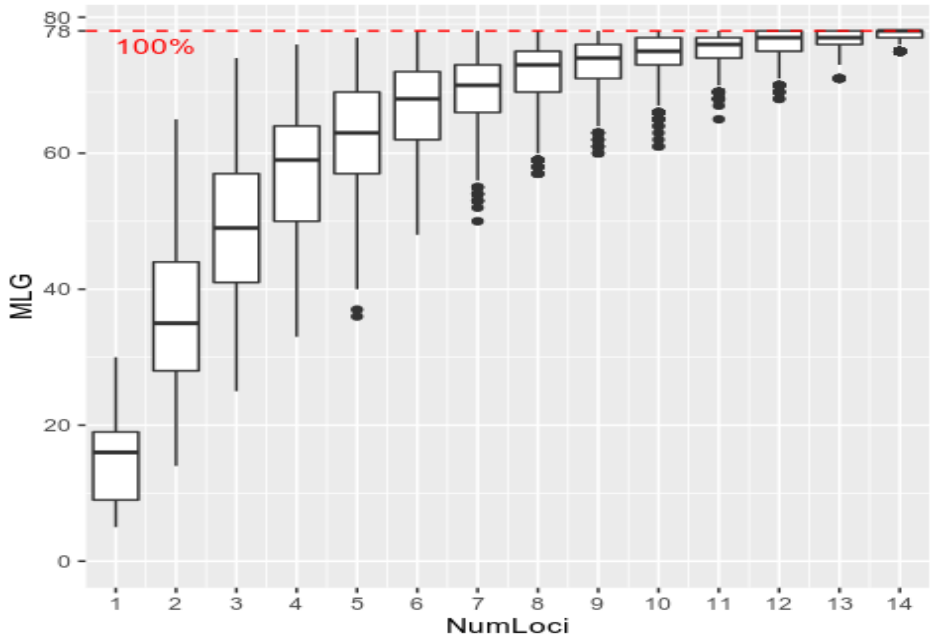


Figure 18: Genotype accumulation curve for 84 isolates of *Pseudomonas savastanoi*. The x-axis represents the number of loci that were randomly sampled 1000 times up to (n-1) loci while the y-axis represents the number of multilocus genotypes observed up to 78 MLGs in the dataset. The dashed lines at the plateau indicate that enough number of loci were used to discriminate between individuals.

The *poppr*() function from the same R package was used to obtain the diversity indexes reported in (TableIII). More specifically, besides the number of individual (N) and MLG per population, the expected MLG (eMLG) represents the number of MLG at the lowest sample size. Simpson's Index (λ), which is a measure of the probability that two randomly selected genotypes are different from each other, indicates that the *Pseudomonas savastanoi* pv. *savastanoi* clade, the more populated in the data set, resulted to be also the more diverse. Genotype Evenness (E.5) measures the genotype abundances within a population. In this case, *Pseudomonas savastanoi* pv. *fraxini* and *Pseudomonas savastanoi* pv. *phaseolicola* resulted to be the ones with the highest genotype diversity. Finally, Nei's diversity is the average genetic diversity per locus, also defined as expected heterozygosity (Hexp), which is highest for *Pseudomonas savastanoi* pv. *savastanoi*.

Table IV: Numbers of strains of each population, the number of multilocus genotypes, The expected number of MLG, Evenness, and diversity indices (Simpson's, Nei's gene diversity and) for each population[5]

Pseudomonas savastanoi **pathovar**	**N**	**MLG**	**eMLG**	**E.5**	**lambda**	**Hexp**
savastanoi	60	55	9.810206	0.905049	0.978889	0.640589
fraxini	11	11	10	1	0.909091	0.415931
nerii	10	9	9	0.9517	0.88	0.383704
phaseolicola	2	2	2	1	0.5	0.133333
retacarpa	1	1	1	NA	0	0
Total	84	78	9.88875	0.923356	0.985544	0.657298

Hierarchical clustering

The hierarchical clustering performed using Bruvo's distance and UPGMA as agglomerative algorithm generated the dendrogram showed in (Fig. 2), with bootstrap values higher than 80%. According to the intrinsic nature of the method, the number of clusters depends on the choice of the threshold distance and thus, in this first step of the analysis, the organization of the groups was observed as it is, keeping the bootstrap values as a key point of the validation.

In the dendrogram, the *Pseudomonas savastanoi* pv. *phaseolicola* strains appear each other very similar, but also the most distant from all the other strains with a 100% bootstrap support, representing a sort of outgroup for the analysis. About the other pathovars, all the Italian strains belonging to the pathovar *nerii* clustered together; the only exception in pathovar *nerii* clustering regards the strain ITM305, which was instead isolated in California, USA, back in 1981. The same happens for the strains of pathovar *fraxini*, which again form a clear cluster without exclusion. These clusters are loosely related to a single strain (SUPP3085) isolated from olive in Japan and the only strain belonging to the pathovar *retacarpa*.

Then, the numerous strains belonging to the pathovar *savastanoi* show a quite scattered distribution along the dendrograms, providing for a large genetic variability. Nonetheless, several clusters are recognizable, often in clear relationship with the plant host, when it is different from olive, or with their geographical origin. The cluster "*savast. C)*", formed by strains isolated from *Jasminum*, or the cluster "*savast.G)*" whose strains were obtained from

[5] https://www.rdocumentation.org/packages/poppr/versions/2.8.3/topics/poppr

Ligustrum, are examples of liaison to host plant species, whilst geographic affinities referable to the country of isolation are clearly evident for clusters "*savast. B)*"-Albania, "*savast. E)*"-Portugal, "*savast. F)*"-Morocco, "*savast. H)*"-Syria, and "*savast. C)*"-California. An independent position in the dendrogram is coherently occupied by single strains from Japan (SUPP3129 and SUPP3085) and from Tunisia (TN177). About Italian strains, it is of peculiar interest the grouping of a strain isolated from a single orchard in Viterbo, Italy, among which the strains isolated from a single tree furtherly group within an additional subcluster.

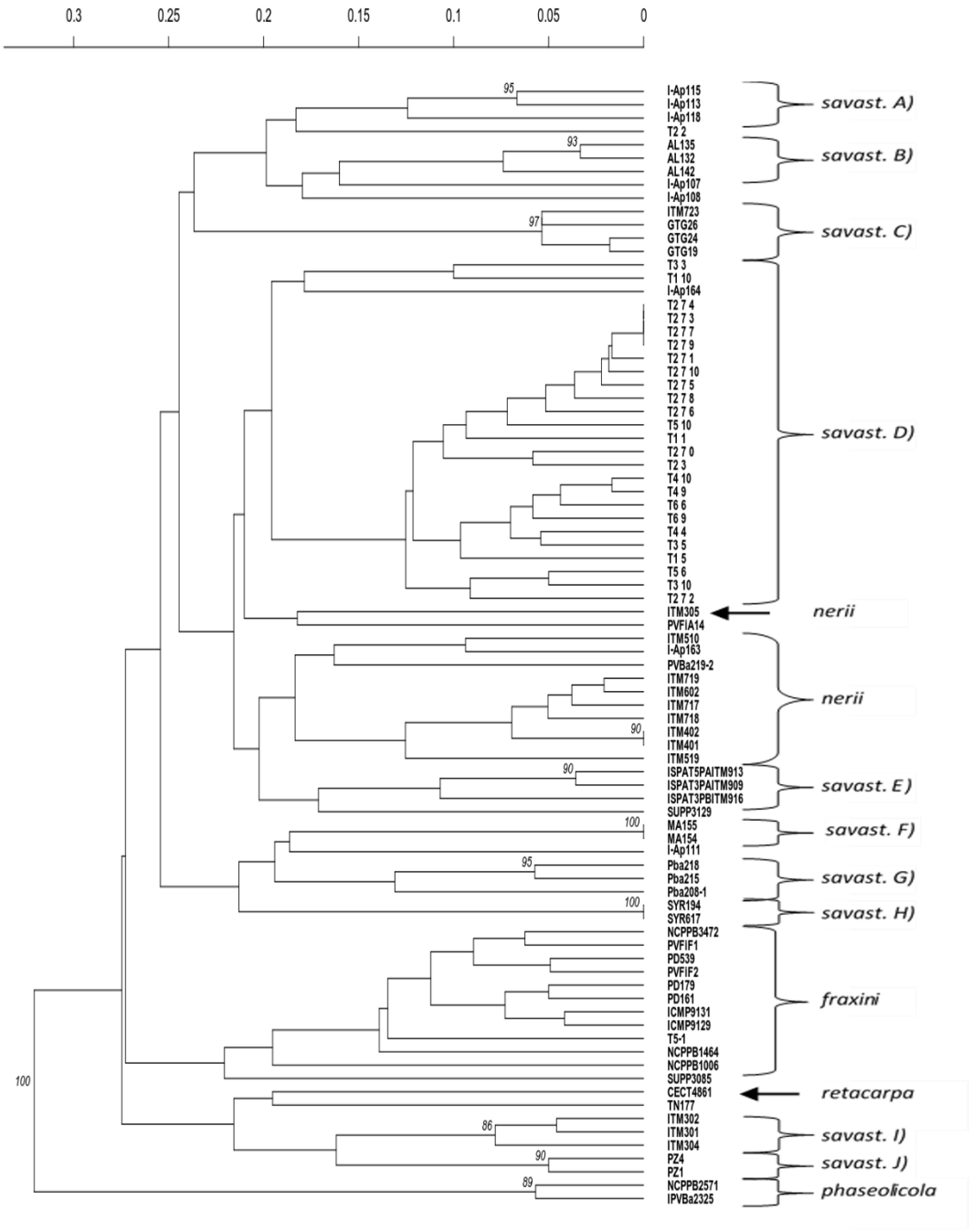


Figure 19: Dendrogram of 84 *Pseudomonas savastanoi* strains obtained based on Bruvo's distance and UPGMA as algorithm, the bootstrap values (for 1,000 replicates) are given at the nodes.

STRUCTURE analysis

The analysis of MLVA data using STRUCTURE software was run with K ranging from 2 to 12, and 10 independent iteration runs for each K. The results were imported on R environment and Evanno's method was applied using *evannoMethodStructure*() function in the R package pophelper (83) in order to obtain the optimal K value, which resulted to be 9, as shown in (Fig.3) (81).

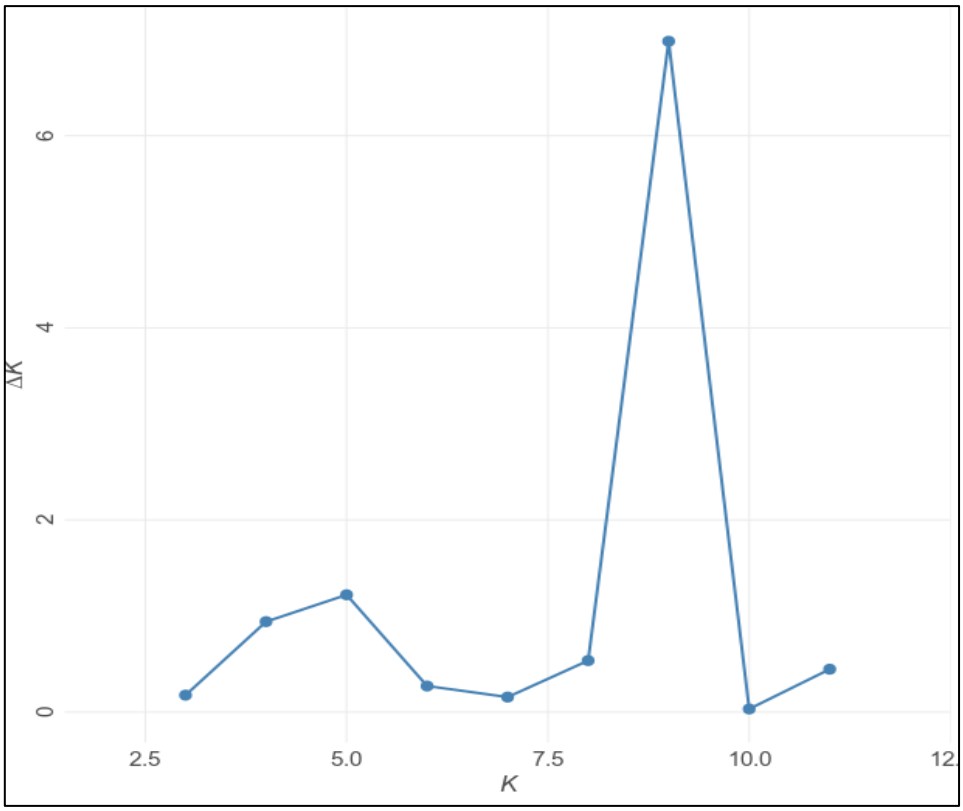


Figure 20: Evanno's graph showing the best ΔK when clustering 9 groups.

Thus, the 10 iterated results from K= 9, combined and merged using CLUMPP, are shown in

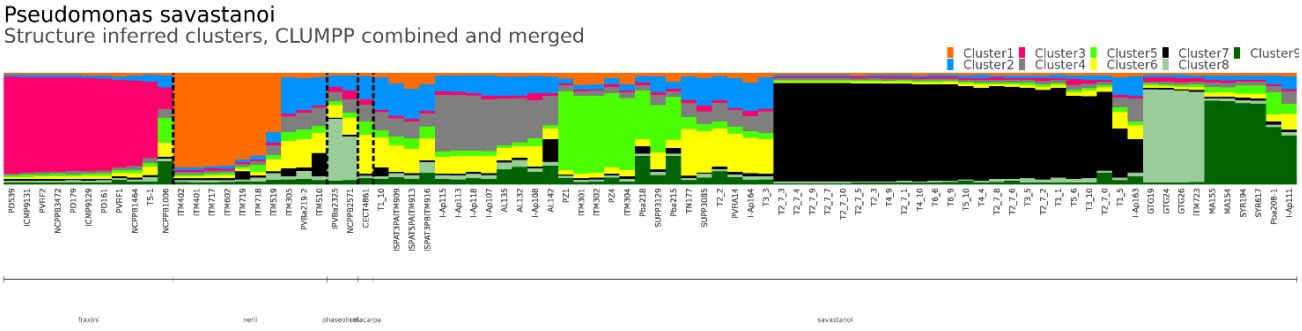


Figure 21: Structure clump output for K = 9 following analysis of 84 isolates of *Pseudomonas savastanoi*, each strain represented by a single vertical line indicating its membership in each of K independent clusters. Pathovars indicated along the picture are separated by vertical black dashed lines.

In this figure, the percentage of assignment to each of the 9 groups as suggested by K analysis, is represented in the proportion of respective colours in individual bars. Here the previous clustering is quite well confirmed, at least for the pathovar *fraxini* (Cluster 3 – dark pink), including the exception of strain NCPPB1006, and the pathovar *nerii* (Cluster1 – orange), in which, however, two strains (ITM510 and PVBa219-2) show a deep admixture of possible clusters, as well as for the strain ITM305, already reported out of the cluster in hierarchical clustering. It has to be noted that here, the two *phaseolicola* strains didn't show a clear fitting to a single cluster as in the previous analysis. The STRUCTURE grouping instead reveals a strong signal of independence for the *savastanoi* strains from Viterbo (Cluster 7 – black), from Greece (Cluster 8 – gray-green) and from Morocco and Syria (Cluster 9 – dark green), from USA and Pomarico, Italy (Cluster 5 – light green). All the others show a more or less significant grouping admixture.

Discriminant analysis of principal components (DAPC)

The DAPC analysis also supported the hypothesis of grouping the *Pseudomonas savastanoi* strains in 9 clusters as the best solution, according to the Bayesian Information Criterion, as showed in (Fig. 5).

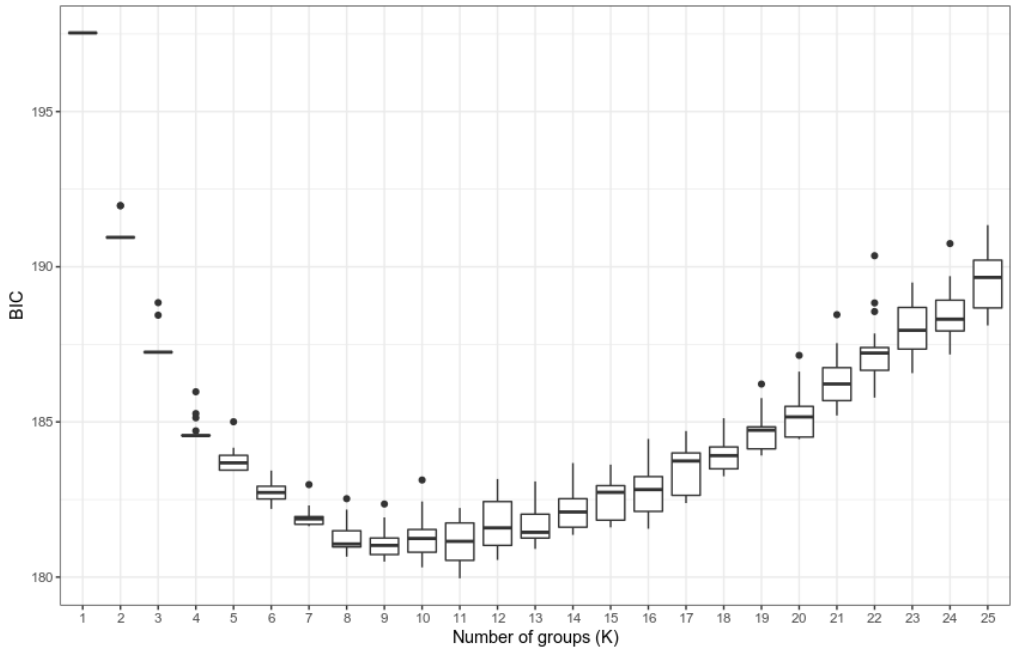


Figure 22: Curve of the BIC trend when K ranges from 2 to 2

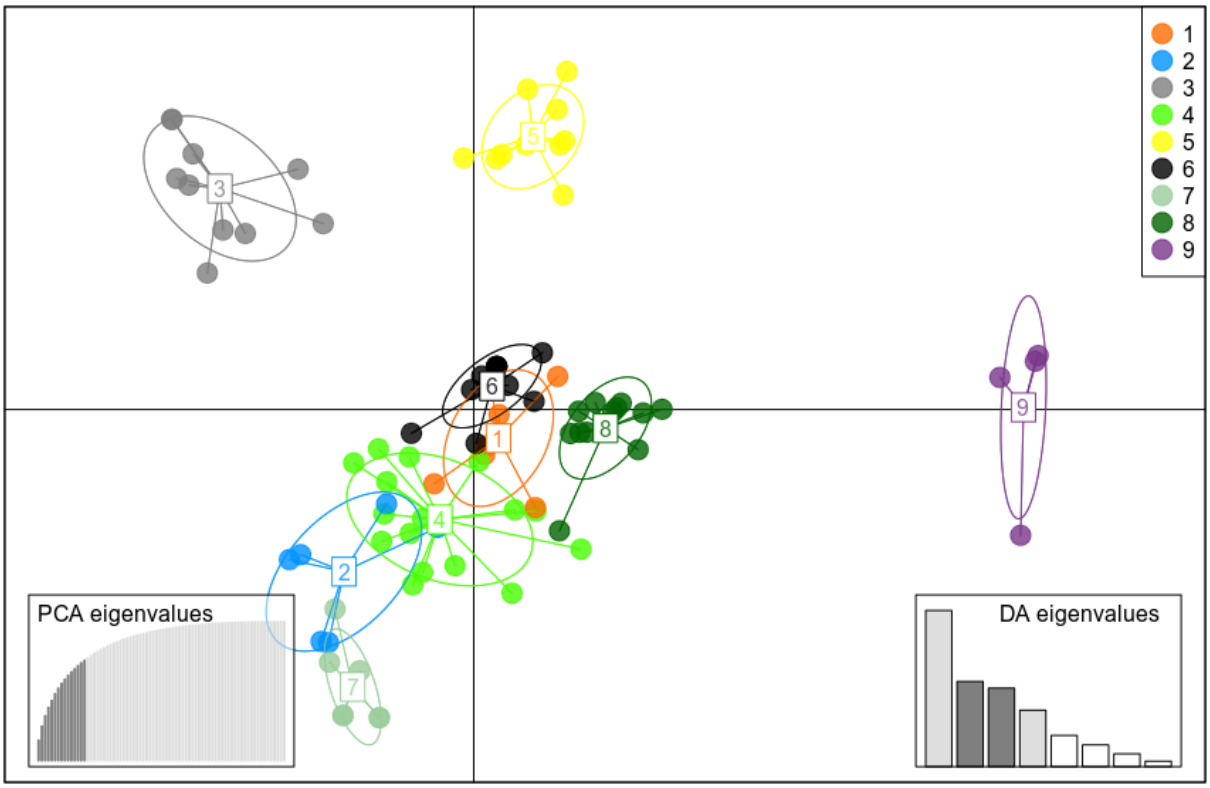


Figure 23: Scatterplot of genotypes in 9 groups according to DAPC. The individuals of the groups are distinguishable as in the present figure.

The resulting scatterplot (Fig.6) shows the distribution of the clusters of the groups on LD axes when 15 principal components were retained and second and third linear discriminants were selected. Once again, the general structure of groups as reported from previous analytical approaches is conserved. Indeed, the strains belonging to pathovar *fraxini* cluster separately from all the others (group 5 – yellow), with the exception of strain NCPPB 1006, as well as pathovar *nerii* (group 3 – grey) and *phaseolicola* (group 2 – light blue). Then, the numerous strains of the pathovar *savastanoi*, even if scattered in many groups, present the usual admixed composition, represented by the overlapping of different ellipses descriptive of the same subgroups identified before.

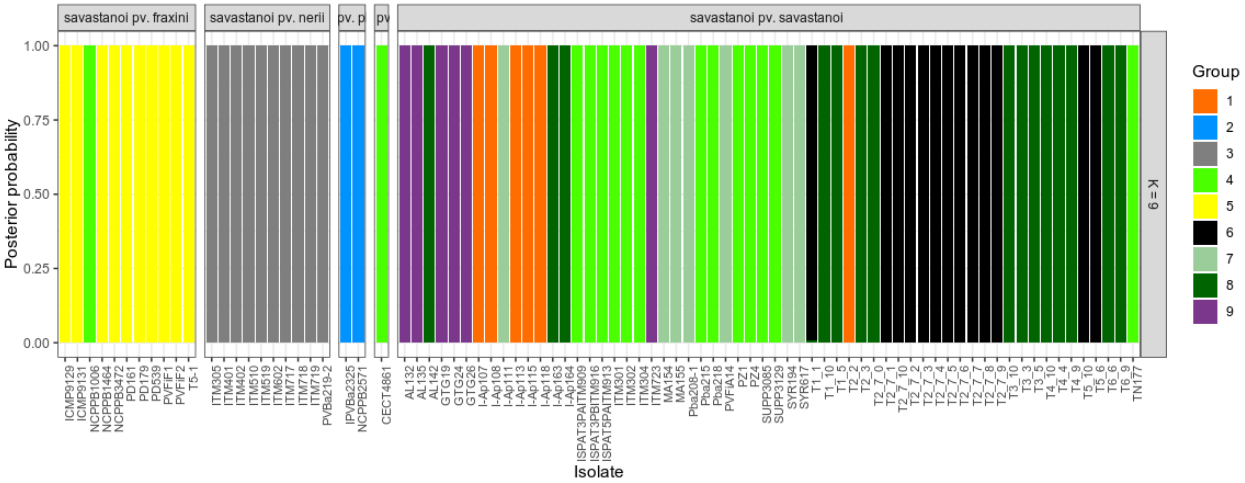


Figure 24: The barplot shows the posterior probability of the assignment of individuals in the 9 groups represented in different colours.

DISCUSSION

Molecular approaches are often used to define the evolutionary history, the host adaptation and the genetic diversity of a pathogen (25–27). Furthermore, they could reveal biological features influencing virulence, pathogenicity, and host specificity of the pathogen, which could, in turn, help us to define sustainable control measures (28,29,84). Of particular significance to understand the diffusion dynamics of plant bacterial pathogens are those molecular tools that allow discriminating between individuals of the same taxon according to host preferences, as well as time and geographic origin. As already mentioned in the Introduction, several methods have been already applied to ascertain the *Pseudomonas savastanoi* population structure.

Nonetheless, in this study, an MLVA assay, a method whose efficacy in resolving pathogenic bacterial epidemics have already been reported, was tested for the first time.

Hence, 14 TRs loci have been identified and an MLVA assay developed to investigate the genetic diversity and to infer the population structure within *Pseudomonas savastanoi* pathovars. The identified loci have proven to be able to provide enough information to disclose relationships among *Pseudomonas savastanoi* pathovars and strains.

As a first observation, the assay seems to be able to assign the samples to respective pathovars, as particularly evidenced for the strains belonging to phaseolicola, neri, and fraxini pathovars. Some exceptions, however, as well as the scattered distribution of strains belonging to pathovar *savastanoi*, are signs that the very high efficacy of this method in revealing differences made it only partially suitable to assess this type of information. Moreover, molecular methods for such type of distinction are already available to the scientific community (85).

When the data are instead analysed as a whole, the MLVA assay was able to provide information about population structure and to further separate strains in groups. Data elaboration was approached in three analytical methods and their congruence was investigated. Interestingly, the preliminary approach of both STRUCTURE, through Evanno's method and DAPC, through the BIC criterion, proposed that nine should be the number of clusters best fitting to the variability of the input data set. This grouping was quite consistent between the two methods, in particular with the groups including strains of the pathovars *phaseolicola*, *fraxini* and *nerii* that were clearly defined, as well as with the group of strains isolated from a single orchard in Viterbo. Then, about the other groups STRUCTURE gave quite confusing results as demonstrated by strong admixture in posterior assignment to the defined groups. DAPC, as expected from its ability to maximize between-groups variance, assigned individuals to groups much more sharply. It was the hierarchical clustering based on Bruvo's distance, producing more detailed results with a higher number of possible clusters, however, that suggested interesting connections of the clusters with the geographical origin of the included strains or with the respective host plant species.

In any case, the overall framework depicted in this study indicates a huge amount of variability in *Pseudomonas savastanoi* in general and more specifically, for the pathovar *savastanoi*. The first explanation for this confidently resides in the heterogeneity of the collection analysed, which included strains of worldwide origin, from 13 different countries, and isolated in a time span of more than 50 years. It is a further confirmation that MLVA, relying on genetic elements

with a particularly high rate of mutation, has its application to the depiction of more specific situations, as, typically, clonal outbreaks and epidemics.

This leads to another key point of this research, which concerns the analysis of strains isolated from knots collected contemporaneously from 10 olive plants in a small orchard (with a surface less than 1 hectare) in Viterbo, Italy, and from different knots of a single plant in the same orchard. In a situation like that, if the disease would have been related to an epidemic from an invasive bacterium, it would have been also logical to assume a substantial genetic homogeneity among its isolates, as in clonal outbreaks. In our case, instead, the level of genetic variability observed, even if less than between strains of diverse geographic origin, was still high, even from knots collected on a single plant. This can be in turn explained considering that *P. savastanoi* pv. *savastanoi* is not an epidemic but an endemic pathogen which is present in Italy for centuries. A long story of coevolution between the pathogen and its host plant has probably led to the differentiation of uncountable diverse genotypes so that even on a single plant, as for plant T2.7, multiple infections by different genotypes normally occurs.

Also, the cluster of strains from this plant is included few strains isolated from other plants in the orchard, as few strains from the single plant cluster with those from other plants. This admixture on a very confined situation could be explained by human-mediated transmission of the bacterial infection from a plant to another during pruning practices by infected tools.

In conclusion, for the first time, an MLVA assay was developed and applied to *Pseudomonas savastanoi*, aiming to investigate genetic relationships among a wide group of strains representatives of pathovars and with widespread geographic origin. The assay proved to be so much sensitive that was possible to distinguish isolates even from a single plant. Consequently, the most conceivable purpose of this method on an endemic pathogen as *Pseudomonas savastanoi* pv. *savastanoi* would be the analysis of very specific situations. Indeed, the method could be useful to investigate many biological facets such as how single genotypes are able to infect and colonize tissues and organs of the same plant, or to evaluate possible interactions between strains colonizing the same tissue, or even to reconstruct micro-evolutive trends during time in single plants or orchards, possibly in connection to changes in environmental conditions or in orchard management

Acknowledgments:

We are gratefully thanking the CIHEAM Mediterranean Agronomic Institute of Bari for the financial support within the HiSpro Project with the Iraqi ministry of agriculture to complete this work.

3.12 REFERENCES OF ARTICLE 2

1. Janse JD. *Pseudomonas syringae* subsp. *savastanoi* (ex Smith) subsp. nov., nom. rev., the bacterium causing excrescences on Oleaceae and Nerium oleander L. Int J Syst Evol Microbiol. 1982;32(2):166–9.

2. Iacobellis NS. Olive knot. Encycl Plant Pathol. 2001;2:713–5.

3. Quesada JM, Penyalver R, López MM. Epidemiology and control of plant diseases caused by phytopathogenic bacteria: the case of olive knot disease caused by *Pseudomonas savastanoi* pv. *savastanoi* . Plant Pathol. 2012;299–326.

4. Ramos C, Matas IM, Bardaji L, Aragón IM, Murillo J, Aragon IM, et al. *Pseudomonas savastanoi* pv. *savastanoi*: some like it knot. Mol Plant Pathol [Internet]. 2012;13(9):998–1009. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22805238

5. Buonaurio R, Moretti C, da Silva DP, Cortese C, Ramos C, Venturi V. The olive knot disease as a model to study the role of interspecies bacterial communities in plant disease. Front Plant Sci [Internet]. 2015;6:434. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26113855

6. Gardan L, Shafik H, Belouin S, Broch R, Grimont F, Grimont PAD. DNA relatedness among the pathovars of *Pseudomonas syringae* and description of *Pseudomonas tremae* sp. nov. and *Pseudomonas cannabina* sp. nov.(ex Sutic and Dowson 1959). Int J Syst Evol Microbiol. 1999;49(2):469–78.

7. Smith EF. Recent studies of the olive-tubercle organism. US Government Printing Office; 1908.

8. Gardan L, Bollet C, Ghorrah MA, Grimont F, Grimont PAD. DNA relatedness among the pathovar strains of *Pseudomonas syringae* subsp. *savastanoi* Janse (1982) and proposal of *Pseudomonas savastanoi* sp. nov. Int J Syst Evol Microbiol. 1992;42(4):606–12.

9. De Los Rios JEG. Retama sphaerocarpa (L.) Boiss., a new host of *Pseudomonas savastanoi* . Phytopathol Mediterr. 1999;54–60.

10. Bradbury JF. Guide to plant pathogenic bacteria. CAB international; 1986.

11. Iacobellis NS, Caponero A, Evidente A. Characterization of *Pseudomonas syringae* ssp. *savastanoi* strains isolated from ash. Plant Pathol. 1998;47(1):73–83.

12. Marchi G, Cinelli T, Surico G. 6. A review on *Pseudomonas savastanoi* genetic traits involved in disease development and in symptom induction. 2014;

13. Moretti C, Vinatzer BA, Onofri A, Valentini F, Buonaurio R. Genetic and phenotypic diversity of Mediterranean populations of the olive knot pathogen, *Pseudomonas savastanoi* pv. *savastanoi* . Plant Pathol. 2017;66(4):595–605.

14. Janse JD. The bacterial disease of ash (Fraxinus excelsior), caused by *Pseudomonas syringae* subsp. *savastanoi* pv. Fraxini II. Etiology and taxonomic considerations. For Pathol. 1981;11(7):425–38.

15. Phytopathology WB-, 1926 undefined. A new bacterial disease of the bean. cabdirect.org [Internet]. [cited 2019 Nov 5]; Available from: https://www.cabdirect.org/cabdirect/abstract/19271100629

16. Völksch B, Weingart H. Comparison of ethylene-producing *Pseudomonas syringae* strains isolated from kudzu (Pueraria lobata) with *Pseudomonas syringae* pv. phaseolicola and *Pseudomonas syringae* pv. glycinea. Eur J Plant Pathol. 1997;103(9):795–802.

17. Dworkin M. The Prokaryotes: Vol. 6: Proteobacteria: Gamma Subclass. Springer Science & Business Media; 2006.

18. Young JM, Dye DW, Bradbury JF, Panagopoulos CG, Robbs CF. A proposed nomenclature and classification for plant pathogenic bacteria. New Zeal J Agric Res. 1978;21(1):153–77.

19. Mugnai L, Giovannetti L, Ventura S, Surico G. The grouping of strains of *Pseudomonas syringae* subsp. *savastanoi* by DNA restriction fingerprinting. J Phytopathol. 1994;142(4):209–18.

20. Bennasar A, Mulet M, Lalucat J, García-Valdés E. PseudoMLSA: a database for multigenic sequence analysis of Pseudomonas species [Internet]. Vol. 10, BMC Microbiology. 2010. Available from: http://www.biomedcentral.com/1471-2180/10/118

21. Grünwald NJ, Goss EM. Evolution and population genetics of exotic and re-emerging pathogens: novel tools and approaches. Annu Rev Phytopathol. 2011;49:249–67.

22. Grunwald NJ, Everhart SE, Knaus BJ, Kamvar ZN. Best Practices for Population Genetic Analyses. Phytopathology [Internet]. 2017;107(9):1000–10. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28513284

23. Hartl DL, Clark AG. Molecular population genetics. Princ Popul Genet. 2007;317(ok):384.

24. Cesbron S, Pothier J, Gironde S, Jacques M-A, Manceau C. Development of multilocus variable-number tandem repeat analysis (MLVA) for *Xanthomonas arboricola* pathovars. J Microbiol Methods. 2014;100:84–90.

25. van Belkum A, Struelens M, de Visser A, Verbrugh H, Tibayrenc M. Role of genomic typing in taxonomy, evolutionary genetics, and microbial epidemiology. Clin Microbiol Rev [Internet]. 2001;14(3):547–60. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11432813

26. Scortichini M, Rossi MP, Salerno M. Relationship of genetic structure of *Pseudomonas savastanoi* pv. *savastanoi* populations from Italian olive trees and patterns of host genetic diversity. Plant Pathol. 2004;53(4):491–7.

27. Wolska K, Szweda P. Genotyping techniques for determining the diversity of microorganisms. In: Genetic diversity in microorganisms. IntechOpen; 2012.

28. Tibayrenc M. Bridging the gap between molecular epidemiologists and evolutionists. Trends Microbiol [Internet]. 2005;13(12):575–80. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16214342

29. Vinatzer BA, Monteil CL, Clarke CR. Harnessing population genomics to understand how bacterial pathogens emerge, adapt to crop hosts, and disseminate. Annu Rev Phytopathol [Internet]. 2014;52:19–43. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24820995

30. Sisto A, Morea M, Baruzzi F, Palumbo G. Differentiation of *Pseudomonas syringae* subsp. *savastanoi* strains isolated from various host plants by restriction fragment length polymorphism. Phytopathol Mediterr. 2002;41(1):63–71.

31. Krid S, Rhouma A, Quesada JM, Penyalver R, Gargouri A. Delineation of *Pseudomonas savastanoi* pv. *savastanoi* strains isolated in Tunisia by random-amplified polymorphic DNA analysis. J Appl Microbiol [Internet]. 2009;106(3):886–94. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19191961

32. Sisto A, Cipriani MG, Tegli S, Cerboneschi M, Stea G, Santilli E. Genetic characterization by fluorescent AFLP of *Pseudomonas savastanoi* pv. *savastanoi* strains isolated from different host species. Plant Pathol. 2007;56(3):366–72.

33. López MM, Quesada JM, Penyalver R, Biosca EG, Caruso P, Bertolini E, et al. Current technologies for Pseudomonas spp. and Ralstonia solanacearum detection and molecular typing. In: *Pseudomonas syringae* pathovars and related pathogens–identification, epidemiology and genomics. Springer; 2008. p. 3–19.

34. Cangelosi GA, Freeman RJ, Lewis KN, Livingston-Rosanoff D, Shah KS, Milan SJ, et al. Evaluation of a high-throughput repetitive-sequence-based PCR system for DNA fingerprinting of *Mycobacterium tuberculosis* and Mycobacterium avium complex strains. J Clin Microbiol [Internet]. 2004;42(6):2685–93. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15184453

35. Zhao S, Poulin L, Rodriguez RL, Serna NF, Liu S-YY, Wonni I, et al. Development of a variable number of tandem repeats typing scheme for the bacterial rice pathogen *Xanthomonas oryzae* pv. *oryzicola*. Phytopathology [Internet]. 2012;102(10):948–56. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22957820

36. Cinelli T, Marchi G, Cimmino A, Marongiu R, Evidente A, Fiori M. Heterogeneity of *Pseudomonas savastanoi* populations infecting *Myrtus communis* in Sardinia (I taly). Plant Pathol. 2014;63(2):277–89.

37. Achtman M. Evolution, population structure, and phylogeography of genetically monomorphic bacterial pathogens. Annu Rev Microbiol [Internet]. 2008;62:53–70. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18785837

38. Grissa I, Bouchon P, Pourcel C, Vergnaud G. On-line resources for bacterial micro-evolution studies using MLVA or CRISPR typing. Biochimie [Internet]. 2008;90(4):660–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17822824

39. Perez-Martinez I, Zhao Y, Murillo J, Sundin GW, Ramos C, Pérez-Martínez I, et al. Global genomic analysis of *Pseudomonas savastanoi* pv. *savastanoi* plasmids. J Bacteriol [Internet]. 2008;190(2):625–35. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17993520

40. Rodriguez-Palenzuela P, Matas IM, Murillo J, Lopez-Solanilla E, Bardaji L, Perez-Martinez I, et al. Annotation and overview of the *Pseudomonas savastanoi* pv. *savastanoi* NCPPB 3335 draft genome reveals the virulence gene complement of a tumour-inducing pathogen of woody hosts. Env Microbiol [Internet]. 2010;12(6):1604–20. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20370821

41. Moretti C, Cortese C, Passos da Silva D, Venturi V, Ramos C, Firrao G, et al. Draft Genome Sequence of *Pseudomonas savastanoi* pv. *savastanoi* Strain DAPP-PG 722, Isolated in Italy from an Olive Plant Affected by Knot Disease. Genome Announc [Internet]. 2014;2 (5). Available from http://www.ncbi.nlm.nih.gov/pubmed/25278521

42. Bartoli C, Lamichhane JR, Berge O, Varvaro L, Morris CE. Mutability in *Pseudomonas viridiflava* as a programmed balance between antibiotic resistance and pathogenicity. Mol Plant Pathol. 2015;16(8):860–9.

43. Poulin L, Grygiel P, Magne M, Gagnevin L, Rodriguez RL, Forero Serna N, et al. New multilocus variable-number tandem-repeat analysis tool for surveillance and local epidemiology of bacterial leaf blight and bacterial leaf streak of rice caused by *Xanthomonas oryzae*. Appl Env Microbiol [Internet]. 2015;81(2):688–98. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25398857

44. Davis Ii EW, Weisberg AJ, Tabima JF, Grunwald NJ, Chang JH. Gall-ID: tools for genotyping gall-causing phytopathogenic bacteria. PeerJ [Internet]. 2016;4:e2222. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27547538

45. Van Belkum A, Scherer S, Van Alphen L, Verbrugh H. Short-sequence DNA repeats in prokaryotic genomes. Microbiol Mol Biol Rev. 1998;62(2):275–93.

46. Lindstedt BA. Multiple-locus variable number tandem repeats analysis for genetic fingerprinting of pathogenic bacteria. Electrophoresis [Internet]. 2005;26(13):2567–82. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15937984

47. Gemayel R, Vinces MD, Legendre M, Verstrepen KJ. Variable tandem repeats accelerate evolution of coding and regulatory sequences. Annu Rev Genet [Internet]. 2010;44:445–77. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20809801

48. Vieira MLC, Santini L, Diniz AL, Munhoz Cde F, Munhoz C de F. Microsatellite markers: what they mean and why they are so useful. Genet Mol Biol [Internet]. 2016;39(3):312–28. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27561112

49. Saunders NJ, Jeffries AC, Peden JF, Hood DW, Tettelin H, Rappuoli R, et al. Repeat-associated phase variable genes in the complete genome sequence of *Neisseria meningitidis* strain MC58. Mol Microbiol. 2000;37(1):207–15.

50. Rocha EPC. DNA repeats lead to the accelerated loss of gene order in bacteria. Trends Genet. 2003;19(11):600–3.

51. Oliveira EJ, Pádua JG, Zucchi MI, Vencovsky R, Vieira MLC. Origin, evolution and genome distribution of microsatellites. Genet Mol Biol. 2006;29(2):294–307.

52. Reinar WB, Knutsen JB, Jentoft S, Tørresen OK, Butenko MA, Jakobsen KS. Evolutionary adaptability linked to length variation in short genic tandem repeats . 2018;

53. Treangen TJ, Abraham A-LL, Touchon M, Rocha EPC. Genesis, effects and fates of repeats in prokaryotic genomes. FEMS Microbiol Rev [Internet]. 2009;33(3):539–71. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19396957

54. Li W, Raoult D, Fournier P-EE. Bacterial strain typing in the genomic era. FEMS Microbiol Rev [Internet]. 2009;33(5):892–916. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19453749

55. Ee CS, Ismail A, Singh KKB. Multiple-locus variable-number tandem repeat analysis (MLVA) for bacterial typing. 2016;

56. Van Belkum A, Scherer S, Van Leeuwen W, Willemse D, Van Alphen L, Verbrugh H. Variable number of tandem repeats in clinical strains of *Haemophilus influenzae*. Infect Immun. 1997;65(12):5017–27.

57. Noller AC, McEllistrem MC, Pacheco AGF, Boxrud DJ, Harrison LH. Multilocus variable-number tandem repeat analysis distinguishes outbreak and sporadic *Escherichia coli* O157: H7 isolates. J Clin Microbiol [Internet]. 2003;41(12):5389–97. Available from: http://www.ncbi.nlm.nih.gov/pubmed/14662916

58. Marsh JW, O'Leary MM, Shutt KA, Pasculle AW, Johnson S, Gerding DN, et al. Multilocus variable-number tandem-repeat analysis for investigation of Clostridium difficile transmission in hospitals. J Clin Microbiol [Internet]. 2006;44(7):2558–66. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16825380

59. Le Flèche P, Jacques I, Grayon M, Al Dahouk S, Bouchon P, Denoeud F, et al. Evaluation and selection of tandem repeat loci for a Brucella MLVA typing assay. BMC Microbiol [Internet]. 2006; 6(1): 9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16469109

60. Della Coletta-Filho H, Takita MA, de Souza AA, Aguilar-Vildoso CI, Machado MA, Coletta-Filho HD, et al. Differentiation of strains of *Xylella fastidiosa* by a variable number of tandem repeat analysis. Appl Env Microbiol [Internet]. 2001;67(9):4091–5. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11526010

61. Verniere C, Bui Thi Ngoc L, Jarne P, Ravigne V, Guerin F, Gagnevin L, et al. Highly polymorphic markers reveal the establishment of an invasive lineage of the citrus bacterial pathogen *Xanthomonas citri* pv. *citri* in its area of origin. Env Microbiol [Internet]. 2014;16(7):2226–37. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24373118

62. Pruvost O, Magne M, Boyer K, Leduc A, Tourterel C, Drevet C, et al. A MLVA genotyping scheme for global surveillance of the citrus pathogen *Xanthomonas citri* pv. citri suggests a worldwide geographical expansion of a single genetic lineage. PLoS One [Internet]. 2014;9(6):e98129. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24897119

63. Bergsma-Vlami M, Martin W, Koenraadt H, Teunissen H, Pothier JF, Duffy B, et al. Molecular typing of Dutch isolates of *Xanthomonas arboricola* pv. *pruni* isolated from ornamental cherry laurel. J Plant Pathol. 2012;94(1sup):1-29-S1. 35.

64. Gironde S, Manceau C. Housekeeping gene sequencing and multilocus variable-number tandem-repeat analysis to identify subpopulations within *Pseudomonas syringae* pv. *maculicola* and *Pseudomonas syringae* pv. *tomato* that correlate with host specificity. Appl Env Microbiol [Internet]. 2012;78(9):3266–79. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22389364

65. N'Guessan C A, Brisse S, Le Roux-Nio A-CC, Poussier S, Kone D, Wicker E, et al. Development of variable number of tandem repeats typing schemes for *Ralstonia solanacearum*, the agent of bacterial wilt, banana Moko disease and potato brown rot. J Microbiol Methods [Internet]. 2013;92(3):366–74. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23376194

66. Zaluga J, Stragier P, Van Vaerenbergh J, Maes M, De Vos P. Multilocus Variable-Number-Tandem-Repeats Analysis (MLVA) distinguishes a clonal complex of *Clavibacter michiganensis* subsp. *michiganensis* strains isolated from recent outbreaks of bacterial wilt and canker in Belgium. BMC Microbiol. 2013;13(1).

67. Buhlmann A, Dreo T, Rezzonico F, Pothier JF, Smits THM, Ravnikar M, et al. Phylogeography and population structure of the biologically invasive phytopathogen *Erwinia amylovora* inferred using minisatellites. Environ Microbiol [Internet]. 2014;16(7):2112–25. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24112873

68. Ciarroni S, Gallipoli L, Taratufolo MC, Butler MI, Poulter RTM, Pourcel C, et al. Development of a Multiple Loci Variable Number of Tandem Repeats Analysis (MLVA) to Unravel the Intra-Pathovar Structure of *Pseudomonas syringae* pv. *actinidiae* Populations Worldwide. PLoS One [Internet]. 2015;10(8):e0135310. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26262683

69. Guinard J, Latreille A, Guerin F, Poussier S, Wicker E. New Multilocus Variable-Number Tandem-Repeat Analysis (MLVA) Scheme for Fine-Scale Monitoring and Microevolution-Related Study of *Ralstonia pseudosolanacearum* Phylotype I Populations. Appl Env Microbiol [Internet]. 2017;83(5). Available from: http://www.ncbi.nlm.nih.gov/pubmed/28003195

70. Penyalver R, García A, Ferrer A, Bertolini E, López MM. Detection of *Pseudomonas savastanoi* pv. *savastanoi* in olive plants by enrichment and PCR. Appl Environ Microbiol. 2000;66(6):2673–7.

71. Benson G. Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res. 1999;27(2):573–80.

72. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990;215(3):403–10.

73. Team RC. R: A language and environment for statistical computing; 2015. 2018.

74. Jombart T. adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics. 2008;24(11):1403–5.

75. Kamvar ZN, Tabima JF, Grünwald NJ. Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. PeerJ. 2014;2:e281.

76. Kamvar ZN, Brooks JC, Grünwald NJ. Novel R tools for analysis of genome-wide population genetic data with emphasis on clonality. Front Genet. 2015;6:208.

77. Dray S, Dufour A-B. The ade4 package: implementing the duality diagram for ecologists. J Stat Softw. 2007;22(4):1–20.

78. Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, et al. Ordination methods, diversity analysis and other functions for community and vegetation ecologists. vegan: Community Ecology Package. 2019.

79. Kassambara A, Mundt F. Factoextra: Extract and visualize the results of multivariate data analyses. 2017. R Packag version. 2016;1(3).

80. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. Genetics. 2000;155(2):945–59.

81. Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol. 2005;14(8):2611–20.

82. Jombart T, Devillard S, Balloux F. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genet. 2010;11(1):94.

83. Francis RM. pophelper: an R package and web app to analyse and visualize population structure. Mol Ecol Resour. 2017;17(1):27–32.

84. Tibayrenc M. Microbial molecular epidemiology: an overview. Methods in molecular biology (Clifton, NJ). 2009;551:1–12.

85. Caugant DA. Molecular epidemiology of microorganisms. Springer; 2009.

3.13 **Supplementary materials of article 2**

Table S1. *Pseudomonas savastanoi* strains used in this study

NAME	POPULATION	HOST	COUNTRY	CITY	YEAR
ITM301	*Pseudomonas savastanoi* pv. *savastanoi*	*Olea europaea*	USA	California	1981
ITM302	*Pseudomonas savastanoi* pv. *savastanoi*	*Olea europaea*	USA	California	1981
ITM304	*Pseudomonas savastanoi* pv. *savastanoi*	*Olea europaea*	USA	California	1981
ISPAT3PAITM909	*Pseudomonas savastanoi* pv. *savastanoi*	*Olea europaea*	Portugal	Evora	2001
ISPAT5PAITM913	*Pseudomonas savastanoi* pv. *savastanoi*	*Olea europaea*	Portugal	Evora	2001
ISPAT3PBITM916	*Pseudomonas savastanoi* pv. *savastanoi*	*Olea europaea*	Portugal	Evora	2001
AL132	*Pseudomonas savastanoi* pv. *savastanoi*	*Olea europaea*	Albania	Durazzo	2006
AL135	*Pseudomonas savastanoi* pv. *savastanoi*	*Olea europaea*	Albania	Durazzo	2006
AL142	*Pseudomonas savastanoi* pv. *savastanoi*	*Olea europaea*	Albania	Fier	2006
SYR617	*Pseudomonas savastanoi* pv. *savastanoi*	*Olea europaea*	Syria	n.d.	1984
SYR194	*Pseudomonas savastanoi* pv. *savastanoi*	*Olea europaea*	Syria	Latakia	2006
MA154	*Pseudomonas savastanoi* pv. *savastanoi*	*Olea europaea*	Morocco	Haouz (Marrakech Rmat)	2006
MA155	*Pseudomonas savastanoi* pv. *savastanoi*	*Olea europaea*	Morocco	Haouz (Marrakech Rmat)	2006
TN177	*Pseudomonas savastanoi* pv. *savastanoi*	*Olea europaea*	Tunisia	Tunisi	2006
PZ1	*Pseudomonas savastanoi* pv. *savastanoi*	*Olea europaea*	Italy	Pomarico,MT	1995
PZ4	*Pseudomonas savastanoi* pv. *savastanoi*	*Olea europaea*	Italy	Pomarico,MT	1995
I-AP107	*Pseudomonas savastanoi* pv. *savastanoi*	*Olea europaea*	Italy	Puglia (Locorotondo, BA)	2005
I-AP108	*Pseudomonas savastanoi* pv. *savastanoi*	*Olea europaea*	Italy	Puglia (Locorotondo, BA)	2005
I-AP111	*Pseudomonas savastanoi* pv. *savastanoi*	*Olea europaea*	Italy	Puglia (Tuturano, BR)	2005
I-AP113	*Pseudomonas savastanoi* pv. *savastanoi*	*Olea europaea*	Italy	Puglia (Monte S. Angelo, FG)	2005
I-AP115	*Pseudomonas savastanoi* pv. *savastanoi*	*Olea europaea*	Italy	Puglia (S. Giovanni Rotondo, FG)	2005
I-AP118	*Pseudomonas savastanoi* pv. *savastanoi*	*Olea europaea*	Italy	Puglia (S. Giovanni Rotondo, FG)	2005
I-AP163	*Pseudomonas savastanoi* pv. *savastanoi*	*Olea europaea*	Italy	Puglia (Valenzano, BA)	2006
I-AP164	*Pseudomonas savastanoi* pv. *savastanoi*	*Olea europaea*	Italy	Puglia (Valenzano, BA)	2006

PVFIA14	*Pseudomonas savastanoi* pv. *savastanoi*		*Olea europaea*	Italy	Italia (Centro)	n.d.
SUPP3085	*Pseudomonas savastanoi* pv. *savastanoi*		*Olea europaea*	Japan	n.d.	n.d.
SUPP3129	*Pseudomonas savastanoi* pv. *savastanoi*		*Olea europaea*	Japan	n.d.	n.d.
PVFIF1	*Pseudomonas savastanoi* pv. *fraxini*		*Fraxinus sp.*	Italy	n.d.	n.d.
PVFIF2	*Pseudomonas savastanoi* pv. *fraxini*		*Fraxinus sp.*	Italy	n.d.	n.d.
PD539	*Pseudomonas savastanoi* pv. *fraxini*		*Fraxinus sp.*	France	n.d.	n.d.
T5-1	*Pseudomonas savastanoi* pv. *fraxini*		*Fraxinus sp.*	France	n.d.	n.d.
NCPPB3472	*Pseudomonas savastanoi* pv. *fraxini*		*Fraxinus sp.*	France	Puy-de-DÃ´me.	1975
ICMP9129	*Pseudomonas savastanoi* pv. *fraxini*		*Fraxinus sp.*	Netherland	Walcheren	1982
ICMP9131	*Pseudomonas savastanoi* pv. *fraxini*		*Fraxinus sp.*	Netherland	Houten	1983
PD161	*Pseudomonas savastanoi* pv. *fraxini*		*Fraxinus sp.*	Netherland	n.d.	1979
PD179	*Pseudomonas savastanoi* pv. *fraxini*		*Fraxinus sp.*	Netherland	n.d.	1979
NCPPB1464	*Pseudomonas savastanoi* pv. *fraxini*		*Fraxinus sp.*	UK	n.d.	1963
NCPPB1006	*Pseudomonas savastanoi* pv. *fraxini*		*Fraxinus sp.*	UK	n.d.	1961
ITM510	*Pseudomonas savastanoi* pv. *fraxini*		*Nerium oleander*	Italy	Sibari	1983
PVBA219-2	*Pseudomonas savastanoi* pv. *fraxini*		*Nerium oleander*	Italy	n.d.	
ITM401	*Pseudomonas savastanoi* pv. *fraxini*		*Nerium oleander*	Italy	TERMOLI,CB	1981
ITM402	*Pseudomonas savastanoi* pv. *fraxini*		*Nerium oleander*	Italy	TERMOLI,CB	1981
ITM519	*Pseudomonas savastanoi* pv. *fraxini*		*Nerium oleander*	Italy	Polignano	1982
ITM602	*Pseudomonas savastanoi* pv. *fraxini*		*Nerium oleander*	Italy	SIRMIONE	1985
ITM717	*Pseudomonas savastanoi* pv. *fraxini*		*Nerium oleander*	Italy	Bitonto, BA,A14	1985
ITM718	*Pseudomonas savastanoi* pv. *fraxini*		*Nerium oleander*	Italy	Bitonto, BA,A15	1985
ITM719	*Pseudomonas savastanoi* pv. *fraxini*		*Nerium oleander*	Italy	Bitonto, BA,A16	1985
ITM305	*Pseudomonas savastanoi* pv. *fraxini*		*Nerium oleander*	USA	CA	1981
GTG19	*Pseudomonas savastanoi* pv. *savastanoi*		*Jasminum sp.*	Greece	n.d.	
GTG24	*Pseudomonas savastanoi* pv. *savastanoi*		*Jasminum sp.*	Greece	n.d.	
GTG26	*Pseudomonas savastanoi* pv. *savastanoi*		*Jasminum sp.*	Greece	n.d.	
ITM723	*Pseudomonas savastanoi* pv. *savastanoi*		*Jasminum sp.*	Greece	n.d.	1984

PBA208-1	*Pseudomonas savastanoi savastanoi*	pv.	*Ligustrum sp.*	Italy	BARI	197?
PBA215	*Pseudomonas savastanoi savastanoi*	pv.	*Ligustrum sp.*	Italy	BARI	197?
PBA218	*Pseudomonas savastanoi savastanoi*	pv.	*Ligustrum sp.*	Italy	BARI	197?
IPVBA2325	*Pseudomonas savastanoi phaseolicola*	pv.	*Phaseolus sp.*		n.d.	
NCPPB2571	*Pseudomonas savastanoi phaseolicola*	pv.	*Phaseolus sp.*	UK	n.d.	1966
CECT4861	*Pseudomonas savastanoi retacarpa*	pv.	*Retama sphaerocarpa*	Spain	Madrid	1999
T1_1	*Pseudomonas savastanoi savastanoi*	pv.	*Olea europaea*	Italy	Viterbo	2019
T1_10	*Pseudomonas savastanoi savastanoi*	pv.	*Olea europaea*	Italy	Viterbo	2019
T2_3	*Pseudomonas savastanoi savastanoi*	pv.	*Olea europaea*	Italy	Viterbo	2019
T3_3	*Pseudomonas savastanoi savastanoi*	pv.	*Olea europaea*	Italy	Viterbo	2019
T3_5	*Pseudomonas savastanoi savastanoi*	pv.	*Olea europaea*	Italy	Viterbo	2019
T3_10	*Pseudomonas savastanoi savastanoi*	pv.	*Olea europaea*	Italy	Viterbo	2019
T4_4	*Pseudomonas savastanoi savastanoi*	pv.	*Olea europaea*	Italy	Viterbo	2019
T4_9	*Pseudomonas savastanoi savastanoi*	pv.	*Olea europaea*	Italy	Viterbo	2019
T4_10	*Pseudomonas savastanoi savastanoi*	pv.	*Olea europaea*	Italy	Viterbo	2019
T5_6	*Pseudomonas savastanoi savastanoi*	pv.	*Olea europaea*	Italy	Viterbo	2019
T5_10	*Pseudomonas savastanoi savastanoi*	pv.	*Olea europaea*	Italy	Viterbo	2019
T6_6	*Pseudomonas savastanoi savastanoi*	pv.	*Olea europaea*	Italy	Viterbo	2019
T6_9	*Pseudomonas savastanoi savastanoi*	pv.	*Olea europaea*	Italy	Viterbo	2019
T1_5	*Pseudomonas savastanoi savastanoi*	pv.	*Olea europaea*	Italy	Viterbo	2019
T2_2	*Pseudomonas savastanoi savastanoi*	pv.	*Olea europaea*	Italy	Viterbo	2019
T2_7_0	*Pseudomonas savastanoi savastanoi*	pv.	*Olea europaea*	Italy	Viterbo	2019
T2_7_1	*Pseudomonas savastanoi savastanoi*	pv.	*Olea europaea*	Italy	Viterbo	2019
T2_7_2	*Pseudomonas savastanoi savastanoi*	pv.	*Olea europaea*	Italy	Viterbo	2019
T2_7_3	*Pseudomonas savastanoi savastanoi*	pv.	*Olea europaea*	Italy	Viterbo	2019
T2_7_4	*Pseudomonas savastanoi savastanoi*	pv.	*Olea europaea*	Italy	Viterbo	2019
T2_7_5	*Pseudomonas savastanoi savastanoi*	pv.	*Olea europaea*	Italy	Viterbo	2019
T2_7_6	*Pseudomonas savastanoi savastanoi*	pv.	*Olea europaea*	Italy	Viterbo	2019

T2_7_7	*Pseudomonas savastanoi savastanoi*	pv.	*Olea europaea*	Italy	Viterbo	2019
T2_7_8	*Pseudomonas savastanoi savastanoi*	pv.	*Olea europaea*	Italy	Viterbo	2019
T2_7_9	*Pseudomonas savastanoi savastanoi*	pv.	*Olea europaea*	Italy	Viterbo	2019
T2_7_10	*Pseudomonas savastanoi savastanoi*	pv.	*Olea europaea*	Italy	Viterbo	2019

Table S2 TRs results for each strain and the loci of *Pseudomonas savastanoi*

Name	TR1 sav	TR2 sav	TR3 sav	TR6 sav	TR7 sav	TR8 sav	TR10 sav	TR11 sav	TR13 sav	TR15 sav	TR16 sav	TR17 sav	TR18 sav	TR19 sav	TR20 sav
ITM301	2	9	5	5	8	3	3	5	4	11	17	2	2	7	5
ITM302	2	9	5	5	5	3	3	5	4	11	16	2	2	7	5
ITM304	2	8	5	5	4	3	3	5	4	11	15	2	2	7	4
ISPAT3PAITM909	3	11	3	5	0	3	4	3	4	14	38	2	3	3	5
ISPAT5PAITM913	3	11	3	5	0	3	3	3	4	14	39	2	3	3	5
ISPAT3PBITM916	3	11	3	5	0	3	4	5	5	14	17	2	3	3	4
AL132	5	5	3	3	5	3	2	3	5	21	12	2	3	5	7
AL135	5	5	3	3	5	3	2	3	5	22	11	2	3	5	7
AL142	5	5	3	3	5	3	2	4	5	19	15	2	3	5	7
SYR617	4	12	6	2	3	3	4	4	5	12	8	2	4	5	7
SYR194	4	12	6	2	3	3	4	4	5	12	8	2	4	5	7
MA154	4	5	3	3	4	3	4	5	4	14	16	3	4	5	7
MA155	4	5	3	3	4	3	4	5	4	14	16	3	4	5	7
TN177	3	8	7	5	4	3	3	7	4	17	3	2	3	7	8
PZ1	2	9	6	5	3	3	3	5	4	15	17	2	2	8	9
PZ4	2	9	6	4	3	3	3	6	4	15	18	2	2	8	9
I-Ap107	5	5	2	9	13	3	2	3	5	21	10	2	3	4	6
I-Ap108	3	5	2	5	5	3	2	3	4	12	16	2	3	4	6
I-Ap111	5	10	3	6	4	3	4	6	4	15	8	2	2	5	7
I-Ap113	5	9	2	4	4	3	2	3	5	27	9	2	2	4	6
I-Ap115	5	8	2	4	4	3	2	3	5	29	11	2	2	4	6
I-Ap118	5	12	2	12	4	3	2	3	5	22	7	2	2	4	6
I-Ap163	3	9	3	6	6	3	2	3	5	18	12	2	3	5	7
I-Ap164	3	7	3	7	7	3	2	4	5	11	4	2	3	5	7

PVFiA14	4	11	3	4	4	3	3	4	5	22	13	2	4	3	3
SUPP3085	4	13	2	5	2	3	3	7	4	14	6	2	4	1	0
SUPP3129	3	9	4	2	5	3	3	5	4	15	7	2	3	3	5
PVFiF1	3	8	3	4	2	3	2	7	2	14	4	2	4	4	0
PVFiF2	3	8	2	4	2	3	2	7	2	14	4	2	3	2	3
PD539	3	5	2	4	2	3	2	7	2	13	4	2	3	2	0
T5-1	3	5	4	4	2	3	1	8	2	13	4	2	3	3	5
NCPPB3472	3	8	2	4	2	3	2	9	2	14	4	2	3	4	7
ICMP9129	3	6	2	3	2	3	2	6	2	16	4	2	3	4	7
ICMP9131	3	6	2	3	2	3	2	8	2	15	4	2	3	4	7
PD161	3	5	2	3	2	3	2	8	2	15	5	2	3	4	5
PD179	3	5	2	3	2	3	2	8	2	14	4	2	3	4	6
NCPPB1464	3	5	2	3	2	3	2	6	4	13	4	2	3	2	5
NCPPB1006	2	8	2	2	3	3	2	5	2	14	8	2	2	3	0
ITM510	0	9	4	6	6	3	2	3	5	3	0	2	0	5	5
PVBa219-2	0	9	4	6	2	3	3	3	5	3	23	2	0	3	4
ITM401	0	7	3	5	2	3	3	6	5	4	0	2	0	5	5
ITM402	0	7	3	5	2	3	3	6	5	4	0	2	0	5	5
ITM519	0	8	3	3	11	3	3	3	5	4	21	2	0	5	5
ITM602	0	11	3	5	2	3	3	4	5	4	0	2	0	5	5
ITM717	0	8	3	5	2	3	3	4	5	4	0	2	0	5	5
ITM718	0	9	3	5	2	3	3	4	5	4	0	2	0	5	4
ITM719	0	12	3	5	2	3	3	4	5	4	0	2	0	5	5
ITM305	0	11	4	3	6	3	3	6	5	3	0	2	0	6	3
GTG19	3	10	1	3	2	2	2	3	5	22	0	2	3	1	4
GTG24	3	10	1	3	2	2	2	3	5	22	0	2	3	1	4
GTG26	3	10	1	3	2	2	1	3	5	22	0	2	3	3	4
ITM723	3	10	1	3	3	2	2	3	5	22	17	2	3	3	4
Pba208-1	4	8	4	6	3	3	4	5	4	15	12	2	4	6	7
Pba215	3	12	4	5	3	3	3	5	4	15	21	2	3	6	7
Pba218	3	14	4	5	3	3	3	5	4	15	16	2	3	6	7
IPVBa2325	0	10	1	0	0	2	48	3	4	2	4	2	0	15	5
NCPPB2571	0	8	2	0	0	2	48	3	4	2	4	2	0	15	5
CECT4861	2	8	2	7	0	3	4	0	4	17	4	2	2	12	0
T1_1	3	12	4	6	6	3	2	4	5	21	19	2	3	6	2
T1_10	3	9	3	7	4	3	3	4	3	20	7	2	3	5	3
T2_3	3	12	3	6	5	3	2	4	5	20	14	2	3	6	2
T3_3	3	7	2	7	4	3	2	4	3	21	9	2	3	5	3
T3_5	3	12	3	6	5	3	2	4	5	17	17	2	3	6	2
T3_10	3	10	2	6	6	3	2	4	5	17	20	2	3	5	2
T4_4	3	15	3	6	5	3	2	4	5	17	15	2	3	6	2
T4_9	3	14	3	6	5	3	2	4	5	17	14	2	3	5	2
T4_10	3	14	3	6	5	3	2	4	5	17	13	2	3	5	2

T5_6	3	10	2	6	6	3	3	4	5	17	19	2	3	4	2
T5_10	3	12	4	6	5	3	3	4	5	17	15	2	4	5	2
T6_6	3	14	4	6	5	3	2	4	5	17	16	2	3	5	2
T6_9	3	14	3	6	6	3	2	4	5	17	15	3	3	5	2
T1_5	3	11	3	6	5	3	2	4	5	17	18	3	3	5	3
T2_2	4	14	2	7	4	3	2	4	5	23	8	2	4	4	5
T2_7_0	3	12	4	6	5	3	3	4	5	19	14	3	3	5	2
T2_7_1	3	12	4	6	6	3	3	4	5	20	16	2	3	5	2
T2_7_2	3	10	4	6	6	3	3	4	5	18	13	2	3	4	2
T2_7_3	3	12	4	6	6	3	3	4	5	20	15	2	3	5	2
T2_7_4	3	12	4	6	6	3	3	4	5	19	15	2	3	5	2
T2_7_5	3	12	3	6	6	3	3	4	5	20	15	2	3	5	2
T2_7_6	4	12	4	6	6	3	3	4	5	18	15	2	3	5	2
T2_7_7	3	12	4	6	6	3	3	4	5	20	15	2	3	5	2
T2_7_8	3	12	3	3	6	3	3	4	5	20	15	2	3	5	2
T2_7_9	3	12	4	6	6	3	3	4	5	17	15	2	3	5	2
T2_7_10	3	12	4	6	6	3	3	4	5	20	14	2	3	5	2

Chapter 4 **: Olive quick decline syndrome**

4.1 **Disease description**

This disease has expanded in a short time to many olive orchards. The different scenarios have been discussed and the University of Bari has investigated the problem. In a study focused on fungi in 2013 conducted on olive orchards in Apulia, results showed fungal species *Phaeoacremonium* spp., *Phaemoniella* in particular, *Pleumostomophora* and *Neofusicoccum*, in addition to galleries of the leopard moth (Zeuzera pyrina) and bark beetles occurring in different parts of the plant (1). In addition to obtaining some of the results above, another study by Nigro and his colleague in 2013 has identified the fungi *Pleurostomophora richardsiae* as capable of causing brown wood streaking (2).

In the same year, the bacterium, *Xylella fastidiosa*, was detected from olive, oleander and almond trees by molecular and double-antibody sandwich enzyme-linked immunosorbent assay (DAS-ELISA) in the infected area (3). Due to the involvement of different causal agents attributed to this disease, such as fungi, leopard moth and *Xylella fastidiosa* , the name “olive quick decline syndrome” was given, corresponding to the Italian "Complesso del Disseccamento Rapido dell’livo" (CoDiRO) (4). The dispute continued about the real causal agents after the identification and detection of the different causal agents (5), until it was finally proved that only the bacterium *Xylella fastidiosa* subsp. *pauca* is the causal agent of the death of the trees (5,6). The pathogen involved has been classified as *Xylella fastidiosa* subsp. *pauca* (CoDiRO) (4).

Researchers continued to investigate the isolated bacterium until the strains in Apulia were assigned to a novel sequence type profile (ST53) close to subsp *pauca* on the basis of MLST analysis (4,7). The presence of insect vectors is the main factor contributing to the spread of the disease over a wide area (8). Since the emergence of the bacterium, researchers have identified additional hosts in south Italy, such as oleander, cherry, myrtle-leaf milkwort (*Polygala myrtifolia*) and coastal rosemary (*Westringia fruticosa*) (9). The disease has spread extremely rapidly in southern Italy, as shown in the maps below.

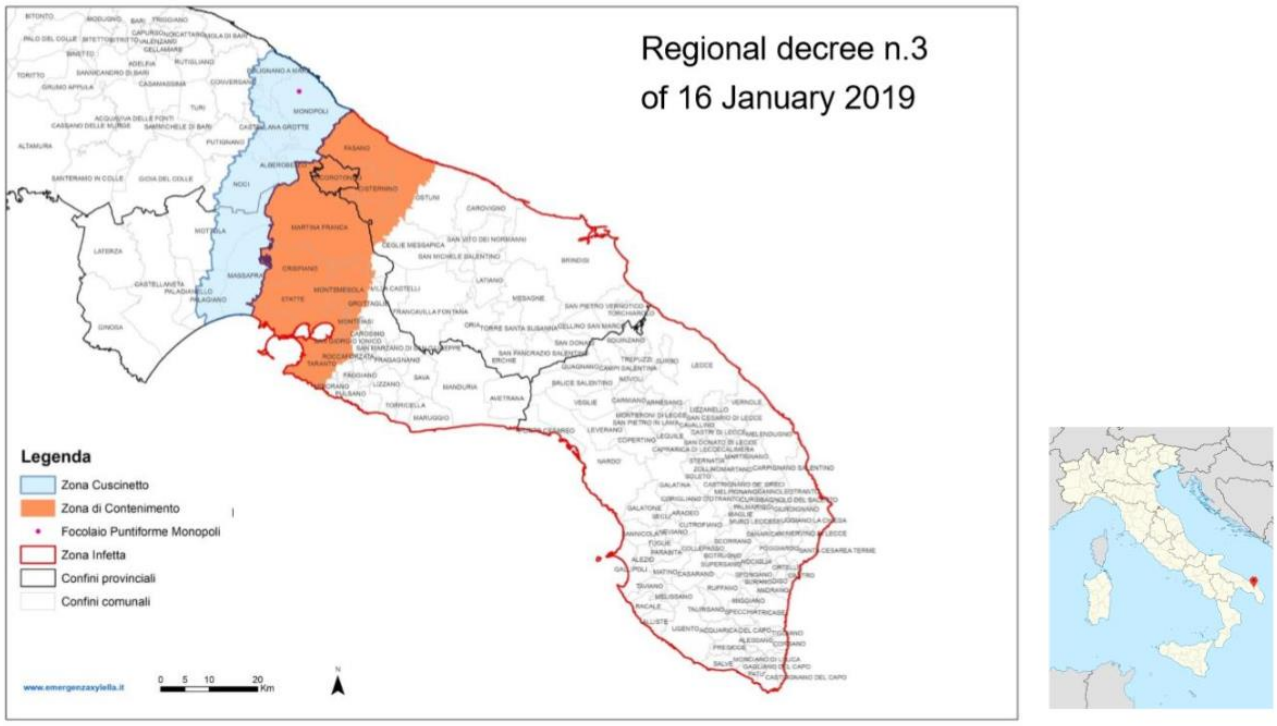


Figure 25: Map of the distribution of Olive Quick Decline Syndrome in the Apulia region of southern Italy (COMMISSION, 2018).

The emerging bacterium poses a great threat due to its spread over the whole area where it is detected primarily on olive trees, to the presence of vectors, and to its wide host range (10). After identification of the pathogen in Apulia, different *Xylella fastidiosa* subspecies have been identified around the world: in France (11), in Germany (2016) where it was eradicated (12), in Spain (13), and in Iran (14). Subsequently, the disease was detected in Brazil, Argentina, Chile (15)

4.2 ***Xylella fastidiosa* :**

The bacterium is gram-negative, rod-shaped, slow-growing, strictly-aerobic, non-flagellated, catalase-positive, oxidase negative, non-pigmented and mesophilic (optimum growth occurs at 26-28°C), GC content of 51.9% (13).

It is a vector-transmitted pathogen, and primary vectors are sharpshooters and froghoppers or spittlebugs (*Cicadellidae*); xylem sap-feeding insects lack a latent period. The bacterium can be transmitted by infected plant material but not via seed (16).

The bacterial cells possess two types of pili, short type I pili and long type IV pili, which allow the bacterium to move both up- and downstream (17). Therefore, the bacterium was detected in all parts of the diseased plants, including xylem vessels, roots, stems, and leaves (18). These types of pili, Type I and Type IV pili, which play different roles in twitching motility, biofilm formation, and cell-cell aggregation, and also have opposite effects on movement and biofilm formation (19). Twitching motility helps the bacterium to achieve intra-plant long-distance movement and colonization, contributing to pathogenicity (20–22). In addition to their role in pathogenicity, biofilm formation and cell-cell aggregation have an important role in biological functions, such as enhanced resistance to the environmental conditions against antimicrobial agents (23). Biofilm formation occurs in the host plant and vector (foregut), in the plant, and in advanced stages, the vessels are blocked by bacterial aggregates, which prevent the water and nutrient movement (24–26). Since *Xylella fastidiosa* requires special media for growth in the laboratory, different media have been developed for this purpose, including PD2, PW, CS20 or BCYE (17). *Xylella fastidiosa* has a wide host range and is associated with a large number of diseases, many of which cause great economic losses in fruit crops and ornamental (27).

4.3 Disease caused by *Xylella fastidiosa*

The first disease caused by this bacterium appeared in the 1880s in United States vineyards (22). The disease was described by Newton B. Pierce, a US professional plant pathologist, and thus became known as “Pierce’s disease” (Pierce, 1892).

In 1890, similar symptoms were also found in Georgia, USA, and this disease was later known as phony peach disease (28). In 1980, the bacterium was found to cause almond leaf scorch (29). It was subsequently reported to cause elm, sycamore, and oak leaf scorch (30). The bacterium was described and cultured for the first time in 1987 (31). The disease began to be identified in other parts of the world. In Brazil *Xylella fastidiosa* was reported as the causal agent of Citrus Variegated Chlorosis (CVC), it was also responsible for the leaf scorching on coffee plants (*Coffea arabica L.)* (32), and was found to be the causal agent of leaf scorch disease of oleander (*Nerium oleander*) (33,34).

The bacterium was also detected on alfalfa (35) avocado, plum, mulberry, periwinkle wilt, and many other host plants (36). Recently the bacterium was detected in olive trees in Italy, and in other hosts in different countries: USA, Spain, France, Germany, Iran, Portugal (16,28,37,38).

At present, 563 plant species are described as infected by *Xylella fastidiosa*; these plant species belong to different host plant genera and botanical families (7,37,39).

4.4 Taxonomy

The taxonomic tree for *Xylella fastidiosa* is as follows:

Domain: Bacteria

Phylum: Proteobacteria

Class: Gammaproteobacteria

Order: *Xanthomonadales*

Family: X*anthomonadaceae*

Genus: *Xylella*

Species: *Xylella fastidiosa*

According to DNA-DNA relatedness and sequencing of the 16S–23S intergenic spacer (ITS), there are three subspecies: *Xylella fastidiosa* subsp. *pierce*, subsp. *multiplex*, and subsp. *pauca* (40). The three species were confirmed by MLST in addition to a suggested fourth subspecies sandyi (41). A fifth subspecies is *morus*, whose sequences are a recombination of subsp. *fastidiosa* and subsp. *multiplex* (42). The four subspecies *fastidiosa, multiplex, pauca,* and *sandyi* have evolved within a distinct geographical range (43). The causal agent of olive quick decline syndrome belonging to subsp. *pauca* strain is CoDiRO, and approximately 380 species and 28 families of plants are considered as its hosts (44). *Xylella fastidiosa* subsp. *pauca* is native to South America and attacks citrus and coffee plants (11). MLST analysis results have assigned the Apulian strain into a new sequence type profile (ST53) (45). The sequence type was extremely close to *Xylella fastidiosa* subsp. *pauca* affecting coffee plants. The same sequence type was obtained by MLST analysis of *Xylella fastidiosa* subsp. *pauca* isolates from Costa Rica hosted by oleander and coffee plants (8). *Xylella fastidiosa* subsp. *pauca* is the causal agent of Citrus Variegated Chlorosis (CVC) (46) and Coffee (47), and the strain originates from South America.

4.5 Symptoms

Leaf tips and margins showed massive browning, and leaf scorching is followed by rapid dieback of shoots, twigs, and branches, which leads to the death of the entire tree (48). Older leaves show the symptoms before young leaves, discolorations of the vascular system are observed, and mummified fruits remain attached to the shoots. Numerous galleries of the leopard moth, *Zeuzera pyrina* occur on the trunks (4,5). These symptoms have been associated

with water and nutrient shortages caused by xylem blockage due to the formation of bacterial biofilms and products of plant defence response, such as gums and tyloses (2,3). It is not necessary for all the symptoms to be present in order to diagnose an infected tree (2). Bacterial colonization blocks the vessels of host plants. Another colonization strategy is a vessel-to-vessel movement, whose failure results in vessel plugging and disease (49). The blocking of xylem could be due to the production of tyloses and polysaccharide-rich gels by plants (50,51). There is evidence that the pathogen can produce cell-wall degrading enzymes and hemolysin-like toxin to kill the plant (52).

Figure 26: Olive quick decline symptoms: (A) leaf tips ; (B) initial symptoms on the trees;(C) branch diebacks; (D) Symptoms om the field different trees.

Until now the Leccino variety shows tolerance to the pathogens according to the experimental studies and observations, also the measurements showed a low bacterial population in this variety compared to (e.g. *Cellina di Nardo, Ogliarola salentina*) (53).

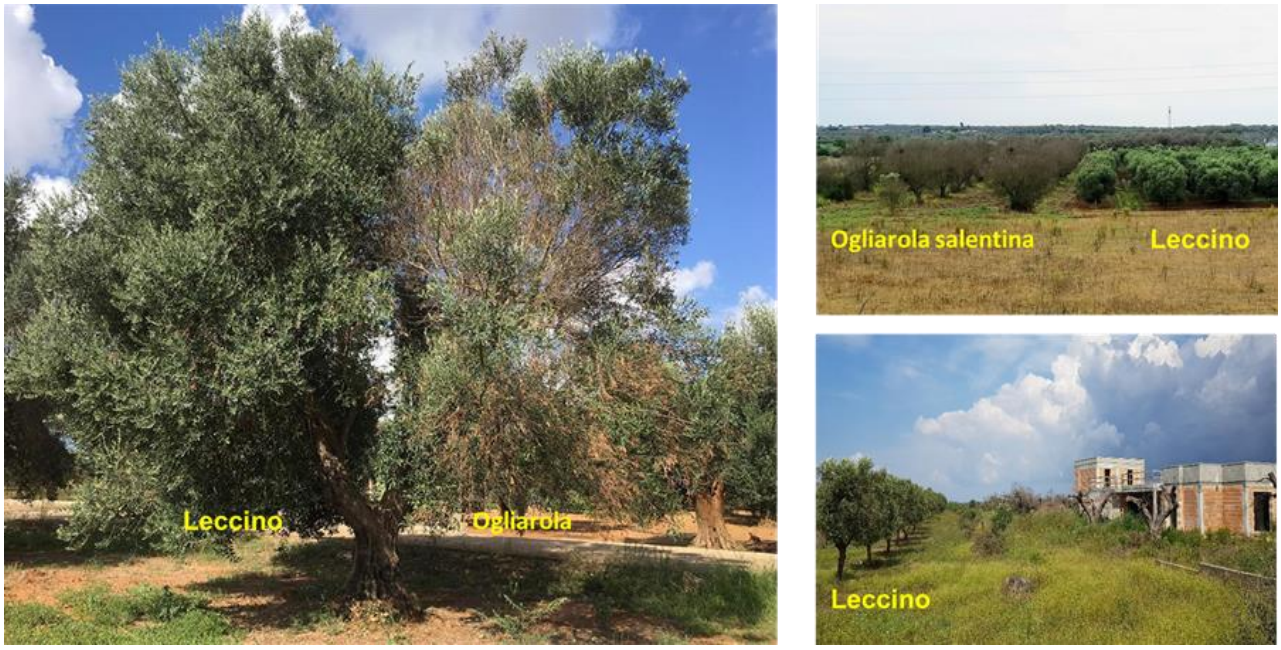


Figure 27: Shows the difference between Olive resistance variety Leccino and *Ogliarola salentina*.

4.6 **Vectors**

It is very important to understand the biology of the vectors as the entrance to understanding the epidemiology of the disease. *Xylella fastidiosa* can be transmitted by several species of suborder *Auchenorrhyncha* of the *Hemiptera*, sharpshooter leafhoppers (*Cicadellidae: Cicadellidae*) and spittlebugs or froghoppers (*Hemiptera: Cercopoidea*), which are xylem-fluid feeders and able to transmit the bacterium (54). These insects lack a latent period and have no transstadial or transovarial transmission of the bacterium (55,56). Sucking mouthparts are a characteristic feature of these insects (mandibular and maxillary stylets) and help the insects to reach the xylem. Xylem fluid feeders ingest large amounts of crude sap and produce large amounts of liquid excretions due to the poor nutritional value of xylem fluid (17).

Insects acquire *Xylella fastidiosa* by feeding on the xylem (less than 2 hours) but *Xylella fastidiosa* does not circulate in their haemolymph, nor does it require a latent period before transmission of the bacterium. Bacteria are pearly attached to the mouthparts and released while the insect feeds on other host plants. The bacteria are restricted to the foregut of the insects where they can multiply and persist (57). The most important vectors identified in North and South America belong to 39 species of *Cicadellinae*, and five spittlebugs (*Aphrophoridae*) are vectors of *Xylella fastidiosa* (58). Of these, the glassy-winged sharpshooter leafhopper (*Homalodisca vitripennis*) is the most important species of the *Xylella fastidiosa* vectors and it has the ability to distribute the bacteria efficiently in a new area (56).

Present in Europe, the Families *Cicadidae, Cicadellidae, Cercopidae, Aphrophoridae* and *Tibicinidae* are xylem-fluid feeding insects and are potential vectors of the bacteria (59). (60) reported that only the green leafhopper, (*Cicadella viridis*), and the meadow spittlebug, (*Philaenus spumarius*) (*Hemiptera: Aphrophoridae*) are potential vectors for *Xylella fastidiosa* in Europe. However, Auchenorrhyncha species were also found in Spain and considered as potential vectors of *Xylella fastidiosa*, for example, the spittlebugs (*Hemiptera: Cercopoidea*) and *Neophileanus sp*. were found near citrus, olive, and grape fields (17).

In Italy, *Philaenus spumarius* is the only vector experimentally able to acquire and transfer the bacterium from/to different host plants, in addition to other hosts that serve as pathogen inoculum sources (61). Besides *Philaenus spumarius*, *Neophilaenus campestris* and *Euscelis lineolatus* were reported as positive for the presence of *Xylella fastidiosa* in southern Italy (8,62,63), while they were reported negative in another study (8). Other studies contain further details of potential *Xylella fastidiosa* vectors in Europe (10).

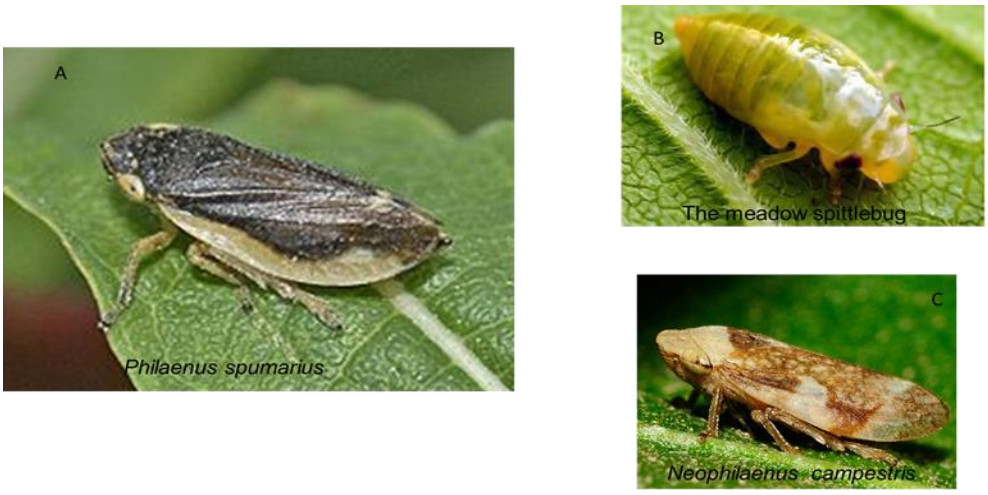


Figure 28: (A) *Philaenus spumarius* is the main vector in southern Italy; (B) the Nymph of meadow spittlebug and (C) *Neophilaenus campestris* a suspected vector.

4.7 The disease processes

Two important factors play an important role in the disease processes; the first is temperature, which can regulate the growth of the pathogen in the host plants since the bacterial population can stay in a stable condition with temperatures between 17 and 34°C. The other factor is wet weather in winter, which helps the survival of high vector populations and favours disease spread in regions with dry summers (40,57,64). The source of inoculum is usually the host plants in the Apulia region, for example, almond, sweet cherry, oleander, broom, *Polygala myrtifolia*, *Westringia fruticose*, *Acacia saligna*, *rosmarinus officinalis*, *Rhamnu salaternus*,

and *Myrtus communis* are the main source of inoculum (30). Another source is the high populations of this spittlebug that colonize olive trees in spring-late summer (4,65).

Regarding this aspect, vectors play an important role in the disease process. The bacteria have two environments to colonize; the first one is the host plant xylem, and the second is the vector foregut. Both habitats form the life cycle of the bacterium and protection to the bacterium from disease management applications (5). Three essential steps are involved in the transmission of bacteria. The first is an acquisition from an infected host plant then attachment and maintenance in the vector's foregut cuticle; finally come detachment and inoculation into a new host (56,66,67). Insects acquire the XF. by feeding on the xylem (less than 2 hours) but the bacterium does not circulate in its haemolymph, nor does it require a latent period before transmission of the bacterium when it moves to the other plant or host (17). When the adult vector acquires the bacteria, it barely attached to its mouthparts, maintained in vector's foregut, and then released to the next host plant again. The bacterium is restricted to the foregut of the insects where it can multiply and persist (58). Despite their ability to transmit the bacteria, nymphs lose this ability at each stage the bacteria moult, which means that the new adult must feed on the infected plant in order to transfer the bacteria (68). When the infected vector feeds on the host plant, the bacteria move systematically within the xylem vessels and extensively colonize the xylem vessels, where they can easily be accessible for acquisition by other insects (68,69). Multiplication of the bacteria within xylem vessels blocks the transportation of mineral nutrients and water (70). *Xylella fastidiosa* can move within the plant through vessels bordered pits after damaging them because these pits do not allow larger objects to pass, only xylem sap (17,39).

This process is very important for the pathogen to move within plant cells (71). In any case, the disease process and its progress are based on the ability of the bacteria to move and spread, forming a community within the plant from the point of infection. Biofilm formation is due to the gathering of bacterial colonies; when it reaches a large amount, it is able to completely block xylem vessels, thus causing a deficiency in water and nutrients (44).

Xylella fastidiosa has different pathogenicity factors for host-specific colonization. *Xylella fastidiosa* has Type II secretion system, and this system is responsible for the colonization of the bacteria on the xylem vessels, and the exportation of exoenzymes that degrade the plant cell wall (52). Another factor is the endoglucanase and protease enzymes used by *Xylella fastidiosa* to degrade the pit membrane, which allows the bacteria to pass via

xylem vessels. These enzymes are controlled by the Type II secretion system (17). The movement of the bacteria is based on Type IV pili, via twitching motility, which allows the bacteria to occupy the healthy vessels, as the movement of the bacteria occurs against a fluid current (53).

Extracellular polysaccharides play an important role in xylem blocking and also play a fundamental role in plant virulence, biofilm formation, and vector transmission (18).

Biofilm formation enhances the *Xylella fastidiosa* bacteria for adaption and tolerance to environmental stresses, such as antibiotics, dehydration, and host defences, in addition to the competition within the host xylem (72,73).

Attachment to host surfaces by cell surface adhesins is an initial step in the formation of a biofilm. The process of forming the biofilm on a xylem surface is initiated by attachment of formerly planktonic cells to a surface, followed by proliferation of the cells and their self-association to form an aggregate population of micro-organisms attached to each other (24–26,51,74). In addition, adhesins play two important roles in adherence to vector tissues, this controlled by Type IV fimbriae and adherence to the xylem cell walls as well as microcolony formation in the xylem vessels and/or bacteria movement in the vascular system (75,76).

Cyclic di-GMP synthase is also required for biofilm formation, plant virulence, and vector transmission (77). Signalling sensor or cell-to-cell communication is required for virulence and insect transmission of *Xylella fastidiosa* (78).

Different genes responsible for the regulation of iron homeostasis involved in pathogenicity in bacteria have been identified with *Xylella fastidiosa* , genes responsible for toxicity, and anti-oxidant response genes (52).

4.8 References

1. Hosni T, Moretti C, Devescovi G, Suarez-Moreno ZR, Fatmi MB, Guarnaccia C, et al. Sharing of quorum-sensing signals and the role of interspecies communities in bacterial plant disease. ISME J [Internet]. 2011;5(12):1857. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21677694

2. Nigro F, Boscia D, Antelmi I, Ippolito A. Fungal species associated with a severe decline of olive in southern Italy. J Plant Pathol. 2013;95(3): ok.

3. Carlucci A, Raimondo Mal, Cibelli F, Phillips Alajl, LOPS F. *Pleurostomophora richardsiae*, *Neofusicoccum parvumand*, *Phaeoacremonium aleophilum* associated with a decline of olives in southern Italy. Phytopathol Mediterr. 2013;52(3):517–27.

4. Saponari M, Boscia D, Nigro F, Martelli GP. Identification of DNA sequences related to *Xylella fastidiosa* in oleander, almond and olive trees exhibiting leaf scorch symptoms in Apulia (Southern Italy). J Plant Pathol. 2013;95(3).

5. Martelli GP, Boscia D, Porcelli F, Saponari M. The olive quick decline syndrome in south-east Italy: a threatening phytosanitary emergency. Eur J plant Pathol. 2016;144(2):235–43.

6. Ciervo M. The olive quick decline syndrome (OQDS) diffusion in Apulia Region: an apparent contradiction according to the agricultural model. Belgeo. 2016;(4).

7. Saponari M, Loconsole G, Almeida R, Coletta-Filho HD. Isolation, genotype and preliminary data on the pathogenicity of *Xylella fastidiosa* CoDiRO strain. In 2013.

8. Elbeaino T, Valentini F, Kubaa RA, Moubarak P, Yaseen T, Digiaro M. Multilocus sequence typing of *Xylella fastidiosa* isolated from olive affected by "olive quick decline syndrome" in Italy. Phytopathol Mediterr. 2014;53(3):533–42.

9. Cavalieri V, Porcelli F. Main insect vectors of *Xylella fastidiosa* in Italy and worldwide. *Xylella fastidiosa* Olive Quick Decline Syndr. 2017;31.

10. Saponari M, Loconsole G, Cornara D, Yokomi RK, De Stradis A, Boscia D, et al. Infectivity and transmission of *Xylella fastidiosa* by *Philaenus spumarius* (Hemiptera: Aphrophoridae) in Apulia, Italy. J Econ Entomol. 2014;107(4):1316–9.

11. D'Onghia AM, Brunel S, Valentini F, D'Onghia AM, Digiaro M, Brunel S, et al. *Xylella fastidiosa* & the Olive Quick Decline Syndrome (OQDS) A serious worldwide challenge for the safeguard of olive trees. IAM Bari CIHEAM (Centre Int Hautes Etudes Agron Méditerranéennes), 20 – p (17 172 A Série Mediterr Semin N° 121, Options Méditerranéennes). 2017;20-p.

12. Eppo. First report of *Xylella fastidiosa* in France. EPPO Glob Database. 2015;(Num. article: 2015/144).

13. Health EP on P, Jeger M, Caffier D, Candresse T, Chatzivassiliou E, Dehnen-Schmutz K, et al. Updated pest categorization of *Xylella fastidiosa* . EFSA J. 2018;16(7):e05357.

14. Amanifar N, Taghavi M, Izadpanah K, Babaei G. Isolation and pathogenicity of *Xylella fastidiosa* from grapevine and almond in Iran. Phytopathol Mediterr. 2014;318–27.

15. COMMISSION E. List of demarcated areas established in the Union territory for the presence of *Xylella fastidiosa* as referred to in Article4 (1) of Decision (EU) 2015/789 -UPDATE10-. 2018; Available from https://ec.europa.eu/food/sites/food/files/plant/docs/ph_biosec_legis_list-demarcated-union-territory_en.pdf

16. Wells JM, Raju BC, Hung H-Y, Weisburg WG, Mandelco-Paul L, Brenner DJ. *Xylella fastidiosa* gen. nov., sp. nov: gram-negative, xylem-limited, fastidious plant bacteria related to *Xanthomonas* spp. Int J Syst Evol Microbiol. 1987;37(2):136–43.

17. Janse JD, Obradovic A. *Xylella fastidiosa* : its biology, diagnosis, control, and risks. J Plant Pathol. 2010; S35–48.

18. Meng Y, Li Y, Galvani CD, Hao G, Turner JN, Burr TJ, et al. Upstream migration of *Xylella fastidiosa* via pilus-driven twitching motility. J Bacteriol [Internet]. 2005;187(16):5560–7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16077100

19. Krivanek AF, Walker MA. Vitis Resistance to Pierce's Disease Is Characterized by Differential *Xylella fastidiosa* Populations in Stems and Leaves. Phytopathology [Internet]. 2005;95(1):44–52. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18943835

20. De La Fuente L, Burr TJ, Hoch HC. Autoaggregation of *Xylella fastidiosa* cells is influenced by type I and type IV pili. Appl Environ Microbiol [Internet]. 2008;74(17):5579–82. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18641157.

21. Li Y, Hao G, Galvani CD, Meng Y, De La Fuente L, Hoch HC, et al. Type I and type IV pili of *Xylella fastidiosa* affect twitching motility, biofilm formation, and cell-cell aggregation. Microbiology [Internet]. 2007;153(Pt 3):719–26. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17322192

22. Chatterjee S, Almeida RPP, Lindow S. Living in two worlds: the plant and insect lifestyles of *Xylella fastidiosa*. Annu Rev Phytopathol [Internet]. 2008;46:243–71. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18422428

23. Shi X, Lin H. Visualization of Twitching Motility and Characterization of the Role of the PilG in *Xylella fastidiosa*. J Vis Exp JoVE [Internet]. 2016;(110). Available from: http://www.ncbi.nlm.nih.gov/pubmed/27166660

24. Marques LLR, Ceri H, Manfio GP, Reid DM, Olson ME. Characterization of biofilm formation by *Xylella fastidiosa* in vitro. Plant Dis. 2002;86(6):633–8.

25. Lorite GS, Rodrigues CM, de Souza AA, Kranz C, Mizaikoff B, Cotta MA. The role of conditioning film formation and surface chemical changes on *Xylella fastidiosa* adhesion and biofilm evolution. J Colloid Interface Sci [Internet]. 2011;359(1):289–95. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21486669

26. Silva MS, De Souza AA, Takita MA, Labate CA, Machado MA. Analysis of the biofilm proteome of *Xylella fastidiosa* . Proteome Sci. 2011;9(1):58.

27. Hopkins DL. *Xylella fastidiosa* and other fastidious bacteria of uncertain affiliation. Lab Guid Identif Plant Pathog Bact 2nd ed(NW Schaad, ed), APS Press St Paul, MN. 1988;201–13.

28. Hopkins DL, Purcell AH. *Xylella fastidiosa* : cause of Pierce's disease of grapevine and other emergent diseases. Plant Dis. 2002;86(10):1056–66.

29. Wells JM, Raju BC, Nyland G. Isolation, culture and pathogenicity of the bacterium causing phony disease of peach. Phytopathology. 1983;73(6):859–62.

30. Davis MJ, Thomson S V, Purcell AH. Etiological role of a xylem-limited bacterium causing Pierce's disease in almond leaf scorch. Phytopathology. 1980;70(472):5.

31. Hearon SS, Sherald JL, Kostka SJ. Association of xylem-limited bacteria with elm, sycamore, and oak leaf scorch. Can J Bot. 1980;58(18):1986–93.

32. De Lima JEO, Miranda VS, Hartung JS, Brlansky RH, Coutinho A, Roberto SR, et al. Coffee leaf scorch bacterium: axenic culture, pathogenicity, and comparison with *Xylella fastidiosa* of citrus. Plant Dis. 1998;82(1):94–7.

33. Lee RF, Derrick KS, Beretta MJG, Chagas CM, Rosetti V. Citrus variegated chlorosis: a new destructive disease of citrus in Brazil. Citrus Ind. 1991;72(10):12–3.

34. Beretta MJG, Lee RF, Barthe GA, Neto JT, Derrick KS, Davis CI. Citrus variegated chlorosis: detection of *Xylella fastidiosa* in symptomless trees. In: International Organization of Citrus Virologists Conference Proceedings (1957-2010). 1993.

35. Purcell AH, Saunders SR, Hendson M, Grebus ME, Henry MJ. Causal role of *Xylella fastidiosa* in oleander leaf scorch disease. Phytopathology. 1999;89(1):53–8.

36. Daugherty MP, Lopes JRS, Almeida RPP. Strain-specific alfalfa water stress induced by *Xylella fastidiosa*. Eur J plant Pathol. 2010;127(3):333–40.

37. Krugner R, Sisterson MS, Chen J, Stenger DC, Johnson MW. Evaluation of olive as a host of *Xylella fastidiosa* and associated sharpshooter vectors. Plant Dis [Internet]. 2014;98(9):1186–93. Available from: http://www.ncbi.nlm.nih.gov/pubmed/30699616

38. Jones P, Comfort D. x factor threat. T Ctry Plan. 2017;375–9.

39. EPPO, Organization) E (European and MPP, EPPO. PM 7/24 (3) *Xylella fastidiosa*. EPPO Bull. 2018;48(2):175–218.

40. EFSA, Authority EFS. Update of the Xylella spp. host plant database. EFSA J. 2018;16(9):e05408.

41. Schaad NW, Postnikova E, Lacy G, Fatmi M, Chang C-J. *Xylella fastidiosa* subspecies: *X. fastidiosa* subsp *piercei*, subsp. nov., *X. fastidiosa* subsp. *multiplex* subsp. nov., and *X. fastidiosa* subsp. *pauca* subsp. nov. Syst Appl Microbiol. 2004;27(3):290–300.

42. Scally M, Schuenzel EL, Stouthamer R, Nunney L. Multilocus sequence type system for the plant pathogen *Xylella fastidiosa* and relative contributions of recombination and point mutation to clonal diversity. Appl Env Microbiol [Internet]. 2005;71(12):8491–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16332839

43. Nunney L, Schuenzel EL, Scally M, Bromley RE, Stouthamer R. Large scale intersubspecific recombination in the plant pathogenic bacterium *Xylella fastidiosa* is associated with the host shift to mulberry. Appl Environ Microbiol. 2014;AEM. 04112-13.

44. Almeida RPP, Nunney L. How do plant diseases caused by *Xylella fastidiosa* emerge? Plant Dis [Internet]. 2015;99(11):1457–67. Available from: http://www.ncbi.nlm.nih.gov/pubmed/30695952

45. Almeida RPP, Nascimento FE, Chau J, Prado SS, Tsai C-WW, Lopes SA, et al. Genetic structure and biology of *Xylella fastidiosa* strains causing disease in citrus and coffee in Brazil. Appl Environ Microbiol [Internet]. 2008;74(12):3690–701. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18424531

46. Nunney L, Schuenzel EL, Scally M, Bromley RE, Stouthamer R. Large-scale intersubspecific recombination in the plant-pathogenic bacterium *Xylella fastidiosa* is associated with the host shift to mulberry. Appl Env Microbiol [Internet]. 2014;80(10):3025–33. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24610840

47. Rossetti V, Garnier M, Bové JM, Beretta M-J-G, Teixeira ARR, Quaggio JA, et al. Présence de bactéries dans le xylème d'orangers atteints de chlorose variégée, une nouvelle maladie des agrumes au Brésil. Comptes rendus l'Académie des Sci Série 3, Sci la vie. 1990;310(8):345–9.

48. Li W-B, Pria Jr WD, Teixeira DC, Miranda VS, Ayres AJ, Franco CF, et al. Coffee leaf scorch caused by a strain of *Xylella fastidiosa* from citrus. Plant Dis. 2001;85(5):501–5.

49. Elmore W, Gillett-Kaufman. *Xylella fastidiosa* and Olive Quick Decline: Symptoms and Identification of an Insect Vectored Pathogen . 2017; Available from: http://edis.ifas.ufl.edu/pdffiles/IN/IN116500.pdf

50. Newman KL, Almeida RPP, Purcell AH, Lindow SE. Use of a green fluorescent strain for analysis of *Xylella fastidiosa* colonization of *Vitis vinifera*. Appl Env Microbiol [Internet]. 2003;69(12):7319–27. Available from: http://www.ncbi.nlm.nih.gov/pubmed/14660381

51. Hoch HC, Burr TJ. Understanding control of *Xylella fastidiosa* cell aggregation: importance in colonization and biofilm development in grapevine and sharpshooter foregut. In: Proceedings of the Pierce's Disease Research Symposium. 2007.

52. Chatterjee S, Wistrom C, Lindow SE. A cell–cell signaling sensor is required for virulence and insect transmission of *Xylella fastidiosa* . Proc Natl Acad Sci. 2008;105(7):2670–5.

53. Dow JM, Daniels MJ. Xylella genomics and bacterial pathogenicity to plants. Int J Genomics. 2000;1(4):263–71.

54. EFSA, Authority EFS, Baù A, Delbianco A, Stancanelli G, Tramontini S. Susceptibility of Olea europaea L. varieties to *Xylella fastidiosa* subsp. pauca ST53: systematic literature search up to 24 March 2017. EFSA J. 2017;15(4):e04772.

55. Almeida RPP, Purcell AH. Transmission of *Xylella fastidiosa* to grapevines by *Homalodisca coagulata* (Hemiptera: Cicadellidae). J Econ Entomol. 2003;96(2):264–71.

56. Redak RA, Purcell AH, Lopes JRS, Blua MJ, Mizell Iii RF, Andersen PC. The biology of xylem fluid–feeding insect vectors of *Xylella fastidiosa* and their relation to disease epidemiology. Annu Rev Entomol. 2004;49(1):243–70.

57. De Jong Y. Fauna europaea version 2.6. Web Serv available online http//www faunaeur org. 2013;

58. Purcell AH. Control of the blue-green sharpshooter and effects on the spread of Pierce's disease of grapevines. J Econ Entomol. 1979;72(6):887–92.

59. Chang CJ, Shih HT, Su CC, Jan FJ. Diseases of Important Crops, a Review of the Causal Fastidious Prokaryotes and Their Insect Vectors. 2012;

60. EFSA, Health EP on P. Scientific Opinion on the risks to plant health posed by *Xylella fastidiosa* in the EU territory, with the identification and evaluation of risk reduction options. EFSA J. 2015;13(1):3989.

61. Lopes JRS, Landa BB, Fereres A. Short communication. A survey of potential insect vectors of the plant pathogenic bacterium *Xylella fastidiosa* in three regions of Spain. Spanish J Agric Res. 2014;12(3):795.

62. Saponari M, Boscia D, Loconsole G, Palmisano F, Savino V, Potere O, et al. New hosts of *Xylella fastidiosa* strain CoDiRO in Apulia. J Plant Pathol. 2014;96(3).

63. Cornara D, Saponari M, Zeilinger AR, de Stradis A, Boscia D, Loconsole G, et al. Spittlebugs as vectors of *Xylella fastidiosa* in olive orchards in Italy. J Pest Sci [Internet]. 2017;90(2):521–30. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28275326

64. Pautasso M, Petter F, Rortais A, Roy A-S. Emerging risks to plant health: a European perspective. Health (Irvine Calif). 2015;50:52.

65. Boscia D, Potere O, Loconsole G, Saponari M, Delle Donne A, Susca L, et al. The possible role of oleander in the epidemiology of *Xylella fastidiosa* in the Salento peninsula. In: Prooceedings "International Symposium on the European Outbreak of *Xylella fastidiosa* in Olive", Gallipoli–Locorotondo, Italy. 2014.

66. Purcell A. Paradigms: examples from the bacterium *Xylella fastidiosa* . Annu Rev Phytopathol [Internet]. 2013;51:339–56. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23682911

67. Nunney L, Vickerman DB, Bromley RE, Russell SA, Hartman JR, Morano LD, et al. Recent evolutionary radiation and host plant specialization in the *Xylella fastidiosa* subspecies native to the United States. Appl Environ Microbiol. 2013;79(7):2189–200.

68. Almeida RPP, Blua MJ, Lopes JRS, Purcell AH. Vector transmission of *Xylella fastidiosa* : applying fundamental knowledge to generate disease management strategies. Ann Entomol Soc Am. 2005;98(6):775–86.

69. Freitag JH. Host range of the Pierce's disease virus of grapes as determined by insect transmission. Phytopathology. 1951;41(10).

70. Hill BL, Purcell AH. Multiplication and movement of *Xylella fastidiosa* within grapevine and four other plants. Phytopathology. 1995;85(11):1368–72.

71. Zimmermann MH. Xylem structure and the ascent of sap. Springer Science & Business Media; 2013.

72. Roper MC, Greve LC, Labavitch JM, Kirkpatrick BC. Detection and visualization of an exopolysaccharide produced by *Xylella fastidiosa* in vitro and in planta. Appl Environ Microbiol. 2007;73(22):7252–8.

73. Killiny N, Martinez RH, Dumenyo CK, Cooksey DA, Almeida RPP. The exopolysaccharide of *Xylella fastidiosa* is essential for biofilm formation, plant virulence, and vector transmission. Mol Plant-Microbe Interact. 2013;26(9):1044–53.

74. Ramey BE, Koutsoudis M, von Bodman SB, Fuqua C. Biofilm formation in plant–microbe associations. Curr Opin Microbiol. 2004;7(6):602–9.

75. Costerton JW, Lewandowski Z, Caldwell DE, Korber DR, Lappin-Scott HM. Microbial biofilms. Annu Rev Microbiol. 1995;49(1):711–45.

76. Feil H, Feil WS, Lindow SE. Contribution of fimbrial and afimbrial adhesins of *Xylella fastidiosa* to attachment to surfaces and virulence to grape. Phytopathology. 2007;97(3):318–24.

77. Lambais MR, Goldman MHS, Camargo LEA, Goldman GH. A genomic approach to the understanding of *Xylella fastidiosa* pathogenicity. Curr Opin Microbiol. 2000;3(5):459–62.

78. Chatterjee S, Killiny N, Almeida RPP, Lindow SE. Role of cyclic di-GMP in *Xylella fastidiosa* biofilm formation, plant virulence, and insect transmission. Mol Plant-Microbe Interact. 2010;23(10):1356–63.

4.9 Article 3

A new inclusive MLVA assay to investigate the genetic variability of *Xylella fastidiosa* with a specific focus on the Italian outbreak.

Angelo Mazzaglia[1], Yaseen Jundi Rahi[1,3], Maria Claudia Taratufolo[1], Marta Tatì[1], Silvia Turco[1], Serena Ciarroni[2], Vincenzo Tagliavento[2], Franco Valentini[3], Anna Maria D'Onghia[3], Giorgio Mariano Balestra[1]

[1]Università della Tuscia, Dipartimento di Scienze Agrarie e Forestali (DAFNE), Via S. Camillo de Lellis snc, 01100 Viterbo, Italia;
[2]Phytoparasite Diagnostics s.r.l., Via S. Camillo de Lellis snc, 01100 Viterbo, Italia;
[3]CIHEAM-Mediterranean Agronomic Institute of Bari, Via Ceglie 9, 70010 Valenzano (Ba), Italy,

Corresponding author: *angmazza@unitus.it*

Abstract

Olive Quick Decline Syndrome (OQDS) by *Xylella fastidiosa* subspecies *pauca* is among the most severe phytopathological emergencies nowadays. In a few years, the Italian outbreak devastated olive groves in Apulia, potentially endangering the entire Mediterranean basin. This research aimed to develop a Multiple Locus VNTR Analysis (MLVA) assay, a molecular tool to distinguish populations of the pathogen. It has been already applied successfully to different *Xylella fastidiosa* subspecies from various plant hosts. Thus, the formerly published TR loci, together with a set of new identification, have been tested *in silico* on the genome of the DeDonno strain. The resulting selection of 37 TR loci was amplified on the genomic DNAs of strains from the Italian outbreak, from representatives of *Xylella fastidiosa* subspecies, and directly on DNA extracted from infected plants. The assay clearly discerned among subspecies or even sequence types (ST), but it also identified variants within the same ST so as to be suitable to provide more detailed information about dynamics and diffusion pathways of the pathogen. Its effective application even on total DNAs extracted from infected tissues of several host plants makes it particularly useful for large screening purposes and for the strengthening of containment measures.

Introduction

Xylella fastidiosa is a Gram-negative phytopathogenic bacterium belonging to the *Xanthomonadaceae* family, which is able to infect and cause diseases on more than 500 plant species. It is transmitted by

several xylem-fluid feeding insect species, especially sharpshooter, froghoppers (Hemiptera: Cicadellidae) and spittlebugs (Hemiptera: Cercopidea)[1]. The bacterium grows in the xylem of the host, where actively multiplies and forms a biofilm slowly occluding xylem vessels, thus causing consequently water stress and nutritional deficiencies[2]. *X. fastidiosa* causes a broad range of symptoms according to the infected host. The first disease attributable to this pathogen was Pierce's disease (PD) on grapevine, but then many others, such as Citrus Variegated Chlorosis (CVC) on citrus[3], leaf scorch diseases on almond, oleander (OLS), coffee (CLS), plum (PLS), pecan, mulberry, pear and other diseases of crops, ornamentals and woody plants[4] were progressively reported. Currently, relying on DNA-DNA hybridization and MLST data[1], six *X. fastidiosa* subspecies have been described, each one specific to a particular range of host plants and a native zone[5–7]. These are the subspecies *fastidiosa*, the subsp. *sandyi* which cause PD and OLS respectively, supposed to have been introduced into the USA from Central America, the subsp. *multiplex*, associated with scorch diseases of a large range of trees in North America and the subsp. *pauca* mostly found in South America on citrus and coffee[8, 9]. Two additional subspecies have been proposed: the subsp. *morus*, which includes isolates infecting mulberry[10], and the subsp. *tashke*, from *Chitalpa tashkentensis* in southwestern USA[11]. The evolutionary history and the geographical distribution of *X. fastidiosa* subspecies in the Americas indicate that the different subspecies evolved in reciprocal geographic isolation, the *X. fastidiosa* subsp. *multiplex* and subsp. *sandyi* in North America, and the subsp. *fastidiosa* and subsp. *pauca* in Central and South America[1], but then the anthropogenic activities have introduced it into new areas. Indeed, in 2013 the pathogen was reported in Taiwan on grapevine[12], in Italy on olive[13], in 2014 in Iran on grapevine and almond[14], in 2015 in France on *Polygala myrtifolia*, in 2016 in the Balearic Islands[9, 15]. Recently, the presence of *X. fastidiosa* subsp. *multiplex* was reported in Tuscany, Italy, in 2018[16] and in Porto, Portugal, in 2019 on ornamental plants[17]. All these new outbreaks raised up the need to accurately estimate the changes of genetic features

to understand the dynamics and evolutionary process of populations, as well as the adaptation to different hosts and environments. In response to this, several molecular technologies have been tested on *Xylella fastidiosa* . These include non-sequence-based methods, such as (RFLP), (RAPD), and (AFLP), but with some limitations related to scarce reproducibility and potential homoplasy of alleles. Then, sequenced-based methods targeting specific regions such as 16S rDNA or 16S-23S internal transcribed spacer (ITS) were applied, with some success, as the new subspecies *tashke* has been identified and proposed by Randall and colleagues in 2009[11] using these approaches. The introduction of foreign *X. fastidiosa* strains in new geographical areas and subsequent recombination with endemic strains have been demonstrated to be relevant in increasing the genetic variability, shifting the target host and thus, inducing new crop diseases[18]. In this regard, multilocus sequence typing (MLST), based on the identification of nucleotide sequence differences in seven housekeeping genes, has been applied to study the evolution of *X. fastidiosa* and its subspecies[5, 19]. Using this method, it was proposed that the new subsp. *morus* originated by intersubspefic homologous recombination from *X. fastidiosa* subsp. *fastidiosa* and *X. fastidiosa* subsp. *multiplex*[20]. Likewise, MLST analysis of *X. fastidiosa* subsp. *pauca* isolated from coffee plants in Costa Rica and subsequently from olive trees in Italy was referred to the ST53 and provided important information about the origin of the outbreak[18, 21]. However, despite being a powerful tool, MLST has some limits: relying on only seven core genome genes could be feeble to distinguish very closely related strains and, moreover, it is not suitable for large scale and routine monitoring because of the costs of sequencing[4, 9, 22, 23]. Then, the whole genome sequencing is clearly the most informative approach but, although extremely powerful, this analysis requires highly skilled personnel, is time-consuming and is the most expensive, making it poorly suitable for fast and large surveys[24]. Instead, tools based on markers with an adequate discrimination power like Simple-Sequence Repeat (SSR), also known as Variable-Number Tandem Repeat (VNTR), 27,28 are a good compromise to analyse genetically homogeneous bacteria. The analysis of VNTR loci is the basis for MLVA (Multiple-Locus VNTR Analysis), where the

number of repeats can be determined by PCR amplification using primers complementary to the well-conserved sequences flanking the tandem repeats. This method is rapid, easy to perform, inexpensive and highly reproducible[24, 25]. Indeed, the MLVA analysis has been often adopted by microbiologists to study the population structure of several human and animal bacterial pathogens, such as *Mycobacterium tuberculosis*, *Yersinia pestis*, *Staphylococcus aureus*, *Salmonella typhimurium*, and *Mycobacterium bovis*[26]. Among plant pathogens, several studies have already been accomplished using SSR markers for *X. fastidiosa* genotyping. Della Coletta Filho and colleagues in 2001[27], described the efficacy of a set of 9 SSR markers in comparison with RAPD; then, in 2005, Lin and colleagues[28] used a genome-wide approach to identify a new set of SSR to evaluate genetic divergence between *X. fastidiosa* isolates from different host species. A similar approach has been used by Montero-Astùa and colleagues in 2007[29] on different plant species from Costa Rica, Brazil, and USA in combination with other techniques to understand the relationship between the strains. Della Coletta-Filho and colleagues in 2014[30] used MLVA to provide information on the genetic diversity of populations in sweet orange, as well as the consequences of vector transmission of *X. fastidiosa* on their structure. Another successful SSR genotyping has been proposed to analyse the seasonal and annual variation in genetic diversity of this bacterium in two almond orchards in California[31]. In the same year, a combined approach SNP-based assay and multilocus SSR markers were attempted to assess the genetic diversity of *X. fastidiosa* subsp. *pauca* infecting citrus and coffee[32]. More recently, the MLVA was also used to demonstrate that *X. fastidiosa* subsp. *pauca* populations from coffee have higher genetic diversity and allelic richness compared with those from citrus[33]. These researches represented the starting point for this study, focused on *X. fastidiosa* subsp. *pauca*, which gained enormous attention after its report on olive in Italy in 2013[34] and its new pathogenic expression was described as Olive Quick Decline Syndrome (OQDS)[35]. It represents an exemplary model of the introduction of an exotic pathogen in an area where a cultivated species, not

coevolved with the pathogen, proves to be defenseless and extremely susceptible to its attack. This, together with the transmission by not host-specific xylem-feeding insects and the difficulty to control a pathogen living within the vascular system of the infected plant, has led to an impressive spread of the disease within few years. Despite massive efforts in containment measures, nowadays thousands of hectares of olive groves in Apulia are harshly affected by the syndrome and the spread is still ongoing. Again, in this situation, a fine-tuned genotyping of the strains responsible for the outbreak is crucial to understand where the disease comes from, how it moves in the infected areas and to monitor if new variants would appear in this scenario or if the type present undergoes to evolutionary forces that can lead to new variants. Today is acknowledged that the strains infecting olive trees, but also other plant species, in Italy belong to the sequence type ST53, and the most plausible origin of CoDiRO (Complesso del Disseccamento Rapido dell' Olivo), the Italian acronym for the disease, refers to strains infecting imported coffee plants (as ornamentals) from Costa Rica[21, 34–37]. It is worth noting that the same ST53 was retrieved in imported plants in both France and Netherlands[7, 9, 36]. Thus, we firstly aimed to check-*in silico* the presence of the VNTR loci reported in literature within the completely edited genome of the DeDonno strain (accession n° CP020870). Then, a further search for new VNTR loci was independently conducted on the same genome, aiming to obtain a final selection of markers to be used in a novel, inclusive MLVA assay capable to generate novel and deeper information about the genetic diversity of this subspecies, with specific reference to the Italian outbreak.

Results

In silico analysis of VNTR loci from literature

The *in-silico* check of the 50 TRs and related primers reported in the literature (S1) on the genome of the DeDonno strain of *X. fastidiosa* subsp. *pauca* evidenced several inconsistencies. First of all, numerous markers resulted to be the same, even if reported in the papers with different names and amplified with different primers; In some cases, the tandem repeat is also reported as reverse and opposite sequence. In details, the SSR20

marker[27] is the same as COSS1 marker[33], the SSR28 marker[27] is the same as marker ASSR-14[28], the SSR30 marker[27] is the same as marker OSSR-19[28], the SSR32 marker[27] is the same as COSSR6 marker[33], the CSSR-17 marker[28] is the same as COSSR3 marker[33], the marker OSSR-9[28] is the same as marker ASSR-20[28], the OSSR-14 marker[28] is the same as CSSR45 marker[33], the OSSR-16 marker[28] is the same as CSSR-20 marker[28], the OSSR-17 marker[28] is the same as to the CSSR-7 marker[28], the CSSR-18 marker[28] is the same as GSSR-6 marker[28], and the marker GSSR-12[28] is the same as marker CSSR42[33]. Thus, in these cases, only one pair of primers was chosen for the amplification, i.e. the ones whose sequences best-fitted *X. fastidiosa* subsp. *pauca* strain DeDonno (loci marked with * in Table S1). Besides these duplications, several additional anomalies have been found in comparison to the DeDonno genome. Specifically for markers from the study of Della Coletta-Filho *et al*. 2001[27], it was not possible to detect the reverse primer for the SSR26 marker (in red in Table S1), the SSR32 marker contains 2 different tandem repeats, both of 8 bp (the reported one is in green) and one SNP in the forward primer, for SSR36 and SSR40 markers, even if the respective primers were detected with few differences, for both the TR sequence was instead not retrievable; finally any of the two primers for the amplification of SSR34 marker was not found. Regarding the markers described in Lin *et al*. 2005[28] (Table S1), it was not possible to find the forward primer of the OSSR-12 marker, the OSSR-19 marker has an SNP in the reverse primer, only a single TR was found for the markers CSSR-4, CSSR-6, GSSR-14, GSSR-15, GSSR-19, GSSR-20, the TR is absent for the markers CSSR-12 and CSSR-13, the primers of the CSSR-16 marker show multiple annealing sequences, a different TR sequence, and a SNP were found in the reverse primer for the marker ASSR-16, one SNP was found in the forward and one in the reverse primer of both the markers ASSR-19 and GSSR-4. No anomalies were instead found in the 7 VNTR loci described in the paper of Francisco *et al*. 2017[33]. After this check, several markers were discarded according to duplication, failure in the detection of the primer sequence or of the repeat sequence. Also, all the differences detected between the primer sequences and the corresponding pairing sequences on the DeDonno genome were accordingly corrected for the synthesis of the primers for this study, as reported in Table 1.

Table 1. Primer pairs for the amplification of the selected 37 VNTR loci and sequence of the respective tandem repeats. TRs 1-3 were obtained from Della Coletta-Filho et al., 2001; TRs 4-18 were obtained from Lin et al., 2005; TRs 19-23 were obtained from Francisco et al., 2015; TRs 24-37 are from this study.

TR	Locus	Forward primer	Reverse primer	TR sequence
1	SSR20	ATGAAGAAGCCAGGATACAT	GCTACACGTGCAACAAC	ATTGCTG
2	SSR21	AACACGGATCAAGCTCATG	GGAACACGCAATAGTAAGA	TGTTATC
3	SSR28	GTAACGCTGTTATCTCAAT	ATTACGCTTCTTATCGCTGT	GTGTGCCT
4	OSSR-9	TAGGAATCGTGTTCAAACTG	TTACTATCGGCAGCAGAC	TTTCCGT
5	OSSR-16	GCAAATAGCATGTACGAC	GTGTTGTGTATGTGTTGG	CTGCTA
6	OSSR-19	GCTGTGAACTTCCATCAATCC	GCAAGTAGGGGTAAATATGAC	CAGGATCA
7	OSSR-20	ATCTGTGCGGCGGTTCTG	CACTTGCGGCGTAGATACTTC	AGGATGCTA
8	CSSR-7	CACAGCGAACAGGCATTG	AGCAACCAAGACGGGAAC	CTGTGC
9	CSSR-10	GCAACCACAAAGCCGCAG	AGCACCTCTTAGCATCACTGG	CAATGA
10	CSSR-18	GTGCTTCCAGAAGTTGTG	GACTGTTCTCTTCGTTCAG	GCCAA
11	CSSR-19	TGCTGTGATTGGAGTTTTGC	TCAAACGAATCTGTCCATCAAG	TGGTGAG
12	ASSR-9	GGTTGTCGGGCTCATTCC	TTGTCACAGCATCACTATTCTC	CAAGTAC
13	ASSR-11	AGAGGCAACGCAGGAACAG	GTGAGTTATATCGGTGCAGCAG	ACGCATC
14	ASSR-12	TGCTCATTGTGGCGAAGG	CGCAACGTGCATTCATCG	GATTCAG
15	ASSR-16	TTAATCAACAACGCTTATCC	TCGCAGTAGCCAGTATGC	GCTCCA
16	ASSR-19	CGCCGACTGTCTATATGAC	TTCGTAGCAATGGCAATGTTG	ACAACG
17	GSSR_4	GCGTTACTGGCGACAAGC	GCTCGT(C)TCCTGACCTGTG	ATCC
18	GSSR_7	ATCATGTCGTGTCGTTTC	CAATAAAGCACCGAATTAGC	GGCAAC
19	COSSR6	TGCTGCGCGATAACCAAGT	CATCCAATCAGCCCTAACCT	GTGATGCG
20	CSSR45	ACAGACATCACCGGCATTG	AATGTCGCTGCCAATCCAT	CACACCGAGATGGAC
21	COSSR4	CAAGGTGACCGCTAGCCTAT	GCTGTCATTGGGTGATGC	CAATACAC
22	COSSR5	ACACTGACACAACAGCCACCA	AATGGTGGGTGTGATGGTTTC	CATACAGA
23	CSSR42	ATTACGCTGATTGGCTGCAT	GTTTCATTACGCGGAACAC	TGTTATC
24	TR4	CATACGGCAGTTCTGTGTCG	CGGGCAAGCTTTTCCCACCC	CAGCGCAT
25	TR5	ATTCCAAGATTTGCGAGTGG	ACGATTCGAACATGGAGGTA	TTCTAG
26	TR6	ACATCGGAGGTAGGCTGTGA	ATTGAAGACCCTTTTCAGCC	CGCTTAT
27	TR7	GGGTTGGGTCTTTTATTTGC	CATTGACTCTCAACCCTGCTAC	GCTGT
28	TR8	GCGGTTTGGTTGTATTGCTT	CTCACATCACGCACCGACGA	GACAGG
29	TR9	GGTGTGCCGTGTACATTGAG	TTGCCATCACCGACACCTCT	ATGATCTGA
30	TR10	CGTGCTGAAGTCTTGCTTGA	ACTTCACCCTACCCTGCATA	GTAACG
31	TR12	AGGGATATAGTGCCGCGATT	TTTTGTGGTCGAACGTGCGG	GGTGTGA
32	TR15	ATGCAGCGGTAGTCCCTCTA	CACGATGCCCACGTAGCAGC	GTGTCG
33	TR18	TGTCATGACCGTGCTTATGG	TGGTGGTCAAGGCAGCGG	CCGCCGCCGTAACCACCG
34	TR19	CTGCCTTGACCACCACCAC	ACAAAGCTCTCTGATCAATCAC	CCACTCCAGCTG
35	TR21	CAGGGTGTATGGCCTGAAGT	CCTACCATCCATGCAGCAAC	CAGCACAT
36	TR23	CAGGAGCCTCCATGAACAAT	AATGATCCTTGCTGGGTCAG	CTTCAAGAG
37	TR24	ATGGCCCAAACATACTCCAA	TGTTCATATCTTGGTCTCAT	GTCCTG

New VNTR loci identification

The searching procedure by Tandem Repeat Finder (TRF; https://tandem.bu.edu/trf/trf.html) [38] has led to the identification of 25 VNTR loci, which have also undergone an *in silico* check to ascertain their position on the DeDonno genome and to verify any correspondence with loci selected from the literature. Again, this correspondence was found for 5 of them, which were consequently discarded. In the end, 45

total VNTR loci were selected (25 from the literature and 20 newly identified) to be used in the following experimental procedures.

PCR amplification of VNTR loci

The first round of PCR was done using only genomic DNAs from two strains, the DeDonno strain from olive and the strain V104 isolated from oleander, to validate the efficacy of the entire MLVA assay. The primer pairs related to newly identified VNTR loci TR3, TR11, TR14, TR16, TR17, and TR25 produced multiple amplicons, indicating the presence of multiple pairing sites for at least one of the two primers and were consequently discarded. Also, the SSR40 locus [27] invariably produced a 133 bp amplicon, corresponding exactly to the sum of the only flanking regions, thus indicating the absence of the tandem repeat, whereas the OSSR-2 locus[28] invariably produced a 181 bp amplicon, smaller than the sum of the flanking regions; both were discarded. Table S1 shows the list of 37 loci and the related final primers used in the amplification step. According to the amplicon sizes, as obtained by capillary electrophoresis, the number of repeats per each locus and per each strain was calculated. Amplification failure was eventually coded as "0". Thus, the haplotype of each individual was defined as the ordered sequence of 37 numbers, as reported in Table S2. Here, it has to be noticed that the amplification of two DNAs obtained from infected tissues of *Prunus dulcis* and *Polygala myrtifolia* in the province of Lecce (Pd_Le2 and Pm_Le10) provided multiple bands in some loci. Due to these anomalies, which cannot be coded properly as input data, the two samples were excluded from the following analysis.

Data analysis

As already introduced, the haplotypes obtained by the amplification of 37 VNTR loci on a total of 51 DNAs are reported in Table S2, that includes the genomic DNAs extracted from 15 strains of the CFBP collection, 9 strains isolated in Apulia and 27 total DNAs as extracted from infected plant samples. In the last 2 lines of the Table S2 are also reported the multiple repeats identified in Pd_Le2 and Pm_Le10, which

were not further considered. We expected that the very high number of loci here analyzed would have allowed to point out even minimal differences between isolates. Indeed, among all the individuals, only 2 haplotypes, obtained from samples of infected *Rhamnus alaternus* plants in the same geographic location, resulted completely identical to each other. The first data analysis was carried out on data concerning only the 15 strains from CFBP collection by hierarchical clustering using Bruvo's distance (Figure 1) to check the effectiveness of the method to assess genetic differences between subspecies and sequence types STs. Noticeably, strains of the same subspecies but belonging to different STs are separate by higher levels of genetic distance than those measured between strains belonging to the same sequence type, i.e. the two strains of subspecies *pauca* belonging to ST74 and two, out of three, belonging to ST53.

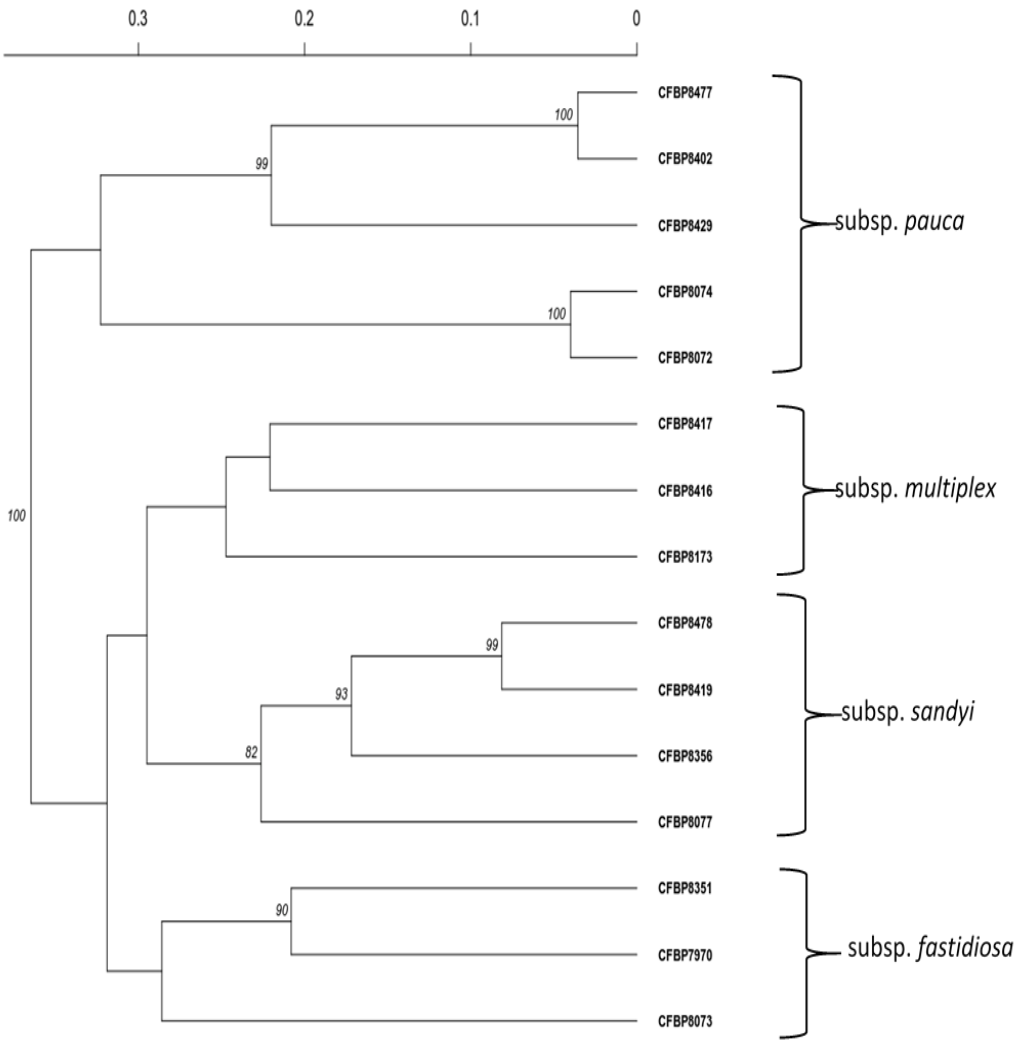


Figure 1: shows the taxonomic position of *Xylella fastidiosa* isolates from worldwide collections. hierarchical clustering using Bruvo's distance. Bootstrap scores (1000 replicates) are displayed at each node.

Then, according to the focus of the paper, the results of VNTR loci amplification of DNAs extracted from strains isolated in Apulia and from whole DNAs extracted from infected tissues of different host plants in the same region were added to the analysis. The hierarchical clustering of all the individuals (Figure 2) maintains the same structure of the previous dendrogram, with the addition of a large cluster that includes all the Italian strains. It has to be noticed that the strain CFBP8429, isolated in 2015 from *Coffea arabica* plants in Angers, France, and reported in the CFBP database as belonging to the ST53 of subspecies *pauca*, resulted significantly distant from all the others ST53 from Italy. Then, the 38 DNAs of Italian origin were analysed independently to better appreciate their relationships, provided that their reciprocal genetic variability decreased considerably. Indeed, an identical number of repetitions was obtained in as many as 18 loci out of 37 (48%), and, among the remaining 19 loci, 11 changed sporadically in few samples, whilst only 7 loci (TR7, TR8, TR12, OSSR-16, OSSR19, ASSR-16, and COSSR-4) showed frequent variations in the numbers of repeats.

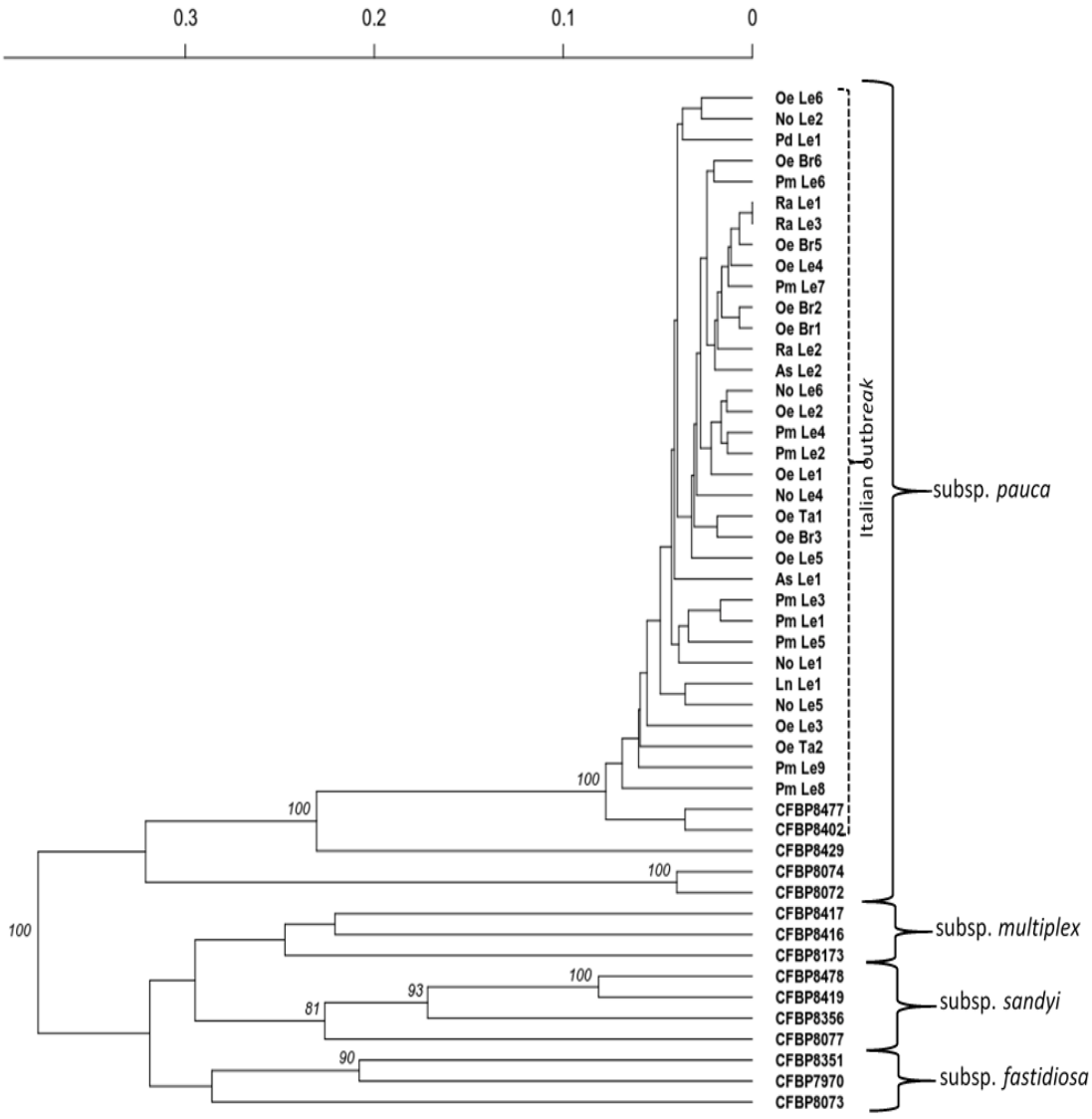


Figure 2: Shows the taxonomic position of *Xylella fastidiosa* isolates from the Italian outbreak. Hierarchical clustering using Bruvo's distance. Bootstrap scores (1000 replicates) are displayed at each node.

The Minimum Spanning Trees in Figure 3 illustrate how these samples are best linked to each other. However, it wasn't possible to appreciate in their grouping any relationship with the species of host plants from which they were obtained (Figure 3a), nor with their specific geographic origin (Figure 3b).

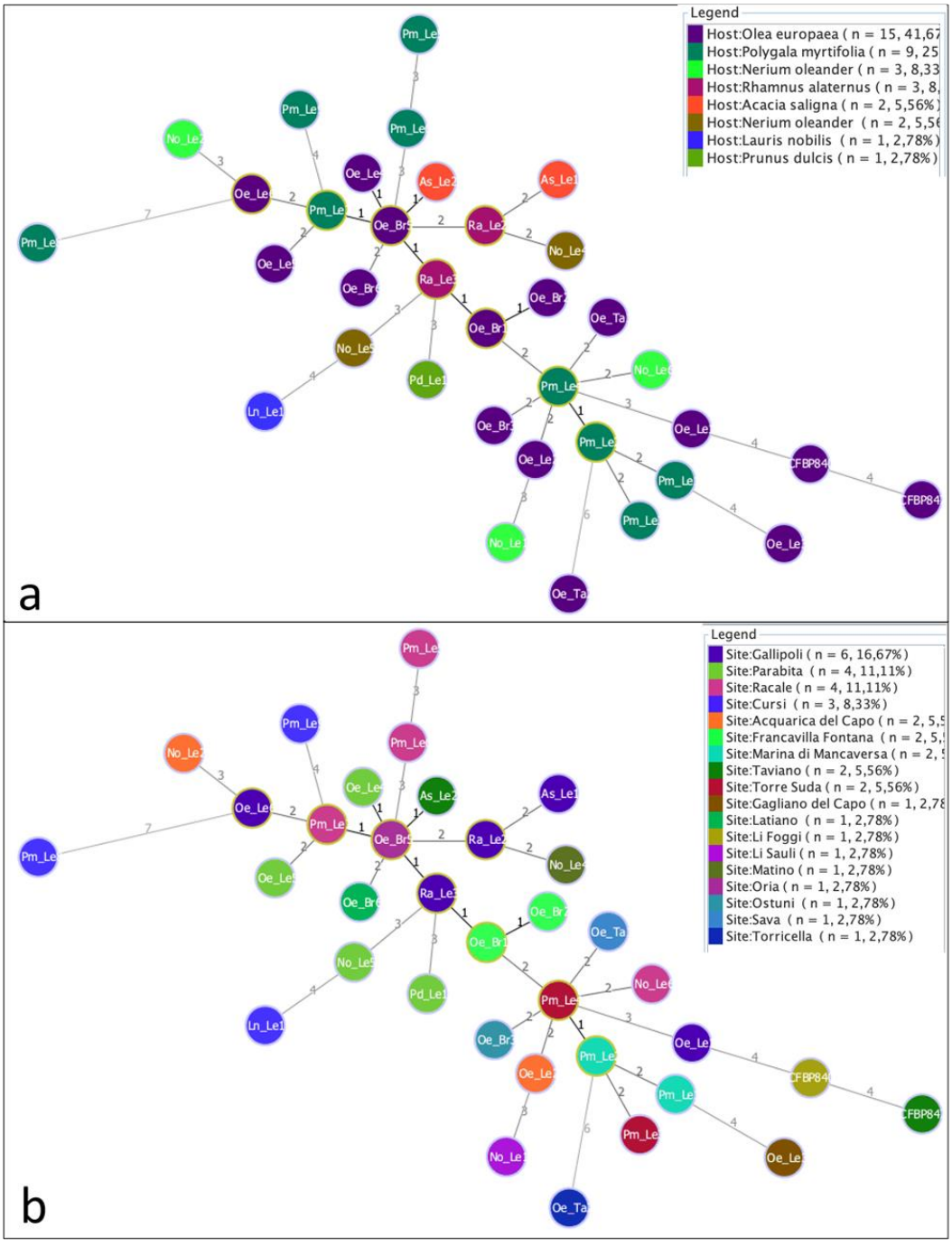


Figure 3: Minimum spanning tree representing the genetic diversity: a) Worldwide strains collection of *Xylella fastidiosa* in relation to its host represented in colors. B) Italian outbreak in the Apulia region representing 18 isolating sites in different colours.

Discussion

Xylella fastidiosa is one of the most feared bacterial plant diseases nowadays. There are several biological features that make its containment and control hardly challenging, first and foremost its ability to colonize an astonishing number of plant hosts, often without visible symptoms, but also the difficulty in its isolation and detection. A crucial issue to understand the spread dynamics is the genotyping of the pathogen, which can be carried out by different molecular techniques. The method of choice is the MLST (MultiLocus Sequence Typing), based on the SNPs in seven constitutive genes[39]. According to this method, the six subspecies of *Xylella fastidiosa* are further categorized in 86 sequence types (ST)[7, 18, 19, 22, 40]. In this categorization, all the strains to date analysed of the Italian outbreak are assigned to the genotype ST53 of *X. fastidiosa* subsp. *pauca*[20, 21, 41]. However, because MLST relies on Sanger sequencing, which in turn implies the isolation of pure colonies, costs and time to obtain the results can be a limit, especially when large numbers of samples are to be analysed[4, 9, 22, 23]. To solve these issues, the potential of a new inclusive MLVA assay was here evaluated. Indeed, several MLVA studies have been already carried out to characterize *X. fastidiosa* isolated from different plants such as grapevine[42], orange[30], almond tree[31] and coffee[33]. Only recently, a study on *X. fastidiosa* subsp. *pauca* strains from olive trees in Brazil were conducted using a panel of 12 SSRs[40]. Therefore, to our knowledge, this is the first study that applies this approach to *X. fastidiosa* isolates from the Italian outbreak. It's worth emphasizing the importance of a preliminary *in silico* analysis of markers when obtained from multiple literature sources. Indeed, the screening accomplished here revealed that several loci from independent studies were, in the end, the same. A first arising indication is that, even if many repeats can be found in a bacterial genome, then the effective loci with proper features for genotyping are not so frequent. Also, the need for a clearer classification of markers to avoid misunderstanding and confusion in their use appears to be urgent. As an example, in the last paper of Safady and colleagues[40] 12 markers from the

literature have been used[28, 33], but 2 of them, i.e. the loci CSSR42 and GSSR12, refer to the same repetition, obviously giving the same, doubled, result. This kind of mistake can negatively affect results and possibly lead to incorrect conclusions. Nevertheless, due to the high number of starting markers considered here, 50 from literature and 25 of new identification, the *in silico* screening and a first round of PCR testing have led to a still conspicuous set of 37 TR loci to be evaluated for their effectiveness in revealing genetic differences among *X. fastidiosa* strains, with a specific attention at the genetic characterization of the subspecies *pauca*. Concerning the results, of course, this MLVA assay is not primarily aimed to resolve phylogeny between subspecies of *Xylella fastidiosa* , due to its intrinsic sensitivity in appreciating minimal differences between individuals. Despite this, it still resulted in proficient in depicting a correct clustering for them. Conversely, this methodology seems to be finer than MLST analysis, due to the sharp distinction among STs evidenced in the results. Interestingly, the strain CFBP 8429, isolated from a coffee plant intercepted in Angers, France in 2015, and belonging to subsp. *pauca* ST53 (as reported in CFBP database), has significant differences in the number of repetitions compared to the other samples of the same subspecies and type ST53, that led to an independent positioning in both hierarchical clustering and MST results. Since this discrepancy could be related to a misidentification of its ST, a larger comparison with isolates of the same subspecies from other geographic regions should be carried out to validate this hypothesis. The discriminatory power of this MLVA assay goes beyond ST detection, being capable to ascertain differences even within single ST. This makes its usage crucial in studying cases like the Italian outbreak, where a highly clonal population of *X. fastidiosa* subsp. *pauca*, belonging to the only ST53 and whose origin is likely attributable to a single infection event, is under investigation. In our data, such clonality is widely confirmed and the information capable to distinguish individuals are relegated to only a few loci. Those VNTR loci could be hereafter selected to put on a devoted assay effective in resolving close genetic relationships in comparison to other genotyping methods. However, the small differences ascertained in this analysis don’t show evidence for specific relationships with the species of the host plant or with the geographic origin of the strains. Most

probably, they account for the first signal of casual mutability of the *Xylella* population in Italy, do not reflecting the consequences of any evolutionary pressure, even though being compatible with the presence of the pathogen in Apulia since about 6 years. Then, in terms of convenience in comparison with the MLST approach, MLVA has some advantages. The first one is that any sequencing procedure is required to obtain genetic fingerprints of bacterial individuals, as instead required by MLST. Moreover, even if the cost and the skill requirements of personnel in performing Sanger sequencing are constantly decreasing, at least for the present MLVA still remains easier to perform, more cost-effective and less time-consuming. Since *Xylella fastidiosa* is very tedious and time-consuming to isolate and culturing, another significant advantage of our MLVA assay regards the possibility to screen directly the DNA extracted from the infected plant material without losing reliability. Furthermore, we hypothesize that this method could diagnose infections by multiple genotypes in the same plant tissue, as likely occurred in the two samples where in some loci, amplicons corresponding to different numbers of repetitions were detected at the same locus, underlining another stimulating benefit of the assay. In conclusion, all the results seem to indicate that this novel MLVA essay has the potentiality to become a reference method for detailed monitoring of the Italian outbreak of *Xylella fastidiosa* subsp. *pauca*, as well as for potentially any other occurrence of *Xylella fastidiosa* epidemics.

Material and methods

All the 24 strains analyzed in this study are reported in Table 2, where subspecies, host plant, geographic origin, time of isolation, and ST classification are also indicated. Fifteen strains (indicated with §) were sourced from the CIRM-CFBP (Collection Française de Bactéries associées aux Plantes). These strains were grown on Buffered Charcoal-Yeast Extract (BCYE) medium at 25˚C for 3-4 weeks; then 100 mg of bacterial cells were collected, and DNA was extracted with the Nucleospin Plant kit (Macherey Nagel) according to the manufacturer's instructions. Similarly, 9 strains (indicated with *), were isolated from different host plants in Apulia and their genomic DNAs were extracted from freshly grown strains at the

Mediterranean Agronomic Institute of Bari (IAMB). The remaining 27 samples (indicated with °) are instead constituted by total DNAs extracted straight from tissues of plants whose infection by *X. fastidiosa* was formerly assessed. These include DNAs from different host plants in various locations in Apulia, i.e. in the provinces of Lecce, Taranto, and Brindisi. All DNAs were checked by q-PCR to confirm their belonging to *X. fastidiosa* according to the protocol described in Harper et al., 2010[43].

Table 2: Bacterial strains of *Xylella fastidiosa* subspecies used in this study

Sample	Subspecies	Host	Country(Region)	Province	Year	ST
CFBP8073§	*fastidiosa/sandyi*	*Coffea canephora*	Mexico	Mexico	2012	ST75
CFBP7970§	*fastidiosa*	*Grapevine*	USA(Florida)	Florida	1987	ST2
CFBP8351§	*fastidiosa*	*Vitis vinifera L.*	USA(California)	Fresno	1993	-
CFBP8077§	*sandyi*	*Nerium oleander*	USA(California)	Orange	1995	ST5
CFBP8356§	*sandyi*	*Coffea arabica*	France	(intercepted)	2015	ST72
CFBP8419§	*sandyi*	*Coffea arabica*	Costarica	Costarica	2015	-
CFBP8478§	*sandyi*	*Coffea arabica*	France	(intercepted)	2015	-
CFBP8173§	*multiplex*	*Prunus sp.*	USA(Georgia)	Georgia	1983	ST41
CFBP8416§	*multiplex*	*Polygala myrtifolia*	France(Corsica)	Propriano	2015	ST7
CFBP8417§	*multiplex*	*Spartium junceum*	France(Corsica)	Alata	2015	ST6
CFBP8429§	*pauca*	*Coffea arabica*	France(Loira)	Angers	2015	ST53
CFBP8072§	*pauca*	*Coffea arabica*	Equador	Equador	2012	ST74
CFBP8074§	*pauca*	*Coffea arabica*	Equador	Equador	2012	ST74
CFBP8402§	*pauca*	*Olea europaea*	Italy(Apulia)	Gallipoli	2014	ST53
CFBP8477§	*pauca*	*Olea europaea*	Italy(Apulia)	Taviano	2015	ST53
Oe_Le1*	*pauca*	*Olea europaea*	Italy(Apulia)	*Gallipoli*	2014	ST53
No_Le1*	*pauca*	*Nerium oleander*	Italy(Apulia)	*Gallipoli*	2016	ST53
Oe_Le2*	*pauca*	*Olea europaea*	Italy(Apulia)	Acquarica del Capo	2017	ST53
No_Le2*	*pauca*	*Nerium oleander*	Italy(Apulia)	Acquarica del Capo	2017	ST53
Oe_Le3*	*pauca*	*Olea europaea*	Italy(Apulia)	Gagliano del Capo	2017	ST53
Pm_Le1*	*pauca*	*Polygala myrtifolia*	Italy(Apulia)	Taviano	2017	ST53
Pm_Le2*	*pauca*	*Polygala myrtifolia*	Italy(Apulia)	Taviano	2017	ST53
Pm_Le3*	*pauca*	*Polygala myrtifolia*	Italy(Apulia)	Racale	2017	ST53
Pm_Le4*	*pauca*	*Polygala myrtifolia*	Italy(Apulia)	Racale	2017	ST53
Oe_Le4°	*pauca*	*Olea europaea*	Italy(Apulia)	Lecce	2018	ST53
Pm_Le5°	*pauca*	*Polygala myrtifolia*	Italy(Apulia)	Lecce	2018	ST53
Pd_Le1°	*pauca*	*Prunus dulcis*	Italy(Apulia)	Lecce	2018	ST53
Ra_Le3°	*pauca*	*Rhamnus alaternus*	Italy(Apulia)	Lecce	2018	ST53
Ra_Le1°	*pauca*	*Rhamnus alaternus*	Italy(Apulia)	Lecce	2018	ST53
Pm_Le6°	*pauca*	*Polygala myrtifolia*	Italy(Apulia)	Lecce	2018	ST53

No_Le6°	*pauca*	*Nerium oleander*	Italy(Apulia)	Lecce	2018	ST53
Oe_Le5°	*pauca*	*Olea europaea*	Italy(Apulia)	Lecce	2018	ST53
Oe_Br5°	*pauca*	*Olea europaea*	Italy(Apulia)	Brindisi	2018	ST53
Oe_Br1°	*pauca*	*Olea europaea*	Italy(Apulia)	Brindisi	2018	ST53
Oe_Br2°	*pauca*	*Olea europaea*	Italy(Apulia)	Brindisi	2018	ST53
Oe_Br6°	*pauca*	*Olea europaea*	Italy(Apulia)	Brindisi	2018	ST53
Oe_Br3°	*pauca*	*Olea europaea*	Italy(Apulia)	Brindisi	2018	ST53
Oe_Ta1°	*pauca*	*Olea europaea*	Italy(Apulia)	Taranto	2018	ST53
Oe_Ta2°	*pauca*	*Olea europaea*	Italy(Apulia)	Taranto	2018	ST53
Oe_Le6°	*pauca*	*Olea europaea*	Italy(Apulia)	Lecce	2018	ST53
As_Le1°	*pauca*	*Acacia saligna*	Italy(Apulia)	Lecce	2018	ST53
Ra_Le2°	*pauca*	*Rhamnus alaternus*	Italy(Apulia)	Lecce	2018	ST53
As_Le2°	*pauca*	*Acacia saligna*	Italy(Apulia)	Lecce	2018	ST53
Pm_Le7°	*pauca*	*Polygala myrtifolia*	Italy(Apulia)	Lecce	2018	ST53
No_Le4°	*pauca*	*Nerium oleander*	Italy(Apulia)	Lecce	2018	ST53
No_Le5°	*pauca*	*Nerium oleander*	Italy(Apulia)	Lecce	2018	ST53
Ln_Le1°	*pauca*	*Lauris nobilis*	Italy(Apulia)	Lecce	2018	ST53
Pm_Le8°	*pauca*	*Polygala myrtifolia*	Italy(Apulia)	Lecce	2018	ST53
Pm_Le9°	*pauca*	*Polygala myrtifolia*	Italy(Apulia)	Lecce	2018	ST53
Pd_Le2°	*pauca*	*Prunus dulcis*	Italy(Apulia)	Lecce	2018	ST53
Pm_Le10°	*pauca*	*Polygala myrtifolia*	Italy(Apulia)	Lecce	2018	ST53

In silico **analysis of Tandem Repeats previously reported on** ***Xylella fastidiosa***

Molecular typing of *X. fastidiosa* by VNTR markers has been already accomplished in the past with significant results. In detail, 9 Tandem Repeats loci have been described by Della Coletta-Filho et al. 2001[27], and used, in comparison with the RAPD method, to evaluate the genetic diversity of Brazilian strains of *X. fastidiosa* subsp. *pauca*, responsible for citrus variegated chlorosis (CVC) disease. Then, 34 VNTR markers have been independently depicted by Lin et al. 2005[28] to estimate genetic diversity among 43 isolates of *X. fastidiosa* collected mainly in California from grape, almond, citrus, and oleander, which resulted particularly effective in resolving differences within genetically homogeneous isolates. Also, to estimate differences in a large set of isolates from citrus and coffee in Brazil, Francisco et al. 2017[33] used 7 new markers. Successively, several studies have been realized, often using selections of the above-reported SSR markers, to genotype *X. fastidiosa* strains in different host-pathogen

interactions[30, 42]. Here, because of the aims of this research, we decided to firstly check-*in silico* the presence of all these 50 markers and their respective primers on the genome of the strain "DeDonno" of *X. fastidiosa* subsp. *pauca*, edited and deposited in GenBank (accession number: CP020870.1), using Nucleotide BLAST (for the tandem repeats) or Primer-BLAST (for the respective primers) tools in NCBI website.

New VNTR loci identification and primer design

Contemporaneously a new search, aimed to identify potential new markers to be added to the analysis, was carried out on the same genome using the TRF program[38] set with the following parameters: 2 match, 7 mismatch, 7 indels as alignment Parameters; 50 as Minimum Alignment Score; 250 as Maximum Period Size; Alignment Size. Then, a further selection of the results obtained was made imposing the following parameters : > 5 as Period Size, >2 as Copy Number; >90% as Percent Matches, Consensus Size as Period Size. This procedure has led to the identification of 25 new Tandem Repeats loci for which suitable primers in respective flanking regions have been designed using Primer3 with default parameters.

PCR amplification of VNTR loci

12.5 All VNTR loci were amplified with single PCR reactions using the primer pairs reported in Table 1; each reaction contained 12 μl of GoTaq G2 Green Master Mix 2x (Promega Corporation, USA), 1 μl of DNA sample (40 ng), 1 μl of primer forward and 1 μl of reverse primer (10 μM final concentration), 9.5 μl of molecular grade Sterile distilled water (SDW) to the final volume of 25 μl. The PCR amplifications were performed with a C1000 thermocycler (Biorad Laboratories Inc., Ca., USA). The following parameters were used for the TR loci identified in this study: initial denaturation for 5 minutes at 95 ˚C; 40 denaturation cycles for 30 s at 95 ˚C, annealing for 30 s at temperatures ranging from 47.9 ˚C to 58 ˚C according to primers requirements, and extension for 36 s at 72 ˚C, plus a final elongation step for 5 minutes at 72 ˚C. For the primers obtained from literature review[27, 28, 33] the respective protocols were

followed.

12.6

Capillary electrophoresis

The QIAxcel capillary electrophoresis system (QIAGEN, Milan, Italy) was used to estimate accurately the amplicon size. The DNA High-Resolution cartridge was used for all samples and the OM800 method was run, as recommended to obtain the maximum precision (2-3 bp maximum error) with amplicons ranging in size from 200 to 500 bp. No template controls (SDW) and size markers were included in each run. The results were analyzed and interpreted using the ScreenGel v.1.6.0 software (QIAGEN), which gives an accurate estimation of both the size and concentration of amplicons. Then, the number of tandem repeats at each VNTR locus was calculated by subtracting the size of the flanking region from the amplicon size and then dividing the remaining by the repeat unit length. In the case of loci with a truncated final repeat, the copy number was rounded down to the previous integer. To evaluate the accuracy of these calculations, a random selection of amplicons was Sanger sequenced, as done in any case of disputable TR number attribution. A final data matrix (Table S2) of 51 samples and 37 number of TR for each locus was produced.

Data elaboration

The string of integer numbers obtained as above described represented the haplotype of each strain under investigation. These were reciprocally compared using two analytical approaches suitable for this type of data when the assessment of genetic relationships among populations of clonal organisms is aimed. Hierarchical clustering and the goeBURST algorithm were independently applied to data. Data were analyzed in two steps: first, only the 15 strains from the CFBP collection were included to check the effectiveness of the method in resolving differences at subspecies and Sequence Type (ST) levels; then, all the other DNAs belonging to the Italian outbreak were included in the analysis.

Hierarchical clustering

The data matrix was imported into R version 3.4.4 (R Core Team, 2018). The genetic distance among individuals was calculated using Bruvo's distance[44], an algorithm particularly suitable for genetic markers as tandem repeats because it deems repeat length into the calculation and is not sensitive to ploidy levels[45–47]. After the distance computation, the hierarchical clustering was obtained by *hclust()* function of the R package stats (R Core Team, 2018) using UPGMA as an agglomerative algorithm. Bruvo's distance was also bootstrapped using the poppr *bruvo.boot()* function and a cut-off threshold of 80% was set. The final dendrogram was visualized with the R package factoextra version 1.0.5[48].

Globally optimized eBURST algorithm (goeBURST – Phyloviz)

For goeBURST analysis, the devoted software Phyloviz 2.0[49] was used. Another successful approach to the analysis of tandem repeat data has been proposed by Francisco and colleagues (2010)[50] in which the efficacy of eBURST algorithm[51] was reinforced with additional rules to elucidate possible patterns of evolution in a dataset. Then, a Minimum Spanning Tree (MST) was constructed representing a tree in which the sum of the distances among all the isolates, as represented by data, is the shortest possible one, according to the general concept that the most probable evolution is the one that requires the lowest modification events.

4.10 REFERENCES OF ARTICLE 3

1. Almeida, R. P. P. & Nunney, L. How do plant diseases caused by *Xylella fastidiosa* emerge? Plant Disease. DOI: 10.1094/ PDIS-02-15-0159-FE (2015).
2. Hopkins, D. L. *Xylella fastidiosa* : Xylem-limited bacterial pathogen of plants. *Annu. Rev. Phytopathol.* DOI: 10.1146/ annurev.py.27.090189.001415 (1989).
3. Purcell, A. Paradigms: examples from the bacterium *Xylella fastidiosa* . *Annu. Rev Phytopathol* **51**, 339–56, DOI: 10.1146/annurev-phyto-082712-102325 (2013).
4. Baldi, P. & La Porta, N. *Xylella fastidiosa* : Host range and advance in molecular identification techniques. *Front. Plant Sci.* **8**, 944–944, DOI: 10.3389/fpls.2017.00944 (2017).
5. Almeida, R. P. P. P. *et al.* Genetic structure and biology of *Xylella fastidiosa* strains causing disease in citrus and coffee in Brazil. *Appl. Environ. Microbiol.* **74**, 3690–3701, DOI: 10.1128/AEM.02388-07 (2008).
6. Harris, J. L. & Balci, Y. Population structure of the bacterial pathogen *Xylella fastidiosa* among street trees in Washington D.C. *PLoS One* **10**, e0121297, DOI: 10.1371/journal.pone.0121297 (2015).
7. Nunney, L. *et al.* Recent evolutionary radiation and host plant specialization in the *Xylella fastidiosa* subspecies native to the United States. *Appl. Environ. Microbiol.* **79**, 2189–2189, DOI: 10.1128/AEM.03208-12 (2013).
8. Schaad, N. W., Postnikova, E., Lacy, G., Fatmi, M. & Chang, C. J. *Xylella fastidiosa* subspecies: *X. fastidiosa* subsp. *fastidiosa* subsp. nov., *X. fastidiosa* subsp. *multiplex* subsp. nov., and *X. fastidiosa* subsp. *pauca* subsp. nov. *Syst Appl Microbiol* **27**, 290–300, DOI: 10.1078/0723-2020-00263 (2004).
9. Denancé, N. *et al.* Several subspecies and sequence types are associated with the emergence of *Xylella fastidiosa* in natural settings in France. *Plant Pathol.* **66**, 1054–1064, DOI: 10.1111/ppa.12695 (2017).
10. Giampetruzzi, A. *et al.* Genome-wide analysis provides evidence on the genetic relatedness of the emergent *Xylella fastidiosa* genotype in Italy to isolates from Central America. *Phytopathology* **107**, 816–827, DOI: 10.1094/ PHYTO-12-16-0420-R (2017).
11. Randall, J. J. *et al.* Genetic analysis of a novel *Xylella fastidiosa* subspecies found in the southwestern United States. *Appl. Environ. Microbiol.* **75**, 5631–5638, DOI: 10.1128/AEM.00609-09 (2009).
12. Su, C. *et al*. *Xylella taiwanensis* sp nov., causing pear leaf scorch disease. *Int. J. Syst. Evol. Microbiol.* **66**, 4766–4771, DOI: 10.1099/ijsem.0.001426 (2016).
13. Saponari, M., Boscia, D., Nigro, F. & Martelli, G. Identification of DNA sequences related to *Xylella fastidiosa* in oleander, almond and olive trees exhibiting leaf scorch symptoms in Apulia (southern Italy). *J. Plant Pathol.* **95**, 668–668 (2013).
14. Amanifar, N., Taghavi, M., Izadpanah, K. & Babaei, G. Isolation and pathogenicity of *Xylella fastidiosa* from grapevine and almond in Iran. *Phytopathol. Mediterr.* DOI: 10.14601/Phytopathol_Mediterr-12647 (2014).
15. Olmo, D. *et al.* First detection of *Xylella fastidiosa* infecting cherry (*Prunus avium*) and *Polygala myrtifolia* plants, in Mallorca Island, Spain. *Plant Dis.* **101**, 1820–1820, DOI: 10.1094/PDI5-04-17-0590-PDN (2017).
16. Marchi, G. *et al.* First detection of *Xylella fastidiosa* subsp. *multiplex* DNA in Tuscany - Italy. *Phytopathol. Mediterr.* **57**, 363–364, DOI: 10.14601/Phytopathol_Mediterr-24454 (2018).
17. EPPO. First report of *Xylella fastidiosa* subsp. *multiplex* in Portugal. Report, European Plant Protection Organization (2019).
18. Nunney, L., Ortiz, B., Russell, S., Sanchez, R. & Stouthamer, R. The complex biogeography of the plant pathogen *Xylella fastidiosa* : Genetic evidence of introductions and subspecific introgression in Central America. *Plos One* **9**, DOI: 10.1371/journal.pone.0112463 (2014).
19. Yuan, X. *et al.* Multilocus sequence typing of *Xylella fastidiosa* causing Pierce's Disease and oleander leaf scorch in the United States. *Phytopathology* DOI: 10.1094/PHYTO-100-6-0601

(2010).

20. Nunney, L., Schuenzel, E., Scally, M., Bromley, R. & Stouthamer, R. Large-scale intersubspecific recombination in the plant-pathogenic bacterium *Xylella fastidiosa* is associated with the host shift to mulberry. *Appl. Environ. Microbiol.* **80**, 3025–3033, DOI: 10.1128/AEM.04112-13 (2014).

21. Loconsole, G. *et al.* Intercepted isolates of *Xylella fastidiosa* in Europe reveal novel genetic diversity. *Eur. J. Plant Pathol.*
146, 85–94, DOI: 10.1007/s10658-016-0894-x (2016).

22. Coletta-Filho, H. D., Francisco, C. S., Lopes, J. R. S., Muller, C. & Almeida, R. P. P. Homologous recombination and *Xylella fastidiosa* host–pathogen associations in South America. *Phytopathology* **107**, 305–312, DOI: 10.1094/ PHYTO-09-16-0321-R (2017).

23. Marcelletti, S. & Scortichini, M. Genome-wide comparison and taxonomic relatedness of multiple *Xylella fastidiosa* strains reveal the occurrence of three subspecies and a new X*ylella* species. *Arch. Microbiol.* **198**, 803–812, DOI: 10.1007/s00203-016-1245-1 (2016).

24. Li, W., Raoult, D. & Fournier, P. Bacterial strain typing in the genomic era. *Fems Microbiol. Rev.* **33**, 892–916, DOI: 10.1111/j.1574-6976.2009.00182.x (2009).

25. Kremer, K. *et al.* Comparison of methods based on different molecular epidemiological markers for typing of M*ycobacterium tuberculosis* complex strains: Interlaboratory study of discriminatory power and reproducibility. *J. Clin. Microbiol.* **37**, 2607–2618 (1999).

26. Lindstedt, B. Multiple-locus variable number tandem repeats analysis for genetic fingerprinting of pathogenic bacteria.
Electrophoresis **26**, 2567–2582, DOI: 10.1002/elps.200500096 (2005).

27. Della Coletta-Filho, H. *et al.* Differentiation of strains of *Xylella fastidiosa* by a variable number of tandem repeat analysis.
Appl. Environ. Microbiol. DOI: 10.1128/AEM.67.9.4091\T1\textendash4095.2001 (2001).

28. Lin, H. *et al.* Multilocus simple sequence repeat markers for differentiating strains and evaluating genetic diversity of
Xylella fastidiosa . *Appl. Environ. Microbiol.* **71**, 4888–4892, DOI: 10.1128/AEM.71.8.4888-4892.2005 (2005).

29. Montero-Astúa, M. *et al.* Isolation and molecular characterization of *Xylella fastidiosa* from coffee plants in Costa Rica. *J. Microbiol.* DOI: 10.1007/s12275-008-0072-8 (2008).

30. Coletta-Filho, H. D., Francisco, C. S. & Almeida, R. P. P. Temporal and spatial scaling of the genetic structure of a vector-borne plant pathogen. *Phytopathology* DOI: 10.1094/PHYTO-06-13-0154-R (2014).

31. Lin, H., Islam, M., Cabrera-La Rosa, J., Civerolo, E. & Groves, R. Population structure of *Xylella fastidiosa* associated with almond leaf scorch disease in the San Joaquin Valley of California. *Phytopathology* **105**, 825–832, DOI: 10.1094/ PHYTO-09-14-0254-R (2015).

32. Montes-Borrego, M., Lopes, J. R. S., Jiménez-Díaz, R. M. & Landa, B. B. Combined use of a new SNP-based assay and multilocus SSR markers to assess genetic diversity of *Xylella fastidiosa* subsp. pauca infecting citrus and coffee plants. *Int. Microbiol.* DOI: 10.2436/20.1501.01.230 (2015).

33. Francisco, C. S., Ceresini, P. C., Almeida, R. P. P. & Coletta-Filho, H. D. Spatial genetic structure of coffee-associated *Xylella fastidiosa* populations indicates that cross-infection does not occur with sympatric citrus orchards. *Phytopathology* **107**, 395–402, DOI: 10.1094/PHYTO-08-16-0300-R (2017).

34. Cariddi, C. *et al.* Isolation of a *Xylella fastidiosa* strain infecting olive and oleander in Apulia, Italy. *J. Plant Pathol.* DOI: 10.4454/JPP.V96I2.024 (2014).

35. Martelli, G., Boscia, D., Porcelli, F. & Saponari, M. The olive quick decline syndrome in south-east Italy: a threatening phytosanitary emergency. *Eur. J. Plant Pathol.* **144**, 235–243, DOI: 10.1007/s10658-015-0784-7 (2016).

36. Bergsma-Vlami, M. *et al.* Assessment of the genetic diversity of *Xylella fastidiosa* in imported ornamental coffea arabica plants. *Plant Pathol.* **66**, 1065–1074, DOI: 10.1111/ppa.12696 (2017).

37. Bleve, G. *et al.* In vitro activity of antimicrobial compounds against *Xylella fastidiosa* , the causal agent of the olive quick decline syndrome in Apulia (Italy). *FEMS Microbiol. Lett.* **365**, DOI:

10.1093/femsle/fnx281 (2017).

38. Benson, G. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res.* **27**, 573–580, DOI: 10.1093/nar/27.2.573 (1999).

39. Maiden, M. *et al.* Multilocus sequence typing: A portable approach to the identification of clones within populations of pathogenic microorganisms. *Proc. Natl. Acad. Sci. United States Am.* **95**, 3140–3145, DOI: 10.1073/pnas.95.6.3140 (1998).

40. Safady, N., Lopes, J., Francisco, C. & Della Coletta, H. Distribution and genetic diversity of *Xylella fastidiosa* subsp. pauca associated with olive quick syndrome symptoms in southeastern Brazil. *Phytopathology* **109**, 257–264, DOI: 10.1094/PHYTO-07-18-0273-FI (2019).

41. EFSA. Statement on diversity of *Xylella fastidiosa* subsp. *pauca* in Apulia. *EFSA J.* **14**, 4542, DOI: 10.2903/j.efsa.2016. 4542 (2016).

42. Lin, H. *et al.* Genetic variation of *Xylella fastidiosa* associated with grapevines in two major viticulture regions in the
United States: California and Texas. *J. Plant Pathol.* **95**, 329–337 (2013).

43. Harper, S., Ward, L. & Clover, G. Development of lamp and real-time PCR methods for the rapid detection of *Xylella fastidiosa* for quarantine and field applications. *Phytopathology* **100**, 1282–1288, DOI: 10.1094/PHYTO-06-10-0168 (2010).

44. Bruvo, R., Michiels, N., D'Souza, T. & Schulenburg, H. A simple method for the calculation of microsatellite genotype distances irrespective of ploidy level. *Mol. Ecol.* **13**, 2101–2106, DOI: 10.1111/j.1365-294X.2004.02209.x (2004).

45. Clark, L. & Jasieniuk, M. POLYSAT: An R package for polyploid microsatellite analysis. *Mol. Ecol. Resour.* **11**, 562–566, DOI: 10.1111/j.1755-0998.2011.02985.x (2011).

46. Kamvar, Z. N., Tabima, J. F. & Grunwald, N. J. Poppr: An R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ* **1**, 1–14, DOI: 10.7717/peerj.281 (2014).

47. Meirmans, P. & Van Tienderen, P. GENOTYPE, and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. *Mol. Ecol. Notes* **4**, 792–794, DOI: 10.1111/j.1471-8286.2004.00770.x (2004).

48. Kassambara, A. & Mundt, F. Factoextra: Extract and visualize the results of multivariate data analyses [R package Factoextra version 1.0.5] (2016).

49. Nascimento, M. *et al.* Phyloviz 2.0: providing scalable data integration and visualization for multiple phylogenetic inference methods. *Bioinformatics* **33**, 128–129, DOI: 10.1093/bioinformatics/btw582 (2017).

50. Francisco, A., Bugalho, M., Ramirez, M. & Carrico, J. Global optimal eburst analysis of multilocus typing data using a graphic matroid approach. *Bmc Bioinforma.* **10**, DOI: 10.1186/1471-2105-10-152 (2009).

51. Feil, E., Li, B., Aanensen, D., Hanage, W. & Spratt, B. eBURST: inferring patterns of evolutionary descent among clusters of related bacterial genotypes from multilocus sequence typing data. *J. Bacteriol.* **186**, 1518–1530, DOI: 10.1128/JB.186.5.1518-1530.2004 (2004).

4.11 **Supplementary material of article 3**

Table S1: Tandem Repeats identified and used in literature

Locus Name	Forward primer	Reverse primer	TR sequence
SSR20	ATGAAGAAGCCAGGATACAT	GCTACACGTGCAACAAC	ATTGCTG
SSR21	AACACGGATCAAGCTCATG	GGAACACGCAATAGTAAGA	TGTTATC
SSR26	CTGTGATCGGTGAATTGA	TCAAGCACACTTCCTACG	GTGTGTGA
SSR28	GC(T)AACGCTGTTATCTCAAT	ATTACGCTTCTTATCGCTGT	GTGTGCCT
SSR30	TACGCTGCAC(G)CTGTCTG(T)	CTGTGAACTTCCATCAATCC	TGATCCTG
SSR36	ATGTCACTCAGGTCAGG	CAGAACCACCGACTG(CTCT)	TGTTGGGG
SSR40	ACCT(G)TGACGACGGATG	TAGGAACTGCTGCTACTGAT	GAAGGCGTA
SSR32	AGATGAACC (-) TCGCCAC	GTACTCATCTGCGATGG	CTGATGTG (GTGATGCG)
SSR34	TGATAGAACTGTTTGACGCATTTG	TCGGGAAGTTTGGGGTGAC	(TTGGGTAG)/(TTGGGTAA)
OSSR-2	TTGCTTCACCATTAGCCTTATC	GGCCGTACAGGACCGATC	ATG
OSSR-9	TAGGAATCGTGTTCAAACTG	TTACTATCGGCAGCAGAC	TTTCCGT
OSSR-12	ACAGTCTGTGTCCGCAATTTG	CAGGCGCAGATAGCATTGATC	AGAGGGTAT
OSSR-14	GGCGTAACGGAGGAAACG	ATGAACACCCGTACCTGG	TGA(G)TCCATCC(T)CT(G)GTG
OSSR-16	GCAAATAGCATGTACGAC	GTGTTGTGTATGTGTTGG	CTGCTA
OSSR-17	AGT(C)ACAGCGAACAGGCATTG	AGCAACCAG(A)GACGGGAAC	TGCCTG
OSSR-19	GCTGTGAACTTCCATCAATCC	GCAAGTAGGGGTAAATG(A)TGAC	CAGGATCA
OSSR-20	ATCTGTGCGGCGGTTCTG	CACTTGCGGCGTAGATACTTC	AGGATGCTA
CSSR-4	AACCCA(C)ATTCTTT-(G)TAATATGTG	TTGCAGCATTAGATATTTGAG	TGCC(A)
CSSR-6	CGCACTGTCATCCATTT(C)AATC	GCTGCTTCATCTAGACGTG	G(C)CTGTA
CSSR-7	CACAGCGAACAGGCATTG	AGCAACCAAGACGGGAAC	CTGTGC
CSSR-10	GCAACCACAAAGCCGCAG	AGCACCTCTTAGCATCACTGG	CAATGA
CSSR-12	TAAGTCCATCACCGAGAAG	AAACGGATTTAGGAACACTC	GAAGGCGTA
CSSR-13	CAATGTCACTCAGGTCAG	TTCTGGAATACATCAAATGC	TGTTGGGG
CSSR-16	CGATCAACCCATTCACTG	GCT(C)CCTATTTGCATGATATTG	GTGGTGGCA
CSSR-17	AGAAGTATTCGCTACGCTACG	GGTGATGATTCAGTTGGTGTTG	CTGATGTG
CSSR-18	GTGCTTCCAGAAGTTGTG	GACTGTTCTCTTCGTTCAG	GCCAA

CSSR-19	TGCTGTGATTGGAGTTTTGC	TCAAACGAATCTGTCCATCAAG	TGGTGAG
CSSR-20	GGTATCGCCTTTGGTTCTGG	GACAACCGACATCCTCATGG	GTAGCA
ASSR-9	GGTTGTCGGGCTCATTCC	TTGTCACAGCATCACTATTCTC	CAAGTAC
ASSR-11	AGAGGCAACGCAGGAACAG	GTGAGTTATATCGGTGCAGCAG	ACGCATC
ASSR-12	TGCTCATTGTGGCGAAGG	CGCAACGTGCATTCATCG	GATTCAG
ASSR-14	TTGACTCAAGGAATAAAAC	GAAAAGAGTGTCAATACG	CTGCGTGC
ASSR-16	TTAATCAACAACGCTTATCC	TCGCAGTAGCCAGTATA(G)C	GCTCCGGTTCTA (GCTCCA)
ASSR-19	CGCCGACTGTCTATG(A)TGAC	TTCC(G)TAGCAATGGCAATGTTG	ACAACG
ASSR-20	TTACTATCGGCAGCAGACG	TGAAGCAATGGTGGATTTAGG	ACAGAAA
GSSR_4	GCGTTACTGGCGACAAA(G)C	GCTCGT(C)TCCTGACCTGTG	ATCC
GSSR_6	TGTTCTCTTCGTTCAGCCAAGC	CGCAGCAGAGCAGCAGTG	CTTGT(G)
GSSR_7	ATCATGTCGTGTCGTTTC	CAATAAAGCACCGAATTAGC	GGCAAC
GSSR_12	TTACGCTGATTGGCTGCATTG	GTCAAACACTGCCTATAGAGCG	TATCTGT
GSSR_14	TTGATGTGCTTTTGCGGTAAG	GACAGG(C)TCCTCTCATTGCG	TCCC(T)GTA
GSSR_15	CCGCAGAGTCCGTTGTA(C)AC	AGCC(T)GACGCACGGTATATC	AGCCTGC
GSSR_19	GCCGATGCAGAACAAGAAC	TCAACTTCGCCACACCTG	GAAAACAAG(C)
GSSR_20	TGGATGGATAGATGATTCAGCC	CGATCAGTGGAGGATGTCTTG	GAACCACT(C)A
COSSR1	GAAACAAGATGGCGGTTGC	CATTTAAACGGGCGGCATA	ATTGCTG
COSSR6	TGCTGCGCGATAACCAAGT	CATCCAATCAGCCCTAACCT	GTGATGCG
CSSR45	ACAGACATCACCGGCATTG	AATGTCGCTGCCAATCCAT	CACACCGAGATGGAC
COSSR4	CAAGGTGACCGCTAGCCTAT	GCTGTCATTGGGTGATGC	CAATACAC
COSSR5	ACACTGACACAACAGCCACCA	AATGGTGGGTGTGATGGTTTC	CATACAGA
COSSR3	AAGTATTCGCTACGCTACGC	GTGTGTTATGTGTGCCATTCGT	CTGATGTG
CSSR42	ATTACGCTGATTGGCTGCAT	GTTTCATTACGCGGAACAC	TGTTATC

Sample	TR4	TR5	TR6	TR7	TR8	TR9	TR10	TR12	TR15	TR18	TR19	TR21	TR23	TR24	SSR20	SSR21	SSR28	CSSR-9	CSSR-16	CSSR-19	CSSR-20	CSSR-7	CSSR-10	CSSR-45	CSSR-18	CSSR-19	ASSR-9	ASSR-11	ASSR-12	ASSR-16	ASSR-19	GSSR-4	GSSR-7	CSSR-42	CCSSR-6	CCSSR-4	CCSSR-5
CFBP8073	13	9	0	3	13	7	11	10	5	6	0	9	2	2	2	2	0	16	11	0	20	7	10	13	8	3	2	9	12	72	2	49	30	1	1	2	1
CFBP7970	4	3	1	3	8	10	10	4	5	3	0	12	5	2	2	0	4	11	9	5	16	12	4	8	10	3	2	8	3	61	2	24	10	7	1	2	1
CFBP8351	4	3	1	3	9	10	11	4	5	6	0	9	7	2	7	2	0	14	14	6	8	11	6	10	10	3	2	9	6	53	2	15	23	11	1	2	1
CFBP8077	13	7	4	3	4	10	4	5	5	3	0	2	1	2	7	2	0	21	13	10	5	12	3	5	7	3	2	7	7	66	2	24	17	19	1	2	1
CFBP8356	13	5	6	3	15	10	4	7	5	3	0	3	1	2	6	2	0	24	15	5	12	15	6	5	5	2	2	4	10	65	3	11	21	7	1	2	1
CFBP8419	13	4	7	3	4	10	4	12	5	3	2	3	1	2	6	2	0	13	5	3	3	15	15	5	4	2	2	4	12	66	2	11	17	21	2	2	1
CFBP8478	14	4	1	3	4	10	4	12	5	3	2	3	1	2	6	2	0	11	5	3	3	15	11	5	4	2	2	4	10	66	3	11	13	21	2	2	1
CFBP8173	13	2	1	3	9	10	7	4	5	6	6	7	1	5	12	2	0	7	8	5	1	16	2	5	22	3	3	15	7	68	12	6	23	3	1	2	1
CFBP8416	17	2	1	3	11	10	4	5	5	3	6	7	1	5	8	2	0	5	22	7	1	9	9	5	7	3	15	7	9	55	12	8	12	1	1	2	1
CFBP8417	16	2	3	3	9	10	5	6	5	3	6	3	1	3	10	2	0	4	8	10	1	9	6	5	7	4	11	12	13	65	9	17	19	1	1	2	1
CFBP8429	4	8	2	9	12	6	7	13	4	4	5	3	8	4	5	2	10	3	7	7	24	11	4	9	2	0	2	2	4	70	3	6	2	1	7	12	5
CFBP8072	4	13	2	12	14	7	4	9	6	3	6	4	1	4	7	2	5	3	12	7	1	6	5	5	11	0	2	4	14	55	2	55	2	13	9	9	1
CFBP8074	4	13	2	14	14	10	9	16	5	3	6	4	1	4	7	2	5	3	12	7	1	6	5	5	13	0	2	4	14	55	2	54	2	13	9	9	1
CFBP8402	4	7	3	27	12	9	9	16	5	3	3	3	8	4	11	2	7	3	27	12	1	9	7	5	2	1	2	2	4	47	3	6	2	7	7	10	6
CFBP8477	4	7	1	23	12	9	9	16	5	3	3	3	8	4	11	2	7	3	27	9	1	9	7	5	2	0	2	2	4	47	3	6	2	7	7	10	6
Oe_Ie1	4	7	3	23	12	9	9	16	5	3	3	3	8	4	11	2	7	3	27	12	1	9	7	8	3	1	2	2	4	47	3	6	2	2	7	10	6
Nq_Ie1	4	7	3	23	13	9	9	17	5	3	3	3	8	4	12	2	7	3	27	10	1	9	7	8	2	1	2	2	4	47	3	6	2	2	7	10	6
Oe_Ie2	4	7	3	23	12	9	9	16	5	3	3	3	8	4	12	2	7	3	27	11	1	9	7	8	2	1	2	2	4	47	3	6	2	2	7	10	6
Nq_Ie2	4	7	3	21	12	9	9	16	5	3	3	3	8	4	12	2	7	3	24	11	1	9	7	8	3	2	2	2	4	46	3	6	2	2	7	10	6
Oe_Ie3	4	7	3	26	12	9	9	16	5	3	3	3	8	4	11	2	8	3	27	10	1	9	7	8	2	1	2	2	4	47	3	6	2	2	7	11	6
Pm_Ie1	4	7	1	23	12	9	9	16	5	3	3	3	8	4	11	2	7	3	27	12	1	9	7	8	2	1	2	2	4	47	3	6	2	2	7	11	6
Pm_Ie2	4	7	3	19	12	9	9	16	5	3	3	3	8	4	12	2	7	3	27	11	1	9	7	8	3	1	2	2	4	47	3	6	2	2	7	10	6
Pm_Ie3	4	7	3	19	14	9	9	16	5	3	3	3	8	4	12	2	7	3	27	11	1	9	7	8	2	1	2	2	4	47	3	6	2	2	7	10	6
Pm_Ie4	4	7	3	24	12	9	9	16	5	3	3	3	8	4	12	2	7	3	27	11	1	9	7	8	3	1	2	2	4	47	3	6	2	2	7	10	6
Oe_Ie4	4	7	3	25	12	9	9	16	5	3	3	3	8	4	12	2	7	3	26	11	1	9	7	8	3	1	2	2	4	46	3	6	2	2	7	10	6
Pm_Ie5	4	7	3	21	13	10	9	16	5	3	3	3	8	4	12	2	7	3	27	11	1	9	7	8	2	1	2	2	4	46	3	6	2	2	7	11	6
Pd_Ie1	4	7	3	14	12	9	9	16	5	3	3	3	8	4	12	2	7	3	23	11	1	9	7	8	3	1	2	2	4	46	3	6	2	3	7	10	6
Ra_Ie3	4	7	3	22	12	9	9	16	5	3	3	3	8	4	12	2	7	3	26	11	1	9	7	8	3	1	2	2	4	46	3	6	2	2	7	10	6
Ra_Ie1	4	7	3	22	12	9	9	16	5	3	3	3	8	4	12	2	7	3	26	11	1	9	7	8	3	1	2	2	4	46	3	6	2	2	7	10	6
Pm_Ie6	4	7	3	23	13	10	9	16	5	3	3	3	8	4	12	2	7	3	27	11	1	9	7	8	3	1	2	2	4	46	3	6	2	2	7	10	6
No_Ie6	4	7	3	23	12	9	8	16	5	3	3	3	8	4	12	2	7	3	27	11	1	9	7	8	3	1	2	2	4	47	3	6	2	2	7	10	6
Oe_Ie5	4	7	3	23	14	9	9	16	5	3	3	3	8	4	12	2	7	3	26	11	1	9	7	8	3	1	2	2	4	47	3	6	2	2	7	9	6
Oe_Br5	4	7	3	23	12	9	9	16	5	3	3	3	8	4	12	2	7	3	26	11	1	9	7	8	3	1	2	2	4	46	3	6	2	2	7	10	6
Oe_Br1	4	7	3	22	12	9	9	16	5	3	3	3	8	4	12	2	7	3	27	11	1	9	7	8	3	1	2	2	4	46	3	6	2	2	7	10	6
Oe_Br2	4	7	3	22	12	9	9	16	5	3	3	3	8	4	12	2	7	3	27	11	1	9	7	8	3	1	2	2	4	46	3	6	2	2	7	11	6
Oe_Br6	4	7	3	23	13	9	9	16	5	3	3	3	8	4	12	2	7	3	26	11	1	9	7	8	3	1	2	2	4	46	3	6	2	2	7	11	6
Oe_Br3	4	7	3	24	13	9	9	16	5	3	3	3	8	4	12	2	7	3	27	11	1	9	7	8	3	1	2	2	4	47	3	6	2	2	7	11	6
Oe_Tp1	4	7	3	24	12	9	9	16	5	3	3	3	8	4	12	2	7	3	24	11	1	9	7	8	3	1	2	2	4	47	3	6	2	2	7	11	6
Oe_Tp2	4	7	3	20	12	8	9	16	5	3	3	3	8	4	12	2	7	3	25	10	1	9	8	8	3	1	2	2	4	47	3	6	2	2	7	12	6
Oe_Ie6	4	7	3	20	12	9	9	16	5	3	3	3	8	4	12	2	7	3	24	11	1	9	7	8	3	1	2	2	4	46	3	6	2	2	7	9	6
Ae_Ie1	4	7	3	24	12	9	9	16	5	3	3	3	8	4	13	2	7	3	25	11	1	9	7	8	3	1	2	2	4	11	3	6	2	2	7	10	6
Ra_Ie2	4	7	3	24	12	9	9	16	5	3	3	3	8	4	12	2	7	3	25	11	1	9	7	8	3	1	2	2	4	46	3	6	2	2	7	10	6
Ae_Ie2	4	7	3	23	12	9	9	16	5	3	3	3	8	4	12	2	7	3	26	11	1	9	7	8	3	1	2	2	4	48	3	6	2	2	7	10	6
Pm_Ie7	4	7	3	23	12	9	9	16	5	3	3	3	8	4	12	2	7	3	26	11	1	9	7	8	3	1	2	2	4	46	3	6	2	2	7	9	6
No_Ie4	4	7	3	24	12	9	9	17	5	3	3	3	8	4	12	2	7	3	28	11	1	9	7	8	3	1	2	2	4	46	3	6	2	2	7	10	6
No_Ie5	4	7	3	21	12	9	9	16	5	3	3	3	8	4	11	2	7	3	26	11	1	9	7	8	3	1	2	2	4	46	3	6	2	2	7	10	6
Un_Ie1	4	7	3	24	12	9	9	14	5	3	3	3	8	4	11	2	8	3	26	12	1	9	7	8	3	1	2	2	4	46	3	6	2	2	7	9	6
Pm_Ie8	4	6	3	20	13	9	10	15	5	3	3	3	8	4	12	2	7	3	23	11	1	9	7	8	2	1	2	2	4	47	3	6	2	2	7	9	6
Pm_Ie9	4	7	3	23	12	8	9	16	5	3	3	3	8	4	12	2	7	3	17	7	1	9	7	8	2	1	2	2	4	46	3	6	2	2	7	9	6
Pd_Ie2	4	7	3	38	8/9	9	9	17/16/19	5	3	3	3	8	4	04/11/12	2	7	3	23	12/11/13	1	9	7	8	3	1	2	2	4	46	3	6	2	2	7	10	6
Pm_Ie10	4	7	3	24/23/27	12	9	9	16	5	3	3	3	8	4	11	2	7	3	25	11	1	9	7	8	2	1	2	2	4	46	3	6	2	2	7	9	6

Table S2: Results of TRs of each strain for each locus

Chapter 5 **: Conclusion**

5.1 **Thesis conclusions**

Bacterial plant pathogens represent one of the main threats to the plants and food due to their control difficulties. In the last two decades, many bacterial plant pathogens have emerged due to different factors such as the active movement of goods around the world, the wide application of agrochemicals and climate change. To understand how populations of these pathogens emerge and adapt to new environments, genetic analysis can represent a useful path that is able to trace pathogens responsible for the epidemics and to understand their evolution, profile, adaption and population structures, aiming to use this information for plant disease control measurements (1). Population genetic analysis depends on molecular methods that have proved their efficiency in detecting and identifying plant pathogens but also to reveal more about genetic diversity (2–4). Here, we developed and used an MLVA molecular genotyping approach for identification, detection, tracking, and study of the microevolutionary history of three different bacterial plant pathogens: *Pseudomonas syringae* pv. *Actinidiae*, the causal agent of the bacterial canker of kiwifruit, *Xylella fastidiosa* subsp. *pauca* that causes the Olive Quick Decline Syndrome (OQDS), both causing sudden and harmful outbreaks in Italy (5,6), and the endemic plant pathogen *Pseudomonas savastanoi*, which causes olive knots on *Olea europea* L. and other similar diseases on other host plants by its pathovars.

Aims of this thesis are:

- ✓ To set up and evaluate a molecular genotyping approach by MLVA (Multiple Locus Variable-Number Tandem Repeat Analysis) to obtain genetic fingerprinting of the three different types of bacterial pathogens.
- ✓ To understand if this method, with each of the pathogens analysed, could be able to differentiate populations within the same species, subspecies, pathovars, biovars.
- ✓ To check if the method matches the required criteria for the detection, identification, and tracing pathogen of disease outbreaks.
- ✓ To find out how this method is suitable in comparing and evaluating genetic differences in both endemic and epidemic bacterial pathogens in order to reveal the genetic structures of their populations.

Three pathogens were used representing three different diseases and based on the results obtained and discussed in this thesis. The conclusion will deal with two aspects, the first one is the **technical point of view of the MLVA method**

- ✓ The few handling steps of an MLVA made it an easy and rapid method to handle genetic analysis and a simple but powerful technique for bacterial genotyping by providing numerical data, which will be accessible and comparable via the internet in a future MLVA database.
- ✓ The method has proved to be more effective than many other classical molecular analyses due to its ability to differentiate among closely related strains with low running costs, and high accuracy.
- ✓ The described method could be considered as the first-line assay for the local outbreaks in case of endemic pathogens, even more in epidemic outbreaks, and in surveillance investigations.

Regarding the biological point of view

- ✓ The ability of the method to discriminate the strains within the same species, subspecies, pathovars, biovars was demonstrated. In addition to its ability to track pathogens origin, often the method assigned the strains to their origin and hosts, especially in case of outbreaks.

Concerning ***Pseudomonas syringae*** pv. ***actinidiae*** results

- ✓ The method correctly assigns the strains to the known biovars using three different analytical methods of the data, Hierarchical clustering, STRUCTURE, and Discriminant analysis of principal components (DAPC).
- ✓ In addition to the assignment of the main five biovars, new groups were identified and remained independent with strong statistical support (Pop1-Pop9). These groups belong to Japan, Korea, and the majority of the groups are from China.
- ✓ We believe that the genetic variability of *Pseudomonas syringae* pv. *actinidiae* was not fully described yet.

Pseudomonas savastanoi

- ✓ The method was able to discriminate all the pathovar, it clustered together isolates mainly according to their hosts and geographic origin.
- ✓ Higher variability among the strains of pathovar *savastanoi* was observed.

- ✓ The method was able to expose the genetic diversity, from a single tree in the field and resulted particularly useful in identification and tracking *P. savastanoi* populations at the local level
- ✓ Multiple infections in a single plant were found in the strains isolated from Viterbo field, which could be explained by human-mediated transmission of the bacterial infection from a plant to another during pruning practices by infected tools.

Xylella fastidiosa

- ✓ The method was able to assign the strains to their subspecies and ascertain differences even within a single sequence type (ST).
- ✓ The analyses of the 38 Italian strains for 37 loci, revealed that 18 loci are identical, 11 changed periodically, and only7 has differences in the numbers of repeats.
- ✓ The strain CFBP8429, isolated from *Coffea arabica* plants in Angers, France reported as belonging to ST53, seems to be significantly distant from all the others ST53 from Italy.
- ✓ Crude DNA of the bacteria isolated from infected plants directly was used successfully, giving a very important advantage in cost and time reduction for large screening purposes and slow-growing bacteria.
- ✓ No correlation between the strains and geographic origin or the host was found.
- ✓ Again, three analytical methods were used to analyse the data: Hierarchical clustering, STRUCTURE, and Discriminant analysis of principal components (DAPC). Hierarchical clustering, and (DAPC), seem to give a clearer result than STRUCTURE

5.2 References

1. Grunwald NJ, Everhart SE, Knaus BJ, Kamvar ZN. Best Practices for Population Genetic Analyses. Phytopathology [Internet]. 2017;107(9):1000–10. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28513284

2. Cangelosi GA, Freeman RJ, Lewis KN, Livingston-Rosanoff D, Shah KS, Milan SJ, et al. Evaluation of a high-throughput repetitive-sequence-based PCR system for DNA fingerprinting of *Mycobacterium tuberculosis* and *Mycobacterium avium* complex strains. J Clin Microbiol [Internet]. 2004;42(6): 2685–93. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15184453

3. Zhao S, Poulin L, Rodriguez RL, Serna NF, Liu S-YY, Wonni I, et al. Development of a variable number of tandem repeats typing scheme for the bacterial rice pathogen *Xanthomonas oryzae* pv. *oryzicola*. Phytopathology [Internet]. 2012;102(10):948–56. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22957820

4. Zaluga J, Stragier P, Van Vaerenbergh J, Maes M, De Vos P. Multilocus Variable-Number-Tandem-Repeats Analysis (MLVA) distinguishes a clonal complex of *Clavibacter michiganensis* subsp. *michiganensis* strains isolated from recent outbreaks of bacterial wilt and canker in Belgium. BMC Microbiol. 2013;13(1).

5. Balestra GM, Mazzaglia A, Spinelli R, Graziani S, Quattrucci A, Rossetti A. Cancro batterico su *Actinidia chinensis*. L'Informatore Agrar. 2008;38:75–6.

6. Saponari M, Loconsole G, Almeida R, Coletta-Filho HD. Isolation, genotype and preliminary data on the pathogenicity of *Xylella fastidiosa* CoDiRO strain. In 2013.