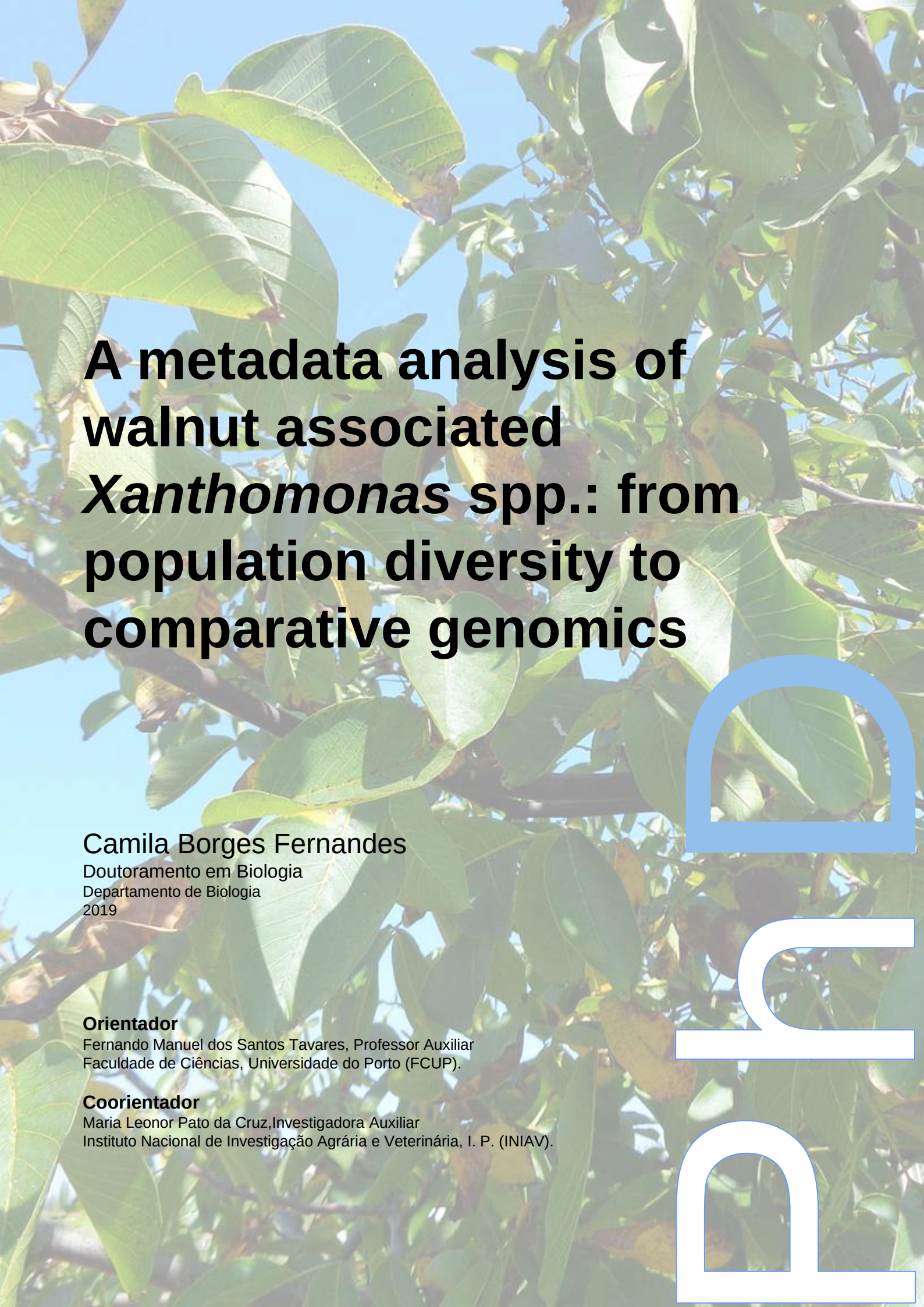

A metadata analysis of walnut associated *Xanthomonas* spp.: from population diversity to comparative genomics

Camila Borges Fernandes
Doutoramento em Biologia
Departamento de Biologia
2019

Orientador
Fernando Manuel dos Santos Tavares, Professor Auxiliar
Faculdade de Ciências, Universidade do Porto (FCUP).

Coorientador
Maria Leonor Pato da Cruz,Investigadora Auxiliar
Instituto Nacional de Investigação Agrária e Veterinária, I. P. (INIAV).

A metadata analysis of walnut associated *Xanthomonas* spp.: from population diversity to comparative genomics

O trabalho apresentado nesta dissertação foi realizado no Centro de Investigação em Biodiversidade e Recursos Genéticos, InBIO Laboratório associado (CiBIO-InBIO) e no Instituto Nacional de Investigação Agrária e Veterinária, I.P. (INIAV, I.P.), no âmbito da bolsa de doutoramento (SFRH/BD/95913/2013) financiada pela Fundação para a Ciência e a Tecnologia (FCT) com o projeto de doutoramento intitulado, "*Xanthomonas arboricola* pv. *juglandis* – unveiling the biogeographic distribution of emergent and acutely virulent *Xaj* genotypes to infer epidemiological patterns and improve phytosanitary control".

A metadata analysis of walnut associated *Xanthomonas* spp.: from population diversity to comparative genomics

Os trabalhos apresentados também tiveram apoio dos projetos de investigação:

- EVOXANT (PTDC/BIA-EVF/3635/2014, POCI-01-0145-FEDER-016600), financiado pelo European Structural & Investment Funds (ESIFs) através da Autoridade de Gestão do Programa Operacional Competitividade e Internacionalização (COMPETE 2020) e pela Fundação para a Ciência e a Tecnologia (FCT).

- European Cooperation in Science and Technology (COST) action CA16107 EuroXanth – "Integrating science on Xanthomonadaceae for integrated plant disease management in Europe".

FCT
Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR

POCH
PROGRAMA OPERACIONAL CAPITAL HUMANO

EuroXanth
COST Action CA16107

Acknowledgments

Four years have passed since I started my PhD and hereby I would like to express my gratitude to all the people that contributed to this dissertation and made me a better scientist and a better person throughout these years:

To my PhD supervisor, Professor Fernando Tavares, thank you for believing in me since the beginning, when we met at the interview for a fellowship. For always inspiring me, believing in the work we set out to accomplish and encouraging me to challenge myself. Thank you to open your lab for me, but most of all, to opening your home and always, together with Perpétua and Antónia, making me feel welcome! It has been an honour to be your PhD student and I hope our paths continue to cross.

To my co-supervisor, Eng. Leonor Cruz, thank you for accepting me in your lab, where you generously received me with open arms and allowed me to perform a substantial amount of my work. For your help and expertise in phytopathology that were essential to my professional progress. Thank you to share your knowledge with me, for your constant availability and support in the last five years.

To Dr. Pedro Albuquerque, for me my second co-supervisor, thank you for all your scientific explanations, leadership in the lab, for being a brilliant person and for teaching me a lot! Thank you for your availability since the beginning and for making sure you would be always present!

To Dr. Paula Sá Pereira, thank you for sharing with me your enthusiasm for science! Thank you for your scientific discussions and helpful suggestions during my thesis. Thank you for cheer me up through the hard times and encouraging words.

To my colleagues that turned friends for life, Ana Chegão, Joana Cruz and Márcia Santos, thank you for your continued support and encouragement through all these years. Thanks for your advices in all matters of life!

To all the members of Microbial Diversity and Evolution group, to Cláudia Serra, Cristina Dias, Leonor Martins and Sofia Martins, thank you for all the support during my stay at Porto and friendship. A special acknowledgement is due to Leonor Martins for her important contributions to this thesis regarding the work developed in Chapter III and Chapter V.

To Eng. Rui de Sousa, co-author in joint publications, thank you for your expertise that were essential to this work and support during surveys on walnut orchards.

To Dr. Jochen Blom and Dr. Joël Pothier, both co-authors in joint publications, thank you for allowing us to take our work one step further. I would like to thank specially Dr. Joël Pothier for welcoming me in your lab in Switzerland and for your guidance in the results presented in Chapter V.

To the team of the Seed Lab (INIAV), Eng. Lurdes Santos and Eng. Paulo Godinho, thank you for the work performed on walnut seedlings (Chapter IV), to Eng. Joaquim Trindade for your kind collaboration in taking the photos of pathogenicity results (Chapter IV and VI) and to all the walnut producers, thank you for the valuable collaborations.

À Isaura Velez, obrigada por ser a nossa Isaura, que sempre está pronta a ajudar e a defender as suas meninas. Obrigada por toda a sua amizade e os momentos de descontração que partilhamos!

À minha família maluca e a todos os meus amigos, obrigada pelo carinho constante.

Aos meus primos, obrigada pelas nossas jurucas que me ajudaram a descontrair e pela vossa amizade de irmãos! #reisdasjurucas

Aos meus queridos pais, obrigada por sempre acreditarem em mim e por apoiarem todas as minhas decisões e os meus sonhos académicos. Espero poder um dia retribuir tudo aquilo que fizeram por mim. Considero-me uma pessoa muito feliz e devo tudo o que sou a vocês! Amo-vos muito!

Às minhas manas, nani e pilili, obrigada por serem minhas amigas e pela nossa união eterna. A mana mais velha estará sempre aqui para vocês! Amo-vos muito!

Ao meu yuyu, obrigada por teres mudado a minha forma de estar na vida e por me teres tornado numa pessoa mais afetuosa.

Ao meu marido, meu melhor amigo e bioquímico de todas as horas, Marco obrigada por todo o teu apoio e por estar sempre ao meu lado! Agradeço a tua compreensão e paciência nos momentos em que estive mais distante. A nossa vida ainda agora começou e acredito que seremos muito felizes! Te amo! ♪ “I found a love for me Darling just dive right in, And follow my lead…♫”

This thesis is dedicated to my parents, my sisters, my husband and my dog.

Thank you for everything, you are my reason for living.

Abstract

Xanthomonas arboricola pv. *juglandis* (*Xaj*) is the etiological agent of the most damaging and widespread diseases affecting walnut trees-growing areas, namely walnut bacterial blight (WBB), brown apical necrosis (BAN) and vertical oozing canker (VOC). These diseases are responsible for major economic losses in the main walnut production regions, associated with the reduction of marketability of walnut fruits, the reduction of walnut orchard productivity and the increase in walnut nursery costs.

Regardless the research efforts carried out in the last decade, there is still a poor understanding of *Xaj* biology, including pathogen diversity, population dynamics and adaptation fitness to host and environment. In a phytodiagnostics setting, detection of *Xaj* is still carried out mainly via culture-based methods, with no specific diagnostic tools available. Molecular identification of *Xaj* strains is limited to the species due to difficulties in finding reliable discriminatory genetic regions at pathovar level. Furthermore, the comprehensive assessment of genetic diversity within and between *Xaj* population is still missing, which delays the development of advanced tools for identification of particularly virulent strains, as well as the implementation of efficient control and prevention measures for *Xaj*.

In this work, the selection and validation of nine taxa-specific DNA markers for *Xaj* (XAJ1 to XAJ9) was carried out using *in silico* approaches. These markers were shown to be useful for both hybridization and PCR-based detection and identification techniques. A multiplex PCR was validated as an efficient culture-independent method for detection of *Xaj* in naturally infected walnut samples, and a qPCR assay was optimized to determine the bacterial load of infected walnut tissues. Furthermore, a high throughput dot blot platform using the nine DNA markers allowed prompt identification of different *Xaj* lineages.

Genetic diversity of *Xanthomonas* isolates obtained from walnut trees in Portugal was assessed, considering thermoclimatic and ecological variables thought to modulate the diseases caused by *Xaj*. The walnut-associated xanthomonads population observed to occur in Portugal was more complex than originally thought, including lineages of *Xaj*, lineages of non-*juglandis X. arboricola* and lineages of non-*arboricola Xanthomonas* sp. Among *Xanthomonas* sp. lineages, the evidence of a pathogenic strain diverging from the nonpathogenic phenotype and commensal lifestyle to which these atypical strains have been described was an interesting finding as a promising model to address pathoadaptations to walnut. Additionally, this work showed that co-colonization of the same plant host by distinct *Xanthomonas* strains is frequent, overcoming the phytopathology paradigm of a single pathogenic strain for each diseased plant. Moreover, the isolation of new *X. arboricola* strains from a new host species, *Carya illinoinensis* of the

Juglandaceae family, and shown to be pathogenic in walnut, was particularly exciting, since calls for the need to review the host range of walnut infective *Xanthomonas arboricola*.

The comparative genomics study carried out allowed to highlight the genomic differences between five *Xanthomonas* strains (CPBF 367, CPBF 424, CPBF 426, CPBF 427 and CPBF 1521) isolated from a single walnut host. Core genome phylogeny and average nucleotide identity of genomes *of Xanthomonas* spp. showed two distinct clusters regarding the five strains, one grouping CPBF 427 and CPBF 1521 with *Xaj* strains, and another grouping together CPBF367, CPBF424 and CPBF426, indicating that these strains belong to a new *Xanthomonas* species. The unfolded genomic differences between the two groups, could translate into different pathogenicity and virulence features that are still being addressed. Major differences regarding the repertoire of type III secretion system and associated effectors genes were observed among *Xaj* and *Xanthomonas* sp. strains. In addition, differential sets of genes encoding chemotaxis related proteins, non-fimbrial adhesins, extracellular enzymes, and genes encoding proteins involved in Type II secretion system and Type IV pilus, were also identified between the five strains.

The genomic data of strains CPBF 367, CPBF 424, CPBF 426, CPBF 427 and CPBF 1521 have already been deposited in genomic databases, which enhance the genomic patrimony for further studies. In fact, the genomes made available include a *Xaj* strain (CPBF 1521), believed to not been exposed to selective pressures caused by phytosanitary treatments, which may be a reference genome to address resistance to chemicals commonly used to control walnut bacterial diseases; and the genome of one walnut-pathogenic *Xanthomonas* sp. strain (CPBF 424), particularly appealing to address pathogenesis within the xanthomonads.

The outcomes of the present dissertation are expected to be a solid contribution for the improvement of *Xaj* detection in plant samples, focused on early detection and identification of virulent strains, and to help phytosanitary services to infer epidemiological patterns and to assess infection progression. Importantly, this work contributed for the thorough characterization of *Xaj* populations in Portugal, providing a comprehensive snapshot of *Xaj* current diversity, infection dynamics and relevant genotypes.

Key-words: *Xanthomonas arboricola*, walnut diseases, detection, identification, DNA markers, multiplex PCR, qPCR, genotyping, dot blot, MLSA, genetic diversity, epidemiology, whole-genome sequencing, comparative genomics.

Resumo

Xanthomonas arboricola pv. *juglandis* (*Xaj*) é o agente etiológico das doenças mais graves que afetam a cultura da nogueira mundialmente, nomeadamente a bacteriose da nogueira, e outras duas patologias que apresentam sintomas característicos como a necrose apical dos frutos (brown apical necroses) e o exsudado dos troncos (vertical oozing canker). Estas doenças são responsáveis pelas principais perdas económicas nas principais regiões produtoras de noz, e estão associadas com a diminuição da qualidade da noz, com a diminuição da produção e com o aumento dos custos de manutenção das nogueiras nos viveiros.

Apesar dos esforços recentes de investigação, pouco se conhece acerca da biologia de *Xaj*, incluindo a diversidade genética, dinâmica populacional, e a sua capacidade de adaptação ao hospedeiro e ao ambiente. No âmbito do diagnóstico fitopatológico, a deteção de *Xaj* está ainda dependente de métodos como o isolamento em cultura pura, sem métodos específicos disponíveis. A identificação molecular de estirpes de *Xaj* é restrita ao nível da espécie devido às dificuldades em encontrar regiões genéticas discriminatórias ao nível do patovar. Além disso, uma avaliação abrangente da diversidade genética presente entre e nas populações de *Xaj* é ainda necessária, o que atrasa o avanço de novas abordagens para a identificação especifica de estirpes virulentas, bem como a implementação de medidas de controlo e de prevenção contra *Xaj*.

Neste trabalho, a seleção e validação de nove marcadores de DNA taxa específicos (XAJ1 ao XAJ9) foi efetuada utilizando abordagens *in silico*. Os marcadores demonstraram ser úteis para aplicação em técnicas de deteção e identificação baseadas em PCR e em hibridação. A técnica de multiplex PCR foi validada como um método eficiente e independente do isolamento em cultura para deteção de *Xaj* em amostras de nogueiras naturalmente infetadas, e a técnica qPCR foi otimizada para determinação da carga bacteriana em tecidos de nogueira infetados. Adicionalmente, a utilização de uma plataforma dot blot de alto rendimento permitiu a identificação de diferentes linhagens de *Xaj*.

A diversidade genética dos isolados de *Xanthomonas* obtidos a partir de nogueiras foi analisada tendo em conta variáveis termoclimáticas e ecológicas que possam ter influenciado as doenças causadas por *Xaj* em Portugal. A diversidade de *Xanthomonas* associadas às nogueiras em Portugal foi mais complexa do que inicialmente se esperava, incluindo linhagens de *Xaj*, linhagens de *X. arboricola* não-*juglandis* e linhagens de *Xanthomonas* sp. não-*arboricola*. Entre as linhagens de *Xanthomonas* sp. a evidência para uma estirpe patogénica, que difere do fenótipo não-patogénico e estilo de vida comensal descrito para as estirpes atípicas, constituiu uma descoberta interessante como um modelo para abordar as patoadaptações à nogueira.

Adicionalmente, este trabalho demonstrou que a co-colonização de um mesmo hospedeiro por estirpes distintas de *Xanthomonas* é frequente, contrariando um dos paradigmas da fitopatologia de uma única estirpe patogénica para uma planta doente. Por outro lado, o isolamento de novas estirpes de *X. arboricola* a partir de uma nova espécie de hospedeiro, *Carya illinoinensis* da família Juglandaceae, mas patogénicas em nogueira, é particularmente interessante, pois implica a revisão dos hospedeiros conhecidos para *Xanthomonas arboricola*.

Os estudos de genómica comparativa permitiram identificar diferenças entre cinco estirpes (CPBF 367, CPBF 424, CPBF 426, CPBF 427 e CPBF 1521) isoladas a partir de um espécime de nogueira. A inferência filogenética obtida com o conjunto de genes presentes em todos os genomas de *Xanthomonas* spp. (core genome), e a análise de identidade nucleotídica (average nucleotide identity) de genomas de *Xanthomonas* spp., permitiu separar estas estirpes em dois ramos distintos, um reunindo CPBF 427 e CPBF 1521 com outras estirpes de *Xaj*, e outro agrupando CPBF 367, CPBF 424 e CPBF 426, a sugerir que estas estirpes pertencem a uma nova espécie de *Xanthomonas*. Esta investigação permitiu identificar diferenças genómicas entre os dois grupos associadas a características de patogenicidade e virulência, que estão a ser caracterizadas. Diferenças particularmente relevantes foram observadas entre as estirpes de *Xaj* e *Xanthomonas* sp. para o conjunto de genes do sistema de secreção tipo III e seus efetores. Adicionalmente, foram também identificadas diferenças entre as cinco estirpes a nível do património de genes que codificam para proteínas associadas a quimiotaxia, adesinas não fimbriais, enzimas extracelulares, e genes que codificam para proteínas relacionadas com o sistema de secreção tipo II e pili tipo IV.

Os genomas das estirpes CPBF 367, CPBF 424, CPBF 426, CPBF 427 e CPBF 1521 foram já depositados em bases de dados e, portanto, aumentam o património genómico para estudos futuros. De facto, os genomas disponibilizados incluem uma estirpe de *Xaj* (CPBF 1521) que se pensa não ter sido exposta a pressões seletivas causadas por tratamentos fitossanitários, e, portanto, pode ser um genoma de referência para abordar a resistência aos produtos químicos comummente utilizados no controlo das doenças bacterianas da noz; e o genoma de uma estirpe de *Xanthomonas* sp. patogénica em nogueira (CPBF 424), particularmente interessante para abordar patogénese em xanthomonads.

Espera-se que os resultados desta dissertação sejam um sólido contributo para melhorar a deteção precoce de *Xaj* em plantas, especialmente direcionado para a deteção rápida e a identificação de estirpes virulentas, no sentido de ajudar os serviços fitossanitários a inferir padrões epidemiológicos e avaliar o progresso da infeção. Mais importante, esta dissertação contribuiu para uma caracterização extensa das populações de *Xaj* em Portugal, nomeadamente em relação à diversidade atual, dinâmica da infeção e para a identificação de genótipos relevantes.

Palavras-chave: *Xanthomonas arboricola*, doenças na nogueira, deteção, identificação, marcadores de DNA, multiplex PCR, qPCR, genotipagem, dot blot, MLSA, diversidade genética, epidemiologia, sequenciação de genomas, genómica comparativa.

Table of contents

List of figures

List of tables

List of abbreviations

ANI – Average Nucleotide Identity
BAN – Brown Apical Necrosis
BLAST – Basic Local Alignment Search Tool
Bp – Base pairs
BRIG – BLAST Ring Image Generator
CDS – Coding DNA Sequence
COG – Clusters of Orthologous Group
CPBF – Coleção Portuguesa de Bactérias Fitopatogénicas
Cq – quantitative Cycle
CTAB – Cetyl Trimethylammonium Bromide
CV – Coefficient of Variation
DIG – Digoxigenin
EDGAR – Efficient Database Framework for Comparative Genome Analyses using BLAST Score Ratios
eggNOG – Database of Orthologous Groups and Functional Annotation
ENA – European Nucleotide Archive
EPPO – European Plant Protection Organization
HGT – Horizontal Gene Transfer
HP – Hybridization Pattern
IS – Insertion Sequence
LOD – Limit of detection
MEGA – Molecular Evolutionary Genetics Analysis
MLSA – Multilocus Sequence Analysis
MLST – Multilocus Sequence Typing
Mm – Mesomediterranean thermoclimatic region
Mt – Mesotemperature thermoclimatic region
NCBI – National Center for Biotechnology Information
ORF – Open Reading frame
qPCR – quantitative real time PCR
Sm – Supramediterranean thermoclimatic region
SPC – Sample Processing Control
T3E – Type Three Effectors
T3SS – Type Three Secretion System

Tm – Thermomediterranean thermoclimatic region

VOC – Vertical Oozing Canker

WBB – Walnut Bacterial Blight

CHAPTER I

General introduction

General introduction

1. The phytopathogenic bacteria *Xanthomonas arboricola* and walnut production, an overview

The plant pathogen species *Xanthomonas arboricola* (Vauterin et al. 1995) is a constant threat to important fruit and nut tree crops (**Figure I.1**), affecting production of stone fruits and nuts, namely almond, hazelnut and walnut. Among these, walnut yield losses caused by *X. arboricola* diseases are currently considered to be a significant problem causing negative socio-economic implications worldwide, and particularly in Europe (Boudon et al. 2005, Strange and Scott 2005, Oerke 2006, Hajri et al. 2010, Palacio-Bielsa et al. 2010, Scortichini 2010, Lamichhane 2014, Lamichhane and Varvaro 2014).

For many years, bacterial walnut diseases caused by *X. arboricola* strains, namely those acknowledged as belonging to pathovar *juglandis* (acronym *Xaj*), were neglected in comparison to diseases severely spread and impacting central European countries, caused by other *X. arboricola* pathovars, particularly pv. *pruni* and pv. *corylina* affecting the stone fruit trees, almond and hazelnut trees, respectively. In fact, *X. arboricola* pv. *pruni* and *X. arboricola* pv. *corylina* are considered pathogens of serious concern, being strictly regulated by the European Union and the European Plant Protection Organization (EPPO) countries as quarantine pests for *Prunus* spp. and *Corylus* spp., as an effort to prevent their spread in the European region (pv. *pruni* EPPO A2 list no. 62, EPPO, 2006; pv. *corylina* EPPO A2 list no. 134, EPPO, 2004; www.eppo.int). In contrast, *Xaj* is considered by EPPO as a non-regulated pest for *Juglans* spp., mainly because it is a widespread pathogen in Europe, affecting the commercial values of marketable nut fruits but without major short-term consequences for walnut growing trees (Belisario et al. 1999, EPPO, 2001).

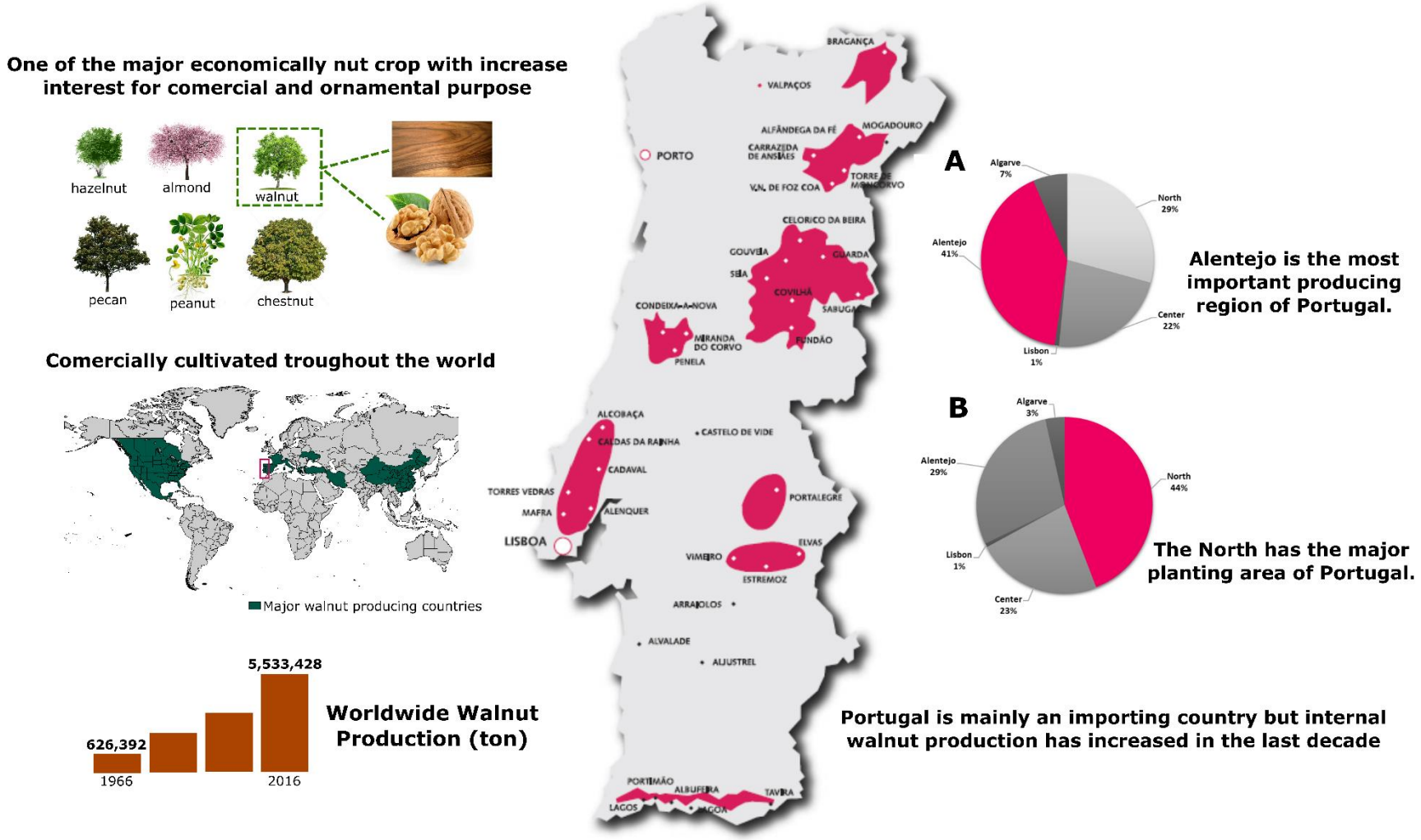

Figure I.1. Global and national distribution of walnut production. The increasing worldwide interest for walnut crop is related with the high economic value of walnut wood and fruits acknowledged for its phytotherapeutic and cosmeceutical properties (Blomhoff et al. 2006, Pereira et al. 2007, Lamichhane 2014). Walnut orchards are mainly centered in temperate climates regions, but can also be found in western south America, northern Africa and Australia (Mikulic-Petkovsek et al. 2011). The main production countries are China, the USA, Mexico, Iran, and several European countries, particularly of the Mediterranean basin, namely Ukraine, Turkey, Italy, Greece, France, Spain and Portugal, which have contributed for the worldwide increase of walnut production (FAO 2016, official data available). Walnut production in Portugal is observed throughout their continental territory (Portugal map adapted from Soares 2005). The most representative production geographic regions are Bragança, Mirandela, Condeixa, Estremoz, Beja and Portalegre. In **A** is represented in percentage the national walnut production for a total of 4299 tons, across the different agricultural regions in 2016. In **B** is represented in percentage the national walnut planting area for a total of 3303 hectares, across different agricultural regions in 2016 (INE 2016, official report).

The recent outbreaks of emerging walnut pathologies caused by *Xaj* (currently known as Walnut Bacterial Blight, WBB; Brown Apical Necrosis, BAN; and Vertical Oozing Canker, VOC) occurring in the main walnut producing countries, including the EPPO region (**Figure I.2**), have enhanced the efforts to understand the biology of *X. arboricola* strains associated with walnut trees. Unsurprisingly, it has been acknowledged as extremely important to increase the knowledge about the environmental and host-related factors modulating these walnut diseases (Lamichhane, 2014). Furthermore, it became obvious the urgency to detail the biology of walnut-associated *Xanthomonas* pathogens, which are probably not-restricted to *Xaj* according to data discussed in this dissertation (Fernandes et al. 2018a, 2018b, 2018d), and are still poorly understood, namely regarding diversity, taxonomy, identification, ecology, genetics and molecular determinants of pathoadaptations to walnut hosts. The scientific knowledge in all these perspectives will be instrumental to understand the epidemiology of walnut diseases caused by xanthomonads, and, consequently, be decisive to make assertive risk assessment analysis and improve phytosanitary practices.

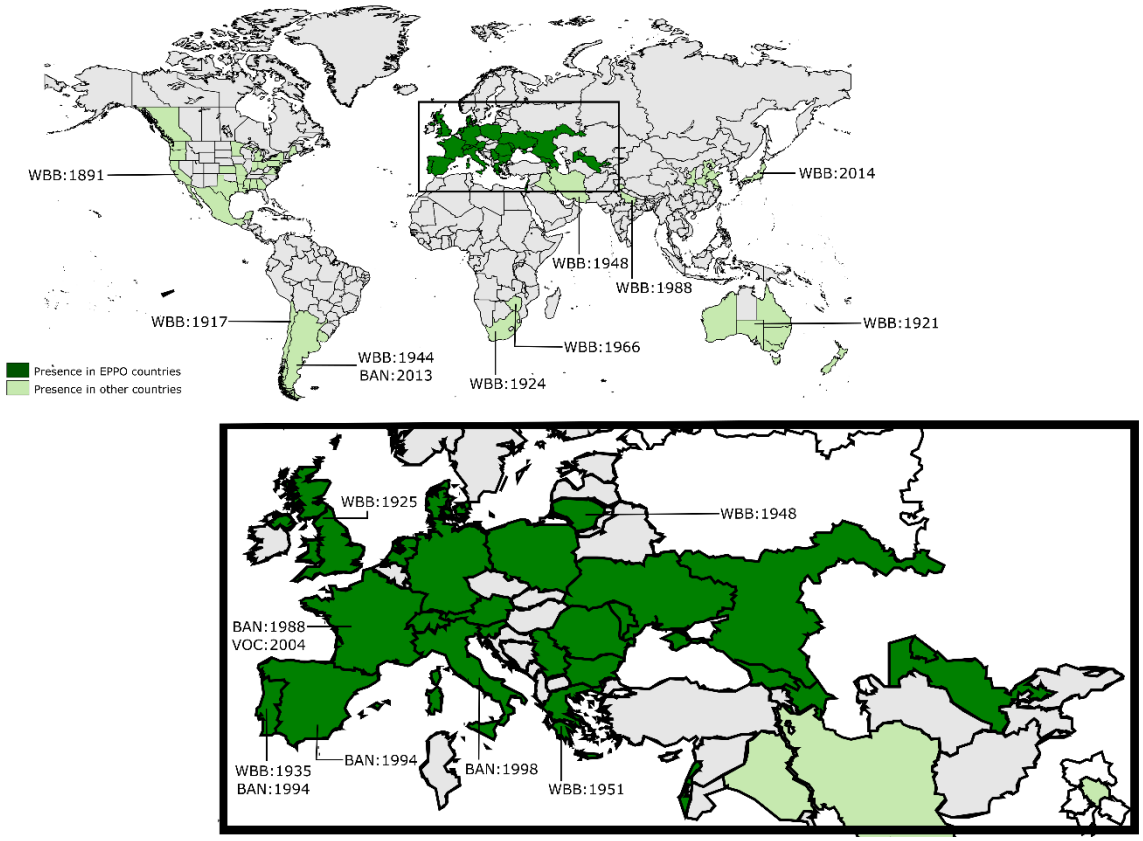

Figure I.2. Worldwide distribution map of walnut diseases (WBB, BAN and VOC) caused by *Xanthomonas arboricola* (adapted and modified from: CABI-EPPO 2001). On top, the dates of the first reports of WBB, BAN and/or VOC diseases are displayed for relevant non-european walnut producing countries, namely Argentina (Flores et al. 2004, Temperini et al. 2017), Australia (Osborn and Samuel 1922), Chile (Esterio and Latorre 1982, Rudolph 1933), India (Adhikari et al. 1988, Sharma and Sharma 1999), Iran (Armani 1977, Shami et al. 2013), Japan (Sawada et al. 2017), South Africa (Doidge 1924, Duplessis and Westhuizen 1995), USA (Smith 1912) and Zimbabwe (Rothwell 1975). The European map below details the years when walnut diseases (WBB, BAN and VOC) were first reported for several countries across europe, namely France (Charlot and Penet 1988, Garcin and Duchesne 2001, Menard et al. 2004, Hajri et al. 2010), Greece (Rouskas et al. 2006, Vagelas et al. 2012), Italy (Belisario et al. 2002), Lithuania (Burokiene and Pulawska 2012), Portugal (Oliveira and Cabral 1942, Martins et al. 1997, Fernandes et al. 2017), Spain (Aletà and Ninot 2002, Arquero et al. 2005, Moragrega and Ozaktan 2010) and United Kingdom (Woemald 1930).

1.1. Walnut disease caused by *Xanthomonas*

Although walnut trees are susceptible to many pests and diseases (Ginzel 2010, Sharma et al. 2012), this plant species has been continuously affected by diseases caused by *Xanthomonas* sp. in general and *Xaj* in particular (Frutos 2010, Lamichhane 2014). As pointed out above, these walnut diseases are widespread around the world (CABI-EPPO 2001), and follow the distribution of walnut trees and the rapid development of commercial *Juglans regia* cultivars, the main species used for cultivation and commercial trade (Smith et al. 1912, Ercisli

et al. 2012, Frutos and López 2012). It is acknowledged that *Xaj* is responsible for a variety of disease symptoms that affect walnut tree health (**Figure I.3**).

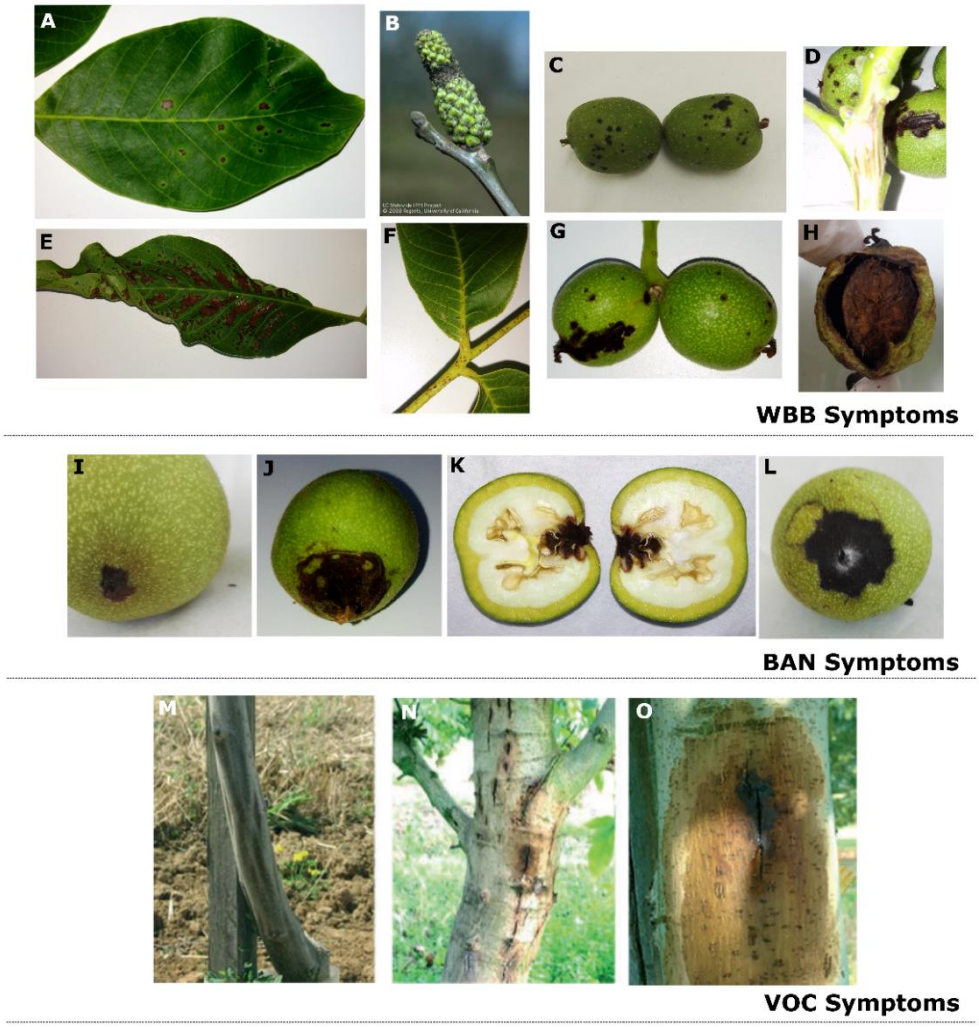

Figure I.3. Common symptoms observed in walnut trees due to *Xaj* infection. **A.** to **H.** Characteristic symptoms of Walnut Bacterial Blight (WBB) disease: symptoms on leaves are characterized as small, circular and water-soaked areas, which can evolve to angular lesions surrounding by a chlorotic halo (**A**) or confluent lesions causing large necrotic zones (**E**). Symptoms typically extend along the petioles and veins of leaflets (**F**). Symptoms on fruits start on the external immature green husk where translucent or water-soaked circular spots can be observed. As the fruits grow, black spots enlarges (**C**) or become depressed expanding on the fruit surface area (**D, G**). Lesions can lead to problems in nut development as shape deformations or cause necrosis on the walnut shell (**H**). Furthermore, infected catkins are deformed and necrotic (**B**) and infected flowers normally abscise (Pierce 1901, Smith et al. 1912, Woemald 1930, Woeste et al. 1992, Frutos and López 2012, Lang and Evans 2010). **I.** to **L.** Characteristic symptoms of Brown Apical Necrosis (BAN) disease: the initial symptoms consist of a small brown non-water-soaked and circular lesion, which appears restricted near the stigma after fruit set (**I**). Symptoms progress as red-brown irregular external lesions (**J**) and can gradually grow within the internal tissues, with dry and hard necrotic lesions from the epicarp to the mesocarp, until reach the kernel (**K**). Symptoms detected at early fruit development stages are linked to drop of immature fruits causing more damage. Dropped fruit can be superficially covered by fungal mycelium (**L**) (Smith et al. 1912, Moragrega and Ozaktan 2010, Moragrega et al. 2010, 2011) **M.** to **O.** Characteristic symptoms of Vertical Oozing Canker (VOC) disease: symptoms observed on trunk and branches consisted in longitudinal deformation (**M**) and the development of vertical cankers (**N**). Early canker infections can evolve leading from brown to black oozing exudates that normally stain the bark in late summer (**O**) (Menard et al. 2004, Hajri et al. 2010, Lamichhane 2014).

The most well-known walnut disease caused by *Xaj* is Walnut Bacterial Bight (WBB), affecting all new growing organs of the year in walnut trees (Pierce 1901, Smith et al. 1912, Belisario et al. 1999) (**Figure I.3 A-H**). WBB was first reported in 1891 and rapidly spread across the walnut-growing regions of California and into other USA states where walnut was commercially grown (Smith et al. 1912). WBB was considered a common disease of walnut, causing reductions in walnut fruit quality and tree vigour up to 70% (Ninot et al. 2002, Solar et al. 2012, Lamichhane 2014).

It was in the late 1990s that *X. arboricola* became the most threatening pathogen affecting walnut, most likely due to changes in walnut cultural management systems (Scotton et al. 2015). As especially observed in Europe, where yield and fruit quality of walnut increased greatly, due to, intensive monoculture orchards with implementation of pruning, fertilization, irrigation, and chemical treatments (Scotton et al. 2015). These practices are common in most agricultural ecosystems found nowadays, leading to a more vulnerable environment for pathogens adaptation and coevolution into more virulent crop pathogens (McDonald and Stukenbrock 2016, Jacques et al. 2016). This could be the case of *Xaj* and the appearance of new emerging diseases on walnut, namely Brown Apical Necrosis (BAN) and Vertical Oozing Canker (VOC).

BAN disease was the name used to describe the necrosis symptoms observed in the stigma scar of walnut fruits and premature fruit drop occurring in managed orchards (Belisario et al. 2002, Moragrega et al. 2011, Scotton et al. 2015) (**Figure I.3 I-L**). It was precisely in Europe that the first consequences of BAN were reported. In Italy, a reduced yield of up to 20% was observed due to walnut fruits dropping before reaching maturity (Belisario et al. 2002). In Spain, total yield loss was documented to occur in some walnut orchards (Aletà and Ninot 2002, Moragrega et al. 2011). Moreover, evidence of premature fruit drop by infection in intensive commercial walnut orchards has also been recorded outside the European continent: in several Australian orchards almost all walnut production was lost due to drop of infected walnut young fruits (Lang et al. 2005, Lang and Evans 2010) and, more recently, Argentine producers also registered losses of up to 40% caused by BAN (Temperini et al. 2017).

VOC is the most recently described walnut disease attributed to X*aj* (**Figure I.3 M-O**). It is acknowledged to affect mainly trunks and branches of old and young walnut trees in orchards, and young walnut trees in nurseries (Hajri et al. 2010). VOC disease has been described as limited to walnut trees settled in France, being known that the disease had progressed rapidly throughout French walnut cultivation regions, since it was first reported in 2004 (Menard et al. 2004). Although, no substantial reports have been available concerning the current situation of VOC disease, future negative economic impacts have been predicted, especially due to the propagation of diseased walnut plants in nurseries potentially leading to a devastation of infected young orchards (Hajri et al. 2010).

1.2. Control of walnut diseases

The effects of *Xaj* and other walnut pathogenic *Xanthomonas* have on walnut orchards are evident worldwide, particularly the reduction of marketable of nuts, the reduction of orchard productivity and the increase in nursery production costs (Radix et al. 1998, Lamichhane 2014). These effects lead to question orchard management practices, since walnut diseases caused by *Xanthomonas spp.*, mainly *Xaj*, are becoming uncontrolled (Moragrega and Ozaktan 2010). Proper cultural practices, including improvement of cropping systems, orchards sanitation, management of crop residues and elimination of potential biotic reservoirs, contribute to prevent or mitigate walnut's infection (Smith et al. 1912, Swings and Civerolo 1993, Agrios 2005, Moragrega and Ozaktan 2010, Lamichhane 2014). However, these improvements are mainly dependent on several key aspects related to bacterial epidemiology that are still largely unknown. The use of resistant walnut cultivars could also be an important control measure of walnut disease-associated to *Xanthomonas* spp. since all available walnut cultivars are susceptible to *Xaj* (Swings and Civerolo, 1993, Belisario et al. 1999, Frutos and López 2012, Vagelas et al. 2012). Nonetheless, the efficient selection of resistant cultivars is particularly challenging as long as epidemiological information is still limited.

Common agrochemical products based on copper compounds, which are usually required to reduce crop losses caused by *Xaj* diseases in walnut orchards (Swings and Civerolo 1993, Smith et al. 1912, Lamichhane 2014, Agrios 2005b), are becoming increasingly ineficient, as shown by the emergence of copper resistant *Xaj* strains (Gardan et al. 1993, Lee et al. 1993, Scortichini et al. 2001, Pereira et al. 2015). Moreover, the widespread use of copper-based compounds in commercial walnut orchards worldwide has led to negative effects on the ecosystem (Radix et al. 1998, Ninot et al. 2002, Agrios 2005b, Ozaktan et al. 2012).

Since efficient alternatives to nonspecific bacteriostatic copper-based treatments are not yet available (Lang et al. 2005), there is an increasing interest for novel disease management strategies based in biological control (McNeil et al. 2001, Ninot et al. 2002, Frampton et al. 2012, Ozaktan et al. 2012, Romero-Suarez et al. 2012, Lamichhane 2014). In fact, validation for the use of bacterial antagonists isolated from healthy walnut trees (Ozaktan et al. 2012), or for the use of lytic bacteriophages specific to *Xaj* is compulsory (Romero-Suarez et al. 2012, Retamales et al. 2016).

2. Taxonomy of *Xanthomonas* spp. and the unfinished classification of *X. arboricola*

Almost exclusively associated with plants, *Xanthomonas* spp. (Dowson 1939) as a whole are able to cause a variety of disease symptoms in a large number of plant hosts, covering around

400 plant species, including monocots and dicots (Leyns et al. 1984, Swings and Civerolo 1993, Bergey et al. 2005). However, the host range of a single *Xanthomonas* species is frequently delimited to a unique or a few plants species from the same botanical family (Jacques et al. 2016). *Xanthomonas* spp. strains display the same or similar biochemical and physiological features, which brings problems concerning the systematics of the genus *Xanthomonas*, since a taxonomy based in phenotyping methods appeared to be unattainable to discriminate *Xanthomonas* species (Baltrus 2016).

The definition of bacterial species within *Xanthomonas* has been a matter of long unsolved debates, and inspired multiple concepts, frequently based on pathogenic phenotypes and symptoms on host plants, as evidenced by the following examples: “new host – new species” by Burkholder and Starr 1948; “genetic species” by Friedman and De Ley 1965; “single species” by Dye and Lelliot 1974; and “pathovar” by Young et al. 1978. Advances in bacterial genotyping methodologies showed that these definitions are inadequate to characterize properly the genetic diversity and evolutionary relationships of *Xanthomonas* (Baltrus 2016). The first transition to a polyphasic system improved *Xanthomonas* systematics and taxonomy, bringing both genomic and phylogenetic data and culminating in the reclassification of xanthomonads in fifteen different *Xanthomonas* species (Vauterin et al. 1995). However, this classification system continued to use the term pathovar, which is not validated by the official rules of the International Code of Nomenclature of Bacteria (Schaad 1987, Lapage et al. 1992). Pathovar, which is used for further differentiation of *Xanthomonas* species at intrasubspecific level, is based exclusively on pathogenic traits, as host specificity, disease symptoms and/or infection mechanisms (Bull et al. 2008).

To date, plant pathologists continue to routinely use the pathovar concept, mainly because classifying bacterial strains having the same phytopathogenic profile as a specific pathovar facilitates the dissemination of important information, namely basic pathogen knowledge, epidemic outbreaks, specific management solutions and control measures (Bull et al. 2008). Nevertheless, a classification dependent on pathogenicity and host range of plant bacterial pathogens has clear limitations. Pathogenicity tests are based on Koch’s postulates (Lelliot and Stead 1987) and hold practical difficulties, such as being laborious, time consuming, dependent on artificial inoculation methods under laboratory set-ups, occasionally resulting in ambiguous results, and limited by the culturability of strains (Baltrus 2006, Jacques et al. 2016, Rademarker et al. 2006). These facts imply that *Xanthomonas* strains, displaying a distinct pathogenicity pattern or isolated from asymptomatic plant hosts, are often difficult to assign into a pathovar (Rademarker et al. 2006).

Xanthomonas arboricola and its infrasubspecific division into pathovars was proposed in 1995 by Vauterin and collaborators during the revision of *Xanthomonas* taxonomy (Vauterin et

al. 1995). Mainly stirred by the well-known host range of *X. arboricola* strains, pathovars were defined as the subspecies level for the group of strains causing disease on the same plant species. Furthermore, *X. arboricola* was considered as a homogeneous pathogenic species group, differing mostly by the range of susceptible hosts and disregarding the mostly unknown asymptomatic hosts (Jacques et al. 2016).

In the classification of Vauterin *et al.* (1995), *Xanthomonas arboricola* was divided into six pathovars: pv. *juglandis* for walnut pathogenic strains; pv. *celebensis*, the banana pathogen; pv. *corylina*, the halzenut pathogen; pv. *poinsettiicola*, the poinsettia pathogen; pv. *populi*, the poplar pathogen; and pv. *pruni*, the stone fruit and almond pathogen. Later, four additional pathovars were proposed: in 2001, the pv. *fragariae*, the causative agent of strawberry bacterial leaf blight (Janse et al. 2001); and in 2015 the pathovars *arracaciae* (the arracacha pathogen), *guizotiae* (the niger pathogen) and *zantedeschiae* (the arum lily pathogen), all three previously classified as pathovars of *X. campestris* (Fischer-Le Saux et al. 2015) (**Figure I.4**).

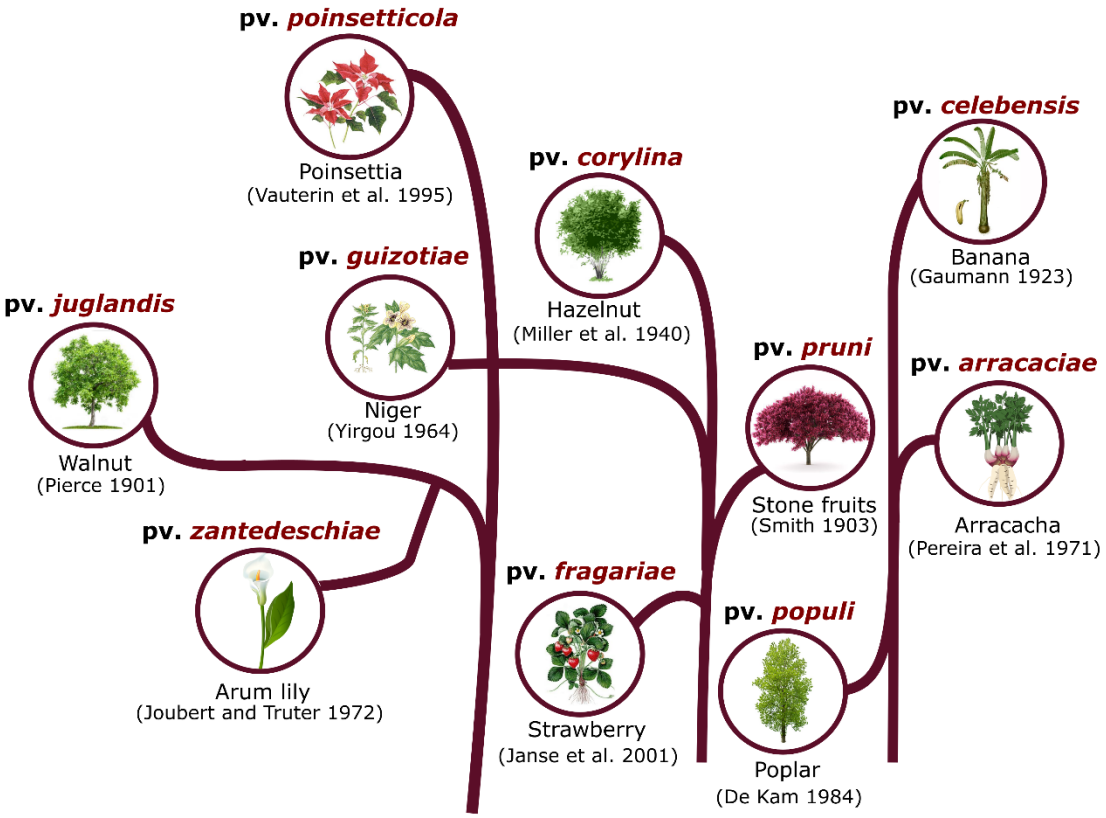

Figure I.4. Representation of the ten pathovars presently accepted for the species *Xanthomonas arboricola*, with reference to the first report and the host plants species, which includes ornamental plants (niger, arum lily, poinsettia and poplar), fruit trees (banana), stone fruit trees (peach, apricot and Japanese plum), nut trees (almond, hazelnut and walnut) and herbaceous plants (arracacha and strawberry).

In general, there are no consistent and simple phenotypic tests for specific identification of *X. arboricola* pathovars. Some phenotypic properties, namely quinate (Lee et al. 1992) and nicotinic acid (Starr 1946) synthesis; xanthomonadin pigmentation (Andrewes et al. 1976); colonial morphology (Gitaitis et al. 1988); and β-glucosidase activity (Hayward 1977), were proposed to differentiate groups of pathovars (Bergey et al. 2005), however, these assays were not pathovar specific. Even though classification of *X. arboricola* strains is strongly dependent on the identification of their hosts (Bergey et al. 2005), the pathovar concept is not assertive, and does not reflect the genomic diversity and phylogeny of strains within this species. Additionally, not all *X. arboricola* pathogens are assigned to pathovars, namely the *X. arboricola* pathogenic strains of cereals, brassicas and sunflower (Ignatov et al. 2010), jujube (Myung et al. 2010a), bell pepper (Myung et al. 2010b), grapevine (Sawada et al. 2011), mango (Qi et al. 2012), onion and chrysanths (Young et al. 2010, Fischer-Le Saux et al. 2015), pecan (Fernandes et al. 2018e, further detailed in chapter VI), and *Arabidopsis thaliana* (Wang et al. 2018). This list also includes *X. arboricola* with unsolved phylogeny and uncertain pathogenicity isolated from poinsettia, strawberry, magnolia, onion, clove, kohekohe, happy wanderer, liquindambar and chrysanths (Vandroemme et al. 2013; Essakhi et al. 2015); nonpathogenic strains isolated from both healthy

and diseased plant organs of walnut trees (Essakhi et al. 2015) and plum trees (Garita-Cambronero et al. 2016a,b); as well as strains pathogenic on barley, which are capable to infect other plant species (Ignatov et al. 2015).

Whole genome sequencing has been used in the last years to propose new *Xanthomonas* species (Rodriguez-R et al. 2012, Triplett et al. 2015, Vicente et al. 2017, López et al. 2018). Particularly, the average nucleotide identity (ANI) has been used to support the proposition of new *Xanthomonas* species, being generally acknowledged as a trustworthy alternative to the DNA-DNA hybridization-based classification proposed in 1995 by Vauterin et al. (Konstantinos and Tiedje 2004, Rodriguez-R et al. 2012, Baltrus, 2016, Vinatze et al. 2017). ANI has been widely used as a complement for taxonomic classification and phylogenetic inferences, and helped to solve the inconsistencies and weaknesses of a classification into pathovars, since it provides comprehensive data to address the whole diversity between members of a single bacterial species (Baltrus 2016, Vinatze et al. 2017). Interestingly, recommendations for the definition of new species supported by ANI values already exist for *Xanthomonas* strains isolated from either stone fruits or walnut trees, which displayed similar symptoms to the diseases caused in these hosts by *X. arboricola* pv. *pruni* and pv. *juglandis*, respectively (López et al. 2018, Fernandes et al. 2018a, 2018b, 2018d - further detailed in Chapter IV and V).

Eventually, many other bacterial strains will be reclassified following the taxonomic changes observed in the *Xanthomonas* genus. Clearly, a reliable classification for *Xanthomonas* spp. is a step forward for the implementation of more precise detection and identification methods, reducing the difficulties presently faced by plant pathologists to diagnose diseases caused by xanthomonads.

3. Plant disease epidemiology and pathogen populations – *X. arboricola* pv. *juglandis* as a case study

Epidemiology in plant pathology is generally defined as the study of diseases in plant populations, taking into account their dynamics in time and space, and also the analysis of biotic and abiotic factors that could strongly influence the disease progression (Milgroom and Peever 2003, Agrios 2005a). Comprehensive knowledge of the interactions that occur between populations of a virulent phytopathogen with susceptible host plants in a favourable environment are essential to accurately predict the progression of plant diseases, the occurrence of new outbreaks and the design of more assertive prevention and control practices (Milgroom and Peever 2003, Agrios, 2005a). However, during the course of this thesis work, it was clear that this complex pathogen-host-environment triangle continues to be poorly integrated on studies of plant disease caused by many economically important phytopathogens (Fernandes et al. 2018a).

Therefore, most of the existing studies make inferences about the genetic diversity of plant pathogen populations solely through the application of numerous molecular techniques, frequently dismissing the importance of ecological traits, and traits related with the host and the environment (Milgroom 2001, Agrious 2005a). A few representative examples can be found on diversity studies of *Xaj* populations (Loreti et al. 2001, Marcelletti et al. 2010, Burokiene and Pulawska 2012, Kaluzna et al. 2014). In these studies, a considerable genetic variability was observed on *Xaj*, which appears to be quite large unlike other *X. arboricola* pathovars (Zaccardeli et al. 1999, Scortichini et al. 2002, Boudon et al. 2005). However, this diversity has not yet been related with environmental factors aside from the host plant or a few strain's features, such as the geographic origin and the year of isolation. This means that the biological significance of this genetic diversity has not been fully reached. Many of the diversity studies have emerged in parallel with advances on genotyping approaches and the development of more polymorphic markers, giving the perception that they were produced to apply new methodologies rather than use the most appropriate method to address important questions (Milgroom 2001). Additionally, these studies are frequently biased since they rely on a sampled population comprised by isolates collected without a planned sampling strategy or by strains from laboratory bacterial collections, which cannot be considered representative. In this context, it is worth mentioning that studies on field populations of bacterial plant pathogens normally consider the selection of a single bacterial isolate in pure culture per tree, loosing adequate epidemiological information. Furthermore, many of these studies are based on a low number of isolates, obtained during a short time period comprised by one growing season, centered on specific walnut orchards and a narrow geographic territory (Du Plessis and Van der Westhuizen 1995, Marcelletti et al. 2010, Hajri et al. 2010, Shami et al. 2013, Ivanovic et al. 2014, Kaluzna et al. 2014, Giovanardi et al. 2015). This implies that the collection of isolates is planned based in the clonal population concept of pathogen populations, which is often acknowledged in plant pathology (Lamichhane and Venturi 2015). Other studies use bacterial collections addressing the genetic diversity of strains across continents, which although informative to get a glimpse into the genetic heterogeneity present in a single bacterial species, may be misleading or inappropriate to make epidemiologic inferences. In these works, the studied bacterial populations are artificial, since they gather a set of worldwide strains unrelated in time and space, neglecting important epidemiological data (Loreti et al. 2001, Scortichini et al. 2001, Barionovi and Scortichini 2008). Not surprisingly, studies covering different types of *Xaj* populations have shown apparently contradictory findings. For instance, while Scortichini *et al.* (2001) showed that *Xaj* strains from a culture collection are genetically grouped according their geographic location, Hajri et al. (2010) assumed that genomic heterogeneity found in *Xaj* field isolates was not related with the

geographic origin. The merit of these studies is to provide evidence for other unforeseen factors beyond the geography of strain's origin.

Future studies of *Xaj* diversity need to direct efforts towards solving these shortcomings in order to select a set of strains representative of a given population complemented by epidemiologically relevant metadata. The choice of isolates is a crucial step as a bad sampling methodology may lead to errors and incorrect conclusions. Therefore, a sampling process with relevant raw genetic material is extremely important for the insights and interpretation that can be made in a given study (Wolf and Knot 1982, Milgroom and Fry 1997, Milgroom, 2001). Knowledge about the pathogen's biology and disease cycle can help to delineate a correct sampling procedure (Milgroom and Fry 1997). Although this knowledge about *Xaj* is limited (**Figure I.5**), the perception that diseases occurs at an ecological and environmental scale, which can influence pathogen's populations, indicates that future genetic population studies should take into account environmental metadata to improve the way in which the collection of isolates is conducted and the genetic data is interpreted (Archie et al. 2009). Therefore, the integration of genotyping and epidemiological metadata in population studies is instrumental to determine pathogen population structure (Milgroom 2001). Moreover, metadata analysis, integrating concepts of evolution, ecology and genetics, could help to determine disease patterns in plant pathology and improve control strategies (Archie et al., 2009, Milgroom 2001). This type of analysis has been successfully implemented in clinical microbiology and is an important framework to answer relevant epidemiological questions (Parkhill and Wren 2011; McMahon and Denaxas 2016).

① Walnut dormant buds have been identified as the most probable source of primary inoculum, where *Xaj* can overwinter and multiply without causing disease symptoms.

② It is during bud opening and developing shoots that *Xaj* probably multiplies on the surface of young leaves.

③ Early-infected leaves are probably sources of secondary inoculum that is probably carried by rainwater to infect other portions of the tree, including developing fruits, or other walnut trees. Dissemination can also probably occur by infected pollen and leaf-feeding insects.

④ It is not known how *Xaj* spread internally into the host tissues for establishment of infection, but it is hypothesized that plant surface breaches occuring naturally or caused by mechanic injuries, could be common entry sites.

⑤ *Xaj* is frequently found in asymptomatic walnut tissues and probably survives for long periods in different overwintering sites of walnut trees that serve as source of new inoculum.

⑥ There is little evidence of epiphytic *Xaj* population found on leaves of a non-host plant (hazelnut) and on symptomless weeds. There is no information about *Xaj* survival in the environment, as in plant debris or in the soil.

⑦ Specific environmental conditions during spring, as warm temperature, rainfall patterns and high relative humidity, may play an important role on *Xaj* dissemination, penetration, multiplication and growth.

Figure I.5. Current knowledge on the lyfe cycle of *Xaj* in walnut hosts and its relevance to set off disease. Sources: Smith et al. 1912, Miller 1934, Ark 1944, Miller 1946, Esterio and Latorre 1982, Mulrean and Schroth 1982, Pruvost and Gardan 1988, Swings and Civerolo 1993, Duplessis and Vanderwesthuizen 1995, Belisario et al. 1999, Garcin and Duchesne 2001, Agrios 2005a, Melotto et al. 2008, Moragrega and Ozaktan 2010, Scortichini 2010, Soltani and Aliabadi 2010, Frutos and López 2012, Solar et al. 2012, Vorholt 2012, Lamichhane 2014, Lindow et al. 2014.

4. Importance of DNA-based methods for *X. arboricola* pv. *juglandis* detection, identification and typing

Nowadays, bacterial detection and identification techniques are not just required to precisely determine the pathogen of interest, but they also need to confidently discriminate at intrasubspecific level, allowing the identification of different lineages; to enable a rapid and early diagnostic of the disease; and to be useful for assessing pathogen load in infected samples. Complementarily, typing techniques are currently essential for an in depth assessment of the pathogen, in order to determine the microbial population structure; and to identify inoculum sources; transmission routes and biotic reservoirs. Ideally, typing techniques should allow efficient phytosanitary surveillance, which demands high-throughput screening, contributing to infer epidemiological patterns related to pathogenesis, host and environmental adaptation, ultimately capable to instruct and evaluate efficient control measures (Louws et al. 1999, Agrios 2005a, Strange and Scott 2005, Belkum et al. 2007).

The conventional methods considered as reliable and efficient for routine diagnosis of *X. arboricola* strains have important limitations and do not reach all the above-mentioned criteria. Particularly, the methods currently applied rely on the culturability of organisms and require additional confirmation to assign strains at the pathovar level (Moragrega et al. 2002, Janse 2010, EPPO 2004, 2006). In this context, the development and validation of new methods for *Xaj* detection and identification are of unquestionable need. The diagnosis of walnut diseases caused by *Xaj* is generally based on the visual examination of the symptoms, followed by direct isolation of the pathogen in semi or non-selective media (Mulrean and Schroth 1981, McGuire et al. 1986, Lelliott and Stead 1987, Schaad 2001, Moragrega et al. 2002, Janse 2010). Classical microbiological methods (Schaad et al. 2001, Janse 2010) can be applied, but no information exists about specific phenotyping methods for pv. *juglandis*, as well as, for the other pathovars of *X. arboricola* species (Moragrega et al. 2002). PCR primers are only available for detection at the genus level (Leite et al. 1994, RST1-RST2 primers) or species level (Pothier et al. 2011, XarbQ-F-XarbQ-R primers). In addition, sequence analysis of genes commonly used to address phylogenies, such as 16S rRNA, or other housekeeping genes, e.g. DNA gyrase subunit B (*gyrB*), frequently used as a molecular identification tools, display a low-resolution to differentiate *X. arboricola* pathovars (Hauben et al. 1997, Parkinson et al. 2007, 2009). Thus, the inoculation of susceptible walnut plants or immature walnut fruits (pathogenicity tests) is still the only procedure available for the allocation of *X. arboricola* strains into pathovar *juglandis* (Moragrega et al. 2002, Janse 2010). Numerous typing technologies have been used for the genetic characterization of *Xaj* populations, namely Amplified Fragment Length Polymorphism-based

analysis (Loreti et al. 2001, Hajri et al. 2010); genomic fingerprinting of repetitive sequences by PCR using REP, ERIC and BOX primers (rep-PCR) (Scortichini et al. 2001, Burokiene and Pulawska 2012, Shami et al. 2013, Ivanovic et al. 2014, Kaluzna et al. 2014, Giovanardi et al. 20159); Pulsed-Field Gel Electrophoresis (Ivanovic et al. 2014); PCR Melting Profile (i.e. PCR MP) (Kaluzna et al. 2014); Multilocus Sequence Typing and Multilocus Sequence Analysis (Marcelletti et al. 2010, Burokiene and Pulawska 2012 - Kaluzna et al. 2014, Giovanardi et al. 2015); and Multilocus Variable-Number Tandem-Repeat Analysis (Essakhi et al. 2015, Giovanardi et al. 2015); however, it is important to recall that these genotyping techniques do not substitute the need for previous identification of strains as belonging to *Xanthomonas arboricola*. Although these techniques provide useful information, they cannot be used to detect pv. *juglandis* strains and the suitability of obtained data for future genotyping studies is highly debatable (Olive and Bean 1999, Shurch and Soolingen 2011, Ranjbar et al. 2014). Most of these molecular typing techniques rely on electrophoretic analysis of polymorphic DNA regions are characterized by their low reproducibility and the inherent difficulties to interpret complex electrophoretic band profiles, therefore strongly limiting the accuracy of those methods. The difficult standardization across laboratories makes unfeasible the analysis of numerous samples directing their utility mostly for short-term studies (Olive and Bean 1999, Parkinson et al. 2007, Ranjbar et al. 2014). Alternatively, the approaches based on sequencing analysis of housekeeping genes (Parkinson et al. 2007, Young et al. 2008), provide higher genotyping resolution but require the analysis of a considerable number of loci to ensure a high discriminatory potential (Ranjbar et al. 2014). Although, a large number of housekeeping gene sequences are now available in public databases, making multilocus sequencing methods (MLSA and MLST) the most preferable approaches for genotyping, the target genes are generally too highly conserved to be directly used for detection and identification of *Xaj* populations. Additionally, these genes are usually not related to pathogenesis or virulence, thus being unsuitable to make direct epidemiological inferences (Belkum et al. 2007).

DNA-based methods have been recurrently proposed as a helpful alternative for diagnostic laboratories and epidemiologic studies (Schaad and Frederick 2002, Albuquerque et al. 2009). Particularly for phytopathology, DNA-based methods have circumvented many of the limitations of the traditional methods for detection purposes (Ward et al. 2004, Vincelli and Tisserat 2008, Albuquerque et al. 2012a, Fang and Ramasamy 2015), showing attributes that could be also advantageous for the improvement of molecular techniques in *Xaj* detection, identification and typing.

4.1. Selection of specific DNA markers

The application of newly developed molecular biology techniques, for use in diagnostics or as epidemiological tools, is usually hampered by the selection of appropriate DNA markers for a target pathogen (Lang et al. 2010). In plant pathology, the difficulty in finding specific and unique genomic regions for a particular target bacterial taxa, favoured the use of different genetic regions for DNA marker design depending on the bacterial species (Lang et al. 2010). In fact, genes commonly used for phylogenetic inferences, namely the 16S and 23S rRNA genes (Ludwig and Schleifer 1994, Ruppitsch et al. 2007); housekeeping genes, particularly *gyrB* (Parkinson et al. 2009) and *rpoD* (Yamoto et al. 2000); and repetitive chromosomal regions, such as REP, BOX and ERIC (Alves et al. 2004); and genes coding for specific bacterial traits (Schönfeld et al. 2003, Kang et al. 2008, Park et al. 2009, Park et al. 2010, Pothier et al. 2011), have all been used in DNA-based detection methods for the detection of different *Xanthomonas* species. However, most of these methods can now be considered outdated, firstly because they are unsuitable for validation against extensive collections of bacteria and secondly due to their limited reliability when targeting a bacterial taxa characterized by a high genotypic diversity (Vieira et al. 2007, Albuquerque et al. 2009, 2012a). This is overly evident for *X. arboricola* species characterized by a broad and genetically heterogeneous number of bacterial strains, for which the available DNA markers are not specific for pathovar discrimination, and surely not reliable to overcome the intrapathovar genetic variability between strains. Moreover, the mentioned DNA markers were not designed to be used simultaneously in the same assay.

The opportunity to select new and improved *Xaj* DNA markers became more attainable in the last few years with the release of the first *Xaj* genome (strain NCPPB 1447, accession number PRJNA84273) and with the advancements of strategies to easily screen different genomic regions as potential informative molecular markers (Schaad and Frederick 2002, Vieira et al. 2007, Albuquerque et al. 2009, 2012a). Indeed, bioinformatics tools and dedicated *in silico* pipelines have been successfully used for the selection of numerous novel DNA markers for diverse bacterial species (Almeida et al. 2013, Albuquerque et al. 2017, Almeida et al. 2018), including several phytopathogens namely *Pseudomonas syringae* pv. *tomato* (Vieira et al. 2007); *Xanthomonas fragariae*, *Xanthomonas axonopodis* pv. *phaseoli* and *Xanthomonas fuscans* subsp. *fuscans* (Albuquerque et al. 2011); *Xanthomonas euvesicatoria* (Albuquerque et al. 2012b); and *Ralstonia solanacearum* (Albuquerque et al. 2015). The mentioned studies demonstrated that the design of highly specific DNA markers relies on an accurate *in silico* prediction on the region's stability and specificity, without the need to have a comprehensive knowledge of the pathogen's biology and metabolism, neither requiring an extensive and laborious experimental validation. Most importantly, these studies emphasized the adequateness

to use these DNA signatures simultaneously in hybridization-based methods increasing the detection reliability of a specific target pathogen, but also to infer the genetic diversity among different strains of the target taxa.

4.2. Implementation of PCR and hybridization-based techniques

As previously stated, the number of studies of phytopathogens focused on PCR and hybridization-based detection, identification and typing are increasing daily, highlighting the advantages of these techniques over conventional methods used in plant pathology (Henso and French 1993, Schena et al. 2004, Palacio-Bielsa 2009, Pelludat et al. 2009, Albuquerque et al. 2012b, Adriko et al. 2014). Plant pathologists adopted PCR-based assays as a highly sensitive, rapid and easy alternative for early detection of phytopathogens directly in infected plant samples, i.e. without the need to isolate the etiologic bacterial agent in pure culture (Verdier et al. 2001, Vincelli and Tisserat, 2008, Palacio-Bielsa 2009, Albuquerque et al. 2012b). PCR assays also contribute to increase the throughput and efficiency of routine diagnosis when amplification and analysis are carried out in the same assay (Schaad and Frederick 2002, Diallo et al. 2009). Multiplex and quantitative PCR approaches allow the simultaneous detection of phytopathogens of distinct species, or different strains of the same species (Glick et al. 2002), as well as, an estimation of pathogen's cell numbers (Schena et al. 2004, Palacio-Bielsa et al. 2011). Alternatively, hybridization-based technologies are commonly used to screen simultaneously a large number of samples and have been currently investigated as a reliable and promising alternative platform for typing workflows (Amann et al. 2001, Vieira et al. 2007, Hull 2014, Albuquerque et al. 2015). In comparison with PCR methods, hybridization methods are a better solution as high throughput assays, and the hybridizing DNA marker reflects the sequence similarity with the target in all its length, contrary to PCR assays that are fully dependent on primers hybridization occurring during the annealing, which may lead to false negative results. With the development of an automatic image analysis application to quantify the hybridization signals, allowing statistical analysis, the main drawback of hybridization-based methods has been overcomed (Belkum et al. 2007, Caridade et al. 2010).

Acknowledging the advances that these tools would represent regarding the detection, identification and typing of walnut-associated *Xanthomonas* in general, and *Xaj* in particular, deep research efforts to improve the early diagnostics of walnut diseases caused by *Xanthomonas* as well as to provide epidemiologic insights, have been developed during this PhD project and are detailed in this dissertation (Fernandes et al. 2017, Martins et al. 2019, Fernandes et al. 2018a, further detailed in chapter II, III and IV, respectively).

5. Genomic era in *Xanthomonas arboricola*

Over the past two decades, high-throughput DNA sequencing technologies allowed major advances in bacterial genomics and opened new opportunities in plant pathology. Almost two decades have passed since the first full genome sequence of a phytopathogenic bacteria strain (*Xylella fastidiosa* 9a5c responsible for citrus variegated chlorosis) was published (Simpson et al. 2000). From this landmark, an endless number of genomes of phytopathogens have been released, including more than one thousand belonging to *Xanthomonas* species (NCBI database, accessed in October 2018). Regarding *X. arboricola*, the first draft genome of a *Xaj* strain (NCPPB 1447, accession number PRJNA84273) was made available in 2012. Presently *X. arboricola* is represented by 72 genomes (**Table I.1**), with the latest genome announcements published in August 2018 (NCBI database, accessed in October 2018).

Table I.1. *Xanthomonas arboricola* genomes currently available.

Strain	Pathogenicity, Significance	Genome accession	Genome Size (Mb)	GC%	Scaffolds	Genes	Proteins	Ref.
X. arboricola								
3004	pathogenic on barley	AZQY	4.76	66.0	132	4171	3829	[1]
CFBP 1022	nonpathogenic isolated from walnut	MDRU	5.09	65.6	43	4385	4155	-
CFBP 7610		MDRV	4.95	65.6	46	4362	4135	-
CFBP 7614		MIGG	4.92	65.5	46	4266	4031	-
CFBP 7622		MIGF	4.75	66.2	218	4146	3845	-
CFBP 7629	nonpathogenic isolated from walnut	MIGI	4.94	65.6	62	4335	4079	-
CFBP 7634	nonpathogenic isolated from walnut	JZEH	4.93	65.6	4	4320	4071	[2]
CFBP 7645	nonpathogenic isolated from walnut	MIGY	4.75	65.5	11	4248	3966	-
CFBP 7651	nonpathogenic isolated from walnut	JZEI	5.03	65.5	6	4405	4149	[2]
CFBP 7652	nonpathogenic isolated from walnut	MIGJ	5.22	65.5	78	4535	4300	-
CFBP 7653	nonpathogenic isolated from walnut	MIGK	4.86	66.1	77	4247	4002	-
CFBP 8130		MIGL	4.76	66.0	78	4115	3892	-
CFBP 8132		MIGM	4.93	65.7	44	4326	4101	-
CFBP 8138		MIGN	4.76	66.0	57	4087	3882	-
CFBP 8142		MIGO	4.84	65.8	55	4193	3967	-
CFBP 8147		MIGP	4.88	65.7	106	4256	4007	-
CFBP 8149		MIGQ	4.87	65.7	33	4242	4031	-
CFBP 8150		MIGR	4.92	65.8	352	4241	3851	-
CFBP 8152		MIGS	4.72	65.5	200	4163	3754	-
CFBP 8153		MIGT	4.83	65.9	78	4146	3938	-
CFBP 6827		MDDX	4.95	65.7	5	4285	4070	-
CFBP 7604		MDDY	4.84	65.8	6	4229	4012	-
CFBP 7697		MDDZ	4.95	65.7	8	4300	4087	-
CITA 124	avirulent isolated from peach	LXKK	4.75	65.8	128	4218	3845	[3]
CITA 14	avirulent isolated from peach	LXIB	4.86	65.6	72	4208	3922	[3]
CITA 44	nonpathogenic isolated from cherry	LJGM	4.76	65.8	71	4151	3757	[4]
FOR_F20	pathogenic on *Arabidopsis thaliana*	PUPZ	5.27	64.3	48	4632	4144	[5]
FOR_F21	pathogenic on *Arabidopsis thaliana*	PUPY	5.27	64.2	11	4696	4102	[5]
FOR_F23	pathogenic on *Arabidopsis thaliana*	PUPX	5.26	64.2	8	4689	4115	[5]
FOR_F26	pathogenic on *Arabidopsis thaliana*	PUPW	5.26	64.2	8	4683	4118	[5]
MEDV_A37	pathogenic on *Arabidopsis thaliana*	PUPR	4.87	65.8	5	4282	3886	[5]
MEDV_P39	pathogenic on *Arabidopsis thaliana*	QREQ	4.88	65.9	19	4301	3795	[5]
MEU_M1	pathogenic on *Arabidopsis thaliana*	QRER	5.11	64.0	27	4483	4039	[5]
NL_P126	pathogenic on *Arabidopsis thaliana*	QREM	4.73	66.4	24	4100	3802	[5]

Strain	Pathogenicity, Significance	Genome accession	Genome Size (Mb)	GC%	Scaffolds	Genes	Proteins	Ref.
PLY_3	pathogenic on *Arabidopsis thaliana*	PUPV	5.26	64.3	7	4710	4140	[5]
PLY_4	pathogenic on *Arabidopsis thaliana*	PUPU	5.25	64.2	13	4766	4030	[5]
PLY_9	pathogenic on *Arabidopsis thaliana*	PUPT	5.27	64.2	11	4721	4155	[5]
pv. *arracaciae*								
CFBP 7407	pathogenic on arracacha	MIGU	5.16	65.2	232	4517	4193	-
pv. *celebensis*								
NCPPB 1630	pathogenic on banana	JPHE	4,98	65.5	7	4301	4045	[6]
NCPPB 1832	pathogenic on banana	JPHC	4,90	65.6	3	4278	4036	[6]
pv. *corylina*								
CFBP 1159	pathogenic on hazelnut	MDEA	5.10	65.5	124	4432	4079	-
CFBP 2565	pathogenic on hazelnut	MDSJ	5.05	65.6	98	4503	4145	-
NCCB 100457	pathogenic on hazelnut	APMC	5.22	65.5	281	4596	4185	[7]
pv. *fragariae*								
CFBP 6762	pathogenic on strawberry	OEQE	4.88	65.7	38	4283	4050	-
CFBP 6771	pathogenic on strawberry	MIGH	4.91	65.9	56	4239	4021	-
CFBP 6773	pathogenic on strawberry	OEQD	4.69	65.9	39	4089	3875	-
LMG 19144	pathogenic on strawberry	OEQF	4.84	65.9	58	4166	3935	-
LMG 19145	pathogenic on strawberry	OEQL	4.90	65.9	39	4252	4020	-
LMG 19146	pathogenic on strawberry	OEQG	4.88	65.7	39	4262	4024	-
pv. *guizotiae*								
CFBP 7408	pathogenic on niger	MDSK	4.93	65.6	161	4390	3955	-
CFBP 7409	pathogenic on niger	MDSL	4.82	65.9	219	4292	3860	-
pv. *juglandis*								
CFBP 2528	pathogenic on walnut causing WBB	JZEF	5.08	65.5	8	4451	4123	[2]
CFBP 7179	pathogenic on walnut causing VOC	JZEG	5.15	65.4	14	4544	4222	[2]
CFBP 8253	pathogenic on walnut	MDSM	4.99	65.6	91	4340	4009	-
CPBF 1521	pathogenic on walnut, not exposed to copper treatments	UIHD	5.19	65.4	56	4636	4212	[8]
DW3F3	pathogenic on walnut	PNRC	5.14	65.4	8	4522	4186	[9]
J303	pathogenic on walnut	LSGZ	5.06	65.5	161	4409	4095	-
NCPPB 1447	pathogenic on walnut	AJTL	5.02	65.4	371	4543	3972	-
Xaj 417	pathogenic on walnut, copper-resistant	CP012251.1	5.21	65.4	1	4532	4186	[10]
CFSAN033077	pathogenic on walnut	LHBK	5.10	65.4	197	4353	4125	[11]
CFSAN033078	pathogenic on walnut	LHBL	5.11	65.6	215	4357	4137	[11]
CFSAN033085	pathogenic on walnut	LHBS	5.11	65.6	196	4351	4131	[11]
CFSAN033086	pathogenic on walnut	LHBT	5.14	65.4	205	4387	4155	[11]
pv. *populi*								
CFBP 3122	pathogenic on poplar	MIGV	4.78	65.4	92	4218	3868	-
CFBP3123	pathogenic on poplar	MDEB	4.65	65.5	28	4214	3918	-
pv. *pruni*								
CFBP 3894	pathogenic on plum	LOMI	5.05	65.4	77	4423	4063	-
IVIA 2626.1	pathogenic on plum	LJGN	5.02	65.4	214	4431	3925	[4]
MAFF301420		BAVC	5.00	65.3	396	5215	4729	-
MAFF301427		BAVD	4.90	65.4	356	5180	4686	-
MAFF311562		BAVB	5.08	65.3	432	5422	4897	-
Xap33	pathogenic on almond	JHUQ	5.10	65.4	474	4236	3759	[12]
pv. *zantedeschiae*								
CFBP 7410	pathogenic on arum lily	MIGW	4.97	65.5	37	4331	4099	-

Reference (Ref.): [1]: Ignatov et al. 2015; [2]: Cesbron et al. 2015; [3]: Garita-Cambronero et al. 2016b; [4]: Garita-Cambronero et al. 2016a; [5]: Wang et al. 2018; [6]: Harrison et al. 2016; [7]: Caballero et al. 2013; [8]: Fernandes et al. 2018c; [9]: Fu et al. 2018; [10]: Pereira et al. 2015; [11]: Higuera et al. 2015; [12]: Garita-Cambronero et al. 2014.

This increasingly amount of publicly available *Xanthomonas* genomes rose in concert with studies of comparative genomics, which is one of the most powerful approaches currently used to unveil the biology of a bacterium (Schneider and Collmer, 2010, Miller et al. 2014). In fact, analysis of xanthomonads genomic sequences has provided major insights to understand

chromosomal structure particularly regarding genome rearrangements likely involved in adaptation, such as genes duplications, genes losses and genes acquisition (Booher et al. 2015, Gordon et al. 2015, Zhang et al. 2015). Expectably, full genome comparative analyses have also facilitated evolutionary studies of the *Xanthomonas* genus, mainly by addressing the evidence of horizontal gene transfer and recombination, translated in hypothesis-driven phylogenies (Rodriguez-R et al. 2012, Aritua et al. 2015, Huang et al. 2015). In addition, comparative genomics of *Xanthomonas* sp. strains has been used to identify strain-specific features and genomic differences to infer determinants of pathogenicity and virulence, which may help to disclose different pathogenic strategies and mechanisms (da Silva et al. 2002, Ryan et al. 2011, Jacobs et al. 2015, Jaques et al. 2016). Specifically in *X. arboricola,* recent comparative genomics studies of pathogenic and nonpathogenic strains provided the first insights into genome dynamics and stability associated with insertion sequences and mobile genetic elements, as well as, shared and unique genomic features, including putative virulence-related factors such as extracellular polysaccharides, plant cell wall-degrading enzymes, adhesins and protein secretion systems (Cesbron et al. 2015, Garita-Cambronero et al. 2016a, 2017). The most remarkable differences observed among *X. arboricola* strains were mainly in gene content of important virulence determinants, particularly the type III secretion system and its secreted effectors (T3SS and T3Es), which are known to create a favourable environment for bacterial plant colonization and infection, and are involved in the inhibition of host plant defences (Buttner and Bonas 2010, Cesbron et al. 2015, Garita-Cambronero et al. 2016a, 2017, 2018, Denance et al. 2016, Jacques et al. 2016).

Presently, it is undeniable that comparative genomic studies of xanthomonads are opening new opportunities providing important insights into bacterial lifestyle biology, host specificity and adaptation processes, including pathogenicity and virulence. For *X. arboricola*, genomic comparisons between strains characterized by distinct pathogenicity and virulence phenotypes generated a long list of putative virulence genes and virulence regulators, that can be useful in functional studies and comparative host-pathogen interaction experiments to investigate the molecular basis of host infection and pathogenicity, as well as disease progression (Schneider and Collmer 2010, Ryan et al. 2011, Garita-Cabronero et al. 2018). We are still at the childhood of comparative genomics of *X. arboricola* and further challenges lie ahead. For example, it would be interesting to determine and analyse the genome sequences of bacterial clones, representative of the diversity within the species, including strains from the same tree showing to be distinctly adapted to different plant niches; strains showing a higher survival rate outside their common host; strains from environmental reservoirs or from putative vectors; or strains revealing a higher evolutionary fitness in the existing subpopulations. Since it is unrealistic and unworthy to sequence the genome of every single isolate to ensure a complete coverage of

genomic determinants of pathogenicity, host and niche-specific adaptation in *X. arboricola*, the main difficulty is to select the strains most representative of the genetic diversity, which, once again, can only be achieved thorough epidemiological studies.

Thesis framework and objectives

This PhD project started to be developed while research efforts performed to contain walnut diseases caused by *Xaj* were focused on the selection of the most resistant walnut cultivars (He et al. 2010, Frutos and López 2012, Vagelas et al. 2012); in the improvement of copper-based treatments (Burchner et al. 2010); and on the development of predictive disease models (Giraud et al. 2010). It was most recently, probably related with the poor knowledge available of the etiological agent, together with the advances of genotyping approaches, that molecular studies have emerged to address the diversity of *Xaj* in order to identify evolutionary trends related with pathogenicity, virulence fitness, dissemination, and biogeographic distribution (Shami et al. 2013, Ignatov et al. 2014, Kaluzna et al. 2014, Giovanardi et al. 2015). However, these efforts have been hampered mainly due to the lack of informative and reliable molecular markers and the use of appropriated molecular methods to disclose *Xaj* population structure, pathogenicity traits, host and environmental adaptations.

The release of the first draft genome of the *Xaj* strain NCPPB 1447 (accession number PRJNA84273) on October 2012, unleashed the opportunity to select and validate new *Xaj*-specific DNA-markers and implement DNA-based methods for detection and genotyping primarily aiming to improve the assessment of *Xaj*, including their diversity, population structure and genomic dynamics, but also suitable for comprehensive phytosanitary surveys of walnut-associated *Xanthomonas*, that may instruct better phytosanitary practices regarding the control of epidemic walnut diseases caused by *Xanthomonas*.

Overall, the main objectives of the work presented in this dissertation were:

i) To select and validate novel *Xaj* pathovar-specific molecular markers;
ii) To implement culture-independent methods for *Xaj* detection and identification;
iii) To implement PCR-based techniques to quantify *Xaj* bacterial load from walnut infected plant tissues;
iv) To characterize the genetic diversity of a large number of field *Xanthomonas* isolates obtained from walnut trees over a three-year period considering epidemiological metadata including, host-related features, thermoclimatic factors and geographic locations;
v) To identify non-walnut hosts as possible plant reservoirs for *Xaj.*
vi) To compare genome sequences of five *Xanthomonas* strains isolated simultaneously (i.e. at the same sampling time point) and longitudinally (i.e. at different sampling time points) from a single diseased walnut tree.

Thesis publications

The current dissertation is composed by the following original articles, published, submitted or in preparation to be submitted, in international scientific journals:

Fernandes C, Albuquerque P, Sousa R, Cruz L, Tavares F. 2017. Multiple DNA markers for identification of *Xanthomonas arboricola* pv. *juglandis* isolates and its direct detection in plant samples. Plant Disease 101 (6): 858-865. DOI: 10.1094/PDIS-10-16-1481-RE (**Chapter II**)

Martins L, **Fernandes C**, Albuquerque P, Tavares F. 2019. Assessment of *Xanthomonas arboricola* pv. *juglandis* bacterial load in infected walnut fruits by qPCR. *Plant Disease*. In Press. DOI:10.1094/PDIS-12-18-2253-RE (**Chapter III**)

Fernandes C, Albuquerque P, Cruz L, Tavares F. 2019. Comprehensive diversity assessment of walnut associated *Xanthomonas arboricola* reveal distinct lineages within the same walnut tree and unveil a new *Xanthomonas* sp. Submitted. (**Chapter IV**)

Fernandes C, Blom J, Pothier JF, Tavares F. 2018. High-quality draft genome sequence of *Xanthomonas* sp. strain CPBF 424, a walnut pathogenic strain with atypical features. Microbiology Resource Announcements 7 (15): e00921-18. DOI: 10.1128/MRA.00921-18 (**Chapter V**)

Fernandes C, Blom J, Pothier JF, Tavares F. 2018. High-quality draft genome sequence of *Xanthomonas arboricola* pv. *juglandis* CPBF 1521, isolated from leaves of a symptomatic walnut tree in Portugal without a past of phytosanitary treatment. Microbiology Resource Announcements 7 (16): e00887-18. DOI: 10.1128/MRA.00887-18. (**Chapter V**)

Fernandes C, Martins L, Blom J, Pothier JF, Tavares F. 2018. Comparative genomic of *Xanthomonas* strains isolated from a single walnut tree host. In Preparation. (**Chapter V**)

Fernandes C, Sousa R, Tavares F, Cruz L. 2018. First report of *Xanthomonas arboricola* causing bacterial blight on pecan trees in Portugal. Plant Disease 102 (12): 2632. DOI: 10.1094/PDIS-03-18-0467-PDN. (**Chapter VI**)

References

Adhikari R, Bora S, Singh S. 1988. *Xanthomonas campestris* pv. *juglandis* - a new report from India. *Current Science* **57**, 728.

Adriko J, Mbega ER, Mortensen CN, Wulff EG, Tushemereirwe WK, Kubiriba J, Lund OS. 2014. Improved PCR for identification of members of the genus *Xanthomonas*. *European Journal of Plant Pathology* **138**, 293-306.

Agrios GN. 2005. Plant Pathology (Fifth Edition). San Diego, California: Academic Press.

Alabouvette C, Olivain C, Steinberg C. 2016. Biological control of plant diseases: the european situation. *European Journal of Plant Pathology* **114**, 329-341.

Albuquerque P, Mendes MV, Santos CL, Moradas-Ferreira P, Tavares F. 2009. DNA signature-based approaches for bacterial detection and identification. *Science of the Total Environment* **407**, 3641-3651.

Albuqerque P, Caridade C, Marçal ARS, Cruz J, Cruz L, Santos CL, Mendes MV, Tavares F. 2011. Identification of *Xanthomonas fragariae*, *Xanthomonas axonopodis* pv. *phaseoli*, and *Xanthomonas fuscans* subsp. *fuscans* with novel markers and using a dot blot platform coupled with automatic data analysis. *Applied and Environmental Microbiology* **77**, 5619-5628.

Albuquerque P, Rodrigues AS, Caridade CMR, Marçal ARS, Tavares F. 2012a. An *in silico* workflow for the selection of novel DNA markers for detection of phytopathogenic bacteria, using *Agrobacterium tumefaciens* as a model. *Journal of Plant Pathology* 94, S1.83-S1.90.

Albuquerque P, Caridade CRM, Rodrigues AS, Marçal ARS, Cruz J, Cruz L, Snatos CL, Mendes MV, Tavares F. 2012b. Evolutionary and experimental assessment of novel markers for detection of *Xanthomonas euvesicatoria* in plant samples. *PLoS ONE* **7**, e37836.

Albuquerque P, Marçal ARS, Caridade CRM, Costa R, Mendes MV, Tavares F. 2015. A quantitative hybridization approach using 17 DNA markers for identification and clustering analysis of *Ralstonia solanacearum*. *Plant Pathology* **64**, 1270-1283.

Albuquerque P, Ribeiro JN, Almeida A, Panshin I, Porfirio A, Vales M, Diniz F, Madeira H, Tavares F. 2017. Application of a dot blot hybridization platform to assess *Streptococcus uberis* population structure in dairy herds. *Frontiers in Microbiology* **8**, 54.

Aletà N, Ninot A. 2002. La bactériose et la nécrose apicale du noyer en Espagne. *InfoNOIX* **19**, 8. [In French]

Almeida A, Albuquerque P, Araujo R, Ribeiro JN, Tavares F. 2013. Detection and discrimination of common bovine mastitis-causing streptococci. *Veterinary Microbiology* **164**, 370-377.

Almeida E, Serra CR, Albuquerque P, Guerreiro I, Oliva-Teles A, Enes P, Tavares F. 2018. Multiplex PCR identification and culture-independent quantification of *Bacillus licheniformis* by qPCR using specific DNA markers. *Food Microbiology* **74**, 1-10.

Alves A, Henriques I, Fragoeiro S, Santos C, Phillips AJL, Correia A. 2004. Applicability of rep-PCR genomic fingerprinting to molecular discrimination of members of the genera *Phaeoacremonium* and *Phaeomoniella*. *Plant Pathology* **53**, 629-634.

Amann R, Fuchs BM, Behrens S. 2001. The identification of microorganisms by fluorescence in situ hybridisation. *Current Opinion in Biotechnology* **12**, 231-236.

Andrewes AG, Jenkins CL, Starr MP, Shepherd MP, Hope H. 1976. Structure of xanthomonadin 1, a novel dibrominated aryl-polyene pigment produced by the bacterium *Xanthomonas juglandis*. *Tetrahedron Letters* **45**, 4023-4024.

Archie EA, Luikart G, Ezenwa O. 2008. Infecting epidemiology with genetics: a new frontier in disease ecology. *Trends in Ecology and Evolution* **24**, 21-30.

Aritua V, Harrison J, Sapp M, Buruchara R, Smith J, Studholme DJ. 2015. Genome sequencing reveals a new lineage associated with lablab bean and genetic exchange between *Xanthomonas axonopodis* pv. *phaseoli* and *Xanthomonas fuscans* subsp. *fuscans*. *Frontries in Microbiology* **6**,1080.

Ark PA. 1944. Pollen as a source of walnut bacterial blight infection. *Phytopathology* **34**, 330-334.

Armani B. 1977. Bacterial blight of walnut in Iran. *Iranian Journal of Plant Pathology* **13**, 14-23 [In Persian]

Arquero O, Lovera M, Rodriguez R, Salguero A, Trapero A. 2005. Characterization and development of necrotic lesions of walnut tree fruits in southern Spain. *Acta Horticulturae* **705**, 457-461.

Baltrus DA. 2016. Divorcing strain classification from species names. *Trends in Microbiology* **24**, 431-439.

Bandi A, Hevesi M, Szani Z, Tóth M. 2015. Assessment of bacterial blight tolerance of persian walnut based on immature nut test. *Notulae Scientia Biologicae* **7**, 5.

Barionovi D, Scortichini M. 2008. Integron variability in *Xanthomonas arboricola* pv. *juglandis* and *Xanthomonas arboricola* pv. *pruni* strains. *FEMS microbiology letters* **288**, 19-24.

Belisario A, Maccaroni M, Corazza L, Balmas V, Valier A. 2002. Occurrence and etiology of brown apical necrosis on Persian (English) walnut fruit. *Plant Disease* **86**, 599-602.

Belisario A, Zoina A, Pezza L, Luongo L. 1999. Susceptibility of species of *Juglans* to pathovars of *Xanthomonas campestris*. *European Journal of Forest Pathology* **29**, 75-80.

Belkum A. Tassios PT, Dijkshoom L, Haeggman S, Cookson B, Fry NK, Fussing V, Green J, Feil E,Gerner-Smidt P, Brisse S, Struelens M. 2007. Guidelines for the validation and application of typing methods for use in bacterial epidemiology. *Clinical Microbiology and Infectious Diseases* **13**, 1-46.

Bergey DH. 2005. *Bergey's manual of systematic bacteriology. Volume 2. The proteobacteria. Part B. The gammaproteobacteria*. New York, New York: Springer.

Blomhoff R, Carlsen MH, Andersen LF, Jacobs DR, Jr. 2006. Health benefits of nuts: potential role of antioxidants. *The British Journal of Nutrition* **96**, S52-S60.

Booher NJ, Carpenter SCD, Sebra RP, Wang L, Salzberg SL, Leach JE, Bogdanove AJ. 2015. Single molecule real-time sequencing of *Xanthomonas oryzae* genomes reveals a dynamic

structure and complex TAL (transcription activator-like) effector gene relationships. *Microbial Genomics* **1**, e000032.

Boudon S, Manceau C, Notteghem JL. 2005. Structure and origin of *Xanthomonas arboricola* pv. *pruni* populations causing bacterial spot of stone fruit trees in Western Europe. *Phytopathology* **95**, 1081-1088.

Bull CT, De Boer SH, Denny TP, Firrao G, Fischer-Le Saux M, Saddler GS, Scortichini M, Stead DE, Takikawa Y. 2008. Demustifying the nomenclature of bacterial plant pathogens. *Journal of Plant Pathology* **90**, 403-417.

Burchner RP, Gilles C, Olson WH, Adaskaveg JE, Lindow SE, Koutsoukis R. 2010 Spray timing and materials for walnut blught (*Xanthomonas campestris* pv. *juglandis*, *Xanthomonas arboricola* pv. *juglandis*) control in Northen California USA. *Acta Horticulturae* **861**, 457-464.

Burkholder WH, Starr MP. 1948. The generic and specific characters of phytopathogenic species of *Pseudomonas* and *Xanthomonas*. *Phytopathology* **38**, 494–502.

Burokiene D, Pulawska J. 2012. Characterization of *Xanthomonas arboricola* pv. *Juglandis* isolated from walnuts in Lithuania. *Journal of Plant Pathology* **94**, S1.23-S1.27.

Buttner D, Bonas U. 2010. Regulation and secretion of *Xanthomonas* virulence factors. *FEMS Microbiology Review* **34**, 107-133.

Caballero JI, Zerillo MM, Snelling J, Boucher C, Tisserat N. 2013 Genome sequence of *Xanthomonas arboricola* pv. *corylina*, isolated from Turkish filbert in Colorado. *Genome announcements* **1**, e00246-13.

CABI-EPPO, 2001. *Xanthomonas arboricola* pv. *juglandis*. Distribution Maps of Plant Diseases, No. 133.

Caridade CRM, Marçal ARS, Albuquerque P, Mendes MV, Tavares F. 2015. Automatic analysis of dot blot images. *Intelligent Automation and Soft Computing* **21**, 1-16.

Cesbron S, Briand M, Essakhi S, Gironde S, Boureau T, Manceau C, Fischer-Le Saux M, Jacques MA. 2015. Comparative genomics of pathogenic and nonpathogenic strains of *Xanthomonas arboricola* unveil molecular and evolutionary events linked to pathoadaptation. *Frontiers in Plant Science* **6**, 1126.

Charlot G, Penet C. 1988. Compte rendu des travaux de recherche et expérimentations entrepris à le station de Chatte en 1988. *Rapport CTIFL 1988 Station de Chatte*, 122-123. [In French]

da Silva AC, Ferro JA, Reinach FC, Farah CS, Furlan LR, Quaggio RB, Monteiro-Vitorello MA, et al. 2002. Comparison of the genomes of two *Xanthomonas* pathogens with differing host specificities. *Nature* **417**, 459-463.

De Kam M. 1984. *Xanthomonas campestris* pv. *populi*, the causal agent of bark necrosis in poplar. *Netherlands Journal of Plant Pathology* **90**, 13-22.

Denancé N, Lahaye T, Noël DL. 2016. Editorial: Genomics and effectomics of the crop killer *Xanthomonas*. *Frontiers in Plant Science* **7**, 71.

Diallo S, Latour X, Groboillot A, Smadja B, Copin P, Orange N, Feuilloley MGJ, Chevalier S. 2009. Simultaneous and selective detection of two major soft rot pathogens of potato:

Pectobacterium atrosepticum (*Erwinia carotovora* subsp. *atrosepticum*) and *Dickeya* spp. (*Erwinia chrysanthemi*). *European Journal of Plant Pathology* **125**, 349-354.

Doidge EM. 1924. A preliminary check list of plant disease occurring in South Africa. *Memoirs of Botany, South Africa* **6**.

Dowson DW. 1939. On the systematic position and generic names of the gram negative bacterial plant pathogens. *Zentralblatt fur Bakteriologie, Parasitenkunde, Infektionskrankheiten und Hygiene* **100**, 177-193.

Duplessis HJ, Vanderwesthuizen TJ. 1995. Identification of *Xanthomonas campestris* pv. *juglandis* from (Persian) english walnut nursery trees in South Africa. *Journal of Phytopathology-Phytopathologische Zeitschrift* **143**, 449-54.

Dye DW, Lelliott RA. 1974. Genus II *Xanthomonas*. Manual of determinative bacteriology. Baltimore: Williams & Wilkins.

EPPO. 2001. *Xanthomonas arboricola* pv. *juglandis*. https://gd.eppo.int/taxon/XANTJU. Accessed October 2018.

EPPO. 2004. *Xanthomonas arboricola* pv. *corylina*. *EPPO bulletin* **34**, 179-81.

EPPO. 2006. *Xanthomonas arboricola* pv. *pruni*. *EPPO bulletin* **36**, 129-33.

Ercisli S, Sayinci B, Kara M, Yildiz C, Ozturk I. 2012. Determination of size and shape features of walnut (*Juglans regia* L.) cultivars using image processing. *Scientia Horticulturae* **133**, 47-55.

Essakhi S, Cesbron S, Fischer-Le Saux M, Bonneau S, Jacques MA, Manceau C, 2015. Phylogenetic and variable-number tandem-repeat analyses identify nonpathogenic *Xanthomonas arboricola* lineages lacking the canonical Type III Secretion System. *Applied and Environmental Microbiology* **81**, 5395-5410.

Esterio MA, Latorre BA. 1982. Potential Sources of Inoculum of *Xanthomonas juglandis* in walnut blight outbreaks. *Journal of Horticultural Science* **57**, 69-72.

Fang Y, Ramasamy RP. 2015. Current and prospective methods for plant disease detection. *Biosensors* **4**, 537-561.

FAO, 2016. FAO statistics. http://faostat.fao.org/site/339/default.aspx. Accessed January 2013.

Fernandes C, Albuquerque P, Sousa R, Cruz L, Tavares F. 2017. Multiple DNA markers for identification of *Xanthomonas arboricola* pv. *juglandis* isolates and its direct detection in plant samples. *Plant Disease* **101,** 858-865.

Fernandes C, Albuquerque P, Cruz L, Tavares F. 2018a. Comprehensive diversity assessment of walnut associated *Xanthomonas arboricola* reveal distinct lineages within the same walnut tree and unveil a new *Xanthomonas* sp.. Submitted.

Fernandes C, Blom J, Pothier JF, Tavares F. 2018b. High-quality draft genome sequence of *Xanthomonas* sp. strain CPBF 424, a walnut pathogenic strain with atypical features. *Microbiology Resource Announcements* **7**, e00921-18.

Fernandes C, Blom J, Pothier JF, Tavares F. 2018c. High-quality draft genome sequence of *Xanthomonas arboricola* pv. *juglandis* CPBF 1521, isolated from leaves of a symptomatic walnut

tree in Portugal without a past of phytosanitary treatment. *Microbiology Resource Announcements* **7**, e00887-18.

Fernandes C, Martins L, Blom J, Pothier JF, Tavares F. 2018d. Comparative genomic characterization of *Xanthomonas* strains isolated from a single walnut tree host. In preparation.

Fernandes C, Sousa R, Tavares F, Cruz L. 2018e. First report of *Xanthomonas arboricola* causing bacterial blight on pecan trees in Portugal. *Plant Disease* **102**, 2632.

Fischer-Le Saux M, Bonneau S, Essakhi S, Manceau C, Jacques MA, 2015. Aggressive emerging pathovars of *Xanthomonas arboricola* represent widespread epidemic clones distinct from poorly pathogenic strains, as revealed by multilocus sequence typing. *Applied and Environmental Microbiology* **81**, 4651-4668.

Flores P, Seta S, Gonzalez M, Coniglio R, Sferco S. 2004. Chemical and varietal management of walnut trees against walnut bacteriosis. *Research Journal of the Faculty of Agricultural Sciences* **5**, 25-31.

Frampton RA, Pitman AR, Fineran PC. 2012. Advances in bacteriophage-mediated control of plant pathogens. *International Journal of Microbiology* **2012**, 326452.

Friedman S, De Ley J- 1965. “Genetic Species” concept in *Xanthomonas*. *Jounal of Bacteriology* **89**, 95-100.

Frutos D, López G. 2012. Search for *Juglans regia* genotypes resistant/tolerant to *Xanthomonas arboricola* pv. *juglandis* in the fremework of COST Action 873. *Journal of Plant Pathology* **94**, 37-46.

Frutos D. 2010. Bacterial diseases of walnut and hazelnut and genetic resources. *Journal of Plant Pathology* **92**, S79-S85.

Fu B, Chen Q, Wei M, Zhu J, Zou L, Li G, Wang L. 2018. Complete genome sequence of *Xanthomonas arboricola* pv. *juglandis* strain DW3F3, isolated from a *Juglans regia* L. bacterial blighted fruitlet. *Genome Announcements* **6**, e00023-18.

Garcin A, Duchesne D. 2001. Walnut blight and apical necrosis. *Acta horticulturae* **544**, 379-387.

Gardan L, Brault T, Germain E. 1993. Copper resistance of *Xanthomonas campestris* pv. *juglandis* in french walnut orchards and its association with conjugative plasmids. *Acta Horticulturae* **311**, 259-65.

Garita-Cambronero J, Sena-Velez M, Palacio-Bielsa A, Cubero J. 2014 Draft genome sequence of *Xanthomonas arboricola* pv. *pruni* strain Xap33, causal agent of bacterial spot disease on almond. *Genome Announcements* **2**, e00440-14.

Garita-Cambronero J, Palacio-Bielsa A, López MM, Cubero J. 2016a. Draft genome sequence for virulent and avirulent strains of *Xanthomonas arboricola* isolated from *Prunus* spp. in Spain. *Standards in Genomic Sciences* **11**, 12.

Garita-Cambronero J, Palacio-Bielsa A, López MM, Cubero J. 2016b. Draft genome sequence of two strains of *Xanthomonas arboricola* isolated from *Prunus persica* which are dissimilar to strains that cause bacterial spot disease on *Prunus* spp. *Genome Announcements* **4**, e00974-16.

Garita-Cambronero J, Palacio-Bielsa A, López MM, Cubero J. 2017. Pan-genomic analysis permits differentiation of virulent and non-virulent strains of *Xanthomonas arboricola* that cohabit *Prunus* spp. and elucidate bacterial virulence factors. *Frontiers in Microbiology* **8**, 573.

Garita-Cambronero J, Palacio-Bielsa A, Cubero J. 2018. *Xanthomonas arboricola* pv. *pruni*, causal agent of bacterial spot of stone fruits and almond: its genomic and phenotypic characteristics in the *X. arboricola* species context. *Molecular Plant Pathology* **19**, 2053-2065.

Gaumann E. 1923. Onderzoekeningen over de bloedziekte der bananen op celebes I & II. . *Mededeelingen van het Instituut voor Plantenziekten* **59**, 47. [In Dutch]

Ginzel MD. 2010. Walnut insects: ecology and control. *Encyclopedia of pest management* **1**, 1-3.

Giraud M, Prunet JP, Chevallier A, Ramain S, Thiriaud V, Santrac I, Bray O. 2010. Study of *Xanthomonas arboricola* pv. *juglandis* population dynamics in French walnut orchards over three years. *Acta Horticulturae* **861**, 439-444.

Gitaitis RD, Hamm JD, Bertrand PF. 1988. Differentiation of *Xanthomonas campestris* pv. *pruni* from other yellow-pigmented bacteria by the refractive quality of bacterial colonies on an agar medium. *Plant Disease* **72**, 416–417.

Glick DL, Coffey CM, Sulzinski MA. 2002. Simultaneous PCR detection of the two major bacterial pathogens of Geranium. *Journal of Phytopathology* **150**, 54–59.

Gordon JL, Lefeuvre P, Escalon A, Barbe V, Curveiller S, Gagnevin L, Pruvost O. 2015. Comparative genomics of 43 strains of *Xanthomonas citri* pv. *citri* reveals the evolutionary events giving rise to pathotypes with different host ranges. *BMC Genomics* **16**,1098.

Hajri A, Meyer D, Delort F, Guillaumes J, Brin C, Manceau C. 2010. Identification of a genetic lineage within *Xanthomonas arboricola* pv. *juglandis* as the causal agent of vertical oozing canker of Persian (English) walnut in France. *Plant Pathology* **59**, 1014-22.

Harrison J, Grant MR, Studholme DJ. 2016. Draft genome sequences of two strains of *Xanthomonas arboricola* pv. *celebensis* isolated from banana plants. *Genome Announcements* **4**, e01705-15.

Hauben L, Vauterin L, Swings J, Moore ERB. 1997. Comparison of 16S ribossomal DNA sequences of all *Xanthomonas* species. *International Journal of Systematic Bacteriology* **47**, 328-335.

Hayward AC. 1977. Occurrence of glycoside hydrolases in plant pathogenic and related bacteria. *Journal of Applied Bacteriology* **43**, 407–411.

He F, Wang H, Zhanf Z, Zhao S, Zhao Y, Gao Y. 2010. Identification of walnut cultivars with AFLP fingerprinting. *Acta Horticulturae* **861**, 151-154.

Henson JM, French RC. 1993. The polymerase chain reaction and plant disease diagnosis. *Annual Review of Phytopathology* **31**, 81-109.

Higuera G, González-Escalone N, Véliz C, VeraF, Romero J. 2015. Draft genome sequences of four *Xanthomonas arboricola* pv. *juglandis* strains associated with walnut blight in Chile. *Genome Announcements* **3**, e01160-15.

Huang CL, Pu PH, Huang HJ, Sung HM, Liaw HJ, Chen YM, Chen CM, Huang MB, et al. 2015. Ecological genomics in *Xanthomonas*: the nature of genetic adaptation with homologous recombination and host shifts. *BMC Genomics* **16**, 188.

Hull R. 2014. Assay, detection and diagnosis of plant viruses. In: Plant virology (Figth edition). London, UK: Academic Press.

Ignatov AN, Kyrova EI, Vinogradova SV, Kamionskaya AM, Schaad NW, Luster DG. 2015. Draft genome sequence of *Xanthomonas arboricola* strain 3004, a causal agent of bacterial disease on Barley. *Genome Announcements* **3**, e01572-14.

INE I.P. 2016. Estatísticas agrícolas 2016. Edição 2017. Instituto Nacional de Estatística. Statistics Portugal. Lisboa- Portugal. [In Portuguese]

Ivanović Ž, Popović T, Janse J, Kojić M, Stanković S, Gavrilović V, Fira D. 2014. Molecular assessment of genetic diversity of *Xanthomonas arboricola* pv. *juglandis* strains from Serbia by various DNA fingerprinting techniques. *European Journal of Plant Pathology* **141**, 133-145.

Jacobs JM, Pesce C, Lefeuvre P, Koebnik R. 2015. Comparative genomics of a cannabis pathogen reveals insight into the evolution of pathogenicity in *Xanthomonas*. *Frontiers in Plant Sciences* **6**, 431.

Jacques MA, Arlat M, Boulanger A, Boureau T, Carrère S, Cesbron S, Chen NWG, Cociancich S, Darrasse A, Denancé N, Fischer-Le Saux M, Gagnevin L, Koebnik R, Lauber E, Noel LD, Pieretti I, Portier P, Pruvost O, Rieux A, Robène MR, Szurek B, Verdier V, Vernière C. 2016. Using ecology, physiology, and genomics to understand host specificity in *Xanthomonas*: French Network on Xanthomonads (FNX). *Annual Review of Phytopathology* **54**, 163-187.

Janse J. 2010. Diagnostic methods for phytopathogenic bacteria of stone fruits and nuts in COST 873. *EPPO Bulletin* **40**, 68-85.

Janse JD, Rossi MP, Gorkink RFJ, Derks JHJ, Swings J, Janssens D, Scortichini M. 2001. Bacterial leaf blight of strawberry (*Fragaria* (x) *ananassa*) caused by a pathovar of *Xanthomonas arboricola*, not similar to *Xanthomonas fragariae* Kennedy & King. Description of the causal organism as *Xanthomonas arboricola* pv. *fragariae* (pv. nov., comb. nov.). *Plant Pathology* **50**, 653-665.

Joubert JJ, Truter SJ. 1972. A variety of *Xanthomonas campestris* pathogenic to *Zantedeschia aethiopica*. *Netherlands Journal of Plant Pathology* **78**, 212-217.

Kałużna M, Pulawska J, Waleron M, Sobiczewski P. 2014. The genetic characterization of *Xanthomonas arboricola* pv. *juglandis*, the causal agent of walnut blight in Poland. *Plant Pathology* **63**, 1404-1416.

Kang MJ, Shim JK, Cho MS, Seol YJ, Hahn JH, Hwang DJ, Park DS. 2008. Specific detection of *Xanthomonas oryzae* pv. *oryzicola* in infected rice plant by use of PCR assay targeting a membrane fusion protein gene. *Journal of Microbiology and Biotechnology* **18**, 1492–1495.

Konstantinidis KT, Tiedje JM. 2005. Genomic insights that advance the species definition for prokaryotes. *Proceedings of the National Academy of Science of the United States of America* **102**: 2567-2572.

Lamichhane JR. 2014. *Xanthomonas arboricola* diseases of stone fruit, almond, and walnut trees: Progress toward understanding and management. *Plant Disease* **98**, 1600-10.

Lamichhane JR, Varvaro L. 2014. *Xanthomonas arboricola* disease of hazelnut: current status and future perspectives for its management. *Plant Pathology* **63**, 243-54.

Lamichhane JR, Venturi V. 2015. Synergisms between microbial pathogens in plant disease complexes: a growing trend. *Frontiers in Plant Science* **6**, 385.

Lang JM, Hamilton JP, Diaz MGQ, Van Sluys MA, Burgos MRG, Cruz CMV, Buell CR, Tisserat NA, Leach JE. 2010. Genomics-based diagnostic marker development for *Xanthomonas oryzae* pv. *oryzae* and *X. oryzae* pv. *oryzicola*. *Plant Disease* **94**, 311-319.

Lang MD, Evans KJ. 2010. Epidemiology and status of walnut blight in Australia. *Journal of Plant Pathology* **92**, S49-S55.

Lang MD, Hills JL, Evans KJ. 2005. Preliminary studies toward managing walnut blight in Tasmania, Australia. *Acta Horticulturae* **705**, 451-6.

Lapage SP, Sneath PHA, Lessel EF, Skerman VBD, Seeliger HPR, Clark WA. 1992. International Code of Nomenclature of Bacteria. Bacterial Code, 1990 Revision. Washington (DC): ASM Press.

Lee YA, Hildebrand DC, Schroth MN. 1992. Use of quinate metabolism as a phenotypic property to identify members of Xanthomonas campestris DNA homology group 6. *Phytopathology* **82**, 971-973.

Lee YA, Schroth MN, Hendson M, Lindow SE, Wang XL, Olson B, Buchner RP, Teviotdale B. 1993. Increased toxicity of iron-amended copper-containing bactericides to the walnut blight pathogen *Xanthomonas campestris* pv. *juglandis*. *Phytopathology* **85**, 1460-1465.

Leite RP, Minsavage GV, Bonas U, Stall RE. 1994. Detection and identification of phytopathogenic *Xanthomonas* strains by amplification of DNA sequences related to the hrp genes of *Xanthomonas campestris* pv. *vesicatoria*. *Applied and Environmental Microbiology* **60**, 1068–1077.

Lelliott RA, Stead DE. 1987. Methods for the diagnosis of bacterial diseases of plants. Methods in Plant Pathology. Oxford, UK: British Society for Plant Pathology by Blackwell Scientific Publications.

Leyns F, Decleene M, Swings JG, Deley J. 1984. The host range of the genus *Xanthomonas*. *Botanical Review* **50**, 308-56.

Lindow S, Olson W, Buchner R. 2014. Colonization of dormant walnut buds by *Xanthomonas arboricola* pv. *juglandis* is predictive of subsequent disease. *Phytopathology* **104**, 1163-1174.

López MM, Lopez-Soriano P, Garita-Cambronero J, Beltrán C, Taghouti G, Pothier P, Cubero J, Fischer-Le Saux M, Marco-Noales E. 2018. *Xanthomonas prunicola* sp. nov., a novel pathogen that affects nectarine (*Prunus persica* var. nectarina) trees. *International Journal of Systematic and Evolution Microbiology* **68**, 1857-1866.

Loreti S, Gallelli A, Belisario A, Wajnberg E, Corazza L. 2001. Investigation of genomic variability of *Xanthomonas arboricola* pv. *juglandis* by AFLP analysis. *European Journal of Plant Pathology* **107**, 583-91.

Louws FJ, Rademaker JLW, Bruijn FJ. 1999. The three Ds of PCR-based genomic analysis of phytobacteria: diversity, detection, and disease diagnosis. *Annual Review of Phytopathology* **37**, 81–125.

Ludwig W, Schleifer KH. 1994. Bacterial phylogeny based on 16S and 23S ribosomal-RNA sequence-analysis. *FEMS Microbiology Reviews* **15**, 155-173.

Marcelletti S, Ferrante P, Scortichini M. 2010. Multilocus sequence typing reveals relevant genetic variation and different evolutionary dynamics among strains of *Xanthomonas arboricola* pv. *juglandis*. *Diversity* **2**, 1205-1222.

Martins JMS, Pinto C, Gomes JA. 1997. Chemical control of bacterial blight of walnut. *Acta Horticulturae* **442**, 367-372.

Martins L, Fernandes C, Albuquerque P, Tavares F. 2018. Assessment of *Xanthomonas arboricola* pv. *juglandis* bacterial load in infected walnut fruits by qPCR. Accepted.

McDonald BA, Stukenbrock EH. 2016. Rapid emergence of pathogens in agro-ecosystems: global threats to agricultural sustainability and food security. *Philosophical Transactions Royal Society B* **371**, 20160026.

McGuire RG, Jones JB, Sasser M. 1986. Tween media for semiselective isolation of *Xanthomonas campestris* pv. *vesicatoria* from soil and plant material. *Plant Disease* **70**, 887-891.

McMahon C, Denaxas S. 2016. A novel framework for assessing metadata quality in epidemiological and public health research settings. Proceedings *AMIA Joint Summits on Translational Science* **2016**, 199-208.

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

Melotto M, Underwood W, He SY. 2008. Role of stomata in plant innate immunity and foliar bacterial diseases. *Annual Review of Phytopathology* **46**, 101-122.

Menard M, Baudry A, Fisher Le Saux M. 2004. First report of bacterial canker of walnut caused by *Brenneria nigrifluens* in France. *Plant Disease* **88**, 220.

Milgroom MG, Fry WE. 1997. Contributions of populations genetics to plant disease epidemiology and management. *Advances in Botanical Research* **24**, 1-25.

Milgroom MG, Peever TL. 2003. Population biology of plan pathogens: The synthesis of plant disease epidemiology and population genetics. *Plant Disease* **97**, 608-617.

Milgroom MG. 2001. The synthesis of genetics and epidemiology: Contributions of population biology in plant pathology. *Journal of Plant pathology* **83**, 57-62.

Miller PW, Bollen WB, Simmons JE, Gross HN, Barss HP. 1940. The pathogen of filbert bacteriosis compared with *Phytomonas juglandis*, the cause of walnut blight. *Phytopathology* **30**, 713-733.

Miller PW. 1934. Walnut blight and its control in the Pacific Northwest USA. *Department of Agriculture* **331**, 13.

Miller PW. 1946. Walnut bacteriosis and its control. Corvallis: Oregon State System of Higher Education, Agricultural Experiment Station.

Miller W, Makova KD, Nekrutenko A, Hardison RC. 2004. Comparative Genomics. *Annual Review of Genomics and Human Genetics* **5**, 15-16.

Mikulic-Petkovsek M, Slatnar A, Veberic R, Stampar F, Solar A. 2011. Phenolic response in green walnut husk after the infection with bacteria *Xanthomonas arboricola* pv. *juglandis*. *Physiological and Molecular Plant Pathology* **76**, 159-65.

Moragrega C, Matias J, Aleta N, Montesinos E, Rovira M. 2011. Apical necrosis and premature drop of Persian (English) walnut fruit caused by *Xanthomonas arboricola* pv. *juglandis*. *Plant Disease* **95**, 1565-1570.

Moragrega C. 2010. Apical necrosis of walnut fruit. Fact Sheet created under the COST 873 frame by STF2.2 Meeting Advances in research on Brown Apical Necrosis of walnut fruit. Etiological conclusions. Girona 15-16 July 2010.

Moragrega C, Ozaktan H. 2010. Apical necrosis of Persian (English) walnut (*Juglans regia*): An update. *Journal of Plant Pathology* **92**, S67-S71.

Moragrega C. 2012. Detection and identification methods and new tests as developed and used in the framework of COST 873 for bacteria pathogenic to stone fruits and nuts: *Xanthomonas arboricola* pv. *juglandis*. *Journal of Plant Pathology* **94**, S1.55-S1.159.

Mulrean EN, Schroth MN. 1982. Ecology of *Xanthomonas campestris* pv. *juglandis* on Persin (English) walnuts. *Phytopathology* **72**, 434-8.

Myung I, Moon SY, Jeong IH, Lee SW, Ra D. 2010a. A new disease, bacterial fruit rot of jujube, caused by *Xanthomonas arboricola* in Korea. *New Disease Reports* **22**, 12.

Myung IS, Jeong IH, Moon SY, Lee SW, Shim HS. 2010b. A new disease, arboricola leaf spot of bell pepper, caused by *Xanthomonas arboricola*. *Plant Disease* **94**, 271.

Ninot A, Aletà N, Moragrega C, Montesinos E. 2002. Evaluation of a reduced copper spraying program to control bacterial blight of walnut. *Plant Disease* **86**, 583-7.

Oerke EC. 2006. Crop losses to pests. *Journal of Agricultural Science* **144**, 31-43.

Olive DM, Bean P. 1999. Principles and applications of methods for DNA-based typing of microbial organisms. *Journal of Clinial Microbiology* **37**, 1661-1669.

Oliveira M, Cabral R. 1942. Doenças bacterianas das plantas diagnosticadas em Portugal. *Revista Agronómica* **XXX**,176-184. [In Portuguese]

Osborn TGB, Samuel G. 1922. Some new records of fungi for South Australia. Part II. Together with a description of a new species of Puccinia. *Transactions of the Royal Society of South Australia* **XLVI**, 166-181.

Ozaktan H, Erdal M, Akkopru A, Aslan E. 2012. Biological control of bacterial blight of walnut by antagonistic bacteria. *Journal of Plant Pathology* **94**, 53-56.

Palacio-Bielsa A, Cambra MA. López MM. 2009. PCR detection and identification of plant-pathogenic bacteria: updated review of protocols (1989–2007). *Journal of Plant Pathology*, **91**, 249–297.

Palacio-Bielsa A, Roselló M, Cambra MA, López MM. 2010. First report on almond in Europe of bacterial spot disease of stone fruits caused by *Xanthomonas arboricola* pv. *pruni*. *Plant Disease* **94**, 786.

Palacio-Bielsa A, Cubero J, Cambra MA, Collados R, Berruete IM, López MM. 2011. Development of an efficient real-time quantitative PCR protocol for detection of *Xanthomonas arboricola* pv. *pruni* in *Prunus* species. *Applied and Environmental Microbiology* **77**, 89-97.

Park DS, Shim JK, Kim JS, Lim CK, Shrestha R, Hahn JH, KIM HG. 2009. Sensitive and specific detection of *Xanthomonas campestris* pv. *vesicatoria* by PCR using pathovar-specific primers based on *rhs* family gene sequences. *Microbiological Research* **164**, 36–42.

Park SY, Lee YS, Koh YJ, Hur JS, Jung, JS. 2010. Detection of *Xanthomonas arboricola* pv. *pruni* by PCR using primers based on DNA sequences related to the hrp genes. *Journal of Microbiology* **48**, 554–558.

Parkhill J, Wren BW. 2011. Bacterial epidemiology and biology - lessons from genome sequencing. *Genome Biolology* **12**, 230-230.

Parkinson N, Aritua V, Heeney J, Cowie C, Bew J, Stead D. 2007. Phylogenetic analysis of *Xanthomonas* species by comparison of partial gyrase B gene sequences. *International Journal of Systematic and Evolutionary Microbiology* **57**, 2881-2887.

Parkinson N, Cowie C, Heeney J, Stead D. 2009. Phylogenetic structure of *Xanthomonas* determined by comparison of *gyrB* sequences. *International Journal of Systematic and Evolutionary Microbiology* **59**, 264-274.

Pelludat C, Duffy B, Frey JE. 2009. Design and development of a DNA microarray for rapid identification of multiple European quarantine phytopathogenic bacteria. *European Journal of Plant Pathology* **125**, 413-423.

Pereira ALG, Paradella FO, Zagatto AG. 1971. Uma nova doença bacteriana da mandioquinha salsa (*Arracacia xanthorrhiza*) causada por *Xanthomonas arracaciae* n. sp. *Arquivos do Instituto Biológico* **38**, 99-108. [In Portuguese]

Pereira JA, Oliveira I, Sousa A*,* Valentão P, Andrade PB, Ferreira IC, Ferreres F, Bento A, Seabra R, Estevinho L. 2007. Walnut (*Juglans regia* L.) leaves: phenolic compounds, antibacterial activity and antioxidant potential of different cultivars. *Food and Chemical Toxicologyl* **45**, 2287-2295.

Pereira UP, Gouran H, Nascimento R, Adaskaveg JE, Goulart LR, Dandekar AM, 2015. Complete genome sequence of *Xanthomonas arboricola* pv. *juglandis* 417, a copper-resistant strain isolated from *Juglans regia* L. *Genome announcements* **3**, e01126-15.

Pierce NB. 1901. Walnut bacteriosis. *Botanical Gazette* **31**, 272-3.

Pothier JF, Pagani MC, Pelludat C, Ritchie DF, Duffy B. 2011. A duplex-PCR method for species and pathovar-level identification and detection of the quarantine plant pathogen *Xanthomonas arboricola* pv. *pruni*. *Journal of Microbiological Methods* **86**, 16–24.

Pruvost O, Gardan L. 1988. Dynamics of epiphytic survival of *Xanthomonas campestris* pathovar *corylina*, *Xanthomonas campestris* pathovar *juglandis*, *Erwinia herbicola* and *Pseudomonas paucimobilis* on hazelnut leaves. *Agronomie* **8**, 925-32.

Qi YX, Pu JJ, Zhang X, Lu Y, Zhang H, Lu YC, Zhang HQ, Xie YX. 2012. First report of bacterial leaf spot of mango caused by a *Xanthomonas arboricola* pathovar in China. *Journal of Plant Pathology*, **94**: S4.85.

Rademarker JLW, Norman DJ, Forster RL, Louws FJ, Schultz MH, Bruijn FJ. 2006. Classification and identification of *Xanthomonas translucens* isolates, including those pathogenic to ornamental asparagus. *Phytopathology* **96**, 876-884.

Radix P, Bastien C, Jay-Allemand C, Charlot G, Seigle-Murandi F. 1998. The influence of soil nature on polyphenols in walnut tissues. A possible explanation of differences in the expression of walnut blight. *Agronomie* **18**, 627-37.

Ranjbar R, Karami A, Farshad S, Giammanco GM, Mammina C. 2014. Typing methods used in the molecular epidemiology of microbial pathogens: a how-to guide. *New Microbiologica* **37**, 1-15.

Retamales J, Vasquez I, Santos L, Segovia C, Ayala M, Alvarado R, Nuñez P, Santander J. 2016. Complete genome sequences of lytic bacteriophages of *Xanthomonas arboricola* pv. *juglandis*. *Genome announcements* **4**, e00336-16.

Rodriguez RLM, Grajales A, Arrieta-Ortiz ML, Salazar C, Restrepo S, Bernal A. 2012. Genomes-based phylogeny of the genus *Xanthomonas*. *BMC Microbiology* **12**, 43.

Romero-Suarez S, Jordan B, Heinemann JÁ. 2012. Isolation and characterization of bacteriophages infecting *Xanthomonas arboricola* pv. *juglandis*, the causal agent of walnut blight disease. *World Journal of Microbiology & Biotechnology* **28**, 1917-1927.

Rothwell A. 1975. A revised list of plant diseases in Rhodesia – Additions, 1966-72. *Kirkia* **10**, 295-307.

Rouskas D, Akrivos J, Zakynthinos G. 2006. Evaluation of 33 walnut (*Juglans regia* L.) seedling selections from mountainous Greek regions. *Agricultural Research* **29**, 23-40.

Rudolph BA. 1933. Bacteriosis (blight) of the English walnut in California and its control. *California Agricultural Experimentation Station Bulletin* **564**, 3-83.

Ruppitsch W, Stoger A, Indra A, Grif K, Schabereiter-Gurtner C, Hirschl A, Allerberger F. 2007. Suitability of partial 16S ribosomal RNA gene sequence analysis for the identification of dangerous bacterial pathogens. *Journal of Applied Microbiology* **102**, 852-859.

Ryan RP, Vorhölter FJ, Potnis N, Jones JB, Van Sluys MA, Bogdanove AJ, Dow JM. 2011. Pathogenomics of *Xanthomonas*: understanding bacterium-plant interactions. *Nature reviews. Microbiology* **9**, 344–355.

Sawada H, Kunugi Y, Watauchi K, Kudo A, Sato T. 2011. Bacterial spot, a new disease of grapevine (*Vitis vinifera*) caused by *Xanthomonas arboricola*. *Japanese Journal of Phytopathology* **77**, 7-22.

Sawada H. Yokosawa S. Uematsu H, Nishiguchi T, Kondo K. 2017. First report of walnut blight caused by WB strains of *Xanthomonas arboricola* pv. *juglandis* in Japan. *Japanese Journal of Phytopathology* **83**, 10–21. [In Japanese]

Schaad NW, Frederick RD. 2002. Real-time PCR and its application for rapid plant disease diagnostics. *Canadian Journal of Plant Pathololog*y **24**, 250–258.

Schaad NW, Jones JB, Chum W. 2001. Laboratory guide for identification of plant pathogenic bacteria. 3rd Ed. St. Paul, MN, USA: APS Press.

Schaad NW. 1987. Problems with the Pathovar Concept. Plant Pathogenic Bacteria. Current Plant Science and Biotechnology in Agriculture, vol 4. Dordrecht, Netherlands: Springer.

Schena L, Nigro F, Ippolito A, Gallitelli D. 2004. Real-time quantitative PCR: a new technology to detect and study phytopathogenic and antagonistic fungi. *European Journal of Plant Pathology* **110**, 893-908.

Schneider DJ, Collmer A. 2010. Studying plant-pathogen interactions in the genomcis era: Beyond molecular Koch's postulates to systems biology. *Annual Review of Phytopathology* **48**, 457-479.

Schurch AC, Soolingen DV. 2012. DNA fingerprinting of *Mycobacterium tuberculosis*: From phage typing to whole-genome sequencing. *Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases* **12**, 602-609.

Scortichini M, Marchesi U, Di Prospero P. 2001. Genetic diversity of *Xanthomonas arboricola* pv. *juglandis* (synonyms: *X. campestris* pv. *juglandis*; *X. juglandis* pv. *juglandis*) strains from different geographical areas shown by repetitive polymerase chain reaction genomic fingerprinting. *Journal of Phytopathology-Phytopathologische Zeitschrift* **149**, 325-32.

Scortichini M, Rossi MP, Marchesi U. 2002. Genetic, phenotypic and pathogenic diversity of *Xanthomonas arboricola* pv. *corylina* strains question the representative nature of the type strain. *Plant Pathology* **51**, 374-381.

Scortichini M. 2010. Epidemiology and predisposing factors of some major bacterial diseases of stone and nut fruit trees species. *Journal of Plant Pathology* **92**, S73-S8.

Scotton M, Bortolin E, Fiorin A, Belisario A. 2015. Environmental and pathogenic factors inducing brown apical necrosis on fruit of english (persian) walnut. *Phytopathology* **105**, 1427-36.

Shami M, Ghasemi A, Ali-Abadi AA, Eskandari A. 2013. Genetic diversity and phylogenetic study of *Xanthomonas arboricola* pv. *juglandis* the causal agent of walnut bacterial blight disease. *Journal of Nuts* **4**, 57-62.

Sharma R, Sharma YP. 1999. Diseases of temperate nut crops. *Diseases of Horticultural Crops: Fruits* 208.

Sharma RM, Pandey MK, Shankar U. 2012. *Pest management in walnut: An overview*. New Delhi, India: New India Publishing Agency.

Simpson AJG, Reinach FC, Arruda P, Abreu FA, Acencio M, Alvarenga R, Alves LM, Araya JE, Baia GS, et al. 2000. The genome sequence of the plant pathogen *Xylella fastidiosa*. *Nature* **406**, 151–157.

Smith EF. 1903. Observations on a hitherto unreported bacterial disease, the cause of which enters the plant through ordinary stomata. *Science* **17**, 456-457.

Smith RE, Smith CO, Ramsey HJ. 1912. Walnut culture in California: walnut blight. Berkeley, California: California Agricultural Experiments Station Publications.

Soares, AP. 2005. Frutos frescos e secos. Anuário Vegetal 2005 – Crop production yearbook. GPPAA – Gabinete de Planeamento e Polítca Agro-Alimentar. Ministério da Agricultura do Desenvolvimento Rural e das Pescas.

Solar A, Jakopic J, Veberic R, Stampar F. 2012. Correlations between *Xanthomonas arboricola* pv. *juglandis* severity and endogenous juglone and phenolic acids in walnut. *Journal of Plant Pathology* **94**, 229-235.

Soltani J, Aliabadi AA. 2010. Genetic resistance to bacterial blight disease in Persian walnut. *European Journal of Plant Pathology* **128**, 65-70.

Stanford AM, Harden R, Parks CR. 2000. Phylogeny and biogeography of *Juglans* (Juglandaceae) based on matK and ITS sequence data. *American Journal of Botany* **87**, 872-882.

Starr MP. 1946. The nutrition of phytopathogenic bacteria I. Minimal nutritive requirements of the genus *Xanthomonas*. *Journal of Bacteriology* **51**, 131-143.

Strange RN, Scott PR. 2005. Plant disease: A threat to global fodd security. *Annual Review of Phytopathology* **43**, 83-116.

Swings JG, Civerolo EL. 1993. *Xanthomonas*. London: Chapman & Hall.

Temperini CV, Pardo AG, Pose GN. 2017. First report of apical necrosis in walnut cultivars from northern argentinean Patagonia. *Journal of Plant Pathology and Microbiology* **8**, 1-6.

Triplett LR, Verdier V, Campillo T, Van Malderghem C, Cleenwerck I, Maes M, Deblais L, Corral R, Koita O, Cottyn B, Leach JE. 2015. Characterization of a novel clade of *Xanthomonas* isolated from rice leaves in Mali and proposal of *Xanthomonas maliensis* sp. nov. *Antonie van Leeuwenhoek* **107**, 869-881.

Vagelas LK, Rumbos CI, Tsiantos J. 2012. Variation in disease development among persian walnut cultivars, selections and crosses when inoculated with *Xanthomonas arboricola* pv. *juglandis* in Greece. *Journal of Plant Pathology* **94**, 57-61.

Vandroemme J, Cottyn B, Pothier JF, Pfluger V, Duffy B, Maes M. 2013. *Xanthomonas arboricola* pv. *fragariae*: what's in a name? *Plant Pathology* **62**, 1123–1131.

Vauterin L, Hoste B, Kersters K, Swings J. 1995. Reclassification of *Xanthomonas*. *International Journal of Systematic Bacteriology* **45**, 472-89.

Verdier V, Ojeda S, Mosquera G. 2001. Methods for detecting the cassava bacterial blight pathogen: A practical approach for managing the disease. *Euphytica* **120**, 103-107.

Vicente JG, Rothwell S, Holub EB, Studholme DJ. 2017. Pathogenic, phenotypic and molecular characterisation of *Xanthomonas nasturtii* sp. nov. and *Xanthomonas floridensis* sp. nov., new species of *Xanthomonas* associated with watercress production in Florida. *International Journal of Systematic and Evolutionary Microbiology* **67**, 3645-3654.

Vieira J, Mendes MV, Albuquerque A, Moradas-Ferreira P, Tavares F. 2007. A novel approach for the identification of bacterial taxa-specific molecular markers. *Letters in Applied Microbiology* **44**, 506-512.

Vinatzer BA, Weisberg AJ, Monteil CL, Elmarakeby HA, Sheppard SK, Health LS. 2017. A proposal for a genome similarity-based taxonomy for plant-pathogenic bacteria that is sufficiently precise to reflect phylogeny, host range, and outbreak affiliation applied to *Pseudomonas syringae sensu lato* as a proof of concept. *Phytopathology* **107**, 18-28.

Vincelli P, Tisserat N. 2008. Nucleic acid-based pathogen detection in applied plant pathology. *Plant Disease* **92**, 660–669.

Vorholt JA. 2012. Microbial life in the phyllosphere. *Nature Reviews. Microbiology* **10**, 828-840.

Wang M, Roux F, Bartoli C, Huard-Chauveau C, Meyer C, Lee H, Roby D, McPeek MS, Bergelson J. 2018. Two-way mixed-effects methods for joint association analysis using both host and pathogen genomes. *Proceedings of the National Academy of Sciences of the United States of Americ*a **115**, E5440-E5449.

Ward E, Foster SJ, Fraaije BA, Mccartney HA. 2004. Plant pathogen diagnostics: immunological and nucleic acid-based approaches. *Annals of Applied Biology* **145**, 1–16.

Woemald H. 1930. Bacterial “blight” of walnuts in Britain. *Annals of Applied Biology* **17**, 59-70.

Woeste KE, Mcgranahan GH, Schroth MN. 1992. Variation among persian walnuts in response to inoculation with *Xanthomonas campestris* pv. *juglandis*. *Journal of the American Society for Horticultural Science* **117**, 527-31.

Wolfe MS, Knott DR. 1982. Populations of plant pathogens: some constraints on analysis of variation in pathogenicity. *Plant Pathology* **31**, 79-90.

Yamamoto S, Kasai H, Arnold DL, Jackson RW, Vivian A, Harayama S. 2000. Phylogeny of the genus *Pseudomonas*: intrageneric structure reconstructed from the nucleotide sequences of *gyrB* and *rpoD* genes. *Microbiology* **146**, 2385–2394.

Yirgou D. 1964. *Xanthomonas guizotiae* sp. nov. on *Guizotia abyssimica*. *Phytopathology* **54**, 1490-1491.

Young JM, Dye DW, Bradbury JF, Panagopoulos CG, Robbs CF. 1978. A proposed nomenclature and classification for plant pathogenic bacteria. *New Zealand Journal of Agricultural Research* **21**, 153-77.

Young JM, Park DC, Shearman HM, Fergier E. 2008. A multilocus sequence analysis of the genus *Xanthomonas*. *Systematic and Applied Microbiology* **31**, 366-377.

Young JM, Wilkie JP, Park DC, Watson DRW. 2010. New Zealand strains of plant pathogenic bacteria classified by multi-locus sequence analysis; proposa of Xanthomonas dyei sp. nov.. *Plant Pathology* 59, 270-281.

Zaccardelli M, Ceroni P, Mazzucchi U. 1999. Amplified fragment length polymorphism fingerprinting of *Xanthomonas arboricola* pv. *pruni*. *Journal of Plant Pathology* **81**, 173–179.

Zhang SL, Shen JG, Shen GH, Sun ZQ, Xu PH, Peng YL, Yang ZR, Sun Q. 2007. Use of a novel multiplex probe array for rapid identification of *Mycobacterium* species from clinical isolates. *World Journal of Microbiology & Biotechnology* **23**, 1779-1788.

Zhang Y, Jalan N, Zhou X, Goss E, Jones JB, Setubal JC, Deng X, Wang N. 2015. Positive selection is the main driving force for evolution of citrus canker-causing *Xanthomonas*. *The ISME Journal* **9**, 2128–2138.

CHAPTER II

Multiple DNA markers for identification of *Xanthomonas arboricola* pv. *juglandis* isolates and its direct detection in plant samples

Chapter II include the following publication:

Fernandes C, Alburquerque P, Sousa R, Cruz L, Tavares F. 2017. Multiple DNA markers for identification of *Xanthomonas arboricola* pv. *juglandis* isolates and its direct detection in plant samples. *Plant Disease*, 101(6):858-865.

Part of the results from this chapter were presented at:

Fernandes C, Albuquerque P, Sousa R, Cruz L, Tavares F. 2017 Detection, identification and diversity of *Xanthomonas arboricola* pv. *juglandis* by multiplex PCR and dot blot hybridization using nine DNA markers. 7th Congress of European Microbiologists (FEMS 2017), Valencia, Spain, July 9-13, 2017.

Fernandes C, Albuquerque P, Sousa R, Cruz L, Tavares F. 2017. Nove marcadores moleculares para a deteção, identificação e tipagem de *Xanthomonas arboricola* pv. *juglandis*. Proteção das Plantas 2017: 2º Simpósio SCAP, 8º Congresso da SPF e 11º Encontro Nacional de Proteção Integrada, Escola Superior Agrária de Santarém, Santarém, Portugal, October 26-27, 2017.

Multiple DNA markers for identification of *Xanthomonas arboricola* pv. *juglandis* isolates and its direct detection in plant samples

Abstract

Xanthomonas arboricola pv. *juglandis* (*Xaj*) is the etiological agent of the walnut (*Juglans regia* L.) bacterial blight (WBB), and has been associated to other walnut emerging diseases namely the brown apical necrosis (BAN), and the vertical oozing canker (VOC), altogether severely affecting the walnut production worldwide. Regardless the research efforts carried out to disclose *Xaj* genetic diversity, reliable molecular methods for rapid identification of *Xaj* isolates and culture-independent detection of *Xaj* in infected plant samples are still missing. In this work we propose nine novel specific DNA markers (XAJ1 to XAJ9) selected by dedicated *in silico* approaches to identify *Xaj* isolates and detect these bacteria in infected plant material. To confirm the efficacy and specificity of these markers, dot blot hybridization was carried out across a large set of xanthomonads. This analysis, which confirmed the pathovar specificity of these markers, allowed to identify four broad-range markers (XAJ1, XAJ4, XAJ6 and XAJ8) and five narrow-range markers (XAJ2, XAJ3, XAJ5, XAJ7, and XAJ9), originating 12 hybridization patterns (HP1 to HP12). No evident relatedness was observed between these hybridization patterns and the geographic origin from which the isolates were obtained. Interestingly, four isolates which clustered together according the *gyrB* phylogenetic analysis (CPBF 1507, 1508, 1514 and 1522) presented the same hybridization pattern (HP11), suggesting that these nine markers might be informative to rapidly discriminate and identify different *Xaj* lineages. Taking into account that a culture-independent detection of *Xaj* in plant material has never been described, a multiplex PCR was optimized using markers XAJ1, XAJ6 and XAJ8. This triplex PCR, besides confirming the dot blot data for each of the 52 *Xaj*, was able to detect *Xaj* in field infected walnut leaves and fruits. Altogether, these nine *Xaj*-specific markers allow conciliating the specificity of DNA-detection assays with typing resolution, contributing to rapidly detect and identify potential emergent and acutely virulent *Xaj* genotypes, infer their distribution, disclose the presence of this phytopathogen on potential alternative host species and improve phytosanitary control.

1. Introduction

Xanthomonas arboricola (Vauterin et al. 1995) is a bacterial species responsible for economically relevant diseases in several important fruit trees (Frutos 2010, Lamichhane 2014). According to host specificity, several pathovars of *Xanthomonas arboricola* are presently acknowledged (Vauterin et al. 1995, Janse et al. 2001, Saux et al. 2015). Among these, pathovar *juglandis* has been considered as one of the most serious threats to walnut production worldwide (Belisario et al. 2002, Hajri et al. 2010, Moragrega and Ozaktan 2010, Lamichhane 2014). *Xanthomonas arboricola* pv. *juglandis* (*Xaj*) was first described in the beginning of the 20th century (Pierce 1901), and it was since identified as the etiological agent of the walnut bacterial blight (WBB), the most well-known disease of walnut (*Juglans* sp.), often recognized by the presence of necrotic spots on leaves and fruits (Smith et al. 1912). *Xaj* has been also frequently associated to the brown apical necrosis (BAN) characterized by apical necrotic lesions near the blossom end of the nut, which has also been associated with other phythopathogens (Belisario et al. 2002, Moragrega and Ozaktan 2010, Moragrega et al. 2011), and the vertical oozing canker (VOC), consisting of unusual symptoms, including vertical cankers, deformations and brown exudates, that appear mainly in the trunk of walnut trees (Hajri et al. 2010). While different parts of the plant can be infected leading to a decrease of tree vigour, the main economic impact of diseases attributed to *Xaj* is the substantial yield losses, due to a premature fruit drop and a decrease of the quality of marketable nuts (Scortichini 2010, Lamichhane 2014).

Occurrence of *Xaj* infection has been observed worldwide (Frutos 2010) with a recent increase in the number of outbreaks reported in different walnut production regions (Hajri et al. 2010, Moragrega et al. 2011, Lamichhane 2014). The impact of more frequent and severe outbreaks affecting the main producing countries is still unknown. Thus, further understanding of pathogen biology and epidemiology is crucial for the effective control of the disease (Scortichini 2010, Lamichhane 2014). Studies have shown that *Xaj* can multiply epiphytically, penetrate and invade walnut trees during favourable conditions mainly through stomata structures, leaf-scars, or plant wounds (Scortichini 2010, Lamichhane 2014). It has also been described that dormant buds are privileged overwintering sites for *Xaj* (Mulrean and Schroth 1982, Lindow et al. 2014). Furthermore, the existence of alternative host reservoirs of *Xaj*, such as asymptomatic plant species, might also be relevant for the dissemination and transmission of *Xaj* to walnut plants (Allen et al. 2009, Lamichhane 2014).

Currently, diagnostic procedures for *Xaj* rely on traditional methods based on bacteria isolation using differential or semi-selective media, species-specific biochemical assays and pathogenicity tests (Janse 2010, Moragrega 2012). Despite the numerous DNA-based typing methods for genetic characterization of *Xaj* populations that have been proposed in recent years

(Hajri et al. 2010, Marcelletti et al. 2010, Burokiene and Pulawska 2012, Kaluzna et al. 2014, Essakhi et al. 2015, Giovanardi et al. 2015, Ivanovic et al. 2015), reliable molecular methods for rapid identification of *Xaj* isolates and culture-independent detection of *Xaj* in infected plant samples are still missing. Nonetheless, there are a few studies that report identification *of Xaj* isolates by PCR amplification of a single marker (Gironde et al. 2009, Burokiene and Pulawska 2012, Moragrega 2012, Kaluzna et al. 2014, Ivanovic et al. 2015), but neither the sequence of the marker nor the primers were made available, making difficult a broad assessment of this method.

To overcome some of these limitations, this study focused on the selection and validation of unique and discriminative *Xaj* genomic regions. These DNA markers, besides providing the tools for culture-independent detection of *Xaj* in infected plant material, and for the reliable identification of isolates, will certainly be useful to disclose the presence of this phytopathogen on potential alternative host species, contributing to detail the life cycle of these bacteria and helping to design the most suitable phytosanitary strategies for prevention and containment of diseases caused by *Xaj*.

2. Material and methods

2.1. Selection of specific DNA markers for *Xanthomonas arboricola* pv. *juglandis*

The identification of *Xaj*-specific DNA regions was carried out through the analysis of all possible ORFs (open reading frames) retrieved from the concatenated contigs of *Xaj* strain NCPPB 1447 draft genome (Accession PRJNA84273) using Geneious v. 7.1.2 software (Biomatters, Auckland, New Zealand). A BLAST analysis (Altschul et al. 1990) was performed in order to infer the *Xaj*-specificity of each ORF. Sequences with the best BLAST specificity results were chosen as the most promising regions for marker design (**Table II.1**).

For DNA marker selection, the *Xaj* specific ORFs were further screened for their putative function and for their suitability to design primer pairs using the Vector NTI 10 software (Invitrogen, Carlsbad, CA) favouring high PCR stringency, in order to prevent unspecific amplification and facilitate multiplexing. These premises resulted in nine putative *Xaj*-specific markers and corresponding primers pairs (XAJ1 to XAJ9), which were chosen for further experimental validation using bacterial reference strains and isolates (**Table II.1**).

Table II.1. *Xanthomonas arboricola* pv. *juglandis* specific markers, corresponding primer-pairs sequences and expected amplicon sizes. For each marker is indicated the best BLAST hits (blastn and wgs).

Markers	Primers	Sequences (5'-3')	Length (bp)	Best BLAST hit (E value/Q coverage)[a]	
				blastn	wgs blast
XAJ1	XAJ1F	GTTGTCAAGGTCGCTACTGGACGC	758	*Xanthomonas campestris* 17	*Xanthomonas arboricola* pv. *celebensis* NCPPB 1832
	XAJ1R	CGACAGGTACCGAAAGTGGACG		(2e-08/12%)	(0/100%)
XAJ2	XAJ2F	AGATCGGGCTGGATGAAGAAGAG	753	*Desulfomicrobium baculatum* DSM 4028	*Xanthomonas perforans* TB15
	XAJ2R	CCAGGATCGCAAGTTCTTCAAGTG		(4e-12/17%)	(0/100%)
XAJ3	XAJ3F	CTCGTCTTTACCGTTCCGTCAAC	785	*Pseudomonas suwonensis* 11-1	*Xanthomonas* sp. Leaf148
	XAJ3R	GCCCAAGGACATTCACACACTC		(6e-16/10%)	(1e-12/21%)
XAJ4	XAJ4F	TCGATCGAATCAGTAGTCACTGGC	748	*Bifidobacterium animalis* subsp. *animalis* ATCC 25527	*Xanthomonas hyacinthi* DSM 19077
	XAJ4R	CGGATCTCTTCATTCGTCAGGTG		(0.26/5%)	(1e-12/21%)
XAJ5	XAJ5F	AGCCTTGGGGTAGCTTTTGAGG	491	*Corynebacterium mustelae* DSM 45274	*Xanthomonas axonopodis* pv. *glycines* CFBP 7119
	XAJ5R	CGTTCTTTGACGGCACTCCC		(0.17/6%)	(0/100%)
XAJ6	XAJ6F	AAGTCAGATGCGAAGCGAAAGG	436	*Ascaris lumbricóides* v1 5 4	*Xanthomonas axonopodis* pv. *glycines* CFBP 7119
	XAJ6R	GCACAGCGGGAAGTAATAGCAAAC		(0.51/6%)	(0/100%)
XAJ7	XAJ7F	ATAACTACCACCAACTTCCTGGGG	766	*Alligator sinensis*	*Xanthomonas arboricola* pv. *celebensis* NCPPB 1832
	XAJ7R	TCTGGGCGAGGTAAGTGATTCTAC		(0.93/4%)	(1e-133/35%)
XAJ8	XAJ8F	GGGGTGAAATCCAGTACCTGTCAC	588	*Monodelphis domestica*	*Xanthomonas campestris* pv. *campestris* Xca5
	XAJ8R	CGCAGAACGTACTTTCGGTACTGG		(2.5/4%)	(8e-122/98%)
XAJ9	XAJ9F	CACGGGGATCAGCTTTTCATAC	443	*Ralstonia solanacearum*	*Xanthomonas hortorum* pv. *carotae* M081
	XAJ9R	GCGTCGGTAACGATGTTTTGTC		(1e-08/16%)	(2e-166/83%)

[a] Q coverage, Query coverage

2.2. Bacterial strains, culture conditions and DNA extraction

The bacterial strains used for validation of the nine putative *Xaj*-specific markers are listed in **Table II.2**, and include 18 strains of *Xaj,* five strains representing the pathovars *celebensis*, *corylina*, *fragariae, populi* and *pruni* of *X. arboricola*, as well as 12 strains belonging to other species of *Xanthomonas*. Bacterial strains were cultured in Yeast Extract Glucose Carbonate (YGC) medium (yeast extract, 5.0 g; glucose, 10.0 g; $CaCO_3$, 30.0 g; agar, 15.0 g; distilled water up to 1.0 L) at 28°C, except for *Xanthomonas fragariae* that was cultured in Yeast-Peptone-Glucose Agar (YPGA) medium (yeast extract, 5.0 g; bacto peptone, 5.0 g; glucose, 10.0 g; agar, 15.0 g; distilled water up to 1.0 L) at 20°C.

DNA was extracted from pure cultures using the EZNA Bacterial DNA Purification kit (Omega Bio-Tek, Norcross, GA, USA), according to the manufacturer's instructions, and quantified using the Qubit® 2.0 Fluorometer HS Assay (Invitrogen, Carlsbad, CA, USA).

Table II.2. List of reference bacterial strains used for validation of *Xanthomonas arboricola* pv. *juglandis* specific markers.

***Xanthomonas* species and pathovars**[a]	**Geographic origin**	**Year of isolation**[b]
X. arboricola pv. *juglandis* CFBP 176	France	1961
X. arboricola pv. *juglandis* CFBP 877	France	1966
X. arboricola pv. *juglandis* CFBP 2564	Italy	1985
X. arboricola pv. *juglandis* CFBP 2632	Spain	1984
X. arboricola pv. *juglandis* CFBP 5252	France	2000
X. arboricola pv. *juglandis* CFBP 6556	Italy	1993
X. arboricola pv. *juglandis* CFBP 6557	Italy	1999
X. arboricola pv. *juglandis* CFBP 7072	Spain	1993
X. arboricola pv. *juglandis* CFBP 7179^{GSS}	France	2002
X. arboricola pv. *juglandis* CFBP 7244	France	1978
X. arboricola pv. *juglandis* LMG 745	Netherlands	1978
X. arboricola pv. *juglandis* LMG 746	United Kingdom	1955
X. arboricola pv. *juglandis* LMG 747$^{TS; GSS}$	New Zealand	1956
X. arboricola pv. *juglandis* LMG 748	New Zealand	1956
X. arboricola pv. *juglandis* LMG 749	New Zealand	1957
X. arboricola pv. *juglandis* LMG 751$^{GSS; *}$	Romania	1962
X. arboricola pv. *juglandis* LMG 752	United Kingdom	1964
X. arboricola pv. *juglandis* LMG 8047	Netherlands	1979
X. arboricola pv. *celebensis* LMG 677^{PRS}	New Zealand	1960
X. arboricola pv. *corylina* LMG 689^{PRS}	United States	1939
X. arboricola pv. *fragariae* LMG 19145 ^{PRS}	Italy	n/a
X. arboricola pv. *populi* CFBP 3123^{PRS}	Netherlands	1979
X. arboricola pv. *pruni* LMG 852^{PRS}	New Zealand	1953
X. axonopodis pv. *citri* LMG 9322^{TS}	United States	1989
X. axonopodis pv. *dieffenbachiae* LMG 695^{PRS}	Brazil	1965
X. axonopodis pv. *phaseoli* LMG 7455^{PRS}	United States	1986
X. campestris pv. *campestris* LMG 568	United Kingdom	1957
X. euvesicatoria LMG 922	United States	1939
X. fragariae LMG 708^{TS}	United States	1960
X. gardneri LMG 962$^{GSS;TS}$	Yugoslavia	1953
X. oryzae pv. *oryzae* LMG 5047^{PRS}	India	1965
X. oryzae pv. *oryzicola* LMG 797^{PRS}	Malaysia	1964
X. perforans NCPPB 4321^{TS}	United States	n/a
X. translucens pv. *translucens* LMG 876$^{GSS; PRS}$	United States	1933
X. vesicatoria LMG 911$^{GSS; TS}$	New Zealand	1955

[a] CFBP, French Collection for Plant-associated Bacteria, Institut National de la Recherche Agronomique, Angers, France; LMG, – Belgian Coordinated Collections of Microorganisms/ LMG Bacteria Collection, Universiteit Gent - Laboratorium voor Microbiologie, Gent, Belgium; NCPPB, National Collection of Plant Pathogenic Bacteria, Fera Science Ltd., York, United Kingdom; *, other collection no: NCPPB 1447; TS, type strain; GSS, genome sequenced strain; PRS, pathovar reference strain. [b] n/a, not available.

2.3. Isolation of bacteria from walnut leaves and fruits with WBB and BAN symptoms

Leaf and fruit samples of walnut trees (*Juglans regia*) showing WBB and/or BAN symptoms were collected during the spring-summer seasons, from six geographic regions of Portugal (Alcobaça, Azeitão, Beja, Estremoz, Loures, and Seia) (**Table II.3**).

Putative *Xaj* were isolated from infected leaves and fruits by using a sterile scalpel to excise plant tissue adjacent to necrotic areas. Sampled plant fragments were disinfected by immersion in 70% ethanol during 30s, thoroughly washed with sterile distilled water (SDW) and then macerated with 5 mL of SDW in extraction bags (Bioreba AG, Reinach, Switzerland). The resulting suspensions and the correspondent 10^{-1} dilutions, were streaked on medium Yeast Extract Dextrose Carbonate (YDC) (yeast extract, 10.0 g; dextrose, 20.0 g; $CaCO_3$, 20.0 g; agar, 15.0 g; distilled water up to 1.0 L) and incubated at 26±2°C for 4 days (Lelliott and Stead 1987).

Characteristic mucoid yellow colonies were streaked on fresh Nutrient Agar medium to ensure purity. All isolates were stored at -80°C at the Portuguese Collection of Phytopathogenic Bacteria (CPBF – Colecção Portuguesa de Bactérias Fitopatogénicas). Bacterial DNA was extracted using EZNA Bacterial DNA Purification kit.

Symptomatic walnut leaves and fruits were additionally used for culture-independent detection of *Xaj* by multiplex PCR. Total DNA was extracted from pelleted cells using 2 mL of the plant suspensions obtained as described above and centrifuged at 16000 g for 10 min. DNA extraction was performed using both the QIAGEN DNeasy Plant Mini kit, following manufacturer's instructions (Qiagen, Hilden, Germany) and the CTAB-based DNA extraction method (Murray and Thompson 1980, Doyle 1991). To confirm *Xaj* as the causal agent of infection in these plant samples, bacterial isolates were obtained following the procedure described above.

Table II.3. List of bacterial isolates collected in Portugal used in this study and the correspondent GenBank accession number (AC) of partial gyrB sequences (835bp).

Isolates[a]	Host plant[b]	Isolation sources[c]	Location, year	Accession number
CPBF 1271	*Juglans regia*	n/a	Azeitão, 2009	KU058325
CPBF 1479	*Juglans regia*	Leaves	Azeitão, 2014	KU058326
CPBF 1480	*Juglans regia*	Fruits	Azeitão, 2014	KU058327
CPBF 1484	*Juglans regia* cv. Hartley	Leaves	Alcobaça, 2014	KU058329
CPBF 1485	*Juglans regia* cv. Hartley	Fruits	Alcobaça, 2014	KU058330
CPBF 1486	*Juglans regia* cv. Lara	Leaves	Alcobaça, 2014	KU058331
CPBF 1487	*Juglans regia* cv. Franquette	Leaves	Alcobaça, 2014	KU058332
CPBF 1489	*Juglans regia* cv. Rego	Leaves	Alcobaça, 2014	KU058333
CPBF 1490	*Juglans regia* cv. Rego	Fruits	Alcobaça, 2014	KU058334
CPBF 1491	*Juglans regia* cv. Corne	Fruits	Alcobaça, 2014	KU058335
CPBF 1492	*Juglans regia* cv. Amigo	Leaves	Alcobaça, 2014	KU058336
CPBF 1496	*Juglans regia* cv. Franquette	Fruits	Alcobaça, 2014	KU058337
CPBF 1497	*Juglans regia* cv. Corne	Leaves	Alcobaça, 2014	KU058338
CPBF 1502	*Juglans regia* cv. Hartley	Leaves	Alcobaça, 2014	KU058339
CPBF 1503	*Juglans regia*	Leaves	Seia, 2014	KU058340
CPBF 1504	*Juglans regia*	Leaves	Seia, 2014	KU058341
CPBF 1505	*Juglans regia*	Fruits	Seia, 2014	KU058342
CPBF 1506	*Juglans regia*	Fruits	Seia, 2014	KU058343
CPBF 1507	*Juglans regia*	Leaves	Beja, 2014	KU058344
CPBF 1508	*Juglans regia*	Leaves	Beja, 2014	KU058345
CPBF 1509	*Juglans regia*	Fruits	Azeitão, 2014	KU058346
CPBF 1510	*Juglans regia*	Fruits	Azeitão, 2014	KU058347
CPBF 1511	*Juglans regia*	Fruits	Azeitão, 2014	KU058348
CPBF 1512	*Juglans regia*	Leaves	Seia, 2014	KU058349
CPBF 1513	*Juglans regia* cv. Hartley	Leaves	Estremoz, 2014	KU058350
CPBF 1514	*Juglans regia* cv. Hartley	Leaves	Estremoz, 2014	KU058351
CPBF 1518	*Juglans regia* cv. Lara	Leaves	Estremoz, 2014	KU058354
CPBF 1519	*Juglans regia* cv. Tulana	Leaves	Estremoz, 2014	KU058355
CPBF 1520	*Juglans regia* cv. Tulana	Leaves	Estremoz, 2014	KU058356
CPBF 1521	*Juglans regia*	Leaves	Loures, 2014	KU058357
CPBF 1522	*Juglans regia* cv. Howard	Leaves	Estremoz, 2014	KU058358
CPBF 1525	*Juglans regia* cv. Lara	Leaves	Estremoz, 2014	KU058359
CPBF 1526	*Juglans regia* cv. Lara	Leaves	Estremoz, 2014	KU058360
CPBF 1527	*Juglans regia* cv. Lara	Leaves	Estremoz, 2014	KU058361

[a] CPBF, Colecção Portuguesa de Bactérias Fitopatogénicas, Instituto Nacional de Investigação Agrária e Veterinária, Oeiras, Portugal; [b] cv., cultivar; [c] n/a, not available.

2.4. Molecular analysis of isolates by *gyrB*

The identification of the bacterial isolates was carried out by amplification and sequencing of the *gyrB* gene as previously described by Young et al. (2008), using the primers XgyrB1F and XgyrB1R and a reaction mixture containing 1X DreamTaq Buffer with 2.0 mM $MgCl_2$ (Fermentas), 0.2 mM of each dNTP (Fermentas), 0.2 µM of each primer, 1U of Dream*Taq* DNA Polymerase (Thermo Scientific) and 25 ng of DNA template. *gyrB* amplicons were sequenced in both strands and compared with *gyrB* sequences available at GenBank database using the BLAST program. To determine the phylogenetic relatedness of these isolates with other *Xanthomonas* spp., a Maximum Likelihood tree was built using the Tamura-Nei model in MEGA 6.0 (Tamura et al. 2013). The analysis was generated from the MUSCLE alignment of 145 *gyrB* sequences (34 sequences of *Xaj* Portuguese isolates and 111 sequences representing different species of *Xanthomonas* sp.).

2.5. Experimental validation of putative *Xaj*-specific markers by dot blot

Validation of the nine putative *Xaj*-specific markers across a large set of *Xanthomonas* spp. strains and isolates was carried out using a dot blot hybridization platform. DNA probes were prepared from the purified amplicons obtained from PCR amplifications using *Xaj* strain LMG 751 (= NCPPB 1447), from which the fully sequenced genome was used in this study to select the DNA markers. A 20 µL PCR reaction mix consisted of 1X DreamTaq Buffer (Fermentas, Ontario, Canada), 0.2 mM of each deoxynucleotide triphosphate (dNTP) (Fermentas), 0.2 µM of each forward and reverse primers (**Table II.1**), 1U of DreamTaq DNA Polymerase (Fermentas) and 25 ng of DNA template. PCR cycling parameters were carried out with the implementation of a first amplification cycle of 5 min at 95°C, followed by 35 cycles of 95°C for 30 s, 61°C for 30 s, and 72°C for 30 s, and a final DNA extension at 72°C for 10 min. The identity of each amplicon, used as probe, was confirmed by sequencing. The obtained PCR products were purified using the illustra GFX GEL Band Purification kit (GE Healthcare, Buckingham-shire, United Kingdom), following the reference protocol available, and sequenced (STAB Vida, Caparica, Portugal).

Probes were labelled with digoxigenin using the DIG-High Prime kit, according to the instructions of the manufacturer (Roche Diagnostics GmbH, Basel, Switzerland). For dot blot assays, 100 ng of heat-denaturated bacterial DNA were spotted into nylon membranes (Roche Diagnostics GmbH, Basel, Switzerland) using a Bio-Dot apparatus (Bio-Rad, Hercules, CA). Each membrane was hybridized overnight at 68°C to ensure high stringency with a final probe concentration of 100 ng/mL, and stringency washes were performed as described in the DIG application manual. Probe-target hybrids were detected with Chemiluminescent alkaline

phosphatase substrate (CDP-Star) reagent (Roche Diagnostics, Basel, Switzerland) and the images were acquired using a Molecular Imager ChemiDoc system (Bio-Rad, Hercules, CA).

2.6. Multiplex PCR validation

A multiplex PCR targeting the most promising *Xaj*-specific markers was optimized in order to validate a method to rapidly identify *Xaj* isolates and for the direct diagnostics of *Xaj* infection in walnut tree samples. XAJ1, XAJ6 and XAJ8, were the chosen markers since they were shown to hybridize to most of the *Xaj* tested and have distinct amplicon lengths of 758 bp, 588 bp and 436 bp, respectively. Multiplex PCR assays were performed on DNA extracted from bacterial strains/isolates and from plant tissue extracts. For chromosomal bacterial DNA the PCR conditions were essentially the same as those used to obtain the DNA dot blot probes with an increase of DreamTaq DNA polymerase to 1.5U per PCR reaction. Concerning DNA extracted from infected plant material PCR reactions were prepared with 2 µL of DNA extracted from infected walnut leaves and fruits using both the QIAGEN DNeasy Plant Mini kit and the CTAB-based DNA extraction as described above, and 1.5U of GoTAQ® G2 Flexi DNA Polymerase (Promega, Madison, USA). For these samples, PCR amplification conditions were of 10 initial PCR cycles at an annealing temperature of 55°C, followed by 30 cycles at an annealing temperature of 61°C. To further confirm multiplex PCR results from infected plant material, the same DNA samples were used as template in PCR reactions using each of the markers individually (XAJ1, XAJ6 and XAJ8). PCR amplification of these three markers was performed as described to prepare the DNA dot blot probes.

Finally, to ensure that multiplex PCR amplification from symptomatic plant material was due to infection by *Xaj*, a confirmatory multiplex PCR was carried out using the *Xaj* isolates obtained from the same infected plant samples. Partial sequencing of *gyrB* confirmed the *Xaj* identity of these isolates. All the 436 bp amplicons corresponding to the marker XAJ6 and obtained from both the plant samples and *Xaj* isolates, were purified using the illustra GFX GEL Band Purification kit and sequenced to confirm their identity.

2.7. Detection limit of multiplex PCR

Sensitivity of multiplex PCR was determined with chromosomal DNA, bacterial suspensions and spiked plant extracts using *Xaj* LMG 747. Chromosomal DNA multiplex PCR conditions were the same as described above for multiplex PCR validation with bacterial DNA template, using 2 µL of each ten-fold DNA dilutions ranging from 100 ng to 10 fg. Multiplex PCR detection limit of cell suspensions was obtained from ten-fold dilutions, prepared in SDW, of *Xaj* LMG 747 cells

grown for 48h at 28°C±2°C on NA medium, within the range of 10^8 to 10^0 CFU/mL. 5 µL of each dilution was used as DNA template for multiplex PCR reactions, which were carried out as described above but increasing the initial denaturation step to 10 min at 95°C to ensure bacterial cell lysis. The *Xaj* LMG 747 suspension of 10^8 CFU/mL was also used to determine the multiplex PCR detection limit in spiked walnut leaf samples. Walnut leaf fragments (about 0.6 g) were macerated as described previously. The resulting extract was used to prepare ten-fold serial dilutions of spiked walnut leaf samples ranging from 10^7 down to 10^0 CFU/mL. From each dilution, DNA was extracted using both CTAB-based method and QIAGEN DNeasy Plant Mini kit as described above. Multiplex PCR was carried out as described for naturally infected plant material.

2.8. Nucleotide sequences accession numbers

DNA sequences corresponding to the nine DNA markers (XAJ1 to XAJ9) were deposited in the NCBI database with the following accession numbers: KU577313 to KU577316 for XAJ1, KU577317 and KU577318 for XAJ2, KU577319 to KU577322 for XAJ3, KU577323 to KU577326 for XAJ4, KU577327 to KU577330 for XAJ5, KU577331 to KU577334 plus KX530957 for XAJ6, KU577335 to KU577337 for XAJ7, KU577338 to KU577341 for XAJ8, and KU577342 to KU577345 for XAJ9. NCBI accession numbers for *gyrB* of *Xaj* Portuguese isolates are shown in **Table II.3**.

3. Results

3.1. *In silico* selection of DNA markers specific for *Xanthomonas arboricola* pv. *juglandis*

The comprehensive BLAST analysis of all ORFs (> 100 bp) inferred from the draft genome of *Xaj* (NCPPB 1447 = LMG751, Accession PRJNA84273) resulted in the selection of nine markers, designated as XAJ1 to XAJ9 which showed low identity (E value > 1e-10, Query coverage < 35%) to sequences of non-xanthomonads species (**Table II.1**). Significant BLAST wgs hits were obtained with *Xanthomonas* species, namely with *X. arboricola* pv. *celebensis*, *X. perforans*, *X. axonopodis* pv. *glycines* and *X. hortorum pv. carotae* (**Table II.1**). Concerning ORF identity, none of the markers could be unequivocally assigned to a well characterized gene of the reference genome (Accession PRJNA84273), being either located in intergenomic regions or in genomic regions coding for hypothetical proteins (Supplementary **Table II.S1**).

3.2. *gyrB* sequencing analysis of *Xaj* isolates obtained from walnut samples

A set of 34 isolates were obtained from field samples of leaves and fruits from walnut trees exhibiting characteristic symptoms of WBB or/and BAN (**Table II.3**). The assignment of all isolates into *Xanthomonas arboricola* was confirmed by *gyrB* gene sequencing analysis. In fact, all isolates showed higher sequence similarity to *X. arboricola*, with a BLAST E value of 0.0 and identity of 99 to 100 %. In addition, the Maximum Likelihood tree showed that within the genus *Xanthomonas*, all the isolates clustered together with other *X. arboricola* pathovars. Moreover, four isolates (CPBF 1507, CPBF 1508, CPBF 1514 and CPBF 1522) were grouped in a different cluster (**Figure II.1**).

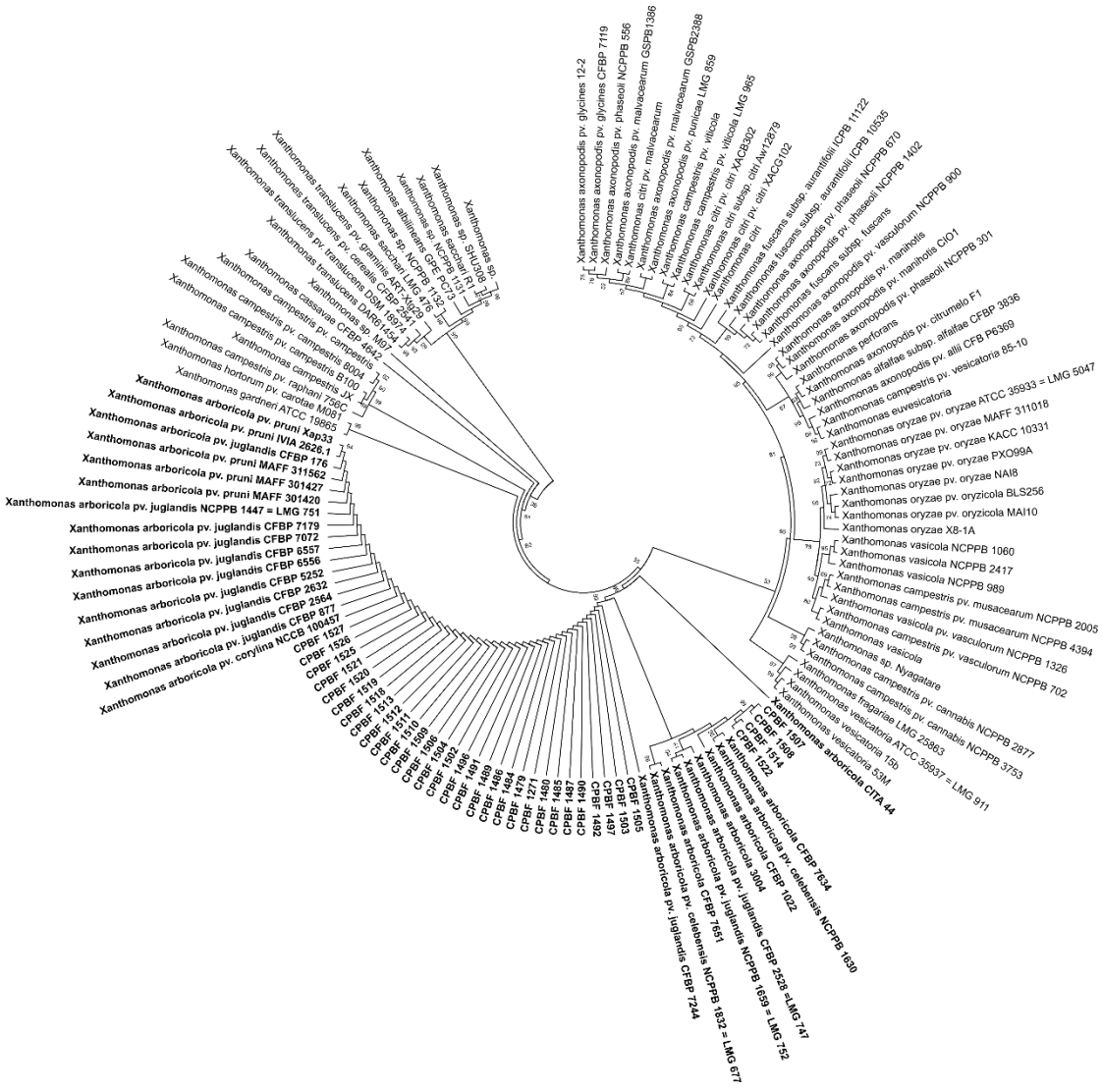

Figure II.1. Phylogenetic analysis based on 145 *gyrB* sequences of *Xanthomonas* species, including *Xanthomonas arboricola* isolates studied in this work. The tree was constructed according the Maximum likelihood method and Tamura-nei model using MEGA 6.06. Bootstrap values higher than 50 are shown. *Xanthomonas arboricola* strains and Portuguese isolates are highlighted in bold.

3.3. Experimental validation of the selected markers by dot blot

The dot blot assays carried out using as probes the nine markers (XAJ1 to XAJ9) obtained from *Xaj* strain LMG 751, confirmed the specificity of the markers to *Xaj* since no hybridization was observed whatever the non-target *Xanthomonas* spp. (**Table II.2**) tested (data not shown), with exception of markers XAJ1 and XAJ7 that hybridized to *X. arboricola* pv. *celebensis* LMG 677 (**Figure II.2**), as predicted by the initial BLAST analysis (**Table II.1**). Furthermore, from the 52 *Xaj* strains/isolates tested, 12 distinct hybridization profiles, designated as HP1 to HP12, were observed (**Figure II.2**). These hybridization patterns varied from *Xaj* hybridizing with the nine markers (HP1), to a *Xaj* strain (CFBP 2632) hybridizing with two markers (HP12). While most of these hybridization patterns were present in two to six *Xaj* strains/isolates (HP1, HP2, HP4, HP6, HP8, HP9, HP10, and HP11), one hybridization pattern was dominant and observed in 15 *Xaj* (HP10) and four hybridization patterns were unique to single *Xaj* strains/isolates (HP3, HP5, HP7, and HP12). Concerning the efficacy of the markers, it is evident that XAJ1, XAJ4, XAJ6 and XAJ8 hybridized to most of the *Xaj* strains and isolates, with emphasis for XAJ1 which hybridized with all the *Xaj* tested. Although the remaining markers were unable to hybridize with some *Xaj* strains/isolates, it is important to acknowledge their utility to obtain distinct hybridization patterns. For instance, the *Xaj* broad spectrum markers XAJ4, XAJ6 and XAJ8, did not hybridize with the group of *Xaj* strains and isolates sharing exactly the same hybridization pattern, i.e. HP9 and HP12 for marker XAJ4, HP11 for marker XAJ6, HP8 and HP11 for marker XAJ8 (**Figure. II.2**).

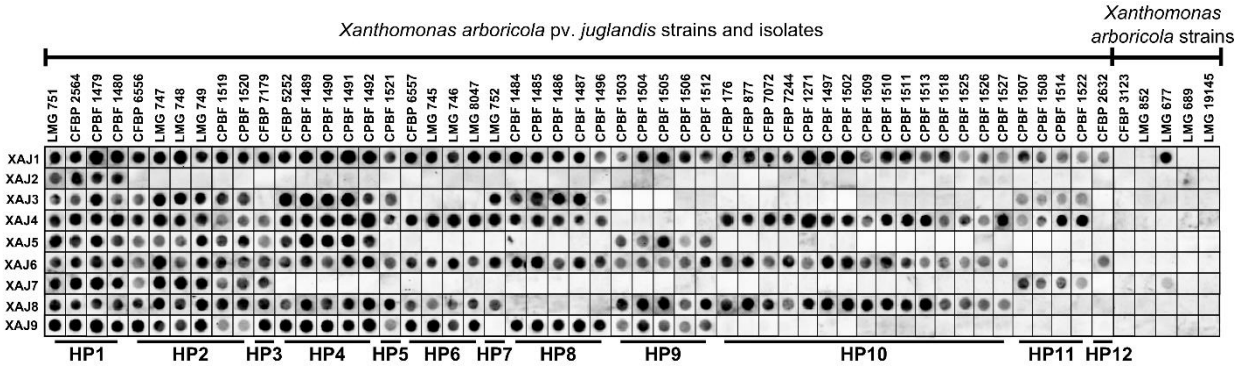

Figure II.2. Dot blot validation of *Xanthomonas arboricola* pv. *juglandis* specific markers (XAJ1 to XAJ9). Markers were tested across 18 *Xaj* reference strains, 34 *Xaj* isolates and 5 non-*juglandis* pathovars of *Xanthomonas arboricola* (**Table II.1** and **Table II.2**). Twelve different hibridization patterns (HP1 to HP12) were identified. The strain LMG 751 was used as positive hybridization control, and thirteen non-*arboricola Xanthomonas* species (**Table II.2**) were used as negative hybridization control (data not shown). DNA probes for the nine markers were prepared from the purified amplicons obtained from PCR amplifications using *Xaj* strain LMG 751 (= NCPPB 1447).

3.4. Multiplex PCR for identification of *Xaj* isolates and direct detection in infected walnut leaves and fruits

Reliable identification of putative *Xaj* isolates by multiplex PCR was optimized for markers XAJ1, XAJ6 and XAJ8. The multiplex PCR was validated with DNA extracted from all the strains and isolates listed in **Table II.2** and **Table II.3**, respectively. The results showed three distinct amplification profiles: amplification of the three markers for most *Xaj* strains/isolates (41 out of 52), amplification of two markers (XAJ1 and XAJ6) for one Xaj strain (CFBP 2632) and five *Xaj* isolates (CPBF 1484, CPBF 1485, CPBF 1486, CPBF 1487 and CPBF 1496), and amplification of a single marker (XAJ1) for four isolates (CPBF 1507, CPBF 1508, CPBF 1514 and CPBF 1522). As expected, whatever the marker, no specific amplification was obtained with any representative of other *Xanthomonas* species and with any other *X. arboricola* pathovars (pv. *corylina*, pv. *fragariae*, pv. *populi* and pv. *pruni*), with exception of pathovar *celebensis* (LMG 677) for which marker XAJ1 was amplified (Supplementary **Figure II.S1**). Interestingly, the larger amplicon (with 1.669 bp) observed for isolate CPBF 1497 was confirmed to correspond to marker XAJ6. Sequencing analysis of this amplicon (accession number KX530957) revealed the insertion of a gene coding for a transposase within marker XAJ6. Altogether these multiplex PCR results corroborated the dot blot patterns obtained with these three markers. In fact, matching results were observed for dot blot patterns and multiplex PCR for all the *Xaj* strains/isolates tested, including the *X. arboricola* pv. *celebensis* (LMG 677) the only non-*juglandis* pathovar strain for which amplification/hybridization of marker XAJ1 was observed.

Multiplex PCR was also able to detect *Xaj* in DNA extracted from naturally infected walnut leaves and fruits (**Figure II.3**). Indeed, marker XAJ6 was efficiently amplified for the 11 plant samples assayed, regardless the DNA extraction procedure and plant organ, i.e. infected leaves (L1 to L5) or fruits (F1 to F6). Marker XAJ1 was amplified for all samples, with exception for F6, and its amplification efficiency was higher for infected fruits samples and using CTAB-based DNA extraction. Marker XAJ8 was clearly the less efficient, with positive amplification for five out of six fruit samples, with no amplification observed for F6. Regarding the leaf samples this marker showed low efficiency in multiplex PCR with no amplification for most of the samples or poorly amplified in sample L1. Faint non-specific PCR bands were also observed for some samples. The efficiency of each of these three markers was higher in single PCR reactions for each the 11 plant samples assayed, particularly for the samples where no amplification was obtained with the multiplex PCR, emphasizing the utility of single PCR reactions to resolve the lower efficiency of multiplex PCR particularly observed with leaf samples and marker XAJ8 (**Figure II.3**).

In parallel, to demonstrate that plant samples were infected by *Xaj*, at least one putative *Xaj* isolate obtained from each of these infected leaves samples (L1 to L5), was screened for the

presence of XAJ6 marker and further identified by *gyrB* sequencing (Supplementary **Figure II.S2**). With exception for the *Xaj* putative isolate obtained from L1 sample, for which no XAJ6 and *gyrB* amplification was obtained, all the other isolates were positive for marker XAJ6 and *gyrB* (**Supplementary Figure II.S2**). Sequence analysis of both XAJ6 and *gyrB* for each of these isolates (L2 to L5), further confirmed their identity as *X. arboricola*. Taken together these results underline the robustness of the multiplex PCR for the direct detection of *Xaj* in symptomatic walnut leaves.

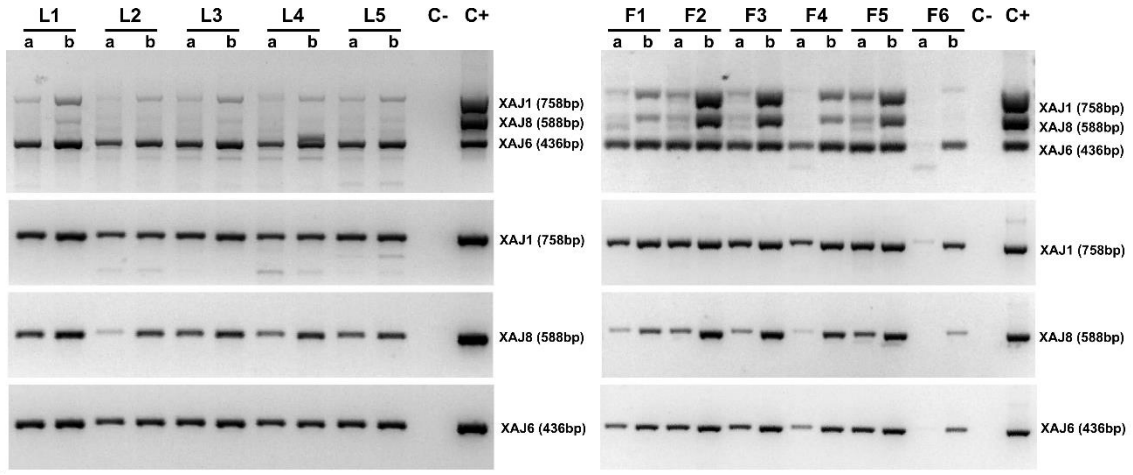

Figure II.3. Multiplex and single PCR reactions for detection of *Xaj* in naturally infected walnut leaves (L1 to L5) and fruits (F1 to F6) using three markers, XAJ1 (758 bp), XAJ6 (436 bp), and XAJ8 (588 bp). Each number represents a different walnut (*Juglans regia*) tree. For each sample two PCR reactions were carried out, one using DNA extracted with the QIAGEN DNeasy Plant Mini kit (a), and the other with a CTAB-based DNA extraction procedure (b). For all the samples the identity of XAJ6 marker was confirmed by sequencing. Strain LMG 751 was used as positive control (C+).

3.5. Multiplex PCR sensitivity

The detection limit of multiplex PCR, targeting the specific *Xaj* markers XAJ1, XAJ6 and XAJ8, determined by serial dilution of chromosomal DNA, bacterial cell suspensions and spiked plant extracts was, respectively, 20 pg DNA (10 pg/µL), 10^4 CFU/mL, and 10^3 CFU/mL (Supplementary **Figure II.S3**).

4. Discussion

Aiming to develop robust and culture-independent DNA-based approaches for identification of *Xaj* isolates and in *planta* detection of *Xaj*, several specific DNA loci were selected based on a thorough BLAST analysis of the first *Xaj* draft genome made available (Accession PRJNA84273). The putative *Xaj*-specific loci were then filtered to infer their suitability as DNA

markers considering their genomic stability and detection by PCR or dot blot hybridization, as previously described (Albuquerque et al. 2012).

The BLAST analysis revealed that the selected markers (XAJ1 to XAJ9), despite the low query coverage for markers XAJ4 and XAJ7, shared similarity with other *Xanthomonas* species (**Table II.1**, wgs blast). Significant BLAST hits were obtained with other plant pathogens never identified from walnut, as *X. arboricola* pv. *celebensis* (a banana plant pathogen), *X. perforans* (associated with tomato bacterial spot disease), *X. axonopodis* pv. *glycines* (causal agent of bacterial pustule of soybeans), *X. campestris* pv. *campestris* (attacks *Brassicas sp.*, responsible for the black rot of crucifers) and *X. hortorum* pv. *carotae* (a bacterial pathogen of carrot). The fact that these *Xanthomonas* spp. do not share the same ecological niche with *Xaj* (Leyns et al. 1984, Hayward 1993), supports the adequacy to use these markers to establish new molecular methodologies of detection and identification. Meanwhile, a complementary BLAST search allowed to reveal the presence of these markers, with exception for marker XAJ2, in the recently release *Xaj* genomes (Supplementary **Table II.S1**) (Cesbron et al. 2015, Higuera et al. 2015, Pereira et al. 2015).

In addition to the BLAST analysis, a validation by dot blot hybridization was carried out in order to confirm the consistency and specificity of the nine markers across a broad range of xanthomonads (**Figure II.2**). All the 18 *Xaj* strains and 34 *Xaj* isolates clustered into 12 different hybridization profiles, varying from hybridization with all the nine markers corresponding to HP1 (strains LMG 751 and CFBP 2564, Portuguese isolates CPBF 1479 and 1480), to hybridization with only two markers corresponding to HP12 (strain CFBP 2632). No evident relatedness was observed between these hybridization patterns and the geographic origin from which the isolates were obtained. Indeed, isolates from the same location displayed different hybridization profiles (e.g. HP4 for CPBF 1489 and HP10 for CPBF 1502 from Alcobaça), and inversely, identical hybridization patterns were found for isolates from distinct locations (e.g. HP10 for the isolates CPBF 1510 from Azeitão and CPBF 1513 from Estremoz). Taken together, these results demonstrate the utility of this dot blot platform to disclose the high genomic diversity of *Xaj*, as reported by several genotyping studies (Loreti et al. 2001, Scortichini et al. 2001, Hajri et al. 2010, Marcelletti et al. 2010, Hajri et al. 2012, Kaluzna et al. 2014, Essakhi et al. 2015, Ivanovic et al. 2015, Giovanardi et al. 2015). Most importantly, the distinct hybridization patterns can be informative to characterize different *Xaj* lineages, which can be particularly convenient before carrying out laborious fine tune methods of genotyping such as Multilocus sequence analysis (MLSA) and Multilocus sequence typing (MLST), or full genome sequencing.

When looking at the dot blot efficacy, it is evident that markers XAJ1, XAJ4, XAJ6 and XAJ8 hybridized with most of the *Xaj* strains tested (**Figure II.2**) and should be primary choices for detection and identification of *Xaj*. Interestingly, the *Xaj* isolates which did not hybridize to

markers XAJ4, XAJ6 and XAJ8 shared identical hybridization profiles, suggesting that the genotype of these markers might be determined by either their presence or absence. In this regard, it is worth mentioning that the four *Xaj* isolates (CPBF 1507, 1508, 1514 and 1522) corresponding to HP11, were PCR positive exclusively to marker XAJ1 in the multiplex assay as expected (Supplementary **Figure II.S1**), and clustered together in the *gyrB* tree (**Figure II.1**).

The specificity of the selected markers to *Xaj* was shown by the fact that none of the nine markers hybridized with the 12 non-*arboricola* strains of *Xanthomonas* spp., or to any of the other five *Xanthomonas arboricola* pathovars studied, with exception for *X. arboricola* pv. *celebensis* LMG 677, which hybridized to markers XAJ1 and XAJ7, confirming the blast wgs analysis (**Table II.1**).

Currently, only a PCR method has been reported for identification of *Xaj* isolates (Gironde et al. 2009). Although this PCR approach has been used in recent studies (Burokiene and Pulawska 2012, Kaluzna et al. 2014, Ivanovic et al. 2015), the method is based on a single DNA marker and the sequence of the primers has not been published, undermining the fully assessment of this method, namely in infected plant material. In this work we propose a triplex PCR using markers XAJ1, XAJ6 and XAJ8, optimized to ensure an efficient, reliable and rapid method to identify *Xaj* isolates, as well as for culture-independent detection of *Xaj* in plant samples. These three markers were chosen due to their different amplicon sizes and their consistency across the different *Xaj* strains (**Figure II.2**). The validation of the multiplex PCR was carried out with DNA extracted from all the *Xaj* strains and isolates, other pathovars of *X. arboricola*, and non-*arboricola* species of *Xanthomonas* included in this work (**Table II.2** and **Table II.3**). The PCR multiplex results fully confirmed the dot blot data, i.e. amplification was observed for any of the three markers, whenever the dot blot was positive for the corresponding marker, and whatever the *Xaj* strain or isolate tested (**Figure II.2**, Supplementary **Figure. II.S1**). It should be noted that a simplex PCR using marker XAJ1 would also be satisfactory for routine detection of *Xaj* in infected plant samples, since this marker is highly specific and present throughout all tested *Xaj* strains. However, the higher detection confidence provided by the use of three independent markers cannot be underestimated, especially when considering *Xaj* diversity.

PCR based methods can be a valuable approach for direct detection of *Xaj* in plant tissues, allowing a rapid and culture-independent detection of *Xaj* in walnuts, relevant for an early diagnosis of WBB or BAN and to determine the bacteria life cycle overwinter. In this study, we were able to amplify the markers XAJ1, XAJ6 and XAJ8 in DNA extracted from naturally infected walnut leaves and fruits indicating the presence of *Xaj* (**Figure II.3**). It has been acknowledged that the quality, purity and quantity of DNA can impair the detection in plant material (López et al. 2003, Demeke and Jenkins 2010). The multiplex efficiency, particularly low for the larger

markers (XAJ1 and XAJ8) could be related with the presence of plant-derived compounds or even treatment derived chemicals that may interfere with the amplification (Schrader et al. 2012). The higher efficiency observed for DNA extracted using the CTAB procedure in comparison with DNeasy, and for fruit samples than for leaves, seems to support this hypothesis. Furthermore, when the three markers were amplified in independent PCR reactions, markers' efficiency increased considerably, whatever the plant material and the DNA extraction procedure (**Figure II.3**). The faint unspecific amplifications observed in the multiplex PCR were likely favoured by the lower annealing temperature, since most of these unspecific bands disappeared in the single PCR amplification of each marker (**Figure II.3**). The detection threshold of multiplex PCR for chromosomal DNA (10 pg/µL) and bacterial cells (10^4 CFU/mL corresponding to 50 cells per reaction) are within the sensitivity range of multiplex PCR for other *Xanthomonas* (Supplementary **Figure II.S3**) (Toth et al. 1998, Pothier et al. 2011, Araujo et al. 2012). In addition, the multiplex PCR detection limit of 10^4 CFU/mL *Xaj* in spiked walnut samples, (i.e. 2 cells per reaction) (Supplementary **Figure II.S3**), demonstrate that the load of *Xaj* cells in field infected walnut leaves and fruits (**Figure II.3**) are well above this detection limit, which strengthens the effectiveness of this method for routine detection and to access *Xaj* in latently infected plant material (Hayward 1974).

In conclusion, this work proposes nine novel molecular markers for identification of *Xaj* isolates and its detection in infected plant samples using a dot blot hybridization assay or multiplex PCR. While four markers (XAJ1, XAJ4, XAJ6, XAJ8) were shown to be particularly effective, since they were detected in most of the *Xaj* tested, regardless their genetic diversity, the other five markers, although clearly specific to *Xaj*, displayed a narrower, but complementary detection range. Beyond the usefulness of these markers to identify and discriminate different *Xaj* lineages, we demonstrated that the broad range markers XAJ1, XAJ6 and XAJ8 might be used in culture-independent approaches. Attending that direct detection of *Xaj* in plant material has never been described, these results are a promising contribution for the early and reliable diagnostics of walnut diseases caused by *Xaj*.

Acknowledgements

This work was co-financed by the European Structural & Investment Funds (ESIFs) through the Operational Competitiveness and Internationalization Programme –COMPETE 2020 and by National Funds through FCT – Foundation for Science and Technology, within the framework of the project EVOXANT (PTDC/BIA-EVF/3635/2014). Pedro Albuquerque was supported by the project 'Genomics Applied to Genetic Resources' co-financed by North Portugal Regional Operational Programme 2007/2013 (ON.2 – O Novo Norte), under the National Strategic

Reference Framework (NSRF), through the European Regional Development Fund (ERDF). Camila Fernandes is supported by a fellowship from the FCT (SFRH/BD/95913/2013). The authors would like to thank the walnut producers who collaborated in this study, Cláudia Serra for helpful discussions and to Paula Sá-Pereira for the fruitful comments made on the manuscript.

References

Albuquerque P, Caridade CMR, Rodrigues AS, Marcal ARS, Cruz J, Cruz L, Santos CL, Mendes MV, Tavares F. 2012. Evolutionary and experimental assessment of novel markers for detection of *Xanthomonas euvesicatoria* in plant samples. *Plos One* **7**, 15.

Allen C, Bent A, Charkowski A. 2009. Underexplored niches in research on plant pathogenic bacteria. *Plant Physiology* **150**,1631-1637.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *Journal of Molecular Biology* **215**, 403-410.

Araújo ER, Costa JR, Ferreira MASV, Quezado-Duval AM. 2012. Simultaneous detection and identification of the *Xanthomonas* species complex associated with tomato bacterial spot using species-specific primers and multiplex PCR. *Journal of applied microbiology* **113**, 1479-1490.

Belisario A, Maccaroni M, Corazza L, Balmas V, Valier A. 2002. Occurrence and etiology of brown apical necrosis on Persian (English) walnut fruit. *Plant Disease* **86**, 599-602.

Burokiene D, Pulawska J. 2012. Characterization of *Xanthomonas arboricola* pv. *juglandis* isolated from walnuts in Lithuania. *Journal of Plant Pathology* **94**, S1. 23-S21. 27.

Cesbron S, Briand M, Essakhi S, Gironde S, Boureau T, Manceau C, Fischer-Le Saux M, Jacques MA. 2015. Comparative genomics of pathogenic and nonpathogenic strains of *Xanthomonas arboricola* unveil molecular and evolutionary events linked to pathoadaptation. *Frontiers in Plant Science* **6**, 1126.

Demeke T, Jenkins GR. 2010. Influence of DNA extraction methods, PCR inhibitors and quantification methods on real-time PCR assay of biotechnology-derived traits. *Analytical and Bioanalytical Chemistry* **396**, 1977-1990.

Doyle J. 1991. DNA protocols for plants. Molecular Techniques in Taxonomy. Heidelberg, Berlin: Springer.

Essakhi S, Cesbron S, Fischer-Le Saux M, Bonneau S, Jacques MA, Manceau C. 2015. Phylogenetic and variable-number tandem-repeat analyses identify nonpathogenic *Xanthomonas arboricola* lineages lacking the canonical Type III Secretion System. *Applied and Environmental Microbiology* **81**, 5395-5410.

Fischer-Le Saux M, Bonneau S, Essakhi S, Manceau C, Jacques MA. 2015. Aggressive emerging pathovars of *Xanthomonas arboricola* represent widespread epidemic clones distinct from poorly pathogenic strains, as revealed by multilocus sequence typing. *Applied and Environmental Microbiology* **81**, 4651-4668.

Frutos D. 2010. Bacterial diseases of walnut and hazelnut and genetic resources. *Journal of Plant Pathology* **92**, S79-S85.

Giovanardi D, Bonneau S, Gironde S, Fischer-Le Saux M, Manceau C, Stefani E. 2015. Morphological and genotypic features of *Xanthomonas arboricola* pv. *juglandis* populations from walnut groves in Romagna region, Italy. *European Journal of Plant Pathology* **145**, 1-16.

Gironde S, Guillaumes J, Manceau C. 2009. Specific detection of *Xanthomonas arboricola* pv. *juglandis* pathogen on walnut. EPPO Conference on Diagnostics and Associated workshops, York, UK.

Hajri A, Meyer D, Delort F, Guillaumes J, Brin C, Manceau C. 2010. Identification of a genetic lineage within *Xanthomonas arboricola* pv. *juglandis* as the causal agent of vertical oozing canker of Persian (English) walnut in France. *Plant Pathology* **59**, 1014-1022.

Hajri A, Pothier JF, Fischer-Le Saux M, Bonneau S, Poussier S, Boureau T, Duffy B, Manceau C. 2012. "Type Three Effector gene distribution and sequence analysis provide new insights into the pathogenicity of plant-pathogenic *Xanthomonas arboricola*. *Applied and Environmental Microbiology* **78**, 371-384.

Hayward AC. 1974. Latent infections by bacteria. *Annual review of phytopathology* **12**, 87-97.

Hayward AC.1993. The hosts of *Xanthomonas*. *Xanthomonas*. London: Chapman & Hall.

Higuera G, Gonzalez-Escalona N, Veliz C, Vera F, Romero J. 2015. Draft genome sequences of four *Xanthomonas arboricola* pv. *juglandis* strains associated with walnut blight in Chile. *Genome announcements* **3**, e01160-15.

Ivanovic Z, Popovic T, Janse J, Kojic M, Stankovic S, Gavrilovic V, Fira D. 2015. Molecular assessment of genetic diversity of *Xanthomonas arboricola* pv. *juglandis* strains from Serbia by various DNA fingerprinting techniques. *European Journal of Plant Pathology* **141**, 133-145.

Janse J. 2010. Diagnostic methods for phytopathogenic bacteria of stone fruits and nuts in COST 873. EPPO bulletin **40**, 68-85.

Janse J, Rossi M, Gorkink R, Derks J, Swings J, Janssens D, Scortichini M. 2001. Bacterial leaf blight of strawberry (*Fragaria* (x) *ananassa*) caused by a pathovar of *Xanthomonas arboricola*, not similar to *Xanthomonas fragariae* Kennedy & King. Description of the causal organism as *Xanthomonas arboricola* pv. *fragariae* (pv. nov., comb. nov.). *Plant Pathology* **50**, 653-665.

Kaluzna M, Pulawska J, Waleron M, Sobiczewski P. 2014. The genetic characterization of *Xanthomonas arboricola* pv. *juglandis*, the causal agent of walnut blight in Poland. *Plant Pathology* **63**, 1404-1416.

Lamichhane JR. 2014. *Xanthomonas arboricola* diseases of stone fruit, almond, and walnut trees: Progress toward understanding and management. *Plant Disease* **98**, 1600-1610.

Lelliott RA, Stead DE. 1987. Methods for the diagnosis of bacterial diseases of plants. Methods in Plant Pathology. Oxford, UK: British Society for Plant Pathology by Blackwell Scientific Publications.

Leyns F, Decleene M, Swings JG, Deley J. 1984. The host range of the genus *Xanthomonas*. *Botanical Review* **50**, 308-356.

Lindow S, Olson W, Buchner R. 2014. Colonization of dormant walnut buds by *Xanthomonas arboricola* pv. *juglandis* is predictive of subsequent disease. *Phytopathology* **104**, 1163-1174.

López MM, Bertolini E, Olmos A, Caruso P, Gorris MT, Llop P, Penyalver R, Cambra M. 2003. Innovative tools for detection of plant pathogenic viruses and bacteria. *International Microbiology* **6**, 233-243.

Loreti S, Gallelli A, Belisario A, Wajnberg E, Corazza L. 2001. Investigation of genomic variability of *Xanthomonas arboricola* pv. *juglandis* by AFLP analysis. *European Journal of Plant Pathology* **107**, 583-591.

Marcelletti S, Ferrante P, Scortichini M. 2010. Multilocus sequence typing reveals relevant genetic variation and different evolutionary dynamics among strains of *Xanthomonas arboricola* pv. *juglandis*. *Diversity* **2**, 1205.

Moragrega C. 2012. Detection and identification methods and new tests as developed and used in the framework of COST 873 for bacteria pathogenic to stone fruits and nuts: *Xanthomonas arboricola* pv. *juglandis*. *Journal of Plant Pathology* **94**, S1.155-S151.159.

Moragrega C, Matias J, Aleta N, Montesinos E, Rovira M. 2011. Apical necrosis and premature drop of Persian (English) walnut fruit caused by *Xanthomonas arboricola* pv. *juglandis*. *Plant Disease* **95**, 1565-1570.

Moragrega C, Ozaktan H. 2010. Apical necrosis of Persian (English) walnut (*Juglans regia*): An update. *Journal of Plant Pathology* **92**, S67-S71.

Mulrean EN, Schroth MN. 1982. Ecology of *Xanthomonas campestris* pv. *juglandis* on Persin (English) walnuts. *Phytopathology* **72**, 434-438.

Murray MG, Thompson WF. 1980. Rapid isolation of high molecular-weight plant DNA. *Nucleic Acids Research* **8**, 4321-4325.

Pereira UP, Gouran H, Nascimento R, Adaskaveg JE, Goulart LR, Dandekar AM. 2015. Complete genome sequence of *Xanthomonas arboricola* pv. *juglandis* 417, a copper-resistant strain isolated from *Juglans regia* L. *Genome announcements* **3**, e01126-15.

Pierce NB. 1901. Walnut bacteriosis. *Botanical Gazette* **31**, 272-273.

Pothier JF, Pagani MC, Pelludat C, Ritchie DF, Duffy B. 2011. A duplex-PCR method for species- and pathovar-level identification and detection of the quarantine plant pathogen *Xanthomonas arboricola* pv. *pruni*. *Journal of microbiological methods* **86**, 16-24.

Schrader C, Schielke A, Ellerbroek L, Johne R. 2012. PCR inhibitors–occurrence, properties and removal. *Journal of applied microbiology* **113**, 1014-1026.

Scortichini M, Marchesi U, Di Prospero P. 2001. Genetic diversity of Xanthomonas arboricola pv. *juglandis* (synonyms: *X. campestris* pv. *juglandis*, *X. juglandis* pv. *juglandis*) strains from different geographical areas shown by repetitive polymerase chain reaction genomic fingerprinting. *Journal of Phytopathology-Phytopathologische Zeitschrift* **149**, 325-332.

Scortichini M. 2010. Epidemiology and predisposing factors of some major bacterial diseases of stone and nut fruit trees species. *Journal of Plant Pathology* **92**, S73-S78.

Smith RE, Smith CO, Ramsey HJ. 1912. Walnut culture in California: walnut blight. Berkeley, California: California Agricultural Experiments Station Publications.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution* **30**, 2725-2729.

Toth IK, Hyman LJ, Taylor R, Birch PRJ. 1998. PCR-based detection of *Xanthomonas campestris* pv. *phaseoli* var. *fuscans* in plant material and its differentiation from *X. c.* pv. *phaseoli*. *Journal of applied microbiology* **85**, 327-336.

Vauterin L, Hoste B, Kersters K, Swings J. 1995. Reclassification of *Xanthomonas*. *International Journal of Systematic Bacteriology* **45**, 472-489.

Young JM, Park DC, Shearman HM, Fargier E. 2008. A multilocus sequence analysis of the genus *Xanthomonas*. *Systematic and Applied Microbiology* **31**, 366-377.

Supplementary material

Table II.S1. Annotation of the nine regions selected as DNA markers throughout the 18 sequenced *Xaj* genomes.

Strain \ Marker	XAJ1	XAJ2	XAJ3	XAJ4	XAJ5	XAJ6	XAJ7	XAJ8	XAJ9
NCPPB 1447	intergenic	Q55_RS23110 (hypothetical protein)	intergenic	Q55_RS0116845 (hypothetical protein)	Q55_RS0117840 (hypothetical protein)	Q55_RS0118740 (hypothetical protein)	intergenic	Q55_RS0121810 (hypothetical protein)	Q55_RS0122690 (hypothetical protein)
417	intergenic	-	-	AKJ12_RS00160 (hypothetical protein)	-	AKJ12_RS16090 (hypothetical protein)	-	AKJ12_RS00090 (hypothetical protein)	-
CFBP 2528	intergenic	-	Fwd primer only	XARJCFBP2528_RS05355 (hypothetical protein)	XARJCFBP2528_RS12000 (hypothetical protein)	XARJCFBP2528_RS15805 (hypothetical protein)	intergenic	XARJCFBP2528_RS05295 (hypothetical protein)	XARJCFBP2528_RS10940 (hypothetical protein)
CFBP 7179	intergenic	-	-	XARJCFBP7179_RS08435 (hypothetical protein)	XARJCFBP7179_RS09700 (hypothetical protein)	XARJCFBP7179_RS16695 (hypothetical protein)	intergenic	XARJCFBP7179_RS08410 (hypothetical protein)	XARJCFBP7179_RS08645 (hypothetical protein)
J303	intergenic	-	Non-contiguous	AXA70_RS02980 (hypothetical protein)	-	AXA70_RS17965 (hypothetical protein)	-	-	AXA70_RS11715 (hypothetical protein)
CFSAN033077	intergenic	-	Non-contiguous	AE920_19320 (hypothetical protein)	-	AE920_10440 (hypothetical protein)	-	AE920_07710 (hypothetical protein)	AE920_04700 (hypothetical protein)
CFSAN033078	intergenic	-	-	AE921_RS18610 (hypothetical protein)	-	AE921_RS08240 (hypothetical protein)	-	AE921_RS02215 (hypothetical protein)	-
CFSAN033079	intergenic	-	-	AE922_RS17765 (hypothetical protein)	-	AE922_RS02770 (hypothetical protein)	-	AE922_RS03600 (hypothetical protein)	-
CFSAN033080	intergenic	-	-	AE923_RS07875 (hypothetical protein)	-	AE923_RS03580 (hypothetical protein)	-	AE923_RS15205 (hypothetical protein)	-
CFSAN033081	intergenic	-	Non-contiguous	AE924_RS07140 (hypothetical protein)	-	AE924_RS16670 (hypothetical protein)	-	AE924_RS12985 (hypothetical protein)	AE924_RS15980 (hypothetical protein)
CFSAN033082	intergenic	-	-	AE925_10800 (hypothetical protein)	-	Non contiguous	-	AE925_RS04700 (hypothetical protein)	-
CFSAN033083	intergenic	-	-	AE926_RS14940 (hypothetical protein)	-	AE926_RS09245 (hypothetical protein)	-	AE926_RS14235 (hypothetical protein)	-
CFSAN033084	intergenic	-	-	AE927_RS19725 (hypothetical protein)	-	AE927_RS11905 (hypothetical protein)	-	AE927_RS05120 (hypothetical protein)	-
CFSAN033085	intergenic	-	-	AE928_RS13880 (hypothetical protein)	-	AE928_RS11390 (hypothetical protein)	-	AE928_RS12415 (hypothetical protein)	-
CFSAN033086	intergenic	-	Non-contiguous	AE929_RS05095 (hypothetical protein)	-	AE929_RS02400 (hypothetical protein)	-	AE929_RS16110 (hypothetical protein)	AE929_RS01845 (hypothetical protein)
CFSAN033087	intergenic	-	-	AE930_RS16010 (hypothetical protein)	-	AE930_RS01095 (hypothetical protein)	-	AE930_RS04125 (hypothetical protein)	-
CFSAN033088	intergenic	-	-	AE931_RS04865 (hypothetical protein)	-	AE931_RS04565 (hypothetical protein)	-	AE931_RS06740 (hypothetical protein)	-
CFSAN033089	intergenic	-	-	AE932_RS19640 (hypothetical protein)	-	AE932_RS11510 (hypothetical protein)	-	AE932_RS17595 (hypothetical protein)	-

-, no similarity has been found; non-contiguous, primers are present in the genome, but do not amplify the DNA fragment of expected size.

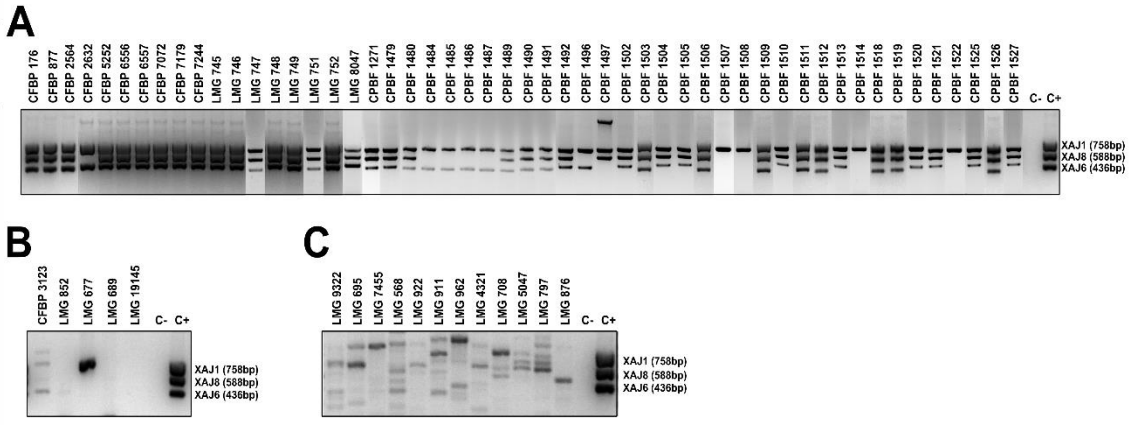

Figure II.S1. Multiplex PCR identification of *Xaj* using the markers, XAJ1 (758 bp), XAJ6 (436 bp) and XAJ8 (588 bp). DNA was extracted from pure cultures of bacterial reference strains (**Table II.1**) and isolates (**Table II.2**). A - Multiplex PCR results obtained for 18 *Xaj* strains, and 34 *Xaj* isolates, B - Multiplex PCR results obtained from 5 non-*juglandis* pathovars of *Xanthomonas arboricola*, C - Multiplex PCR results obtained from twelve non-*arboricola Xanthomonas* species. *Xaj* strains LMG 751 and LMG 747 were used as positive control (C+).

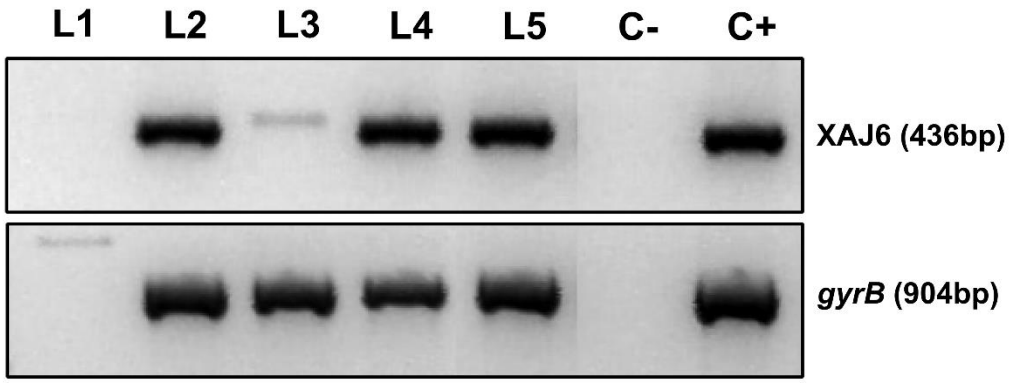

Figure II.S2. PCR amplification of *gyrB* (904 bp) and XAJ6 marker (436 bp) of five isolates obtained from each sample of the naturally infected walnut leaves (L1 to L5). The identity of the *gyrB* and XAJ6 PCR products were confirmed by sequencing. Strain LMG 751 was used as positive control (C+).

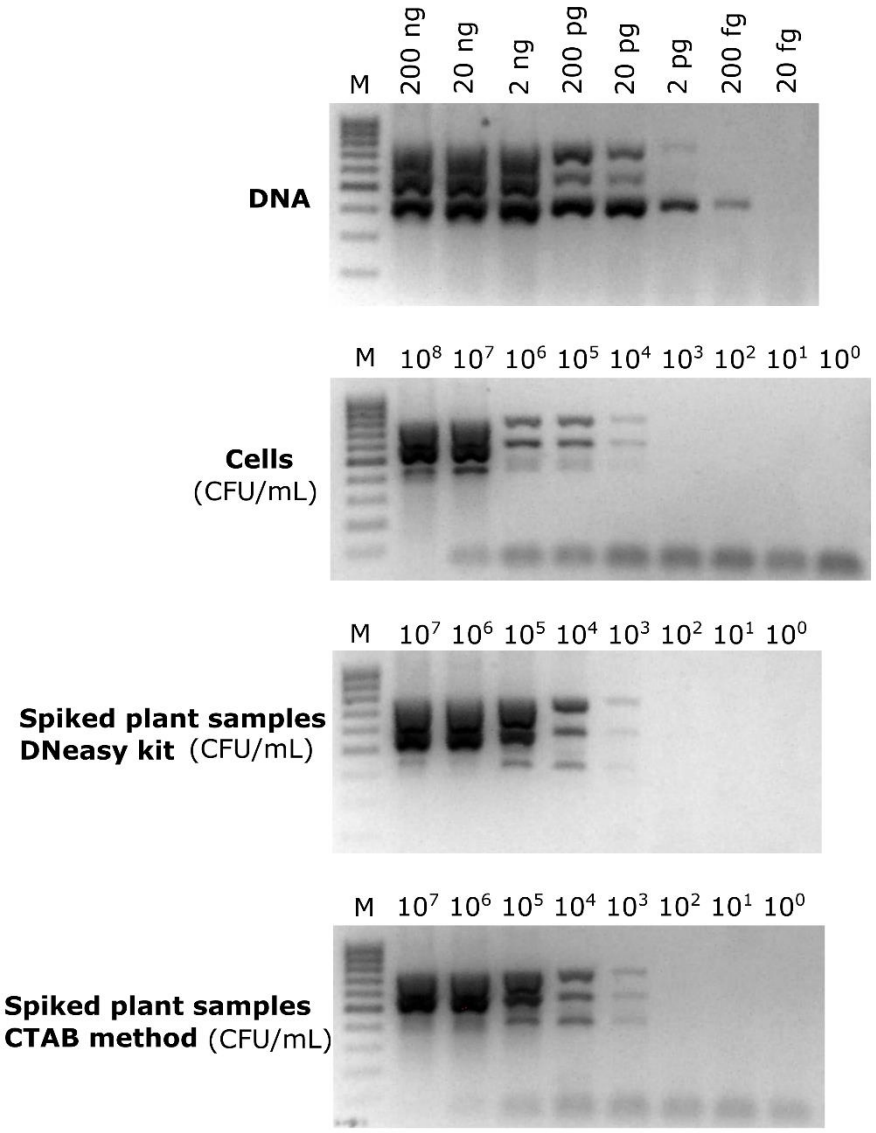

Figure II.S2. Detection limit of multiplex PCR (XAJ1 - 758 bp, XAJ6 - 436 bp and XAJ8 - 588 bp) using *Xaj* strain LMG 747. DNA serial dilutions ranging from 200 ng to 20 fg, Serial dilutions of bacterial cell suspensions ranging from 10^8 CFU/mL to 10^0 CFU/mL, Serial dilutions of spiked plant samples ranging from 10^7 CFU/mL to 10^0 CFU/mL after DNA extraction by QIAGEN DNeasy Plant Mini kit and CTAB-based method. M - Molecular marker 100 bp DNA Ladder (Solis BioDyne, Tartu, Estonia).

CHAPTER III

Assessment of *Xanthomonas arboricola* pv. *juglandis* bacterial load in infected walnut fruits by qPCR

Chapter III include the following publication:

Martins L, Fernandes C, Alburquerque P, Tavares F. 2019. Assessment of *Xanthomonas arboricola* pv. *juglandis* bacterial load in infected walnut fruits by qPCR. *Plant Disease*. (In press)

Part of the results from this chapter were presented at:

Martins L, Fernandes C, Albuquerque P, Tavares F. 2018. Development of a qPCR method to determine *Xanthomonas arboricola* pv. *juglandis* load in infected walnut samples. 6th *Xanthomonas* Genomics Conference. 2nd Annual EuroXanth Conference, Leopoldina, Halle (Saale), Germany, July 18-21, 2018.

Assessment of *Xanthomonas arboricola* pv. *juglandis* bacterial load in infected walnut fruits by qPCR

Abstract

Xanthomonas arboricola pv. *juglandis* (*Xaj*) is the etiologic agent of important walnut (*Juglans regia* L.) diseases, causing severe fruit drop and high economic losses in walnut production regions. Rapid diagnostics and knowledge of bacterial virulence fitness are key to hinder disease progression and to apply timely phytosanitary measures. This work describes a *Xaj*-specific real-time quantitative PCR (qPCR) using *Xaj* specific DNA markers to quantify *Xaj* bacterial load in walnut infected plant tissues. Method validation was achieved using calibration curves obtained with serial dilutions of *Xaj* chromosomal DNA and standard curves obtained from walnut samples spiked with *Xaj* cells. High correlation (R^2>0.990 and R^2>0.995) and low limit of detection (35 chromosomes/qPCR reaction and 2.7 CFU/qPCR reaction) were obtained for both markers considering the calibration and standard curves, respectively. Assessment of qPCR repeatability, reproducibility, and specificity allowed us to demonstrate the reliability and consistency of the method. Furthermore, *in planta* quantification of *Xaj* bacterial load using infected walnut fruit samples showed a higher detection resolution compared with standard PCR detection. By allowing quantifying the virulence fitness of distinct *Xaj* strains in planta, the proposed qPCR method may contribute to assertive risk assessment of walnut diseases caused by *Xaj*, and ultimately help to improve phytosanitary practices.

1. Introduction

Xanthomonas arboricola pv. *juglandis* (*Xaj*) is the etiological agent of walnut bacteria blight (WBB), the most important bacterial disease of walnut (*Juglans regia* L.) (Frutos 2010). *Xaj* has also been associated with brown apical necrosis (Moragrega et al. 2011) and vertical oozing canker diseases (Hajri et al. 2010), altogether causing major economic losses attributable to its worldwide distribution and persistence (Frutos and López 2012, Lamichhane 2014).

In the last decade, several studies emphasized the genetic diversity of *Xaj* populations, contributing to improve our knowledge on the epidemic behavior of this pathogen (Hajri et al. 2012, Cesbron et al. 2014, Ivanovic et al. 2014, Fischer-Le Saux et al. 2015). In addition, nonpathogenic *X. arboricola* lineages were isolated from symptomatic walnut trees. These genetically diverse strains were shown to group separately from the well-defined clusters of pathogenic *Xaj* strains (Cesbron et al. 2015, Essakhi et al. 2015), also supporing differences in virulence (Giovanardi et al. 2016). Considering the high genetic diversity of *X. arboricola* strains and the isolation of related nonpathogenic bacteria, there is a need to characterize walnut-colonizing xanthomonads, taking into account their pathogenicity and virulence, to make robust risk assessments and contribute to implementation of timely and scientifically based phytosanitary practices.

Pathogenicity and virulence are often inaccurately used as synonyms and are sometimes inconsistently used in the scientific literature, leading to debates regarding the most appropriate definitions of these terms (Bos and Parlevliet 1995, Casadevall and Pirofski 1999, Thomas and Elkinton 2004, Shapiro-Ilan et al. 2005). Despite these disputes, the thorough revision made by Shapiro-Ilan et al. (2005) provides a coherent understanding of these two concepts, defining pathogenicity as the ability to cause disease, whereas virulence refers to pathogen fitness or its aptitude to colonize the host. Thus, although pathogenicity is a qualitative trait, virulence refers to pathogen's capabilities to thwart the host's defenses, ultimately leading to a disease phenotype. Aligned with these definitions, some authors define virulence as inversely proportional to the number of microorganisms required to trigger infection (Casadevall and Pirofski 2001, Shapiro-Ilan et al. 2005). These and other related concepts frequently used in phytopathology were revisited in a thoughtful pedagogic essay addressing the above definitions (Surico 2013). Whatever the semantic understanding of pathogenicity and virulence, the interest to quantifying the bacterial load in infected tissues in order to determine the pathogen fitness of a given infective strain is irrefutable. Currently, most of the studies to determine virulence of phytopathogenic bacteria are based on a gradient scale of symptoms (Frutos and López 2012, Bandi et al. 2015). Although this procedure may provide a semiquantitative measure of virulence and might distinguish a poorly virulent strain from a highly virulent strain, it is dependent on a

visual or phytopathometrical evaluation of the symptoms and not suitable to accurately quantify the bacterial load.

Quantitative PCR (qPCR) is a frequently used method in environmental microbiology to measure target DNA from environmental samples, using a standard curve plotted from known concentrations of target DNA (Pérez et al. 2013). Despite the existence of numerous real-time PCR methods established for several phytopathogens, the quantification of plant pathogenic microorganisms in infected plant material or in environmental matrices such as soils, still lacks adequate optimization to assist plant disease management practices (Lievens et al. 2006, Alemu 2014).

Real-time PCR protocols have been widely used to detect and identify several members of the *Xanthomonas* genus, namely *X. campestris* (Berg et al. 2006), *X. fragariae* (Weller et al. 2007, Vandroemme et al. 2008), *X. axonopodis* cv. *citri* (Golmohammadi et al. 2007), *X. oryzae* pv. *oryzae* (Cho et al. 2011), *X. arboricola* pv. *pruni* (Palacio-Bielsa et al. 2011, 2015) *X. campestris* pv. *campestris* (Laala et al. 2015) and *X. axonopodis* pv. *allii* (Robène et al. 2015). Despite the importance of these contributions, the qPCR potential beyond detection and identification of *Xanthomonas* species remains poorly explored. One exception has been a qPCR method developed to distinguish resistant and susceptible cultivars based on quantification of *X. albilineans* (Garces et al. 2014).

Advances concerning a framework to select taxon-specific DNA markers (i.e. DNA signatures) successfully validated for PCR-based detection of *Xanthomonas* species in infected plant samples (Albuquerque et al. 2012) further empower the development of qPCR approaches to determine the bacterial load in several host plants. Recently, nine *X. arboricola* pv. *juglandis*-specific DNA markers (XAJ1-XAJ9) were selected and validated in 2017 (Fernandes et al. 2017). Three of these markers (XAJ1, XAJ6, and XAJ8) were shown to efficiently detect *Xaj* in naturally infected walnut leaves and fruits using a multiplex PCR approach (Fernandes et al. 2017), raising the opportunity to use these DNA markers to measure *Xaj* virulence fitness from symptomatic walnut leaves and fruits.

In this work, we developed a qPCR method using two *Xaj* specific DNA markers able to quantify *Xaj* bacterial load from infected walnut plant tissues and measure virulence of distinct *Xaj* strains. This method may contribute to rapid identification of highly virulent *Xaj* lineages or strains in walnut orchards, which is essential to take timely phytosanitary action and perform epidemiological risk assessments of *Xaj*-causing diseases.

2. Material and methods

2.1. Bacterial cultures and DNA extraction

The bacterial strains used in this study included 25 *Xanthomonas arboricola* pv. *juglandis* strains, seven *Xanthomonas arboricola* strains not belonging to the *juglandis* pathovar and 15 strains of other *Xanthomonas* species (**Table III.1**). Single colonies were grown in Nutrient broth (2 g/litre of yeast extract, 5 g/litre of peptone, 5 g/litre of NaCl, 0.45 g/litre of KH_2PO_4, 2.29 g/litre of $Na_2HPO_4 \cdot 12H_2O$) and DNA was extracted using the E.Z.N.A. Bacterial DNA kit (Omega Bio-Tek, Norcross, GA), following the manufacturer's instructions. DNA quantification was carried out using the Qubit 2.0 fluorometer (Invitrogen, Carlsbad, CA).

Table III.1. List of bacteria strains used in this study

Strain[z]	Geographic origin	Year of isolation
Xanthomonas arboricola* pv. *juglandis		
CFBP 176	France	1961
CFBP 877	France	1966
CFBP 2564	Italy	1985
CFBP 2632	Spain	1984
CFBP 5252	France	2000
CFBP 6556	Italy	1993
CFBP 6557	Italy	1999
CFBP 7072	Spain	1993
CFBP 7179^{GSS}	France	2002
CFBP 7244	France	1978
LMG 745	The Netherlands	1978
LMG 746	United Kingdom	1955
LMG 747$^{T;GSS}$	New Zealand	1956
LMG 748	New Zealand	1956
LMG 749	New Zealand	1957
LMG 751$^{GSS;*}$	Romania	1962
LMG 752	United Kingdom	1964
LMG 8047	The Netherlands	1979
CPBF 1480	Azeitão	2014
CPBF 1486	Alcobaça	2014
CPBF 1492	Alcobaça	2014
CPBF 1502	Alcobaça	2014
CPBF 1527	Estremoz	2014
CPBF 100	Ponte da Barca	2015
CPBF 1514	Estremoz	2014
***X. arboricola* strains not belonging to the *juglandis* pv**		
X. arboricola pv. *populi* CFBP 3123^{PRS}	The Netherlands	1979
X. arboricola pv. *pruni* LMG 852^{PRS}	New Zealand	1953
X. arboricola pv. *celebensis* LMG 677^{PRS}	New Zealand	1960
X. arboricola pv.*corylina* LMG 689^{PRS}	United States	1939
X. arboricola pv. *fragariae* LMG 19145^{PRS}	Italy	1993
X. arboricola CPBF 122	Ponte da Barca	2015
X. arboricola CPBF 1530	Estremoz	2014
Other *Xanthomonas* strains		
X. axonopodis pv. *citri* LMG 9322^{T}	United States	1989
X. axonopodis pv. *dieffenbachiae* LMG 695 ^{PRS}	Brazil	1965

Strain[z]	Geographic origin	Year of isolation
X. axonopodis pv. *phaseoli* LMG 7455^{PRS}	United States	1986
X. campestris pv. *campestris* LMG 568	United Kingdom	1957
X. euvesicatoria LMG 922	United States	1939
X. fragariae LMG 708^{T}	United States	1960
X. gardneri LMG 962$^{GSS;T}$	Yugoslavia	1953
X. oryzae pv. *oryzae* LMG 5047^{PRS}	India	1965
X. oryzae pv. *oryzicola* LMG 797^{PRS}	Malaysia	1964
X. perforans NCPPB 4321^{T}	United States	1991
X. translucens pv. *translucens* LMG 876$^{GSS;PRS}$	United States	1933
X. vesicatoria LMG 911$^{GSS;T}$	New Zealand	1955
Xanthomonas sp. CPBF 367	Loures	2016
Xanthomonas sp. CPBF 98	Ponte da Barca	2015
Xanthomonas sp. CPBF 1488	Alcobaça	2014

[z]CFBP :French Collection for Plant-associated Bacteria Institut National de la Recherche Agronomique, Angers, France. LMG: Belgian Coordinated Collections of Microorganisms/LMG Bacteria Collection, Universiteit Gent- Laboratorium voor Microbiologie, Gent Belgium. NCPPB: National Collection of Plant Pathogenic Bacteria, Fera Science Ltd., York, United Kingdom.CPBF: Colecção Portuguesa de Bactérias Fitopatogéncias, Oeiras, Portugal *: other collection no: NCPPB 1447. T, type strain; GSS genome sequenced strain; PRS, pathovar reference strain.

2.2. Walnut fruits inoculation with *Xaj* and bacterial DNA extraction

Immature walnut fruits from a single walnut tree (phenophases Gf + 30 and Gf + 45, Mikulic-Petkovsek et al. 2011) without symptomatic lesions were used in this study. The fruits were disinfected by immersion in 70% ethanol for 30 s, and were thoroughly washed with sterile distilled water. Five groups, of four walnut fruits each were tested. The fruits were inoculated by immersion in an *Xaj* LMG 747^{T} bacterial suspension of approximately 10^{8} CFU/ml up to 30 min at room temperature with slow shaking. As negative controls, the fruits were immersed in sterile distilled water for 30 min, following the same conditions as described above. After inoculation, the fruits were incubated under controlled environmental conditions (16-h/8-h photoperiod and a temperature of 24°C/18°C) in sterile individual plastic containers, to ensure high humidity. After 2 weeks, all fruits inoculated with *Xaj* LMG 747^{T} showed clear disease symptoms. One fruit from each group was then divided into eight identical parts by two longitudinal sections passing through the fruit apical scar (micropile) followed by an equatorial section (**Figure III.1**). For DNA extraction, the outer pericarp layer (external fruit green husk containing the lesions) of each selected fruit fraction was excised with a sterile scalpel, the weight of tissue recovered was registered, and the fraction was then macerated with 5 ml of sterile distilled water in extraction bags (Bioreba AG, Reinach, Switzerland). Then, 2 ml of the extract was centrifuged at 16,000 × g for 10 min to obtain a bacterial pellet used for DNA extraction using the QIAGEN DNeasy Plant Mini kit (Qiagen, Hilden, Germany), following the manufacturer's instructions.

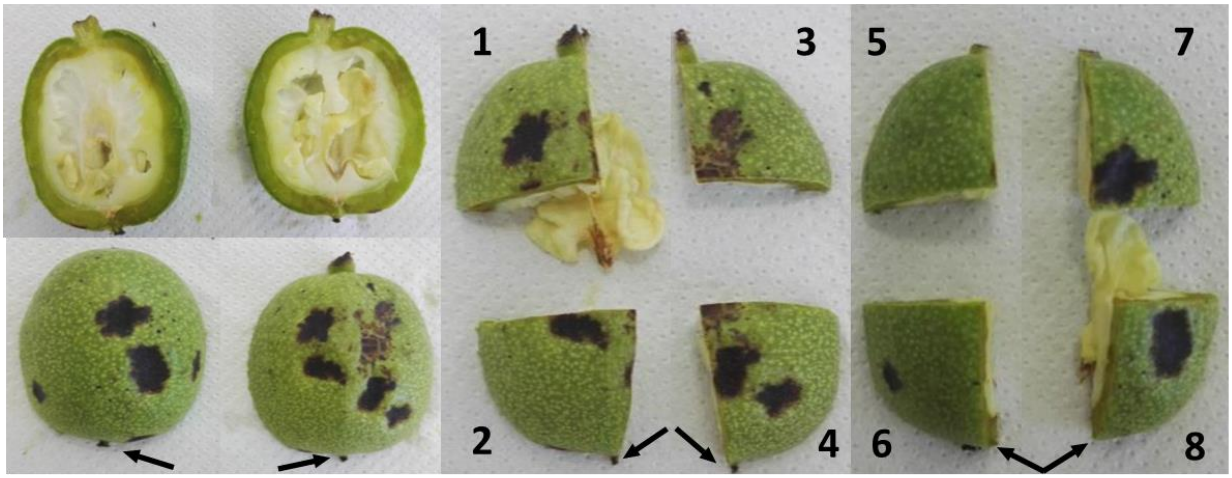

Figure III.1. Walnut fruit symptoms of *Xaj* infection. Each fruit was divided in eight parts (1-8). Arrows indicate the fruit apical scar.

2.3. Selection of *Xaj* specific DNA markers for qPCR

Eighteen fully sequenced genomes of *Xaj* were surveyed for the presence of *Xaj*-specific DNA markers (XAJ1-XAJ9) previously characterized by Fernandes et al. (2017): NCPPB 1447 (NZ_AJTL00000000.1), CFBP 2528 (NZ_JZEF00000000.1), CFBP 7179 (JZEG01000001.1), CFSAN033077 (NZ_LHBK00000000.1), CFSAN033081 (NZ_LHBO00000000.1), CFSAN033086 (NZ_LHBT00000000.1), J303 (NZ_LSGZ00000000.1), 417 (CP012251.1), CFSAN033078 (NZ_LHBL00000000.1), CFSAN033079 (NZ_LHBM00000000.1), CFSAN033080 (NZ_LHBN00000000.1), CFSAN033082 (NZ_LHBP00000000.1), CFSAN033083 (NZ_LHBQ00000000.1), CFSAN033084 (NZ_LHBR00000000.1), CFSAN033085 (NZ_LHBS00000000.1), CFSAN033087 (NZ_LHBU00000000.1), CFSAN033088 (NZ_LHBV00000000.1), CFSAN033089 (NZ_LHBW00000000.1) (Supplementary **Figure III. S1**). The occurrence of these nine *Xaj*-specific markers was further evaluated in 108 *Xaj* isolates (Fernandes et al. 2018) (Supplementary **Figure III. S2**). The matrices assembled based on the data gathered showed that four markers (XAJ1, XAJ4, XAJ6 and XAJ8) were present in most of the *Xaj* strains, but only markers XAJ1 and XAJ6 were present in all *Xaj* analysed and therefore were chosen as the most promising for qPCR.

2.4. PCR-based confirmation of walnut fruit's infection by *Xaj*

Infection of fruits by *Xaj* LMG 747^T was confirmed by PCR using the primer pairs XAJ1F/XAJ1R and XAJ6F/XAJ6R to target XAJ1 and XAJ6 DNA markers as previously described (Fernandes et al. 2017) (**Table III.2**). PCR reactions were carried out in 20 µl reactions containing 1X DreamTaq Buffer (Thermo Scientific, Vilnius, Lithuania), 0.2 mM of each dNTP

(Thermo Scientific), 0.2 µM of each primer (STAB Vida, Lisbon, Portugal), 1U of DreamTaq DNA Polymerase (Thermo Scientific) and 2 µl of DNA extracted from infected fruit fractions. PCR was performed with an initial 5 min denaturation at 95 °C, followed by 35 cycles of 30 s at 95 °C, 30 s at 61 °C, and 30 s extension at 72 °C, with a final extension step of 10 min at 72 °C. PCR products were separated on 0.8% agarose gels stained with GreenSafe Premium (Nzytech, Lisbon, Portugal).

Table III.2. *Xanthomonas arboricola* pv. *juglandis* specific markers used in this study[Z].

DNA marker	Primer	Sequence (5'-3')	Amplicon length (bp)	Source
XAJ1	XAJ1qPCR_Fwd	AGACGCGTCACTTCGGAAAGAG	100	This study
	XAJ1qPCR_Rev	CCAGCCTGTTCTAGGTGGTTTG		
	XAJ1F	GTTGTCAAGGTCGCTACTGGACGC	758	Fernandes et al. 2017
	XAJ1R	CGACAGGTACCGAAAGTGGACG		
XAJ6	XAJ6qPCR_Fwd	AGAGATTGATGAGCGATGCGAG	101	This study
	XAJ6qPCR_Rev	GCACAGCGGGAAGTAATAGCAA		
	XAJ6F	AAGTCAGATGCGAAGCGAAAGG	436	Fernandes et al. 2017
	XAJ6R	GCACAGCGGGAAGTAATAGCAAAC		

[Z]F = forward and R = Reverse

2.5. qPCR calibration curves with XAJ1 and XAJ6 markers

Two new primer pairs (XAJ1qPCR_Fwd/XAJ1qPCR_Rev, and XAJ6qPCR_Fwd/ XAJ6qPCR_Rev) targeting a smaller fragment of *Xaj*-specific markers XAJ1 (100 bp) and XAJ6 (101 bp) were designed for qPCR amplification, (**Table III.2**), using the Geneious software (version7.1.2; Biomatters, Auckland, New Zealand). Optimal annealing temperatures were determined based on the melting curves using both pairs of primers (Supplementary **Figure III. S3**). qPCR calibration curves were constructed using serial dilutions (1 to 10^{-7}) of *Xaj* LMG 747^T DNA by plotting the quantification cycle (Cq) versus the number of DNA marker copies. Each dilution was assessed in triplicate and the number of *Xaj* chromosomes per microliter was calculated using the Science Primer copy number calculator for real-time PCR (http://scienceprimer.com). Each real-time PCR reaction contained 1x Sso Advanced Universal SYBR Green Supermix (Bio-Rad, Hercules, CA), 0.35 µM of each primer, 1µl of each DNA template dilution, and nuclease-free water to a total volume of 20µl. Reactions were carried out in hard-shell PCR plates (96 well, white shell/clear well), sealed with Microseal B Adhesive Sealer (Bio-Rad). qPCR cycle conditions were as follows: initial denaturation step of 98 °C for 3 min, 39 cycles of 98 °C for 10 s and 63 °C for 30 s, and a final melting curve analysis from 65 °C to 95

°C, with 0.5 °C increments for 5 s per step. A nontemplate control was included in each run. Reactions were carried out in a CFX Connect™ Real-Time Detection System (Bio-Rad) and results were analyzed using with the CFX Manager™ Software (version 3.1; Bio-Rad).

2.6. Validation of qPCR with spiked walnut leaf samples

Spiked walnut leaf samples were used for qPCR validation and inferred the presence of possible PCR inhibitors in plant extracts. *Xaj* LMG 747^T cells were grown in nutrient agar (NA) medium (Difco Laboratories, Detroit, MI) to an optical density of 0.135 at λ = 600, corresponding to approximately 10^8 CFU/ml as estimated by plating serial dilutions of the bacterial culture in NA at 28 °C. Spiked walnut leaf samples were prepared by mixing *Xaj* LMG 747^T cells recovered from 2 ml of each dilution (10^7 CFU/ml to 10^0 CFU/ml) after a centrifugation at 7,000 × *g* for 5 min with 2 ml of bacteria-free plant extract. DNA from each spiked sample (i.e. sample processing control [SPC]) was extracted using the QIAGEN DNeasy Plant Mini kit as mentioned above and amplified in triplicate to construct a calibration curve for each marker (XAJ1 and XAJ6). The limit of detection (LOD) was determined as the CFU value of the dilution corresponding to the last linear point of the standard curve (Marancik and Wiens 2013). qPCR reactions and data analysis were carried out as described in the previous section.

2.7. qPCR repeatability, reproducibility and analytical specificity

qPCR repeatability (intra-assay variance) and reproducibility (interassay variance) were assessed according to Bustin et al. (2009). Intra-assay variation was calculated considering three replicates in a single run/plate and interassay variation was calculated using three independent experiments with three replicates each (**Table III.3** and **Table III.4**). Coefficients of variation (CVs) were calculated according to the formula $CV = standard\ deviation\ Cq\ values/average\ Cq$ (Robène et al. 2015).

Analytical specificity evaluates the detection of target DNA relatively to nontarget DNA sequence (Bustin et al. 2009) and was validated by taking into consideration the Cq value variations and/or differences between target (i.e. *Xanthomonas arboricola* pv. *juglandis*) and nontarget bacteria (i.e. other *X. arboricola* strains not belonging to the *juglandis* pathovar and other *Xanthomonas* species). Accordingly, for this analysis, DNA from reference *Xanthomonas* strains isolated from different geographic regions was used, including 25 *Xanthomonas arboricola* pv. *juglandis* strains, seven *X. arboricola* strains not belonging to the *juglandis* pathovar, and 15 strains of other *Xanthomonas* species (**Table III.5**). DNA extraction and quantification was carried out as described above, and 20 ng of DNA was used for each qPCR

reaction. Melting curves to confirm primer pairs specificity for both markers (XAJ1 and XAJ6) and for each tested strain were determined in duplicate (Supplementary **Figure III. S4**). qPCR reaction efficiency was calculated according to the following curve fitting method (Kralik and Ricchi 2017: $\% \, Efficiency = (10^{(-1/m)} - 1) \times 100$, where *m* is the slope of the linear regression equation of the standard curve.

Table III.3. DNA marker XAJ1 repeatability and reproducibility of qPCR assay[y]

Samples[z]	CFU/µl	Intra-assay CV				Inter-assay CV			
		Nº assays	Mean Cq	SD	% CV	Nº assays	Mean Cq	SD	% CV
SPC0	**$2.70x10^5$**	**3**	**17.41**	**0.10**	**0.58**	**3**	**17.22**	**0.23**	**1.36**
SPC-1	**$2.70x10^4$**	**3**	**20.73**	**0.14**	**0.68**	**3**	**20.50**	**0.27**	**1.32**
SPC-2	**$2.70x10^3$**	**3**	**23.73**	**0.14**	**0.58**	**3**	**23.59**	**0.19**	**0.79**
SPC-3	**$2.70x10^2$**	**3**	**27.36**	**0.17**	**0.60**	**3**	**27.01**	**0.35**	**1.31**
SPC-4	**$2.70x10^1$**	**3**	**30.55**	**0.43**	**1.39**	**3**	**30.35**	**0.55**	**1.80**
SPC-5	**$2.70x10^0$**	**3**	**35.36**	**0.25**	**0.72**	**3**	**34.20**	**1.26**	**3.67**
SPC-6	$2.70x10^{-1}$	3	35.68	n/a	n/a	3	35.48	0.57	1.62
SPC-7	$2.70x10^{-2}$	3	n/a	n/a	n/a	3	35.71	n/a	n/a

[y] CV = coefficient of variation, Cq = quantification cycle, SPC = sample processing control, and n7a = not amplified.
[z] The SPC dilutions within the linear range of the standard curve are highlighted in bold.

Table III.4. DNA marker XAJ6 repeatability and reproducibility of the qPCR assay[y]

Samples[z]	CFU/µl	Intra-assay CV				Inter-assay CV			
		Nº assays	Mean Cq	SD	% CV	Nº assays	Mean Cq	SD	% CV
SPC0	**$2.70x10^5$**	**3**	**16.79**	**0.03**	**0.16**	**3**	**16.85**	**0.26**	**1.54**
SPC-1	**$2.70x10^4$**	**3**	**20.06**	**0.12**	**0.60**	**3**	**20.03**	**0.22**	**1.08**
SPC-2	**$2.70x10^3$**	**3**	**23.31**	**0.10**	**0.43**	**3**	**23.31**	**0.24**	**1.05**
SPC-3	**$2.70x10^2$**	**3**	**26.67**	**0.04**	**0.14**	**3**	**26.66**	**0.14**	**0.54**
SPC-4	**$2.70x10^1$**	**3**	**29.84**	**0.18**	**0.61**	**3**	**29.78**	**0.37**	**1.23**
SPC-5	**$2.70x10^0$**	**3**	**31.66**	**0.14**	**0.43**	**3**	**31.81**	**0.49**	**1.55**
SPC-6	$2.70x10^{-1}$	3	34.89	0.53	1.53	3	33.80	1.02	3.02
SPC-7	$2.70x10^{-2}$	3	35.05	0.23	0.64	3	34.00	1.10	3.22

[y] CV = coefficient of variation, Cq = quantification cycle, SPC = sample processing control, and n7a = not amplified.
[z] The SPC dilutions within the linear range of the standard curve are highlighted in bold.

Table III.5. Quantitative PCR assay specificity for XAJ1 and XAJ6 markers: Comparative analysis of quantification cycle (Cq) values obtained for target groups (*Xanthomonas arboricola* pv. *juglandis* strains), and nontarget groups (other *X. arboricola* strains and other *Xanthomonas* strains)[z]

Strain	XAJ1		XAJ6	
	Cq	Mean ± SD	Cq	Mean ± SD
X. arboricola* pv. *juglandis				
CFBP 176	13.76		13.43	
CFBP 2564	14.06		13.42	
CFBP 2632	13.74		13.46	
CFBP 5252	13.81		13.76	
CFBP 6556	13.68		13.28	
CFBP 6557	13.22		13.09	
CFBP 7072	13.69		13.07	
CFBP 7179	13.66		13.16	
CFBP 7244	13.91		13.56	
CFBP 877	13.92		13.60	
LMG 745	14.47		13.80	
LMG 746	13.91	13.38 ± 1.00	13.70	13.12 ± 0.84
LMG 747	11.02		11.56	
LMG 748	13.22		12.68	
LMG 749	13.67		13.11	
LMG 751	13.36		12.73	
LMG 752	13.20		12.92	
LMG 8047	13.75		13.45	
CPBF 1480	10.17		9.89	
CPBF 1486	14.31		13.8	
CPBF 1492	13.69		13.77	
CPBF 1502	13.33		13.09	
CPBF 1527	13.96		13.51	
CPBF 100	13.43		13.12	
CPBF 1514	11.63		32.98	
***X. arboricola* strains not belonging to the *juglandis* pathovar**				
X. arboricola pv. *populi* CFBP 3123	29.56		29.21	
X. arboricola pv. *pruni* LMG 852	n/a		n/a	
X. arboricola pv. *celebensis* LMG 677	12.22		35.56	
X. arboricola pv. *corylina* LMG 689	19.96		37.69	
X. arboricola pv. *fragariae* LMG 19145	35.75		n/a	
X. arboricola CPBF 122	32.48		31.47	
X. arboricola CPBF 1530	37.63		36.00	
Other *Xanthomonas* species				
X. axonopodis pv. *citri* LMG 9322	35.32		31.97	
X. axonopodis pv. *phaseoli* LMG 7455[PRS]	36.78		39.78	
X. campestris pv. *campestris* LMG 568	33.03		34.41	
X. fragariae LMG 708	37.82		n/a	
X. gardneri LMG 962	34.84		34.84	
X. vesicatoria LMG 911	35.56		n/a	
X. axonopodis pv. *dieffenbachiae* LMG 695	n/a		35.70	
X. euvesicatoria LMG 922	n/a		35.56	
X. oryzae pv. *oryzae* LMG 5047[PRS]	n/a		n/a	
X. oryzae pv. *oryzicola* LMG 797	n/a		35.58	
X. perforans NCPPB 4321	n/a		35.1	
X. translucens pv. *translucens* LMG 876	n/a		36.71	
Xanthomonas sp. CPBF 98	12.94		29.96	
Xanthomonas sp. CPBF 1488	31.04		30.36	
Xanthomonas sp. CPBF 367	37.66		36.30	

[z] CFBP = French Collection for Plant-Associated Bacteria, Institut National de la Recherche Agronomique, Angers, France; LMG = Belgian Coordinated Collections of Microorganisms/LMG Bacteria Collection, Universiteit Gent Laboratorium voor Microbiologie, Gent, Belgium; CPBF = Colecção Portuguesa de Bactérias Fitopatogéncias, Oeiras, Portugal; n/a = not amplified; PRS, pathovar reference strain; and NCPPB = National Collection of Plant Pathogenic Bacteria, Fera Science Ltd., York, United Kingdom.

2.8. qPCR quantification of *Xaj* in infected walnut fruits

To estimate the number of *Xaj* cells by qPCR in infected walnut fruit tissues, standard curves obtained from serial dilutions of SPCs for both XAJ1 and XAJ6 markers were generated as described above.

The qPCR reaction for each *Xaj*-infected one-eighth fruit part contained 1x Sso Advanced Universal SYBR Green Supermix (Bio-Rad), 0.35 µM of each primer, 2µl template DNA, and nuclease-free water to a total volume of 20 µl. qPCR reactions were carried out in triplicate and in parallel with the SPCs standard curves. CFX Manager™ Software (version v3.1) was used to analyzed the data and to calculate the efficiency of the qPCR reaction and the correlation coefficient (R^2) of SPC standard curves. The absolute copy number for each marker in each fruit sample was calculated considering that each marker (XAJ1 and XAJ6) occur as single copy locus in *Xaj* genomes and using the following formula: $Copy\ number = 10^{((Cq-b)/m)}$. This formula is derived from the linear regression equation of the standard curve: $Cq = m(\log\ copy\ number) + b$. Where Cq is the quantification cycle value obtained for the tested sample, *m* is the slope, and *b* is the y-intercept of the linear regression equation of the calibration curve. The results were expressed as the number of *Xaj* chromosome or the cell number (*Xaj* load) per gram of fruit fresh weight for each one-eighth fruit part. *Xaj* bacteria load per fruit was determined as the average *Xaj* load per gram fresh weigh obtained for the eight fruit parts and the corresponding standard deviation.

2.9. Statistical analysis

IBM SPSS Statistics software (version 25) was used to perform all statistical analyses. Two different tests were performed on qPCR data to compare the Cq values obtained: the Student *t* test was carried out to compare Cq values between two markers, whereas one-way analysis of variance was used for the comparison of Cq values between fruits, and the means were compared using Tukey test at $P<0.05$ level.

3. Results

3.1. Selection of *Xaj*-specific DNA markers for qPCR

Nine *Xaj*-specific DNA markers (XAJ1-XAJ9) were previously described and characterized by Fernandes et al. (2017). The occurrence of these nine markers was investigated in 18 fully sequenced genomes of *Xaj* strains isolated from walnut plant organs. XAJ1, XAJ4 and XAJ6

were present in all genomes, whereas XAJ8 was not present in one out of the 18 genomes analyzed. Although the other markers were *Xaj*-specific, they were absent from some *Xaj* strains: marker XAJ9 was present in seven genomes, XAJ7 and XAJ5 in three genomes and XAJ2 in only one genome. These results were further validated in a collection of 108 *Xaj* recently tested for the presence of XAJ markers by dot blot hybridization (Fernandes et al. 2018). The data reveal that XAJ1 was present in 108 isolates (100%), XAJ2 in 19 isolates (18%), XAJ3 in 57 isolates (53%), XAJ4 in 102 isolates (94%), marker XAJ5 in 53 isolates (49%), XAJ6 in 108 isolates (100%), XAJ7 in 34 isolates (31%), XAJ8 in 100 isolates (93%), and XAJ9 in 70 isolates (65%). Altogether, these data allowed to identify markers XAJ1 and XAJ6 as the most reliable markers to detect and identify *Xaj* and were chosen for further analysis.

A BLAST analysis of the XAJ1 and XAJ6 marker sequences, obtained from *Xaj* strain NCPPB 1447 (accession number PRJNA84273), showed that the best nontarget BLAST hit for qPCR marker XAJ1 (100bp) was obtained with *X. campestris* pv. *campestris* strain 17 (e-value = 2e-04, query coverage = 51%). Concerning the XAJ6 qPCR marker (101 bp) the best nontarget blastn hit was obtained with *X. axonopodis* pv. *phaseoli* strain ISO18C8 (e-value = 5.4, query coverage = 14%).

3.2. qPCR calibration curves for markers XAJ1 and XAJ6

After selection of *Xaj*-specific markers XAJ1 and XAJ6 and design of qPCR primers pairs, two qPCR calibration curves for each marker were generated using 10-fold dilutions of *Xaj* LMG 747^T chromosomal DNA and 10-fold dilutions of DNA extracted from walnut samples spiked with *Xaj* LMG 747^T bacteria (SPCs). The qPCR calibration curves obtained with chromosomal DNA revealed an efficiency of 98.7 and 96.4% for XAJ1 and XAJ6, respectively, and an R^2 value of 0.990 for both markers (**Figure III.2**). Taking into account that *in silico* analyses showed that both markers occur as a single copy locus within the chromosome of the sequenced *Xaj*, the standard curves allowed estimating of an LOD of 35 chromosomes/qPCR reaction for both markers. The SPC standard curves obtained from three independent experiments with three replicates each resulted in an average efficiency of 97.6 ±6.4 and 111.8 ±0.5% and R^2 values of 0.999 and 0.995 for XAJ1 and XAJ6, respectively (**Figure III.3**). The LOD obtained for SPC was determined as 2.7 CFU/qPCR reaction for both markers.

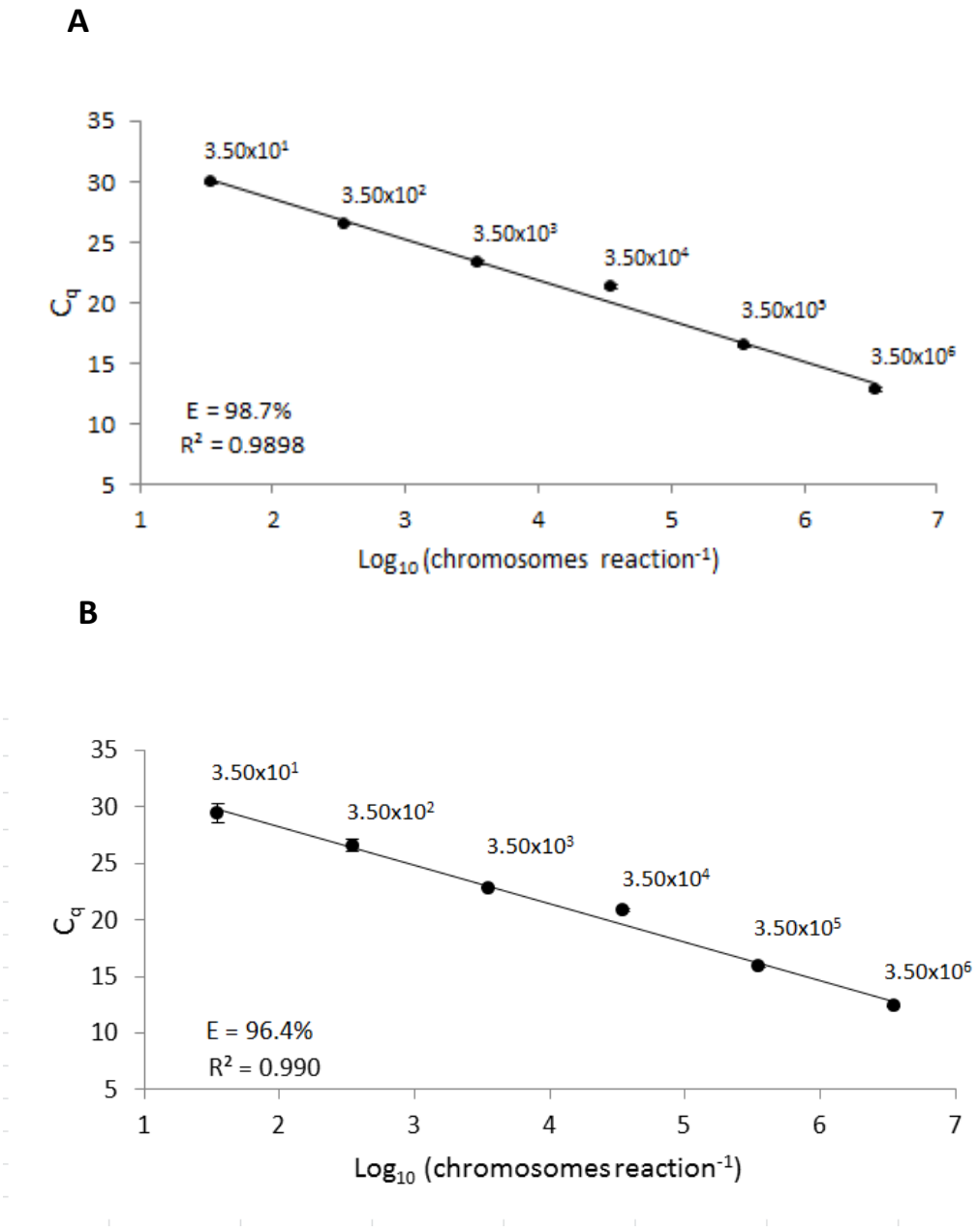

Figure III.2. Calibration curves for **A**, XAJ1 and **B**, XAJ6 DNA markers obtained from serial dilutions of *Xanthomonas arboricola* pv. *juglandis* strain LMG 747 DNA. The Linear portion of the curve range from 3.5 x 10^6 to 35 chromosomes/µL. The efficiency (E) of qPCR assays was 98.7 and 96.4% for XAJ1 and XAJ6, respectively. The correlation efficiency (R^2) was 0.990 for XAJ1 and 0.990 for XAJ6. Standard deviations at each point are represented in the figure. Cq = quantification cycle.

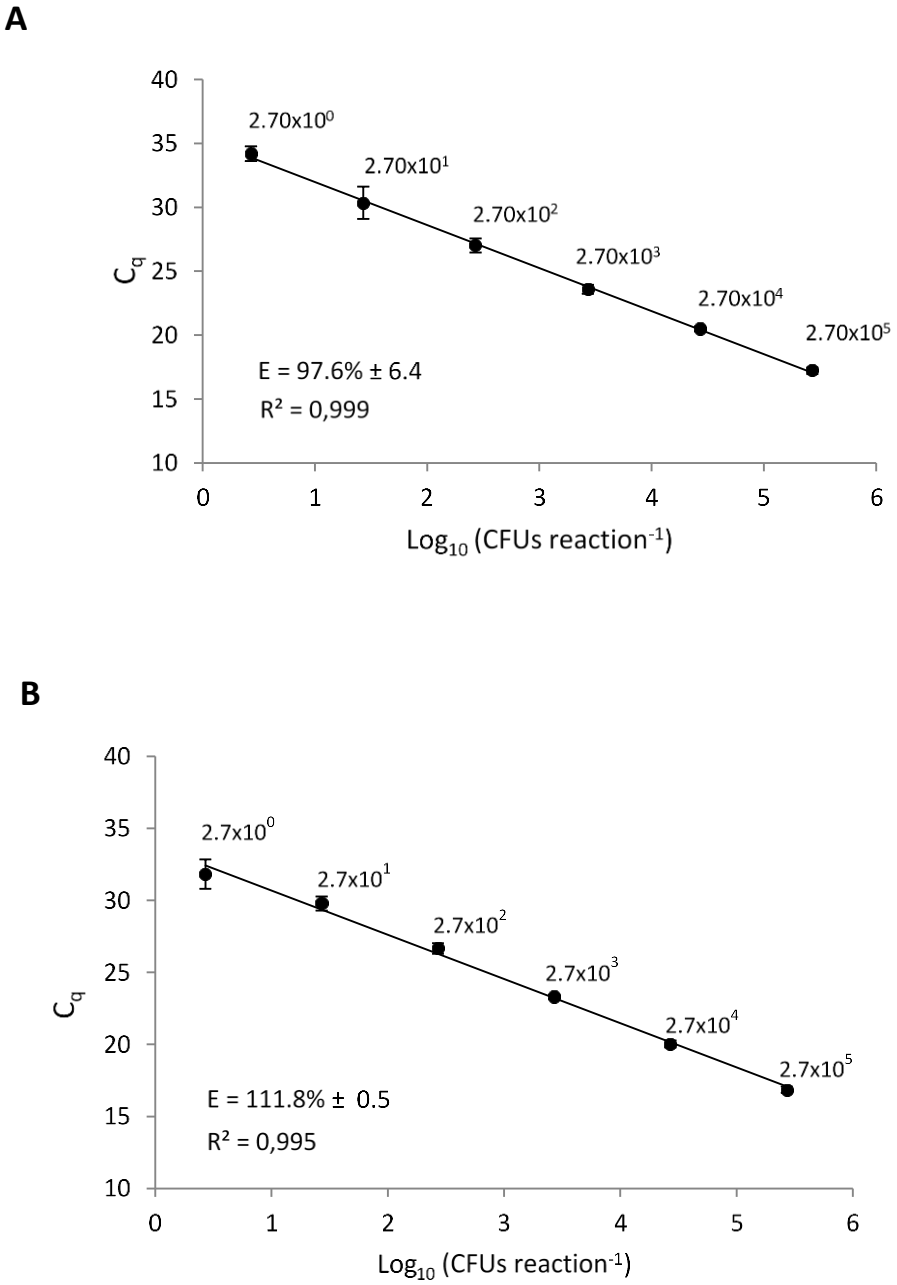

Figure III.3. Standard curves for **A**, XAJ1 and **B**, XAJ6 DNA markers obtained from serial dilutions of sample processing control (SPC) samples. The linear portion of the curve range from SPC-0 (2.7 x 10^5 CFU/µL) to SPC-5 (2.7x10^0 CFU/µL). The efficiency (E) of qPCR assays was 97.6 ± 6.4% for XAJ1 and 111.8 ±0.5% for XAJ6. The correlation efficiency (R^2) was 0.999 for XAJ1 and 0.995 for XAJ6. Standard deviations at each point are represented in the figure. Cq = quantification cycle.

3.3. qPCR repeatability, reproducibility and specificity

Repeatability (i.e., intra-assay variability) of SPCs revealed a CV of Cq values that ranged from a minimum of 0.58% (SPC-0) to a maximum of 1.39% (SPC-4) for the XAJ1 marker and from a minimum of 0.16% (SPC-0) to a maximum of 0.61% (SPC-4) for the XAJ6 marker (**Table III.3** and **Table III.4**). Regarding reproducibility (i.e., interassay variability), the CV values of Cq ranged from a minimum of 0.79 (SPC-2) to a maximum of 3.67% (SPC-5) for XAJ1 and from a minimum of 0.54 (SPC-3) to a maximum of 1.55% (SPC-5) for XAJ6, taking into account three independent experiments (**Table III.3** and **Table III.4**).

Assessment of qPCR specificity for both markers showed that all tested *Xaj* strains gave identical average Cq values for XAJ1 (13.38 ± 1.00) and XAJ6 (13.12 ± 0.84) (**Table III.5**). With only *Xaj* isolate CPBF 1514, the Cq value was 32.98 for XAJ6 but 11.63 for marker XAJ1, which is within the Cq average obtained for this marker with all the other *Xaj* strains tested. On the contrary, the Cq values of *X. arboricola* strains not belonging to the *juglandis* pathovar doubled the Cq values obtained for *Xaj* strains (>29.21) or showed no amplification for both markers, with the exception for *X. arboricola* pv. *celebensis* LMG 677 and *X. arboricola* pv. *corylina* LMG 689, which showed Cq values of 12.22 and 19.96 for XAJ1, respectively. In addition, the Cq values for strains of other *Xanthomonas* species were at least twofold higher (>29.96) than the average Cq values for *X. arboricola* pv. *juglandis* or revealed no amplification for both markers, except for *Xanthomonas* sp. CPBF 98 with a Cq value of 12.94 for XAJ1 (**Table III.5**).

3.4. PCR detection and qPCR quantification of *Xaj* in infected walnut fruits

To confirm walnut fruit infection with *Xaj* and dismiss the occurrence of possible inhibitory PCR compounds, a preliminary detection PCR was carried out using DNA extracted from walnut fruit epicarp tissues. The results showed PCR amplification for both markers (XAJ1 and XAJ6) in at least one of the one-eighth parts of the five tested fruits (fruits A to E) (**Figure III.4**), although a higher PCR efficiency was generally obtained with XAJ1. Interestingly, for some of the one-eighth fruit samples and regardless of the symptomatic evidence of infection, no amplification was observed whatever the marker, as shown with one-eighth fruit parts B2, B6 and B8 from fruit B (**Figure III.4**).

When DNA extracted from each of the one-eighth parts of the five studied walnut fruits was used as a template for qPCR, full detection efficiency was achieved, even with the fruit parts for which standard PCR using the same markers was negative (**Figure III.4**, **Table III.6**). Furthermore, the Cq values obtained for XAJ1 and XAJ6 were consistent between each other and among the different fruit parts, because no statistically significant differences were observed

between the Cq average values obtained for both markers within the same fruit (**Table III.6**). When comparing the average Cq values obtained for both markers and the five fruits studied, only fruit A was shown to be significantly different relative to the Cq obtained for the other fruits (**Table III.6**).

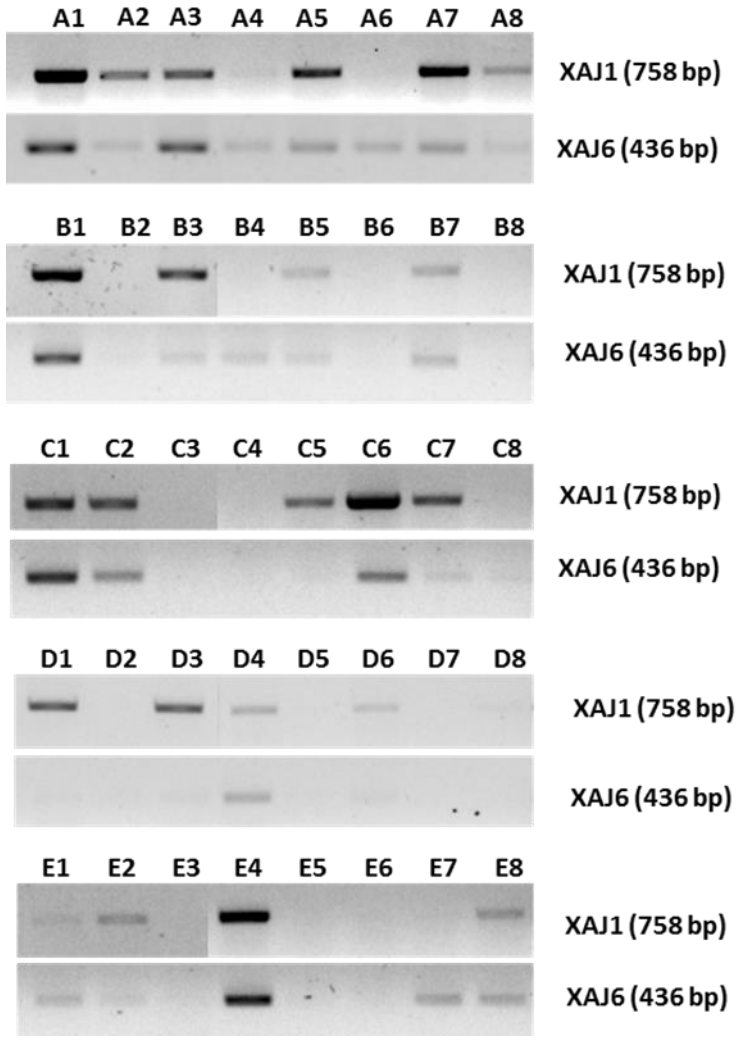

Figure III.4. PCR analysis of infected walnut fruits with DNA markers XAJ1 (758 bp) and XAJ6 (436 bp) for detection of *Xanthomonas arboricola* pv. *juglandis*. bp = bare pair.

Table III.6. Quantification cycle (Cq) values of qPCR quantification of infected walnut fruits with DNA markers XAJ1 and XAJ6[v]

	XAJ1				XAJ6			
	Cq[w]	Mean ± SD[x,y]	Bacterial load ($\times10^3$)[z]	Mean ($\times10^3$)[x]	Cq[w]	Mean ± SD[x,y]	*Xaj* load ($\times10^3$)[z]	Mean ($\times10^3$)[x]
Fruit A								
	24.3		46.10		23.7		36.7	
	28.6		8.90		27.9		5.47	
	25.3		45.0		24.8		33.6	
	28.3	27.01 ± 1.98[a]	9.86	21.1 ± 16.9	27.5	26.36 ± 1.85[a]	6.55	15.2 ± 13.4
	26.4		19.2		25.7		14.0	
	28.2		7.63		27.6		4.76	
	25.2		28.0		24.8		17.9	
	29.8		4.01		28.9		2.52	
Fruit B								
	26.6		21.6		26.8		15.4	
	31.9		2.20		31.7		1.55	
	30.8		4.60		30.7		3.10	
	27.8	30.88 ± 2.66[b]	11.2	6.22 ± 6.98	26.4	30.48 ± 2.6[b]	15.0	5.31 ± 6.13
	31.8		2.99		31.4		2.23	
	33.5		1.60		33.2		1.09	
	30.2		4.29		30.4		2.97	
	34.4		1.28		33.2		1.18	
Fruit C								
	25.8		23.7		26.0		18.4	
	29.0		8.00		29.2		5.32	
	35.3		0.537		35.4		0.265	
	36.1	31.00 ± 3.49[b]	0.553	6.38 ± 8.05	35.2	31.12 ± 3.24[b]	0.416	4.43 ± 6.20
	31.8		1.51		32.4		0.728	
	28.0		11.8		28.5		7.29	
	30.7		2.52		31.1		1.38	
	31.3		2.40		31.2		1.59	
Fruit D								
	31.7		2.57		32.2		1.28	
	32.0		1.73		32.2		0.975	
	30.6		3.44		30.9		1.99	
	29.0	32.31 ± 2.03[b]	5.12	2.20 ± 1.51	29.2	32.51 ± 1.87[b]	3.39	1.27 ± 1.01
	33.9		1.11		34.0		0.567	
	32.2		2.15		32.6		1.09	
	35.5		0.491		35.1		0.340	
	33.6		1.03		33.9		0.502	
Fruit E								
	29.8		4.71		29.7		3.34	
	31.7		2.54		31.3		1.89	
	33.4		1.25		33.2		0.803	
	27.6	31.5 ± 2.49[b]	18.7	5.36 ± 5.78	27.6	31.9 ± 2.21[b]	14.0	3.94 ± 4.33
	34.2		1.58		33.8		1.06	
	34.8		1.57		33.6		1.44	
	29.9		6.28		29.8		4.65	
	30.6		6.21		30.5		4.26	

[v] The *Xanthomonas arboricola* pv. *juglandis* quantity is expressed as chromosomes per gram of fresh weight walnut fruit tissue.
[w] Cq values from qPCR reactions are mean values from samples run in triplicate (technical replicates).
[x] Data are means ± SD of Cq values obtained from all fruit parts for each fruit.
[y] The independent-samples t test was used to determine statistically significant differences between the Cq values obtained with markers XAJ1 and XAJ6 for each
fruit. One-way analysis of variance was applied to ascertain statistically significant differences between Cq values for each fruit for the same marker. Values
followed by different superscript letters indicate statistically significant differences in Cq among different fruits for $P < 0.05$.
[z] Bacterial load of *X. arboricola* pv. *juglandis*.

4. Discussion

Considerable advances were made in the last decade concerning amplification-based methods for detection of phytopathogenic bacteria. Specifically, the improvements associated with real-time PCR and LAMP technologies, the access to extensive genomic data, which is instrumental for identifying specific DNA markers and designing better primers, and the progress obtained with more efficient thermostable DNA polymerases led to major advances in culture-independent methods of phytopathogen detection.

Currently, real-time PCR is acknowledged as cost-effective, high throughput, and capable of providing fast and accurate diagnostics of a broad range of diseases caused by bacteria, as reviewed by Mirmajlessi et al. (2015). Because of its efficiency and specificity, real-time PCR is presently considered the “gold standard” method for phytopathogen detection (Alemu 2014). In fact, qPCR has been used for a panoply of solutions that are beyond detection of phytopathogenic bacteria in plant samples, namely to evaluate the growth of a pathogen in resistance studies with transgenic plants (Reynoird et al. 1999, Hanke et al. 2003), after application of fireblight control agents such as prohexadione calcium (Costa et al. 2001), and in breeding programs (Norelli et al. 2003), and to estimate the extent of contamination with *E. amylovora* on fruits (Van der Zwet and Bell 1990). Despite these advances, not many qPCR protocols have been proposed for quantifying phytopathogenic microorganisms in general and bacteria in particular. A major difficulty in optimizing qPCR protocols to detect and quantify phytopathogenic bacteria in plant samples is mainly attributable to the difficulty in identifying short species-specific DNA signatures suitable for qPCR (i.e., generally <200 bp). This is particularly pertinent for phytopathogens such as *X. arboricola* pv. *juglandis*, which are characterized by high genetic diversity and closely related to nonpathogenic *Xanthomonas* species.

Fernandes et al. (2017) identified and validated nine *Xaj*-specific DNA markers (XAJ1 to XAJ9), using a comprehensive BLAST analysis of all open reading frames (>100 bp) retrieved from the draft genome of *Xaj* (NCPPB 1447, LMG 751), filtered for high and exclusive similarity toward *Xaj*. The occurrence of these DNA markers in a genetically diverse collection of *Xaj* strains, including 18 *Xaj* genomes and 108 field isolates, allowed identification of markers XAJ1 and XAJ6 as being present in all 126 *Xaj* strains analyzed and therefore as promising targets for qPCR. Although the other markers *Xaj* were specific, they were absent from some *Xaj* strains, which could lead to false-negative qPCR results. Interestingly, XAJ1 was not present in the two fully sequenced nonpathogenic strains CFBP 7634 and 7651, whereas XAJ6 was present in the nonpathogenic CFBP 7634 but not in 7651, which raises the question of the importance for pathogenicity and virulence of the genomic regions were these markers are located. Although marker XAJ1 was assigned to an intergenic region, marker XAJ6 is part of a putative coding

region annotated as a hypothetical protein (Fernandes et al. 2017), which currently makes it difficult to relate these markers with genetic determinants of pathogenicity and virulence.

Conventional PCR amplification of markers XAJ1 (758 bp) and XAJ6 (436 bp) from DNA extracted from infected fruit samples did not reveal the presence of PCR inhibitors co-extracted during DNA extraction, confirming results previously reported (Fernandes et al. 2017). Furthermore, although it was not possible to detect the presence of each marker in all one-eighth fruit parts of the five fruits assayed by conventional PCR, higher efficiency was generally observed for XAJ1 compared with XAJ6, as described by Fernandes et al. (2017).

Aiming to optimize a qPCR method to quantify *Xaj* in infected walnut plant samples and to make these markers suitable to ensure a high qPCR efficiency, a set of primer pairs was designed to amplify a smaller region of each marker (namely, 100 bp for XAJ1 and 101 bp for XAJ6). These qPCR targets are within the size range recommended for qPCR (Bustin et al. 2009). The calibration curves obtained resulted in an efficiency > 96% and a R^2 value of 0.99 for both markers, regardless of the DNA template (i.e., *Xaj* chromosomal DNA or DNA extracted from walnut spiked samples), indicating high stability of the qPCR assay (Johnson et al. 2013). The calibration curves allowed us to determine the LOD as 35 chromosomes/qPCR reaction for both markers, using serial dilutions of *Xaj* chromosomal DNA and considering that both markers occur as single copies in the *Xaj* genome, whereas the LOD for spiked samples was as low as 2.7 CFU/qPCR reaction. qPCR repeatability (i.e., intra-assay variation) and reproducibility (i.e., interassay variation) of three independent experiments revealed maximum CV values of 1.39 and 3.67% for intra-assay and interassay variation, respectively, for both markers. These data indicate consistent calibration curves, which is particularly important to accurately determine the bacterial load of environmental samples (Almeida et al. 2018).

Aside from the technical validations addressed above, good qPCR specificity is essential to ensure high confidence, particularly when the microbial target is in a complex matrix such as plant samples, frequently characterized by rich microbiota (Brader et al. 2017). In addition, the high genetic diversity of *X. arboricola*, characterized by several pathovars and the existence of closely related other *Xanthomonas* species, calls for particular attention to achieve a reliable validation of a qPCR procedure for *Xaj*. In this regard, specificity assays were carried out in 25 *Xaj* strains, seven *X. arboricola* strains not belonging to the *juglandis* pathovar, and 15 strains of other *Xanthomonas* species.

Blastn analysis of XAJ qPCR markers showed that the best off-target values were obtained with *X. campestris* pv. *campestris* strains 17 (e-value = 2e-04, query coverage = 51%) for XAJ1, and with *X. axonopodis* pv. *phaseoli* strains ISO18C8 (e-value = 5.4, query coverage = 14%) for XAJ6. Because these bacteria are unlikely to share the same ecological niche as *X. arboricola* pv. *juglandis*, it is not expected that they might lead to false-positive results when assaying

infected walnut plant samples. Importantly, given that the present qPCR method is based on two markers (XAJ1 and XAJ6) and that the off-targets are different for both markers, the risk of a false-positive result is negligible.

Regarding the experimental specificity validation, the Cq values obtained for the target *Xaj* (XAJ1 ~ 13.38 ± 1.00 and XAJ6 ~ 13.12 ± 0.84) contrasted with the lack of amplification or statistically different higher Cq values recorded for nontarget species. Interestingly, two of the tested strains (*Xaj* CPBF 1514 and *Xanthomonas* sp. CPBF 98) showed Cq values for XAJ1 within the average range obtained for other *Xaj* strains (11.63 and 12.94, respectively) but high Cq values for XAJ6 (32.98, and 36.71, respectively), characteristic of off-target xanthomonads. These results confirm the presence of XAJ1 and absence of XAJ6 as shown by dot blot hybridization of these two strains (Fernandes et al. 2017). Multilocus sequence analysis genotyping studies showed that both strains are outside the main clusters of *Xaj* and closer to other *X. arboricola* strains not belonging to the *juglandis* pathovar or atypical and nonpathogenic walnut *X. arboricola* isolates (Cesbron et al. 2015, Essakhi et al. 2015, Fernandes et al. 2018).

When comparing qPCR and PCR results of infected walnut fruit parts, the former method was shown to be more efficient in *Xaj* detection than conventional PCR, because it was able to detect bacteria in fruit samples where no amplification was observed by conventional PCR (**Figure III.4**).

In this work, we validated a qPCR approach using two DNA markers able to determine the bacterial load of *Xaj* in infected plant material. Quantification of *Xaj* infecting host plant tissues within a small time frame will allow investigators to measure the virulence fitness of different strains and will support the adequacy of this method to evaluate the tolerance and/or resistance of distinct walnut genotypes to diseases caused by Xaj, altogether contributing to improved risk assessment analysis. Taking into account the genetic diversity of *Xaj* and its broad distribution, the rapid identification of highly virulent *Xaj* strains will be particularly useful to inform the implementation of suitable phytosanitary containment measures.

Acknowledgments

The authors would like to thank Leonor Cruz for lab support.

References

Albuquerque P, Caridade CM, Rodrigues AS, Marcal AR, Cruz J, Cruz L, Santos CL, Mendes MV, Tavares, F. 2012. Evolutionary and experimental assessment of novel markers for detection of *Xanthomonas euvesicatoria* in plant samples. *PLoS ONE* **7**, e37836.

Alemu C. 2014. Real-Time PCR and its application in plant disease diagnostics. *Advances in Life Science and Technology* **27**, 39-49.

Almeida E, Serra CR, Albuquerque P, Guerreiro I, Teles AO, Enes P, Tavares F, 2018. Multiplex PCR identification and culture-independent quantification of *Bacillus licheniformis* by qPCR using specific DNA markers. *Food Microbiology* **74**, 1-10.

Bandi A, Hevesi M, Szani Z, Magdolna T. 2015. Assessment of bacterial blight tolerance of Persian walnut based on immature nut test. *Notulae Scientia Biologicae* **7**, 253-257.

Berg T, Tesoriero L, Hailstones DL. 2006. A multiplex real-time PCR assay for detection of *Xanthomonas campestris* from brassicas. *Letters in Applied Microbiology* **42**, 624-630.

Bos L, Parlevliet J. 1995. Concepts and terminology on plant/pest relationships: toward consensus in plant pathology and crop protection. *Annual Review of Phytopathology* **33**, 69-102.

Brader G, Compant S, Vescio K, Mitter B, Trognitz F, Ma LJ, Sessitsch A. 2017. Ecology and genomic insights into plant-pathogenic and plant-nonpathogenic endophytes. *Annual Review of Phytopathology* **55**, 61-83.

Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL, Vandesompele J, Wittwer CT. 2009. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clinical Chemistry* **55**, 611-622.

Casadevall A, Pirofski L. 1999. Host-pathogen interactions: redefining the basic concepts of virulence and pathogenicity. *Infection and Immunity* **67**, 3703-3713.

Casadevall A, Pirofski L. 2001. Host-Pathogen interactions: The attributes of virulence. *The Journal of Infectious Diseases* **184**, 337-344.

Cesbron S, Pothier J, Gironde S, Jacques MA, Manceau C. 2014. Development of multilocus variable-number tandem repeat analysis (MLVA) for *Xanthomonas arboricola* pathovars. *Journal of Microbiological Methods* **100**, 84-90.

Cesbron S, Briand M, Essakhi S, Gironde S, Boureau T, Manceau C, Fischer-Le Saux M, Jacques MA. 2015. Comparative genomics of pathogenic and nonpathogenic strains of *Xanthomonas arboricola* unveil molecular and evolutionary events linked to pathoadaptation. *Frontiers in Plant Science* **6**,1126.

Cho MS, Kang MJ, Kim CK, Seol YJ, Hahn JH, Park SC, Hwang DJ, Ahn TY, Park DH, Lim CK. 2011. Sensitive and specific detection of *Xanthomonas oryzae* pv. *oryzae* by real-time bio-PCR using pathovar-specific primers based on an rhs family gene. *Plant Disease* **95**, 589-594.

Costa G, Andreotti C, Bucchi F, Sabatini E, Bazzi C, Malaguti S, Rademacher W. 2001. Prohexadione-Ca (Apogee®): Growth regulation and reduced fire blight incidence in pear. *HortScience* **36**, 931-933.

Essakhi S, Cesbron S, Fischer-Le Saux M, Bonneau S, Jacques MA, Manceau C. 2015. Phylogenetic and variable-number tandem-repeat analyses identify nonpathogenic *Xanthomonas arboricola* lineages lacking the canonical Type III Secretion System. *Applied and Environmental Microbiology* **81**, 5395-5410.

Fernandes C, Albuquerque P, Cruz L, Tavares F. 2018. Genotyping and epidemiological metadata provides new insights into population structure of *Xanthomonas* isolated from walnut trees. bioRxiv:397703.

Fernandes C, Albuquerque P, Sousa R, Cruz L, Tavares F. 2017. Multiple DNA Markers for identification of *Xanthomonas arboricola* pv. *juglandis* isolates and its direct detection in plant samples. *Plant Disease* **101**, 858-865.

Fischer-Le Saux M, Bonneau S, Essakhi S, Manceau C, Jacques MA. 2015. Aggressive emerging pathovars of *Xanthomonas arboricola* represent widespread epidemic clones distinct from poorly pathogenic strains, as revealed by Multilocus Sequence Typing. *Applied and Environmental Microbiology* **81**, 4651-4668.

Frutos D. 2010. Bacterial diseases of walnut and hazelnut and genetic resources. *Journal of Plant Pathology* **92**, S79-S85.

Frutos D, López G. 2012. Search for *Juglans regia* genotypes resistant/tolerant to *Xanthomonas arboricola* pv. *juglandis* in the framework of Cost action 873. *Journal of Plant Pathology* **94** (1, Supplement), S1.37-S31.46.

Garces FF, Gutierrez A, Hoy JW. 2014. Detection and quantification of *Xanthomonas albilineans* by qPCR and potential characterization of sugarcane resistance to leaf scald. *Plant Disease* **98**, 121-126.

Giovanardi D, Bonneau S, Gironde S, Fischer-Le Saux M, Manceau C, Stefani E, 2016. Morphological and genotypic features of *Xanthomonas arboricola* pv. *juglandis* populations from walnut groves in Romagna region, Italy. *European Journal of Plant Pathology* **145**, 1-16.

Golmohammadi M, Cubero J, Peñalver J, Quesada J, López M, Llop P. 2007. Diagnosis of *Xanthomonas axonopodis* pv. *citri*, causal agent of citrus canker, in commercial fruits by isolation and PCR-based methods. *Journal of Applied Microbiology* **103**, 2309-2315.

Hajri A, Meyer D, Delort F, Guillaumes J, Brin C, Manceau C. 2010. Identification of a genetic lineage within *Xanthomonas arboricola* pv. *juglandis* as the causal agent of vertical oozing canker of Persian (English) walnut in France. *Plant Pathology* **59**,1014-1022.

Hajri A, Pothier JF, Fischer-Le Saux M, Bonneau S, Poussier S, Boureau T, Duffy B, Manceau C. 2012. Type three effector gene distribution and sequence analysis provide new insights into the pathogenicity of plant-pathogenic *Xanthomonas arboricola*. *Applied and Environmental Microbiology* **78**, 371-384.

Hanke V, Geider K, Richter K. 2003. Transgenic apple plants expressing viral EPS-depolymerase: evaluation of resistance to the phytopathogenic bacterium *Erwinia amylovora*. In: Plant Biotechnology 2002 and Beyond. Springer, 153-157.

Ivanovic Z, Popovic T, Janse J, Kojic M, Stankovic S, Gavrilovic V, Fira D. 2014. Molecular assessment of genetic diversity of *Xanthomonas arboricola* pv. *juglandis* strains from Serbia by various DNA fingerprinting techniques. *European Journal of Plant Pathology* **141**, 133-145.

Johnson G, Nolan T, Bustin SA. 2013. Real-time quantitative PCR, pathogen detection and MIQE. In: PCR Detection of Microbial Pathogens. Springer, 1-16.

Kralik P, Ricchi M. 2017. A basic guide to real time PCR in microbial diagnostics: Definitions, parameters, and everything. *Frontiers in Microbiology* **8**, 108.

Laala S, Bouznad Z, Manceau C. 2015. Development of a new technique to detect living cells of *Xanthomonas campestris* pv. *campestris* in crucifers seeds: the seed-qPCR. *European Journal of Plant Pathology* **141**, 637-646.

Lamichhane JR. 2014. *Xanthomonas arboricola* diseases of stone fruit, almond, and walnut trees: Progress toward understanding and management. *Plant Disease* **98**, 1600-1610.

Lievens B, Brouwer M, Vanachter AC, Cammue BP, Thomma BP. 2006. Real-time PCR for detection and quantification of fungal and oomycete tomato pathogens in plant and soil samples. *Plant Science* **171**, 155-165.

Marancik DP, Wiens GD. 2013. A real-time polymerase chain reaction assay for identification and quantification of *Flavobacterium psychrophilum* and application to disease resistance studies in selectively bred rainbow trout *Oncorhynchus mykiss*. *FEMS microbiology letters* **339**, 122-129.

Mikulic-Petkovsek M, Slatnar A, Veberic R, Stampar F, Solar A. 2011. Phenolic response in green walnut husk after the infection with bacteria Xanthomonas arboricola pv. juglandis. *Physiological and Molecular Plant Pathology* **76**, 159-165.

Mirmajlessi SM, Destefanis M, Gottsberger RA, Mänd M, Loit E. 2015. PCR-based specific techniques used for detecting the most important pathogens on strawberry: a systematic review. *Systematic reviews* **4**, 9.

Moragrega C, Matias J, Aleta N, Montesinos E, Rovira M. 2011. Apical necrosis and premature drop of Persian (English) walnut fruit caused by *Xanthomonas arboricola* pv. *juglandis*. *Plant Disease* **95**, 1565-1570.

Norelli JL, Jones AL, Aldwinckle HS. 2003. Fire blight management in the twenty-first century: using new technologies that enhance host resistance in apple. *Plant Disease* **87**, 756-765.

Palacio-Bielsa A, Cubero J, Cambra MA, Collados R, Berruete IM, López MM. 2011. Development of an efficient real-time quantitative PCR protocol for detection of *Xanthomonas arboricola* pv. *pruni* in *Prunus* species. *Applied and Environmental Microbiology* **77**, 89-97.

Palacio-Bielsa A, Lopez-Soriano P, Buhlmann A, van Doom J, Pham K, Cambra MA, Berruete IM, Pothier JF, Duffy B, Olmos A, Lopez MM. 2015. Evaluation of a real-time PCR and a loop-mediated isothermal amplification for detection of *Xanthomonas arboricola* pv. *pruni* in plant tissue samples. *Journal of Microbiological Methods* **112**, 36-39.

Pérez LM, Fittipaldi M, Adrados B, Morató J, Codony F. 2013. Error estimation in environmental DNA targets quantification due to PCR efficiencies differences between real samples and standards. *Folia Microbiologica* **58**, 657-662.

Reynoird J, Mourgues F, Norelli J, Aldwinckle H, Brisset M, Chevreau E. 1999. First evidence for improved resistance to fire blight in transgenic pear expressing the attacin E gene from *Hyalophora cecropia*. *Plant Science* **149**, 23-31.

Robène I, Perret M, Jouen E, Escalon A, Maillot MV, Chabirand A, Moreau A, Laurent A, Chiroleu F, Pruvost O. 2015. Development and validation of a real-time quantitative PCR assay to detect *Xanthomonas axonopodis* pv. *allii* from onion seed. *Journal of Microbiological Methods* **114**, 78-86.

Shapiro-Ilan DI, Fuxa JR, Lacey LA, Onstad DW, Kaya HK. 2005. Definitions of pathogenicity and virulence in invertebrate pathology. *Journal of Invertebrate Pathology* **88**, 1-7.

Surico G. 2013. The concepts of plant pathogenicity, virulence/avirulence and effector proteins by a teacher of plant pathology. *Phytopathologia Mediterranea*, 399-417.

Thomas SR, Elkinton JS. 2004. Pathogenicity and virulence. *Journal of Invertebrate Pathology* **85**,146-151.

Van der Zwet T, Bell R. 1990. Fire blight susceptibility in *Pyrus germplasm* from Eastern Europe. *HortScience* **25**, 566-568.

Vandroemme J, Baeyen S, Van Vaerenbergh J, De Vos P, Maes M. 2008. Sensitive real-time PCR detection of *Xanthomonas fragariae* in strawberry plants. *Plant Pathology* **57**, 438-444.

Weller S, Beresford-Jones N, Hall J, Thwaites R, Parkinson N, Elphinstone J. 2007. Detection of *Xanthomonas fragariae* and presumptive detection of *Xanthomonas arboricola* pv. *fragariae*, from strawberry leaves, by real-time PCR. *Journal of Microbiological Methods* **70**, 379-383.

Supplementary material

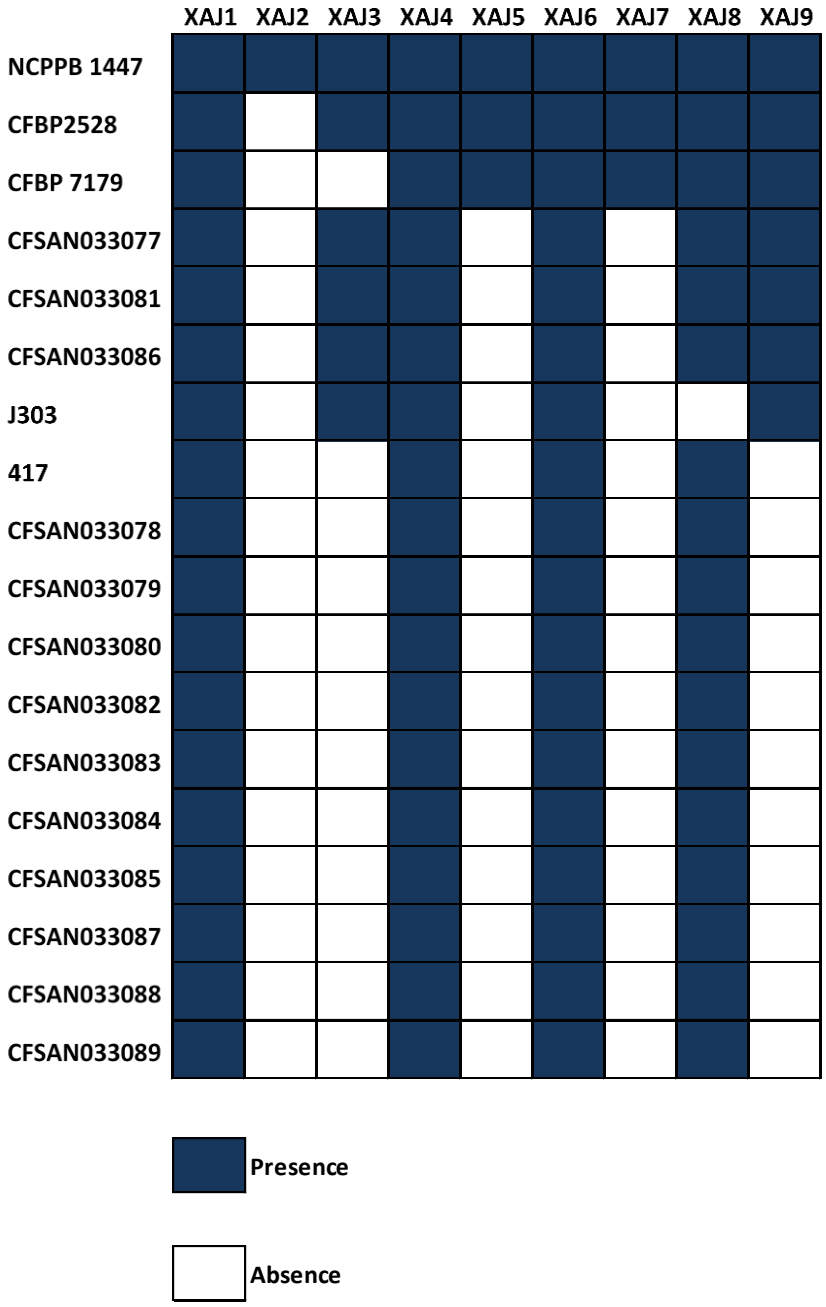

Figure III.S1. Distribution of XAJ markers throughout 18 sequenced *Xaj* genomes.

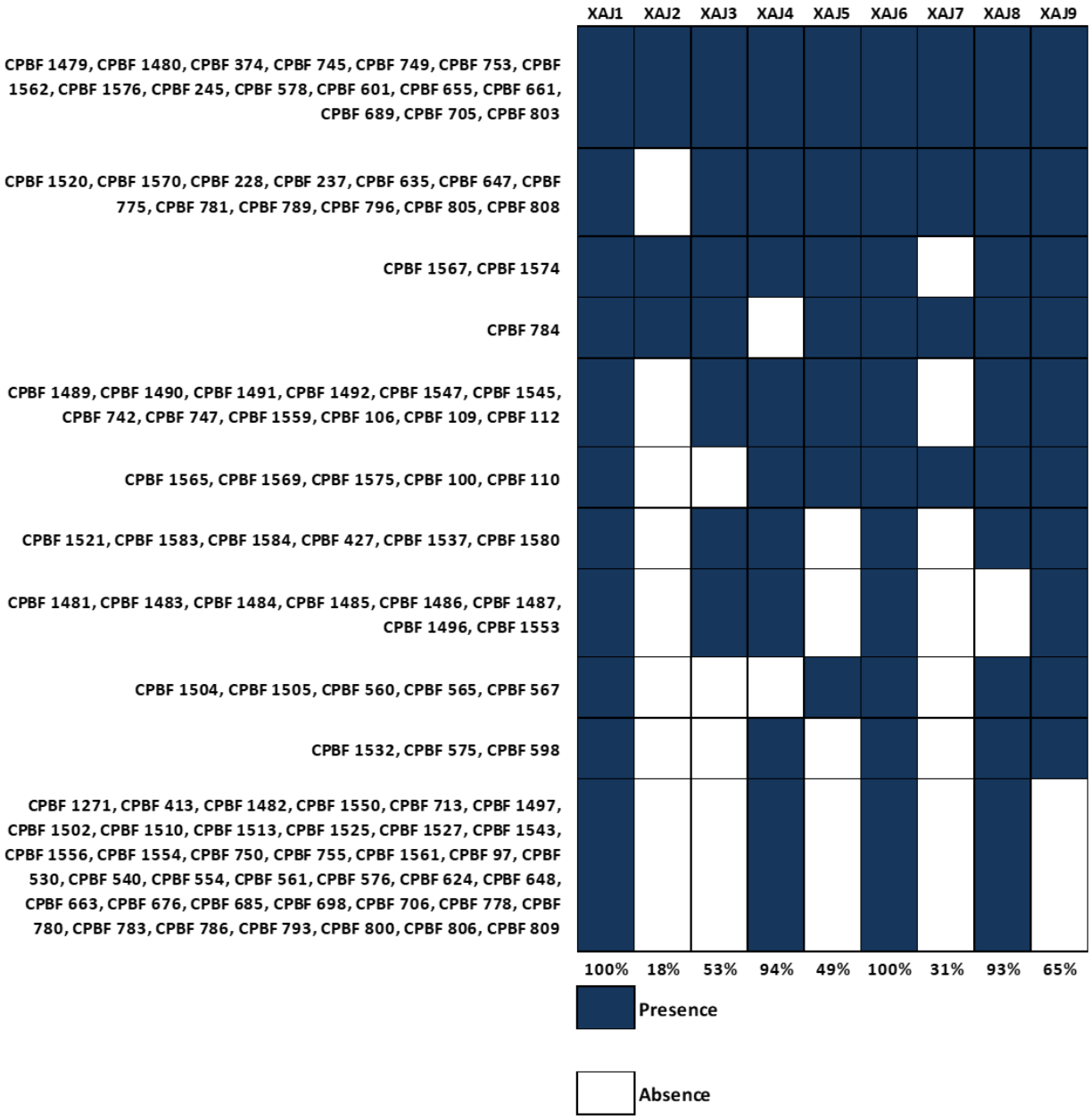

Figure III.S2. Distribution of XAJ markers present in 108 CPBF *Xaj* isolates.

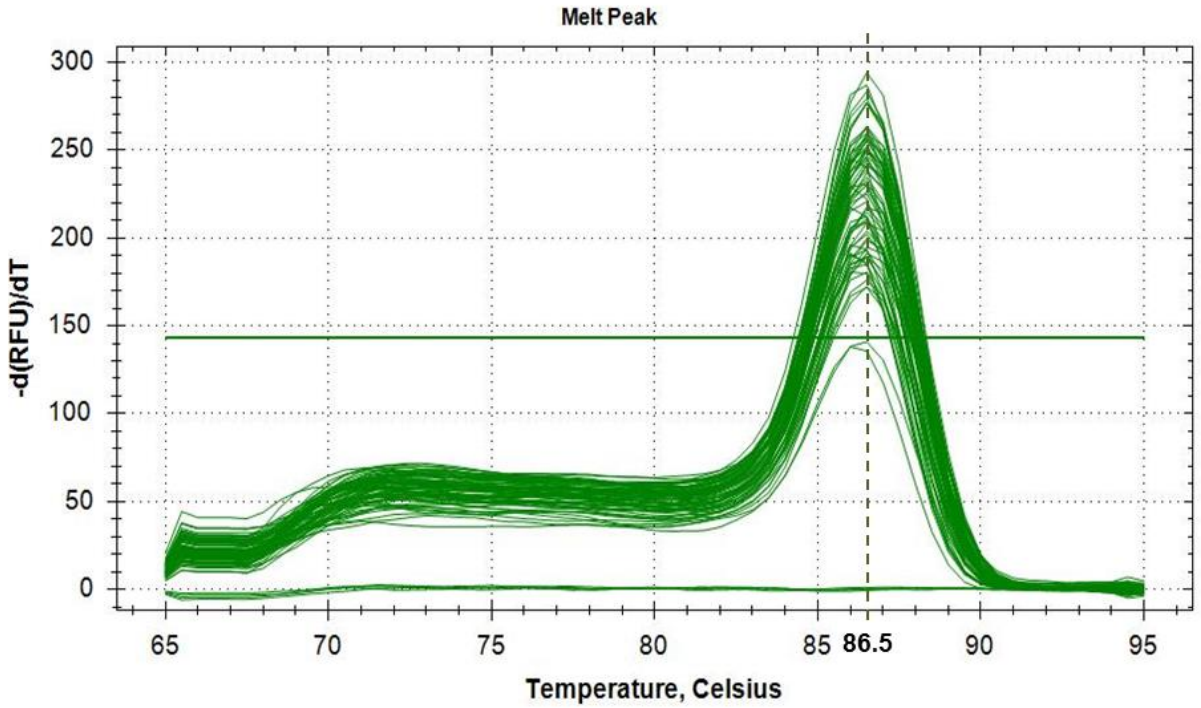

Figure III. S3. Melting curves to determine the optimal qPCR annealing temperature (Ta) using XAJ1qPCR and XAJ6qPCR primers. A gradient qPCR was performed with six 10-fold DNA dilutions of *Xaj* LMG 751. The melting temperature (Tm) for curves was 86.5 °C.

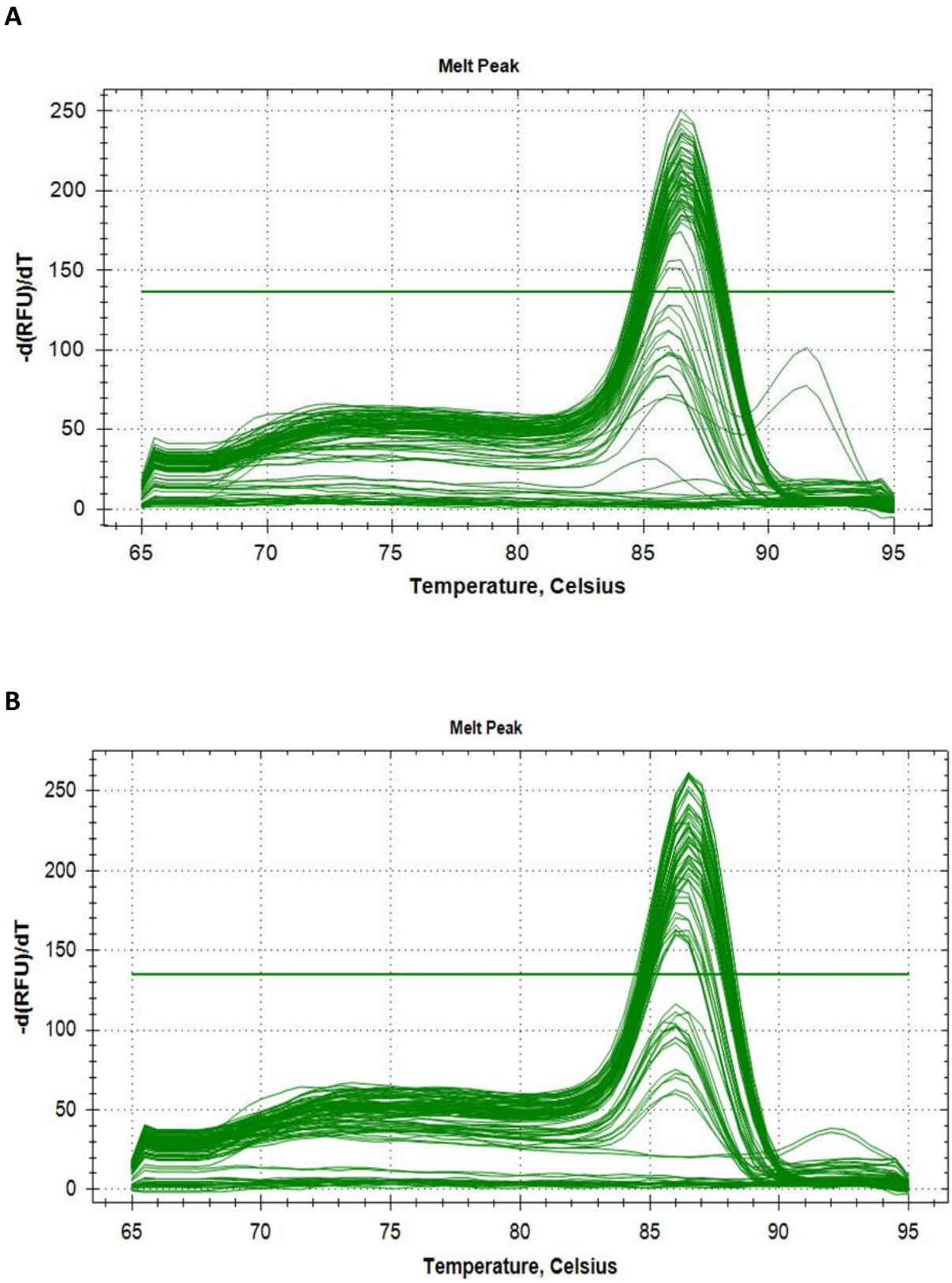

Figure III. S4. Melting curves for specificity assay for XAJ1qPCR (A) and XAJ6qPCR (B) primers with bacterial strains indicated in Table III. 5.

CHAPTER IV

Comprehensive diversity assessment of walnut associated *Xanthomonas arboricola* reveal distinct lineages within the same walnut tree and unveil a new *Xanthomonas* sp.

Chapter IV included the following publication:

Fernandes C, Albuquerque P, Cruz L, Tavares F. 2018. Comprehensive diversity assessment of walnut associated *Xanthomonas arboricola* reveal distinct lineages within the same walnut tree and unveil a new *Xanthomonas* sp.. Submitted.

Part of the results from this chapter were presented at:

Albuquerque P, Fernandes C, Cruz L, Tavares F. 2017 Diversity of *Xanthomonas arboricola* pv. *juglandis* in Portugal evoques a cosmopolitan dispersion. 7th Congress of European Microbiologists (FEMS 2017), Valencia, Spain, July 9-13, 2017.

Fernandes C, Albuquerque P, Sousa R, Santos L, Godinho-Ferreira P, Cruz L, Tavares F. 2017. Estudo da diversidade de *Xanthomonas arboricola* pv. *juglandis* em Portugal revela um padrão de distribuição cosmopolita. Proteção das Plantas 2017: 2º Simpósio SCAP, 8º Congresso da SPF e 11º Encontro Nacional de Proteção Integrada, Escola Superior Agrária de Santarém, Santarém, Portugal, October 26-27, 2017.

Fernandes C, Albuquerque P, Cruz L, Tavares F. *Xanthomonas arboricola* pv. *juglandis*: an endemic walnut pathogen in Portugal. Microbiotec'17, Universidade Católica Portuguesa, Porto, Portugal, December 7-9, 2017.

Fernandes C, Albuquerque P, Cruz L, Tavares F. 2017. *Xanthomonas arboricola* pv. *juglandis*: an endemic walnut pathogen in Portugal. 1st Annual Conference of the EuroXanth COST Action, Instituto Pedro Nunes, Coimbra, Portugal, December 13-15, 2017.

Comprehensive diversity assessment of walnut associated *Xanthomonas arboricola* reveal distinct lineages within the same walnut tree and unveil a new *Xanthomonas* sp.

Abstract

Xanthomonas arboricola pv. *juglandis* (*Xaj*) is the etiological agent of walnut diseases causing high economic losses on walnut production worldwide. This phytopathogen is widely spread in walnut producing regions and has a considerable genetic diversity. Using a comprehensive sampling methodology to disclose the diversity of walnut colonizing *Xaj* in Portugal this work provides new insights on xanthomonads population diversity on walnut trees. Genetic diversity was assessed by multilocus sequence analysis (MLSA) and dot blot hybridization patterns on 131 *Xanthomonas* isolates obtained from 64 walnut trees considering epidemiological metadata such as year of isolation, distinct bioclimatic regions, production regimes, and host-related features. The results showed that the majority of isolates (112 out of 131) were split in 17 lineages of *Xaj*, while the other nineteen isolates, clustered in four MLSA groups which do not include *Xaj* strains, were represented by three lineages of non-*juglandis X. arboricola*, and 11 lineages of non-*arboricola Xanthomonas* sp., including lineages suggestive of a new *Xanthomonas* species. Furthermore, distinct *Xaj*, *X. arboricola* and *Xanthomonas* sp. were isolated from the same walnut tree sample, suggesting possible genetic admixture within the same walnut host tree. Assessment of type III effector genes and pathogenicity assays revealed that non-*arboricola Xanthomonas* lineage tested were nonpathogenic on walnut, with exception for *Xanthomonas* sp. CPBF 424, which differs from the nonpathogenic and commensal phenotype that have been associated to atypical strains found on walnut trees. Altogether, these findings might contribute to understand the sympatric lifestyle of walnut-associated xanthomonads and address *Xanthomonas* pathoadaptations to walnut.

1. Introduction

Xanthomonas is a gammaproteobacteria genus belonging to Xanthomonaceae family, which includes soil dwelling bacteria and important phytopathogenic species, often composed of several host-specific pathovars (Vauterin et al. 1995, Rademarker et al. 2005). *X. arboricola* pv. *juglandis* (*Xaj*) (Pierce 1901) is the only xanthomonads described as a pathogen on walnut trees, acknowledged as the etiological agent of walnut pathologies known as the "Walnut Bacterial Blight (WBB); the "Brown Apical Necrosis" (BAN); and the "Vertical Oozing Canker" (VOC), all causing considerable economic losses (Smith et al. 1912, Hajri et al. 2010, Moragrega et al. 2011). *Xaj* is frequently isolated from symptomatic walnut leaves and fruits, but also from other walnut plant organs without apparent visible lesions, namely dormant walnut buds and catkins (Mulrean and Schroth 1982, Lindow et al. 2014), and from non-plant materials, such as orchard machinery (Giovanardi et al. 2015). This pathogen has a cosmopolitan distribution and diseased trees are commonly observed throughout walnut-growing regions (Frutos 2010, Lamichhane 2014). Genotyping studies have found a considerable genetic diversity in *Xaj*, which is reportedly high when compared with the diversity described for the other *X. arboricola* pathovars (Loreti et al. 2001, Barionovi and Scortichini 2008, Hajri et al. 2010, Lamichhane 2014, Fischer-Le Saux et al. 2015). This feature, together with evidence of genomic trade-offs within the species (Barionovi and Scortichini 2008, Hajri et al. 2010, 2012, Essakhi et al. 2015, Merda et al. 2016), evokes an opportunistic pathogen, even though evidence for environmental reservoirs of *Xaj* remains poorly characterized (Esterio and Latorre 1982, Pruvost and Gardan 1988, Rudolpht 1993, Lamichhane 2014). These data are further supported by the sympatric occurrence of nonpathogenic strains of *X. arboricola* on walnut hosts (Essakhi et al. 2015, Merda et al. 2016), described as phylogenetically heterogeneous, and grouping separately from the well-defined clusters of pathogenic *Xaj* strains (Essakhi et al. 2015). Assessment of *Xaj* in France (Hajri et al. 2010), Italy (Giovanardi et al. 2015) and Serbia (Ivanovic et al. 2014) showed large genetic diversity among isolates and proposed the existence of diverse *Xaj* populations. Furthermore, the genetic differences found between *Xaj* strains collected from VOC and WBB symptoms, suggests the presence of distinct genetic lineages within *Xaj* populations (Hajri et al. 2010).

Different factors have been hypothesized to influence *Xaj* population diversity, namely geographical location (Loreti et al. 2001, Scortichini et al. 2001); origin of plant propagation material (Hajri et al. 2010, Giovanardi et al. 2015); adaptation to particular environmental conditions (Scortichini et al. 2001, Kaluzna et al. 2014); genome flexibility or pathogen virulence (Scortichini et al. 2001, Ivanovic et al. 2014, Kaluzna et al. 2014); or even selective pressure by the host plant (Hajri et al. 2010, Marcelletti et al. 2010). Regardless their valuable contributions, these studies were either based on a low number of bacterial isolates, often obtained without a

planned sampling strategy, or based on a set of *Xaj* strains from worldwide collections, overlooking important metadata such as date, plant host traits, and climatic features which are essential to determine epidemiological patterns (Parkhill and Wren 2011, McMahon and Denaxas 2016). In fact, to understand *Xaj* population structure, it is of utmost importance to conciliate in a single study comprehensive genotyping analysis of a coherent set of isolates with insightful metadata.

This study aimed to characterize the genetic diversity of xanthomonads isolates obtained from walnut trees in the main production regions of Portugal, over a three-year period. Sampling strategy also considered epidemiological factors that potentially contribute to modulate *Xaj* population, as distinct bioclimatic regions, different walnut cultivars, production regimes (orchards vs. isolated trees) and plant organs. Furthermore, to address xanthomonads diversity within the same walnut specimen, several isolates obtained from the same walnut host tree were analyzed. Multilocus sequence analysis (MLSA) and dot blot hybridization patterns (HP) were used to determine the genetic diversity of isolates. These data were complemented by monitoring the presence/absence of four informative type III effector genes (T3E) observed to differentiate nonpathogenic *X. arboricola* and pathogenic *Xaj* clusters (Essakhi et al. 2015) and by pathogenicity tests of new lineages of non-*arboricola Xanthomonas*. Ultimately, the genotyping and epidemiological metadata displayed in this research may contribute to better elucidate *Xanthomonas* populations within walnut trees.

2. Materials and methods

2.1. Bacterial isolation from different walnut plant organs

A total of 94 plant samples were collected from different organs of 64 symptomatic walnut trees (*Juglans regia*) distributed throughout Portugal (66 leaflets samples, 17 fruiting shoot samples, four secondary branch samples, and seven terminal branches for collection of six bud samples and one catkin sample). Sampling occurred between April and October (Leaflets from June to October; Fruiting shoots from June to September; Secondary branches in June and July; Terminal branches in April and September). Necrotic lesions characteristic of WBB or BAN symptoms were observed in all leaves, fruits and branches sampled, and all buds and catkins samples collected were asymptomatic.

Sampling were done in 14 important Portuguese geographic locations for walnut production, and along four bioclimatic regions, characterized by distinct thermoclimatic parameters (Rivas-Martínez et al. 2004): mesotemperature - Mt (Ponte da Barca and Ponte de Lima locations), mesomediterranean - Mm (Carrazeda de Ansiães, Baião, Guarda and Seia locations),

supramediterranean - Sm (Alcobaça, Beja, Bombarral, Estremoz, Leiria and Loures locations) and thermomediterranean - Tm (Azeitão and Portimão locations). Sampled walnuts were either isolated trees or walnut hosts found in orchards established with different French, American or/and Portuguese walnut cultivars, including F1 hybrids (**Table IV.1**). Apart from one sample collected in 2009, all trees were sampled between 2014 and 2016, with six trees being sampled more than once in different years (**Table IV.1**). Multiple samples of different organs were also collected at the same sampling date from 14 walnut trees (**Table IV.1**).

Sample preparation for bacterial isolation was carried out differently for symptomatic and asymptomatic material: i) for symptomatic leaves, fruits and branches, plant tissues adjacent to necrotic areas were first excised using a sterile scalpel (Fernandes et al. 2017); ii) for asymptomatic buds and catkins, either single terminal buds, axillary bud groups or catkins groups of the same branch were excised also using a sterile scalpel (Mulrean and Schroth 1982; Lindow et al. 2014). Bacterial isolation was carried out as procedure detailed in Fernandes et al. (2017). Briefly, each sample of excised plant material was disinfected by immersion in 70% ethanol followed by washing with sterile distilled water (SDW) and them macerated with SDW in extraction bags. The resulting suspensions, and correspondent dilutions, were streaked on yeast extract dextrose carbonate medium and characteristic mucoid yellow colonies of *Xanthomonas* were selected for growth on nutrient agar (NA) medium. One to three isolates were selected per sample and stored at -80°C at the Portuguese Collection of Phytopathogenic Bacteria (CPBF - Colecção Portuguesa de Bactérias Fitopatogénicas, Oeiras, Portugal).

Table IV.1 Walnut trees included in this study and the correspondent epidemiological information and molecular results obtained.

Tree	Host plant	Geographic areas (bioclimatic regions[a])	Year	Plant samples	CPBF[b] isolates	MLSA	HP	Lineage
Jr#01	*Juglans regia*	Azeitão (Tm)	2009	Leaves	1271	V	HP10	Lin12
			2014	Leaves	1479	I	HP1	Lin01
			2014	Fruits	1480	I	HP1	Lin01
			2016	Buds	374	I	HP1	Lin01
					413	V	HP10	Lin12
Jr#02	*Juglans regia* cv. Amigo	Alcobaça (Sm)	2014	Leaves	1481	VI	HP8	Lin17
				Fruits	1482	V	HP10	Lin12
				Branches	1483	VI	HP8	Lin17
Jr#03	*Juglans regia* cv. Hartley	Alcobaça (Sm)	2014	Leaves	1484	VI	HP8	Lin17
				Fruits	1485	VI	HP8	Lin17
			2015	Leaves	1550	V	HP10	Lin12
Jr#04	*Juglans regia* cv. Lara	Alcobaça (Sm)	2014	Leaves	1486	VI	HP8	Lin17
Jr#05	*Juglans regia* cv. Franquette	Alcobaça (Sm)	2014	Leaves	1487	VI	HP8	Lin17
				Fruits	1496	VI	HP8	Lin17
Jr#06	*Juglans regia* cv. Serr	Alcobaça (Sm)	2014	Leaves	1488	X	HP0	Lin24
			2015	Leaves	1553	VI	HP8	Lin17
			2016	Leaves	713	V	HP10	Lin12
Jr#07	*Juglans regia* cv. Rego	Alcobaça (Sm)	2014	Leaves	1489	III	HP4	Lin09
				Fruits	1490	III	HP4	Lin09
Jr#08	*Juglans regia* cv. Corne	Alcobaça (Sm)	2014	Leaves	1497	IV	HP10	Lin11
				Fruits	1491	III	HP4	Lin09

Tree	Host plant	Geographic areas (bioclimatic regions[a])	Year	Plant samples	CPBF[b] isolates	MLSA	HP	Lineage
Jr#09	*Juglans regia* cv. Amigo	Alcobaça (Sm)	2014	Leaves	1492	III	HP4	Lin09
Jr#10	*Juglans regia* cv. Hartley	Alcobaça (Sm)	2014	Leaves	1502	IV	HP10	Lin11
Jr#11	*Juglans regia*	Seia (Mm)	2014	Leaves	1504	I	HP9	Lin05
				Fruits	1505	I	HP9	Lin05
Jr#12	*Juglans regia*	Beja (Sm)	2014	Leaves	1508	IX	HP11	Lin21
Jr#13	*Juglans regia*	Azeitão (Tm)	2014	Fruits	1510	V	HP10	Lin12
Jr#14	*Juglans regia* cv. Hartley	Estremoz (Sm)	2014	Leaves	1513	V	HP10	Lin12
Jr#15	*Juglans regia* cv. Hartley	Estremoz (Sm)	2014	Leaves	1514	IX	HP11	Lin21
Jr#16	*Juglans regia* cv. Cisco	Estremoz (Sm)	2014	Leaves	1515	VIII	HP16	Lin19
Jr#17	*Juglans regia* cv. Tulane	Estremoz (Sm)	2014	Leaves	1520	I	HP2	Lin03
Jr#18	*Juglans regia*	Loures (Sm)	2014	Leaves	1521	VI	HP5	Lin14
			2015	Leaves	1586	VIII	HP16	Lin20
				Fruits	1583	VI	HP5	Lin14
					1584	VI	HP5	Lin14
			2016	Buds	414	X	HP18	Lin31
					427	VI	HP5	Lin14
				Buds	367	X	HP18	Lin31
				Buds	424	X	HP17	Lin30
				Buds	426	X	HP0	Lin25
Jr#19	*Juglans regia* cv. Howard	Estremoz (Sm)	2014	Leaves	1522	IX	HP11	Lin21
Jr#20	*Juglans regia* cv. Lara	Estremoz (Sm)	2014	Leaves	1525	V	HP10	Lin12
Jr#21	*Juglans regia* cv. Lara	Estremoz (Sm)	2014	Leaves	1527	V	HP10	Lin12
Jr#22	*Juglans regia* cv. Pedro	Estremoz (Sm)	2014	Leaves	1530	VIII	HP16	Lin19
Jr#23	*Juglans regia*	Portimão (Tm)	2015	Fruits	1532	VI	HP6	Lin15
					1535	VI	HP5	Lin14
Jr#24	*Juglans regia*	Portimão (Tm)	2015	Leaves	1537	VI	HP5	Lin14
Jr#25	*Juglans regia* cv. Amigo	Alcobaça (Sm)	2015	Branches	1540	III	HP4	Lin09
				Leaves	1543	V	HP10	Lin12
Jr#26	*Juglans regia* cv. Hartley	Alcobaça (Sm)	2015	Leaves	1547	III	HP4	Lin09
				Fruits	1545	III	HP4	Lin09
			2016	Leaves	742	III	HP4	Lin09
					745	I	HP1	Lin01
				Fruits	747	III	HP4	Lin09
					749	I	HP1	Lin01
Jr#27	*Juglans regia* cv. Hartley	Alcobaça (Sm)	2015	Leaves	1556	V	HP10	Lin12
					1559	III	HP4	Lin09
				Fruits	1554	V	HP10	Lin12
			2016	Leaves	750	V	HP10	Lin12
					753	I	HP1	Lin01
				Fruits	755	IV	HP10	Lin11
Jr#28	*Juglans regia*	Bombarral (Sm)	2015	Leaves	1561	V	HP10	Lin12
					1562	I	HP1	Lin01
Jr#29	*Juglans regia* cv. Chandler	Bombarral (Sm)	2015	Leaves	1565	II	HP3	Lin07
Jr#30	*Juglans regia* cv. Chandler	Bombarral (Sm)	2015	Leaves	1569	II	HP3	Lin07
				Fruits	1567	I	HP13	Lin02
Jr#31	*Juglans regia* cv. Chandler	Bombarral (Sm)	2015	Leaves	1570	I	HP2	Lin03
Jr#32	*Juglans regia* cv. Chandler	Bombarral (Sm)	2015	Leaves	1574	I	HP13	Lin02
					1575	II	HP3	Lin07
Jr#33	*Juglans regia*	Guarda (Mm)	2015	Fruits	1576	I	HP1	Lin01
Jr#34	*Juglans regia*	Loures (Sm)	2015	Leaves	1580	VI	HP5	Lin14
Jr#35	*Juglans regia*	Loures (Sm)	2015	Buds	78	X	HP0	Lin27
				Catkins	75	X	HP0	Lin28
Jr#36	*Juglans regia* cv. Franquette	Ponte da Barca (Mt)	2015	Leaves	100	II	HP3	Lin07
					97	V	HP10	Lin12
					98	IX	HP14	Lin22
Jr#37	*Juglans regia* cv. Franquette	Ponte da Barca (Mt)	2015	Leaves	105	IX	HP11	Lin21
					106	III	HP4	Lin10
					108	IX	HP15	Lin23

Tree	Host plant	Geographic areas (bioclimatic regions[a])	Year	Plant samples	CPBF[b] isolates	MLSA	HP	Lineage
Jr#38	*Juglans regia* cv. Lara	Ponte da Barca (Mt)	2015	Leaves	109	III	HP4	Lin10
					110	II	HP3	Lin07
					112	III	HP4	Lin09
Jr#39	*Juglans regia* cv. Lara	Ponte da Barca (Mt)	2015	Leaves	122	VII	HP0	Lin18
Jr#40	*Juglans regia* cv. Lara	Ponte de Lima (Mt)	2015	Leaves	228	III	HP2	Lin08
Jr#41	*Juglans regia* cv. Lara	Ponte de Lima (Mt)	2015	Leaves	237	III	HP2	Lin08
					239	III	HP4	Lin09
Jr#42	*Juglans regia* cv. Franquette	Ponte de Lima (Mt)	2015	Leaves	245	I	HP1	Lin01
Jr#43	*Juglans regia*	Loures (Sm)	2015	Leaves	268	X	HP17	Lin29
Jr#44	*Juglans regia*	Carrazeda de Ansiães (Mm)	2016	Leaves	530	IV	HP10	Lin11
					540	IV	HP10	Lin11
Jr#45	*Juglans regia*	Carrazeda de Ansiães (Mm)	2016	Branches	554	V	HP10	Lin12
Jr#46	*Juglans regia*	Carrazeda de Ansiães (Mm)	2016	Leaves	560	I	HP9	Lin05
					561	V	HP10	Lin12
Jr#47	*Juglans regia*	Carrazeda de Ansiães (Mm)	2016	Leaves	565	I	HP9	Lin05
				Branches	567	I	HP9	Lin05
Jr#48	*Juglans regia*	Carrazeda de Ansiães (Mm)	2016	Leaves	575	VI	HP6	Lin16
Jr#49	*Juglans regia*	Carrazeda de Ansiães (Mm)	2016	Leaves	576	V	HP10	Lin12
					578	I	HP1	Lin01
Jr#50	*Juglans regia*	Carrazeda de Ansiães (Mm)	2016	Leaves	598	V	HP6	Lin13
					601	I	HP1	Lin01
					606	X	HP0	Lin26
Jr#51	*Juglans regia*	Carrazeda de Ansiães (Mm)	2016	Leaves	624	V	HP10	Lin12
Jr#52	*Juglans regia*	Carrazeda de Ansiães (Mm)	2016	Leaves	635	I	HP2	Lin03
					647	I	HP2	Lin03
					648	V	HP10	Lin12
					655	I	HP1	Lin01
Jr#53	*Juglans regia*	Carrazeda de Ansiães (Mm)	2016	Leaves	661	I	HP1	Lin01
					663	V	HP10	Lin12
					676	V	HP10	Lin12
Jr#54	*Juglans regia*	Baião (Mm)	2016	Leaves	685	V	HP10	Lin12
					689	I	HP1	Lin01
Jr#55	*Juglans regia*	Leiria (Sm)	2016	Leaves	698	V	HP10	Lin12
Jr#56	*Juglans regia*	Leiria (Sm)	2016	Leaves	705	I	HP1	Lin01
					706	V	HP10	Lin12
				Fruits	710	I	HP1	Lin01
Jr#57	*Juglans regia* F[1] hybrid	Alcobaça (Sm)	2016	Leaves	775	II	HP2	Lin06
					778	IV	HP10	Lin11
Jr#58	*Juglans regia* F[1] hybrid	Alcobaça (Sm)	2016	Leaves	780	IV	HP10	Lin11
					781	II	HP2	Lin06
Jr#59	*Juglans regia* F[1] hybrid	Alcobaça (Sm)	2016	Leaves	783	IV	HP10	Lin11
					784	I	HP19	Lin04
Jr#60	*Juglans regia* F[1] hybrid	Alcobaça (Sm)	2016	Leaves	786	IV	HP10	Lin11
					789	II	HP2	Lin06
Jr#61	*Juglans regia* F[1] hybrid	Alcobaça (Sm)	2016	Leaves	793	IV	HP10	Lin11
					796	II	HP2	Lin06
Jr#62	*Juglans regia cv. Sunland*	Alcobaça (Sm)	2016	Leaves	800	IV	HP10	Lin11
					803	I	HP1	Lin01
Jr#63	*Juglans regia cv. Refoios 1*	Alcobaça (Sm)	2016	Leaves	805	II	HP2	Lin06
					806	IV	HP10	Lin11
Jr#64	*Juglans regia cv. Refoios 2*	Alcobaça (Sm)	2016	Leaves	808	II	HP2	Lin06
					809	IV	HP10	Lin11

[a] Mt – mesotemperature thermoclimatic region; Mm – mesomediterranean thermoclimatic region; Sm – supramediterranean thermoclimatic region; Tm – thermomediterranean thermoclimatic region (Rivas-Martínez et al. 2004). [b] CPBF, Colecção Portuguesa de Bactérias Fitopatogénicas, Instituto Nacional de Investigação Agrária e Veterinária, Oeiras, Portugal.
F[1] hybrids between two walnut cultivars. Walnut American cultivars: Amigo, Chandler, Cisco, Hartley, Howard, Pedro, Serr, Sunland, Tulane; Walnut French cultivars: Corne, Franquette, Lara; Portuguese cultivars: Refoios 1, Refoios 2, Rego.
The coloured boxes represents the *Xanthomonas* distinct lineages isolated from the same sample of walnut organs (leaves, fruits, and buds).

2.2. Growth conditions of bacterial pure cultures and DNA extraction

The whole set of xanthomonads walnut isolates (**Table IV.1**) and *X. arboricola* strains used in this work were cultured at 28°C on YGC medium (5 g liter^{-1} yeast extract, 10 g liter^{-1} glucose, 30 g liter^{-1} $CaCO_3$, 15 g liter^{-1} agar). Genomic DNA from pure bacterial cultures was extracted using the EZNA Bacterial DNA Purification kit (Omega Bio-Tek, Norcross, GA), following the manufacturer's instructions, and quantified using the Qubit 2.0 Fluorometer HS Assay (Invitrogen, Carlsbad, CA).

2.3. Multilocus sequence analysis and dot blot hybridization

MLSA was carried out using the concatenated sequences of four housekeeping gene fragments: *acnB* (aconitase), *fyuA* (tonB-dependent receptor), *gyrB* (DNA gyrase subunit B) and *rpoD* (RNA polymerase sigma factor). Primer pairs used for *acnB* amplification (684 bp) were described by Parkinson et al. (2007) and for *fyuA* (724 bp), *gyrB* (904 bp) and *rpoD* (915 bp) by Young et al. (2008). The PCR reaction mixture (total volume of 40 µL) contained 1X DreamTaq Buffer with 2.0 mM $MgCl_2$ (Fermentas, Ontario, Canada), 0.2 mM of each dNTP (Fermentas), 0.2 µM of each forward and reverse primers (STAB Vida, Caparica, Portugal), 1U of DreamTaq DNA Polymerase (Fermentas) and 25 ng of bacterial DNA. PCR conditions for the four genes, were 95°C for 5 min, followed by 30 cycles of 95°C for 30 s, 54°C for 30 s, 72°C for 45 s, and a final extension step of 72°C for 10 min. PCR products were purified using the illustra GFX GEL Band Purification kit (GE Healthcare, Buckingham-shire, United Kingdom), according to the manufacturer's instructions, and sequenced on both strands (STAB Vida). Consensus nucleotide sequences obtained for each gene fragments were aligned, trimmed and concatenated using the Geneious v. 9.1.7 software (Biomatters, Auckland, New Zealand). The 131 concatenated sequences (513 bp of *acnB*, 640 bp of *fyuA*, 828 bp of *gyrB* and 793 bp of *rpoD*) obtained from the Portuguese isolates were used to build a Maximum Likelihood tree based on the General Time Reversible (GTR+G+I) model in MEGA 7.0 (Kumar et al. 2016). To account for the known *Xanthomonas* genomic diversity, 32 additional *X. arboricola* strains were included in the analysis together with 89 *Xanthomonas* spp. strains, for which all *acnB*, *fyuA*, *gyrB* and *rpoD* sequences were available in the GenBank database.

Dot blot assays were performed as described in a previous study using nine *Xaj* specific DNA markers (XAJ1 to XAJ9) (Fernandes et al. 2017). Briefly, 100 ng of bacterial DNA were bound to a nylon hybridization transfer membrane (Roche Diagnostics GmbH, Basel, Switzerland) using a Bio-Dot microfiltration unit (Bio-Rad, Hercules, CA). Hybridization was carried out overnight at

68°C to ensure high stringency, with each of the nine DIG-labelled probes (XAJ1 to XAJ9). Dot blot images were acquired using a Molecular Imager ChemiDoc system (Bio-Rad, Hercules, CA).

2.4. Type three effector genes assessed by dot blot assays

The presence of four type three effector genes (T3E), *xopR*, *avrBs2*, *xopF1* and *xopN*, was assessed by dot blot hybridization assays. Genes *xopR*, *avrbs2*, *xopF1* have been considered to be ubiquitous T3E in strains of *X. arboricola*, whereas *xopN* has been suggested to be normally associated with *X. arboricola* strains from pathovars *juglandis*, *pruni* and *corylina* (Hajri et al. 2012). Moreover, the distribution of these T3Es genes have been referenced to differentiate pathogenic from nonpathogenic strains of *X. arboricola* isolated from walnut trees (Essakhi et al. 2015). PCR primers used for preparation of T3E DNA probes were previously described (Hajri et al. 2012). Partial sequences of the four T3E genes were obtained for *Xaj* strain LMG 751 using the PCR reaction conditions described above. PCR amplifications were performed with one cycle of 5 min at 95°C, followed by 35 cycles of 35 s at 95°C, 60 s at 60°C, 60 s at 72°C and a final step of 10 min at 72°C. Each DNA amplicon obtained (303 bp of *xopR*, 850 bp of *avrBs2*, 779 bp of *xopF1* and 864 bp of *xopN)* was purified with the illustra GFX GEL Band Purification kit, and sequenced (STAB Vida) to confirm its identity. The DIG-High Prime kit (Roche Diagnostics GmbH, Basel, Switzerland) was used for probe labelling, following the reference protocol available and a final probe concentration of 100 ng/ml was used in dot blot assays performed as described above. In addition to the 35 walnut xanthomonads isolates, one nonpathogenic strain of *X. arboricola* (CFBP 1022) and three *Xaj* reference strains (CFBP 176, LMG 747 and LMG 751) were also included in each dot blot assay.

2.5. Pathogenicity assays

Juglans regia cv. Hartley seedlings were used for determination of pathogenicity of selected *Xanthomonas* sp. isolates (CPBF 75, CPBF 367, CPBF 424, CPBF 1488 and CPBF 1514). After 30 days of cold stratification treatment at 3-5°C to break dormancy, *Juglans regia* seeds were sown in sterilized sand substrate and germinated during 60 days at alternated temperatures, 16 h day at 30°C and 8 h night at 20°C (ISTA 1999). Walnut plantlets were then maintained in a climatic chamber under controlled environmental conditions of 16-hour photoperiod (16 h of light at 24°C and 8 h of darkness at 18°C).

Bacterial inoculations were performed when walnut plantlets had at least four young leaves fully expanded. Three plantlets were used for each isolate tested. Inoculum suspensions, prepared with SDW, were obtained from pure cultures grown on NA medium at 28 ± 2°C for 48

h. Bacterial suspensions were adjusted to a concentration of approximately 1 x 10^8 CFU ml^{-1} and confirmation of bacterial inoculum concentration was carried out by plating serial decimal dilutions on NA medium, with viable cell counting made 48 h after incubation. Plantlets were inoculated by spraying with a manual atomizer until runoff and kept in closed polyethylene bags for 48 h to promote bacterial infection, under the same temperatures and photoperiod conditions mentioned above. Plastic bags were then opened and plants maintained during four weeks for development of symptoms. Walnut plantlets sprayed with SDW were used as negative controls. Positive controls were performed by spraying a suspension of the reference type strain *Xaj* LMG 747 and *Xaj* isolate CPBF 1480 using the same concentration of viable cells. In order to fulfil Koch's postulates (Lelliott and Stead 1987), reisolation was performed from leaves presenting necrotic spots.

2.6. Accession number(s)

GenBank accession numbers corresponding to *acnB*, *fyuA*, *gyrB* and *rpoD* sequences of xanthomonads CPBF isolates is available as supplemental material (Supplementary **Table IV.S1**).

3. Results

3.1. Bacterial isolates obtained from walnut trees

A total of 131 isolates displaying yellow-pigmented colonies typical of *Xanthomonas* species were obtained. The majority of isolates were collected from symptomatic plant material (94%, 122/131), mostly from leaf samples (97 isolates, 74%), whereas nine (6%, 9/131) isolates were obtained from asymptomatic plant organs (**Table IV.1**). More than one isolate was recovered from 38 trees, either from the same sampling occurrence (i.e. the same plant in the same date) or from the same walnut tree at different sampling dates (**Table IV.1**).

3.2. MLSA revealed that *Xanthomonas* walnut isolates were grouped into ten distinct clusters

All walnut Portuguese isolates (CPBF) were assigned into the genus *Xanthomonas* and were clustered in ten different groups (Clusters I to X, **Figure IV.1**). Most isolates (85.5%, 112/131) were grouped in clusters I to VI together with strains of *X. arboricola* pv. *juglandis* and other closely related strains belonging to pathovars *pruni* and *corylina*. The other 19 (14,5%, 19/131)

isolates belong to clusters VII to X. Isolates from clusters VII and VIII were closely related with *X. arboricola* CFBP 7634 and *X. arboricola* CFBP 7651, previously described as nonpathogenic strains by Essakhi et al. (2015); with *X. arboricola* CITA 124, *X. arboricola* CITA 44 and *X. arboricola* CITA 14, previously described as avirulent strains by Garita-Cambronero et al. (2016, 2017) and with strain NCPPB 1630 belonging to the poorly characterized pathovar *celebensis* (Harrison et al. 2016) and *X. arboricola* 3004 isolated from barley and with uncertain pathogenicity (Ignatov et al. 2015). Isolates of groups IX and X formed two different clusters (bootstrap values > 95%) distant from the other isolates and strains of *X. arboricola* (**Figure IV.1**). Moreover, MLSA clustering data showed a clear separation between the six isolates from cluster IX and the nine isolates from cluster X with the other strains of the MLSA tree representatives of different *Xanthomonas* species, namely *X. albilineans*, *X. alfalfae*, *X. axonopodis*, *X. sacchari*, *X. campestris*, *X. cassava*e, *X. citri*, *X. euvesicatoria*, *X. floridensis*, *X. fragariae*, *X. fuscans*, *X. gardneri*, *X. hortorum*, *X. oryzae*, *X. perforans*, *X. prunicola*, *X. translucens*, *X. vasicola*.

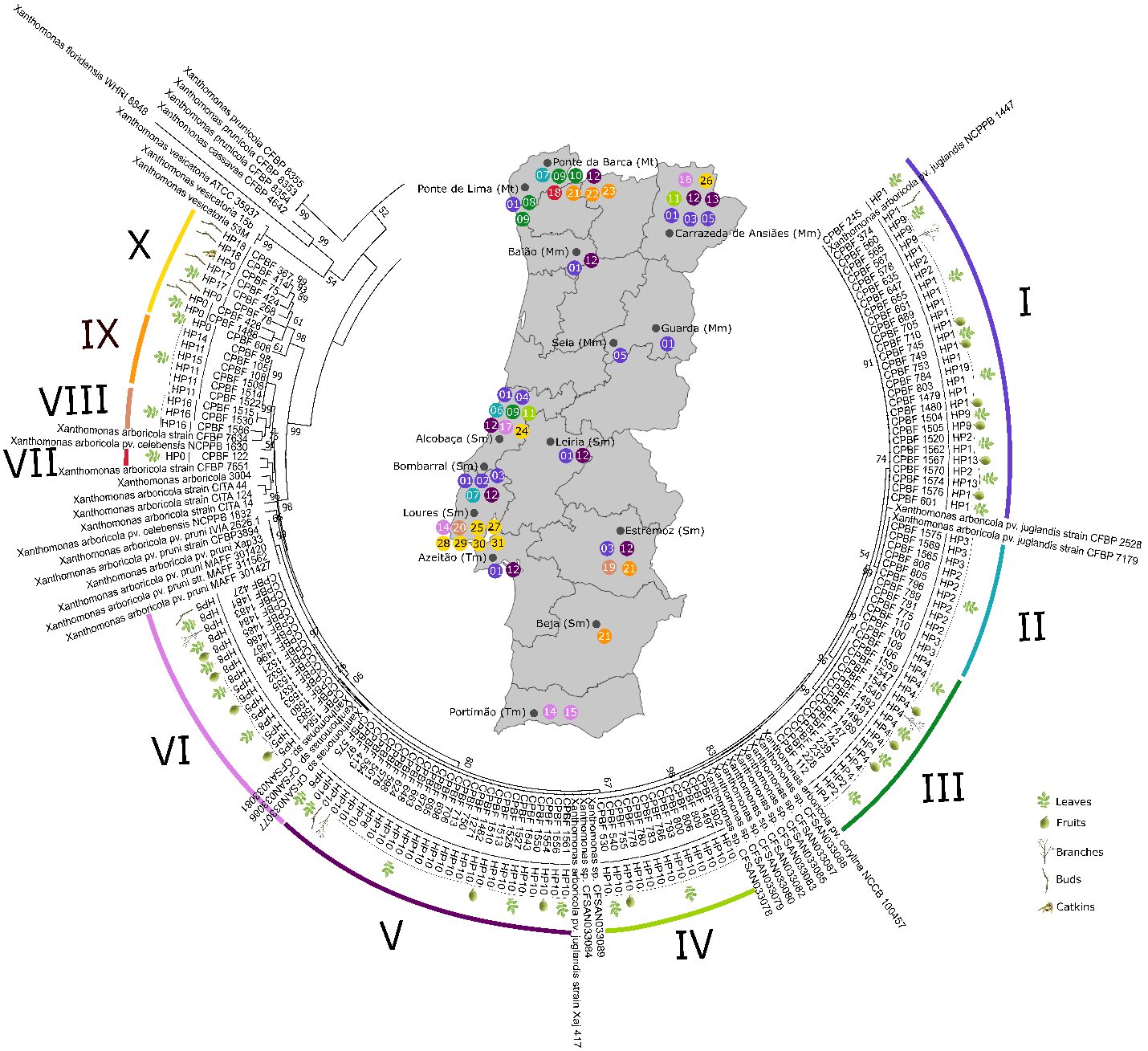

Figure IV.1. MLSA analysis and dot blot genotyping coupled with metadata including geographic locations, bioclimatic regions, and walnut host organs for the 131 *Xanthomonas* walnut isolates used in this work. The Maximum Likelihood tree was based on the nucleotide alignments of 252 concatenated sequences (2774 bp) of *acnB*, *gyrB*, *fyuA* and *rpoD* genes, using the General Time Reversible (GTR+G+I) model. Bootstrap values higher than 50 are shown. The tree was edited using MEGA 7.0 (Kumar et al. 2016) and the principal results are showed. Distinct MLSA clusters (I to X) of *Xanthomonas* isolates are highlighted with different colours. For all the isolates, the respective hybridization patterns (HP0 to HP19) is shown, as well as the plant organ from which they were isolated. The map in the centre highlights the fourteen geographic locations sampled, and details the respective thermoclimatic classification: Mt – mesotemperature; Mm – mesomediterranean; Sm – supramediterranean; Tm – thermomediterranean (Rivas-Martínez et al. 2004). The coloured circles indicate the different lineages found for each sampled region.

3.3. Diversity of isolates assessed using *Xaj* specific DNA markers

A total of 18 different hybridization patterns (HP), identified for the nine *Xaj* specific markers (XAJ1 to XAJ9) were obtained for the 131 isolates (**Figure IV.1** and **Figure IV.2**). The most representative was HP10, identified in 38 isolates (29%, 38/131). The less frequent were HP14, HP15 and HP19, only identified in one isolate each, followed by HP6, HP11, HP16, HP17, HP18 found in less than five isolates (**Figure IV.1**). Among these, HP16 (CPBF 1515, CPBF 1586 and CPBF 1530), HP17 (CPBF 268 and CPBF 424) and HP18 (CPBF 414 and CPBF 367) correspond to hybridization patterns limited to a single marker, XAJ9, XAJ2 and XAJ1, respectively. Six out of 131 isolates (CPBF 75, CPBF 78, CPBF 122, CPBF 426, CPBF 606 and CPBF 1488) provided negative hybridization results for all markers (pattern HP0, **Figure IV.1** and **Figure IV.2**).

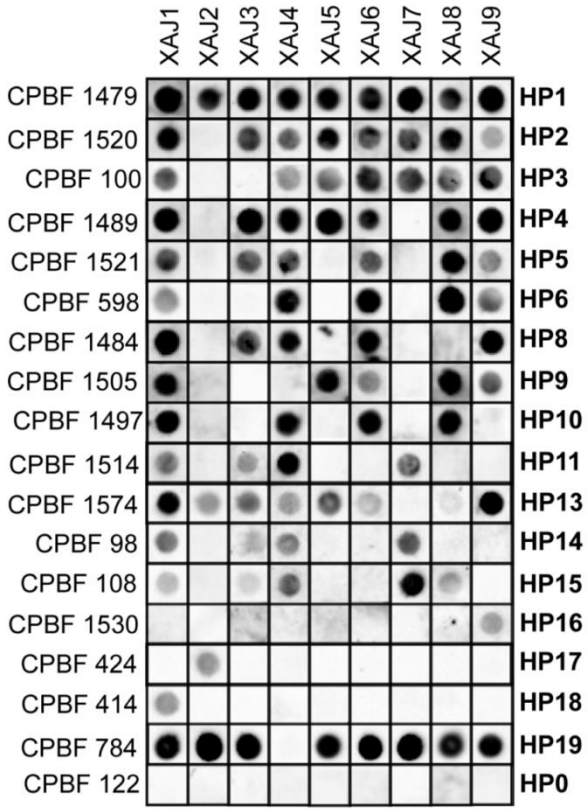

Figure IV.2. Dot blot matrix summarizing the 18 different hybridization patterns (HP0 to HP19) obtained with the 131 isolates using nine *Xanthomonas arboricola* pv. *juglandis* specific markers (XAJ1 to XAJ9). Strain LMG 751 was used as positive control for all dot blot assays.

3.4. Dot blot and MLSA data allowed to distinguish 31 lineages including *Xaj*, non-*juglandis X. arboricola* and non-*arboricola Xanthomonas* sp. isolates

Considering the MLSA for the 131 isolates, 20 unique concatenated sequences were obtained with 403 polymorphisms in 199 nucleotide sites (data not shown). The additional discrimination obtained with dot blot analysis for the XAJ1 to XAJ9 markers, resulted in 18 different HPs, allowing identify distinct lineages within the same MLSA cluster. Five distinct clonal lineages were observed in Group I (Lin01 to Lin05); two lineages in Group II (Lin06 and Lin07); three lineages in Group III (Lin08 to Lin10); one lineage in Group IV (Lin11); two lineages in Group V (Lin12 and Lin13); four lineages in Group VI (Lin14 to Lin17); one lineage in Group VII (Lin18); two lineages in Group VIII (Lin19 and Lin20); three lineages in Group IX (Lin21 to Lin23); and eight lineages in Group X (Lin24 to Lin) (**Table IV.1**). Altogether, compiling MLSA data and dot blot hybridization allowed distinguishing 31 lineages within the xanthomonads populations found in walnut trees (**Table IV.1**). Lin01 to Lin17 were represented by *Xaj* isolates (n=112), Lin18 to Lin20 were represented by *X. arboricola* isolates (n=4) and Lin21 to Lin31 were represented by non-*arboricola Xanthomonas* sp. isolates (n=15).

Xaj lineages Lin01 (comprised by 17 isolates) and Lin12 (comprised by 25 isolates) were the most representative. Moreover, four more lineages of *Xaj*, including Lin01 and Lin12, were always collected within walnut trees that were sampled in consecutive years (Lin1 and Lin12 from Jr#01 tree; Lin09 from Jr#26 tree; Lin12 from Jr#27 tree; and Lin14 from Jr#18 tree, **Table IV.1**). Furthermore, two *Xaj* lineages (Lin06 and Lin 17) were highly prevalent in walnut orchards located in Alcobaça. In fact, while Lin06 was found in six out of eight walnut trees sampled in one orchard, Lin17 was found in all five walnut trees sampled in another orchard. In the other hand, most of lineages of *X. arboricola* (Lin18, Lin19 and Lin20) and *Xanthomona*s sp. (Lin22, Lin23, Lin24, Lin25, Lin26, Lin27, Lin28, Lin29, Lin30 and Lin31) were represented by only one or two isolates of this population.

None of the epidemiological factors (year of isolation, bioclimatic regions, walnut cultivars, production regimes, host walnut specimen and plant organs) were observed to be linked with specific lineages. Nevertheless, it is worth to notice that *Xanthomonas* sp. lineages closely related and belonging to MLSA cluster X (Lin25, Lin27, Lin28, Lin30 and Lin31), were found to be overrepresented in walnut bud and catkin samples.

3.5. Isolates of distinct *Xaj*, *X. arboricola* and *Xanthomonas* sp. genetic lineages were found in the same walnut tree

From the 64 trees analysed in this work, more than one isolate was retrieved from 37 trees sampled at the same date. From these, 32 trees allowed to recover two or more distinct bacterial lineages from a single sampling event (**Table IV.1**). The majority of these isolates were obtained from the same leaflets samples of walnut trees (23 trees: Jr#26, Jr#27, Jr#28, Jr#32, Jr#36, Jr#37, Jr#38, Jr#41, Jr#46, Jr#49, Jr#50, Jr#52, Jr#53, Jr#54, Jr#56, Jr#57, Jr#58, Jr#59, Jr#60, Jr#61, Jr#62, Jr#63, Jr#64). Isolates were also obtained from the same fruit samples and bud samples of walnut trees (Jr#23, Jr#26 and Jr#01, Jr#18). Interestingly, all six trees that were analysed in consecutive years (Jr#01, Jr#03, Jr#06, Jr#18, Jr#26, Jr#27), consistently allowed recovering different lineages (**Table IV.1** and **Figure IV.1**).

3.6. Assessment of type III effector genes (*xopR*, *avrbs2*, *xopF1* and *xopN*), revealed *Xanthomonas* sp. lineage with novel T3E profile

The presence of T3E genes (*xopR*, *avrbs2*, *xopF1* and *xopN*) was assessed on representative isolates of each main cluster constituted by *Xaj* lineages (clusters I to VI, 16 isolates) and on all *X. arboricola* lineages (4 isolates) belonging to clusters VII and VIII and *Xanthomonas* sp. lineages (15 isolates) from cluster IX and X (**Table IV.1** and **Figure IV.1**). Among CPBF *Xaj* isolates, positive hybridization dots were obtained for all four T3E genes studied. *xopN* was not detected in all the 19 isolates belonging to *X. arboricola* and *Xanthomonas* sp. lineages. Ten of these isolates, belonged to clusters VII, VIII and IX, hybridized with *xopR*, *avrBs2* and *xopF1* specific probes. Two *Xanthomonas* sp. isolates of cluster X (CPBF 78 and CPBF 424) hybridized with *xopR* and *xopF1* and the remaining seven isolates of cluster X only hybridized with *xopR* gene (**Figure IV.3**).

The T3E gene patterns obtained for the 35 isolates assayed could be compared with data previously reported by Essakhi et al. (2015), with exception of two isolates (CPBF 78 and CPBF 424), which presented a novel profile of T3E genes (**Figure IV.3** and Supplementary **Figure IV.S1**).

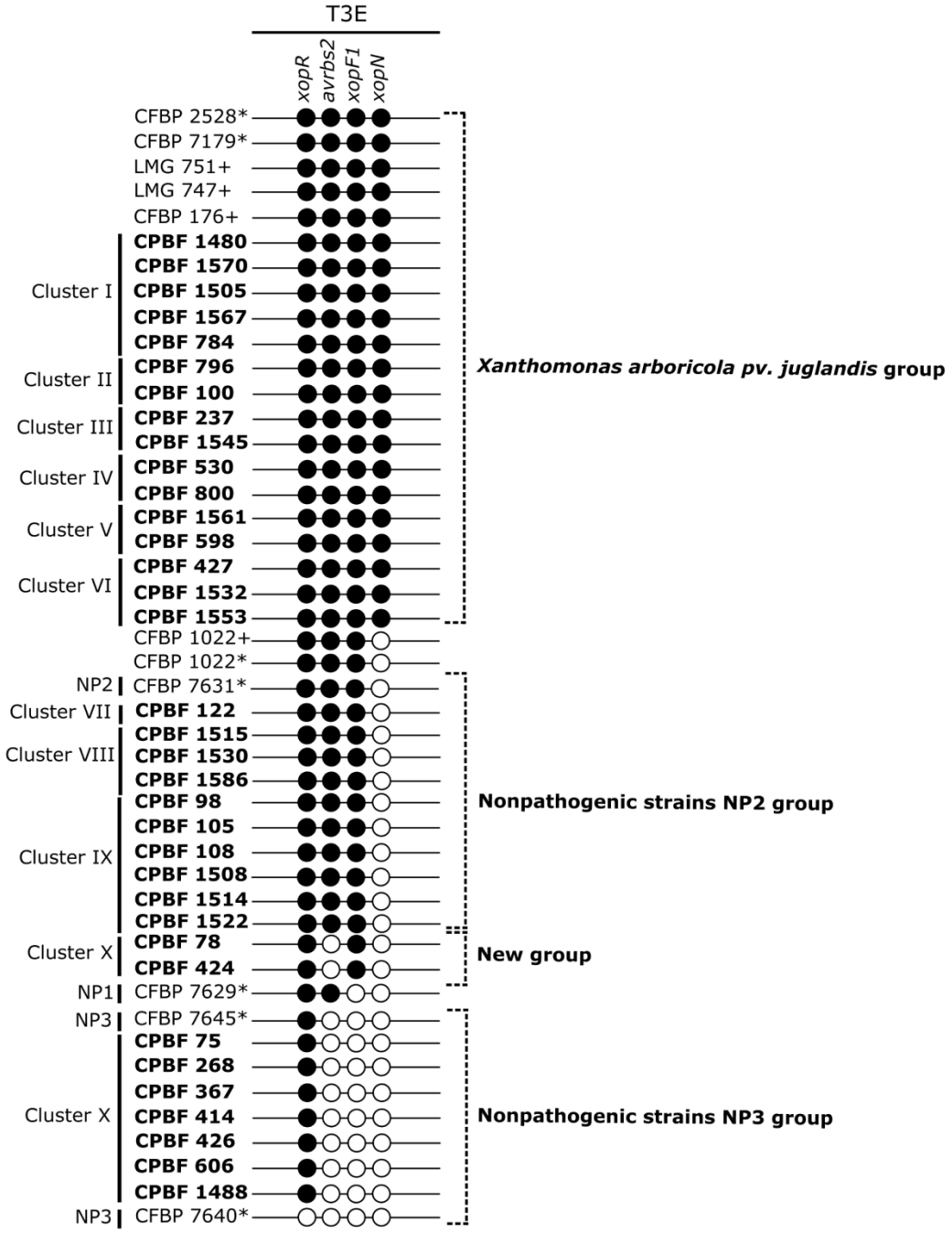

Figure IV.3. Dot blot matrix reporting the presence/absence of four type three effector genes (T3E - *xopR*, *avrBs2*, *xopF1* and *xopN*) in *Xanthomonas* strains and isolates representing the ten MLSA clusters (I to X) and the 18 hybridization patterns (HP0 to HP19). Black circles represent positive dot blot hybridization to the target gene and white circles represent negative dot blot hybridization to the target gene. Sixteen isolates belonging to MLSA clusters I to VI and all isolates grouped in MLSA clusters VII to X were screened. *Xanthomonas arboricola* pv. *juglandis* (*Xaj*) reference strains LMG 751, LMG 747, CFBP 176 and *Xanthomonas arboricola* CFBP 1022 were used as controls (highlighted with +). Dot blot images are supplied in Fig. S1 in the supplemental material. For comparison, the distribution of *xopR*, *avrBs2*, *xopF1* and *xopN* genes, described by Essakhi et al. (2015) for six strains is highlighted with an asterisk (*), including two *Xaj* strains (CFBP 2528 and CFBP 7179) and four strains isolated from walnut and belonging to three nonpathogenic groups (NP1 – strains CFBP 1022, CFBP 7629; NP2 – strain CFBP 7631; and NP3 – strain CFBP 7645).

3.7. Non-*arboricola Xanthomonas* isolate can cause bacterial leaf spots on walnut

Pathogenicity assays were conducted in order to verify the ability of five *Xanthomonas* sp. isolates with distinct composition of T3Es (CPBF 75, CPBF 367, CPBF 424, CPBF 1488 and CPBF 1514) to cause typical disease symptoms on *Juglans regia*. Seven days after inoculation of *Xaj* strain LMG 747, bacterial necrotic spots were observed on leaves of walnut plantlets. Similar symptoms were detected for *Xaj* isolate CPBF 1480 and for the non-*arboricola* *Xanthomonas* isolate CPBF 424 (**Figure IV.4**). No symptoms were observed using *Xanthomonas* sp. isolates CPBF 75, CPBF 367, CPBF 1488 and CPBF 1514 (data not shown), suggesting they are nonpathogenic on walnut trees. Re-isolation of yellow mucoid bacterial colonies followed by sequencing analysis of *gyrB* and *fyuA* partial sequences (data not shown) confirmed that *Xanthomonas* sp. isolate CPBF 424, is pathogenic on *Juglans regia*.

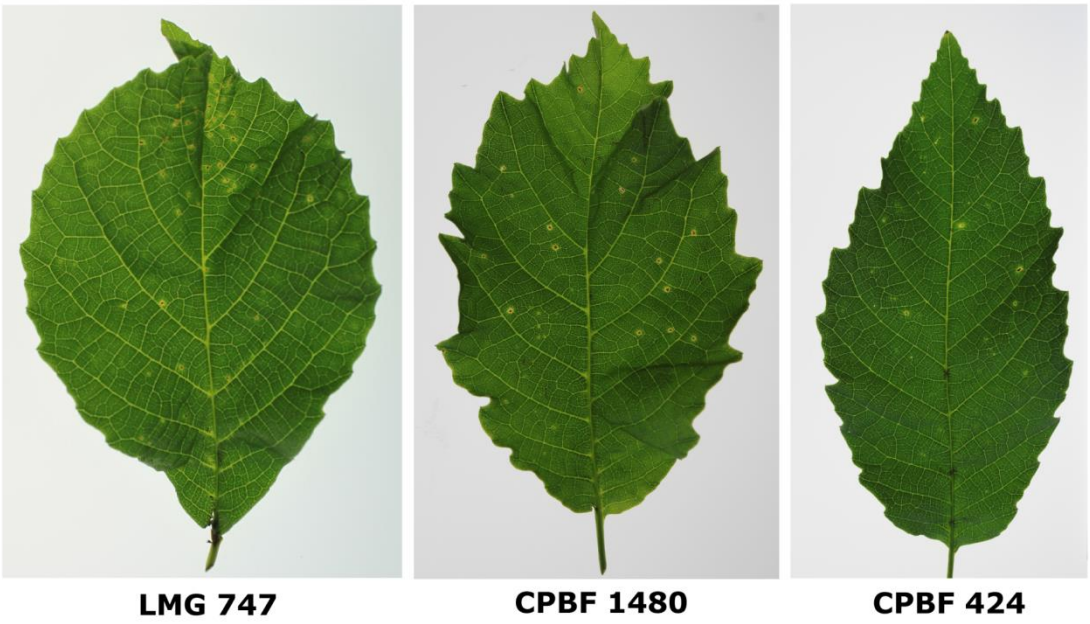

Figure IV.4. Pathogenicity assays showing symptoms on walnut leaves developed four weeks after inoculation. LMG 747 (type strains of *X. arboricola*) and CPBF 1480 (isolate from MLSA cluster I and HP1) were used as positive controls. Walnut leaves inoculated with isolate CPBF 424 (MLSA cluster X) showed similar bacterial necrotic spots as the positive controls. No symptoms were observed for *Xanthomonas* sp. isolates CPBF 75, CPBF 367, CPBF 1488 (MLSA cluster X) and CPBF 1514 (MLSA cluster IX) (data not shown).

4. Discussion

Walnut tree is acknowledged as a host plant species for *X. arboricola* providing a permanent niche for the pathovar *juglandis*, which are walnut pathogens characterized by a high genetic diversity (Loreti et al. 2001, Scortichini et al. 2001, Hajri et al. 2010, Marcelletti et al. 2010, Ivanovic et al. 2014, Kaluzna et al. 2014, Giovanardi et al. 2015). Regardless these contributions, most of these studies were based either on established collections of strains, or on a limited number of field isolates apparently obtained through a random sampling strategy. To characterize comprehensively the population diversity of *Xanthomonas* walnut pathogens, the present research followed a sampling plan taking into account metadata, which might had influenced disease epidemiology in Portugal. Currently, WBB and BAN disease symptoms are commonly observed in walnut Portuguese orchards and in dispersed walnut trees across the country, affecting crop yield drastically. Although is acknowledged that *Xaj* is the main pathogenic agent of walnut diseases all over the country, information about genetic diversity of this pathogen, its dissemination and epidemiology remain unknown, impairing the implementation of scientifically informed phytosanitary measures.

Portugal is a walnut production country characterized by distinct bioclimatic regions and a long historical record of WBB disease symptoms. The first scientific evidence for walnuts infected with *Xaj* date back to 1935 (Oliveira and Cabral 1942). The authors mentioned that high humidity favoured the infection and suggested that both environmental and meteorological conditions could influence disease development (Oliveira and Cabral 1942). Bioclimatic features and culture practices have also been suggested to influence not only the epidemiology and prevalence of walnut disease, but also the bacterial population structure (Scortichini et al. 2001, Kaluzna et al. 2014, Giovanardi et al. 2015). Accordingly, the 131 isolates of this work were collected from leaves, fruits, branches, buds and catkins of different cultivars of diseased walnut trees mainly located in geographic regions with high walnut production and characterized by distinct bioclimatic conditions and different management strategies. Being noteworthy the most relevant walnut cultivation regions of the country, the regions of Trás-os-Montes (e.g. Carrazeda de Ansiães location) with intensive summer drought and winter severity, i.e. mesomediterranean bioclimatic region, characterized by a walnut production largely based on isolated walnut trees; the regions of Alentejo (e.g. Beja and Estremoz locations) dominated by regularly warm temperatures and dry climate, i.e. supramediterranean bioclimatic region; and the regions of Minho (e.g. Ponte da Barca and Ponte de Lima) with high annual precipitation and relatively mild summers, i.e. mesotemperature bioclimatic region; being both (Alentejo and Minho) major walnut producing areas in Portugal characterized by orchards planted with selected *Juglans regia* cultivars.

Isolates diversity was assessed by MLSA considering four housekeeping genes (*acnB*, *fyuA*, *gyrB* and *rpoD*), following an approach previously used to determine the phylogeny of *Xanthomonas* spp. (Parkinson et al. 2007, Young et al. 2008). In fact, the large number of sequences available at the GenBank database, the discriminatory power of each of these housekeeping genes for strains from pathovar *juglandis* and its wide utilization in transversal studies on *Xaj* (Marcelletti et al. 2010; Ivanovic et al. 2014; Kaluzna et al. 2014; Giovanardi et al. 2015), makes MLSA a suitable method to rapid infer diversity and identify deeply branching clusters within bacterial populations (Gevers et al. 2005). Furthermore, MLSA has been particularly useful for differentiating *X. arboricola* isolates obtained from the same host plant (Essakhi et al. 2015, Garita-Cambronero et al. 2016, 2017). In this study, ten different MLSA clusters were defined (I to X, **Figure IV.1**) with most of the isolates being distributed in clusters I to VI, which also contained several fully sequenced *Xaj* strains, namely *Xaj* 417 (Pereira et al. 2015); NCPPB 1447; CFBP 2528 and CFBP 7179 (Cesbron et al. 2015); CFSAN033077-89 (Higuera et al. 2015). The remaining isolates, clustered in groups VII to X (**Figure IV.1**), were clearly distinguishable from the main *Xaj* MLSA clusters. Clusters VII and VIII, closely related to other *X. arboricola,* as nonpathogenic or avirulent strains (Essakhi et al. 2015, Garita-Cambronero et al. 2016, 2017), comprise isolates belong to the species *X. arboricola* isolated from symptomatic leaves. Cluster IX and X were grouped outside any other cluster defined by *X. arboricola*. Cluster IX is formed by a homogeneous group of six *Xanthomonas* sp. isolated from symptomatic leaves and cluster X is the group of *Xanthomonas* sp. mainly isolated from asymptomatic material (bud and catkin samples) showing the highest genetic variation. To further elucidate the genetic diversity of all 131 isolates, dot blot hybridization assays using *Xaj*-specific DNA markers XAJ1 to XAJ9 (Fernandes et al. 2017), allowed to identify eighteen distinct hybridization patterns (**Figure IV.2**), and distinguish isolates with very low levels of sequence divergence among housekeeping genes, within the same MLSA cluster (**Figure IV.1**). The genotyping appraisal of a considerable number of isolates confirm the utility of these markers for the identification of different *Xaj* lineages, as suggested previously (Fernandes et al. 2017). Furthermore, these molecular markers were shown to be useful as genetic tools on the characterization of all xanthomonads population found on walnut trees. Although a perfect match between specific hybridization patterns and MLSA groups was not observed, it was clear that all the HPs corresponding to none (HP0) or a single positive *Xaj* specific markers (HP16, HP17, HP18) were represented by *X. arboricola* and *Xanthomonas* sp. isolates from more divergent MLSA groups. Moreover, when MLSA and dot blot hybridization patterns were combined, it was possible to enhance the genotyping resolution and disclose at least 31 *Xanthomonas* lineages, among the 131 isolates associated with walnut trees.

Seventeen lineages of *Xaj*, three lineages of *X. arboricola* and eleven lineages of non-*arboricola Xanthomonas* (*Xanthomonas* sp.) defined the walnut-associated xanthomonads population observed to occur in Portugal. *Xaj* lineages Lin01 and Lin12 appeared to be the most prevalent over time and well stablished in the country, and Lin06 and Lin17 showed to be persistent in orchards likely under pressure of specific conditions (environment, climatic or culture practices), making these *Xaj* lineages as the main troublesome currently occurring in Portugal. Further studies will be needed to determine the incidence of these lineages, evaluate their virulence, and to identify specific adaptation traits that might explain their broad dissemination.

The majority of *X. arboricola* and *Xanthomonas sp.* genetic lineages showed to be the less frequent lineages on population. These groups probably reflect part of the recombinant network proposed for *X. arboricola* population structure (Merda et al. 2016). Interestingly, the most divergent group of *Xanthomonas* sp. lineages appeared to be preferentially associated to walnut catkins and buds than with mature leaves and fruits, indicating a preferential plant-organ colonization pattern for these lineages. These results aligned with previous studies suggesting that selective pressures on bud organs could favouring the emergence of distinct lineages (Hajri et al. 2010, Marcelletti et al. 2010, Merda et al. 2016). The identification of non-*arboricola Xanthomonas* lineages, also showed that bacterial xanthomonads found on walnut trees is more complex than originally thought, with possible interspecies interactions (Lamichhane and Venturi 2015) beyond the dynamics already proposed between nonpathogenic and pathogenic *X. arboricola* strains (Merda et al. 2016). Furthermore, walnut trees could be infected simultaneously by different *Xaj* strains, causing distinct symptoms at different plant organs (WBB, BAN and VOC) (Belisario et al. 2002, Hajri et al. 2010). These data indicate that sympatry of walnut colonizing *Xanthomonas* may play an important role on the emergence of new lineages (Essakhi et al. 2015, Merda et al. 2016). In the present study, we gathered evidence suggesting that sympatric association is not occasional, since was found in 26 of the 64 walnut trees sampled, as shown by the coexistence of different xanthomonads lineages infecting leaves, fruits and buds of the same walnut tree (**Table IV.1**). Further investigations are important to determine if the xanthomonads diversity within the same walnut host plant is a mix of bacterial populations colonizing evenly the same plant organ, or if it results from the co-colonization of dominant versus lessened xanthomonads populations.

When studying the presence of the T3Es genes *xopR*, *avrBs2*, *xopF1* and *xopN*, suitable to distinguish groups of nonpathogenic and pathogenic strains, according Essakhi et al. (2015), we could assign most of the 19 isolates evaluated from cluster VII to X to nonpathogenic groups, whereas CPBF 424 (and CPBF 78) displayed a different pattern of T3E genes (presence of *xopR* and *xopF1*; absence of *avrbs2* and *xopN*) when compared with *X. arboricola* pathogenic strains (**Figure IV.3**). Although a core set of T3Es genes has been described as essential for

pathogenicty in *X. arboricola*, variations in T3E repertoire is considered normal (Hajri et al. 2012, Essakhi et al. 2015, Merda et al. 2016). For instance, the *xopR*, *xopF1* and *avrBs2* genes have been pointed out as belonged to the set of T3E ubiquitous and also ancestors of *X. arboricola* species (Hajri et al. 201, Merda et al. 2016), being also identified in nonpathogenic strains (Essakhi et al. 2015, Merda et al. 2016), whereas *xopN* has been characterized as conserved in *Xanthomonas* genomes, but in *X. arboricola* species was described inside the group of variable T3Es, being absent in some pathovars (Hajri et al. 2012). Interestingly, in this work we isolated a pathogenic non-*arboricola Xanthomonas* (CPBF 424 isolate) having a unique pattern of T3E and characterized by the absence of T3Es commonly present on *Xaj* pathogenic strains. The recently sequenced genome of *Xanthomonas* sp. CPBF 424 (Fernandes et al. 2018a) allowed to confirm the T3E profile of this strains and revealed and average nucleotide identity below 95% to any of the other Xanthomonas species which is suggestive of a new species (Fernandes et al. 2018b). This finding reporting strain CPBF 424 as the first non-*arboricola Xanthomonas* pathogenic on walnut, might be particularly important to disclose determinants of virulence and contribute to better understand the evolution of pathogenicity in walnut *Xanthomonas* pathogens.

5. Conclusions

Following an extensive sampling of walnut associated *Xanthomonas* during three consecutive years, in distinct walnut producing regions characterized by diverse bioclimatic regions and different walnut production practices, this study investigated the distribution of *Xaj* in Portugal and its genetic diversity. Comprehensive genotyping analyses allowed to identify the most prevalent *Xaj* lineages, the possible emergence of new *Xaj* lineages and disclose non-infective *X. arboricola* strains and a non-*arboricola* pathogenic *Xanthomonas* sp., which might provide new insights to elucidate *Xanthomonas* pathoadaptations and speciation.

Acknowledgments

This work was co-financed by the European Structural & Investment Funds (ESIFs) through the Operational Competitiveness and Internationalization Programme – COMPETE 2020 and by National Funds through FCT – Fundação para a Ciência e a Tecnologia, within the framework of the project EVOXANT (PTDC/BIA-EVF/3635/2014, POCI-01-0145-FEDER-016600). CF is supported by a fellowship from the FCT (SFRH/BD/95913/2013).

The authors would like to thank the walnut producers who collaborated in this study, Lurdes Santos and Paulo Godinho for the work performed on walnut seedlings and Joaquim Trindade for taking the photos of pathogenicity results.

References

Barionovi D, Scortichini M. 2008. Integron variability in *Xanthomonas arboricola* pv. *juglandis* and *Xanthomonas arboricola* pv. *pruni* strains. *FEMS Microbiology Letters* **288**,19-24.

Belisario A, Maccaroni M, Corazza L, Balmas V, Valier A, 2002. Occurrence and etiology of brown apical necrosis on Persian (English) walnut fruit. *Plant Disease* **86**, 599-602.

Cesbron S, Briand M, Essakhi S, Gironde S, Boureau T, Manceau C, Fischer-Le Saux M, Jacques MA. 2015. Comparative genomics of pathogenic and nonpathogenic strains of *Xanthomonas arboricola* unveil molecular and evolutionary events linked to pathoadaptation. *Frontiers in Plant Science* **6**, 1126.

Essakhi S, Cesbron S, Fischer-Le Saux M, Bonneau S, Jacques MA, Manceau C. 2015. Phylogenetic and variable-number tandem-repeat analyses identify nonpathogenic *Xanthomonas arboricola* lineages lacking the canonical Type III Secretion System. *Applied and Environmental Microbiology* **81**, 5395-5410.

Esterio MA, Latorre BA. 1982. Potential sources of inoculum of *Xanthomonas juglandis* in walnut blight outbreaks. *Journal of Horticultural Science* **57**, 69-72.

Fernandes C, Albuquerque P, Sousa R, Cruz L,Tavares F. 2017. Multiple DNA markers for identification of *Xanthomonas arboricola* pv. *juglandis* isolates and its direct detection in plant samples. *Plant Disease* **101**, 858-865.

Fernandes C, Blom J, Pothier JF, Tavares F. 2018a. High-quality draft genome sequence of *Xanthomonas* sp. strain CPBF 424, a walnut-pathogenic strain with atypical features. *Microbiology Resource Announcements* **7** (15), e00921-18.

Fernandes C, Pothier JF, Tavares F. 2018b. Whole-genome sequencing of distinct *Xanthomonas arboricola* lineages isolated from a single walnut tree host, abstr poster 9, p 88. Abstr 6th *Xanthomonas* Genomics Conference and 2nd Annual EuroXanth Conf, Halle (Saale), Germany. https://doi.org/10.13140/RG.2.2.13768.37128.

Fischer-Le Saux M, Bonneau S, Essakhi S, Manceau C, Jacques MA. 2015. Aggressive emerging pathovars of *Xanthomonas arboricola* represent widespread epidemic clones distinct from poorly pathogenic strains, as revealed by multilocus sequence typing. *Applied and Environmental Microbiology* **81**, 4651-4668.

Frutos D. 2010. Bacterial diseases of walnut and hazelnut and genetic resources. *Journal of Plant Pathology* **92**, S79-S85.

Garita-Cambronero J, Palacio-Bielsa A, López MM, Cubero J. 2016. Draft genome sequence for virulent and avirulent strains of *Xanthomonas arboricola* isolated from *Prunus* spp. in Spain. *Standards in Genomic Sciences* **11**, 12.

Garita-Cambronero J, Palacio-Bielsa A, Lopez MM, Cubero J. 2017. Pan-genomic analysis permits differentiation of virulent and non-virulent strains of *Xanthomonas arboricola* that cohabit *Prunus* spp. and elucidate bacterial virulence factors. *Frontiers in Microbiology* **8**, 573.

Gevers D, Cohan FM, Lawrence JG, Spratt BG, Coenye T, Feil EJ, Stackebrandt E, Van de Peer Y, Vandamme P, Thompson FL, Swings J. 2005. Re-evaluating prokaryotic species. *Nature Reviews Microbiology* **3**, 733-739.

Giovanardi D, Bonneau S, Gironde S, Fischer-Le Saux M, Manceau C, Stefani, E. 2015. Morphological and genotypic features of *Xanthomonas arboricola* pv. *juglandis* populations from walnut groves in Romagna region, Italy. *European Journal of Plant Pathology* **145**,1-16.

Hajri A, Meyer D, Delort F, Guillaumes J, Brin C, Manceau C. 2010. Identification of a genetic lineage within *Xanthomonas arboricola* pv. *juglandis* as the causal agent of vertical oozing canker of Persian (English) walnut in France. *Plant Pathology* **59**, 1014-1022.

Hajri A, Pothier JF, Fischer-Le Saux M, Bonneau S, Poussier S, Boureau T, Duffy B, Manceau C. 2012. Type three effector gene distribution and sequence analysis provide new insights into the pathogenicity of plant-pathogenic *Xanthomonas arboricola*. *Applied and Environmental Microbiology* **78**, 371-84.

Harrison J, Grant MR, Studholme DJ**.** 2016. Draft genome sequences of two strains of *Xanthomonas arboricola* pv. *celebensis* isolated from banana plants. *Genome Announcements* **4**, e01705-15.

Higuera G, Gonzalez-Escalona N, Veliz C, Vera F, Romero J. 2015. Draft genome sequences of four *Xanthomonas arboricola* pv. *juglandis* strains associated with walnut blight in Chile. *Genome Announcements* **3**, e01160-15.

Ignatov AN, Kyrova EI, Vinogradova SV, Kamionskaya AM, Schaad NW, Luster DG. 2015. Draft genome sequence of *Xanthomonas arboricola* strain 3004, a causal agent of bacterial disease on barley. *Genome Announcements* **3**, e01572-14.

International Seed Testing Association. 1999. International rules for seed testing rules 1999. Seed Science & Technology 27 supplement.

Ivanovic Z, Popovic T, Janse J, Kojic M, Stankovic S, Gavrilovic V, Fira D. 2014. Molecular assessment of genetic diversity of *Xanthomonas arboricola* pv. *juglandis* strains from Serbia by various DNA fingerprinting techniques. *European Journal of Plant Pathology* **141**, 133-145.

Kaluzna M, Pulawska J, Waleron M, Sobiczewski P. 2014. The genetic characterization of *Xanthomonas arboricola* pv. *juglandis*, the causal agent of walnut blight in Poland. *Plant Pathology* **63**, 1404-1416.

Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution* **33**, 1870-1874.

Lamichhane JR. 2014. *Xanthomonas arboricola* diseases of stone fruit, almond, and walnut trees: Progress toward understanding and management. *Plant Disease* **98**, 1600-1610.

Lamichhane JR, Venturi V. 2015. Synergisms between microbial pathogens in plant disease complexes: a growing trend. *Frontiers in Plant Science* **6**, 385.

Lelliott RA, Stead DE. 1987. Methods for the diagnosis of bacterial diseases of plants. Methods in Plant Pathology. Oxford, UK: British Society for Plant Pathology by Blackwell Scientific Publications.

Lindow S, Olson W, Buchner R. 2014. Colonization of dormant walnut buds by *Xanthomonas arboricola* pv. *juglandis* is predictive of subsequent disease. *Phytopathology* **104**, 1163-1174.

Loreti S, Gallelli A, Belisario A, Wajnberg E, Corazza L. 2001. Investigation of genomic variability of *Xanthomonas arboricola* pv. *juglandis* by AFLP analysis. *European Journal of Plant Pathology* **107**, 583-591.

Marcelletti S, Ferrante P, Scortichini M. 2010. Multilocus sequence typing reveals relevant genetic variation and different evolutionary dynamics among strains of *Xanthomonas arboricola* pv. *juglandis*. *Diversity* **2**, 1205-1222.

McMahon C, Denaxas S. 2016. A novel framework for assessing metadata quality in epidemiological and public health research settings. Proceedings *AMIA Joint Summits on Translational Science* **2016**, 199-208.

Merda D, Bonneau S, Guimbaud JF, Durand K, Brin C, Boureau T, Lemaire C, Jacques Ma, Fischer-Le Saux M. 2016. Recombination-prone bacterial strains form a reservoir from which epidemic clones emerge in agroecosystems. *Environmental Microbiology Reports* **8**, 572-581.

Moragrega C, Matias J, Aleta N, Montesinos E, Rovira M. 2011. Apical necrosis and premature drop of Persian (English) walnut fruit caused by *Xanthomonas arboricola* pv. *juglandis*. *Plant Disease* **95**, 1565-1570.

Mulrean EN, Schroth MN. 1982. Ecology of *Xanthomonas campestris* pv. *juglandis* on Persin (English) walnuts. *Phytopathology*, **72**, 434-438.

Oliveira M, Cabral R. 1942. Doenças bacterianas das plantas diagnosticadas em Portugal. *Revista Agronómica* **XXX**,176-184. [In Portuguese]

Parkhill J, Wren BW. 2011. Bacterial epidemiology and biology - lessons from genome sequencing. *Genome Biolology* **12**, 230-230.

Parkinson N, Aritua V, Heeney J, Cowie C, Bew J, Stead D. 2007. Phylogenetic analysis of *Xanthomonas* species by comparison of partial gyrase B gene sequences. International *Journal of Systematic and Evolutionary Microbiology* **57**, 2881-2887.

Pereira UP, Gouran H, Nascimento R, Adaskaveg JE, Goulart LR, Dandekar AM. 2015. Complete genome sequence of *Xanthomonas arboricola* pv. *juglandis* 417, a copper-resistant strain isolated from *Juglans regia* L. *Genome Announcements* **3**, e01126-15.

Pierce NB. 1901. Walnut bacteriosis. *Botanical Gazette* **31**, 272-273.

Pruvost O, Gardan L. 1988. Dynamics of epiphytic survival of *Xanthomonas campestris* pathovar *corylina*, *Xanthomonas campestris* pathovar *juglandis*, *Erwinia herbicola* and *Pseudomonas paucimobilis* on hazelnut leaves. *Agronomie* **8**, 925-932.

Rademaker JL, Louws FJ, Schultz MH, Rossbach U, Vauterin L, Swings J, de Vruijn FJ. 2005. A comprehensive species to strain taxonomic framework for *Xanthomonas*. *Phytopathology* **95**, 1098-111.

Rivas-Martínez S, Penas A, Díaz T E. Data from: Bioclimatic map of Europe – Thermoclimatic belts. Phytosociological Research Center. 2004
http://www.globalbioclimatics.org/form/tb_map/MS30W180.htm

Rudolpht R. 1993. Infection of the plant by *Xanthomonas*. *Xanthomonas*. London: Chapman & Hall.

Scortichini M, Marchesi U, Di Prospero P. 2001. Genetic diversity of *Xanthomonas arboricola* pv. *juglandis* (synonyms: *X. campestris* pv. *juglandis*; *X. juglandis* pv. *juglandis*) strains from different geographical areas shown by repetitive polymerase chain reaction genomic fingerprinting. *Journal of Phytopathology-Phytopathologishe Zeitschrift* **149**, 325-332.

Smith RE, Smith CO, Ramsey HJ. 1912. Walnut culture in California: walnut blight. Berkeley, California: California Agricultural Experiments Station Publications.

Vauterin L, Hoste B, Kersters K, Swings J. 1995. Reclassification of *Xanthomonas*. *International Journal of Systematic Bacteriology* **45**, 472-489.

Young JM, Park DC, Shearman HM, Fargier E. 2008. A multilocus sequence analysis of the genus *Xanthomonas*. *Systematic and Applied Microbiology* **31**, 366-377.

Supplementary material

Table IV.S1. Genbank accession number of partial *acnB*, *fyuA*, *gyrB* and *rpoD* sequences of *Xanthomonas* CPBF isolates used for MLSA analysis in this study

Isolates[a] (CPBF)	*acnB* Accession number	*fyuA* Accession number	*gyrB* Accession number	*rpoD* Accession number
75	KY745907	KY746011	KY746115	KY746217
78	KY745908	KY746012	KY746116	KY746218
97	KY745909	KY746013	KY746117	KY746219
98	KY745910	KY746014	KY746118	KY746220
100	KY745911	KY746015	KY746119	KY746221
105	KY745912	KY746016	KY746120	KY746222
106	KY745913	KY746017	KY746121	KY746223
108	KY745914	KY746018	KY746122	KY746224
109	KY745915	KY746019	KY746123	KY746225
110	KY745916	KY746020	KY746124	KY746226
112	KY745917	KY746021	KY746125	KY746227
122	KY745918	KY746022	KY746126	KY746228
228	KY745919	KY746023	KY746127	KY746229
237	KY745920	KY746024	KY746128	KY746230
239	KY745921	KY746025	KY746129	KY746231
245	KY745922	KY746026	KY746130	KY746232
268	KY745923	KY746027	KY746131	KY746233
367	KY745924	KY746028	KY746132	KY746234
374	KY745925	KY746029	KY746133	KY746235
413	KY745926	KY746030	KY746134	KY746236
414	KY745927	KY746031	KY746135	KY746237
424	KY745928	KY746032	KY746136	KY746238
426	KY745929	KY746033	KY746137	KY746239
427	KY745930	KY746034	KY746138	KY746240
530	KY745931	KY746035	KY746139	KY746241
540	KY745932	KY746036	KY746140	KY746242
554	KY745933	KY746037	KY746141	KY746243
560	KY745934	KY746038	KY746142	KY746244
561	KY745935	KY746039	KY746143	KY746245
565	KY745936	KY746040	KY746144	KY746246
567	KY745937	KY746041	KY746145	KY746247
575	KY745938	KY746042	KY746146	KY746248
576	KY745939	KY746043	KY746147	KY746249
578	KY745940	KY746044	KY746148	KY746250
598	KY745941	KY746045	KY746149	KY746251
601	KY745942	KY746046	KY746150	KY746252
606	KY745943	KY746047	KY746151	KY746253
624	KY745944	KY746048	KY746152	KY746254
635	KY745945	KY746049	KY746153	KY746255
647	KY745946	KY746050	KY746154	KY746256
648	KY745947	KY746051	KY746155	KY746257
655	KY745948	KY746052	KY746156	KY746258
661	KY745949	KY746053	KY746157	KY746259
663	KY745950	KY746054	KY746158	KY746260
676	KY745951	KY746055	KY746159	KY746261
685	KY745952	KY746056	KY746160	KY746262
689	KY745953	KY746057	KY746161	KY746263
698	KY745954	KY746058	KY746162	KY746264
705	KY745955	KY746059	KY746163	KY746265
706	KY745956	KY746060	KY746164	KY746266
710	KY745957	KY746061	KY746165	KY746267
713	KY745958	KY746062	KY746166	KY746268

Isolates[a] (CPBF)	*acnB* Accession number	*fyuA* Accession number	*gyrB* Accession number	*rpoD* Accession number
742	KY745959	KY746063	KY746167	KY746269
745	KY745960	KY746064	KY746168	KY746270
747	KY745961	KY746065	KY746169	KY746271
749	KY745962	KY746066	KY746170	KY746272
750	KY745963	KY746067	KY746171	KY746273
753	KY745964	KY746068	KY746172	KY746274
755	KY745965	KY746069	KY746173	KY746275
775	KY745966	KY746070	KY746174	KY746276
778	KY745967	KY746071	KY746175	KY746277
780	KY745968	KY746072	KY746176	KY746278
781	KY745969	KY746073	KY746177	KY746279
783	KY745970	KY746074	KY746178	KY746280
784	KY745971	KY746075	KY746179	KY746281
786	KY745972	KY746076	KY746180	KY746282
789	KY745973	KY746077	KY746181	KY746283
793	KY745974	KY746078	KY746182	KY746284
796	KY745975	KY746079	KY746183	KY746285
800	KY745976	KY746080	KY746184	KY746286
803	KY745977	KY746081	KY746185	KY746287
805	KY745978	KY746082	KY746186	KY746288
806	KY745979	KY746083	KY746187	KY746289
808	KY745980	KY746084	KY746188	KY746290
809	KY745981	KY746085	KY746189	KY746291
1271	KU058273	KU058299	KU058325	KU058363
1479	KU058274	KU058300	KU058326	KU058364
1480	KU058275	KU058301	KU058327	KU058365
1481	KY745982	KY746086	KY746190	KY746292
1482	KY745983	KY746087	KY746191	KY746293
1483	KU058276	KU058302	KU058328	KU058366
1484	KU058277	KU058303	KU058329	KU058367
1485	KU058278	KU058304	KU058330	KU058368
1486	KU058279	KU058305	KU058331	KU058369
1487	KU058280	KU058306	KU058332	KU058370
1488	KY986729	KY986728	KY986727	KY986726
1489	KU058281	KU058307	KU058333	KU058371
1490	KU058282	KU058308	KU058334	KU058372
1491	KU058283	KU058309	KU058335	KU058373
1492	KU058284	KU058310	KU058336	KU058374
1496	KU058285	KU058311	KU058337	KU058375
1497	KY745984	KY746088	KU058338	KY746294
1502	KU058286	KU058312	KU058339	KU058376
1504	KU058287	KU058313	KU058341	KU058377
1505	KU058288	KU058314	KU058342	KU058378
1508	KU058289	KU058315	KU058345	KU058379
1510	KU058290	KU058316	KU058347	KU058380
1513	KU058291	KU058317	KU058350	KU058381
1514	KU058292	KU058318	KU058351	KU058382
1515	KY745985	KY746089	KU058352	KY746295

[a] CPBF, Colecção Portuguesa de Bactérias Fitopatogénicas, Instituto Nacional de Investigação Agrária e Veterinária, Oeiras, Portugal.

	1	2	3	4	5	6	7	8	9	10	11	12
A	C-	CPBF 1480	CPBF 1570	CPBF 1505	CPBF 1567	CPBF 784	CPBF 796	CPBF 100	CPBF 237	CPBF 1545	CPBF 530	C-
B	CPBF 800	CPBF 1561	CPBF 598	CPBF 427	CPBF 1532	CPBF 1553	CPBF 122	CPBF 1530	CPBF 1515	CPBF 1586	CPBF 1508	CPBF 1514
C	CPBF 1522	CPBF 105	CPBF 98	CPBF 108	CPBF 1488	CPBF 426	CPBF 606	CPBF 268	CPBF 78	CPBF 75	CPBF 424	CPBF 367
D	C+	CPBF 414	CFBP 1022	CFBP 176	LMG 747							C+

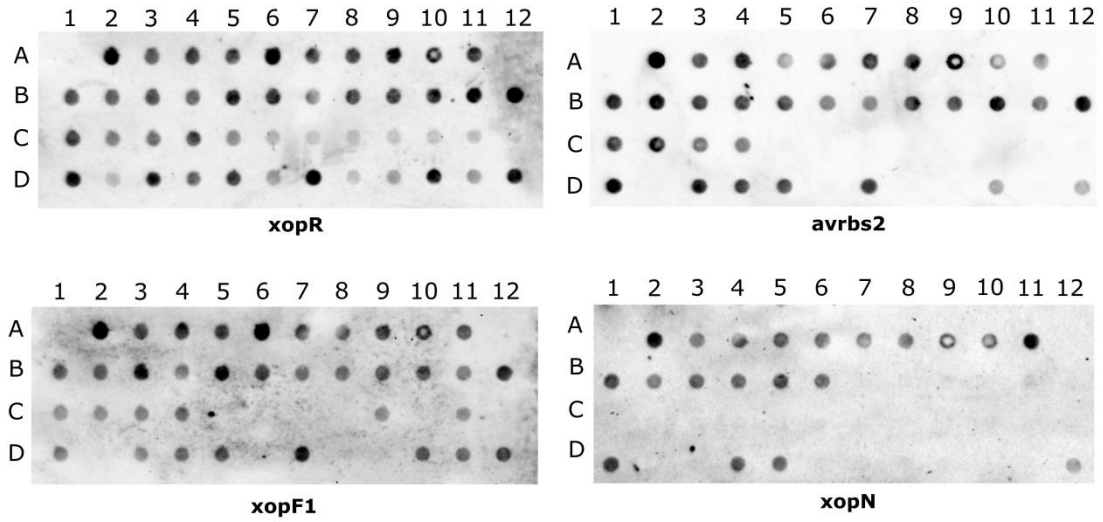

Figure IV.S1. Dot blot hybridization results using type three effector genes probes. Each probe (*xopR* 303bp, *avrbs2* 850bp, *xopF1* 779bp and *xopN* 864bp) were obtained from purified amplicon of PCR using *Xanthomonas arboricola* pv. *juglandis* strain LMG 751. Dot blot assays were performed for selected isolates grouped in MLSA clusters I to VI and for all isolates grouped in MLSA clusters VII to X (**Table IV.1** and **Figure IV.1**). The strain LMG 751 and Tris-EDTA buffer were used as positive (C+) and negative (C-) controls, respectively. D6, D7, D8, D9, D10 and D11 positions correspond to results found for isolates not included in the present work

CHAPTER V

Whole-genome comparison of distinct *Xanthomonas* lineages isolated from a single walnut tree host

Chapter V include the following publications:

Fernandes C, Blom J, Pothier JF, Tavares F. 2018. High-quality draft genome sequence of *Xanthomonas* sp. strain CPBF 424, a walnut pathogenic strain with atypical features. Microbiology Resource Announcements 7 (15): e00921-18.

Fernandes C, Blom J, Pothier JF, Tavares F. 2018. High-quality draft genome sequence of *Xanthomonas arboricola* pv. *juglandis* CPBF 1521, isolated from leaves of a symptomatic walnut tree in Portugal without a past of phytosanitary treatment. Microbiology Resource Announcements 7 (16): e00887-18.

Fernandes C, Martins L, Blom J, Pothier JF, Tavares F. 2019. Comparative genomic of *Xanthomonas* strains isolated from a single walnut tree host. In preparation.

Part of the results from this chapter were presented at:

Fernandes C, Pothier JF, Tavares F. 2018. Whole-genome sequencing of distinct *Xanthomonas arboricola* lineages isolated from a single walnut tree host. 6th *Xanthomonas* Genomics Conference. 2nd Annual EuroXanth Conference, Leopoldina, Halle (Saale), Germany, July 18-21, 2018.

Fernandes C, Martins L, Blom J, Pothier JF, Tavares F. 2018. Insights into sympatric lifestyle of walnut *Xanthomonas* isolates through comparative genomics. TiBE 2018 – Host Parasite Interection. CIBIO-InBIO, Campus de Vairão, December 5-7, 2018.

High-quality draft genome sequence of *Xanthomonas arboricola* pathovar *juglandis* CPBF 1521, isolated from leaves of a symptomatic walnut tree in Portugal without a past of phytosanitary treatment

Abstract

Here, we report the draft genome sequence of *Xanthomonas arboricola* pathovar *juglandis* CPBF 1521, isolated from symptomatic leaves of an ornamental walnut in a public site in Portugal without any record of phytosanitary treatment. This isolate may constitute a genomic reference of a wild-type strain in comparative genomics studies.

Announcement

Xanthomonas arboricola pathovar *juglandis* (Gammaproteobacteria class, Xanthomonadales order, Xanthomonadaceae family) is a threatening and important pathogen of the principal commercial nut trees Persian walnut and English walnut (*Juglans regia* L.) (Frutos 2010, Lamichhane 2014). Diseases caused by *X. arboricola* pv. *juglandis* have been demonstrated by the development of several symptoms, namely the presence of necrotic lesions on leaves and fruits, the presence of external apical necrosis near the blossom end evolving into fruit necrosis, and the presence of vertical cankers, brown to black exudates, and distortions on trunks (Smith et al. 2012, Moragrega and Ozaktan 2010, Hajri et al. 2010). Not surprisingly, *X. arboricola* pv. *juglandis* is responsible for increasing losses in walnut production resulting in a negative economic impact for walnut crop regions in many countries worldwide (Frutos 2010, Moragrega et al. 2011, Lamichhane 2014).

The present announcement reports the whole-genome sequence of a *X. arboricola* pv. *juglandis* strain, CPBF 1521, isolated in October 2014 from the leaves of an ornamental *J. regia* specimen in a public site in Loures, Portugal, showing typical symptoms of walnut bacterial blight, and for which no phytosanitary treatments were applied. This set of features suggests that this strain has not been exposed to selective pressures caused by phytosanitary treatments, such as copper-based compound sprays, making this genomic data set particularly interesting for comparative genomics studies.

X. arboricola pv. *juglandis* CPBF 1521 was obtained from infected leaf samples as previously described (Fernandes et al. 2017) and was grown on M2 medium (yeast extract, 2 g liter^{-1}; Bacto peptone, 5 g liter^{-1}; NaCl, 5 g liter^{-1}; KH2PO4, 0.45 g liter^{-1}; and Na_2HPO_4 $12H_2O$,

2.39 g liter^{-1}) at 28°C for 48 h with shaking (100 rpm). The EZNA bacterial DNA purification kit (Omega Bio-Tek, Norcross, GA) was used for DNA extraction. Standard genomic library preparation and sequencing was carried out with at the GATC BiotechAG (Konstanz, Germany) using an Illumina HiSeq platform with 2 × 150-bp paired-end reads. Raw sequence data with approximately 10,113,730 reads were assembled *de novo* using MIRA version 4.0 (Chevreux et al. 1999) in accurate mode with standard settings. The set of contigs obtained was reassembled using SeqMan Pro from the Lasergene genomics package version 12.1.0 (DNAStar, Madison, WI) with Pro Assembler parameters and reads were mapped using SeqMan NGen with standard settings to check for inconsistencies. Contigs were ordered according to the related reference genome of *X. arboricola* pv. *juglandis* (strain CFBP 2528; GenBank accession number NZ_JZEF01000000) (Cesbron et al. 2015) using the Move Contigs function in Mauve version 20150226 build 10 (Darling et al. 2004, 2010). Genome annotation was performed with Prokka version 1.12 (Seemann 2014) based on de novo discovery of genes using a *Xanthomonas* genus database.

The *X. arboricola* pv. *juglandis* CPBF 1521 genome contains 56 contigs with an N50 value of 173,159 bp, has a total size of 5,194,740 bp, and a G+C content of 65.41%. A total of 4,494 coding sequences (CDS) with 4 rRNAs and 53 tRNAs were found using Prokka. Initial comparative analysis using EDGAR version 2.0 (Blom et al. 2016) identified 3,984 CDS shared between CPBF 1521 and *X. arboricola* pv. *juglandis* CFBP 2528 that was used as a reference genome. Moreover, a high average nucleotide identity value (98%) was obtained between these two strains. This newly sequenced genome expands the existing genomic data for *X. arboricola* pv. *juglandis* (Cesbron et al. 2015, Higuera et al. 2015, Pereira et al. 2015, Fu et al. 2018). Further genomic comparison studies are being conducted to identify unique genomic features of *X. arboricola* pv. *juglandis* CPBF 1521.

Data availability

This whole-genome shotgun project has been deposited in European Nucleotide Archive (ENA) under the BioProject accession number PRJEB27248 and in GenBank under the accession number UIHD00000000. The version described in this paper is the first version, UIHD01000000.

Acknowledgments

This work was cofinanced by the European Structural and Investment Funds (ESIFs) through the Operational Competitiveness and Internationalization Program, COMPETE 2020,

and by national funds through Fundação para a Ciência e a Tecnologia (FCT), within the framework of the project EVOXANT (reference numbers PTDC/BIA-EVF/3635/2014 and POCI-01-0145-FEDER-016600). This article is based on work from European Cooperation in Science and Technology (COST) action CA16107 EuroXanth, supported by COST. Camila Fernandes was supported by a fellowship (SFRH/BD/95913/2013) from FCT and a Short Term Scientific Mission (STSM) from COST action CA16107. The EDGAR platform is funded by BMBF grant number FKZ 031A533. We thank Leonor Cruz for lab support.

References

Blom J, Kreis J, Spänig S, Juhre T, Bertelli C, Ernst C, Goesmann A. 2016. EDGAR 2.0: an enhanced software platform for comparative gene content analyses. *Nucleic Acids Research* **44**, W22-W28.

Cesbron S, Briand M, Essakhi S, Gironde S, Boureau T, Manceau C, Fischer-Le Saux M, Jacques MA. 2015. Comparative genomics of pathogenic and nonpathogenic strains of *Xanthomonas arboricola* unveil molecular and evolutionary events linked to pathoadaptation. *Frontiers in Plant Science* **6**,1126.

Chevreux B, Wetter T, Suhai S. 1999. Genome sequence assembly using trace signals and additional sequence information. Computer science and biology. Proceedings of the German Conference on Bioinformatics, GCB 99, Hannover, Germany, 45–56.

Darling AC, Mau B, Blattner FR, Perna NT. 2004. Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Research* **14**, 1394-1403.

Darling AE, Mau B, Perna NT. 2010. Progressive Mauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS One* **5**, e11147.

Fernandes C, Albuquerque P, Sousa R, Cruz L, Tavares F. 2017. Multiple DNA markers for identification of *Xanthomonas arboricola* pv. *juglandis* isolates and its direct detection in plant samples. *Plant Disease* **101**, 858-865.

Frutos D. 2010. Bacterial diseases of walnut and hazelnut and genetic resources. *Journal of Plant Pathology* **92**, S79-S85.

Fu B, Chen Q, Wei M, Zhu J, Zou L, Li G, Wang L. 2018. Complete genome sequence of *Xanthomonas arboricola* pv. *juglandis* strain DW3F3, isolated from a *Juglans regia* L. bacterial blighted fruitlet. *Genome Announcements* **6**, e00023-18.

Higuera G, González-Escalona N, Véliz C, Vera F, Romero J. 2015. Draft genome sequences of four *Xanthomonas arboricola* pv. *juglandis* strains associated with walnut blight in Chile. *Genome Announcements* **3**, e01160-15.

Hajri A, Meyer D, Delort F, Guillaumes J, Brin C, Manceau C. 2010. Identification of a genetic lineage within *Xanthomonas arboricola* pv. *juglandis* as the causal agent of vertical oozing canker of Persian (English) walnut in France. *Plant Pathology* **59**, 1014-1022.

Lamichhane JR. 2014. *Xanthomonas arboricola* diseases of stone fruit, almond, and walnut trees: progress toward understanding and management. *Plant Disease* **98**, 1600-1610.

Moragrega C, Matias J, Aleta N, Montesinos E, Rovira M. 2011. Apical necrosis and premature drop of Persian (English) walnut fruit caused by *Xanthomonas arboricola* pv. *juglandis*. *Plant Disease* **95**,1565-1570.

Moragrega C, Ozaktan H. 2010. Apical necrosis of Persian (English) walnut (*Juglans regia*): an update. *Journal of Plant Pathology* **92**, S67-S71.

Pereira UP, Gouran H, Nascimento R, Adaskaveg JE, Goulart LR, Dandekar AM. 2015. Complete genome sequence of *Xanthomonas arboricola* pv. *juglandis* 417, a copper-resistant strain isolated from *Juglans regia* L. *Genome Announcements* **3**, e01126-e01115.

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

Smith RE, Smith CO, Ramsey HJ. 1912. Walnut culture in California: walnut blight. California Agricultural Experiments Station Publications, Berkeley, CA.

High-quality draft genome sequence of *Xanthomonas* sp. strain CPBF 424, a walnut-pathogenic strain with atypical features

Abstract

We report here the draft genome sequence of *Xanthomonas* sp. strain CPBF 424, isolated from a diseased walnut tree. Multilocus sequence analysis showed that this walnut-pathogenic isolate is located between the nonpathogenic *X. arboricola* and *X. prunicola* clusters. These features make this strain a promising reference to use to disclose new genetic determinants of pathogenesis.

Announcement

The *Xanthomonas arboricola* species complex includes numerous phytopathogenic bacteria comprising different pathovars capable of infecting a wide range of plants (Vauterin et al. 1995, Janse et al. 2001, Fischer-Le Saux et al. 2015) and causing severe disease symptoms and serious economic losses in important crops (Lamichhane 2014). Recently, particular attention has been given to *X. arboricola*-related strains shown to be phylogenetically distinct from pathogenic *X. arboricola* pathovar strains (Essakhi et al. 2015, Jacques et al. 2016, Garita-Cambronero et al. 2017, López et al. 2018). *Xanthomonas* sp. strain CPBF 424 was isolated in April 2016 from asymptomatic dormant buds of a diseased walnut tree in Loures, Portugal, with common symptoms of walnut bacterial blight. Multilocus sequence analysis (MLSA) of the concatenated partial sequences of the atpD (750 bp), dnaK (759 bp), efp (339 bp), fyuA (684 bp), glnA (675 bp), gyrB (735 bp), and rpoD (586 bp) genes confirmed the strain's identity as *Xanthomonas* sp., revealing that strain CPBF 424 is located between the nonpathogenic *X. arboricola* and *X. prunicola* clusters and diverges from *Xanthomonas arboricola* pv. *juglandis* strains, i.e., walnut-pathogenic bacteria, and from other *X. arboricola* pathovars (Fernandes et al. 2018a). Pathogenicity tests on walnut plantlets further showed that CPBF 424 is pathogenic to walnut trees (Albuquerque et al. 2017, Fernandes et al. 2018b), making this strain particularly appealing to provide new insights into xanthomonad pathoadaptations.

Here, we make available the whole-genome sequence of *Xanthomonas* sp. strain CPBF 424.

Xanthomonas sp. strain CPBF 424 was grown on bacterial culture medium M2 (yeast extract, 2 g liter^{-1}; Bacto peptone, 5 g liter^{-1}; NaCl, 5 g liter^{-1}; KH_2PO_4, 0.45 g liter^{-1}; Na_2HPO_4 $12H_2O$, 2.39 g liter^{-1}) at 28°C and 100 rpm for 48 h. DNA was extracted for sequencing using the EZNA bacterial DNA purification kit (Omega Bio-tek, Norcross, GA). Genomic library preparation and genome sequencing were outsourced to GATC Biotech, AG (Konstanz, Germany) and conducted using an Illumina HiSeq platform with 2×150-bp paired-end reads, which resulted in 12,672,550 reads of raw sequence data with a sequencing coverage of 776×. De novo genome assembly was obtained with MIRA version 4.0 (Chevreux et al. 1999) using standard settings in accurate mode. This was followed by contig reassembly using SeqMan Pro from the Lasergene genomics package version 12.1.0 (DNAStar, Madison, WI) with Pro assembler parameters and read mapping using SeqMan NGen with standard settings to check for inconsistencies (overlapping contig extremities with no or low coverage with paired read inconsistencies). A total of five irregularities were found on which contigs were broken open as per the initial de novo assembly. Contigs were ordered using the Move Contigs function in Mauve 20150226 version 10 (Darling et al. 2004, 2010) according to the genome of *X. arboricola* pv. *juglandis* CFBP 2528 (GenBank accession number NZ_JZEF00000000) (Cesbron et al. 2015). Automatic genome annotation was performed with a *Xanthomonas* genus database using the Prokka software tool version 1.12 (Seemann 2014).

The *Xanthomonas* sp. CPBF 424 genome had a total size of 4,896,146 bp and a G+C content of 65.89% represented by 10 contigs with an N50 value of 1,029,447 bp. The genome of CPBF 424 is estimated to be composed of 4,143 coding sequences (CDS), including 58 tRNAs and 4 rRNAs. Preliminary analysis with the EDGAR version 2.0 platform (Blom et al. 2016) allowed us to detect 3,502 coding sequences that are shared between CPBF 424 and *X. arboricola* pv. *juglandis* CFBP 2528, which was used as the reference genome.

The whole-genome sequence of strain CPBF 424 may contribute to elucidating new walnut pathoadaptations within the genus *Xanthomonas*.

Data availability

This whole-genome shotgun project has been deposited in DDBJ/ENA/GenBank under the BioProject accession number PRJEB27248 (SRA accession number ERR2767968), and the sequence accession number is UIHB00000000. The version described in this paper is version UIHB01000000.

Acknowledgments

This work was cofinanced by the European Structural & Investment Funds (ESIFs) through the Operational Competitiveness and Internationalization Programme—COMPETE 2020 and by national funds through Fundação para a Ciência e Tecnologia (FCT) within the framework of the project EVOXANT (PTDC/BIA-EVF/3635/2014, POCI-01-0145-FEDER-016600). This article is based on work from European Cooperation in Science and Technology (COST) action CA16107 EuroXanth, supported by COST. C.F. was supported by a fellowship from FCT (SFRH/BD/95913/2013) and a Short Term Scientific Mission (STSM) from COST action CA16107. The EDGAR platform is funded by Federal Ministry of Education and Research (BMBF) grant FKZ 031A533. We thank Leonor Cruz for lab support.

References

Albuquerque P, Fernandes C, Cruz L, Tavares F. 2017. Diversity of *Xanthomonas arboricola* pv. *juglandis* in Portugal evoques a cosmopolitan dispersion, abstr FEMS7-1173, P-487. Abstr 7th Congress of European Microbiologists. Federation of European Microbiological Societies, Valencia, Spain.

Blom J, Kreis J, Spänig S, Juhre T, Bertelli C, Ernst C, Goesmann A. 2016. EDGAR 2.0: an enhanced software platform for comparative gene content analyses. *Nucleic Acids Research* **44**, W22–W28.

Cesbron S, Briand M, Essakhi S, Gironde S, Boureau T, Manceau C, Fischer-Le Saux M, Jacques MA. 2015. Comparative genomics of pathogenic and nonpathogenic strains of *Xanthomonas arboricola* unveil molecular and evolutionary events linked to pathoadaptation. *Frontiers in Plant Science* **6**, 1126.

Chevreux B, Wetter T, Suhai S. 1999. Genome sequence assembly using trace signals and additional sequence information. Computer science and biology. Proceedings of the German Conference on Bioinformatics, GCB 99, Hannover, Germany, 45–56.

Darling AC, Mau B, Blattner FR, Perna NT. 2004. Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Research* **14**, 1394-1403.

Darling AE, Mau B, Perna NT. 2010. progressiveMauve: Multiple genome alignment with gene gain, loss and rearrangement. *PLoS One* **5**, e11147.

Essakhi S, Cesbron S, Fischer-Le Saux M, Bonneau S, Jacques MA, Manceau C. 2015. Phylogenetic and variable-number tandem-repeat analyses identify nonpathogenic *Xanthomonas arboricola* lineages lacking the canonical type III secretion system. *Applied and Environmental Microbiology* **81**, 5395-5410.

Fernandes C, Pothier JF, Tavares F. 2018a. Whole-genome sequencing of distinct *Xanthomonas arboricola* lineages isolated from a single walnut tree host, abstr poster 9, p 88. Abstr 6th *Xanthomonas* Genomics Conference and 2nd Annual EuroXanth Conf, Halle (Saale), Germany.

Fernandes C, Albuquerque P, Cruz L, Tavares F. 2018b. Genotyping and epidemiological metadata provides new insights into population structure of *Xanthomonas* isolated from walnut trees. bioRxiv:397703.

Fischer-Le Saux M, Bonneau S, Essakhi S, Manceau C, Jacques MA. 2015. Aggressive emerging pathovars of *Xanthomonas arboricola* represent widespread epidemic clones distinct from poorly pathogenic strains, as revealed by multilocus sequence typing. *Applied and Environmental Microbiology* **81**, 4651-4668.

Garita-Cambronero J, Palacio-Bielsa A, López MM, Cubero J. 2017. Pangenomic analysis permits differentiation of virulent and non-virulent strains of *Xanthomonas arboricola* that cohabit *Prunus* spp. and elucidate bacterial virulence factors. *Frontiers in Microbiology* **8**, 573.

Jacques M-A, Arlat M, Boulanger A, Boureau T, Carrère S, Cesbron S, Chen NWG, Cociancich S, Darrasse A, Denancé N, Fischer-Le Saux M, Gagnevin L, Koebnik R, Lauber E, Noël LD, Pieretti I, Portier P, Pruvost O, Rieux A, Robène I, Royer M, Szurek B, Verdier V, Vernière C. 2016. Using ecology, physiology, and genomics to understand host specificity in *Xanthomonas*. *Annual Review of Phytopathology* **54**,163-187.

Janse J, Rossi M, Gorkink R, Derks J, Swings J, Janssens D, Scortichini M. 2001. Bacterial leaf blight of strawberry (Fragaria (x) ananassa) caused by a pathovar of *Xanthomonas arboricola*, not similar to *Xanthomonas fragariae* Kennedy & King. Description of the causal organism as *Xanthomonas arboricola* pv. *fragariae* (pv. nov., comb. nov.). *Plant Pathology* **50**, 653-665.

Lamichhane JR. 2014. *Xanthomonas arboricola* diseases of stone fruit, almond, and walnut trees: Progress toward understanding and management. *Plant Disease* **98**, 1600-1610.

López MM, Lopez-Soriano P, Garita-Cambronero J, Beltrán C, Taghouti G, Portier P, Cubero J, Fischer-Le Saux M, Marco-Noales E. 2018. *Xanthomonas prunicola* sp. nov., a novel pathogen that affects nectarine (*Prunus persica* var. *nectarina*) trees. *International Journal of Systematic and Evolution Microbiology* **68**, 1857.

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

Vauterin L, Hoste B, Kersters K, Swings J. 1995. Reclassification of *Xanthomonas*. *International Journal of Systematic Bacteriology* **45**, 472-489.

Comparative genomic of *Xanthomonas* strains isolated from a single walnut tree host

Abstract

Xanthomonas arboricola pv. *juglandis* (*Xaj*) cause serious bacterial diseases on walnut trees worldwide, affecting negatively walnut cultivation and production. This phytopathogen has a particular genetic diversity within the species *X. arboricola* and the occurrence of distinct *Xanthomonas* lineages within the same walnut tree specimen was reported by independent studies. More recently, it was shown that these consortium of walnut associated *Xanthomonas* include both pathogenic and nonpathogenic strains and non-*arboricola Xanthomonas* species. The ecological, evolutionary and pathogenicity implications of this co-colonization is still poorly understood. In order to unveil niche-specific adaptations, the genome of five *Xanthomonas* strains (CPBF 367, CPBF 424, CPBF 426, CPBF 427, and CPBF 1521) isolated from a single walnut tree in Loures (Portugal) were sequenced. Core genome phylogeny allowed separating these isolates in two distinct clusters, one grouping CPBF 427 and CPBF 1521 with *Xaj* strains, and another clustering together CPBF 367, CPBF 424 and CPBF 426. The average nucleotide identity (ANI) of genomes of the strains CPBF 427 and CPBF 1521 showed ANI values higher than 97% with *Xaj* strains, while CPBF 367, CPBF 424 and CPBF 426, although sharing an ANI>98% with each other, had a low ANI value (<90%) to other *Xanthomonas* species analysed, suggesting that these strains belong to a new *Xanthomonas* species. These two clusters of CPBF strains are herein designated as CPBF *Xaj* and CPBF *Xanthomonas* sp., respectively. General genomic features showed that CPBF 427 and CPBF 1521 possess a genome similar characteristics to other *Xaj* strains, regarding its size, CDSs and richness of IS elements, while CPBF 367, CPBF 424 and CPBF 426 strains have a reduced genome showing a difference of over 400 CDSs to *Xaj* genomes, and a considerable reduction of the number of IS elements comparatively to *Xaj* strains. Whole genome comparisons revealed remarkable genomic differences between the two distinct groups of CPBF strains, which could translate into different pathogenicity and virulence features. Ongoing genomics studies showed differences regarding the repertoire of type three secretion system (T3SS) and associated effectors genes (T3E) genes. While CPBF *Xaj* strains showed a complete T3SS and a set of T3E genes characteristic of typical *Xaj* pathogenic strains, as expected, CPBF 367 and CPBF 426 strains were deficient for T3SS and for most of T3E as shown by nonpathogenic *Xanthomonas*. Interestingly, strain CPBF 424, which belong to CPBF *Xanthomonas* sp. and was shown to be pathogenic in walnut,

have a complete T3SS and an enriched T3E pattern. Beyond the distinct profile of T3SS and T3E genes, differences between CPBF *Xaj* isolates and CPBF *Xanthomonas* sp. are extended to other genetic determinants, putatively linked to pathogenicity and virulence. In fact, when compared with *Xaj*, the CPBF *Xanthomonas* sp. strains hold an exclusive set of homologous encoding for chemotaxis related proteins, pectate lyase, pectinesterase, and XspN from T2SS, suggesting distinct pathoadaptations. Most interestingly, is the presence of genes encoding for a putatively functional type IV pilus in strain CPBF 424 which pathogenicity in walnut has been demonstrated. Altogether, the distinct genomic repertoire of *Xanthomonas* sp. CPBF isolates may be particularly useful to address emergence and evolution of pathogenicity in walnut-associated X*anthomonas*.

1. Introduction

Xanthomonas is a genus of gammaproteobacteria (Dowson 1939) which include numerous species acknowledged as important plant-associated bacteria with capacity to cause disease in a wide range of plant species, including important agricultural crops (Leyns et al. 1984, Swings and Civerolo 1993). The genus comprises 35 phytopathogenic species that are subdivided at the infrasubspecific level into distinct pathogenicity groups, also known as pathovars, according to their high degree of host specificity, disease symptoms and infection mechanisms (Young et al. 1978, Vauterin et al. 2000, Meyer and Bogdanove 2009, Ryan et al. 2011). Within this genus, *Xanthomonas arboricola* (Vauterin et al. 1995) is responsible for severe diseases in important stone fruits and nut trees (Lamichhane 2014). Pathogenic strains of X*. arboricola* belonging to pathovar *juglandis* (*Xaj*) (Pierce 1901) cause serious diseases on walnut, namely Walnut Bacterial Blight (WBB), Brown Apical Necrosis (BAN) and Vertical Oozing Canker (VOC), altogether responsible for major drops in walnut production and extended damages in walnut nurseries, leading to major economic losses (Smith 1912, Belisario et al. 2002, Hajri et al. 2010).

Comparatively to other *Xanthomonas arboricola* pathovars, *Xaj* is characterized by a larger genetic diversity recorded throughout geographically distinct walnut cultivation regions (Ivanovic et al. 2014, Kaluzna et al. 2014, Fischer-Le Saux et al. 2015, Giovanardi et al. 2015). Genotyping and population studies of walnut-associated xanthomonads allowed to identify *Xaj* pathogenic strains responsible for walnut bacterial diseases (Hajri et al. 2010) and nonpathogenic strains of *X. arboricola* shown to be asymptomatic on walnut (Essakhi et al. 2015). Based on a large body of evidence, carefully reviewed by Buttner and Bonas (2010), it is currently acknowledged that *Xanthomonas* pathogenesis is dependent on an array of pathogenicity and virulence factors. Starting with the importance of type three secretion system (T3SS) and type three effectors (T3E), but also including other type secretion systems, adhesins, extracellular polysaccharides (EPS), lipopolysaccharides (LPSs), and plant cell wall hydrolytic enzymes, among others, under the control of tight transcriptional and post-transcriptional regulatory mechanisms that allowed infective bacteria to adhere, invade the plant host tissues and to multiply and overcome the plant defence mechanisms (Buttner and Bonas 2010).

Comparative genomics of pathogenic and nonpathogenic strains of *Xanthomonas arboricola* has been instrumental to unveil genetic determinants of pathogenicity and virulence (Hajri et al. 2011, Cesbron et al. 2015, Garita-Cambronero et al. 2018) and to provide insights into evolutionary events linked to pathoadaptation, as emphasised by Cesbron et al. (2015). T3SS and T3E have been pointed out as major players of pathogenicity, the different repertoire of virulence-associated genes between pathogenic and nonpathogenic *X. arboricola* strains suggested that pathogenicity and virulence is determined by a complex network of genetic

determinants which is not fully understood. Furthermore, the abundance of IS elements in pathogenic *Xaj* strains, has been suggested to contribute for the genomic plasticity by homologous recombination, capable to respond to niche adaptations and ensure the evolutionary fitness of these pathogens (Cesbron et al. 2015).

The discovery of pathogenic and nonpathogenic walnut isolates of a likely new species of *Xanthomonas* addressed in this work, raises further questions regarding new pathoadaptations to walnut (Fernandes et al. 2018a, b). Furthermore, the fact that these CPBF *Xanthomonas* sp. strains were isolated from the same walnut host tree, together with characteristic *Xaj* strains (CPBF 427, CPBF 1521), suggests a sympatric lifestyle which may contribute to unveil genetic trade-offs related to pathogenicity and virulence of *Xanthomonas* in walnut. In this study we carried out a comprehensive comparative genomics study of five walnut-associated *Xanthomonas* strains, i.e. two *Xaj* strains (CPBF 417 and CPBF 1521) and three *Xanthomonas* sp. (CPBF 367, CPBF 424, CPBF 426), isolated from a single diseased walnut tree, to unveil genetic determinants of pathogenicity and virulence and disclose niche-specific adaptations, which may contribute to further understand the ecology and evolution of walnut-associated *Xanthomonas*.

2. Material and methods

2.1. *Xanthomonas* strains used in this study

Following an extensive survey to characterize the genetic diversity of xanthomonad populations obtained from walnut hosts, we isolated distinct *Xanthomonas* strains from 38 walnut trees, either from the same sampling occurrence (i.e. the same plant in the same date) or from the same walnut tree at different sampling dates (Fernandes et al. 2018a). From this collection of walnut-associated *Xanthomonas*, five isolates (CPBF 367, CPBF 424, CPBF 426, CPBF 427 and CPBF 1521) with distinct genotypes and obtained from a single walnut tree host located in the region of Loures, Portugal, were chosen for genome sequencing. CPBF 1521 was isolated in October 2014 from leaf spots whereas the other four strains (CPBF 367, CPBF 424, CPBF 426 and CPBF 427) were isolated in April 2016 from different asymptomatic buds. Based on MLSA analysis of *acnB*, *fyuA*, *gyrB*, *rpoD* genes (Young et al. 2008; Parkinson et al. 2007) and *Xaj* multiple DNA markers (Fernandes et al. 2017), CPBF 427 and CPBF 1521 were identified as *Xaj* strains, herein generally designated as CPBF *Xaj* strains, whereas CPBF 367, CPBF 424 and CPBF 426 were identified as belonging to a non-*arboricola Xanthomonas* species, herein generally designated as CPBF *Xanthomonas* sp. strains (Fernandes et al. 2018a). Pathogenicity tests carried out with *Xanthomonas* sp. strains CPBF 424 and CPBF 367, revealed two distinct

pathogenicity phenotypes. While CPBF 424 caused symptoms of WBB on walnut leaflets characterized by small necrotic spots surrounded by chlorotic halos, CPBF 367 was shown to be nonpathogenic on walnut (Fernandes et al. 2018a).

2.2. Bacterial DNA extraction, genome sequencing and annotation

Library preparation, genome sequencing and annotation were performed as previously described (Fernandes et al. 2018b, c). Bacteria were grown during 48 h on M2 medium, at 28°C with shaking at 100 rpm, and bacterial DNA extraction was carried out using the EZNA Bacterial DNA Purification kit (Omega Bio-Tek, Norcross, GA). Illumina Hi Seq sequencing system platform of 2x150 bp paired-end reads was used for whole genome sequencing. *De novo* genome assemblies were obtained with MIRA v. 4.0. (Chevreux et al. 1999) and automatic annotation was processed using Prokka v. 1.12 (Seeman 2014).

The high-quality draft genome sequence of *Xanthomonas* strains CPBF 367, CPBF 424, CPBF 426, CPBF 427 and CPBF 1521 were deposited at Genbank under the accession numbers UNRN00000000.1, UIHB00000000.1, UNRM00000000.1, UNRO00000000.1 and UIHD00000000.1, respectively.

2.3. Bacterial taxonomic classification according the Average Nucleotide Identity (ANI)

The Average Nucleotide Identity (ANI) based on MUMmer (ANIm) algorithm (Pritchard et al. 2016) was used to determine the genetic distance between each of the five bacterial genomes (CPBF 367, CPBF 424, CPBF 426, CPBF 427 and CPBF 1521) and 18 genomes of *X. arboricola,* including five distinct pathovars, and 29 genomes of *Xanthomonas* spp. available at NCBI genomes database (*X. arboricola* 3004, CITA 14, CITA 44, CFBP 7634; *X. arboricola* pv. *celebensis* NCPPB 1630, NCPPB 1832; *X. arboricola* pv. *corylina* NCCB 100457; *X. arboricola* pv. *fragariae* CFBP 6773, LMG 19145; *X. arboricola* pv. *juglandis* CFBP 2528, DW3F3, J303, NCPPB 1447, *Xaj* 417; *X. arboricola* pv. *pruni* CFBP 3894, IVIA 2626.1, MAFF311562, MAFF301420; *X. cassavae* CFBP 4642; *X. fragariae* FaP21, LMG 25863, NBC2815, PD885, PD5205; *X. gardneri* ATCC 19865, ICMP7383, JS749-3; *X. hortorum* B07-007, M081).

2.4. The core, accessory and strain-specific genome of *Xanthomonas* CPBF strains

The efficient database framework for comparative genome analyses using BLAST score Ratios (EDGAR v. 2.0, Blom et al. 2009, Blom et al. 2016) was used for comparative genomics of the five CPBF genomes used in this study.

A phylogenomics tree of 42 *Xanthomonas* genomes, including CPBF 367, CPBF 424, CPBF 426, CPBF 427 and CPBF1521, was generated by the Neighbour-Joining method based on 1228 core genes using the EDGAR platform. Briefly, the core genome was first calculated by a pairwise comparison criteria completely automated, with thresholds based on BLAST score ratio values (SRVs) (Lerat et al. 2003, Blom et al. 2009). Then, alignments of each core gene set were generated using MUSCLE (Edgar et al. 2004). The Neighbor-Joining algorithm, as implemented in the PHYLIP package, was used to construct the tree.

Pools of genes shared between the CPBF 367, CPBF 424, CPBF 426, CPBF 427 and CPBF 1521 genomes and strain-specific features were determined by a Venn diagram. This comparative analysis was also performed in EDGAR by pairwise comparison based on SRVs (Lerat et al. 2003, Blom et al. 2009).

2.5. Broad genome comparisons using BRIG

BLAST Ring Image Generator (BRIG, Alikhan et al. 2011) was used to visualize the similarities between the five *Xanthomonas* CPBF strains genomes, using *Xanthomonas arboricola* pv. *juglandis* CPBF 2528 as the reference genome. The circular BLAST comparison was determined with blastn (Altschul et al. 1990).

2.6. Classification of COGs functional categories using eggNOG

The database eggNOG 4.5 (Huerta-Cepas et al. 2016) of clusters of orthologous groups (COGs) was used for fast and automatic characterization of protein sequences from CPBF genomes using the eggNOG-mapper. COGs functional categories were obtained for all protein sequences of the CPBF 367, CPBF 424, CPBF 426, CPBF 427 and CPBF 1521 genomes.

2.7. Homologous of pathogenicity and virulence-associated proteins

A database was created with protein sequences previously reported to have an important involvement in the pathogenesis and virulence of *Xanthomonas* pathogens to the host plant. All genomic sequences of the five CPBF genomes were aligned using BLAST analysis (tblastn, Altschul et al. 1990) against the created database. To ensure a high confidence, the criteria to identify protein homologous in CPBF genomes were an e-value threshold of $1e^{-10}$, a sequence similarity with a query length similarity cut-off ≥40%, and a sequence identity with a cut-off ≥ 70%. The search for homologous was performed using as query sequences proteins of xanthan biosynthesis identified by Lee et al. (2005) and Vorholter et al. (2008); a list of proteins of the

flagellar system from *X. campestris* pv. *vesicatoria* 85-10 (Darrasse et al. 2013) and from *X. fragariae* LMG 25863 (AJRZ00000000.1, NCBI database), as well as proteins of the *rpf* gene cluster for regulation of pathogenicity factors (*X. campestris*, Crossman 2004; *X. fragariae* LMG 25863, AJRZ00000000.1, NCBI database). Homologous of chemotactic and methyl-accepting chemotaxis proteins, proteins involved in the biosynthesis of quorum sensing signals and non-fimbrial adhesins, were identified using as query the protein sequences previously used in a study with *X. arboricola* genomes (Garita-Cambronero et al. 2018). The presence or absence of homologous associated with components of the different secretion systems were also predicted, using as query sequences proteins of the type II secretion system (T2SS) (Filloux 2004, Szczesny et al. 2010) and related hemilcellulolytic, cellulolytic, pectolytic enzymes, lipases and proteases (Garita-Cambronero et al. 2018); proteins of the type III secretion system (T3SS) (Hajri et al. 2012; *X. fragariae* LMG 25863 AJRZ00000000.1, NCBI database) and related effectors (*Xanthomonas* Resource; White et al. 2009; Hajri et al. 2012, *X. fragariae* LMG25863, AJRZ00000000.1, NCBI database); proteins of the type IV secretion system and type IV pilus (Dunger et al. 2016, Garita-Cambronero et al. 2018); and proteins of the type VI secretion system (Shrivastava and Mande 2008).

3. Results

3.1. Taxonomic affiliation and phylogenetic analysis

The taxonomic affiliation of the five *Xanthomonas* walnut isolates (CPBF 367, CPBF 424, CPBF 426, CPBF 427 and CPBF 1521) was carried out through the analysis of the ANI with 18 genomes of *X. arboricola,* including five distinct pathovars and 29 genomes of *Xanthomonas* spp. strains (**Figure V.1**). Strains CPBF 427 and CPBF 1521 showed ANI values higher than 95% in comparison with reference genomes of *X. arboricola* strains, and higher than 97% to *X. arboricola* genomes belonging to pathovar *juglandis* (**Figure V.1A**). Conversely, strains CPBF 367, CPBF 424 and CPBF 426, which revealed an ANI ≥ 98% between each other, showed an ANI ≤ 94% when compared to *X. arboricola* strains (**Figure V.1A**), and an ANI ≤ 90% with other *Xanthomonas* species such as *X. cassavae*, *X. gardneri*, *X. hortorum* and *X. fragariae* (**Figure V.1B**).

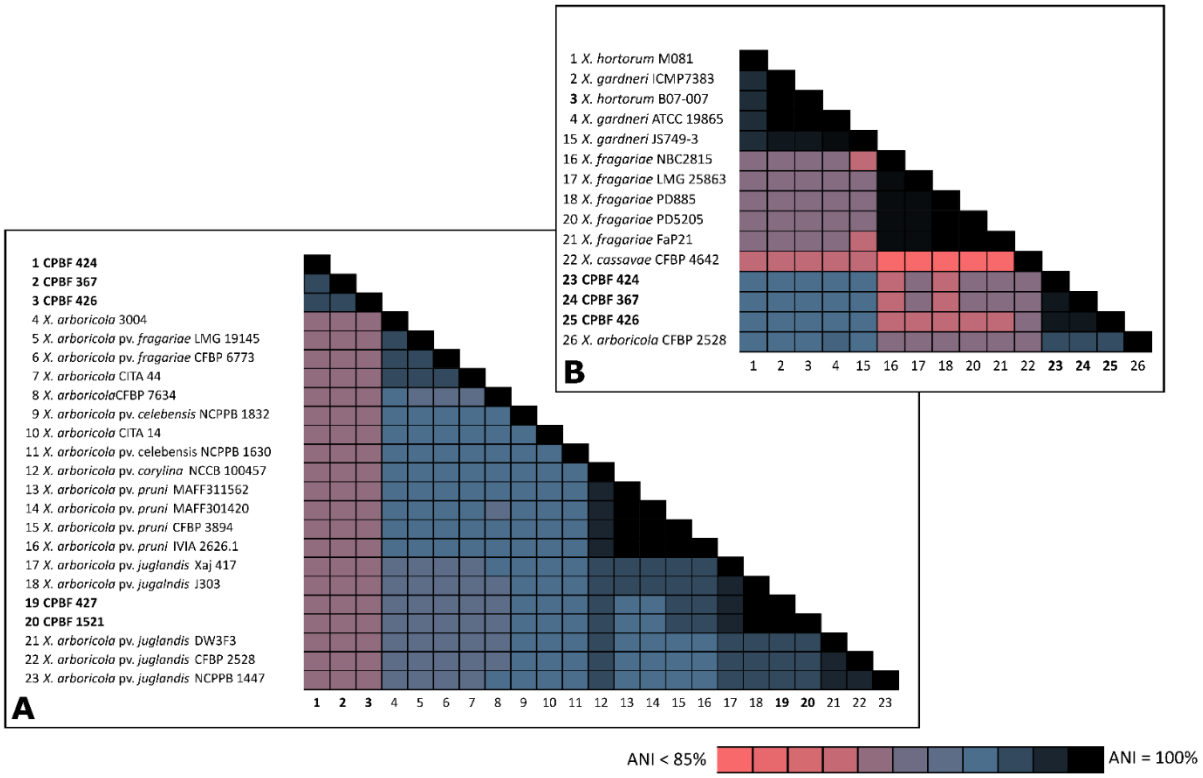

Figure V.1. Average nucleotide identity determined for 34 *Xathomonas* genomes, including the five CPBF genomes of this study (CPBF 367, CPBF 424, CPBD 426, CPBF 427 and CPBF 1521). Percent ANI values were calculated through the multiple alignment of shared genes using pyANI v. 0.2.0 (Pritchard et al., 2016) and are represented followed a colour gradient. ANI percentages higher than 95% are considered to belong to the same species (Konstantinidis and Tiedje 2005). **A.** ANI values for the five CPBF strains and different pathovars of *X. arboricola*. **B.** ANI values for the three CPBF 367, CPBF424 and CPBF 426 strains and CPBF 426 strains belonging to five *Xanthomonas* species (*X. hortorum*, *X. gardneri*, *X. fragariae*, *X. cassavae* and *X. arboricola*).

These results are aligned with the core genome phylogenomics tree, where these five strains are separated in two clusters (**Figure V.2**). Strains CPBF 427 and CPBF 1521 clustered together with *Xaj* strains, whereas CPBF 367, CPBF 424 and CPBF 426 form a distinct group, separated from the other *Xanthomonas* species.

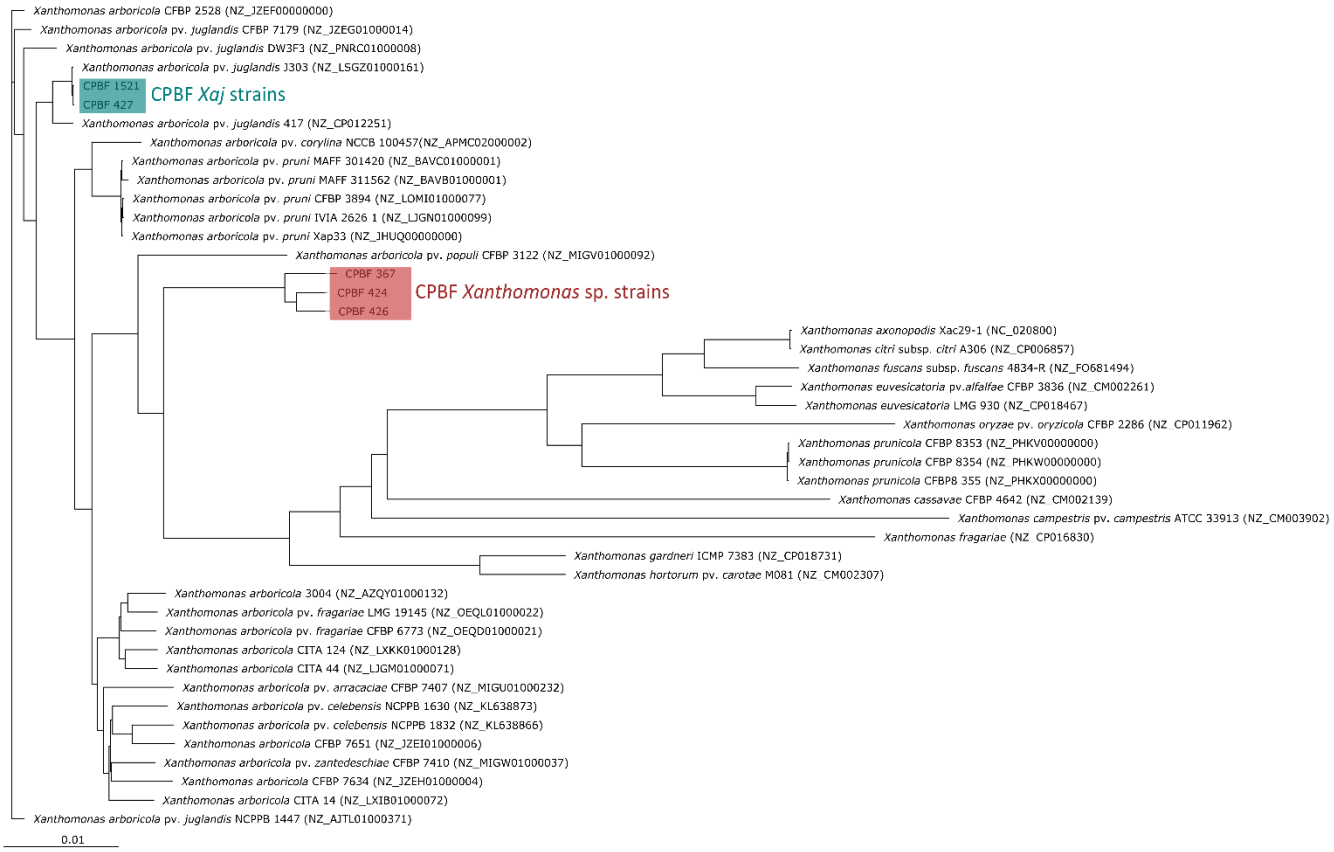

Figure V.2. Phylogenomics tree of 44 *Xanthomonas* spp. genomes, including the five CPBF strains. The tree based on concatenated sequences of the core genome of each *Xanthomonas* spp. strains was obtained using EDGAR. Highlighted with different colours are the five CPBF strains, clustered separately. Blue – group of CPBF *Xaj* strains. Red – group of CPBF *Xanthomonas* sp. strains.

3.2. General genomics features of *Xaj* and *Xanthomonas* sp. strains

The genomes of CPBF 427 and CPBF 1521 had a size of approximately 5.19 MB, whereas strains CPBF 367, CPBF 424 and CPBF 426 have slightly smaller genomes, i.e. less than 5.0 MB. A similar G+C content of 65% was obtained for all the five genomes.

Major differences were observed for the genomic features related to genetic mobile elements. In fact, while a total of 61 and 60 IS elements were annotated in CPBF 427 and CPBF 1521 genomes, respectively, the number of IS elements present in CPBF 367, CPBF 424 and CPBF 426 genomes were five, ten, and seven, respectively (**Table V.1**). Representatives of the IS3 and IS200 families were identified in all genomes, however in CPBF 427 and CPBF 1521 genomes, many of the members of the IS3 family were present in multicopies. Furthermore, CPBF 427 and CPBF 1521 genomes also include IS elements belonging to other families, namely the IS4, IS30, IS110 and IS1595. Notably, CPBF 426 strain harboured other IS elements from IS701 and IS5 families.

Table V.1. General genomic features of the five CPBF strains analysed in this study.

General features	*Xanthomonas* sp. strains			*Xaj* strains	
	CPBF 367	CPBF 424	CPBF 426	CPBF 427	CPBF 1521
Genome size (bp)	4,956,382	4,896,146	4,894,012	5,190,560	5,194,740
No. contigs	22	10	11	57	56
N50 (bp)	687,415	1,029,447	730,188	178,455	173,159
G+C (%)	65.81	65.89	65.87	65.41	65.41
Plasmids	0	0	0	0	0
No. Total CDS	4094	4034	4049	4354	4368
IS elements	5	10	7	61	60
CDS shared with CFBP 2528	3476	3502	3500	3983	3985
tmRNA	1	1	1	1	1
tRNA	56	58	53	53	53
rRNA	4	4	4	4	4
Misc_RNA	48	46	45	71	68
Assembly no.	GCA_900537245	GCA_900476395	GCA_900537265	GCA_900537235	GCA_900476315
WGS accession	UNRN00000000.1	UIHB00000000.1	UNRM00000000.1	UNRO00000000.1	UIHD00000000.1

The total number of protein-coding sequences (CDSs) of the five CPBF genomes studied was proportional to the size of the genomes and revealed that CPBF 427 and CPBF 1521 genomes share more CDSs with the genome of *Xaj* strain CFBP 2528, chosen as reference, than with the genomes of strains CPBF 367, CPBF 424 and CPBF 426 (**Table V.1**). Additionally, a BLAST Ring Image Generator (BRIG) analysis allowed highlighting major differences between the five CPBF genomes and *Xaj* strain (**Figure V.3**). The circular whole-genomes alignment revealed that the absence of some loci, comparatively with the reference *Xaj* strain CFBP 2528 is visibly more pronounced in CPBF *Xanthomonas* sp. strains (CPBF 367, CPBF 424 and CPBF 426), than in CPBF *Xaj* strains (CPBF 427 and CPBF 1521), further supporting a higher similarity between these two CPBF strains and the reference genome. In general, the circular whole genome alignments revealed also a similar pattern shared by the strains CPBF 367, CPBF 424 and CPBF 426. The genomic regions associated to the asymmetric genomic contents observed between *Xaj* CFBP 2528 and CPBF strains is coupled with considerable shifts in CG content and CG skew (**Figure V.3**).

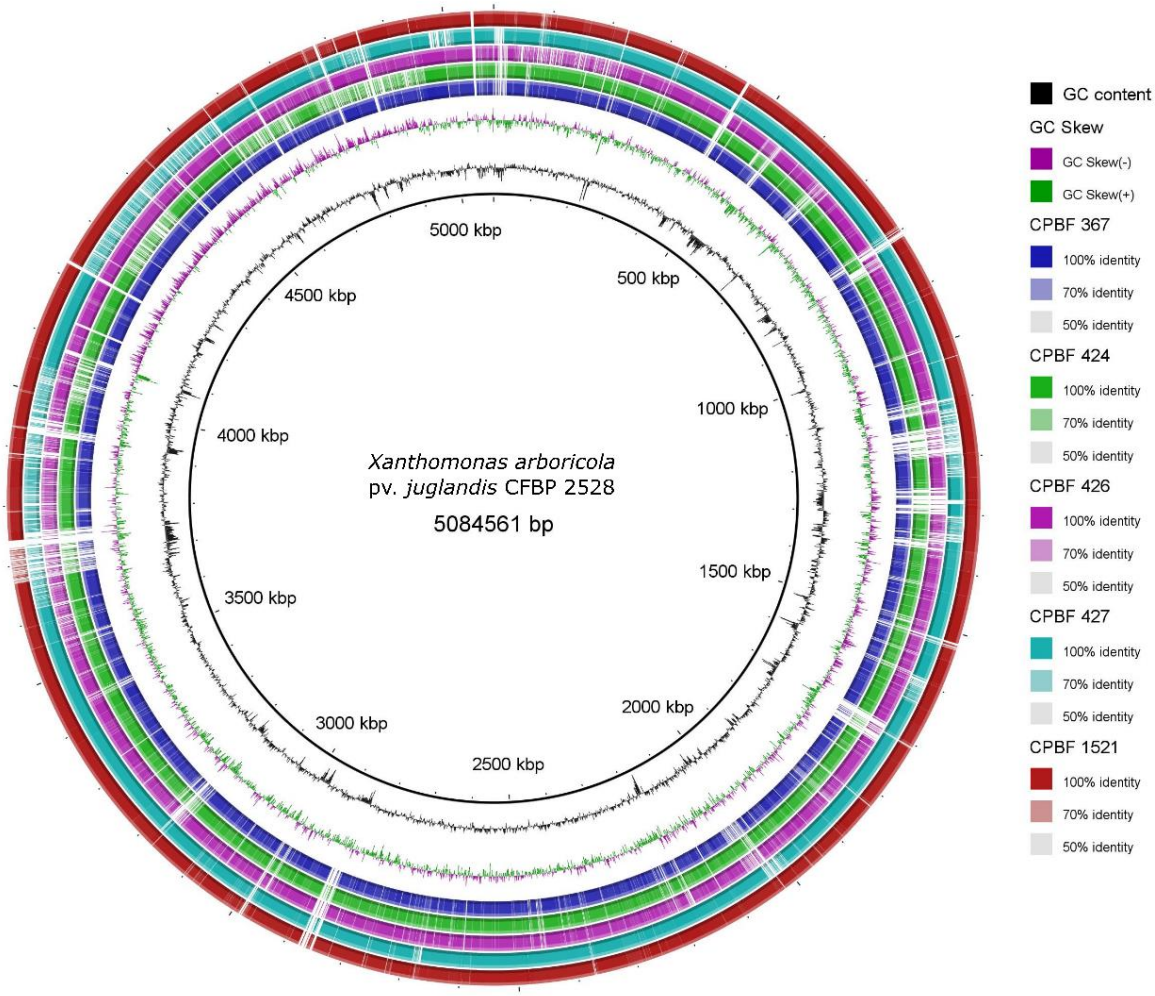

Figure V.3. Multiple genome comparison of CPBF genomes with the reference genome sequence from *Xaj* CFBP 2528 obtained by BRIG blast analysis. Whole genome alignments were performed using blastn. Each outer ring represents one of the genomes, CPBF 1521 (red ring), CPBF 427 (light blue ring), CPBF 426 (purple ring), CPBF 424 (green ring), CPBF 367 (blue ring). The second and third innermost rings shown GC content and GC skew, respectively. The inner black ring corresponds to *Xaj* CFBP 2528 genome used as reference. BLAST identity of 50%, 70% and 100% are represented in each CPBF genome ring according a colour gradient scale showed on the right of the image. Genomic regions of CFBP 2528 with no blastn matches appeared as blank spaces across CPBF genome rings.

3.3. Pangenome analysis

To disclose the gene pool of xanthomonads found in a single walnut tree host, we estimated the total number of CDSs corresponding to non-redundant genes shared between the five CPBF strains studied, generally known as the pangenome. The results showed that the pangenome of CPBF strains is of 5412 CDSs, from which 63% (i.e. 3421 CDSs) belong to the core genome, i.e. the set of genes common to the five CPBF strains (**Figure V.4**). The remaining 2091 CDSs, which correspond to the accessory genome, represent the differential gene content between these strains, which include all strain-specific CDSs and genes shared by two or more CPBF strains. Consistently with the above-mentioned results, CPBF *Xaj* strains (CPBF 427 and CPBF 1521)

share 4131 CDSs of the pangenome (i.e. 76%), while CPBF *Xanthomonas* sp. strains (CPBF 367, CPBF 424 and CPBF 426) share 3618 CDSs (i.e. 67%) of the pangenome. Moreover, CPBF *Xaj* strains share exclusively 710 CDSs, while CPBF *Xanthomonas* sp. strains share exclusively 197 CDSs. Taking the intersection from both CPBF *Xaj* strains with the walnut pathogenic CPBF *Xanthomonas* sp. (strain CPBF 424), it was possible to identify 50 CDSs belonging to the accessory genome. Regarding the strain-specific genomes, from a total of 873 CDSs, the CPBF *Xanthomonas* sp. strains showed the majority of strain-specific genes, with 320 unique CDSs for CPBF 367, 274 unique CDSs for CPBF 426, and 220 unique CDSs for CPBF 424. Most of these predicted singletons were assigned as hypothetical proteins. The remaining 59 strain-specific CDSs were found in the genomes of *Xaj* strains CPBF 1521 (39 unique CDSs) and CPBF 427 (20 unique CDSs).

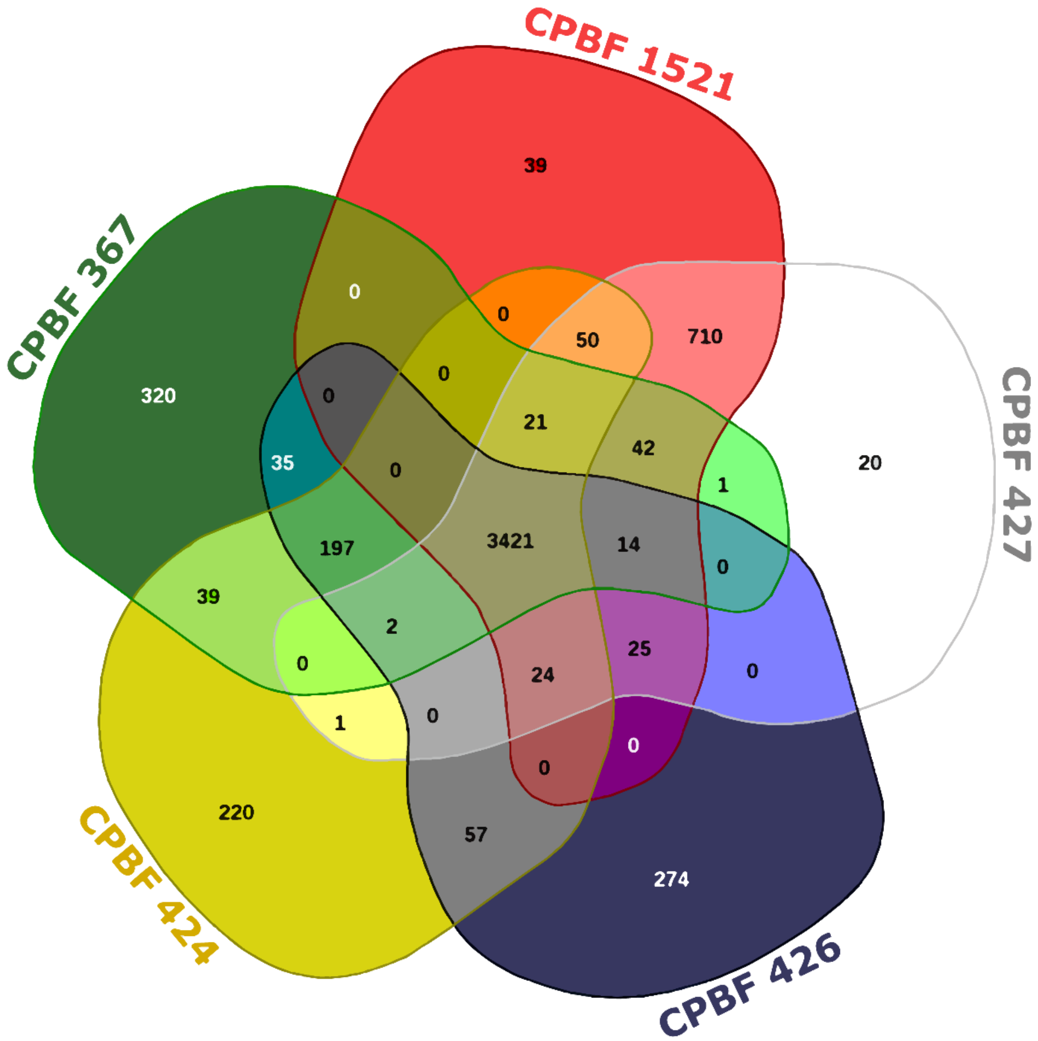

Figure V.4. Pangenome of CPBF genomes represented through a Venn diagram. The core genome is given by the interception of the five CPBF strains and corresponds to the number of orthologous CDSs shared by the five genomes, (3421 CDSs). Strain specific CDSs are also represented in the periphery of the diagram (320 for CPBF 367, 274 for CPBF 426, 220 for CPBF 424, 20 for CPBF 427 and 39 for CPBF 1521). The remaining combinations represent the number of orthologous shared between two to four CPBF genomes. Venn diagram was created using the EDGAR platform (EDGAR v. 2.0, Blom et al. 2009, Blom et al. 2016).

3.4. COGs functional categories

Functional analysis based in clusters of orthologous groups (COGs) were assigned for the set of CDSs shared by the group of the CPBF *Xaj* strains (CPBF 427 and CPBF 1521) and CPBF *Xanthomonas* sp. strains (CPBF 367, CPBF 424 and CPBF 426) (**Table V.2**). About 2531 out of 4131 CDSs shared by CPBF *Xaj* strains (i.e. 61%) and 2369 out of 3618 CDSs shared by CPBF *Xanthomonas* sp. strains (i.e. 65%) were attributed to one specific COG category of eggNOG database. The CDSs to which was not possible to accurately assign a COG category, i.e. 1600 out of 4131 CDSs for CPBF *Xaj* strains and 1249 out of 3618 CDSs for CPBF *Xanthomonas* sp. strains, were assigned to poorly characterized functional categories, including proteins not assigned or classified as unknown function (category COG S).

Several COGs categories, which include genes putatively related with pathogenicity and virulence were predominantly represented in the core genome, namely genes involved in biosynthetic pathways of xanthan, lipopolysaccharide (LPS) and flagella, chemotaxis, and hydrolytic enzymes, such as peptidases and cellulases (**Figure V.5.**). The genes encoding for the type II and type IV secretion systems, respectively T2SS and T4SS, are also allocated to the core genome (**Figure V.5.**). Concerning the functional analysis of the CDSs common to the accessory genome of both CPBF taxa, only 185 out 710 CDSs exclusive to both CPBF *Xaj* strains, and 80 CDSs out of the 197 CDSs exclusive to the three CPBF *Xanthomonas* sp. strains, were assigned to a COG category, which means that 74% and 59% of CDSs belonging respectively to CPBF *Xaj* and CPBF *Xanthomonas* sp. are poorly characterized group, including proteins not assigned or classified in a COG category with unknown function (category COG S). (**Figure V.5**). When analysing the set of 50 CDSs belonging to the accessory genome and shared by both CPBF *Xaj* strains and the pathogenic *Xanthomonas* sp. strain CPBF 424, it was noticeable the predominance of COG U to which the T3SS and T3E genes are assigned (**Figure V.5**).

Table V.2. COG categories distribution for the sets of orthologous CDSs corresponding to the core genome of the five CPBF strains.

	Shared by *Xaj* nº of genes (%)	**Shared by *Xanthomonas sp.*** nº of genes (%)
Information Storage and Processing Categories		
COG A - RNA processing and modification	1 (0.24)	1 (0.28)
COG J - Translation, ribosomal structure and biogenesis	164 (3.96)	163 (4.50)
COG K - Transcription	216 (5.16)	197 (5.44)
COG L - DNA replication, recombination, and repair	170 (4.11)	122 (3.37)
Cellular Processes and Signalling Categories		
COG D - Cell cycle control, cell division, and chromosome partitioning	30 (0.73)	28 (0.77)
COG M - Cell wall/membrane/envelope biogenesis	218 (5.28)	209 (5.78)
COG N - Cell motility	50 (1.21)	38 (1.05)
COG O - Posttranslational modification, protein turnover, chaperones	160 (3.87)	152 (4.20)
COG T - Signal transduction mechanisms	201 (4.86)	195 (5.39)
COG U - Intracellular trafficking, secretion, and vesicular transport	77 (1.86)	66 (1.82)
COG V - Defense mechanisms	66 (1.60)	53 (1.46)
Metabolism Categories		
COG C - Energy production and conversion	181 (4.38)	184 (5.08)
COG E - Amino acid transport and metabolism	230 (5.57)	218 (6.02)
COG F - Nucleotide transport and metabolism	77 (1.86)	75 (2.07)
COG G - Carbohydrate transport and metabolism	202 (4.89)	195 (5.38)
COG H - Coenzyme transport and metabolism	107 (2.59)	106 (2.93)
COG I - Lipid transport and metabolism	108 (2.61)	111 (3.07)
COG P - Inorganic ion transport and metabolism	218 (5.27)	209 (5.78)
COG Q - Secondary metabolites biosynthesis, transport and catabolism	55 (1.33)	47 (1.30)
Poorly characterized		
COG S - Function unknown	1124 (27.90)	957 (26.40)
Not in COGS	449 (10.80)	292 (8.07)

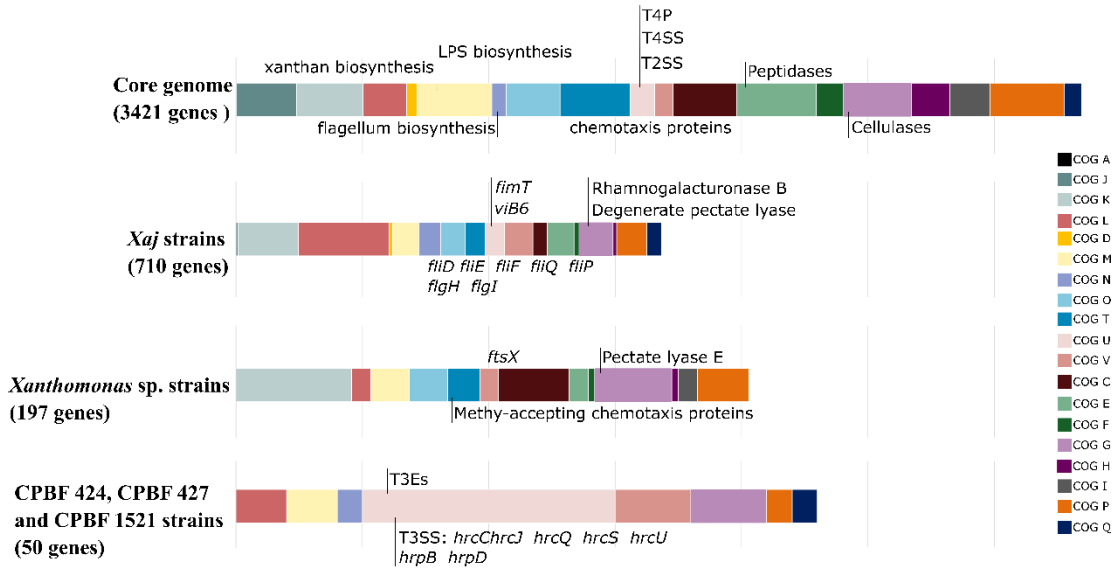

Figure V.5. Representation of the percentage of COG functional categories of CDSs corresponding to the core genome of five CPBF strains (from a total of 3421 CDSs); of CDSs shared exclusively by CPBF *Xaj* strains (CPBF 427 and CPBF 1521, from a total of 710 CDSs); of CDSs shared exclusively by CPBF *Xanthomonas* sp. strains (CPBF 367, CPBF 424, and CPBF 426, from a total of 197 CDSs); and of CDSs shared exclusively by CPBF *Xaj* strains and the pathogenic *Xanthomonas* sp. CPBF 424 strain (from a total of 95 CDSs). COG categories are represented by bars of different colours corresponding to the percentage of each COG category to the total number of CDSs for each group of CPBF strains. CDSs putatively linked to pathogenicity and virulence are highlighted for the matching COG category.

3.5. Pathogenic and virulence-related factors of CPBF strains

In addition to the functional analysis, the profile of *Xanthomonas* pathogenicity and virulence factors characterized in previous studies were determined for the five CPBF strains (CPBF 367, CPBF 424, CPBF 426, CPBF 427 and CPBF 1521). The results show a well-defined set of genes associated with infection processes, but most importantly the results also show notable differences between the consortium of the five CPBF strains isolated from the same walnut host.

All five xanthomonads CPBF strains were found to share numerous gene associated with pathogenesis, namely genes for the biosynthesis of the xanthan polysaccharide biosynthesis (operon *gumBCDEFGHIJKLMN*); the lipopolysaccharides biosynthesis; the flagellar system; the regulatory *rpf* cluster of pathogenicity factor synthesis; the *xps* gene of T2SS (*xpsD*, *E*, *F*, *G*, *H*, *I*, *J*, *K*, *L*, *M*); homologous of the T4P, several T4SS genes (*virB1*, *virB2*, *virB3*, *virB4*, *virB6*, *virB8*, *virB9*, *virB10*, *virB11* and *virD4*); as well as genes from the T6SS (**Supplementary material: Best BLAST results_Virulence genes in xls format**). Regarding the T2SS, *xpsN* gene was only present on the CPBF *Xanthomonas* sp. strains (CPBF 367, CPBF 424 and CPBF 426), and five T4P-related genes (*pilY1*, *pilX*, *pilW*, *pilV* and *fimT*) were found to be exclusively present in

the pathogenic *Xanthomonas* sp. strain CPBF 424 (**Table V.3, Supplementary material: Best BLAST results_Virulence genes in xls format**).

Concerning the presence of non-fimbrial adhesins it is noticeable that the *Xantomonas* sp. CPBF 424 and CPBF 426 strains did not harbour homologous genes for *fhaB1* and *fhaB2,* which encode filamentous hemagglutinin-related proteins, regardless the fact that all CPBF strains share at least six homologous to non-fimbrial adhesins described for *X. campestris* pv. *vesicatoria* 85-10 and *X. oryzae* pv. *oryzae* KACC 10331 (**Table V.3, Supplementary material: Best BLAST results_Virulence genes in xls format**). Additionally, no major differences were observed between the five strains studied regarding genes encoding proteins associated to *Xanthomonas* sensing and chemotaxis mechanisms, with exception for a methyl-accepting chemotaxis protein of which a homolog was present in all three CPBF *Xanthomonas* sp. strains but not in CPBF *Xaj* strains. (**Table V.3, Supplementary material: Best BLAST results_Virulence genes in xls format**). The main differences between CPBF *Xaj* strains (CPBF 427 and CPBF 1521) and *Xanthomonas* sp. strains (CPBF 367, CPBF 424, and CPBF 426) were observed in the profile of extracellular enzymes associated with T2SS (**Table V.3**). Homologous of *X. campestris* pv. *campestris* ATCC 33913 pectolytic enzymes, particularly a pectate lyase E (NP_635517.1) and a pectinesterase (NP_635516.1) were only identified in the genomes of the three *Xanthomonas sp.* strains (CPBF 367, CPBF 424 and CPBF 426). On the contrary, homologous of a degenerate pectate lyase (AAM37225.1), an endoglucanase (AAM38359.1), a rhamnogalacturonase B (AAM38348.1), and a polygalacturonase (NP_637621.1) were present in CPBF *Xaj* and not in CPBF *Xanthomonas* sp. strains. Furthermore, homologous of xylosidases (NP_636552.2, NP_639444.2) were found in the two CPBF *Xaj* strains and in the pathogenic *Xanthomonas* sp. strain CPBF 424.

Table V.3. Comparison of the occurrence of homologous associated with chemotaxis, non-fimbrial adhesins, type II secretion system (T2SS), extracellular enzymes, and type IV pilus (T4P), between CPBF Xaj strains (CPBF 427, CPBF 1521) and CPBF *Xanthomonas* sp. strains (CPBF 367, CPBF 424, and CPBF 426).

Protein (Accession number)	Gene name	*Xanthomonas* sp. strains			*Xaj* strains	
		CPBF 367	CPBF 424	CPBF 426	CPBF 427	CPBF 1521
Chemotaxis related proteins:						
Methyl-accepting chemotaxis protein (CAJ23615.1)	XCV1938	X	X	X	-	-
Non-fimbrial adhesins:						
Filamentous hemagglutinin-related protein (CAJ23537.1)	*fhaB1* XCV1860	X	-	-	X	X
Filamentous hemagglutinin-related protein (CAJ23538.1)	*fhaB2* XCV1861	X	-	-	X	X
Type II secretion system (T2SS):						
General secretion pathway protein XpsN (NP_636061.1)	*xpsN* XCC0669	X	X	X	-	-
Extracellular enzymes:						
Xylosidase/arabinosidase (NP_636552.2)	*xylB* XCC1178	-	X	-	X	X
Xylosidase/arabinosidase (NP_637752.1)	*xsa* XCC2398	-	-	-	X	X
Xylosidase/arabinosidase (NP_639444.2)	*xylB* XCC4105	-	X	-	X	X
Endoglucanase (AAM38359.1)	*bcsZ* XAC3516	-	-	-	X	X
Polygalacturonase (NP_637621.1**)**	*pglA* XCC2266	-	-	-	X	X
Pectate lyase E **(**NP_635517.1)	XCC0122	X	X	X	-	-
Degenerated pectate lyase (AAM37225.1)	*pel* XAC2373	-	-	-	X	X
Pectinesterase (NP_635516.1)	XCC0121	X	X	X	-	-
Rhamnogalacturonase B (AAM38348.1)	*rhgB* XAC3505	-	-	-	X	X
Type IV pilus (T4P):						
PilY1 protein (WP_011051753.1)	*pilY1* XAC2665	-	X	-	-	-
PilX protein (WP_011051754.1)	*pilX* XAC2666	-	X	-	-	-
PilW protein (WP_040107776.1)	XAC2667	-	X	-	-	-
Pre-pilin leader sequence (WP_011051756.1)	*pilV* XAC2668	-	X	-	-	-
Pre-pilin like leader sequence (WP_011051757.1)	*fimT* XAC2669	-	X	-	-	-

The distinct patterns of Type III Secretion System (T3SS) and Type 3 Effectors (T3E) homologous present in CPBF *Xaj* strains and CPBF *Xanthomonas* sp. strains seems to reflect the pathogenicity phenotype. In fact, the pathogenic CPBF *Xaj* strains (CPBF 427 and CPBF 1521) and the pathogenic *Xanthomonas* sp. CPBF 424 have a similar and extensive set of T3SS, while the two-other *Xanthomonas* sp. strains (CPBF 367 and CPBF 426) lack several *hrp* genes and *hrc* genes of T3SS, with exception for *hrpG, hrpN, hrcQ and hrpX* (**Figure V.6**, **Supplementary material: Best BLAST results_Virulence genes in xls format**).

Considering the T3E, both *Xaj* strains CPBF 427 and CPBF 1521 possess 23 homologous from the 24 genes previously identified in *Xanthomonas sp.* (*avrBs1, avrBs2*, *avrBs3, avrXccA1*, *avrXccA2*, *hpaA*, *hrpW*, *xopA*, *xopAW, xopAZ*, *xopC, xopF1*, *xopG*, *xopK*, *xopL*, *xopM, xopN*, *xopQ*, *xopR*, *xopV*, *xopX*, *xopZ xopZ1,* and *xopZ2*) (*Xanthomonas* Resource, White et al. 2009, Hajri et al. 2012, Cesbron et al. 2015, Essakhi et al. 2015). From this set of 24 T3Es, eight

homologous of proteins were found in CPBF *Xanthomonas* sp. strains, although not evenly distributed between the three strains. Indeed, *xopAZ* and *xopR* were present in the three *Xanthomonas* sp. genomes (CPBF 367, CPBF 424 and CPBF 426), as well as in CPBF *Xaj* strains (CPBF 427 and CPBF 1521), a *xopF1* homolog was found in CPBF 424 and CPBF 426 but not in CPBF 367, and homologous of *hpaA*, *hrpW*, *xopA*, and *xopM* were identified only in *Xanthomonas* sp. CPBF 424, but not in *Xanthomonas* sp. CPBF 367 and CPBF 426, and *xopZ2* was the only homolog exclusively found in strain CPBF 424 (**Figure V.6**, **Supplementary material: Best BLAST results_Virulence genes in xls format**).

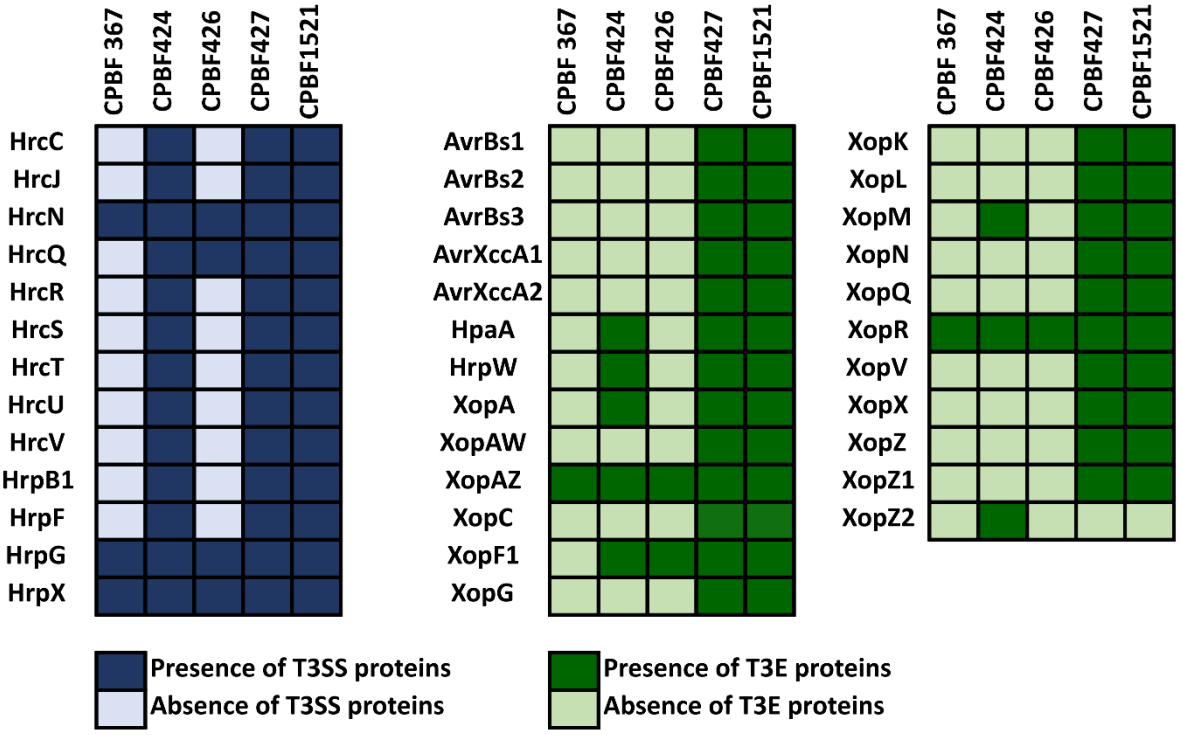

Figure V.6. Distribution of protein homologous of the type III secretion system (T3SS) and type III effectors (T3Es) among the five CPBF strains (CPBF 367, CPBF 424, CPBF 426, CPBF 427 and CPBF 1521). Dark blue corresponds to the presence of T3SS protein homologous, and dark green corresponds to the presence of T3E homologous. The light blue and green means absence of T3SS and T3E homologous, respectively.

4. Discussion

The occurrence of distinct *Xanthomonas* populations colonizing the same host plant has been documented (Jacques et al. 2016). In walnut and stone fruit trees, besides the presence of *X. arboricola* strains belonging to the pathovars characteristic of these two host species, namely pv. *juglandis* in walnut and pv. *pruni* in stone fruit trees, it has been reported the isolation of yellow-pigmented xanthomonads that do not form a phylogenetically coherent group with the pathogenic

strains of *X. arboricola* pathovars (Essakhi et al. 2015, Garita-Cambronero et al. 2016, Jacques et al. 2016). Interestingly these distinct strains were mostly represented by nonpathogenic lineages of *X. arboricola* (Essakhi et al. 2015, Garita-Cambronero et al. 2016). More recently, independent studies mentioned the existence of strains belonging to new *Xanthomonas* species (Fernandes et al. 2018a, López et al. 2018), including the newly pathogenic *Xathomonas prunicola* isolated from *Prunus persica* var. nectarine causing similar symptoms as *X. arboricola pv. pruni* on *Prunus* spp. (López et al. 2018). These findings raise the need to understand the role played in the pathosystems by these bacteria taxonomically assigned to different species and characterized by distinct genotypes. Currently the ecological, evolutionary and pathogenicity implications of this co-colonization is not understood, but is hardly refutable that this knowledge is needed for improvement of efficient phytosanitary practices and the design of appropriate management strategies (Quibod et al. 2016).

In this study, five *Xanthomonas* strains (CPBF 367, CPBF 424, CPBF 426, CPBF 427 and CPBF 1521) isolated from the same walnut tree, and revealing distinct genotypes were chosen for genome sequencing. Previous MLSA analysis coupled with dot blot hybridization using nine DNA markers (Fernandes et al. 2018a), allowed to divide these five CPBF strains in two distinct groups, one formed by CPBF 427 and CPBF 1521 strains, which cluster with *Xaj* strains, and another group formed by strains CPBF 367, CPBF 424 and CPBF 426, which are closer to nonpathogenic *Xanthomonas arboricola* frequently referred as atypical *X. arboricola* than to *Xaj* (Essakhi et al. 2015).

In order to provide a more in-depth characterization of these five CPBF strains, capable to clarify their taxonomic affiliation and disclose differential genomic contents putatively related with pathogenicity, virulence and other specific niche adaptations, the genome of these CPBF strains was sequenced, assembled, annotated, and deposited at the European Nucleotide Archive (ENA) and uploaded at EDGAR (Blom et al. 2009, Blom et al. 2016) for comprehensive comparative genomics. Both the Average Nucleotide Identity (ANI) and a core-genome phylogenetic analysis with a total of 44 *Xanthomonas* genomes, confirmed that strains CPBF 427 and CPBF 1521 are within the cluster formed by *Xaj* strains and were therefore designated as CPBF *Xaj* strains. The other three strains CPBF 367, CPBF 424 and CPBF 426, have ANI values below the threshold proposed by Konstantinidis and Tiedje (2005) for distinct species (i.e. 95%), in comparison with any of the other *Xanthomonas* species included in the study, and form a distinct clade in the phylogenomics tree, suggesting that CPBF 367, CPBF 424 and CPBF 426 belong to a new *Xanthomonas* species, herein designated as CPBF *Xanthomonas* sp. strains.

The general genomic features further sustain the existence of two distinct species. In fact, CPBF *Xaj* strains (CPBF 427 and CPBF 1521) differ in their genome size from the other three strains isolated from the same walnut host and considered belonging to a new *Xanthomonas*

species (CPBF 367, CPBF 424 and CPBF 426) (**Table V.1**). The genome size of CPBF *Xaj* strains (5.19 Mb for CPBF 427 and CPBF 1521) is roughly equal to the genome size reported for other sequenced *Xaj* genomes, namely *Xaj* 417 (NZ_CP012251), NCPPB 1447 (AJTL0100000000), J303 (LSGZ0100000000), DW3F3 (PNRC0100000000, Fu et al. 2018), CFBP 2528 and CFBP 7179 (JZEF0100000000 and JZEG0100000000, Cesbron et al. 2015), or CFSAN033077 and CFSAN033080 (Accession no. LHBK0100000000 and LHBN0100000000, Higuera et al. 2015). The genomes of *Xanthomonas* sp. strains CPBF 367, CPBF 424 and CPBF 426 were considerably smaller, presenting values closer to the nonpathogenic or avirulent strains of *X. arboricola*, i.e. with uncertain pathogenicity or belonging to non-*juglandis*, non-*pruni* and non-*corylina* pathovars, with genome sizes inferior than 5Mb (Garita-Cambronero et al. 2016, Ignatov et al. 2015, Harrison et al. 2016). When analysing the pangenome, defined according to Rouli et al. (2015) by the whole gene repertoire of the studied strains, and illustrated by a Venn diagram (**Figure V.4**), besides the presence of a large core genome shared by the five CPBF strains, which comprise genes associated with basic biological aspects of the *Xanthomonas* genus, as phenotypic traits (Medini et al. 2005), it was possible to discern a considerable number of genes shared exclusively by the two CPBF *Xaj* strains (710 genes) and 197 genes shared exclusively by CPBF *Xanthomonas* sp. strains. The analysis of these gene sets specific to each of these two walnut associated bacteria formed by CPBF *Xaj* strains and CPBF *Xanthomonas* sp. strains suggests the presence of genes encoding for proteins associated with biochemical functions probably conferring selective advantages, such as adaptation to different niches, pathogenicity, or colonization to a new host (Medini et al. 2005).

The number and diversity of insertion sequences (IS) were also different across the two groups of strains. *Xaj* strains (CPBF 427 and CPBF 1521) possess approximately six-fold more IS elements in their genomes than *Xanthomonas* sp. strains (CPBF 367, CPBF 424 and CPBF 426). A similar observation was reported previously for *Xanthomonas* strains associated with walnut, in which *Xaj* strains was found to have 10 times more IS elements than nonpathogenic strains of *X. arboricola* (Cesbron et al. 2015). Furthermore, the five CPBF genomes share IS families which are common in other *Xanthomonas* spp. genomes (Bogdanove et al. 2011, Ryan et al. 2011, Cesbron et al. 2015). The CPBF 427 and CPBF 1521 *Xaj* strains have mostly elements of IS3 family in their genomes, which was similar to previously observed in other *Xaj* genomes (Cesbron et al. 2015). The differences evident in this data call attention to investigate in the near future their importance in the evolution of CPBF *Xanthomonas* sp. strains. It is generally acknowledged that IS elements are major contributors for genomic diversity among species of the genus *Xanthomonas*, and consequently of *Xanthomonas* evolution (Meyer and Bogdanove 2009, Ryan et al. 2011). In fact, IS elements can generate genome modifications as

rearrangements, deletions, and inversions that can lead to changes of gene composition (Meyer and Bogdanove 2009).

The search for orthologous of the gene datasets that belongs to CPBF *Xaj* strains and to *Xanthomonas* sp. CPBF strains, using the eggNOG database (Huerta-Cepas et al. 2016), allowed to underline the differences about the distribution in the COG categories and provide some insights into distinct categories of these two groups. Focusing mainly on the identification of genomic determinants of pathogenicity and virulence, the intention was to provide evidences about the functional significance of this consortium of walnut-associated *Xanthomonas* occurring in the same host specimen. Furthermore, the 710 genes shared exclusively by *Xaj* strains, the 197 genes shared exclusively by *Xanthomonas* sp. strains, and the 50 genes shared exclusively by *Xaj* strains and the pathogenic *Xanthomonas* sp. CPBF 424, were also scrutinized for the presence of orthologous know to be involved in pathogenicity and virulence of xanthomonads.

Regarding the core genome, not surprisingly, numerous CDSs are common to both groups of strains, with emphasis for COG categories responsible for morphology, physiology, energy acquisition and most cellular functions (Meyer and Bogdanove 2009, Ryan et al. 2011). Furthermore, previous studies highlight that COG categories related to metabolic pathways of nucleotides, carbohydrates, cofactors and vitamins, as well as, amino acids and lipids are conserved in the gammaproteobacterial group (Nourdin-Galindo et al. 2017). The core genome also includes common components related to *Xanthomonas* virulence factors, which were also found in CPBF *Xaj* and CPBF *Xanthomonas* sp. strains. This analysis evokes the conservation of the molecular basis of virulence related with host adaptation or tissue specificity: Xanthan is the most characteristic exopolysaccharide (EPS) of *Xanthomonas* spp. and has been shown to be implicated in bacterial epiphytic survival, growth and biofilm formation (Katzen et al. 1998, Vorholter et al. 2008, Buttner and Bonas 2010). Lipopolysaccharides (LPSs) are major components of the bacterial outer membrane, protects bacteria from unfavourable environmental conditions contributing to the restrictive membrane permeability but also induce plant defence responses (Dow et al. 2000, Buttner and Bonas 2010). Adhesins, as non-fimbrial adhesins and bacterial type IV pilus (T4P), are involved with interaction and aggregation mechanisms related to biofilm formation and bacteria attachment to specific sites of the host surface (Buttner and Bonas 2010, Berne et al. 2015). Another cell component implicated in surface attachment is flagella, which play a key role during bacterial colonization and it was already related to the development and dispersion of biofilm and virulence in a *Xanthomonas* species (Malamud et al. 2011, Newman et al. 2013). The *rpf* gene cluster is mostly related to the control of proteases, endoglucanases, EPS and other known important pathogenicity factors (Dow et al. 2000). Bacterial receptors as methyl-accepting chemotaxis proteins (MCPs) are involved in *Xanthomonas* chemotaxis signal transduction and cell motility towards more favourable niches,

being important to bacterial movement and survival in a changing environment (Wadhams and Armitage 2004). Protein secretion systems, as type II secretion system (T2SS), type III secretion system (T3SS), type IV secretion system (T4SS) and type VI secretion system (T6SS) are responsible for the translocation of effector proteins, toxins, extracellular degradative enzymes or DNA into the plant cell cytosol that might contribute to the host–pathogen interaction (Buttner and Bonas 2010). In fact, most of these pathogenesis and virulence-associated gene clusters that are found in *Xanthomonas* spp. genomes (Ryan et al. 2011) are also present in the genomes of CPBF *Xaj* and CPBF *Xanthomonas* sp. strains, namely homologous of the xanthan EPS biosynthetic pathway, LPSs biosynthesis, flagellar system, *rpf* cluster of pathogenicity factor system, and of several genes associated with T2SS and T4P, as well as, genes described to be associated with T6SS.

On the contrary, some genetic determinants of virulence were showed to be group-specific or even strain-specific, when the accessory genome was considered. For example, CPBF *Xanthomonas sp.* strains evidenced an exclusive presence of *xpsN* gene, which encodes the XpsN protein of type II secretion system. Proteins of the *xps* system were shown to be associated with virulence of *Xanthomonas* species as *X. campestris*, *X. oryzae* and *X. euvesicatoria* (Ryan et al. 2011). Furthermore, *Xanthomonas* sp. strain CPBF 424 harbour a set of genes that encodes proteins associated to type IV pilus (PilY1, PilX, PilW, PilV, FimT), some of these proteins are considered primary structures of the T4P pilin subunits. Indeed, T4P could play an important role in pathogenesis of various species of *Xanthomonas*, in some cases it is thought that this system may have a role in plant colonization (Dunger et al. 2016, Hersemann et al. 2017). It was also possible identify homologous specific to CPBF *Xaj* strains encoding for proteins that are missing in CPBF *Xanthomonas* sp. strains. Homologous of extracellular enzymes, as xylosidase, endoglucanase, polygalacturonase, degenerate pectate lyase, rhamnogalacturonase B, were only detected in the CPBF *Xaj* strains CPBF 427 and CPBF 1521, both pathogenic for walnut.

Not surprisingly, major genomic differences between the CPBF strains were observed for T3SS and T3E homologous. In *Xanthomonas* spp. T3SS is crucial for pathogenicity being able to translocate effector proteins that have a key role in bacterial proliferation in host tissues and development of disease symptoms (Buttner and Bonas 2010). Accordingly, T3SS differentiate the pathogenic CPBF strains (CPBF 424, CPBF 427 and CPBF 1521) from the nonpathogenic strain CPBF 367. Indeed, noticeable differences were observed in some of the structural and regulatory components of T3SS, including all nine hypersensitive response conserved genes (*hrc* – *hrcN*, *hrcV*, *hrcU*, *hrcQ*, *hrcC*, *hrcJ*, *hrcR*, *hrcS*, *hrcT*) and some of the genes from the cluster of hypersensitive response and pathogenicity (*hrp* – *hrpB1*, *hrpF*, *hrpX*, *hrpG*). This result is aligned with previous studies emphasizing a highly conserved T3SS of the Hrp2 family commonly found in *Xanthomonas* group, including of *X. arboricola* strains belonging to pathovars *pruni*,

corylina, *juglandis*, *celebensis* and *fragariae* (Hajri et al. 2012). The two pathogenic CPBF *Xaj* strains (CPBF 427 and CPBF 1521) and the pathogenic *Xanthomonas* sp. strain CPBF 424, showed to have a T3SS profile comprised by homologous of genes also described in pathogenic strains of *X. arboricola* (Cesbron et al. 2015, Garita-Cambronero et al. 2016, 2017, 2018). The other two CPBF *Xanthomonas* sp. strains (CPBF 367, CPBF 426), miss some of the components from Hrp2 family, which suggests that they may be impaired for pathogenicity and, although CPBF 426 has not been tested for pathogenicity, strain CPBF 367 was shown to be nonpathogenic on walnut. These two strains harboured regulators genes of T3SS, as *hrpX* and *hrpG* (Guo et al. 2011, Jacobs et al. 2015), but the genes that composed the macromolecular structure of the T3SS were absent, as the *hrpF* gene known to be involved in the translocation of effector proteins into the host cell (Cesbron et al. 2015).

The prediction of type III effectors homologous in the five *Xanthomonas* CPBF strains analysed in this work was carried out by BLAST using a database with candidate T3Es of the *Xanthomonas* group. The search included the T3E genes *xopR*, *avrBs2*, *xopK*, *xopL*, *xopN*, *xopQ*, *xopX*, and *xopZ*, reported for several sequenced *Xanthomonas* spp. (Ryan et al. 2011). Homologous of the T3Es which have been reported as confined to some *Xanthomonas* species, were also investigated, namely homologous of *xopB*, *xopE2*, *xopF1*, *xopV* and *xopZ* which are described to be normally present in pathogenic strains that cause disease on hazelnut, stone fruit trees and walnut (Garita-Cambronero et al. 2017); homologous of *xopAL1*, *xopAC*, *xopAD* and *xopAL2* associated with *X. campestris* strains that infect cruciferous plants (Ryan et al. 2011); and homologous of *xopE3* and *xopAI* found on citrus pathogens *X. citri* pv. *citri* and *X. fuscans* subsp. *aurantifolii* (Ryan et al. 2011). The data showed that the repertoire of T3Es of *Xanthomonas* sp. CPBF 424 strain differs markedly from the CPBF *Xaj* and the other two CPBF *Xanthomonas* sp. strains (CPBF 367 and CPBF 426). Although, this pathogenic strain has shown to possess homologous for eight known effectors (*hpaA*, *hrpW*, *xopA*, *xopAZ, xopF1*, *xopM*, *xopR*, and *xopZ2*), from which only three (*xopAZ*, *xopR* and *xopF1*) were identified in the other two CPBF *Xanthomonas* sp. (CPBF 367, and CPBF 426), no homologous were found for most of the genes acknowledge as important for the pathogenicity of *Xanthomonas* species, namely *avrbs2, avrbs2*, *avrbs3, avrXccA1, avrXccA2, xopAW, xopC, xopG, xopK*, *xopL*, *xopN*, *xopQ*, *xopV, xopX*, *xopZ, xopZ1*, which, excluding *xopZ2,* were all present in both CPBF *Xaj* strains (CPBF 427, and CPBF 1521). This data suggests that these T3E homologous are not absolutely needed for *Xanthomonas* pathogenicity in walnut. Interestingly, members of a few other *Xanthomonas* species, as *X. albilineans*, *X. campestris* pv. *armoraciae*, X. *campestris* pv. *raphanin* also lack most of these conserved xanthomonads T3Es (Jalan et al. 2011, Ryan et al. 2011, Bogdanove et al. 2011).

Efforts have been made by Cesbron et al. (2015) to elucidate the molecular basis of pathogenesis processes of *Xaj*, and to provide insights into the mechanisms associated with the emergence of *Xaj* pathogenic strains, by comprehensive comparative genomics of four *X. arboricola* strains isolated from walnut trees differing on their pathogenicity (Cesbron et al. 2015). Genomic differences related to virulence between pathogenic *Xaj* and nonpathogenic *X. arboricola* strains were found. Notably, pathogenic *Xaj* strains contained a conserved cluster of *hrp/hrc* genes encoding a T3SS, which were absent in one of the nonpathogenic *X. arboricola* strains (Cesbron et al. 2015). In addition, the number of virulence-associated proteins of T3SS, i.e. T3Es, showed to be different in pathogenic and nonpathogenic *X. arboricola* strains, and clearly reduced in nonpathogenic *X. arboricola* strains (Cesbron et al. 2015). Recently, a nonpathogenic *X. arboricola* pv. *pruni* (*Xap*)-look-a-like strain obtained from *Prunus* was characterized and its genome compared to pathogenic *Xap* strains. This study also revealed differences in the profile of virulence related genes, such as the genes related to the T3SS and T3Es (Garita-Cambronero et al. 2017).

Regardless the contributions of these comparative genomics studies focused on pathogenicity and virulence related genes, to highlight the importance of T3SS and T3E genes, dedicated functional studies are absolutely needed to identify which genes are essential to ensure a successful infection. The current availability of nonpathogenic walnut-associated *Xanthomonas* sp. as CPBF 367 and CPBF 426, and the *Prunus* nonpathogenic *Xap*-look-a-like strain, are particularly valuable to address questions regarding the evolution of pathogenicity of *Xanthomonas*.

Beyond the incontestably importance of T3SS and T3E emphasised above, the existence of other genetic determinants of pathogenicity and virulence in *Xanthomonas* cannot be dismissed, such as the role of other secretion systems, the weight of chemotaxis for a pathogenic phenotype, the importance of non-fimbrial adhesins for bacterial attachment, and the plant cell wall hydrolytic activity of bacterial extracellular enzymes. Although, as emphasized above, some of this genetic patrimony are present in the core genome, the comparison of gene homologous coding for some of these features in pathogenic CPBF *Xaj* and CPBF *Xanthomonas* sp. strains, which include a pathogenic (CPBF 424) and a nonpathogenic (CPBF 367) strain could not clearly be assigned to a pathogen phenotype. Comparative studies between pathogenic *Xaj* and nonpathogenic *X. arboricola* strains has shown a differential repertoire of genes encoding chemotaxis related proteins and proteins related with type I secretion system, type II secretion system, type IV secretion system (Cesbron et al. 2015). Furthermore, a differential repertoire of non-fimbrial adhesins involved in different functions related to bacterial attachment to the host surface were found in all strains, which homologous of non-fimbrial adhensis *fhaB* probably associated with the bacteria colonization were only identified in pathogenic strains (Cesbron et al. 2015). In the

same way, homologous of proteins involved in the biogenesis of type IV pilus were observed but the absence of PilA, PliX and/or PilV proteins in the pathogenic and nonpathogenic genomes could be an indication of the absence of bacterial surface filaments in all strains (Cesbron et al. 2015).

In the present work an extensive comparison between the genomes of five walnut-associated strains isolated from the same walnut specimen and belonging to two species (*Xaj* and *Xanthomonas* sp.), including pathogenic (CPBF 427, CPBF 1521, CPBF 424) and a nonpathogenic strain (CPBF 367) phenotypes, was carried out firstly, to disclose genetic determinants of pathogenicity and virulence, and secondly, to provide insights about niche-specific adaptations that could inform about the role played by each of these strains in the co-colonization of walnut. With exception for the presence of T3SS and some T3E, the data gathered suggests a pattern of homologous putatively related with the pathogenicity, virulence and niche-specific adaptations variable between the five CPBF strains that need to be addressed in functional studies to determine their importance in walnut diseases caused by *Xanthomonas*.

Acknowledgments

This work was co-financed by the European Structural & Investment Funds (ESIFs) through the Operational Competitiveness and Internationalization Programme – COMPETE 2020 and by National Funds through FCT – Fundação para a Ciência e a Tecnologia, within the framework of the project EVOXANT (PTDC/BIA-EVF/3635/2014, POCI-01-0145-FEDER-016600). This article is based on work from European Cooperation in Science and Technology (COST) action CA16107 EuroXanth, supported by COST. C.F. was supported by a fellowship from FCT (SFRH/BD/95913/2013) and a Short Term Scientific Mission (STSM) from COST action CA16107. The EDGAR platform is funded by Federal Ministry of Education and Research (BMBF) grant FKZ 031A533.

References

Alikhan NF, Petty NK, Zakour NLB, Beatson SA. 2011. BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. *BMC genomics* **12**, 402.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *Journal of Molecular Biology* **215**, 403-410.

Belisario A, Maccaroni M, Corazza L, Balmas V, Valier A, 2002. Occurrence and etiology of brown apical necrosis on Persian (English) walnut fruit. *Plant Disease* **86**, 599-602.

Berne C, Ducret A, Hardy GG, Brun YV. 2015. Adhesins involved in attachment to abiotic surfaces by gram-negative bacteria. *Microbial Spectr* **3**, 1-45.

Blom J, Albaum SP, Doppmeier D, Puhler A, Vorholter FJ, Zakrzewski M, Goesmann A. 2009. EDGAR: a software framework for the comparative analysis of prokaryotic genomes. BMC *Bioinformatics* **10**, 154.

Blom J, Kreis J, Spanig S, Juhre T, Bertelli C, Ernst C, Goesmann A, 2016. EDGAR 2.0: an enhanced software platform for comparative gene content analyses. *Nucleic Acids Research* **44**, W22-W28.

Bogdanove AJ, Koebnik R, Lu H, Furutani A, Angiuoli SV, Patil PB, Van Sluys MA, Ryan RP, Meyer DF, Han SW, Aparna G, Brendel VP, Rabinowicz PD, Leach JE, White FFm, Salzberg SL. 2011. Two new complete genome sequences offer insight into host and tissue specificity of plant pathogenic *Xanthomonas* spp. *Journal of Bacteriology* **193**, 5450-5464.

Buttner D, Bonas U. 2010. Regulation and secretion of *Xanthomonas* virulence factors. *FEMS Microbiology Review* **34**, 107-133.

Cesbron S, Briand M, Essakhi S, Gironde S, Boureau T, Manceau C, Fischer-Le Saux M, Jacques MA, 2015. Comparative genomics of pathogenic and nonpathogenic strains of *Xanthomonas arboricola* unveil molecular and evolutionary events linked to pathoadaptation. *Frontiers in Plant Science* **6**, 1126.

Chevreux B, Wetter T, Suhai S. 1999. Genome sequence assembly using trace signals and additional sequence information, p. 45-56. In Wingender E. Computer Science and Biology: proceedings of the German Conference on Bioinformatics, GCB 99, Hannover, Germany.

Crossman L, Dow JM. 2004. Biofilm formation and dispersal in *Xanthomonas campestris*. *Microbes and infection* **6**, 623-629.

Darrasse A, Carrère S, Barbe V, Boureau T, Arrieta-Ortiz ML, Bonneau S, Fouteau, S. 2013. Genome sequence of *Xanthomonas fuscans* subsp. *fuscans* strain 4834-R reveals that flagellar motility is not a general feature of xanthomonads. *BMC genomics* **14**, 761.

Dow JM, Daniels MJ, Feng JX, Tang JL, Barber CE. 2000. Novel genes involved in the regulation of pathogenicity factor production within the rpf gene cluster of *Xanthomonas campestris*. *Microbiology* **146**, 885–891.

Dowson DW. 1939. On the systematic position and generic names of the gram negative bacterial plant pathogens. Zentralblatt fur Bakteriologie, Parasitenkunde, Infektionskrankheiten und Hygiene 100, 177-193.

Dunger G, Llontop E, Guzzo CR, Farah CS. 2016. The *Xanthomonas* type IV pilus. *Current opinion in microbiology* **30**, 88-97.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32**, 1792-1797.

Essakhi S, Cesbron S, Fischer-Le Saux M, Bonneau S, Jacques MA, Manceau C, 2015. Phylogenetic and variable-number tandem-repeat analyses identify nonpathogenic *Xanthomonas arboricola* lineages lacking the canonical Type III Secretion System. *Applied and Environmental Microbiology* **81**, 5395-410.

Fernandes C, Albuquerque P, Sousa R, Cruz L, Tavares F. 2017. Multiple DNA markers for identification of *Xanthomonas arboricola* pv. *juglandis* isolates and its direct detection in plant samples. *Plant Disease* **101**, 858-865.

Fernandes C, Albuquerque P, Cruz L, Tavares F. 2018a. Genotyping and epidemiological metadata provides new insights into population structure of *Xanthomonas* isolated from walnut trees. bioRxiv:397703.

Fernandes C, Blom J, Pothier JF, Tavares F. 2018b. High-quality draft genome sequences of *Xanthomonas* sp. strain CPBF 424, a walnut pathogenic strain with atypical features. *Microbiology Resource Announcements* **7**, e00921-18.

Fernandes C, Blom J, Pothier JF, Tavares F. 2018c. High-quality draft genome sequences of *Xanthomonas arboricola* pv. *juglandis* CPBF 1521, isolated from leaves of a symptomatic walnut tree in Portugal without a past of phytosanitary treatment. *Microbiology Resource Announcements* **7**, e00887-18.

Filloux A. 2004. The underlying mechanisms of type II protein secretion. *Biochimica et Biophysica Acta* **1694**, 163-179.

Fischer-Le Saux M, Bonneau S, Essakhi S, Manceau C, Jacques MA, 2015. Aggressive emerging pathovars of *Xanthomonas arboricola* represent widespread epidemic clones distinct from poorly pathogenic strains, as revealed by multilocus sequence typing. *Applied and Environmental Microbiology* **81**, 4651-4668.

Fu B, Chen Q, Wei M, Zhu J, Zou L, Li G, Wang L. 2018. Complete genome sequence of *Xanthomonas arboricola* pv. *juglandis* strain DW3F3, isolated from *Juglans regia* L. bacterial blighted fruitlet. *Genome Announcements* **6**, e00023-18.

Garita-Cambronero J, Palacio-Bielsa A, López MM, Cubero J. 2016. Draft genome sequence for virulent and avirulent strains of *Xanthomonas arboricola* isolated from *Prunus* spp. in Spain. *Standards in Genomic Sciences* **11**, 12.

Garita-Cambronero J, Palacio-Bielsa A, Lopez MM, Cubero J. 2017. Pan-genomic analysis permits differentiation of virulent and non-virulent strains of *Xanthomonas arboricola* that cohabit *Prunus* spp. and elucidate bacterial virulence factors. *Frontiers in Microbiology* **8**, 573.

Garita-Cambronero J, Palacio-Bielsa A, Cubero J, 2018. *Xanthomonas arboricola* pv. *pruni*, causal agent of bacterial spot of stone fruits and almond: its genomic and phenotypic characteristics in the *X. arboricola* species context. *Molecular Plant Pathology* **19**, 2053-2065.

Giovanardi D, Bonneau S, Gironde S, Fischer-Le Saux M, Manceau C, Stefani E. 2015. Morphological and genotypic features of *Xanthomonas arboricola* pv. *juglandis* populations from walnut groves in Romagna region, Italy. *European Journal of Plant Pathology* **145**, 1-16.

Guo Y, Figueiredo F, Jones J, Wang N. 2011. HrpG and HrpX play global roles in coordinating different virulence traits of *Xanthomonas axonopodis* pv. *citri*. *Molecular Plant-Microbe Interactions: MPMI* **24**, 649-661.

Hajri A, Meyer D, Delort F, Guillaumes J, Brin C, Manceau C. 2010. Identification of a genetic lineage within *Xanthomonas arboricola* pv. *juglandis* as the causal agent of vertical oozing canker of Persian (English) walnut in France. *Plant Pathology* **59**, 1014-22.

Hajri A, Pothier JF, Fischer-Le Saux M, Bonneau S, Poussier S, Boureau T, Manceau C. 2012. Type three effector gene distribution and sequence analysis provide new insights into the pathogenicity of plant-pathogenic *Xanthomonas arboricola*. *Applied and Environment Microbiology* **78**, 371-384.

Harrison J, Grant MR, Studholme DJ**.** 2016. Draft genome sequences of two strains of *Xanthomonas arboricola* pv. *celebensis* isolated from banana plants. *Genome Announcements* **4**, e01705-15.

Hersemann L, Wibberg D, Blom J, Goesmann A, Widmer F, Vorholter FJ, Koliker. 2017. Comparative genomics of host adaptive traits in *Xanthomonas translucens* pv. *graminis*. *BMC Genomics* **15**, 35.

Higuera G, González-Escalone N, Véliz C, Vera F, Romero J. 2015. Draft genome sequences of four *Xanthomonas arboricola* pv. *juglandis* strains associated with walnut blight in Chile. *Genome Announcements* **3**, e01160-15.

Huerta-Cepas J, Szklarczyk D, Forslund K, Cook H, Heller D, Walter MC, Rattei T, Mende DR, Sunagawa S, Kuhn M, Jensen, LJ, von Mering C, Bork P. 2016. eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic acids research* **44**, D286-D293.

Ignatov AN, Kyrova EI, Vinogradova SV, Kamionskaya AM, Schaad NW, Luster DG. 2015. Draft genome sequence of *Xanthomonas arboricola* strain 3004, a causal agent of bacterial disease on barley. *Genome Announcements* **3**, e01572-14.

Ivanović Ž, Popović T, Janse J, Kojić M, Stanković S, Gavrilović V, Fira D. 2015. Molecular assessment of genetic diversity of *Xanthomonas arboricola* pv. *juglandis* strains from Serbia by various DNA fingerprinting techniques. *European Journal of Plant Pathology* **141**, 133-145.

Jacobs JM, Pesce C, Lefeuvre P, Koebnik R. 2015. Comparative genomics of a cannabis pathogen reveals insight into the evolution of pathogenicity in *Xanthomonas*. *Frontiers in Plant Science* **6**, 431.

Jacques MA, Arlat M, Boulanger A, Boureau T, Carrère S, Cesbron S, Chen NWG, Cociancich S, Darrasse A, Denancé N, Fischer-Le Saux M, Gagnevin L, Koebnik R, Lauber E, Noel LD, Pieretti I, Portier P, Pruvost O, Rieux A, Robène MR, Szurek B, Verdier V, Vernière C. 2016. Using ecology, physiology, and genomics to understand host specificity in *Xanthomonas*: French Network on Xanthomonads (FNX). *Annual Review of Phytopathology* **54**, 163-187.

Jalan N, Aritua V, Kumar D, Yu F, Jones JB, Graham JH, Setubal JC, Wang N. 2011. Comparative genomic analysis of *Xanthomonas axonopodis* pv. *citrumelo* F1, which causes citrus bacterial spot disease, and related strains provides insights into virulence and host specificity. *Journal of Bacteriology* **193**, 6342-6357.

Kałużna M, Pulawska J, Waleron M, Sobiczewski P. 2014. The genetic characterization of *Xanthomonas arboricola* pv. *juglandis*, the causal agent of walnut blight in Poland. *Plant Pathology* **63**, 1404-1416.

Katzen F, Ferreiro DU, Oddo CG, Ielmini MV, Becker A, Puhler A, Ielpi L. 1998. *Xanthomonas campestris* pv. *campestris gum* mutants: Effects on xanthan biosynthesis and plant virulence. *Journal of Bacteriology* **180**, 1607-1617.

Konstantinidis KT, Tiedje JM. 2005. Genomic insights that advance the species definition for prokaryotes. *Proceedings of the National Academy of Science of the United States of America* **102**: 2567-2572.

Lamichhane JR. 2014. *Xanthomonas arboricola* diseases of stone fruit, almond, and walnut trees: Progress toward understanding and management. *Plant Disease* **98**, 1600-1610.

Lee BM, Park YJ, Park DS, Kang HW, Kim JG, Song ES, Park IC, Yoon UH, Hahn JH, Koo BS, Lee GB, Kim H, Park HS, Yoon KO, Kim JH, Jung CH, Koh NH, Seo JS, Go SJ. 2005. The genome sequence of *Xanthomonas oryzae* pathovar *oryzae* KACC10331, the bacterial blight pathogen of rice. *Nucleic acids research* **33**, 577-586.

Lerat E, Daubin V, Moran N. 2003. From gene trees to organismal phylogeny in prokaryotes: the case of the gamma-Proteobacteria. *PLoS Biology* **1**, e19.

Leyns F, Decleene M, Swings JG, Deley J. 1984. The host range of the genus *Xanthomonas*. *Botanical Review* **50**, 308-56.

López MM, Lopez-Soriano P, Garita-Cambronero J, Beltrán C, Taghouti G, Pothier P, Cubero J, Fischer-Le Saux M, Marco-Noales E. 2018. *Xanthomonas prunicola* sp. nov., a novel pathogen that affects nectarine (*Prunus persica* var. nectarina) trees. *International Journal of Systematic and Evolution Microbiology* **68**, 1857-1866.

Malamud F, Torres PS, Roeschlin R, Rigano LA, Enrique R, Bonomi HR, Castagnaro AP, Marano MR, Vojnov AA. 2011. The *Xanthomonas axonopodis* pv. *citri* flagellum is required for mature biofilm and canker development. *Microbiology* **157**, 819-829.

Medini D, Donati C, Tettelin H, Masignani V, Rappuoli R. 2005. The microbial pan-genome. *Current Opinion in Genetics & Development* **15**, 589-594.

Meyer DF, Bogdanove AJ. 2009. Genomics-driven advances in *Xanthomonas* biology. In: Jackson R, ed. Plant Pathogenic Bacteria: Genomics and Molecular Biology. Norfolk: Caister Academic Press, 147-161.

Newman MA, Sundelin T, Nielsen JT, Erbs G. 2013. MAMP (microbe-associated molecular pattern) triggered immunity in plants. *Frontiers in Plant Science* **4**, 1–14.

Nourdin-Galindo G, Sánchez P, Molina CF, Espinoza-Rojas DA, Oliver C, Ruiz P, Vargas-Chacoff L, Cárcamo JG, Figueroa JE, Mancilla M, Maracaja-Coutinho V, Yanez AJ. 2017. Comparative pan-genome analysis of *Piscirickettsia salmonis* reveals genomic divergences within Genogroups. *Frontiers in Cellular and Infection Microbiology* **7**, 459.

Parkinson N, Aritua V, Heeney J, Cowie C, Bew J, Stead D. 2007. Phylogenetic analysis of *Xanthomonas* species by comparison of partial gyrase B gene sequences. International *Journal of Systematic and Evolutionary Microbiology* **57**, 2881-2887.

Pierce NB, 1901. Walnut bacteriosis. *Botanical Gazette* **31**, 272-273.

Pritchard L, Glover RH, Humphris S, Elphinstone JG, Toth IK. 2016. Genomics and taxonomy in diagnostics for food security: soft-rotting enterobacterial plant pathogens. *Analytical Methods* **8**, 12-24.

Quibod IL, Perez-Quintero A, Booher NJ, Dossa GS, Grande G, Szurek B, Cruz CV, Bogdanove AJ, Oliva R. 2016. Effector diversification contributes to *Xanthomonas oryzae* pv. *oryzae* phenotypic adaptation in a semi-isolated environment. *Scientific Reports* **6**, 34137.

Rouli L, Merhej V, Fournier PE, Raouly D. 2015. The bacterial pangenome as a new tool for analysing pathogenic bacteria. *New Microbe and New Infect* **7**, 72-85.

Ryan RP, Vorholter FJ, Potnis N, Jones JB, Van Sluys MA, Bogdanove AJ, Maxwell Dow J. 2011. Pathogenomics of *Xanthomonas*: understanding bacterium–plant interactions. *Nature reviews* **9**, 344-355.

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

Shrivastava S, Mande SS. 2008. Identification and functional characterization of gene components of Type VI Secretion system in bacterial genomes. *PloS one* **3**, e2955.

Smith RE, Smith CO, Ramsey HJ, 1912. Walnut culture in California : walnut blight. Berkeley, California: California Agricultural Experiments Station Publications.

Swings JG, Civerolo EL. 1993. *Xanthomonas*. London: Chapman & Hall.

Szczesny R, Jordan M, Schramm C, Schulz S, Cogez V, Bonas U, Büttner D. 2010. Functional characterization of the Xcs and Xps type II secretion systems from the plant pathogenic bacterium *Xanthomonas campestris* pv. *vesicatoria*. *New Phytologist* **187**, 983-1002.

Vauterin L, Hoste B, Kersters K, Swings J. 1995. Reclassification of *Xanthomonas*. *International Journal of Systematic Bacteriology* **45**, 472-89.

Vauterin L, Rademarker J, Swings J. 2000. Synopsis on the taxonomy of the genus *Xanthomonas*. Phytopathology **90**, 677-682.

Vorhölter FJ, Schneiker S, Goesmann A, Krause L, Bekel T, Kaiser O, Linke B, Parschkowski T, Ruckert C, Shmid J, Sidhu VK, Sieber V, Tauch A, Watt SA, Weisshaar B, Becker A, Niehaus K, Puhler A. 2008. The genome of *Xanthomonas campestris* pv. *campestris* B100 and its use for the reconstruction of metabolic pathways involved in xanthan biosynthesis. *Journal of biotechnology* **134**, 33-45.

Wadhams GH, Armitage JP. 2004. Making sense of it all: bacterial chemotaxis. *Nature Reviews Molecular Cell Biology* **5**, 1024–1037.

White FF, Potnis N, Jones JB, Koebnik R. 2009. The type III effectors of *Xanthomonas*. *Molecular plant pathology* **10**, 749-766.

Young JM, Dye DW, Bradbury JF, Panagopoulos CG, Robbs CF. 1978. A proposed nomenclature and classification for plant pathogenic bacteria. *New Zealand Journal of Agricultural Research* **21**, 153-77.

Young JM, Park DC, Shearman HM, Fargier E. 2008. A multilocus sequence analysis of the genus *Xanthomonas*. *Systematic and Applied Microbiology* **31**, 366-377.

Supplementary material

The supplementary material of Chapter V is available in one excel file named Best BLAST results_virulence genes.

CHAPTER VI

First report of *Xanthomonas arboricola* causing bacterial blight on pecan trees

Chapter VI include the following publication:

Fernandes C, Sousa R, Tavares F, Cruz L. 2018. First report of *Xanthomonas arboricola* causing bacterial blight on pecan trees in Portugal. Plant Disease 102 (12): 2632.

First report of *Xanthomonas arboricola* causing bacterial blight on pecan trees

Disease note

Pecan (*Carya illinoinensis*) is an economically important nut-producing tree in North America and widely cultivated in many countries (Thompson and Conner 2012), with a smaller commercial relevance in Portugal. During field surveys performed in June 2014 and July 2016, necrotic lesions were observed on leaves of pecan trees in an orchard in Alcobaça, Portugal. The lesions were similar to those on walnut (*Juglans regia*) caused by *Xanthomonas arboricola* pv. *juglandis*. Symptoms were characterized by small brown spots that coalesced to form necrotic areas surrounded by chlorotic halos (Supplementary **Figure VI. S1**). No damage was observed on fruits under these field conditions. Lesions from leaves collected during the two surveys were plated on yeast extract dextrose calcium carbonate medium and incubated for four days at 26°C. Yellow, mucoid colonies typical of *Xanthomonas* spp. were then streaked on a fresh nutrient agar medium to obtain pure colonies. Isolates were able to induce a hypersensitivity response on *Nicotiana benthamiana* leaves 72 h after inoculation (Supplementary **Figure VI. S2**) (Lelliott and Stead 1987). A portion of the *gyrB* gene from two isolates (CPBF 765, Genbank accession no. MG897454 and CPBF 1494, MG897455) was sequenced with primers XgyrB1F and XgyrB1R (Young et al. 2008). Both isolates were assigned to *X. arboricola* showing 98 to 99% identity with sequences from *X. arboricola* pathovars. Pathogenicity tests were performed on *C. illinoinensis* and *J. regia* 'Hartley', both members of the family Juglandaceae. For each of the two isolates, three plantlets from each species, with at least four fully expanded leaves, were inoculated by spraying with bacterial suspensions of 10^8 CFU/mL, OD 0.1. Plants inoculated with *X. arboricola* pv. *juglandis* LMG 747^T were used as positive controls and plants sprayed with sterile distilled water as negative controls. All plantlets were kept under a 16-h photoperiod (24°C day and 18°C night) and observed daily for four weeks. Spots on walnut and pecan leaves were first observed one week after inoculation with CPBF 765 and CPBF 1494. Disease developed progressively forming small necrotic areas surrounded by chlorotic halos. Similar symptoms were produced on the positive controls but absent on the negative controls. Bacterial colonies showing (Supplementary **Figure VI. S3**) characteristics similar to the original isolates were re-isolated from both plant species tested. In addition, multiplex PCR for detection of *X. arboricola* pv. *juglandis* (Fernandes et al. 2017) was negative for both strains. The genetic relationship of CPBF 765 and CPBF 1494 was further compared with representative strains of *X. arboricola* by

multilocus sequence analysis using partial sequences of seven genes (Fischer-Le Saux et al. 2015): *atpD*, Genbank accession nos. MG897444 - MG897445; *dnaK*, MG897446 - MG897447; *efp*, MG897448 - MG897449; *fyuA*, MG897450 - MG897451; *glnA*, MG897452 - MG897453; *gyrB*, MG897454 - MG897455 and *rpoD*, MG897456 - MG897457. *X. arboricola* CPBF 765 and CPBF 1494 were genetically distant from monophyletic clusters of *Xanthomonas* pathovars *pruni*, *corylina*, *juglandis* and *populi*. The isolates did not cluster together, making allocation to pathovar level unlikely (Supplementary **Figure VI. S4**). CPBF 1494 was more closely related to *X. arboricola* CFBP 1022 isolated from *J. regia*, whereas CPBF 765 was more closely related to *X. arboricola* CFBP 4021 and *X. arboricola* pv. *celebensis* CFBP 7150 isolated from *Magnolia* spp. and *Musa* spp., respectively (Fischer-Le Saux et al. 2015). This first report of *X. arboricola* causing leaf blight symptoms on pecan trees is a warning to countries with important pecan-growing areas.

Acknowledgments

Work co-financed by the European Structural & Investment Funds (ESIFs) through the Operational Competitiveness and Internationalization Programme – COMPETE 2020 and by National Funds through FCT – Foundation for Science and Technology within the framework of the project EVOXANT (PTDC/BIA-EVF/3635/2014, POCI-01-0145-FEDER-016600). Camila Fernandes is supported by a fellowship from FCT (SFRH/BD/95913/2013).

References

Fernandes C, Albuquerque P, Sousa R, Cruz L, Tavares F. 2017. Multiple DNA markers for identification of *Xanthomonas arboricola* pv. *juglandis* isolates and its direct detection in plant samples. *Plant Disease* **101**, 858-865.

Fischer-Le Saux M, Bonneau S, Essakhi S, Manceau C, Jacques MA. 2015. Aggressive emerging pathovars of *Xanthomonas arboricola* represent widespread epidemic clones distinct from poorly pathogenic strains, as revealed by multilocus sequence typing. *Applied and Environmental Microbiology* **81**, 4651-4668.

Lelliott RA, Stead DE. 1987. Methods for the diagnosis of bacterial diseases of plants. Methods in Plant Pathology. Oxford, UK: British Society for Plant Pathology by Blackwell Scientific Publications.

Thompson TE, Conner PJ. 2012. Pecan. Fruit Breeding. Boston, MA, USA: Springer.

Young JM, Park DC, Shearman HM, Fargier E. 2008. A multilocus sequence analysis of the genus *Xanthomonas*. *Systematic and Applied Microbiology* **31**, 366-77.

Supplementary material

Figure VI.S1. Natural symptoms of bacterial leaf spot of pecan trees observed in Alcobaça Portugal.

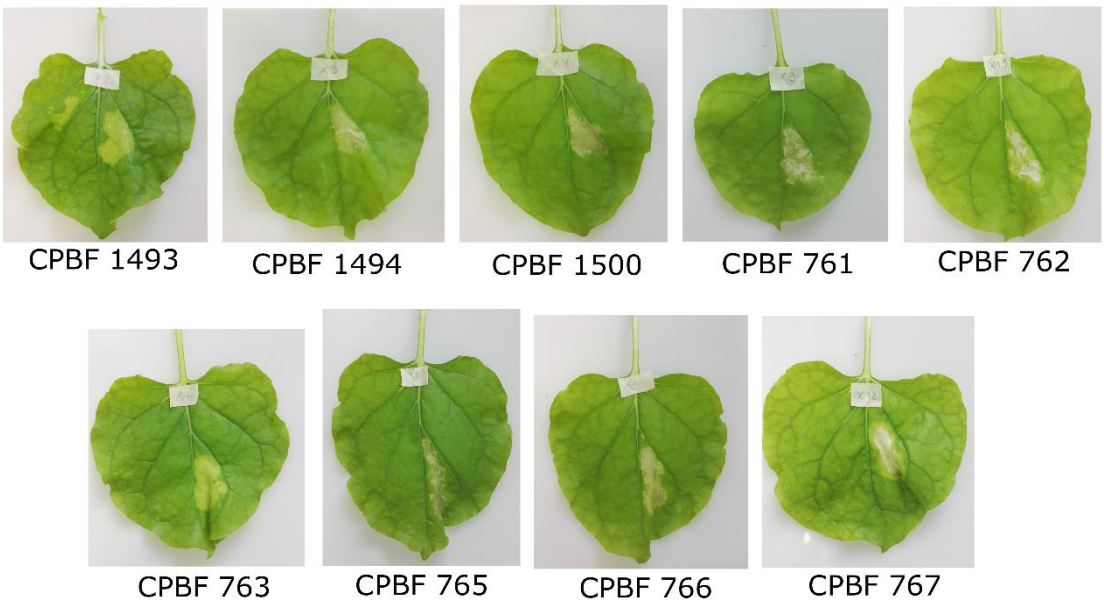

Figure VI. S2. Hypersensitivity response symptoms on leaves of *Nicotiana benthamiana* after 72 h after inoculation with the nine isolates obtained from diseased leaves of *Carya illionensis.*

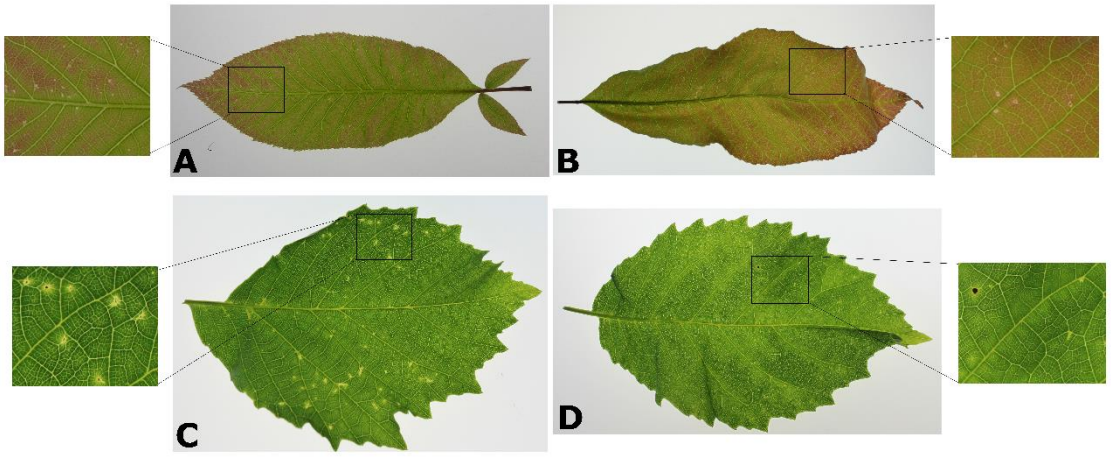

Figure VI. S3. Results from pathogenicity tests registered four weeks after inoculation. Small leaf lesions characterized by spots with chlorotic halos produced by strain CPBF 765 on pecan (A) and walnut (C) leaflets and by strains CPBF 1494 on pecan (B) and walnut (G) leaflets.

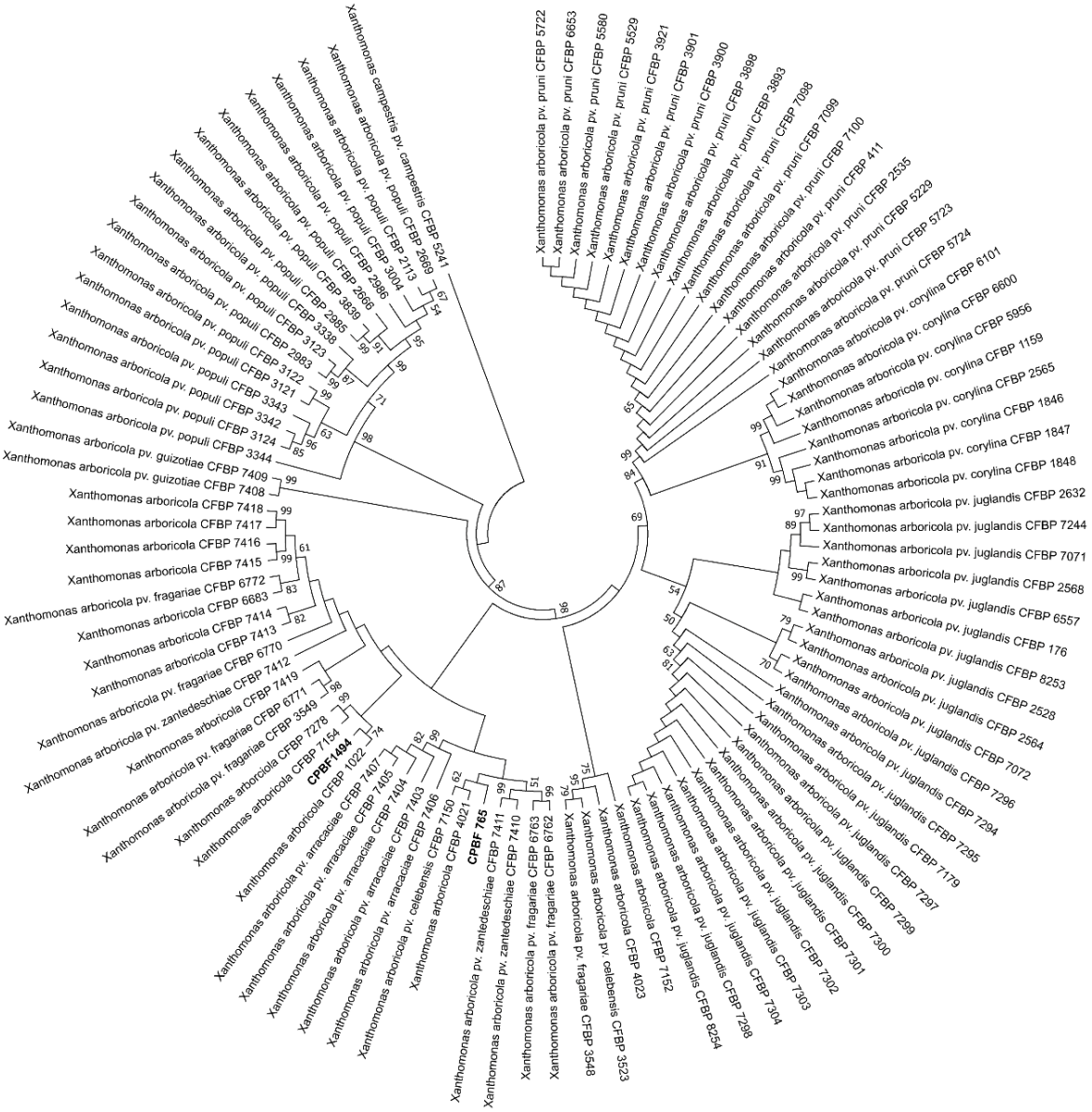

Figure VI.4. Maximum Likelihood tree based on 100 concatenated sequences of genes *atpD*, *dnaK*, *efp*, *fyuA*, *glnA*, *gyrB*, and *rpoD*, using the Tamura-Nei model in MEGA 7.0. The tree includes *X. arboricola* CPBF 765 and CPBF 1494 (highlighted in bold) and sequences of representative strains of *X. arboricola* used in a previous phylogenetic study (Fischer-Le Saux et al. 2015). Bootstrap values higher than 50 are shown.

CHAPTER VII

General Discussion

General discussion

Xaj is currently the most important bacterial pathogen of walnut and a major threat to all walnut cultivation regions worldwide. The severity pattern of walnut diseases caused by *Xaj* seems to have changed over the last decades, probably due to diverse unknown factors related to epidemiologic and environmental determinants, as well as to host and pathogen features. Indeed, until the end of the 1990s, *Xaj* remained a pathogen known to affect fruit quality but without significant losses for walnut crops. When *Xaj* emerged as etiological agent of new walnut pathologies, namely BAN and VOC, responsible for the reduction of orchard's productivity and an increase in nursery production costs, the pathogen triggered an immediate and higher socio-economic impact. These negative effects have been observed mainly in Euro-Mediterranean regions, where frequent disease outbreaks caused by *Xaj* were documented (Frutos 2010, Moragrega et al. 2011, Lamichhane 2014). Consequently, numerous studies have been conducted to understand the biology of this walnut pathogen, aiming to mitigate the disease impacts on walnut yields and help to implement scientifically informed containment practices (Hajri et al. 2010, Kaluzna et al. 2014, Lamichhane 2014, Giovanardi et al. 2015, Ivanovic et al. 2015). Although the knowledge gathered lead to improvements regarding disease management practices, and to a better understanding of pathogen´s diversity based on state-of-the-art molecular approaches (Hajri et al. 2010, Kaluzna et al. 2014, Giovanardi et al. 2015), not much is known about the underlying causes of walnut diseases caused by *Xaj*.

The main objective of this PhD project was to identify new molecular markers and propose new molecular approaches to gain further insights into the diversity, population dynamics, pathogenicity traits, as well as host and environmental adaptation of *Xaj*. These data should ultimately contribute to detail epidemiological patterns of walnut diseases caused by *Xaj*. The project framework consisted in the identification and validation of *Xaj* specific DNA markers for culture-independent *Xaj* detection and genotyping, complemented with the use and optimization of frontline molecular techniques (multiplex PCR, qPCR and dot blot), to accomplish the following:

i) Propose a workflow to allow the fast screening of numerous suspected plant samples, which is essential for an extensive, rapid and cost-effective phytosanitary inspection; also capable to unveil the presence of the pathogen outside its main host.

ii) Identify the pathogen and determine distinct genotypes through the presence/absence profile of taxa-specific markers and of virulence loci.

iii) Disclose epidemiological profiles taking into account metadata, namely of thermoclimatic regions, and host-related features.

iii) Quantify the load of infective bacteria as an accurate measure of its virulence,

iv) Disclose new genetic determinants of virulence and adaptation through comparative genomics of isolates colonizing the same walnut host tree.

In this chapter the major findings of this Thesis, detailed in chapters II to VI, are summarized, emphasizing the main scientific contributions and, their relevance towards the improvement of phytosanitary practices and to instruct future research.

1. Improving *Xaj* detection, identification and typing

When this PhD project started, there were no available specific DNA markers to detect *Xaj* with exception for a pair of primers which was not publicly disclosed to the scientific community (Gironde et al. 2009). Thus, the main ambition of this work was to identify and validate a set of novel molecular markers specific for pathovar *juglandis* that could be used in PCR and hybridization detection-based approaches.

The search to find promising DNA markers, capable of discriminating pathovar *juglandis* from all the other *X. arboricola* pathovars, was a challenge. Firstly, due to the phylogenetic ambiguity within the species *X. arboricola*, i.e. between the different pathovars, and secondly, due to the high genetic diversity within pathovar *juglandis* (Vauterin et al. 1995, Fischer-Le Saux et al. 2015). In fact, pathovar *juglandis* is characterized by an unusual genetic diversity in comparison with other closely related *X. arboricola* pathovars (Zaccardeli et al. 1999, Boundon et al. 2005, Fischer-Le Saux et al. 2015). Members of the *X. arboricola* species, regardless the pathovar, have a high percentage of DNA-DNA hybridization homology (79±15%, Vauterin et al. 1995), suggesting that all pathovars are probably derived from a common ancestor. Having this in mind, the chosen DNA markers had to be pathovar specific to discriminate *Xaj* from closely related pathovars and, if possible, capable of discriminating between different *Xaj* strains, which add genotyping resolution.

The release of the first *Xaj* genome (strain NCPPB 1447; Accession PRJNA84273) allowed the selection of several putative taxa-specific DNA markers. These genomic regions were selected by carrying out a whole genome BLAST screening, complemented by bioinformatics assessment according to Albuquerque et al (2012b), as fully detailed in Chapter II (Fernandes et al. 2017). This work lead to the identification of nine loci putatively suitable as DNA markers, considering their genomic stability and suitability for PCR and hybridization methods of detection (Chapter II, Fernandes et al. 2017). The follow up laboratorial validation confirmed the

specificity of these DNA markers, designated as XAJ1 to XAJ9, and most importantly their efficiency in culture-independent PCR detection methods. Aside from the major contribution of a set of *Xaj* specific DNA markers, previously unavailable, these DNA markers allowed to increase the reliability of *Xaj* detection and allowed to design assays with genotyping potential, due to the recognition of hybridization patterns capable of discriminating distinct *Xaj* strains. Beyond their immediate utility for robust *Xaj* detection from plant material, these *Xaj*-specific markers may be useful to validate other detection methodologies, such as LAMP, or informative to carry out screening diagnostics of walnut diseases caused by particularly virulent *Xaj* lineages. In this Thesis, we proposed three molecular approaches, namely multiplex-PCR, qPCR, and dot blot hybridization, to tackle different questions and aiming to build up tools for future research.

In Chapter II (Fernandes et al. 2017), the rationale to implement a multiplex PCR method using three markers (XAJ1, XAJ6 and XAJ8) was to validate an efficient culture-independent *Xaj* detection assay, suitable to be used in leaves and fruits of suspected diseased walnut samples allowing a fast detection of *Xaj*. Attending that direct detection of *Xaj* in plant material has never been described, multiplex PCR is a simple procedure, not requiring advanced technical and analytical skills, and available to any basic plant pathology laboratory. The low detection limit allows to detect low numbers of the bacterial pathogen, even in asymptomatic plant material, which might be particularly important for timely diagnostics and for a watchful surveillance of *Xaj* outside the walnut host, namely in complex environment samples (soil, contaminated water, plant debris), in agricultural equipment, and in asymptomatic plant species or putative animal vectors. Attending to the high specificity, efficiency and consistency of the proposed multiplex PCR based in three broad range markers (XAJ1, XAJ6 and XAJ8) (Fernandes et al. 2017), we believe that this method can be a reliable alternative to the standard culture-dependent detection methods of *Xaj* currently used.

In chapter III (Martins et al. 2018), the optimization and validation of a quantitative real-time PCR (qPCR) procedure to estimate the load of *Xaj* in plant samples is described. Although real-time PCR has been widely used as a detection and identification technique for phytopathogens, mainly because of its obvious advantage regarding efficiency and rapidity, its potential to quantitatively determine the number of infectious bacterial cells in plant samples has been disregarded. The main objective of this work was to develop a method to quantify the bacterial load of *Xaj* in samples of diseased walnut fruits. Using two *Xaj* broad range DNA markers (XAJ1 and XAJ6) in order to increase the reliability of the procedure, the method described showed to be suitable for rapid and direct quantification of *Xaj* in walnut leaves and fruits and adequate for a high sample throughput. The method showed to have a detection limit estimated as 2.7 CFU/qPCR reaction, therefore capable to determine the degree of *Xaj* colonization on walnut organs and suitable to monitor the disease progression, particularly during

the fruit development, when *Xaj* growth could vary as a result of environmental factors such as humidity and temperature (Oliveira and Cabral 1942, Giraud et al. 2010). Furthermore, the availability of a method capable of quantifying the population size of infective *Xaj* strains on walnut plant samples, is a powerful tool to evaluate the virulence fitness of different *Xaj* strains in walnut plant tissues, and a contribution to identify particularly virulent *Xaj* pathotypes. The qPCR method proposed, besides its diagnostic potential for *Xaj* associated walnut diseases, allows evaluating the colonization of *Xaj* strains as a measure of virulence fitness in walnut infected organs, and is a step forward to make disease risk assessments, and consequently instruct assertive containment practices and efficient phytosanitary treatments.

Notwithstanding the advantages of multiplex PCR and qPCR as complementary strategies for a reliable detection and quantification of *Xaj* in plant samples, their utility to further discriminate rapidly different *Xaj* genotypes is limited. Both methods are based on *Xaj* specific broad range DNA markers (XAJ1, XAJ6, and XAJ8), i.e. DNA markers generally present in all *Xaj* strains regardless their genotype. The remaining six DNA markers (XAJ2, XAJ3, XAJ4, XAJ5, XAJ7, and XAJ9), although shown to be *Xaj* specific were not present in all *Xaj* strains and isolates assayed, and therefore unsuitable to be used as single DNA markers in *Xaj* detection assays, due to the risk of false negative results. Despite this limitation, when taken together, the nine *Xaj*-specific DNA markers showed to be a resourceful aproach to discriminate different *Xaj* strains. In order to analyse the nine DNA markers simultaneously on a large number of strains, we describe in Chapter II a dot blot hybridization platform as a rapid and reliable technique for validation of the DNA markers (Fernandes et al. 2017). Due the different patterns of presence/absence of the nine DNA-markers in distinct *Xaj* strains, we propose in Chapter IV the use of this dot blot hybridization platform as a fast and high-throughput genotyping method (Fernandes et al. 2018a). In fact, dot blot hybridization using the nine markers in a large collection of strains and isolates allowed to discriminate 20 distinct hybridization patterns (HP). Furthermore, distinct HP were identified within the same MLSA cluster based on four housekeeping genes (*acnB*, *fyuA*, *gyrB*, and *rpoD*), contributing to an enhanced genotyping efficiency, capable of distinguishing bacterial lineages of closely related xanthomonads associated to walnut trees (Fernandes et al. 2017, Fernandes et al. 2018a). Dot blot was shown to provide fast and reproducible typing results, being comparatively cheaper for laboratory usage than other hybridization techniques, such as DNA microarrays (Ehrenreich 2006). Furthermore, the existence of an algorithm for automatic data analyses of dot blot images (Caridade et al. 2010), facilitates inter-laboratorial and long-term comparison studies, allowing the detection and identification of bacterial xanthomonads populations on walnut with the intention to follow new emerging pathogenic lineages, recent outbreaks and new routes of transmission.

2. Xanthomonas arboricola pv. *juglandis* in Portugal, current situation

Portuguese walnut orchards are currently threatened by various pests, namely aphids (*Chromaphis juglandicola*, *Myzus persicae*, *Panaphis juglandis*) (Cecilio and Ilharco 1997, Blackman and Eastop 2018), walnut husk fly (Coutinho 2014) and walnut codling moth (Assunção 1998). Several diseases are also a major threat, particularly root and crown rot of walnut (Moreira and Coutinho 2017), walnut anthracnose caused by *Gnomonia leptostyla* (Moureira and Coutinho 2008a), walnut bacterial blight and brown apical necrosis (Assunção 1999, Moureira and Coutinho 2008b).

Walnut bacterial blight and brown apical necrosis are the most relevant diseases of walnut in Portugal, however "bacteriose da nogueira" is the name commonly used to report the presence of both necrotic spots on leaves and fruits and the apical necrosis lesions near the blossom end of the fruits. Therefore, a differentiation between symptoms associated to WBB and BAN diseases has not been described, which probably led to difficulties in reporting cases of BAN disease. Indeed, available data suggest that BAN was occurring in Portugal before 2000 (Martins et al. 1997, Moureira and Coutinho 2008), but the first evidence of BAN disease in Portugal is firstly documented in this work (Chapter II, Fernandes et al. 2017).

The extensive sampling carried out during this PhD project and detailed in Chaper IV, covered distinct thermoclimatic regions in Portugal and particularly across the main walnut producing regions, taking into account host-related metadata. This methodology allowed to uncover the real and worrying epidemic status of walnut diseases caused by *Xanthomonas* in Portugal. The walnut plantations found across the country are located in different geographic locations and exposed to distinct thermoclimatic conditions, characterized by distinct temperature and humidity regimes, known to be particularly important for *Xaj* epidemiology (Oliveira and Cabral 1942, Giraud et al. 2010). Walnut trees were found either dispersed or growing in walnut orchards, frequently composed of a single cultivar or formed by a mix of different cultivars. Sampled walnut trees included traditionally used cultivars, particularly "Amigo" and "Lara", or cultivars more recently introduced as is the case of "Hartley", "Chandler", and "Franquette". Adding to these variables, the use of phytosanitary treatments was generally neglected, while different management practices were found in walnut orchards, as well as in old orchards non mobilized. Whatever the sampled orchard, walnut trees showing common symptoms of WBB and BAN diseases were frequently observed. Furthermore, despite the different prophylactic measures applied on managed orchards, including proper cultural practices and/or application of copper-based compounds, infected trees showed low yields of poor-quality nuts or an intensive premature fruit drop, dramatically affecting productivity, leading to a general concern among Portuguese walnut producers.

The in-depth study to determine the population diversity of *Xaj* found in walnut growing regions of Portugal detailed in Chapter IV (Fernandes et al. 2018a), involved the largest number of xanthomonads isolates ever reported in a single study, including isolates obtained from the same walnut tree. Additionally, this study considered the potential factors that influence *Xaj* dissemination and population diversity. The results showed that *Xaj* are widespread in Portugal, whatever the thermoclimatic or geographic location and apparently independent of walnut host features. Besides the wake-up call for the urgency to implement effective phytosanitary measures, this study provides insights into population structure and infrasubspecific diversity of *Xaj*, confirming the high genetic diversity of pathovar *juglandis*, and showing that *Xaj* pathogenic lineages are likely a fraction of walnut-associated xanthomonads, frequently co-colonizing the same walnut tree. In addition, evidence was gathered dismissing the generally acknowledged paradigm of a homogeneous within-host population of *Xaj*. In fact, the etiology of *Xanthomonas*-associated walnut diseases seems to involve a consortium of distinct *Xanthomonas* lineages, not necessarily belonging to *Xanthomonas arboricola* species, or pathovar *juglandis*.

Overall, this study (Chapter IV, Fernandes et al. 2018a) is a solid foundation to identify persistent, dominant or acutely virulent walnut-associated *Xanthomonas* lineages linked with recent epidemics of a given walnut production region. Additionally, the results of this study (Fernandes et al. 2018a) emphasized the importance of a well-planned sampling methodology combined with informative metadata to properly characterize the epidemiology of infectious plant diseases.

3. Comparative genomics of a consortium of five xanthomonads isolated from a single walnut tree

The evidence of mixed xanthomonads lineages occurring on approximately half of the symptomatic walnut trees sampled documented in **Table IV.1** of Chapter IV (Fernandes et al. 2018a), raised questions about the role played by each of different *Xanthomonas* lineages co-colonizing the same tree and even the same plant organ. To comprehensively disclose lineage-specific traits that may provide clues about the functional meaning of a *Xanthomonas* consortium in walnut infection, we sequenced the genomes of five CPBF *Xanthomonas* isolates (CPBF 367, CPBF 424, CPBF 426, CPBF 427, and CPBF 1521) obtained from the same walnut tree (Fernandes et al. 2018b, 2018c, 2018d). Moreover, these isolates were assigned to four distinct lineages according to MLSA and dot blot hybridization, namely lineages L31, L30, L25, for strains CPBF 367, 424 and 426, respectively, and lineage L14 for strains CPBF 427 and 1521, which were isolated in two different years from buds and leaves, respectively. Although, a few genomic comparison studies of *X. arboricola* have been published (Cesbron et al. 2015, Garita-

Cambronero et al. 2018), these papers focused on strains epidemiologically unrelated and unsuitable to make assumptions about *Xanthomonas* diversity within the same walnut host specimen.

The work presented in Chapter V provides the general genome description of strains CPBF 424 and CPBF 1521 (Fernandes et al. 2018b, Fernandes et al. 2018c) and the preliminary findings regarding comparative genomics analyses of the five sequenced *Xanthomonas* isolates (Fernandes et al. 2018d).

The whole-genome sequence of the five isolates analysed with EDGAR (Blom et al. 2009, 2016), a resourceful bioinformatics platform for comparative genomics, allowed to infer the taxonomic positioning of the five sequenced *Xanthomonas* isolates, in comparison with 25 other sequenced *Xanthomonas arboricola*, including five pathovar *juglandis* strains, and 12 non-*arboricola Xanthomonas*. Furthermore, the establishment of the core genome from the five CPBF *Xanthomonas* isolates and of the accessory genome, either strain-specific or shared by a subset of isolates, facilitated the search for strain-specific genomic features with emphasis for virulence-related loci or strain-specific adaptive traits. This knowledge might help to hypothesize about strain-specific functions that may confer adaptive advantageous in a *Xanthomonas* consortium sharing the same niche.

MLSA genotyping showed that sequenced isolates cluster in two distinct groups (CPBF 427 and CPBF 1521 were allocated to MLSA cluster VI, while CPBF 367, CPBF 424, and CPBF 426 were assigned to MLSA cluster X, Fernandes et al. 2018a). General genomic features revealed that CPBF 427 and CPBF 1521 possess genomes matching the genomes of other *Xaj* strains regarding its size, CDSs and richness of IS elements. On the contrary, CPBF 367, CPBF 424 and CPBF 426 genomes are considerable smaller, with over 400 less CDSs in total, and intriguingly a six-fold reduction in the number of IS elements. These results were further confirmed by whole-genome phylogeny of the core genome from 44 *Xanthomonas*, and by Average Nucleotide Identity (ANI) analyses which suggested that, while CPBF 427 and CPBF 1521 are, in fact, *Xaj*, the isolates CPBF 367, CPBF 424, and CPBF 426 belong to a new *Xanthomonas* species (Fernandes et al. 2018b, 2018c, 2018d).

The comparative genomics analysis of the five sequenced isolates revealed that the most striking differences were related with T3SS and T3E. As expected, isolates CPBF 427 and 1521 have a pattern of T3SS and T3E genes similar to what is found in other *Xaj* genomes (Cesbron et al. 2015), contrary to isolates CPBF 367 and 426, which are deficient for the majority of the T3SS and T3E coding genes, similar to atypical nonpathogenic strains of *Xanthomonas* sp. previously described (Cesbron et al. 2015, Garita-Cambronero et al. 2017). Intriguingly, CPBF 424 seems to have a functional T3SS and holds a narrow gene repertoire of T3E homologs in comparison with other *Xaj* (Hajri et al. 2012, Cesbron et al. 2015, Essakhi et al. 2015, Fernandes

et al. 2018d). Walnut infection assays showed that this isolate is pathogenic for walnut (Fernandes et al. 2018a), contrary to what has been observed for other atypical walnut-associated *Xanthomonas*, frequently considered nonpathogenic and described as epiphytic and commensal (Essakhi et al. 2015). Taken together, the genomic features of these three isolates (CPBF 367, 424, and 426), their phylogenomic and taxonomic closeness, and the pathogen phenotype of isolate CPBF 424 in walnut, makes these strains particularly appealing to elucidate the evolutionary hypothesis of pathogenicity in walnut, to uncover the genetic footprints of adaptation, and to address speciation in *Xanthomonas*. Ultimately, these studies may lead to design new markers for an effective and broad surveillance of walnut-associated *Xanthomonas* or to screen specific virulence genes. These features, coupled with epidemiologic metadata, will certainly contribute to deepen the knowledge concerning the epidemiology of walnut diseases caused by *Xanthomonas*, therefore providing robust scientific grounds to take meaningful phytosanitary actions.

4. Non-*Xaj* walnut-associated *Xanthomonas*

In addition to the dozens of isolates belonging to *X. arboricola* pathovar *juglandis*, the extensive sampling carried out during the course of this PhD project also allowed to isolate *Xanthomonas* strains positioned in new MLSA clusters, namely cluster IX and X, and *Xanthomonas* isolated from a non-walnut host (*Carya illionensis*, Family Juglandaceae). This set of strains included *Xanthomonas* isolates obtained from walnut, but not belonging to *Xanthomonas arboricola* species (Chapters IV and V, Fernandes et al. 2018a, 2018d), and strains isolated from pecan trees showing disease symptoms similar to the ones found in walnuts, which cannot be identified as pathovar *juglandis* (Chapter VI, Fernandes et al., 2018e). Some of these xanthomonads have been shown to induce typical symptoms in walnut leaves, indicating that they are pathogenic *Xanthomonas* (Fernandes et al. 2018a, 2018e). Furthermore these studies indicate that the diversity of *Xanthomonas* associated with walnut is higher than what has been described and, as a whole, probably represents a species complex characterized by different pathoadaptations, which likely include pathogenic strains from other xanthomonads species and pathogenic strains causing disease in other host species. Although the importance of these recently described pathogenic strains is unknown, they should be considered a new threat to walnut production. Ongoing studies of these strains are expected to bring new insights about pathoadaptation specificities.

In Chapter IV a deep genotyping analysis including MLSA and dot blot hybridization with nine DNA markers allowed to disclose walnut-associated *Xanthomonas* clustering separately from *Xaj* (Fernandes et al. 2018a). The whole-genome analyses described in Chapter V, indicate

that strains CPBF 367, CPBF 424, and CPBF 426 belong to new *Xanthomonas* species, and include a walnut-pathogenic strain (CPBF 424) (Fernandes et al. 2018b, 2018d). The polyphasic study that followed the genomic analysis of these *Xanthomonas* isolates (Chapter V, Fernandes et al. 2018d) and required for the description of a new bacterial species (Wayne et al. 1987, Stackebrandt et al. 2002, López et al. 2018) is still underway. These strains are being characterized for phenotypic features, including physiological and biochemical assays. In addition, the deposit of strains in European culture collections is already in progress and extended quality controls are being performed, including the analysis of the 16S rRNA and housekeeping genes sequences.

During the surveys on walnut trees, another species of the Juglandaceae family (Pecan trees, *Carya illionensis*) was found to be a host for *X. arboricola* strains (Chapter VI, Fernandes et al. 2018e). This is not surprising since *X. arboricola* strains, although generally considered to cause diseases on different tree species according the pathovar, have been isolated from ornamental and herbaceous plants (Jacques et al. 2016), which suggests a host range expansion of this species. In Chapter VI we showed that *X. arboricola* strains isolated from pecan trees were able to cause bacterial blight symptoms on leaves of both pecan and walnut plantlets, indicating that these strains can infect non-walnut hosts. Interestingly, other *X. arboricola* isolates collected from pecan trees showed to be pathogenic on walnut but not on their host of isolation (data not shown). These results showed that strains pathogenic to walnut could be isolated from leaves of pecan trees, suggesting a possible survival of walnut-pathogenic *X. arboricola* strains in other plant species without causing disease. These preliminary results highlight that alternative hosts might be implicated in the transmission of *Xaj* strains. A detailed comparative genomics study of *X. arboricola* strains obtained from pecan tree hosts with *Xaj* obtained from walnut will be useful to determine the relation of these strains and give insights into host adaptation.

Final remarks and future perspectives

In 2014, when this PhD research project was initiated the knowledge regarding *Xaj* was limited and mainly centred on studies to characterize the etiology of walnut diseases caused by *Xanthomonas arboricola*, to select for walnut cultivars resistant to *Xaj*, and to address the diversity of *Xaj*. This PhD Thesis gathers the findings obtained during my PhD project framed within an updated knowledge and recent scientific developments, which I believe resulted in valuable advances: i) on DNA-based detection and typing methodologies of *Xaj;* ii) on the diversity of walnut-associated *Xanthomonas* taking into account metadata thought to be important to understand the epidemiology of *Xanthomonas* causing walnut diseases; iii) the discovery of a new *Xanthomonas* species; and iv) to uncover genomic determinants of *Xanthomonas* pathoadaptations.

Altogether, these contributions established the grounds for further research, namely:

To uncover the lifecycle of walnut-associated *Xanthomonas*. Molecular methods employing the novel DNA markers specific for *Xaj* have proven to be reliable tools for culture-independent detection, and for quantitative identification and typing of isolates from diseased walnut hosts. Therefore, new opportunities are now available to search for walnut-associated *Xanthomonas* in environmental samples outside the walnut host, including alternative plant host species that may be important reservoirs; to differentiate strains and measure their adaptation fitness to walnut; to track the sources of disease outbreaks and to detail population structure.

To develop an epidemiology framework of walnut diseases caused by *Xanthomonas* using the comprehensive metadata gathered in this thesis. The long-term diversity study of *Xanthomonas* carried out across distinct thermoclimate regions and geographic locations, taking into account several walnut host features, was instrumental to disclose the presence of non-*arboricola Xanthomonas* pathogenic on walnut; to emphasize the importance of an extensive sampling; to show that co-colonization of the same walnut host tree is frequent and generally underestimated; and to review the environmental and host-related factors modulating the epidemiology walnut diseases caused by xanthomonads.

To unveil pathoadaptation determinants of *Xanthomonas* through comparative genomics. Whole genome sequencing studies of a consortium of five *Xanthomonas* strains isolated from the same walnut specimen, allowed to gather strong evidence for a new *Xanthomonas* species.

The variations observed among the five strains particularly regarding genome size and number of IS elements, raise questions about the likelihood of recombination events occurring within these sympatric populations of *Xanthomonas* and its adaptative and evolutionary significance, that deserve careful attention. Comparative genomics studies focused on pathogenicity and virulence related genes, provided new insights regarding pathoadaptations and highlighted the importance of T3SS and T3E, that need to be addressed in dedicated functional studies. Furthermore, the genome sequences will allow to design strain-specific and functional DNA markers which will be instrumental to better understand adaptation and evolutionary fitness, and consequently help to identify acutely virulent *Xanthomonas* strains, with implications regarding the improvement of efficient phytosanitary practices.

Ultimately, by the reasons emphasized above, this PhD dissertation is certainly a solid contribution to implement scientifically informed disease management and phytosanitary control measures aiming to decrease walnut production losses caused by xanthomonads infections.

References

Albuquerque P, Caridade CRM, Rodrigues AS, Marçal ARS, Cruz J, Cruz L, Snatos CL, Mendes MV, Tavares F. 2012b. Evolutionary and experimental assessment of novel markers for detection of *Xanthomonas euvesicatoria* in plant samples. *PLoS ONE* **7**, e37836.

Assunção A. 1998. Pragas da nogueira: O bichado da noz. Divisão de Vitivinicultura e Fruticultura. Ficha Técnica 66. DRAEDM Direcção Regional de Agricultura de Entre-Douro e Minho. [In Portuguese]

Assunção A. 1999. Bacteriose da nogueira (*Xanthomonas campestris* pv. *juglandis*). Estação EXP. De Viticultura e Fruticultura. Ficha Técnica 79. DRAEDM Direcção Regional de Agricultura de Entre-Douro e Minho. [In Portuguese]

Blackman RL, Eastop VF. 2018. Aphids on the world's plants. An online identification and information guide. http://influentialpoints.com/Gallery/Aphids_on_walnut_Juglans.htm. Accessed October 2018.

Blom J, Albaum SP, Doppmeier D, Puhler A, Vorholter FJ, Zakrzewski M, Goesmann A. 2009. EDGAR: a software framework for the comparative analysis of prokaryotic genomes. *BMC Bioinformatics* **10**, 154.

Blom J, Kreis J, Spanig S, Juhre T, Bertelli C, Ernst C, Goesmann A. 2016. EDGAR 2.0: an enhanced software platform for comparative gene content analyses. *Nucleic Acids Research* **44**, W22-W28.

Boudon S, Manceau C, Notteghem JL. 2015. Structure and origin of Xanthomonas arboricola pv. pruni populations causing bacterial spot of stone fruit trees in Western Europe. *Phytopathology* **95**, 1081-1088.

Caridade CRM, Marçal ARS, Albuquerque P, Mendes MV, Tavares F. 2015. Automatic analysis of dot blot images. *Intelligent Automation and Soft Computing* **21**, 1-16.

Cecilio A, Ilharco F. 1997. The control of walnut aphid, *Chromaphis juglandicola* (Homoptera: Aphidoidea) in walnut orchards in Portugal. *Acta Horticulturae* **442**, 399-406.

Cesbron S, Briand M, Essakhi S, Gironde S, Boureau T, Manceau C, Fischer-Le Saux M, Jacques MA. 2015. Comparative genomics of pathogenic and nonpathogenic strains of *Xanthomonas arboricola* unveil molecular and evolutionary events linked to pathoadaptation. *Frontiers in Plant Science* **6**, 1126.

Coutinho C. 2014. A mosca da casca verde da noz (*Rhagoletis completa* Cresson). Textos de divulgação técnica da Estação de Avisos de Entre Douro e Minho nº9_2014 (II Série). Ministério da Agricultura e do Mar/ DRAP-Norte/ Divisão de Apoio ao Setor Agroalimentar/ Estação de Avisos de Entre Douro e Minho. [In Portuguese]

Ehrenreich A. 2006. DNA microarray technology for the microbiologist: an overview. *Applied Microbiology and Biotechnology* **73**, 255-273.

Essakhi S, Cesbron S, Fischer-Le Saux M, Bonneau S, Jacques MA, Manceau C, 2015. Phylogenetic and variable-number tandem-repeat analyses identify nonpathogenic Xanthomonas arboricola lineages lacking the canonical Type III Secretion System. *Applied and Environmental Microbiology* **81**, 5395-410.

Fernandes C, Albuquerque P, Sousa R, Cruz L, Tavares F. 2017. Multiple DNA markers for identification of *Xanthomonas arboricola* pv. *juglandis* isolates and its direct detection in plant samples. *Plant Disease* **101,** 858-865.

Fernandes C, Albuquerque P, Cruz L, Tavares F. 2018a. Genotyping and epidemiological metadata provides new insights into population structure of Xanthomonas isolated from walnut trees. Submitted.

Fernandes C, Blom J,Pothier JF, Tavares F. 2018b. High-quality draft genome sequence of *Xanthomonas* sp. strain CPBF 424, a walnut pathogenic strain with atypical features. *Microbiology Resource Announcements* **7**, e00921-18.

Fernandes C, Blom J,Pothier JF, Tavares F. 2018c. High-quality draft genome sequence of *Xanthomonas arboricola* pv. *juglandis* CPBF 1521, isolated from leaves of a symptomatic walnut tree in Portugal without a past of phytosanitary treatment. *Microbiology Resource Announcements* **7**, e00887-18.

Fernandes C, Martins L, Blom J, Pothier JF. Tavares F. 2018d. Comparative genomic characterization of Xanthomonas strains isolated from a single walnut tree host. In preparation.

Fernandes C, Sousa R, Tavares F, Cruz L. 2018e. First report of *Xanthomonas arboricola* causing bacterial blight on pecan trees in Portugal. *Plant Disease* **102**, 2632.

Fischer-Le Saux M, Bonneau S, Essakhi S, Manceau C, Jacques MA. 2015. Aggressive emerging pathovars of *Xanthomonas arboricola* represent widespread epidemic clones that are distinct from poorly pathogenic strains, as revealed by multilocus sequence typing. *Applied and environmental microbiology* **81**, 4651-4668.

Frutos D. 2010. Bacterial diseases of walnut and hazelnut and genetic resources. *Journal of Plant Pathology* **92**, S79-S85.

Garita-Cambronero J, Palacio-Bielsa A, López MM, Cubero J. 2017. Pan-genomic analysis permits differentiation of virulent and non-virulent strains of *Xanthomonas arboricola* that cohabit *Prunus* spp. and elucidate bacterial virulence factors. *Frontiers in Microbiology* **8**, 573.

Garita-Cambronero J, Palacio-Bielsa A, Cubero J. 2018. *Xanthomonas arboricola* pv. *pruni*, causal agent of bacterial spot of stone fruits and almond: its genomic and phenotypic characteristics in the *X. arboricola* species context. *Molecular Plant Pathology* **19**, 2053-2065.

Giovanardi D, Bonneau S, Gironde S, Fischer-Le Saux M, Manceau C, Stefani E. 2015. Morphological and genotyping features of *Xanthomonas arboricola* pv. *juglandis* populations from walnut groves in Romagna region, Italy. *European Journal of Plant Pathology* **145**, 1-16.

Giraud M, Prunet JP, Chevallier A, Ramain S, Thiriad V, Santrac I, Bray O. 2010. Study of *Xanthomonas arboricola* pv. *juglandis* population dynamics in french walnut orchards over three years. *Acta Horticulturae* **861**, 439-444.

Gironde S, Guillaumes J, Manceau C. 2009. Specific detection of *Xanthomonas arboricola* pv. *juglandis* pathogen on walnut. EPPO Conference on Diagnostics and Associated workshops, York, U.K.

Hajri A, Meyer D, Delort F, Guillaumes J, Brin C, Manceau C. 2010. Identification of a genetic lineage within *Xanthomonas arboricola* pv. *juglandis* as the causal agent of vertical oozing canker of Persian (English) walnut in France. *Plant pathology* **59**, 1014-1022.

Hajri A, Pothier JF, Fischer-Le Saux M, Bonneau S, Poussier S, Boureau T, Poussier S, Boureau T, Duffy B, Manceau, C. 2012. Type three effector gene distribution and sequence analysis provide new insights into the pathogenicity of plant-pathogenic *Xanthomonas arboricola*. *Applied and Environmental Microbiology* **78**, 371-384.

Ivanović Ž, Popović T, Janse J, Kojić M, Stanković S, Gavrilović V, Fira D. 2015. Molecular assessment of genetic diversity of *Xanthomonas arboricola* pv. *juglandis* strains from Serbia by various DNA fingerprinting techniques. *European Journal of Plant Pathology* **141**, 133-145.

Jacques MA, Arlat M, Boulanger A, Boureau T, Carrère S, Cesbron S, Chen NWG, Cociancich S, Darrasse A, Denancé N, Fischer-Le Saux M, Gagnevin L, Koebnik R, Lauber E, Noel LD, Pieretti I, Portier P, Pruvost O, Rieux A, Robène MR, Szurek B, Verdier V, Vernière C. 2016. Using ecology, physiology, and genomics to understand host specificity in *Xanthomonas*: French Network on Xanthomonads (FNX). *Annual Review of Phytopathology* **54**, 163-187.

Kałużna M, Pulawska J, Waleron M, Sobiczewski P. 2014. The genetic characterization of *Xanthomonas arboricola* pv. *juglandis*, the causal agent of walnut blight in Poland. *Plant Pathology* **63**, 1404-1416.

Lamichhane JR. 2014. *Xanthomonas arboricola* diseases of stone fruit, almond, and walnut trees: Progress toward understanding and management. *Plant Disease* **98**, 1600-1610.

López MM, Lopez-Soriano P, Garita-Cambronero J, Beltrán C, Taghouti G, Portier P, Cubero J, Fischer-Le Saux M, Marco-Noales E. 2018. *Xanthomonas prunicola* sp. nov., a novel pathogen that affects nectarine (*Prunus persica* var. *nectarina*) trees. *International Journal of Systematic and Evolutionary Microbiology* **68**, 1857-1866.

Martins JMS, Pinto C, Gomes JA. 1997. Chemical control of bacterial blight of walnut. *Acta Horticulturae* **442**, 367-372.

Martins L, Fernandes C, Albuquerque P, Tavares F. 2018. Accurately determination of *Xanthomonas arboricola* pv. *juglandis* in infected walnut fruits by qPCR. Accepted.

Moragrega C, Matias J, Aletà N, Montesinos E, Rovira M. 2011. Apical necrosis and premature drop of Persian (English) walnut fruit caused by *Xanthomonas arboricola* pv. *juglandis*. *Plant Disease*, **95**, 1565-1570.

Moureira JFG, Coutinho C. 2008a. Nogueira. Antracnose da nogueira. Avisos Agrícolas. Circular nº1, 30.01.08. Direção Regional de Agricultura e Pescas do Norte. Estação de Avisos de Entre Doutro e Minho. [In Portuguese]

Moureira JFG, Coutinho C. 2008b. Bacteriose da nogueira *Xanthomonas campestris* pv. *juglandis.* Divisão de proteção e controlo fitossanitário. DRAP Norte Direcção Regional de Agricultura e Pescas do Norte. [In Portuguese]

Moureira JFG, Coutinho C. 2017. Nogueira – Doença da Tinta na Nogueira (*Phytophthora cinnamoni*). Avisos Agrícolas. Circular nº 01/2017. Direção Regional de Agricultura e Pescas do Norte. Estação de Avisos de Entre Doutro e Minho. [In Portuguese]

Oliveira M, Cabral R. 1942. Doenças bacterianas das plantas diagnosticadas em Portugal. *Revista Agronómica* **XXX**,176-184. [In Portuguese]

Stackebrandt E, Frederiksen W, Garrity GM, Grimont PA, Kampfer P, Maiden MC, Nesme X, Rosselló-Mora R, Swings J, Truper HG, Vauterin L, Ward AC, Whitman WB. 2002. Report of the Ad Hoc Committee for the re-evaluation of the species definition in bacteriology. *International Journal of Systematic Bacteriology* **52**, 1043-1047.

Vauterin L, Hoste B, Kersters K, Swings J. 1995. Reclassification of *Xanthomonas*. *International Journal of Systematic Bacteriology* **45**, 472-89.

Wayne LG, Brenner DJ, Colwell RR, Grimont PAD, Kandler O, Krichevsky MI, Moore LH, Moore WEC, Murray RGE, Stackebrandt E, Starr MP, Truper HG. 1987. Report of the Ad Hoc Committee on Reconciliation of Approaches to Bacterial Systematics. *International Journal of Systematic Bacteriology* **37**, 463-464.

Zaccardelli M, Ceroni P, Mazzucchi U. 1999. Amplified fragment length polymorphism fingerprinting of *Xanthomonas arboricola* pv. *pruni*. *Journal of Plant Pathology* **81**, 173-179.