UNIVERSIDAD POLITÉCNICA DE MADRID

Escuela Técnica Superior de Ingeniería Agronómica, Alimentaria y de Biosistemas

GENÓMICA COMPARATIVA DE CEPAS DE *Xanthomonas arboricola*
ASOCIADAS A *Prunus* spp.: CARACTERIZACIÓN DE LOS
PROCESOS DE INFECCIÓN DE LA MANCHA BACTERIANA DE
FRUTALES DE HUESO Y ALMENDRO.

Tesis Doctoral

Jerson Martín Garita Cambronero

Licenciado en Biología

2016

Departamento de Biotecnología

Escuela Técnica Superior de Ingeniería Agronómica, Alimentaria y de
Biosistemas

Universidad Politécnica de Madrid

Tesis Doctoral

Genómica comparativa de cepas de *Xanthomonas arboricola* asociada a
Prunus spp.: Caracterización de los procesos de infección de la mancha
bacteriana de frutales de hueso y almendro.

Autor

Jerson M. Garita Cambroner

Licenciado en Biología

Director

Jaime Cubero Dabrio

Dr. en Ciencias Biológicas

Directora

María Milagro López González

Dr. Ingeniero Agrónomo

UNIVERSIDAD POLITÉCNICA DE MADRID

Tribunal nombrado por el Magfco. y Excmo. Sr. Rector de la Universidad Politécnica de Madrid el día de de 201 .

Presidente

Secretario

Vocal

Vocal

Vocal

Suplente

Suplente

Realizado el acto de defensa de la tesis el día de de 2017 en la Escuela Técnica Superior de Ingenieros Agrónomos de Madrid.

Calificación:

El Presidente

El Secretario

Los Vocales

A Olga Cecilia, José Ángel y Andrés

“All in all it`s just another brick in the Wall”

(Pink Floyd)

Reconocimientos

Este trabajo se ha realizado en el Departamento de Protección Vegetal del Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA), bajo la dirección del Dr. Jaime Cubero Dabrio y de la Dra. María M. López González.

Para este estudio se contó además con la colaboración del Instituto Valenciano de Investigaciones Agrarias (IVIA) y del Centro de Investigación y Tecnología Agroalimentaria de Aragón (CITA).

En este trabajo han colaborado las siguientes personas:

Dr. Jaime Cubero Dabrio (INIA)

Dra. María M. López González (IVIA)

Dra. Ana Palacio Bielsa (IVIA)

Dra. Marta Sena Vélez (IVIA)

Dra. Pilar Sabuquillo Castrillo (INIA)

Elisa Ferragud Capó (INIA)

Ana Ruíz Padilla (INIA)

Isabel M. Berruete (CITA)

Jerson M. Garita Cambroner, contó para el desarrollo de esta tesis con un contrato de Formación de Profesorado Universitario (FPU 1000/2012), financiado por el Ministerio de Educación, Cultura y Deporte del Gobierno de España.

Este trabajo fue financiado por los siguientes proyectos: RTA2011-00140-C03-02, RTA2014-00018-C02-01 y XAPDIAG EUPHRESCO Project II.

Agradecimientos

Agradezco al Dr. Jaime Cubero Dabrio y a la Dra. María M. López González por haberme dado la oportunidad de formar parte de sus grupos de investigación, así como el haber aceptado dirigir este trabajo. Gracias además por haber compartido su conocimiento, experiencia y amistad conmigo durante estos años. No tengo palabras para expresar todo mi agradecimiento. Gracias infinitas.

A Ana Palacio-Bielsa por su amistad, toda su ayuda y su aporte en este trabajo, además por todas las críticas constructivas y la información que me ha dado durante todos estos años.

A mis compañeros de laboratorio y amigos, Elisa, Gema, Maite, Marta, Cristina, Pilar, Iray, Ana y Carlos. Gracias por vuestra compañía y por echarme una mano siempre que lo necesité en todo aspecto, incluso en lo personal, además por las discusiones y aportes que siempre ayudaron a que mi vida y trabajo mejorara. Qué habría sido de mi sin vosotros?.

A mi familia costarricense, Olga, José, Angie, Stuarth y Lubin por las palabras de ánimo y muestras de cariño que siempre me dieron desde el otro lado del mundo.

A mi familia española-ecuatoriano-francesa: Andrés, Angélica, María, Ángel, Bea, Mario, Diego, Flor y Nubia. Sin duda alguna lo mejor que España me ha dado. Quién iba a pensar que tenía que “atravesar el charco” para encontrar a las personas que más feliz me han hecho en mi vida, por suerte os encontré.

A Miguel Cambra e Iabel Berruete, por su amistad y por toda la ayuda que me dieron durante este trabajo. Por enseñarme la enfermedad en campo, por las fotos de los síntomas y por enviarme las cepas y estar pendientes de echarme una mano siempre que lo necesité.

A Pablo López y Javier Peñalver por su amistad y ayuda con las cepas usadas del IVIA y por la información de las mismas.

A mis colegas y amigos de la Universidad de Costa Rica Axel, Vanessa y Ethel, gracias por estar siempre pendientes de mí y brindarme su ayuda y amistad.

Al sistema de educación pública del gobierno de Costa Rica y del Gobierno de España quienes han permitido que hasta este momento haya podido estudiar toda mi vida gracias a un sistema de educación gratuito y a diversas becas durante mi formación universitaria y de posgrado.

“Gracias a la vida que me ha dado tanto”.

Principales abreviaturas

ADN: Ácido desoxirribonucleico

AFLP: Polimorfismo en la longitud de fragmentos amplificados

ARNr: ARN ribosomal

BOX-PCR: "BOX-A1R-based repetitive extragenic palindromic-PCR"

CDS= regiones codificantes de proteínas

COG: "Cluster" de grupos ortólogos

di-GMP: diguanilato monofosfato cíclico

dpi: días post inoculación

DSF: factor de señal difusible

ELISA= Ensayo por inmuno adsorción ligado a enzimas

EPPO: "European and Mediterranean Plant Protection Organization"

ERIC-PCR: "Enterobacterial repetitive intergenic consensus-PCR"

ETI: "effector-triggered immunity"

FAFLP: AFLP fluorescente

GMP: Guanosin monofosfato

hpi: horas post inoculación

ITS: Espacio del transcrito interno

Kpb: Kilo pares de bases

LAMP= Amplificación isotérmica de ácidos nucleicos

MIGS: Información mínima requerida para la secuencia de un genoma

MLSA: Análisis de secuencias multilocus ("Multilocus sequence analysis")

MLVA: Análisis del número variable de secuencias repetidas en tandem en múltiples loci ("Multilocus variable-number tandem repeats analysis")

NaCl: cloruro de sodio

°C: Grado centígrado

ORF: "Open reading frame"

PCR: reacción en cadena de la polimerasa

pH: grado de acidez

PTI: "PAMP-triggered immunity"

QTL= Locus de un carácter cuantitativo

REP-PCR: "Repetitive extragenic palindromic-PCR"

T2SS: Sistema de secreción tipo 2

T3SS: Sistema de secreción tipo 3

T4SS: Sistema de secreción tipo 4

T5SS: Sistema de secreción tipo 5

T6SS: Sistema de secreción tipo 6

TAL: "Transcription activator like effector"

tRNA: Ácido ribonucleico de transferencia

UFC o CFU= Unidades formadoras de colonias

VNTR: Número Variable de repeticiones en tándem

Tabla de Contenido

<u>Reconocimientos</u>	<u>ix</u>
<u>Agradecimientos</u>	<u>xi</u>
<u>Principales abreviaturas</u>	<u>xiii</u>
Resumen	xxi
Abstract	xxiv
Capítulo 1: Introducción	1
<hr/>	
<u>1.1 Descripción y Diversidad en el género <i>Xanthomonas</i>.</u>	<u>4</u>
<u>1.2 Mecanismos Asociados al Desarrollo de Enfermedad en <i>Xanthomonas</i>.</u>	<u>7</u>
<u>1.3 Mecanismos Iniciales de Infección en <i>Xanthomonas</i>.....</u>	<u>8</u>
<u>1.3.1 Adherencia al tejido vegetal y formación de biopelículas.</u>	<u>8</u>
<u>1.3.2 Adhesinas fimbriales y no fimbriales asociadas a la adherencia bacteriana.....</u>	<u>11</u>
<u>1.3.3 El pilus tipo IV y su relación con la adherencia y motilidad bacteriana.</u>	<u>12</u>
<u>1.3.4 Adhesinas no fimbriales y otros componentes asociados a la adherencia y formación de biopelículas bacterianas.</u>	<u>14</u>
<u>1.3.5 Sistemas bacterianos de quorum sensing.....</u>	<u>15</u>
<u>1.3.6 Otros sistemas sensores asociados a la percepción bacteriana de las condiciones ambientales.....</u>	<u>18</u>
<u>1.3.7 La quimiotaxis como mecanismo de respuesta ambiental y su relación con el movimiento celular.....</u>	<u>21</u>
<u>1.3.8 Estructura y función del flagelo bacteriano.....</u>	<u>23</u>
<u>1.3.9 Tipos de movimiento bacteriano relacionados con el flagelo.....</u>	<u>26</u>
<u>1.4 Sistemas de secreción y proteínas efectoras asociadas a procesos tardíos de infección en <i>Xanthomonas</i>.....</u>	<u>30</u>
<u>1.4.1 Sistema bacteriano de secreción tipo II.....</u>	<u>31</u>
<u>1.4.2 Sistema bacteriano de secreción tipo III.</u>	<u>33</u>
<u>1.4.3 Sistema Bacteriano de secreción tipo IV.....</u>	<u>39</u>
<u>1.4.4 Sistemas bacteriano de secreción tipo V.</u>	<u>42</u>
<u>1.4.5 Sistema bacteriano de secreción tipo VI.</u>	<u>44</u>

1.5 <i>Xanthomonas arboricola</i> como agente causal de enfermedad en diversos grupos de plantas de interés económico.	46
1.5.1 Descripción de la especie <i>X. arboricola</i>	46
1.5.2 Diversidad genética de <i>X. arboricola</i>	46
1.5.3 Factores de virulencia de <i>X. arboricola</i>	48
1.5.4 <i>Xanthomonas arboricola</i> pv. <i>corylina</i> : agente causal de la podredumbre bacteriana del avellano.	49
1.5.5 <i>Xanthomonas arboricola</i> pv. <i>juglandis</i> : Agente causal de la podredumbre bacteriana, la necrosis apical y la chancrosis del nogal.	51
1.5.6 <i>Xanthomonas arboricola</i> pv. <i>pruni</i> : agente causal de la mancha bacteriana en frutales del género <i>Prunus</i>	54
1.5.7 Diversidad genética y estructura poblacional:	54
1.5.8 Interacción huésped-patógeno-ambiente en el contexto de la mancha bacteriana de los frutales de hueso y el almendro:	56
1.5.8.1 El patógeno.	56
1.5.8.2 El huésped.	57
1.5.8.3 El ambiente.	60
1.5.9 Distribución geográfica e impacto económico de <i>X. arboricola</i> pv. <i>pruni</i> :	61
1.5.10 Sintomatología asociada a la mancha bacteriana de los frutales del género <i>Prunus</i> :	63
1.5.11 Epidemiología de la mancha bacteriana de los frutales del género <i>Prunus</i> :	65
1.5.12 Diagnóstico y control de la mancha bacteriana de los frutales del género <i>Prunus</i> :	66
1.5.13 Condición fitosanitaria de <i>X. arboricola</i> pv. <i>pruni</i> :	70
Capítulo 2: Secuenciación y características generales del genoma de cinco cepas de <i>X. arboricola</i> aisladas de <i>Prunus</i> spp. en España.	71
2.1 Introducción:	73
2.2 Draft genome sequence of <i>Xanthomonas arboricola</i> pv. <i>pruni</i> strain Xap33, causal agent of bacterial spot disease on almond.	78
2.2.1 Abstract	78
2.2.2 Genome announcement	78
2.2.3 Nucleotide sequence accession numbers.	79
2.2.4 Acknowledgments.	79

<u>2.3 Draft genome sequence for virulent and avirulent strains of <i>Xanthomonas arboricola</i> isolated from <i>Prunus</i> spp. in Spain.....</u>	<u>81</u>
<u>2.3.1 Abstract</u>	<u>81</u>
<u>2.3.2 Keywords.....</u>	<u>81</u>
<u>2.3.3 Introduction</u>	<u>81</u>
<u>2.3.4 Organism information</u>	<u>83</u>
<u>2.3.4 Genome sequencing information</u>	<u>87</u>
<u>2.3.5 Genome properties</u>	<u>89</u>
<u>2.3.6 Insights from the genome sequence</u>	<u>92</u>
<u>2.3.7 Conclusions</u>	<u>94</u>
<u>2.3.8 Acknowledgements.....</u>	<u>94</u>
<u>2.4 Draft genome sequence of two strains of <i>Xanthomonas arboricola</i> isolated from <i>Prunus persica</i> which are dissimilar to strains that cause bacterial spot disease on <i>Prunus</i> spp.</u>	<u>95</u>
<u>2.4.1 Abstract</u>	<u>95</u>
<u>2.4.2 Genome announcement</u>	<u>95</u>
<u>2.4.3 Accession numbers.....</u>	<u>96</u>
<u>2.4.4 Acknowledgments.....</u>	<u>97</u>

Capítulo 3: Comparación genómica y caracterización fenotípica de cepas patógenas y no patógenas de *X. arboricola* revela aspectos relacionados con el proceso de infección de la mancha bacteriana de los frutales de hueso.....71

<u>3.1 Abstract</u>	<u>101</u>
<u>3.2 Introduction</u>	<u>101</u>
<u>3.3 Results</u>	<u>103</u>
<u>3.3.1 Molecular characterization using multilocus sequence analysis (MLSA), genome based phylogeny and genome comparison.....</u>	<u>103</u>
<u>3.3.2 Carbon sources utilization and chemotaxis profile.....</u>	<u>107</u>
<u>3.3.3 Flagella, fimbrial and non-fimbrial adhesins associated with motility and attachment in <i>X. arboricola</i>.</u>	<u>110</u>
<u>3.3.4 Pathogenicity tests and genomic components associated with late stages of infection in <i>Prunus</i>-associated strains.....</u>	<u>114</u>
<u>3.4 Discussion.....</u>	<u>118</u>
<u>3.5 Materials and Methods</u>	<u>125</u>
<u>3.5.1 Bacterial strains and culture conditions</u>	<u>125</u>

3.5.2 Multi Locus Sequence Analysis (MLSA).	126
3.5.3 Genomic Analysis	127
3.5.4 Carbon sources utilization analysis and environmental sensors profile	128
3.5.5 Chemotaxis assay and repertoire of methyl accepting chemotaxis proteins	129
3.5.6 Motility in solid and semisolid surfaces and surfactant activity	130
3.5.6 Pathogenicity tests on <i>Prunus</i> spp.	131
3.6 Acknowledgments	133
3.7 Author contributions	133
Capítulo 4: El análisis pan genómico ha permitido diferenciar cepas no patógenas y patógenas de <i>Xanthomonas arboricola</i> que cohabitan en <i>Prunus</i> spp. así como elucidar los factores de la virulencia bacteriana.	101
4.1 Abstract	137
4.2 Introducción	138
4.3 Materials and Methods	139
4.3.1 Bacterial strains and classification using multilocus sequence analysis	139
4.3.2 Study of the type III secretion system and type III effectors gene repertory:	144
4.3.3 Pathogenicity tests:	145
4.3.4 Genome sequencing and comparison:	146
4.3.5 A molecular tool to differentiate <i>X. arboricola</i> pv. <i>pruni</i> from atypical strains of <i>X. arboricola</i> associated with <i>Prunus</i> spp.:	149
4.3.6 Specificity of the real-time PCR test:	149
4.3.7 Sensitivity of the real-time PCR test:	150
4.4 Results	151
4.4.1 Characterization of atypical strains of <i>X. arboricola</i> associated with <i>Prunus</i> spp.:	151
4.4.2 T3SS and T3Es repertoire in <i>Xap</i>-look-a-like strains:	152
4.4.3 Pathogenicity of the <i>Xap</i>-look-a-like strains:	153
4.4.4 General features of whole genomes of <i>X. arboricola</i> strains CITA 14 and CITA 124:	155
4.4.5 Genes associated with pathogenicity in <i>X. arboricola</i> strains:	159
4.4.6 The real-time PCR test for <i>xopE3</i> permitted to differentiate <i>Xap</i> from atypical strains of <i>X. arboricola</i> isolated on <i>Prunus</i> spp.:	163

<u>4.5 Discussion.....</u>	<u>165</u>
<u>4.6 Acknowledgments.....</u>	<u>173</u>
Capítulo 5: Discusión general	136
Capítulo 6: Conclusiones	136
Bibliografía.....	1
Material Suplementario.....	136
<hr/>	
<u>Material suplementario, capítulo 3:</u>	<u>227</u>
<u>Material suplementario, capítulo 4:</u>	<u>230</u>

Resumen

Xanthomonas arboricola es una especie de bacterias asociada a plantas. Dentro de la misma se han descrito nueve grupos infraespecíficos o patovares causantes de enfermedad en diversas especies vegetales. Asimismo, se han identificado tanto cepas patógenas como no patógenas que no se incluyen en ninguno de estos grupos. Entre los grupos definidos como fitopatógenos se encuentra *X. arboricola* pv. *pruni*, agente causal de la mancha bacteriana de los frutales de hueso y el almendro. Este patógeno de aparición reciente en España, está considerado como organismo de cuarentena en la Unión Europea.

En este trabajo se ha caracterizado, tanto molecular como fenotípicamente, una serie de cepas de *X. arboricola* aisladas de *Prunus* spp. Un análisis de secuencias multilocus determinó que junto a las cepas de *X. arboricola* pv. *pruni*, existían, cohabitando en los mismos huéspedes, algunas cepas de la misma especie que no pertenecían a ninguno de los grupos que se integran en *X. arboricola*. Además, según la caracterización fenotípica basada en aspectos relacionados con la infección bacteriana, estas cepas no presentaban la capacidad para producir enfermedad en *Prunus* spp.

La caracterización inicial de las cepas de *Xanthomonas* aisladas de *Prunus* spp. fue aprovechada para seleccionar una cepa patógena de almendro y otra de ciruelo, CITA 33 e IVIA 2626.1 respectivamente, y tres cepas no patógenas, CITA 14, aislada de *P. mahaleb*, así como CITA 44 y CITA 124, aisladas de melocotonero. El estudio de genómica comparativa de estas cepas permitió identificar factores que podrían estar relacionados con el desarrollo de la enfermedad de la mancha bacteriana de los frutales del género *Prunus*.

El análisis filogenético basado en el “core genome” de la especie, así como un análisis de similitud basado en la presencia de los genes del pan genoma entre las diversas cepas de *X. arboricola*, permitió corroborar que las cepas no patógenas formaban un grupo filogenético diferente al del patovar *pruni*, junto con otras cepas de baja virulencia o no patógenas aisladas de diferentes plantas huésped.

Además, la comparación de los diferentes genomas permitió encontrar diferencias en los genes relacionados con mecanismos implicados en los estados iniciales de la infección, tales como transportadores TonB-dependientes (TBDTs), sensores del sistema regulador de dos componentes (STCRs) y proteínas aceptoras de grupos metilo (MCPs). Asimismo, se observó en el patovar *pruni* la presencia de adhesinas no fimbriales, como FhaB2, que no estaban presentes en las cepas no patógenas. De la misma manera, se encontró variación en cuanto al contenido de genes asociados al pilus tipo IV entre ambos grupos bacterianos.

Respecto a la estructura del flagelo, tanto las cepas patógenas como las no patógenas presentaron el mismo perfil en aquellos genes relacionados con su estructura y biogénesis, sin embargo, se observó un cambio en la secuencia de aminoácidos de la proteína flagelina de *X. arboricola* pv. *pruni*, la cual podría estar relacionada con la respuesta de la bacteria frente a los mecanismos de defensa de la planta.

De igual manera, se encontraron diferencias entre las cepas patógenas y no patógenas en caracteres asociados con etapas tardías de la infección. Una de ellas fue la presencia de diferentes perfiles de genes codificantes de enzimas degradadoras de la pared celular. Otra diferencia encontrada entre ambos grupos de organismos fue la referente al sistema de secreción tipo III y sus efectores. Las cepas patógenas, identificadas como pertenecientes al patovar *pruni*, presentaron todos los genes asociados a los componentes estructurales y de regulación de este sistema de secreción, así como un total de 21 efectores tipo III. Sin embargo, dos de las cepas no patógenas CITA 44 y CITA 124, no presentaron sistema de secreción tipo III, ni ninguno de los efectores tipo III descritos en *Xanthomonas*. La cepa no patógena CITA 14 presentó los componentes génicos para un sistema de secreción tipo III funcional pero se identificó solamente un repertorio de seis efectores.

Se seleccionó además el gen *xopE3*, que solamente está presente en el patovar *pruni* y se desarrolló un protocolo de PCR tiempo real que permitió aumentar la especificidad del protocolo de diagnóstico para la mancha bacteriana de los frutales de hueso y el almendro, al permitir diferenciar de manera específica y sensible las cepas causantes de la enfermedad de aquellas cepas no patógenas que cohabitan los frutales del género *Prunus*.

Este estudio ha permitido conocer una serie de componentes genéticos que podrían estar asociados con el desarrollo de la enfermedad de la mancha bacteriana de los frutales del género *Prunus*. De la misma manera ha permitido profundizar en el conocimiento sobre la diversidad biológica de *X. arboricola* y ha generado datos que pueden marcar las pautas de cómo ha podido ser la evolución de la patogenicidad y especificidad de huésped en esta especie de bacterias. Los datos generados han permitido además el desarrollo de un nuevo método de diagnóstico de esta enfermedad más preciso basado en PCR en tiempo real.

Abstract

Xanthomonas arboricola is a species of plant-associated bacteria, conformed by at least nine infraspecific groups or pathovars. The species encompasses groups of bacteria with diverse plant host range, as well as pathogenic and non-pathogenic strains that does not belong to any of these nine pathovars. Within the plant pathogenic groups, *X. arboricola* pv. *pruni*, which is the causal agent of bacterial spot disease of stone fruits and almond, has been recently determined as present in Spain and it is regulated as a quarantine pathogen in the European Union.

In this work, a group of Spanish strains of *X. arboricola*, isolated from *Prunus* spp. has been characterized using molecular and phenotypic approaches. An initial molecular characterization, conducted by using a multilocus sequence type analysis, determined that there were two groups of strains that cohabited in these hosts. The first one was composed by strains of *X. arboricola* pv. *pruni* and the second one by atypical strains that did not belong to any of the nine well-established pathovars of *X. arboricola*. In addition, according to the phenotypic analysis, none of the atypical strains were able to cause disease on *Prunus* spp.

Based on these results, two pathogenic strains of *X. arboricola* pv. *pruni*, CITA 33 and IVIA 2626.1, isolated from peach and plum respectively, and three non-pathogenic strains of *X. arboricola*, CITA 14, isolated from *P. mahaleb* and CITA 44 and CITA 124, both isolated from peach, were selected for whole genome sequencing. The genome comparative analysis of these strains permitted to find those genomic characters that could be associated with the development of bacterial spot disease of stone fruits and almond.

A phylogenetic analysis based in the core genome of *X. arboricola* as well as a similarity analysis based on the pan-genome of this species, corroborated that the atypical non-pathogenic strains formed cluster which was different to the one composed by the strains of the pathovar *pruni*. These non-pathogenic strains clustered in a group with other non-pathogenic or low-pathogenic strains isolated from a wide range of plant hosts.

Additionally, according to the gene content comparison, variants in the genes related to the different stages of the disease process were found. For instance, both groups showed a different repertoire of genes associated with the sensing of the environmental stimuli, such as those related to the Ton-B dependent transporters (TBDTs), the sensors of the two-component regulatory system (STCRs) and the methyl accepting chemotaxis proteins (MCPs). Beside this, in *X. arboricola* pv. *pruni*, some non-fimbrial adhesins, like FhaB2, as well as some components related to the type IV pilus that were absent in the non-pathogenic strains, were found.

According to the gene content associated with the flagella structure and regulation, no differences among gene profile of both groups of organisms were elucidated, however, in the pathovar *pruni*, a variant in the flagellin protein sequence. This variant could be associated with the ability of this bacterial group to avoid the initial steps of the plant defence process.

In the same way, variations in the gene content of some key characters related to late stages of the disease process were found. One of the differences was the dissimilar profile of cell wall degrading enzymes between pathogenic and non-pathogenic strains. Moreover, the most important difference observed between these two groups was found in the gene content associated with the type III secretion system and its related effectors. Plant pathogenic strains of *X. arboricola* pv. *pruni* harboured all the genetic components related to the structure and regulation of the type III secretion system as well as in a repertoire of 21 type III effectors. Nevertheless, non-pathogenic strains CITA 44 and CITA 124 did not harbour this secretion system or none of the type III effectors described in *Xanthomonas*. In the case of the non-pathogenic strain CITA 14, this contained all the genes related to a functional type three secretion system but a reduced profile of effectors composed by only six genes was determined.

In addition, the gene *xopE3*, which is present only in *X. arboricola* pv *pruni*, was selected and used to design a real-time PCR protocol which permitted to improve the specificity of the available protocol to identify the bacteria associated to the spot disease of stone fruit and almond. This protocol allowed to perform a specific and sensitive differentiation of the disease-associated strains from those non-pathogenic strains that cohabited on *Prunus* spp.

This study provides a global overview of the genetic components that could be associated with the development of the bacterial spot disease of *Prunus*. At the same time, it reveals new insights in some other elements of the biology of this bacterial group such as its biological diversity and the potential process associated with the evolution of pathogenicity and host specificity. All the data generated in this investigation has also permitted to develop a new and more accurate real time PCR diagnostic tool for this disease.

Capítulo 1: Introducción

El dominio Bacteria está conformado por organismos unicelulares procariotas, con una membrana compuesta por lípidos, predominantemente diésteres de diacilglicerol y con ribosomas que contienen un ARN ribosomal (ARNr) de tipo eubacteriano (Woese *et al.*, 1990; Winker & Woese, 1991). Estos organismos son ubicuos, altamente abundantes y no solo contribuyen al flujo de nutrientes en el planeta, sino que además constituyen una proporción significativa de los nutrientes existentes en la biomasa viva. Debido a que los estudios sobre bacterias se han basado principalmente en aquellas que son cultivables, solamente se conoce aproximadamente el 1% de su diversidad total, que se especula podría alcanzar los 10 billones de especies. Recientemente, y como consecuencia de los estudios de metagenómica se ha avanzado considerablemente en el conocimiento sobre la diversidad taxonómica y metabólica de estos organismos en el ecosistema (Horner-Devine *et al.*, 2004; Xu, 2006).

Desde el punto de vista del estudio de la actividad microbiana y su relación con el ambiente y con otros organismos, se considera que las bacterias pueden actuar como descomponedores de materia orgánica, como patógenos de plantas y animales, así como desarrollar asociaciones mutualistas o incluso ser depredadores de otros microorganismos (Xu, 2006).

Se han descrito variedad de interacciones entre estos microorganismos y las plantas, algunas de ellas positivas y otras negativas. Las bacterias pueden potenciar el crecimiento y estimular la respuesta de defensa de las plantas ante otros organismos (Glick, 2015; Moreno-hagelsieb *et al.*, 2016), asimismo, también son capaces de penetrar en la planta como endófitos sin causar daño a su huésped e incluso desarrollar relaciones simbióticas bacteria-planta (Rosenblueth & Martínez-Romero, 2006; Gourion *et al.*, 2015).

En contraposición a las interacciones previamente mencionadas, algunos grupos bacterianos contenidos en Actinobacteria, Mollicutes y Proteobacteria (Loria *et al.*, 2006; Marcone, 2014; van der Wolf & De Boer, 2015), han desarrollado una relación ecológica de tipo parasítica con las plantas, siendo causantes de un gran número de enfermedades vegetales que pueden llegar a tener un efecto devastador en el rendimiento de los cultivos (Tarkowski & Vereecke, 2014).

De acuerdo con “The International Society of Plant Pathology Committee on the Taxonomy of Plant Pathogenic Bacteria (ISPP-CTPPB)” (Bull *et al.*, 2010, 2012, 2014),

existen descritos al menos 39 géneros bacterianos que contienen organismos patógenos en huéspedes vegetales, destacando dentro de estos como los más diversos, los géneros *Pseudomonas* y *Xanthomonas*.

1.1 Descripción y Diversidad en el género *Xanthomonas*.

El género *Xanthomonas* (Dominio *Bacteria*, Phylum *Proteobacteria*, Clase *Gammaproteobacteria*, Orden *Lysobacterales*, Familia *Lysobacteraceae*) (Dowson, 1939; Williams & Kelly, 2013; Naushad *et al.*, 2015; Oren & Garrity, 2015), es un grupo de bacterias con forma de bacilos y Gram negativas con un tamaño de entre 0,5 a 0,6 μm de ancho por 1,0 a 2,9 μm de largo. Fenotípicamente, según la reclasificación y descripción realizada por Vauterin y colaboradores (1995), las colonias bacterianas de este género se caracterizan por ser de color amarillo, tener borde liso y una consistencia mucosa. Tanto el color como la consistencia de las colonias se deben principalmente a la secreción de xanthomonadina y del exopolisacárido xanthano. Las células de *Xanthomonas*, suelen además, ser móviles y presentan un único flagelo polar.

En cuanto a su actividad enzimática, las bacterias de este género presentan actividad catalasa, pero no actividad oxidasa ni ureasa, ni son capaces de producir indol ó acetoina, o reducir el nitrato a nitrito. Al ser organismos quimiorganotrófos, utilizan variedad de compuestos como fuentes de carbono o nitrógeno pero no L-ramnosa, adonitol, sorbosa, D-sorbitol, meso-eritritol, glicina, L-glutamina y L-asparagina. *Xanthomonas* presenta la capacidad de oxidar los siguientes compuestos presentes en el sistema de microplaca Biolog GN, D-fructuosa, α -D-glucosa, D-manosa, metilpiruvato y el ácido α -ketoglúutámico. Asimismo, presentan la incapacidad de oxidar la α -ciclodextrina, adonitol, D-arabitol, meso-eritritol, meso-inositol, xilitol, ácido D-glucosamínico, ácido γ -hidroxibutírico, ácido itacónico, ácido sebásico, L-ornitina, ácido L-pirolglutámico, D-serina, DL-carnitina, ácido γ -aminobutírico, feniletilamina, putrescina, 2-aminoetanol y 2,3 butanodiol.

Las bacterias de este género no son capaces de crecer en un pH menor a 4,5 y su rango de crecimiento se encuentra entre los 4 y los 37 °C (Vauterin *et al.*, 1995). Su crecimiento se ve inhibido a concentraciones de NaCl superiores al 6% y de glucosa al 30%, así como al 0,01% de verde de metilo, 0,01% de tionina, 0,01% de acetato, 0,1% de trifenitetrazólio cloruro. Los miembros de este género suelen ser susceptibles a la

eritromicina y a la tetraciclina. Producen espermidina en cantidades detectables y algunas cepas además cadaverina. Por último, el género *Xanthomonas* contiene nueve ácidos grasos, de los cuales tres, el 11:0 iso, el 11:0 iso 3OH y el 13:0 iso 3OH son característicos del género, por lo que son utilizados para su diferenciación de otros grupos bacterianos (Vauterin *et al.*, 1995).

El género *Xanthomonas* es un grupo de bacterias, principalmente fitopatógenas, que es filogenéticamente complejo y que por su importancia económica ha sido ampliamente estudiado. Los estudios filogenéticos en *Xanthomonas* han sido abordados utilizando multitud de estrategias basadas en diversos caracteres. Entre ellas se pueden mencionar el uso de caracteres fenotípicos, perfiles bioquímicos o de ácidos grasos, así como la utilización de una serie de técnicas de caracterización molecular basadas en la amplificación en cadena de la polimerasa o PCR, como lo son el BOX-PCR, el ERIC-PCR y la REP-PCR (Vauterin *et al.*, 2000).

Recientemente, el avance en las técnicas de secuenciación de ácidos nucleicos ha permitido la generación de bases de datos de secuencias de regiones específicas que permiten identificar cepas desconocidas por comparación con las previamente clasificadas. La secuenciación de ácidos nucleicos se han revelado como una herramienta rápida y eficiente de clasificación, por ejemplo utilizando los datos disponibles para la subunidad beta de la DNA girasa (*gyrB*) (Parkinson *et al.*, 2009).

Ha sido de gran utilidad para la taxonomía de este grupo el desarrollo de esquemas de análisis de secuencias multilocus (MLSA), basado en las secuencias parciales de los genes constitutivos asociados a la chaperona DnaK, el receptor tipo TonB, FyuA, la subunidad beta de la DNA girasa, GyrB y el factor sigma de la RNA polimerasa, RpoD (Young *et al.*, 2008). El MLSA es actualmente el método más utilizado ya que tiene una alta capacidad de diferenciación entre las especies de *Xanthomonas*, equivalente a la hibridación DNA-DNA y además ha permitido, de manera relativamente fácil, la descripción de nuevas especies de *Xanthomonas* (Young *et al.*, 2010).

Más recientemente, gracias a la disponibilidad de genomas completos para este género, se han llevado a cabo análisis filogenéticos basados en la secuencia concatenada de los genes que conforman el “core genome”, que corresponde a las secuencias de aquellos genes compartidos por todos los genomas que se analizan (Blom *et al.*, 2009; Rodríguez-R *et al.*, 2012; Cesbron *et al.*, 2015). Este método no es ampliamente

utilizado hoy en día por la dificultad que aún entraña la obtención de genomas completos, pero sin duda en un futuro se convertirá en una metodología de gran importancia.

El género *Xanthomonas* contiene actualmente un total de 29 especies bacterianas principalmente asociadas a plantas y muy raramente encontradas en otros ambientes (Hayward, 1993; Bull *et al.*, 2010, 2012; Jacques *et al.*, 2016). Este grupo de bacterias se caracteriza por presentar poca variación fenotípica entre las diversas especies, lo cual contrasta con su alta capacidad para infectar una gran diversidad de huéspedes. Las especies de este género bacteriano se han encontrado en al menos 124 especies de monocotiledóneas y 268 especies de dicotiledóneas (Jacques *et al.*, 2016) afectando a cultivos de gran interés económico y alimentario para la humanidad (Bull *et al.*, 2010, 2012, 2014). Por ejemplo *X. axonopodis*, *X. campestris* y *X. oryzae* están consideradas, entre los principales problemas bacterianos para la agricultura a nivel mundial (Mansfield *et al.*, 2012).

Las diversas cepas descritas dentro del género *Xanthomonas*, presentan como característica su restringida gama de huéspedes (Jacques *et al.*, 2016). Además se han descrito grupos sub-infraespecíficos con una gama de huésped aún más limitada, la cual en algunos casos como en *X. oryzae*, llega a presentarse a nivel de tejido dentro del mismo huésped (Niño-Liu *et al.*, 2006). Se ha creado una clasificación basada en el término patovar, término que según “The Committee on the Taxonomy of Plant Pathogenic Bacteria” (Vauterin *et al.*, 2000), se refiere a una cepa o un grupo de cepas con características fenotípicas similares y diferentes a nivel sub-infraespecífico de otras cepas de la misma especie o subespecie, que presentan una patogenicidad distinta en una o más plantas huésped. Actualmente, se han descrito patovares en al menos 14 especies de *Xanthomonas*, las cuales contienen entre dos (*X. alfalfae*, *X. gardneri*, *X. melonis* y *X. sacchari*) y 56 (*X. axonopodis*) patovares (Jacques *et al.*, 2016).

La diversidad conocida en el género *Xanthomonas* sin embargo puede estar subestimada puesto que se ha prestado mayor interés por aquellas cepas que tienen capacidad de producir enfermedad (Young *et al.*, 2010). Sin embargo, hay evidencia sobre la existencia de cepas oportunistas, las cuales no causan ningún tipo de enfermedad aparente. Asimismo, en múltiples estudios se han encontrado cepas no patogénicas de *Xanthomonas* en un gran número de plantas huésped con y sin síntomas de enfermedad.

Al caracterizar estas cepas, se ha observado que forman parte de poblaciones heterogéneas y muchas de ellas no pueden ser clasificadas en ninguna de las categorías taxonómicas existentes en este género, constatándose la amplia diversidad de este grupo bacteriano hasta ahora subestimada (Vauterin *et al.*, 1996; Fischer-Le Saux *et al.*, 2015; Jacques *et al.*, 2016).

1.2 Mecanismos Asociados al Desarrollo de Enfermedad en *Xanthomonas*.

Las especies del género *Xanthomonas* son capaces de producir enfermedad en prácticamente todos los tejidos de la planta. Aunque son más importantes en los tejidos y órganos aéreos, tales como troncos, tallos, hojas, yemas, brotes, flores, frutos y semillas (Jacques *et al.*, 2016). Además, algunas especies son capaces de penetrar e infectar a través de las raíces, como es el caso de *X. campestris* pv. *musacearum*, agente causal de la marchitez del plátano (Mwangi *et al.*, 2007). Las especies de este género pueden presentar una alta afinidad por determinados tejidos, identificándose como patógenos vasculares a aquellos que infectan los elementos del xilema, por ejemplo *X. albilineans* y *X. oryzae* pv. *oryzae* o patógenos mesofílicos, aquellos que infectan los espacios intercelulares del tejido del mesófilo, como por ejemplo *X. citri* subsp. *citri* y *X. oryzae* pv. *oryzicola*. Además hay casos como el de *X. manihotis*, especie capaz de desarrollar infección tanto a nivel del mesófilo como a nivel de tejidos vasculares (Ryan *et al.*, 2011).

Los daños que las bacterias del género *Xanthomonas* producen en el huésped, se manifiestan en gran variedad de síntomas, entre los que se incluyen chancros, pústulas, podredumbres, manchas y estriados foliares, marchitez, hipertrofia, hiperplasia e incluso la muerte del huésped (Ryan *et al.*, 2011; Jacques *et al.*, 2016).

Según Pfeilmeier y colaboradores (2016), en las plantas, las infecciones bacterianas se inician por lo general con una etapa epífita, en la cual el organismo debe ser capaz de adaptarse a las condiciones que presenta el medio y sobrevivir el tiempo suficiente para continuar el proceso de infección. Estos mecanismos iniciales incluyen la activación de una serie de cascadas de transducción como respuesta a los estímulos ambientales en las poblaciones bacterianas. Posteriormente, los patógenos pueden migrar desde la superficie vegetal al interior de la planta en procesos mediados por quimiotaxis,

penetrando por aperturas naturales, tipo estomas, hidátodos, lenticelas y nectarios o a través de heridas. Una vez en la zona del apoplasto, las bacterias utilizan una serie de sistemas de secreción para liberar enzimas degradadoras, fitotoxinas y proteínas efectoras que desempeñan diversas funciones asociadas a la patogénesis. Durante todo este proceso, las bacterias deben de enfrentar y superar los sistemas de defensa con los que cuenta la planta.

En la infección por *Xanthomonas* están implicados numerosos mecanismos que en esta memoria se han dividido en aquellos asociados al inicio de la infección, como son, la capacidad de adherencia a los tejidos vegetales, la formación de biopelículas, la percepción de los estímulos ambientales, la motilidad y la quimiotaxis y en aquellos responsable de las fases más avanzadas del proceso infectivo, tales como la secreción de enzimas degradadoras de pared celular y de efectores por los diversos sistemas de secreción presentes en *Xanthomonas*. En los apartados siguientes se ha hecho un resumen de dichos procesos.

1.3 Mecanismos Iniciales de Infección en *Xanthomonas*.

1.3.1 Adherencia al tejido vegetal y formación de biopelículas.

Las bacterias fitopatógenas, y entre ellas las del género *Xanthomonas*, producen gran cantidad de polisacáridos extracelulares, que pueden estar formados por un solo tipo de carbohidrato o por complejos de varios azúcares ordenados en unidades repetidas (Denny, 1995). En *Xanthomonas*, es especialmente importante y característica la producción del xanthano. Estructuralmente, este compuesto está formado por una columna vertebral de unidades de D-glucosa unidas entre sí mediante enlaces β -1,4; a esta columna vertebral se unen cadenas laterales formadas por manosa-(β -1,4)-ácido glucurónico-(β -1,2)-manosa. Dichas cadenas laterales se unen de manera alterna a los residuos de glucosa que conforman la columna vertebral por medio de enlaces α -1,3. Los residuos de manosa están acetilados y piruvilados en sitios específicos (Figura 1) (Dunger *et al.*, 2007; Hamman, 2010). La síntesis, la polimerización y la secreción de este polisacárido extracelular se lleva a cabo mediante un operón conformado por 12 genes denominados genes *gum* (*gumB-gumM*), los cuales están altamente conservados entre las diferentes especies del género (Dunger *et al.*, 2007).

Figura 1. A) Estructura química del xanthano (Hamman, 2010). B) representación esquemática del operón *gum* de la especie *X. campestris* (Da Silva *et al.*, 2001).

Además, al igual que todas las bacterias Gram negativas contienen como parte de su membrana externa una capa de lipopolisacáridos, compuestos de un lípido A y un núcleo de polisacáridos que contiene en su extremo unidades de antígeno O (Hori & Matsumoto, 2010). En *Xanthomonas* los lipopolisacáridos están generalmente compuestos de ácido urónico, glucosa, manosa y 2-keto-3-dioxitanato, y además contienen fosfato orgánico y carbohidratos adicionales como ramnosa, xilosa, fucosa y galactosa, en algunas de las especies del género (Volk, 1966). La biosíntesis de lipopolisacáridos en *X. campestris* pv. *campestris* está asociada a 15 genes (Vorhölter *et al.*, 2001), cuyo porcentaje de identidad varía entre las diferentes especies del género. Dicha variación puede estar asociada a la patogenicidad y gama de huésped a través del antígeno O del lipopolisacárido, específico para cada patovar o cepa y con una de sus funciones descrita como barrera en contra de las toxinas vegetales (da Silva *et al.*, 2002; Patil *et al.*, 2007) (Figura 2).

Figura 3. Representación esquemática de la estructura de una biopelícula y de las diferentes etapas que llevan a su formación en *X. fuscans* subsp. *aurantifolii*. 1) y 2) Acceso de la bacteria a la superficie y adhesión inicial causada por polisacáridos, apéndices adhesivos y adhesinas secretadas, 3) y 4) Adherencia irreversible y formación de una biopelícula madura, etapas llevadas a cabo principalmente por polisacáridos extracelulares, flagelo, estructuras tipo pili y DNA extracelular, 5) salida de la biopelícula y paso de las células a fase planctónica, dirigida principalmente por cambios ambientales percibidos por las células bacterianas que forman parte de la matriz (Moreira *et al.*, 2010; Kostakioti *et al.*, 2013).

Debido a que la formación de biopelículas se ha asociado a la capacidad de desarrollar infección, y debido a que los polisacáridos extracelulares forman parte primordial de su estructura, es por lo que tanto a este tipo de polisacáridos como a las biopelículas se les considera factores de virulencia bacterianos (Hori & Matsumoto, 2010). En el caso de las *Xanthomonas*, se ha estudiado la actividad del polisacárido extracelular xanthano en la formación de biopelículas y su relación con la virulencia en *X. citri* subsp. *citri*. Una mutación en el gen *gumB*, en esta especie, eliminó la formación de este polisacárido y la cepa mutante presentó una menor capacidad de adherencia a superficies tanto inertes como a la superficie foliar de lima. Mediante microscopía se determinó que el mutante además era incapaz de desarrollar una biopelícula con estructura compleja y sufría una drástica disminución de su virulencia en comparación con la cepa silvestre (Rigano *et al.*, 2007).

1.3.2 Adhesinas fimbrilares y no fimbrilares asociadas a la adherencia bacteriana.

Entre los factores determinantes que promueven la adhesión y colonización bacteriana, además de los polisacáridos extracelulares, juegan un papel importante las proteínas superficiales de la célula bacteriana, estudiadas principalmente en bacterias patógenas de animales. Estas proteínas una vez producidas en la bacteria deben ser translocadas a

la membrana externa celular o bien ser liberadas por procesos no activos como la lisis celular. De los nueve sistemas de secreción descritos en bacterias Gram negativas, solamente el sistema de secreción tipo VI no ha sido asociado con la secreción de proteínas involucradas en la adhesión bacteriana o a la formación de biopelículas (Chagnot *et al.*, 2013).

En bacterias fitopatógenas, las proteínas asociadas a la adhesión, denominadas adhesinas, se pueden dividir en dos tipos, fimbriales y no fimbriales. Las adhesinas fimbriales son aquellas que tienen una estructura de filamento proteico estructuralmente relacionado con el sistema de secreción tipo II. Uno de los principales ejemplares de este tipo de adhesinas es el pilus tipo IV), fundamental en diversas actividades de la célula bacteriana (Büttner & Bonas, 2010).

1.3.3 El pilus tipo IV y su relación con la adherencia y motilidad bacteriana.

Como se ha mencionado, la adhesión bacteriana es promovida en parte por filamentos conocidos como pili o fimbrias que se extienden desde la superficie bacteriana y están compuestos mayoritariamente por una sola proteína conocida como pilina. Entre los diferentes tipos de pili descritos, los pili tipo IV son los que se encuentran más extendidos entre las diversas especies bacterianas, tanto en los Gram negativos como en los Gram positivos (Figura 4) (Pelicic, 2008). Esta estructura polimérica se encuentra relacionada a diversas actividades bacterianas además de la adherencia, participando también en la motilidad sobre superficies sólidas, la formación de microcolonias y de biopelículas, la señalización celular y la captación de DNA proveniente de procesos como la transformación natural o la unión a bacteriófagos. Los pili tipo IV son considerados por sí mismos como un factor de virulencia, ya que su disrupción en la mayoría de especies Gram negativas lleva asociada una reducción de la virulencia (Craig *et al.*, 2004).

Figura 4. A) Microfotografía de fuerza atómica que muestra el pilus tipo IV (señalado con una flecha amarilla) de *Xylella fastidiosa* cepa Temecula. B) Representación esquemática de la estructura del pilus tipo IV nombrando las proteínas estructurales mínimas necesarias para su formación. Las proteínas se nombran según la nomenclatura utilizada para *Pseudomonas aeruginosa* (Burdman *et al.*, 2011).

Los pili tipo IV, son estructuras delgadas (5-8 nm de ancho), largas (varios micrómetros de longitud), flexibles y extremadamente fuertes, formadas por miles de subunidades de pilina. Estas proteínas, suelen variar entre especies, pero todas parecen conservar una región N-terminal metilada, un residuo hidrofóbico en la posición 25 del extremo N-terminal y un puente disulfuro en el extremo carboxilo-terminal. Las pilinas además, según su secuencia de aminoácidos y su longitud, se dividen en los tipos IVa y IVb, siendo el tipo IVa el que presentan las especies bacterianas Gram negativas que causan enfermedad en las plantas (Craig *et al.*, 2004).

El estudio del pilus tipo IV se ha basado principalmente en el modelo de la bacteria *Pseudomonas aeruginosa*. En cuanto a la estructura de los pili, los componentes mínimos para que genere esta estructura macromolecular son aquellos codificados en los genes *pilA*, *pilB*, *pilC*, *pilD*, *pilQ* y *pilT*. El filamento de este tipo de pilus está conformado por subunidades de la pilina PilA, sintetizadas como prepilinas que son cortadas y metiladas en su extremo N-terminal por la proteína PilD. Las subunidades de PilA son ensambladas por la proteína de la membrana citoplasmática PilC. Posteriormente el filamento es conducido fuera de la membrana exterior por la secretina PilQ. Por último las ATPasas PilB y PilT se encargan del ensamblaje y desensamblaje del filamento, que dan lugar a la extensión y retracción del pili respectivamente (Burdman *et al.*, 2011) (Figura 4).

En el caso de *Xanthomonas*, estructuralmente, se conoce que junto con PilC, las proteínas PilM, PilN, PilO y PilP, conforman una plataforma estructural en la membrana interna en la cual se da el ensamblaje del filamento. Asimismo, se ha determinado que en adición a PilA, otras pilinas menores pueden ser agregadas al filamento. Estas pilinas modulan diversas características como la especificidad de la adhesión, la longitud del filamento o el balance entre retracción y extensión de los pili. Se ha determinado que el número de genes asociados a la formación del pilus tipo IV en *Xanthomonas* se encuentra entre 23, en *X. translucens* pv. *undulosa* y 31 en la mayor parte de especies del género (Dunger *et al.*, 2016).

En cuanto a la relación del pilus tipo IV con la adherencia y la formación de biopelículas, estudios funcionales llevados a cabo principalmente en la especie *X. citri* subsp. *citri*, han demostrado que mutaciones en *pilA*, *pilB*, *pilZ* y *fimX*, conllevan a la incapacidad de las cepas para formar biopelículas estables así como una notable disminución en la capacidad de adherirse a la superficie foliar. Por otro lado, estas cepas al ser infiltradas en el tejido foliar, presentaron una virulencia similar al fenotipo silvestre, demostrando así la importancia del pilus tipo IV en las etapas iniciales de la infección. Resultados similares se han obtenido en mutantes de *pilA* en *X. fuscans* subsp. *fuscans* (Darsonval *et al.*, 2009; Dunger *et al.*, 2014).

1.3.4 Adhesinas no fimbriales y otros componentes asociados a la adherencia y formación de biopelículas bacterianas.

Junto con las adhesinas fimbriales, las adhesinas no fimbriales actúan también en los procesos de adhesión. A diferencia de las anteriores, las adhesinas no fimbriales están constituidas por una sola proteína y se ha propuesto que participan principalmente en la adhesión inicial de las células bacterianas a la superficie y no en las etapas posteriores de agregación celular (Das *et al.*, 2009). Las adhesinas no fimbriales se incluyen dentro de los autotransportadores triméricos del sistema de secreción tipo V y como sustratos del sistema de secreción de dos componentes (Büttner & Bonas, 2010).

Los estudios de genómica comparativa han encontrado variedad de genes que podrían codificar proteínas asociadas con la adhesión superficial (da Silva *et al.*, 2002; Moreira *et al.*, 2010). En *Xanthomonas*, se han realizado estudios funcionales principalmente en las adhesinas FhaB, las homólogas a XadA, XadB y el ortólogo de la proteína YapH de

Yersinia spp. Estos estudios, demuestran una variación en cuanto al papel que juegan estas adhesinas en la virulencia; en *X. oryzae* pv. *oryzae* mutaciones en *xadA* y *xadB* dan lugar a una disminución significativa de la capacidad de adhesión, penetración y virulencia de las bacterias en el tejido foliar del arroz, mientras que mutantes en *yapH* presentan solamente una disminución parcial en estas propiedades (Das *et al.* 2009). Sin embargo, en *X. fuscans* subsp. *fuscans*, YapH parece jugar un papel primordial en la capacidad de las bacterias para adherirse a superficies bióticas y abióticas así como para la formación de biopelículas, mientras que mutaciones en los genes homólogos a *xadA* y en el gen *fhaB* solo producen un fenotipo con una disminución parcial en la capacidad de las células a adherirse a las semillas (Darsonval *et al.*, 2009).

Adicionalmente, otro componente importante de la matriz que conforma la biopelícula, tanto de bacterias Gram positivas como Gram negativas, es el DNA extracelular. Este componente en especies como *P. aeruginosa* y *Bacillus cereus*, se ha mostrado fundamental en la estabilidad de las biopelículas, la resistencia a antibióticos y como fuente de fósforo, carbono y nitrógeno. Recientemente se ha descrito la presencia de DNA extracelular en la matriz de las biopelículas de *X. citri* subsp. *citri*, demostrándose su función asociada principalmente con los estadios iniciales de la formación de estas estructuras así como con su estabilidad (Sena-Vélez *et al.*, 2016).

Los diversos factores que juegan un papel en la adherencia y la formación de biopelículas, reflejan el carácter dinámico y no estático de estos procesos. Dependiendo de las condiciones ambientales que rodean a las bacterias, éstas se adhieren y agregan sobre las superficies del huésped, o se desagregan pasando a la forma planctónica para dirigirse a sitios con condiciones más favorables para proseguir el proceso infeccioso (Kostakioti *et al.*, 2013). Estos procesos implican a nivel celular sistemas de percepción medioambiental y de motilidad, los cuales son descritos a continuación.

1.3.5 Sistemas bacterianos de quorum sensing.

Muchos de los procesos que desencadenan los mecanismos de patogénesis en las bacterias, son dependientes de los niveles de población bacteriana y están sometidas a respuestas, tanto a nivel inter como intraespecífico. En bacterias, este fenómeno, se ha denominado quorum sensing y se basa en señalización química. Generalmente, a estas señales químicas se las conoce como autoinductores y el aumento en su concentración

es directamente proporcional al aumento en la densidad poblacional. Cuando el aumento de los autoinductores alcanza el umbral mínimo estimulador la expresión génica de la bacteria se ve alterada y esto se produce ante diversas condiciones ambientales. La respuesta es desarrollada de manera sincronizada a nivel poblacional por lo que la colonia bacteriana llega incluso a actuar como un organismo pluricelular (Waters & Bassler, 2005).

En bacterias Gram negativas, los autoinductores más extendidos entre los diversos grupos filogenéticos son los conocidos como homoserin lactonas. En el género *Xanthomonas* sin embargo, se produce un autoinductor diferente, el cis-11-metil-2-ácido dodecenoico, conocido también como factor de señalización difusible (DSF). Todas las moléculas de este tipo son sintetizadas por la proteína RpfF, la cual se encuentra codificada en el cluster génico *rpfA-rpfG*. Además del DSF, en *X. campestris* pv. *campestris*, se ha identificado la presencia de otro factor difusible denominado DF, el cual está asociado al gen *pigB*. Ambos factores de señalización se encuentran ampliamente conservados y producidos en la mayor parte de las especies de este género bacteriano (Papenfort & Bassler, 2016).

La producción de DSF por RpfF, está regulada por un sistema de dos componentes que percibe y transduce la señal del DSF y que está compuesto por las proteínas RpfC y RpfG (He & Zhang, 2008). En la señalización mediada por el DSF también juegan un importante papel el mensajero secundario di-GMP cíclico y el regulador transcripcional global Clp (Ryan *et al.*, 2011; Srivastava & Waters, 2012). En *X. campestris* pv. *campestris*, a niveles de baja densidad bacteriana y baja concentración de DSF, la proteína RpfC (con actividad quinasa), interactúa con RpfF (DSF sintasa), lo cual genera una disminución en la cantidad de DSF. A bajas poblaciones bacterianas, el dominio REC de RpfG se encuentra desfosforilado, dando lugar a que el dominio HD-GYP (encargado de degradar el di-GMP cíclico) se encuentre enzimáticamente inactivo, provocando así que las concentraciones celulares del di-GMP cíclico sean altas. Por otro lado, cuando la densidad poblacional aumenta, RpfC percibe el DSF se autofosforila y se produce la liberación de RpfF, una vez liberada esta proteína se sintetiza más DSF. Además, RpfC fosforila el dominio receptor de RpfG, generando en esta enzima la activación del dominio HD-GYP, que degrada el di-GMP cíclico y reduce su concentración a nivel celular. Esta reducción en la concentración del di-GMP cíclico

estimula la activación de diversidad de procesos celulares, principalmente a través de la transcripción del factor Clp (He & Zhang, 2008; Srivastava & Waters, 2012) (Figura 5).

La delección del gen asociado a la producción de DSF, así como de los genes del sistema de dos componentes, produce una reducción en la virulencia de *X. campestris* pv. *campestris*. Por ejemplo, mutantes en *rpfF* ven alterada su expresión en 165 que son considerados los genes “core” del regulón DSF. De éstos, el 80% ven aumentada su expresión ante la presencia del DSF mientras que en el otro 20% la reprimen. Estos genes se encuentran asociados al menos a 12 categorías funcionales, dentro de las cuales destacan, en el ámbito de la patogénesis, aquellos relacionados con la producción de enzimas extracelulares, la síntesis y secreción de lipopolisacáridos y polisacáridos extracelulares, la biosíntesis del flagelo, la motilidad, la quimiotaxis, la respuesta hipersensible y el sistema de secreción tipo III (He & Zhang, 2008).

Figura 5. Representación esquemática del proceso de regulación asociado al sistema de comunicación intercelular quorum sensing, basado en la especie *X. campestris* pv. *campestris* y conservado en las demás especies del género *Xanthomonas* (Srivastava & Waters, 2012).

1.3.6 Otros sistemas sensores asociados a la percepción bacteriana de las condiciones ambientales.

Además del sistema de dos componentes asociado con el fenómeno del quorum sensing, las bacterias presentan otros sistemas de regulación de dos componentes, que son esenciales para la percepción y adaptación de los microorganismos al ambiente. Estos sistemas se encuentran muy conservados a través de los diversos grupos de organismos y se caracterizan por tener una proteína perceptora con actividad histidin quinasa, y una proteína reguladora citoplasmática que actúa como un factor de transcripción. Estructuralmente, los sensores del sistema regulador de dos componentes, son por lo general una proteína homodimérica integrada en la membrana cuyo dominio sensor se encuentra contenido entre dos segmentos de la membrana, y un dominio transmisor que se encuentra localizado en el citoplasma. Algunos de estos receptores histidin quinasa presentan variaciones con respecto a esta estructura básica, pudiendo tener su dominio incluido dentro de la membrana, o ser completamente citoplasmático (Cheung & Hendrickson, 2010) (Figura 6).

La fase correspondiente a la percepción de la señal por la proteína sensora, es una de las etapas del mecanismo de regulación de dos componentes que menos se conoce. La dificultad estriba en el hecho de que cada proteína sensora parece que puede detectar una amplia variedad de señales. El tipo de señales ambientales que detecta la bacteria a través de este sistema es muy variable, pudiendo ser por ejemplo niveles de pH, oxígeno o dióxido de carbono, compuestos de carbono, sustancias quimioatrayentes, productos finales de rutas metabólicas, nutrientes, osmolaridad, luz y diferentes moléculas procedentes del huésped. Una vez detectadas estas señales y transducidas a la proteína reguladora, ésta dispara cambios en la expresión génica en aspectos celulares tales como el metabolismo del carbono, el transporte y asimilación de nutrientes, respuestas al estrés, motilidad, quorum sensing, homeostasis y virulencia (Runyen-Janecky, 2014).

Figura 6. Representación esquemática de la organización básica de los diversos dominios (dominio transmembrana (TM), dominio sensor y dominio transmisor) encontrados en un receptor del tipo histidin quinasa (Cheung & Hendrickson, 2010).

En el caso particular del género *Xanthomonas*, debido a la capacidad de las especies de este género para adaptarse a gran variedad de ambientes, se ha sugerido que este grupo cuenta con un amplio número de genes asociados a proteínas que interactúan en sistemas reguladores de dos componentes. Estas observaciones han sido corroboradas mediante la detección de dominios específicos, determinándose que en los genomas disponibles se pueden encontrar al menos 632 genes hipotéticamente relacionados. A nivel de genomas individuales, el número sigue siendo elevado ya que se han predicho entre 92 y 121 genes (Qian *et al.*, 2008).

Además del sistema de regulación de dos componentes descrito previamente para quorum sensing en *Xanthomonas*, también se ha descrito funcionalmente el sistema de dos componentes denominado RavS/RavR, el cual modula los niveles de di-GMP cíclico mediante la actividad fosfodiesterasa que tiene la proteína RavR en su dominio GGDEF-EAL. La histidin kinasa RavS, que contiene dos sensores tipo PAS asociados a la detección de los niveles bajos de oxígeno, activa a RavR, la cual actúa como factor de transcripción de al menos 206 genes, entre los cuales se encuentran los asociados a la proteína Clp, los genes del sistema de secreción tipo III, enzimas extracelulares y la producción de lipopolisacáridos o polisacáridos extracelulares. La coincidencia entre este sistema regulador de dos componentes y el sistema regulado por RpfC/RpfG en cuanto a la actividad de la proteína Clp, sugiere que la síntesis de factores de virulencia en *Xanthomonas* está controlada tanto por quorum sensing como por la disminución en las concentraciones de oxígeno (Büttner & Bonas, 2010).

En cuanto a los receptores del sistema regulador de dos componentes, en *Xanthomonas* se ha estudiado un repertorio de siete sensores en las especies *X. axonopodis*, *X. campestris*, *X. citri*, *X. oryzae* y *X. vasicola*, encontrándose distinto para cada especie e incluso para patovares de la misma especie, sugiriéndose que estos receptores pueden estar asociados con la especificidad por el huésped (Mhedbi-Hajri *et al.*, 2011).

Las bacterias, además de estos sensores, presentan otro grupo de proteínas de membrana asociadas a la detección y transporte de quelatos de hierro (denominados sideróforos), vitamina B12, complejos de níquel y carbohidratos. Este tipo de transportadores se denominan transportadores TonB-dependientes. Este sistema, transporta sustancias de manera activa para lo que utiliza como recurso energético el complejo proteico TonB, ExbB, ExbD, el cual se encuentra en la membrana interna. Desde el punto de vista

estructural y funcional, todos los transportadores tonB-dependientes presentan un dominio N-terminal que se ubica dentro de un dominio del tipo barril beta ubicado en el extremo C-terminal. Además, cerca del extremo N-terminal presenta una región denominada caja TonB, la cual se sitúa hacia el periplasma pero se mantiene replegada dentro del dominio barril beta en ausencia de ligando. Una vez se une el ligando a esta región, se produce un cambio conformacional que expone la caja TonB, que interacciona subsecuentemente con la proteína TonB, dándose así el transporte a través de la membrana interna con ayuda de una proteína periplásmica de unión y un transportador tipo ABC (Noinaj *et al.*, 2010) (Figura 7).

Figura 7. Representación esquemática del transporte y regulación de sideroforos férricos mediante transportadores TonB-dependientes (Noinaj *et al.*, 2010).

En relación al estudio de este tipo de transportadores en *Xanthomonas*, se ha determinado de manera predictiva que el repertorio de estos transportadores en los diversos genomas es alto, por ejemplo en *X. campestris* se han predicho al menos 72, muchos de los cuales se hipotetiza que pueden estar asociados con la adquisición de carbohidratos. Entre éstos, el transportador TonB-dependiente SuxA presenta alta afinidad por la sacarosa y se ha visto que actúa en conjunción con un transportador interno de membrana, una amilosucrasa y una proteína reguladora. Todos éstos forman parte de un mismo locus, el cual ha sido denominado CUT. Bacterias mutantes para este locus han mostrado un fenotipo poco virulento de este patógeno sobre

Arabidopsis thaliana, demostrando así la relación de estos transportadores con el desarrollo de la enfermedad producida por *X. campestris* (Blanvillain *et al.*, 2007).

Este tipo de asociación, además ha sido comprobado en *X. citri* subsp. *citri*, en la cual el transportador TonB-dependiente *XAC4131* ralentiza el desarrollo de la respuesta hipersensible en tabaco. Asimismo, al ser la bacteria cultivada en medio inductor del sistema de secreción tipo III (genes *hrp*), *XAC4131*, se controla indirectamente la actividad del regulón *hrpG* a través de los genes *rpoE* y *XAC4131* (Aini *et al.*, 2010). Por otro lado, al igual que en los sensores de dos componentes, el estudio comparativo de la presencia de 27 transportadores TonB-dependientes, ha demostrado la existencia de una gran variación a nivel de especie, e incluso a nivel de patovar, por lo que no se descarta su asociación con la afinidad por un determinado huésped en los estadios iniciales de infección (Mhedbi-Hajri *et al.*, 2011).

1.3.7 La quimiotaxis como mecanismo de respuesta ambiental y su relación con el movimiento celular.

Uno de los sistemas más estudiados de respuesta como consecuencia de la percepción de una determinada condición ambiental es la quimiotaxis, que regula la motilidad bacteriana. La quimiotaxis es la base del movimiento de las células bacterianas hacia regiones que contienen una mayor concentración de sustancias químicas beneficiosas, o hacia sitios con una menor concentración de sustancias químicas tóxicas (Wadhams & Armitage, 2004). Tanto las moléculas atrayentes como las repelentes son detectadas por quimiorreceptores citosólicos o unidos a la membrana, llamados proteínas quimiotácticas aceptoras de grupos metilo o MCPs (Jacques *et al.*, 2016).

Desde el punto de vista estructural y funcional, las MCPs se han descrito como homodímeros compuesto casi completamente por estructuras tipo alfa hélice (Wadhams & Armitage, 2004). La estructura del dominio periplásmico de las MCPs suele ser muy variable entre los diversos grupos bacterianos poseyendo además un amplio espectro de ligandos que son detectados por estos quimiorreceptores. El proceso por el cual el sistema quimiotáctico regula el movimiento bacteriano se inicia con la unión del ligando al dominio periplásmico de la MCP, lo que altera las interacciones que se dan entre las hélices transmembrana y permite propagar la señal a través de la membrana bacteriana. En su región citoplasmática las MCPs se encuentran asociadas a dos

proteínas CheA y CheW. La proteína CheW, cuya estructura y función está muy conservada entre las distintas especies, es la encargada de transducir la señal entre la MCP y la proteína CheA. CheA, que es un dímero que interactúa con 3 dímeros de la MCP y dos dominios de CheW, tiene la capacidad de autofosforilarse y aumenta su actividad en respuesta a la disminución de la unión quimioatrayente, o por un incremento en la unión de un quimiorrepelente al sensor MCP.

A su vez, otras proteínas citosólicas CheY y CheB compiten por unirse al dominio P2 de CheA generándose una cascada de fosforilación desde el residuo His48 de CheA al residuo Asp57 de CheY o Asp56 de CheB. La fosforilación de CheY, que es más afin y rápida que la de CheB, permite que esta interactue con la proteína FliM, que forma parte del complejo proteico que conforma el motor flagelar. Esta unión da lugar a un movimiento del flagelo en dirección de las manecillas del reloj y hace que la célula realice un tumbo y cambie de dirección. Este proceso acaba con la señal de finalización dada por la proteína CheZ la cual defosforila a CheY, por lo que la actividad de CheZ se asocia directamente con la concentración del CheY fosforilado que haya en el citoplasma (Figura 8).

Figura 8. Diagrama esquemático del sistema de quimiorrepción de *Escherichia coli* (Wadhams & Armitage, 2004).

En el género *Xanthomonas*, la expresión de proteínas asociadas a quimiotaxis durante el inicio de la infección ha sido estudiado en *X. citri* subsp. *citri*. En este caso se ha comprobado la expresión de 26 proteínas relacionadas a la quimiotaxis y 11 proteínas relacionadas con la biogénesis y regulación del flagelo durante los primeros estadios de la infección. En cuanto a las MCPs, de las 21 proteínas predichas para este genoma, solamente de dos no se demostrado su expresión. En relación a proteínas asociadas a la transducción de la señal quimiotáctica, se ha detectado la expresión de CheA, CheB, CheR, CheV, CheW, CheY y CheZ solamente cuando las bacterias son inoculadas en el huésped. Este hecho demuestra que la maquinaria proteica asociada a la respuesta quimiotáctica es potenciada solamente cuando los sensores detectan determinados metabolitos que son propios de la planta. Es por tanto un paso esencial para la activación del flagelo en el espacio intercelular, y lleva a un proceso de colonización más efectivo. Basado en el perfil proteómico a lo largo del tiempo, se ha observado una activación previa adaptativa de las células mediante la expresión de las MCPs y la posterior transducción de la señal por medio de las proteínas Che, lo que genera seguidamente una expresión al unísono de los genes de regulación y actividad flagelar, los genes asociados al pili tipo IV, la producción de xanthano y la secreción de efectores del sistema de secreción tipo III (Moreira *et al.*, 2015).

Al igual que lo observado para los otros sensores ambientales mencionados previamente, se ha observado una variación en el número y el repertorio de MCPs presente en las diversas especies y patovares del género *Xanthomonas*, siendo este más amplio en las cepas de la especie *X. axonopodis*. Esta variación incluso a nivel intraespecífico, indica que además de estar relacionados con el inicio de la infección al detectar los metabolitos del hospedero, cumplen una función en cuanto a la gama de huésped de las diversas cepas (Mhedbi-Hajri *et al.*, 2011; Jacques *et al.*, 2016).

1.3.8 Estructura y función del flagelo bacteriano.

Como se describe en apartados anteriores, el sistema de regulación quimiotáctica está íntimamente relacionado con la actividad del flagelo. Según Rossez y colaboradores (2015) esta organela, presenta dos actividades primordiales durante el inicio de la colonización bacteriana, la primera de éstas se relaciona con el movimiento. El flagelo está directamente implicado en el paso de las células desde un estado sésil a un estado mótil en las etapas maduras de las biopelículas. Además, en condiciones en las que

disminuye la concentración de nutrientes el flagelo permite que células individuales flageladas actúen como pioneras en busca de ambientes más favorables, potenciando la formación de microcolonias.

En cuanto a la adherencia, se ha confirmado en algunos grupos de bacterias fitopatógenas como *P. syringae*, y en bacterias fijadoras de nitrógeno como *Azospirillum brasilense*, la necesidad del flagelo en procesos iniciales de adherencia. Además se ha descrito que el flagelo puede intervenir en esta etapa de diversas maneras, una de ellas estaría asociada con la longitud de la organela, ya que ésta al ser muy larga, permitiría un anclaje inicial inespecífico al tejido del huésped. Posteriormente, la rotación del flagelo en las cercanías del tejido del huésped puede generar fuerzas que propicien una interacción más cercana entre las membranas de ambos organismos. Por último, el flagelo, al estar formado por gran cantidad de unidades repetidas de la proteína flagelina, podría generar uniones iónicas no específicas con estructuras o proteínas de la célula del huésped, que aunque son de baja afinidad, son muy numerosas.

Desde una perspectiva estructural, el flagelo es un filamento delgado y largo, de hasta 20 μm , que sobresale desde el cuerpo celular. Esta organela se conforma de tres estructuras, el cuerpo basal, conocido también como el motor flagelar, el filamento, el cual actúa como un propulsor y la estructura denominada garfio, que sirve de unión entre las dos anteriores (Chevance & Hughes, 2008).

El motor flagelar es la estructura proteica más compleja que se conoce en las células bacterianas. Las proteínas que lo conforman abarcan desde el citoplasma hasta el exterior de la célula, atravesando la membrana celular. El motor flagelar convierte el flujo de iones que atraviesa la membrana citoplasmática en un torque que hace que el flagelo rote, una rotación en contra del sentido de las manecillas del reloj permite a la bacteria dirigirse hacia adelante, mientras que en sentido contrario permite a la célula dar tumbos al azar que reorientan su dirección (Wadhams & Armitage, 2004).

El motor flagelar ha sido extensamente estudiado y se ha visto que está compuesto por más de 20 proteínas que se organizan en varias estructuras (Chen *et al.*, 2011), (Figura 9). La primera de éstas es el anillo M-S el cual se encuentra en la membrana interna y esta formado solamente por la proteína FliF. Seguidamente se encuentra el aparato exportador que se encarga de exportar a través de la membrana las proteínas que

conforman la estructura tipo varilla, el garfio y el filamento. Este aparato consiste en seis proteínas transmembrana (FlhA, FlhB, FliO, FliP, FliQ, FliR) y tres proteínas solubles (FliH, FliI, FliJ). Alrededor del Anillo M-S y del aparato exportador se ubica el anillo C, que se encuentra estructurado por las proteínas FliG, FliM y FliN y está asociado a la exportación, la generación del torque y el cambio en la dirección rotacional. Seguidamente, en la parte superior del anillo S y actuando como un eje de transmisión del torque desde el anillo M-S al garfio y al filamento, se encuentra la estructura tipo varilla (“rod structure”), la cual está conformada por FliE, FlgB, FlgC, FlgF y FlgG. Sobre esta estructura se encuentran los anillos L y P, que sirven como eje de unión de la estructura a través de las capas de peptidoglicano y polisacárido extracelular; las dos proteínas que lo conforman FlgH y FlgI son transportadas al periplasma a través del sistema tipo Sec. Sobre el anillo C y rodeando el M-S se encuentra el complejo iniciador, el cual junto con las proteínas del anillo C convierte el flujo de H^+ o Na^+ en un torque que da energía al motor, este complejo está formado por las proteínas MotA y MotB.

Figura 9. A) Representación esquemática de la estructura y de los componentes de flagelo de *Salmonella enterica* (Chevance & Hughes, 2008). B) Imagen de microscopía electrónica y representación esquemática de los componentes estructurales que conforman el motor flagelar de *Escherichia coli* (Chen *et al.*, 2011).

Finalmente, el motor flagelar está unido al filamento mediante el garfio que se compone de las proteínas FlgE, FlgD, FliK, FlgK y FlgL. Una vez que se han exportado las proteínas de la estructura tipo varilla y del garfio, este último se va ensamblando y al alcanzar un tamaño aproximado de 50 nm, se produce una interacción entre la proteína del garfio FliK, que actúa como una vara de medir molecular, y la proteína del sistema transportador FlhB, lo cual genera un cambio de afinidad del sistema transportador, haciendo que sea más afín a la exportación de las proteínas FliC, FliD y FliM, que son las que conforman el filamento del flagelo (Chevance & Hughes, 2008) (Figura 9).

1.3.9 Tipos de movimiento bacteriano relacionados con el flagelo.

Dentro de los diversos tipos de movimiento descritos en bacterias, dos de ellos están asociados a la utilización del flagelo, el movimiento tipo “swimming” y el movimiento tipo “swarming”.

El primero de estos, es un movimiento compartido por un gran número de grupos procariotas y se refiere al movimiento bacteriano en un ambiente líquido o viscoso. El movimiento tipo “swimming” (Figura 10) es individual y se produce generalmente de forma directa como consecuencia de su activación por un mecanismo de la quimiotaxis. Cuando el sistema quimiotáctico percibe un aumento en la concentración de un atrayente, el sentido de rotación de los flagelos se da en contra de las manecillas del reloj y la célula se propulsa en dirección y sentido del recurso atrayente. Cuando el flagelo cambia el sentido de la rotación se produce un tumbos en la célula bacteriana y un cambio de dirección en su desplazamiento, un cambio posterior en el sentido de la rotación, hace que las células vuelvan a moverse mediante “swimming” hacia adelante. Ante sustancias atrayentes, los tumbos de las células son suprimidos lo cual permite a la célula ir directamente hacia el recurso (Girgis *et al.*, 2007).

A diferencia del movimiento tipo “swimming”, el “swarming” no es un movimiento asociado a células individuales sino que es un movimiento bacteriano rápido y multicelular a través de una superficie (Figura 11). Este tipo de movimiento se ha encontrado únicamente limitado a algunos grupos bacterianos presentes en Firmicutes, Alfa proteobacteria y Gamma proteobacteria. En algunos casos las células bacterianas, durante este proceso, adquieren características fenotípicas particulares, tales como un aumento en su tamaño o la pluriflagelación (Kearns, 2010).

Figura 10. A) Representación esquemática del comportamiento de las células bacterianas de *Escherichia coli* durante el proceso de movimiento tipo “swimming” (Terashima *et al.*, 2008). B) Detalle del fenotipo asociado al movimiento tipo swimming en medio de cultivo semisólido (medio MMA al 0,3% de Agar) de la bacteria *Xanthomonas citri* subsp. *citri* (Sena-Vélez *et al.*, 2015).

Además del flagelo, para que se dé el movimiento tipo swarming, también son requeridas la síntesis y secreción de sustancias surfactantes (surfactina, serrawetina ó ramnolípidos) por parte de las células. Estas moléculas tienen como función reducir la tensión entre la superficie y la célula, permitiendo así su propagación. La producción de surfactantes está mediada por quorum sensing, por lo que son sintetizados y secretados únicamente cuando hay una alta densidad poblacional bacteriana. En algunas bacterias como *E. coli*, se ha observado la existencia de “swarming” pero no de sustancias surfactantes, para estos casos, se ha comprobado que la misma actividad atribuida a éstos es desarrollada de manera análoga por los lipopolisacáridos de la membrana bacteriana (Kearns, 2010). En bacterias del género *Pseudomonas*, se ha demostrado que además del flagelo y del surfactante es requerido el pilus tipo IV para realizar este tipo de movimiento superficial (Kohler *et al.*, 2000).

Se han sido descritos mediante estudios genómicos y fenotípicos, tanto la composición como regulación de la síntesis del flagelo, así como la existencia de movimiento dependiente del flagelo, en *Xanthomonas*. Estudios de genómica comparativa (da Silva *et al.*, 2002) permitieron conocer los genes asociados a las rutas biosintéticas del flagelo y del sistema quimiotáctico en *Xanthomonas*. Además la descripción de nuevos genomas de este género (Jacobs *et al.*, 2015) han permitido determinar que están organizados en cuatro grandes grupos de genes (*fle*, *flg*, *flh* y *fli*) en una región de alrededor de 150 Kpb. Uno de estos grupos presenta además varios genes asociados a MCPs que se encuentran en tandem, el número de MCPs en esta región varía dependiendo de la especie de *Xanthomonas*. Estudios funcionales además han

demostrado que en *Xanthomonas* el regulador de la biosíntesis del flagelo es la proteína FleQ. Bacterias mutantes en su gen codificante presentan fenotipos aflagelados y sin capacidad de movimiento. La actividad reguladora de FleQ se ha comprobado directamente sobre la expresión de los genes *fliE*, *fliQ*, *fliL*, *flgG*, *flgB*, *flgH* y *flhF* (Hu *et al.*, 2005).

Figura 11. A) diversos fenotipos de colonias bacterianas durante el movimiento tipo “swarming” (de derecha a izquierda, *Bacillus subtilis*, *Proteus mirabilis*, *Pseudomonas aeruginosa*) (Kearns, 2010). B) Diferenciación morfológica celular entre células de *P. aeruginosa* realizando movimiento tipo “swarming” (izquierda) y células en fase estática (derecha) (Kohler *et al.*, 2000).

Por otro lado, estudios desarrollados sobre una librería de 6000 mutantes de *X. citri* subsp. *citri*, han permitido identificar que además de los genes asociados con la estructura y biogénesis del flagelo, en esta especie, la mutación en 11 genes (XAC0019, XAC0533, XAC0618, XAC0695, XAC0804, XAC1043, XAC1081, XAC1098, XAC1495, XAC2113, XAC2583) produjo bacterias con fenotipos incapaces de desarrollar movimiento tipo “swimming”, por lo que se deduce que son genes que actúan de manera directa o indirecta sobre la funcionalidad del flagelo (Malamud *et al.*, 2013). Los estudios con mutantes han permitido determinar que la actividad flagelar en *X. oryzae* pv. *oryzae* además de estar relacionada con los receptores ambientales tipo MCP, también se asocia a otro tipo de estímulos ambientales, sistemas reguladores de dos componentes conectados a otras actividades propias de la patogénesis en esta

especie como la producción de polisacáridos extracelulares. Estos sistemas de dos componentes asociados a la regulación de la actividad del flagelo corresponden a los relacionados con las proteínas RpfC/RpfG, PdeK/PdeR, ColS/ColR, StoS, SreK, SreR y SreS (Zheng *et al.*, 2016).

La constatación de papel que juega el sistema de dos componentes RpfC/RpfG, como uno de los reguladores de la actividad flagelar y por lo tanto del movimiento bacteriano, demuestra una vez más, la clara relación entre el sistema de quorum sensing y la motilidad, tanto en *X. citri* subsp. *citri* (Malamud *et al.*, 2011) como en *X. campestris* (Jeong *et al.*, 2008). En ambos modelos se ha observado que mutantes en componentes asociados al quorum sensing como los genes *rpfB*, *rpfC*, *rpfF* y *rpfG* dan lugar a fenotipos con una disminución significativa en la capacidad para moverse. Se ha determinado además, que las proteínas que interconectan ambos sistemas en *X. campestris*, son las codificadas por los CDS XC_0249 y XC_0420, ambas con actividad sintasa para del diGMP-cíclico.

La necesidad del flagelo como organela relacionada con el proceso de patogénesis se ha demostrado en *Xanthomonas* (Malamud *et al.*, 2011) a partir de mutantes de los genes *fliC* (flagelina) y *flgE* (garfio del flagelo) que presentando ausencia de flagelo, veían reducida significativamente la motilidad tipo “swimming”, la formación de biofilm, la adherencia y la virulencia en *X. citri* subsp. *citri*. Estas mismas variaciones fenotípicas han sido también descritas en *X. oryzae* al mutar los genes *flgD* y *flgE* (Yin *et al.*, 2011).

En cuanto al movimiento tipo swarming, en el género *Xanthomonas*, a pesar de considerarse inicialmente que las especies *X. citri* y *X. oryzae* presentaban este tipo de movimiento, actualmente se postula que en realidad corresponde al movimiento tipo “sliding” que se define como la translocación pasiva sobre una superficie potenciada por el crecimiento de la colonia bacteriana y que se produce tanto en células flageladas como aflageladas (Malamud *et al.*, 2011). Se ha observado la capacidad de realizar movimiento tipo swarming en la especie *X. perforans*, pero únicamente asociado a cambios ambientales provocados por otras especies bacterianas de su entorno, como *Proteus mirabilis*, con capacidad para este tipo de movimiento (Hagai *et al.*, 2014; Venieraki *et al.*, 2016).

Por otra parte, el movimiento tipo swimming ha sido observado en varias especies de *Xanthomonas*. Estudios recientes han revelado la existencia de variación en este

fenotipo entre diversas cepas de la misma especie dependiendo de la gama de huésped que las mismas presenten (Sena-Vélez *et al.*, 2015). Es así, como para *X. citri* subsp. *citri*, se ha descrito que en condiciones que asemejan el apoplásto de los cítricos, aquellas cepas menos virulentas y con una gama de huésped restringida a lima mejicana presentan mayor capacidad para moverse mediante este tipo de movimiento que aquellas cepas más virulentas y con una gama de huésped más amplia.

Los mecanismos iniciales de infección descritos a lo largo de este apartado están relacionados con procesos de supervivencia de la bacteria y establecimiento de su población en el huésped. Una vez establecida ésta, se produce la expresión y secreción de enzimas y proteínas efectoras que interactúan directamente sobre los procesos fisiológicos de la planta.

1.4 Sistemas de secreción y proteínas efectoras asociadas a procesos tardíos de infección en *Xanthomonas*.

Las bacterias poseen una serie de estructuras macromoleculares encargadas de secretar gran variedad de sustratos, entre ellos se encuentran pequeñas moléculas, proteínas y DNA. Los sustratos secretados desarrollan actividades que son esenciales para la supervivencia de las bacterias ante condiciones ambientales adversas así como para desencadenar mecanismos relacionados con la patogenicidad. Dependiendo del sistema de secreción, la molécula secretada puede mantenerse asociada a la membrana externa, secretada al ambiente extracelular o ser inyectada directamente dentro de la célula del huésped. La mayoría de los patógenos requieren la combinación de diversos sistemas de secreción para colonizar de manera exitosa un determinado organismo huésped (Costa *et al.*, 2015).

El género *Xanthomonas* presenta al menos seis sistemas de secreción y dos vías por las cuales se secretan enzimas degradadoras y proteínas efectoras al ambiente extracelular (Büttner & Bonas, 2010).

1.4.1 Sistema bacteriano de secreción tipo II.

El sistema de secreción tipo II (T2SS) se encuentra ampliamente distribuido entre las bacterias Gram negativas, tanto patógenas como no patógenas. Este sistema se encarga de secretar proteínas desde el periplasma hasta el exterior de la célula y sus sustratos corresponden a enzimas hidrolíticas y toxinas (Costa *et al.*, 2015). Se compone por una estructura macromolecular que contiene entre 12 y 15 proteínas que por lo general están codificadas en un mismo operón. Se ha sugerido que la estructura del T2SS se expande a través de la membrana interna y externa de la célula en donde cuatro diferentes estructuras pueden ser distinguidas, el pseudopilus, el complejo de la membrana externa, la plataforma de la membrana interna y la ATPasa de secreción (Korotkov *et al.*, 2012) (Figura 12).

Figura 12. Representación esquemática del sistema de secreción tipo II en bacterias Gram negativas, la nomenclatura de las proteínas se basa en la vía general de secreción (Korotkov *et al.*, 2012).

En especies como *X. campestris* pv. *campestris* y *X. campestris* pv. *vesicatoria*, se ha demostrado que enzimas degradadoras, como las xylanases, proteasas y lipasas, que son inicialmente consideradas sustratos del sistema de secreción tipo II (T2SS), siguen siendo secretadas incluso en mutantes defectivos para este sistema de secretor. Esto es

posible debido a que en estas especies, este grupo de enzimas se puede secretar también a través de vesículas de la membrana externa (Solé *et al.*, 2015). El T2SS se relaciona con patogénesis en algunos grupos de bacterias fitopatógenas como *Dickeya*, *Erwinia*, *Pseudomonas*, *Ralstonia*, *Xanthomonas* y *Xylella* (Jha *et al.*, 2005). En estas especies, se considera que las enzimas secretadas por el T2SS contribuyen a la degradación de la pared celular. El T2SS podría facilitar el ensamble de apéndices extracelulares asociados a la secreción de proteínas de virulencia, como lo es el sistema de secreción tipo III (T3SS) (Büttner & Bonas, 2010). Un hallazgo a favor de esta hipótesis es el hecho de que la regulación de este sistema de secreción en *X. citri* subsp. *citri* se encuentra en parte regulado por el regulón HrpG, el cual es además un regulador clave para la expresión del T3SS en *Xanthomonas* (Yamazaki *et al.*, 2008).

Gracias a la secuenciación y comparación de genomas en *Xanthomonas*, se conoce que este género presenta uno (*X. oryzae*) o dos (*X. campestris* pv. *campestris*, *X. campestris* pv. *vesicatoria* y *X. citri* subsp. *citri*) T2SS, los cuales son conocidos como Xcs y Xps (Büttner & Bonas, 2010). El organismo de este grupo en el que más se ha estudiado este sistema secretor es *X. campestris* pv. *vesicatoria* en el cual se encuentra tanto el T2SS Xcs (formado por el cluster *xcsC-xcsN*) como el Xps (formado por el cluster *xpsD-xpsN*). Funcionalmente, se ha demostrado que el sistema Xps es promotor del desarrollo de la enfermedad y además contribuye a la translocación de proteínas efectoras secretadas por el T3SS. Los análisis transcriptómicos indican que este grupo génico es activado por HrpG y HrpX, que constituyen los reguladores del T3SS. Por otro lado, con respecto al sistema Xcs, mutantes en los componentes de este grupo génico no presentan un cambio significativo en cuanto a su virulencia, consecuencia probablemente de su complementación por los genes homólogos al sistema Xps (Szczesny *et al.*, 2010).

En estos grupos bacterianos, se ha demostrado como sustratos del T2SS una serie de enzimas con actividad pectato liasa, celulasa, endoglucanasa, proteasa, α -amilasa, xilanasa, cisteín proteasa, cellobiosidasa y lipasa/esterasa (Jha *et al.*, 2005; Szczesny *et al.*, 2010). En las especies *X. campestris* pv. *campestris*, *X. campestris* pv. *vesicatoria* y *X. oryzae* pv. *oryzae*, se ha demostrado mediante análisis de fenotipos mutantes que la ausencia de estas proteínas secretadas da lugar a una disminución de la virulencia de las cepas.

1.4.2 Sistema bacteriano de secreción tipo III.

Además del T2SS, uno de los principales factores de virulencia descrito en bacterias Gram negativas, es el sistema de secreción tipo III (T3SS), cuyo origen evolutivo es compartido con el flagelo. Se encuentra ampliamente distribuido en la naturaleza, tanto en organismos no patógenos como patógenos de plantas, insectos y vertebrados (Galan & Wolf-Watz, 2006). Esta estructura macromolecular se especializa en la transferencia de proteínas efectoras hacia el citoplasma o la membrana plasmática del huésped. Una vez dentro, estos efectores modulan o minan diversas funciones celulares específicas, lo cual promueve la colonización e invasión bacteriana (Costa *et al.*, 2015).

En bacterias fitopatógenas, el T3SS ha demostrado ser esencial para la patogénesis bacteriana y la inducción de respuesta hipersensible en interacciones planta patógeno compatibles e incompatibles, respectivamente. Este sistema secretor se encuentra codificado en el “cluster” génico denominado *hrp*, el cual está compuesto por más de 20 genes organizados en diferentes unidades transcripcionales (Büttner & Bonas, 2010).

Nueve de los más de 20 componentes del T3SS identificados se encuentran conservados tanto en patógenos de plantas como de animales. Son los denominados como genes *hrc* y conforman el aparato secretor que se encuentra en las membranas interna y externa. Ambos aparatos son estructuras proteicas con forma de anillo que están presumiblemente conectadas entre sí mediante una estructura periplásmica cilíndrica. Esta estructura como un todo conforma el cuerpo basal del T3SS. El cuerpo basal a su vez se encuentra asociado a un pilus extracelular el cual forma en su extremo un traslocón que permite la interacción con la membrana plasmática de la célula del huésped. La energía requerida tanto para el acoplamiento como para el despliegue de los sustratos del T3SS, es decir los componentes extracelulares del propio sistema y sus proteínas efectoras, es probablemente provista por una ATPasa asociada al cuerpo basal del T3SS en su extremo citoplasmático (Büttner, 2012) (Figura 13).

Figura 13. Representación esquemática de la estructura del sistema de secreción tipo III, presente en bacterias Gram negativas (Büttner, 2012).

Como ya se ha mencionado, entre los sustratos translocados a través del T3SS, se encuentran las proteínas denominadas efectores. El término efector se refiere a un grupo de sustratos del T3SS que ejercen una función en el interior de la célula del huésped, alterando una amplia variedad de procesos celulares y fisiológicos asociados a la patogénesis. En patógenos vegetales, incluyen a las proteínas que pueden actuar como factores de virulencia, promoviendo la enfermedad o como factores de avirulencia cuando activan las defensas del huésped. La supresión de la respuesta inmune de la planta juega un papel primordial en el establecimiento y multiplicación bacteriana en el huésped (Büttner & Bonas, 2010). La mayoría de los efectores descritos en patógenos vegetales han sido designados como proteínas de avirulencia o proteínas Avr. Otro grupo de efectores han sido descritos al observar su capacidad de ser translocados por el T3SS, estos son designados de diferente manera según la especie bacteriana, pero en el caso de las *Xanthomonas* se les conoce como proteínas Xop (Alfano & Collmer, 2004; Bogdanove *et al.*, 2011; Feng & Zhou, 2012).

Las proteínas efectoras son reconocidas mediante una señal de 20-30 aminoácidos localizada por lo general en el extremo N-terminal. A pesar de que este extremo de la proteína no suele permanecer conservado entre los diversos efectores, la mayoría contienen una composición o un patrón específico. Además, mediante ensayos de fusión

Las proteínas codificadas por *hrpB* y *hrpE*, se ha demostrado que además de ser secretadas, y formar parte del apéndice extracelular del T3SS, son esenciales para que se dé la secreción de otras proteínas a través de este sistema. Estudios realizados a partir de mutantes para el “cluster” *hrp* en *X. campestris* pv. *vesicatoria*, han permitido demostrar que los genes denominados *hpa* (genes asociados a *hrp*), contribuyen en la patogénesis pero no son esenciales para ello. Por ejemplo, el gen *hpaA*, cuya proteína es secretada, es necesario para una rápida y efectiva respuesta hipersensible por parte de la planta, mientras que la proteína HpaB, que permanece en el citoplasma, es requerida para el control de la exportación de las demás proteínas a través del T3SS (Gürlebeck *et al.*, 2006). Por otro lado, mutantes en el cluster *hrp* de *X. citri* subsp. *citri*, han permitido reconocer a la proteína HrcN como la ATPasa que da la energía necesaria para que se lleve a cabo este sistema de transporte proteico (Gottig *et al.*, 2010) (Figura 15).

Figura 15. Representación gráfica del sistema de secreción tipo III en *X. citri* subsp. *citri* (Gottig *et al.*, 2010).

El T3SS de cada una de las especies de *Xanthomonas*, secreta un grupo diferente de proteínas efectoras en el citoplasma de su huésped. Mediante análisis bioinformático así como en trabajos de laboratorio, se han encontrado alrededor de 60 efectores ó posibles efectores del T3SS en este género bacteriano (White *et al.*, 2009). El repertorio varía

según la especie, habiéndose encontrado entre 24 efectores en *X. citri* subsp. *citri* hasta 37 en la especie *X. oryzae* pv. *oryzae* (Büttner & Bonas, 2010).

Los efectores de *Xanthomonas* se encuentran divididos en subgrupos según sea su origen evolutivo, algunos ejemplos de estos subgrupos son los efectores de la familia XopJ o los de la familia AvrBs3/PthA. El repertorio de efectores de este género refleja ser producto de su adquisición mediante transferencia horizontal, y muchos de éstos parecen proceder de otras especies fitopatógenas como *Pseudomonas syringae*, *Ralstonia solanacearum* y *Acidovorax avenae* (Büttner & Bonas, 2010).

Funcionalmente, se ha determinado en varias especies de *Xanthomonas* que algunos de sus efectores del T3SS disminuyen la respuesta de defensa de la planta, así como la expresión de genes relacionados con la defensa de la planta (Figura 16). Por ejemplo, se ha determinado que los efectores XopD, XopN y XopX reducen la PTI (“PAMP-Triggered Immunity”) e inhiben la activación de rutas metabólicas relacionadas con la ruta del ácido salicílico y la deposición de callosa, entre otros. La actividad de los efectores, no solamente se ha visto asociada a la supresión de la PTI del huésped, sino también a la respuesta de la planta asociada a la detección mediante su sistema de defensa de los efectores del patógeno (ETI) (“Effector-Triggered Immunity”), la cual da como resultado una rápida respuesta hipersensible. En *Xanthomonas*, la supresión de la ETI se ha visto asociada por ejemplo al efector AvrBsT de *X. oryzae* pv. *oryzae* (Boch & Bonas, 2010).

Por otro lado, determinados efectores del T3SS en este género bacteriano han sido asociados con un aumento en la virulencia así como en la sintomatología de la enfermedad. El ejemplo más visible ha sido el descrito en *X. citri* subsp. *citri* y en *X. oryzae* pv. *oryzae* en los cuales mutaciones en el gen asociado al efector relacionado con la activación de la transcripción (efectores tipo TAL) (“transcription activator-like”), *pthA*, genera una marcada reducción en la capacidad del patógeno para dispersarse y de provocar síntomas propios de la enfermedad. Un efecto semejante ha sido descrito en *X. campestris* pv. *vesicatoria* con el efector tipo TAL AvrBs3 (Boch & Bonas, 2010). En esta misma especie, aquellas cepas que no contienen XopD o XopN en su repertorio de efectores, presentan en la planta poblaciones significativamente menores así como muestran una reducción en la necrosis y la senescencia de la planta en etapas tardías de la infección (Kim *et al.*, 2013). Además, se ha visto en diversas especies *Xanthomonas*,

una disminución de la virulencia en cepas con mutaciones en los genes que codifican para los efectores XopAE, XopAH, XopJ, XopX y XopZ (Song & Yang, 2010).

Figura 16. Representación gráfica del modelo molecular de como actúan los efectores del sistema de secreción tipo III de *Xanthomonas* una vez son secretados en el citoplasma de la célula del hospedero. En este caso se representa a través del efector XopD y del efector asociado a la activación de la transcripción (efectores tipo TAL) AvrBs3 (Kay & Bonas, 2009).

Para el género *Xanthomonas* también se han llevado a cabo estudios sobre el proceso de regulación asociado al T3SS y sus efectores. A pesar de que el sistema de quorum sensing se ha comprobado que está involucrado en la regulación de los genes *hrp*, no es único para la expresión de los mismos. La regulación de este grupo génico depende de un regulador clave para la patogénesis en este género bacteriano, HrpG, asociado a la expresión de los genes del T3SS de sus efectores y de al menos 29 sustratos asociados al T2SS. La expresión de este regulador se produce cuando la bacteria entra en el apoplasto o bien cuando se cultiva en medio mínimo (Guo *et al.*, 2011).

HrpG activa la expresión de *hrpX*, el cual tiene un activador transcripcional tipo Ara-C que se une a una región conservada del genoma conocida como región inducible por la planta o “PIP-box” que está presente en la región promotora de la mayoría de genes inducidos por HrpG. Sin embargo, también induce la expresión de otro grupo de genes que no contienen el PIP-box por lo que esta región parece no ser indispensable para la inducción por parte de HrpX. Mutantes tanto de *hrpG* y *hrpX* generan fenotipos no patogénos (Büttner & Bonas, 2010). Además de estos dos reguladores, la expresión de *hrpC* y *hrpE* es controlada por el sistema de dos componentes ColR/ColS, lo que sugiere que diversas rutas metabólicas están implicadas en la regulación del “cluster” *hrp* (Yan & Wang, 2011).

Junto con los reguladores mencionados, la proteína HpaS que es kinasa sensora que forma parte de un sistema de dos componentes con el regulador HrpG, tiene un papel primordial. Estudios mutacionales en *hpaS* indican una disminución de la fosforilación de HrpG que implica una disminución de su actividad. HpaS no solamente conforma un sistema regulador de dos componentes con HrpG, sino que también con la proteína HpaR2, es requerida para expresar diversos factores de virulencia pero no para generar una respuesta hipersensible (Li *et al.*, 2014; Jacobs *et al.*, 2015).

1.4.3 Sistema Bacteriano de secreción tipo IV.

Junto con los dos sistemas de secreción descritos previamente, otro sistema asociado a la patogénesis bacteriana es el sistema de secreción tipo IV (T4SS). Este sistema macromolecular encontrado tanto en Gram negativos como Gram positivos permite la traslocación tanto de DNA como de proteínas hacia el citoplasma de células procariotas y eucariotas. El T4SS está implicado en la conjugación de plásmidos de DNA, por lo que tiene un rol primordial en la transferencia de genes de resistencia a antibióticos (Costa *et al.*, 2015).

El T4SS más profundamente caracterizado es el codificado por el plásmido Ti de *Agrobacterium tumefaciens*, que consta de 11 proteínas, VirB1-VirB11 y la proteína VirD4. Estructuralmente, VirB3, VirB6, VirB7, VirB8, VirB9 y VirB10 conforman el andamiaje del sistema así como el aparato de translocación, mientras que VirB2 y VirB5 forman el pilus extracelular. VirB1, que es una transglicosilasa lítica, se encarga de degradar la capa de peptidoglicano y además es requerido para el ensamble del pilus.

Por último, las proteínas VirB4, VirB11 y VirD4 son ATPasas cuya función es dar energía al sistema (Darbari & Waksman, 2015) (Figura 17).

Figura 17. Representación esquemática de los componentes y de la estructura del sistema de secreción tipo IV VirB/D4 de *Agrobacterium tumefaciens*, así como la representación de los operones *virB* y *virD* en los cuales se codifican sus componentes (Christie *et al.*, 2014).

Además del T4SS representado por el sistema VirB/D4, existen otros T4SS. Todos ellos se clasifican en tres grandes grupos. El primero es el T4SS IVa y contiene al sistema descrito previamente. El segundo es el IVb, se encuentra representado por el complejo DoT/Icm de *Legionella pneumophila* y un tercer grupo designado como “otros tipos de T4SS”, incluye a aquellos que se encuentran codificados en islas genómicas, por ejemplo el T4SS *tfc* que se encuentra en la isla genómica ICE *Hin1056* de *Haemophilus influenzae*, dedicado principalmente a la transferencia horizontal de material genético (Nagai & Kubori, 2011; Darbari & Waksman, 2015).

La translocación de efectores por parte del T4SS ha sido estudiada principalmente en bacterias Gram negativas y consiste en el transporte de factores de virulencia hacia el citosol de la célula eucariota del huésped. A través de estas proteínas efectoras, el patógeno es capaz de inducir cambios fisiológicos en la célula del huésped que permiten

el establecimiento de la relación, ya sea patogénica o simbiótica entre ambos organismos. En cuanto a bacterias asociadas a plantas que contienen el sistema VirB/D4, se han identificado una serie de efectores del T4SS asociados con la virulencia y el desarrollo de la sintomatología, como el caso de las proteínas VirE2, VirE3, VirF y VirD5 de *A. tumefaciens* directamente relacionadas con la formación de agallas. Para la especie *Mesorhizobium meliloti*, se ha demostrado que los efectores Msi059 y Msi061, una vez secretados en el citoplasma del hospedero, atraen determinadas proteínas con las que conforman un complejo de ubiquitinación y proteólisis (Christie *et al.*, 2005).

La gran cantidad de proyectos de secuenciación en las diversas especies del género *Xanthomonas*, han permitido identificar loci homologos al sistema VirB/D4 en varias especies del género, tales como, *X. albilineans*, *X. campestris*, *X. citri*, *X. fuscans* y *X. oryzae*. Sin embargo, algunas especies como *X. citri* subsp. *citri* presentan un megaplásmido que codifica para un segundo T4SS que probablemente esté relacionado con la movilización del plásmido. Asimismo, algunas especies, como en *X. campestris* pv. *vesicatoria*, se encuentran homólogos para el sistema Dot/Icm, propio de *Legionella* (Potnis *et al.*, 2011; Souza *et al.*, 2011).

Una característica particular del T4SS virB/D4 de *Xanthomonas*, que corresponde a la proteína VirB7, es el no presentar similitud con las proteínas de la familia VirB7 halladas en otros grupos bacterianos, sin embargo, esta proteína presenta la misma función, lo cual permite que forme complejos con la proteína VirB9 (Souza *et al.*, 2011). Se ha demostrado que la unión de estas dos proteínas entre si, así como su interacción con VirB10, en *X. citri* subsp. *citri*, es esencial para el ensamblaje y la funcionalidad de este T4SS (Oliveira *et al.*, 2016). Además, estudios recientes (Souza *et al.*, 2015), demuestran que en *X. citri* subsp. *citri*, el homólogo a VirD4, que tiene como función el reclutamiento de los efectores que serán translocados, interactúa con una serie de toxinas que son utilizadas por *Xanthomonas* para matar a otras células bacterianas de su entorno, y así poseer una ventaja competitiva en presencia de otras comunidades bacterianas mixtas.

Estudios relacionados con la inducción de la expresión de 18 ORFs asociados a los dos T4SS que existen en *X. citri* subsp. *citri* indican que estos se ven reprimidos cuando la bacteria infecta las hojas de los cítricos, por lo que al parecer puede que estos sistemas de secreción no se vean implicados en la patogénesis sobre el huésped (Jacob *et al.*,

2014). Por otro lado se observó que en la especie *X. campestris* pv. *campestris*, un mutante carente de T4SS presentaba una respuesta hipersensible más débil que el fenotipo silvestre en una planta no huésped (Guo-feng *et al.*, 2015). Además, la expresión del T4SS en el fenotipo silvestre se inducía en condiciones de resistencia a níquel, peróxido de hidrógeno y fenol, por lo que podría ser que en esta especie el T4SS se encuentre relacionado con las etapas de contacto y reconocimiento propias del inicio de la infección.

1.4.4 Sistemas bacteriano de secreción tipo V.

En cuanto a los sistemas de secreción tipo V y tipo VI, estos han sido menos estudiados en bacterias fitopatógenas en general. El sistema de secreción tipo V (T5SS) o también denominado sistema de autotransporte, tiene como característica única, que las proteínas que se secretan contienen 3 o 4 dominios; el dominio transportador en el C-terminal que forma el poro de secreción en la membrana externa de la bacteria, el dominio de unión, el dominio pasajero, que contiene el dominio funcional de la proteína autotransportada y en algunos casos se presenta un cuarto dominio proteasa el cual escinde la proteína una vez haya atravesado el poro de la membrana externa. Las proteínas secretadas por el T5SS atraviesan la membrana interna de la célula mediante el sistema Sec, por lo que presentan un péptido señal en el N-terminal que se escinde una vez es transportado por este sistema (Green & Mecsas, 2016) (Figura 18).

Este sistema de secreción se subdivide en varios tipos, T5SSVa-T5SSVe, e independientemente del subtipo, todos son capaces de translocar grandes proteínas las cuales adoptan una estructura de tipo barril β . Las funciones de las proteínas secretadas por el T5SS incluyen serin proteasas, lipasas, citotoxinas y adhesinas, las cuales colectivamente influyen en procesos de agregación y formación de biofilm y en la virulencia (Costa *et al.*, 2015).

En cuanto a la presencia de este sistema secreción en bacterias fitopatógenas, se han identificado muchos sustratos candidatos de ser secretados por el T5SS, sin embargo, no se ha comprobado de manera funcional que estos tengan relación con la virulencia de los patógenos. En *Dickeya dadantii*, se ha informado sobre una adhesina con homología de secuencia a las secretadas por el T5SS Vb que es necesaria para la adhesión a hojas, mientras que un ortólogo a esta adhesina descritos en *Xanthomonas* y *Xylella*,

denominado XatA, se ha demostrado que es necesario para la agregación celular (Chang *et al.*, 2014).

Figura 18. Representación gráfica del sistema de secreción tipo V o sistema de autotransporte. Este sistema consiste en una proteína con varios dominios dentro de los cuales destacan el dominio transportador y el dominio pasajero (Chang *et al.*, 2014).

Propiamente, en el caso de las *Xanthomonas*, gracias a la comparación de secuencias genómicas, se han encontrado diversas proteínas que teóricamente podrían ser sustratos del T5SS. Asimismo en el genoma de diversas especies como *X. campestris* pv. *vesicatoria*, *X. euvesicatoria*, *X. gardneri*, *X. oryzae* pv. *oryzae* y *X. perforans*, se han encontrado homólogos a adhesinas que pueden tener un rol importante en el inicio de la infección, tal es el caso de la adhesina FhaB, la cual se ha visto como necesaria para la adhesión al tejido foliar así como para la supervivencia en apoplasto y la formación de biopelículas. Además, en *X. oryzae* pv. *oryzae*, las adhesinas XadA y XadB se ha demostrado que promueven la virulencia al potenciar la colonización de la hoja y la entrada al apoplasto a través de los hidatodos. En esta misma especie, la proteína YpH se ha demostrado involucrada en la virulencia en etapas en las que el patógeno migra a través de los vasos del xilema. Por último, Al lado de *fhaB* en el genoma de

Xanthomonas, se encuentra el operón *hms* que es homólogo al operón *pga* de *E. coli*, por lo que podrían ser también sustratos del T5SS (Potnis *et al.*, 2011).

1.4.5 Sistema bacteriano de secreción tipo VI.

En cuanto al último sistema de secreción descrito en *Xanthomonas*, el sistema de secreción tipo VI (T6SS), es una estructura dedicada a la translocación de efectores tóxicos dentro de células procariotas y eucariotas, por lo que podría tener un papel importante tanto en la virulencia como en la competencia entre poblaciones bacterianas. El T6SS, se encuentra distribuido principalmente en Proteobacteria y se compone de 13 componentes conservados esenciales y un número variable de componentes accesorios, que se encuentran codificados por lo general en islas de patogenicidad. La transcripción de estos genes se ha visto que es activada cuando el patógeno entra en contacto con el huésped (Costa *et al.*, 2015).

Los componentes esenciales del T6SS, se encuentran en el “cluster” génico *tssA-tssM*. TssJ, TssL y TssM conforman el complejo de la membrana. La proteína TssK une a este complejo con el complejo extracelular del T6SS, el cual se forma a partir de una plataforma proteica constituida por las proteínas VgrG y TssE. Posteriormente, se encuentra la estructura extracelular, conformada por unidades de la proteína Hcp, que es homóloga a la cola tubular de los bacteriófagos. TssB y TssC, forman heterodímeros que rodean a la estructura tubular formando una envoltura. El mecanismo de secreción, se cree que ha de ser similar al mecanismo de contracción presentado por las colas de los fagos. La proteína ClpV que es una ATPasa podría ser la encargada de desensamblar la envoltura formada por TssB-TssC generando así la contracción del aparato (Costa *et al.*, 2015; Green & Mecsas, 2016) (Figura 19).

Los efectores del T6SS presentan muchas formas y funciones, dentro de las cuales destacan su actuación en contra de la pared celular y la membrana del huésped eucariota o de otras bacterias. Muchos efectores del T6SS se encuentran codificados junto a un gen que provee de inmunidad contra el efector, por lo que de esta manera se previene la auto-intoxicación de la célula bacteriana (Green & Mecsas, 2016).

Figura 19. Representación esquemática del sistema de secreción tipo VI en bacterias Gram negativas (Costa *et al.*, 2015).

En el caso del género *Xanthomonas*, la evidencia experimental del rol que desempeña este sistema de secreción en la virulencia de las especies es escasa. En *X. populi* y *X. codiaii*, aquellas cepas que no contienen homólogos para los componentes esenciales del T6SS presentan un fenotipo avirulento, por lo que se especula que podría estar relacionado con la patogénesis en estos organismos (Shrivastava & Mande, 2008). Por otro lado, mediante un estudio de genómica comparativa se ha determinado que *X. axonopodis*, *X. campestris* pv. *vesicatoria* y *X. perforans* contienen dos T6SS de los tipos A y C, mientras que *X. euvesicatoria* solamente presenta uno del tipo C. *X. campestris* pv. *campestris* y *X. gardneri* no presentan ninguna copia de este sistema de secreción (Potnis *et al.*, 2011). Finalmente, los dos patovares de *X. oryzae*, presentan tres copias del T6SS, dos de ellas semejantes a los tipos A y B presentes en otras *Xanthomonas* y el tercero es más semejante al T6SS presente en *R. solanacearum*.

1.5 *Xanthomonas arboricola* como agente causal de enfermedad en diversos grupos de plantas de interés económico.

1.5.1 Descripción de la especie *X. arboricola*.

X. arboricola se diferencia metabólicamente de las otras especies de *Xanthomonas* por su capacidad para metabolizar los compuestos de carbono, dextrina, glicógeno, Tween 80, N-acetil-D-glucosamina, celobiosa, L-fucosa, D-galactosa, gentibiosa, maltosa, D-psicosa, D-trealosa, monometilsuccinato, ácido acético, DL-ácido láctico, ácido succínico, ácido bromosuccínico, ácido succinámico, L-ácido glutámico, ácido-glicil-L-glutámico, y glicerol. Además, por la incapacidad de metabolizar los sustratos de carbono N-acetil-D-galactosamina, L-arabinosa, D-manitol, β -metil-D-glucosida, D-rafinosa, L-ramnosa, ácido fórmico, ácido-D-glacturónico, ácido-D-glucónico, ácido-D-glucurónico, ácido-p-hidroxifenil acético, ácido α -ketovalérico, ácido quínico, ácido-D-sacárido, glucuronamida, L-fenilalanina, timidina y glucosa 6 fosfato. El contenido de G+C de esta especie está entre 66 y 67% y la cepa tipo de la especie es la cepa LMG747 (=ATCC 49083=CFBP 2528=ICMP 35=NCPPB 411) de *X. arboricola* pv. *juglandis* (Vauterin *et al.*, 1995).

X. arboricola es reconocida como una especie fitopatógena, dentro de la cual se describen diversos grupos sub-infraespecíficos o patovares con una determinada gama de huéspedes. Estos grupos son, *X. arboricola* pv. *arracaciae*, *X. arboricola* pv. *celebensis*, *X. arboricola* pv. *corylina*, *X. arboricola* pv. *fragariae*, *X. arboricola* pv. *guizotiae*, *X. arboricola* pv. *juglandis*, *X. arboricola* pv. *populi*, *X. arboricola* pv. *pruni* y *X. arboricola* pv. *zantesdeschiae*. Además, sin ser incluidas en ningún patovar, se han descrito cepas patogénicas en *Zizyphus jujuba*, *Capsicum annum* y *Vitis vinifera* (Fischer-Le Saux *et al.*, 2015).

1.5.2 Diversidad genética de *X. arboricola*.

La estructura genética de *X. arboricola* y sus relaciones filogenéticas a nivel intraespecífico han sido objeto de discusión durante mucho tiempo. En los últimos años, se ha podido profundizar en este tema por la utilización de herramientas moleculares con mayor capacidad de resolución filogenética. Ha sido el caso de los estudios basados

en el análisis de secuencias multilocus (MLSA), así como el análisis del número de secuencias repetidas en tándem en múltiples loci (MLVA) (Cesbron *et al.*, 2014; Essakhi *et al.*, 2015; Fischer-Le Saux *et al.*, 2015).

El análisis MLSA basado en seis genes constitutivos, *atpD*, *dnaK*, *efp*, *glnA*, *gyrB* y *rpoD* y un gen que codifica para una proteína transmembrana, *fyuA*, permite formar grupos filogenéticos claramente separados para cada uno de los patovares descritos. Los patovares *corylina*, *juglandis* y *pruni*, que representan los grupos más virulentos y económicamente más relevantes, son los que están genéticamente más relacionados y cada uno de ellos conforma un único complejo clonal. Esta cercana relación corrobora lo que previamente había sido observado según análisis de rep-PCR y de hibridación DNA-DNA. Esta característica junto con otros caracteres como el origen geográfico de estos patógenos, todos descritos en el siglo XX en Estados Unidos, así como un repertorio semejante de efectores tipo III, podrían ser indicativos de un mismo origen común para estos tres patovares provienen (Fischer-Le Saux *et al.*, 2015).

Además, se apunta que el patovar *populi*, agente causal de bacteriosis en chopo, es el más distante genéticamente de todos los patovares descritos en *X. arboricola* (Fischer-Le Saux *et al.*, 2015). También se ha observado una gran variación entre las cepas de los patovares *celebensis*, agente causal de bacteriosis en banana y *fragariae*, asociado a bacteriosis en fresa. Ambos grupos presentan diversos clusters dentro de cada patovar. La diversidad genética de este grupo además muestra que ha existido recombinación interespecífica en el gen *atpD* del patovar *guizotiae* y el gen *rpoD* del patovar *populi*, pero no ha sido posible determinar exactamente cuál ha sido la especie de *Xanthomonas* con la que ha habido recombinación. Por último, una serie de cepas aisladas son ubicadas dentro de los grupos a los que se suponía debían pertenecer según su huésped original. Éstas corresponden a cepas avirulentas o poco caracterizadas que se clasifican como *X. arboricola* pero no presentan afiliación a ningún patovar de la especie.

En cuanto a los estudios de diversidad genética basados en MLVA, utilizando un grupo de 17 repeticiones variables en tándem (VNTRs) (Cesbron *et al.*, 2014), se ha podido observar la estructura poblacional de la especie, contemplando cepas patogénicas descritas para cada uno de los patovares y diversas cepas no patógenas aisladas por ejemplo en nogal. En general, los resultados obtenidos mediante esta técnica de genotipado, han corroborado la agrupación y diferenciación de los diversos patovares,

tal como se ha descrito con los estudios llevados a cabo mediante MLSA. Además, permiten determinar que las cepas avirulentas del nogal se agrupan fuera de estos “clusters” formando tres grupos basales en la especie que corresponden a linajes claramente separados (Essakhi *et al.*, 2015).

1.5.3 Factores de virulencia de *X. arboricola*.

Referente a los estudios relacionados con la caracterización de factores de virulencia para la especie *X. arboricola*, aún son escasos y han consistido básicamente en determinar mediante PCR la presencia del T3SS, así como de los efectores del T3SS presentes en siete de los patovares de la especie (Hajri *et al.*, 2012). Recientemente, se han realizado estudios de genómica comparativa que han permitido obtener información sobre otros factores de virulencia que podrían estar involucrados en la patogenicidad de *X. arboricola* pv. *juglandis*, y que serán discutidos más adelante en el capítulo I de este trabajo.

Los principales componentes asociados tanto a la estructura como a la regulación del T3SS de *Xanthomonas* fueron analizados mediante PCR para las cepas patógenas de los patovares *celebensis*, *corylina*, *fragariae*, *juglandis*, *poisentiicola* y *pruni*. Asimismo, el repertorio de efectores tipo III, mostró ser mayor en los patovares más virulentos de la especie (*corylina*, *juglandis* y *pruni*) presentando entre 18 y 22 efectores. Por otro lado, los patovares *celebensis*, *poisentiicola* y *fragariae* mostraron el repertorio más pequeño de *X. arboricola* con tan solo 6 efectores. Todas las cepas analizadas en este estudio, independientemente de su patovar, presentaron un grupo común de efectores codificados por los genes *avrBs2*, *xopF1*, *xopA*, *hrpW*, *hpaA* y *xopR*. Los patovares más virulentos, compartieron además los efectores codificados por *xopN*, *xopX*, *xopZ*, *xopQ*, *xopK* y *xopV*. Los efectores *xopG*, *xopAF* y *xopE2* fueron compartidos por los patovares *corylina* y *pruni*. Por último, los patovares más virulentos presentaron algunos efectores únicos para cada uno de ellos, este es el caso de *xopE3* en el patovar *pruni*, *xopB* en el patovar *juglandis* y *avrBs3* en el patovar *corylina*.

Se ha descrito que la variación observada en el repertorio de efectores del T3SS entre las cepas de un mismo patovar es muy baja (Hajri *et al.*, 2012). Sin embargo estas diferencias se han encontrado por ejemplo, en el patovar *corylina* en el cual las cepas aisladas de *Coryllus avellana*, contienen el efector *xopH* mientras que las cepas aisladas

de *C. maxima* no. En el patovar *juglandis*, estas diferencias fueron observadas entre los miembros de los dos linajes que presenta este grupo. Así, el efector *xopB* solamente se encontró en aquellas cepas que producen el chancro del nogal (denominado linaje VOC), mientras que el efector *xopAH* se encontraba únicamente en el linaje asociado a la bacteriosis del nogal (denominado WB) (Hajri *et al.*, 2012).

Como ya ha sido mencionado, dentro de la especie *X. arboricola*, los patovares *corylina*, *juglandis* y *pruni*, presentan una mayor virulencia y esto está asociada a su vez con un mayor impacto económico. En Europa estos organismos son clasificados como organismos de cuarentena (EU directiva 2000/29/CE y EPPO lista A2) (Lamichhane, 2014; Lamichhane & Varvaro, 2014) Durante la última década, los focos de enfermedad causados por estos patovares han aumentado a nivel mundial y por ello se ha potenciado su estudio con el objetivo de contener avance en las diversas regiones productoras y evitar así futuras epidemias.

Para evitar o solucionar posibles cuadros epidémicos es necesario recabar la mayor información posible referente al patógeno y los factores vinculados con su patogénesis, la interacción de éste con su huésped y las condiciones ambientales asociadas a la epidemiología de las enfermedades causadas por cada uno de estos grupos intraespecíficos. A continuación, se presentará una breve descripción de las enfermedades provocadas por estos tres patovares, dando un mayor énfasis al patovar *pruni* y su relación con los huéspedes del género *Prunus*.

1.5.4 *Xanthomonas arboricola* pv. *corylina*: agente causal de la podredumbre bacteriana del avellano.

X. arboricola pv. *corylina*, agente causal de la podredumbre bacteriana del avellano, fue descrita por primera vez en Estados Unidos a inicios del siglo XX. Actualmente esta enfermedad se encuentra distribuida a nivel mundial y en Europa se encuentra clasificada como organismo de cuarentena según la lista A2 de la EPPO. El principal huésped de esta bacteria es el avellano (*C. avellana*), sin embargo, también puede encontrarse en otros huéspedes de este género como son *C. colurna*, *C. maxima* y *C. pontica* (Scortichini, 2004) (Figura 20).

Figura 20. Mapa de distribución de *X. arboricola* pv. *corylina* a nivel mundial según la “European and Mediterranean Plant Protection Organization” (tomado de EPPO actualizado a fecha 13 de octubre de 2015).

La sintomatología de esta enfermedad es posible observarla en la mayor parte de órganos y tejidos de la parte aérea del árbol. Por lo general incluye afecciones en brotes, hojas, ramas, tronco y frutos. En las hojas se producen manchas necróticas poligonales y acuosas que empiezan siendo amarillentas y opacas pero que con el tiempo se tornan rojizas o marrones y se presentan principalmente en la zona de goteo. En los tallos y en las ramas suelen aparecer manchas acuosas de color verde oscuro que con el tiempo se tornan marrones. Por lo general se acaba produciendo la muerte de las hojas y de las ramillas, causando graves pérdidas económicas, puesto que es en éstas donde se desarrolla el fruto. A nivel del tronco, suelen formarse chancros de color marrón que forman motas a lo largo del tronco aunque a veces son difíciles de ver, a menos que se retire la corteza. Por último, los frutos suelen presentar manchas necróticas acuosas pero son superficiales, por lo que no causan mayor daño a la avellana (Scortichini, 2004; Lamichhane & Varvaro, 2014) (Figura 21).

Los principales productores de este fruto seco en el mundo son Turquía, Italia, España y Estados Unidos. En este último país, este patógeno es considerado el principal problema bacteriano para el avellano. La mayor parte de las pérdidas se dan en plantaciones con árboles jóvenes de entre uno y cuatro años, edad en la cual pueden provocar hasta un 10% de muerte del huésped. Los árboles por encima de esta edad raramente mueren pero suelen tener pérdidas de hasta el 10% de los frutos debido a la muerte de las

ramillas. Se han registrado pérdidas en Francia de hasta 250 000 plantas jóvenes y en diversas plantaciones se ha registrado la muerte de hasta un 100% de los árboles menores de cuatro años junto con una proporción muy baja de árboles mayores a esta edad (Lamichhane & Varvaro, 2014).

Para el control químico de esta enfermedad se utilizan principalmente compuestos basados en cobre, aunque resultan ser poco efectivos una vez que la bacteria penetra el tejido de la planta pero pueden ser usados como métodos preventivos. Es necesario por tanto el uso de otras técnicas tales como la poda de zonas infectadas y la eliminación del material vegetal, así como la correspondiente desinfección de las herramientas de trabajo y el material de poda. Además, se aconseja sellar las heridas provocadas con materiales antibacterianos basados en cobre. El método más efectivo que se ha usado para el control de la enfermedad hasta el momento es la prevención, basada en el uso de material vegetal libre del patógeno, así como una política de vigilancia y detección temprana (Lamichhane & Varvaro, 2014).

Figura 21. Sintomatología propia de la mancha bacteriana del avellano, causada por *X. arboricola* pv. *corylina* (Lamichhane & Varvaro, 2014).

1.5.5 *Xanthomonas arboricola* pv. *juglandis*: Agente causal de la podredumbre bacteriana, la necrosis apical y la chancrosis del nogal.

X. arboricola pv. *juglandis*, es el agente causal de la mancha bacteriana, la necrosis apical y la chancrosis del nogal. Esta enfermedad, descrita en Estados Unidos a inicios del siglo XX, actualmente se encuentra distribuida a nivel mundial (Lamichhane, 2014)

(Figura 22). *X. arboricola* pv. *juglandis* no es considerado patógeno de cuarentena en la Unión Europea y su único huésped es la especie *Juglans regia* (Marcelletti *et al.*, 2010).

Estudios de variabilidad genética poblacional para este patógeno, inicialmente desarrollados con F-AFLPs (Hajri *et al.*, 2010) y corroborados posteriormente mediante MLSA y MLVA (Essakhi *et al.*, 2015; Fischer-Le Saux *et al.*, 2015), han demostrado la existencia de dos linajes, uno de ellos asociado a la podredumbre bacteriana y la necrosis apical, denominado WB y otro relacionado con la chancrosis del nogal, denominado VOC (Hajri *et al.*, 2010).

Figura 22. Mapa de distribución de *X. arboricola* pv. *juglandis* a nivel mundial según la “European and Mediterranean Plant Protection Organization” (tomado de EPPO actualizado a fecha 17 de marzo de 2015).

La sintomatología propia de esta enfermedad incluye la deformación de las hojas y la presencia de manchas necróticas en las mismas, chancros necróticos en las ramillas y el tronco que exudan durante el verano creando manchas en la madera. La severidad de la enfermedad puede ser muy alta viéndose reducida la producción de fruta debido a su caída o disminuida su calidad como consecuencia de la decoloración de la nuez. En etapas tardías, la necrosis apical, no se debe solamente a la actividad bacteriana sino también a la presencia de algunos hongos oportunistas de los géneros *Alternaria* y *Fusarium* (Scortichini, 2010) (Figura 23).

El inóculo primario de este patógeno lo conforman las infecciones latentes en brotes, las cicatrices de las hojas y los chancros. Una vez se inicia una nueva etapa vegetativa, las

bacterias pasan del estado saprófito al estado parasítico y penetran en el interior del tejido vegetal, ya sea por los estomas o por heridas. Las infecciones suelen ser más severas en la primavera, cuando hay mayor cantidad de precipitaciones (Lamichhane, 2014). Se ha visto predisposición a infecciones graves en suelos arenosos ácidos y con deficiencias en fósforo, calcio y magnesio (Scortichini, 2010).

El control de esta enfermedad se basa en el uso de diversas estrategias, empezando por el uso de material vegetal libre del patógeno. Una vez establecido el agente causal, se recomienda realizar aplicaciones de compuestos cúpricos, la eliminación de las zonas infectadas del árbol y la optimización de la fertilización y el riego. Se recomienda además eliminar las malas hierbas presentes en las plantaciones, que podrían servir como reservorios de la bacteria. En relación al uso de control biológico, se han documentado cepas bacterianas del género *Pseudomonas* que son capaces de reducir la severidad de la enfermedad entre un 41 y un 77% en árboles jóvenes (Ozaktan *et al.*, 2012; Lamichhane, 2014).

Figura 23. Sintomatología propia de la mancha bacteriana y la necrosis apical del nogal (panel superior) y del chancro del nogal (panel inferior), causado por *X. arboricola* pv. *juglandis* (Ozaktan *et al.*, 2012; Lamichhane, 2014).

1.5.6 *Xanthomonas arboricola* pv. *pruni*: agente causal de la mancha bacteriana en frutales del género *Prunus*.

X. arboricola pv. *pruni* (Vauterin *et al.*, 1995) (sinónimo, *X. campestris* pv. *pruni* Smith; *X. campestris* Dowson), es el agente causal de la mancha bacteriana de los frutales de hueso y el almendro, enfermedad descrita en Estados Unidos en 1903 afectando ciruelo japonés (Boudon *et al.*, 2005). A continuación, se tratarán diversos aspectos sobre este organismo.

1.5.7 Diversidad genética y estructura poblacional:

A pesar de que la enfermedad de la mancha bacteriana se encuentra presente en Italia desde 1920 y en Rumanía desde 1941 (Boudon *et al.*, 2005), no es hasta finales del siglo XX cuando se inician los estudios dirigidos a la caracterización molecular de sus poblaciones y su diversidad genética en Europa.

En un primer trabajo, se llevó a cabo el análisis de la diversidad poblacional en una colección de cepas italianas, incluyendo también algunas provenientes de otros continentes (Zaccardelli *et al.*, 1999). El perfil de AFLPs mostrado por las 109 cepas analizadas, permitió distinguir 12 diferentes grupos que presentaban una similitud muy alta del 90-99%.

Posterior a este estudio, Boudon y colaboradores (2005), determinaron mediante análisis de la secuencia del ITS del operón *rrn* y mediante MLSA utilizando los genes *atpD*, *dnaK*, *efp* y *glnA* que la secuencia de nucleótidos era idéntica para todas las cepas de *X. arboricola* pv. *pruni* independientemente del lugar y la fecha de aislamiento, lo cual evidenciaba una muy alta similitud dentro de este grupo filogenético. Por otro lado utilizando FAFLPs, se determinó la existencia de 14 genotipos, pero la mayoría de cepas, representantes de cinco países y cuatro continentes pertenecían a un único grupo, lo cual corroboraba la baja diversidad genética media observada en el patovar ($0,29 \pm 0,11\%$). Se encontró además, que la diversidad genotípica era mayor en las cepas europeas ($H= 0,8$) que en aquellas cepas aisladas de Estados Unidos ($H= 0,48$).

Además, estudios basados en el perfil de integrones y casetes génicos y BOX-PCR, llevados a cabo para este patógeno (Barionovi & Scortichini, 2006), han permitido profundizar más en los estudios sobre la variabilidad de las cepas. Se encontraron dos patrones de amplificación de los casetes, que sin embargo no presentaban ninguna relación a otra característica asociada a las cepas, como pueden ser su origen geográfico o su huésped original. De igual manera, el genotipado basado en la amplificación de secuencias repetidas de los elementos BOX mostró poca variabilidad entre las cepas, además ésta no concordaba con los resultados obtenidos mediante la técnica anterior. Estos resultados podrían indicar que a pesar de la uniformidad en los perfiles de casete génico, existen diferentes linajes dentro de este patovar.

Más recientemente, en nuevo intento de observar la composición poblacional basado en un esquema de MLSA utilizando los genes *atpD*, *dnaK*, *efp*, *fyuA*, *glnA*, *gyrB* y *rpoD* (Fischer-Le Saux *et al.*, 2015), determinó que todas las cepas del patovar *pruni* que se analizaron, formaban un grupo monofilético con un bajo grado de polimorfismo, corroborando que este grupo conforma un grupo casi clonal.

Dado que todos estos métodos han presentado poca resolución para poder analizar la variación genética existente dentro del patovar, en especial para determinar diferencias entre cepas con una reciente separación espacio-temporal, se ha procedido a realizar el genotipado de este patovar mediante el uso del MLVA, al tratarse éste de un método eficiente para analizar la diversidad de patógenos bacterianos monomórficos (López-Soriano *et al.*, 2016). Se ha desarrollado un esquema de MLVA basado en el genoma de *X. arboricola* pv. *pruni* con el fin de diferenciar los diferentes patovares de la especie, utilizando un total de 26 loci VNTR polimórficos que se encuentran presentes en todas las cepas de la especie (Essakhi *et al.*, 2015).

Utilizando un esquema basado en seis loci VNTR polimórficos, se estudiaron 25 cepas del patovar *pruni* aisladas de *P. laurocerasus* en Holanda. La diversidad genética observada mediante esta técnica fue baja puesto que solo se lograron identificar dos diferentes haplotipos y tal como se ha observado con otras técnicas de genotipado usadas en este patovar, no fue posible entablar relación alguna entre el grupo de MLVA y su origen geográfico (Bergsma-Vlami *et al.*, 2012).

Recientemente, haciendo uso de esta misma técnica, se ha estudiado la diversidad genética de las cepas de *X. arboricola* pv. *pruni* aisladas en España (López-Soriano *et*

al., 2016). Un total de 239 cepas españolas y 25 cepas de referencia procedentes de diversas regiones geográficas fueron analizadas utilizando un esquema de MLVA basado en 23 loci polimórficos. El esquema incluía la combinación de dos marcadores moleculares, los mini y micro satélites. Se logró identificar un total de 119 haplotipos de MLVA, los cuales se agruparon en 18 “cluster” génicos, cada uno agrupando a aquellos haplotipos que se diferenciaban en menos de 4 loci de secuencias repetidas. Estos haplotipos, también se agruparon en 16 complejos clonales y 34 singletons. Al comparar el perfil de MLVA de las 239 cepas de España con las 25 cepas procedentes de diferentes regiones del mundo, se observó una clara diferencia entre ambos grupos, donde las cepas españolas presentaron variaciones entre 4 y 15 loci. El hecho de que al menos el 50% de las cepas analizadas presentaran diferentes haplotipos, demostró que el MLVA es una técnica apropiada para analizar molecularmente la variación entre cepas que han tenido una historia evolutiva muy próxima. Finalmente, se determinó que la introducción de este patovar en España no es consecuencia de un único evento en el tiempo sino que ha sido introducido en varias ocasiones y por al menos cuatro “cluster” génicos diferentes.

1.5.8 Interacción huésped-patógeno-ambiente en el contexto de la mancha bacteriana de los frutales de hueso y el almendro:

De manera clásica, la interacción entre los componentes que permiten el desarrollo de una enfermedad es representada mediante el llamado triángulo de la enfermedad, en el cual los vértices representan a la planta, el patógeno y el ambiente. En este apartado, se tratará cada uno de los componentes que intervienen en el triángulo de la mancha bacteriana de los frutales del género *Prunus* así como sus interacciones.

1.5.8.1 El patógeno

En cuanto al patógeno, la información existente sobre los diversos factores asociados a la patogénesis es muy reducida. Con respecto a los mecanismos iniciales de la infección, solamente se han estudiado aspectos asociados al lipopolisacárido y a la goma xanthano, mientras que en relación a etapas más tardías de la infección, se ha determinado mediante PCR el repertorio de efectores tipo III de este patovar.

Con respecto a la estructura del polisacárido específico O que forma parte del lipopolisacárido, ésta ha sido estudiada en la cepa NCPPB 416 de este patovar (Molinaro *et al.*, 2003), observándose que se encuentra compuesta por tres monosacáridos, glucosa, ramnosa y xilosa. El espectro observado por resonancia magnética nuclear ha permitido dilucidar una estructura muy irregular en la cual se da la repetición de cinco unidades distintas.

La producción de la goma xanthano en esta especie ha sido estudiada en al menos 48 cepas que presentan una producción variable. Todos los polímeros producidos presentan un comportamiento pseudoplástico con una viscosidad que aumenta de manera proporcional a la concentración de manosa. La mayor viscosidad de este polisacárido se observó después de 24 horas de fermentación (Moreira *et al.*, 2001). Además, en la composición química del xanthano en esta bacteria intervienen los compuestos glucosa, ramnosa, manosa y ácido glucurónico (Borges & Vendruscolo, 2008).

Por último en cuanto a los factores de virulencia estudiados en *X. arboricola* pv. *pruni*, las cepas virulentas de esta especie presentan un T3SS del tipo *hrc/hrp* típico de *Xanthomonas*, el cual está compuesto por los genes *hrcC*, *hrcJ*, *hrcN*, *hrcR*, *hrcS*, *hrcT*, *hrcU*, *hrcV*, *hrpB1*, *hrpD5* y *hrpF*. Mientras que el perfil de efectores tipo III para esta bacteria fitopatógena está compuesto por al menos 21 efectores identificados todos ellos mediante PCR. Los efectores presentes en el patovar *pruni* son: *avrBs2*, *avrXccA1*, *avrXccA2*, *hpaA*, *hrpW*, *xopA*, *xopAF*, *xopAH*, *xopAI*, *xopE2*, *xopE3*, *xopF1*, *xopG*, *xopK*, *xopL*, *xopN*, *xopQ*, *xopR*, *xopV*, *xopX* y *xopZ* (Hajri *et al.*, 2012).

1.5.8.2 El huésped.

Con respecto al segundo vértice del triángulo de la enfermedad de esta bacteriosis, los huéspedes naturales para este patógeno corresponden a frutales, ornamentales y plantas silvestres del género *Prunus*. Dentro de las especies cultivadas, aquellas en las que esta bacteria produce mayores daños son, el albaricoquero (*P. armeniaca*), el almendro (*P. amygdalus*), el ciruelo (*P. salicina* y *P. japonica*) y sus correspondientes híbridos, así como el melocotonero (*P. persica*) y sus correspondientes híbridos. Otros huéspedes en los cuales se ha identificado la presencia de este patógeno son, el endrino (*P. domestica*), el cerezo (*P. avium*), el guindo (*P. cerasus*), el lauroceraso (*P. laurocerasus*), el melocotón silvestre chino (*P. davidiana*), *P. buergeriana*, *P. crassipes*

y *P. donarium*. Por lo general los patrones portainjertos utilizados son resistentes a la enfermedad (EPPO, 2003; Stefani, 2010; EFSA, 2014).

A pesar de que la gran mayoría de especies del género *Prunus* se encuentran en las regiones templadas del hemisferio norte, en el continente europeo la producción de albaricoque, almendro, ciruelo, melocotón y nectarinas se da principalmente en la región mediterránea. En este ámbito geográfico, España es el primer productor de almendro (561,5 Tm/1000 Ha) y albaricoquero (48 Tm/1000 Ha), el segundo productor de melocotón (51,9 Tm/ 1000 Ha) y nectarino (25,7 Tm/1000 Ha) y el tercer productor europeo de cerezo (24,1 Tm/1000 Ha) y de ciruelo (19,1 Tm/ 1000 Ha) (EFSA, 2014). Además, España, es el tercer productor mundial de melocotón y nectarino y el segundo productor mundial de almendro. Debido a lo anterior, *X. arboricola* pv. *pruni* representa una importante amenaza para la fruticultura y su impacto económico puede llegar a ser muy alto (Palacio-Bielsa *et al.*, 2014).

La resistencia por parte del huésped a *X. arboricola* pv. *pruni* ha sido estudiada en albaricoquero, ciruelo y melocotonero, tanto en condiciones de invernadero como en campo, encontrándose una variación en la susceptibilidad entre las distintas especies. Sin embargo, aún los huéspedes menos susceptibles presentan síntomas de la mancha bacteriana, no existiendo una resistencia completa. La existencia de un mapa físico del melocotonero y la secuenciación de su genoma, ha permitido generar mapas de diversos loci de carácter cuantitativo (QTL) relacionados con la resistencia a patógenos, los cuales gracias a la alta sintenia presente entre las diversas especies de *Prunus* pueden extrapolarse a otras especies (Socquet-Juglard *et al.*, 2013a).

Se ha estudiado además la resistencia a *Xanthomonas arboricola* pv. *pruni* en una población de ciruelo derivada del cruce entre las variedades “Harostar” x “Rouge de Mauves” con el fin desarrollar un análisis de QTLs para la resistencia a este patógeno. Como resultado, se encontró un QTL con un efecto mayoritario sobre la variación fenotípica observada. Éste se encontró en el grupo de ligamiento 5 del parental “Rouge de Mauves” que segrega junto con el microsatélite UDAp-452. Al dividir la población analizada según los alelos observados para este microsatélite, se observó que aquellos con el alelo favorable presentaron una reducción del 35,6% en la severidad de la enfermedad (Socquet-Juglard *et al.*, 2013a).

El uso de variedades susceptibles se encuentra entre los factores críticos para la aparición de brotes severos y la distribución a larga escala de la enfermedad. Los cultivares más susceptibles para el ciruelo son, Calita, Angeleno, Frontier, Santa Rosa, the “Black” group, Anna y T.C. Sun. Mientras que los cultivares más susceptibles para el melocotonero corresponden a Summer Rich, O’Henry, Rich Lady, Red Valley, Elegant Lady, Big Top, Flavorcrest, Flavortop y Cassiopea (Scortichini, 2010).

En relación a la interacción de los huéspedes del género *Prunus* con *X. arboricola* pv. *pruni*, los estudios realizados con microscopía electrónica (Plessis, 1984), inmunolocalización y la histocitopatología (Aarouf *et al.*, 2008), han permitido determinar cómo se da el ingreso y la colonización del tejido por parte de la bacteria tanto en ciruelo como en melocotonero.

En el caso del ciruelo (Plessis, 1984), se estableció que *X. arboricola* pv. *pruni* infecta las ramillas de manera sistémica a través de los peciolo de las hojas infectadas. El patógeno fue observado en los tejidos del xilema de los peciolo de las hojas inoculadas y además se encontró en el tejido del córtex de las ramillas en la zona adyacente a los peciolo infectados. De igual manera, se determinó que la bacteria es capaz de dirigirse de manera sistémica hasta el tejido meristemático de los brotes. Posteriormente, en melocotonero (Aarouf *et al.*, 2008), se demostró que el patógeno es capaz de ingresar al tejido del mesófilo a través de los estomas, colonizando el espacio intercelular, degradando los componentes de la pared celular, colapsando las células y dando lugar a lesiones necróticas. El análisis de secciones de la nervadura central tomadas a diversas distancias desde la lámina infectada, permitió observar la presencia de la bacteria en los tubos cribosos y en el xilema lo que podría sugerir el tráfico sistémico del patógeno.

La respuesta de la planta ante la infección del *X. arboricola* pv. *pruni*, ha sido estudiada en hojas de melocotonero. Inicialmente se procedió a analizar los factores de respuesta asociados a etileno (Sherif *et al.*, 2012), los cuales corresponden a factores de transcripción asociados al desarrollo y la inmunidad de las plantas. Se estudió el rol desempeñado por cinco de estos factores de transcripción durante el desarrollo de la mancha bacteriana. Todos éstos, correspondieron a activadores transcripcionales y dos de ellos (Pp-ERF2.b y Pp-ERF2.c) fueron localizados núcleo celular. La inducción cinética de estas proteínas, después de la inoculación con *X. arboricola* pv. *pruni*, estudiada mediante PCR cuantitativa, demostró que todos incrementaron su expresión

de manera rápida y muy marcada en aquellos cultivares de melocotonero más resistentes. Estos activadores transcripcionales también aumentaron su expresión al tratar las plantas con etileno o ácido jasmónico, lo que sugiere una actividad sinérgica de ambas rutas metabólicas durante el proceso de defensa contra *X. arboricola* pv. *pruni*.

Posteriormente se analizó el efecto del patógeno sobre el huésped mediante el examen del transcriptoma del melocotonero (Socquet-Juglard *et al.*, 2013b) a las 2 h y 12 h posteriores a la inoculación. De los 19.781 genes conocidos en el melocotonero, 34 y 263 fueron diferencialmente expresados en cada uno de los tiempos analizados respectivamente. Aquellos genes asociados a procesos metabólicos y respuesta a estrés fueron reprimidos a 2 h, mientras que a 12 h fueron los que presentaron el mayor grado de expresión. Dentro de éstos, han sido de particular interés aquellos asociados a receptores de PAMPs, como por ejemplo los de la familia RPM1. Otros genes relacionados con la fotosíntesis, la reorganización de la pared celular y las rutas metabólicas hormonales fueron también diferencialmente expresados. Por último, una serie de genes con función desconocida también sufren cambios en su expresión, por lo que son dianas interesantes para desarrollar futuros estudios sobre la defensa en *Prunus* a este patógeno bacteriano.

1.5.8.3 El ambiente.

Desde el punto del ambiente, el suelo parece jugar un papel importante en la severidad observada en esta enfermedad en los diferentes huéspedes. Se considera que suelos arenosos y arcillosos son factores que predisponen para la aparición de brotes severos de la mancha bacteriana (Scortichini, 2010). Los síntomas de esta enfermedad, no se han desarrollado cuando los árboles han sido plantados en suelos arenosos con una mezcla de limo y vermiculita (Zehr *et al.*, 1996). De la misma manera se ha visto un incremento considerable de la susceptibilidad a la enfermedad por parte del huésped cuando este se encuentra en suelos infestados con el nematodo *Crictonemella xenoplax* (Shepard *et al.*, 1999).

En cuanto a las condiciones ambientales, la influencia de la congestión hídrica, el período de mojadura foliar y la temperatura, han sido estudiadas en melocotonero cultivar Suwanee a 24 y 30 °C. Se determinó, que la congestión hídrica en la hoja

seguida por un periodo de mojadura foliar, con una humedad relativa cercana al 100%, por al menos 36 h y en suelos arenosos, conforman las condiciones óptimas para el desarrollo de síntomas. La aparición de síntomas a 24 °C fue más lenta que a 30 °C. Para que se desarrollaran síntomas severos a temperaturas de 24 °C, se requirió de periodos de mojadura foliar superiores a 18 h (Zehr *et al.*, 1996).

1.5.9 Distribución geográfica e impacto económico de *X. arboricola* pv. *pruni*:

X. arboricola pv. *pruni* se encuentra distribuida a nivel mundial en las principales zonas productoras de frutales del género *Prunus* (Figura 24). En Europa, los casos más severos de la mancha bacteriana se han encontrado en zonas de la región mediterránea, como en el centro y este de España, el sureste de Francia y el valle del río Po al norte de Italia (EFSA, 2014). Además, por países, se ha informado la presencia de este patógeno en Italia, donde es considerado endémico y ha sido catalogado como enfermedad emergente en Alemania, Bélgica, España, Francia, Holanda, Suiza y diversos países de Europa del Este (López-Soriano *et al.*, 2016). La dispersión de este patógeno desde una zona geográfica a otra se debe principalmente al movimiento de material vegetal (portainjertos o plantas injertadas) principalmente con infecciones latentes. A corta distancia el patógeno se disemina fácilmente gracias a las lluvias abundantes y las tormentas acompañadas de fuertes vientos (EFSA, 2014).

España es uno de los principales productores de fruta del género *Prunus*. Se dedican alrededor de 562.616 Ha a la producción de almendro, 76.730 a la producción de melocotón y nectarina, 24.304 Ha a la producción de cereza, 19.226 Ha a la producción de albaricoque y 18.489 Ha a la de ciruelo. *X. arboricola* pv. *pruni* se encontró por primera vez en el año 2002 en Badajoz en ciruelo japonés y posteriormente en el año 2004 en nectarina y ciruelo japonés en viveros de plantones en Valencia. En el año 2006 fue identificada en almendro en Alicante y en el año 2008 en melocotonero y ciruelo en Zaragoza y Huesca, así como sobre melocotonero y nectarina en Lérida. En el año 2009, fue encontrado en almendro en Navarra, Zaragoza y Huesca y finalmente en el año 2010 fue informada su presencia en Mallorca afectando ciruelo japonés. En todos los casos se implementaron programas de erradicación por lo que actualmente se sigue considerando una enfermedad emergente y de cuarentena (Roselló *et al.*, 2012; Palacio-Bielsa *et al.*, 2014).

Los síntomas asociados a este patógeno son la defoliación, la formación de chancros en el tronco y eventualmente se puede provocar la muerte del árbol, así como producirse pérdidas en la producción o la obtención de fruta no apta para la venta, todo ello conlleva cuantiosas pérdidas en los cultivos (EFSA, 2014). En Estados Unidos se informó en 1932 que las pérdidas en plantaciones de ciruelo japonés podían estar entre el 25 y 75% de la fruta producida (Stefani, 2010). En Uruguay, en años con condiciones ambientales favorables para la enfermedad, las pérdidas en variedades sensibles de melocotonero pueden llegar a superar el 50% y si se consideran las pérdidas generadas por la imposibilidad de comercializar el fruto sintomático, éstas pueden llegar a superar el 80% en variedades susceptibles de nectarina (Palacio-Bielsa *et al.*, 2015a).

Figura 24. Mapa de distribución de *X. arboricola* pv. *pruni* a nivel mundial según la “European and Mediterranean Plant Protection Organization” (tomado de EPPO actualizado a fecha 30 de setiembre de 2016)

En Europa, una estimación de lo que significaría económicamente la mancha bacteriana ha determinado que las pérdidas que se producirían serían de hasta 11.200 euros por hectárea en el caso de la variedad Golden o 9.500 euros en el caso de la variedad Angeleno (Stefani, 2010). En España, sólo se han estimado las pérdidas en almendro en Aragón, las cuales en el año 2013 oscilaron entre 23 y 47%, y dependían éstas principalmente de las condiciones ambientales y a el uso de cultivares sensibles (Palacio-Bielsa *et al.*, 2014, 2015a).

1.5.10 Sintomatología asociada a la mancha bacteriana de los frutales del género *Prunus*:

De manera general, tanto los tejidos herbáceos, como hojas, brotes, frutos, así como los tejidos lignificados, ramillas, ramas y el tronco; pueden ser afectados por este patógeno. Sin embargo, no se han observado daños en las flores. Los síntomas en nectarinos, melocotoneros y almendros suelen observarse principalmente en las hojas y los frutos, mientras que los chancros en las ramas y el tronco suelen observarse en ciruelo (Lamichhane, 2014).

Los síntomas foliares se inician como pequeñas manchas de color verde pálido que derivan hacia manchas necróticas poligonales que por lo general se encuentran delimitadas por las nervaduras de las hojas y pueden estar rodeadas por un halo clorótico (Palacio-Bielsa *et al.*, 2014). El mayor daño suele observarse en el punto de goteo de la hoja, tanto en melocotonero como en nectarino. Desde el punto de goteo hacia el peciolo de la hoja es común observar un gradiente de color marrón-amarillo-verde y en infecciones severas se puede producir defoliación. En el caso del almendro se pueden observar los mismos síntomas, excepto que no se produce clorosis de la hoja ni se da la defoliación (Figuras 25 y 26).

Los síntomas en fruto se describen como manchas pardas rodeadas de un halo clorótico. En etapas iniciales de la infección, estas manchas pueden extenderse y necrosarse, produciendo hundimiento del mesocarpio y la secreción de exudados conforme el fruto crece. En el almendro los síntomas suelen ser muy característicos y conspicuos, estos se inician como manchas oscuras y hundidas en el mesocarpio que presentan frecuentemente la producción de exudados. Al deshidratarse el mesocarpio, la zona de las heridas sobresale de la superficie del fruto y se quedan adheridas al endocarpio. Se suele producir la caída prematura del fruto o éste permanecer adherido al árbol tras la cosecha. Las infecciones severas pueden generar también problemas a nivel de la semilla, ya sea deshidratándola o provocando manchas de color pardo (Palacio-Bielsa *et al.*, 2014).

Figura 25. Sintomatología propia de la mancha bacteriana causada por *X. arboricola* pv. *pruni* tanto en hojas (A y B) y frutos (C, D y E) de melocotonero. Hojas (F) y fruto (G) de albaricoquero así como la presencia de chancros en el tronco de ciruelo (H) (Lamichhane, 2014).

Figura 26. Sintomatología propia de la mancha bacteriana provocada por *X. arboricola* pv. *pruni* en hojas (A) y frutos (B, C y D) de almendro (Lamichhane, 2014).

1.5.11 Epidemiología de la mancha bacteriana de los frutales del género *Prunus*:

X. arboricola pv. *pruni* sobrevive el invierno en brotes latentes desde los cuales una vez iniciada la etapa vegetativa del ciclo productivo y los brotes empiezan a crecer, cambia a fase parasítica y coloniza el tejido foliar y los frutos. La bacteria penetra al interior del tejido foliar a través de los estomas o mediante heridas producidas por las labores agrícolas, la caída de hojas, las heladas, el granizo o el viento. Para que la infección sea exitosa, se requiere que los tejidos se encuentren en estado de congestión hídrica, haya temperaturas entre 19 y 30 °C y una elevada humedad relativa. En el caso del ciruelo, los chancros de los troncos y ramas también pueden actuar como fuentes de inóculo primario de este patógeno. Muy pocas infecciones se producen en condiciones ambientales secas y cálidas (Scortichini, 2010; Lamichhane, 2014; Palacio-Bielsa *et al.*, 2014).

Como resultado de esta primera infección se producen las lesiones de primavera, a partir de los cuales y mediante la producción de exudado la bacteria se disemina a otras regiones del árbol dando lugar a las lesiones de verano. En el otoño las infecciones se dan en los brotes latentes que serán los sitios donde la bacteria persista durante el invierno. A corta distancia, entre árboles de una misma plantación, la bacteria se disemina mediante el efecto aerosol generado por la lluvia y el viento, aunque también mediante la maquinaria agrícola. Se ha determinado que la bacteria es capaz de permanecer en el suelo en los restos vegetales, por lo que no se pueden descartar como fuentes de inóculo primario (Scortichini, 2010; Palacio-Bielsa *et al.*, 2014) (Figura 27).

Figura 27. Ciclo de la enfermedad de la mancha bacteriana de los frutales del género *Prunus* provocada por *X. arboricola* pv. *pruni* (Palacio-Bielsa *et al.*, 2014).

1.5.12 Diagnóstico y control de la mancha bacteriana de los frutales del género *Prunus*:

Al igual que en la mayoría de las bacteriosis de plantas, el control de la mancha bacteriana no es una tarea fácil, por lo que es necesario recurrir a una estrategia de control integrado que incluya y combine diferentes métodos. Para evitar la dispersión del patógeno hacia zonas libres del mismo es necesario realizar una detección temprana y para ello es imprescindible contar con sistemas de detección y diagnóstico eficaces que permitan determinar la presencia del patógeno tanto en material vegetal sintomático como asintomático.

Para el diagnóstico de *X. arboricola* pv. *pruni* se han desarrollado diversas herramientas. Se han descrito al menos dos medios semiselectivos para este patógeno, el primero de ellos, el XCP1, permite diferenciar las colonias del patógeno, ya que éstas se observan de color amarillo, mucoides y rodeadas de un halo translúcido como producto

de la degradación del Tween 80 (Palacio-Bielsa *et al.*, 2012). Un segundo medio de cultivo semiselectivo es el XPSM (Civerolo *et al.*, 1982) que permite observar al patógeno después de 6-7 días a 27 °C como colonias de color gris claro, convexas y opacas. Realizar el diagnóstico únicamente basado en la utilización de estos medios de cultivo no es recomendable ya que se han observado fenotipos semejantes con otras bacterias del género.

En cuanto al uso de métodos serológicos, se aconseja el uso de antisueros y kits disponibles comercialmente para este patógeno, tanto para su uso en inmunofluorescencia como para ensayos de tipo ELISA. El uso de estas técnicas ha sido validado tanto en ciruelo como en melocotonero. Sin embargo, el hecho de que los anticuerpos disponibles sean antisueros policlonales, hace que la detección no sea específica para *X. arboricola* pv. *pruni*. Además, los antisueros disponibles presentan reacción cruzada con otras especies de *Xanthomonas*, por lo que su uso solo se aconseja para determinar de manera rápida, si la muestra contiene bacterias de este género bacteriano, pero no para un diagnóstico preciso de la mancha bacteriana (López *et al.*, 2010).

En relación a los métodos de detección moleculares, se han desarrollado protocolos de PCR, PCR en tiempo real y de amplificación isotérmica de ácidos nucleicos (LAMP). Los protocolos basados en PCR convencional, han sido ampliamente utilizados basados en los cebadores Y17CoF/Y17CoR, que amplifican un fragmento de 943 pb, correspondiente a una región parcial del gen hipotético *ftsX* que parece corresponder a un transportador tipo ABC. A pesar de que permite identificar a *X. arboricola* pv. *pruni*, el método no resulta ser lo suficientemente sensible, por lo que no es aconsejado para ser usado con muestras vegetales asintomáticas. Además también se ha informado de su reacción cruzada con *X. arboricola* pv. *corylina* (López *et al.*, 2010).

Otro protocolo utilizado para detectar *X. arboricola* pv. *pruni* mediante PCR convencional es el basado en los cebadores XapF/XapR, los cuales amplifican una región de 243 pb asociada al cluster *hrp* de *Xanthomonas* (Park *et al.*, 2010). Según sus autores, este protocolo resulta ser específico para el patovar *pruni* y ha sido útil para detectar el patógeno tanto en hojas como en frutas naturalmente infectadas. Este método tiene una sensibilidad de hasta 3 pg de DNA bacteriano puro, lo cual corresponde aproximadamente a 300 células.

También han sido desarrollado un método de PCR convencional utilizando la variante de PCR multiplex (Pothier *et al.*, 2011a). Este método se basa en el uso de un primer par de cebadores que concuerdan con los descritos previamente para la amplificación de una región de 943 pb del gen *ftsX*, y un segundo par de cebadores (XarbQ-F/XarbQ-R) que amplifica una región de 402 pb del gen *qumA* asociado al metabolismo del quinato. Al utilizar ambas parejas de cebadores en modo multiplex, ambas regiones se amplifican en todas las cepas del patovar *pruni*, pero también en las cepas de *X. arboricola* pv. *corylina*. La sensibilidad de este método a partir de cultivo puro o de muestras inoculadas se encuentra en el rango de 5×10^2 UFC/ml.

Asimismo, se han descritos dos métodos de PCR tiempo real, el primero de ellos se basa en el uso de cebadores así como una sonda TaqMan sintetizados a partir de la región del gen *ftsX* descrita previamente (Palacio-Bielsa *et al.*, 2011, 2015b). Los cebadores Xap-2F/Xap-2R así como la sonda Xap-2p, permiten amplificar un fragmento de 72 pb. Este método, resulta ser altamente específico y sensible y ha sido validado por diversos laboratorios tanto en muestras sintomáticas como asintomáticas de melocotonero, almendro y ciruelo; llegándose a detectar en planta hasta 10^2 UFC/ml. Este protocolo es de los más utilizados actualmente, sin embargo, presenta reacción cruzada con *X. arboricola* pv. *corylina*.

Otro método de PCR tiempo real pero basado en la tecnología SYBR Green es el desarrollado a partir de los cebadores 29F/29R, que amplifican una región de 344 pb. Este método resulta ser específico para *Xanthomonas arboricola* pv. *pruni* y presenta una sensibilidad de 10^3 UFC/ml en cultivo puros y de 0.1 pg de DNA a partir de DNA bacteriano puro. Debido a que la sensibilidad observada en muestras vegetales naturalmente infectadas no fue lo suficientemente alta, se aconseja para esos casos utilizarlo en su versión de Bio-PCR, en el cual se añade un paso previo a la reacción de PCR en tiempo real consistente en un enriquecimiento de la población bacteriana mediante su cultivo en el medio semiselectivo XPSM (Ballard *et al.*, 2011).

Finalmente, entre los métodos moleculares, también se ha desarrollado un estrategia de detección basada en la tecnología LAMP (Bühlmann *et al.*, 2013). Este método presenta una sensibilidad semejante a la obtenida en el método de PCR tiempo real basado en la tecnología de sondas TaqMan descrito previamente (Palacio-Bielsa *et al.*, 2015b), por lo

que resulta ser un método rápido, específico y sensible para la detección de la mancha bacteriana y el diagnóstico de *X. arboricola* pv. *pruni*.

Además de un diagnóstico rápido y eficiente, para el control de esta enfermedad, también se aconseja el uso de material vegetal con baja susceptibilidad a este patógeno. Los cultivares que se han encontrado, sin embargo, no presentan las características deseadas para la producción de fruta de calidad, por lo que en general, se recomienda utilizar dentro de lo posible, cultivares diferentes a los descritos como más susceptibles en Francia e Italia (Palacio-Bielsa *et al.*, 2014).

En relación al control químico, de modo preventivo, se aconseja utilizar componentes químicos basados en el cobre, sin embargo el uso excesivo de este tipo de materiales podrían llevar a la generación de resistencia por parte del patógeno (Lamichhane, 2014). Cuando el patógeno ya ha sido detectado en el área de producción, se aconseja que los tratamientos con cobre se den en otoño, una vez se dé la caída de las hojas, para evitar que las bacterias pasen el invierno en las heridas de las hojas o en los brotes latentes. En primavera se aconseja también su aplicación con el fin de disminuir las infecciones primarias (Palacio-Bielsa *et al.*, 2014).

Entre las medidas agronómicas a tomar en cuenta para el control de la mancha bacteriana de los frutales de hueso y del almendro también se encuentran la selección de suelos aptos para el cultivo, ya que zonas de suelo arenoso y arcilloso o con capacidad de retener mucha humedad pueden favorecer el desarrollo de la enfermedad. Además, se aconseja podar durante el invierno aquellas zonas del árbol que se han visto afectadas, así como los chancros de ramas y troncos. Todos estos restos vegetales, así como los que se encuentren en el suelo deben ser eliminados y la herramienta de trabajo debe ser convenientemente desinfectada al realizar las tareas de poda (Palacio-Bielsa *et al.*, 2014).

En cuanto al control biológico de este patógeno, se han descrito varias cepas bacterianas epífitas con efecto antagonista sobre *X. arboricola* pv. *pruni*. En experimentos realizados *in vitro*, una cepa de *P. fluorescens* fue seleccionada debido a su mayor potencial inhibitorio. Durante dos diferentes ciclos productivos la cepa biocontroladora seleccionada dio lugar a una reducción en la severidad de la enfermedad en plantaciones de melocotonero y nectarino analizadas (Biondi *et al.*, 2009). Además de esta cepa, también se ha utilizado la cepa de *P. aeruginosa* LV. El sobrenadante liofilizado de

cultivos de esta bacteria aplicado a una concentración de 450 µg/ml sobre las plantas en condiciones de invernadero, produjo una reducción considerable de la severidad de la enfermedad. Además, sobre el patógeno, se observó un efecto sobre la producción de exopolisacárido y generó cambios en su morfología celular (Silva Vasconcellos *et al.*, 2014). De la misma manera, se ha comprobado la capacidad controladora de dos cepas no patógenas de *X. campestris* (AZ98101 y AZ98106), en ensayos de campo sobre diversos cultivares de melocotonero. De manera general, se observó una reducción significativa en aquellas parcelas inoculadas con cualquiera de las dos cepas, por lo que se han considerado como buenos candidatos para el control de *X. arboricola* pv. *pruni* (Kawaguchi, 2014). El uso de fagos para el control de esta bacteria también ha sido descrito (Randhawa & Civerolo, 1986). En ensayos sobre hoja, al aplicar 5 µl de una suspensión de los fagos purificados previamente de cultivos del patógeno en la zona de inoculación, producían una reducción de la población de bacteria tanto sobre la hoja como en el interior de la misma. Sin embargo, se observó que cerca del 50% de la población bacteriana era resistente al fago.

1.5.13 Condición fitosanitaria de *X. arboricola* pv. *pruni*:

X. arboricola pv. *pruni* agente causal de la mancha bacteriana de los frutales de hueso y el almendro, así como de especies ornamentales y silvestres del género *Prunus*, es considerado organismo de cuarentena según la legislación fitosanitaria de la Unión Europea (Directiva de la EU 2009/29/CE) y según la “European and Mediterranean Plant Protection Organization” (EPPO). Esta información se encuentra recogida en la lista A2 de la EPPO de organismos para los cuales se recomienda regulación (Roselló *et al.*, 2012; Lamichhane, 2014).

Capítulo 2: Secuenciación y características generales del genoma de cinco cepas de *X. arboricola* aisladas de *Prunus* spp. en España.

Draft genome sequence of *X. arboricola* pv. *pruni* strain Xap 33, Causal agent of bacterial spot disease on almond.

J. Garita-Cambronero¹, M. Sena-Vélez¹, A. Palacio-Bielsa², J. Cubero¹.

¹Departamento de Protección Vegetal, Laboratorio Bacteriología, Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA), Madrid, Spain. ²Centro de Investigación y Tecnología Agroalimentaria de Aragón (CITA), Zaragoza, Spain.

Genome Announcements 2014 May-Jun; 2(3): e00440-14. DOI: 10.1128/genomeA.00440-14

Draft genome sequence for virulent and avirulent strains of *Xanthomonas arboricola* isolated from *Prunus* spp. in Spain.

J. Garita-Cambronero¹, A. Palacio-Bielsa², M. M. López³, J. Cubero¹.

¹Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA), Madrid, Spain. ²Centro de Investigación y Tecnología Agroalimentaria de Aragón, Zaragoza, Spain. ³Instituto Valenciano de Investigaciones Agrarias, Valencia, Spain.

Standards in Genomic Sciences 2016 11:12. DOI: 10.1186/s40793-016-0132-3

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.

J. Garita-Cambronero¹, A. Palacio-Bielsa², M. M. López³, J. Cubero¹.

¹Departamento de Protección Vegetal, Instituto Nacional de Investigación y Tecnología Agraria, Madrid, Spain. ²Centro de Investigación y Tecnología Agroalimentaria de Aragón, Instituto Agroalimentario de Aragón-IA2 (CITA-Universidad de Zaragoza), Zaragoza, Spain. ³Instituto Valenciano de Investigaciones Agrarias, Valencia, Spain.

Genome Announcements 4(5):e00974-16. DOI:10.1128/genomeA.00974-16.

2.1 Introducción:

La secuenciación y los análisis de genomas bacterianos han cambiado drásticamente desde su inicio en el año 1995, hasta el punto que actualmente, gracias al avance de la tecnología disponible y la reducción en su coste, se han convertido en prácticas habituales en muchos laboratorios de microbiología. Como resultado de este avance, actualmente se cuenta con decenas de miles de genomas bacterianos disponibles (30.000 públicamente accesibles en el “National Center for Biotechnology Information” (NCBI) hasta 2015), los cuales sin duda, han contribuido a obtener toda una nueva visión sobre el mundo de las bacterias. La abundancia de datos disponibles ha provocado que la dificultad haya cambiado de obtener la secuencia de un genoma bacteriano, a la capacidad de ensamblar, analizar y tener un buen control de las secuencias obtenidas (Land *et al.*, 2015).

El primer genoma secuenciado de una bacteria fitopatógena fue el de *Xylella fastidiosa*, agente causal de la clorosis variegada de los cítricos (Simpson *et al.*, 2000). En *Xanthomonas*, la era de la secuenciación de genomas se inició en el año 2002 con la obtención del genoma completo de *X. citri* subsp. *citri* 306 y de *X. campestris* pv. *campestris* ATCC 33913 (da Silva *et al.*, 2002). Este estudio pionero, junto con la secuenciación del genoma de *X. oryzae* pv. *oryzae* KACC 10331 (Lee *et al.*, 2005) marcaron el inicio de la genómica comparativa en este género bacteriano. Estos trabajos iniciales permitieron la identificación de un amplio número de genes hipotéticamente relacionados con patogenicidad, convirtiendo a estos tres organismos en auténticos modelos para estudios de genómica funcional de *Xanthomonas*. Asimismo, fue un punto de partida para los estudios posteriores de genómica comparativa entre las diferentes especies del género.

Actualmente (octubre de 2016), existen disponibles en las bases de datos del NCBI al menos 521 genomas de *Xanthomonas* spp. de éstos, 72 son secuencias completas del genoma que cuentan con una anotación de sus genes curada manualmente y 21 genomas se encuentran completamente secuenciados pero su anotación ha sido realizada de manera automática. Por último, 428 genomas se encuentran en estado de “draft”, es decir, son menos precisos que un genoma completo, pero generalmente presentan una

alta calidad y al menos 90% de nucleótidos secuenciados. La secuencia de estos 428 genomas se encuentra representada en diversos segmentos o “contigs” (Collins, 1998).

El contar con este tipo de información genómica, tanto de cromosomas completos como incompletos, resultan de gran valor en actividades como el desarrollo de análisis filogenéticos más completos y precisos que permiten clasificar los diversos grupos bacterianos. Asimismo, es posible estudiar la diversidad genómica a nivel intraespecífico, por ejemplo, mediante el estudio de su pan genoma, lo cual permite determinar aquellos genes asociados con funciones biológicas básicas para la vida y el fenotipo de la especie, así como aquellos genes que contribuyen en la diversidad de la especie. Muchos de estos genes no son esenciales para la vida pero confieren ventajas selectivas como pueden ser la adaptación a un determinado nicho, la resistencia a antibióticos o la capacidad de colonizar un nuevo huésped (Tettelin *et al.*, 2008; Land *et al.*, 2015).

Este tipo de estudios se han llevado a cabo en el género *Xanthomonas* mediante la comparación de 10 genomas que representaban a tres especies del género. De esta manera se pudo conocer que el pan genoma de este grupo de cepas consiste en 12.951 secuencias codificantes (CDS), de las cuales 2.156 CDS representan el “core genome”, es decir, son compartidas por todas las cepas. El estudio filogenético basado en este grupo de genes permitió además determinar que la cepa de *X. campestris* pv. *vesicatoria* 85-10, en realidad es filogenéticamente distante al complejo de *X. campestris*, por lo que se propuso renombrarla como una nueva especie, *X. vesicatoria* (Blom *et al.*, 2009).

Posteriormente (Rodriguez-R *et al.*, 2012), se realizó un análisis filogenético basado en 989 grupos de genes orthologos compartidos por las especies *X. albilineans*, *X. axonopodis*, *X. campestris*, *X. euvesicatoria*, *X. fuscans*, *X. oryzae* y *X. vasicola*. Esta reconstrucción filogenética muestra que en general, el género *Xanthomonas* no es monofilético y cada una de las especies analizadas representa un clado diferenciado de los demás, excepto las especies *X. axonopodis*, *X. euvesicatoria* y *X. fuscans*, las cuales conforman un solo grupo. Por otro lado, en el caso de *X. albilineans*, ésta quedó contemplada en un grupo basal tanto de *Xanthomonas* como de *Xylella*, por lo que debería considerarse un género distinto con el fin de que la taxonomía del grupo concuerde con su filogenia.

En la especie *X. arboricola*, los trabajos de genómica comparativa se iniciaron con el uso de la hibridación genómica basada en el uso de microarreglos de ADN (Mayer *et al.*, 2011). Mediante esta tecnología, se comparó el contenido del genoma de *X. arboricola* pv. *pruni* 109 al hibridarlo con un chip de microarreglos de ADN de las especies *X. campestris* pv. *campestris* B100 y *X. campestris* pv. *vesicatoria* 85-10. Se obtuvo inicialmente, que la cepa del patovar *pruni* presenta una similitud de entre 93,3 y 94,8% con estas especies y comparten al menos un total de 2.826 CDS. Además, este estudio permitió identificar la presencia en el genoma de *X. arboricola* pv. *pruni* de 101 genes asociados con la producción de xanthano, la biosíntesis y el uso de azúcares, diversas exoenzimas, genes asociados a la estructura del T3SS, genes asociados a la regulación del quorum-sensing, así como algunos receptores del sistema TonB.

El estudio genómico de *X. arboricola*, basado en la secuenciación y análisis de sus genomas, se inicia en el año 2012 con la secuenciación del genoma de *X. arboricola* pv. *juglandis*, cepa NCPPB 1447, posteriormente, se han publicado, según las bases de datos del NCBI a noviembre de 2016, un total de 24 genomas así como la secuencia del plásmido pXap41 de *X. arboricola* pv. *pruni* (Pothier *et al.*, 2011c).

Del total de genomas disponibles para *X. arboricola*, cinco han sido secuenciados como parte de este trabajo y son tratados más adelante en este mismo capítulo. Los restantes 19 genomas, se dividen en dos cepas de *X. arboricola* pv. *celebensis*, NCPPB 1630 y NCPPB 1638 (Harrison *et al.*, 2016), causantes de la marchitez de la banana; uno de *X. arboricola* pv. *corylina* (Ibarra Caballero *et al.*, 2013), agente causal de la podredumbre bacteriana del avellano; nueve genomas de *X. arboricola* pv. *juglandis*, cepas CFBP 2528, CFBP 7179, J303, NCPPB 1447, Xaj2, Xaj4.1, Xaj43a y XajA3 (Higuera *et al.*, 2015; Pereira *et al.*, 2015; Cesbron *et al.*, 2015) causantes tanto de la necrosis apical como del chancro del nogal. Cuatro genomas de *X. arboricola* pv. *pruni*, cepas CFBP 3894, MAFF 301420, MAFF301427 y MAFF 311562 (López-Soriano *et al.*, 2016), causantes de la mancha bacteriana de los frutales del género *Prunus* y por último, tres cepas de *X. arboricola* que no pertenecen a ninguno de los nueve patovares descritos en esta especie. De estos últimos, la cepa 3004 (Ignatov *et al.*, 2015), ha sido descrita como causante de enfermedad en cebada, mientras que las cepas CFBP 7634 y CFBP 7651, aisladas de nogal, son consideradas no patógenas (Cesbron *et al.*, 2015).

A excepción de las cepas patógenas (CFBP 2528 y CFBP 7179) y no patógenas aisladas de nogal (CFBP 7634 y CFBP 7651), en las que se ha realizado un análisis más profundo, de todas las demás cepas, solamente se conocen las características generales sobre su secuenciación, ensamblaje y anotación inicial. Sobre ellas no existe más información que la presente en las bases de datos del NCBI o en publicaciones, realizadas a manera de nota corta con el respectivo anuncio genómico (Tabla 1).

En cuanto a las cuatro cepas de nogal mencionadas previamente (Cesbron *et al.*, 2015), el análisis comparativo de sus genomas permitió conocer que las cepas patógenas presentan un rango más amplio de elementos genéticos móviles que las cepas no patógenas y por ejemplo, la cepa CFBP 7179, contiene un elemento integrativo y conjugativo que le confiere resistencia al cobre. Se ha determinado además que el repertorio de efectores del T3SS es mayor en las cepas patógenas que en las no patógenas y por ejemplo para la cepa CFBP 7634, se encontró que no presentaba T3SS y contenía solamente ocho potenciales efectores, mientras que las cepas patógenas contenían entre 24 y 25. De igual manera sucedió con el T4SS el cual parece no estar completo en ninguna de las cepas estudiadas y sus potenciales efectores en cepas patógenas se presentaron en un mayor número (17-18 efectores), en comparación con aquellas no patógenas (14 efectores). Entre ambos grupos de cepas, según este mismo estudio, también se dieron diferencias en relación al repertorio de enzimas degradadoras de pared y en el correspondiente a las adhesinas no fimbriales, dentro de las cuales FhaB y YadA, solamente estaban presentes en las cepas no patógenas.

En este capítulo de la tesis, se presentarán las características generales asociadas al proyecto de secuenciación del genoma de cinco cepas de *X. arboricola* aisladas de *Prunus* spp. en España. Los resultados de este trabajo inicial y exploratorio se han dado a conocer a través de tres publicaciones en las cuales se describe el trabajo realizado de forma global, dirigido a dilucidar aspectos moleculares asociados a la patogénesis en *X. arboricola*. Las diversas publicaciones se muestran en orden cronológico y son la base sobre la cual se desarrolla el análisis genómico que se presentará en capítulos posteriores de este trabajo.

Tabla 1. Características de los genomas disponibles en la base de datos del NCBI para la especie *X. arboricola*.

Taxon	Cepa	GenBank acceso	Tamaño (Mb)	GC%	Genes	Proteínas
<i>Xanthomonas arboricola</i>	CITA 14	LXIB01	486.444	65,6	4.061	3.870
<i>Xanthomonas arboricola</i>	CITA 44	LJGM01	476.048	65,8	4.002	3.728
<i>Xanthomonas arboricola</i>	CITA 124	LXKK01	475.224	65,8	4.086	3.798
<i>Xanthomonas arboricola</i>	CFBP 7634	JZEH01	493.518	65,6	4.110	4.006
<i>Xanthomonas arboricola</i>	CFBP 7651	JZEI01	503.201	65,5	4.191	4.086
<i>Xanthomonas arboricola</i>	3004	AZQY01	47.659	66,0	3.997	3.775
<i>Xanthomonas arboricola</i> pv. <i>celebensis</i>	NCPBP 1630	JPHE01	498.021	65,5	4.113	3.977
<i>Xanthomonas arboricola</i> pv. <i>celebensis</i>	NCPBP 1832	JPHC01	490.802	65,6	4.081	3.967
<i>Xanthomonas arboricola</i> pv. <i>corylina</i>	NCCB 100457	APMC02	522.658	65,5	4.365	4.089
<i>Xanthomonas arboricola</i> pv. <i>juglandis</i>	Xaj 417	CP012251.1	521.894	65,4	4.358	4.188
<i>Xanthomonas arboricola</i> pv. <i>juglandis</i>	J303	LSGZ01	506.642	65,5	4.262	4.040
<i>Xanthomonas arboricola</i> pv. <i>juglandis</i>	CFBP 2528	JZEF01	508.448	65,5	4.263	4.095
<i>Xanthomonas arboricola</i> pv. <i>juglandis</i>	CFBP 7179	JZEG01	515.746	65,4	4.351	4.194
<i>Xanthomonas arboricola</i> pv. <i>juglandis</i>	NCPBP 1447	AJTL01	502.168	65,4	4.308	3.921
<i>Xanthomonas arboricola</i> pv. <i>juglandis</i>	Xaj2	LHBK01	510.122	65,4	4.266	4.081
<i>Xanthomonas arboricola</i> pv. <i>juglandis</i>	Xaj43a	LHBT01	514.414	65,4	4.314	4.115
<i>Xanthomonas arboricola</i> pv. <i>juglandis</i>	XajA3	LHBS01	511.898	65,6	4.283	4.093
<i>Xanthomonas arboricola</i> pv. <i>juglandis</i>	Xaj4.1	LHBL01	511.626	65,6	4.285	4.102
<i>Xanthomonas arboricola</i> pv. <i>pruni</i>	CITA 33	JHUQ01	510.066	65,4	4.222	3.720
<i>Xanthomonas arboricola</i> pv. <i>pruni</i>	IVIA 2626.1	LJGN01	502.767	65,4	4.253	3.849
<i>Xanthomonas arboricola</i> pv. <i>pruni</i>	CFBP 3894	LOMI01	505.907	65,4	4.228	3.967
<i>Xanthomonas arboricola</i> pv. <i>pruni</i>	MAFF301420	BAVC01	500.276	65,3	4.315	3.601
<i>Xanthomonas arboricola</i> pv. <i>pruni</i>	MAFF301427	BAVD01	490.544	65,4	5.180	4.686
<i>Xanthomonas arboricola</i> pv. <i>pruni</i>	MAFF311562	BAVB01	508.549	65,3	5.422	4.897

2.2 Draft genome sequence of *Xanthomonas arboricola* pv. *pruni* strain Xap33, causal agent of bacterial spot disease on almond.

2.2.1 Abstract

We report the annotated genome sequence of *Xanthomonas arboricola* pv. *pruni* strain Xap33 (CITA 33), isolated from almond leaves showing bacterial spot disease symptoms in Spain. The availability of this genome sequence will aid our understanding of the infection mechanism of this bacterium as well as its relationship to other species of the same genus.

2.2.2 Genome announcement

Xanthomonas arboricola pv. *pruni* (synonym *Xanthomonas campestris* pv. *pruni*) is a Gram-negative plant-pathogenic bacterium that causes bacterial spot disease of stone fruits and almond, one of the major threats of *Prunus* fruit crops. Symptoms occur on leaves, fruits, and twigs, ranging from necrotic angular lesions on leaves and sunken lesions on fruits to cankers on twigs. *X. arboricola* pv. *pruni* can be very damaging when severe infections occur on highly susceptible cultivars (Stefani, 2010; Roselló *et al.*, 2012; Socquet-Juglard *et al.*, 2012; Tjou-Tam-Sin *et al.*, 2012). This bacterium is considered a quarantine pathogen in the European Union phytosanitary legislation and in the European and Mediterranean Plant Protection Organization (EPPO) (Anonymous, 2000; EPPO, 2003). Here we report the draft genome sequence of *X. arboricola* pv. *pruni* strain Xap33 (CITA33), isolated from symptomatic almond leaves (*Prunus amygdalus* cv. Guara) cultivated in Aragón, Spain.

The genome of *X. arboricola* pv. *pruni* strain Xap33 was sequenced by the Ion Torrent Personal Genome Machine (PGM) (Life Technologies) instrument (579.53 Mb; 95-fold coverage). De novo assembly using CLC Genomics Workbench 7.0 (CLC bio, Denmark) produced a total length of 5,104,864 bp (G+C content, 65.43%) placed in 501 contigs with an N50 of 34,438 bp and a maximum contig length of 126,681 bp. The annotation for the Xap33 genome sequence was conducted using the automated Bacterial Annotation System (BASys) (Van Domselaar *et al.*, 2005), the Rapid Annotation using Subsystem Technology server (RAST) (Aziz *et al.*, 2008), and the

Prokaryotic Genome Annotation Pipeline (Tatusova *et al.*, 2013). RNAmmer 1.2 and tRNAscan-SE 1.21 were used to predict rRNAs and tRNAs, respectively (Lowe & Eddy, 1997; Lagesen *et al.*, 2007).

The annotation for the Xap33 genome sequence detected 4,348 coding sequences, 3 rRNAs, and 50 tRNAs, representing 436 subsystems compared to 461 subsystems found in *Xanthomonas axonopodis* pv. *citri* strain 306, which is the most closely related bacterium based on nucleotide similarity in the RAST database. 16S rRNA comparison analysis performed with BLASTn (Morgulis *et al.*, 2008) detected 99 to 100% identity to other *Xanthomonas* strains, including the *X. arboricola* pv. *juglandis* strain LMG 747. In addition, whole-genome nucleotide comparison using BLASTn against 214 complete and draft genomes from several *Xanthomonas* strains determined that the Xap33 genome sequence has 99% identity to the draft genome sequences of *Xanthomonas arboricola* pv. *pruni* strains MAFF301420, MAFF301427, and MAFF31562. Furthermore, in order to check the robustness of the Xap33 genome sequence, a multilocus sequence analysis using partial sequences of the housekeeping genes *dnaK*, *fyuA*, *gyrB*, and *rpoD* (Young *et al.*, 2008; Kaluzna *et al.*, 2014) clustered the strain Xap33 in the *X. arboricola* pv. *pruni* phylogroup, revealing a high similarity (99.89%) to the type strain *X. arboricola* pv. *pruni* ICMP 51 (CFBP 2535). The genome information presented here will allow genetic analysis and manipulation of *X. arboricola* pv. *pruni* Xap33, as well as future comparative studies with other *Xanthomonas arboricola* pathovars. This new information will help in understanding the molecular mechanisms in the plant-pathogen interactions related to bacterial spot disease on almond.

2.2.3 Nucleotide sequence accession numbers.

This whole-genome shotgun project has been deposited in DDBJ/EMBL/GenBank under the accession no. JHUQ00000000. The version described in this paper is version JHUQ01000000.

2.2.4 Acknowledgments

This work was supported financially by the Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA) project RTA2011-00140-C03-02. J. Garita-

Cambronero held a Ph.D. fellowship from the Spanish Government (Ministerio de Educación, Cultura y Deporte; fellowship FPU12/01000).

We thank Carla Clemente and Magdalena Lewicka from the STAB VIDA Pvt., Ltd., NGS Laboratory for their assistance with sequencing and analyses

2.3 Draft genome sequence for virulent and avirulent strains of *Xanthomonas arboricola* isolated from *Prunus* spp. in Spain

2.3.1 Abstract

Xanthomonas arboricola is a species in genus *Xanthomonas* which is mainly comprised of plant pathogens. Among the members of this taxon, *X. arboricola* pv. *pruni*, the causal agent of bacterial spot disease of stone fruits and almond, is distributed worldwide although it is considered a quarantine pathogen in the European Union. Herein, we report the draft genome sequence, the classification, the annotation and the sequence analyses of a virulent strain, IVIA 2626.1, and an avirulent strain, CITA 44, of *X. arboricola* associated with *Prunus* spp. The draft genome sequence of IVIA 2626.1 consists of 5,027,671 bp, 4,720 protein coding genes and 50 RNA encoding genes. The draft genome sequence of strain CITA 44 consists of 4,760,482 bp, 4,250 protein coding genes and 56 RNA coding genes. Initial comparative analyses reveals differences in the presence of structural and regulatory components of the type IV pilus, the type III secretion system, the type III effectors as well as variations in the number of the type IV secretion systems. The genome sequence data for these strains will facilitate the development of molecular diagnostics protocols that differentiate virulent and avirulent strains. In addition, comparative genome analysis will provide insights into the plant-pathogen interaction during the bacterial spot disease process.

2.3.2 Keywords

Xanthomonas arboricola, *Prunus* spp., Stone fruits, Bacterial spot disease, Plant pathogenic bacteria.

2.3.3 Introduction

Xanthomonas arboricola (Vauterin *et al.*, 1995) are plant associated bacteria in nine pathovars with a diverse range of biotic relationships (Bull *et al.*, 2010; Fischer-Le Saux *et al.*, 2015). Within this taxon, plant pathogenic strains with non-pathogenic strains have been described. Bacterial spot of *Prunus* spp. (*X. arboricola* pv. *pruni*), bacterial blight of *Juglans* spp. (*X. arboricola* pv. *juglandis*) and *Corylus* spp. (*X. arboricola* pv.

corylina) are among the most harmful diseases of these tree hosts. These bacterial diseases are distributed worldwide and the causal bacteria are regulated in several countries including the European Union, where *X. arboricola* pv. *pruni* is a quarantine pathogen (Scortichini *et al.*, 2002; Lamichhane, 2014).

Within the pathovars, *X. arboricola* pv. *pruni* is a major threat to cultivated, exotic and ornamental *Prunus* species. This bacterium has been identified as a pathogen of *P. armeniaca*, *P. avium*, *P. buergeriana*, *P. cerasus*, *P. crassipes*, *P. davidiana*, *P. domestica*, *P. donarium*, *P. dulcis*, *P. laurocesasus*, *P. mume*, *P. persica* and *P. salicina* (EFSA, 2014). During the last decade, some local outbreaks of bacterial spot in Spain have been reported on almond, peach, nectarine and plum (Palacio-Bielsa *et al.*, 2014). For initial characterization of the bacterial strains isolated from Spanish outbreaks of bacterial spot, we performed a polyphasic study based on a multilocus sequence analysis, as well as some phenotypic characters (Garita-Cambronero *et al.*, 2013). After the characterization that showed the presence of different molecular and phenotypic variants, selected strains were analysed to assess the differences at the whole genome level.

Genome sequencing of *X. arboricola* strains has been completed for five strains isolated from walnut, three from peach, two from *Musa* sp., one from almond (Garita-Cambronero *et al.*, 2014), one from barley (Ignatov *et al.*, 2015) and one from Turkish hazel (Ibarra Caballero *et al.*, 2013). Genome sequencing includes the plasmid pXap41 (Pothier *et al.*, 2011c), present in the *X. arboricola* pv. *pruni* strains. All these sequences have been deposited in the NCBI database. Four genome sequences are available for pathogenic strains from *Prunus*, identified as *X. arboricola* pv. *pruni*. However, with the exception of the strain CITA 33 isolated from almond (*P. amygdalus*, syn. *P. dulcis*) in Spain (Garita-Cambronero *et al.*, 2014), no detailed information about features of those genomes have been published. In the same way, there are no sequenced strains isolated from Japanese plum (*P. salicina*) or cherry rootstock (*P. mahaleb*). In addition, no avirulent strain of *X. arboricola* from *Prunus* spp. has been analysed at the whole-genome level. The occurrence of avirulent strains is of particular importance for a quarantine pathogen like *X. arboricola* pv. *pruni* with respect to accurate diagnosis of virulent strains.

Herein we present draft genome sequences for two *X. arboricola* strains: an avirulent strain, CITA 44, isolated from *P. mahaleb*, and *X. arboricola* pv. *pruni* strain, IVIA 2626.1, isolated from *P. salicina* cv. Fortuna, which differs from other sequenced strains in phenotypical features and virulence on several hosts (Garita-Cambroneró *et al.*, 2014). The genome analysis of these two strains as well as comparison with other related strains should provide insight into the genetics of the pathogenesis process in *X. arboricola* strains associated with the bacterial spot disease of stone fruits and almond.

2.3.4 Organism information

Classification and features:

Strain CITA 44 was isolated in 2009 from asymptomatic leaves of Santa Lucía SL-64 cherry rootstock (*P. mahaleb*) in a nursery located in the north-eastern Spanish region of Aragón. This strain showed flagella associated swarming and swimming motility on 0.5 % agar PYM plates and 0.3 % agar MMA plates, respectively. Additionally, strain CITA 44 showed type IV pili associated twitching motility in the interstitial surface between 1.5 % agar PYM layer and the plastic plate surface. According to the atomized oil assay (Burch *et al.*, 2010), this strain produced surfactant compounds on 1.5 % agar LB plates after 24 h at 27 °C. In accordance with a detached leaf assay, conducted with a cotton swap damped with 1×10^8 CFU/ml, on almond cv. Ferraduel, apricot cv. Canino, peach cv. Calanda and European plum (*P. domestica*) cv. Golden Japan, *X. arboricola* strain CITA 44 did not cause bacterial spot symptoms at 28 days post inoculation (dpi). Despite this lack of symptoms, the bacterium could be re-isolated after such period.

X. arboricola pv. *pruni* strain IVIA 2626.1 was isolated from symptomatic leaves of Japanese plum (*P. salicina* cv. Fortune) in the southwestern Spanish region of Extremadura in 2002. This strain showed swarming, swimming and twitching type motility as well as production of surfactant compounds in the same culture conditions described above for strain CITA 44. In addition, according to the detached leaf assay described previously, strain IVIA 2626.1 was able to produce bacterial spot symptoms on almond, peach and European plum but not on apricot after 28 dpi.

Classification of the strains was performed using an MLSA approach based on the partial sequences of the housekeeping genes *atpD*, *dnaK*, *efP*, *fyuA*, *glnA*, *gyrB* and *rpoD* of the strains CITA 44 and IVIA 2626.1 as well as related strains of *X. arboricola* (Fischer-Le Saux *et al.*, 2015). Nucleotide sequences were aligned with ClustalW and both ends of each alignment were trimmed (*atpD* 750 bp, *dnaK* 759 bp, *efP* 339 bp, *fyuA* 753 bp, *glnA* 675 bp, *gyrB* 735 bp and *rpoD* 756 bp) and concatenated to a total length sequence of 4,620 nucleotide positions. The phylogenetic tree was constructed using the maximum likelihood method implemented in MEGA 6.0 (Tamura *et al.*, 2013) using 1,000 bootstrap re-samplings. According to the phylogenetic analysis, strain CITA 44 belongs to the species *X. arboricola*, nevertheless, this strain could not be associated to any of the pathovars of this species. The concatenated sequence similarity among this strain and the other *X. arboricola* strains analysed varied from 97.08 % to 98.79 %. In contrast, strain IVIA 2626.1 was clustered in a group with the pathotype strain *X. arboricola* pv. *pruni* CFBP 2535, isolated from *P. salicina* in New Zealand, with a sequence similarity of 100%.

X. arboricola CITA 44 and *X. arboricola* pv. *pruni* IVIA 2626.1 strains are Gram-negative, non-sporulating, rod-shaped, motile cells with a single polar flagellum. Rod-shaped cells of CITA 44 are approximately 0.6 µm in width and 1.4–2.5 µm in length. Rod-shaped cells of IVIA 2626.1 are approximately 0.7 µm in width and 1.7–2.5 µm in length. These strains formed 2.0–3.0 mm colonies within 48 h at 27 °C on YPGA 1.5 % agar plates (Trébaol *et al.*, 2001). Both strains formed mucoid, circular, yellow colonies with a convex elevation and an entire margin (Fig. 1). Strains CITA 44 and IVIA 2626.1 grew in the nutritive culture media PYM (Siciliano *et al.*) and LB (Macwilliams & Liao, 2006), as well as in the minimal medium A (Sena-Vélez *et al.*, 2015). According to the Biolog GN2 system, both strains metabolized α-D-glucose, α-keto glutamic acid, bromosuccinic acid, D-cellobiose, D-fructose, D-mannose, D-psicose, D-threulose, glycyl-L-glutamic acid, L-glutamic acid, L-serine, pyruvic acid methyl ester, succinic acid, succinic acid mono-methyl-ester, sucrose and Tween 40. The carbon compound D-saccharic acid was only utilized by strain CITA 44. Dextrin and L-proline were only metabolized by strain IVIA 2626.1. In addition to this analysis, strain CITA 44 hydrolysed casein and starch, while strain IVIA 2626.1 did not (Table 1).

Minimum information about genome sequence (Field *et al.*, 2008) of *X. arboricola* strain CITA 44 and *X. arboricola* pv. *pruni* strain IVIA 2626.1, as well as their phylogenetic position, are provided in Table 1 and Fig. 2.

Fig. 1. Images of *X. arboricola* CITA 44 (up) and *X. arboricola* pv. *pruni* IVIA 2626.1 (down) cells using contrast-phase microscopy (left) and the appearance of the colony morphology after 48 h growing on YPGA agar medium at 27 °C (right). Flagella was stained (left) as described previously (Kraiselburd *et al.*, 2013).

Fig. 2. Phylogenetic tree highlighting the position of two *X. arboricola* strains (shown in bold) relative to the pathotype strains (PT) of *X. arboricola*. *X. citri* subsp. *citri* str. 306 (Gabriel *et al.*, 1989; Euzéby, 2007) was used as an outgroup. The tree was built based on the comparison of concatenated nucleotide sequences of seven housekeeping genes (*atpD*, *dnaK*, *efP*, *fyuA*, *glnA*, *gyrB* and *rpoD*) (Fischer-Le Saux *et al.*, 2013).

et al., 2015). Sequences were first aligned and concatenated. The phylogenetic tree was constructed by using MEGA 6.0 software (Tamura *et al.*, 2013) with Maximum Likelihood method based on Tamura-Nei model. Bootstrap values (1,000 replicates) are shown at the branch points. GenBank accession number of *X. citri* subsp. *citri* str. 306 genome sequence is shown in parenthesis; accession numbers associated to the housekeeping loci of the pathotype strains can be found in a previous study (Fischer-Le Saux *et al.*, 2015).

Table 1. Classification and general features of two *Xanthomonas arboricola* strains according to the MIGS recommendation (Field *et al.*, 2008) published by the Genomic Standards Consortium (Field *et al.*, 2011).

MIGS ID	Property	Term	Evidence code ^a
	Classification	Domain: <i>Bacteria</i>	TAS [54]
		Phylum: <i>Proteobacteria</i>	TAS [55]
		Class: <i>Gammaproteobacteria</i>	TAS [56–58]
		Order: <i>Lysobacterales</i>	TAS [57, 59, 60]
		Family: <i>Lysobacteraceae</i>	TAS [57, 58, 60]
		Genus: <i>Xanthomonas</i>	TAS [1]
		Species: <i>Xanthomonas arboricola</i>	IDA
		Strain: CITA 44, IMA 2626.1	IDA
	Gram stain	Negative	TAS [61]
	Cell shape	Rod-shaped	IDA
	Motility	Motile	IDA
	Sporulation	Non-sporulating	IDA
	Temperature range	4-37 °C	TAS [1]
	Optimum temperature	27 °C	IDA
	pH range; Optimum	7.5-8.5	TAS [61]
	Carbon source	α-D-glucose, α-keto glutaric acid, bromosuccinic acid, D-cellobiose, D-fructose, D-mannose, D-psicose, D-saccharic acid (only strain CITA 44), D-threulose, Dextrin (only strain IMA 2626.1), glycyl-L-glutamic acid, L-glutamic acid, L-proline (only strain IMA 2626.1), L-serine, pyruvic acid methyl ester, succinic acid, succinic acid mono-methyl ester, Sucrose, tween 40	IDA
	Energy metabolism	Chemoorganotrophic	TAS [1]
MIGS-6	Habitat	Plants	IDA TAS [1]
MIGS-6.3	Salinity	0-6.0 % NaCl	TAS [1]
MIGS-10	Extrachromosomal elements	None in CITA 44, one in IMA 2626.1	IDA, TAS [12]
MIGS-22	Oxygen requirement	Aerobic	IDA
MIGS-15	Biotic relationship	Epiphyte and endophyte	TAS [1]
MIGS-14	Pathogenicity	CITA 44 is avirulent; IMA 2626.1 is virulent on almond, peach and European plum	IDA
	Host	Mahaleb cherry (<i>P. mahaleb</i>) (CITA 44) and plum (<i>P. salicina</i>) (IMA 2626.1)	IDA
	Host taxa ID	129217 (CITA 44) and 88123 (IMA 2626.1)	
	Isolation source	Leaf	IDA
MIGS-4	Geographic location	Spain	IDA
MIGS-5	Sample collection	2002 (IMA 2626.1) and 2009 (CITA 44)	IDA
MIGS-4.1	Latitude	Unknown	NAS
MIGS-4.2	Longitude	Unknown	NAS
MIGS-4.4	Altitude	Unknown	NAS

^aEvidence codes - IDA: Inferred from Direct Assay; TAS: Traceable Author Statement (i.e., a direct report exists in the literature); NAS: Non-traceable Author Statement (i.e., not directly observed for the living, isolated sample, but based on a generally accepted property for the species, or anecdotal evidence). These evidence codes are from the Gene Ontology project [62]

1: (Vauterin *et al.*, 1995), 12: (Pothier *et al.*, 2011c), 54: (Woese *et al.*, 1990), 55: (Garrrity *et al.*, 2005a), 56: (Garrrity *et al.*, 2005b), 57: (Euzéby, 2004), 58: (Williams & Kelly, 2013), 59: (Christensen & Cook, 1978), 60: (Naushad *et al.*, 2015), 61: (Van den Mooter & Swings, 1990), 62: (Ashburner *et al.*, 2000).

2.3.4 Genome sequencing information

Genome Project history

X. arboricola strain CITA 44 and *X. arboricola* pv. *pruni* strain IVIA 2626.1 were selected for comparative whole sequencing analysis as *X. arboricola* strains isolated from *Prunus* spp. with several different phenotypic characters including virulence. Comparative genomics among the avirulent strain CITA 44 and the available *Prunus*-pathogenic strains including IVIA 2626.1 should be useful for identifying the molecular determinants associated with pathogenesis as well as those associated with host resistance and for diagnostic characterization of *X. arboricola* strains causing bacterial spot of *Prunus* spp. Whole Genome Shotgun Projects have been deposited at DDBJ/EMBL/GenBank under the accession numbers LJGM000000000 and LJGN000000000. The versions described in this paper are versions LJGM010000000 and LJGN010000000. Table 2 summarizes the project information and its association with MIGS.

Table 2. Project information

MIGS ID	Property	Term/Strains	
		CITA 44	IVIA 2626.1
MIGS 31	Finishing quality	Draft	Draft
MIGS 28	Libraries used	One 400 bp Ion Torrent library	One 400 bp Ion Torrent library
MIGS 29	Sequencing platforms	Ion Torrent PGM	Ion Torrent PGM
MIGS 31.2	Fold coverage	198x	92x
MIGS 30	Assemblers	MIRA 4.0	MIRA 4.0
MIGS 32	Gene calling method	Glimmer 3.0 that used in the RAST pipeline	Glimmer 3.0 that used in the RAST pipeline
	Locus Tag	AN651	AN652
	Genbank ID	LJGM000000000	LJGN000000000
	GenBank Date of Release	06-October-2015	06-October-2015
	GOLD ID	Gp0124696	Gp0124697
	BIOPROJECT	PRJNA294649	PRJNA294655
MIGS 13	Source Material Identifier	CITA 44	IVIA 2626.1
	Project relevance	Agricultural, Environmental, Biotechnology, Plant-Bacteria Interaction	Agricultural, Environmental, Biotechnology, Plant-Bacteria Interaction

Growth conditions and genomic DNA preparation

X. arboricola strain CITA 44 and *X. arboricola* pv. *pruni* strain IVIA 2626.1 are deposited and available at the bacterial collections of the Instituto Valenciano de Investigaciones Agrarias (IVIA, Valencia, Spain) and the Centro de Investigación y

Tecnología Agroalimentaria de Aragón (CITA, Zaragoza, Spain). Both strains were streaked on 1.5 % agar LB plates and were grown for 48 h at 27 °C. A single colony of each strain was inoculated separately in 30 ml of LB broth and grown on an orbital shaker for 24 h at 27 °C. DNA from pure bacterial cultures was extracted using a QIAamp DNA miniKit (Qiagen, Barcelona, Spain) according to the manufacturer instructions. DNA quality and quantity were determined by 1 % agarose gel electrophoresis, as well as using the Qubit flurometer (Invitrogen) according to the Quant-it dsDNA BR Assay Kit (Invitrogen) manufacturer instructions, and by a spectrophotometry (NanoDrop 2000 spectrophotometer, Thermo Scientific). A 2.0 µg/µl aliquot of 200 ng/µl sample was submitted for the sequencing.

Genome sequencing and assembly

The draft genome sequences for strains CITA 44 and IVIA 2626.1 were generated at the STAB VIDA Next Generation Sequencing Laboratory (Caparica, Portugal) using the Ion Torrent sequencing technology. Draft genome assembly of strain CITA 44 was based on 3,060,638 usable reads with a total base number of 948,933,067. The mean read length was 361.70 ± 93.50 and the mode read length was 385 bp. The draft genome assembly of IVIA 2626.1 was based on 2,317,319 reads, with a total base number of 461,361,072. The mean read length and the mode read length for this strain were 201.80 ± 85.30 bp and 241 bp, respectively. Genomic assemblies were constructed using MIRA 4.0 (Chevreux *et al.*, 1999). From the total of contigs generated, only those with a contig size above 500 bp and an average coverage above 99 in the case of CITA 44, and 40, in the case of IVIA 2626.1 were considered significant. Finally, 71 contigs (N50 = 120,981 bp; largest contig = 352,479 bp; average coverage = 198X) were generated for strain CITA 44 and for strain IVIA 2626.1, 214 contigs (N50 = 47,650; largest contig = 115,385; average coverage = 92X) were generated.

Genome annotation

The assembled draft genome for both strains was annotated using the RAST platform and the gene-caller GLIMMER 3.02 (Delcher, 1999; Aziz *et al.*, 2008). RNAmmer version 1.2 (Lagesen *et al.*, 2007) and tRNAscan-SE version 1.21 (Schattner *et al.*,

2005) were used to predict rRNAs and tRNAs, respectively. Signal peptides and transmembrane domains were determined using the SignalP 4.1 server (Petersen *et al.*, 2011) and the TMHMM server version 2.0 (Krogh *et al.*, 2001), respectively. Assignment of genes to the COG database (Tatusov, 2000) and Pfam domains (Finn *et al.*, 2014) was performed with the NCBI conserved domain database using an expected value threshold of 0.001 (Marchler-Bauer *et al.*, 2014).

Major structural components associated with the flagellum (Blocker *et al.*, 2003; Chevance & Hughes, 2008), the type IV pilus (Dunger *et al.*, 2014), the type III secretory system (White *et al.*, 2009; Büttner & Bonas, 2010; Abby & Rocha, 2012) and the type III effectors (White *et al.*, 2009; Hajri *et al.*, 2012) as well as the type IV secretory system and effectors (Christie *et al.*, 2014; Waksman & Orlova, 2014; Guglielmini *et al.*, 2014), were identified in the draft genome sequence for each strain. Initially, the query of those genes was based on the coding sequence regions automatically annotated by RAST, and were confirmed using the BLASTn and BLASTx tools available at NCBI. Those components which were not automatically annotated were found in the genome sequence using the progressive Mauve alignment method (Darling *et al.*, 2010). Nucleotide sequences of the genes used for these alignments were obtained from other xanthomonads in the NCBI gene database. Finally, the nucleotide sequence of the aligned regions was analysed using the BLAST approaches mentioned above. Those sequences with query coverage and identity percentage higher than 90 % were annotated. Additionally, the core components of the T3SS and T4SS were searched using the T346Hunter application (Martínez-García *et al.*, 2015). T3Es and T4Es genes were predicted using the Effective database (Jehl *et al.*, 2011) after selection of the “gram-” parameter as organism type and the “plant set” parameter as classification module, and the SecReT4 tool (Bi *et al.*, 2013), respectively. All the predicted genes were corroborated and annotated according to the BLAST parameters mentioned above.

2.3.5 Genome properties

The draft genome sequence of *X. arboricola* strain CITA 44 was 4,760,482 bp in length with an average GC content of 65.8 %, which is similar to that for other genomes of this

species (65.4 to 66.0 %) reported in the NCBI genome database. For this strain, 4,306 genes were predicted and 4,250 were determined as protein coding genes. From these protein coding genes, 3,330 genes were assigned to a putative function and the remaining 920 were designated as hypothetical proteins. This strain presented 3 rRNA and 53 tRNA genes. In the case of the *X. arboricola* pv. *pruni* strain IVIA 2626.1, the draft genome sequence was 5,027,671 bp in length with an average GC content of 65.4 %, which is the same as for other strains of *X. arboricola* pv. *pruni* according to the NCBI database. A total of 4,770 genes were predicted and, among them, 4,720 were predicted as protein coding genes with 69.17 % assigned to a function and 30.83 % designated as hypothetical proteins. 50 RNA genes (3 rRNA and 47 tRNA genes) were predicted for this strain. The properties and characteristics associated with these genomes are presented in Table 3. The classification of the predicted protein coding genes into COG functional categories (Galperin *et al.*, 2015) is summarized in Fig. 3 and Table 4.

Fig. 3. Graphical circular representation of the draft genome of *X. arboricola* CITA 44 and *X. arboricola* pv. *pruni* IVIA 2626.1. The contigs of both strains were ordered by Mauve (Darling *et al.*, 2004) using the genome sequence of *X. campestris* pv. *campestris* ATCC 33913 (Dowson, 1939; Skerman *et al.*, 1980) as the reference. COG categories were assigned to genes by NCBI's conserved domain database (Marchler-Bauer *et al.*, 2014). The circular map was constructed using CGView (Stothard & Wishart, 2005). From outside to center: Genes on forward strand (colored by COG categories); genes on reverse strand (colored by COG categories); GC content; GC skew.

Table 3. Genome statistics

Attribute	Strain			
	CITA 44		IMA 2626.1	
	Value	% of total	Value	% of total
Genome size (bp)	4,760,482	100.00	5,027,671	100.00
DNA coding (bp)	3,992,937	83.88	4,295,592	84.44
DNA G + C (bp)	3,134,520	65.80	3,288,794	65.40
Total genes	4306	100.00	4770	100.00
Protein coding genes	4250	98.62	4720	98.95
RNA genes	56	1.38	50	1.05
Pseudo genes	0	0.00	0	0.00
Genes with function prediction	3330	78.35	3265	69.17
Genes assigned to COGs	3137	73.81	3237	68.58
Genes with Pfam domains	3337	78.51	3433	72.73
Genes with signal peptides	526	12.37	545	11.55
Genes with transmembrane helices	1121	26.37	1221	25.86
CRISPR repeats	1	-	1	-

Table 4. Number of genes associated with general COG functional categories.

Code	Strain				Description
	CITA 44		IMA 2626.1		
	Value	% age	Value	% age	
J	218	5.13	217	6.70	Translation, ribosomal structure and biogenesis
A	1	0.02	2	0.06	RNA processing and modification
K	193	4.54	199	6.15	Transcription
L	111	2.61	127	3.92	Replication, recombination and repair
B	1	0.02	1	0.03	Chromatin structure and dynamics
D	35	0.82	40	1.23	Cell cycle control, cell division, chromosome partitioning
V	63	1.48	68	2.10	Defense mechanisms
T	211	4.96	208	6.42	Signal transduction mechanisms
M	225	5.29	235	7.26	Cell wall/membrane biogenesis
N	114	2.68	119	3.67	Cell motility
Z	2	0.05	2	0.06	Cytoskeleton
W	2	0.05	2	0.06	Extracellular structures
U	69	1.62	81	2.50	Intracellular trafficking and secretion
O	160	3.76	171	5.28	Posttranslational modification, protein turnover, chaperones
C	184	4.33	173	5.34	Energy production and conversion
G	220	5.18	215	6.64	Carbohydrate transport and metabolism
E	225	5.29	239	7.38	Amino acid transport and metabolism
F	76	1.79	76	2.35	Nucleotide transport and metabolism
H	153	3.60	144	4.45	Coenzyme transport and metabolism
I	156	3.67	156	4.82	Lipid transport and metabolism
P	218	5.13	212	6.55	Inorganic ion transport and metabolism
Q	57	1.34	63	1.95	Secondary metabolites biosynthesis, transport and catabolism
R	223	5.25	223	6.89	General function prediction only
S	213	5.01	223	6.89	Function unknown
X	7	0.16	41	1.27	Mobilome: prophages, transposons
-	1113	26.19	1483	31.42	Not in COGs

The total is based on the total number of protein coding genes in the annotated genome

2.3.6 Insights from the genome sequence

Based on the phenotypic differences between CITA 44 and IVIA 2626.1 strains, selected genes associated with motility and pathogenicity were analysed (Table 5). No differences were observed for the structural components associated with bacterial flagella. A total of 30 out of the 31 components described for this organelle were identified (Chevance & Hughes, 2008), but neither of the two strains contained a homolog of the *flhE* gene. Regarding the 27 components associated with type IV pilus biogenesis and regulation in *Xanthomonas* (Dowson, 1939; Skerman *et al.*, 1980; Dunger *et al.*, 2014), *fimX*, *pilD*, *pilE*, *pilL* and *pilW* genes were absent in strain CITA 44, whereas strain IVIA 2626.1 sequence did not contain homologs for *fimX* and *pilL* genes.

In the genus *Xanthomonas*, 24 structural and regulatory components of the T3SS have been determined. They are present in the *hrp* gene cluster which is regulated by the master regulons HrpG and HrpX (Guo *et al.*, 2011). Strain CITA 44 did not contain any of the 24 components of this gene cluster except two coding sequences which correspond to *hrpG* and *hrpX* homologs. The absence of T3SS has also been reported for another *X. arboricola* strain isolated from barley as well as for *X. cannabis* (Ignatov *et al.*, 2015; Jacobs *et al.*, 2015). The absence of the genes *hrcC*, *hrcJ*, *hrcN*, *hrcR*, *hrcS*, *hrcT*, *hrcU*, *hrcV*, *hrpB1*, *hrpD5* and *hrpF* was corroborated by conventional PCR as previously described (Hajri *et al.*, 2012). In the case of strain IVIA 2626.1, 22 out of the 24 components, as well as homologs for the two master regulons were present, but no homologs for *hpaF* and *hrpB5* were found. Homologs for these two genes were also absent in all the genome sequences of *X. arboricola* publicly available. Sixty T3Es described in genus *Xanthomonas* were absent in strain CITA 44 and absence of 21 of them, identified in *X. arboricola* pv. *pruni*, was corroborated by conventional PCR using specific primers (Hajri *et al.*, 2012). On the other hand, strain IVIA 2626.1 contained 22 T3Es, 21 of them were described previously in other *X. arboricola* pv. *pruni* strains (Hajri *et al.*, 2012). In addition to these effectors, a homolog of *xopAQ* was found. Both strains contained all 12 components associated with *Agrobacterium tumefaciens* (Conn, 1942; Skerman *et al.*, 1980) VirB/VirD4 T4SS (Guglielmini *et al.*, 2014). Additionally, strain IVIA 2626.1 harbored a gene cluster homologous to the type four conjugation cluster (*tfc*). This cluster is composed by 24 genes associated with the

expression of a conjugative pilus which is involved in the propagation of genomic islands (Juhas *et al.*, 2007). In strain IVIA 2626.1, 17 out of the 24 genes associated with the T4SS were found and, within them, *tfc2*, *tfc4*, *tfc12*, *tfc14*, *tfc16*, *tfc22* and *tfc23* were identified as the core components required for the functioning of this T4SS (Juhas *et al.*, 2007).

An additional feature of the *X. arboricola* pv. *pruni* sequence is the presence of the plasmid pXap41 (41,102 Kbp) (Pothier *et al.*, 2011c). This plasmid is exclusively in *X. arboricola* pv. *pruni* strains and is associated with virulence because it contains some T3Es such as XopE3. Genome alignment of the plasmid pXap41 nucleotide sequence and the draft genome sequence for strain IVIA 2626.1 showed a region of 41.1 Kbp which was 99.90 % similar to the pXap41 plasmid of *X. arboricola* pv. *pruni* strain CFBP 5530. Conversely, no sequence region in the strain CITA 44 draft genome was similar to this plasmid. Negative results in the amplification of the genes *repA1*, *repA2* and *mobC* associated with pXap41 (Pothier *et al.*, 2011c) confirmed the absence of this plasmid in strain CITA 44.

Table 5. Molecular components putatively involved in motility and pathogenesis

	Shared by CITA 44 ^a and IMA 2626.1	Absent in CITA 44 and IMA 2626.1	Unique in IMA 2626.1
Flagella	<i>flgB</i> , <i>flgC</i> , <i>flgD</i> , <i>flgE</i> , <i>flgF</i> , <i>flgG</i> , <i>flgH</i> , <i>flgI</i> , <i>flgK</i> , <i>flgL</i> , <i>flgM</i> , <i>flgN</i> , <i>flhA</i> , <i>flhB</i> , <i>fliC</i> , <i>fliD</i> , <i>fliE</i> , <i>fliF</i> , <i>fliG</i> , <i>fliH</i> , <i>fliJ</i> , <i>fliK</i> , <i>fliL</i> , <i>fliM</i> , <i>fliN</i> , <i>fliO</i> , <i>fliP</i> , <i>fliQ</i> , <i>fliR</i> , <i>fliK</i>	<i>flhE</i>	-
Type IV pilus	<i>fimT</i> , <i>pilA</i> , <i>pilB</i> , <i>pilC</i> , <i>pilF</i> , <i>pilG</i> , <i>pilH</i> , <i>pilJ</i> , <i>pilM</i> , <i>pilN</i> , <i>pilO</i> , <i>pilP</i> , <i>pilQ</i> , <i>pilR</i> , <i>pilS</i> , <i>pilT</i> , <i>pilU</i> , <i>pilV</i> , <i>pilX</i> , <i>pilY1</i> , <i>pilZ</i>	<i>fimX</i> , <i>pilL</i>	<i>pilD</i> , <i>pilE</i> , <i>pilW</i>
Type III Secretion System	<i>hrpG</i> , <i>hrpX</i>	<i>hpaF</i> , <i>hrpB5</i>	<i>hpa1</i> , <i>hpa2</i> , <i>hpaB</i> , <i>hpaF</i> , <i>hpaP</i> , <i>hrcC</i> , <i>hrcJ</i> , <i>hrcN</i> , <i>hrcQ</i> , <i>hrcR</i> , <i>hrcS</i> , <i>hrcT</i> , <i>hrcU</i> , <i>hrcV</i> , <i>hrpB1</i> , <i>hrpB2</i> , <i>hrpB4</i> , <i>hrpB5</i> , <i>hrpB7</i> , <i>hrpD5</i> , <i>hrpD6</i> , <i>hrpE</i> , <i>hrpF</i>
Type III effectors	-	<i>avrBs1</i> , <i>avrBs3</i> , <i>xopAA</i> , <i>xopAB</i> , <i>xopAC</i> , <i>xopAD</i> , <i>xopAE</i> , <i>xopAG</i> , <i>xopAJ</i> , <i>xopAK</i> , <i>xopAL1</i> , <i>xopAL2</i> , <i>xopAM</i> , <i>xopAO</i> , <i>xopAP</i> , <i>xopAQ</i> , <i>xopAR</i> , <i>xopAS</i> , <i>xopAT</i> , <i>xopB</i> , <i>xopC1</i> , <i>xopD</i> , <i>xopE1</i> , <i>xopF2</i> , <i>xopH</i> , <i>xopI</i> , <i>xopJ1</i> , <i>xopJ2</i> , <i>xopJ3</i> , <i>xopJ4</i> , <i>xopJ5</i> , <i>xopO</i> , <i>xopP</i> , <i>xopT</i> , <i>xopU</i> , <i>xopW</i> , <i>xopY</i> , <i>xopZ2</i>	<i>avrBs2</i> , <i>avrXccA1</i> , <i>hpaA</i> , <i>hprW</i> , <i>xopA</i> , <i>xopAF</i> , <i>xopAH</i> , <i>xopAI</i> , <i>xopAQ</i> , <i>xopE2</i> , <i>xopE3</i> , <i>xopF1</i> , <i>xopG</i> , <i>xopK</i> , <i>xopL</i> , <i>xopN</i> , <i>xopQ</i> , <i>xopR</i> , <i>xopV</i> , <i>xopX</i> , <i>xopZ</i>
Type IV Secretion System	<i>virB1</i> , <i>virB2</i> , <i>virB3</i> , <i>virB4</i> , <i>virB5</i> , <i>virB6</i> , <i>virB7</i> , <i>virB8</i> , <i>virB9</i> , <i>virB10</i> , <i>virB11</i> , <i>virD4</i>	<i>tfc1</i> , <i>tfc7</i> , <i>tfc11</i> , <i>tfc17</i> , <i>tfc18</i> , <i>tfc20</i> , <i>tfc21</i>	<i>tfc2</i> , <i>tfc3</i> , <i>tfc4</i> , <i>tfc5</i> , <i>tfc6</i> , <i>tfc8</i> , <i>tfc9</i> , <i>tfc10</i> , <i>tfc12</i> , <i>tfc13</i> , <i>tfc14</i> , <i>tfc15</i> , <i>tfc16</i> , <i>tfc19</i> , <i>tfc22</i> , <i>tfc23</i> , <i>tfc24</i>

^aCITA 44 did not present any unique component putatively involved in the analysed features

2.3.7 Conclusions

Here we report and describe the draft genome sequence for two *X. arboricola* strains, CITA 44 and IVIA 2626.1, isolated from *Prunus* in Spain and associated with bacterial spot of stone fruits and almond by PCR protocols for identification of this pathovar (Pagani, 2004; Palacio-Bielsa *et al.*, 2011). The phenotype of these two strains varied for motility and virulence. Initial genomic analysis identified several differences associated with motility (Type IV pilus) and virulence (T3SS, T3Es and T4SS), including the presence of the putative virulence plasmid pXap41 only in *X. arboricola* pv. *pruni* IVIA 2626.1 and the absence of the T3SS, T3Es and the plasmid pXap41 in the avirulent strain CITA 44. All these features make the avirulent strain a candidate for comparative studies to elucidate the molecular processes associated with the plant host interaction and virulence for strains of *X. arboricola* on *Prunus* species. Likewise, comparative genomic studies with related strains could provide target sequences for design of molecular diagnostics for the different pathovars of *X. arboricola*, as well as to differentiate between virulent and avirulent strains. Further functional studies will also provide insights into the pathogenesis process for *X. arboricola* strains associated with bacterial spot of stone fruits and almond.

2.3.8 Acknowledgements

This work was supported by the Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA) grants projects RTA2011-00140-C03-02, RTA2014-00018-C02-01 and XAPDIAG EUPHRESCO II project. JGC held a PhD fellowship from the Spanish Government (Ministerio de Educación, Cultura y Deporte fellowship FPU12/01000). We like to express our gratitude to Dr. James H. Graham from the University of Florida, Citrus Research and Education Center (CREC), for the scientific and English revision of this manuscript.

2.4 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.

2.4.1 Abstract

The draft genome sequences of two strains of *Xanthomonas arboricola*, isolated from asymptomatic peach trees in Spain, are reported here. These strains are avirulent and do not belong to the same phylogroup as *X. arboricola* pv. *pruni*, a causal agent of bacterial spot disease of stone fruits and almonds.

2.4.2 Genome announcement

Xanthomonas arboricola is a Gram-negative species that comprises nine subspecific phylogroups (Vauterin *et al.*, 1995; Fischer-Le Saux *et al.*, 2015). In addition, some strains that differ from the pathovars *X. arboricola* pv. *pruni* and *X. arboricola* pv. *juglandis*, isolated from the plant genera *Prunus* and *Juglans*, have also been described. These strains are unable to cause disease because they lack some essential characteristics related to pathogenesis, such as the type III secretion system and related effectors (Cesbron *et al.*, 2015; Garita-Cambronero *et al.*, 2016a).

The two new genomes presented here, which belong to the avirulent strains CITA 14 and CITA 124, were compared to the genomes of other *X. arboricola* strains (Ibarra Caballero *et al.*, 2013; Garita-Cambronero *et al.*, 2014; Higuera *et al.*, 2015; Ignatov *et al.*, 2015; Harrison *et al.*, 2016) and could be helpful in understanding the pathogenesis and evolution in this species. Moreover, the genomes will be useful for generating accurate diagnostic tools to differentiate the quarantine pathogen *X. arboricola* pv. *pruni* from other nontarget strains, thereby avoiding unnecessary control measures that could result in high economic losses as occurred with other xanthomonads (Graham & Gottwald, 1991; Jalan *et al.*, 2011).

Strains CITA 14 and CITA 124 were isolated from asymptomatic peach trees (*Prunus persica*) in Spain and were deposited in the Centro de Investigación y Tecnología Agroalimentaria de Aragón (CITA) in Zaragoza, Spain. To confirm their identification,

a multilocus sequence analysis using partial sequences of the housekeeping genes *dnaK*, *fyuA*, *gyrB*, and *rpoD* was performed (Young *et al.*, 2008). By means of such analysis, these strains were not clustered into any of the established pathovars of *X. arboricola*. CITA 14 and CITA 124 were enclosed in a group with CITA 44, another atypical strain of *X. arboricola* isolated from the rootstock Santa Lucía SL-64 (*Prunus mahaleb*) (Garita-Cambronerio *et al.*, 2016a), which shows sequence identities of 98.3% with strain CITA 14 and 98.9% with strain CITA 124.

Genome sequencing was performed using Ion Torrent sequencing technology (PGM, Life Technologies) at STAB VIDA, Caparica, Portugal (733.06 Mb, 100-fold coverage for CITA 14 and 403.99 Mb, 50-fold coverage for CITA 124). De novo assembly was conducted using CLC Genomics Workbench version 8.5.1 (CLC bio, Denmark) and MIRA version 4.0 (Chevreux *et al.*, 1999) and the subsequent merging of the two assemblies was done using Sequencher version 5.4 (Gene Codes Corporation, USA). Total lengths of 4,746,212 bp (65.60% G+C content) and 4,752,241 bp (65.8% G+C content) were generated and placed in 72 (N50, 108,424 bp; maximum contig length, 265,354 bp) and 128 contigs (N50, 60,818 bp; maximum contig length 189,085) for CITA 14 and CITA 124, respectively.

Annotations were conducted using the NCBI Prokaryotic Genome Annotation Pipeline version 3.1 (Angiuoli *et al.*, 2008) and the Rapid Annotations using Subsystem Technology server (Aziz *et al.*, 2008). The annotation for strain CITA 14 detected 3,974 coding sequences, four rRNAs, and 53 tRNAs, representing 439 subsystems; meanwhile, for strain CITA 124, 3,797 coding sequences, three rRNAs, and 50 tRNAs, representing 442 subsystems, were detected. A detailed genome comparative analysis will be addressed in a future publication.

2.4.3 Accession numbers.

These whole-genome shotgun projects have been deposited in DDBJ/EMBL/GenBank under the accession numbers LXIB000000000 for strain CITA 14 and LXKK000000000 for strain CITA 124. The versions described in this paper are the first versions, LXIB01000000 and LXKK01000000.

2.4.4 Acknowledgments

This work was supported financially by the Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA) project RTA2014-00018-CO2-01. J. Garita-Cambronero held a Ph.D. fellowship from the Spanish Government (Ministerio de Educación, Cultura y Deporte; fellowship FPU12/01000).

Capítulo 3: Comparación genómica y caracterización fenotípica de cepas patógenas y no patógenas de *X. arboricola* revela aspectos relacionados con el proceso de infección de la mancha bacteriana de los frutales de hueso.

Comparative genomic and phenotypic characterization of pathogenic and non-pathogenic strains of *Xanthomonas arboricola* reveals insights into the infection process of bacterial spot disease of stone fruits.

J. Garita-Cambronero¹, A. Palacio-Bielsa², M. M. López³, J. Cubero¹.

¹Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA), Madrid, Spain. ²Centro de Investigación y Tecnología Agroalimentaria de Aragón, Zaragoza, Spain. ³Instituto Valenciano de Investigaciones Agrarias, Valencia, Spain.

PLoS ONE 11(8): e0161977. doi:10.1371/journal.pone.0161977.

3.1 Abstract

Xanthomonas arboricola pv. *pruni* is the causal agent of bacterial spot disease of stone fruits, a quarantinable pathogen in several areas worldwide, including the European Union. In order to develop efficient control methods for this disease, it is necessary to improve the understanding of the key determinants associated with host restriction, colonization and the development of pathogenesis. After an initial characterization, by multilocus sequence analysis, of 15 strains of *X. arboricola* isolated from *Prunus*, one strain did not group into the pathovar *pruni* or into other pathovars of this species and therefore it was identified and defined as a *X. arboricola* pv. *pruni* look-a-like. This non-pathogenic strain and two typical strains of *X. arboricola* pv. *pruni* were selected for a whole genome and phenotype comparative analysis in features associated with the pathogenesis process in *Xanthomonas*. Comparative analysis among these bacterial strains isolated from *Prunus* spp. and the inclusion of 15 publicly available genome sequences from other pathogenic and non-pathogenic strains of *X. arboricola* revealed variations in the phenotype associated with variations in the profiles of TonB-dependent transporters, sensors of the two-component regulatory system, methyl accepting chemotaxis proteins, components of the flagella and the type IV pilus, as well as in the repertoire of cell-wall degrading enzymes and the components of the type III secretion system and related effectors. These variations provide a global overview of those mechanisms that could be associated with the development of bacterial spot disease. Additionally, it pointed out some features that might influence the host specificity and the variable virulence observed in *X. arboricola*.

3.2 Introduction

Xanthomonas arboricola (Vauterin *et al.*, 1995) is a species of Gram negative, rod-shaped bacteria exclusively associated with plants. Most of the strains of this species cause diseases on several herbaceous and woody plants of agricultural interest. Beside these, some other strains have been identified as non-pathogenic, saprophytic or opportunistic pathogens. Based on the host specialization of the pathogenic strains, nine pathovars have been recently proposed (Fischer-Le Saux *et al.*, 2015).

X. arboricola pv. *pruni*, causal agent of bacterial spot disease on at least 13 species of the genus *Prunus*, is considered one of the most economically important pathogens within *X. arboricola*. This pathogen, which is classified as a quarantinable organism in the European Union, causes damages on leaves, fruits, twigs, branches and trunks of the trees (EFSA, 2014). Damages in 25-75% of the peach fruits in orchards in the USA have been reported (Stefani, 2010).

Previous molecular typing analysis of a selection of strains isolated from *Prunus* has demonstrated the low diversity on *X. arboricola* pv. *pruni*, which forms a monophyletic group comprised of a unique clonal complex regardless of the host, the continent or the year of isolation, corresponding to a pandemic lineage that stays as a monomorphic group during evolution (Boudon *et al.*, 2005; Fischer-Le Saux *et al.*, 2015).

Next generation sequencing approaches have provided valuable information related to the genomics of the genus *Xanthomonas*, allowing understanding of the role of several pathogenic factors as well as the key determinants of bacterial adaptation and host restriction (Denancé *et al.*, 2016). Since 2012, genome sequencing projects have started for *X. arboricola*, generating at least one complete genome and 15 draft genome sequences, comprising meaningful information for understanding pathogenesis in this species and for the improvement of diagnostic tools addressed to disease prevention. Comparative genomic studies in *X. arboricola* are attracting increased interest in this species and have started to bring important results. For instance, in a study on *X. arboricola* pv. *juglandis*, causal agent of bacterial blight on walnut (*Juglans* spp.), a group of non-pathogenic strains isolated from *J. regia* were identified and classified as phylogenetically distant from the pathogenic strains. Complete genome analysis of these atypical strains revealed differences from pathogenic strains in features related to the initial steps of bacterial infection, such as those connected to structural components of the flagellar system, non-fimbrial adhesins and chemosensors. Furthermore, the profile of type III effectors was correlated with the capacity to produce disease on walnut (Essakhi *et al.*, 2015; Cesbron *et al.*, 2015).

Previous studies of plant-pathogen interaction in other models such as *X. citri* (Vernière *et al.*, 2003; Gordon *et al.*, 2015), *X. campestris* (Kamoun & Kado, 1990; da Silva *et al.*, 2002; Crossman & Dow, 2004) and *X. oryzae* (Qian *et al.*, 2008; Triplett *et al.*, 2011), have demonstrated that the study of phenotypic and genotypic features associated with

bacterial sensing and attachment, chemotaxis, motility, xanthan production, biofilm organization, the metabolism of carbon sources, as well as secretion of virulence factors, were basic for improving the knowledge of the bacterial ability to penetrate the plant tissue and to cause disease in a specific host range of plants.

The general genome features of two *X. arboricola* pv. *pruni* pathogenic strains, isolated from almond (*Prunus amygdalus*, syn. *P. dulcis*) and Japanese plum (*Prunus salicina*), as well as one of a non-pathogenic strain isolated from Santa Lucía SL-64 rootstock (*Prunus mahaleb*) were described in a previous work (Garita-Cambronero *et al.*, 2014, 2016a). Herein, our goal was to associate the genomic content of these xanthomonads with the phenotypic features of *X. arboricola* pv. *pruni* that causes bacterial spot of stone fruits. Differences in key aspects associated with bacterial sensing, motility, attachment and secretion of cell-wall degrading enzymes as well as type III effectors were determined in pathogenic and non-pathogenic strains of *X. arboricola*.

3.3 Results

3.3.1 Molecular characterization using multilocus sequence analysis (MLSA), genome based phylogeny and genome comparison.

As an initial approach to characterize a group of strains phenotypically similar to *Xanthomonas* isolated from *Prunus* in Spain, one *Xanthomonas* strain isolated from asymptomatic leaves of *P. mahaleb* (strain CITA 44), and 14 *Xanthomonas* strains isolated from *Prunus* spp. with symptoms of bacterial spot disease, were typed at four loci (*dnaK*, *fyuA*, *gyrB* and *rpoD*) in order to determine their taxonomic position (Young *et al.*, 2008) (Table 1).

Analysis of the concatenated sequence (2,858 nucleotide positions) revealed a percent of similarity from 96.86 to 100% within the group encompassed by six *X. arboricola* reference strains and the 15 strains isolated from *Prunus* spp. Despite this, strain CITA 44, according to the maximum likelihood analysis, could not be consistently associated with any of the reference strains (Fig 1). The remaining 14 strains isolated from symptomatic hosts were clustered into a group with all the reference strains of *X.*

arboricola pv. *pruni*. Comparative analysis between these two intra-specific groups revealed a total of 43 nucleotide variations in the concatenated sequence of CITA 44 (11 variable nucleotide sites for *dnaK*, nine for *fyuA*, 16 for *gyrB* and seven for *rpoD* genes), or a sequence similarity of 98.49% between this strain and all the other *X. arboricola* isolated from *Prunus*. These nucleotide changes could be associated with silent mutations, due to the fact that the translated amino acid sequence of each locus showed a 100% of sequence coverage and 99-100% of sequence identity to several strains of *X. arboricola* according to the Blastp results.

Table 1. *Xanthomonas arboricola* strains used in this study.

Strain ^a	Pathovar	Host	Year of isolation	Origin Country (province)
<i>X. arboricola</i>				
CITA 44		<i>Prunus mahaleb</i> (rootstock Santa Lucía SL-64)	2009	Spain (Aragón)
ICMP 1488 [*]	<i>celebensis</i>	<i>Musa acuminata</i>	1960	New Zealand
CFBP 1846 [*]	<i>corylina</i>	<i>Corylus avellana</i>	1975	France
ICMP 5726 [*]	<i>corylina</i>	<i>Corylus maxima</i>	1939	United States
ICMP 35 [*]	<i>juglandis</i>	<i>Juglans regia</i>	1956	New Zealand
IVIA 2113 [*]	<i>juglandis</i>	<i>Juglans regia</i>	2010	Spain (Badajoz)
CFBP 3123 ^{PT*}	<i>populi</i>	<i>Populus x euramericana</i> cv Robusta	1979	Netherlands
10.0343 [*]	<i>pruni</i>	<i>Prunus laurocerasus</i> cv. Caucasica	2010	Netherlands
10.0400 [*]	<i>pruni</i>	<i>Prunus laurocerasus</i> cv. Grüner Teppich	2010	Netherlands
10.0439 [*]	<i>pruni</i>	<i>Prunus laurocerasus</i> cv. Zabeliana	2010	Netherlands
CFBP 3894 ^{PT*}	<i>pruni</i>	<i>Prunus salicina</i>	1953	New Zealand
CFBP 5530 [*]	<i>pruni</i>	<i>Prunus persica</i>	1989	Italy
CFBP 5724 [*]	<i>pruni</i>	<i>Prunus amygdalus</i>	Unknown	United States
CITA 4	<i>pruni</i>	<i>Prunus persica</i> cv. Catherine	2004	Spain (Valencia)
CITA 9	<i>pruni</i>	<i>Prunus persica</i> cv. Merrill O'Henry	2008	Spain (Aragón)
CITA 11	<i>pruni</i>	<i>Prunus persica</i> cv. Richard Lady	2008	Spain (Aragón)
CITA 33	<i>pruni</i>	<i>Prunus amygdalus</i> cv. Guara	2009	Spain (Aragón)
CITA 46	<i>pruni</i>	<i>Prunus persica</i> cv. Summer Lady	2009	Spain (Navarra)
CITA 70 ^{**}	<i>pruni</i>	<i>P. persica</i> x <i>P. amygdalus</i> (rootstock Garnem (GxN15)	2010	Italy
IVIA 2626.1	<i>pruni</i>	<i>Prunus salicina</i> cv. Fortuna	2002	Spain (Badajoz)
IVIA 2647-1-2	<i>pruni</i>	<i>Prunus salicina</i> cv. Larry Ann	2002	Spain (Badajoz)
IVIA 2647-3-1	<i>pruni</i>	<i>Prunus salicina</i> cv. Friar	2002	Spain (Badajoz)
IVIA 2826-1	<i>pruni</i>	<i>Prunus salicina</i> cv. Anne Gold	2003	Spain (Valencia)
IVIA 2832-10	<i>pruni</i>	<i>Prunus salicina</i> cv. Angeleno	2003	Spain (Valencia)
IVIA 3161-2	<i>pruni</i>	<i>Prunus amygdalus</i> cv. Rumbeta	2006	Spain (Alicante)
IVIA 3162	<i>pruni</i>	<i>Prunus amygdalus</i> cv. Rumbeta	2006	Spain (Alicante)
IVIA 3487-1	<i>pruni</i>	<i>Prunus salicina</i>	2008	Spain (Aragón)

^a CITA, Centro de Investigación y Tecnología Agroalimentaria de Aragón, Zaragoza, Spain; ICMP, International Collection of Microorganisms from Plants, Auckland, New Zealand; CFBP, Collection Française de Bactéries Phytopathogènes, Angers, France; IVIA, Instituto Valenciano de Investigaciones Agrarias, Valencia, Spain.

^{PT} Pathotype strains.

^{*} Strains characterized previously as *Xanthomonas arboricola* and used in this study as control during the molecular characterization.

^{**} Interception in a Spanish nursery (Aragón) of plants produced in an Italian nursery.

[†] Sequences from these type strains were obtained from a previous work [18], and used as reference for the MLSA in this paper.

Fig 1. Maximum likelihood tree of concatenated nucleotide sequences for partial sequences of the genes *dnaK*, *fyuA*, *gyrB* and *rpoD* of selected strains of *X. arboricola* isolated from *Prunus*. Target strains are in bold. For comparative purposes other *X. arboricola* strains are included. *X. citri* subsp. *citri* strain ICMP 24 was considered as outgroup. Bootstrap values (1,000 replicates) are indicated below and above the branches.

Fig 2. Phylogenetic analysis of 18 strains of *X. arboricola* based on the core genome sequence of the species (2,525 CDS). Sequences were aligned using MAFFT and Maximum likelihood analysis was carried out using RaxML. Bootstrap values (1,000 replicates) are presented above or below the branches. The strain of *X. campestris* pv. *campestris* ATCC 33913 was used as an outgroup.

Phylogenetic analysis of the 18 *X. arboricola* strains, based on the core genome sequence (S1 and S2 Tables), showed that the three strains that cause disease on stone fruits and almond (CITA 33, IVIA 2626.1 and MAFF 301420) clustered according to the pathovar classification, as observed previously using the multilocus sequence analysis (MLSA) approach. In the same manner, strain CITA 44 did not cluster with the other xanthomonads isolated from *Prunus* nor to any of the well-established pathovars described for *X. arboricola*. Instead, CITA 44 was included in a clade along with the strain 3004 of *X. arboricola*, recently described as the causal agent of disease on barley (*Hordeum vulgare*) (Ignatov *et al.*, 2015) (Fig 2).

After automatic annotation of 18 genome sequences publicly available of *X. arboricola* (Ibarra Caballero *et al.*, 2013; Garita-Cambronero *et al.*, 2014, 2016a; Higuera *et al.*, 2015; Ignatov *et al.*, 2015; Pereira *et al.*, 2015; Cesbron *et al.*, 2015; Harrison *et al.*, 2016), 3,927 protein coding sequences (CDS) were predicted for CITA 44, whereas in the draft genome sequence of the other *X. arboricola* strains that did not represent a well-established pathovar (strains 3004, CFBP 7634 and CFBP 7651), a total of 5,221 different CDS were determined. For the strains of the pathovars *celebensis*, *juglandis* and *pruni*, 4,485; 5,700 and 5,048 CDS were predicted, respectively (S1 Table). The core genome sequence of the analyzed strains was comprised by at least 2,525 CDS (S2 Table). CITA 44 shared 3,387 CDS with the three strains of the pathovar *pruni* (Fig 3A, S1 Fig) and 3,720 CDS with all the strains of *X. arboricola* (Fig 3B). This strain shared the highest number of CDS with strain 3004 of *X. arboricola* (3,485 CDS), not classified in any pathovar, whereas the lowest number of CDS was shared with the strain MAFF 301420 of the pathovar *pruni* (3,181 CDS). Unique CDS for strains CITA 44, CITA 33 and IVIA 2626.1 were predicted. CITA 44 showed 206 exclusive CDS, 55 of which were classified into 18 clusters of orthologous groups (COG) functional categories, being predominantly those related to cell motility and carbohydrate transport and metabolism. Signal peptide cleavage sites and transmembrane helices were predicted only in 19 and 33 CDS, respectively. In the case of the *X. arboricola* pv. *pruni* strains, 149 unique CDS were found, 54 of them were classified into 20 COG functional categories, with a predominant presence of CDS related to cell wall/membrane biogenesis and amino acids transport and metabolism. Additionally, 17 CDS showed signal peptide cleavage sites and transmembrane helices were predicted in 33 CDS (Fig 3C, S3 Table).

Fig 3. Venn diagram representing the CDS shared among *X. arboricola*. CDS shared by *Prunus*-pathogenic strains and CITA 44 (A), CDS shared by CITA 44 and *X. arboricola* (B). The distribution of the unique CDS found in the draft genome sequences of CITA 44, CITA 33 and IVIA 2626.1 that have been classified into the COGs functional categories are also represented (C). Different letters represent the standardized code of the COGs functional categories.

3.3.2 Carbon sources utilization and chemotaxis profile.

The profile of carbon sources metabolized by CITA 44 was determined using the BIOLOG GN2 microplate system and compared to that obtained from the strains of the pathovars *corylina*, *juglandis*, *populi* and *pruni* (Table 1). Seventeen carbon sources compounds over 95 were utilized by CITA 44, meanwhile 45 carbon sources were not utilized and 33 showed variable reactions and, consequently, were considered as not informative. On the other hand, the profile observed in 20 strains isolated from *Prunus* and identified as *X. arboricola* pv. *pruni*, as well as the patterns observed in the remaining strains of *X. arboricola*, were different compared to CITA 44, as represented in the dendrogram obtained from the similarity analysis (Fig 4). All the strains of the

pathovar *pruni*, unlike CITA 44, were able to metabolize dextrin and proline and unable to metabolize D-saccharic acid.

In order to determine the chemotactic effect of 18 carbon compounds on CITA 44 and on two reference strains of *X. arboricola* pv. *pruni*, a microtiter plate chemotaxis assay was conducted (Sena-Vélez, 2015). Strain CITA 44 was attracted only by serine (200 mM), and repelled by citric acid (10 mM), galacturonic acid (10 mM), glucuronic acid (10 mM), arginine (10 mM), cysteine (10 mM) and galactose (10 mM). No chemotactic activity was recorded toward the remaining compounds assayed. Strains CFBP 5530 and CITA 33 showed a different chemotactic pattern compared to CITA 44. The first two strains were attracted to glycerol (2%), maltose (10 mM) and serine (10 mM and 200 mM), and repelled by citric acid (10 mM), galacturonic acid (10 mM), glucuronic acid (10 mM) and cysteine (10 mM). Detailed chemotactic patterns for each strain are shown in Table 2.

Fig 4. Carbon compounds utilization profile of *X. arboricola*. Biolog GN2 profile for CITA 44 (A), comparative cluster analysis including carbon compounds profile from *X. arboricola* isolates (B). Cluster analysis was constructed based on 63 informative substrates. Data were computed using the UPGMA model. Reliability of the tree was determined by the cophenetic correlation index ($r > 0.98$).

Table 2. Chemotactic profile of 18 carbon compounds of three representative strains of *Xanthomonas arboricola* isolated from *Prunus* spp.

Carbon compounds	CFBP 5530	CITA 33	CITA 44
Alanine 10 mM	+	0	0
Alanine 250 mM	+	0	0
Arginine 10 mM	0	-	-
Arginine 100 mM	+	0	0
Citric acid 10 mM	-	-	-
Cysteine 10 mM	-	-	-
Fructose 10 mM	+	0	0
Galactose 10 mM	+	0	-
Galacturonic acid 10 mM	-	-	-
Glucose 10 mM	0	0	0
Glucuronic acid 10 mM	-	-	-
Glycerol 0.2%	0	0	0
Glycerol 2%	+	+	0
Leucine 10 mM	0	0	0
Leucine 150 mM	0	+	0
Maltose 10 mM	+	+	0
Mannitol 0.2%	0	0	0
Serine 10 mM	+	+	0
Serine 200 mM	+	+	+
Sodium citrate 10 mM	+	0	0
Succinic acid 10 mM	0	0	0
Sucrose 10 mM	0	+	0
Xylose 10 mM	0	0	0

+, Chemotactic activity
 -, Chemorepellent activity
 0, No chemotactic activity

In addition to the phenotypic variants observed in the features mentioned above, genome comparison also revealed variants in the profile of the environmental receptors studied within this species. *X. arboricola* strains harbored 14 of the 28 TonB-dependent transporters (TBDTs) analyzed (Mhedbi-Hajri *et al.*, 2011) (S4 Table), and from these, only two homologs to the TBDTs encoded by the loci XAC3620 (outer membrane receptor FepA) and XCC3595 (ferric pseudobactin receptor), described in *X. citri* and *X. campestris*, respectively, were shared by all the *X. arboricola* strains.

All strains isolated from *Prunus* spp., both pathogenic and non-pathogenic, harbored eight TBDTs. The presence of homologous sequences to the TonB-dependent receptor loci XCC1719 and XAC3077, and the absence of XCC0304 and XCC2867 in the strains of *X. arboricola* pv. *pruni*, differentiated them from the non-pathogenic strain CITA 44. The distribution of the TBDTs revealed that none of these variations was unique for those strains that inhabit the *Prunus* hosts. Nevertheless, the TBDT *cirA* (XAC3077) was found only in those strains that caused disease on stone fruits and Turkish hazel (*Corylus colurna*).

Regarding the distribution of sensors of two-component regulatory system (STCRS) (Mhedbi-Hajri *et al.*, 2011), 58 orthologous CDS of 86 genes previously described in *Xanthomonas* were found in *X. arboricola*. All the analyzed strains shared 44 STCRS (S4 Table), and CITA 33, CITA 44 and IVIA 2626.1 have 53 STCRS in common. Homologous CDS to a diguanylate cyclase of *X. citri* (XAC2804) was only present in the pathogenic strains, meanwhile homologues of four STCRS (XAC1819, XAC2804, XAC0136 and XAC1345) were only found in CITA 44. Additionally, as compared to other *X. arboricola* strains, CITA 33 and IVIA 2626.1 presented two unique STCRS homologous to the loci XAC0136 and XAC1345 of *X. citri*, respectively (Fig 5).

Out of 26 methyl-accepting chemotaxis proteins (MCPs) previously described (Mhedbi-Hajri *et al.*, 2011), 22 were found in the genome sequences of *X. arboricola* (S4 Table). The number of MCPs varied from 16 in strain CFBP 7634 to 22 in strains 3004, CFBP 7651 and NCPPB 1630; all the *X. arboricola* strains shared a MCPs pattern composed by 13 genes (S4 Table). *Prunus*-associated strains shared 20 MCPs, pathogenic strains differed from CITA 44 in the existence of a CDS orthologous to the MCP XCV1933 and the absence of the MCP XCV1938, both described in *X. campestris* pv. *vesicatoria*. No MCPs could be detected as specific for a particular host plant, however, orthologous CDS for three MCPs described in *X. citri* and *X. campestris* pv. *vesicatoria* (XAC3768, XCV1952 and XCV1938) were absent only in the strain that causes disease on Turkish hazel (S4 Table).

X. arboricola genome sequences harbored 15 to 17 genes homologous to the chemotactic related *che* genes (Mhedbi-Hajri *et al.*, 2011) (S4 Table). Pathogenic strains of *Prunus* had 17 homologous CDS for these genes, and CITA 44 had 16 CDS. CITA 44 did not show homologous CDS to the locus XAC2447 (*cheW*) described in *X. citri*, which was present in all the remaining genome sequences of *X. arboricola* (Fig 5).

3.3.3 Flagella, fimbrial and non-fimbrial adhesins associated with motility and attachment in *X. arboricola*.

Swarming motility was evaluated in all the 21 *X. arboricola* strains isolated from *Prunus* (Table 3). Two kinds of swarmer colony phenotype were observed. Dendritic pattern, encompassed by several tendrils that extended away from a central colony, was observed in CITA 44 as well as in other 13 strains of the pathovar *pruni* after 12-15

hours post inoculation (hpi). Despite this, CITA 44 showed some unique characteristic such as the presence of shorter and wider tendrils in the swarming colony (Fig 6A). The second colony morphology was characterized by a circular shape shown by seven strains of *X. arboricola* pv. *pruni* (Fig 6A). Light and electron microscopy observations discarded the existence of hyper-flagellated bacterial cells in the swarming colony for all strains, including CITA 44 (S2 Fig).

Fig 5. Graphical circular representation of the CDS associated with pathogenesis in *X. arboricola* isolated from *Prunus*. The contigs of the draft genome sequence of strains CITA 33, CITA 44 and IVIA 2626.1 were arranged by Mauve (Darling *et al.*, 2004), using the genome sequence of *X. arboricola* pv. *juglandis* Xaj 417 as the reference. The circular map was constructed using CGview. From outside to center: CDS on forward strand, CDS on reverse strand, GC content and GC skew.

Table 3. Motility profiles of 21 strains of *Xanthomonas arboricola* on semisolid surfaces.

Strain	Pathovar	Swarming ^a	Surfactant activity ^b	Swimming ^c
100.343	<i>pruni</i>	17.15±6.22*	3.06±0.42	0.58±0.10
100.400	<i>pruni</i>	4.91±2.82*	3.64±0.74	0.55±0.05
100.439	<i>pruni</i>	30.25±14.84*	1.63±0.42	0.55±0.05
CFBP 3894	<i>pruni</i>	4.30±1.14°	1.59±0.48	0.31±0.03
CFBP 5530	<i>pruni</i>	3.32±0.66°	2.87±0.32	0.65±0.13
CFBP 5724	<i>pruni</i>	4.78±1.70°	2.18±0.33	1.25±0.05
CITA 4	<i>pruni</i>	13.47±7.04*	6.62±0.93	0.50±0.05
CITA 9	<i>pruni</i>	17.78±7.66*	1.94±0.56	0.65±0.05
CITA 11	<i>pruni</i>	4.24±1.54*	4.22±1.22	0.55±0.05
CITA 33	<i>pruni</i>	18.72±9.62*	2.80±0.66	0.60±0.05
CITA 44	-	13.42±6.02*	3.69±0.54	2.42±0.10
CITA 46	<i>pruni</i>	22.16±11.91*	8.54±0.16	0.50±0.05
CITA 70	<i>pruni</i>	15.30±4.33*	5.44±0.79	0.52±0.08
IVIA 2626.1	<i>pruni</i>	2.75±0.83°	4.44±0.76	0.30±0.05
IVIA 2647-1-2	<i>pruni</i>	2.32±0.42°	3.94±0.78	0.55±0.05
IVIA 2647-3-1	<i>pruni</i>	23.55±14.67*	3.91±0.64	0.40±0.05
IVIA 2826-1	<i>pruni</i>	17.98±8.23*	3.55±0.45	0.45±0.05
IVIA 2832-10	<i>pruni</i>	21.45±7.41*	2.94±0.52	0.55±0.05
IVIA 3161-2	<i>pruni</i>	3.15±1.04°	2.97±0.73	0.45±0.05
IVIA 3162	<i>pruni</i>	2.71±0.55°	0.52±0.10	0.78±0.08
IVIA 3487-1	<i>pruni</i>	5.89±2.39*	11.98±0.89	0.50±0.00

^a Mean ± SD of the area (cm²) covered by the swarmer colony on PYM 0.5% agar plates after 24 hpi

* Dendritic swarming phenotypes shown by the analyzed strains.

° Circular swarming phenotypes shown by the analyzed strains.

^b Mean ± SD of the ratio between the radius of the bright halo and the area of the colony observed 24 hpi after atomizing mineral oil over the bacterial colony.

^c Mean ± SD of the radius (cm) of the swimming colony after 72 hpi on MMA 0.3% agar plates.

Fig 6. Flagella and type IV pilus mediated motility on semisolid and solid surfaces in some representative strains of *Xanthomonas arboricola*. Swarming dendritic (A1 and A2) and circular (A3) phenotypes were determined on PYM 0.5% agar plates after 24 hpi. Variable swimming phenotypes (B1, B2 and B3) observed after 72 hpi on MMA 0.3% agar plates. Twitching type motility (C1, C2 and C3) was observed after 72 hpi on polystyrene plate surface after inoculation through PYM 1.5% agar (40X). Each motility assay was conducted in three independent experiments with three replicates.

Surfactant activity was evaluated for the 21 strains mentioned above according to the atomized oil assay (Burch *et al.*, 2010). A bright halo, associated with a change in the surface tension, was observed around the colonies of all the strains tested. Surfactant production, estimated as the ratio between the radius of the bright halo and the area of the colony, was variable among strains (Table 3). CITA 44 showed a mean ratio of 3.69 ± 0.54 (mean $\pm$ standard deviation), meanwhile strains of the pathovar *pruni* showed a mean ratio of 3.94 ± 2.62 . Strains IVIA 3162 and IVIA 3847–1 showed mean ratios of 0.52 ± 0.10 and 11.98 ± 0.89 , being the lowest and the highest activity identified, respectively (Table 3).

The swimming ability of these strains was also assayed on 0.3% agar MMA plates. After inoculation, all the strains showed a circular and turbid growth around the inoculating point (Fig 6B). This activity, recorded as the radius of the bacterial colony, was variable among the strains (Table 3). CITA 44 presented the highest activity with an average of the radius of the halo of 2.42 ± 0.10 , meanwhile the mean for pathogenic strains was 0.65 ± 0.46 , with a maximum value of 1.25 ± 0.05 for CFBP 5724 and a minimum value of 0.30 ± 0.05 for IVIA 2626.1.

Twitching motility, which is carried out by type IV pilus (Mattick, 2002), was observed in all the strains on the plastic surface of the culture plate, after crystal violet staining (Fig 6C). Microscopic observation at the edge of the colony revealed a twitching zone

composed by bacterial rafts moving away from the colony, creating a concentric pattern preceded by a lattice like network and microcolonies.

Genome analysis performed to characterize the presence of the major structural components of the two motility and adherence structures described above (Chevance & Hughes, 2008; Mhedbi-Hajri *et al.*, 2011), revealed that 33 (*X. arboricola* 3004 and *X. arboricola* pv. *corylina*) to 35 (*X. arboricola* pv. *pruni* and *X. arboricola* CFBP 7634 and CFBP 7651) flagellar components were found in *X. arboricola*. The genomes in these species shared 29 CDS (S4 Table). The strains isolated from *Prunus* presented 34 (CITA 44) and 35 (CITA 33 and IVIA 2626.1) structural components; CITA 44 did not show homologous sequence to *fliD* (XAC1974), which is present in all the analyzed strains (Fig 5, S4 Table). Despite the similar pattern of flagellar components observed among the *X. arboricola* strains, variants in the amino acid sequence of the flagellin protein were observed between pathogenic strains (CFBP 7179, CITA 33, IVIA 2626.1, MAFF 301420 and NCCB 100457), which showed the identical protein WP_039814449.1, and the non-pathogenic or low-pathogenic strains (3004, CFBP 7634, CFBP 7651, NCPPB 1630 and CITA 44), which showed the identical protein WP_024939608.1. Both proteins showed 354 identical amino acids of a total of 399 (pairwise identity of 88.7%). A remarkable difference, associated with the potential to infect in other xanthomonads, is the substitution of aspartic acid by valine in the amino acid position 43 of the N-terminal region of the flagellin gene in the pathovars *corylina*, *juglandis* and *pruni* (S3 Fig).

Regarding the structural and regulatory components of the type IV pilus, 22 out of 31 CDS homologous to the components described in *X. citri* subsp. *citri* strain 306 (Dunger *et al.*, 2016) were found in the nine genome sequences of *X. arboricola*; 16 of these CDS were shared by all the strains (S4 Table). The other *Prunus*-associated strains presented the same components, including CITA 44. The remaining strains of *X. arboricola* showed a similar pattern, with the exception of NCPPB 1630, CFBP 7634 and CFBP 7651, which did not have homologous sequence to the locus XAC0259 of *X. citri*. No sequence associated with *pilF* was observed in strain CFBP 7634 and none homologues of *pilQ* was found in strain 3004 (S4 Table).

Regarding the presence of the genes associated with type IV pilus, that are present in most of the *Xanthomonas* species, the cluster *pilB*, *C*, *D*, *R*, *S* was found in all the

strains, as well as the cluster that encodes the minor pilins, *pilE*, *V*, *W*, *X*, *Y1* and *fimT*, that showed a sequence identity lower than 80.0%. Presence of *pilE*, *pilV*, *pilY1* and *fimT* was variable among the strains, meanwhile *pilW* and *pilX* were found in all the genome sequences. Prepilin *pilA* was found in all the strains of *X. arboricola*, despite its identity to the locus XAC3505 of *X. citri* was lower than 80.0% (S4 Table).

In addition to the fimbrial adhesins described above, the repertoire of non-fimbrial adhesins (Mhedbi-Hajri *et al.*, 2011) of *X. arboricola* comprised five genes (S4 Table). All the strains shared homologous CDS to *panB* (XAC1816), *yapH* (XAC2151) and *xadA* (XCV3670) of *X. citri* and *X. campestris* pv. *vesicatoria*. *Prunus*-pathogenic strains harbored an adhesin homologous to the locus XAC3672 of *X. citri* which was absent in CITA44 (Fig 5; S4 Table). Finally, the hemagglutinin encoded by the locus XAC444 of *X. citri* was found only in those strains that were able to colonize hosts of the genus *Corylus*, *Juglans* and *Prunus*.

3.3.4 Pathogenicity tests and genomic components associated with late stages of infection in *Prunus*-associated strains.

Detached leaf assay was carried out by inoculating 21 *X. arboricola* strains, isolated from *Prunus* spp., on almond (cv. Ferraduel), apricot (cv. Canino), peach (cv. Calanda) and European plum (cv. Golden Japan). CITA 44 did not cause bacterial spot disease symptoms on almond, apricot, peach or plum 28 days post inoculation (dpi). Significant differences ($p < 0.05$) in the virulence among strains of the pathovar *pruni* were shown (Table 4). These strains were able to induce disease symptoms on almond, peach and European plum but none of them caused clear disease symptoms on apricot. Representatives of this assay are shown in Fig 7.

The secretory pathway components, and plant cell wall-degrading enzymes, were also analyzed in *X. arboricola* genomes (S4 Table). Regarding the encoding elements for the two type II secretory systems (T2SS) described in *Xanthomonas* (Filloux, 2004; Szczesny *et al.*, 2010), all strains harbored the 11 genes of the *xps* cluster but, in the case of the *xcs* cluster, only CDS homologous to *xcsD*, *xcsE*, *xcsF*, *xcsG* and *xcsJ* were found. The remaining seven components of this secretory system were automatically annotated, but amino acid sequence analysis showed for all of them an identity lower than 80% when compared with the *xcs* cluster described in *X. campestris* pv. *campestris* (Szczesny *et al.*, 2010).

Table 4. Differences in the virulence of 24 *Xanthomonas arboricola* strains isolated from *Prunus* on detached leaves of almond, apricot, peach and European plum.

Strain	Pathovar	Almond ^a	Apricot	Peach	European plum
10.0343	<i>pruni</i>	69.21 ± 12.97 _b DEF	0 _a A	79.27 ± 14.13 _b DE	65.00 ± 20.94 _b C
10.0439	<i>pruni</i>	55.04 ± 13.21 _b BCDEF	0 _a A	69.09 ± 35.54 _b CDE	65.23 ± 13.10 _b C
10.0400	<i>pruni</i>	60.59 ± 7.17 _b CDEF	0 _a A	62.82 ± 26.55 _b CDE	57.35 ± 26.97 _b BC
CFBP 3894	<i>pruni</i>	54.11 ± 13.57 _c BCDEF	0 _a A	32.94 ± 29.19 _{bc} BC	24.12 ± 33.64 _{ab} ABC
CFBP 5530	<i>pruni</i>	24.74 ± 20.37 _b ABC	0 _a A	76.70 ± 14.45 _c CDE	57.23 ± 22.96 _c BC
CFBP 5724	<i>pruni</i>	61.65 ± 15.72 _b CDEF	0 _a A	60.23 ± 16.91 _b CDE	46.38 ± 29.31 _b BC
CITA 4	<i>pruni</i>	36.36 ± 27.88 _b BCDE	0 _a A	52.80 ± 34.04 _b BCDE	40.92 ± 31.23 _b BC
CITA 9	<i>pruni</i>	63.09 ± 13.47 _b DEF	0 _a A	50.28 ± 32.30 _b BCDE	42.36 ± 19.17 _b BC
CITA 11	<i>pruni</i>	60.52 ± 18.77 _b CDEF	0 _a A	48.18 ± 29.78 _b BCDE	44.95 ± 20.58 _b BC
CITA 33	<i>pruni</i>	56.30 ± 15.37 _b BCDEF	0 _a A	72.96 ± 22.19 _b CDE	56.87 ± 25.41 _b BC
CITA 44	-	0 _a A	0 _a A	0 _a A	0 _a A
CITA 46	<i>pruni</i>	65.93 ± 11.94 _c DEF	0 _a A	77.90 ± 20.09 _c DE	39.55 ± 22.71 _b BC
CITA 70	<i>pruni</i>	61.62 ± 14.56 _b CDEF	0 _a A	42.99 ± 28.09 _b BCDE	50.83 ± 20.71 _b BC
IVIA 2647-1-2	<i>pruni</i>	51.74 ± 15.80 _b BCDEF	0 _a A	73.05 ± 19.87 _b CDE	55.21 ± 24.10 _b BC
IVIA 2647-3-1	<i>pruni</i>	74.41 ± 22.75 _c EF	0.03 ± 0.08 _a A	69.38 ± 20.59 _c CDE	47.17 ± 18.42 _b BC
IVIA 2826-1	<i>pruni</i>	57.87 ± 16.05 _b BCDEF	0.60 ± 1.67 _a A	60.52 ± 26.40 _c CDE	33.11 ± 18.42 _b ABC
IVIA 2832-10	<i>pruni</i>	65.07 ± 16.01 _b DEF	0 _a A	81.49 ± 13.81 _c E	53.32 ± 14.05 _b BC
IVIA 3161-2	<i>pruni</i>	71.72 ± 16.47 _b DEF	0 _a A	75.29 ± 18.70 _b CDE	40.92 ± 33.80 _b BC
IVIA 3162	<i>pruni</i>	54.62 ± 14.67 _{bc} BCDEF	0 _a A	70.60 ± 28.90 _c CDE	41.97 ± 28.26 _b BC
IVIA 3847-1	<i>pruni</i>	52.98 ± 33.83 _b BCDEF	0 _a A	69.61 ± 12.97 _b CDE	51.80 ± 23.05 _b BC
IVIA2626.1	<i>pruni</i>	75.38 ± 15.99 _c F	0 _a A	59.80 ± 25.27 _c CDE	32.26 ± 26.27 _b ABC

^a Mean ± SD of the percentage of symptomatic leaf area 28 dpi using a cotton swab damped with bacterial inoculum.

Different capital letters indicate statistically significant differences among the assayed strains on each host. Non capital letters indicate significant differences of each strain on four assayed hosts ($p = 0.05$).

Fig 7. Schematic representation of bacterial spot symptoms observed on almond, apricot, peach and European plum 28 dpi. Detached leaves were inoculated with a sterile cotton swap damped with 1×10^8 CFU/mL of bacteria or with 10 mM MgCl₂ as negative control and maintained on 0.5% water agar plates.

Variability in the profile of type II secreted virulence factors was found in *X. arboricola* (S4 Table). A total of 11 pectolytic enzymes (da Silva *et al.*, 2002; Vandroemme *et al.*, 2013) were found in this species and only the polygalacturonase, encoded by the locus XCC3459, and the rhamnogalacturonan acetylsterase, encoded by XCC0154 in *X. campestris*, were shared by all the strains. With respect to the strains isolated from *Prunus*, they shared, in addition, one pectate lyase (XCC2815), one polygalacturonase (XCC2266) and one rhamnogalacturonase (XAC3505). CITA 44 had homologous sequences to a pectate lyase (XCC0112), a pectin methylesterase (XCC0121) and a pectinmethylesterase (XCC2265) that were absent in strains CITA 33 and IVIA 2626.1, meanwhile these two pathogenic strains had a homologous sequence to the degenerated pectate lyase of *X. citri* (XAC2373) which was absent in CITA 44 (Fig 5).

Variation in the profile of cellulolytic enzymes (da Silva *et al.*, 2002) was also found in *X. arboricola*, 14 of these degrading molecules were present in this species but only nine of them were shared by all the strains. None of the enzymes were unique in those strains isolated from *Prunus* spp. Pathogenic strains differed from CITA 44 in the presence of a homologous CDS to a beta-glucosidase (XCC1775), and in the absence of the cellulases encoded by the loci XAC3516 and XCC2387, which were present only in CITA 44 (Fig 5, S4 Table).

Moreover, 11 CDS homologous to hemicellulolytic enzymes, previously reported for *Xanthomonas* (da Silva *et al.*, 2002), were found in *X. arboricola* and eight of these were shared by all the strains (S4 Table). CITA 33 and IVIA 2626.1 presented a xylosidase which was absent in CITA 44 (Fig 5, S4 Table). Those strains of *X. arboricola* which belongs to the pathovars *corylina*, *juglandis* and *pruni* harbored a xylosidase/arabinoxidase enzyme which was absent in all the non-pathogenic strains of *X. arboricola* and in those strains with a lower pathogenic ability, such as the one described from the pathovar *celebensis*. Finally, the lipase virulence factor, LipA (Subramoni *et al.*, 2010), was found in all the genomes analyzed. The *gum* gene cluster, associated with the biosynthesis of xanthan in *Xanthomonas* (Vorhölter *et al.*, 2008), was analyzed. None of the strains presented homologous sequences to *gumG*, and CITA 44 and 3004 did not show homologous sequences to the *gumF*. (S4 Table).

Other crucial pathogenic elements analyzed were the T3SS components and the type III effectors (Bonas *et al.*, 1989; da Silva *et al.*, 2002; White *et al.*, 2009; Bogdanove *et al.*,

2011; Potnis *et al.*, 2011; Guo *et al.*, 2011; Hajri *et al.*, 2012; Vandroemme *et al.*, 2013; Li *et al.*, 2014) in the nine genome sequences of *X. arboricola* (S4 Table). All the strains harbored CDS that encoded for HpaR2, HpaS, HrpG and HrpX, which are involved in the regulation of the T3SS, and for many of the T3Es and some cell-wall degrading enzymes (Jacobs *et al.*, 2015). However, the strains 3004, CFBP 7634 and CITA 44, as described before (Ignatov *et al.*, 2015; Cesbron *et al.*, 2015; Garita-Cambronero *et al.*, 2016a), did not present the components of the T3SS that have been found in other *Xanthomonas*. Besides the master regulons mentioned above, all the remaining *X. arboricola* strains encompassed and shared 14 hypersensitive response and pathogenicity genes (*hpa2*, *hpaB*, *hrcC*, *hrcJ*, *hrcN*, *hrcR*, *hrcS*, *hrcT*, *hrcU*, *hrcV*, *hrpB1*, *hrpB2*, *hrpB5*, *hrpD6*) (S4 Table).

The T3E profile of *X. arboricola* composed of 22 virulence effectors (*avrBs2*, *avrXccA2*, *hpaA*, *hrpW*, *xopA*, *xopAF*, *xopAH*, *xopAI*, *xopAQ*, *xopB*, *xopE2*, *xopE3*, *xopF1*, *xopG*, *xopK*, *xopL*, *xopN*, *xopQ*, *xopR*, *xopV*, *xopX* and *xopZ1*). The non-pathogenic strain CITA 44 from *Prunus*, and the pathogen of barley strain 3004, did not have any T3Es. The remaining non-pathogenic strains, CFBP 7634 and CFBP 7651, harbored two (*avrBs2* and *xopR*) and six (*avrBs2*, *hpaA*, *hrpW*, *xopA*, *xopF1* and *xopR*) T3Es, respectively (Cesbron *et al.*, 2015) (S4 Table). In addition, the pathogenic strain NCPPB 1630 from *Musa* sp. showed six T3Es, the pathogenic strain CFBP 7179 from *J. regia* showed 16 T3Es and the strain NCCB 100457, that causes disease on *C. colurna*, showed 19 T3Es. *X. arboricola* pv. *pruni* strains CITA 33 from *P. amygdalus* and IVIA 2626.1 from *P. salicina* harbored 21 T3Es. All the strains, regardless of their hosts, shared six T3Es (*avrBs2*, *hpaA*, *hrpW*, *xopA*, *xopF1* and *xopR*). The effector *xopB* was unique in the strain that causes disease on *Juglans*, meanwhile *xopAQ*, which has not been previously reported in *X. arboricola*, and *xopE3*, were only found in those strains that cause disease on *Prunus*. CDS homologous to these two T3Es were specifically present in the plasmid pXap41, which is unique in the strains of the pathovar *pruni* (Pothier *et al.*, 2011c) (Fig 8).

3.4 Discussion

MLSA has been shown to be useful to characterize strains of *X. arboricola* (Boudon *et al.*, 2005; Young *et al.*, 2008; Burokiene & Pulawska, 2012; Kaluzna *et al.*, 2014; Fischer-Le Saux *et al.*, 2015). In this work, we utilized a MLSA scheme proposed by Young and collaborators (2008), which is based in the analysis of partial sequences of the genes *dnaK*, *fyuA*, *gyrB* and *rpoD*. Here, we have concluded that for 14 *X. arboricola* strains isolated from *Prunus*, the selected MLSA scheme was a good approach to characterize and discriminate the members of the pathovar *pruni* from atypical or commensal strains of *X. arboricola*.

Fig 8. Nucleotide sequence similarity among the sequence of the exclusive plasmid of *X. arboricola* pv. *pruni* pXap41 and the draft genome sequences of nine strains of *X. arboricola*. Blastn was used for the comparative sequence analysis with an expected value threshold of 0.001. Circular graphic representation was constructed using the BLAST Ring Image Generator (BRIG) tool. See legend for information related to each ring of the map.

Classification of CITA 44 as an atypical strain of *X. arboricola* was corroborated by the phylogenetic analysis conducted with 2,525 genes identified as the components of the core genome of this species. This atypical strain isolated from *Prunus* did not present the same phylogenetic origin as the pathogenic strains classified as pathovar *pruni*, but it was more similar to strain 3004 of *X. arboricola*, which did not cluster within any of the well-established pathovars of this taxon. Even though CITA 44 was closely related to 3004, this strain did not produce symptoms on inoculated barley in assays performed in our group. This result is similar to a previous work on *X. arboricola* strains isolated from walnut that were more similar to strains isolated from *Musa* sp. than to those of the pathovar *juglandis* isolated from walnut (Cesbron *et al.*, 2015).

Moreover, the study of these atypical strains, found on barley, walnut and here in *P. mahaleb*, open the discussion about the origin and evolution of pathogenicity in this species as it has been proposed previously for the genus *Xanthomonas* (Jacobs *et al.*, 2015). Based in the current data, it is not possible to determine if these non-pathogenic strains were predecessors of the pathogenic groups or the result of the loss of their pathogenicity. Further exploratory studies regarding some other *Prunus*-associated strains with other phenotypic and genotypic variants are needed to elucidate the evolutionary process of pathogenicity in this bacterial species.

In addition, this study contributes to our understanding of the diversity of *X. arboricola* associated with stone fruit trees and almond. Moreover, it also provides information to develop tools to identify and discriminate pathogenic and non-pathogenic *Xanthomonas* strains found in *Prunus* spp. This precise characterization of atypical strains of *X. arboricola* will avoid bacterial identification mistakes, as occurred sometimes in other bacterial models which implied unnecessary control measures and resulted in high economic losses (Graham & Gottwald, 1991; Jalan *et al.*, 2011).

The phylogenetic variation among strains of *X. arboricola* isolated from *Prunus* concurred with the existence of differences in several phenotypic features including pathogenicity. Additionally, genomic information generated from xanthomonads isolated from *Prunus* (Garita-Cambronero *et al.*, 2014, 2016a) and from other *X. arboricola* (Ibarra Caballero *et al.*, 2013; Higuera *et al.*, 2015; Ignatov *et al.*, 2015; Pereira *et al.*, 2015; Cesbron *et al.*, 2015; Harrison *et al.*, 2016), has been used to reveal those features which could be involved in the disease process.

Carbon source utilization profiles, as expected (Vauterin *et al.*, 1995; Boudon *et al.*, 2005; Fischer-Le Saux *et al.*, 2015), showed high homogeneity among strains of the pathovar *pruni*. Nevertheless, the profile of carbon sources utilization shown by CITA 44 was different to the one presented by the other strains as well as the one provided in the description of *X. arboricola* species (Vauterin *et al.*, 1995). This disparity from the original metabolic description of the species has been also described for some strains of other *Xanthomonas* species, such as *X. vesicatoria* (Stoyanova *et al.*, 2014) and *X. campestris* pv. *campestris* (Massomo *et al.*, 2003). The discordant profile of CITA 44 could be due to the fact that the original description of the species *X. arboricola* was based on strains of seven pathogenic pathovars which presented a high intra-pathovar homogeneity (Vauterin *et al.*, 1995; Fischer-Le Saux *et al.*, 2015) and did not consider a wider ecological diversity including the atypical non-pathogenic strains currently of great research interest.

Initial stages of bacterial adaptation and host colonization encompass a series of sensors and receptors that detect stimuli, providing to the cells the information of their biotic and abiotic environment which triggers a series of processes such as the cell motility, chemotaxis, quorum sensing, biofilm formation and many other cellular events (Qian *et al.*, 2008; Pieretti *et al.*, 2012). At this stage, a group of protein complexes denominated TBDTs, STCRS and MCPs play a crucial role (Mhedbi-Hajri *et al.*, 2011).

The repertoire of TBDTs in *X. arboricola*, which are bacterial outer membrane proteins associated with the transport of different substrates including carbohydrates (Noinaj *et al.*, 2010), was extensive compared to other species of this genus (Mhedbi-Hajri *et al.*, 2011; Vandroemme *et al.*, 2013), resulting similar to the one observed in *X. campestris* and other epiphytic xanthomonads (Mhedbi-Hajri *et al.*, 2011). Additionally, this repertoire is in accordance with the one previously observed in two not publicly available genome sequences of *X. arboricola* pv. *fragariae* (strains LMG 19145 and LMG 19146) and one of *X. arboricola* pv. *pruni* (strain LMG 25862) (Vandroemme *et al.*, 2013). This high number of TBDTs has been associated in other xanthomonads with the need for carbohydrates scavenging in variable conditions encountered in epiphytic niches (Blanvillain *et al.*, 2007).

Bacterial sensing and chemotaxis are essential components of the initial infection processes (Wadhams & Armitage, 2004). Regarding STCRS proteins, which are part of

the dominant molecular mechanism by which unicellular organisms respond to environmental stimuli, the large repertoire observed in *Xanthomonas arboricola* was similar to that observed in *X. campestris* (Mhedbi-Hajri *et al.*, 2011), and it was in accordance with the number of STCRS observed in most of the complete genome sequences of *Xanthomonas*, with the exception of those species which inhabit in restricted niches such as *X. albilineans* (Qian *et al.*, 2008; Pieretti *et al.*, 2012). Evaluation of TBDTs and STCRS content revealed differences between CITA 44 and *X. arboricola* pv. *pruni* strains, despite of the general low variability observed among the different *X. arboricola* strains evaluated.

Beside this, CITA 44, compared to strains of the pathovar *pruni*, also showed a distinct chemotactic pattern, being influenced by just a few compounds. This kind of intraspecific chemotactic variability has been also observed in other xanthomonads models like *X. citri* subsp. *citri* with dissimilar pathogenic ability (Sena-Vélez, 2015).

In order to explain these chemotactic behaviors, MCP content was evaluated. MCPs sense beneficial or toxic environmental compounds and transduce the signal to the cytoplasm by CheW protein, causing changes in the flagella direction and rotation speed, which finally guides bacterial cells to favourable environments (Wadhams & Armitage, 2004; Moreira *et al.*, 2015). Here, we have found that the core repertoire of the MCPs in nine genome sequences of *X. arboricola* included at least 13 chemoreceptors. Differences in the MCPs content were observed among the genome-sequenced strains and could be associated with their host range. Differences in MCPs content were found between CITA 44 and other xanthomonads from *Prunus* spp. Moreover, a variation in the *cheW* locus was also identified.

Although functional analyses are necessary to confirm the role of these sensors at initial stages of the infection process, our results point out that these processes and the genes involved may mark the diverse behaviour of the different strains. Moreover, the very initial stages of the bacterial-host interaction described above could trigger other molecular routes which involve components such as the flagella and the type IV pilus, that are related not only to motility on liquid or solid environments, but also with attachment to the host or the development of biofilm structures (Kearns, 2010; Malamud *et al.*, 2011; Dunger *et al.*, 2016). Motility on solid and semisolid surfaces, which are controlled by flagella or type IV pilus, was also variable among the assayed

strains of *X. arboricola* isolated from *Prunus*. This variability among different strains in swimming, swarming and twitching motility have been shown in other xanthomonads such as *X. citri* (Sena-Vélez *et al.*, 2015), as well as in *X. arboricola* strains isolated from walnut (Cesbron *et al.*, 2015). Strain CITA 44 showed the higher ability to swim that maybe connected to a more restricted niche to survive and higher requirements to locate it. Similarly, this enhanced ability to swim has been described in other bacterial models like *X. citri* subsp. *citri*, which less virulent strains were described to have a higher swimming ability (Sena-Vélez *et al.*, 2015). In relation to the surface motility in *X. arboricola*, two different colony phenotypes were observed; the circular one matched with the one observed for other species such as *X. oryzae* and *X. citri* (Shen *et al.*, 2001; Gottig *et al.*, 2009), described previously as independent of flagella and defined as sliding type motility instead of swarming (Malamud *et al.*, 2011). Contrary to this, some other strains showed dendritic swarmer colonies, which had not been previously described in *Xanthomonas* but confirmed in other bacterial species such as *Pseudomonas aeruginosa* as a real swarming type motility (Kohler *et al.*, 2000). In addition, these strains showed other swarming-related features such as the production of surfactants and a rapid outward migration (Verstraeten *et al.*, 2008). Again, CITA 44 showed a different pattern to that of *X. arboricola* pv. *pruni*, presenting a colony with an intermediate phenotype between dendritic and circular and also an intermediate surfactant production.

Genome analysis of the flagellum components in CITA 44 revealed an interesting amino acid substitution in the N-terminal region of the flagellin that was also found in other non-pathogenic xanthomonads. Previously, Cesbron and collaborators (2015) reported an amino acid polymorphism (Asp-43/Val-43) in the flagellin domain flg22 among pathogenic and non-pathogenic strains of *X. arboricola* isolated from walnut. Here we report that this variation is not only present in *X. arboricola* pv. *juglandis* strains but it is also found in the other two major pathogenic pathovars of this species, *corylina* and *pruni*. In *X. campestris*, strains with Val-43 in flg22 were not detected by the flagellin sensing 2 kinase (FLS2) of *Arabidopsis*, therefore they did not elicit the pathogen-associated molecular (PAMP)-triggered immunity and, additionally, these strains were more virulent than those with an Asp-43 residue in flg22 (Sun *et al.*, 2006). The genomic analysis also revealed the absence of *fliD* in CITA 44. In several bacterial species the absence of *fliD* was associated with non-flagellated and non-motile cells

(Ikeda *et al.*, 1996; Kim *et al.*, 1999). This is not the case of CITA 44, which showed a single polar flagellum like *X. arboricola* pv. *pruni* strains (Garita-Cambronero *et al.*, 2016a).

The type IV pilus is an important structure related to the movement across the surface, adhesion, microcolony formation, secretion of proteases and colonization factors, being a key pathogenesis factor (Craig & Li, 2008). In general, strains isolated from *Prunus*, despite their pathogenicity, harbor all the four subcomplexes that permit the biogenesis and function of this structure (Dunger *et al.*, 2016); additionally these components were demonstrated as functional in the twitching motility assay.

As the type IV pilus, non-fimbrial adhesins are involved in bacterial attachment to host surfaces; in *X. oryzae*, they play an important role in pathogenesis and are particularly involved at the initial stage of leaf attachment and penetration into the host (Ray *et al.*, 2002; Das *et al.*, 2009). *X. arboricola*, shows various combinations of adhesins among the different strains studied. This variability could be associated to the different bacterial-plant compatible interactions of the different strains and hosts (Ryan *et al.*, 2011). Once more, CITA 44 lacked one of the genes involved in adhesion synthesis which was present in *X. arboricola* pv. *pruni*. Further functional studies in this way could be valuable to define the definitive role of these adhesins in the pathogenicity and host specificity observed in the pathovars of *X. arboricola*.

After the very initial stages of pathogenesis described above, some virulence genes related to host colonization, multiplication and development of symptoms are expressed. T2SS permits the export of proteins from the bacterial cell and it is involved in the translocation of degradative enzymes which causes damage to the host cells and tissues (Cianciotto, 2005). Most species of *Xanthomonas* encode the components of two T2SS. One of them is Xps, which contributes to bacterial virulence by the secretion of xylanases and proteases in *X. campestris* and *X. oryzae* (Szczesny *et al.*, 2010). In *X. arboricola*, all the components of the *xps* gene cluster were conserved. As in all the *Xanthomonas* species, *X. arboricola* showed a core repertoire of cell wall degrading enzymes which is formed by at least 30 CDS. Beside this, a high variation in the repertoire of these enzymes was also found among the analyzed strains. These variations could be related to the different cell wall composition of the hosts or tissues and the requirements to produce symptoms during the infection (Ryan *et al.*, 2011).

Differences in the profile of degrading enzymes were identified among the *Prunus*-pathogenic strains and other strains of *X. arboricola*, including CITA 44.

Another gene cluster associated with pathogenesis in *Xanthomonas* is the xanthan producing *gum* gene cluster which is composed by 12 genes. Deletion of this group of genes inhibits the disease development in *X. campestris* on *Arabidopsis* and *Nicotiana* (Yun *et al.*, 2006). In *X. arboricola*, this cluster was conserved, with the exception of *gumG* that was absent in all the sequenced strains and the O-acetyltransferases encoded by *gumF* lacking in strains CITA 44 and 3004. Nevertheless, the absence of these two genes has been demonstrated to not affect the xanthan polymerization (Katzen *et al.*, 1998).

T3SS, as well as the T3Es, plays an essential role in pathogenicity once bacteria have penetrated the host tissue and could be related to host specificity in *Xanthomonas* (White *et al.*, 2009). As occurred in other *Xanthomonas* species, all the *X. arboricola* strains harbored orthologs CDS to the genes that encode HpaR2, HpaS, HrpG and HrpX, which are involved in the regulation of T3SS, T3Es as well as other pathogenic factors such as those encoded by *pehA* and *pehD* genes (Jacobs *et al.*, 2015). This was also previously reported in the strains isolated from walnut as well as for strains 3004, CITA 44 and IVIA 2626.1 (Jacobs *et al.*, 2015; Cesbron *et al.*, 2015; Garita-Cambronero *et al.*, 2016a).

It is important to note the localization in *X. arboricola* of an ortholog to the T3E, called *xopAQ*, in the plasmid pXap41, next to other two virulence associated proteins, *xopE3* and *mltB* (Pothier *et al.*, 2011c). This effector has been previously found in the chromosome of *X. citri* and *X. gardneri* and a similar sequence has been found in *Ralstonia solanacearum* (Potnis *et al.*, 2011; Jalan *et al.*, 2013). Nucleotide sequence analysis revealed a high pairwise nucleotide identity (99.7%) to the *xopAQ* sequence of *X. citri* subsp. *citri* Aw12879, as well as a putative plant-inducible promoter box (PIP-box) sequence 67 bps upstream the start codon, but this conserved sequence associated with the regulon HrpX showed a variant in one nucleotide (S4 Fig). The presence of at least three T3Es in pXap41 reinforces the hypothesis that this plasmid may contribute to the virulence of this pathovar as well as to the specialization of the pathovar *pruni* towards hosts of the *Prunus* genus (Pothier *et al.*, 2011c). CITA 44 did not show *xopAQ* or any of the T3Es found, which is consistent with the absence of pXap41 and its non-

pathogenic character. Currently, functional studies on the effect of pXap41 and its related T3Es are being conducted in our group to confirm their role in the pathogenicity and the host specificity of the pathovar *pruni* strains.

In conclusion, this study of strains with diverse virulence range, and particularly strain CITA 44, has provided an opportunity to elucidate essential mechanisms in the host-bacteria interaction in *Xanthomonas arboricola* pv. *pruni*. In addition, it is useful to understand the evolution of the pathogenicity in this species. Furthermore, the comparative analysis reported here highlights several differences regarding to key genotypic and phenotypic features of initial and late stages of the infection process. The different TBDT, STCR, MCP profiles, the genomic variation in the components of flagellin, the type IV pilus and in cell-wall degrading enzymes repertoire, and the alteration in main virulence factors provide information that contributes to explain the phenotypic differences observed between pathogenic and non-pathogenic strains. Although the work presented shows a global overview of mechanisms involved in virulence, further functional studies are needed to test and improve the understanding of the role of all the virulence-related features mentioned here.

3.5 Materials and Methods

3.5.1 Bacterial strains and culture conditions

X. arboricola strain CITA 44, isolated from asymptomatic leaves of *P. mahaleb* in the Spanish region of Aragón, and fourteen bacterial strains of *X. arboricola* pv. *pruni* isolated from trees showing bacterial spot symptoms in Spain and Italy, were used in this study (Table 1). From these fourteen strains, three were isolated from almond [*Prunus amygdalus* Batsch, syn. *P. dulcis* (Miller) D. A. Webb], four from peach [*Prunus persica* (L.) Batsch], six from Japanese plum (*Prunus salicina* Lindley) and one from the hybrid rootstock Garnem (GxN15) (*P. persica* x *P. amygdalus*). Reference strains of *X. arboricola* pv. *pruni* (10.0343, 10.0400, 10.439, CFBP 3894, CFBP 5530, CFBP 5724), *X. arboricola* pv. *corylina* (CFBP 1846), *X. arboricola* pv. *populi* (CFBP 3123) and *X. arboricola* pv. *juglandis* (IVIA 2113) were also included in this study (Table 1). Strains were routinely cultured on 1.5% agar plates of Luria Bertani (LB) or in LB broth at 27°C for 48 h. All *Xanthomonas* strains are conserved in the collections

of Centro de Investigación y Tecnología Agroalimentaria de Aragón (CITA, Aragón, Spain) and Instituto Valenciano de Investigaciones Agrarias (IVIA, Valencia, Spain) and no specific permission was required for their use.

3.5.2 Multi Locus Sequence Analysis (MLSA).

Bacterial DNA was extracted from cultures in LB broth obtained after 24 h incubation at 27°C, using a QIAamp DNA miniKit according to the manufacturer's instructions (Qiagen, Barcelona, Spain). DNA was used for PCR or stored at -20°C until further use. Degenerate primers, previously determined as useful in MLSA analysis conducted in *Xanthomonas* (Young *et al.*, 2008; Kaluzna *et al.*, 2014), were used for PCR amplification of partial sequences of the housekeeping genes *dnaK*, *fyuA*, *gyrB* and *rpoD*. PCR amplifications were carried out in a 50 µL volume containing 1X PCR buffer (10 mM Tris-HCl, 50 mM KCl, 0.1% Triton X-100 [pH 9.0]); 0.2 µM of each primer; 1.25 U Taq DNA polymerase (Biotools, Madrid, Spain); 0.2 mM each dNTP (Biotools Madrid, Spain); 1.5 mM MgCl₂ and 1.0 µg/µL of DNA template. All PCR reactions were performed in an ABI 2720 thermal cycler (Applied Biosystems, Foster Urban district, CA, USA) with an initial denaturation at 94°C for 5 min, 40 cycles of denaturation at 94°C for 1 min, annealing at 55°C for 1 min and extension at 72°C for 2 min, and a final extension step at 72°C for 10 min. PCR products were visualized under UV light in 2% agarose gels stained with ethidium bromide and purified with the Wizard SV Gel and PCR Clean-up System Kit (Promega Corporation, Madison, USA). PCR products were sequenced at STAB VIDA (Lisbon, Portugal), and edited using BioEdit Sequence Alignment Editor (Hall, 2011). Additionally, sequences of the housekeeping genes used for MLSA analysis from *X. arboricola* pathovars *celebensis* (ICMP 1488), *corylina* (ICMP 5726) and *juglandis* (ICMP 35) as well as *X. axonopodis* pv. *citri* strain ICMP 24, included as outgroup, were obtained from the National Center for Biotechnology Information database (NCBI) (<http://www.ncbi.nlm.nih.gov>).

Nucleotide sequences were aligned with ClustalW version 1.83 (Goujon *et al.*, 2010) using default parameters. Both ends of each alignment were trimmed to the following sizes: *dnaK*, 864 positions; *fyuA*, 607 positions; *gyrB*, 631 positions and *rpoD*, 756 positions. Nucleotide sequences from all strains analyzed here and from those *Xanthomonas* spp. available in databases (Young *et al.*, 2008) were aligned and

concatenated to give a total length of 2,858 nucleotide positions. The programs jModelTest 0.1.1 (Posada, 2008) and MEGA 5.05 (Tamura *et al.*, 2011) were used to determine the best model of evolution for Maximum Likelihood analysis (ML) based in the akaike information criterion (AIC) (Posada & Buckley, 2004). For the concatenated gene dataset the model selected was TN93+G. Maximum likelihood trees, using 1,000 bootstrap re-samplings, were generated in MEGA 5.05.

Nucleotide sequences were deposited in GenBank. Accession numbers for the partial sequences of the genes used in this study are: KR054426 to KR054449 for *dnaK*; KR054450 to KR054473 for *fyuA*; KR054474 to KR054497 for *gyrB* and KR054498 to KR054521 for *rpoD*.

3.5.3 Genomic Analysis

Based in the MLSA analysis, three *X. arboricola* strains, CITA 33 (Synonymy Xap 33), CITA 44 and IVIA 2626.1, were selected for complete genome sequencing and genome features for these strains have been previously announced (Garita-Cambronero *et al.*, 2014, 2016a). Nucleotide sequences of the draft genome sequences of these three strains (Deposited at DDBJ, EMBL, GenBank under the accession numbers JHUQ000000000 for CITA 33, LJGM000000000 for CITA 44, and LJGN000000000 for IVIA 2626.1), as well as nucleotide sequences of three strains of *X. arboricola* (3004, CFBP 7634 and CFBP 7651) (Ignatov *et al.*, 2015; Cesbron *et al.*, 2015), two strains of *X. arboricola* pv. *celebensis* (NCPBP 1630 and NCPBP 1832) (Harrison *et al.*, 2016), one strain of *X. arboricola* pv. *corylina* (NCCB 100457) (Ibarra Caballero *et al.*, 2013), eight strains of *X. arboricola* pv. *juglandis* (CFBP 2528, CFBP7179, NCPBP 1447, XajA3, Xaj2, Xaj4-1, Xaj43a and Xaj417) (Higuera *et al.*, 2015; Pereira *et al.*, 2015; Cesbron *et al.*, 2015) and one strain of *X. arboricola* pv. *pruni* (MAFF 301420) were obtained from the NCBI's genome database (S1 Table). Genome scaffolds presented in each one of the bacterial genomes mentioned above were arranged and oriented by Mauve (Darling *et al.*, 2004) using the complete genome sequence of *X. arboricola* pv. *juglandis* strain Xaj417 (Pereira *et al.*, 2015) as reference.

For the purpose of homogeneity for further comparison, all the genome sequences were annotated using Prokka (Seemann, 2014). GFF3 archives generated by Prokka were used as the input to determine the core and the dispensable accessory genes in the analyzed genomes using Roary (Page *et al.*, 2015). The online tool available at bioinformatics.psb.ugent.be, was used to generate the Venn diagram to compare the genome composition of the *Prunus*-isolated strains with the remaining subspecific groups of *X. arboricola*. Signal peptides and transmembrane domains for the unique protein coding sequences (CDS) of *X. arboricola* CITA 44 and *X. arboricola* pv. *pruni* CITA 33 or IVIA 2626.1, were determined using the signalP 4.1 server and the TMHMM server version 2.0 (Krogh *et al.*, 2001; Petersen *et al.*, 2011). Beside this, the assignment of these genes to the clusters of orthologous groups (COG) database (Tatusov, 2000) was performed with the NCBI's conserved domain database using an expected value of 0.001 (Marchler-Bauer *et al.*, 2014).

The core genome sequence obtained for each strain was aligned using MAFFT (Katoh *et al.*, 2002) for further phylogenetic analysis. Subsequently, a maximum likelihood tree, using 1,000 bootstrap resamplings, was constructed to accurately determine the phylogenetic position of the atypical strain isolated from *Prunus*, CITA 44, within *X. arboricola*. Maximum likelihood tree was carried out using the RaxML tool (Stamatakis, 2014) and the dendrogram obtained was visualized using Dendroscope (Huson *et al.*, 2007). *X. campestris* pv. *campestris* strain ATCC 33913 was used as an outgroup in the analysis.

3.5.4 Carbon sources utilization analysis and environmental sensors profile

To prepare the inocula, bacterial strains used in the MLSA analysis (Table 1) were cultured in LB plates and resuspended in sterile phosphate-buffered saline (PBS) (OD₆₀₀=0.3). Suspensions (150 µL) were inoculated into each well of the Biolog GN2 microplates (Biolog Inc., USA). Microplates were read at 570 nm using a Labsystems Multiskan RC spectrophotometer (Fisher Scientific, Waltham, USA) after 24 h incubation at 27°C (Vernière *et al.*, 1993; Vauterin *et al.*, 1995). Three independent assays were performed, each including two microplates per strain and three reads per well. Means from all the reads were performed and used in further analysis.

Three substrate utilization levels were defined based in the percentage of utilization referred to water control (Maharjan *et al.*, 2007): positive (+, > 160%); borderline, considered as non-informative ($\pm$, 130-160%) and negative (-, < 130%).

The carbon sources utilization profile of strain CITA 44 was determined and compared with those obtained for strains of the pathovars *corylina*, *juglandis*, *populi* and *pruni*; results were converted to a binary form by scoring the metabolism observed for each compound as 0 (no catabolism of the carbon source) and 1 (catabolism of the carbon source). Similarity for pairs was calculated according to the Jaccard's coefficient and was subjected to Unweight Pair Group Method with Arithmetic Mean (UPGMA) cluster analysis. Finally, the reliability of the similarity trees was determined using the cophenetic correlation index. All the analysis were computed on NTSYS 2.11T (Exeter Software, Setauket, NY).

Profiles of sensors of two-component regulatory system (STCRS) and TonB-dependent transporters (TBDTs) for *X. arboricola* strains isolated from *Prunus* (CITA 33, CITA 44 and IVIA 2626.1) and six other *Xanthomonas* (3004, CFBP 7634, CFBP 7651, NCPPB 1630, NCCB 100457 and CFBP 7179) (S1 Table) were determined based in the analysis of the complete genomes and the search of homologous sequences for 86 STCRS and 28 TBDTs previously described in other *Xanthomonas* species (Mhedbi-Hajri *et al.*, 2011). Sequence homology searches were conducted using the Blast tool according to the protocol proposed by the NCBI (NCBI, 2016), using the nucleotide sequence that encodes each one as the input. Those CDS which presented an identity and a query coverage percentage over 80% were considered as orthologous genes in the target genomes of *X. arboricola*.

3.5.5 Chemotaxis assay and repertoire of methyl accepting chemotaxis proteins

Eighteen carbon compounds that induce variation in the chemotactic response in other *Xanthomonas* (Sena-Vélez, 2015) were evaluated to determine differences in the chemotactic profile between the atypical strain CITA 44 and *X. arboricola* pv. *pruni* strains CITA 33 and CFBP 5530, for which complete genome sequence had been already obtained (Pothier *et al.*, 2011b; Garita-Cambronero *et al.*, 2014, 2016a). The chemotactic compounds tested were: alanine (10 and 250 mM), arginine (10 and 100 mM), citric acid (10 mM), cysteine (10 mM), fructose (10 mM), galactose (10 mM),

galacturonic acid (10 mM), glucose (10 mM), glucuronic acid (10 mM), glycerol (0.2 and 2%), leucine (10 and 150 mM), maltose (10 mM), mannitol (0.2%), serine (10 and 200 mM), sodium citrate (10 mM), succinic acid (10 mM), sucrose (10 mM) and xylose (10 mM).

The chemotactic effect of the carbon compounds was evaluated according to a microtiter plate chemotaxis assay previously developed (Sena-Vélez, 2015). Briefly, 10 μ L tips containing 5 μ L of the carbon source tested were inserted into 48 of a 96 wells plate previously inoculated with 200 μ L of bacterial suspension in 10 mM MgCl₂ (10^8 CFU/mL). The number of bacteria that moved into the tip during one hour was estimated by means of serial dilutions of the tip content plated on 1.5% agar LB plates. Significant differences ($p < 0.05$) between the means from each carbon source tested and the negative control (10 mM MgCl₂) were determined by an analysis of variance (ANOVA) according to the Student-Newman-Keuls method. Statistical analyses were performed using STATGRAPHICS Plus v.5.1 (Manugistics Inc. Rockville Maryland, USA). A carbon source was considered as a chemoattractant when the average number of bacteria contained in the tip (6 replicates in 2 independent assays) was significantly higher to the blank control (10 mM MgCl₂), and considered as a chemorepellent when the average was significantly lower ($p < 0.05$).

Orthologous genes for the 26 methyl accepting chemotaxis proteins (MCPs) as well as for 18 specific-chemotaxis *che* genes (Mhedbi-Hajri *et al.*, 2011), described in *X. campestris* pv. *campestris*, *X. campestris* pv. *vesicatoria*, *X. citri* subsp. *citri* and *X. oryzae*, were searched in the genome sequences of nine *X. arboricola* strains as mentioned above. Those CDS with more than 80% of identity and more than 80% of sequence coverage containing MCPs periplasmic domains were considered as orthologous genes.

3.5.6 Motility in solid and semisolid surfaces and surfactant activity

Swarming, swimming and twitching motility assays were conducted for strain CITA 44 as well as for 20 strains classified as *X. arboricola* pv. *pruni* by MLSA. Bacterial cultures in exponential growth phase were centrifuged at 6,350 g for 15 min, then washed and resuspended in 10 mM MgCl₂ (OD₆₀₀=1.0). To analyze swarming motility, 0.5% agar PYM swarming plates (peptone 0.5%, yeast extract 0.3%, malt

extract 0.3%, glucose 1.0%) were inoculated with 10 μ L of bacterial suspension. For swimming motility assay, cultures resuspended in 10 mM MgCl₂ were centrifuged at 6,350 g during 15 min and the pellets were inoculated using a sterile toothpick in the center of semisolid 0.3% agar minimal medium A (MMA) plates (K₂HPO₄ 0.7%, KH₂PO₄ 0.3%, MgSO₄·7H₂O 0.01%, (NH₄)₂SO₄ 0.1%, sodium citrate 0.005%, glycerol 0.2%). For twitching assay, bacterial cultures were prepared as mentioned above, and then inoculated with a sterile toothpick through a 5 mm PYM 1.5% agar layer to the bottom of the plate. After 72 h, culture medium was removed and the bottom of the plate was stained with 0.3% crystal violet for 15 min and washed using sterile distilled water.

Surfactant production, which is closely related to bacterial motility on solid surfaces, was also assessed according to the atomized oil assay (Burch *et al.*, 2010). Briefly, bacteria were inoculated using a sterile toothpick on 1.5% agar LB plates, and after 24 hpi, a fine mist of mineral oil droplets were sprayed on the colony. Bacterial strains that instantaneously showed a bright halo around the colony were recorded as surfactant producers.

Plates from swarming motility and surfactant activity assays were incubated at 27°C during 24 h, whilst swimming and twitching plates were incubated for 72 h. Images from all the plates were recorded. All assays were performed in three independent experiments with three replicates each time.

Homologous CDS to the major structural components associated with flagellum (Blocker *et al.*, 2003; Chevance & Hughes, 2008; Mhedbi-Hajri *et al.*, 2011) as well as the fimbrial adhesin type IV pilus (Dunger *et al.*, 2016) and a group of non-fimbrial adhesins (Mhedbi-Hajri *et al.*, 2011) were identified in the three genomes of *X. arboricola* isolated from *Prunus* and in other six *X. arboricola* strains as described above.

3.5.6 Pathogenicity tests on *Prunus* spp.

Bacteria were grown on 1.5% agar LB plates and incubated at 27°C for 48 h, then a single bacterial colony was resuspended in 30 mL of LB broth and incubated at 27°C on a rotary shaker for 24 h. After incubation, cultures were centrifuged at 6,350 g for 15

min and washed three times with 10 mM MgCl₂. Finally, cultures were adjusted to 10⁸ CFU/mL (OD₆₀₀= 0.1) and used to inoculate detached leaves of different *Prunus* species.

Young, fully expanded leaves from greenhouse-grown plants of almond (cv. Ferraduel), apricot (cv. Canino), peach (cv. Calanda) and European plum (cv. Golden Japan) were collected, brought to the laboratory and washed three times in sterile distilled water. After surface sterilization with 70% ethanol and 0.05% sodium hypochlorite, the leaves were rinsed three times with sterile distilled water and dried on a hood. Three leaves per host were selected randomly for being inoculated with each strain in three independent assays.

Surface sterilized leaves were placed on 0.5% water agar plates and the abaxial surfaces were gently inoculated using a sterile cotton swab damped with the bacterial inoculum. Sealed plates were incubated in a grow chamber adjusted at 27°C, 80-90% relative humidity and 12 h photoperiod. Inoculated leaves were digitally recorded at 28 dpi and the percentage of symptomatic area per host and strain was quantified using the ImageJ 1.45s software. Samples, inoculated with sterile 10 mM MgCl₂, were used as blank control for comparison.

Additionally to the genomic analysis related to the early events in the bacteria-host interaction described below, homologous CDS to several genes associated with later stages of pathogenesis, like the type III secretory system (T3SS) and the type III effectors (T3Es), were searched in the genome of the nine strains of *X. arboricola* according to the percentage of the sequence length and identity with the amino acid sequence of the genes previously associated with these features (Bonas *et al.*, 1989; Hajri *et al.*, 2009; White *et al.*, 2009; Büttner & Bonas, 2010; Bogdanove *et al.*, 2011; Potnis *et al.*, 2011; Guo *et al.*, 2011; Abby & Rocha, 2012). In the same manner, orthologous CDS to other remarkable processes and features associated with virulence such as the quorum sensing (He & Zhang, 2008), the xanthan biosynthesis (Vorhölter *et al.*, 2008), the type II secretion system (Filloux, 2004; Szczesny *et al.*, 2010) as well as the cellulolytic, hemicellulolytic, pectolytic and lipases enzymes, previously described in other species of *Xanthomonas* (da Silva *et al.*, 2002; Subramoni *et al.*, 2010; Vandroemme *et al.*, 2013; Nascimento *et al.*, 2016), were searched in the genome

sequences of the nine strains of *X. arboricola*. Only those CDS with a percentage of coverage and identity over 80.0% were considered as orthologous sequences.

Finally, the presence of the plasmid pXap41 (Pothier *et al.*, 2011c), putatively involved in virulence in *X. arboricola* pv. *pruni*, was searched in the analyzed genomes based in the nucleotide sequence similarity and graphically represented using the BLAST Ring Image Generator (BRIG) tool (Alikhan *et al.*, 2011); blastn was used for the sequence comparative analysis with an expected value threshold of 0.001. All the CDS associated with pathogenesis, mentioned above, were represented in a circular genome map using CGView (Grant & Stothard, 2008). For this purpose, contigs of the draft genome sequence of CITA 33, CITA 44 and IVIA 2626.1 were arranged by Mauve (Darling *et al.*, 2004) using the circularized genome sequence of *X. arboricola* pv. *juglandis* strain Xaj 417 as the reference (Pereira *et al.*, 2015).

3.6 Acknowledgments

We would like to thank to Elisa Ferragud and Isabel M. Berruete for technical assistance on the phenotypic features evaluated as well as to Pilar Sabuquillo for assistance on searching orthologous genes.

3.7 Author contributions

Conceptualization: JGC, JC, AP, MML. Methodology: JGC, JC. Software: JGC. Validation: JGC, JC, AP. Formal analysis: JGC, JC. Investigation: JC, JGC Resources: JC, AP, ML. Data curation: JGC. Writing (original draft preparation): JGC, JC. Writing (review and editing): JGC, JC, AP, MML. Visualization: JGC. Supervision: JC, AP, MML. Project administration: JC, APB. Funding acquisition: JC, AP, MML.

Capítulo 4: El análisis pan genómico ha permitido diferenciar cepas no patógenas y patógenas de *Xanthomonas arboricola* que cohabitan en *Prunus* spp. así como elucidar los factores de la virulencia bacteriana.

Pan-genomic analysis permitted to differentiate non-pathogenic and pathogenic strains of *Xanthomonas arboricola* that cohabit *Prunus* spp. and elucidate bacterial virulence factors.

J. Garita-Cambroner¹, A. Palacio-Bielsa², M. M. López³, J. Cubero¹.

¹Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA), Madrid, Spain. ²Centro de Investigación y Tecnología Agroalimentaria de Aragón, Zaragoza, Spain. ³Instituto Valenciano de Investigaciones Agrarias, Valencia, Spain.

Enviado a Frontiers in Microbiology.

4.1 Abstract

X. arboricola is a plant-associated bacterial species which comprises nine pathovars that cause disease on several economically important plant hosts. One of the most virulent pathovars within this genus is *X. arboricola* pv. *pruni* (*Xap*), the causal agent of bacterial spot disease of stone fruits and almond. Recently, a non-pathogenic *Xap*-look-alike strain isolated from *Prunus* was characterized and its genome compared to pathogenic strains of *Xap*. This study revealed differences in the profile of virulence factors between pathogenic and non-pathogenic strains. For instance, the non-pathogenic strain did not contain genes related to the type III secretion system (T3SS), type III effectors (T3Es) nor the putative virulence-related plasmid pXap41. The existence of this atypical strain arouses several questions associated with the abundance, the pathogenicity and the evolutionary context of *X. arboricola* on *Prunus* hosts. After an initial characterization of a collection of xanthomonads strains isolated from *Prunus* bacterial spot outbreaks in Spain during the past decade, six *Xap*-look-alike strains, that did not clustered with the pathogenic strains of *Xap* according to a multi locus sequence analysis, were identified. Pathogenicity of these strains were analysed and based in this character, the genome sequence of two *Xap*-look-alike strains, CITA 14 and CITA 124, selected by their non-pathogenicity characteristic, were obtained and compared to those available genomes of the *Prunus*-associated strains of *X. arboricola*. Differences were found among the pathogenic and non-pathogenic strains in several characters related to the pathogenesis process. Additionally, a pan-genomic analysis that included the available genomes of *X. arboricola*, revealed that the atypical strains associated with *Prunus* were related to a group of non-pathogenic or low pathogenic strains isolated from a wide host range. The repertoire of the genes related to T3SS and T3Es varied among the strains of this cluster and those strains related to the most pathogenic pathovars of the species, *corylina*, *juglandis* and *pruni*. This variability provides information about the potential evolutionary process associated to the acquisition of pathogenicity and host specificity in *X. arboricola*. Finally, based in the genomic differences observed between pathogenic and non-pathogenic strains, a sensitive and specific real-time PCR protocol to identify and detect *Xap* strains was designed; this method avoids miss-identifications due to atypical strains of *X. arboricola* that can cohabit *Prunus*.

Keywords: Stone fruits, Almond, Comparative genomics, Bacterial spot disease.

4.2 Introducción

Xanthomonas arboricola is a species of plant pathogenic bacteria associated with a wide range of hosts plants (Vauterin *et al.*, 1995). Due to their distinctive pathogenicity towards a delimited host range as well as to other genomic determinants, this species has been divided at infrasubspecific level in at least nine different pathovars (Vauterin *et al.*, 1995; Fischer-Le Saux *et al.*, 2015). Among these bacterial groups, the pathovars *corylina*, *juglandis* and *pruni*, which cause disease on nut stone fruit trees and almond, are considered the most relevant (Lamichhane & Varvaro, 2014). Several outbreaks of the pathovar *pruni*, which is regulated by quarantine policies in the European Union and the possibility of future epidemics and spread to disease-free producing zones, have potentiated the efforts to understand the molecular diversity of the species (EFSA, 2014).

Very recent studies conducted by multilocus sequence typing and genome-wide based techniques, have provided a significant increase in the knowledge associated with the genetic structure and diversity of the strains of *X. arboricola*. As occurred in other *Xanthomonas* species (White *et al.*, 2009; Hajri *et al.*, 2012; Jacobs *et al.*, 2015), most significant genomic differences associated with virulence in *X. arboricola* strains are related to type three secretion system (T3SS) and type three effectors (T3Es) predicted by the available genomes (Ibarra Caballero *et al.*, 2013; Garita-Cambronero *et al.*, 2014, 2016a; Essakhi *et al.*, 2015; Higuera *et al.*, 2015; Ignatov *et al.*, 2015; Pereira *et al.*, 2015; Cesbron *et al.*, 2015; Harrison *et al.*, 2016).

Besides this, phylogenetic analysis based in the core genome sequence of *X. arboricola* (Cesbron *et al.*, 2015; Garita-Cambronero *et al.*, 2016b), revealed that three non-pathogenic strains did not group together with pathogenic strains isolated from walnut stone fruit or almond. Instead, these non-pathogenic strains, isolated from walnut and mahaleb cherry (*Prunus mahaleb*), were comprised in a group with several other low pathogenic strains, such as the low pathogenic strains of the pathovar *celebensis*, a pathogen of banana (*Musa* spp.) (Harrison *et al.*, 2016), or with the strain 3004 of *X. arboricola* isolated from barley (Ignatov *et al.*, 2015).

The existence of non-pathogenic strains has aroused the concern on how abundant are these strains and the possibility of their misidentification as pathogenic strains by the

current diagnostic approaches. Moreover, they could be useful to obtain some clues related to the evolution of pathogenesis within *X. arboricola*. Recalling all these recent advances, our goal was to deepen the characterization of the genomic features of three atypical strains isolated from *Prunus* spp., in order to determine how these key features associated with pathogenesis varied among atypical and pathogenic strains of *X. arboricola*, as well as to determine if these variants could be used to design precise molecular tools to differentiate these two groups when they are cohabiting the same host.

4.3 Materials and Methods

4.3.1 Bacterial strains and classification using multilocus sequence analysis

A selection of 31 previously characterized strains of *X. arboricola* (Garita-Cambronero *et al.*, 2016b) from the pathovars *corylina*, *juglandis*, *populi* and *pruni* were utilized. Besides, 40 strains phenotypically similar to *Xanthomonas*, collected during the Spanish outbreaks of bacterial spot disease of stone fruits as well as from routine screenings performed on Spanish nurseries (Table 1), were screened for identification as *Xanthomonas arboricola* pv. *pruni* (*Xap*). All the bacterial strains listed in Table 1 are available in the bacterial collections from the Instituto Valenciano de Investigaciones Agrarias (IVIA, Valencia, Spain) and the Centro de Investigación y Tecnología Agroalimentaria de Aragón (CITA, Aragón, Spain).

Bacterial strains were routinely cultured on Luria Bertani (LB) 1.5 % agar plates or in LB broth at 27 °C for 48 h. The commensal bacterial strains, isolated from *Prunus* and used in this study (Table 1), were identified to genus level based on the partial sequence of the 16S rDNA gene according to a method described previously (Lagacé *et al.*, 2004).

For *Xap* classification, the real-time PCR for an ABC transporter developed by Palacio-Bielsa and collaborators (Palacio-Bielsa *et al.*, 2011, 2015b) was used. Due to the fact, demonstrated in previous studies, that the presence of the plasmid pXap41 (Pothier *et al.*, 2011c) differentiated between *Xap* and the atypical strain CITA 44 of *X. arboricola*

isolated from *Prunus* (Garita-Cambronero *et al.*, 2016b), the presence of this plasmid was also tested by using a previously reported multiplex PCR protocol based on the plasmid-related genes *repA1*, *repA2* and *mobC* (Pothier *et al.*, 2011c). Those strains that showed a positive result in the real-time assay (Palacio-Bielsa *et al.*, 2011) but did not show amplification for the three genes associated with pXap41 (Pothier *et al.*, 2011c), were considered as *Xap*-look-a-like strains, and were further identified according to a multilocus sequence typing scheme (MLSA) based in the partial sequences of the housekeeping genes *dnaK*, *fyuA*, *gyrB* and *rpoD* (Young *et al.*, 2008), which was used to characterize and discriminate the members of the pathovar *pruni* from the atypical or commensal strain CITA 44 of *X. arboricola* (Garita-Cambronero *et al.*, 2016b).

Additionally, sequences of these housekeeping genes from the non-pathogenic *X. arboricola* strain CITA 44 and sequences from *X. arboricola* pathovars *corylina* (CFBP1159=ICMP 5726 and CFBP 1846), *celebensis* (CFBP3523=ICMP 1488), *juglandis* (CFBP2528=ICMP 35 and IVIA 2113), *populi* (CFBP 3123) and *pruni* (CFBP 2535=ICMP 51, CFBP 5530, Xap 33=CITA 33 and IVIA 2626.1), as well as *X. citri* subsp. *citri* strain CFBP 2525=ICMP 24 included as outgroup, were obtained from the National Center for Biotechnology Information database (NCBI) (<http://www.ncbi.nlm.nih.gov>).

Purified PCR products were sequenced at STAB VIDA (Lisbon, Portugal), and edited using Geneious (Kearse *et al.* 2012). Obtained nucleotide sequences were aligned with ClustalW version 1.83 (Hall, 2011) using default parameters. Both ends of each alignment were trimmed to the following sizes: *dnaK*, 842 positions; *fyuA*, 601 positions; *gyrB*, 631 positions and *rpoD*, 759 positions. Then, they were aligned and concatenated to give a total length of 2,836 nucleotide positions. For the analysis of the concatenated gene dataset, Tamura-Nei (TN93) model was selected and maximum likelihood trees, using 1,000 bootstrap re-samplings, were generated using MEGA 6.0 software (Tamura *et al.*, 2013)

Nucleotide sequences were deposited in GenBank. Accession numbers for the partial sequences of the genes used in this study are: KX357115 to KX357120 for *dnaK*; KX357133 to 357138 for *fyuA*; KX357127 to KX357132 for *gyrB* and KX357121 to KX357126 for *rpoD*.

Table 1. Bacterial strains used in this study.

Taxa	Strain	Origin	Host	ftsx	xopE3	pXap41
<i>Xanthomonas arboricola</i>						
<i>X. arboricola</i> pv. <i>pruni</i> *	100.343	Zuidwolde (Netherlands)	<i>Prunus laurocerasus</i> cv. Caucasicus	+	+	+
<i>X. arboricola</i> pv. <i>pruni</i> *	100.400	Zuidwolde (Netherlands)	<i>Prunus laurocerasus</i> cv. Grüner Teppich	+	+	+
<i>X. arboricola</i> pv. <i>pruni</i> *	100.439	Elburg (Netherlands)	<i>Prunus laurocerasus</i> cv. Zabeliana	+	+	+
<i>X. arboricola</i> pv. <i>pruni</i> *	CFBP 3894 ^{PT}	New Zealand	<i>Prunus salicina</i>	+	+	+
<i>X. arboricola</i> pv. <i>pruni</i> *	CFBP 5530	Italy	<i>Prunus persica</i>	+	+	+
<i>X. arboricola</i> pv. <i>pruni</i> *	CFBP 5724	United States	<i>Prunus amygdalus</i>	+	+	+
<i>X. arboricola</i> pv. <i>pruni</i> *	CITA 4	Aragón (Spain)	<i>Prunus persica</i> cv. Catherine	+	+	+
<i>X. arboricola</i> pv. <i>pruni</i> *	CITA 9	Zaragoza (Spain)	<i>Prunus persica</i> cv. Merrill O'Henry	+	+	+
<i>X. arboricola</i> pv. <i>pruni</i> *	CITA 11	Huesca (Spain)	<i>Prunus persica</i> cv. Richard Lady	+	+	+
<i>X. arboricola</i> pv. <i>pruni</i> *	CITA 33	Huesca (Spain)	<i>Prunus amygdalus</i> cv. Guara	+	+	+
<i>X. arboricola</i> pv. <i>pruni</i> *	CITA 46	Navarra (Spain)	<i>Prunus persica</i> cv. Summer Lady	+	+	+
<i>X. arboricola</i> pv. <i>pruni</i> *	CITA 70	Zaragoza (Spain)	<i>Prunus persica</i> x <i>Prunus amygdalus</i>	+	+	+
<i>X. arboricola</i> pv. <i>pruni</i>	CITA 154	Zaragoza (Spain)	<i>Prunus amygdalus</i>	+	+	+
<i>X. arboricola</i> pv. <i>pruni</i> *	IVIA 2626.1	Badajoz (Spain)	<i>Prunus salicina</i> cv. Fortuna	+	+	+
<i>X. arboricola</i> pv. <i>pruni</i>	IVIA 2626.3	Badajoz (Spain)	<i>Prunus salicina</i> cv. Fortuna	+	+	+
<i>X. arboricola</i> pv. <i>pruni</i>	IVIA 2626.6	Badajoz (Spain)	<i>Prunus salicina</i> cv. Fortuna	+	+	+
<i>X. arboricola</i> pv. <i>pruni</i>	IVIA 2626.7	Badajoz (Spain)	<i>Prunus salicina</i> cv. Fortuna	+	+	+
<i>X. arboricola</i> pv. <i>pruni</i>	IVIA 2647.1.3	Badajoz (Spain)	<i>Prunus salicina</i> cv. Larry Ann	+	+	+
<i>X. arboricola</i> pv. <i>pruni</i> *	IVIA 2647.1-2	Badajoz (Spain)	<i>Prunus salicina</i> cv. Larry Ann	+	+	+
<i>X. arboricola</i> pv. <i>pruni</i>	IVIA 2647.1-13	Badajoz (Spain)	<i>Prunus salicina</i> cv. Larry Ann	+	+	+
<i>X. arboricola</i> pv. <i>pruni</i>	IVIA 2647.3-2	Badajoz (Spain)	<i>Prunus salicina</i> cv. Friar	+	+	+
<i>X. arboricola</i> pv. <i>pruni</i> *	IVIA 2647.3-1	Badajoz (Spain)	<i>Prunus salicina</i> cv. Friar	+	+	+
<i>X. arboricola</i> pv. <i>pruni</i>	IVIA 2649.10	Badajoz (Spain)	<i>Prunus salicina</i> cv. Friar	+	+	+
<i>X. arboricola</i> pv. <i>pruni</i>	IVIA 2649.3	Badajoz (Spain)	<i>Prunus salicina</i> cv. Friar	+	+	+
<i>X. arboricola</i> pv. <i>pruni</i>	IVIA 2649.7	Badajoz (Spain)	<i>Prunus salicina</i> cv. Friar	+	+	+
<i>X. arboricola</i> pv. <i>pruni</i>	IVIA 2758.2	Badajoz (Spain)	<i>Prunus salicina</i>	+	+	+
<i>X. arboricola</i> pv. <i>pruni</i>	IVIA 2758.3	Badajoz (Spain)	<i>Prunus salicina</i>	+	+	+
<i>X. arboricola</i> pv. <i>pruni</i> *	IVIA 2826.1	Valencia (Spain)	<i>Prunus salicina</i> cv. Anna Gold	+	+	+
<i>X. arboricola</i> pv. <i>pruni</i>	IVIA 2826.10	Valencia (Spain)	<i>Prunus persica</i> cv. Zephir	+	+	+

(Table1 continuation)

Taxa	Strain	Origin	Host	ftsx	xopE3	pXap41
<i>X. arboricola</i> pv. <i>pruni</i>	IVIA 2826.11	Valencia (Spain)	<i>Prunus persica</i> cv. Zephir	+	+	+
<i>X. arboricola</i> pv. <i>pruni</i>	IVIA 2826.3	Valencia (Spain)	<i>Prunus salicina</i> cv. Anna Gold	+	+	+
<i>X. arboricola</i> pv. <i>pruni</i>	IVIA 2826.4	Valencia (Spain)	<i>Prunus salicina</i> cv. Anna Gold	+	+	+
<i>X. arboricola</i> pv. <i>pruni</i>	IVIA 2826.5	Valencia (Spain)	<i>Prunus salicina</i> cv. Anna Gold	+	+	+
<i>X. arboricola</i> pv. <i>pruni</i>	IVIA 2826.6	Valencia (Spain)	<i>Prunus salicina</i> cv. Anna Gold	+	+	+
<i>X. arboricola</i> pv. <i>pruni</i>	IVIA 2826.9	Valencia (Spain)	<i>Prunus persica</i> cv. Zephir	+	+	+
<i>X. arboricola</i> pv. <i>pruni</i> *	IVIA 2832.10	Valencia (Spain)	<i>Prunus salicina</i> cv. Angeleno	+	+	+
<i>X. arboricola</i> pv. <i>pruni</i>	IVIA 2832.17	Valencia (Spain)	<i>Prunus salicina</i> cv. Larry Ann	+	+	+
<i>X. arboricola</i> pv. <i>pruni</i>	IVIA 2832.19	Valencia (Spain)	<i>Prunus salicina</i> cv. Larry Ann	+	+	+
<i>X. arboricola</i> pv. <i>pruni</i>	IVIA 2832.21	Valencia (Spain)	<i>Prunus persica</i> cv. Zephir	+	+	+
<i>X. arboricola</i> pv. <i>pruni</i>	IVIA 2832.24	Valencia (Spain)	<i>Prunus salicina</i> cv. Anna Gold	+	+	+
<i>X. arboricola</i> pv. <i>pruni</i>	IVIA 2832.26	Valencia (Spain)	<i>Prunus persica</i> cv. 58CC70	+	+	+
<i>X. arboricola</i> pv. <i>pruni</i>	IVIA 2832.30	Valencia (Spain)	<i>Prunus persica</i> x <i>P. davidiana</i>	+	+	+
<i>X. arboricola</i> pv. <i>pruni</i>	IVIA 2832.5	Valencia (Spain)	<i>Prunus persica</i> x <i>P. davidiana</i>	+	+	+
<i>X. arboricola</i> pv. <i>pruni</i> *	IVIA 3161.2	Alicante (Spain)	<i>Prunus amygdalus</i> cv. Rumbeta	+	+	+
<i>X. arboricola</i> pv. <i>pruni</i> *	IVIA 3162	Alicante (Spain)	<i>Prunus amygdalus</i> cv. Rumbeta	+	+	+
<i>X. arboricola</i> pv. <i>pruni</i>	IVIA 3177.1-6	Alicante (Spain)	<i>Prunus amygdalus</i> cv. Rumbeta	+	+	+
<i>X. arboricola</i> pv. <i>pruni</i>	IVIA 3177.3-4	Alicante (Spain)	<i>Prunus amygdalus</i> cv. Rumbeta	+	+	+
<i>X. arboricola</i> pv. <i>pruni</i>	IVIA 3177.3-8	Alicante (Spain)	<i>Prunus amygdalus</i> cv. Rumbeta	+	+	+
<i>X. arboricola</i> pv. <i>pruni</i>	IVIA 3181.3-1	Alicante (Spain)	<i>Prunus amygdalus</i> cv. Rumbeta	+	+	+
<i>X. arboricola</i> pv. <i>pruni</i>	IVIA 3181.3-3	Alicante (Spain)	<i>Prunus amygdalus</i> cv. Rumbeta	+	+	+
<i>X. arboricola</i> pv. <i>pruni</i> *	IVIA 3487.1	Huesca (Spain)	<i>Prunus armeniaca</i>	+	+	+
<i>X. arboricola</i> pv. <i>pruni</i>	IVIA 4490.1	Teruel (Spain)	<i>Prunus amygdalus</i> cv. Rumbeta	+	+	+
<i>X. arboricola</i> pv. <i>pruni</i>	IVIA 4491.1	Huesca (Spain)	<i>Prunus amygdalus</i> cv. Rumbeta	+	+	+
<i>X. arboricola</i> pv. <i>pruni</i>	IVIA 4493	Huesca (Spain)	<i>Prunus amygdalus</i> cv. Rumbeta	+	+	+
<i>X. arboricola</i> -look-a-like	CITA 14	Zaragoza (Spain)	<i>Prunus persica</i> cv. Luciana	+	-	-
<i>X. arboricola</i> -look-a-like	CITA 42	Zaragoza (Spain)	<i>Prunus mahaleb</i> rootstock (Santa Lucía SL-64)	+	-	-
<i>X. arboricola</i> -look-a-like*	CITA 44	Zaragoza (Spain)	<i>Prunus mahaleb</i> rootstock (Santa Lucía SL-64)	+	-	-
<i>X. arboricola</i> -look-a-like	CITA 49	Zaragoza (Spain)	<i>Prunus</i> sp.	+	-	-
<i>X. arboricola</i> -look-a-like	CITA 51	Zaragoza (Spain)	<i>Prunus amygdalus</i>	+	-	-
<i>X. arboricola</i> -look-a-like	CITA 124	Badajoz (Spain)	<i>Prunus persica</i>	+	-	-

(Table1 continuation)

Taxa	Strain ^a	Origin	Host	<i>ftsx</i>	<i>xopE3</i>	pXap41
<i>X. arboricola</i> -look-a-like	CITA 149	Guipúzcoa (Spain)	<i>Prunus laurocerasus</i>	+	-	-
<i>X. arboricola</i> pv. <i>corylina</i> *	CFBP 1846	France	<i>Corylus avellana</i>	+	-	-
<i>X. arboricola</i> pv. <i>corylina</i> *	IVIA 3978	Spain	<i>Corylus avellana</i>	+	-	-
<i>X. arboricola</i> pv. <i>corylina</i> *	RIPF-X08	Poland	<i>Corylus avellana</i>	-	-	-
<i>X. arboricola</i> pv. <i>fragariae</i> *	CFBP 6771 ^{PT}	Italy	<i>Fragaria x ananassa</i>	-	-	-
<i>X. arboricola</i> pv. <i>juglandis</i> *	IVIA 2113	Badajoz (Spain)	<i>Juglans regia</i>	-	-	-
<i>X. arboricola</i> pv. <i>juglandis</i> *	Xaj-2	Zaragoza (Spain)	<i>Juglans regia</i>	-	-	-
<i>X. arboricola</i> pv. <i>juglandis</i> *	Xaj-3	Zaragoza (Spain)	<i>Juglans regia</i>	-	-	-
<i>X. arboricola</i> pv. <i>juglandis</i> *	Xaj-4	Zaragoza (Spain)	<i>Juglans regia</i>	-	-	-
<i>X. arboricola</i> pv. <i>juglandis</i> *	Xaj-5	Zaragoza (Spain)	<i>Juglans regia</i>	-	-	-
<i>X. arboricola</i> pv. <i>populi</i> *	CFBP 3123 ^{PT}	Netherlands	<i>Populus x euroamericana</i> cv. Robusta	-	-	-
Other pathogenic species						
<i>Xanthomonas</i> spp.						
<i>X. axonopodis</i>	Xp-2	Navarra (Spain)	<i>Pelargonium</i> sp.	-	+	-
<i>X. fuscans</i> subsp. <i>fuscans</i>	NCPPB 381	Canada	<i>Phaseolus vulgaris</i>	-	+	-
<i>X. fuscans</i> subsp. <i>fuscans</i>	IVIA 151835DA	La Rioja (Spain)	<i>Phaseolus vulgaris</i>	-	+	-
<i>X. campestris</i>	Xca-1	Teruel (Spain)	<i>Brassica oleraceae</i>	-	-	-
<i>X. campestris</i>	Xca-2	Tarragona (Spain)	<i>Brassica oleraceae</i>	-	-	-
<i>X. campestris</i>	IVIA 2734.1	Spain	<i>Brassica mathis</i>	-	+	-
<i>X. citri</i> subsp. <i>citri</i>	306	Brazil	<i>Citrus sinensis</i>	+	+	-
<i>X. citri</i> subsp. <i>citri</i>	IVIA 2889-1	Argentina	<i>Citrus sinensis</i>	-	+	-
<i>X. citri</i> subsp. <i>citri</i>	IVIA 3026-1	Spain	<i>Citrus sinensis</i>	-	+	-
<i>X. hortorum</i>	IVIA 1575.1	Spain	<i>Pelargonium peltatus</i>	-	-	-
<i>X. vesicatoria</i>	IVIA 3619-1	Spain	<i>Capsicum annum</i>	-	+	-
<i>Agrobacterium</i>						
<i>A. tumefaciens</i>	Agrob-1 B1-360-1 (biovar 1)	NA	NA	-	-	NA
<i>A. tumefaciens</i>	Agrob-9	Zaragoza (Spain)	<i>Prunus persica</i>	-	-	NA
<i>Agrobacterium</i> sp.	IVIA 1245-80	Spain	<i>Prunus persica</i> cv. Adafuel	-	-	NA
<i>Agrobacterium</i> sp.	IVIA 2304-10	Spain	<i>Prunus persica</i> cv. Red Candel	-	-	NA
<i>Pantoea</i>						
<i>Pantoea</i> sp.	IVIA 2261-1	Spain	<i>Olea europea</i> cv. Picual	-	-	NA
<i>Pseudomonas</i>						
<i>P. syringae</i> pv. <i>syringae</i>	IVIA 3514-2	Spain	<i>Citrus sinensis</i> cv. Valencia Late	-	-	NA
<i>P. syringae</i> pv. <i>syringae</i>	Psy-3	Zaragoza (Spain)	<i>Prunus avium</i>	-	-	NA
<i>P. syringae</i> pv. <i>syringae</i>	Psy-13	Zaragoza (Spain)	<i>Prunus amygdalus</i>	-	-	NA

(Table1 continuation)

Taxa	Strain ^a	Origin	Host	<i>ftsx</i>	<i>xopE3</i>	pXap41
Commensal strains on <i>Prunus</i> spp.						
<i>Curtobacterium</i> sp.	EP-2.2	Puigmoreno (Teruel, Spain)	<i>Prunus amygdalus</i> cv. Guara	-	-	NA
<i>Curtobacterium</i> sp.	EP-18.1	Puigmoreno (Teruel, Spain)	<i>Prunus amygdalus</i> cv. Guara	-	-	NA
<i>Microbacterium</i> sp.	EP-16.1	Puigmoreno (Teruel, Spain)	<i>Prunus amygdalus</i> cv. Guara	-	-	NA
<i>Pantoea</i> sp.	EP-14.1	Puigmoreno (Teruel, Spain)	<i>Prunus amygdalus</i> cv. Guara	-	-	NA
<i>Pseudomonas</i> sp.	EP-17.1	Puigmoreno (Teruel, Spain)	<i>Prunus amygdalus</i> cv. Guara	-	-	NA
<i>Pseudomonas</i> sp.	21/14-12.B8	Caspe (Zaragoza, Spain)	<i>Prunus persica</i> cv. Guayox 30	-	-	NA
<i>Sphingomonas</i> sp.	EP-16.2	Puigmoreno (Teruel, Spain)	<i>Prunus amygdalus</i> cv. Guara	-	-	NA
<i>Terrabacter</i> sp.	EP-16.6	Puigmoreno (Teruel, Spain)	<i>Prunus amygdalus</i> cv. Guara	-	-	NA

PT:Pathotype strain; NA: not available information, +: Positive PCR result; -: Negative PCR result.

^aCFBP, Collection Française de Bactéries Phytopathogènes, Angers, France; IVIA, Instituto Valenciano de Investigaciones Agrarias, Valencia, Spain; CITA, Centro de Investigación y Tecnología Agroalimentaria de Aragón, Zaragoza, Spain; ICMP, International Collection of Microorganisms from Plants, Auckland, New Zealand.

*Strains characterized previously as *Xanthomonas arboricola* and used in this study as control during the molecular characterization.

4.3.2 Study of the type III secretion system and type III effectors gene repertory

Six strains isolated from *Prunus* spp. and classified as *Xap*-look-a-like, as well as the pathogenic *Xap* strains CITA 33 and CFBP 5530, and the *Prunus*-non-pathogenic strain CITA 44, were typed by PCR for 11 genes related to structural and regulatory components of the type III secretory system and for 21 genes coding for the type III effectors predicted in *Xap* (Hajri *et al.*, 2012; Garita-Cambronero *et al.*, 2016b). PCR reactions were performed according to the conditions proposed previously (Hajri *et al.*, 2012) with the exception of the T3SS genes, *hrpD5* and *hrpF*, and the T3Es genes *hpaA*, *xopAQ* and *xopZ*, for which new sets of primers were designed based in orthologues available in databases for *X. arboricola* using Primer-BLAST designing tool (Table 2). PCR amplifications with the primers for *hrpD5*, *hrpF*, *hpaA*, *xopAQ* and *xopZ* were performed in 20 µl of PCR reaction containing 1X PCR buffer (10 mM Tris-HCl, 50 mM KCl, 0.1% Triton X-100 [pH 9.0]); 0.5 µM of each primer; 0.25 U *Taq*

DNA polymerase (Biotools, Madrid, Spain); 0.2 mM each dNTP (Biotools Madrid, Spain); 1.5 mM MgCl₂ and 1.0 µg/µl of DNA template. PCR conditions consisted in an initial denaturation step of 94 °C for 2 min, 30 cycles of denaturation at 94 °C for 1 min, annealing at 60 °C for 1 min, extension at 72 °C for 2 min and a final extension step at 72 °C for 10 min.

Table 2. PCR primers used to amplify a partial region of some genes associated with type III secretion system (T3SS) and type III effector (T3E) genes in *X. arboricola*.

T3SS/T3E gene	Forward primer	Reverse primer	Fragment size (pb)
<i>hpaA</i>	ATGATCCGGCGCATTTTCG	GCGATGCTGACCCGGC	269
<i>hrpD5</i>	ATCGAGGTGGATGCAGATGG	CGGCAGGGAAGTCAGGTG	795
<i>hrpF</i> *	TCTACCTCTGACGGATGACG	GTCGCCCTGCGAGCC	516
<i>hrpF</i> [‡]	TCTACCTCTGACGGATGACG	GGTCGGCAAAGTCGTAGAGG	947
<i>xopAQ</i>	ATCGGGAGACACAGGGTGTA	CTTCTGAGGTAGCGGAC	146
<i>xopZ</i>	CATTCGTCGCGGATCAACAC	GAAAGCCGGAAGGATGTCT	196

*Primers used to amplify the ortholog of *hrpF* in pathogenic strains of *X. arboricola* pathovars *corylina*, *juglandis* and *pruni*.

[‡]Primers used to amplify the ortholog of *hrpF* in *X. arboricola* pv. *celebensis* and non-pathogenic strains of *X. arboricola*.

4.3.3 Pathogenicity tests

Pathogenicity tests on *Hordeum vulgare*, *Nicotiana benthamiana*, *N. tabacum*, *P. persica* (rootstock GF-305) and tomato (*Solanum lycopersicum*) were carried out for the six *Xap*-look-a-like strains as well as for the pathogenic strain of *Xap* CITA 33 and the non-pathogenic strain of *X. arboricola* CITA 44. Bacterial suspensions were prepared by inoculating a single bacterial colony on 30 ml of LB broth culture medium and maintained during 12-14 h at 26 °C on a rotatory shaker at 200 rpm. Subsequently, cultures were centrifuged at 6,350 g for 15 min and the pellet resuspended in sterile phosphate buffered saline (PBS pH= 7.5) to a final concentration of 1x10⁶ colony forming units (CFU)/ml. Bacterial suspensions were infiltrated in three leaves per plant using a syringe without needle and sterile PBS was used as blank control. All the infiltrated plants were kept in a grown chamber with high humidity, 16 h of light and 8 h of darkness at 26 °C and 22 °C, respectively. Infiltration results were addressed and graphically recorded at zero, seven, 14 and 21 days post inoculation (dpi). After 21 dpi, the infiltrated leaves were macerated on sterile distilled water and tenfold dilutions of the macerated were plated on YPGA (0.5% yeast extract, 0.5% bactopectone, 1.0%

glucose, 2.0% agar) plates, supplemented with 250 mg/l of cycloheximide. Colonies showing a *Xanthomonas*-like phenotype (mucoid, yellow colonies with a convex elevation and an entire margin) were selected and confirmed as *Xap*-look-a-like using a real time PCR protocol indicated above (Palacio-Bielsa *et al.*, 2011). Additionally, for *P. persica* rootstock GF-305, colonies were counted in order to determine the bacterial concentration in the inoculated tissue at the end of the assay (Ah-You *et al.*, 2007).

4.3.4 Genome sequencing and comparison

From the six *Xap*-look-a-like strains analyzed previously, one representative of each cluster, according to the MLSA analysis (Table 1; Figure 1), was selected for genome sequencing. Genome sequencing conditions and features for the genome sequence of CITA 44 have been discussed in a previous paper (Garita-Cambronero *et al.*, 2016b). In addition, in this study the genome features of the *Xap*-look-a-like strains CITA 14 and CITA 124 are described. For these two strains, the genome sequencing and assembly conditions have been previously announced and deposited at DDBJ, EMBL, GenBank databases under the accession numbers LXIB000000000 for CITA 14 and LXKK000000000 for CITA 124 (Garita-Cambronero *et al.*, 2016c).

The assembled draft genome sequence of strains CITA 14 and CITA 124 were automatically annotated using the NCBI's prokaryotic annotation pipeline (Tatusova *et al.*, 2013). Signal peptides and transmembrane domains were predicted using the signalP 4.1 (Petersen *et al.*, 2011) and the TMHMM 2.0 (Krogh *et al.*, 2001) servers, respectively. The assignment of genes to each cluster orthologous group (COG) and its Pfam domain was performed with the NCBI's conserved domain database using an expected value threshold of 0.001 (Marchler-Bauer *et al.*, 2014). The circular genome maps of the draft genome sequences of *X. arboricola* strains CITA 14 and CITA 124 representing the COG categories of the genes, were constructed using CGView (Stothard & Wishart, 2005). The contigs of both strains were arranged by Mauve (Darling *et al.*, 2004, 2010) using the complete genome sequence of *X. arboricola* pv. *juglandis* Xaj417 as the reference.

The genome sequence variation among CITA 14, CITA 124 and the publicly available genomes of *X. arboricola* (strains 3004, CFBP 7634, CFBP 7651 and CITA 44) (Ignatov *et al.*, 2015; Cesbron *et al.*, 2015; Garita-Cambronero *et al.*, 2016a), *X.*

arboricola pv. *celebensis* (NCPBPB 1630 and NCPBPB 1832) (Harrison *et al.*, 2016), *X. arboricola* pv. *corylina* (NCCB 100457) (Ibarra Caballero *et al.*, 2013), *X. arboricola* pv. *juglandis* (CFBP 2528, CFBP 7179, Xaj2 and Xaj417) (Higuera *et al.*, 2015; Pereira *et al.*, 2015; Cesbron *et al.*, 2015) and *X. arboricola* pv. *pruni* (CITA 33, IVIA 2626.1, MAFF 301420 and MAFF 301427) (Garita-Cambronero *et al.*, 2014, 2016a), were determined by using a sequence-based approach and a sequence content approach (Snipen & Ussery, 2010).

The sequence-based approach, which consists in the comparison among genome sequences based on sequence alignment and evolutionary analysis based in the shared protein coding sequences (CDS) were determined using Roary (Page *et al.*, 2015). Contigs of the 17 genome sequences of *X. arboricola* were ordered by Mauve as mentioned above. Afterwards, all the genome sequences were automatically annotated using PROKKA (Seemann, 2014) and used as the input for the search of shared homologous genes among the studied strains.

The concatenated sequences of the genes that composed the core genome sequence of the 17 analyzed strains were aligned using the PRANK alignment algorithm (Löytynoja & Goldman, 2008) and subsequently, a maximum likelihood tree using 1,000 bootstrap resamplings, was constructed to determine the phylogenetic position of the strains CITA 14 and CITA 124 within *X. arboricola*. Maximum likelihood tree was performed with RaxML (Stamatakis, 2014) and the dendrogram obtained was visualized using Dendroscope (Huson *et al.*, 2007).

Additionally, the gene-content comparison, which is based in determining the distance among the 17 genomes using the presence/absence of the genes contained in the pangenome of *X. arboricola*, were performed using the R implemented package for microbial-pangenomics micropan (Snipen & Liland, 2015). The CDS were obtained from nucleotide genome sequences of the 17 strains mentioned above available in the NCBI database using Prodigal v2.6.1 (Hyatt *et al.*, 2010). To determine the similarity of the proteins within and across the genomes, a reciprocal all-against-all BLAST search was performed using the blastAll package. BLAST distance between sequences was obtained using the bDist package. A hierarchical clustering was performed using bClust, and a complete linkage function was selected with a liberal threshold of 0.80. A similarity matrix, with the number of protein sequences contained in each cluster for

each genome, was constructed using the panMatrix function and Jaccard distance was calculated. The matrix was used to perform a principal component analysis for showing how the genomes are distributed in the space according to the two first principal components which revealed the dominant differences between them, and this was computed using the panpca and plotScores functions. Additionally, the similarity of the analyzed genomes was represented in a pan-genome tree computed according to the method described previously by (Snipen & Ussery, 2010), using the panTree function. Tree construction was based in the distance between genomes according to the Manhattan distance. Bootstrap values were calculated by re-sampling the columns of the similarity matrix and the re-clustering of these data, therefore, the bootstrap value represented was the percentage of the re-sampled trees that showed a similar node.

Additionally, the profile of gene clusters (potential groups of orthologous genes) associated with tonB-dependent transporters (TBDTs), the sensors of the two-component regulatory system (STCRs), the methyl accepting chemotaxis proteins (MCPs), the components of the flagella, the type IV pilus, the non-fimbrial adhesins, the components associated to the production of xanthan, the quorum-sensing regulation, the repertoire of cell-wall degrading enzymes, the components of the type II, III and IV secretion systems, as well as the type three effectors, previously reported in *Xanthomonas* spp. (da Silva *et al.*, 2002; Filloux, 2004; Vorhölter *et al.*, 2008; Wang *et al.*, 2008; Chevance & Hughes, 2008; He & Zhang, 2008; White *et al.*, 2009; Subramoni *et al.*, 2010; Szczesny *et al.*, 2010; Mhedbi-Hajri *et al.*, 2011; Potnis *et al.*, 2011; Ryan *et al.*, 2011; Guo *et al.*, 2011; Hajri *et al.*, 2012; Vandroemme *et al.*, 2013; Li *et al.*, 2014; Guglielmini *et al.*, 2014; Cesbron *et al.*, 2015; Dunger *et al.*, 2016; Nascimento *et al.*, 2016) were determined for strains CITA 14 and CITA 124.

The presence of the putatively virulence-associated plasmid pXap41 (Pothier *et al.*, 2011c) was also evaluated in the analyzed genomes based in the nucleotide sequence similarity, and graphically represented using the BLAST Ring Image Generator (BRIG) tool (Alikhan *et al.*, 2011) and blastn was used for the sequence comparative analysis with an expected value threshold of 0.001.

4.3.5 A molecular tool to differentiate *X. arboricola* pv. *pruni* from atypical strains of *X. arboricola* associated with *Prunus* spp.

In order to discriminate *Xap* from other pathovars, a partial sequence of the *xopE3* gene that is enclosed in pXap41 plasmid, described as specific for *Xap* (Pothier *et al.*, 2011c), was used. Sequences of the *xopE3* gene available in GenBank database from strains CITA 33 (GenBank locus tag DK27_00095), IVIA 2626.1 (AN652_04270), MAFF 301420 (XPR_2580), MAFF 301427 (XPN_1257) and CFBP 5530 (XAP_pXAP410005) were aligned with ClustalW and the consensus sequence, generated using Bioedit v. 7.2.5, was used as template for *xopE3* primers and probe design using the ABI PRISM Primer Express software v. 2 (Applied Biosystems, Foster City, CA). Specificity of the primers was firstly evaluated *in silico* using the Primer-BLAST tool available at NCBI. Graphical representation of the probe and the primers hybridization was performed in a set of *X. arboricola* genome sequences using the BRIG software as previously mentioned.

Real-time PCRs were conducted using the same conditions described previously for reactions using primers for ABC transporter (Palacio-Bielsa *et al.*, 2011). Briefly, reactions were conducted in a total volume of 25 µl containing, 12.5 µl of GoTaq probe qPCR MasterMix (Promega), 0.4 µM of each primer and 150 nM of TaqMan probe and 5 µl of sample. Real-time PCR amplifications were conducted in an ABI 7500 Fast thermocycler (Applied Biosystems, Foster city, CA) and consisted of an initial denaturation step of 95 °C for 5 min followed by 45 cycles, each one of 1 min at 95 °C and 1 min at 59 °C.

4.3.6 Specificity of the real-time PCR test

Specificity of the real-time PCR test was assessed in 99 bacterial strains, which comprised 54 strains of *Xap*, seven strains of *Xap*-look-a-like, ten strains from other pathovars of *X. arboricola*, 11 strains from other species of *Xanthomonas*, eight strains from other phytopathogenic bacteria and nine strains from the natural microbiota of *Prunus* spp. (Table 1). Bacterial suspensions of 10⁸ CFU/ml were treated at 95 °C during 10 min and used for real-time PCR reactions and sterile distilled water was used

as negative control. Additionally, a real-time PCR protocol previously described (Palacio-Bielsa *et al.*, 2011) was also applied on all the 99 bacterial strains tested for *xopE3* gene.

4.3.7 Sensitivity of the real-time PCR test

Serial dilutions from 10^8 to 10 CFU/ml of a 48 h LB broth culture, of strain CITA 33 were prepared on sterile distilled water and heat-treated (95 °C for 10 min) for real-time PCR reactions. Additionally, a serial dilution of pure bacterial DNA, extracted using the QIamp DNA miniKit (Qiagen) as mentioned above, ranging from 10^8 pg/ μ l to 0.001 pg/ μ l was also prepared on sterile distilled water. For amplification, 5 μ l of each dilution was used as template, according to the protocol described previously. Seven replicates of each sample were evaluated in each experiment and the experiment was repeated in three independent assays. Appropriate negative controls containing no bacteria or no DNA were subjected to the same procedure. The limit of detection of the test was defined as the lowest target amount giving positive results in at least 15 of the 21 total reactions tested in the three independent assays (Caraguel *et al.*, 2011). Analysis of variance was used to test for differences in the threshold cycles (C_{Ts}) at each bacterial concentration in the three independent assays. Statistical analyses were performed by using Statgraphics Plus v.5.1 software. The amplification efficiency of the protocol for each kind of sample was calculated as described previously (Palacio-Bielsa *et al.*, 2011, 2015b). Linear regression curves representing the C_{Ts} of each reaction were plotted against the logarithmic values of bacterial or DNA concentration. The slope of the curves (k) was used to determine the amplification efficiency (E) according to the equation $E = 10^{[-1/k]}$, where $E = 2$ corresponded to 100% efficiency.

4.4 Results

4.4.1 Characterization of atypical strains of *X. arboricola* associated with *Prunus* spp.

A total of 40 strains isolated from *Prunus* spp. and phenotypically similar to *Xanthomonas arboricola* were initially identified as *Xap* by using a real time PCR protocol (Palacio-Bielsa *et al.*, 2015b). Additionally, the plasmid pXap41, which is considered a unique feature of *Xap*, was not detected by PCR amplification of the genes *repA1*, *repA2* and *mobC* (Pothier *et al.*, 2011c), from strains CITA 14, CITA 42, CITA 49, CITA 51, CITA 124 and CITA 149 and, therefore, they were not considered as *Xap* but *Xap*-look-a-like strains.

In order to further characterization, six *Xap*-look-a-like strains mentioned above and the reference non-pathogenic strain CITA 44 (Garita-Cambronero *et al.*, 2016a), were analyzed using a MLSA scheme based in partial sequences of the genes *dnaK*, *fyuA*, *gyrB* and *rpoD* (Young *et al.*, 2008). The Maximum likelihood analysis of the concatenated sequences revealed that none of the *Xap*-look-a-like strains was consistently clustered with any of the reference strains that belong to the described pathovars of *X. arboricola*. Contrary to this, they were distributed in three separated clusters, one composed by strains CITA 14 and CITA 149, another composed by strains CITA 44 and CITA 49, and a third one composed by strains CITA 42, CITA 51 and CITA 124. These clusters were located in a basal phylogenetic position with respect to most of the strains used as reference. Sequence analysis of the concatenated sequence (2,836 nucleotide positions) revealed a mean similarity of $98.30 \pm 0.2\%$ between the *Xap*-look-a-like strains and the remaining ten strains of *X. arboricola* (Figure 1).

Figure 1. Maximum likelihood tree of concatenated sequences of the genes *dnak*, *fyuA*, *gyrB* and *rpoD* of non-pathogenic *X. arboricola* strains isolated from *Prunus* spp. For comparative purposes pathogenic strains of *X. arboricola* pv. *pruni* isolated from *Prunus* spp. and *X. arboricola* strains isolated from other hosts are included. *X. citri* subsp. *citri* was used as outgroup. Bootstrap values of 1,000 replicates are represented above or below the branches. Selected strains for subsequent whole genome sequencing are in bold.

4.4.2 T3SS and T3Es repertoire in *Xap*-look-a-like strains

Conventional PCR typing of 32 genetic determinants of the T3SS and its related T3Es brought out a variable repertoire in the seven *Xap*-look-a-like strains. The structural and regulatory components of the T3SS were only detected in strains CITA 14 and CITA 149, which harbored all the 11 components tested. In the same manner, only strains CITA 14 and CITA 149 harbored five and two T3Es, respectively. On the other hand, CITA 42, CITA 49, CITA 51 and CITA 124 did not harbor any of these genes. As expected, *Xap* strains CITA 33 and CFBP 5530 showed positive amplification from all the analyzed genes, meanwhile the non-pathogenic strain CITA 44 resulted negative (Table 3).

4.4.3 Pathogenicity of the *Xap*-look-a-like strains

In addition to the variability found in the T3SS components and its related effectors, the ability of the *Xap*-look-a-like strains to cause disease symptoms after bacterial infiltration on leaves of barley, *N. benthamiana*, *N. tabacum*, tomato and the susceptible peach rootstock GF-305 was evaluated. None of the assayed strains were able to cause disease symptoms on barley, with the exception of strain CITA 14, which caused necrosis and chlorosis in the infiltrated zone. Strains CITA 14 and *Xap* strain CITA 33 were able to cause necrosis and chlorosis in *N. benthamiana* and *N. tabacum*. On the contrary, strains CITA 42, CITA 49, CITA 51, CITA 124, CITA 149 and the *Prunus*-non-pathogenic strain CITA 44 only showed a chlorosis effect after 21 dpi. In tomato, all the assayed strains, with the exception of CITA 44, were able to cause necrosis and, in most of the cases, the necrotic area was surrounded by a chlorotic halo. Infiltration on peach rootstock GF-305 showed that only CITA 44 did not cause damage on the leaves; however, the remaining strains CITA 14, CITA 42, CITA 49, CITA 51, CITA 124 and CITA 149 caused necrosis in the infiltrated area after 7 dpi, but these necrotic spots did not expand beyond this zone and were different to typical bacterial spot symptoms.

On the other hand, the *Xap* strain CITA 33, caused necrosis in the infiltrated area and these necrotic zones were surrounded by a chlorotic halo (Supplementary Figure 1). In addition, in GF-305, variation in bacterial populations on the infiltrated leaves was determined after 21 dpi and all the *Xap*-look-a-like strains, as well as strain CITA 44, showed a reduction in their bacterial populations leading to concentrations equal or lower than 10^5 CFU/ml. On the other hand, for strain CITA 33, the bacterial concentration increased from 10^6 to 10^{10} CFU/ml by the end of the assay (Table 3). Positive results in real-time PCR analysis of the isolated colonies after 21 dpi, using the standardized protocol for *Xap* detection (Palacio-Bielsa *et al.*, 2011, 2015b), corroborated that the re-isolated strains corresponded to the same inoculated at the beginning of the study.

Table 3. Components of the type three secretion system, repertoire of the type three effectors, presence of the plasmid pXap41 and pathogenicity of *X. arboricola* strains isolated from *Prunus* spp.

Gene/Strains		CITA 14	CITA 42	CITA 44	CITA 49	CITA 51	CITA 124	CITA 149	CFBP 5530 ^P	CITA 33 ^P
Components of the type III secretion system	<i>hrcC</i>									
	<i>hrcJ</i>									
	<i>hrcN</i>									
	<i>hrcR</i>									
	<i>hrcS</i>									
	<i>hrcT</i>									
	<i>hrcU</i>									
	<i>hrcV</i>									
	<i>hrpB1</i>									
	<i>hrpD5</i>									
	<i>hrpF</i>									
Type III effectors	<i>avrBs2</i>									
	<i>avrXccA2</i>									
	<i>hpaA</i>									
	<i>hrpW</i>									
	<i>xopA</i>									
	<i>xopAF</i>									
	<i>xopAH</i>									
	<i>xopAI</i>									
	<i>XopAQ</i>									
	<i>xopE2</i>									
	<i>xopE3</i>									
	<i>xopF1</i>									
	<i>xopG</i>									
	<i>xopK</i>									
	<i>xopL</i>									
	<i>xopN</i>									
	<i>xopQ</i>									
	<i>xopR</i>									
	<i>xopV</i>									
	<i>xopX</i>									
	<i>xopZ</i>									
pXap41	<i>repA1</i>									
	<i>rpoA2</i>									
	<i>mobC</i>									
Pathogenicity	<i>H. vulgare</i>	N, C	NS	NS	NS	NS	NS	NS	ND	NS
	<i>N. benthamiana</i>	N, C	C	C	MC	C	MC	C	ND	N, C
	<i>N. tabacum</i>	N	C	C	C	C	C	C	ND	N
	<i>S. lycopersicum</i>	N	MN, C	NS, MC	N, C	N, C	N, C	N, C	ND	N, C
	<i>P. persica</i> (GF 305)	MN	N	NS	N	N	MN	N	ND	N, C
	CFU/ml 21 dpi	0-10 ⁵	10 ¹ -10 ⁵	10 ² -10 ⁴	10 ² -10 ⁴	10 ³ -10 ⁵	10 ³ -10 ⁵	10 ³ -10 ⁵	ND	10 ⁶ -10 ¹⁰
	Real-time PCR*	+	+	+	+	+	+	+	ND	+

Positive/negative PCR amplification are represented in grey or white respectively; N: necrosis; C: chlorosis; NS: Not visible symptoms; MN: mild necrosis; MC: mild chlorosis; ND: no data. P: *Prunus* pathogenic strains of *X. arboricola* pv. *pruni* (*Xap*), the remaining tested strains were considered as atypical *Xap*-look-alike strains. *: according to the protocol described by Palacio-Bielsa *et al.* 2011; 2015b.

4.4.4 General features of whole genomes of *X. arboricola* strains CITA 14 and CITA 124

According to the results obtained in the MLSA analysis, one member of each one of the three clusters (CITA 44, CITA 14 and CITA124) observed for the *Xap*-look-a-like strains, was selected for whole genome sequencing (Figure 1). The draft genome sequence of the non-pathogenic strain CITA 44 has been reported and analyzed previously (Garita-Cambronero *et al.*, 2016b). Moreover, main sequencing and structural features of CITA 14 and CITA 124 genome sequences have been previously announced (Garita-Cambronero *et al.*, 2016c).

Draft genome sequence of CITA 14 was 4,864,444 bp in length with an average GC content of 65.6%; 3,870 over 3,974 genes predicted were identified as protein coding genes and a putative function was assigned to 2,991 of them (Supplementary Table 1). Additionally, this strain presented 4 rRNAs and 53 tRNAs. In the case of CITA 124, its draft genome sequence was 4,752,241 bp in length with an average GC content of 65.8%. A total of 4,004 genes were predicted and, among them, 3,798 were identified as protein coding genes with a putative function assigned to 2,838 of them. In addition, 3 rRNA and 50 tRNA genes were predicted (Table 4). From the total of genes with a predicted function, a COG functional category was assigned to 3,090 and 2,668 genes in CITA 14 and CITA 124, respectively. Those genes related to the amino acid and carbohydrate transport and metabolism as well as those associated with translation and ribosomal structure and biogenesis were predominant for both strains (Figure 2). In addition, CITA 14 and CITA 124 presented 3,320 and 2,668 genes, respectively, with a protein domain in the Pfam database. Finally, a total of 942 and 625 genes of CITA 14 presented transmembrane helices and peptide cleavage signals, respectively. Meanwhile, in CITA 124, 954 genes with transmembrane helices and 569 genes with peptide cleavage sites were predicted (Table 4, Supplementary Table 1).

A comparative gene content analysis was performed among the genome sequences of the two non-pathogenic *Xap*-look-a-like strains, CITA 14 and CITA 124, and 15 genome sequences of other pathogenic and non-pathogenic strains of *X. arboricola*. As result, a total of 7,074 potential groups of homologous genes were found in the 17 analyzed genomes, from which 2,714 were shared by all the *X. arboricola* strains and comprised the core group of orthologous genes. CITA 14 and CITA 124 presented 76

and 124 unique cluster genes, respectively, which distinguished them from the other 15 strains (Figure 3A; Supplementary Table 2). Strains CITA 14 and CITA 124, the pathogenic *Xap* strains CITA 33 and IVIA 2626.1 and the non-pathogenic CITA 44, all isolated from *Prunus* spp. in Spain, shared 3,103 groups of homologous genes. A total of 889 cluster genes were found only in the non-pathogenic strains, meanwhile 708 cluster genes were found in the two Spanish strains of *Xap* (Figure 3B). Additionally, 236 CDS, 17 CDS and 131 CDS, were unique for the pathovar *corylina*, *juglandis* and *pruni*, respectively (Supplementary Table 2).

The mean of the gene content similarity among the analyzed genomes according to the Jaccard distance distribution was 0.23, which means that the analyzed genomes shared a mean of 77% of their gene content, meanwhile, the remaining 23% is unique for each one of them (Supplementary Figure 2A).

Figure 2. Graphical circular representation of the draft genome of the non-pathogenic of *X. arboricola* strains CITA 14 and CITA 124. The contigs were arranged by Mauve, using the genome sequence of *X. arboricola* pv. *juglandis* strain Xaj417 as reference. COG categories were assigned to predicted genes using the NCBI's conserved domain database. Circular map was constructed using CGview. From outside to center: Genes on forward strand; genes on reverse strand; GC content; GC skew.

Table 4. Genome sequence information and statistics of the non-pathogenic strains of *X. arboricola* strains CITA 14 and CITA 124, isolated from *Prunus*.

Property/Attribute	CITA 14	CITA 124
	Value	Value
Sequencing platform	Ion Torrent PGM	Ion Torrent PGM
Fold coverage	100x	50x
Assemblers	CLC and MIRA 4.0	CLC and MIRA 4.0
Genome annotation	NCBI-PGAP	NCBI-PGAP
Locus tag	A7D01	A7D35
Genbank ID	LXIB000000000	LXKK000000000
Genome size (bp)	4,864,444	4,752,241
DNA G+C (%)	65.60	65.80
Total genes	4,061	4,086
Protein coding genes	3,870	3,798
RNA genes	87	82
Pseudo genes	104	206
Genes with function prediction	2,991	2,838
Genes assigned to COGs	3,090	2,668
Genes with Pfam domains	3,320	2,668
Genes with signal peptides	625	569
Genes with transmembrane helices	942	954
CRISPR repeat unit	1	0

Figure 3. Potential groups of orthologous genes present in *X. arboricola*. Core, shell and cloud groups of orthologous genes shared by 17 genome sequences of *X. arboricola* (A). Venn diagram showing the groups of orthologous genes share by five genome sequences of pathogenic (CITA 33 and IVIA 2626.1) and non-pathogenic (CITA 14, CITA 44 and CITA 124) strains of *X. arboricola* isolated from *Prunus* spp. (B).

Based in the gene cluster content of each genome, a principal component analysis showed that only 41.0% of the total difference is due to the variation by the two first principal components (Supplementary Figure 2B). Three distinct clusters were elucidated, one of them was comprised by the non-pathogenic strains (CITA 14, CITA 44, CITA 124, CFBP 7634 and CFPB 7651) and the strains with low-pathogenic activity (3004, NCPPB 1630 and NCPPB 1832) which cause disease on barley and banana. The pathogenic strains of the pathovar *pruni* formed another cluster that comprised two subgroups, one formed by the strains isolated in Spain (CITA 33 and IVIA 2626.1) and another by the strains MAFF 301420 and MAFF 301427 isolated in Japan. Strains from the pathovar *juglandis* (CFBP 2528, CFBP 7179, Xaj2 and Xaj417) formed the third group. Finally, the strain NCCB 100457 from the pathovar *corylina* tended to group together with the strains isolated from *Juglans* spp. (Supplementary Figure 2B).

The difference in the gene content cluster was illustrated using a pan-genome tree (Snipen & Liland, 2015) after computing the distance among the genomes using the Manhattan distance algorithm. The pan-genome tree for the 17 analyzed genomes (Figure 4) showed the same clustering organization that was visualized previously with the principal component analysis. Besides this, a division of the cluster comprised by the low-pathogenic and non-pathogenic strains was shown. A first group was composed by those strains that harbored components of the T3SS and T3Es, isolated from banana (NCPPB 1630 and NCPPB 1832), walnut (CFBP 7634 and CFBP 7651) and peach (CITA 14). A total of 10 CDS differentiated this group from all the remaining clusters observed. The second group was comprised by the strain 3004 isolated from barley, and the strains CITA 44 and CITA 124 isolated from *Prunus*, and 27 CDS differentiated this cluster from all the other analyzed strains (Supplementary Table 2). The same strains grouping and distribution was obtained using a Maximum likelihood phylogenetic analysis based in the concatenated sequence of the genes that comprised the core genome sequence of the 17 analyzed strains according to a sequence based methodology recently described (Page *et al.*, 2015) (Supplementary Figure 3).

Figure 4. Pangenome tree for pathogenic and non-pathogenic strains of *Xanthomonas arboricola*. Tree construction was based in the distance between genomes according to the Manhattan distance. Bootstrap values over 50% are showed at the branch points.

4.4.5 Genes associated with pathogenicity in *X. arboricola* strains:

In addition to the gen content comparison, the profiles of genetic components associated with pathogenesis were determined for CITA 14 and CITA 124 and compared to those in other strains of *X. arboricola* isolated from *Prunus* in Spain, especially in the two pathogenic strains of *Xap*, CITA 33 and IVIA 2626.1, and in the non-pathogenic strain of *X. arboricola* CITA 44.

Regarding the profiles of cell degrading enzymes, a total of ten genes that encoded for pectolytic enzymes were found in the five strains. CITA 14 and CITA 124 showed seven and eight of these genes, respectively. For these enzymes, two orthologs, NP_635517.1 and NP_635516.1, were shared by the non-pathogenic strains, meanwhile the degenerated pectate lyase AAM37225.1 was only found in *Xap* strains CITA 33 and

IVIA 2626.1 (Supplementary Table 2). Regarding the profile of cellulolytic enzymes, nine enzymes were shared over 11 genes found in all the strains. In this case, only the presence of the cellulolase AAM38359, described in *X. citri* subsp. *citri* 306, differentiated non-pathogenic strains from *Xap* (Supplementary Table 3). In the case of the hemilcellulolytic enzymes, a total of 11 orthologs were found in the *Prunus*-associated strains; pathogenic strains of *Xap* were differentiated from the non-pathogenic strains due to the presence of the genes that encoded the xylanase NP_638385.1 and the xylosidase/arabinosidase NP_637752.1, both described previously in *X. campestris* pv. *campestris* strain ATCC 33913. Finally, orthologous genes for the virulence associated lipases NP_638797.1 and AAO29541.1 from *X. campestris* pv. *campestris* ATCC 33913 and *Xylella fastidiosa* strain Temecula, were found in the five genomes (Supplementary Table 3).

The profiles of genes related to sensing and chemotaxis varied among the analyzed strains. Four of the five Spanish *Prunus*-isolated strains presented the same gene profile for those genes associated with chemotaxis. However, CITA 124 did not harbor homologous genes to *cheD* (AAM36751.1), *cheZ* (AAM36793.1) and *cheA* (AAM36792.1) described in *X. citri* subsp. *citri* 306. Variants in other sensing mechanisms such as TBDTs were found in the *Prunus*-associated strains of *X. arboricola*; from the 17 TBDTs encoding genes found, those homologues of the proteins NP_635515.1, NP_635700.1, NP_635699.1 and NP_639391.1, initially described in *X. campestris* pv. *campestris* ATCC 33913, differentiated non-pathogenic strains from *Xap*. Additionally, a large repertoire of 60 genes associated with STCRs was found in the analyzed genomes and, from them, 55 were shared for all the strains isolated from *Prunus*. In addition, the STCRs AAM36681.1, AAM35218.1, AAM37649.1 and NP_637535.1 were only present in the non-pathogenic strains CITA 14, CITA 44 and CITA 124. Finally, from a total of 26 MCPs genes, 11 were found in all *Prunus*-associated strains, but the absence of an ortholog to CAJ23610.1, described in *X. campestris* pv. *vesicatoria* 85-10, in strains CITA 14 and CITA 124 differentiated them from the remaining strains (Supplementary Table 3).

Besides to those genes related to environmental sensing, variations in some other genes associated with the initial steps of the pathogenesis process, such as motility, attachment, biopolimerization of the xanthan gum and the inter-cellular cross-talk process controlled by the quorum-sensing system, were also found (Supplementary

Table 3). Pathogenic and non-pathogenic strains of stone fruit and almond shared 35 orthologs associated with molecular components of the flagellar system. The exception to this was the non-pathogenic strain CITA 124, which did not have homologous genes to the flagellar components of *X. citri* subsp. *citri* 306, *flhF* (AAM36797.1), *fliH* (AAM36814.1), *fliJ* (AAM36812.1) and *motB* (AAM38537.1). In addition, an interesting polymorphism was observed in the flagellin protein, encoded by *fliC*, of non-pathogenic strains. In CITA 14 and CITA 124, this problem was identical to protein WP_024939608.1, which has been previously associated with all the non-pathogenic strains of *X. arboricola* or with low pathogenic strains of the pathovar *celebensis*. Meanwhile, pathogenic strains of *Xap* harbored a flagellin protein identical to protein WP_039814449.1, which present a substitution of aspartic acid for valine in the amino acid 45 of the N-terminal region that has been associated to pathogenic strains in other species of *Xanthomonas* (Sun *et al.*, 2006; Cesbron *et al.*, 2015).

Another bacterial structure related to motility, as well as to attachment, is the type IV pilus. The pathogenic *Xap* strains CITA 33 and IVIA 26262.1 harbored 25 orthologs to the 31 genes described in *X. citri* subsp. *citri* meanwhile the non-pathogenic strains CITA 14, CITA 44 and CITA 124 were differentiated for the absence of orthologs to the genes *fimA* (AAM38084.1), *fimT* (AAM37516.1), *pilV* (AAM37515.1), *pilW* (AAM37514.1), *pilX* (AAM37513.1) and *pilY1* (AAM37512.1). With regard to the attachment function carried out by the non-fimbrial adhesins, most of the non-pathogenic strains shared five of the six genes found, with the exception of CITA 14, which did not harbor a homolog to *fhaB1* of *X. campestris* pv. *vesicatoria* 85-10 (CAJ23537.1). Presence of a homologous gene to *fhaB2* (CAJ23538.1) of *Xanthomonas campestris* pv. *vesicatoria* differentiated *Xap* from non-pathogenic strains.

Pathogenic and non-pathogenic strains of *X. arboricola* isolated from *Prunus* shared the same profile of xanthan-associated genes, which are involved in bacterial attachment and biofilm formation. None of the strains had homologous sequences to *gumG* (NP_637802.1) which is found in other xanthomonads (Lee *et al.*, 2005). Regarding quorum sensing system, which is associated with the regulation of the pathogenic activity, from the 12 genes related to this process in *Xanthomonas* (He & Zhang, 2008), all the analyzed strains, with the exception of CITA 14, harbored the same gene pattern conformed by 11 genes. In addition, CITA 14 harbored an ortholog to the

transcriptional regulator NP_636589.1 described in *X. campestris* pv. *campestris* ATCC 33913 (Supplementary Table 3).

Bacterial type II, III and IV secretory systems (T2SS, T3SS and T4SS), which are related to the secretion of proteins and DNA, also play a crucial role in pathogenesis (Ryan *et al.*, 2011). Regarding to T2SS, CITA 14 and CITA 124 presented 19 and 18 orthologs, respectively, of the 23 genes associated with *xcs* and *xps* T2SS gene clusters described in *Xanthomonas* (Filloux, 2004; Szczesny *et al.*, 2010). The only difference among non-pathogenic strains and *Xap* was the presence in the latter of an ortholog to the gene *xcsK* (NP_638764.1) (Supplementary Table 3).

Pronounced differences were observed among pathogenic and non-pathogenic strains regarding the gene profile associated with T3SS and its related effectors, as well as with T4SS. T3SS-related gene profile in CITA 14 was comprised by 24 orthologs of the 28 T3SS described in *Xanthomonas*, and this profile was the same observed in pathogenic strains of *X. arboricola* (Cesbron *et al.*, 2015; Garita-Cambronero *et al.*, 2016b). On the other hand, CITA 124 only harbored four T3SS-related genes (*hpaS*, *hpaR2*, *hrpG* and *hrpX*) which correspond to the regulons of this secretory system (Jacobs *et al.*, 2015). For this strain as for the non-pathogenic strain CITA 44, none of the genes that composed the macromolecular structure of the T3SS were found (Supplementary Table 3). Regarding to the T3SS related effectors, from the 61 T3Es described in *Xanthomonas*, the genome sequence of CITA 14 presented a total of six orthologs to the genes *avrBs2*, *hpaA*, *hrpW*, *xopA*, *xopF1* and *xopR*, meanwhile genome sequence of CITA 124 did not present any of these effectors. Moreover *X. arboricola* strains isolated from *Prunus* showed variants in the number of T4SSs. Most of the strains, regardless of their pathogenic activity, contained ten of the 12 components associated with the VirB/VirD4 T4SS of *Agrobacterium tumefaciens* (Christie, 2004). In addition, the absence of orthologs to the core components associated with the type IV conjugation cluster *tfc*, described in *Haemophilus influenzae*, in the non-pathogenic strains CITA 14 and CITA 124 differentiated them from the *Xap* strains.

Finally, comparative sequence analysis among the nucleotide sequence of the plasmid pXap41 and the draft genome sequence corroborated the absence of this plasmid in both non-pathogenic strains (Figure 5).

Figure 5. Presence of the plasmid pXap41 in pathogenic and non-pathogenic strains of *X. arboricola*. Comparative sequence analysis was performed using Blastn with an expected value threshold of 0.001 and graphically represented by the BRIG tool. Each concentric circle represents one of the analysed genomes.

4.4.6 The real-time PCR test for *xopE3* permitted to differentiate *Xap* from atypical strains of *X. arboricola* isolated from *Prunus* spp.

In silico analysis of the primers XopE3F (5'-TCAGCGATCACGCATCCA-3'), XopE3R (5'-CGCACCAGATCGACAAACAC-3') and the probe XopE3p (5'-CATGCGCAGGCCGCACAT-3'), indicated that they were able to amplify in *X. arboricola* the gene *xopE3* only in those sequences from the pathovar *pruni*. Sequence analysis on the available complete genome sequences of *X. arboricola* showed that *xopE3* was only present in those strains of the pathovar *pruni* (Figure 6). In addition, this set of primers and the designed probe were also able to amplify the *XopE3* gene in other species of *Xanthomonas* such as *X. citri* subsp. *citri* 306 and *X. fuscans* subsp. *fuscans* strain 4834-R (Supplementary Figure 4).

Besides the nucleotide sequence-based analysis, the specificity of the real-time PCR assay was conducted by testing the protocol on the bacterial strains listed in Table 1. Among the *X. arboricola* strains, only those determined previously as *Xap*, by the

presence of the plasmid pXap41 and with a positive result for the standardized real-time protocol that amplify a gene *ftsX* related to the ABC transporter in *Xap* (Palacio-Bielsa *et al.*, 2011, 2015b), presented consistent positive results. None of the seven *Xap*-look-a-like strains (CITA 14, CITA 42, CITA 44, CITA 49, CITA 51, CITA 124 and CITA 149) were amplified. Undesired specific PCR results for *xopE3* were observed from three strains of *X. axonopodis* (Xp-2, NCPPB 381 and IVIA 151835DA), one strain of *X. campestris* (IVIA 2734-1), three strains of *X. citri* subsp. *citri* (306, IVIA 2889-1 and IVIA 3026-1) and the strain 3619-1 of *X. vesicatoria* (Table 1). When the analyzed strains were amplified using the real-time PCR protocol for the ABC transporter-associated gene *ftsX* (Palacio-Bielsa *et al.*, 2011), positive results were obtained with all the strains of *Xap*, the seven strains of *Xap*-look-a-like, two strains of *X. arboricola* pv. *corylina* (CFBP 1846 and IVIA 3978) and the strain of *X. citri* subsp. *citri* 306. Double positive PCR results, using *xopE3* or ABC primers, were only observed for all the *Xap* strains but also for the strain 306 of *X. citri* subsp. *citri* (Table 1).

Figure 6. *In silico* representation of the presence of *xopE3* and *ftsX* and the hybridization zone for the primers and probes used for real-time PCR amplification in *Xanthomonas arboricola*. Comparative sequence analysis was performed using Blastn with an expected value of 0.001. The circular graphic has been constructed using BRIG. Each concentric circle represents one of the analysed genomes.

No significant differences were found among the three independent assays conducted to determine the sensitivity of the real-time PCR protocol to amplify *xopE3* using heat-treated cells or purified DNA as samples. Calibration curves, obtained from serial dilutions of heat-treated cells of *Xap* strain CITA 33, demonstrated that the real-time PCR assay showed a sensitivity of 10 CFU/ml or 100 pg/μl of DNA, with a PCR efficiency of 2.2 ± 0.22 or 1.8 ± 0.03 for bacterial cells or purified DNA from strain CITA 33, respectively (Figure 7).

Figure 7. Calibration curves for detection of *xopE3* in *Xanthomonas arboricola* pv. *pruni*. Calibration curves have been obtained from dilution series of purified DNA (A) and bacterial cells (B) of *X. arboricola* pv. *pruni* strain CITA 33. Real time PCR amplification was performed in three independent assays using the primers XopE3F/R and the TaqMan probe XopE3p.

4.5 Discussion

The results obtained in the initial characterization of the *X. arboricola* strains isolated from *Prunus* spp. pointed out that one of the most widely used real-time PCR protocol for detecting *Xap* (Palacio-Bielsa *et al.*, 2011 2015b) was not able to differentiate bacterial strains of this pathovar from those atypical strains of the same species, which are part of the *Prunus* microbiota. Actually, *in silico* analysis, based in the nucleotide sequence comparison among the available genome sequences of *X. arboricola* and the target genomic regions proposed for the identification of *Xap* in a variety of other published PCR protocols (Park *et al.*, 2010; Pothier *et al.*, 2011a) (Supplementary Figure 5), demonstrated that none of them could be able to discriminate between these two groups of *Prunus*-associated strains.

The MLSA analysis conducted with the housekeeping genes *dnak*, *fyuA*, *gyrB* and *rpoD*, resulted useful to characterize typical and atypical strains of *X. arboricola* as proposed in recent articles (Essakhi *et al.*, 2015; Garita-Cambronero *et al.*, 2016b) and corroborated the existence of genomic variants among the atypical strains of *X. arboricola* isolated from *Prunus* spp.

An initial evaluation of the pathogenic activity of the seven atypical strains of *X. arboricola* detected by MLSA revealed variations in their virulence; for instance, all of them cause necrosis on the susceptible peach rootstock GF-305, but after 21 dpi their populations in the inoculated leaves decreased reflecting a non-compatible plant-microbe interaction described also for other *Xanthomonas* (Ah-You *et al.*, 2007). Therefore, these strains must be considered non-pathogenic in this host. The interaction of these atypical strains with other hosts plants showed differences among the strains and hosts; for example, all atypical strains, with exception of CITA 44, showed necrotic spot in the infiltrated tomato leaf-zone, meanwhile only CITA 14 was able to cause necrosis on *Nicotiana* spp. These variations concurred with the results obtained previously in a lineage of non-pathogenic strains of *X. arboricola* isolated from *Juglans regia* (Essakhi *et al.*, 2015). As in that study, the lineage of non-pathogenic *X. arboricola* strains isolated from *Prunus* spp. (Hajri *et al.*, 2012; Cesbron *et al.*, 2015; Garita-Cambronero *et al.*, 2016b) showed a non-canonical T3SS and T3Es repertoire in CITA 14 and CITA 149, or the absence of T3SS and T3Es, as in strains CITA 42, CITA 49, CITA 51 and CITA 124.

A global overview of these results led us to the question why, even in the absence of the canonical T3SS or the T3Es described in such strains (White *et al.*, 2009), some of them were able to cause hypersensitive response on *Nicotiana* spp., tomato and the peach rootstock GF-305, meanwhile others like strain CITA 44 did not cause apparent effect on the assayed hosts.

Due to the fact that this variants could not be clarified only based in the PCR typing of the components for the T3SS and its related effectors, a more in-depth analysis based on other pathogenicity determinants that could play a role in this plant-microbe interaction was needed. Consequently, a whole-genome comparative analysis was performed on the strains CITA 14, CITA 44 and CITA 124, which are representatives of the three MLSA clusters that enclosed non-pathogenic strains isolated from *Prunus*.

Whole genome sequencing of these three strains permitted to accurately infer their phylogenetic position within *X. arboricola*. After a presence/absence comparative analysis of the groups of orthologous genes found in the pangenome of *X. arboricola*, it was possible to explore the spatial distribution of the strains according to the two principal components that contributed more to the variability among these genomes. This explorative analysis permitted to infer a clear pathovar-based clustering of the strains, as reported previously in other strains of *X. arboricola* isolated from *Juglans regia* with non-canonical T3SS. These non-pathogenic strains, CITA 14 and CITA 124, isolated from *P. persica* were closely related to those strains of *X. arboricola* that do not cause disease (CFBP 7634, CFPB 7651 and CITA 44) (Cesbron *et al.*, 2015; Garita-Cambronero *et al.*, 2016b), or have a low pathogenic ability (3004, NCPPB 1630 and NCPPB 1832) (Ignatov *et al.*, 2015; Harrison *et al.*, 2016). Additionally, a similarity analysis based in the content of all the groups of orthologous genes, as well as a phylogenetic analysis based in the concatenated nucleotide sequences of all the genes shared by the studied strains, have corroborated the assignment of CITA 14 and CITA 124 to a cluster that included the strains mentioned above, which is located in a basal phylogenetic position within the species *X. arboricola*.

Besides this, a clear division of these strains in two subgroups could be established, one cluster composed for those strains that harbored the components for a canonical or non-canonical T3SS as well as some T3Es, and a second cluster composed by those that did not have neither this secretory system nor its related effectors. Considering the distribution of the T3SS and related T3Es within the subinfraspecific groups, and based in the hypothesis of evolution of the virulence gene and regulator acquisition previously proposed for *Xanthomonas* (Jacobs *et al.*, 2015), a possible hypothesis about how evolution has worked on the acquisition of T3SS-associated virulence in this species could be proposed.

X. arboricola, as a member of the phylogenetic group 2 of *Xanthomonas* (Jacobs *et al.*, 2015), harbor homologues of the key regulators of the T3SS and T3Es such as *hpaR2*, *hpaS*, *hrpG* and *hrpX*, as well as other pathogenic-related features like *pehA* and *pehD*, and promoters such as the PIP-boxes associated with HrpX and, as suggested by Jacobs and colleagues (2015), it is possible, that could evolve to adapt to these master regulons. Actually, preceding genome analyses on *X. arboricola* (Cesbron *et al.*, 2015; Garita-Cambronero *et al.*, 2016b) have shown the presence of all these molecular components

in all the available genomes of the species and, in the same manner, in our work they have been found in strains CITA 14 and CITA 124. Later on, an ancestral basal group of non-pathogenic strains, such as the one conformed by CITA 44 and CITA 124, could acquire a T3SS and some T3Es, but only those that presented minimum core group of T3Es, which in *X. arboricola* seems to be comprised by six genes (Hajri *et al.*, 2012), could surpass the plant immunity and developed a compatible plant-pathogen interaction. This could be the case of strains from the pathovar *celebensis*, which are within the sequenced strains, the pathogenic group with a lower pathogenic activity and had the smallest repertoire of T3Es. The subsequent acquisition of accessory effectors by some strains, as could be the case of the T3Es *xopE3* and *xopAQ* in pathovar *pruni*, identical to those found in other *Xanthomonas* spp., could alter the host range of the species. These events allowed some strains to become specialized in a specific host, as could be the case of the three most pathogenic pathovars, *corylina*, *pruni* and *juglandis*, of *X. arboricola*.

Genome comparative analysis also showed variants among CITA 14, CITA 124 and the available genomes of *X. arboricola* in a large list of genes that have been associated with different stages of the pathogenic process in *Xanthomonas* spp. (Suplmental Table 3). On one hand, in these two strains, slight differences, related to environmental sensing such as the MCPs, TBDTs and the STCRs, were found. On the other hand, major differences were found in those features associated with the flagellin protein sequences as well as with the molecular components of the type IV pilus. In other xanthomonads, the flagellin polymorphism observed here between pathogenic and non-pathogenic strains of *Prunus* has been associated to the ability of the plant to detect this microbial associated feature and to trigger the plant immune response associated to the initial stages of the plant-pathogen interaction (Sun *et al.*, 2006).

In *Xanthomonas*, the type IV pilus seems to play an important role in bacterial host-interaction and pathogenesis, in twitching motility, in the formation of mature biofilms and in the interaction with bacteriophages (Dunger *et al.*, 2016). In *X. arboricola* all the described non-pathogenic strains and strains CITA 14 and CITA 124 (Cesbron *et al.*, 2015; Garita-Cambronero *et al.*, 2016b) showed a gene arrangement similar to the one previously observed in *X. translucens* pv. *undulosa* strain Xtu 4699, which is characterized by the absence of homologues of *fimA*, *fimT*, *pilV*, *pilW*, *pilX* and *pilYI* (Dunger *et al.*, 2016). In the *Prunus* non-pathogenic strain CITA 44, the absence of

these minor pilins does not alter the twitching type motility (Garita-Cambronero *et al.*, 2016b). From all the variants found in the molecular components of this macromolecular structure, only mutants in the orthologue of *pilYI* have shown a reduction in virulence in the non-vascular pathogen *Xanthomonas oryzae* pv. *oryzicola* (Burdman *et al.*, 2011). In all the pathogenic pathovars of *X. arboricola*, included the strain NCPPB 1630 of the pathovar *celebensis*, homologues of the minor pilins mentioned above were found, but all of them showed a percentage of identity lower than 80% with respect their orthologues in *Xanthomonas citri* subsp. *citri* 306 (Dunger *et al.*, 2014).

As reported in previous studies (Essakhi *et al.*, 2015; Cesbron *et al.*, 2015; Garita-Cambronero *et al.*, 2016b), remarkable differences have been found among pathogenic and non-pathogenic strains of *X. arboricola* with respect to the T3SS and T3Es. Non-pathogenic strains of this species were separated in two different groups, one composed for those strains described in *X. arboricola* (CITA 14, CFBP 7651, NCPPB 1630 and NCPPB 1832), isolated from banana, stone fruits or walnut, that harbored the molecular components of the T3SS and shared the six core T3Es, *avrBs2*, *hpaA*, *hrpW*, *xopA*, *xopF1* and *xopR*. One exception to this group was strain CFBP 7634, isolated from walnut, which only harbored two of the T3Es, *xopR* and *avrBs2*, and did not showed homologues for the T3SS (Cesbron *et al.*, 2015). To perform pathogenicity tests of the strains isolated from *Junglans* and *Prunus* on banana in tropical conditions could be interesting to determine if those closely related strains are able to cause disease on this host. A second group, comprised by the non-pathogenic strains CITA 44 and CITA 124 isolated from *Prunus* spp. and the pathogenic strain 3004 isolated from barley, was characterized by the absence of T3SS and T3Es. Due to the fact that these strains are closely related according to the phylogenetic analysis, it was tested if strains CITA 44 and CITA 124 caused disease on barley, but negative results pointed out that the ability of strain 3004 to cause disease on this crop could be related to other molecular features that are not shared among this strain and strains CITA 44 and CITA 124.

In addition to the flagellin polymorphism mentioned above and its possible role on the plant immune response, a recent study on *X. euvesicatoria* described 17 T3Es that inhibit the plant immunity triggered by the domain flg22 in *Arabidopsis thaliana* (Popov *et al.*, 2016) and, from these, the T3Es *xopB*, *xopE2*, *xopF1*, *xopL*, *xopN*, *xopV*, *xopX* and *xopZ* have been predicted in *X. arboricola*. Those pathogenic strains that

cause disease on hazelnut, stone fruits and walnut presented seven of these T3Es, meanwhile the non-pathogenic CFBP 7651, CITA 14, and the banana-pathogenic strains NCPPB 1630 and NCPPB 1832, only harbored the T3E *xopF1*. Further functional studies comprising these PTI inhibitors will be useful to understand their role to sidestep the initial plant defense mechanisms and the development of a compatible plant-pathogen interaction. For these purposes, the use of non-pathogenic strains, such as CITA 14, could be useful to determine if these T3Es are playing a key role inhibiting the PTI in *Prunus*.

Regarding the type IV secretion systems, which are molecular structures adapted to translocate large molecules, like proteins or protein-DNA complexes, through multiple cell membranes (Guglielmini *et al.*, 2014), the VirB/VirD4 system has been found in all the strains of *X. arboricola*. Nevertheless, the profile of proteins associated with this T4SS was almost the same in all the strains, with the exception of the homologues of VirB6, which is scattered within the members of the *X. arboricola*. According to the studies of mutants of *virB6* in *A. tumefaciens*, this gene is essential for the biogenesis of the T pilus and the secretion channel (Jakubowski *et al.*, 2003). But in *X. citri* subsp. *citri* and in *X. campestris* pv. *campestris*, this T4SS has been described as not playing a main role in virulence (Jacob *et al.*, 2014) because the presence of the complete core of components for its expression in *X. arboricola* varies among strains, regardless of the pathogenic activity. Therefore, VirB/VirD4 secretion system may not be essential for virulence in this species. Despite this, functional experiments with T4SS-deleted mutants are required for testing this hypothesis.

Evidence of a group of genes putatively related to the *tfc* T4SS of *H. influenza*, which is related to bacterial conjugation, have been obtained after searching for homologues using the Blast tool from the NCBI, and also corroborated using the web-based tool for prediction of T4SS-related genes T346Hunter (Martínez-García *et al.*, 2015); but the identity of the putative orthologous genes found in *X. arboricola* showed an amino acid sequence identity lower than 80% for all the genes. Despite of this, the presence of this group of genes, annotated as integrating conjugative elements, varied among *X. arboricola* strains and were only present in those pathogenic organisms from the pathovars *juglandis* and *pruni*.

As a final result of this comparative analysis, the absence of the recalcitrant plasmid pXap41 observed by the multiplex-PCR approach (Pothier *et al.*, 2011c) was corroborated in CITA 14 and CITA124, and as proposed previously, it was only found in members of the pathovar *pruni*. Presence of this plasmid has been useful not only to differentiate such pathovar from the other pathovars of the species as proposed previously (Pothier *et al.*, 2011c), but also to distinguish pathogenic strains of *Xap* from other strains of *X. arboricola* that coexist in the same plants hosts. In addition to this feature, the pangenomic analysis pointed out a series of unique genes for each infrasubspecific group of *X. arboricola* that could be interesting targets for developing new precise diagnostic tools.

Due to the fact that this plasmid contains at least three virulence factors, *xopAQ*, *xopE3* and *mltB*, which are unique in *Xap*, it is confirmed to be a good target for conducting studies of host specialization in the *Xap-Prunus* relationship. Additionally, it is useful as a genomic marker to differentiate *Xap* from all the other members of the species, especially from those non-pathogenic strains found in *Prunus* hosts.

In this article, the use of pXap41, specifically a partial sequence of the virulence-associated gene *xopE3*, was explored for designing a sensitive and specific real-time PCR-based test to differentiate *Xap* from other *X. arboricola* strains. As shown here, the developed test was highly sensitive on both heat-treated bacterial cells and purified DNA, but showed unwanted positive amplification in *X. axonopodis*, *X. campestris*, *X. citri*, *X. fuscans* and *X. vesicatoria*. To our knowledge, there is not record of the presence of these species causing disease on *Prunus*, and consequently to find one of them in natural conditions on these hosts could be only possible as a fortuitous event. The previously developed real-time PCR for detecting *Xap* (Palacio-Bielsa *et al.*, 2015b), as well as the other PCR-based methods designed for this purposes, with the exception of the Bio-PCR protocol proposed by Ballard and colleagues (2011), were not able to differentiate those members of the pathovars *corylina* and *pruni* (Supplementary Figure 5). But the most important problem was that as has been shown here, the methods described were not able to differentiate *Xap* from non-pathogenic strains of *Prunus*. Therefore, based in our results we propose for *Prunus*-isolated strains this real-time PCR amplification protocol that could be used in conjunction with the method proposed by Palacio-Bielsa and collaborators (2011) for routine detection and identification and diagnosis of this quarantine pathogen responsible for bacterial spot of

stone fruits and almond. A combined result of both tests gives a precise identification of the xanthomonads detected in the plant. If both test result positive, the organism tested could be identified as *Xap* and, on the other hand if the target organism shows positive results only for the ABC-method it could be designated as member of the *Xap*-look-a-like group. Both the multiplex conventional PCR described by Pothier and colleagues (2011c) and the proposed combination of two real-time PCR protocols proposed here are suitable for differentiate *Xap* strains. However, the latter offers advantages since it allows detecting *Xap* from plant material (including asymptomatic samples) (Peñalver *et al.*, 2016), whereas the protocol proposed by Pothier and collaborators (2011) has only been assayed using pure bacterial cultures. Xanthomads group associated to *Prunus* spp. requires further taxonomic analyses for more accurate description of the taxonomic status of the different strains. Exploration of the transcriptome and the metabolome of such strains could also help in identifying factors contributing to their diversity.

There is an scarce number of pangenomes of species of plant pathogenic bacteria now available, but this type of studies previously performed in other bacterial species have shown that the genomes of few species are not sufficient and that sequences from multiple genomes are required. As whole, the pangenome of *X. arboricola* and the characterization of the atypical *X. arboricola* strains found on *Prunus* spp., as well as their use to study the genomic diversity of *X. arboricola*, has revealed and corroborated the existence of a distinct phylogenetical basal lineage of this species which is associated with a wide host range. From the strains including in this group, those considered as low-pathogenic seemed to cause disease in two species of monocotyledon plants (banana and barley). After an extensive comparative analysis of those virulence-related genes, it was determined that this bacterial lineage slightly differed from those which are considered as highly pathogenic in several features associated with the initial or later stages of the pathogenicity process. To improve the knowledge on the pathogenic ability and diversity of this species will eventually open the way for the development of innovative control strategies to the diseases they cause.

4.6 Acknowledgments

We would like to thank to Elisa Ferragud, Ana Ruíz Padilla and Isabel M. Berruete for technical assistance. This work was supported financially by the Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA) project RTA2014-00018-CO2-01. J. Garita-Cambronero held a Ph.D. fellowship from the Spanish Government (Ministerio de Educación, Cultura y Deporte; fellowship FPU12/01000).

Capítulo 5: Discusión general

Como resultado de los diversos brotes de *X. arboricola* pv. *pruni* que han surgido en España en los últimos años (Roselló *et al.*, 2012; Palacio-Bielsa *et al.*, 2014), se aislaron una serie de cepas bacterianas que presentaban colonias con un fenotipo típico de *Xanthomonas*. Dichas cepas fueron inicialmente identificadas, utilizando un protocolo estandarizado, como *X. arboricola* pv. *pruni* (Palacio-Bielsa *et al.*, 2011, 2015b). Una selección de estas cepas, representativa de los huéspedes y las regiones geográficas de España en las que se encontró este patógeno, fue utilizada para realizar una caracterización tanto fenotípica como molecular de caracteres que se relacionan con los procesos de infección de las bacterias fitopatógenas. El objetivo final era conocer, aunque fuera parcialmente, aquellos factores que podrían estar vinculados al desarrollo de la mancha bacteriana de frutales de hueso y el almendro.

Una primera caracterización de las cepas realizada siguiendo un esquema de MLSA basado en los genes *dnaK*, *fyuA*, *gyrB* y *rpoD* (Young *et al.*, 2008), definido útil para caracterizar cepas de *X. arboricola* pv. *juglandis* (Kaluzna *et al.*, 2014), confirmó que la mayor parte de las cepas aisladas se identificaban como *X. arboricola* pv. *pruni*. Además, el método de caracterización demostró ser lo suficientemente resolutivo para identificar una cepa aislada de *Prunus mahaleb*, CITA 44, dentro de esta especie pero filogenéticamente diferente a los nueve patovares descritos en *X. arboricola* (Garita-Cambronero *et al.*, 2016a,b). Este tipo de cepas atípicas habían sido ya descritas mediante MLSA, en otras especies de plantas, siendo la mayoría de ellas identificadas dentro de la especie *X. arboricola* (Fischer-Le Saux *et al.*, 2015; Jacques *et al.*, 2016).

La caracterización, de las cepas aisladas de *Prunus* spp. se acompañó de una caracterización fenotípica tanto de las cepas identificadas como patovar *pruni* como de la cepa atípica CITA 44. Este análisis incluyó diversos caracteres asociados con los estadios iniciales de la infección en planta, como son la capacidad de metabolizar diversos compuestos de carbono y la respuesta quimiotáctica hacia algunos de ellos, la capacidad de movimiento tanto en superficie (“swarming” y “twitching”) como en medio líquido o semisólido (“swimming”), la capacidad de secretar sustancias surfactantes y finalmente su virulencia en diversos huéspedes del género *Prunus* (Garita-Cambronero *et al.*, 2016b). Una primera característica fenotípica encontrada en las cepas de *X. arboricola* estudiadas fue su capacidad motil asociada tanto a flagelo como a pilus tipo IV. Estos tipos de movimiento habían sido observados en otras especies como *X. citri* (Sena-Vélez, 2015), pero en el caso de motilidad mediada por

flagelo, únicamente se había descrito el movimiento tipo “swimming”. En este trabajo se describe por primera vez para *Xanthomonas* otro movimiento dependiente de flagelo que se asocia a colonias de tipo dendrítico en medio de cultivo sólido, tanto en cepas de *X. arboricola* pv. *pruni* (por ejemplo la cepa CITA 33) como en la cepa atípica CITA 44. Este fenotipo es típico del movimiento tipo “swarming”, el cual generalmente está conectado con la capacidad de agregación y formación de biopelículas en superficie, tal y como ocurre en *X. arboricola* pv. *pruni* (Sabuquillo & Cubero, 2016).

Otro aspecto fenotípico importante observado en este análisis, ha sido la variación en cuanto a la virulencia de las cepas aisladas en los diversos huéspedes analizados. El hecho de que ninguna de las cepas estudiadas produjera infección en albaricoquero (*P. armeniaca*) hace pensar en una posible súper-especialización de las cepas de *X. arboricola* de este huésped que podrían presentar unas características especiales que no han sido analizadas en este trabajo y que sin embargo serían objeto interesante de trabajos futuros. Por otro lado, el efecto observado en los diferentes huéspedes estudiados permitió determinar que las diferencias encontradas a través de la caracterización molecular en la cepa atípica CITA 44 iban asociadas a su incapacidad para desarrollar síntomas de enfermedad, por lo que se consideró como una cepa no patógena semejante a *X. arboricola* pv. *pruni* (Garita-Cambronero *et al.*, 2016b)

Además, se observaron diferencias entre ambos tipos de cepas en cuanto al perfil de metabolismo de compuestos de carbono y el patrón quimiotáctico. En el caso de la cepa CITA 44, la utilización de estos compuestos no solamente fue diferente al observado en las cepas del patovar *pruni*, sino que también varió con respecto al descrito para la especie *X. arboricola* (Vauterin *et al.*, 1995), dejando entrever que la descripción inicial de la especie, basada únicamente en cepas virulentas, ha podido subestimar la diversidad de la misma tanto filogenéticamente como metabólicamente. El presente trabajo aporta información que aumenta el conocimiento sobre la verdadera diversidad de esta especie bacteriana.

En base a los datos obtenidos en esta caracterización inicial, tres cepas de *X. arboricola* fueron seleccionadas con el objetivo de secuenciar sus genomas y mediante análisis genómico determinar aquellos factores que pudiesen estar asociados con la patogénesis bacteriana. En primer lugar se obtuvo la secuencia correspondiente al genoma completo de la cepa CITA 33 de *X. arboricola* pv. *pruni*, aislada de *P. amygdalus* var. Guara.

Este genoma constituye el primero públicamente disponible para este patovar y para el cual se describieron las características generales. (Garita-Cambronero *et al.*, 2014). Posteriormente, se publicaron las características de los genomas de una cepa patógena en ciruelo (*P. salicina*) IVIA 2626.1, que difería de CITA 33 en aspectos fenotípicos como la motilidad superficial, y de la cepa no patógena CITA 44 (Garita-Cambronero *et al.*, 2016a).

El aumento de las enfermedades provocadas por *X. arboricola* a nivel mundial hace necesario el desarrollo de nuevas y eficientes estrategias de control, para lo cual es necesario comprender los mecanismos de patogénesis, al menos de los patovares más virulentos como son, *corylina*, *juglandis* y *pruni* (Lamichhane, 2014; Lamichhane & Varvaro, 2014). En los últimos años esta necesidad ha conducido a la realización de numerosos estudios genómicos de la especie (Ibarra Caballero *et al.*, 2013; Higuera *et al.*, 2015; Ignatov *et al.*, 2015; Pereira *et al.*, 2015; Cesbron *et al.*, 2015; Harrison *et al.*, 2016). La disponibilidad de dicha información genómica ha permitido en la realización en este trabajo de un primer análisis comparativo entre las tres cepas españolas de *X. arboricola*, aisladas de *Prunus* spp., y otras 15 cepas de *X. arboricola* aisladas en plantas de los géneros *Corylus*, *Hordeum*, *Juglans* y *Prunus* (Garita-Cambronero *et al.*, 2016b).

En esta comparativa se ha podido profundizar en el conocimiento sobre la filogenia de estas cepas al poder ser realizado un análisis filogenético basado en secuencias concatenadas de los genes compartidos por las 18 cepas analizadas. Este análisis confirmó el resultado obtenido mediante MLSA, situando la cepa CITA 44 en un clado diferente al de las cepas patógenas de *Prunus*, presentando mayor similitud con cepas poco virulentas aisladas de cebada y banana, así como con cepas no patógenas aisladas de nogal. Este resultado concuerda con los obtenidos para dos cepas no patógenas de nogal (CFBP 7634 y CFBP 7651) en trabajos anteriores. Dichas cepas se consideraron diferentes a las cepas del patovar *Juglandis* según el análisis de MLSA basado tanto en un esquema compuesto por siete genes como en el análisis del “core genome” (Fischer-Le Saux *et al.*, 2015; Cesbron *et al.*, 2015).

Una vez determinada de forma precisa la posición taxonómica de las cepas analizadas, en el presente trabajo, se abordó el estudio de genes potencialmente relacionados con patogénesis en los diversos grupos de *X. arboricola*, haciendo énfasis en las cepas

aisladas de *Prunus*, y teniendo en cuenta aquellos descritos en otras especies de *Xanthomonas* así como en otras especies bacterianas (da Silva *et al.*, 2002; Filloux, 2004; Vorhölter *et al.*, 2008; Wang *et al.*, 2008; Chevance & Hughes, 2008; He & Zhang, 2008; White *et al.*, 2009; Subramoni *et al.*, 2010; Szczesny *et al.*, 2010; Mhedbi-Hajri *et al.*, 2011; Potnis *et al.*, 2011; Ryan *et al.*, 2011; Guo *et al.*, 2011; Hajri *et al.*, 2012; Vandroemme *et al.*, 2013; Li *et al.*, 2014; Guglielmini *et al.*, 2014; Cesbron *et al.*, 2015; Dunger *et al.*, 2016; Nascimento *et al.*, 2016).

El análisis comparativo permitió identificar diferencias existentes entre cepas patógenas y no patógenas de *Prunus* con respecto a mecanismos asociados con los primeros estadios de la infección, tales como son la percepción de estímulos ambientales, reflejado en el análisis de la presencia de genes relacionados con los transportadores TonB-dependientes, los sensores del sistema regulatorio de dos componentes, las proteínas aceptoras del grupo metilo, las proteínas asociadas a la ruta metabólica de la quimiotaxis o el quorum sensing así como genes relacionados con la adhesión bacteriana y la motilidad mediante flagelo o el pilus tipo IV. Además, se encontraron diferencias también en aspectos propios de las etapas más tardías de la infección bacteriana, como los sistemas de secreción tipo II, III y IV y sus respectivas enzimas o proteínas efectoras de cada uno de ellos.

Como parte de estos resultados, la diferencia más significativa entre las cepas patógenas y la cepa no patógena aisladas de *Prunus* fue la ausencia del sistema de secreción tipo III (T3SS) y de sus proteínas efectoras (T3Es), así como la ausencia del plásmido hipotéticamente asociado a virulencia, pXap41, en la cepa CITA 44. Este tipo de variaciones han sido descritas en otras especies de *Xanthomonas* (Jacobs *et al.*, 2015). En *X. arboricola* además se ha descrito el genoma de la cepa 3004 que, a pesar de no presentar ni T3SS ni sus efectores asociados, es capaz de producir enfermedad en cebada (Ignatov *et al.*, 2015). También en el caso de las dos cepas no patógenas de *X. arboricola* aisladas de nogal, se han descrito diferencias en el T3SS y sus proteínas secretadas; mientras que la cepa CFBP 7634 no presentó T3SS y contenía un número reducido de efectores, la cepa CFBP 7651 contenía todos los genes para un T3SS funcional y un número reducido de efectores (Cesbron *et al.*, 2015). En cuanto al plásmido pXap41, se corroboró, como había sido descrito previamente (Pothier *et al.*, 2011c), que es exclusivo de *X. arboricola* pv. *pruni* por lo que resulta útil como

herramienta para diferenciar las cepas del patovar *pruni* de aquellas cepas no patógenas presentes en los frutales del género *Prunus*.

El haber encontrado y caracterizado una cepa de *X. arboricola* en *Prunus* filogenéticamente diferente a las cepas del patovar *pruni*, incapaz de producir infección e imposible de diferenciar mediante el protocolo de PCR tiempo real recomendado para el diagnóstico de la mancha bacteriana de los frutales de hueso y almendro (Palacio-Bielsa *et al.*, 2015b), planteó una serie de preguntas, cuyas respuestas son abordadas en el capítulo 4 de este trabajo. ¿Cómo de comunes son estas cepas atípicas?; si el caso de la cepa CITA 44 no es un caso aislado, ¿todas las cepas atípicas son no patógenas en *Prunus*?, de ser así ¿es la ausencia del T3SS y de sus efectores un factor común a todas las cepas atípicas?; ¿es posible diferenciar de alguna manera estas cepas atípicas de aquellas que pertenecen al patovar *pruni*? Y por último, ¿los perfiles variables del T3SS y de sus efectores en *X. arboricola* podría dar alguna pauta sobre la posible evolución de la patogenicidad en esta especie?.

Con el fin de obtener datos que ayudaran a dar respuesta a estas cuestiones, se realizó una prospección dentro de colecciones españolas disponibles de este patógeno. La selección se abordó utilizando el protocolo de PCR multiplex descrito para determinar la presencia de plásmido pXap41 (Pothier *et al.*, 2011c). Se encontró que además de la cepa CITA 44, seis cepas no presentaron amplificación para las tres regiones de este plásmido analizadas, por lo que se consideraron cepas atípicas.

Un posterior análisis mediante MLSA corroboró que estas cepas no pertenecían al mismo grupo filogenético que las cepas del patovar *pruni*. Concordando con lo descrito para otras cepas de *X. arboricola* aisladas de diferentes huéspedes (Essakhi *et al.*, 2015; Fischer-Le Saux *et al.*, 2015). Se hizo un estudio sobre la virulencia de estas cepas en huéspedes pertenecientes, tanto al género *Prunus* como a otros géneros vegetales, y se encontró que ninguna de ellas era patógena en *Prunus* y solo algunas de ellas inducían algún tipo de síntoma en otros huéspedes. Asimismo, se estudió el perfil de genes asociados al T3SS y sus efectores mediante PCR. Al igual que lo observado en las cepas de nogal (Cesbron *et al.*, 2015), las cepas atípicas aisladas en *Prunus* no presentaron T3SS ni efectores o bien presentaron un T3SS completo pero un número reducido de efectores (Essakhi *et al.*, 2015).

Con el fin de determinar si la variación observada en cuanto a la virulencia se asociaba solamente a los perfiles de genes correspondientes al T3SS, o además estaba relacionado con otro tipo de factores relacionados con la patogénesis, se realizó la secuenciación del genoma de dos cepas atípicas CITA 14, en la que se habían detectado tanto T3SS como T3Es y CITA 124, que no contenía T3SS ni T3Es. Ambas cepas junto con la cepa CITA 44 eran representativas de la diversidad de cepas atípicas encontradas en este estudio. El estudio pan genómico situó la cepas en su contexto taxonómico, confirmándose que no pertenecían al patovar *pruni*. El análisis de las secuencias relacionadas a la patogenicidad, de estos genomas ha permitido hipotetizar sobre los posibles pasos evolutivos que llevaron al desarrollo de la patogenicidad en *X. arboricola*, basándose en la hipótesis sobre el origen de la patogenicidad en *Xanthomonas* propuesta por Jacobs y colaboradores (2015).

Brevemente, esta hipótesis plantea que en determinado momento de la historia evolutiva del género *Xanthomonas*, aquellas especies que pertenecen al denominado grupo taxonómico 2, dentro del cual se encuentra *X. arboricola*, adquirieron los reguladores maestros asociados a patogénesis HpaR2, HpaS, HrpG y HrpX. Esta adquisición pudo provocar que los genomas evolucionaran para adaptarse a dichos reguladores. Un ejemplo sería la aparición de secuencias tipo “PIP-Box” en los genomas para adaptarse a HrpX. En *X. arboricola*, un ancestro basal no patógeno, como podrían ser CITA 44 y CITA 124, pudo haber adquirido los genes asociados al T3SS, así como algunos efectores mínimos necesarios para causar una relación compatible entre la planta y la bacteria. Este grupo mínimo de efectores podría estar constituido por aquellos seis compartidos por todos los patovares patógenos y presentes en grupos de baja virulencia como el patovar *celebensis*. La adquisición posterior de otros efectores tipo III por algunas cepas, pudo alterar la gama de huésped de la especie y generar la especialización por huésped, como puede ser el caso actual de los patovares, *corylina*, *juglandis* y *pruni*.

Finalmente, el análisis pan genómico realizado en este trabajo ha proporcionado información sobre una serie de genes únicos para cada uno de los grupos intraespecíficos analizados. La secuencia de uno de estos genes, el *xopE3*, se ha utilizado para el diseño de una estrategia de detección por PCR en tiempo real de *X. arboricola* pv. *pruni* que evita la identificación errónea de cepas filogenéticamente cercanas

El trabajo desarrollado en esta tesis doctoral aporta información valiosa desde el punto de vista fenotípico y genómico de la especie *X. arboricola* y ahonda en los factores implicados en la patogenicidad de los diversos grupos que conforman esta especie, enfatizando en aquellos caracteres relacionados con el desarrollo de la mancha bacteriana en los frutales del género *Prunus*. Los datos generados en este trabajo, obtenidos mayormente desde estrategias genómicas, deberían ser el punto de partida de futuros estudios que confirmen o no, desde un punto de vista funcional, las hipótesis planteadas en torno al papel que juegan los genes identificados en el trabajo.

Solo el conocimiento profundo de los mecanismos implicados en los procesos de patogénesis permitirá el futuro desarrollo de nuevos métodos de control de enfermedades, siendo especialmente importante en aquellas de reciente aparición y con graves consecuencias en agricultura, como es el caso de la mancha bacteriana de los frutales de hueso y el almendro.

Capítulo 6: Conclusiones

1. La técnica de caracterización molecular mediante MLSA ha permitido determinar la existencia de cepas atípicas de *X. arboricola* asociadas a diversas especies de *Prunus*. Estas cepas no fueron agrupadas con ningún otro patovar de los nueve descritos en esta especie.
2. Las cepas atípicas, aisladas de *Prunus* spp., presentaron fenotipos diferentes a las cepas patógenas de *X. arboricola* pv. *pruni*, especialmente con respecto a su patogenicidad y virulencia, no desarrollando infección en frutales del género *Prunus*.
3. La cepa no patógena de *X. arboricola* CITA 44 presentó diferencias fenotípicas en cuanto a la utilización de compuestos de carbono, quimiotaxis y motilidad con respecto a las cepas de *X. arboricola* pv. *pruni*.
4. Por primera vez en el género *Xanthomonas* se ha descrito el movimiento tipo swarming asociado a la formación de colonias de tipo dendrítico en medio de cultivo sólido, tanto en cepas patógenas como no patógenas de *X. arboricola* aisladas de *Prunus* spp.
5. Las cepas no patógenas de *Xanthomonas* aisladas de *Prunus* no presentaron los genes asociados a la estructura del sistema de secreción tipo III ni a sus efectores o bien, sólo un número reducido de éstos. Además, según el análisis por PCR de los genes repA1, repA2, y mobC, estas cepas no poseen el plásmido pXap41, potencialmente asociado a virulencia en cepas del patovar *pruni*.
6. Se ha obtenido la secuencia completa de dos cepas de *X. arboricola* pv. *pruni*: CITA 33 (Xap33), aislada de almendro. E IVIA 2626.1 aislada de ciruelo, además de las cepas atípicas CITA 44, aislada de *P. mahaleb* y CITA 14 y CITA 124, aisladas de melocotonero. Todas ellas son representativas de los grupos obtenidos después de análisis genómicos y fenotípicos.
7. El análisis filogenético basado en el “core genome” de *X. arboricola*, así como el análisis de similitud basado en los genes del pan genoma de esta especie, permitieron corroborar que las cepas atípicas de *X. arboricola* conforman un grupo basal en la especie, el cual contiene cepas avirulentas o cepas poco virulentas, como las descritas para el patovar *celebensis* o la cepa 3004 patógena de cebada

8. El estudio realizado durante este trabajo representa el primer análisis de genómica comparativa de la especie *X. arboricola*, y en él se ha hecho un especial énfasis en las diferencias entre los genomas de cepas patógenas y cepas no patógenas aisladas de *Prunus*, lo que ha permitido dilucidar potenciales factores de virulencia asociados a la mancha bacteriana de los frutales de hueso y el almendro.
9. En cuanto a los transportadores dependientes de TonB (TBDTs), asociados al transporte de diversas sustancias, 17 genes fueron encontrados en las cepas secuenciadas, sin embargo se identificó la presencia exclusiva de tres de estos genes en las cepas no patógenas aisladas de *Prunus* spp.
10. Con respecto al sistema regulador de dos componentes (STCRs), relacionados con la respuesta bacteriana a cambios en estímulos ambientales, las cepas patógenas aisladas de *Prunus* presentaron dos que son únicos para *X. arboricola* pv. *pruni*. Por otro lado, las cepas no patógenas aisladas de *Prunus* presentaron cuatro loci exclusivos.
11. Se observaron además diferencias en cuanto a las proteínasceptoras de grupos metilo (MCPs), asociadas a la percepción de señales químicas implicadas en quimiotaxis. La presencia de una MCP presente en las cepas patógenas del patovar *pruni* diferenció a este grupo del de las cepas no patógenas.
12. En cuanto a la estructura y regulación del flagelo, en general, se ha observado un perfil génico muy similar entre las cepas de *X. arboricola*, sin embargo se encontró un polimorfismo en la flagelina que diferenciaba a las cepas de los patovares más virulentos de aquellas cepas menos virulentas o no patógenas. Este polimorfismo correspondió a un cambio de ácido aspártico por valina en las cepas de los patovares *corylina*, *juglandis* y *pruni*.
13. Las cepas patógenas presentaron un perfil génico diferente al de las cepas no patógenas de *Prunus* en relación al pilus tipo IV. En las cepas no patógenas, se encontró un patrón semejante al de *X. translucens* pv. *undulosa* en el cual los homólogos a *fimA*, *fimT*, *pilV*, *pilW*, *pilX* y *pilY1* de *X. citri*, no se encontraron o presentaron un porcentaje de identidad y cobertura menor al 80%.
14. El repertorio de adhesinas no fimbriales de las cepas patógenas de *Prunus* se diferenció del de las cepas no patógenas en la presencia del gen homólogo a *fhaB2* de *X.*

campestris pv. *vesicatoria*. Además, la hemaglutinina homóloga a XAC0444 de *X. citri*, solamente se encontró en los tres patovares más virulentos de la especie.

15. El análisis de los procesos asociados a etapas tardías de la patogénesis demostró que todas las cepas, independientemente de su patogenicidad, presentaron el sistema de secreción tipo II Xps, asociado con virulencia. Sin embargo, se encontró variación en el perfil de enzimas secretadas por este sistema. Las cepas de *X. arboricola* pv. *pruni* presentaron una pectato liasa, una xilanasa y una xilosidasa que no estaban presentes en las cepas no patógenas aisladas de *Prunus* spp.

16. Todas las cepas de *X. arboricola* presentaron genes homólogos a los componentes del sistema de secreción VirB/VirD4 de *A. tumefaciens*. Además, solo las cepas de los patovares más virulentos, *corylina*, *juglandis* y *pruni*, presentaron homólogos al sistema *tfc* descrito en *H. influenzae*.

17. La principal diferencia entre cepas patógenas y no patógenas de la especie *X. arboricola* se encontró en la presencia del sistema de secreción tipo III y de sus efectores asociados. En el caso de las cepas aisladas de *Prunus*, todas las cepas presentaron homólogos para los reguladores de este sistema, HpaS, HpaR2, HrpG y HrpX. Las cepas no patógenas CITA 44 y CITA 124, no presentaron homólogos para los genes que conforman este sistema de secreción así como ninguno de sus efectores asociados. Asimismo, la cepa CITA 124 presentó homólogos para todos los genes requeridos para la expresión del T3SS y se diferenció de las cepas patógenas del patovar *pruni* al presentar únicamente seis efectores de tipo III.

18. Las cepas patógenas de *X. arboricola* pv. *pruni* presentaron el mayor número de efectores tipo III de la especie *X. arboricola*. Los efectores XopE3 y XopAQ, predicho por primera vez para esta especie en este trabajo, son únicos para este patovar.

19. La identificación del gen *xopE3* como específico de las cepas causantes de la mancha bacteriana de los frutales de hueso y almendro, ha permitido el desarrollo de un protocolo de PCR tiempo real que ha mejorado la precisión en el diagnóstico de esta enfermedad y permite diferenciar las cepas patógenas y no patógenas de *X. arboricola* que cohabitan *Prunus* spp.

20. El conjunto de resultados de este trabajo abre la puerta a futuros estudios funcionales que confirmarán o no la implicación de los genes identificados aquí

presumiblemente involucrados en las diferentes etapas de infección en la mancha bacteriana de los frutales de hueso y el almendro. Los datos obtenidos han permitido además la mejora del método de diagnóstico de esta enfermedad por PCR.

Bibliografía

1. Aarrouf J, Garcin A, Lizzi Y, El Maâtaoui M, 2008. Immunolocalization and histocytopathological effects of *Xanthomonas arboricola* pv. *pruni* on naturally infected leaf and fruit tissues of peach (*Prunus persica* L. Batsch). *Journal of Phytopathology* **156**, 338–345.
2. Abby SS, Rocha EPC, 2012. The non-flagellar type III secretion system evolved from the bacterial flagellum and diversified into host-cell adapted systems. *PLoS genetics* **8**, e1002983.
3. Ah-You N, Gagnevin L, Chiroleu F, Jouen E, Neto JR, Pruvost O, 2007. Pathological variations within *Xanthomonas campestris* pv. *mangiferaeindicae* support its separation into three distinct pathovars that can be distinguished by amplified fragment length polymorphism. *Phytopathology* **97**, 1568–1577.
4. Aini LQ, Hirata H, Tsuyumu S, 2010. A TonB-dependent transducer is responsible for regulation of pathogenicity-related genes in *Xanthomonas axonopodis* pv. *citri*. *Journal of General Plant Pathology* **76**, 132–142.
5. Alfano JR, Collmer A, 2004. Type III secretion system effector proteins : Double Agents in Bacterial Disease and plant defense. *Annual review of phytopathology* **42**, 385–414.
6. Alikhan N-F, Petty NK, Ben Zakour NL, Beatson SA, 2011. BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. *BMC genomics* **12**, 402.
7. Angiuoli S V, Gussman A, Klimke W *et al.*, 2008. Toward an online repository of Standard Operating Procedures (SOPs) for (meta)genomic annotation. *Omics : a journal of integrative biology* **12**, 137–41.
8. Anonymous, 2000. Council directive 2000/29/EC of 8 May 2000 on protective measures against the introduction into the community of organism harmful to plants or plant products and against their spread within the community. *Off. J. Eur. Commun* **L169**, 1–112

9. Ashburner M, Ball CA, Blake JA *et al.*, 2000. Gene Ontology: tool for the unification of biology. *Nat Genet* **25**, 25–29.
10. Aziz RK, Bartels D, Best AA *et al.*, 2008. The RAST Server: rapid annotations using subsystems technology. *BMC genomics* **9**, 75.
11. Ballard E, Dietzgen R, Sly L, Gouk C, Horlock C, Fegan M, 2011. Development of a Bio-PCR protocol for the detection of *Xanthomonas arboricola* pv. *pruni*. *Plant Disease* **95**, 1109–1115.
12. Barionovi D, Scortichini M, 2006. Assessment of integron gene cassette arrays in strains of *Xanthomonas fragariae* and *X. arboricola* pvs. *fragariae* and *pruni*. *Journal of Plant Pathology* **88**, 279–284.
13. Bergsma-Vlami M, Martin W, Koenraadt H *et al.*, 2012. Molecular typing of Dutch isolates of *Xanthomonas arboricola* pv. *pruni* isolated from ornamental cherry laurel. *Journal of Plant Pathology* **94**, S1.29-S1.35.
14. Bi D, Liu L, Tai C, Deng Z, Rajakumar K, Ou H-Y, 2013. SecReT4: a web-based bacterial type IV secretion system resource. *Nucleic acids research* **41**, D660-5.
15. Biondi E, Dallai D, Brunelli A, Bazzi C, Stefani E, 2009. Use of a bacterial antagonist for the biological control of bacterial leaf/fruit spot of stone fruits. *IOBC Bulletin* **43**, 277–281.
16. Blanvillain S, Meyer D, Boulanger A *et al.*, 2007. Plant carbohydrate scavenging through TonB-dependent receptors: A feature shared by phytopathogenic and aquatic bacteria. *PLoS ONE* **2**, e244.
17. Blocker A, Komoriya K, Aizawa S-I, 2003a. Type III secretion systems and bacterial flagella: insights into their function from structural similarities. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 3027–30.
18. Blom J, Albaum SP, Doppmeier D *et al.*, 2009. EDGAR: a software framework for the comparative analysis of prokaryotic genomes. *BMC bioinformatics* **10**, 154.

19. Boch J, Bonas U, 2010. *Xanthomonas* AvrBs3 family-type III effectors: discovery and function. *Annual review of phytopathology* **48**, 419–36.
20. Bogdanove AJ, Koebnik R, Lu H *et al.*, 2011. Two new complete genome sequences offer insight into host and tissue specificity of plant pathogenic *Xanthomonas* spp. *Journal of Bacteriology* **193**, 5450–5464.
21. Bonas U, Stall RE, Staskawicz B, 1989. Genetic and structural characterization of the avirulence gene *avrBs3* from *Xanthomonas campestris* pv. *vesicatoria*. *Molecular & general genetics* **218**, 127–136.
22. Borges CD, Vendruscolo CT, 2008. Xanthan synthesized by strains of *Xanthomonas campestris* pv *pruni*: production, viscosity and chemical composition. *Bioscience Journal* **23**, 67-73.
23. Boudon S, Manceau C, Nottéghem J-L, 2005. Structure and origin of *Xanthomonas arboricola* pv. *pruni* populations causing bacterial spot of stone fruit trees in western Europe. *Phytopathology* **95**, 1081–1088.
24. Bühlmann A, Pothier JF, Tomlinson JA *et al.*, 2013. Genomics-informed design of loop-mediated isothermal amplification for detection of phytopathogenic *Xanthomonas arboricola* pv. *pruni* at the intraspecific level. *Plant Pathology* **62**, 475–484.
25. Bull CT, de Boer SH, Denny TP *et al.*, 2010. Comprehensive list of names of plant pathogenic bacteria, 1980-2007. *Journal of Plant Pathology* **92**, 551–592.
26. Bull CT, De Boer SH, Denny TP *et al.*, 2012. List of new names of plant pathogenic bacteria (2008-2010). *Journal of Plant Pathology* **94**, 21–27.
27. Bull CT, De Boer SH, Denny TP *et al.*, 2014. List of new names of plant pathogenic bacteria (2008-2010). *Journal of Plant Pathology* **94**, 21–27.
28. Burch AY, Shimada BK, Browne PJ, Lindow SE, 2010. Novel high-throughput detection method to assess bacterial surfactant production. *Applied and environmental microbiology* **76**, 5363–72.
29. Burdman S, Bahar O, Parker JK, de la Fuente L, 2011. Involvement of type IV pili in pathogenicity of plant pathogenic bacteria. *Genes* **2**, 706–735.

30. 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.
31. Büttner D, 2012. Protein Export According to Schedule: Architecture, Assembly, and Regulation of Type III Secretion Systems from Plant- and Animal-Pathogenic Bacteria. *Microbiology and Molecular Biology Reviews* **76**, 262–310.
32. Büttner D, Bonas U, 2010. Regulation and secretion of *Xanthomonas* virulence factors. *FEMS Microbiology Reviews* **34**, 107–133.
33. Caraguel CGB, Stryhn H, Gagne N, Dohoo IR, Hammell KL, 2011. Selection of a cutoff value for real-time polymerase chain reaction results to fit a diagnostic purpose: analytical and epidemiologic approaches. *Journal of Veterinary Diagnostic Investigation* **23**, 2–15.
34. Casabuono A, Petrocelli S, Ottado J, Orellano EG, Couto AS, 2011. Structural analysis and involvement in plant innate immunity of *Xanthomonas axonopodis* pv. *citri* lipopolysaccharide. *Journal of Biological Chemistry* **286**, 25628–25643.
35. Cesbron S, Briand M, Essakhi S *et al.*, 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.
36. 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.
37. Chagnot C, Zorgani M a, Astruc T, Desvaux M, 2013. Proteinaceous determinants of surface colonization in bacteria: bacterial adhesion and biofilm formation from a protein secretion perspective. *Frontiers in microbiology* **4**, 303.

38. Chang JH, Desveaux D, Creason AL, 2014. The ABCs and 123s of bacterial secretion systems in plant pathogenesis. *Annual review of phytopathology* **52**, 317–45.
39. Chen S, Beeby M, Murphy GE *et al.*, 2011. Structural diversity of bacterial flagellar motors. *The EMBO journal* **30**, 2972–81.
40. Cheung J, Hendrickson WA, 2010. Sensor domains of two-component regulatory systems. *Current Opinion in Microbiology* **13**, 116–123.
41. Chevance FF V, Hughes KT, 2008. Coordinating assembly of a bacterial macromolecular machine. *Nature reviews. Microbiology* **6**, 455–65.
42. Chevreux B, Wetter T, Suhai S, 1999. Genome sequence assembly using trace signals and additional sequence information. In: Wingender E, ed. *Computer Science and Biology: Proceedings of the German Conference on Bioinformatics (GCB) 99*. Hannover, Germany: GBF-Braunschweig, Department of Bioinformatics, 45–56.
43. Christensen P, Cook FD, 1978. *Lysobacter*, a New Genus of Nonhitting, Gliding Bacteria with a High Base Ratio. *International Journal of Systematic Bacteriology* **28**, 367–393.
44. Christie PJ, 2004. Type IV secretion: the *Agrobacterium* VirB/D4 and related conjugation systems. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research* **1694**, 219–234.
45. Christie PJ, Atmakuri K, Krishnamoorthy V, Jakubowski S, Cascales E, 2005. biogenesis, architecture, and function of bacterial type IV secretion systems. *Annual review Microbiology* **59**.
46. Christie PJ, Whitaker N, González-Rivera C, 2014. Mechanism and structure of the bacterial type IV secretion systems. *Biochimica et biophysica acta* **1843**, 1578–1591.
47. Cianciotto NP, 2005. Type II secretion: A protein secretion system for all seasons. *Trends in Microbiology* **13**, 581–588.

48. Civerolo E, Sasser M, Helkie C, Burbage D, 1982. Selective medium for *Xanthomonas campestris* pv. *pruni*. *Plant Disease* **66**, 39–43.
49. Collins FS, 1998. New goals for the U. S. Human Genome Project: 1998-2003. *Science* **282**, 682–689.
50. Conn HJ, 1942. Validity of the Genus *Alcaligenes*. *Journal of Bacteriology* **44**, 353–360.
51. Costa TRD, Felisberto-Rodrigues C, Meir A *et al.*, 2015. Secretion systems in Gram-negative bacteria: structural and mechanistic insights. *Nature Reviews Microbiology* **13**, 343–359.
52. Craig L, Li J, 2008. Type IV pili: paradoxes in form and function. *Current opinion in structural biology* **18**, 267–77.
53. Craig L, Pique ME, Tainer JA, 2004. Type IV pilus structure and bacterial pathogenicity. *Nature Reviews Microbiology* **2**, 363–378.
54. Crossman L, Dow JM, 2004. Biofilm formation and dispersal in *Xanthomonas campestris*. *Microbes and infection* **6**, 623–9.
55. Darbari V, Waksman G, 2015. Structural biology of bacterial Type IV secretion systems. *Annual review of biochemistry* **84**, 603–29.
56. Darling ACE, Mau B, Blattner FR, Perna NT, 2004. Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Research* **14**, 1394–1403.
57. Darling AE, Mau B, Perna NT, 2010. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PloS one* **5**, e11147.
58. Darsonval A, Darrasse A, Durand K, Bureau C, Cesbron S, Jacques M-A, 2009. Adhesion and fitness in the bean phyllosphere and transmission to seed of *Xanthomonas fuscans* subsp. *fuscans*. *Molecular plant-microbe interactions* **22**, 747–757.

59. Das A, Rangaraj N, Sonti R V, 2009. Multiple adhesin-like functions of *Xanthomonas oryzae* pv. *oryzae* are involved in promoting leaf attachment, entry, and virulence on rice. *Molecular plant-microbe interactions* **22**, 73–85.
60. Delcher A, 1999. Improved microbial gene identification with GLIMMER. *Nucleic Acids Research* **27**, 4636–4641.
61. Denancé N, Lahaye T, Noël LD, 2016. Editorial: Genomics and effectomics of the crop killer *Xanthomonas*. *Frontiers in plant science* **7**, 71.
62. Denny TP, 1995. Involvement of bacterial polysaccharides in plant pathogenesis. *Annual Review of Phytopathology* **33**, 173–197.
63. Van Domselaar GH, Stothard P, Shrivastava S *et al.*, 2005. BASys: a web server for automated bacterial genome annotation. *Nucleic Acids Research* **33**, 455–459.
64. Dowson DW, 1939. On the systematic position and generic names of the Gram negative bacterial plant pathogens. *Zentralblatt für Bakteriologie, Parasitenkunde und Infektionskrankheiten*, **100**, 177–193.
65. Dunger G, Guzzo CR, Andrade MO, Jones JB, Farah CS, 2014. *Xanthomonas citri* subsp. *citri* type IV pilus is required for twitching motility, biofilm development, and adherence. *Molecular plant-microbe interactions : MPMI* **27**, 1132–47.
66. Dunger G, Llontop E, Guzzo CR, Farah CS, 2016. The *Xanthomonas* type IV pilus. *Current Opinion in Microbiology* **30**, 88–97.
67. Dunger G, Relling VM, Tondo ML *et al.*, 2007. Xanthan is not essential for pathogenicity in citrus canker but contributes to *Xanthomonas* epiphytic survival. *Archives of Microbiology* **188**, 127–135.
68. EFSA, 2014. Scientific opinion on pest categorisation of *Xanthomonas arboricola* pv. *pruni* (Smith, 1903). *EFSA Journal* **12**, 1–25.
69. European and Mediterranean Plant Protection Organization (EPPO), 2003. Data sheets on quarantine pests. *Xanthomonas arboricola* pv. *pruni*. 1-5

70. Essakhi S, Cesbron S, Fischer-Le Saux M, Bonneau S, Jacques M-A, Manceau C, 2015. Phylogenetic and VNTR analysis identified non-pathogenic lineages within *Xanthomonas arboricola* lacking the canonical type three secretion system. *Applied and Environmental Microbiology* **81**, 5395-5410.
71. Euzéby J, 2004. Validation of publication of new names and new combinations previously effectively published outside the IJSEM. *International Journal of Systematic and Evolutionary Microbiology* **54**, 1909–1910.
72. Euzéby J, 2007. List of new names and new combinations previously effectively, but not validly, published. *International Journal of Systematic and Evolutionary Microbiology* **57**, 893–897.
73. Feng F, Zhou JM, 2012. Plant-bacterial pathogen interactions mediated by type III effectors. *Current Opinion in Plant Biology* **15**, 469–476.
74. Field D, Amaral-Zettler L, Cochrane G *et al.*, 2011. The genomic standards consortium. *PLoS biology* **9**, e1001088.
75. Field D, Garrity G, Gray T *et al.*, 2008. The minimum information about a genome sequence (MIGS) specification. *Nature Biotechnology* **26**, 541–547.
76. Filloux A, 2004. The underlying mechanisms of type II protein secretion. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1694**, 163–179.
77. Finn RD, Bateman A, Clements J *et al.*, 2014. Pfam: the protein families database. *Nucleic acids research* **42**, D222-30.
78. 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.
79. Gabriel DW, Kingsley MT, Hunter JE, Gottwald T, 1989. Reinstatement of *Xanthomonas citri* (ex Hasse) and *X. phaseoli* (ex Smith) to species and

- reclassification of all *X. campestris* pv. *citri* strains. *International Journal of Systematic and Evolutionary Microbiology* **39**, 14–22.
80. Galan JE, Wolf-Watz H, 2006. Protein delivery into eukaryotic cells by type III secretion machines. *Nature* **444**, 567–573.
 81. Galperin MY, Makarova KS, Wolf YI, Koonin E V, 2015. Expanded microbial genome coverage and improved protein family annotation in the COG database. *Nucleic acids research* **43**, D261-9.
 82. 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.
 83. Garita-Cambronero J, Palacio-Bielsa A, López MM, Cubero J, 2016b. Comparative genomic and phenotypic characterization of pathogenic and non-pathogenic strains of *Xanthomonas arboricola* reveals insights into the infection process of bacterial spot disease of stone fruits. *PLoS ONE* **11**, e0161977.
 84. Garita-Cambronero J, Palacio-Bielsa A, López MM, Cubero J, 2016c. 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.
 85. Garita-Cambronero J, Sena-Vélez 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.
 86. Garita-Cambronero J, Sena-Vélez M, Sabuquillo P *et al.*, 2013. Early steps in the infection process in two *Xanthomonas* spp. models: chemotaxis and biofilm formation. *Acta Phytopathologica Sinica* **43**, 419.
 87. Garrity GM, Bell J, Lilburn T, 2005a. *Phylum XIV . Proteobacteria phyl. Nov.* In (GM Garrity, D Brenner, N Krieg, J Staley, Eds,) *Bergey's Manual of Systematic Bacteriology*. Volume 2 (Part B). 2nd Ed. New York: Springer.

88. Garrity GM, Bell J, Lilburn T, 2005b. *Class III: Gammaproteobacteria Class. nov.* In (D Brenner, N Krieg, J Staley, G Garrity, Eds,) *Bergey's Manual of Systematic Bacteriology*. Volume 2 (Part B). 2nd Ed. New York: Springer.
89. Girgis HS, Liu Y, Ryu WS, Tavazoie S, 2007. A comprehensive genetic characterization of bacterial motility. *PLoS Genetics* **3**, 1644–1660.
90. Glick BR, 2015. Beneficial plant-bacterial interactions. *Beneficial Plant-Bacterial Interactions*, 1–243.
91. Gordon JL, Lefeuvre P, Escalon A *et al.*, 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.
92. Gottig N, Garavaglia BS, Garofalo CG *et al.*, 2010. Mechanisms of infection used by *Xanthomonas axonopodis* pv. *citri* in citrus canker disease. *Current Research, Technology and Education Topics in Applied Microbiology and Microbial Biotechnology*, 196–204.
93. Gottig N, Garavaglia BS, Garofalo CG, Orellano EG, Ottado J, 2009. A filamentous hemagglutinin-like protein of *Xanthomonas axonopodis* pv. *citri*, the phytopathogen responsible for citrus canker, is involved in bacterial virulence. *PLoS ONE* **4**, e4358.
94. Goujon M, McWilliam H, Li W *et al.*, 2010. A new bioinformatics analysis tools framework at EMBL-EBI. *Nucleic Acids Research* **38**, W695–W699.
95. Gourion B, Berrabah F, Ratet P, Stacey G, 2015. *Rhizobium*-legume symbioses: The crucial role of plant immunity. *Trends in Plant Science* **20**, 186–194.
96. Graham JH, Gottwald TR, 1991. Research perspectives on eradication of citrus bacterial diseases in Florida. *Plant Disease* **75**, 1193.
97. Grant JR, Stothard P, 2008. The CGView Server: a comparative genomics tool for circular genomes. *Nucleic acids research* **36**, 181–184.
98. Green ER, Mecsas J, 2016. Bacterial secretion systems – An overview. *Microbiology spectrum* **4**, 10.1128.

99. Guglielmini J, Néron B, Abby SS, Garcillán-Barcia MP, de la Cruz F, Rocha EPC, 2014. Key components of the eight classes of type IV secretion systems involved in bacterial conjugation or protein secretion. *Nucleic acids research* **42**, 5715–27.
100. Guo-feng J, Lei-Lei W, Zheng-chun Z *et al.*, 2015. Type IV secretion system of *Xanthomonas campestris* pv. *campestris* is induced in MMX and related to stress resistance. *Acta Phytopathologica Sinica* **45**, 486–493.
101. 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**, 649–61.
102. Gürlebeck D, Thieme F, Bonas U, 2006. Type III effector proteins from the plant pathogen *Xanthomonas* and their role in the interaction with the host plant. *Journal of Plant Physiology* **163**, 233–255.
103. Hagai E, Dvora R, Havkin-Blank T *et al.*, 2014. Surface-motility induction, attraction and hitchhiking between bacterial species promote dispersal on solid surfaces. *The ISME journal* **8**, 1147–51.
104. Hajri A, Brin C, Hunault G *et al.*, 2009. A «repertoire for repertoire» hypothesis: Repertoires of type three effectors are candidate determinants of host specificity in *Xanthomonas*. *PLoS ONE* **4**, e6632.
105. Hajri A, Meyer D, Delort F, Guillaumès 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.
106. Hajri A, Pothier JF, Saux MF Le *et al.*, 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.
107. Hall T, 2011. BioEdit: an important software for molecular biology. *GERF Bulletin Biosciences* **2**, 60–61.

108. Hamman JH, 2010. Chitosan based polyelectrolyte complexes as potential carrier materials in drug delivery systems. *Marine Drugs* **8**, 1305–1322.
109. 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.
110. Hayward AC, 1993. The hosts of *Xanthomonas*. In (Swings JG, Civerolo EL) *Xanthomonas*. London, United Kingdom: Chapman & Hall. pp. 1-119.
111. He Y-W, Zhang L-H, 2008. Quorum sensing and virulence regulation in *Xanthomonas campestris*. *FEMS microbiology reviews* **32**, 842–57.
112. 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.
113. Hori K, Matsumoto S, 2010. Bacterial adhesion: From mechanism to control. *Biochemical Engineering Journal* **48**, 424–434.
114. Horner-Devine MC, Carney KM, Bohannon BJM, 2004. An ecological perspective on bacterial biodiversity. *Proceedings. Biological sciences* **271**, 113–122.
115. Hu R-M, Yang T-C, Yang S-H, Tseng Y-H, 2005. Deduction of upstream sequences of *Xanthomonas campestris* flagellar genes responding to transcription activation by FleQ. *Biochemical and Biophysical Research Communications* **335**, 1035–1043.
116. Huson DH, Richter DC, Rausch C, DeZulian T, Franz M, Rupp R, 2007. Dendroscope: an interactive viewer for large phylogenetic trees. *BMC bioinformatics* **8**, 460.
117. Hyatt D, Chen G-L, Locascio PF, Land ML, Larimer FW, Hauser LJ, 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC bioinformatics* **11**, 119.

118. Ibarra Caballero J, 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.
119. Ignatov AN, Kyrova EI, Vinogradova S V, 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.
120. Ikeda T, Oosawa K, Hotani H, 1996. Self-assembly of the filament capping protein, FliD, of bacterial flagella into an annular structure. *Journal of molecular biology* **259**, 679–686.
121. Jacob TR, De Laia ML, Moreira LM *et al.*, 2014. Type IV secretion system is not involved in infection process in citrus. *International Journal of Microbiology* **2014**, 763575.
122. 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.
123. Jacques M-A, Arlat M, Boulanger A *et al.*, 2016. Using ecology, physiology, and genomics to understand host specificity in *Xanthomonas*. *Annual Review of Phytopathology* **54**, 163–187.
124. Jakubowski SJ, Krishnamoorthy V, Christie PJ, 2003. *Agrobacterium tumefaciens* VirB6 protein participates in formation of VirB7 and VirB9 complexes required for type IV secretion. *Journal of Bacteriology* **185**, 2867–2878.
125. Jalan N, Aritua V, Kumar D *et al.*, 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.
126. Jalan N, Kumar D, Yu F, Jones JB, Graham JH, Wang N, 2013. Complete genome sequence of *Xanthomonas citri* subsp. *citri* strain Aw12879, a

- restricted-host-range citrus canker-causing bacterium. *Genome Announcements* **1**, e00235-13.
127. Jehl M-A, Arnold R, Rattei T, 2011. Effective-a database of predicted secreted bacterial proteins. *Nucleic acids research* **39**, D591-5.
 128. Jeong K-S, Lee S-E, Han J-W *et al.*, 2008. Virulence reduction and differing regulation of virulence genes in rpf mutants of *Xanthomonas oryzae* pv. *oryzae*. *The Plant Pathology Journal* **24**, 143–151.
 129. Jha G, Rajeshwari R, Sonti R V, 2005. Bacterial type two secretion system secreted proteins: double-edged swords for plant pathogens. *Molecular plant-microbe interactions* **18**, 891–898.
 130. Juhas M, Crook DW, Dimopoulou ID *et al.*, 2007. Novel type IV secretion system involved in propagation of genomic islands. *Journal of Bacteriology* **189**, 761–771.
 131. 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.
 132. Kamoun S, Kado CI, 1990. Phenotypic switching affecting chemotaxis, xanthan production, and virulence in *Xanthomonas campestris*. *Applied and environmental microbiology* **56**, 3855–60.
 133. Katoh K, Misawa K, Kuma K, Miyata T, 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic acids research* **30**, 3059–3066.
 134. Katzen F, Ferreiro DU, Oddo CG *et al.*, 1998. *Xanthomonas campestris* pv. *campestris* gum mutants: effects on xanthan biosynthesis and plant virulence. *Journal of Bacteriology* **180**, 1607–1617.
 135. Kawaguchi A, 2014. Genetic diversity of *Xanthomonas arboricola* pv. *pruni* strains in Japan revealed by DNA fingerprinting. *Journal of General Plant Pathology* **80**, 366–369.

136. Kay S, Bonas U, 2009. How *Xanthomonas* type III effectors manipulate the host plant. *Current Opinion in Microbiology* **12**, 37–43.
137. Kearns DB, 2010. A field guide to bacterial swarming motility. *Nature reviews. Microbiology* **8**, 634–44.
138. Kim JS, Chang JH, Chung S Il, Yum JS, 1999. Molecular cloning and characterization of the *Helicobacter pylori* *fliD* gene, an essential factor in flagellar structure and motility. *Journal of Bacteriology* **181**, 6969–6976.
139. Kim JG, Stork W, Mudgett MB, 2013. *Xanthomonas* type III effector XopD desumoylates tomato transcription factor SLERF4 to suppress ethylene responses and promote pathogen growth. *Cell Host and Microbe* **13**, 143–154.
140. Kohler T, Curty LK, Barja F, van Delden C, Pechere J-C, 2000. Swarming of *Pseudomonas aeruginosa* is dependent on cell-to-cell signaling and requires flagella and pili. *Journal of Bacteriology* **182**, 5990–5996.
141. Korotkov K V, Sandkvist M, Hol WGJ, 2012. The type II secretion system: biogenesis, molecular architecture and mechanism. *Nature Reviews Microbiology* **10**, 336–351.
142. Kostakioti M, Hadjifrangiskou M, Hultgren SJ, 2013. Bacterial biofilms: development, dispersal, and therapeutic strategies in the dawn of the postantibiotic era. *Cold Spring Harbor perspectives in medicine* **3**, 1–23.
143. Kraiselburd I, Daurelio LD, Tondo ML *et al.*, 2013. The LOV protein of *Xanthomonas citri* subsp. *citri* plays a significant role in the counteraction of plant immune responses during citrus canker. *PloS ONE* **8**, e80930.
144. Krogh A, Larsson B, von Heijne G, Sonnhammer EL, 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *Journal of molecular biology* **305**, 567–80.
145. Lagacé L, Pitre M, Jacques M, Roy D, 2004. Identification of the bacterial community of maple sap by using amplified ribosomal DNA (rDNA) restriction analysis and rDNA sequencing. *Applied and Environmental Microbiology* **70**, 2052–2060.

146. Lagesen K, Hallin P, Rødland EA, Staerfeldt H-H, Rognes T, Ussery DW, 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic acids research* **35**, 3100–8.
147. Lamichhane JR, 2014. *Xanthomonas arboricola* diseases of stone fruit, almond, and walnut trees: progress toward understanding and management. *Plant Disease* **98**, 1600–1610.
148. Lamichhane JR, Varvaro L, 2014. *Xanthomonas arboricola* disease of hazelnut: current status and future perspectives for its management. *Plant Pathology* **63**, 243–254.
149. Land M, Hauser L, Jun S-R *et al.*, 2015. Insights from 20 years of bacterial genome sequencing. *Functional & integrative genomics* **15**, 141–61.
150. Lee BM, Park YJ, Park DS *et al.*, 2005. The genome sequence of *Xanthomonas oryzae* pathovar *oryzae* KACC10331, the bacterial blight pathogen of rice. *Nucleic Acids Research* **33**, 577–586.
151. Li RF, Lu GT, Li L *et al.*, 2014. Identification of a putative cognate sensor kinase for the two-component response regulator HrpG, a key regulator controlling the expression of the *hrp* genes in *Xanthomonas campestris* pv. *campestris*. *Environmental Microbiology* **16**, 2053–2071.
152. López-Soriano P, Boyer K, Cesbron S *et al.*, 2016. Multilocus variable number of tandem repeat analysis reveals multiple introductions in Spain of *Xanthomonas arboricola* pv. *pruni*, the causal agent of bacterial spot disease of stone fruits and almond. *PLoS ONE* **11**, e0163729.
153. López MM, Roselló M, Palacio-Bielsa A, 2010. Diagnosis and detection of the main bacterial pathogens of stone fruit and almond. *Journal of Plant Pathology* **92**.
154. Loria R, Kers J, Joshi M, 2006. Evolution of plant pathogenicity in *Streptomyces*. *Annual review of phytopathology* **44**, 469–487.

155. Lowe TM, Eddy SR, 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Research* **25**, 955–964.
156. Löytynoja A, Goldman N, 2008. Phylogeny-aware gap placement prevents errors in sequence alignment and evolutionary analysis. *Science (New York, N.Y.)* **320**, 1632–5.
157. Macwilliams MP, Liao M, 2006. Luria broth (LB) and Luria agar (LA) media and their uses protocol. In ASM MicrobeLibrary. American Society for Microbiology.
<http://www.asmscience.org/content/education/protocol/protocol.3031>.
158. Maharjan RP, Seeto S, Ferenci T, 2007. Divergence and redundancy of transport and metabolic rate-yield strategies in a single *Escherichia coli* population. *Journal of bacteriology* **189**, 2350–8.
159. Malamud F, Homem RA, Conforte VP *et al.*, 2013. Identification and characterization of biofilm formation-defective mutants of *Xanthomonas citri* subsp. *citri*. *Microbiology (United Kingdom)* **159**, 1911–1919.
160. Malamud F, Torres PS, Roeschlin R *et al.*, 2011. The *Xanthomonas axonopodis* pv. *citri* flagellum is required for mature biofilm and canker development. *Microbiology* **157**, 819–829.
161. Mansfield J, Genin S, Magori S *et al.*, 2012. Top 10 plant pathogenic bacteria in molecular plant pathology. *Molecular plant pathology* **13**, 614–29.
162. 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.
163. Marchler-Bauer A, Derbyshire MK, Gonzales NR *et al.*, 2014. CDD: NCBI’s conserved domain database. *Nucleic acids research* **43**, D222–6.
164. Marcone C, 2014. Molecular biology and pathogenicity of phytoplasmas. *Annals of Applied Biology* **165**, 199–221.

165. Martínez-García PM, Ramos C, Rodríguez-Palenzuela P, 2015. T346Hunter: A novel web-based tool for the prediction of type III, type IV and type VI secretion systems in bacterial genomes. *PLoS ONE* **10**, e0119317.
166. Massomo SMS, Nielsen H, Mabagala RB, Mansfeld-Giese K, Hockenhull J, Mortensen CN, 2003. Identification and characterisation of *Xanthomonas campestris* pv. *campestris* strains from Tanzania by pathogenicity tests, biolog, rep-PCR and fatty acid methyl ester analysis. *European Journal of Plant Pathology* **109**, 775–789.
167. Mattick JS, 2002. Type IV pili and twitching motility. *Annual review of microbiology* **56**, 289–314.
168. Mayer L, Vendruscolo CT, Silva WP da, Vorhölter FJ, Becker A, Pühler A, 2011. Insights into the genome of the xanthan-producing phytopathogen *Xanthomonas arboricola* pv. *pruni* 109 by comparative genomic hybridization. *Journal of Biotechnology* **155**, 40–49.
169. Mhedbi-Hajri N, Darrasse A, Pigné S *et al.*, 2011. Sensing and adhesion are adaptive functions in the plant pathogenic xanthomonads. *BMC evolutionary biology* **11**, 67.
170. Molinaro A, Evidente A, Cantore P *et al.*, 2003. Structural determination of a novel O-chain polysaccharide of the lipopolysaccharide from the bacterium *Xanthomonas campestris* pv. *pruni*. *European Journal of Organic Chemistry*, 2254–2259.
171. Van den Mooter M, Swings J, 1990. Numerical analysis of 295 phenotypic features of 266 *Xanthomonas* strains and related strains and an improved taxonomy of the genus. *International journal of systematic bacteriology* **40**, 348–369.
172. Moreira LM, Almeida NF, Potnis N *et al.*, 2010. Novel insights into the genomic basis of citrus canker based on the genome sequences of two strains of *Xanthomonas fuscans* subsp. *aurantifolii*. *BMC Genomics* **11**, 238.
173. Moreira LM, Facincani AP, Ferreira CB *et al.*, 2015. Chemotactic signal transduction and phosphate metabolism as adaptive strategies during citrus

- canker induction by *Xanthomonas citri*. *Functional and Integrative Genomics* **15**, 197–210.
174. Moreira AS, Vendruscolo JLS, Gil-Turnes C, Vendruscolo CT, 2001. Screening among 18 novel strains of *Xanthomonas campestris* pv *pruni*. *Food Hydrocolloids* **15**, 469–474.
 175. Moreno-hagelsieb G, Santoyo G, Moreno-hagelsieb G, Orozco-mosqueda C, 2016. Santoyo et al 2016. Plant growth-promoting bacterial endophytes. *Microbiological Research* **118**, 92-99.
 176. Morgulis A, Coulouris G, Raytselis Y, Madden TL, Agarwala R, Schäffer AA, 2008. Database indexing for production MegaBLAST searches. *Bioinformatics* **24**, 1757–64.
 177. Morris CE, Monier J-M, 2003. The ecological significance of biofilm formation by plant-associated bacteria. *Annual Review of Phytopathology* **41**, 429–453.
 178. Mwangi M, Bandyopadhyay R, Ragama P, Tushemereirwe WK, 2007. Assessment of banana planting practices and cultivar tolerance in relation to management of soilborne *Xanthomonas campestris* pv *musacearum*. *Crop Protection* **26**, 1203–1208.
 179. Nagai H, Kubori T, 2011. Type IVB secretion systems of *Legionella* and other gram-negative bacteria. *Frontiers in Microbiology* **2**, 1–12.
 180. Nascimento R, Gouran H, Chakraborty S *et al.*, 2016. The type II secreted lipase/esterase LesA is a key virulence factor required for *Xylella fastidiosa* pathogenesis in grapevines. *Scientific reports* **6**, 18598.
 181. Naushad S, Adeolu M, Wong S, Sohail M, Schellhorn HE, Gupta RS, 2015. A phylogenomic and molecular marker based taxonomic framework for the order Xanthomonadales: proposal to transfer the families Algiphilaceae and Solimonadaceae to the order Nevskiales ord. nov. and to create a new family within the order Xanthomonadales, . *Antonie van Leeuwenhoek* **107**, 467–485.

182. National Center for Biotechnology Information (NCBI), 2016. How to: find a homolog for a gene in another organism. <https://www.ncbi.nlm.nih.gov/guide/howto/find-homolog-gene>.
183. Niño-Liu DO, Ronald PC, Bogdanove AJ, 2006. *Xanthomonas oryzae* pathovars: model pathogens of a model crop. *Molecular Plant Pathology* **7**, 303–324.
184. Noinaj N, Guillier M, Barnard TJ, Buchanan SK, 2010. TonB-dependent transporters: regulation, structure, and function. *Annual review of microbiology* **64**, 43–60.
185. Oliveira LC, Souza DP, Oka GU *et al.*, 2016. VirB7 and VirB9 interactions are required for the assembly and antibacterial activity of a type IV secretion system. *Structure* **24**, 1707–1718.
186. Oren A, Garrity GM, 2015. Notification of changes in taxonomic opinion previously published outside the IJSEM. *International Journal of Systematic and Evolutionary Microbiology* **65**, 2028–2029.
187. Ozaktan H, Erdal M, Akkopru a, Aslan E, 2012. Biological control of bacterial blight of walnut by antagonistic bacteria. *Journal of Plant Pathology* **94**, S1.53-S1.56.
188. Pagani M, 2004. An ABC transporter protein and molecular diagnoses of *Xanthomonas arboricola* pv. *pruni* causing bacterial spot of stone fruits.
189. Page AJ, Cummins CA, Hunt M *et al.*, 2015. Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* **31**, 3691–3693.
190. Palacio-Bielsa A, Beruete I, López MM *et al.*, 2015a. La mancha bacteriana de los frutales de hueso y del almendro (*Xanthomonas arboricola* pv. *pruni*) en España y Sudamérica. *Phytoma* **271**, 21–28.
191. Palacio-Bielsa A, Cambra MA, Cubero J, Garita-Cambronero J, Roselló M, López MM, 2014. La mancha bacteriana de los frutales de hueso y del almencro (*Xanthomonas arboricola* pv. *pruni*), una grave enfermedad emergente en España. *Phytoma* **259**, 36–42.

192. 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.
193. Palacio-Bielsa A, López-Soriano P, Bühlmann A *et al.*, 2015b. 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**, 36–9.
194. Palacio-Bielsa A, Pothier J, Roselló M, Duffy B, López MM, 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. *Journal of Plant Pathology* **94**, S1.135-S1.146.
195. Pappenfort K, Bassler BL, 2016. Quorum sensing signal-response systems in Gram-negative bacteria. *Nature reviews. Microbiology* **14**, 576–88.
196. 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.
197. 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.
198. Patil PB, Bogdanove AJ, Sonti R V, 2007. The role of horizontal transfer in the evolution of a highly variable lipopolysaccharide biosynthesis locus in xanthomonads that infect rice, citrus and crucifers. *BMC evolutionary biology* **7**, 243.
199. Pelicic V, 2008. Type IV pili: e pluribus unum? *Molecular Microbiology* **68**, 827–837.
200. Peñalver J, Navarro I, López MM, Marco-Noales E, 2016. Optimización de protocolo de detección de *Xanthomonas arboricola* pv. *pruni* en frutales de hueso y almendro. In: *XVIII Congreso de la Sociedad Española de Fitopatología*. Palencia, España, 321.

201. 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.
202. Petersen TN, Brunak S, von Heijne G, Nielsen H, 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature methods* **8**, 785–6.
203. Petrocelli S, Tondo ML, Daurelio LD, Orellano EG, 2012. Modifications of *Xanthomonas axonopodis* pv. *citri* lipopolysaccharide affect the basal response and the virulence process during citrus canker. *PLoS ONE* **7**, e40051.
204. Pfeilmeier S, Caly DL, Malone JG, 2016. Bacterial pathogenesis of plants: future challenges from a microbial perspective: challenges in bacterial molecular plant pathology. *Molecular plant pathology* **7**, 1298–1313.
205. Pieretti I, Royer M, Barbe V *et al.*, 2012. Genomic insights into strategies used by *Xanthomonas albilineans* with its reduced artillery to spread within sugarcane xylem vessels. *BMC genomics* **13**, 658.
206. Plessis H Du, 1984. Scanning electron microscopy of *Xanthomonas campestris* pv . *pruni* in plum petioles and buds. *Phytopathologische Zeitschrift* **109**, 277–284.
207. Popov G, Fraiture M, Brunner F, Sessa G, 2016. Multiple *Xanthomonas euvesicatoria* type III effectors inhibit flg22-Triggered immunity. *Molecular plant-microbe interactions* **29**, 651–660.
208. Posada D, 2008. jModelTest: phylogenetic model averaging. *Molecular biology and evolution* **25**, 1253–6.
209. Posada D, Buckley TR, 2004. Model selection and model averaging in phylogenetics: advantages of akaike information criterion and bayesian approaches over likelihood ratio tests. *Systematic biology* **53**, 793–808.
210. Pothier JF, Pagani MC, Pelludat C, Ritchie DF, Duffy B, 2011a. 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.
211. Pothier JF, Smits THM, Blom J *et al.*, 2011b. Complete genome sequence of the stone fruit pathogen *Xanthomonas arboricola* pv. *pruni*. *Phytopathology* **101**, S144–S145.
 212. Pothier JF, Vorhölter FJ, Blom J *et al.*, 2011c. The ubiquitous plasmid pXap41 in the invasive phytopathogen *Xanthomonas arboricola* pv. *pruni*: complete sequence and comparative genomic analysis. *FEMS Microbiology Letters* **323**, 52–60.
 213. Potnis N, Krasileva K, Chow V *et al.*, 2011. Comparative genomics reveals diversity among xanthomonads infecting tomato and pepper. *BMC genomics* **12**, 146.
 214. Qian W, Han Z-J, He C, 2008. Two-component signal transduction systems of *Xanthomonas* spp.: a lesson from genomics. *Molecular plant-microbe interaction* **21**, 151–161.
 215. Randhawa P, Civerolo E, 1986. Interaction of *Xanthomonas campestris* pv. *pruni* with pruniphage and epiphytic bacteria on detached peach leaves. *Phytopathology* **76**, 549–553.
 216. Ray SK, Rajeshwari R, Sharma Y, Sonti R V., 2002. A high-molecular-weight outer membrane protein of *Xanthomonas oryzae* pv. *oryzae* exhibits similarity to non-fimbrial adhesins of animal pathogenic bacteria and is required for optimum virulence. *Molecular Microbiology* **46**, 637–647.
 217. Rigano L a, Siciliano F, Enrique R *et al.*, 2007. Biofilm formation, epiphytic fitness, and canker development in *Xanthomonas axonopodis* pv. *citri*. *Molecular plant-microbe interactions* **20**, 1222–1230.
 218. Rodriguez-R LM, Grajales A, Arrieta-Ortiz M, Salazar C, Restrepo S, Bernal A, 2012. Genomes-based phylogeny of the genus *Xanthomonas*. *BMC Microbiology* **12**, 43.

219. Roselló M, Santiago R, Palacio-Bielsa A *et al.*, 2012. Current status of bacterial spot of stone fruits and almond caused by *Xanthomonas arboricola* pv. *pruni* in Spain. *Journal of Plant Pathology* **94**, 15–21.
220. Rosenblueth M, Martínez-Romero E, 2006. Bacterial endophytes and their interactions with hosts. *Molecular plant-microbe interactions: MPMI* **19**, 827–837.
221. Rossez Y, Wolfson EB, Holmes A, Gally DL, Holden NJ, 2015. Bacterial flagella: twist and stick, or dodge across the kingdoms. *PLoS Pathogens* **11**, e1004483.
222. Runyen-Janecky L, 2014. Transduction of environmental signals by prokaryotic two-component regulatory systems. *Molecular life Sciences*, 1-7.
223. Ryan RP, Vorhölter F-J, Potnis N *et al.*, 2011. Pathogenomics of *Xanthomonas*: understanding bacterium-plant interactions. *Nature reviews. Microbiology* **9**, 344–355.
224. Sabuquillo P, Cubero J, 2016. Composición de la matriz extracelular de biofilms en *Xanthomonas arboricola* pv. *pruni*. In: *XVIII Congreso de la Sociedad Española de Fitopatología*. Palencia, España, 101.
225. Schattner P, Brooks AN, Lowe TM, 2005. The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. *Nucleic acids research* **33**, W686-9.
226. Scortichini M, 2004. *Xanthomonas arboricola* pv. *corylina*. *EPPO Bulletin* **34**, 179–181.
227. 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.
228. 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.

229. Seemann T, 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30**, 2068–2069.
230. Sena-Vélez M, 2015. Mecanismos implicados en las etapas iniciales de infección en la cancrrosis de los cítricos provocada por *Xanthomonas citri* subsp. *citri*. Doctoral dissertation. Politechnic University of Madrid. <http://oa.ipm.es/68610/>
231. Sena-Vélez M, Redondo C, Gell I *et al.*, 2015. Biofilm formation and motility of *Xanthomonas* strains with different citrus host range. *Plant Pathology* **64**, 767–775.
232. Sena-Vélez M, Redondo C, Graham JH, Cubero J, 2016. Presence of extracellular DNA during biofilm formation by *Xanthomonas citri* subsp. *citri* strains with different host range. *PLoS ONE* **11**, 1–17.
233. Shen Y, Chern M, Silva FG, Ronald P, 2001. Isolation of a *Xanthomonas oryzae* pv. *oryzae* flagellar operon region and molecular characterization of *flhF*. *Molecular plant-microbe interactions* **14**, 204–13.
234. Shepard DP, Zehr EI, Bridges WC, 1999. Increased susceptibility to bacterial spot of peach trees growing in soil infested with *Criconemella xenoplax*. *Plant Disease* **83**, 961–963.
235. Sherif S, El-Sharkawy I, Paliyath G, Jayasankar S, 2012. Differential expression of peach ERF transcriptional activators in response to signaling molecules and inoculation with *Xanthomonas campestris* pv. *pruni*. *Journal of Plant Physiology* **169**, 731–739.
236. Shrivastava S, Mande SS, 2008. Identification and functional characterization of gene components of type VI secretion system in bacterial genomes. *PLoS ONE* **3**, e2955.
237. Siciliano F, Torres P, Sendín L *et al.* Analysis of the molecular basis of *Xanthomonas axonopodis* pv. *citri* pathogenesis in *Citrus limon*. *Electronic Journal of Biotechnology* **9**, 200-204.

238. da Silva a CR, Ferro J a, Reinach FC *et al.*, 2002. Comparison of the genomes of two *Xanthomonas* pathogens with differing host specificities. *Nature* **417**, 459–463.
239. Da Silva FR, Vettore AL, Kemper EL, Leite A, Arruda P, 2001. Fastidious gum: the *Xylella fastidiosa* exopolysaccharide possibly involved in bacterial pathogenicity. *FEMS Microbiology Letters* **203**, 165–171.
240. Silva Vasconcellos FC Da, de Oliveira AG, Lopes-Santos L *et al.*, 2014. Evaluation of antibiotic activity produced by *Pseudomonas aeruginosa* LV strain against *Xanthomonas arboricola* pv. *pruni*. *Agricultural Sciences* **5**, 71–76.
241. Simpson AJ, Reinach FC, Arruda P *et al.*, 2000. The genome sequence of the plant pathogen *Xylella fastidiosa*. The *Xylella fastidiosa* consortium of the organization for nucleotide sequencing and analysis. *Nature* **406**, 151–159.
242. Skerman VBD, Sneath PHA, McGowan V, 1980. Approved lists of bacterial names. *International Journal of Systematic and Evolutionary Microbiology* **30**, 225–420.
243. Snipen L, Liland KH, 2015. microman: an R-package for microbial pan-genomics. *BMC Bioinformatics* **16**, 79.
244. Snipen L, Ussery DW, 2010. Standard operating procedure for computing pangenome trees. *Standards in Genomic Sciences* **2**, 135–141.
245. Socquet-Juglard D, Duffy B, Pothier JF, Christen D, Gessler C, Patocchi A, 2013a. Identification of a major QTL for *Xanthomonas arboricola* pv. *pruni* resistance in apricot. *Tree Genetics and Genomes* **9**, 409–421.
246. Socquet-Juglard D, Kamber T, Pothier JF *et al.*, 2013b. Comparative RNA-seq analysis of early-infected peach leaves by the invasive phytopathogen *Xanthomonas arboricola* pv. *pruni*. *PloS one* **8**, e54196.
247. Socquet-Juglard D, Patocchi A, Pothier JF, Christen D, Duffy B, 2012. Evaluation of *Xanthomonas arboricola* pv. *pruni* inoculation techniques to

- screen for bacterial spot resistance in peach and apricot. *Journal of Plant Pathology* **94**, S1.91-S1.96.
248. Solé M, Scheibner F, Hoffmeister A-K *et al.*, 2015. *Xanthomonas campestris* pv. *vesicatoria* secretes proteases and xylanases via the Xps type II secretion system and outer membrane vesicles. *Journal of Bacteriology* **197**, 2879–2893.
 249. Song C, Yang B, 2010. Mutagenesis of 18 type III effectors reveals virulence function of XopZ (PXO99) in *Xanthomonas oryzae* pv. *oryzae*. *Molecular plant-microbe interactions* **23**, 893–902.
 250. Souza DP, Andrade MO, Alvarez-Martinez CE, Arantes GM, Farah CS, Salinas RK, 2011. A component of the Xanthomonadaceae type IV secretion system combines a VirB7 motif with a NO domain found in outer membrane transport proteins. *PLoS Pathogens* **7**, e1002031.
 251. Souza DP, Oka GU, Alvarez-Martinez CE *et al.*, 2015. Bacterial killing via a type IV secretion system. *Nature Communications* **6**, 6453.
 252. Srivastava D, Waters CM, 2012. A tangled web: regulatory connections between quorum sensing and cyclic Di-GMP. *Journal of Bacteriology* **194**, 4485–4493.
 253. Stamatakis A, 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313.
 254. Stefani E, 2010. Economic significance and control of bacterial spot/canker of stone fruits caused by *Xanthomonas arboricola* pv. *pruni*. *Journal of Plant Pathology* **92**, S1.99-S1.103.
 255. Stothard P, Wishart DS, 2005. Circular genome visualization and exploration using CGView. *Bioinformatics* **21**, 537–9.
 256. Stoyanova M, Vancheva T, Moncheva P, Bogatzevska N, 2014. Differentiation of *Xanthomonas* spp. causing bacterial spot in Bulgaria based on Biolog system. *International journal of microbiology*, 495476.

257. Subramoni S, Suárez-Moreno ZR, Venturi V, 2010. Lipases as pathogenicity factors of plant pathogens. In (Timmis KN, eds) *Handbook of hydrocarbon and lipid microbiology*. Springer, 3269–3277.
258. Sun W, Dunning FM, Pfund C, Weingarten R, Bent AF, 2006. Within-species flagellin polymorphism in *Xanthomonas campestris* pv *campestris* and its impact on elicitation of *Arabidopsis* FLAGELLIN SENSING2-dependent defenses. *The Plant cell* **18**, 764–79.
259. Szczesny R, Jordan M, Schramm C *et al.*, 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.
260. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S, 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular biology and evolution* **28**, 2731–9.
261. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S, 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. *Molecular biology and evolution* **30**, 2725–9.
262. Tarkowski P, Vereecke D, 2014. Threats and opportunities of plant pathogenic bacteria. *Biotechnology Advances* **32**, 215–229.
263. Tatusov RL, 2000. The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Research* **28**, 33–36.
264. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Ciufo S, Li W, 2013. Prokaryotic Genome Annotation Pipeline. In (J Beck, D Benson, J Coleman, et al., Eds,) NCBI Handbook. Bethesda (MD), US: National Center for Biotechnology Information. <https://www.ncbi.nlm.nih.gov/books/NBK174280/?report=reader>
265. Terashima H, Kojima S, Homma M, 2008. Flagellar Motility in Bacteria. Structure and Function of Flagellar Motor. *International Review of Cell and Molecular Biology* **270**, 39-85.

266. Tettelin H, Riley D, Cattuto C, Medini D, 2008. Comparative genomics: the bacterial pan-genome. *Current Opinion in Microbiology* **11**, 472–477.
267. Tjou-Tam-Sin NNA, van de Bilt JLJ, Bergsma-Vlami M *et al.*, 2012. First report of *Xanthomonas arboricola* pv. *pruni* in ornamental *Prunus laurocerasus* in the Netherlands. *Plant Disease* **96**, 759.
268. Trébaol G, Manceau C, Tirilly Y, Boury S, 2001. Assessment of the genetic diversity among strains of *Xanthomonas cynarae* by randomly amplified polymorphic DNA analysis and development of specific characterized amplified regions for the rapid identification of *X. cynarae*. *Applied and Environmental Microbiology* **67**, 3379–3384.
269. Triplett LR, Hamilton JP, Buell CR *et al.*, 2011. Genomic analysis of *Xanthomonas oryzae* isolates from rice grown in the united states reveals substantial divergence from known *X. oryzae* pathovars. *Applied and Environmental Microbiology* **77**, 3930–3937.
270. Vandroemme J, Cottyn B, Baeyen S, De Vos P, Maes M, 2013. Draft genome sequence of *Xanthomonas fragariae* reveals reductive evolution and distinct virulence-related gene content. *BMC genomics* **14**, 829.
271. Vauterin L, Hoste B, Kersters K, Swings J, 1995. Reclassification of *Xanthomonas*. *International Journal of Systematic Bacteriology* **45**, 472–489.
272. Vauterin L, Rademaker J, Swings J, 2000. Synopsis on the taxonomy of the genus *Xanthomonas*. *Phytopathology* **90**, 677–682.
273. Vauterin L, Yang P, Alvarez A *et al.*, 1996. Identification of non-pathogenic *Xanthomonas* strains associated with plants. *Systematic and Applied Microbiology* **19**, 96–105.
274. Venieraki A, Tsalgatidou PC, Georgakopoulos DG, Dimou M, Katinakis P, 2016. Swarming motility in plant-associated bacteria. *Hellenic Plant Protection Journal* **9**, 16–27.

275. Vernière CJ, Gottwald TR, Pruvost O, 2003. Disease development and symptom expression of *Xanthomonas axonopodis* pv. *citri* in various citrus plant tissues. *Phytopathology* **93**, 832–43.
276. Vernière C, Pruvost O, Civerolo EL, Gambin O, Jacquemoud-Collet JP, Luisetti J, 1993. Evaluation of the Biolog substrate utilization system to identify and assess metabolic variation among strains of *Xanthomonas campestris* pv. *citri*. *Applied and environmental microbiology* **59**, 243–9.
277. Verstraeten N, Braeken K, Debkumari B *et al.*, 2008. Living on a surface: swarming and biofilm formation. *Trends in microbiology* **16**, 496–506.
278. Volk W, 1966. Cell wall lipopolysaccharides from *Xanthomonas* species. *Journal of Bacteriology* **91**, 39–42.
279. Vorhölter FJ, Niehaus K, Pühler A, 2001. Lipopolysaccharide biosynthesis in *Xanthomonas campestris* pv. *campestris*: A cluster of 15 genes is involved in the biosynthesis of the LPS O-antigen and the LPS core. *Molecular Genetics and Genomics* **266**, 79–95.
280. Vorhölter F-J, Schneiker S, Goesmann A *et al.*, 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.
281. Wadhams GH, Armitage JP, 2004. Making sense of it all: bacterial chemotaxis. *Nature reviews. Molecular cell biology* **5**, 1024–1037.
282. Waksman G, Orlova E V, 2014. Structural organisation of the type IV secretion systems. *Current Opinion in Microbiology* **17**, 24–31.
283. Wang L, Rong W, He C, 2008. Two *Xanthomonas* extracellular polygalacturonases, PghAxc and PghBxc, are regulated by type III secretion regulators HrpX and HrpG and are required for virulence. *Molecular Plant-Microbe Interactions* **21**, 555–563.

284. Waters CM, Bassler BL, 2005. Quorum sensing: cell-to-cell communication in bacteria. *Annual Review of Cell and Developmental Biology* **21**, 319–346.
285. White FF, Potnis N, Jones JB, Koebnik R, 2009. The type III effectors of *Xanthomonas*. *Molecular Plant Pathology* **10**, 749–766.
286. Williams KP, Kelly DP, 2013. Proposal for a new class within the phylum Proteobacteria, Acidithiobacillia classis nov., with the type order Acidithiobacillales, and emended description of the class Gammaproteobacteria. *International Journal of Systematic and Evolutionary Microbiology* **63**, 2901–2906.
287. Winker S, Woese CR, 1991. A definition of the domains Archaea, Bacteria and Eucarya in terms of small subunit ribosomal RNA characteristics. *Systematic and applied microbiology* **14**, 305–310.
288. Woese CR, Kandler O, Wheelis ML, 1990. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proceedings of the National Academy of Sciences of the United States of America* **87**, 4576–4579.
289. van der Wolf J, De Boer SH, 2015. Phytopathogenic bacteria. In (Lutenberg B, eds) *Principles of Plant-Microbe Interactions*. Switzerland. Springer International Publishing, 65–77.
290. Xu J, 2006. Microbial ecology in the age of genomics and metagenomics: concepts, tools, and recent advances. *Molecular Ecology* **15**, 1713–1731.
291. Yamazaki A, Hirata H, Tsuyumu S, 2008. HrpG regulates type II secretory proteins in *Xanthomonas axonopodis* pv. *citri*. *Journal of General Plant Pathology* **74**, 138–150.
292. Yan Q, Wang N, 2011. The ColR/ColS two-component system plays multiple roles in the pathogenicity of the citrus canker pathogen *Xanthomonas citri* subsp. *citri*. *Journal of Bacteriology* **193**, 1590–1599.

293. Yin F, Zhao Y, Liu C *et al.*, 2011. Analysis of the *flgD* and *flgE* genes regulated by diffusible signal factor in *Xanthomonas oryzae* pv. *oryzicola*. *Acta microbiologica Sinica* **51**, 891–897.
294. Young JM, Park D-C, Shearman HM, Fargier E, 2008. A multilocus sequence analysis of the genus *Xanthomonas*. *Systematic and applied microbiology* **31**, 366–77.
295. Young JM, Wilkie JP, Park DC, Watson DRW, 2010. New Zealand strains of plant pathogenic bacteria classified by multi-locus sequence analysis; proposal of *Xanthomonas dyei* sp. nov. *Plant Pathology* **59**, 270–281.
296. Yun MH, Torres PS, El Oirdi M *et al.*, 2006. Xanthan induces plant susceptibility by suppressing callose deposition. *Plant physiology* **141**, 178–87.
297. Zaccardelli M, Ceroni P, Mazzucchi U, 1999. Amplified fragment length polymorphism fingerprinting of *Xanthomonas arboricola* pv. *pruni*. *Journal of Plant Pathology* **81**, 173–179.
298. Zehr EI, Shepard DP, Bridges Jr WC, 1996. Bacterial spot of peach as influenced by water congestion, leaf wetness duration, and temperature. *Plant Disease* **80**, 339–341.
299. Zheng D, Yao X, Duan M *et al.*, 2016. Two overlapping two-component systems in *Xanthomonas oryzae* pv. *oryzae* contribute to full fitness in rice by regulating virulence factors expression. *Scientific Reports* **6**, 22768.

Material Suplementario

Material suplementario, capítulo 3:

S1 Fig. Nucleotide genome sequence comparison of *X. arboricola* pv. *pruni* strains CITA 33 and IVIA 2626.1 and the non-pathogenic strain CITA 44. Genome sequences were compared against each other based on Blastn results and represented as a circular map using the CGview tool. From outside to center: strain CITA 33, strain IVIA 2626.1, strain CITA 44, GC content, GC skew + and GC skew -. doi:10.1371/journal.pone.0161977.s001

S2 Fig. Transmission electron microscopy of *X. arboricola* strain CITA 44. Negative staining showed monoflagellated cell from the edge of the dendritic swarming colony 24 hpi in 0.5% PYM agar plates. doi:10.1371/journal.pone.0161977.s002

S3 Fig. Amino acid alignment of two variants of the FliC found in *X. arboricola*. Sequence alignment, performed using ClustalW, showed the amino acid change in the position number 43 of the amino terminal region of FliC. doi:10.1371/journal.pone.0161977.s003

S4 Fig. Nucleotide alignment of type III effector *xopAQ* and the PIP-box found in *X. arboricola* pv. *pruni*. Sequence alignment, performed using ClustalW, showed slight variation among *X. arboricola* and other *Xanthomonas* (A) as well as a variant in PIP-box sequence (B). doi:10.1371/journal.pone.0161977.s004

Nota: Las tablas referentes al material suplementario se encuentran disponibles de manera digital en los siguientes enlaces:

S1 Table. Genome statistics of 18 *X. arboricola* strains according to the automatic annotation performed using Prokka.

doi:10.1371/journal.pone.0161977.s005

S2 Table. Annotated protein sequences shared by 18 pathogenic and non-pathogenic strains of *X. arboricola*.

doi:10.1371/journal.pone.0161977.s006

S3 Table. Unique protein coding sequences (CDS) and related features found in the draft genome sequence of *X. arboricola* strain CITA 44 and *X. arboricola* pv. *pruni* strains CITA 33 and IVIA 26262.1.

doi:10.1371/journal.pone.0161977.s007

S4 Table. Orthologous protein coding sequences (CDS) of nine strains of *X. arboricola* associated with pathogenesis.

doi:10.1371/journal.pone.0161977.s008

Material suplementario, capítulo 4:

Supplementary figure 1. Schematic representation of bacterial-caused symptoms on *P. persica* (GF-305), *H. vulgare*, *N. benthamiana*; *N. tabacum* and *S. lycopersicum* 21 dpi. Leaves were infiltrated with 10^8 CFU/mL of bacteria or with sterile phosphate saline buffer (PBS) as negative control.

Supplementary figure 2. Comparative statistics among pathogenic and non-pathogenic strains of *X. arboricola* based in the distribution of the ortholog clusters of gene contained in the analysed genome sequences. (A) Histogram representing the Jaccard distance distribution among the analysed genomes. (B) Principal component analysis based in the ortholog clusters gene present in the pangenome of *X. arboricola* showing how the genomes are located in the space spanned by the two first principal components.

- *X. arboricola* + *X. arboricola* pv. *celebensis*
- *X. arboricola* pv. *pruni*
- *X. arboricola* pv. *juglandis*
- *X. arboricola* pv. *corylina*

Supplementary figure 3. Phylogenetic analysis of 17 strains of *X. arboricola* based on the core genome sequence (2,714 potential groups of homologous genes) and representation of the distribution of the potential ortholog cluster genes of the pangenome (7,074) within the analysed genome sequences. Sequences were aligned using PRANK and maximum likelihood analysis was carried out using RaxML. Bootstrap values (1,000 replicates) are presented above or below the branches.

Supplementary figure 4. Primers and TaqMan probe designed to amplify *xopE3* using real time PCR in *Xanthomonas* spp. Xap and CFBP 5530: *X. arboricola* pv. *pruni*; Xcc: *X. citri* subsp. *citri*, Xff: *X. fuscans* subsp. *fuscans*.

Supplementary figure 5. *In silico* representation of the hybridization zone for the primers used in two PCR amplification protocols published to identify *X. arboricola* pv. *pruni*. Comparative sequence analysis was performed using Blastn with an expected value of 0.001. The circular graphic has been constructed using BRIG. Each concentric circle represents one of the analysed genomes. Full references of the two methods are available in the article.

Nota: Las tablas referentes al material suplementario se encuentran disponibles de manera digital en los siguientes enlaces:

Supplementary Table 1. Genome statistics of 17 pathogenic and non-pathogenic strains of *X. arboricola*.

<https://drive.google.com/open?id=0B7HDVNVyy8PUVjdweW1MMjFybE0>

Supplementary Table 2. Unique clusters of orthologous genes encoded in the genome sequence of the non-pathogenic strains of *X. arboricola* CITA 14 and CITA 124 as well as those encoded in the pathogenic strains of the pathovars *corylina*, *juglandis* and *pruni*.

<https://drive.google.com/open?id=0B7HDVNVyy8PUdjRUOURLUW9RbkU>

Supplementary Table 3. Orthologous protein coding sequences (CDS) of seventeen strains of *Xanthomonas arboricola* associated with pathogenesis.

<https://drive.google.com/open?id=0B7HDVNVyy8PUejJVMkZFdWJBmM>