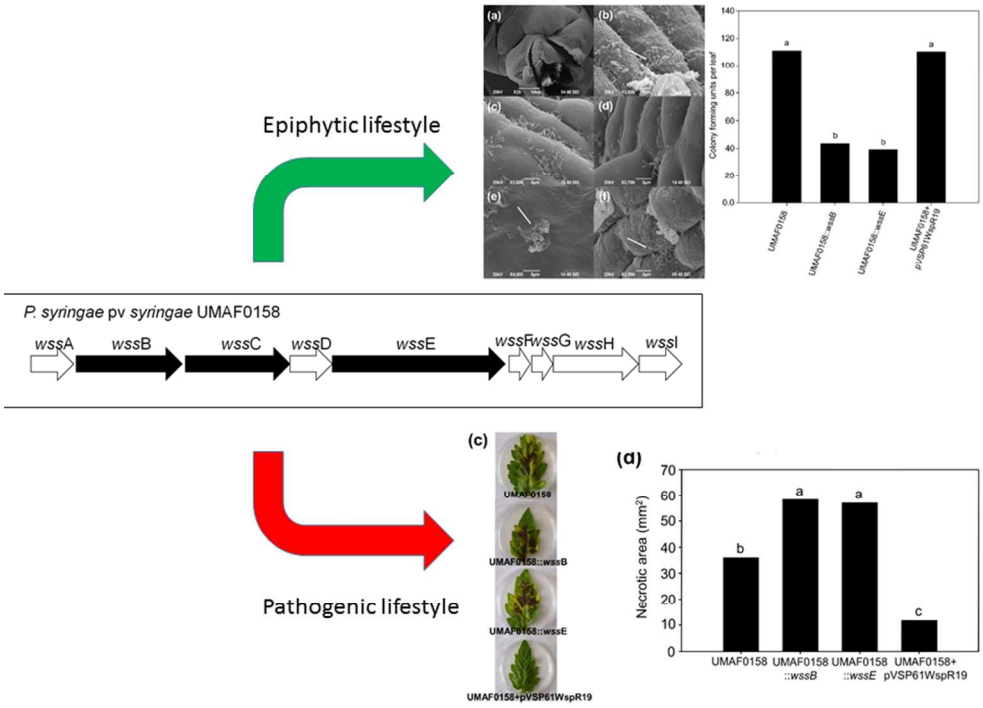
OXFORD UNIVERSITY PRESS | FEMS Microbiology Ecology

http://mc.manuscriptcentral.com/fems

Cellulose production in Pseudomonas syringae pv. syringae: a compromise between epiphytic and pathogenic lifestyles

Journal:	*FEMS Microbiology Ecology*
Manuscript ID:	FEMSEC-14-11-0635.R3
Manuscript Type:	Research Paper
Date Submitted by the Author:	19-Jun-2015
Complete List of Authors:	Arrebola, Eva; Universidad de Málaga, Microbiology Carrion, Victor; Universidad de Málaga, Microbiology Gutierrez-Barranquero, Jose; Universidad de Málaga, Microbiology Pérez-García, Alejandro; University of Malaga, Departament of Microbiology Rodriguez-Palenzuela, Pablo; Universidad Politecnica de Madrid, Bioquimica Cazorla, Francisco; Universidad de Málaga, Microbiología de Vicente, Antonio; Universidad de Málaga, Microbiología
Keywords:	Virulence, Biofilm, Phyllosphere, Woody plants, Bacterial EPS



254x190mm (96 x 96 DPI)

Cellulose production in *Pseudomonas syringae* pv. *syringae*: a compromise between epiphytic and pathogenic lifestyles

Eva Arrebola[1], Víctor J. Carrión[2], José Antonio Gutiérrez-Barranquero[2], Alejandro Pérez-García[2], Pablo Rodríguez-Palenzuela[3], Francisco M. Cazorla[2] and Antonio de Vicente[2]

[1]Instituto de Hortofruticultura Subtropical y Mediterránea "La Mayora", Universidad de Málaga, Consejo Superior de Investigaciones Científicas, Estación Experimental La Mayora, Algarrobo-Costa, 29750 Málaga, Spain; [2]Instituto de Hortofruticultura Subtropical y Mediterránea "La Mayora", Universidad de Málaga, Consejo Superior de Investigaciones Científicas, Departamento de Microbiología, Facultad de Ciencias, Campus de Teatinos, 29071 Málaga, Spain; [3]Centro de Biotecnología y Genómica de Plantas UPM-INIA. Parque Científico y Tecnológico de la Universidad Politécnica de Madrid. Campus de Montegancedo, 28223 Pozuelo de Alarcón, Madrid, Spain.

Corresponding author: Antonio de Vicente, E-mail: adevicente@uma.es

Keywords: virulence, biofilm, phyllosphere, woody plants, bacterial EPS

Running title: Cellulose is involved in the epiphytic fitness of UMAF0158

Abstract

Genome sequencing and annotation have revealed a putative cellulose biosynthetic operon in the strain *Pseudomonas syringae* pv. *syringae* UMAF0158, the causal agent of bacterial apical necrosis. Bioinformatics analyses and experimental methods were used to confirm the functionality of the cellulose biosynthetic operon. In addition, the results showed the contribution of the cellulose operon to important aspects of *P. syringae* pv. *syringae* biology, such as the formation of biofilms and adhesion to mango leaf surface, suggesting that this operon increases epiphytic fitness. However, based on the incidence and severity of the symptoms observed in tomato leaflets, cellulose expression reduces virulence, as cellulose-deficient mutants increased the area of necrosis, whereas the cellulose-overproducing strain decreased the area of necrosis compared to the wild type. In conclusion, the results of this study show that the epiphytic and pathogenic stages of the *P. syringae* pv. *syringae* UMAF0158 lifestyle are intimately affected by cellulose production.

Introduction

Cellulose is a homo-polysaccharide consisting of glucose units linked by β1-4 glycosidic bonds and is the most abundant organic polymer present on Earth (Delmer & Amor, 1995). Cellulose is produced by many bacterial species including *Gluconacetobacter xylinus* (*Acetobacter xylinus*) (Coucheron *et al.*,

1991). The polymer has been identified as a primary component of the extracellular matrix produced by a diversity of biofilm-forming bacteria (Le Quéré & Ghigo, 2009; Zogaj *et al.*, 2003; Lu *et al.*, 2012), including environmental and plant pathogen pseudomonads (Ude *et al.*, 2006; Robertson *et al.*, 2013). The expression of large amounts of cellulose, like many other exopolysaccharides, may protect cells in biofilms and lead to biocide resistance (O'Toole et al. 2000). Furthermore, bacterial cellulose has been shown to be required for cell adhesion to host tissues (Matthysse *et al.*, 1981).

Cellulose-based biofilm formation has also been associated with the virulence of several human and animal diseases (Fux *et al.*, 2005; Yildiz, 2007) such as the enteric bacteria *Escherichia coli* (Weiss-Muszkat *et al.*, 2010) and *Salmonella enterica* (Lu *et al.*, 2012). Pellicle formation by the plant pathogen *Dickeya dadantii* is dependent on cellulose production (Jahn *et al.*, 2011; Pringent-Combaret *et al.*, 2012) encoded by the bacterial cellulose synthase (*bcs*) operon (Wong *et al.*, 1990). Furthermore, cellulose production in *D. dadantii* favours plant surface colonisation and a mutation in the *bcs* operon also decreases resistance to chlorine stresses (Pringent-Combaret *et al.*, 2012). *Pseudomonas putida* mt2 has a similar *bcs* operon involved in cell-surface and cell-cell interactions necessary for biofilm formation and which also contributes to rhizosphere fitness (Nielsen *et al.*, 2011). *Pseudomonas fluorescens* SBW25 also encodes a *bsc*-like *wss* operon, and an over-producing cellulose mutant known as the "Wrinkly Spreader" colonises the air-liquid interface of static microcosms through the formation of self-supporting biofilms (Spiers *et al.*, 2002; Spiers *et al.*, 2003; Spiers, 2014). Cellulose has also been identified as a

matrix component in other *Pseudomonas* spp. biofilms, including *P. syringae* (Ude *et al.*, 2006).

P. syringae is a wide host range pathogen, affecting both herbaceous and woody plants and causing serious diseases in crops (Hirano & Upper, 1990, Cazorla *et al.*, 1998; Kennelly *et al.*, 2007; Gutiérrez-Barranquero *et al.*, 2011). Recently we sequenced the complete genome of *P. syringae* pv. *syringae* UMAF0158 (under submission; GenBank CP005970), isolated from the necrotic tissue of a mango tree (Cazorla *et al.*, 1998); among other traits, we have identified a gene cluster orthologous to the *P. fluorescens* SBW25 *wss* operon. Because cellulose production is directly involved in biofilm formation and is directly or indirectly associated with other aspects of plant-pathogenic bacteria, the aim of our present study is to clarify the role of cellulose in this woody plant pathogen.

Materials and Methods

Bacterial strains and growth conditions

For standard maintenance, *P. fluorescens* SBW25 (Spiers *et al.*, 2002), *P. syringae* pv. *syringae* UMAF0158 (Arrebola *et al.,* 2003) and B728a (Hirano & Upper, 2000) were grown in King's medium B (KB) (King et al., 1954) and incubated at 28ºC for 48 h. UMAF0158::*wss*B and UMAF0158::*wss*E were maintained with 50 µg mL^{-1} kanamycin. The conditions used for plant experiments are specified in the corresponding sections.

Bioinformatics

Database searches were performed using the NCBI (http://www.ncbi.nlm.nih.gov) and ASAP (http://asap.ahabs.wisc.edu/asap/ASAP1.htm) websites. Homology searches and analyses of conserved domains were performed using the NCBI BLAST and the Pfam database (http://pfam.sanger.ac.uk). A complementary search of the protein pattern was performed using the PROSITE™ database (http://prosite.expasy.org). GC analyses, promoter, terminator and IS analyses, and primer design were performed using a GC calculator (http://www.genomicsplace.com/gc_calc.html), SoftBerry (http://linux1.softberry.com/berry.phtml), IS Finder (http://www-is.biotoul.fr) and Primer3 (http://primer3.sourceforge.net).

Phylogenetic Analysis

Phylogenetic analysis of different *Pseudomonas spp*. strains was performed using *E. coli* K-12, *S. typhimurium* LT2 and *E. amylovora* 273 as out groups (Supplementary material, Table S3). Sequences were aligned using ClustalW (Larkin *et al.*, 2007). The cellulose synthase subunits *wssB*, *wssC* and *wssE* (Spiers *et al.*, 2002) and the housekeeping genes *fruK*, *gapA*, *gltA*, *pgi*, *recA*, *rpoA*, *rpoB* and *rpoD* were used for phylogenetic comparison. Phylogenetic analysis was performed using MEGA 5.0 software (Tamura *et al.*, 2007, 2011). Phylogenetic trees were constructed using maximum likelihood fits based on a data-specific model (Nei & Kumar, 2000).

Strain manipulation and molecular assays

Insertional inactivation mutagenesis of *P. syringae* UMAF0158 was used to suppress cellulose production by inserting disruption vectors into the different ORFs of the cellulose operon via single-crossover homologous recombination. The PCR reaction, pCR2.1-TOPO® (Invitrogen Life Tech, USA) cloning and plasmid purification were performed using standard procedures and primers are listed in the Supplementary material, Table S5. Plasmids were transformed into the wild-type strain UMAF0158 using electroporation. The correct insertions of disruption vectors were analysed by PCR and the bacterial growth curves were obtained in KB to confirm similarities with the wild type (data not shown).

The wild-type strain UMAF0158 was transformed with the kanamycin-resistant, self-replicating plasmid pVSP61-WspR19 to produce a strain over-producing cellulose (Ude *et al.*, 2006).

Total RNA was isolated (Castillo *et al.*, 2007) and RT-PCR was performed using the Titan OneTube RT-PCR system (Roche Diagnostics, Basel, Switzerland) and primers are listed in the Supplementary material, Table S5. The transcription start point for the *wss* operon was determined using the 5'-RACE method (Maruyama *et al.* 1995; Filiatrault *et al.*, 2010, 2011; Carrión *et al.*, 2012).

Microscopy

Bacterial suspensions were made in sterile distilled water and adjusted to 10^8 cfu mL^{-1}, and then were sprayed onto mango buds and left for three days on plants maintained in a greenhouse. Mango buds were fixed in 2.5% glutaraldehyde and dehydrated by increasing the ethanol concentration. The

samples were maintained in 100% ethanol until processing for Scanning Electronic Microscopy (SEM). At least three replicates for every sample were observed.

In vitro analysis of cellulose production

Strains were grown in KB medium supplemented with $FeCl_3$ (0.25 g L^{-1}) to avoid siderophore production and 10 μM Calcofluor (SIGMA Fluorescent Whitener 28) for 3 days at 28ºC before visualization and images analysis (See Supplementary material and methods section). The pictures were processed using Visilog 5.0 software (Noesis Vision Inc., France). The colour images were transformed to grey scale pictures. The measured area for each colony was 17202 pixels (0.1388 mm^2). At least three colonies per image and six images per strain and experiment were analysed. Five independent experiments were performed to obtain the cellulose production results. The values were obtained as arbitrary units (a.u.), defined as the value obtained according to a grey scale ranging from 0.0 (pure black) to 255 (pure white) in a 0.1388 mm^2 area.

Biofilm quantification

Liquid cultures were adjusted to 10^8 cfu mL^{-1} and incubated at 28ºC overnight in 96-well microtiter plates to assay biofilm formation. Biofilm quantification were performed as previously described (Peeters *et al.*, 2008). The absorbance of the eluted Crystal violet was measured at 595 nm. At least 24 wells per strain and experiment were used, and four independent experiments were performed to obtain the biofilm results.

Adhesion assay on mango leaves

For adhesion experiments, bacterial suspensions were adjusted to 10^8 cfu mL^{-1}. Drops (10 µL) of each strain were inoculated onto mango leaves wiped with 70% ethanol. After 30 min the leaves were washed with sterile water and the inoculated region removed for proccessing. The leaf pieces were placed into sterile bags with 1 mL of sterile water, homogenised for 3 min, and plated onto KB plates to determine bacterial numbers. Three plates per strain were used, with three replicates per experiment and three independent experiments performed.

Pathogenesis and competitive index evaluation

The assayed strains were inoculated onto detached tomato leaflets (*Solanum lycopersicum* L. cv. Hellfrucht-Früstamm) to evaluate virulence, as previously described (Arrebola *et al.*, 2007, 2009; Carrión *et al.*, 2014). The appearance of necrotic symptoms was monitored by using visual analysis to evaluate disease incidence (number of inoculated points showing necrotic symptoms equal or greater than 0.5 cm in diameter). The total necrotic area per leaflet induced at the last day (10^{th} day) of the experiment (severity) was determined from six leaflets over three independent experiments using the image analysis software Visilog 5.0 (Noesis Vision Inc., France). In addition, two inoculated leaflets were used every day for seven days to estimate the total bacterial population in every strain used.

Additional experiments to obtain a competitive index were performed using tomato leaflets maintained *in vitro*. A bacterial suspension of the wild type UMAF0158 strain was mixed at 1:1 ratio with cellulose-defective mutants, and

the cellulose overproducer strain suspensions was adjusted to 10^8 cfu mL^{-1}, and then processed similarly to the pathogenicity test. The leaflets were processed after 1 h to observe the initial symptom points and after 7 days to observe the final symptom points. Three leaflets per strain and experiment and three independent experiments were performed to obtain competitive index results.

Statistical Analysis

Statistical analyses were performed using IBM.SSPS 19 software (IBM® Company, Armonk, NY). One factor ANOVA was used for the analysis of the means with p=0.05, homogeneity test data were analysed by the Levene test, and statistic descriptive effects homogeneity was analysed by Brown-Forsythe Welch test, with post hoc analysis by LSD α= 0.05 and a 95% confidence interval. The statistical analysis of incidence of necrotic symptoms was performed using SAS9.2 software (SAS Institute Inc., Cary, NC, USA) in the Enterprise Guide 4.2 program with generalised linear model analysis.

Results

Bioinformatics analysis of a cellulose biosynthetic operon from *P. syringae* pv. *syringae* UMAF0158.

The complete genomic sequences of UMAF0158 containing the cellulose synthase operon have been deposited in GenBank under the accession number CP005970 (Martínez-García *et al.* Direct Submission). A gene cluster of 14,642 bp (chromosome section 4684145–4696189 bp; corresponding genes from PsyrMG_20805 to PsyrMG_20845), consisting of nine genes, was annotated

with putative functions associated with cellulose production and acetylation (Table 1; Supplementary material Table S1).

The *wss* cluster located in *P. syringae* pv. *syringae* UMAF0158 had an arrangement similar to the cellulose operons of other *Pseudomonas* strains (Fig. 1). Further sequence analysis using the cellulose operons reported in *Pseudomonas spp.* strains showed high sequence similarity in all ORFs (Supplementary material, Table S2). The only exceptions were *P. syringae* pv. *aesculi* 2250, which showed no similarity with the *wssB* and *wssI* genes, and *P. fluorescens* SBW25, which showed the lowest similarity with the whole *wss* operon at the nucleotide level.

To obtain insight into the evolutionary history of the cellulose operon of *Pseudomonas* spp., three concatenated genes (*wssB*, *wssC* and *wssE*), corresponding to three of the four cellulose synthase subunits, were used to construct phylogenetic trees. *E. coli*, *S. typhimurium* and *Erwinia amylovora* were used as out-groups (Fig. 2a; Supplementary material, Table S3). A phylogenetic tree was also constructed using eight different housekeeping genes (Fig. 2b). The phylogenetic relationship of the cellulose genes did not reveal important variations with respect to the phylogenetic relationship of the housekeeping, but notably, the absence/presence of the cellulose cluster was distributed among pathogens and non-pathogens (Fig. 2).

Additionally, the sequences of the cellulose operon and flanking regions were compared between *P. syringae* UMAF0158, *P. syringae* pv *tomato* DC3000, *P. syringae* pv. *aesculi* 2250 and *P. savastanoi* pv. *savastanoi* NCPPB3335, and two *P. syringae* strains, that lacked cellulose operons, pv. *syringae* B728a and pv. *phaseolicola* 1448A. The results showed identical localisation in all cellulose

operon-containing *P. syringae* strains. Notably, in the strains lacking cellulose operons, the equivalent genomic region showed a sequence with high identity to the flanking regions of the cellulose operon of UMAF0158, suggesting that this region represents the remnants of a cellulose operon. Indeed, in *P. syringae* pv. *phaseolicola* 1448A, this region showed 83% identity with a non-coding region upstream of *wssA* and 72% identity with a 116-bp region at the 3'-end of *wssI* in the strain UMAF0158. Similarly, *P. syringae* pv. *syringae* B728a contains a hypothetical protein (Psyr_0881) with 89% identity to the 3'-end (24% of the total sequence) of *wssI* in the strain UMAF0158.

Operon structure of the cellulose cluster in the *P. syringae* pv. *syringae* UMAF0158 strain.

RT-PCR amplifying 17 fragments across the *wss* operon confirmed that the genes were co-transcribed as a single polycistronic RNA transcript (Fig. 3a). Bioinformatics analysis of the non-coding sequence upstream of *wssA* revealed -10 and -35 boxes corresponding to a putative promoter, and the RpoD sigma transcription factor was also detected (Fig. 3b). In addition, an analysis of the potential insertion sequences (IS) was performed for this 246-nucleotide DNA fragment, showing eight different sequences (Fig. 3b; Supplementary material, Table S4). 5'-RACE analysis of the promoter revealed nucleotide +1 73 bp upstream of the first ATG codon of *wssA*, and a Shine-Dalgarno (SD) sequence (Chen *et al.*, 1994) located 9 bp upstream of the start codon (Fig. 3b). A 363-nucleotide DNA sequence located downstream of *wssI* was identified as a terminator sequence with a typical terminator hairpin secondary structure. Further RT-PCR analysis confirmed that the transcription of this operon stopped

after the *wssI* gene (Fig. 3c). Finally, thirteen different IS sequences were found at the end of the *wss* operon of UMAF0158 (Fig. 3c; Supplementary material, Table S4).

P. syringae pv. _syringae_ UMAF0158 cellulose production and its role in biofilm formation and plant surface colonisation.

To confirm the functionality of the *wss* operon in UMAF0158, mutants in two of the essential conserved genes were constructed: UMAF0158::*wssB*, and UMAF0158::*wss*E, both of which are critical subunits in the cellulose synthase complex (Spiers *et al.* 2013; Spiers, 2014). As expected, these two mutants were defective in cellulose production (Fig. 4). Additionally, UMAF0158 transformed with the plasmid pVS61-WsR19, containing *wsp*R19 from *P. fluorescens* SBW25 (Ude *et al.*, 2006), had a cellulose overproduction phenotype. Clear differences were seen in cellulose production between wild-type UMAF0158, cellulose defective mutants and the overproducing strain (Fig. 4a), confirming that cellulose production required a functional *wss* operon.

In addition, the role of cellulose in biofilm formation was also analysed. A relationship between the presence and absence of the *wss* operon, cellulose production and biofilm formation was observed by Crystal violet staining of bacteria adhering to microtiter plates in an *in vitro* assay of biofilm formation (Fig. 4b). Scanning electron microscopy of colonising bacteria on mango buds also showed differences among the strains with or without active *wss* operons, consistent with the results described above, confirming matrix production in epiphytic colonisation (Fig. 5). Notably, it was possible to observe micro-colonies of the wild-type and overproducer UMAF0158 strains immersed in the

extracellular matrix. In contrast, the cellulose-defective mutants showed much less matrix production (Fig. 5).

To evaluate the influence of cellulose production on the *P. syringae* pv. *syringae* UMAF0158 epiphytic lifestyle, adhesion experiments on mango leaves were performed (Fig. 6). This showed that the amount of bacteria recovered from the surface of mango leaves were significantly higher in strains with an active *wss* operon. Altogether, these results confirmed the functionality of the *wss* operon and the role of cellulose production in biofilm formation and leaf colonisation.

Involvement of cellulose production in the virulence of *P. syringae* pv. *syringae* UMAF0158.

Evaluation of the virulence was performed on tomato leaflets (Arrebola *et al*., 2007, 2009), which is a more reliable plant model for pathogenicity. Growth curves for strains with or without an active *wss* operon were obtained from inoculated tomato leaflets. Bacterial counts displayed similar growth patterns (Fig. 7a). The number of necrotic points developed (incidence) in the cellulose defective mutants was not significantly different from that observed with the wild type. However, the overproducer strain had a delayed onset and a lower number of necrotic symptoms (Fig. 7b). To complete the virulence study, the overall necrotic area (severity) for each strain was also estimated in these experiments (Fig. 7c and d). The results demonstrated significant differences between the defective mutants, the overproducer and the wild-type strains. Indeed, both cellulose-defective mutants had the highest amount of necrotic areas and the overproducer strain showed the lowest amount of necrotic area, (Fig. 7d), suggesting the negative influence of cellulose production on the

virulence of this bacterium. Moreover, the competitive indices of the cellulose-defective mutants and overproducer strain were determined with respect to the wild-type strain (Fig. 8). Only the overproducer strain had a significantly lower competitive index, in agreement with the low virulence shown in the pathogenicity test.

Discussion

The annotation of the completely mapped genome of this strain and the bioinformatics analysis confirmed the sequence similarity of the identified cellulose biosynthetic operon with other available and characterised cellulose operons (Wong *et al.*, 1990; Sofia *et al.*, 1994; Zogaj *et al.*, 2001, Spiers *et al.*, 2002; Ude *et al.*, 2006; Smits *et al.*, 2010) and provided strong support for an operon-like structural organisation which was subsequently confirmed, as well as its functionality. Sequence analysis of this operon showed motifs for polymer acetylation in *P. syringae* UMAF0158, which provides more stability and robustness to the produced cellulose (Spiers & Rainey, 2005). In *P. fluorescens* SBW25, four genes involved in cellulose acetylation (*wssFGHI*) have been reported (Spiers, 2014). The orthologous genes in UMAF0158 have been also identified. To the best of our knowledge, SBW25, DC3000 and UMAF0158 are the only pseudomonads to be identified as having *wssFGHI*, which makes the partial acetylation of cellulose by these strains as unique (Spiers *et al.,* 2013). The phylogenetic relationships and their locations in the genomes among *Pseudomonas* spp*. wss* operons were compared, showing a common ancestor even for *P. syringae* strains that do not currently contain the *wss* operon. The presence of a large number of insertion sequences (IS) in the cellulose operon

flanking regions suggest the potential instability of the *wss* operon (Coucheron *et al.*, 1991) and present an explanation for why the *wss* operon is not present in all *P. syringae* pathovars and strains.

To determine whether cellulose production of UMAF0158 is involved in plant colonisation, the phenotypes of two cellulose production-defective mutants, a cellulose overproducer strain and the wild-type strain were compared. The results showed the clear involvement of cellulose in biofilm formation and plant surface adhesion, consistent with previous studies (Matthysse *et al.*, 1981; Gal *et al.*, 2003; Spiers *et al.* 2003; Le Quéré & Ghigo, 2009; Nielsen *et al.*, 2011). The overproducer strain drove increased biofilm formation, resulting from the increased expression of polysaccharides and secreted adhesion proteins, promoting auto-aggregation (O'Toole *et al.*, 2000). This cellulose overproduction occurs through the unregulated expression of *wsp*R19 *in trans*, which produces an increase in the c-di GMP levels, causing higher activation of *wss* transcription (Spiers *et al.*, 2002) as well as increased cellulose expression and biofilm formation (Spiers *et al.*, 2003; Ude *et al.*, 2006; Giddens *et al.*, 2007; Nielsen *et al.*, 2011). These results support the involvement of cellulose in the epiphytic fitness of *P. syringae* UMAF0158, similar to *P. fluorescens* SWB25 during plant root colonisation (Gal *et al.*, 2003). However, the fact that UMAF0158 cells are able to stick on the mango leaf surface during the epiphytic phase contradict the data found with B728a by Yu et al. (2013). These authors found that flagellar motility was favoured during the epiphytic stage through a global transcriptome profiling. These results demonstrate the diversity of strategies in the same pathovar to live on/in the host plant; this could even explain why B728a lacks the *wss* operon.

P. syringae UMAF0158 is also a pathogenic bacterium, and cellulose production could influence its pathogenicity, similar to *Enterobacteriaceae* (Lu *et al.*, 2012). The analysis of the symptoms did not show evident differences among the cellulose-defective mutants and the wild-type strain; however, the mutants had a significant increase in the induced necrotic area (severity) compared with the wild-type strain. In contrast, the overproducer strain showed a significant decrease. The assessment of these two parameters, incidence and severity, demonstrated that the virulence of this pathogen is higher in the absence of cellulose production, because mobility was not increased in these mutants (data not shown). Notably, when the amount of cellulose in the exopolysaccharide layer is high (overproducer strain), the virulence of these bacteria is significantly decreased. Additionally, the competitive index revealed that only the overproduction of cellulose resulted in a reduction in competitiveness with wild-type bacteria during the infection of plant tissues, supporting the previous results. These results suggest that the produced cellulose material is not a virulence factor but could even act negatively on virulence. The cellulose overproducer cells could be attached more strongly in the epiphytic biofilm and it could result in less cells leaving biofilm to roam through the apoplast to cause disease.

Consistent with these data, the presence or absence of cellulose production in different strains of *P. syringae* pathovars may determine the lifestyle of the strain studied, favouring each strain for the prevalence of the epiphytic or pathogenic lifestyles. Thus, a highly specialised pathogenic strain that does not produce cellulose would be favoured. Bacteria that cause disease on plants are typically much better colonisers of the plant than those that do not (Hirano &

Upper, 2000). The plant host of *P. syringae* pv. *syringae* UMAF0158 is the mango tree (*Mangifera indica* L.), producing a typical disease known as bacterial apical necrosis of mango. This disease is characterised by epiphytic bacterial survival during the spring and summer and is made evident by symptom development during the autumn and winter (Cazorla *et al.*, 1998). We recently described a new phylotype including primarily pv. *syringae* strains adapted to mango trees and other woody hosts (Gutiérrez-Barranquero *et al.*, 2013). It is likely that this bacterium maintained cellulose production as an adaptation and specialisation during leaf and bud colonisation, although these adaptation reduced virulence; this observation implies an adjusted regulation process of the *wss* operon depending on the stage of life of the bacterium on the plant host, whereby cellulose production is induced in the epiphytic phase and repressed during the pathogenic phase. In fact, the second messenger c-di-GMP is a known regulator between the sessile and mobile phases in a huge range of bacterial species (Römling *et al.,* 2013). This relevant role of cellulose production in the *P. syringae* ecology could help to understand also its role in the biofilm formation, surface colonization and ecology of other cellulose-producing bacteria.

Acknowledgments

This study was supported by funding from Consejería de Innovación, Ciencia y Empresa, Secretaría General de Universidades, Investigación y Tecnología, Junta de Andalucía, Spain (Incentivos a Proyectos de Excelencia P07-AGR-2471 and P12-AGR-1473) and, cofinanced by FEDER funds (EU). This work was developed by E. Arrebola hired in the CSIC by the program mode JAEDoc

"Junta para la Ampliación de Estudios" and cofinanced by ESF. Thank you to Dr. Andrew J. Spiers from Abertay University, Dundee (UK) for providing the plasmids pVSP61 and pVSP61-WspR19 for cellulose overproduction. Thank you to Dr. Cayo J. Ramos for his helpful advice and suggestions. Thank you to Manuela Vega Sánchez from SCAI of the Universidad de Málaga, for her assistance in image analysis. We also thank the anonymous reviewers for their critical reading and editing, which has significantly improved this manuscript.

References

Arrebola E, Cazorla FM, Codina JC, Gutiérrez-Barranquero JA, Pérez-García A & de Vicente A (2009) Contribution of mangotoxin to the virulence and epiphytic fitness of *Pseudomonas syringae* pv. *syringae*. *Int Microbiol* **12:** 87-95.

Arrebola E, Cazorlla FM, Durán VE, Rivera E, Olea F, Codina JC, Pérez-García A & de Vicente A (2003) Mangotoxin: a novel antimetabolite toxin produced by *Pseudomonas syringae* inhibiting ornithine/arginine biosynthesis. *Physiol Mol Plant Pathol* **63:** 117-127.

Arrebola E, Cazorla FM, Romero D, Pérez-García A & de Vicente A (2007) A nonribosomal peptide synthetase gene of *Pseudomonas syringae* pv. *syringae* is involved in mangotoxin biosynthesis and is required for full virulence *Mol Plant-Microbe Interact* **20:** 500-509.

Carrión VJ, Arrebola E, Cazorla FM, Murillo J & de Vicente A (2012) The *mbo* operon is specific and essential for biosynthesis of mangotoxin in *Pseudomonas syringae*. *PLoS one* **7:** e36709.

Carrión VJ, Van der Voort M, Arrebola E, Gutiérrez-Barranquero JA, de Vicente A, Raaijmakers JM & Cazorla FM (2014) Mangotoxin production of *Pseudomonas syringae* pv. *syringae* is regulated by MgoA. *BMC Microbiol* **14:** 46.

Castillo T, Ramos LJ, Rodríguez-Herva JJ, Fuhrer T, Sauer U & Duque E (2007) Convergent peripheral pathway catalyze initial glucose catabolism in *Pseudomonas putida*: genomic and flux analysis. *J Bacteriol* **189:** 5142-5152.

Cazorla FM, Torés JA, Olalla L, Pérez-García A, Farré JM & de Vicente A (1998) Bacterial apical necrosis of mango in southern Spain: A disease caused by *Pseudomonas syringae* pv. *syringae*. *Phytopathology* **88:** 614-620.

Chen H, Bjerknes M, Kumar R & Jay E (1994) Determination of the optimal aligned spacing between the Shine-Dalgarno sequence and the translation initiation codon of *Escherichia coli* mRNAs. *Nucleic Acids Res* **22:** 4953-4957.

Coucheron DH (1991) An *Acetobacter xylinum* insertion sequence element associated with inactivation of cellulose production. *J Bacteriol* **173:** 5723-5731.

Delmer DP & Amor Y (1995) Cellulose biosynthesis. *Plant Cell* **7:** 987-1000.

Filiatrault MJ, Stodghill PV, Myers CR, Bronstein PA, Butcher BG, et al. (2011) Genome wide identification of transcriptional start sites in plants pathogen *Pseudomonas syringae* pv. tomato str. DC3000. *PLoS one* **6:** e29335.

Filiatrault MJ, Stodghill PV, Myers CR, Bronstein PA, Moll S, Lindeberg M, et al. (2010) Transcriptome analysis of *Pseudomonas syringae* identifies new

genes, noncoding RNAs, and antisense activity. *J Bacteriol* **192:** 2359-2372.

Fux C, Costerton JW, Steward PS & Stoodley P (2005) Survival strategies of infectious biofilms. *Trends Microbiol* **13:** 34-40.

Gal M, Preston GM, Massey, RC, Spiers AJ & Rainey PB (2003) Genes encoding a cellulosic polymer contribute toward the ecological success of *Pseudomonas fluorescens* SBW25 on plant surface. *Mol Ecol* **12:** 3109-3121.

Giddens SR, Jackson RW, Moon CD, Jacobs MA, Zhang XX, Gehrig SM & Rainey PB (2007) Mutational activation of niche-specific genes provides insight into regulatory networks and bacterial function in a complex environment. *PNAS* **104:** 18247-18252.

Gutiérrez-Barranquero JA, Arrebola E, Bonilla N, Sarmiento D, Cazorla FM & de Vicente A (2011) Environmentally friendly treatment alternatives to Bordeaux mixture for controlling bacterial apical necrosis (BAN) of mango. *Plant Pathol* **61:** 665-676.

Gutiérrez-Barranquero JA, Carrión VJ, Murillo J, Arrebola E, Arnold DL, Cazorla FM & de Vicente A (2013) A *Pseudomonas syringae* diversity survey reveals a differentiated phylotype of the pathovar *syringae* associated with the mango host and mangotoxin production. *Phytopathology* **103:** 1115-1129.

Hirano SS & Upper CD (1990). Population biology and epidemiology of *Pseudomonas syringae*. *Annu Rev Phytopathol* **28:** 155-177.

Hirano SS & Upper CD (2000) Bacteria in the leaf ecosystem with emphasis on *Pseudomonas syringae*, a pathogen, ice nucleus and epiphyte. *Microbiol Mol Biol Rev* **64:** 624-653.

Jahn CE, Selimi DA, Barak JD & Charkowski AO (2011) The *Dickeya dadantii* biofilm matrix consists of cellulose nanofibres, and is an emergent property dependent upon the type iii secretion system and the cellulose synthesis operon. Microbiology **157:** 2733-2744.

Kennelly MM, Cazorla FM, de Vicente A, Ramos C & Sundin GW (2007) *Pseudomonas syringae* diseases in fruit trees. Progress toward understanding and control. *Plant Dis* **91:** 4-17.

King EO, Ward M & Raney DE (1954) Two simple media for the demonstration of pyocyanin and fluorescein. *J Lab Clin Med* **44:** 301-307.

Larkin MA, Blackshields G, Brown NP, et al. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* **23:** 2947-2948.

Le Quéré B & Ghigo JM (2009) BcsQ is an essential component of the *Escherichia coli* cellulose biosynthesis apparatus that localizes at the bacterial cell pole. *Mol Microbiol* **72:** 724-740.

Lu Y, Chen S, Dong H, Sun H, Peng D & Liu X (2012) Identification of genes responsible for biofilm formation or virulence in *Salmonella enterica* serovar *pullorum*. *Avian Dis* **56:** 134-143.

Maruyama IN, Rakow TL & Maruyama HI (1995) cRNA: a simple method for identification of the 5' end of mRNA. *Nucleic Acids Res* **23:** 3796-3797.

Matthysse AG, Holmes KV & Gurlitz RG (1981) Elaboration of cellulose fibrils by *Agrobacterium tumefaciens* during attachment to carrot cells. *J Bacteriol* **145:** 583-595.

Nei M & Kumar S (2000) Molecular evolution and phylogenetics. Oxford University Press, New York (333 pp).

Nielsen L, Li X & Halverson LJ (2011) Cell-cell and cell-surface interactions mediated by cellulose and a novel exopolysaccharide contribute to *Pseudomonas putida* biofilm formation and fitness under water limiting conditions. *Environ Microbiol* **13:** 1342-1356.

O'Toole G, Kaplan HB & Kolter R (2000) Biofilm formation as microbial development. *Annu Rev Microbiol* **54:** 49-79.

Peeters E, Nelis, HJ & Coenye T (2008) Comparison of multiple methods for quantification of microbial biofilms grown in microtiter plates. *J Microbiol Meth* **72:** 157-165.

Pringent-Combaret C, Zghidi-Abouzi O, Effantin G, Lejeune P, Reverchon S & Nasser W (2012) The nucleoid-associated protein Fis directly modulates the synthesis of cellulose, an essential component of pellicles-biofilm in the phytopathogenic bacterium *Dickeya dadantii*. *Mol Micro* **86:** 172-186.

Robertson M, Hapca SM, Moshynets O & Spiers AJ (2013) Air-liquid interface biofilm formation by psychrotrophic pseudomonads recovered from spoilt meat. *A Van Lee J Microb* **103:** 251-259.

Römling U, Galperin MY & Gomelsky M (2013) Cyclic di-GMP: the first 25 years of a universal bacterial second messenger. *Microbiol Mol Biol Rev* **77:** 1-52.

Smits TH, Rezzonico F, Kamber T, Blom J, Goesmann A, Frey JE & Duffy B (2010) Complete genome of the fire blight pathogen *Erwinia amylovora* CFBP1430 and comparison to other *Erwinia spp*. *Mol Plant-Microbe Interact* **23:** 384-393.

Sofia HJ, Burland V, Daniels DL, Plunket G, III & Blattner FR (1994). Analysis of *Escherichia coli* genome. V. DNA sequence of the region from 76.0 to 81.5 minutes. *Nucleic Acids Res* **22:** 2576-2586.

Spiers, AJ (2014) A mechanistic explanation linking adaptive mutation, niche, change, and fitness advantage for the wrinkly spreader. *Int J Evol Biol* http://dx.doi.org/10.1155/2014/675432.

Spiers AJ, Bohannon J, Gehrig SM & Rainer PB (2003) Biofilm formation at the air-liquid interface by the *Pseudomonas fluorescens* SBW25 wrinkly spreader requires an acetylated form of cellulose. *Mol Microbiol* **50:** 15-27.

Spiers AJ, Kahn SG, Bohannon J, Travisano M & Rainey, PB (2002) Adaptative divergence in experimental populations of *Pseudomonas fluorescens*. I Genetic and phenotypic bases of wrinkly fitness. *Genetics* **161:** 33-46.

Spiers AJ, Deeni YY, Folorunso AO, Koza A, Moshynets O & Zawadzki K (2013) Cellulose expression in *Pseudomonas fluorescens* SBW25 and other environmental pseudomonads. *Cellulose-Medical, Pharmaceutical and Electronic Applications* (Theo van Ven & Louis Godbout, eds), pp. 1-26. InTech Europe, Rijeka, Croatia.

Spiers AJ & Rainey PB (2005) The Pseudomonas fluorescens SBW25 wrinkly spreader biofilm requires attachment factor, cellulose fibre and LPS interactions to maintain strength and integrity. *Microbiol* **151:** 2829-2839.

Tamura K, Dudley J, Nei M & Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* **24:** 1596-1599.

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. *Mol Biol Evol* **28:** 2731-2739.

Ude S, Arnold DA, Moon CD, Timms-Wilson T & Spiers AJ (2006) Biofilm formation and cellulose expression among diverse environmental *Pseudomonas* isolates. *Environ Microbiol* **8:** 1997-2011.

Weiss-Muszkat M, Shakh D, Zhou Y, Pinto R, Belausoz E, Chapman MR & Sela S (2010) Biofilm formation by and multicellular behavior of *Escherichia coli* O55:H7, an atypical enteropathogenic strain. *Appl Environ Microbiol* **76:** 1545-1554.

Wong HC, Fear AL, Calhoon RD, Eichinger GH, Mayer R, et al. (1990) Genetic organisation of the cellulose synthase operon in *Acetobacter xylinum*. *Proc Natl Acad Sci USA* **87:** 8130-8134.

Yildiz FH (2007) Processes controlling the transmission of bacterial pathogens in the environment. *Res Microbiol* **158:** 195-202.

Yu X, Lund SP, Scott RA, Greenwald JW, Records AH, et al. (2013) Transcriptional responses of *Pseudomonas syringae* to growth in epiphytic versus apoplastic leaf sites. *PNAS* **14:** E425-E434.

Zogaj X, Brokranz W, Nimtz M & Römling U (2003) Production of cellulose and curli fimbriae by members of the family *Enterobacteriaceae* isolated from the human gastrointestinal tract. *Infect Immun* **71:** 4151-4158.

Zogaj X, Nimtz M, Rohde M, Bokranz W & Römling U (2001) The multicellular morphotypes of *Salmonella typhimurium* and *Escherichia coli* procedure cellulose as the second component of the extracellular matrix. *Mol Microbiol* **39:** 1452-1463.

Figure Legends

Fig. 1. Genomic organisation of bacterial cellulose biosynthesis gene clusters. The cellulose biosynthetic operons of *Gluconacetobacter xylinus* strain (I) (*bcs*), *Escherichia coli* strain K-12, *Salmonella typhimurium* strain LT2 (*yhj*), *Erwinia amylovora* strain ATCC49946 (putative operon *bcs*), *Pseudomonas fluorescens* strain SBW25, *Pseudomonas syringae* pv. *tomato* strain DC3000 (*wss*), and *Pseudomonas syringae* pv. *syringae* strain UMAF0158 are schematically shown. Each operon contains three genes related to the essential subunits of cellulose synthase, represented as black arrows. The figure has been adapted from Spiers *et al.*, (2002).

Fig. 2. *Pseudomonas* spp. strain phylogenetic analysis. *Escherichia coli* K-12, *Salmonella typhimurium* LT2 and *Erwinia amylovora* 273 have been used as out groups. The evolutionary history was inferred using the maximum likelihood method based on a previously described data-specific model (Nei & Kumar, 2000). (a) Phylogenetic tree using three concatenated genes for cellulose synthase subunits (*wssB*, *wssC* and *wssE*). (b) Phylogenetic tree using eight concatenated housekeeping genes *fruK*, *gapA*, *gltA*, *pgi*, *recA*, *rpoA*, *rpoB* and *rpoD*. The strains with the conserved cellulose operon are marked in a grey background. The *P. syringae* Cit7 strain (marked with asterisk) was not included in the cellulose phylogenetic tree because a contiguous and complete *wss* operon was not found within this draft genome, despite evidence of its presence.

Fig. 3. The cellulose operon of *Pseudomonas syringae* pv. *syringae* strain UMAF0158. (a) Organisation of the nine genes (grey arrows) that constitute the *wss* operon in UMAF0158: *wss*ABCDEFGHI. Mutations in *wss*B and *wss*E are

represented with a black triangle (▼). The promoter (P) and terminator (T) locations are indicated in the non-coding region. Double-headed arrows show the overlapping amplifications along the cellulose operon. Agarose gels displaying the PCR products obtained from genomic DNA or mRNA (RT-PCR) of wild-type UMAF0158. (b) The nucleotide sequence (246 bp) located upstream of *wssA* was analysed, showing a putative promoter (*in silico* -35 and -10 boxes indicated with solid-lined boxes) and a putative transcriptional element for *rpo*D16 (indicated with a dashed-line box). The transcript initiation site is indicated as nucleotide +1 with a black point under the nucleotide, and the Shine-Dalgarno (SD) sequence is indicated in bold letters. The insertion sequences (IS) analysed *in silico* are indicated as bold red letters (Table S3). (c) The terminal region of the cellulose operon located downstream of *wss*I, depicting the secondary structure and its location of the putative Rho-independent terminator sequence. The insertion sequences (IS) are indicated as bold red letters (Table S3). Diagram of the experiment designed to confirm the functionality of the cellulose operon terminator: the amplicon sizes and primer directions are indicated. Agarose electrophoresis gels show the results of the RT-PCR experiments. Hyperladder I (Bioline) was used in these experiments.

Fig. 4. Cellulose production and biofilm formation by *Pseudomonas syringae* UMAF0158, defective mutants and derivative overproducer-strain. (a) Determination of cellulose production measured as fluorescence intensity through calcofluor staining. The fluorescence intensity is given as arbitrary units (a.u.), defined as the value obtained from the image analysis according to a grey scale ranging from pure black to pure white. (b) Biofilm formation was

estimated as the absorbance at 595 nm after 15 min of crystal violet staining. Both experiments were performed using *P. syringae* pv. *syringae* UMAF0158 containing the vector pVSP61 and *P. syringae* pv. *syringae* B728 and *P. fluorescens* SBW25 as controls for cellulose production. Different letters represent significant differences, p=0.05. Error bars show standard deviation.

Fig. 5. Scanning electronic microscopy of mango buds treated with different bacterial suspensions. (a) Non-inoculated mango bud overview, (b) Microcolonies (white arrow) of *P. syringae* pv. *syringae* UMAF0158, (c-d) Cellulose defective mutants UMAF0158::*wssB* and UMAF0158::*wssE*, (e-f) microcolonies (white arrow) of cellulose-overproducing bacteria (UMAF0158 + pVSP61-WspR19).

Fig. 6. Bacterial cell counts obtained from adhesion experiments on mango leaves. Drops of bacterial suspension were deposited onto mango leaves, and after incubation for 30 min, the leaves were softly washed, and the adhered cells were recovered and counted. In this experiment, wild type *Pseudomonas syringae* pv. *syringae* UMAF0158, cellulose-defective mutants UMAF0158::*wss*B and UMAF0158::*wss*E, and cellulose-overproducing strain UMAF0158 + pVSP61-WspR19 were assayed. Different letters represent significant differences, p=0.05. Error bars show standard deviation.

Fig. 7. Analysis of cellulose as a putative virulence factor of *Pseudomonas syringae* pv. *syringae* UMAF0158. (a) Growth time course of wild type *P. syringae* pv. *syringae* UMAF0158 (●), the cellulose-defective mutants UMAF0158::*wss*B (○) and UMAF0158::*wss*E (▼) and the cellulose-overproducing strain UMAF0158+pVSP61-WspR19 (Δ) inoculated onto tomato

leaflets after piercing and maintained *in vitro* for ten days at 22ºC under a 16-h photoperiod. (b) The development of necrotic symptoms in tomato leaflets inoculated with the four strains. The incidence of necrotic symptoms is represented as an accumulative number of inoculated points developing necrotic areas higher than 5 mm in diameter. The symptoms were monitored and counted for 0, 3, 6 and 10 days, obtaining a total of 108 inoculated points from each strain. Significant differences from wild-type UMAF0158 are indicated with an asterisk (*).Error bars show standard deviation. (c) Representative pictures of tomato leaflets maintained *in vitro* and inoculated with the four strains, showing symptoms developed at 10 days post-inoculation. (d) Severity of necrotic symptoms: the total necrotic area (mm^2) per leaflet with 6 inoculated points in 3 independent experiments was measured 10 days after inoculation. Six inoculated leaflets in each independent experiment were analysed, and 3 independent experiments were performed. Different letters represent significant differences p=0.05 according to analysis of variance. Error bars show standard deviation.

Fig. 8. The competitive index (CI) score for *Pseudomonas syringae* pv. *syringae* UMAF0158::*wssB*, UMAF0158::*wssE* and UMAF0158 + pVSP61-WspR19 relative to the co-inoculated wild type *P. syringae* pv. *syringae* UMAF0158. Three leaflets per strain were analysed, and three independent experiments were performed to obtain the competitive index values. A CI score of 1 denotes no difference compared with the wild type. Additional analysis of the average data from three independent experiments was performed using analysis of variance (p=0.05). Different letters indicate statistical significance. Error bars show standard deviation.

100

Table 1. Genes comprising the cellulose biosynthetic operon located in *Pseudomonas syringae* pv. *syringae* UMAF0158

Gene Id	Name[a]	Sequence size bp	GC %	Predicted Function	Sequence size aa	Conserved Domains[b]
PsyrMG_20805	*wssA*	1140	62.2	Cellulose synthase associated positioning subunit	379	ParA, MinD, YhjQ: cellulose synthase operon protein.
PsyrMG_20810	*wssB*	2220	61.2	Cellulose synthase catalytic subunit and c-di-GMP binding protein	739	CESA_CelA_like, PilZ domain, YhhN super family, Glyco_transf_GTA_type super family. Cellulose_synthase_UDP-forming.
PsyrMG_20815	*wssC*	2256	61.7	Cellulose synthase regulator protein	751	BcsB superfamily, possible c-di-GMP binding site
PsyrMG_20820	*wssD*	1212	64.7	Endo-1,4-D-glucanase	403	Glyco_hydro_8: Glycosyl hydrolase family 8
PsyrMG_20825	*wssE*	3876	64.4	Cellulose synthase subunit BcsC	1291	Cellulose synthase subunit C BcsC TPR superfamily (protein interaction)
PsyrMG_20830	*wssF*	666	61.3	Cellulose synthase-associated acetylation subunit	221	SGNH_hydrolase or GDSL_hydrolase superfamily
PsyrMG_20835	*wssG*	675	64.9	Cellulose synthase-associated acetylation subunit	224	AlgF_like, Cas6_I-E superfamily
PsyrMG_20840	*wssH*	1416	63.0	Cellulose synthase-associated acetylation subunit	471	MBOAT superfamily
PsyrMG_20845	*wssI*	1107	61.6	Cellulose synthase-associated acetylation subunit	368	AlgX_N_like

[a] Nomenclature of orthologous genes in *Pseudomonas fluorescens* SBW25
[b] Domains found at nucleotide and amino acids sequence by searching in Pfam (http://pfam.sanger.ac.uk) and NCBI (http://www.ncbi.nlm.nih.gov) data bases

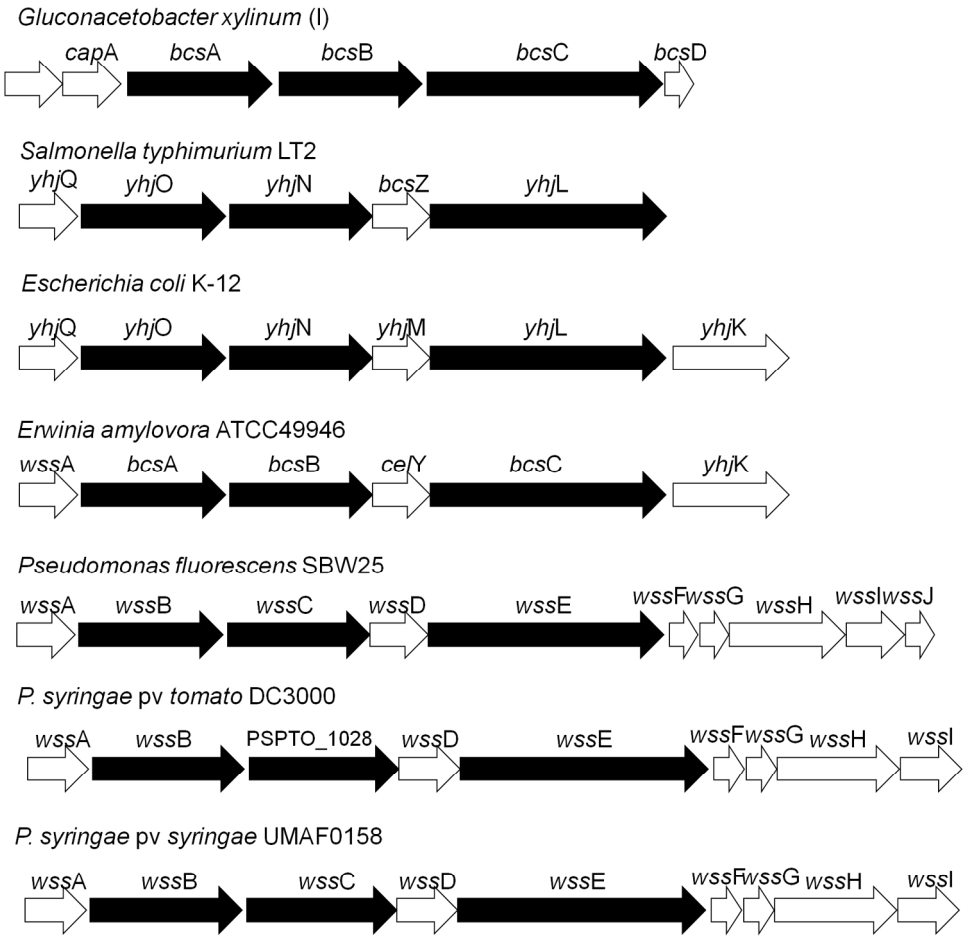


Fig. 1. Genomic organisation of bacterial cellulose biosynthesis gene clusters. The cellulose biosynthetic operons of Gluconacetobacter xylinus strain (I) (bcs), Escherichia coli strain K-12, Salmonella typhimurium strain LT2 (yhj), Erwinia amylovora strain ATCC49946 (putative operon bcs), Pseudomonas fluorescens strain SBW25, Pseudomonas syringae pv. tomato strain DC3000 (wss), and Pseudomonas syringae pv. syringae strain UMAF0158 are schematically shown. Each operon contains three genes related to the essential subunits of cellulose synthase, represented as black arrows. The figure has been adapted from Spiers et al., (2002).
897x904mm (55 x 55 DPI)

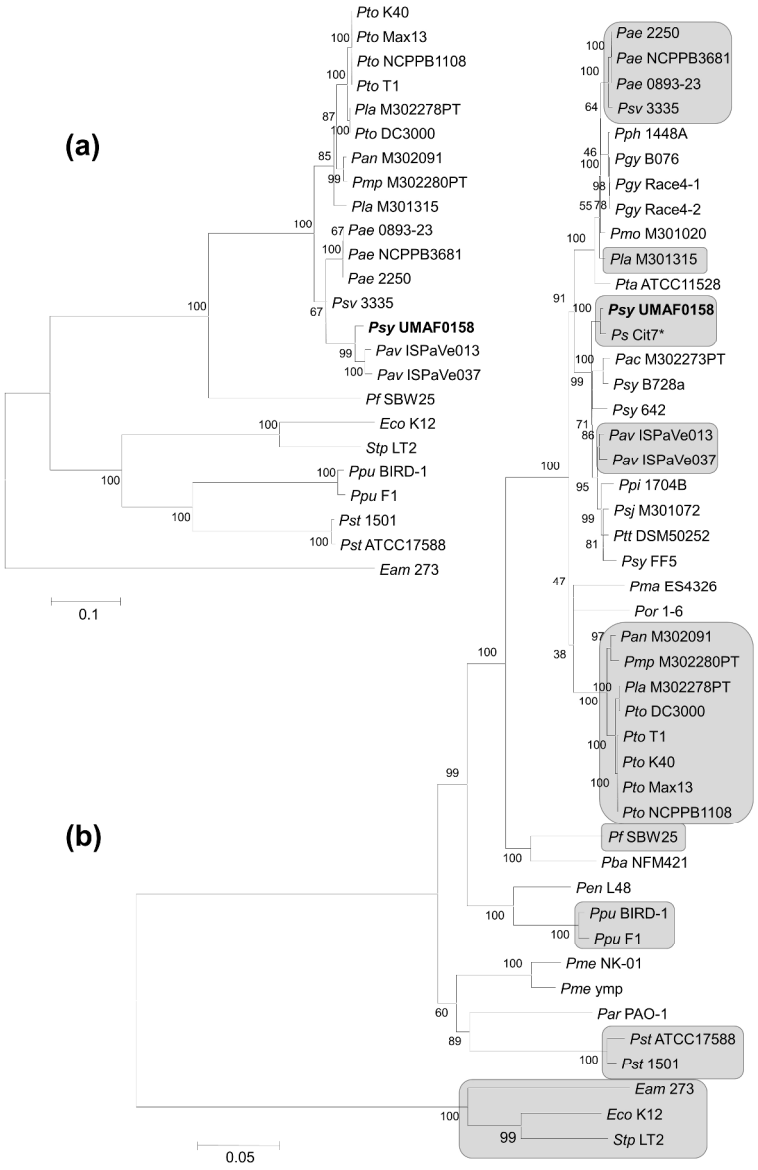


Fig. 2. Pseudomonas spp. strain phylogenetic analysis. Escherichia coli K-12, Salmonella typhimurium LT2 and Erwinia amylovora 273 have been used as out groups. The evolutionary history was inferred using the maximum likelihood method based on a previously described data-specific model (Nei & Kumar, 2000). (a) Phylogenetic tree using three concatenated genes for cellulose synthase subunits (wssB, wssC and wssE). (b) Phylogenetic tree using eight concatenated housekeeping genes fruK, gapA, gltA, pgi, recA, rpoA, rpoB and rpoD. The strains with the conserved cellulose operon are marked in a grey background. The P. syringae Cit7 strain (marked with asterisk) was not included in the cellulose phylogenetic tree because a contiguous and complete wss operon was not found within this draft genome, despite evidence of its presence.
898x1273mm (96 x 96 DPI)

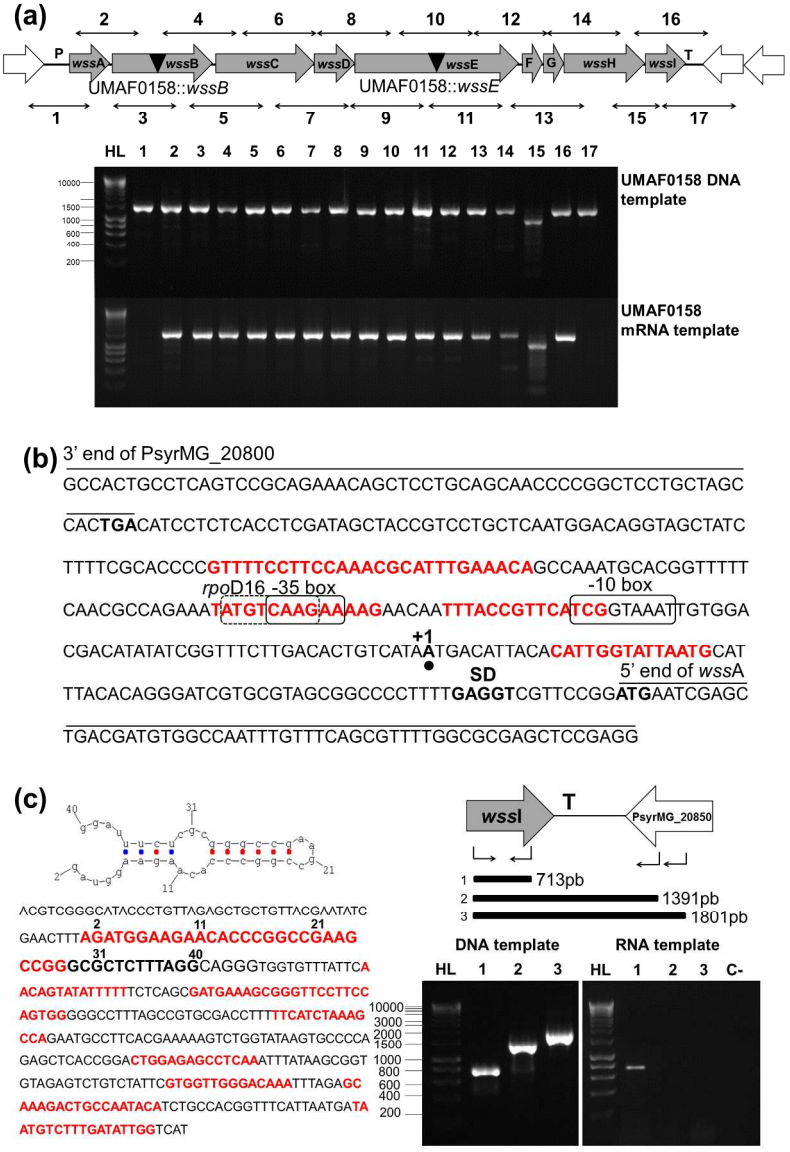


Fig. 3. The cellulose operon of Pseudomonas syringae pv. syringae strain UMAF0158. (a) Organisation of the nine genes (grey arrows) that constitute the wss operon in UMAF0158: wssABCDEFGHI. Mutations in wssB and wssE are represented with a black triangle (▼). The promoter (P) and terminator (T) locations are indicated in the non-coding region. Double-headed arrows show the overlapping amplifications along the cellulose operon. Agarose gels displaying the PCR products obtained from genomic DNA or mRNA (RT-PCR) of wild-type UMAF0158. (b) The nucleotide sequence (246 bp) located upstream of wssA was analysed, showing a putative promoter (in silico -35 and -10 boxes indicated with solid-lined boxes) and a putative transcriptional element for rpoD16 (indicated with a dashed-line box). The transcript initiation site is indicated as nucleotide +1 with a black point under the nucleotide, and the Shine-Dalgarno (SD) sequence is indicated in bold letters. The insertion sequences (IS) analysed in silico are indicated as bold red letters (Table S3). (c) The terminal region of the cellulose operon located downstream of wssI, depicting the secondary structure and its location of the putative Rho-independent terminator sequence. The insertion sequences (IS) are indicated as bold red letters (Table S3). Diagram of the experiment designed to confirm

the functionality of the cellulose operon terminator: the amplicon sizes and primer directions are indicated. Agarose electrophoresis gels show the results of the RT-PCR experiments. Hyperladder I (Bioline) was used in these experiments.
888x1328mm (55 x 55 DPI)

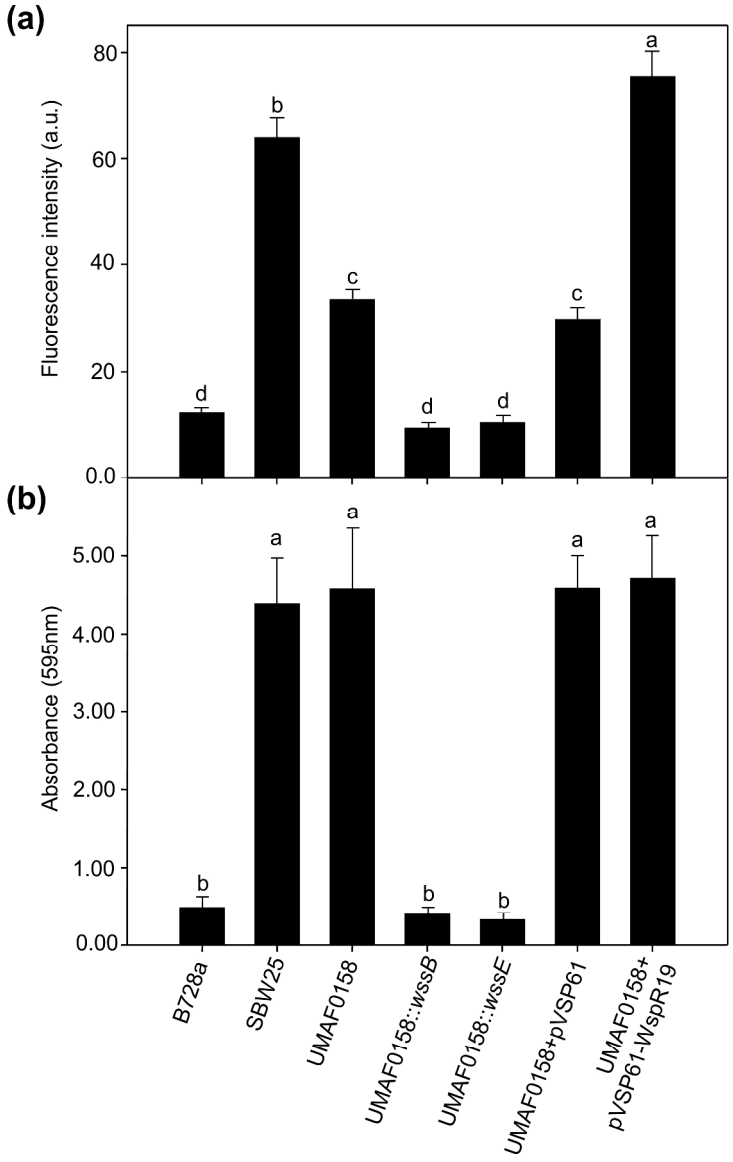


Fig. 4. Cellulose production and biofilm formation by Pseudomonas syringae UMAF0158, defective mutants and derivative overproducer-strain. (a) Determination of cellulose production measured as fluorescence intensity through calcofluor staining. The fluorescence intensity is given as arbitrary units (a.u.), defined as the value obtained from the image analysis according to a grey scale ranging from pure black to pure white. (b) Biofilm formation was estimated as the absorbance at 595 nm after 15 min of crystal violet staining. Both experiments were performed using P. syringae pv. syringae UMAF0158 containing the vector pVSP61 and P. syringae pv. syringae B728 and P. fluorescens SBW25 as controls for cellulose production. Different letters represent significant differences, p=0.05. Error bars show standard deviation.
737x1148mm (96 x 96 DPI)

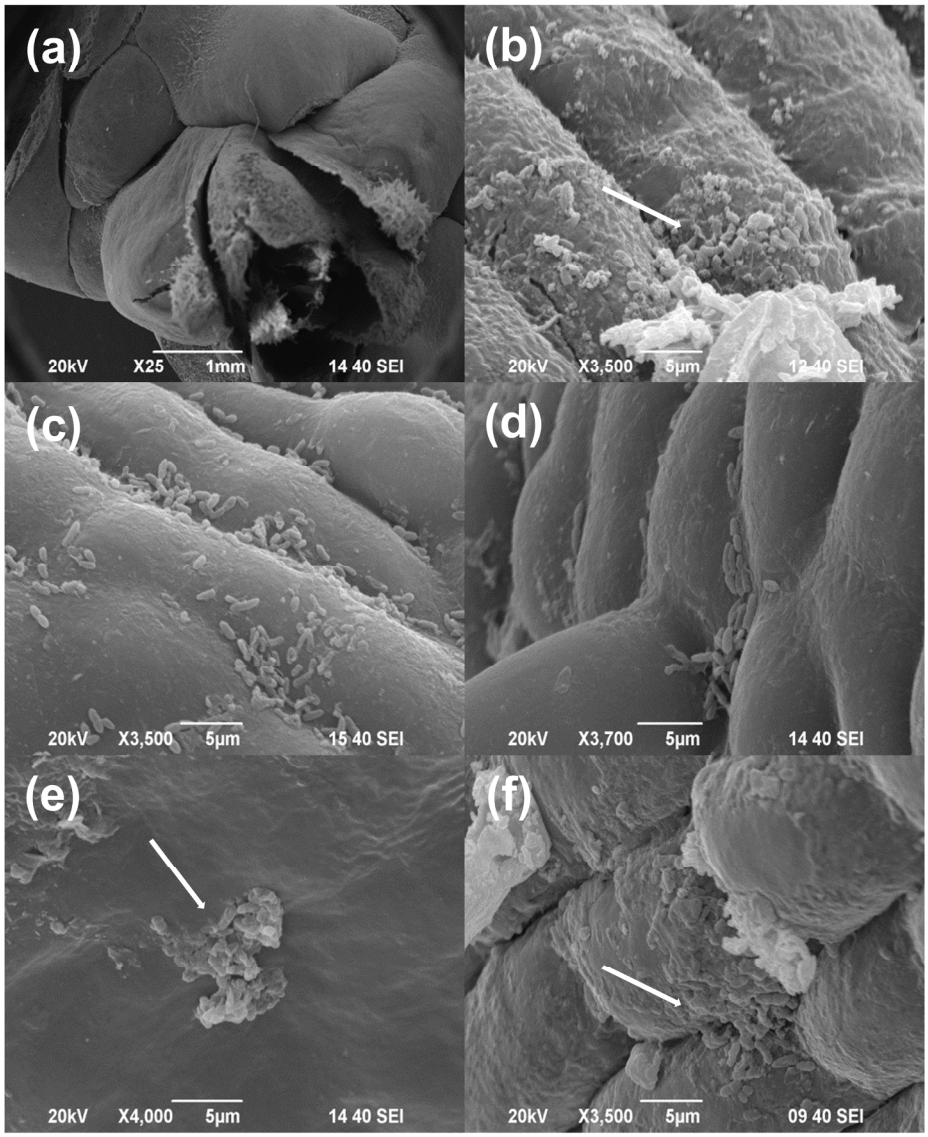


Fig. 5. Scanning electronic microscopy of mango buds treated with different bacterial suspensions. (a) Non-inoculated mango bud overview, (b) Microcolonies (white arrow) of P. syringae pv. syringae UMAF0158, (c-d) Cellulose defective mutants UMAF0158::wssB and UMAF0158::wssE, (e-f) microcolonies (white arrow) of cellulose-overproducing bacteria (UMAF0158 + pVSP61-WspR19).
651x788mm (96 x 96 DPI)

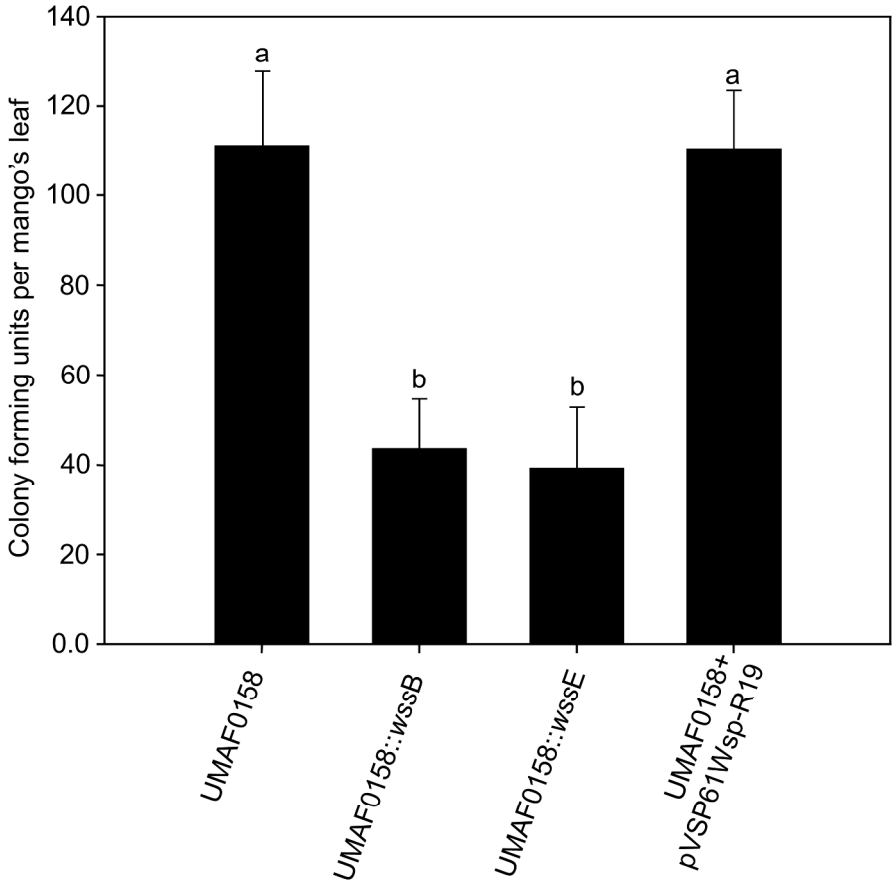


Fig. 6. Bacterial cell counts obtained from adhesion experiments on mango leaves. Drops of bacterial suspension were deposited onto mango leaves, and after incubation for 30 min, the leaves were softly washed, and the adhered cells were recovered and counted. In this experiment, wild type Pseudomonas syringae pv. syringae UMAF0158, cellulose-defective mutants UMAF0158::wssB and UMAF0158::wssE, and cellulose-overproducing strain UMAF0158 + pVSP61-WspR19 were assayed. Different letters represent significant differences, p=0.05. Error bars show standard deviation.
773x740mm (96 x 96 DPI)

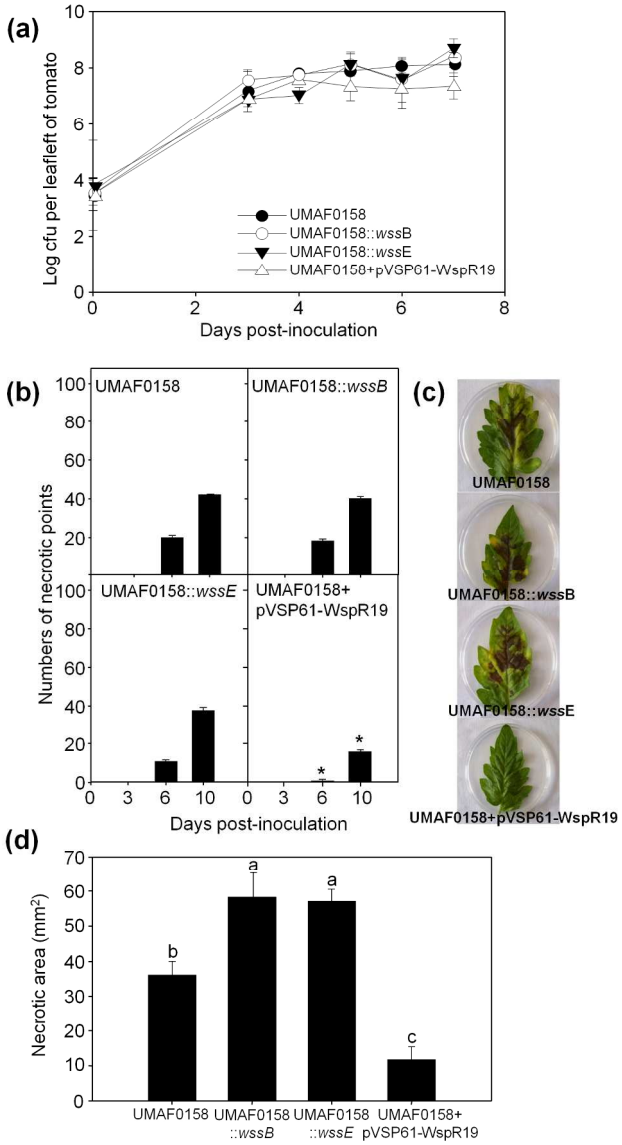


Fig. 7. Analysis of cellulose as a putative virulence factor of Pseudomonas syringae pv. syringae UMAF0158. (a) Growth time course of wild type P. syringae pv. syringae UMAF0158 (●), the cellulose-defective mutants UMAF0158::wssB (○) and UMAF0158::wssE (▼) and the cellulose-overproducing strain UMAF0158+pVSP61-WspR19 (Δ) inoculated onto tomato leaflets after piercing and maintained in vitro for ten days at 22ºC under a 16-h photoperiod. (b) The development of necrotic symptoms in tomato leaflets inoculated with the four strains. The incidence of necrotic symptoms is represented as an accumulative number of inoculated points developing necrotic areas higher than 5 mm in diameter. The symptoms were monitored and counted for 0, 3, 6 and 10 days, obtaining a total of 108 inoculated points from each strain. Significant differences from wild-type UMAF0158 are indicated with an asterisk (*).Error bars show standard deviation. (c) Representative pictures of tomato leaflets maintained in vitro and inoculated with the four strains, showing symptoms developed at 10 days post-inoculation. (d) Severity of necrotic symptoms: the total necrotic area (mm2) per leaflet with 6 inoculated points in 3 independent experiments was measured 10 days after inoculation. Six inoculated leaflets in each independent experiment were analysed, and 3 independent

experiments were performed. Different letters represent significant differences p=0.05 according to analysis of variance. Error bars show standard deviation.
780x1374mm (55 x 55 DPI)

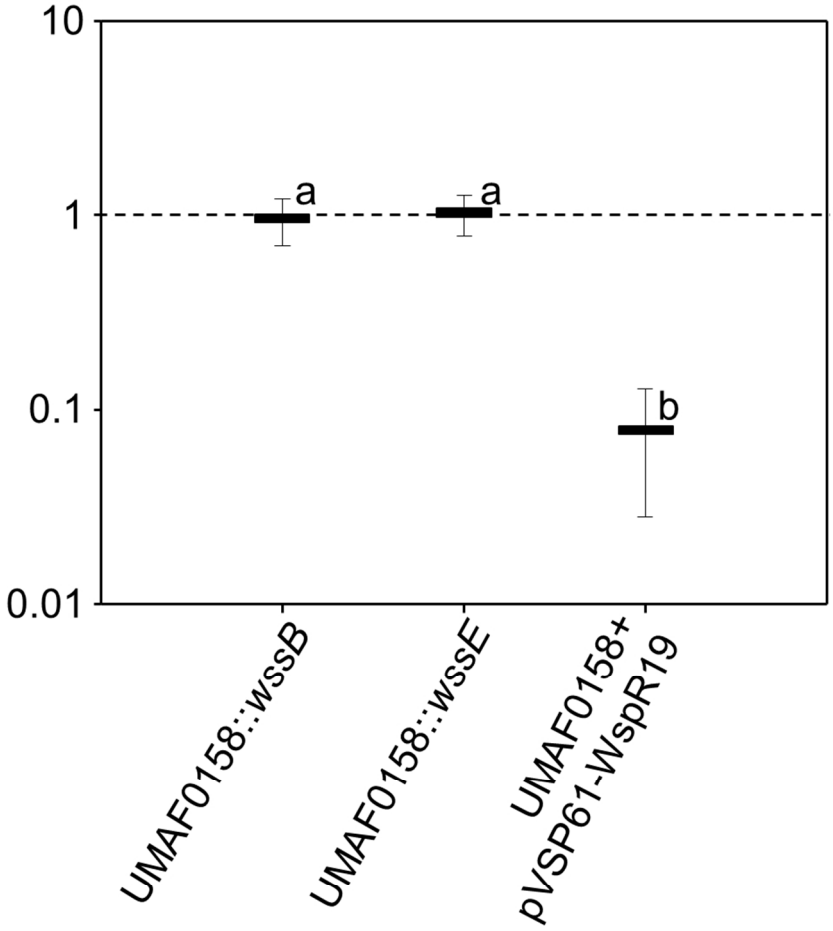


Fig. 8. The competitive index (CI) score for Pseudomonas syringae pv. syringae UMAF0158::wssB, UMAF0158::wssE and UMAF0158 + pVSP61-WspR19 relative to the co-inoculated wild type P. syringae pv. syringae UMAF0158. Three leaflets per strain were analysed, and three independent experiments were performed to obtain the competitive index values. A CI score of 1 denotes no difference compared with the wild type. Additional analysis of the average data from three independent experiments was performed using analysis of variance (p=0.05). Different letters indicate statistical significance. Error bars show standard deviation.
530x554mm (55 x 55 DPI)

Supplemental Materials and Methods

Bacterial strains and growth conditions

For standard maintenance, *P. fluorescens* SBW25 (Spiers *et al.*, 2002), *P. syringae* pv. *syringae* UMAF0158 (Arrebola *et al.,* 2003) and B728a (Hirano & Upper, 2000) were grown in King's medium B (KB) (King et al., 1954) and incubated at 28ºC for 48 h. Defective mutants UMAF0158::*wss*B and UMAF0158::*wss*E were maintained with 50 µg mL^{-1} kanamycin. The conditions used for plant experiments are specified in the corresponding sections.

Phylogenetic Analysis

Phylogenetic analysis of different *Pseudomonas spp*. strains was performed using *E. coli* K-12, *S. typhimurium* LT2 and *E. amylovora* 273 as out groups (Supplementary material, Table S3). The selected sequences were analysed using ContigExpress in Vector NTI Advance 10 software to generalise the sequence size. Thereafter, the group sequences of the genes selected for phylogenetic analysis were joined for multilocus treatment and aligned using ClustalW software (Larkin *et al.*, 2007). The cellulose synthase subunits *wssB*, *wssC* and *wssE* (Spiers *et al.*, 2002) and the housekeeping genes *fruK*, *gapA*, *gltA*, *pgi*, *recA*, *rpoA*, *rpoB* and *rpoD* from the selected strains were used for phylogenetic comparison. Phylogenetic analysis of the tested strains was performed using MEGA 5.0 software (Tamura *et al.*, 2007, 2011). Phylogenetic trees were constructed using maximum likelihood fits based on a data-specific model (Nei & Kumar, 2000) of 24 different nucleotide-substitution models. Evolutionary distances were computed using the best model according to the

previous assessment, including the General Time Reversible method using discrete Gamma distribution (GTR+G) for the cellulose multilocus, and the Tamura-Nei method using discrete Gamma distribution and assuming that a certain fraction of the site was evolutionarily invariable (TN93+G+I) for the housekeeping multilocus. The bootstrap consensus trees were inferred from 1000 replicates, and branches corresponding to partitions reproduced in less than 50% bootstrap replicates were collapsed. The tree with the highest log likelihood (-46080.6247) is shown for the cellulose phylogenetic tree, and (-40441.1534) is shown for the housekeeping phylogenetic tree. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were automatically obtained. When the number of common sites was < 100 or less than one fourth of the total number of sites, the maximum parsimony method was used, otherwise the BIONJ method with the MCL distance matrix was used. Evolutionary analyses were conducted in MEGA5 (Tamura *et al.*, 2007). The cellulose multilocus tree has been drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 24 nucleotide sequences. The 1st+2nd+3rd+Noncoding codon positions were included. All positions containing gaps and missing data were eliminated. There were a total of 5210 positions in the final dataset. A discrete Gamma distribution was used for the housekeeping multilocus to model evolutionary rate differences among sites (5 categories (+G, parameter = 0.6416). The rate variation model allowed for some sites to be evolutionarily invariable ([+I], 28.5912% sites). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 45 nucleotide sequences. The 1st+2nd+3rd+Noncoding codon

positions were included. All positions containing gaps and missing data were eliminated. A total of 5588 positions were included in the final dataset. The *P. syringae* Cit7 strain (marked with an asterisk in Fig. 2b) was not included in the cellulose phylogenetic tree (Fig. 2a) because a contiguous and complete biosynthetic operon was not found within this draft genome, despite evidence of its presence.

Strain manipulation and molecular assays

Insertional inactivation mutagenesis of *P. syringae* UMAF0158 was used to suppress cellulose production by inserting disruption vectors into the different ORFs of the cellulose operon via single-crossover homologous recombination. The PCR using sequence-specific primers was performed to construct integrative plasmid DNA fragments of the two different genes (*wssB* and *wssE*). The PCR reaction, pCR2.1-TOPO® (Invitrogen Life Tech, USA) cloning and plasmid purification were performed using standard procedures and primers are listed in the Supplementary material, Table S5. Plasmids were transformed into the wild-type strain UMAF0158 using electroporation. Both types of mutants (UMAF0158::*wssB* and UMAF0158::*wssE*) were analysed by PCR using in and out primers of the cloned sequence locating the correct insertion of the disruption vectors. Additionally, bacterial growth curves were obtained in KB broth to confirm similarities with the wild type (data not shown).

The wild-type strain UMAF0158 was transformed with the kanamycin-resistant self-replicating plasmid pVSP61-WspR19 to produce an over-producing cellulose strain. UMAF0158 was also transformed with the vector pVSP61 as a control. (Ude *et al.*, 2006).

Total RNA was isolated from a bacterial culture as previously described (Castillo *et al.*, 2007). The RNA concentration was determined using a Nanodrop ND-1000 (NanoDrop Technologies Wilmington, DE) and brought to 50 ng μL^{-1}. The RNA sample integrity was assessed through agarose gel electrophoresis. RT-PCR was performed with 100 ng RNA in a final volume of 50 µL using the Titan OneTube RT-PCR system, according to the manufacturer's instructions (Roche Diagnostics, Basel, Switzerland), with primers designed to include the neighbouring ends of the analysed genes and listed in the Supplementary material, Table S5. Positive control reactions, containing DNA isolated from the corresponding bacteria strain, were included in every assay. The transcription start point for the *wss* operon was determined using the 5'-RACE method (Maruyama *et al.* 1995; Filiatrault *et al.*, 2010, 2011; Carrión *et al.*, 2012). The synthesis of single-stranded cDNA was performed using Total DNA-free RNA, which was obtained from cultures grown in KB broth for 48 h at 28°C. One microgram of this RNA was used as a template to synthesise first-strand cDNA using a cDNA synthesis kit (SMART™ RACE cDNA Amplification Kit, Clontech, Takara Bio Company© 2012, USA) and a gene-specific oligonucleotide primer designed to anneal within the coding region of the gene (Primers listed in Supplementary material, Table S5). The reactions were performed for 90 min at 42°C and subsequently diluted 10-fold in water. A total of 1 µl of these dilutions was added to 20 µl of PCR mixture. The cycling profile included 5 cycles for 30 s at 94°C; 3 min at 72°C; 5 cycles for 30 s at 94°C; 30 s at 70°C; 3 min 72°C; 25 cycles for 30 s at 94°C; 30 s at 68°C; and a final extension for 3 min at 72°C. The amplification products were cloned

into the pGEM®-T Easy Vector (Promega Corporation, Madison, WI) and confirmed through sequencing.

Microscopy

The bacterial cultures were incubated for 48 h at 28ºC in KB medium. Bacterial suspensions were made in sterile distilled water and adjusted to 10^8 cfu mL^{-1}. The suspensions were sprayed onto mango buds (*Mangifera indica* L. var Osteen) and left for three days on plants maintained in a greenhouse. Mango buds were fixed in 2.5% glutaraldehyde in phosphate buffer for 24 h at 4ºC in the dark. Finally, the samples were dehydrated by increasing the ethanol concentration from 0% to 100%, and the samples were kept in each ethanol solution for at least 20 min. The samples were maintained in 100% ethanol until processing for Scanning Electronic Microscopy (SEM). At least three replicates for every sample were observed.

***In vitro* analysis of cellulose production**

Strains were grown in KB medium supplemented with $FeCl_3$ (0.25 g L^{-1}) to avoid siderophore production. Calcofluor (SIGMA Fluorescent Whitener 28) was added to the medium at 10 µM for fluorescent staining of the colonies. The bacteria were grown for 3 days at 28ºC then subsequently observed and photographed using a binocular microscope with a 2x objective, 640x480 focus and 1 s exposure time under UV light. The pictures in Tiff format were processed using Visilog 5.0 software (Noesis Vision Inc., France). The colour images were transformed to grey scale pictures, whereby the fluorescent color is transformed to white color; thus, highly fluorescent colonies will be closer to

white in color than colonies with little fluorescence. Using the image processor, the amount of white in the colonies were measured and transformed to data. The measured area for each colony was 17202 pixels (0.1388 mm^2). At least three colonies per image and six images per strain and experiment were analysed. Five independent experiments were performed to obtain the cellulose production results. The values were obtained as arbitrary units (a.u.), defined as the value obtained according to a grey scale ranging from 0.0 (pure black) to 255 (pure white) in a 0.1388 mm^2 area.

Biofilm quantification

Liquid cultures of UMAF0158, the cellulose-defective mutants UMAF0158::*wssB* and UMAF0158::*wssE*, and the overproducing strain (overproducer type) were adjusted to 10^8 cfu mL^{-1} and incubated at 28ºC overnight to assay biofilm formation. The negative control was cultured using KB broth. Biofilm formation and quantification were performed as previously described (Peeters *et al.*, 2008). Briefly, one hundred microliters of each strain culture were pipetted into several wells of a 96-well microtiter plate and incubated at 28ºC for 24 h to facilitate cell attachment and biofilm formation. The culture was discarded after incubation to remove any loosely associated or planktonic bacteria. For biofilm fixation, 100 μL of methanol was added to each well of a 96-well microtiter plate. After 15 min, the methanol was removed, and the plate was air-dried. The biofilms were quantified using a crystal violet assay. The wells were stained with 100 μL of 0.5% crystal violet and incubated at room temperature for 15 min, followed by washing under running tap water. The bound crystal violet was released after adding 150 μL of 33% of acetic acid

(SIGMA-Aldrich, Bandai, Fukushima, Japan). The 96-well microtiter plates were measured at 595 nm using a micro-plate reader. At least 24 wells per strain and experiment were used, and four independent experiments were performed to obtain the biofilm results.

Adhesion assay on mango leaves

For adhesion experiments, bacterial suspensions using KB plates from strains grown for 48 h at 28ºC were adjusted to 10^8 cfu mL^{-1}. Drops (10 µL) of each strain were inoculated onto mango leaves wiped with 70% ethanol and painted onto dividing lines to avoid mixtures. After 30 min, the leaves were carefully washed with sterile water, cut from the mango tree and dissected according to the painted lines for processing. The leaf pieces were placed into sterile bags with 1 mL of sterile water and homogenised for 3 min. One hundred microliters from every bag was plated onto a KB plates to determine bacterial numbers. Three plates per strain were used, with three replicates per experiment and three independent experiments performed to obtain the adhesion results.

Pathogenesis and competitive index evaluation

Wild-type UMAF0158, the cellulose-defective mutants UMAF0158::*wssB* and UMAF0158::*wssE,* and the overproducer strains were inoculated onto detached tomato leaflets (*Solanum lycopersicum* L. cv. Hellfrucht-Früstamm) to evaluate virulence, as previously described (Arrebola *et al.*, 2007, 2009; Carrión *et al.*, 2014). Bacterial suspensions from exponentially growing cultures were adjusted to 10^8 cfu mL^{-1}. Detached leaflets were inoculated with six 10 µL drops of bacterial suspension at six different points on each leaflet. For inoculation, the

leaves were pierced with bacterial droplets using a sterile entomological pin. The leaflets were maintained in Murashige & Skoog (MS) media at 22 ºC for a 16-h photoperiod. Six tomato leaflets were used for each strain and each independent experiment. Non-infected detached leaflets inoculated with sterile distilled water were included in all experiments as a control. These experiments were repeated three times.

The development of necrotic symptoms at the inoculation points was determined using two different methods. First, the appearance of necrotic symptoms was monitored for 0, 3, 6 and 10 days using visual analysis to evaluate disease incidence (number of inoculated points with symptoms of necrosis), considering points with areas equal or greater than 0.5 cm in diameter to be necrotic. Second, the total necrotic area per leaflet induced by the inoculated strains on the last day (10^{th} day) of the experiment (severity) was determined from six leaflets over three independent experiments using the image analysis software Visilog 5.0 (Noesis Vision Inc., France).

In addition, two inoculated leaflets were used every day for seven days to estimate the total bacterial density in every strain used. The tomato leaflets were homogenised in 10 ml of sterile water, and the bacteria were serially diluted and counted on KB plates after incubation at 28ºC for 48 h.

Additional experiments to obtain a competitive index were performed using tomato leaflets maintained *in vitro*. A bacterial suspension of the wild type UMAF0158 strain was mixed at a 1:1 ratio with cellulose-defective mutants, and the cellulose overproducer strain suspensions was adjusted to 10^8 cfu mL^{-1}. Similar to the pathogenicity, six drops of the mixture were inoculated on the tomato leaflets at six different places and maintained in MS medium at 22ºC

during the experiment. The leaflets were processed after 1 h to observe the initial symptom points and after 7 days to observe the final symptom points. The same suspension from tomato leaflet processing was inoculated in KB medium with and without kanamycin (50μg mL^{-1}). The titers were calculated for every strain at initial and final points. The competitive indices were obtained after dividing the analysed strain by the wild type at each point (initial and final), and dividing again the data obtained from the final point per the data from the initial point in every mutant and overproductive strain. Three leaflets per strain and experiment and three independent experiments were performed to obtain competitive index results.

Statistical Analysis

Statistical analyses of *in vitro* cellulose production, the biofilm quantification adhesion assay on mango leaves, the competitive index and analysis of the necrotic area, were performed using IBM.SSPS 19 software (IBM® Company, Armonk, NY). One factor ANOVA was used for the analysis of the means with p=0.05: homogeneity test data were analysed by the Levene test, and statistic descriptive effects homogeneity was analysed by the Brown-Forsythe Welch test, with post hoc analysis by LSD α= 0.05 and a 95% confidence interval. The statistical analysis of incidence of necrotic symptoms was performed using SAS9.2 software (SAS Institute Inc., Cary, NC, USA) in the Enterprise Guide 4.2 program with generalised linear model analysis.

Table S1. Search of domain patterns in PROSITE data base of the nine genes which constitute the *wss* operon located in *P. syringae* pv. *syringae* UMAF0158

Genes	PROSITE Pattern[a]	Positions in aa the sequence
PsyrMG_20805	**ATP_GTP_A, PATTERN[b].** [AG]-x(4)-G-K-[ST]	129,
	MYRISTYL, PATTERN[c]. G-{EDRKHPFYW}-x(2)-[STAGCN]-{P}	78, 82, 132, 176, 194
	CK2_PHOSPHO_SITE, PATTERN[d]. [ST]-x(2)-[DE]	33, 50, 66
	LEUCINE_ZIPPER, PATTERN. L-x(6)-L-x(6)-L-x(6)-L	242
	PKC_PHOSPHO_SITE, PATTERN[e]. [ST]-x-[RK]	75, 120, 128
PsyrMG_20810	**ASN_GLYCOSYLATION, PATTERN[f].** N-{P}-[ST]-{P}	3, 313, 504
	RGD, PATTERN[g]. R-G-D	219
	MYRISTYL, PATTERN. G-{EDRKHPFYW}-x(2)-[STAGCN]-{P}	27, 28, 121, 333, 401, 405, 627, 641
	AMIDATION, PATTERN[h]. x-G-[RK]-[RK]	199
	CK2_PHOSPHO_SITE, PATTERN. [ST]-x(2)-[DE]	106, 167, 218, 237, 328, 339, 599, 626, 661, 719
	PKC_PHOSPHO_SITE, PATTERN. [ST]-x-[RK]	32, 36, 96, 290, 506, 667, 719, 729
	CAMP_PHOSPHO_SITE, PATTERN[i]. [RK](2)-x-[ST]	325, 731
PsyrMG_20815	**ASN_GLYCOSYLATION, PATTERN.** N-{P}-[ST]-{P}	168, 213, 357, 399, 400, 408, 424
	RGD, PATTERN. R-G-D	686
	MYRISTYL, PATTERN. G-{EDRKHPFYW}-x(2)-[STAGCN]-{P}	31, 41, 63, 82, 220, 227, 383, 433, 660
	TYR_PHOSPHO_SITE, PATTERN. [RK]-x(2,3)-[DE]-x(2,3)-Y	544
	CK2_PHOSPHO_SITE, PATTERN. [ST]-x(2)-[DE]	35, 153, 170, 293, 359, 378, 393, 406, 432, 487, 492, 531, 566, 577, 642, 673
	PKC_PHOSPHO_SITE, PATTERN. [ST]-x-[RK]	2, 293, 341, 378, 675, 739
PsyrMG_20820	**ASN_GLYCOSYLATION, PATTERN.** N-{P}-[ST]-{P}	63, 118, 193
	MYRISTYL, PATTERN. G-{EDRKHPFYW}-x(2)-[STAGCN]-{P}	17, 69, 73, 260,272, 290, 310, 347
	TYR_PHOSPHO_SITE, PATTERN. [RK]-x(2,3)-[DE]-x(2,3)-Y	357
	CK2_PHOSPHO_SITE, PATTERN. [ST]-x(2)-[DE]	31, 65, 120, 126, 128, 149, 255, 300
	PKC_PHOSPHO_SITE, PATTERN. [ST]-x-[RK]	59, 149, 221, 229, 244, 388
	CAMP_PHOSPHO_SITE, PATTERN. [RK](2)-x-[ST]	9
PsyrMG_20825	**ASN_GLYCOSYLATION, PATTERN.** N-{P}-[ST]-{P}	574, 939, 1054, 1231
	RGD, PATTERN. R-G-D	476

	MYRISTYL, PATTERN. G-{EDRKHPFYW}-x(2)-[STAGCN]-{P}	11, 73, 267, 268, 401, 419, 672, 701, 706, 761, 833, 835, 845, 866, 867, 882, 978, 985, 989, 990, 991, 994, 1003, 1005, 1052, 1076, 1080, 1146, 1217, 1222, 1227, 1247, 1251
	TYR_PHOSPHO_SITE, PATTERN. [RK]-x(2,3)-[DE]-x(2,3)-Y	680
	CK2_PHOSPHO_SITE, **PATTERN**. [ST]-x(2)-[DE]	198, 211, 346, 385, 392, 448, 499, 792, 949, 953, 998, 1070, 1092, 1118
	PKC_PHOSPHO_SITE, PATTERN. [ST]-x-[RK]	45, 222, 416, 458, 469, 679, 738, 822, 897, 979, 1000, 1056, 1182
	CAMP_PHOSPHO_SITE, PATTERN. [RK](2)-x-[ST]	4
PsyrMG_20830	**ASN_GLYCOSYLATION, PATTERN.** N-{P}-[ST]-{P}	31
	MYRISTYL, PATTERN. G-{EDRKHPFYW}-x(2)-[STAGCN]-{P}	43, 46, 121, 130, 180, 187, 199
	CK2_PHOSPHO_SITE, PATTERN. [ST]-x(2)-[DE]	19, 48, 169
	PKC_PHOSPHO_SITE, PATTERN. [ST]-x-[RK]	55
PsyrMG_20835	**ASN_GLYCOSYLATION, PATTERN.** N-{P}-[ST]-{P}	184
	MYRISTYL, PATTERN. G-{EDRKHPFYW}-x(2)-[STAGCN]-{P}	82, 129, 142
	CK2_PHOSPHO_SITE, PATTERN. [ST]-x(2)-[DE]	119, 217
	PKC_PHOSPHO_SITE, PATTERN. [ST]-x-[RK]	64, 79, 119, 147
PsyrMG_20840	**ASN_GLYCOSYLATION, PATTERN**. N-{P}-[ST]-{P}	435
	MYRISTYL, PATTERN. G-{EDRKHPFYW}-x(2)-[STAGCN]-{P}	101, 254, 299, 303, 319, 333, 337, 394, 396
	CK2_PHOSPHO_SITE, PATTERN. [ST]-x(2)-[DE]	224, 273
	PKC_PHOSPHO_SITE, PATTERN. [ST]-x-[RK]	201, 307, 404
PsyrMG_20845	**MYRISTYL, PATTERN**. G-{EDRKHPFYW}-x(2)-[STAGCN]-{P}	8, 18, 58, 198, 292, 310, 320
	LEUCINE_ZIPPER, PATTERN. L-x(6)-L-x(6)-L-x(6)-L	19
	PKC_PHOSPHO_SITE, PATTERN. [ST]-x-[RK]	28, 121

a: Search protein patterns on the amino acid sequence template based on PROSITE™ database (http://prosite.expasy.org/)
b: ATP/GTP binding site motif A (P-Loop)
c: N-myristoylation site
d: Casein kinase II phosphorylation site
e: Protein kinase C phosphorylation site
f: N-glycosylation site (glycosylation of asparagine)
g: Cell attachment sequence
h: Amidation site
i: cAMP and cGMP-dependent protein kinase phosphorylation site

Table S2. Sequence coverage (Cov) and identity (Ident) of genes belonging to cellulose operon located in *Pseudomonas syringae* pv.*syringae* UMAF0158 with the orthologous genes located in sequenced *Pseudomonas spp.* strains. The three common genes present in every cellulose operon (Fig. 1), are colored in grey.

Specie	*wssA*		*wssB*		*wssC*		*wssD*		*wssE*		*wssF*		*wssG*		*wssH*		*wssI*	
Strains	**Cov**	**Ident**	**Cov**	**Ident**	**Cov**	**Ident**	**Cov**	**Ident**	**Cov**	**Ident**	**Cov**	**Ident**	**Cov**	**Ident**	**Cov**	**Ident**	**Cov**	**Ident**
Pseudomonas fluorescens																		
SBW25	-	Nss[c]	63	79	49	72	-	Nss	-	Nss	74	75	-	Nss	99	82	-	Nss
Pseudomonas savastanoi																		
NCPPB3335	85	83	100	94	100	98	100	95	93	86	100	89	92	85	97	88	82	87
Pseudomonas syringae																		
pv. *actinidiae*																		
M302091	67	82	100	88	95	88	78	84	100	85	100	89	92	84	100	88	83	91
pv. *aesculi*																		
0893-23	77	82	100	94	100	98	100	95	100	87	100	89	92	84	100	87	81	87
NCPPB3681	85	88	100	94	98	98	61	97	83	85	100	89	92	84	87	88	82	87
2250	77	82	-	Nss	100	98	100	95	99	87	100	89	92	85	97	87	-	Nss
pv *lachrymans*																		
M301315	67	83	100	88	100	88	83	84	100	85	100	89	92	85	100	87	81	87
M302278PT	-	Nss	100	88	100	88	91	86	31	88	100	89	92	84	100	86	76	87
pv. *morsprunorum*																		
M302280PT	67	82	100	88	100	88	96	85	100	85	100	88	92	85	100	88	93	87
pv. *tomato*																		
DC3000	75	83	100	88	98	88	91	86	92	85	100	89	92	84	100	86	77	85
K40	75	83	100	88	98	88	96	86	98	85	100	89	100	85	100	87	92	85
Max13	75	83	100	88	98	88	96	86	98	85	100	89	100	85	100	87	92	85
NCPPB1108	75	83	100	88	98	88	96	86	98	85	100	89	100	85	100	87	92	85
T1	75	83	100	88	97	88	96	86	93	85	100	89	100	85	100	87	92	85

[a] Nss: Not significant identity.

8

Table S3. Bacterial strain abbreviation used in phylogenetic trees. The lifestyle classification, host from which the corresponding strain was isolated, GenBank accession number and the corresponding reference is also indicated.

Bacterial Name	Abbreviation	Classification	Host	GenBank[a]	References[b]
Erwinia amylovora	Eam ATCC49946	Pathogen	Pear/apple	NC_013971.1	Sebaihia *et al.*, 2010
Escherichia coli	Eco K12	Saprophyte	Human	NC_000913.3	Riley *et al.*, 2006
Salmonella typhimurium	Stp LT2	Pathogen	Human	NC_003197.1	McClelland *et al.,* 2001
Pseudomonas sp					
P. aeruginosa	Par PAO-1	Pathogen	Human	NC_002516.2	Stover *et al.,* 2000
P. brassicacearum	Pba NFM421	Epiphyte	Rape plant	NC_015379.1	Ortet *et al.*, 2011
P. entomophila	Pen L48	Pathogen	*Drosophila melanogaster*	NC_008027.1	Vodovar *et al.* 2006
P. mendocina	Pme ymp	Pathogen	Human	NC_009439.1	Copeland *et al.,* 2007
P. putida	Ppu BIRD-1	Saprophyte	Soil	NC_017530.1	Matilla *et al.*, 2011
	Ppu F1	Saprophyte	Soil	NC_009512.1	Copeland *et al.*, 2007
P. stutzeri	Pst ATCC17588	Pathogen	Human	NC_015740.1	Chen *et al.,* 2011
	Pst DSM4166	Epiphyte	Yellow Indian grass	NC_017532.1	Yu *et al.* 2011
P. fluorescens	Pf SBW25	Saprophyte	Soil	NC_012660.1	Silby *et al.*, 2009
P. savastanoi	Psv 3335	Pathogen	Olive	GCA_000164015.1	Rodríguez-Palenzuela *et al.,* 2010
P. syringae					
	Ps Cit7	Epiphyte	Citrus fruit	GCA_000145825.1	Baltrus *et al.,* 2011
pv. aceris	Pac M302273	Pathogen	Maple tree	GCA_000145925.1	Baltrus *et al.,* 2011
pv. actinidiae	Pan M302091	Pathogen	Kiwi	GCA_000145865.1	Baltrus *et al.,* 2011
pv. aesculi	Pae 2250	Pathogen	Horse Chestnut	GCA_000163275.1	Green *et al.* 2010
	Pae NCPPB3681	Pathogen	Horse Chestnut	GCA_000163255.1	Green *et al.* 2010
	Pae 0893-23	Pathogen	Horse Chestnut	GCA_000145685.1	Baltrus *et al.,* 2011
pv. aptata	Ptt DSM50252	Pathogen	Sugar Beet	GCA_000145905.1	Baltrus *et al.,* 2011
pv. glycinea	Pgy B076	Pathogen	Soybean	GCA_000187045.2	Qi *et al.*, 2011
	Pgy Race 4	Pathogen	Soybean	GCA_000143005.1	Baltrus *et al.,* 2011
	Pgy Race 4-2	Pathogen	Soybean	GCA_000187065.2	Qi *et al.*, 2011
pv. japónica	Psj M301072	Pathogen	Barley	GCA_000145785.1	Baltrus *et al.,* 2011
pv. lachrymans	Pla M301315	Pathogen	Cucumber	GCA_000146005.1	Baltrus *et al.,* 2011
	Pla M302278PT	Pathogen	Cucumber	GCA_000145885.1	Baltrus *et al.,* 2011

pv. maculicola	Pma ES4326	Pathogen	Radish	GCA_000145845.1	Baltrus *et al.*, 2011
pv. mori	Pmo 301020	Pathogen	Mulberry	GCA_000145765.1	Baltrus *et al.*, 2011
pv. morsprunorum	Pmp M302280PT	Pathogen	Plum	GCA_000145745.1	Baltrus *et al.*, 2011
pv. oryzae	Por 1-6	Pathogen	Rice	GCA_000156995.1	Baltrus *et al.*, 2011
pv. phaseolicola	Pph 1448A	Pathogen	French Bean	NC_005773.3	Joardar *et al.*, 2005
pv. pisi	Ppi 1704B	Pathogen	Pea	GCA_000145805.1	Baltrus *et al.*, 2011
pv. syringae	Psy642	Epiphyte	unknown	GCA_000177515.1	Clarke *et al.*, 2010
	Psy B728a	Pathogen	Snap Bean	NC_007005.1	Feil *et al.*, 2005
	Psy FF5	Pathogen	Ornamental Pear	GCA_000163315.2	Sohn *et al.*, 2012
	Psy UMAF0158	Pathogen	Mango tree	CP005970	Martínez-García *et al.*, 2013
pv. tabaci	Pta ATCC11528	Pathogen	Tobacco	GCA_000159835.2	Studholme et al., 2009
pv. tomato	Pto DC3000	Pathogen	Tomato	NC_004578.1	Buell *et al.*, 2003
	Pto K40	Pathogen	Tomato	GCA_000177455.1	Vinatzer *et al.*, 2010
	Pto Max13	Pathogen	Tomato	GCA_000177475.1	Vinatzer *et al.*, 2010
	Pto NCPPB1108	Pathogen	Tomato	GCA_000177495.1	Vinatzer *et al.*, 2010
	Pto T1	Pathogen	Tomato	GCA_000172895.1	Almeida *et al.*, 2009

a) GenBank reference sequence accession or assembly accession number, b) publication or GenBank deposit reference corresponding to accession number

References[b]

Almeida NF, Yan S, Lindeberg M, Studholme DJ, Schneider DJ,Condon B, et al. (2009) A draft genome sequence of *Pseudomonas syringae* pv. tomato T1 reveals a type III effector repertoire significantly divergent from that of *Pseudomonas syringae* pv. tomato DC3000. *Mol Plant Microbe Interact* **22:** 52-62.

Baltrus DA, Nishimura MT, Romanchuk A, Chang JH, Mukhtar MS, et al. (2011) Dynamic evolution of pathogenicity revealed by sequencing and comparative genomics of 19 *Pseudomonas syringae* isolates. *PLoS Pathog* **7:** E1002132.

Buell CR, Joardar V, Lindeberg M, Selengut J, Paulsen IT, et al. (2003) The complete genome sequence of the Arabidopsis and tomato pathogen *Pseudomonas syringae* pv. tomato DC3000. *Proc Natl Acad Sci USA* **100:** 10181-10186.

Chen M, Yan Y, Zhang W, Lu W, Wang J, Ping S & Lin M (2011) Complete genome sequence of the type strain *Pseudomonas stutzeri* CGMCC 1.1803. *J Bacteriol* **193:** 6095

Clarke CR, Cai R, Studholme DJ, Guttman DS & Vinatzer BA (2010) *Pseudomonas syringae* strains naturally lacking the classical *P. syringae* hrp/hrc locus are common leaf colonizers equipped with an atypical type III secretion system. *Mol Plant Microbe Interact* **23:** 198-210.

Copeland A, Lucas S, Lapidus A, Barry K, Glavina del Rio T, et al. (2007) Complete sequence of *Pseudomonas mendocina* ymp (Direct Submission). Submitted (17-APR-2007) US DOE Joint Genome Institute, 2800. Mitchell Drive B100, Walnut Creek, CA 94598-1698, USA.

Copeland A, Lucas S, Lapidus A, Barry K, Detter JC, et al. (2007) Complete sequence of *Pseudomonas putida* F1 (Direct Submission). Submitted (04-JUN-2007) National Center for Biotechnology Information, NIH, Bethesda, MD 20894, USA.

Feil H, Feil WS, Chain P, Larimer F, DiBartolo G, et al. (2005) Comparison of the complete genome sequences of *Pseudomonas syringae* pv. syringae B728a and pv. tomato DC3000. *Proc Natl Acad Sci USA* **102:** 11064-11069.

Green S, Studholme DJ, Laue BE, Dorati F, Lovell H, et al. (2010) Comparative genome analysis provides insights into the evolution and adaptation of *Pseudomonas syringae* pv. aesculi on *Aesculus hippocastanum*. *PLoS ONE* **5:** E10224.

Joardar V, Lindeberg M, Jackson RW, Selengut J, Dodson R, et al. (2005) Whole-genome sequence analysis of *Pseudomonas syringae* pv. phaseolicola 1448A reveals divergence among pathovars in genes involved in virulence and transposition. *J Bacteriol* **187:** 6488-6498.

Martinez PM, Arrebola E, Carrion VJ, Perez-Garcia A, Cazorla FM, et al. (2013) *Pseudomonas syringae* strain UMAF0158 (Direct Submission). Submitted (23-MAY-2013) Departamento de Microbiologia, Universidad de Malaga, Facultad de Ciencias. Campus de Teatinos s/n, Malaga, Malaga 29071, Spain.

Matilla MA, Pizarro-Tobias P, Roca A, Fernandez M, Duque E, et al. (2011) Complete genome of the plant growth-promoting rhizobacterium *Pseudomonas putida* BIRD-1. *J Bacteriol* **193:** 1290.

McClelland M, Sanderson KE, Spieth J, Clifton SW, Latreille P, et al. (2001) Complete genome sequence of *Salmonella enterica* serovar Typhimurium LT2. *Nature* **413:** 852-856.

Ortet P, Barakat M, Lalaouna D, Fochesato S, Barbe V, et al. (2011) Complete genome sequence of a beneficial plant root-associated bacterium, *Pseudomonas brassicacearum*. *J Bacteriol* **193:** 3146.

Qi M, Wang D, Bradley CA & Zhao Y (2011) Genome sequence analyses of *Pseudomonas savastanoi* pv. glycinea and subtractive hybridization-based comparative genomics with nine pseudomonads. *PLoS ONE* **6:** E16451.

Riley M1, Abe T, Arnaud MB, Berlyn MK, Blattner FR, et al. (2006) *Escherichia coli* K-12: a cooperatively developed annotation snapshot—2005. *Nucleic Acids Res* **34:**1-9.

Rodríguez-Palenzuela P, Matas IM, Murillo J, López-Solanilla E, Bardaji L, et al. (2010) Annotation and overview of the *Pseudomonas savastanoi* pv. savastanoi NCPPB 3335 draft geno,e reveals the virulence gene complement of a tumor-inducing pathogenof woody hosts. *Environ Microbiol* **12:** 1604-1620.

Sebaihia M, Bocsanczy AM, Biehl BS, Quail MA, Perna NT, et al. (2010) Complete genome sequence of the plant pathogen *Erwinia amylovora* strain ATCC 49946. *J Bacteriol* **192:** 2020-2021.

Silby MW, Cerdeno-Tarraga AM, Vernikos GS, Giddens SR, Jackson RW, et al. (2009) Genomic and genetic analyses of diversity and plant interactions of Pseudomonas fluorescens. *Genome Biol* **10:** R51.

Sohn KH, Jones J. & Studholme DJ (2012) Draft genome sequence of *Pseudomonas syringae* pathovar syringae strain FF5, causal agent of stem tip dieback disease on ornamental pear.

Stover CK, Pham XQ, Erwin AL, Mizoguchi SD, Warrener P, et al. (2000) Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen. *Nature* **406:** 959-964.

Studholme DJ, Ibanez SG, MacLean D, Dangl JL, Chang JH & Rathjen JP (2009) A draft genome sequence and functional screen reveals the repertoire of type III secreted proteins of *Pseudomonas syringae* pathovar tabaci 11528. *BMC Genom* **10:** 395.

Vinatzer BA, Yan S & Lewis J (2010) Microevolution of *Pseudomonas syringae* pv. tomato (Direct Submission). Submitted (26-JAN-2010) Plant Pathology, Physiology and WeedScience, Virginia Polytechnic Institute and State University, 410 Price Hall, Virginia Tech, Blacksburg, VA 24061, USA

Vodovar N, Vallenet D, Cruveiller S, Rouy Z, Barbe V, et al. (2006) Complete genome sequence of the entomopathogenic and metabolically versatile soil bacterium *Pseudomonas entomophila*. *Nat Biotechnol* **24:** 673-679.

Yu H, Yuan M, Lu W, Yang J, Dai S, et al. (2011) Complete genome sequence of the nitrogen-fixing and rhizosphere-associated bacterium *Pseudomonas stutzeri* strain DSM4166. *J Bacteriol* **193:** 3422-3423.

Table S4. Insertion sequences (IS) found in non-coding DNA region where were located promoter and terminator. The IS were found by bioinformatics analysis using: IS Finder (http://www-is.biotoul.fr).

PROMOTER	IS Sequences	TERMINATOR	IS Sequences
ISBcen13	caaacgcatttgaaa	ISAcl1	gaacacccggccgaagccg
ISCac1	cattggtattaatg	ISAcma15	gatgaaagcgggtt
ISMhu11	gttttccttccaaa	ISCpe3	agatggaagaacacc
ISFnu8	tatgtcaagaaaag	ISBpu1	ttcatctaaagcca
ISBth5	acgcatttgaaaca	ISNpu13	gtggttgggacaaa
ISFnu6	tatgtcaagaaaag	ISRin1	gcaaagactgccaataca
ISPfl1	tttaccgttcatcg	ISAtu6	gttccttccagtgg
IS231L	acgcatttgaaaca	ISStr1	ccggccgaagccgg
		ISSlsp1	taatgtctttgatattgg
		ISBun1	aacagtatattttt
		ISRpa2	agaacacccggccgaagc
		ISAav1	ctggagagcctcaa
		IS1131	gttccttccagtgg

Table S5. DNA primers used in this study.

PRIMERS	SEQUENCE 5'→3'	USE
Construction of insertional mutants using integrative plasmids (pCR2.1-TOPO)		
wssB-F	CCGACCGTCGATGTGTTTAT	Amplify a *wss*B fragment (794bp) that cloned into pCR2.1 TOPO vector was used to construct insertional mutant in *wss*B
wssB-R	AGGCCGTAAAAGAAGTGCAG	
wssE-F	CCTGAAGGATGCGGTAAAAA	Amplify a *wss*B fragment (826bp) that cloned into pCR2.1 TOPO vector was used to construct insertional mutant in *wss*B
wssE-R	CAGGAGTACTTGCGGATCGT	
Allocation *wss* genes in mRNA by RT-PCR (see Fig. 3a)		
Cel-1F	GTGATTGCCTGGCTGACC	Amplify overlapping sequence between AMN-31 and *wss*A (1367bp)
Cel-1R	GAGATCGATCGCCAGTGTCT	
Cel-2F	CACGCGTAACTGATCAATCG	Amplify overlapping sequence between *wss*A and *wss*B (1351bp)
Cel-2R	GTCCAGCCAGTTGTCGAAAT	
Cel-3F	TTGTCATGACCAACCTGTCG	Amplify *wss*B internal sequence (1373bp)
Cel-3R	AAGGTGCGGCAACACATAG	
Cel-4F	GTGGTGTGGCCACTGAAAC	Amplify overlapping sequence between *wss*B and *wss*C (1313bp)
Cel-4R	ACAACGCACAAGCCAACAG	
Cel-5F	GGCACTCGTTTCGATGATCT	Amplify overlapping sequence between *wss*B and *wss*C (1341bp)
Cel-5R	AACCGGAGACGTTGAGTTTG	
Cel-6F	CAACGCACTGGTATTGCTCA	Amplify *wss*C internal sequence (1288bp)
Cel-6R	TACCTTGTCGCTGCTGCTC	
Cel-7F	GACCGGTGTGTCGGTTATTC	Amplify overlapping sequence between *wss*C and *wss*D (1315bp)
Cel-7R	AACCCAATCGGCGATAAAG	
Cel-8F	GACCTGTGGATCGTCTACGC	Amplify overlapping sequence between *wss*D and *wss*E (1359bp)
Cel-8R	AAAACAACGCCAGAATCGAG	
Cel-9F	CGCAGACTTACTTGGCTCGT	Amplify *wss*E internal sequence (1377bp)
Cel-9R	ACAGCAACGCACTGGTGTAG	
Cel-10F	AGGCGGGCTATCGTCAAGT	Amplify *wss*E internal sequence (1309bp)
Cel-10R	GCGAACGGATTGGACGAC	
Cel-11F	GTCCGGACGACATTCAGG	Amplify *wss*E internal sequence (1310bp)
Cel-11R	ACTCTGGTCGTTGGCGTAAC	
Cel-12F	AGAGGGCCTCAAGGCTGATA	Amplify overlapping sequence between *wss*E and *wss*F (1334bp)
Cel-12R	CAGGAACTGCGACATCAACT	
Cel-13F	GATCAGCAAGAAGGTCGATTG	Amplify overlapping sequence between *wss*F, *wss*G and *wss*H (1367bp)
Cel-13R	ACCACGGTCAGCAGCATGT	
Cel-14F	CGACCTGTTCAAGTCGGTTG	Amplify overlapping sequence between *wss*G and *wss*H (1296bp)
Cel-14R	CTTGGAACAGCAGGCAGATT	
Cel-15F	CGTACTCGCTGCAACTGTTC	Amplify overlapping sequence between *wss*H and *wss*I (1340bp)
Cel-15R	TACGCAGGTAAGCTGTTTCG	
Cel-16F	GGCGAAGTGACTCACCAACT	Amplify overlapping sequence between *wss*H, *wss*I and 3'end (1269bp)
Cel-16R	AAATTTGTCCCAACCACGAA	
Cel-17F	TGTGGCGATCGAAAAGTTAG	Amplify overlapping sequence between *wss*I and AMN-42 (1299bp)
Cel-17R	CGGAAAGAATGCTGGACCT	

Allocation transcriptional start point (5' RACE analysis)		
GSPcel-1	TCAGGCACCGATTGATCAGTTACGC	Reverse amplification 130bp from ATG of *wss*A
GSPcel-2	CGTACTTTTACCCACACCGCCCTTG	Reverse amplification 386bp from ATG of *wss*A
GSPcel-3	ACCAGCCAGTGTGGATCGTTTTCCT	Reverse amplification 677bp from ATG of *wss*A
GSPcel-4	CGCAGCGTCAATTCCAGATTCAGAC	Forward amplification 452bp to TAA of *wss*I
GSPcel-5	ACGACCTGTTCGGCGACAGTAACCT	Forward amplification 253bp to TAA of *wss*I
GSPcel-6	GTCTACTTCGACAACCCGGCCTTCC	Forward amplification 98bp to TAA of *wss*I