Identification of protein secretion systems and type III effectors in wood-associated bacteria of the genus *Xylophilus*

Zoe Roux[1], Naama Wagner[2], Laurent Brottier[1], Tal Pupko[2], and Ralf Koebnik[1,*]

[1] PHIM, Université de Montpellier, Cirad, INRAE, Institut Agro, IRD, Montpellier, France

[2] Shmunis School of Biomedicine and Cancer Research, Tel Aviv University, Israel

**Author for Correspondence*: Ralf Koebnik, Plant Health Institute of Montpellier, Institut de Recherche pour le Développement, Montpellier, France;

phone: +33-467-416228; email: koebnik@gmx.de

ABSTRACT

To shed light into the biology of bacteria belonging to the genus *Xylophilus*, including the grapevine pathogen *Xylophilus ampelinus*, we scrutinized all available genomes of the genus for the presence of type III secretion and flagellar systems. We found three different flagellar systems in the genus, one of which was present in all twelve strains with good-quality genome sequences were available. The other two flagellar systems were only detected in one or two strains. We also identified two types of type III secretion systems, likely under control of the AraC-type transcriptional activator HrpX. One system with resemblance to systems from plant-pathogenic bacteria was only found in the grapevine pathogen. The other system was found in three strains of *Xylophilus*, all isolated from plant material. We predicted genes that are co-regulated with the type III secretion systems, as supported by the presence of strongly conserved HrpX-binding promoter elements. We identified about 40 type III effectors in the grapevine pathogen with homologs in plant pathogenic bacteria. In contrast, a rhododendron flower isolate had only two type III effector gene candidates with conserved HrpX-binding promoter elements but many genes without homologs beyond the species. Finally, we predicted and confirmed three novel effector candidates from *X. ampelinus* to contain a functional type III secretion signal using an AvrBs1 reporter approach. The presence of type III effectors suggests that effector-triggered immunity may exist in grapevine or non-host plants and that strategies targeting type III effectors for resistance engineering may contribute to suitable control measures.

INTRODUCTION

Xylophilus is an understudied bacterial genus within the order *Burkholderiales*, family *Comamonadaceae*. These bacteria are primarily known for their association with plants, particularly with woody tissues, which is reflected in the genus name (from Greek: *xýlon* – wood; philus – friend, loving) (https://lpsn.dsmz.de/, accessed on 08/07/2025; Freese et al. 2026). For instance, the type strain of the genus, belonging to the species *Xylophilus ampelinus* (formerly known as *Xanthomonas ampelina*), has been isolated from a diseased grapevine plant (Willems et al. 1987a, 1987b). Bacteria belonging to another proposed species, *Xylophilus rhododendri*, have been isolated from the flowers of a *Rhododendron* plant (Lee et al. 2020). Other *Xylophilus* bacteria without species designation have been identified in metagenomic samples from woody tissues, e.g., from almond and blueberry (Anguita-Maeso et al. 2022; Chacón et al. 2022).

Xylophilus bacteria are of particular interest to plant pathologists and viticulturists due to their potential to damage valuable crops. One of the most notable species in this genus is the pathogen *X. ampelinus*, causing bacterial necrosis in grapevines (*Vitis vinifera*). Symptoms include discoloration of young shoots, necrotic spots on leaves, reddish-brown streaks on shoots, cracks, cankers, wilting and dieback (EFSA Panel on Plant Health 2014). Disease severity depends on cultivar susceptibility and environmental conditions (Peros et al. 1995). The bacterium enters plants through wounds or natural openings and colonizes xylem vessels, forming biofilms. Later, also phloem and cambial tissues were found to be infected. The resulting disruption of water transport causes tissue decay (EPPO 2009).

Although *X. ampelinus* is relatively rare compared to other grapevine pathogens, its impact in the affected regions can be serious. It is listed as a quarantine pathogen in several countries due to its potential to spread and damage viticulture industries (EPPO 2025). Understanding and monitoring *X. ampelinus* is critical to maintaining healthy vineyards and ensuring the sustainability of grape production, particularly in Europe where it has been of most concern in the past.

Many plant-associated bacteria of the phylum Pseudomonadota, including Alphaproteobacteria (e.g., *Bradyrhizobium*, *Mesorhizobium*, *Sinorhizobium*), Betaproteobacteria (e.g., *Acidovorax*, *Burkholderia*, *Paracidovorax*, *Papaburkholderia*, *Ralstonia*) and Gammaproteobacteria (e.g., *Pantoea*, *Pseudomonas*, *Xanthomonas*) rely on a specialized protein secretion system for their interaction with the plant, be it a beneficial or a pathogenic relationship (Büttner 2016; Schreiber et al. 2021; Teulet et al. 2022). The type III secretion system (T3SS) injects effector proteins, called type III effectors (T3Es), into the host cells to the benefit of the bacteria, thereby favoring the colonization (infection or symbiotic) process. Historically, the T3SS has first been characterized in human pathogens, focusing on species of *Yersinia*, *Salmonella* and *Shigella* (Cornelis 2002).

Based on sequence similarities, T3SSs have been classified into seven major classes, reflecting the bacteria from which they were first characterized (Chlamy, Hrp1/Hrc1, Hrp2/Hrc2, SPI-1 [Inv/Mxi/Spa], SPI-2 [Esc/Ssa], Rhizo [Rhc], Ysc) (Cornelis 2006; Abby and Rocha 2012; Zboralski et al. 2022). The T3SS is a multi-component protein complex which contains 13 conserved proteins in different copies, namely the ATPase SctN, the sorting platform consisting of SctK, SctL and SctQ, the inner membrane secretion apparatus consisting of SctR, SctS, SctT, SctU and SctV, the IM ring consisting of SctD and SctJ, the inner rod protein SctI, the outer membrane secretin SctC, and the needle filament SctF (Abby and Rocha 2012), according to the unified nomenclature (Wagner and Diepold 2020). Nine of them (SctJLNQRSTUV) have clear homologs in the flagellar basal body, and two (SctP and SctF) have functional and structural counterpart analogs in the flagellum. Compelling evidence suggests that the non-flagellar T3SS evolved from the flagellum (Abby and Rocha 2012; Diepold and Armitage 2015; Worrall et al. 2023). The two gate proteins SctU (and its flagellar homolog FlhB) and SctV (and its flagellar homolog FlhA) constitute the most conserved components of the T3SS. As such, they are considered as a signature for T3SS presence (or remnants thereof) and allow classification of the non-flagellar T3SS into seven classes (Gazi et al. 2012; Zboralski et al. 2022).

Different T3SS classes exhibit distinct regulatory mechanisms governing gene expression (Büttner 2012). For instance, the Hrp1/Hrc1 system is controlled by the extracytoplasmic function (ECF) sigma factor HrpL, which is thought to direct the RNA polymerase to the '*hrp* box' in the promoters of T3SS genes (Kvitko and Collmer 2023). The *hrpL* gene itself is regulated by a heterohexamer of HrpR and HrpS in strains of *Pseudomomas syringae* or by a HrpS homohexamer

in the Enterobacterales (Kvitko and Collmer 2023). The other plant-associated T3SS, Hrp2/Hrc2, is under the control of the two-component response regulator HrpG. It is thought that HrpG is post-translationally activated by protein phosphorylation by a hitherto unknown component, leading to transcriptional activation of the *hrpX* gene (Wengelnik et al. 1996). The AraC-type regulator HrpX (in *Acidovorax* and *Xanthomonas*)/HrpB (in *Burkholderia* and *Ralstonia*) then binds to a conserved promoter element, the *hrpII*/PIP box, thus regulating T3SS and T3E genes (Koebnik et al. 2006). Activation of HrpX-controlled genes strictly relies on the presence of a PIP box and a properly spaced –10 promoter motif (Cunnac et al. 2004a; Tsuge et al. 2005; Furutani et al. 2006; Lipscomb and Schell 2011). Many, if not most HrpX/HrpB-regulated genes encode components of the T3SS machinery, T3Es, or secreted hydrolases, such as endoglucanases, proteases, esterases, and pectinases (Noël et al. 2001; Cunnac et al. 2004b; Occhialini et al. 2005; Tamura et al. 2005; Lipscomb and Schell 2011; Jiménez-Guerrero et al. 2020; Teper et al. 2021). The presence of HrpX-regulated promoters, identified either experimentally or in silico, is used to predict T3E genes, which can then be confirmed and characterized in more detail (Noël et al. 2001; Furutani et al. 2006; Jiménez-Guerrero et al. 2020).

In order to better understand the host-pathogen molecular interactions within the genus *Xylophilus*, we aimed at predicting and comparing T3SSs and candidate T3E proteins at a genome-wide level. Our work provides new insights into the evolution of T3SSs in the genus *Xylophilus* and the presence of T3Es in these species. For three candidate T3Es, we provide evidence that they are bona fide effectors likely translocated by the T3SS and injected into grapevine plants.

MATERIAL AND METHODS

Biological material

Bacterial strains and growth conditions

Strains of *Escherichia coli* (DH10B) and *Xanthomonas campestris* (8000 and its mutants in the T3E *avrBs1* and the type III secretion gene *hrcV*) were used in this study (Durfee et al. 2008; Xu et al. 2008; Zhao et al. 2013). The *X. campestris* strains were cultivated at 28 °C on solid PSA (Peptone-Sucrose-Agar) medium (peptone 10 g/L, sucrose 10 g/L, agar 16 g/L, glutamic acid 1 g/L).

Escherichia coli cells were grown at 37 °C on LB (lysogenic broth) agar plates (LB Lennox agar 35 g/L; Euromedex, Souffelweyersheim, France) or in liquid LB medium (tryptone 10 g/L, yeast extract 5 g/L, NaCl 10 g/L). Antibiotics were added, when needed, to the media at the following final concentrations: 100 mg/L for rifampicin and 5 mg/L for tetracycline.

Plasmids

The plasmid pBBR1MCS-3::avrBs1$_{59-445}$ is a derivative of the plasmid pBBR1MCS-3. It encodes an N-terminally truncated variant of AvrBs1 from *X. campestris*. This variant lacks its own type III secretion signal and is used to construct translational reporter gene fusions with presumed type III secretion signals from candidate T3E genes (Kovach et al. 1995; Zhao et al. 2013). Reporter fusions were constructed by GenScript (Piscataway Township, NJ, USA) from synthetic DNA that consisted of the 100-base-pair noncoding sequence upstream of the translation initiation codon, as well as the 300-base-pair coding sequence that would include the N-terminal type III secretion signal from the candidate genes. The helper plasmid pRK2013 was used to transfer plasmids via conjugation from *E. coli* to *Xanthomonas* by triparental mating (Figurski and Helinski 1979).

Plant material

The pepper cultivar ECW-10R, which triggers a hypersensitive response upon perception of AvrBs1 by the corresponding *Bs1* resistance gene product, was used for plant inoculation experiments (Kousik and Ritchie 1999). Plants were grown in a greenhouse with natural daylight cycles at 27°C ± 5°C and 60% relative humidity.

Bioinformatic analyses

Genomic resources

Genomic resources, such as genome sequences, genome annotations, predicted open reading frames (ORFs) and proteomes, were retrieved from the NCBI Datasets resource (https://www.ncbi.nlm.nih.gov/datasets/) (Sayers et al. 2025). Quality and correct taxonomic classifications were evaluated using the Type (Strain) Genome Server (TYGS; https://tygs.dsmz.de/) and the Microbial Genomes Atlas (MiGA; https://gateway.microbial-genomes.org/) (Rodriguez-R et al. 2018; Freese et al. 2026). Genome-wide average nucleotide

identities (ANI) were calculated using FastANI version 1.3 (Jain et al. 2018), as implemented at the French Galaxy server (https://usegalaxy.fr/) (Galaxy Community 2024).

Sequence comparisons and phylogenetic trees

Sequence similarity searches and pairwise sequence comparisons were performed at the National Center for Biotechnology Information (NCBI) using the BLAST and Needleman-Wunsch algorithm, respectively (Needleman and Wunsch 1970; Altschul et al. 1990; Sayers et al. 2025). Multiple sequence alignments were generated with MUSCLE (Edgar 2004), as implemented at the EMBL's European Bioinformatics Institute (EMBL-EBI) (Thakur et al. 2025). To compare T3SS gene clusters, the corresponding regions were extracted from the annotated GenBank files using the Artemis genome browser (Rutherford et al. 2000). Clinker was used for the comparison and visualization of genomic regions (Gilchrist and Chooi 2021).

Phylogenetic trees based on sequence comparisons were generated at Phylogeny.fr (https://www.phylogeny.fr/), with default parameters (Dereeper et al. 2008). Core genome-wide comparisons were established using the M1CR0B1AL1Z3R suite (https://microbializer.tau.ac.il/) (Shimony et al. 2025). For better visualization of the trees, we used the Interactive Tree of Life (iTOL) tool (Letunic and Bork 2024).

Prediction of protein secretion systems and type III effectors

Protein secretion systems were predicted using TXSScan, as implemented in the Galaxy platform at the Institut Pasteur (https://galaxy.pasteur.fr/root?tool_id=toolshed.pasteur.fr/repos/odoppelt/txsscan/TXSScan/1.0.2) (Abby et al. 2024). The type IV and type VI secretion system were also predicted using SeCreT4 and SeCreT6 (https://bioinfo-mml.sjtu.edu.cn/SecReT4/; https://bioinfo-mml.sjtu.edu.cn/SecReT6/), respectively (Bi et al. 2013; Zhang et al. 2023).

To predict T3Es, we employed Effectidor II, a pan-genomic AI-based algorithm for the prediction of T3Es (https://effectidor.tau.ac.il) (Wagner et al. 2025). The input consisted of predicted ORFs and GFF3 annotations, as available from NCBI, targeting the genomes of five strains that harbor a Hrp2/Hrc2-related T3SS (assembly accession numbers GCA_003217575.1, GCA_024832295.1, GCA_033546935.1, GCA_009906855.1, and GCA_026951875.1 for strains CECT7646, CFBP 1192,

GOD-11R, KACC 21265, and Kf1, respectively). The predicted proteomes of strains ASV27, bin23, GW821-FHT01B05, Leaf220 and SP210_2 were used as a negative training set (assembly accession numbers GCF_016428875.1, GCA_036946225.1, GCA_038961845.1, GCA_001421705.1 and GCF_913776965.1, respectively) to represent strains that do not encode a Hrp2/Hrc2-related T3SS. Search options included prediction of the type III secretion signal and searching for the presence of a PIP-box motif in the upstream region of all ORFs. All other parameters were left at their default settings.

Other bioinformatic tools (Prediction of protein domains and structural modeling of predicted type III effectors)

Protein domains and sequence motifs were predicted using InterProScan (https://www.ebi.ac.uk/interpro/) (Jones et al. 2014). Three-dimensional structures of proteins were predicted by AlphaFold3 (https://alphafoldserver.com/welcome) (Abramson et al. 2024). FoldSeek was used to identify structural homologs (https://search.foldseek.com/) (van Kempen et al. 2024). To reduce complexity of datasets, groups of homologous sequences were clustered using CD-Hit at usegalaxy.fr (Galaxy Version 4.8.1) (Li et al. 2001). Signal sequences for secretion via the Sec or Tat pathway were predicted with SignalP 6.0 (https://services.healthtech.dtu.dk/services/SignalP-6.0/) (Teufel et al. 2022). GeNomad in its Galaxy version 1.11.1 was used to identify mobile genetic elements, such as prophages (Camargo et al. 2024).

Plant inoculation experiments

Eight to twelve weeks old plants of the pepper cultivar ECW-10R were used for plant inoculation experiments (Kousik and Ritchie 1999). Bacteria were grown on PSA medium, resuspended in 10 mM $MgCl_2$ and diluted to an optical density at 600 nm wavelength of 0.1. A needleless 1-mL plastic syringe was used to inoculate the bacterial solution from the lower side of the pepper leaf into the apoplast. Each bacterial solution was inoculated on at least two different leaves from at least two different plants. Experiments were performed at least two times. Plant reactions were scored over a period of seven days after the inoculation.

RESULTS

Clarification of taxonomic status of publicly available genome sequences

In order to obtain an overview of the presence of virulence-associated secretion systems in the genus *Xylophilus*, we first extracted all genome sequences that are assigned to this bacterial genus from NCBI (https://www.ncbi.nlm.nih.gov/datasets/genome/). NCBI GenBank lists 19 genome sequences for this genus (accessed on 29-JUL-2025), six for the species *X. ampelinus*, one for the species *X. rhododendri*, and twelve without species assignment (Table 1).

1 **Table 1.** *Xylophilus* genome sequences at NCBI GenBank.

Organism name at NCBI	Strain/Isolate	GenBank assembly	RefSeq assembly	Genome accession number	Number of contigs	Size (bp)	BioProject	BioSample	Reference
X. ampelinus	CECT7646	GCA_003217575.1	GCF_003217575.1	QJTC01	66	3,731,505	PRJNA463320	SAMN09074800	
X. ampelinus	CFBP 1192	GCA_024832295.1	GCF_024832295.1	JAMOFZ01	85	3,736,570	PRJNA843606	SAMN28766069	Portier et al. 2022
X. ampelinus	Xylophilus_ampelinus_ BgEED09[(i)]	GCA_901875635.1	GCF_901875635.1[§]	CABFMN01[#]	161	6,174,221	PRJEB32184	SAMEA5664384	
X. ampelinus	bin23[(i)]	GCA_036946225.1		JAQIJO01	14	3,596,058	PRJNA922445	SAMN32653142	
X. ampelinus	CCH5-B3	GCA_001556675.1	GCF_001556675.1[§]	LSJH01[#]	1406	6,019,991	PRJNA299404	SAMN04299458	Soto-Giron et al. 2016
X. ampelinus	ERR594292_binner13_ Refined_23[(i)]	GCA_964011595.1		CAXGPN01[#]	932	5,771,620	PRJEB51168	SAMEA115421838	
X. rhododendri	KACC 21265	GCA_009906855.1	GCF_009906855.1	CP047650.1, CP047651.1	2	5,873,400	PRJNA600143	SAMN13783577	Lee et al. 2020
Xylophilus sp.	ASV27	GCA_016428875.1	GCF_016428875.1[§]	JAELTO01[#]	469	4,734,944	PRJNA682646	SAMN17004937	Goyal et al. 2022
Xylophilus sp.	cluster_DBSCAN_ round5_1[(i)]	GCA_009914555.1	GCF_009914555.1	WNDY01	129	4,706,822	PRJNA531449	SAMN12995593	Waterworth et al. 2020
Xylophilus sp.	GOD-11R	GCA_033546935.1	GCF_033546935.1	CP137854.1	1	5,127,500	PRJNA1034763	SAMN38071784	
Xylophilus sp.	Go_Prim_bin_55[(i)]	GCA_017990095.1		JAGNON01	437	2,320,559	PRJNA524094	SAMN18119707	Schneider et al. 2021
Xylophilus sp.	Gw_Inlet_bin_57[(i)]	GCA_018006615.1		JAGMXZ01	431	1.897,017	PRJNA524094	SAMN18119261	Schneider et al. 2021
Xylophilus sp.	Gw_Prim_bin_50[(i)]	GCA_018005875.1		JAGMZJ01	323	1,282,324	PRJNA524094	SAMN18119294	Schneider et al. 2021
Xylophilus sp.	Gw_UH_bin_252[(i)]	GCA_017989255.1		JAGNQF01	338	1,400,660	PRJNA524094	SAMN18119505	Schneider et al. 2021
Xylophilus sp.	GW821-FHT01B05	GCA_038961845.1	GCF_038961845.1[§]	CP152408.1[#]	1	5,832,500	PRJNA1101085	SAMN40985315	Biggs et al. 2024
Xylophilus sp.	Kf1	GCA_026951875.1		WPBX01[#]	1939	5,678,042	PRJNA593051	SAMN13451138	
Xylophilus sp.	Leaf220	GCA_001421705.1	GCF_001421705.1	LMKO01	28	4,483,623	PRJNA297956	SAMN04151686	Bai et al. 2015
uncultured *Xylophilus* sp.	SP51_3_metabat2_ genome_mining.45[(i)]	GCA_913777525.1		CAJYSJ01	494	3,038,960	PRJEB45634	SAMEA8944104	Su et al. 2022
uncultured *Xylophilus* sp.	SP210_2_metabat2_ genome_mining.41[(i)]	GCA_913776965.1	GCF_913776965.1	CAJYQL01	136	4,051,675	PRJEB45634	SAMEA8944525	Su et al. 2022
Xenophilus aerolatus	DSM 19424	GCA_029968055.1		JASCNX01[#]	11	5,117,920	PRJNA967166	SAMN34578380	Kim et al. 2010
Xenophilus arseniciresistens	YW8	GCA_028023875.1	GCF_028023875.1	JAQIPB01	27	5,017,807	PRJNA922719	SAMN32673052	Li et al. 2014
Xenophilus azovorans	DSM 13620	GCA_000745855.1	GCF_000745855.1	JQKD01	137	6,529,337	PRJNA221063	SAMN02745804	Blümel et al. 2001

2 Strain isolates Xylophilus_ampelinus_BgEED09, CCH5-B3 and ERR594292_binner13_Refined_23 do not belong to the genus *Xylophilus*. Sample bin23 appears to belong to the genus *Xylophilus* but
3 not to the species *X. ampelinus*.
4 (i) Warning at NCBI: contaminated; §, RefSeq assembly suppressed
5

The Type strain Genome Server (TYGS; https://tygs.dsmz.de) was applied to confirm the taxonomic classification (Freese et al. 2026) (accessed on 30-JUL-2025) (Supplementary Table S1a). Only the two sequences of strains CECT7646 and CFBP 1192, both corresponding to the same type strain, which was isolated from a necrotic grapevine plant in Crete, Greece, in 1966, were reported to belong to the species *X. ampelinus*. Strain KACC 21265 was confirmed to belong to the species *X. rhododendri*. Eight genome sequences (strains/isolates ASV27, bin23, Cluster_DBSCAN, GOD-11R, GW821-FHT01B05, Leaf220, SP51_3, SP210_2) were classified as a potential new species of the genus *Xylophilus*. Three of the genome sequences (strains/isolates BgEED, CCH5-B3, ERR594292) were classified as a potential new species of the genus *Xenophilus*. Four isolates with surprisingly small genome sizes, originating from metagenome assemblies from wastewater treatment plants in Germany (Schneider et al. 2021), were classified as potential new species of the genera *Acidovorax* (Go_Prim_bin_55, Gw_Prim_bin_50, Gw_UH_bin_252) or *Macromonas* (Gw_Inlet_bin_57). Lastly, the genome sequence of strain Kf1 was flagged as potentially unreliable identification result due to potential sequence contamination, belonging to the species *Staphylococcus epidermidis*.

The TypeMat genome-assisted taxonomic classifier at the Microbial Genomes Atlas (MiGA; https://uibk.microbial-genomes.org/) was used to complement this analysis (accessed on 30-JUL-2025) (Supplementary Table S1b). Except for isolate Cluster_DBSCAN, which was classified as a new species of the genus *Paracidovorax*, all classifications with respect to *Xylophilus* and *Xenophilus* were confirmed. The four wastewater metagenome assemblies were flagged as intermediate-quality assemblies (in contrast to all the other sequences, which were of high or excellent quality) and assigned to the genera *Acidovorax* (Go_Prim_bin_55, Gw_Prim_bin_50) and *Variovorax* (Gw_UH_bin_252). Surprisingly, strain Kf1 was reported as a new species of the genus *Xylophilus*, which might be due to the fact that MiGA only considers contigs with at least 1,000 bp for classification. In the case of strain Kf1, only 116 of 1,939 contigs were considered and sequence information corresponding to 1,823 contigs (775,918 bp) was omitted, which might explain the classification as a *Staphylococcus* strain by TYGS. Hence, ten genome sequences were classified as *Xylophilus* by both tools, one was only supported by TYGS (Cluster_DBSCAN) and another was only supported by MiGA (Kf1).

To clarify the sequence relationships among the various genomes, genome-wide ANI were calculated (Jain et al. 2018) (Supplementary Table S2a). Four genome sequences, originating from the wastewater metagenome assemblies, were omitted due to poor quality and apparent incompleteness as evidenced by their small sizes and quality metrics (Supplementary Table S1b). This analysis suggests that the twelve *Xylophilus* strains, including Cluster_DBSCAN and Kf1, belong to ten different species. CFBP 1192 and CECT7646, with genome sequences sharing more than 99.99% sequence identity, are both equivalent of the *X. ampelinus* type strain of the species held in two different collections. The genome sequences of isolates SP51_3 and SP210_2 were more than 97.8% identical, suggesting that they both correspond to the same *Xylophilus* species. The three sequences that appeared to belong to the genus *Xenophilus* (strains/isolates BgEED, CCH5-B3, ERR594292) were more than 98.5% identical to each other, indicating that they belong to the same species. The most closely related of these three strains is *Xenophilus aerolatus*, whose type strain, DSM 19424, shares more than 91.9% sequence identity with their genomes.

Digital DNA-DNA hybridization (DDH) was performed to complement the ANI calculations (Meier-Kolthoff et al. 2013) (Supplementary Table S2b). This analysis confirmed the conclusion that the twelve *Xylophilus* genomes correspond to 10 species. DDH of the genome sequences of isolates SP51_3 and SP210_2 was estimated with 86%, significantly above the generally accepted species boundary of 70% (Chun et al. 2018). Likewise, strains/isolates BgEED, CCH5-B3, ERR594292 had pairwise DDH values of at least 96% to each other and were most similar to the *X. aerolatus* type strain DSM 19424 with 45-46% DDH values. Taken together, these analyses established a well-curated dataset of twelve genome sequences from eleven different strains for further analyses.

Phylogenetic tree and presence of flagellar and type III secretion genes

Since many Gram-negative plant-associated bacteria rely on a T3SS for pathogenicity or symbiosis, and because most of the *Xylophilus* strains were found to be associated with plants, we hypothesized that these bacteria encode such a secretion system in their genomes. The genome sequences were scrutinized for the presence of T3SS-related sequences based on two of the most conserved proteins, SctV (a homolog of the flagellar FlhA) and SctU (a homolog of the flagellar FlhB) (Gazi et al. 2012; Tampakaki 2014; Zboralski et al. 2022). One representative sequence from each of the seven T3SS classes was used as the query, as well as one flagellar

homologous sequence, i.e., seven SctU and FlhB sequences, and seven SctV and FlhA sequences (Figure 1). BLASTP searches against the predicted proteomes of eleven *Xylophilus* strains identified 19 hits for SctU/FlhB and 19 hits for SctV/FlhA. Each strain contained one to three paralogs of each protein. Of the 19 hits, four represent paralogous doublets derived from the two *X. ampelinus* type strain genomes. Since the genome sequence of isolate SP51_3 was not annotated, we used TBLASTN to find homologs of SctU and SctV.

A multiple sequence alignment of the 20 *Xylophilus* SctV homologs revealed that they fall into five groups (Supplementary Data S1). A phylogenetic tree reconstructed based on the multiple sequence alignment of the 20 sequences as well as representatives of flagellar and T3SS sequences (Zboralski et al. 2022), further supported that these sequences form five clades, which we tentatively called SctV1, SctV2, FlhA1, FlhA2 and FlhA3 (Figure 1). Members of both SctV clades aligned within the Hrp2/Hrc2 class of SctV proteins. Members of clade FlhA1 and the FlhA2 sequence clustered within the flagellar class. Finally, the two FlhA3 sequences formed a sister clade of the flagellar class, suggesting that they belong to flagellar systems. A similar result was obtained when comparing representative SctU/FlhB sequences with the homologs from *Xylophilus* (Supplementary Figure S1).

To examine the distribution of the various flagellar and secretion systems among the *Xylophilus* strains, a core-genome based phylogenetic tree was computed using the M1CR0B1AL1Z3R suite (Shimony et al. 2025) (Figure 2). All strains were found to possess the type-1 flagellum comprising FlhA1 and FlhB1. The second flagellar type with FlhA2 and FlhB2 was only found in strain GW821-FHT01B05. The distribution of the third flagellar type with FlhA3 and FlhB3 was patchy and did not follow phylogenetic descent. The Hrp2/Hrc2 T3SS with SctU1 and SctV1 was only found in the *X. ampelinus* strain (CECT7646/CFPB 1192). The second Hrp2/Hrc2 T3SS type was present in a lineage comprising the strains Kf1, GOD-11R, and KACC 21265.

We also evaluated the presence of two additional multi-protein secretion systems, the type 4 and type 6 secretion systems (T4SS, T6SS). The presence of these systems was predicted using SeCreT4 and SeCreT6 (Bi et al. 2013; Zhang et al. 2023). The i3-class T6SS was found in the same clade as the type-2 T3SS, whereas the i4-class T6SS was found in three genetic lineages comprising DBSCAN, bin23, and the two similar isolates SP51_3 and SP210_2. Gene clusters

encoding T4SS were found in five strains, DBSCAN, Kf1, GW821-FHT01B05, ASV27 and CECT7646/CFPB 1192 and their presence-absence pattern suggests multiple gain and loss events of this system along the phylogenetic tree (Fig. 2). These results were confirmed using the TXSScan program (Supplementary Table S3) (Abby et al. 2024).

Hrp2/Hrc2-class T3SSs are under transcriptional control of HrpX, a AraC-type regulator that binds to a conserved plant-inducible promoter (PIP) box. When a matching -10 promoter element is in a proper distance of 30 to 32 base pairs, genes are transcribed, most of them encoding components of the T3SS, T3Es and type II-secreted cell wall-degrading enzymes. Indeed, perfect PIP patterns with the consensus sequence TTCGB-N_{15}-TTCGB-$N_{30\text{-}32}$-YANNNT were found in the promoters of the Hrp2/Hrc2 T3SS gene clusters, in front of *sctD* and *sctU* in strain CECT7646 and in front of *sctQ* in strain KACC 21265 (Table 2). Nine of the 16 PIP patterns in the genome of strain CECT7646 were found upstream of T3E genes (*hopQ*, *ripC*, *ripE*, *ripAC*, *ripAQ*, *xopD*, *xopZ*, *xopAI, xopBG*) and one upstream of a pectate lyase gene (*ripW*). Two genes, DFQ15_RS08770 and DFQ15_RS08785, had only a homolog in *Sphingomonas* sp. NCPPB 2930 in a region with a similar T3SS.

The predicted PIP-box harboring gene repertoire in strain KACC 21265 was different from that of strain CECT7646. Among its 32 genes encoding hypothetical proteins were two candidate T3E genes. Protein QHI99106.1 (locus tag GT347_14590) was 22.90% identical to XopN (WP_011348010.1) over 69% of its sequence, and protein QHI99786.1 (locus tag GT347_18445) was 26.04% identical to RipL (WP_014630780.1) over 80% its sequence. Notably, 15 genes were found to encode members of strain-specific protein families with unknown function, with up to 13 paralogs (Supplementary Table S4; Supplementary Data S2). All of them are annotated as hypothetical proteins. We used InterProScan to shed light on their potential functions. However, no known protein domain was found for any of them. A signal sequence for secretion was predicted for two members of the 13 homologs of protein family 4. For 13 additional hypothetical proteins beyond the eight protein families, we did not find any homolog using BLASTP searches at NCBI GenBank (Table 2).

Table 2. Candidate HrpX-regulated promoters of *X. ampelinus* strain CECT7646 and *X. rhododendri* strain KACC 21265.

Locus_tag	Old	PIP box sequence[a]	Distance 1 (bp)[b]	–10 motif	Distance 2 (bp)[c]	Homolog/ Remark[d]
Strain		**CECT7646**				
DFQ15_RS16870	N.A.	**ttcgt**cagcctgcacatcgg**ttcgt**	30	**ta**atg**t**	29	XopBG
DFQ15_RS01080	DFQ15_101225	**ttcgc**gctgctggcccagct**ttcgc**	31	**ca**cac**t**	59§	HopQ
DFQ15_RS02885	DFQ15_102289	**ttcgc**tccgtcatgatcgac**ttcgt**	31	**ca**cac**t**	95	XopZ
DFQ15_RS05435	DFQ15_10577	**ttcgt**tctcataccgcccgg**ttcgg**	31	**ca**cgg**t**	-525	Antisense
DFQ15_RS17705	DFQ15_105103	**ttcgt**ttgtcatgtcacatc**ttcgc**	30	**ta**tca**t**	236	RipC
DFQ15_RS05805	DFQ15_10628	**ttcgg**ccccgcgcccacggc**ttcgt**	31	**ta**agg**t**	85§	RipAQ
DFQ15_RS07770	N.A.	**ttcgc**tttggccgcgaaata**ttcgc**	30	**ca**tcc**t**	33	XopD
DFQ15_RS08655	DFQ15_1117	**ttcgt**ttcgcgcaacatcgg**ttcgc**	30	**ca**tcc**t**	45	RipW
DFQ15_RS08715	DFQ15_11120	**ttcgc**aaatcatgaaccggc**ttcgc**	31	**ta**gtc**t**	74§	SctD
DFQ15_RS08730	DFQ15_11123	**ttcgc**accgtgggccaccgc**ttcgg**	30	**ta**gat**t**	66	SctU
DFQ15_RS08770	DFQ15_11131	**ttcgg**tattcatgcaagttc**ttcgt**	30	**ta**gcc**t**	30	Hpa3
DFQ15_RS08785	DFQ15_11134	**ttcgc**aaatcgcggatcgcc**ttcgc**	30	**ta**ttc**t**	66	Hpa2
DFQ15_RS10885	DFQ15_11536	**ttcgt**atagcagggaatatg**ttcgc**	31	**ca**cca**t**	120	RipAC
DFQ15_RS11495	DFQ15_11681	**ttcgg**ctacgcgcatgctgt**ttcgc**	31	**ca**aga**t**	102	XopAI
DFQ15_RS14030	DFQ15_12414	**ttcgc**ccggacgcatggcac**ttcgg**	31	**ca**atc**t**	74§	RipE
DFQ15_RS15225	DFQ15_12925	**ttcgc**caggcagtcgggcgg**ttcgg**	31	**ca**ccg**t**	-1172	Within CDS
Strain		**KACC 21265**				
GT347_RS00670	GT347_00670	**ttcgg**ctgggtttgacggac**ttcgc**	31	**ta**cgc**t**	28	2 paralogs
GT347_RS00990	GT347_00990	**ttcgt**cgcacggcagcagcc**ttcgt**	31	**ta**gag**t**	27	
GT347_RS02290	GT347_02290	**ttcgt**gagaaacggtggcca**ttcgg**	31	**ca**cag**t**	30	9 paralogs
GT347_RS02345	GT347_02345	**ttcgg**ggcgaacaacagtga**ttcgc**	31	**ca**cag**t**	31	9 paralogs
GT347_RS02675	GT347_02675	**ttcgt**gcccgcgcggtcgtt**ttcgc**	30	**ta**ggc**t**	46	No homologs
GT347_RS03520	GT347_03520	**ttcgc**cttcggtcctgcagc**ttcgg**	31	**ca**agc**t**	265	
GT347_RS04255	GT347_04255	**ttcgc**atggaacgggcggca**ttcgt**	31	**ca**gag**t**	30	5 paralogs
GT347_RS04260	GT347_04260	**ttcgc**aataaacaagcggca**ttcgg**	31	**ca**aag**t**	31	5 paralogs
GT347_RS04265	GT347_04265	**ttcgc**gagggccagggagct**ttcgt**	31	**ca**aag**t**	32	5 paralogs
GT347_RS05075	GT347_05075	**ttcgt**aatgcacgggggata**ttcgc**	30	**ca**aag**t**	30	5 paralogs
GT347_RS09625	GT347_09625	**ttcgc**ttccgctgcgcggtc**ttcgc**	30	**ca**aag**t**	29	No homologs
GT347_RS10730	GT347_10730	**ttcgc**ctcggcacacggata**ttcgc**	32	**ca**aaa**t**	36	
GT347_RS12520	GT347_12520	**ttcgc**gtccggcccgaggca**ttcgg**	31	**ca**gca**t**	49	No homologs
GT347_RS14590	GT347_14590	**ttcgg**gtgattcgatctgca**ttcgg**	31	**ca**tct**t**	29	XopN
GT347_RS15795	GT347_15795	**ttcgc**cgcggcacggccggc**ttcgc**	31	**ca**tcg**t**	33	13 paralogs
GT347_RS15815	GT347_15815	**ttcgc**cgcagcggcagcggc**ttcgc**	31	**ca**tct**t**	38	13 paralogs
GT347_RS15845	GT347_15845	**ttcgc**cacgggggtgccggc**ttcgc**	31	**ca**tcg**t**	30	13 paralogs
GT347_RS15855	GT347_15855	**ttcgc**tcgggcgtggcggtt**ttcgc**	30	**ca**tga**t**	35	2 paralogs
GT347_RS15860	GT347_15860	**ttcgt**ttgcccagcctgttg**ttcgg**	32	**ca**cca**t**	145§	2 paralogs
GT347_RS17215	GT347_17215	**ttcgc**cgcggcgtgccgggt**ttcgg**	30	**ca**tat**t**	38	No homologs
GT347_RS17325	GT347_17325	**ttcgg**ccgtgggcccaggca**ttcgt**	31	**ca**gca**t**	33	No homologs
GT347_RS17385	GT347_17385	**ttcgg**tcgggatgaccgtct**ttcgg**	32	**ta**gcg**t**	33§	2 paralogs
GT347_RS18105	GT347_18105	**ttcgg**tttcaccgacgggat**ttcgc**	31	**ca**cat**t**	82	No homologs
GT347_RS18300	GT347_18300	**ttcgc**ctgcggccgtgcgtt**ttcgg**	30	**ca**aca**t**	246	8 paralogs
GT347_RS18380	GT347_18380	**ttcgc**gggatgcggagagcg**ttcgt**	31	**ca**gag**t**	31	No homologs
GT347_RS18405	GT347_18405	**ttcgg**cggcggcggcgcggg**ttcgc**	31	**ca**ccg**t**	41	
GT347_RS18445	GT347_18445	**ttcgt**ggcccggggcgcgga**ttcgc**	32	**ta**cct**t**	7 (73§)	RipL
GT347_RS19325	GT347_19325	**ttcgc**gcgcggccgcagcgc**ttcgc**	30	**ca**aga**t**	25	No homologs
GT347_RS19630	GT347_19630	**ttcgc**attcgccgccacaac**ttcgc**	31	**ga**tca**t**	40	No homologs
GT347_RS19635	GT347_19635	**ttcgc**ctggcatcgcaaggt**ttcgg**	31	**ta**cgc**t**	32	No homologs
GT347_RS19640	GT347_19640	**ttcgc**atgcggtgacgacgc**ttcgc**	30	**ta**cgc**t**	47	SctQ
GT347_RS23280	GT347_23280	**ttcgg**tcgcggccccggtgc**ttcgg**	31	**ca**cag**t**	42	No homologs
GT347_RS23660	GT347_23660	**ttcgc**caccgcgcacggtgc**ttcgg**	31	**ca**ggc**t**	31	2 paralogs
GT347_RS24100	GT347_24100	**ttcgg**acgaaggccgacgac**ttcgc**	32	**ta**ggc**t**	175	No homologs
GT347_RS26130	GT347_26130	**ttcgg**cgcggctgcgggcca**ttcgt**	31	**ca**cgc**t**	7 (46§)	No homologs

[a] PIP box half-sites and conserved nucleotides of the –10 promoter motif are shown in bold.

[b] Distance in base pairs between the end of the PIP box and the –10 promoter motif.

[c] Distance in base pairs between the end of the –10 promoter motif and the predicted translational start codon. Cases where the structural annotation was revised, resulting in a new start codon, are indicated by a § sign.
[d] Homologs of effectors (Xop, Rip) as well as components or helpers of the T3SS (Sct, Hpa) are indicated. For 28 CDS downstream of the PIP box, no homologs were found or paralogs were only found in the same strain. One gene is encoded in antisense orientation to the PIP box. In another gene the PIP box is found within the CDS. These two cases are likely false positives and may not belong to the HrpX regulon.

Comparative genomic organization of flagellar and type III secretion gene clusters within *Xylophilus* strains

To better classify the flagellar systems and the T3SSs, we compared their genomic organization. To this end, we manually curated the structural and functional annotation of the genomic regions where the *sctV* and *sctU* genes were found, using BLAST and InterProScan searches.

Strain GW821-FHT01B05 appears to encode two types of flagella in large gene clusters. The gene cluster with *flhA1*, which was found in all *Xylophilus* strains, spans 55 genes in strain GW821-FHT01B05, among them two flagellin genes, three *flh* genes (ABF), 13 *flg* genes (ABCDEFGHIJKLM), 18 *fli* genes (ADEFGHIJKLMNOPQRST), two *mot* genes (AB) and seven chemotaxis genes (ABDRWYZ) (Supplementary Figure S2). The second gene cluster with *flhA2* contains 38 genes with only one flagellin gene, two *flh* genes (AB), 15 *flg* genes (ABCDEFGHIJ$_2$KLMN), 18 *fli* genes (ADEFGHIJK$_2$LMNPQRST), and two *mot* genes (AB). This strain contains six additional chemotaxis genes (ABDRWY), which are however encoded in another genomic location and it remains to be determined whether they correspond to the FlhA2 flagellar system.

Another candidate flagellar system, which includes *flhA3*, was found in strains ASV27 and KACC 21265. This system lacks several of the *flg*, *flh* and *fli* genes. For this system, no flagellin could be identified. Candidate genes for nine Flg proteins (ABCDEFGHI), two Flh proteins (AB) and twelve Fli proteins (AEFHIJMNOPQR) were predicted. Some of the missing genes, such as *fliG*, *fliK* and *fliL*, might correspond to the hypothetical proteins that are encoded in the gene cluster (Supplementary Figure S2).

Bona fide T3SSs were found in five strains (Figure 3). For comparison, we included the T3SSs of the Hrp2/Hrc2-representative strains that were used for classification of the SctV/FlhA and

SctU/FlhB proteins (Figure 1; Supplementary Figure S1). The two *X. ampelinus* genomes were found to encode a T3SS with 13 *sct* genes (CDJKLNPQRSTUV), which have a genomic organization that is similar to the ones found in clade 1-xanthomonads (*Xanthomonas hyacinthi*, *Xanthomonas theicola*, *Xanthomonas translucens*), in the *Ralstonia solanacearum* species complex, but also in *Collimonas fungivorans*, *Paraburkholderia andropogonis* and *Uliginosibacterium gangwonense* (Pesce et al. 2017). Notably, the *Xanthomonas hrpB* (eight genes, including *sctJ*, *sctK*, *sctL*, *sctN* and *sctT*), *hrpC* (three genes, *sctU*, *sctV* and *sctP*) and *hrpD* (four genes, including *sctQ*, *sctR* and *sctS*) operons are present (Weber et al. 2007). As in all these bacteria, the AraC-type transcriptional activator HrpX is encoded within the T3SS gene cluster and likely forms an operon with *sctC*. A second HrpX paralog with 53% sequence identity is encoded at the border of the T3SS gene cluster. The gene upstream of *hrpB1* appears to encode a protein that is distantly related to the putative translocon HrpF/NolX, as supported by PSI-BLAST analysis.

The *Xylophilus* strains KACC 21265, Kf1 and GOD11-R were found to encode another T3SS, including 13 conserved *sct* genes. Their genetic organization is different from that of systems found in other plant pathogens, such as *Acidovorax*, *Pseudomonas*, *Ralstonia* and *Xanthomonas*. For instance, *hrpB7*, which is encoded between *sctN* and *sctT* in *Ralstonia* and *Xanthomonas* and upstream of *sctT* in *Acidovorax*, is encoded downstream of *hrpX*. Also, *sctP*, which is usually encoded downstream on *sctV*, is here encoded downstream of *sctT*. No homolog of the putative translocons HrpF/NolX or HpaT was identified. And none of the *Xylophilus* T3SS gene clusters encoded an ortholog of the HrpG response regulator. In addition, we found no evidence for the presence of an HrpG ortholog elsewhere in the genome.

Prediction of candidate type III effectors in *Xylophilus*

Candidate effectors were predicted using Effectidor II, a pan-genomic AI-based web server (Wagner et al. 2025). Prediction was targeted for the five genome sequences in which T3SSs were found, i.e., strains CECT7646, CFBP 1192, GOD11R, KACC 21265, and Kf1. Effectidor II optionally utilizes information from related strains without T3SSs (ORFs shared between strains with and without T3SSs are less likely to encode effectors). The proteomes of strains ASV27, bin23, GW821-FHT01B05, Leaf220 and SP210_2 were used for this purpose.

For *X. ampelinus* with the SctV1-type T3SS, Effectidor II predicted 36 orthology groups with scores higher than 0.25 (Table 3). Two groups, OG_3401 (NBM50_08240) and OG_1546 (DFQ15_10953), are bona fide orthologs which were separated due to differences in the assembly and structural annotation of the genome sequences. The NBM50_08240 protein variant contains a 52-aa tandem repeat, which is only present once in the DFQ15_1095 variant. The NBM50_08240 protein variant has an N-terminal extension of 378 amino acids compared to the DFQ15_1095 variant. Three orthology groups, OG_3358 (NBM50_04355), OG_3350 (NBM50_03010) and OG_3346 (NBM50_02550), were only structurally annotated in the genome sequence of strain CFBP 1192. Among the 35 bona fide orthology groups, 30 had homologs in a T3E database of plant-pathogenic bacteria, assembled from three individual databases for *Pseudomonas* (Hop's), *Ralstonia* (Rip's) and *Xanthomonas* (Xop's) (Lindeberg et al. 2005; Peeters et al. 2013; Costa et al. 2024) (Table 3). Altogether, the 31 effectors fall into 28 classes, with AvrB, RipTPS and XopD being represented twice.

The situation was much different for the SctV2-type T3SS, which was found in strains GOD-11R, KACC 21265 and Kf1. Effectidor II predicted only six orthology groups with scores higher than 0.25 (Table 4). Four orthology groups had homologs in the T3E database (HopAJ, RipA, RipL, RipTPS), and each strain had two or three of these T3E candidates.

Table 3. Prediction of type III-secreted proteins for SctV1-type T3SS of *X. ampelinus* using Effectidor II. Best hits with at least 25% coverage are shown. Candidate effectors with a corresponding PIP pattern (Table 2) are indicated. Corresponding effector classes, as revealed by BLASTP searches against a plant-pathogen T3E database, are listed along with the coverage and sequence identity.

OG	Score	Annotation	CECT7646	CFBP 1192	PIP pattern	Effector class	Cov.	Ident.
OG_1112	0.899	hypothetical protein	DFQ15_105103	NBM50_05555	Yes	RipC	100	67
OG_671	0.898	hypothetical protein, type III secretion system effector protein	DFQ15_10390	NBM50_03725		XopN	83	60
OG_2572	0.886	hypothetical protein	DFQ15_12045	NBM50_13085		HopD	75	26
OG_222	0.832	inosine-uridine preferring nucleoside hydrolase, nucleoside hydrolase	DFQ15_101225	NBM50_02130	Yes	HopQ	100	66
OG_1577	0.819	hypothetical protein, inositol hexakisphosphate	DFQ15_10984	NBM50_08090		XopH	97	64
OG_1600	0.817	hypothetical protein	DFQ15_109107	NBM50_07980		XopX	83	63
OG_576	0.801	hypothetical protein	DFQ15_102289	NBM50_01715	Yes	XopZ	90	41
OG_297	0.800	avirulence protein, hypothetical protein	DFQ15_1023	NBM50_00260		AvrB	96	64
OG_1753	0.794	hypothetical protein	DFQ15_11138	NBM50_09320		RipA	56	25

OG_1576	0.787	hypothetical protein, secretion protein EspV	DFQ15_10983	NBM50_08095		AvrA	94	62
OG_2604	0.735	hypothetical protein, type III effector HopG1	DFQ15_12111	NBM50_13615		XopAG	90	57
OG_944	0.731	hypothetical protein	DFQ15_104122	NBM50_04820		XopV	82	62
OG_1116	0.720	hypothetical protein	DFQ15_105107	NBM50_05530		/	/	/
OG_1115	0.703	hypothetical protein	DFQ15_105106	NBM50_05540		/	/	/
OG_1575	0.695	avirulence protein AvrBs1, type III secretion AvrA family effector	DFQ15_10982	NBM50_08100		AvrBs1	89	59
OG_2287	0.687	ADP-ribosyltransferase domain-containing protein, NAD:arginine ADP-ribosyltransferase	DFQ15_11681	NBM50_12035	Yes	XopAI	100	53
OG_1573	0.684	Fic family protein, leucine rich repeat (LRR) protein	DFQ15_10980	NBM50_08110		XopAC	74	32
OG_3401	0.678	C48 family peptidase		NBM50_08240		XopD	97	47
OG_864	0.673	cation transport regulator ChaC, gamma-glutamylcyclotransferase	DFQ15_10441	NBM50_05235		RipAY	45	43
OG_3308	0.671	hypothetical protein	DFQ15_1541	NBM50_17140		/	/	/
OG_1103	0.671	avirulence protein, hypothetical protein	DFQ15_10594	NBM50_05600		AvrB	75	49
OG_92	0.638	lytic murein transglycosylase	DFQ15_10194	NBM50_02800		HopAJ	96	47
OG_1546	0.549	type III effector protein XopD	DFQ15_10953			XopD	84	47
OG_906	0.453	trehalose 6-phosphate synthase, trehalose-6-phosphate synthase	DFQ15_10483	NBM50_05020 (pseudogene)		RipTPS	91	57
OG_1203	0.332	hypothetical protein	DFQ15_10675	NBM50_06370		XopL	92	33
OG_2795	0.327	hypothetical protein	DFQ15_12414	NBM50_14825	Yes	RipE	44	29
OG_3358	0.321	hypothetical protein		NBM50_04355		XopO	87	45
OG_3350	0.317	hypothetical protein		NBM50_03010		XopBG	79	42
OG_2335	0.293	YopT-type cysteine protease domain-containing protein, YopT-type cysteine protease-like protein	DFQ15_11735	NBM50_12290		HopAR	83	31
OG_1323	0.292	glycoside hydrolase family 28 protein, polygalacturonase	DFQ15_10769	NBM50_06990		/	/	/
OG_1938	0.289	E3 ligase-like protein (putative virulence factor), hypothetical protein	DFQ15_112108	NBM50_09720		RipV	64	38
OG_1158	0.281	hypothetical protein	DFQ15_10628	NBM50_06585	Yes	RipAQ	65	33
OG_1549	0.273	alpha,alpha-trehalose-phosphate synthase (UDP-forming)	DFQ15_10956	NBM50_08225		RipTPS	99	48
OG_2162	0.257	hypothetical protein, leucine rich repeat (LRR) protein	DFQ15_11536	NBM50_11465	Yes	RipAC	27	32
OG_1722	0.256	pectate lyase, pectate lyase-like protein	DFQ15_1117	NBM50_09160	Yes	RipW	61	37
OG_3346	0.254	hypothetical protein		NBM50_02550		/	/	/

Table 4. Prediction of type III-secreted proteins for SctV2-type T3SS of *Xylophilus* using Effectidor II. Best hits with at least 25% coverage are shown. Corresponding effector classes, as revealed by BLASTP searches against a plant-pathogen T3E database, are listed along with the coverage and sequence identity.

OG	Score	Annotation	GOD-11R	KACC 21265	Kf1	Effector class	Cov.	Ident.
OG_92	0.638	lytic murein transglycosylase	R9X41_13525	GT347_03245	GN316_20430	HopAJ	98	49
OG_9901	0.285	hypothetical protein			GN316_16015	RipA	83	30
OG_1549	0.273	alpha,alpha-trehalose-phosphate synthase (UDP-forming)	R9X41_19915	GT347_17110	GN316_10830	RipTPS	99	48
OG_5313	0.259	hypothetical protein	R9X41_16315			/	/	/
OG_7634	0.255	hypothetical protein		GT347_22755		/	/	/
OG_7234	0.251	hypothetical protein		GT347_17980		RipL	67	32

Confirmation and structural modeling of predicted type III effectors

We next aimed to experimentally verify the secretion signal of novel effector candidates. Based on the high Effectidor II scores, BLASTP and InterProScan analyses, we focused on three orthology groups, OG_1116, OG_1115 and OG_3308, with Effectidor II scores of 0.720, 0.703 and 0.671, respectively (Table 3). Members of the other orthology groups had a substantially lower Effectidor II score (below 0.3) and had either no homologs beyond *Xylophilus* (OG_5313, OG_7634) or well-defined homologs that are not considered to be T3Es (OG_1323, OG_3346) (Tables 3 and 4). Members of OG_1323 are annotated as a glycoside hydrolase family 28 protein or a polygalacturonase, proteins that are typically secreted by the type II secretion pathway. Indeed, SignalP predicts a Sec/SPI-type Signal Peptide (Sec/SPI) with a likelihood of 0.999 for the 629-aa protein PYE78419 with the locus ID DFQ15_10769. OG_3346 has close homologs with more than 60% sequence identity that are annotated as cytosine-specific methyltransferases and that are encoded on *Acidovorax* bacteriophages (UYL85549.1, UYL85448.1). This observation prompted us to consider that the locus NBM50_02550 may belong to a prophage. To test this, we used geNomad to assess whether this locus is part of a mobile genetic element (Camargo et al. 2024), which was indeed the case. NBM50_02550 corresponds to one out of 58 ORFs that are encoded on a predicted 46-kb prophage (region 142964 to 188749 of contig QJTC01000001.1).

For these reasons, we did not further study the orthology groups OG_5313, OG_7634, OG_1323, and OG_3346.

Three orthology groups, OG_1116 (DFQ15_105107, NBM50_05530), OG_1115 (DFQ15_105106, NBM50_05540) and OG_3308 (DFQ15_1541, NBM50_17140), corresponded to proteins that are similar to each other. However, a multiple sequence alignment revealed that the proteins belonging to OG_3308 lack the C terminus, which can be explained by the fact that the gene is located at the end of the contigs in both strains, CFBP 1192 (contig ID: JAMOFZ010000052.1) and CECT7646 (contig ID: QJTC01000054.1). The full-length versions of OG_1116 and OG_1115 have amino acids sequences that are 58% identical to each other. The C-terminally truncated version of OG_3308 is 54% identical to OG_1116 and 80% identical to OG_1115.

In total, there are twelve homologous protein variants deposited at NCBI, corresponding to the two strains, CECT7646 and CFBP 1192, each with three paralogs, and each paralog either annotated by the submitters and re-annotated by NCBI (RefSeq versions predicted with GeneMarkS-2+). Their multiple sequence alignment revealed that three different translation start codons were used (Supplementary Data S3). Based on sequence comparisons and the presence of a putative Shine-Dalgarno sequence, GGAG, with a distance of six base pairs to the ATG translation initiation codon, we considered the protein versions starting with MKDLYPF (Met-Lys-Asp-Leu-Tyr-Pro-Phe) as the most likely ones. The G+C content of the coding sequences is 56.7% for OG_1115, 56.9% for OG_1116 and 53.8 for OG_3308, compared to an overall G+C content of the genome of 67.8%. Such atypical G+C content can be taken as evidence for horizontal gene transfer, a process that is known to be involved in shaping the T3E repertoires of bacterial pathogens (Dillon et al. 2019).

We employed a reporter fusion system where the candidate effector's type III secretion signal is translationally fused the defense-triggering domain of AvrBs1, a T3E from *Xanthomonas campestris* (Zhao et al. 2013). Our chimeric constructs consisted of the 100 base pairs upstream of the most likely translation initiation codon and the first 100 codons of the CDS, fused to the *avrBs1* CDS covering codons 59 to 445. When introduced into the *X. campestris* reporter strain 8004Δ*avrBs1*, bacteria expressing AvrBs1 reporter fusions triggered a *Bs1*-dependent hypersensitive response in pepper plants of the cultivar ECW-10R (Figure 4). In contrast, plants

showed no reaction upon infiltration with bacteria that expressed the chimeric proteins but were mutated in the essential T3SS gene *hrcV*, thus demonstrating that the N-terminal 100 amino acids functioned as a type III secretion/translocation signal.

To gain insight into the function and conservation of the novel T3Es, we used BLASTP to identify homologous sequences in the ClusteredNR database at NCBI GenBank. Sequences corresponding to OG_1115 and OG_1116 had 70 unique hits with an e-value lower than 10^{-10} (Supplementary Table S5). Most of these sequences were annotated as hypothetical proteins and they originated from *Pseudomonas savastanoi*, *Pseudomonas amygdali*, *Acidovorax* sp., *Pseudomonas syringae*, *Xanthomonas fragariae*, *Xanthomonas hydrangeae*, *Paraburkholderia aspalathi*, *Ralstonia* sp. and *Bradyrhizobium* sp., thus including both pathogenic as well as symbiotic bacteria. Notably, a few of the hits were annotated as C48 family peptidase or Ulp1 family isopeptidase. We then applied CD-Hit with a sequence identity threshold of 0.5 and a word size of 2 to cluster these sequences and established a set of 21 representative sequences. For functional predictions, these sequences were scanned at InterProScan (Supplementary Table S6). Five of the sequences had hits with InterPro entries corresponding to the Ulp1 protease family. However, the portions of OG_1115 and OG_1116 that aligned with these sequences did not correspond to the Ulp1 protease motifs.

To obtain further insight into the molecular function of the novel T3Es, we modeled the 3D structure of the two full-length versions with AlphaFold3 (Abramson et al. 2024) (Supplementary Figure S3). The predicted template modeling (pTM) scores were 0.39 for OG_1115 and 0.36 for OG_1116. Notably, the N-termini were largely predicted to be disordered (135 aa for OG_1115 and 109 aa for OG_1116), as was the case for the InterProScan predictions (regions 24-67 and 88-124 for OG_1115, region 16-116 for OG_1116), or they would include alpha-helices with very low confidence (pIDDT scores < 50). Structural predictions focusing on the apparently folded C-terminal domains, $OG_1115_{149-636}$ and $OG_1116_{117-598}$, improved the overall confidence of the predictions slightly to 0.42 and 0.40, respectively.

The predicted models were then used to find structural homologs, using FoldSeek (van Kempen et al. 2024). This way, a few T3Es (e.g., HopAV1 from *Pseudomonas amygdali*, HsvG and HsvB from *Pantoea agglomerans*), but also several ubiquitin-like protease domain-containing proteins

from rhizobia (*Mesorhizobium*, *Bradyrhizobium*, *Rhizobium* and *Sinorhizobium*) were found (Supplementary Table S7). Most hits in other plants pathogens beyond *Xylophilus* had lower e-values (between 2.3 x 10^{-59} and 1.7 x 10^{-11}) than hits in symbiotic bacteria (e-values above 3.6 x 10^{-12}). Thus, additional experimental work is needed to better elucidate the function(s) of these newly identified effectors.

DISCUSSION

Based on genomic data, we aimed to perform a genus-wide prediction of type III secretion and flagellar systems in *Xylophilus*, including the grapevine pathogen *Xylophilus ampelinus*. First, we curated the data available in NCBI to resolve taxonomic uncertainties. We found 19 genome sequences at NCBI GenBank, six of which were annotated as *X. ampelinus* and one of which was annotated as *X. rhododendri*. Based on average nucleotide identities, however, only two strains (CECT7646, CFBP 1192) corresponding to the same original isolate belong to the species *X. ampelinus*, while the genomes of BgEED09, bin23, CCH5-B3, and ERR594292 do not belong to this species. Based on ANI and DDH, isolates BgEED09, CCH5-B3 and ERR594292 do not even belong to the genus *Xylophilus*, since they are genetically closer to three type strains of the genus *Xenophilus* (Blümel et al. 2001; Kim et al. 2010; Li et al. 2014). Only the isolate bin23 appears to belong to the genus *Xylophilus* and may represent a new, yet uncharacterized species within this genus. Altogether, the remaining 16 genome sequences would correspond to ten different species, among which only two, *X. ampelinus* and *X. rhododendri*, have been formally described (Willems et al. 1987b; Lee et al. 2020).

Since many Gram-negative pathogenic and symbiotic bacteria rely on a T3SS and its set of secreted effectors for their interaction with eukaryotic organism, we scrutinized all twelve good-quality genome sequences of the genus *Xylophilus* for the presence of such a protein secretion system and the evolutionary related flagellar system. Firstly, we found three different flagellar systems in the genus, one of which was present in all twelve strains and is thus vertically inherited. The other two flagellar systems were only detected in one or two strains, and it remains to be analyzed whether the systems were lost in the other strains of the genus or if they were acquired by horizontal gene transfer in this subset of strains.

Based on sequence similarities and supported by synteny, we identified two types of T3SSs, both of which are classified as Hrp2/Hrc2-type T3SSs. Genes of this secretion system are under transcriptional control of the AraC-type transcriptional activator HrpX in *Acidovorax* and *Xanthomonas* and its ortholog HrpB in *Burkholderia* and *Ralstonia* (Genin et al. 1992; Wengelnik et al. 1996). We identified homologs of HrpX/HrpB in the T3SS gene clusters of *Xylophilus*, and we also identified the HrpX/HrpB-binding motif, the *hrpII*/PIP box (Cunnac et al. 2004a; Koebnik et al. 2006), in the promoters of several T3SS operons in *Xylophilus*, suggesting a similar regulatory mechanism to that in the well-studied plant pathogens mentioned above. However, we were unable to identify a bona fide ortholog of the upstream regulatory component HrpG (Teper et al. 2021), although each of the various *Xylophilus* genomes encodes dozens of response regulators. Notably, we found two paralogs of HrpB/HrpX genetically linked to the T3SS gene cluster in *X. ampelinus*, one of which is encoded immediately upstream of *sctC* and the other, which has 53% sequence identity, is encoded seven genes downstream of *hpaB*. It remains to be determined how transcription of the *hrpB*/*hrpX* genes is controlled in *Xylophilus*.

T3Es, which are secreted by the T3SS, need assistance in entering the eukaryotic target cell. For Hrp2/Hrc2-type T3SS, two types of putative translocons have been described, HrpF/NolX and HpaT (Büttner et al. 2002; Pesce et al. 2017). The gene upstream of the *X. ampelinus hrpB1* gene appears to encode a protein that is distantly related to the putative translocon HrpF/NolX, as supported by PSI-BLAST analysis and structural modeling (data not shown). However, we were unable to identify a putative translocon protein for the other T3SS, which is present in GOD11-R, Kf1, and KACC 21265, and it remains to be determined whether this T3SS injects effector proteins into eukaryotic target cells.

The presence of HrpB/HrpX homologs prompted us to predict, using a very conservative approach, genes that are very likely under transcriptional control of HrpB/HrpX and thus co-regulated with the T3SS genes. The power of this approach is underlined by the fact that four out of 16 predicted promoters in *X. ampelinus* control the expression of components or helpers of the T3SS (Hpa2, Hpa3, SctD, SctU) and another ten promoters control candidate T3E genes (*hopQ*, *ripC*, *ripE*, *ripW*, *ripAC*, *ripAQ*, *xopD*, *xopZ*, *xopAI*, *xopBG*). Two additional patterns were found within a CDS or in the antisense direction of a gene and likely represent false positives. For

X. rhododendri, 35 HrpB/HrpX-regulated promoters were predicted, among which three are found in a region of 1,040 bp upstream of the *sctQ* CDS and two belong to genes with remote sequence homology to T3Es (XopN, RipL). Surprisingly, the remaining thirty promoters would all control the expression of hypothetical proteins, most of them without homologs beyond *X. rhododendri* and many of them belonging to gene families with several, sometimes distantly related, paralogs. It is unknown whether they correspond to T3Es or if they are involved in the interaction with other organisms.

Using a pan-genomic AI-based algorithm, we predicted candidate T3Es for *X. ampelinus* and *X. rhododendri*. In this way, we identified 31 homologs of Hop, Rip and Xop proteins from *Pseudomonas*, *Ralstonia* and *Xanthomonas* in *X. ampelinus*, which collectively form a large repertoire of candidate T3Es. This situation resembles other highly virulent plant-pathogenic bacteria, which use their effector repertoires, also called the effectome or effectorome, to successfully infect their host plants (Landry et al. 2020; Bundalovic-Torma et al. 2022). By contrast, we only predicted four Hop or Rip homologs in *X. rhododendri*, which raises questions about its pathogenic potential. Indeed, strain KACC 21265 was isolated from a rhododendron flower and there is no evidence that it would cause symptoms on rhododendron or other plants.

Three of the predicted T3Es of *X. ampelinus*, for which no homolog was found in T3E databases of *Pseudomonas*, *Ralstonia* and *Xanthomonas*, were found to possess a functional type III secretion signal that enabled the secretion and translocation of an AvrBs1 reporter protein from *Xanthomonas* into pepper plants. This shows that heterologous expression in a gamma proteobacterium can be used to validate candidate effectors from a beta proteobacterium. There are few examples for such trans-class secretion experiments. For instance, *Xanthomonas euvesicatoria* pv. *euvesicatoria* (syn. *X. campestris* pv. *vesicatoria*) was able to secrete PopA, a type III-secreted protein from *R. solanacearum* (Rossier et al. 1999). While these secretion experiments were tested for secretion using the same class of T3SS, here the Hrp2/Hrc2 class, other studies have shown that T3Es can even be secreted by a different T3SS class, at least when overexpressed. For instance, *X. euvesicatoria* pv. *euvesicatoria* could secrete AvrB, a Hrc1-class type III-secreted avirulence protein from *P. syringae* pv. *glycinea*, and YopE, a Ysc-class type III-secreted cytotoxin of the mammalian pathogen *Yersinia pseudotuberculosis* (Rossier et al. 1999). Similarly, Olaf Schneewind and colleagues reported that the Hrc1-class T3SS of *Erwinia*

chrysanthemi recognized the secretion signals of the Ysc-class secreted proteins YopE and YopQ from *Yersinia*, and that the Hrc1-class secreted avirulence proteins AvrB and AvrPto from *P. syringae* could be secreted by the Ysc-class T3SS of *Yersinia* (Anderson et al. 1999). Such successful reciprocal secretion of proteins by different classes of T3SSs may suggest universal recognition of secretion signals. However, due to the lack of systematic studies and the fact that unsuccessful examples of heterologous secretion attempts are unlikely to be published, it remains unclear whether these examples are the exception or the rule. The discovery of effector- or effector-class-specific chaperones and translocation motifs suggests that effectors may not be universally recognized or translocated (Scheibner et al. 2018; Drehkopf et al. 2022; Pintor et al. 2025). Nevertheless, even if reciprocal secretion is not the rule, it provides a useful and straightforward strategy to test candidate effectors, as shown in this study.

The three candidate effectors from *X. ampelinus*, which were experimentally confirmed to contain a functional N-terminal type III secretion/translocation signal, share structural similarity with proteins of other plant-associated bacteria, including both pathogens (*Acidovorax* sp., *Pantoea agglomerans*, *P. amygdali*, *P. syringae*, *Ralstonia* sp., *Xanthomonas fragariae*) as well as symbionts (*Bradyrhizobium* sp.). Some of the proteins with similar structure are annotated as Ulp1 family isopeptidases or C48 family peptidases, known to be T3Es. Yet, there is no evidence that the *X. ampelinus* effectors have such a peptidase activity because the peptidase domains of these effectors do not align with the *X. ampelinus* amino acid sequence. The other homologs, which are annotated as hypothetical proteins, might also correspond to T3Es. Further studies need to determine how these effectors contribute to the infection of grapevine plants. Even more work is needed to understand the function of the other Hrp2/Hrc2-class T3SS, its candidate effectors and co-regulated gene families, as exemplified by the *X. rhododendri* strain KACC 21265.

ACKNOWLEDGEMENTS

We thank Sebastien Santini (CNRS/AMU IGS UMR7256, France) and the PACA Bioinfo platform for the availability and management of the phylogeny.fr website used to generate phylogenetic trees (Figure 1, Supplementary Figure S1).

This work was supported by the Chateaubriand fellowships program of the French government to NW. NW was supported in part by a fellowship from the Edmond J. Safra Center for Bioinformatics at Tel Aviv University.

LITERATURE CITED

- Abby SS, Rocha EP (2012). The non-flagellar type III secretion system evolved from the bacterial flagellum and diversified into host-cell adapted systems. PLoS Genet. 8: e1002983. doi: 10.1371/journal.pgen.1002983
- Abby SS, Denise R, Rocha EPC (2024). Identification of protein secretion systems in bacterial genomes using MacSyFinder version 2. Methods Mol. Biol. 2715: 1-25. doi: 10.1007/978-1-0716-3445-5_1
- Abramson J, Adler J, Dunger J, Evans R, Green T, Pritzel A, Ronneberger O, Willmore L, Ballard AJ, Bambrick J, Bodenstein SW, Evans DA, Hung CC, O'Neill M, Reiman D, Tunyasuvunakool K, Wu Z, Žemgulytė A, Arvaniti E, Beattie C, Bertolli O, Bridgland A, Cherepanov A, Congreve M, Cowen-Rivers AI, Cowie A, Figurnov M, Fuchs FB, Gladman H, Jain R, Khan YA, Low CMR, Perlin K, Potapenko A, Savy P, Singh S, Stecula A, Thillaisundaram A, Tong C, Yakneen S, Zhong ED, Zielinski M, Žídek A, Bapst V, Kohli P, Jaderberg M, Hassabis D, Jumper JM (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature 630: 493-500. doi:10.1038/s41586-024-07487-w
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990). Basic local alignment search tool. J. Mol. Biol. 215: 403-410. doi: 10.1016/S0022-2836(05)80360-2
- Anderson DM, Fouts DE, Collmer A, Schneewind O (1999). Reciprocal secretion of proteins by the bacterial type III machines of plant and animal pathogens suggests universal recognition of mRNA targeting signals. Proc. Natl. Acad. Sci. U. S. A. 96: 12839-12843. doi: 10.1073/pnas.96.22.12839
- Anguita-Maeso M, Ares-Yebra A, Haro C, Román-Écija M, Olivares-García C, Costa J, Marco-Noales E, Ferrer A, Navas-Cortés JA, Landa BB (2022). *Xylella fastidiosa* Infection reshapes microbial composition and network associations in the Xylem of almond trees. Front. Microbiol. 13: 866085. doi: 10.3389/fmicb.2022.866085

- Bai Y, Müller DB, Srinivas G, Garrido-Oter R, Potthoff E, Rott M, Dombrowski N, Münch PC, Spaepen S, Remus-Emsermann M, Hüttel B, McHardy AC, Vorholt JA, Schulze-Lefert P (2015). Functional overlap of the *Arabidopsis* leaf and root microbiota. Nature 528: 364-369. doi: 10.1038/nature16192
- Bi D, Liu L, Tai C, Deng Z, Rajakumar K, Ou HY (2013). SecReT4: a web-based bacterial type IV secretion system resource. Nucleic Acids Res. 41: D660-D665. doi: 10.1093/nar/gks1248
- Biggs BW, Price MN, Lai D, Escobedo J, Fortanel Y, Huang YY, Kim K, Trotter VV, Kuehl JV, Lui LM, Chakraborty R, Deutschbauer AM, Arkin AP (2024). High-throughput protein characterization by complementation using DNA barcoded fragment libraries. Mol. Syst. Biol. 20: 1207-1229. doi: 10.1038/s44320-024-00068-z
- Blümel S, Busse HJ, Stolz A, Kämpfer P (2001). *Xenophilus azovorans* gen. nov., sp. nov., a soil bacterium that is able to degrade azo dyes of the Orange II type. Int. J. Syst. Evol. Microbiol. 51: 1831-1837. doi: 10.1099/00207713-51-5-1831
- Bundalovic-Torma C, Lonjon F, Desveaux D, Guttman DS (2022). Diversity, evolution, and function of *Pseudomonas syringae* effectoromes. Annu. Rev. Phytopathol. 60: 211-236. doi: 10.1146/annurev-phyto-021621-121935
- Büttner D (2012). Protein export according to schedule: architecture, assembly, and regulation of type III secretion systems from plant- and animal-pathogenic bacteria. Microbiol. Mol. Biol. Rev. 76: 262-310. doi: 10.1128/MMBR.05017-11
- Büttner D (2016). Behind the lines – actions of bacterial type III effector proteins in plant cells. FEMS Microbiol. Rev. 40: 894-937. doi: 10.1093/femsre/fuw026
- Büttner D, Nennstiel D, Klüsener B, Bonas U (2002). Functional analysis of HrpF, a putative type III translocon protein from *Xanthomonas campestris* pv. *vesicatoria*. J. Bacteriol. 184: 2389-2398. doi: 10.1128/JB.184.9.2389-2398.2002
- Camargo AP, Roux S, Schulz F, Babinski M, Xu Y, Hu B, Chain PSG, Nayfach S, Kyrpides NC (2024). Identification of mobile genetic elements with geNomad. Nat. Biotechnol. 42:1303-1312. doi: 10.1038/s41587-023-01953-y
- Chacón FI, Sineli PE, Mansilla FI, Pereyra MM, Diaz MA, Volentini SI, Poehlein A, Meinhardt F, Daniel R, Dib JR (2022). Native cultivable bacteria from the blueberry microbiome as novel potential biocontrol agents. Microorganisms 10: 969. doi: 10.3390/microorganisms10050969
- Chun J, Oren A, Ventosa A, Christensen H, Arahal DR, da Costa MS, Rooney AP, Yi H, Xu XW, De Meyer S, Trujillo ME (2018). Proposed minimal standards for the use of genome data for

the taxonomy of prokaryotes. Int. J. Syst. Evol. Microbiol. 68: 461-466. doi: 10.1099/ijsem.0.002516

- Cornelis GR (2002). The *Yersinia* Ysc-Yop 'type III' weaponry. Nat. Rev. Mol. Cell Biol. 3: 742-752. doi: 10.1038/nrm932
- Cornelis GR (2006). The type III secretion injectisome. Nat. Rev. Microbiol. 4: 811-825. doi: 10.1038/nrmicro1526
- Costa J, Pothier JF, Bosis E, Boch J, Kölliker R, Koebnik R (2024). A community-curated DokuWiki resource on diagnostics, diversity, pathogenicity, and genetic control of xanthomonads. Mol. Plant Microbe Interact. 37: 347-353. doi:10.1094/MPMI-11-23-0184-FI
- Cunnac S, Boucher C, Genin S (2004a). Characterization of the *cis*-acting regulatory element controlling HrpB-mediated activation of the type III secretion system and effector genes in *Ralstonia solanacearum*. J. Bacteriol. 186: 2309-2318. doi: 10.1128/JB.186.8.2309-2318.2004
- Cunnac S, Occhialini A, Barberis P, Boucher C, Genin S (2004b). Inventory and functional analysis of the large Hrp regulon in *Ralstonia solanacearum*: identification of novel effector proteins translocated to plant host cells through the type III secretion system. Mol. Microbiol. 53: 115-128. doi: 10.1111/j.1365-2958.2004.04118.x
- Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, Dufayard JF, Guindon S, Lefort V, Lescot M, Claverie JM, Gascuel O (2008). Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res. 36: W465-W469. doi: 10.1093/nar/gkn180
- Diepold A, Armitage JP (2015). Type III secretion systems: the bacterial flagellum and the injectisome. Philos. Trans. R. Soc. Lond. B Biol. Sci. 370: 20150020. doi: 10.1098/rstb.2015.0020
- Dillon MM, Almeida RND, Laflamme B, Martel A, Weir BS, Desveaux D, Guttman DS (2019). Molecular evolution of *Pseudomonas syringae* type III secreted effector proteins. Front. Plant Sci. 10: 418. doi: 10.3389/fpls.2019.00418
- Drehkopf S, Otten C, Büttner D (2022). Recognition of a translocation motif in the regulator HpaA from *Xanthomonas euvesicatoria* is controlled by the type III secretion chaperone HpaB. Front. Plant Sci. 13: 955776. doi: 10.3389/fpls.2022.955776
- Durfee T, Nelson R, Baldwin S, Plunkett G 3rd, Burland V, Mau B, Petrosino JF, Qin X, Muzny DM, Ayele M, Gibbs RA, Csörgo B, Pósfai G, Weinstock GM, Blattner FR (2008). The complete genome sequence of *Escherichia coli* DH10B: insights into the biology of a laboratory workhorse. J. Bacteriol. 190: 2597-2606. doi:10.1128/JB.01695-07

- Edgar RC (2004). MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics 5: 113. doi: 10.1186/1471-2105-5-113
- EFSA Panel on Plant Health (2014). Scientific Opinion on the pest categorisation of *Xylophilus ampelinus* (Panagopoulos) Willems et al. EFSA Journal 12: 3921. doi: 10.2903/j.efsa.2014.3921
- EPPO (2009). PM 7/96 (1): *Xylophilus ampelinus*. Bulletin OEPP/EPPO Bulletin 39: 403-412. doi: 10.1111/j.1365-2338.2009.02328.x
- EPPO (2025). *Xylophilus ampelinus* (XANTAM). EPPO Global Database. Available online: https://gd.eppo.int (accessed on 29 July 2025).
- Figurski DH, Helinski DR (1979). Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in trans. Proc. Natl. Acad. Sci. U. S. A. 76: 1648-1652. doi:10.1073/pnas.76.4.1648
- Freese HM, Meier-Kolthoff JP, Sardà Carbasse J, Afolayan AO, Göker M (2026). TYGS and LPSN in 2025: a Global Core Biodata Resource for genome-based classification and nomenclature of prokaryotes within DSMZ Digital Diversity. Nucleic Acids Res. 54: D884-D891. doi: 10.1093/nar/gkaf1110
- Furutani A, Nakayama T, Ochiai H, Kaku H, Kubo Y, Tsuge S (2006). Identification of novel HrpXo regulons preceded by two cis-acting elements, a plant-inducible promoter box and a -10 box-like sequence, from the genome database of *Xanthomonas oryzae* pv. *oryzae*. FEMS Microbiol. Lett. 259: 133-141. doi: 10.1111/j.1574-6968.2006.00265.x
- Galaxy Community (2024). The Galaxy platform for accessible, reproducible, and collaborative data analyses: 2024 update. Nucleic Acids Res. 52: W83-W94. doi: 10.1093/nar/gkae410
- Gazi AD, Sarris PF, Fadouloglou VE, Charova SN, Mathioudakis N, Panopoulos NJ, Kokkinidis M (2012). Phylogenetic analysis of a gene cluster encoding an additional, rhizobial-like type III secretion system that is narrowly distributed among *Pseudomonas syringae* strains. BMC Microbiol. 12: 188. doi: 10.1186/1471-2180-12-188
- Genin S, Gough CL, Zischek C, Boucher CA (1992). Evidence that the *hrpB* gene encodes a positive regulator of pathogenicity genes from *Pseudomonas solanacearum*. Mol. Microbiol. 6: 3065-3076.
- Gilchrist CLM, Chooi YH (2021). clinker & clustermap.js: automatic generation of gene cluster comparison figures. Bioinformatics 37: 2473-2475. doi: 10.1093/bioinformatics/btab007

- Goyal A, Bittleston LS, Leventhal GE, Lu L, Cordero OX (2022). Interactions between strains govern the eco-evolutionary dynamics of microbial communities. Elife 11: e74987. doi: 10.7554/eLife.74987
- Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S (2018). High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. Nat. Commun. 9: 5114. doi: 10.1038/s41467-018-07641-9
- Jiménez-Guerrero I, Pérez-Montaño F, Da Silva GM, Wagner N, Shkedy D, Zhao M, Pizarro L, Bar M, Walcott R, Sessa G, Pupko T, Burdman S (2020). Show me your secret(ed) weapons: a multifaceted approach reveals a wide arsenal of type III-secreted effectors in the cucurbit pathogenic bacterium *Acidovorax citrulli* and novel effectors in the *Acidovorax* genus. Mol. Plant Pathol. 21: 17-37. doi: 10.1111/mpp.12877
- Jones P, Binns D, Chang HY, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, Pesseat S, Quinn AF, Sangrador-Vegas A, Scheremetjew M, Yong SY, Lopez R, Hunter S (2014). InterProScan 5: genome-scale protein function classification. Bioinformatics 30: 1236-1240. doi: 10.1093/bioinformatics/btu031
- Kim SJ, Kim YS, Weon HY, Anandham R, Noh HJ, Kwon SW (2010). *Xenophilus aerolatus* sp. nov., isolated from air. Int. J. Syst. Evol. Microbiol. 60: 327-330. doi: 10.1099/ijs.0.013185-0
- Koebnik R, Krüger A, Thieme F, Urban A, Bonas U (2006). Specific binding of the *Xanthomonas campestris* pv. *vesicatoria* AraC-type transcriptional activator HrpX to plant-inducible promoter boxes. J. Bacteriol. 188: 7652-7660. doi:10.1128/JB.00795-06
- Kousik CS, Ritchie DF (1999). Development of bacterial spot on near-isogenic lines of bell pepper carrying gene pyramids composed of defeated major resistance genes. Phytopathology 89: 1066-1072. doi: 10.1094/PHYTO.1999.89.11.1066
- Kovach ME, Elzer PH, Hill DS, Robertson GT, Farris MA, Roop RM 2nd, Peterson KM (1995). Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. Gene 166: 175-176. doi:10.1016/0378-1119(95)00584-1
- Kvitko BH, Collmer A (2023). Discovery of the Hrp type III secretion system in phytopathogenic bacteria: how investigation of hypersensitive cell death in plants led to a novel protein injector system and a world of inter-organismal molecular interactions within plant cells. Phytopathology 113: 626-636. doi: 10.1094/PHYTO-08-22-0292-KD

- Landry D, González-Fuente M, Deslandes L, Peeters N (2020). The large, diverse, and robust arsenal of *Ralstonia solanacearum* type III effectors and their in planta functions. Mol. Plant Pathol. 21: 1377-1388. doi: 10.1111/mpp.12977
- Lee SA, Heo J, Kim TW, Sang MK, Song J, Kwon SW, Weon HY (2020). *Xylophilus rhododendri* sp. nov., isolated from flower of royal azalea, *Rhododendron schlippenbachii*. Curr. Microbiol. 77: 4160-4166. doi: 10.1007/s00284-020-02194-x
- Letunic I, Bork P (2024). Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. Nucleic Acids Res. 52: W78-W82. doi: 10.1093/nar/gkae268
- Li W, Jaroszewski L, Godzik A (2001). Clustering of highly homologous sequences to reduce the size of large protein databases. Bioinformatics 17: 282-283. doi: 10.1093/bioinformatics/17.3.282
- Li QF, Sun LN, Kwon SW, Chen Q, He J, Li SP, Zhang J (2014). *Xenophilus arseniciresistens* sp. nov., an arsenite-resistant bacterium isolated from soil. Int. J. Syst. Evol. Microbiol. 64: 1926-1931. doi: 10.1099/ijs.0.061515-0
- Lindeberg M, Stavrinides J, Chang JH, Alfano JR, Collmer A, Dangl JL, Greenberg JT, Mansfield JW, Guttman DS (2005). Proposed guidelines for a unified nomenclature and phylogenetic analysis of type III Hop effector proteins in the plant pathogen *Pseudomonas syringae*. Mol. Plant Microbe Interact. 18: 275-282. doi: 10.1094/MPMI-18-0275
- Lipscomb L, Schell MA (2011). Elucidation of the regulon and cis-acting regulatory element of HrpB, the AraC-type regulator of a plant pathogen-like type III secretion system in *Burkholderia pseudomallei*. J. Bacteriol. 193: 1991-2001. doi: 10.1128/JB.01379-10
- Meier-Kolthoff JP, Auch AF, Klenk HP, Göker M (2013). Genome sequence-based species delimitation with confidence intervals and improved distance functions. BMC Bioinformatics 14: 60. doi: 10.1186/1471-2105-14-60
- Needleman SB, Wunsch CD (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins. J. Mol. Biol. 48: 443-453. doi: 10.1016/0022-2836(70)90057-4
- Noël L, Thieme F, Nennstiel D, Bonas U (2001). cDNA-AFLP analysis unravels a genome-wide *hrpG*-regulon in the plant pathogen *Xanthomonas campestris* pv. *vesicatoria*. Mol. Microbiol. 41: 1271-1281. doi: 10.1046/j.1365-2958.2001.02567.x
- Occhialini A, Cunnac S, Reymond N, Genin S, Boucher C (2005). Genome-wide analysis of gene expression in *Ralstonia solanacearum* reveals that the *hrpB* gene acts as a regulatory

switch controlling multiple virulence pathways. Mol. Plant Microbe Interact. 18: 938-949. doi: 10.1094/MPMI-18-0938

- Peeters N, Carrère S, Anisimova M, Plener L, Cazalé AC, Genin S (2013). Repertoire, unified nomenclature and evolution of the Type III effector gene set in the *Ralstonia solanacearum* species complex. BMC Genomics 14: 859. doi: 10.1186/1471-2164-14-859
- Peros JP, Berger G, Ride M (1995). Effect of grapevine cultivar, strain of *Xylophilus ampelinus* and culture medium on in vitro development of bacterial necrosis. Vitis 34: 189-190. doi: 10.5073/vitis.1995.34.189-190
- Pesce C, Jacobs JM, Berthelot E, Perret M, Vancheva T, Bragard C, Koebnik R (2017). Comparative genomics identifies a novel conserved protein, HpaT, in proteobacterial type III secretion systems that do not possess the putative translocon protein HrpF. Front. Microbiol. 8: 1177. doi: 10.3389/fmicb.2017.01177
- Pintor K, Balinovic A, Malengo G, Klepin C, Fleck M, Glatter T, Endesfelder U, Diepold A (2025). In situ analysis of type III secretion chaperone proteins indicates a cytosolic handover of virulence effectors. FEMS Microbes 6: xtaf010. doi: 10.1093/femsmc/xtaf010
- Portier P, Taghouti G, Bertrand PE, Briand M, Dutrieux C, Lathus A, Fischer-Le Saux M (2022). Analysis of the diversity of *Xylophilus ampelinus* strains held in CIRM-CFBP reveals a strongly homogenous species. Microorganisms 10: 1531. doi: 10.3390/microorganisms10081531
- Rodriguez-R LM, Gunturu S, Harvey WT, Rosselló-Mora R, Tiedje JM, Cole JR, Konstantinidis KT (2018). The Microbial Genomes Atlas (MiGA) webserver: taxonomic and gene diversity analysis of Archaea and Bacteria at the whole genome level. Nucleic Acids Res. 46: W282-W288. doi: 10.1093/nar/gky467
- Rossier O, Wengelnik K, Hahn K, Bonas U (1999). The *Xanthomonas* Hrp type III system secretes proteins from plant and mammalian bacterial pathogens. Proc. Natl. Acad. Sci. U. S. A. 96: 9368-9373. doi: 10.1073/pnas.96.16.9368
- Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, Barrell B (2000). Artemis: sequence visualization and annotation. Bioinformatics 16: 944-945. doi: 10.1093/bioinformatics/16.10.944
- Sayers EW, Beck J, Bolton EE, Brister JR, Chan J, Connor R, Feldgarden M, Fine AM, Funk K, Hoffman J, Kannan S, Kelly C, Klimke W, Kim S, Lathrop S, Marchler-Bauer A, Murphy TD, O'Sullivan C, Schmieder E, Skripchenko Y, Stine A, Thibaud-Nissen F, Wang J, Ye J, Zellers E,

Schneider VA, Pruitt KD (2025). Database resources of the National Center for Biotechnology Information in 2025. Nucleic Acids Res. 53: D20-D29. doi: 10.1093/nar/gkae979

- Schneider D, Zühlke D, Poehlein A, Riedel K, Daniel R (2021). Metagenome-assembled genome sequences from different wastewater treatment stages in Germany. Microbiol. Resour. Announc. 10: e0050421. doi: 10.1128/MRA.00504-21
- Schreiber KJ, Chau-Ly IJ, Lewis JD (2021). What the wild things do: mechanisms of plant host manipulation by bacterial type III-secreted effector proteins. Microorganisms 9: 1029. doi: 10.3390/microorganisms9051029
- Scheibner F, Hartmann N, Hausner J, Lorenz C, Hoffmeister AK, Büttner D (2018). The type III secretion chaperone HpaB controls the translocation of effector and noneffector proteins from *Xanthomonas campestris* pv. *vesicatoria*. Mol. Plant Microbe Interact. 31: 61-74. doi: 10.1094/MPMI-06-17-0138-R
- Shimony Y, Dotan E, Wygoda E, Wagner N, Lyubman I, Ecker N, Durante G, Mishan G, Chang JH, Avram O, Pupko T (2025). M1CR0B1AL1Z3R 2.0: an enhanced web server for comparative analysis of bacterial genomes at scale. Nucleic Acids Res. 53: W369-W375. doi: 10.1093/nar/gkaf413
- Soto-Giron MJ, Rodriguez-R LM, Luo C, Elk M, Ryu H, Hoelle J, Santo Domingo JW, Konstantinidis KT (2016). Biofilms on hospital shower hoses: characterization and implications for nosocomial infections. Appl. Environ. Microbiol. 82: 2872-2883. doi: 10.1128/AEM.03529-15
- Su P, Wicaksono WA, Li C, Michl K, Berg G, Wang D, Xiao Y, Huang R, Kang H, Zhang D, Cernava T, Liu Y (2022). Recovery of metagenome-assembled genomes from the phyllosphere of 110 rice genotypes. Sci. Data 9: 254. doi: 10.1038/s41597-022-01320-7
- Tampakaki AP (2014). Commonalities and differences of T3SSs in rhizobia and plant pathogenic bacteria. Front. Plant Sci. 5: 114. doi: 10.3389/fpls.2014.00114
- Tamura N, Murata Y, Mukaihara T (2005). Isolation of *Ralstonia solanacearum hrpB* constitutive mutants and secretion analysis of *hrpB*-regulated gene products that share homology with known type III effectors and enzymes. Microbiology (Reading) 151: 2873-2884. doi: 10.1099/mic.0.28161-0
- Teper D, Pandey SS, Wang N (2021). The HrpG/HrpX regulon of xanthomonads – an insight to the complexity of regulation of virulence traits in phytopathogenic bacteria. Microorganisms 9: 187. doi: 10.3390/microorganisms9010187

- Teufel F, Almagro Armenteros JJ, Johansen AR, Gíslason MH, Pihl SI, Tsirigos KD, Winther O, Brunak S, von Heijne G, Nielsen H (2022). SignalP 6.0 predicts all five types of signal peptides using protein language models. Nat. Biotechnol. 40: 1023-1025. doi: 10.1038/s41587-021-01156-3
- Teulet A, Camuel A, Perret X, Giraud E (2022). The versatile roles of type III secretion systems in rhizobium-legume symbioses. Annu. Rev. Microbiol. 76: 45-65. doi: 10.1146/annurev-micro-041020-032624
- Thakur M, Brooksbank C, Finn RD, Firth HV, Foreman J, Freeberg M, Gurwitz KT, Harrison M, Hulcoop D, Hunt SE, R Leach A, Levchenko M, Marques D, McDonagh EM, Mithani A, Parkinson H, Perez-Riverol Y, Perova Z, Sarkans U, Tirunagari S, Tzampatzopoulou E, Venkatesan A, Vizcaino JA, Wingfield B, Zdrazil B, McEntyre J (2025). EMBL's European Bioinformatics Institute (EMBL-EBI) in 2024. Nucleic Acids Res. 53: D10-D19. doi: 10.1093/nar/gkae1089
- Tsuge S, Terashima S, Furutani A, Ochiai H, Oku T, Tsuno K, Kaku H, Kubo Y (2005). Effects on promoter activity of base substitutions in the *cis*-acting regulatory element of HrpXo regulons in *Xanthomonas oryzae* pv. *oryzae*. J. Bacteriol. 187: 2308-2314. doi: 10.1128/JB.187.7.2308-2314.2005
- van Kempen M, Kim SS, Tumescheit C, Mirdita M, Lee J, Gilchrist CLM, Söding J, Steinegger M (2024). Fast and accurate protein structure search with Foldseek. Nat. Biotechnol. 42: 243-246. doi:10.1038/s41587-023-01773-0
- Wagner N, Baumer E, Lyubman I, Shimony Y, Bracha N, Martins L, Potnis N, Chang JH, Teper D, Koebnik R, Pupko T (2025). Effectidor II: a pan-genomic AI-based algorithm for the prediction of type III secretion system effectors. Bioinformatics 41: btaf272. doi: 10.1093/bioinformatics/btaf272
- Wagner S, Diepold A (2020). A unified nomenclature for injectisome-type type III secretion systems. Curr. Top. Microbiol. Immunol. 427: 1-10. doi: 10.1007/82_2020_210
- Waterworth SC, Flórez LV, Rees ER, Hertweck C, Kaltenpoth M, Kwan JC (2020). Horizontal gene transfer to a defensive symbiont with a reduced genome in a multipartite beetle microbiome. mBio 11: e02430-19. doi: 10.1128/mBio.02430-19
- Weber E, Berger C, Bonas U, Koebnik R (2007). Refinement of the *Xanthomonas campestris* pv. *vesicatoria hrpD* and *hrpE* operon structure. Mol. Plant Microbe Interact. 20: 559-567. doi: 10.1094/MPMI-20-5-0559

- Wengelnik K, Van den Ackerveken G, Bonas U (1996). HrpG, a key *hrp* regulatory protein of *Xanthomonas campestris* pv. *vesicatoria* is homologous to two-component response regulators. Mol. Plant Microbe Interact. 9: 704-712. doi: 10.1094/mpmi-9-0704
- Willems A, Gillis M, Kersters K, van den Broecke L, de Ley J (1987a). The taxonomic position of *Xanthomonas ampelina*. EPPO Bull. 17: 237-240. doi: 10.1111/j.1365-2338.1987.tb00033.x
- Willems A, Gillis M, Kersters K, Van Den Broecke L, De Ley J (1987b). Transfer of *Xanthomonas ampelina* Panagopoulos 1969 to a new genus, *Xylophilus* gen. nov., as *Xylophilus ampelinus* (Panagopoulos 1969) comb. nov. Int. J. Syst. Bacteriol. 37: 422-430. doi: 10.1099/00207713-37-4-422
- Worrall LJ, Majewski DD, Strynadka NCJ (2023). Structural insights into type III secretion systems of the bacterial flagellum and injectisome. Annu. Rev. Microbiol. 77: 669-698. doi: 10.1146/annurev-micro-032521-025503
- Xu RQ, Blanvillain S, Feng JX, Jiang BL, Li XZ, Wei HY, Kroj T, Lauber E, Roby D, Chen B, He YQ, Lu GT, Tang DJ, Vasse J, Arlat M, Tang JL (2008). AvrAC$_{Xcc8004}$, a type III effector with a leucine-rich repeat domain from *Xanthomonas campestris* pathovar *campestris* confers avirulence in vascular tissues of *Arabidopsis thaliana* ecotype Col-0. J. Bacteriol. 190: 343-355. doi:10.1128/JB.00978-07
- Zboralski A, Biessy A, Filion M (2022). Bridging the gap: type III secretion systems in plant-beneficial bacteria. Microorganisms 10: 187. doi: 10.3390/microorganisms10010187
- Zhang J, Guan J, Wang M, Li G, Djordjevic M, Tai C, Wang H, Deng Z, Chen Z, Ou HY (2023). SecReT6 update: a comprehensive resource of bacterial type VI secretion systems. Sci. China Life Sci. 66: 626-634. doi: 10.1007/s11427-022-2172-x
- Zhao S, Mo WL, Wu F, Tang W, Tang JL, Szurek B, Verdier V, Koebnik R, Feng JX (2013). Identification of non-TAL effectors in *Xanthomonas oryzae* pv. *oryzae* Chinese strain 13751 and analysis of their role in the bacterial virulence. World J. Microbiol. Biotechnol. 29: 733-744. doi: 10.1007/s11274-012-1229-5

Legend to the tables and figures

Table 1. *Xylophilus* genome sequences at NCBI GenBank.

Table 2. Candidate HrpX-regulated promoters of *X. ampelinus* strain CECT7646 and *X. rhododendri* strain KACC 21265.

Table 3. Prediction of type III-secreted proteins for SctV1-type T3SS of *X. ampelinus* using Effectidor II.

Table 4. Prediction of type III-secreted proteins for SctV2-type T3SS of *Xylophilus* using Effectidor II.

Figure 1. Phylogenetic tree of representative SctV/FlhA proteins and homologs from *Xylophilus*. *Xylophilus* sequences are boxed with the three flagellar proteins in yellow (type 1), orange (type 2) and red (type 3), and the Sct proteins in green (T3SS#1) and blue (T3SS#2). Strain CFBP 1192 was omitted because it is a clone of the *X. ampelinus* type strain CECT7646. Sequences that were used to identify SctV/FlhA homologs in proteomes of *Xylophilus* are indicated by a blue dot.

Figure 2. Presence of flagellar and secretion systems in strains of *Xylophilus*. The phylogenetic tree was calculated from the core proteome using M1CR0B1AL1Z3R, and is based on 644 proteins with 231,120 aligned amino acids. All nodes are supported with 100% bootstrap values. The genome of the *Xenophilus* type strain DSM 13620 was used to root the tree. Presence of flagellar and secretion systems, as deduced from homology to the conserved *sctV/flhA* gene and supported by TXSScan, SeCreT4 and SeCreT6 analyses, is indicated by green squares and absence by open squares.

Figure 3. Comparison of *Xylophilus* type III secretion gene clusters. Corresponding regions, extracted from GenBank files, were compared and visualized with Clinker. Two classes of type III secretion gene clusters were observed for *Xylophilus*, as supported by synteny and sequence similarities. Conserved *sct* genes are labeled in the same color as the genes. Blocks with near-

perfect synteny are boxed. The following genomic regions are depicted, from top to bottom: QU43_RS64645 (*sctV*) to QU43_RS64560 (*sctQ*) (*Burkholderia cenocepacia* J2315; NC_011001.1), Acav_0545 (*hpa2*) to ACAV_0510 (*hrpX*) (*Acidovorax avenae* ATCC 19860; CP002521.1), XOO0080 (*hpa2*) to XOO0110 (*xopAE*) (*Xanthomonas oryzae* MAFF 311018; AP008229.1), BFN94_RS13565 (*hpa2*) to BFN94_RS21845 (*xopM*) (*Xanthomonas translucens* DSM 18974; NZ_LT604072.1), RRSL_PS12385 (*sctC*) to RRSL_RS12490 (*hpaB*) (*Ralstonia solanacearum* UW551; NZ_AAKL01000026.1), DFQ15_RS08660 (*sctT*) to DFQ15_RS08795 (*hrpX2*) (*Xylophilus ampelinus* CECT7646; NZ_QJTC01000011.1), GN316_15560 (*sctC*) to GN316_15670 (*hrpB7*) (*Xylophilus* sp. Kf1; WPBX01000006.1), GT347_RS19740 (*sctC*) to GT347_RS19620 (*hrpB7*) (*Xylophilus rhododendri* KACC 21265; NP_CP047650.1), R9X41_RS04700 (*sctC*) to R9X41_RS04585 (*hrpB7*) (*Xylophilus* sp. GOD-1R; NZ_CP137854.1).

Figure 4. Translocation assays of novel T3E AvrBs1 reporter protein fusions. Bacterial *Xanthomonas* cells were inoculated into leaves of pepper cv. ECW-10R using needleless 1-mL plastic syringe. Strain 8004, which carries a copy of the avrBs1 gene and can elicit HR on non-host pepper cv. ECW-10R (Xu et al. 2008), was the positive control. Strains 8004Δ*avrBs1*, 8004Δ*hrcV* and mock inoculation with the buffer served as negative controls. The results presented are from a representative experiment, with similar results obtained in all other experiments.

Supplementary Figure S1. Phylogenetic tree of representative SctU/FlhB proteins and homologs from *Xylophilus*. *Xylophilus* sequences are boxed with the three flagellar proteins in yellow (type 1), orange (type 2) and red (type 3), and the Sct proteins in green (T3SS#1) and blue (T3SS#2). Strain CFBP 1192 was omitted because it's a clone of strain CECT7646. Sequences that were used to identify SctV/FlhA homologs in proteomes of *Xylophilus* are indicated by a blue dot.

Supplementary Figure S2. Comparison of *Xylophilus* flagellar gene clusters. Corresponding regions, extracted from GenBank files, were compared and visualized with Clinker. Three types of flagellar gene clusters were observed, as supported by synteny and sequence similarities. Flagellar genes are indicated as follows: flagellin genes in yellow, *flg* genes in red, *flh* genes in purple, *fli* genes in blue, *mot* genes in green and *che* genes in orange.

Supplementary Figure S3. Three-dimensional models of the tertiary structure of members of orthology groups OG_1115 and OG_1116. Snapshots of the predicted structure, coloured according to the pIDDT confidence scores, and the expected position errors are shown.

Supplementary Table S1. (a) Taxonomic classification of *Xylophilus* genome sequences using TYGS. (b) Taxonomic classification of *Xylophilus* genome sequences using MiGA.

Supplementary Table S2. (a) Genetic relationships between *Xylophilus* genomes based on genome-wide average nucleotide identities. (b) Genetic relationships between *Xylophilus* genomes based on digital DNA-DNA hybridization.

Supplementary Table S3. Predicted protein secretion systems in *Xylophilus* genomes using TXSScan.

Supplementary Table S4. HrpX-regulated protein families in *Xylophilus rhododendri*.

Supplementary Table S5. BLASTP hits of novel T3Es (OG_1115 and OG_1116).

Supplementary Table S6. InterPro hits of representation homologs of novel T3Es (OG_1115 and OG_1116).

Supplementary Table S7. FoldSeek hits of AlphaFold structural models of novel T3Es (OG_1115 and OG_1116).

Supplementary Data S1. Multiple sequence alignment of SctV/FlhA homologs in *Xylophilus*.

Supplementary Data S2. HrpX-regulated protein families in *Xylophilus rhododendri*.

Supplementary Data S3. Multiple sequence alignments of novel T3Es (OG_1115, OG_1116, OG_3308).

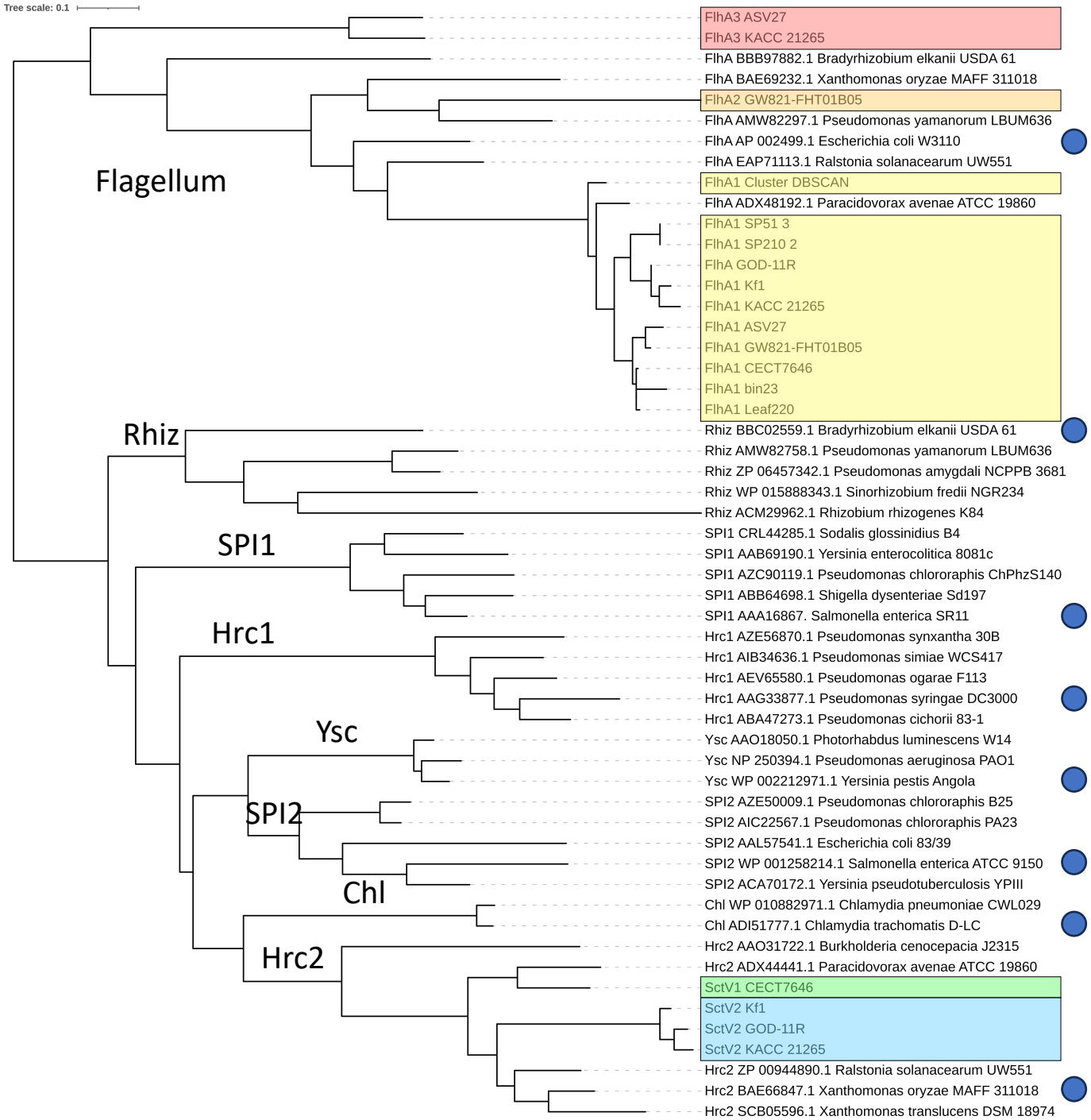
Tree scale: 0.1
Flagellum
Rhiz
SPI1
Hrc1
Ysc
SPI2
Chl
Hrc2
FlhA3 ASV27
FlhA3 KACC 21265
FlhA BBB97882.1 Bradyrhizobium elkanii USDA 61
FlhA BAE69232.1 Xanthomonas oryzae MAFF 311018
FlhA2 GW821-FHT01B05
FlhA AMW82297.1 Pseudomonas yamanorum LBUM636
FlhA AP 002499.1 Escherichia coli W3110
FlhA EAP71113.1 Ralstonia solanacearum UW551
FlhA1 Cluster DBSCAN
FlhA ADX48192.1 Paracidovorax avenae ATCC 19860
FlhA1 SP51 3
FlhA1 SP210 2
FlhA GOD-11R
FlhA1 Kf1
FlhA1 KACC 21265
FlhA1 ASV27
FlhA1 GW821-FHT01B05
FlhA1 CECT7646
FlhA1 bin23
FlhA1 Leaf220
Rhiz BBC02559.1 Bradyrhizobium elkanii USDA 61
Rhiz AMW82758.1 Pseudomonas yamanorum LBUM636
Rhiz ZP 06457342.1 Pseudomonas amygdali NCPPB 3681
Rhiz WP 015888343.1 Sinorhizobium fredii NGR234
Rhiz ACM29962.1 Rhizobium rhizogenes K84
SPI1 CRL44285.1 Sodalis glossinidius B4
SPI1 AAB69190.1 Yersinia enterocolitica 8081c
SPI1 AZC90119.1 Pseudomonas chlororaphis ChPhzS140
SPI1 ABB64698.1 Shigella dysenteriae Sd197
SPI1 AAA16867. Salmonella enterica SR11
Hrc1 AZE56870.1 Pseudomonas synxantha 30B
Hrc1 AIB34636.1 Pseudomonas simiae WCS417
Hrc1 AEV65580.1 Pseudomonas ogarae F113
Hrc1 AAG33877.1 Pseudomonas syringae DC3000
Hrc1 ABA47273.1 Pseudomonas cichorii 83-1
Ysc AAO18050.1 Photorhabdus luminescens W14
Ysc NP 250394.1 Pseudomonas aeruginosa PAO1
Ysc WP 002212971.1 Yersinia pestis Angola
SPI2 AZE50009.1 Pseudomonas chlororaphis B25
SPI2 AIC22567.1 Pseudomonas chlororaphis PA23
SPI2 AAL57541.1 Escherichia coli 83/39
SPI2 WP 001258214.1 Salmonella enterica ATCC 9150
SPI2 ACA70172.1 Yersinia pseudotuberculosis YPIII
Chl WP 010882971.1 Chlamydia pneumoniae CWL029
Chl ADI51777.1 Chlamydia trachomatis D-LC
Hrc2 AAO31722.1 Burkholderia cenocepacia J2315
Hrc2 ADX44441.1 Paracidovorax avenae ATCC 19860
SctV1 CECT7646
SctV2 Kf1
SctV2 GOD-11R
SctV2 KACC 21265
Hrc2 ZP 00944890.1 Ralstonia solanacearum UW551
Hrc2 BAE66847.1 Xanthomonas oryzae MAFF 311018
Hrc2 SCB05596.1 Xanthomonas translucens DSM 18974

Figure 1

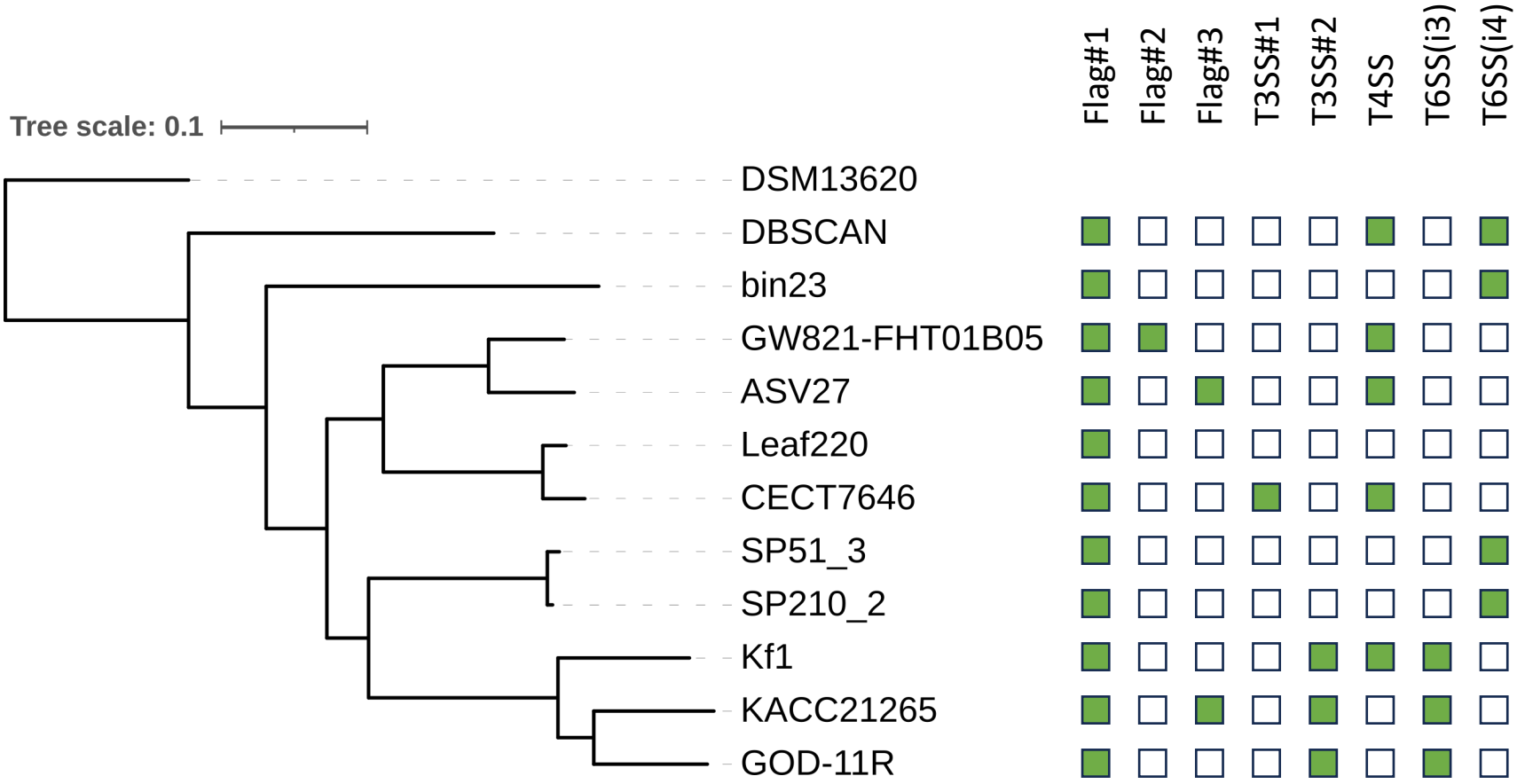
Tree scale: 0.1
Flag#1
Flag#2
Flag#3
T3SS#1
T3SS#2
T4SS
T6SS(i3)
T6SS(i4)
DSM13620
DBSCAN
bin23
GW821-FHT01B05
ASV27
Leaf220
CECT7646
SP51_3
SP210_2
Kf1
KACC21265
GOD-11R

Figure 2

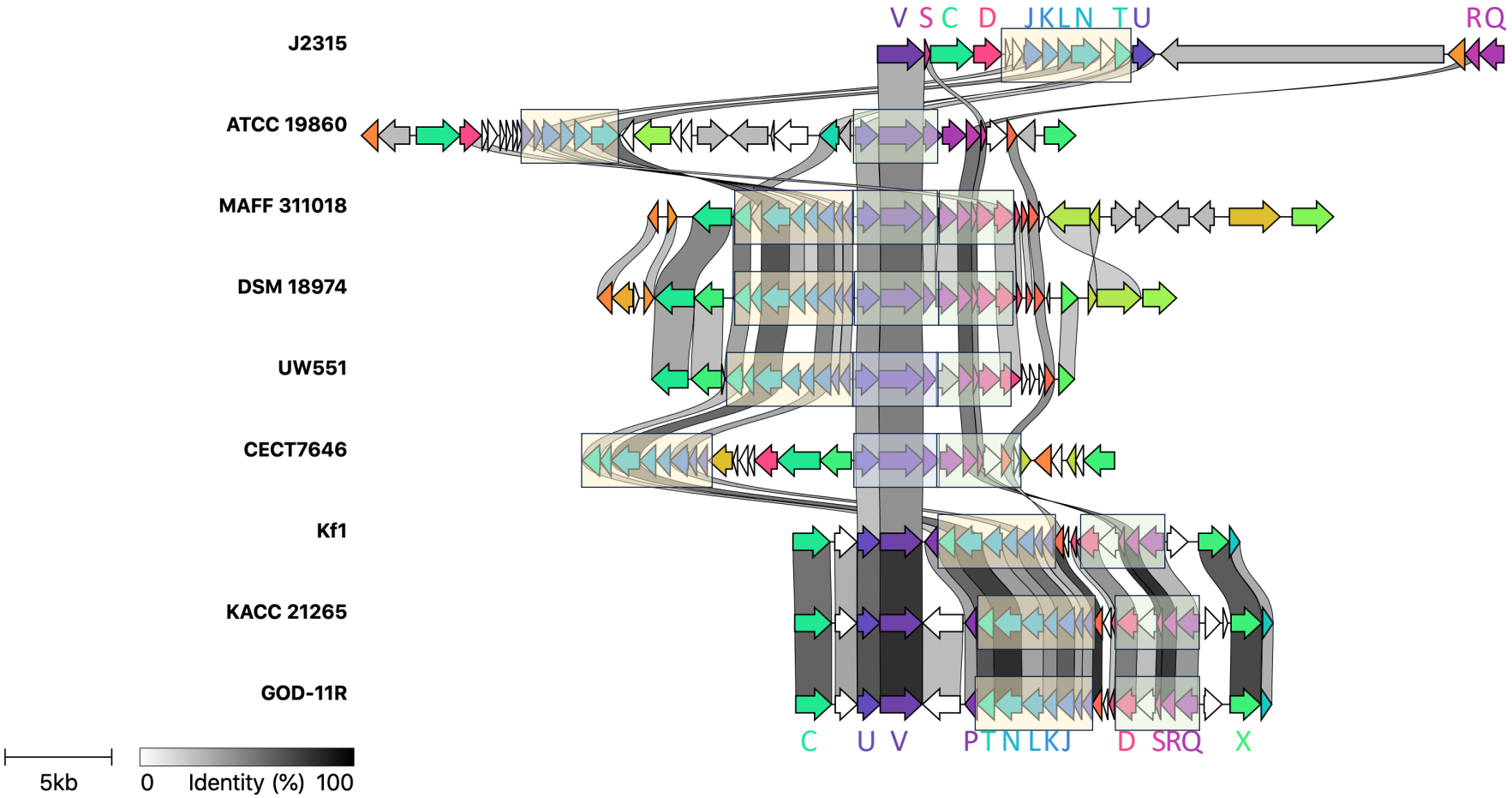
V S C D J K L N T U
R Q
J2315
ATCC 19860
MAFF 311018
DSM 18974
UW551
CECT7646
Kf1
KACC 21265
GOD-11R
C U V P T N L K J D S R Q X
5kb
0 Identity (%) 100

Figure 3

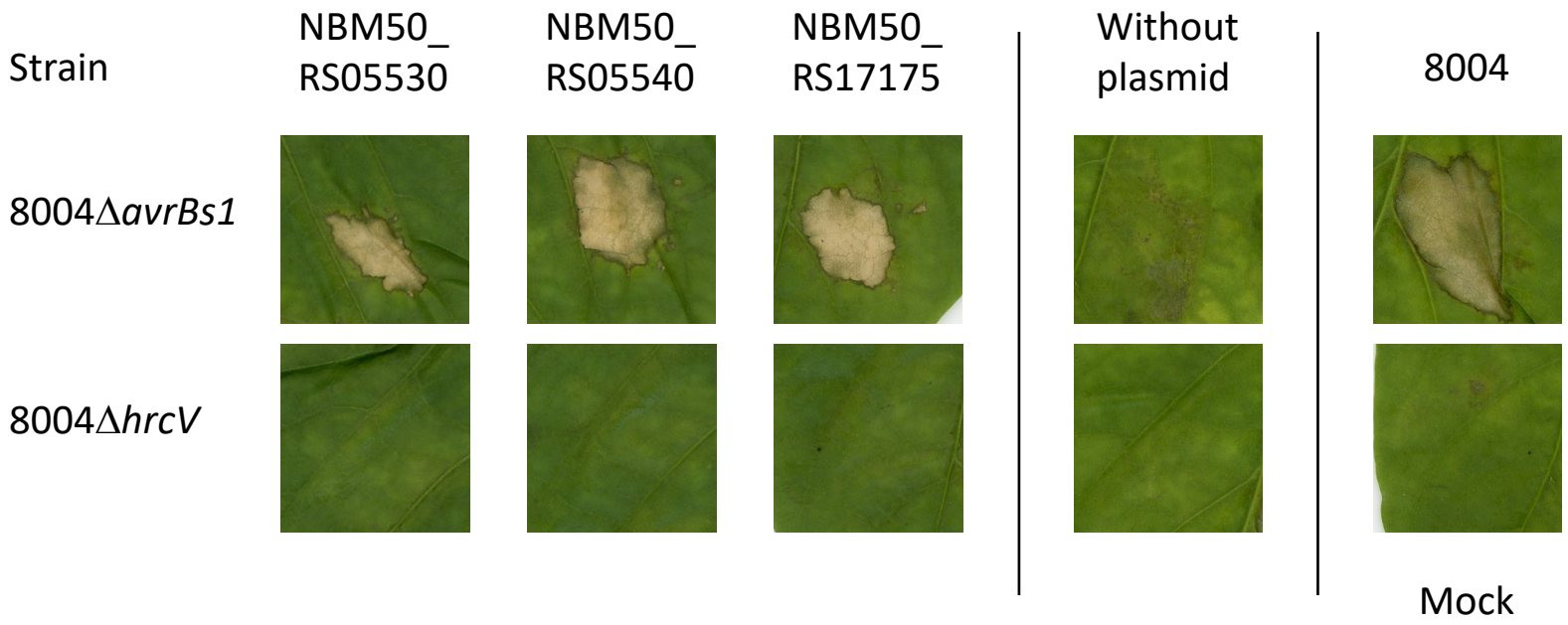
Strain
NBM50_ RS05530
NBM50_ RS05540
NBM50_ RS17175
Without plasmid
8004
8004Δ*avrBs1*
8004Δ*hrcV*
Mock

Figure 4