Infection, Genetics and Evolution 113 (2023) 105486

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Genetic rearrangements in *Pseudomonas amygdali* pathovar *aesculi* shape coronatine plasmids

Tue Kjærgaard Nielsen [a,1], Caroline S. Winther-Have [a,b,1], Iben Margrete Thomsen [c], Robert W. Jackson [d], Mojgan Rabiey [d], Rosanna Catherine Hennessy [a], Frederik Bak [a], Witold Kot [a], Mette Haubjerg Nicolaisen [a], Alexander Byth Carstens [a], Lars Hestbjerg Hansen [a,*]

[a] Department of Plant and Environmental Sciences, University of Copenhagen, 1871 Frederiksberg C, Denmark
[b] Globe Institute, University of Copenhagen, 1350 Copenhagen, Denmark
[c] Department of Geosciences and Natural Resource Management, University of Copenhagen, 1958 Frederiksberg C, Denmark
[d] School of Biosciences and the Birmingham Institute of Forest Research, University of Birmingham, Birmingham B15 2TT, United Kingdom

ARTICLE INFO

Keywords:
Plant pathogens
Pseudomonas
Phytotoxins
Coronatine
Virulence
Plasmid
Mobile genetic elements
Nanopore sequencing

ABSTRACT

Plant pathogenic *Pseudomonas* species use multiple classes of toxins and virulence factors during host infection. The genes encoding these pathogenicity factors are often located on plasmids and other mobile genetic elements, suggesting that they are acquired through horizontal gene transfer to confer an evolutionary advantage for successful adaptation to host infection. However, the genetic rearrangements that have led to mobilization of the pathogenicity genes are not fully understood. In this study, we have sequenced and analyzed the complete genome sequences of four *Pseudomonas amygdali* pv. *aesculi* (*Pae*), which infect European horse chestnut trees (*Aesculus hippocastanum*) and belong to phylogroup 3 of the *P. syringae* species complex. The four investigated genomes contain six groups of plasmids that all encode pathogenicity factors. Effector genes were found to be mostly associated with insertion sequence elements, suggesting that virulence genes are generally mobilized and potentially undergo horizontal gene transfer after transfer to a conjugative plasmid. We show that the biosynthetic gene cluster encoding the phytotoxin coronatine was recently transferred from a chromosomal location to a mobilizable plasmid that subsequently formed a co-integrate with a conjugative plasmid.

1. Introduction

The *Pseudomonas syringae* species complex is comprised of 15 known species of phytopathogens (Gomila et al., 2017), including *P. amygdali*. The species complex is widely used as a model for studying plant-microbe interactions, bacterial pathogenicity, and microbial molecular ecology (Xin et al., 2018). Strains belonging to the *P. syringae* complex harbor multiple biosynthetic gene clusters encoding phytotoxins and other virulence genes. However, it is unknown how these genes evolve and whether they are shared between pathovars (pv). Such information would provide valuable insight into the evolution and specialization of pathovars towards their hosts.

The *P. syringae* species complex is split into 13 phylogroups and there are >60 different pathogenic varieties, each attacking a specific group of plant hosts which includes economically important crops e.g. tomato, bean, kiwi, and mango (Xin et al., 2018). *P. syringae* complex members have dynamic genomes with a variety of different virulence factors that make them successful pathogens (Dillon et al., 2019; Ruinelli et al., 2019). They all have type III protein secretion systems (T3SS) that can penetrate plant cell walls and plasma membranes to enable injection of virulence factors (type III secreted effectors; T3SEs) directly into the host cell cytosol (Khan et al., 2018). Conversely, other toxins, such as phaseolotoxin and coronatine, are only found in subsets of pathovars (Baltrus et al., 2011; O'Brien et al., 2011; Xin et al., 2018).

Abbreviations: AF, Aligned fraction; ANI, Average nucleotide identity; CFA, coronafacic acid; CMA, coronamic acid; COR, Coronatine; HCBC, Horse chestnut bleeding canker; IS, Insertion sequence; JA-IIe, (+)-7-isojasmonoyl-L-isoleucine; MGE, Mobile genetic element; MOB, Plasmid mobilization; MPF, Mating-pair formation; Pae, *Pseudomonas amygdali* pv. *aesculi*; PFP, pPT23A family plasmids; PG, polygalacturonase; Pv., Pathovar; T3SE, Type III secreted effector; T4SS, Type IV secretion system.

* Corresponding author at: University of Copenhagen, Department of Plant and Environmental Sciences, Thorvaldsensvej 40, 1871 Frederiksberg C, Denmark.
E-mail address: lhha@plen.ku.dk (L.H. Hansen).

[1] TKN and CSWH contributed equally to this study.

https://doi.org/10.1016/j.meegid.2023.105486
Received 2 May 2023; Received in revised form 23 June 2023; Accepted 1 August 2023
Available online 2 August 2023

P. amygdali pv. *aesculi* (*Pae*) of the *P. syringae* complex is the causative agent of horse chestnut bleeding canker (HCBC) (Webber et al., 2008; Green et al., 2009). Bleeding canker disease caused by *Phytophthora* was first observed in the UK during the early 1970s, but since then a bacterial canker agent (*Pae*) has spread through central and northwestern Europe causing a severe epidemic (Brasihr and Strouts, 1976; Green et al., 2009; James et al., 2020; La Porta et al., 2023). Infected European horse chestnut trees develop lesions in the bark, oozing brown or black liquid, chlorotic foliage inducing crown dieback, and formation of necrotic phloem, thereby impairing water and nutrient transport which can lead to branch and tree death (Webber et al., 2008). Strains of *Pae* produce several phytotoxins encoded by biosynthetic gene clusters, most notably the coronatine (COR) cluster. The main mode of action of coronatine involves manipulation of the plant immune defense system by mimicking the structure of the active phytohormone (+)-7-iso-jasmonoyl-L-isoleucine (JA-IIe) that regulates plant stress response (Fonseca et al., 2009; Ma et al., 2007; Zhao et al., 2005). The COR cluster consists of two subcomponents, coronafacic acid (CFA) and coronamic acid (CMA), that are linked through an amide bond. COR manipulates the plant defense system by binding to the receptor complex COR insensitive 1, which activates JA-IIe-signaling and starts a cascade that leads to inhibition of salicylic acid signaling, due to antagonistic cross-talk between JA-IIe and salicylic acid-signaling (Geng et al., 2014; Zheng et al., 2012). Salicylic acid is vital for activating and regulating plant defenses against pathogens, including closure of the stomata (Melotto et al., 2006). In tomato, it has been shown that COR affects the transcription of plant hosts coupled with the inhibition of chlorophyll biosynthesis and significant changes to jasmonic acid synthesis and phenylpropane metabolism (Zhang et al., 2021). Hypothetically, the presence of COR-encoding genes in *Pae* presents an entry pathway for the pathogens into leaves by manipulating the stomata (Geng et al., 2014).

Strains of the *P. syringae* species complex are known to carry one or more plasmids, most of which belong to the pPT23A family plasmids (PFP) that share the plasmid replication gene *repA* and origin of replication (Gutiérrez-Barranquero et al., 2017; Zhao et al., 2005). The widespread occurrence of PFPs in the *P. syringae* species complex implies evolutionary importance and success of this plasmid family. This is supported by the array of virulence factors and host fitness enhancing components found to be encoded by the plasmids, aiding the pathogens in host interactions (Gutiérrez-Barranquero et al., 2017; Sesma et al., 1998; Zhao et al., 2005). However, many genes encoding virulence factors, including the important COR, have not yet been shown to be mobilized on plasmids. PFPs can harbor two classes of type IV secretion systems (T4SS), i.e. the mating-pair formation (MPF) systems MPF_T and MPF_I (Guglielmini et al., 2014; Gutiérrez-Barranquero et al., 2017; Ma et al., 2007). Despite the differences in gene content, all PFPs encode a conserved replication protein RepA, which can vary slightly in sequence, enabling the plasmids to suppress incompatibility (Bardaji et al., 2017; Sesma et al., 1998). Incompatibility between multiple co-existing PFPs is furthermore overcome by sequence variation in an antisense RNA encoded in the PFP plasmid replication control modules (Bardaji et al., 2017). PFPs often have large, conserved regions and show a mosaic structure with high rates of recombination (Gutiérrez-Barranquero et al., 2017; Sesma et al., 1998; Stavrinides and Guttman, 2004). About a third of the genetic material in PFPs is constituted by insertion sequence (IS) elements that enhance plasmid plasticity (Alarcón-Chaidez et al., 1999; Bardaji et al., 2017; O'Brien et al., 2011).

This study used comparative genomics to characterize how important virulence genes are associated with mobile genetic elements (MGEs), including multiple classes of IS elements and conjugative plasmids, allowing for ongoing and rapid adaptation towards increased pathogenicity. We hypothesized that virulence genes, such as phytotoxin coronatine, in *Pae* are mobilized from a chromosomal location to several mobilizable and conjugative plasmids. This leads to ongoing evolution of *Pae* pathovars with frequent genetic rearrangements, helping adaption for successful host infection.

2. Materials and methods

2.1. Environmental strain isolation and DNA extraction

P. amygdali pv. *aesculi* isolates 1804 and 2203 are novel isolates sampled from Kongens Have (55.6853° N, 12.5798° E) in Copenhagen, Denmark, from the inner bark of an infected lesion of HCBC on a European horse chestnut tree (*Aesculus hippocastanum)*. Isolates 6617 and 2250 were previously isolated from necrotic phloem of a diseased horse chestnut tree in Scotland in 2006 and 2008, respectively (Green et al., 2010).

For strain isolations, 1 g of bark sample was vortexed in 1× phosphate-buffered saline buffer and incubated for 48 h in 20 °C. All strains were cultured on King's B media, selective for *Pseudomonas* (King et al., 1954). Single non-fluorescent colonies were picked and inoculated in liquid LB for 48 h, shaking 250 rpm at 20 °C. From broth overnight culture, genomic DNA was extracted using the Genomic Mini AX Bacteria kit (A&A Biotechnology, Gdansk, Poland) according to the manufacturer's instructions. Quantity and quality control of the sampled DNA was estimated using a Qubit 2.0 fluorometer and Nanodrop ND-1000 spectrophotometer, respectively (Thermo Fisher Scientific, Waltham, MA, USA).

2.2. Library building and sequencing

For complete genome assemblies of the four *P. amygdali* pv. *aesculi* strains, separate sequencing libraries were built for both Illumina and Nanopore platforms. Sequencing libraries for the Illumina® NextSeq® system, were prepared using the Illumina® NEXTERA XT™ kit and sequenced on Illumina® NextSeq® system V2, MID, 2 × 150 cycles with paired-ended chemistry.

For Nanopore sequencing, the extracted DNA content was adjusted to 53.3 ng/μl of DNA using PCR-grade water for dilution. The Nanopore libraries, were made using the Rapid Barcoding kit (SQK-RBK004) and sequenced using the Oxford Nanopore Technologies MinION (Oxford, UK) platform with flow cell type FLO-MIN106D. Basecalling of raw Nanopore data was performed with Guppy (v. 4.2.2) with "high accuracy" model.

2.3. Assembly, annotation, and gene prediction

Illumina reads were trimmed for adapter sequences and poor quality using Trim Galore (v. 0.6.4) (Krueger, 2016), using default parameters. Nanopore reads were similarly trimmed with Porechop (v. 0.2.4) with default parameters. Initial hybrid assemblies were performed with Unicycler (v. 0.4.8) (Wick et al., 2017). However, due to the presence of several large repeated sequences within the genomes, Unicycler did not yield complete assemblies (contigs not estimated to be circular). To complete the genomes, as evaluated by obtaining circular contigs, assemblies from Unicycler were consolidated with assemblies from long-read assemblers Canu (Koren et al., 2017), Flye (Kolmogorov et al., 2019), and Raven (Vaser and Šikić, 2021), since some assemblers may outperform others on complex long repeats. The consolidated assemblies were verified for correctness by mapping Illumina and Nanopore reads and checking contiguity using minimap2 (v. 2.17) (Li, 2018), Bandage (v. 0.8.1) (Wick et al., 2015), and CLC Genomics Workbench (v. 21.0.3) (Qiagen, Aarhus, Denmark). Final assemblies were polished with both trimmed Illumina and Nanopore reads using the "unicycler_polish" module from Unicycler (Wick et al., 2017) that applies multiple rounds of Pilon (Walker et al., 2014) and Racon (Vaser et al., 2017). The

Table 1
Information on the complete genomes of the four strains of *P. amygdali* pv. *aesculi*.

Strain	Chr. Size (Mbp)	Plasmids/MGEs (kbp) - MOB/MPF	Year and country of isolation	Reference
P. amygdali pv. *aesculi* 2203	6.11	pPae2203CORI (181) - MOB_{PQ}/MPF_T	2020 DK	This study
		pPae2203Z (73) - MOB_{P/MPF_I}		
		pPae2203P (70) - MOB_P		
		pPae2203D (60) - MOB_{P/MPF_I}		
		pPae2203X (50) - MOB_P		
P. amygdali pv. *aesculi* 1804	6.11	pPae1804COR (115) - MOB_{PQ}	2020 DK	This study
		pPae1804ZI (137) - MOB_{PQ}/MPF_{TI}		
		pPae1804P (70) - MOB_P		
		pPae1804D (58) - MOB_{P/MPF_I}		
		pPae1804X (50) - MOB_P		
P. amygdali pv. *aesculi* 2250	6.12	pPae2250COR (116) - MOB_{PQ}	2008 UK	Green et al. (2010). Complete genome from this study.
		pPae2250Z (73) - MOB_P/MPF_I		
		pPae2250P (70) - MOB_P		
		pPae2250D (58) - MOB_P/MPF_I		
		pPae2250X (50) - MOB_P		
		pPae2250I (66) - MOB_{PQ}/MPF_T		
P. amygdali pv. *aesculi* 6617	6.14	pPae6617COR (115) - MOB_{PQ}	2006 UK	Green et al. (2010). Complete genome from this study.
		pPae6617Z (73) - MOB_P/MPF_I		
		pPae6617P (71) - MOB_P		
		pPae6617D (58) - MOB_P/MPF_I		
		IME6617 (424)		
		Tnp*HopAB1* (15)		
		pPae6617X (absent in assembly; found with mapping) - MOB_P		

complete genomes were assessed with BUSCO (v. 5.4.7) (Manni et al., 2021), using the "bacteria_odb10" BUSCO dataset.

Taxonomy was assigned using GTDB-Tk (v. 2.1.0) (Chaumeil et al., 2019) and plasmids were classified with MOB-suite (v. 3.0.3) (Robertson and Nash, 2018) and COPLA (Redondo-Salvo et al., 2021). Assembled and polished complete genomes were annotated with Prokka (v. 1.14.6) (Seemann, 2014). Gene predictions of regions of interest were confirmed using BLASTX (Altschul et al., 1990). Effector proteins potentially involved in virulence were identified by downloading the Hop effector protein database from pseudomonas-syringae.org (downloaded June 12th 2022) and finding homologous genes in the genomes with BLASTP (min. 80% ID and query coverage), using protein sequences predicted by Prokka. Pfam 34.0 (Mistry et al., 2021) was used for further inquiries into genes annotated as hypothetical carried by plasmids. Identification of secondary metabolites biosynthesis gene clusters was achieved using antiSMASH (Blin et al., 2021). ISfinder (Siguier, 2006) was used for identification of insertion sequences (IS elements). Promoters were identified using CNNPromoter_b with *Escherichia coli* as model organism (Softberry inc.) (Umarov and Solovyev, 2017).

2.4. Comparative genomics

Similar strains with complete genomes for comparative genomics were identified with BLASTN search, using plasmids of the four horse chestnut pathogens (Table 1) as queries.

Chromosomes and plasmids were compared for average nucleotide identity (ANI) with FastANI (v. 1.2) (Jain et al., 2018) using the options '–fragLen 30' to adjust the length of compared fragments to 30 bp. The ANI and aligned fraction (AF) were imported into R (v. 4.1.3) (R Core Team, R.D.C, 2016). The ANI and AF values were filtered to a minimum of 80% before plotting in a heatmap with the 'pheatmap' package (Kolde, 2019), with hierarchical clustering of the replicons. Genome rearrangements were identified using the multiple genome alignment program MAUVE (Darling et al., 2010), with progressive alignment setting. Sequence comparisons of COR regions were performed using Clinker (Gilchrist and Chooi, 2021) with subsequent manual text annotation performed using InkScape (Inkscape, 2020).

3. Results and discussion

3.1. Genome assembly and general features

To understand the genetic rearrangements of pathogenicity genes in *Pae*, four strains were whole genome sequenced and their plasmids were investigated. New European *Pae* strains 2203 and 1804 were isolated and sequenced for the first time from bleeding canker of horse chestnut trees, whereas both *Pae* 6617 and *Pae* 2250 have been previously whole genome sequenced by Green et al., 2010 with Illumina, resulting in draft genomes. In this study, completely resolved genomes of these are generated, allowing for comparative genomics on complete replicons. Illumina data resulted in a sequence depth of 51× to 168×, while Nanopore resulted in a depth of 96× to 107× with read length N50s of 9.6 to 25 kbp for the four strains (more details on sequencing data in Supplementary Table 1). The assembled complete genomes ranged in size from 6.11 to 6.14 Mbp, which is close to the mean 6 Mbp for the *P. syringae* complex (Gomila et al., 2017). All four complete genomes were found to be of very high quality, with 124/124 BUSCO groups of universal orthologous genes identified as "Complete and single-copy BUSCOs". They contained five to six mobilizable and conjugative plasmids (Table 1) with MOB_Q and/or MOB_P relaxases and MPF_T or MPF_I-type IV secretion systems, as determined with MOB-Suite (Robertson and Nash, 2018). A high number (two to six) of coexisting PFP plasmids has been observed before in the *P. syringae* species group (Bardaji et al., 2017; Gomila et al., 2017). All plasmids had toxin-antitoxin

systems (identified with BLASTX (Altschul et al., 1990); results not shown), that increases plasmid stability through addiction (Harms et al.,

2018).

The ANI between the four complete genomes was >99.9%. They furthermore had an ANI of 98.35–98.43% (aligned fraction 95%) with *P. amygdali* (acc. Number GCF_002699855.1) and only 89.5% with *P. syringae* (acc. Number GCF_000507185.2), as determined with GTDB-Tk (Chaumeil et al., 2019). This shows that the four main strains in this study are highly similar and are of the *P. amygdali* species. Using the multiple genome alignment program MAUVE (Darling et al., 2004), the four complete genomes showed high synteny on the chromosome level (results not shown), whereas higher diversity and recombination could be observed between the plasmids. All plasmids carried a *repA* gene that was approx. 90.5% similar to *repA* in the model PFP plasmid pPT23A. The coronatine-encoding plasmids (e.g. pPae2203CORI and pPae1804-COR) carried an additional, second *repA* that was 88.5% similar to that of pPT23A. It has previously been shown that *P. syringae* strains can carry several highly related coexisting plasmids that display altered in-compatibility behavior (Bardaji et al., 2017; Sesma et al., 1998). This might explain our observations of multiple, seemingly related plasmids within the same strains. The identified plasmids were named according to the strain they were found in and the important feature they encode, e.g. PFP-family plasmid pPae1804X is found in *Pae* 1804 and harbors the *hopX1* gene; all the other plasmids are also PFP-family plasmids. Some plasmids display inter-strain size variation (Table 1), where the most drastic size variations appear to be due to co-integrations between two plasmids, as discussed below.

Strain 2250 was sequenced in parallel to this study at University of Birmingham (NCBI acc. Numbers CP123794-CP123798) and it was interestingly found that the 58 kbp plasmid pPae2250D was missing in this variant (results not shown). As this plasmid is present in all four strains in this study, we believe that the plasmid was lost in the UK variant, as has been observed before when bacterial strains are shared between labs (Landgraf et al., 2006; Nielsen et al., 2021). This is supported by the plasmid gel analysis in Green et al. (2010), showing the presence of six plasmids for strain 2250 (including pPae2250D), with sizes largely corresponding to those reported in Table 1. The same gel from Green et al. shows that the 66 kbp pPae2250I plasmid is missing in strain 6617, agreeing with our finding (Table 1).

3.2. Virulence factors associated with plasmids

Multiple genes encoding the Hop and Avr T3SE proteins, known as virulence factors in *P. syringae* pathovars (Dillon et al., 2019), were identified in the four *Pae* genomes (Fig. 1; amino acid similarities to database proteins: 93.6–100%). These T3SEs are scattered across the chromosomes and plasmids. The chromosomes encode 24 different *hop* and *avr* genes, with the UK isolates 6617 and 2250 having two copies of

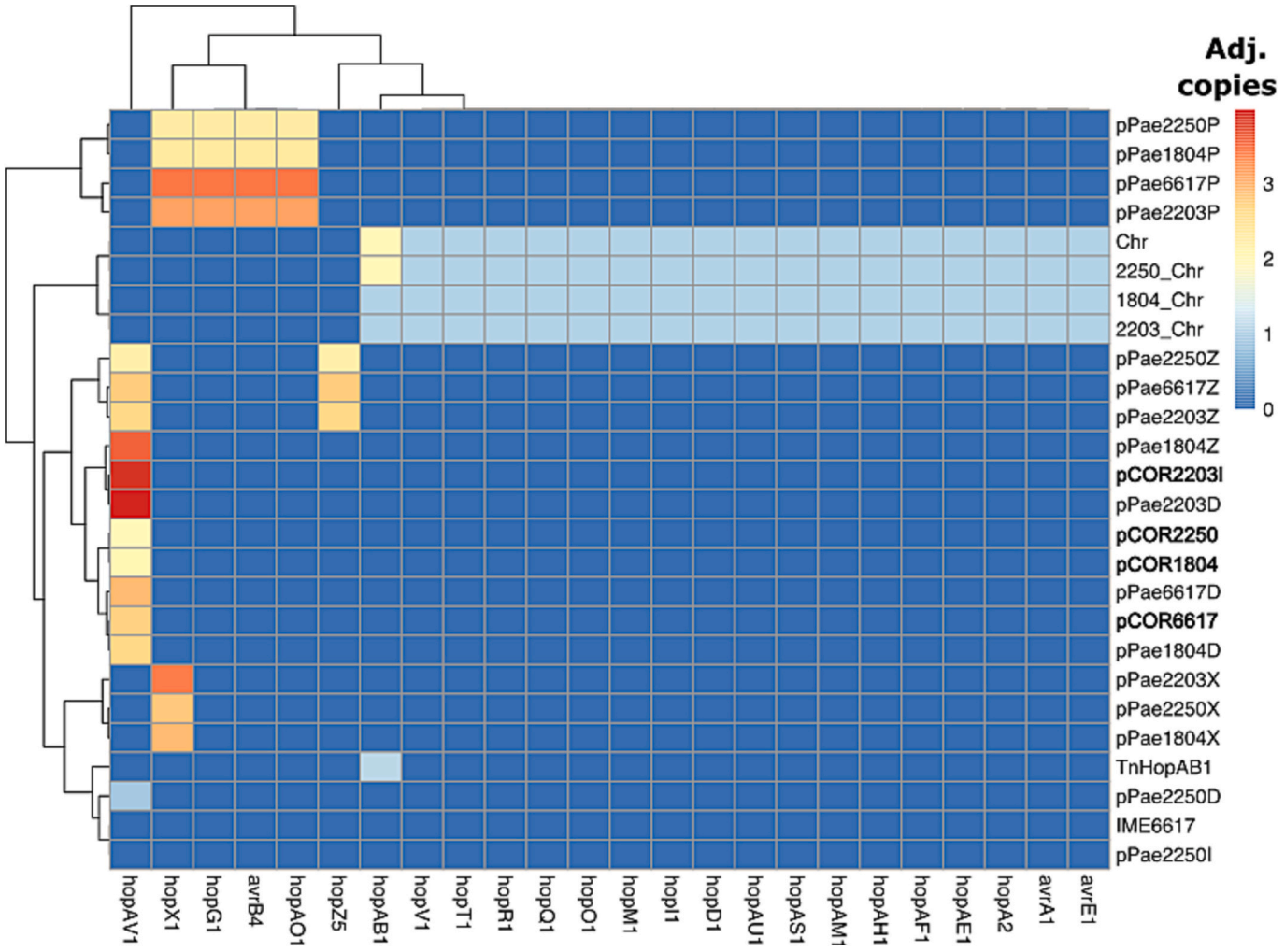

Fig. 1. Heatmap showing the presence or absence of effector proteins on replicons in *P. amygdali* pv. *aesculi*. Effector proteins were predicted using BLASTP (Altschul et al., 1990), coupled with the Hop effector protein database from pseudomonas-syringae.org. The color of tiles indicates the predicted number of genes adjusted for the replicon copy numbers in Table 2. Rows and columns are clustered with hierarchical clustering in the pheatmap R package (Kolde, 2019). The COR plasmids in four *Pae* in this study are in bold font. Coronatine is not included in the figure.

Table 2
Characteristics of plasmids and extrachromosomal elements in four *P. amygdali* pv. *aesculi* strains analyzed by MOB-suite (v. 3.0.3) (Robertson and Nash, 2018) and COPLA (Redondo-Salvo et al., 2021). The listed elements were identified as independent, circular elements in genome assemblies.

MGE group[a]	Size range (kbp)[b]	MOB/MPF[c]	Relevant features[d]	Copy number[e]
COR (coronatine)	115–181	$MOBP_Q$/MPF_T	CFA, CMA, and *corRSP* clusters	2.0–2.8 (3.9)
Z (*hopZ5*)	73–137	$MOB_{P(Q)}$/$MPF_{I(T)}$	*hopZ5, hopAV1*	2.2–3.6
P (Polygalacturonase)	70–71	MOB_P	Polygalacturonase encoding gene, *hopG1*, *hopX1*	2.3–3.5
D (D,D-dipeptide)	58–60	MOB_V/MPF_I	D,D-dipeptide ABC transporter encoding gene.	0.9–4
X (*hopX1*)	50	MOB_P	*hopX1*	2.8–3.5
I (Integrates with COR and Z plasmids)	66	$MOBP_{PQ}$/MPF_T	Formed cointegrate plasmid with COR plasmid in strain 2203 and with Z plasmid in strain 1804, as supported by Nanopore data. It is an independent plasmid in strain 2250 but absent in strain 6617.	3.3
IME6617	424	MOB_P	Integrated into chromosomes of all chestnut pathogens. Also exists as independent, circular molecule in strain 6617, as verified by assembly and read mapping. The copy number is based on extrachromosomal copies.	1.0
Tn*hopAB1*	15		Carries the T3SE *hopAB1* gene. Integrated into chromosomes of all four chestnut pathogens. Also exists as circular transposition intermediate molecule in strain 6617, as verified by read mapping across ends of molecule. The copy number is based on extrachromosomal copies whereby a copy exists in the chromosome but in some cells an extra copy exists extrachromosomally.	1.1

a) MGE groups are named after encoded trait, not Inc. group, b) the size range of MGEs in kbp, c) the plasmid mobilization (MOB) and/or mating pair formation (MPF) group predicted with MOB-suite (Robertson and Nash, 2018), d) other relevant features and notes, e) Copy number is calculated from the read depth of the MGE normalized to the read depth of chromosome. For elements IME6617 and Tn*hopAB1*, the copy number is calculated based on the sequencing depth subtracted by the chromosome depth, to account for both the chromosomally integrated and extrachromosomal states.

hopAB1 and three copies of *hopAV1* (27 total *hop/avr* genes in each strain). Remarkably, 21 out of the 27 effector genes are located directly adjacent to an IS element or have maximum two genes between the effector gene and an IS element. This may indicate that the effector genes have recently been, and likely continue to be, mobilized and that they have the potential to be shared across different *P. syringae* phylogroups, as shown previously (Dillon et al., 2019; Ruinelli et al., 2019). In particular, the DEDD-type IS*Psy16* and DDE-type IS*Psy17* transposases are associated with effector genes and are found in proximity with 10 out of the 27 total *hop/avr* genes. Other IS elements associated with effector genes in the four *Pae* strains include IS*52*, IS*5376*, IS*801*, IS*Psy31* among others. The chromosomal effector genes are scattered and do not appear in apparent clusters or operons, suggesting that their mobilization and selection are under individual mechanisms. The scattering of effector genes may be common but the current number of complete genome sequences from the *P. syringae* complex is too limited to conclude on this. All four *Pae* isolates harbor large mobilizable plasmids encoding the COR biosynthetic gene clusters for CFA and CMA (Table 2). These COR plasmids show a high degree of similarity between the strains (Fig. 2).

The putative conjugative group Z plasmids (Tables 1 and 2) carries a gene encoding for a C55 peptidase HopZ5 which in soybean promotes infection by suppressing biosynthesis of the polyphenolic compound isoflavone (Zhou et al., 2011). However, in *Pae* strain 1804, the Z group plasmid (pPae1804Z) appears to have recombined with a conjugative pPae2250I-like plasmid and has thereby apparently lost the *hopZ5* gene (Fig. 1). The resulting co-integrate is named pPae1804ZI. All the Z group plasmids further harbor a gene encoding a different effector, HopAV1.

Group P plasmids all harbor polygalacturonase (PG) encoding genes, which encode a similar pectinase that breaks down pectin, a major component of the plant cell wall's middle lamella (Sénéchal et al., 2014). The plant cell wall is a major target of pathogens, and PG has been linked to enhanced plant susceptibility towards pathogens, as the breakdown of the cell wall eases colonization (De Lorenzo and Ferrari, 2002; Wang et al., 2017). The expression of PG is considered a primary virulence factor in numerous pathogens (Huang and Allen, 2000; Rodriguez-Palenzuela et al., 1991). In a study by Nowell and colleagues (Nowell et al., 2016), in which comparative genomics was used to identify genes associated with infection of woody hosts in *P. syringae*, the pectinase pectin lyase was identified as being one of three key genes unique to *Pae*. We found that PG was solely encoded on mobilizable plasmids, and that group P plasmids were present in all four *Pae* genomes, suggesting that PG enhances *Pae* pathogenicity. All the investigated group P plasmids also encode the T3SE genes *hopX1*, *hopG1*, *avrB4*, and *hopAO1* (Fig. 1), in addition to PG. HopG1 targets mitochondria and promotes disease (Block et al., 2010), and HopX1 has been linked to cell death in specific in *Arabidopsis thaliana* genotypes, although the specific interaction remains poorly understood (Gimenez-Ibanez et al., 2014; Nimchuk et al., 2007).

Group D plasmids all encode a D,D-dipeptide ABC transporter system (Fig. 1). A recent study by Yan et al. (2020), found that ABC transporter systems might play a role in *P. syringae* virulence. They found that the substrate binding transport system directly connected host-derived acidic amino acids to transcriptional regulation of genes encoding the T3SS, thereby playing a central role in *P. syringae* virulence (Yan et al., 2020). The D group plasmids furthermore encode the *hopAV1* virulence factor, like COR plasmids and group Z plasmids. The HopAV1 proteins are >97.9% similar (amino acid) between the COR, Z, and D plasmids, but it is not known whether the variants are redundant or perform different niche functions.

The group X plasmids carry another T3SE HopX1-encoding gene that is 100% similar (amino acid) to that on group P plasmids, suggesting redundancy unless transcription is different from the X and P plasmids. The X plasmid was for unknown reasons missing in the assembly of strain 6617 but was manually confirmed to be present by mapping of reads (Table 1) to the X plasmid from strain 2203 (pPae2203X) with a

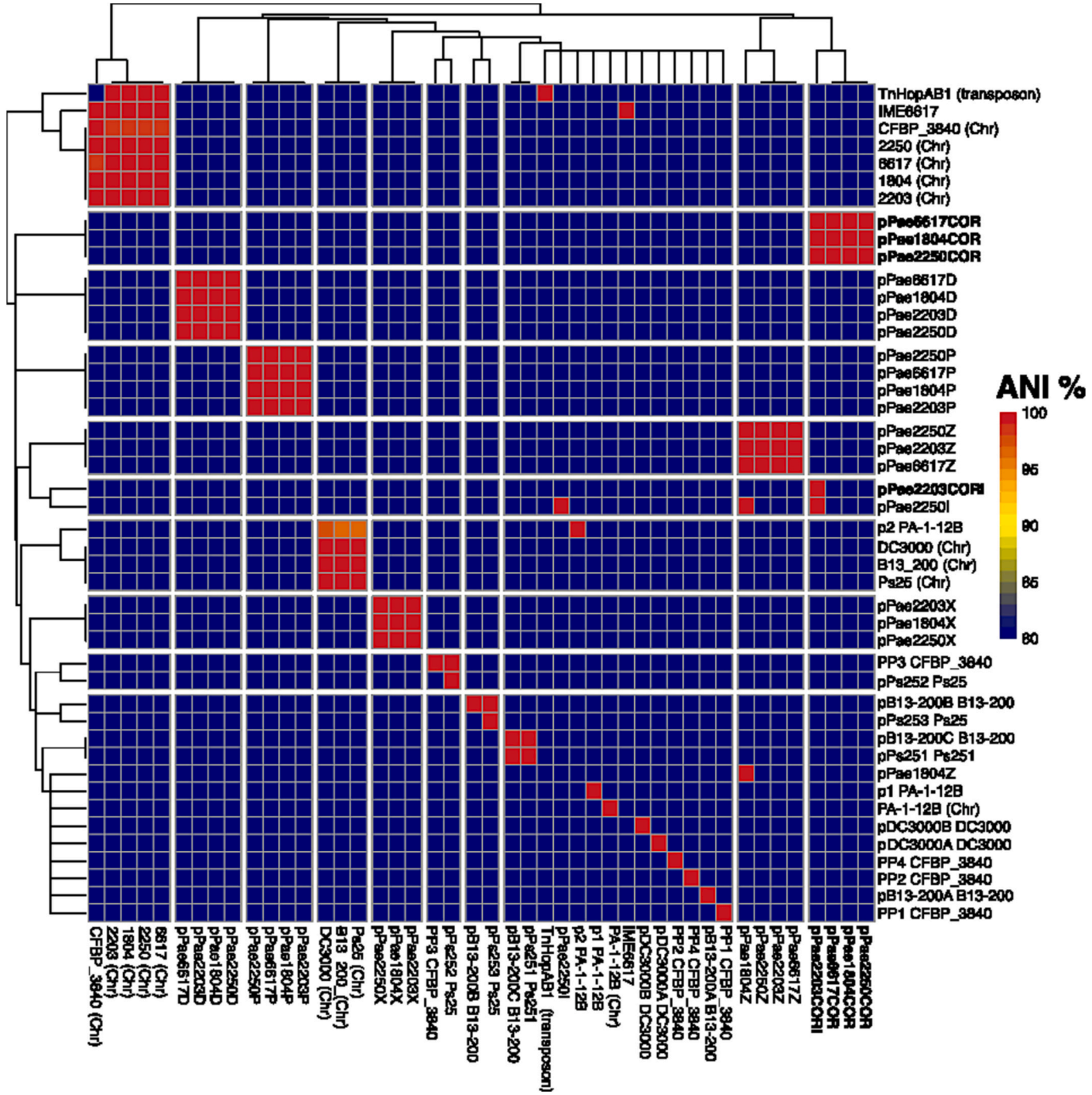

Fig. 2. Average nucleotide identity (ANI) values between replicons from this study, as well as similar reference strains. Similar strains with complete genomes for comparative genomics were identified with BLASTN search, using plasmids of the four horse chestnut pathogens (Table 1) as queries. Strains used for comparison are *P. syringae* CFBP3840 (acc. Number GCF_900235814.1), *P. syringae* B13–200 (acc. Number GCF_002966565.1), *P. syringae* Ps25 (acc. Number GCF_004323805.1), *P. syringae* DC3000 (acc. Number GCF_000007815.1), *P. tremae* PA-1-12B (acc. Number GCF_023278115.1). Reference strains ANI values were calculated with FastANI (Jain et al., 2018) and heatmap was constructed with 'pheatmap' package in R (Kolde, 2019), with hierarchical clustering of rows and columns. The same replicons appear on both axes, but the order and clustering are different since e.g. ANI of pPae2203CORI vs. pPae2250COR is not the same as pPae2250COR vs pPae2203CORI, due to the size difference from co-integration with the 66 kbp MPF_T I group conjugative plasmid in pPae2203CORI (Table 1). Only ANI values higher than 80% and with aligned fractions higher than 80% are shown. The coronatine-encoding plasmids are highlighted in bold.

calculated copy number of 3.3 which is within the same range as the X plasmid in the other three strains. The I group plasmid (e.g. pPae2250I) was similarly investigated in strain 6617 with read mapping but was confirmed to be absent. This shows the importance of recovering unassembled contigs.

In strain 2203 (DK), the pPae2203CORI plasmid is a co-integrate molecule with the 66 kbp MPF_T I group conjugative plasmid that is seen as an independent plasmid (pPae2250I) in UK strain 2250 but is absent in UK strain 6617 (Table 2, Fig. 2). In DK strain 1804, the pPae2250I plasmid has co-integrated with the Z group plasmid, to form

a 137 kbp co-integrate plasmid pPae1804ZI (Table 2, Fig. 2). As such, the COR plasmid in strain 2203 (pPae2203CORI) and the Z plasmid in strain 1804 (pPae1804ZI) now carry a T4SS from the co-integration events with the I group plasmid, likely increasing the mobility of these plasmids. These co-integrates are discussed further below.

3.3. Genome comparison reveals extrachromosomal MGEs

Additional genomes of *Pseudomonas* strains also harboring COR clusters displaying a high similarity to those identified in the four strains of this study (DK strains 2203 and 1804; UK strains 2250 and 6617) were downloaded and ANI comparisons made (Fig. 2). The complete chromosomes of the four horse chestnut pathogens sequenced in this study were found to be very similar to each other and to the chromosome of cherry pathogen *P. syringae* pv. *morsprunorum* CFBP 3840 from France (Ruinelli et al., 2019), belonging to *P. syringae* phylogroup 3. A 424 kbp potential integrative and mobilizable element (abbreviated IME; IME6617) was found integrated in the chromosome of the four chestnut pathogens (Table 2), as well as in strain CFBP 3840 (Fig. 2). In *Pae* strain 6617, this potential IME was uncovered by assembly and read mapping to constitute an independent circular molecule in a copy number of 1 per chromosome copy (Table 2), as well as integrated into the chromosome where it is flanked by inverted IS*3* family IS*51* elements. While IME6617 carries a MOB_P relaxase, it is not known whether the extrachromosomal IME can mobilize and integrate into new bacterial host chromosomes. Within this potential IME, two biosynthetic gene clusters encoding potential antibiotics were predicted (antiSMASH) in addition to potential genes required for motility (not shown).

A 15 kbp transposon harboring a T3SE HopAB1 protein-encoding gene is found to be independent and circular (as reported by Flye assembler) in strain 6617 at a copy number of 1.1 per chromosome copy, based on assembly and read mapping validation. The same transposon is integrated into the chromosome in 6617 and the other three chestnut pathogens (Tn*hopAB1* in Table 2 and Fig. 2). The observation of a circular form of the transposon may reflect an excision and circularization of the IS in some of the cells in the population, as seen for other mobile genetic elements (Nielsen et al., 2017). On the chromosome, the transposon has an IS*5* transposase in one end and an IS*110* transposase in the other. It is unknown which of the two is responsible for forming the circular intermediate molecule, as both appear in Tn*hopAB1*. The transposon is missing from the closely related strain CFBP 3840 (Fig. 2). In other *P. syringae* strains, HopAB has been shown to be injected into the host plant to inhibit its immune system and promote further infection (Wei et al., 2015).

3.4. Genetic rearrangements shape coronatine plasmids

It was previously shown that coronatine-encoding genes are found in only a few members of distinct *P. syringae* phylogroups (PG1, PG3, PG4, and PG5) (Ruinelli et al., 2019), suggesting that COR might be generally associated with mobile genetic elements, due to its mosaic distribution across phylogroups. Our results support this plasmid-borne distribution of COR genes. Based on the work presented here, we propose evolution of the COR clusters from a chromosomal position to several mobilizable and conjugative plasmids (Fig. 3). The order of genetic regions in Fig. 3 represents the simplest and thus most likely progression of genetic rearrangement, although this may not prove true when more complete *Pae* genomes become available. However, a multiple sequence alignment of the major coronafacic acid polyketide synthase genes, *cfa6* and *cfa7*, supports this order of rearrangements (Fig. 3 insert), assuming that the rates of major genetic rearrangements and within-gene mutations are approximately similar. On the chromosome of type strain *P. syringae* pv. *tomato* DC3000 (Fig. 3A), the CFA and CMA genes are located in two distinct loci, separated by 8 ORFs (26 kbp) encoding various IS elements and a non-ribosomal peptide synthetase that is thought not to be related to coronatine but shows similarities to iron-chelating myxochelin MxcG. The putative insertion of the unrelated non-ribosomal peptide synthetase and flanking IS elements suggests that the two COR loci were originally located adjacent to each other, as seen in the chromosome of *P. tremae* PA-1-12B (Fig. 3B). It has furthermore been suggested that COR genes were recently introduced in DC3000 (Baltrus et al., 2011). However, only a possibly degenerated or truncated transposase gene (IS*481*) upstream of the first gene in the CFA cluster encoding coronafacic acid synthetase ligase (Cfl) could be found in this study, suggesting that a potential IS-mediated transfer is not recent. Further upstream of the CFA cluster and immediately upstream of the CMA cluster are two

Fig. 3. Comparisons between coronatine genetic regions made with Clinker and annotated in InkScape. Only selected, relevant genes and features are annotated. Asterisk in *P. syringae* DC3000 chromosome sequence represents 29 ORFs that were removed to fit the sequence in the fig. A phylogenetic tree (insert) based on the cfa6–7 genes supports the order of sequences. The tree was constructed in CLC Genomics Workbench 22 (Qiagen, Aarhus, Denmark), using UPGMA algorithm and the Jukes-Cantor distance model.

copies of an IS*pssy* transposase that is 81% similar to that upstream of CFA in *P. tremae* PA-1-12B (Fig. 3B), suggesting that these elements may have been involved in rearranging chromosomal CFA clusters.

In the chromosome of *P. tremae* PA-1-12B, the CFA cluster is located immediately downstream of an IS*Psy1* element that has likely been involved in mobilizing the COR cluster to plasmids, as seen in pB13–200A from *P. syringae* B13–200 (Fig. 3C). On plasmid pB13–200A, several other IS elements (IS*Psy19*, IS*1182*, IS*Psy1*, IS*52*, and IS*Psy2*) are associated with the COR cluster, likely enabling increased mobilization. Furthermore, pB13–200A is a conjugative plasmid, complete with plasmid replication gene *repA*, a *tra* operon, and a relaxase *mobA* gene (Fig. 3C), enabling conjugative transfer of the COR cluster. The T4SS *tra* operon and *mobA* loci of pB13–200A has been rearranged on plasmid pPs252 from strain Ps25 (Fig. 3D). The pPs252 plasmid furthermore displays two repeated regions, R1 and R2, of which R1' contains a truncated version of the coronafacic acid polyketide synthase I (*cfa6*). Plasmid pPs252 likely represents a co-integrate with another plasmid, as a *repA* gene is found here. Furthermore, another T4SS is encoded on pPs252 that is likely a protein secretion T4SS, as there is no associated DNA relaxase. This is supported by a COPLA (Redondo-Salvo et al., 2021) analysis that only predicts a conjugative T4SS for the *tra* genes that are shared with pB13–200A (not shown). This co-integrate plasmid has likely originated from a plasmid similar to pPP3 from strain CFBP 3840 (Fig. 3E). We did not find a *mob* relaxase or a conjugation T4SS on plasmid pPP3 and is thus likely to depend on other plasmids in strain CFBP 3840 for horizontal transfer. The *cfa6* gene on pPP3 appears to be truncated, but we suggest that this may be an erroneous annotation, as there are no mutations that may account for this, when compared to *cfa6* from the other strains (not shown).

While there are certainly intermediate states of the COR cluster evolution that is not reflected in Fig. 3 and the order of genetic rearrangements is yet unclear, the comparison of COR-related regions in the chromosomes of strains *P. syrinage* DC3000 and *P. tremae* PA-1-12B and the plasmids pB13–200A, pPs252, and pPP3 indicates that the COR cluster has been mobilized from a chromosomal position to conjugative plasmids and is associated with several IS elements, likely forming transposons (Fig. 3A-E). It is not currently known what represents the ancestral state of the COR cluster, but a mobilization from a chromosome to a plasmid, as shown in Fig. 3, is the simplest and thus likely course of events. We have included all available sequences from complete bacterial genomes with the COR cluster, based on a BLASTN search, in Fig. 3. More thorough comparative genomics analyses will hopefully be possible in the future, as more complete genomes become available.

The genetic surroundings of the described COR clusters display high degree of variation, suggesting an ongoing evolution with frequent rearrangements. On the other hand, plasmids from the four distinct *P. amygdali* isolates from UK and Denmark in this study display a much higher conservation (Fig. 3F-I). These plasmids (pPae2250COR, pPae6617COR, pPae1804COR, and pPae2203CORI) display a very high number of IS elements (20–27% of CDS) and all share a similar overall structure. An inversion of an IS-rich region separates the DK plasmids from the UK plasmids (Fig. 3G-H). Furthermore, pPae2203CORI is a co-integrate with the conjugative I group plasmid pPae2250I that exists as an independent plasmid in *Pae* strain 2250 (Fig. 3I), converting the mobilizable COR plasmid to a conjugative plasmid in strain 2203 (Figs. 2–3; Tables 1–2). pPae2250I has likewise integrated pPae1804Z in *Pae* strain 1804 (Fig. 2; Tables 1–2), forming pPae1804ZI.

We identified the *corRSP* two-component regulatory system within the *cma* cluster in all investigated plasmids and chromosomes (Fig. 3). This two-component system is known to regulate COR expression in a temperature-dependent manner with highest levels at 18 °C in *P. savastanoi* pv. *glycinea* PG4180 (Xie et al., 2020). The transmembrane histidine protein kinase CorS senses temperature and regulates expression of COR through transphosphorylation of CorR that in turn activates COR biosynthesis. Strain *P. syringae* DC3000 does not display temperature-dependent expression of COR, although CorS is required for COR expression (Xie et al., 2020). We identified a single adenine insertion in the *corS* gene of DC3000 compared to strain PG4180 and all other strains in this study (not shown). This suggests that DC3000 *corS* contains a frameshift mutation leading to loss of function, whereas other strains in this study likely display temperature-dependent COR expression. The region encoding the short *corP* gene shows variation between the compared strains, resulting in variation in lengths. However, we could not detect mutations that could account for this, suggesting that the differences are rather due to gene prediction software errors. Furthermore, the *cfa* and *cma* clusters have an upstream promoter in the flanking IS*Psy1* and IS*Psy2* genes, suggesting potentially increased expression of the COR toxin.

3.5. Ongoing genomic evolution of *P. amygdali pv. aesculi* with regards to pathogenicity

This study describes a high degree of genome plasticity in the four *Pae* pathogens. They harbor a high number of PFP plasmids, which can combine and form co-integrates that increases the mobility of important virulence factors. As described above, a coronatine phytotoxin-encoding mobilizable plasmid was found to have integrated with a putative conjugative plasmid (pPae2203CORI), leading to potential increased horizontal transfer of this virulence factor that likely originates from a chromosomal location. Similarly, the gene encoding the T3SE HopAB1 was found within a transposon that actively forms extrachromosomal molecules and can thus transpose to e.g. conjugative plasmids within the bacterial host. The genome plasticity and mobility of virulence genes may be a consequence of the likely recent introduction of *Pae* to Europe from India (Green et al., 2010). Based on genomic evidence, it was proposed that only a single introduction event has happened (Green et al., 2010), explaining the high sequence similarity between *Pae* isolates. The plasmid variation described in this study shows that *Pae* continues to evolve as a plant pathogen and that intraspecies genomic rearrangements leading to potential horizontal gene transfer of virulence genes is important for continuous adaptation. *Pae* are known as tree trunk pathogens causing HCBC, but the conserved presence of coronatine-encoding genes hypothetically suggests that these bacteria enter the tree tissue via leaves through manipulation of the stomata (Geng et al., 2014).

CRediT authorship contribution statement

Tue Kjærgaard Nielsen: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Visualization, Writing – original draft, Writing – review & editing. **Caroline S. Winther-Have:** Data curation, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review & editing. **Iben Margrete Thomsen:** Methodology, Writing – original draft, Writing – review & editing. **Robert W. Jackson:** Investigation, Resources, Validation, Writing – review & editing. **Mojgan Rabiey:** Investigation, Resources, Validation, Writing – review & editing. **Rosanna Catherine Hennessy:** Methodology, Writing – original draft, Writing – review & editing. **Frederik Bak:** Methodology, Writing – original draft, Writing – review & editing. **Witold Kot:** Methodology, Writing – original draft, Writing – review & editing. **Mette Haubjerg Nicolaisen:** Methodology, Writing – original draft, Writing – review & editing. **Alexander Byth Carstens:** Methodology, Writing – original draft, Writing – review & editing. **Lars Hestbjerg Hansen:** Funding acquisition, Investigation, Methodology, Resources, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

The authors declare no competing interests.

Data availability

The genome sequences are available under the following GenBank accession numbers: CP123204-CP123226 under the BioProject PRJNA954292.

Acknowledgements

This study was funded by a University of Copenhagen grant to boost integration of research into teaching (awarded to LHH and TKN), which has enabled this study. The authors would like to thank Dr. Sabrine Dhaouadi (University of Birmingham) with sharing information on the parallel sequencing efforts of strain *P. amygdali* pv. *aesculi* 2250.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.meegid.2023.105486.

References

Alarcón-Chaidez, F.J., Peñaloza-Vázquez, A., Ullrich, M., Bender, C.L., 1999. Characterization of plasmids encoding the Phytotoxin Coronatine in *Pseudomonas syringae*. Plasmid 42, 210–220. https://doi.org/10.1006/plas.1999.1424.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2.

Baltrus, D.A., Nishimura, M.T., Romanchuk, A., Chang, J.H., Mukhtar, M.S., Cherkis, K., Roach, J., Grant, S.R., Jones, C.D., Dangl, J.L., 2011. Dynamic evolution of pathogenicity revealed by sequencing and comparative genomics of 19 *Pseudomonas syringae* isolates. PLoS Pathog. 7, e1002132 https://doi.org/10.1371/journal.ppat.1002132.

Bardaji, L., Añorga, M., Ruiz-Masó, J.A., del Solar, G., Murillo, J., 2017. Plasmid replicons from Pseudomonas are natural chimeras of functional, exchangeable modules. Front. Microbiol. 8 https://doi.org/10.3389/fmicb.2017.00190.

Blin, K., Shaw, S., Kloosterman, A.M., Charlop-Powers, Z., van Wezel, G.P., Medema, M. H., Weber, T., 2021. antiSMASH 6.0: improving cluster detection and comparison capabilities. Nucleic Acids Res., gkab335 https://doi.org/10.1093/nar/gkab335.

Block, A., Guo, M., Li, G., Elowsky, C., Clemente, T.E., Alfano, J.R., 2010. The *Pseudomonas syringae* type III effector HopG1 targets mitochondria, alters plant development and suppresses plant innate immunity. Cell. Microbiol. 12, 318–330. https://doi.org/10.1111/j.1462-5822.2009.01396.x.

Brasihr, C.M., Strouts, R.G., 1976. New records of Phytophthora on trees in Britain. For. Pathol. 6, 129–136. https://doi.org/10.1111/j.1439-0329.1976.tb00517.x.

Chaumeil, P.-A., Mussig, A.J., Hugenholtz, P., Parks, D.H., 2019. GTDB-Tk: a toolkit to classify genomes with the genome taxonomy database. Bioinformatics. https://doi.org/10.1093/bioinformatics/btz848.

Darling, A.C.E., Mau, B., Blattner, F.R., Perna, N.T., 2004. Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res. https://doi.org/10.1101/gr.2289704.

Darling, A.E., Mau, B., Perna, N.T., 2010. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. PLoS One 5, e11147. https://doi.org/10.1371/journal.pone.0011147.

De Lorenzo, G., Ferrari, S., 2002. Polygalacturonase-inhibiting proteins in defense against phytopathogenic fungi. Curr. Opin. Plant Biol. 5, 295–299. https://doi.org/10.1016/S1369-5266(02)00271-6.

Dillon, M.M., Almeida, R.N.D., Laflamme, B., Martel, A., Weir, B.S., Desveaux, D., Guttman, D.S., 2019. Molecular evolution of *Pseudomonas syringae* type III secreted effector proteins. Front. Plant Sci. 10.

Fonseca, S., Chini, A., Hamberg, M., Adie, B., Porzel, A., Kramell, R., Miersch, O., Wasternack, C., Solano, R., 2009. (+)-7-iso-Jasmonoyl-L-isoleucine is the endogenous bioactive jasmonate. Nat. Chem. Biol. 5, 344–350. https://doi.org/10.1038/nchembio.161.

Geng, X., Jin, L., Shimada, M., Kim, M.G., Mackey, D., 2014. The phytotoxin coronatine is a multifunctional component of the virulence armament of *Pseudomonas syringae*. Planta 240, 1149–1165. https://doi.org/10.1007/s00425-014-2151-x.

Gilchrist, C.L.M., Chooi, Y.-H., 2021. Clinker & clustermap.js: automatic generation of gene cluster comparison figures. Bioinformatics 37, 2473–2475. https://doi.org/10.1093/bioinformatics/btab007.

Gimenez-Ibanez, S., Boter, M., Fernández-Barbero, G., Chini, A., Rathjen, J.P., Solano, R., 2014. The bacterial effector HopX1 targets JAZ transcriptional repressors to activate Jasmonate signaling and promote infection in Arabidopsis. PLoS Biol. 12, e1001792.

Gomila, M., Busquets, A., Mulet, M., García-Valdés, E., Lalucat, J., 2017. Clarification of taxonomic status within the *Pseudomonas syringae* species group based on a Phylogenomic analysis. Front. Microbiol. 8.

Green, S., Laue, B., Fossdal, C.G., A'Hara, S.W., Cottrell, J.E., 2009. Infection of horse chestnut (*Aesculus hippocastanum*) by *Pseudomonas syringae* pv. *aesculi* and its detection by quantitative real-time PCR. Plant Pathol. 58, 731–744. https://doi.org/10.1111/j.1365-3059.2009.02065.x.

Green, S., Studholme, D.J., Laue, B.E., Dorati, F., Lovell, H., Arnold, D., Cottrell, J.E., Bridgett, S., Blaxter, M., Huitema, E., Thwaites, R., Sharp, P.M., Jackson, R.W., Kamoun, S., 2010. Comparative genome analysis provides insights into the evolution and adaptation of *Pseudomonas syringae* pv. *aesculi* on *Aesculus hippocastanum*. PLoS One 5, e10224. https://doi.org/10.1371/journal.pone.0010224.

Guglielmini, J., Néron, B., Abby, S., Garcillán-Barcia, M., la Cruz, F., Rocha, E., 2014. Key components of the eight classes of type IV secretion systems involved in bacterial conjugation or protein secretion. Nucleic Acids Res. 42, 5715–5727. https://doi.org/10.1093/nar/gku194.

Gutiérrez-Barranquero, J.A., Cazorla, F.M., de Vicente, A., Sundin, G.W., 2017. Complete sequence and comparative genomic analysis of eight native *Pseudomonas syringae* plasmids belonging to the pPT23A family. BMC Genomics 18, 365. https://doi.org/10.1186/s12864-017-3763-x.

Harms, A., Brodersen, D.E., Mitarai, N., Gerdes, K., 2018. Toxins, targets, and triggers: an overview of toxin-antitoxin biology. Mol. Cell 70, 768–784. https://doi.org/10.1016/j.molcel.2018.01.003.

Huang, Q., Allen, C., 2000. Polygalacturonases are required for rapid colonization and full virulence of *Ralstonia solanacearum* on tomato plants. Physiol. Mol. Plant Pathol. 57, 77–83. https://doi.org/10.1006/pmpp.2000.0283.

Inkscape, 2020. Inkscape Project.

Jain, C., Rodriguez-R, L.M., Phillippy, A.M., Konstantinidis, K.T., Aluru, S., 2018. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. Nat. Commun. https://doi.org/10.1038/s41467-018-07641-9.

James, S.L., Rabiey, M., Neuman, B.W., Percival, G., Jackson, R.W., 2020. Isolation, characterisation and experimental evolution of phage that infect the horse chestnut tree pathogen, *Pseudomonas syringae* pv. *aesculi*. Curr. Microbiol. 77, 1438–1447. https://doi.org/10.1007/s00284-020-01952-1.

Khan, M., Seto, D., Subramaniam, R., Desveaux, D., 2018. Oh, the places they'll go! A survey of phytopathogen effectors and their host targets. Plant J. 93, 651–663. https://doi.org/10.1111/tpj.13780.

King, E.O., Ward, M.K., Raney, D.E., 1954. Two simple media for the demonstration of pyocyanin and fluorescin. J. Lab. Clin. Med. 44, 301–307.

Kolde, R., 2019. pheatmap: Pretty Heatmaps. R package version 1.0.12.

Kolmogorov, M., Yuan, J., Lin, Y., Pevzner, P.A., 2019. Assembly of long, error-prone reads using repeat graphs. Nat. Biotechnol. 37, 540–546. https://doi.org/10.1038/s41587-019-0072-8.

Koren, S., Walenz, B.P., Berlin, K., Miller, J.R., Bergman, N.H., Phillippy, A.M., 2017. Canu: scalable and accurate long-read assembly via adaptive k -mer weighting and repeat separation. Genome Res. 27, 722–736. https://doi.org/10.1101/gr.215087.116.

Krueger, F., 2016. Trim Galore. Babraham Bioinformatics https://doi.org/http://www.bioinformatics.babraham.ac.uk/projects/trim galore/.

La Porta, N., Hietala, A.M., Baldi, P., 2023. Bacterial diseases in forest trees. In: Forest Microbiology. Elsevier, pp. 139–166. https://doi.org/10.1016/B978-0-443-18694-3.00001-8.

Landgraf, A., Weingart, H., Tsiamis, G., Boch, J., 2006. Different versions of *Pseudomonas syringae* pv. tomato DC3000 exist due to the activity of an effector transposon. Mol. Plant Pathol. 7, 355–364. https://doi.org/10.1111/j.1364-3703.2006.00343.x.

Li, H., 2018. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics 34, 3094–3100. https://doi.org/10.1093/bioinformatics/bty191.

Ma, Z., Smith, J.J., Zhao, Y., Jackson, R.W., Arnold, D.L., Murillo, J., Sundin, G.W., 2007. Phylogenetic analysis of the pPT23A plasmid family of *Pseudomonas syringae*. Appl. Environ. Microbiol. 73, 1287–1295. https://doi.org/10.1128/AEM.01923-06.

Manni, M., Berkeley, M.R., Seppey, M., Simão, F.A., Zdobnov, E.M., 2021. BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. Mol. Biol. Evol. 38, 4647–4654. https://doi.org/10.1093/molbev/msab199.

Melotto, M., Underwood, W., Koczan, J., Nomura, K., He, S.Y., 2006. Plant stomata function in innate immunity against bacterial invasion. Cell 126, 969–980. https://doi.org/10.1016/j.cell.2006.06.054.

Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G.A., Sonnhammer, E.L.L., Tosatto, S.C.E., Paladin, L., Raj, S., Richardson, L.J., Finn, R.D., Bateman, A., 2021. Pfam: the protein families database in 2021. Nucleic Acids Res. 49, D412–D419. https://doi.org/10.1093/nar/gkaa913.

Nielsen, T.K., Rasmussen, M., Demanèche, S., Cecillon, S., Vogel, T.M., Hansen, L.H., 2017. Evolution of sphingomonad gene clusters related to pesticide catabolism revealed by genome sequence and mobilomics of *Sphingobium herbicidovorans* MH. Genome Biol. Evol. https://doi.org/10.1093/gbe/evx185.

Nielsen, T.K., Horemans, B., Lood, C., T'Syen, J., van Noort, V., Lavigne, R., Ellegaard-Jensen, L., Hylling, O., Aamand, J., Springael, D., Hansen, L.H., 2021. The complete genome of 2,6-dichlorobenzamide (BAM) degrader Aminobacter sp. MSH1 suggests a polyploid chromosome, phylogenetic reassignment, and functions of plasmids. Sci. Rep. 11, 18943. https://doi.org/10.1038/s41598-021-98184-5.

Nimchuk, Z.L., Fisher, E.J., Desveaux, D., Chang, J.H., Dangl, J.L., 2007. The HopX (AvrPphE) family of *Pseudomonas syringae* Type III effectors require a catalytic triad and a novel N-terminal domain for function. Mol. Plant-Microbe Interact. 20, 346–357. https://doi.org/10.1094/MPMI-20-4-0346.

Nowell, R.W., Laue, B.E., Sharp, P.M., Green, S., 2016. Comparative genomics reveals genes significantly associated with woody hosts in the plant pathogen *Pseudomonas syringae*. Mol. Plant Pathol. 17, 1409–1424. https://doi.org/10.1111/mpp.12423.

O'Brien, H.E., Thakur, S., Guttman, D.S., 2011. Evolution of plant pathogenesis in *Pseudomonas syringae*: a genomics perspective. Annu. Rev. Phytopathol. 49, 269–289. https://doi.org/10.1146/annurev-phyto-072910-095242.

R Core Team, R.D.C, 2016. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. https://doi.org/10.1017/CBO9781107415324.004.

Redondo-Salvo, S., Bartomeus-Peñalver, R., Vielva, L., Tagg, K.A., Webb, H.E., Fernández-López, R., de la Cruz, F., 2021. COPLA, a taxonomic classifier of plasmids. BMC Bioinform. 22, 390. https://doi.org/10.1186/s12859-021-04299-x.

Robertson, J., Nash, J.H.E., 2018. MOB-suite: software tools for clustering, reconstruction and typing of plasmids from draft assemblies. Microb. Genom. https://doi.org/10.1099/mgen.0.000206.

Rodriguez-Palenzuela, P., Burr, T.J., Collmer, A., 1991. Polygalacturonase is a virulence factor in *Agrobacterium tumefaciens* biovar 3. J. Bacteriol. 173, 6547–6552. https://doi.org/10.1128/jb.173.20.6547-6552.1991.

Ruinelli, M., Blom, J., Smits, T.H.M., Pothier, J.F., 2019. Comparative genomics and pathogenicity potential of members of the *Pseudomonas syringae* species complex on Prunus spp. BMC Genomics 20, 172. https://doi.org/10.1186/s12864-019-5555-y.

Seemann, T., 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30, 2068–2069. https://doi.org/10.1093/bioinformatics/btu153.

Sénéchal, F., Wattier, C., Rustérucci, C., Pelloux, J., 2014. Homogalacturonan-modifying enzymes: structure, expression, and roles in plants. J. Exp. Bot. 65, 5125–5160. https://doi.org/10.1093/jxb/eru272.

Sesma, A., Sundin, G.W., Murillo, J., 1998. Closely related plasmid replicons coexisting in the Phytopathogen *Pseudomonas syringae* show a mosaic organization of the replication region and altered incompatibility behavior. Appl. Environ. Microbiol. 64, 3948–3953. https://doi.org/10.1128/AEM.64.10.3948-3953.1998.

Siguier, P., 2006. ISfinder: the reference Centre for bacterial insertion sequences. Nucleic Acids Res. 34, D32–D36. https://doi.org/10.1093/nar/gkj014.

Stavrinides, J., Guttman, D.S., 2004. Nucleotide sequence and evolution of the five-plasmid complement of the phytopathogen Pseudomonas syringae pv. maculicola ES4326. J. Bacteriol. 186, 5101–5115. https://doi.org/10.1128/JB.186.15.5101-5115.2004.

Umarov, R., Solovyev, V., 2017. Recognition of prokaryotic and eukaryotic promoters using convolutional deep learning neural networks. PLOS ONE 12, e0171410. https://doi.org/10.1371/journal.pone.0171410.

Vaser, R., Šikić, M., 2021. Time- and memory-efficient genome assembly with raven. Nat. Comput. Sci. 1, 332–336. https://doi.org/10.1038/s43588-021-00073-4.

Vaser, R., Sović, I., Nagarajan, N., Šikić, M., 2017. Fast and accurate de novo genome assembly from long uncorrected reads. Genome Res. 27, 737–746. https://doi.org/10.1101/gr.214270.116.

Walker, B.J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., Cuomo, C.A., Zeng, Q., Wortman, J., Young, S.K., Earl, A.M., 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS One. https://doi.org/10.1371/journal.pone.0112963.

Wang, X., Hou, S., Wu, Q., Lin, M., Acharya, B.R., Wu, D., Zhang, W., 2017. <scp>IDL</scp> 6- <scp>HAE</scp> / <scp>HSL</scp> 2 impacts pectin degradation and resistance to *Pseudomonas syringae* pv tomato <scp>DC</scp> 3000 in Arabidopsis leaves. Plant J. 89, 250–263. https://doi.org/10.1111/tpj.13380.

Webber, J.F., Parkinson, N.M., Rose, J., Stanford, H., Cook, R.T.A., Elphinstone, J.G., 2008. Isolation and identification of *Pseudomonas syringae* pv. *aesculi* causing bleeding canker of horse chestnut in the UK. Plant Pathol. 57, 368. https://doi.org/10.1111/j.1365-3059.2007.01754.x.

Wei, H.-L., Chakravarthy, S., Mathieu, J., Helmann, T.C., Stodghill, P., Swingle, B., Martin, G.B., Collmer, A., 2015. *Pseudomonas syringae* pv. Tomato DC3000 Type III secretion effector Polymutants reveal an interplay between HopAD1 and AvrPtoB. Cell Host Microbe 17, 752–762. https://doi.org/10.1016/j.chom.2015.05.007.

Wick, R.R., Schultz, M.B., Zobel, J., Holt, K.E., 2015. Bandage: interactive visualization of de novo genome assemblies. Bioinformatics 31, 3350–3352. https://doi.org/10.1093/bioinformatics/btv383.

Wick, R.R., Judd, L.M., Gorrie, C.L., Holt, K.E., 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput. Biol. https://doi.org/10.1371/journal.pcbi.1005595.

Xie, Y., Liu, W., Shao, X., Zhang, W., Deng, X., 2020. Signal transduction schemes in *Pseudomonas syringae*. Comput. Struct. Biotechnol. J. 18, 3415–3424. https://doi.org/10.1016/j.csbj.2020.10.039.

Xin, X.-F., Kvitko, B., He, S.Y., 2018. *Pseudomonas syringae*: what it takes to be a pathogen. Nat. Rev. Microbiol. 16, 316–328. https://doi.org/10.1038/nrmicro.2018.17.

Yan, Q., Rogan, C.J., Pang, Y.-Y., Davis, E.W., Anderson, J.C., 2020. Ancient co-option of an amino acid ABC transporter locus in *Pseudomonas syringae* for host signal-dependent virulence gene regulation. PLoS Pathog. 16, e1008680 https://doi.org/10.1371/journal.ppat.1008680.

Zhang, Z.C., He, B., Sun, S., Zhang, X., Li, T., Wang, H.H., Xu, L.R., Afzal, A.J., Geng, X. Q., 2021. The phytotoxin COR induces transcriptional reprogramming of photosynthetic, hormonal and defence networks in tomato. Plant Biol. 23, 69–79. https://doi.org/10.1111/plb.13239.

Zhao, Y., Ma, Z., Sundin, G.W., 2005. Comparative genomic analysis of the pPT23A plasmid family of *Pseudomonas syringae*. J. Bacteriol. 187, 2113–2126. https://doi.org/10.1128/JB.187.6.2113-2126.2005.

Zheng, X., Spivey, N.W., Zeng, W., Liu, P.-P., Fu, Z.Q., Klessig, D.F., He, S.Y., Dong, X., 2012. Coronatine promotes *Pseudomonas syringae* virulence in plants by activating a signaling cascade that inhibits salicylic acid accumulation. Cell Host Microbe 11, 587–596. https://doi.org/10.1016/j.chom.2012.04.014.

Zhou, H., Lin, J., Johnson, A., Morgan, R.L., Zhong, W., Ma, W., 2011. *Pseudomonas syringae* Type III effector HopZ1 targets a host enzyme to suppress Isoflavone biosynthesis and promote infection in soybean. Cell Host Microbe 9, 177–186. https://doi.org/10.1016/j.chom.2011.02.007.