

ORIGINAL ARTICLE

Journal of
Phytopathology

WILEY

Dressing like a pathogen: Comparative analysis of different *Pseudomonas* genomospecies wearing different features to infect *Corylus avellana*

Silvia Turco  | Leonardo Zuppante | Mounira Inas Drais  | Angelo Mazzaglia

Dipartimento di Scienze Agrarie e Forestali, Università degli Studi della Tuscia, Viterbo, Italy

Correspondence

Angelo Mazzaglia, Dipartimento di Scienze Agrarie e Forestali, Università degli Studi della Tuscia, Via S. Camillo de Lellis snc, 01100 Viterbo, Italy.
Email: angmazza@unitus.it

Funding information

Ministero dell'Istruzione, dell'Università e della Ricerca, Grant/Award Number: Law232/2016; Università degli Studi della Tuscia

Abstract

During the last decade, the European hazelnut (*Corylus avellana*) has been threatened by several bacterial and fungal infections. In particular, at least four different *Pseudomonas* species have developed the capability to converge into the same host. However, one of these species, *Pseudomonas avellanae*, may infect other hosts as well. Here, we newly provide the complete genome sequence of two *Pseudomonas avellanae* strains obtained by a combination of long Oxford Nanopore Technologies reads (ONT) and short Illumina reads. These have been included in a comparative analysis with all the genomes of *Pseudomonas* strains known to infect *Corylus avellana* available on the NCBI database, and the genomes of *Pseudomonas avellanae* strains infecting other host plants, like *Prunus*, *Actinidia* and from the river epilithon. The *Pseudomonas* strains clustered according to their taxonomy, host and geographic origins. The identification of specific traits related to each population cluster could give us insights regarding the host–pathogen interaction and the host range specificity. Indeed, the *Pseudomonas avellanae* strains infecting hazelnut have a peculiar set of three type III secretion effectors, while the *Pseudomonas avellanae* infecting *Prunus* and *Actinidia* are carrying the genomic WHOP island, necessary for woody-host infection. The *syringae* cluster, instead, contains more phytotoxins, the ice nucleation cluster, but less effectors. Thus, here, we confirmed that the convergence into the same host was possible due to different unrelated mechanisms of infection.

KEYWORDS

bacterial canker, comparative genomics, hazelnut, hybrid assembly, *Pseudomonas*

1 | INTRODUCTION

Corylus avellana (hazelnut) is an important shrub species of the *Betulaceae* family. By covering a surface of 89,913 ha spread among Campania, Latium, Sicily and Piedmont regions, Italy is the second

largest world hazelnut producer, with an overall yield of 140,563 t recorded in 2020 (ISTAT, 2021). Among other pathogens, hazelnut is threatened by a group of bacteria, all belonging to the genus *Pseudomonas*, i.e., *Pseudomonas avellanae* (Pav), *Pseudomonas syringae* pv. *avellanae* (Psav), *Pseudomonas syringae* pv. *coryli*

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2022 The Authors. *Journal of Phytopathology* published by Wiley-VCH GmbH.

(Psc) and *Pseudomonas syringae* pv. *syringae* (Pss) (Marcelletti & Scortichini, 2015), which may cause severe damage to the plants that lead to production drops and thus, important economic losses.

The disease related to these pathogens is known as bacterial canker of hazelnut, which was reported for the first time in Greece in 1976 and associated at the time to a new *Pseudomonas* pathovar named *Pseudomonas syringae* pv. *avellanae* (Psallidas & Panagopoulos, 1979, Psallidas 1993). The syndrome is associated with the development of cankers on the stems, the discolouration of twigs and stems, with progressive tip branches dieback which eventually would spread to the rest of the plant (Psallidas & Panagopoulos, 1979, Psallidas, 1984). The mechanism of infection requires leaf scars through which the pathogen enters the tree during the fall season and systemically colonizes the plant (Scortichini & Lazzari, 1996). The occurrence of the same bacterium, associated with the bacterial canker and named as 'Moria del nocciolo', was lately reported in 1994 in Central Italy (Scortichini & Tropiano, 1994), while a less virulent situation was reported in Northwest Italy (Piedmont region) and in Campania and Sardinia regions as well (Scortichini & Morone 1997).

However, significant differences between the causal agent of the canker of hazelnut and the *Pseudomonas syringae* group were observed, leading to the reclassification of *Pseudomonas syringae* pv. *avellanae* that was elevated to new species and named *Pseudomonas avellanae* (Janse et al., 1996). Further studies proved that the strains from Northern Greece were closely related to those isolated in Central Italy but also distinguishable from them by repetitive PCR (Scortichini et al., 1998). When compared to the strains isolated from Northern Italy, a higher heterogeneity was observed, suggesting the presence of two distinct clusters of bacteria infecting hazelnut in Italy. The first cluster resulted to be more virulent, it was potentially restricted to hazelnut and was related to *Pseudomonas avellanae* (Scortichini et al. 2002). The second cluster was detected in Piedmont, Campania, Latium, Sicily and Sardinia inducing mild symptoms on plants and classified as *Pseudomonas syringae* pv. *coryli* (Scortichini et al. 2002, 2005). Nonetheless, further taxonomic studies allowed us to distinguish the third cluster of aggressive strains isolated only in Italy that was classified as *Pseudomonas syringae* pv. *avellanae* (Scortichini et al., 2013). A fourth cluster constituted by *Pseudomonas syringae* pv. *syringae* strains was isolated in Campania, Latium, Piedmont, Sardinia and Sicily regions and, together with *Pseudomonas syringae* pv. *coryli*, is able to induce only mild symptoms, limited to twig dieback (Scortichini et al., 2002; Loreti et al., 2009).

Hence, these four groups of pseudomonads have been characterized and distinguished by several molecular methods like rep-PCR, MLST analysis and 16S-23S rRNA genotyping and no recombination events between them were found (Marcelletti & Scortichini, 2015). Despite the peculiar traits identified either in *Pseudomonas avellanae*, belonging to genomospecies 8, or within the pathovar *syringae*, *coryli* and *avellanae* strains, all from genomospecies 1, they converged onto hazelnut as host (Marcelletti & Scortichini, 2015).

On the other side, some *Pseudomonas avellanae* strains, previously identified as *Pseudomonas syringae* pv. *morsprunorum* race 2, have been isolated from *Prunus avium* as well (Hulin et al., 2018,

2020). This apparent contradiction could be explained by the observation that the host-range determinants of *Pseudomonas* selectivity are mainly linked to virulence-related factors like the type III secretion system and their relative secreted effectors, phytotoxins, WHOP (Woody hosts and *Pseudomonas*) and woody associated enzymes, as reported in a recent study by Dillon et al. (2021).

This work aims to further clarify the differences between the four *Pseudomonas* clusters that have converged into *Corylus avellana* and highlight their peculiarities by performing a comparative genomic analysis among all the available genome sequences of *Pseudomonas avellanae*, *Pseudomonas syringae* pv. *coryli*, *syringae* and *avellanae* strains isolated from Greece and Italy. Furthermore, two additional *Pseudomonas avellanae* strains, NCPPB 4222 from Italy and NCPPB 3491 from Greece, have also been sequenced using both Illumina and Nanopore sequencing technologies and included in the analysis. Plus, to give a broader picture of the host-specificity and geographic distribution, the genomes of *Pseudomonas avellanae* strains isolated in the United Kingdom and New Zealand from *Prunus avium*, of a Japanese strain isolated from *Actinidia chinensis*, previously known as *Pseudomonas syringae* pv. *actinidiae*, and of a strain isolated from river epilithon (a rock-attached biofilms composed of algae, diatoms, rotifers, bacteria and nematodes) in Palisade, USA (Morris et al., 2007), were included in the analysis as well.

2 | MATERIALS AND METHODS

2.1 | DNA isolation and genome sequencing

Two *Pseudomonas avellanae* strains from in-house collection have been selected for genome sequencing and downstream analysis: the Italian NCPPB 4222 and the Greek NCPPB 3491. An aliquot of the stock stored at -80°C was revitalized and grown on King's B (KB) medium (King et al., 1954) at 27°C . High molecular weight (HMW) DNA was extracted following the protocol of Mayjonade et al. (2016), with few modifications. Briefly, a suspension of liquid culture with an OD above 3 was precipitated by centrifugation at $12,000g$ for 1 min and the pellet was resuspended in $600\mu\text{l}$ of lysis buffer (1% PVP 40, 1% sodium metabisulfite, 0.5 M sodium chloride, 100mM Tris 1 M pH 8, 50mM EDTA 0.5 M, 1.25% SDS 20%) and $4\mu\text{l}$ of RNase A (100mg/ml QIAGEN). The solution was incubated at 50°C for 30 min, and then, $600\mu\text{l}$ of lysis buffer was added. After 20 more min at 50°C , $400\mu\text{l}$ of potassium acetate 5 M were added to the solution and incubated on ice for 1 min. To get phase separation, the solution was centrifuged for 20 min at $5000g$, the upper phase was transferred into a new tube and $500\mu\text{l}$ of cold isopropanol was added. After 5 min on ice, the DNA was precipitated at $10,000g$ for 10 min. The isopropanol was removed, and the pellet was washed twice with 70% ethanol. Lastly, the DNA was resuspended in a $50\mu\text{l}$ elution buffer. The DNA was quantified through Invitrogen Qubit fluorimeter (Thermo Fisher Scientific) and its purity was evaluated in a 0.5% agarose gel electrophoresis run and with Thermo Scientific Multiskan GO (Thermo Fischer Scientific).

Finally, 500ng of both *Pseudomonas avellanae* strains were sequenced with the Illumina NovaSeq 6000 S2 using paired-end sequencing at Novogene (Cambridge, United Kingdom). Three microgram of DNA, instead, were sequenced onto MinION Mk1C device (ONT) using two Flongle flow cell R9.4.1 (ONT), after End Repair and dA tailing using NEBNext Ultra II End Repair/dA Tailing Module (New England BioLabs) and library preparation using the SQK-LSK109 kit (ONT).

2.2 | Genome assembly and annotation

The quality of the paired-end Illumina reads was evaluated using FastQC (Andrews, 2010). Guppy v4.3.4 (ONT) within MinKNOW v. 21.02.2 (ONT) installed on the MK1C device, was used for the base calling of the long ONT reads, while MinKNOW itself provided the report with the summary run and sequencing stats. Both Illumina and ONT reads were used as input data for the de novo hybrid assembly performed with Unicycler v0.4.9b using default parameters (Wick et al., 2017). The assembly performance was evaluated using QUAST 5.0 (Gurevich et al. 2013), and a graphical representation of the assembled genomes and related plasmids was obtained using CGview (Petkau et al., 2010). Genome annotation was done using Prokka v1.14.5 (Seemann, 2014).

2.3 | Phylogenetic and comparative genomics analysis

For phylogenetic and comparative genomics analysis, the complete or draft genomes of 15 strains of *Pseudomonas avellanae* and six *Pseudomonas syringae* strains known to infect hazelnut were downloaded from NCBI genomes and listed in Table 1.

For a phylogenetic MLST analysis, the sequences of seven house-keeping genes (*gapA*, *gltA*, *gyrB*, *rpoD*, *acnB*, *pfkB* and *pgi*) were retrieved from each strain under comparison, concatenated and aligned using Muscle (Edgar, 2004). The alignment was used as input to build an ML tree with RAXML-HPC, using the GTRCATI algorithm as a substitution model and 1000 bootstraps (Stamatakis, 2014). The tree was then visualized with FigTree v1.4.4, and the picture was adjusted with Inkscape. For an unbiased comparison, all the genomes listed in Table 1 were annotated with Prokka v1.14.5 (Seemann, 2014) and reordered with MAUVE (Darling et al., 2004) towards *Pseudomonas avellanae* BPIC 631, *Pseudomonas syringae* pv. *avellanae* ISPAve013 and *Pseudomonas syringae* pv. *coryli* strain NCPPB 4273 and *Pseudomonas syringae* pv. *syringae* CRAFRU12 as reference genomes, respectively. The reordered genomes were then aligned by progressive MAUVE algorithm with default parameters (Darling et al., 2004).

2.4 | Pangenome analysis

Roary v1.7.7 (Page et al., 2015) and Anvi'o v.7 (Eren et al., 2015) were used for a separated pangenome analysis. Roary provided the

core genome alignment that was used to construct an ML phylogenetic tree with RAXML-HPC, the GTRCATI algorithm as substitution model and 1000 bootstraps (Stamatakis 2014). Instead, within Anvi'o all the genomes were also annotated, besides Prokka, using NCBI's Clusters of Orthologous Groups (COG) database (Tatusov et al., 2000), Pfam (El-Gebali et al., 2019). The pangenome was created using NCBI blastp for amino acid sequence similarity search, the default minbit heuristic set to 0.5 and the MCL inflation parameter set to 10. Average Nucleotide Identity (ANI) was evaluated within Anvi'o using ANIm and the functional enrichment analysis function was used to identify gene clusters peculiar to either defined clusters, such as the phylogenetic groups or the host origin.

2.5 | Identification of virulence-related genes

To identify the presence of virulence-related features, such as Type II, III, IV and VI secretion systems, type III secretion effectors, phytohormones, ice nucleation factors and toxins, the sequences of the features were blasted using NCBI blastn (e-value $1e-5$) towards the genomes under exam. Each feature was considered present if it showed a minimum 70% of identifying and at least 80% of query coverage and the results were used to create a presence/absence binary matrix that was imported into the R environment v3.4.4, where the pheatmap package v1.0.12 (Kolde, 2019) was used to plot the heatmap.

The sequences of the type III secretion effectors were retrieved from Dillon et al. (2019), while the secretion systems, virulence-related features, phytohormones, toxins and WHOP genes were obtained from Rodriguez-Palenzuela et al. (2010), Caballo-Ponce et al. (2017) and from the annotated genome of *Pseudomonas avellanae* R2leaf strain.

3 | RESULTS

3.1 | Hybrid de novo assembly and genomes annotation

High molecular weight DNA was extracted from the Italian strain *Pseudomonas avellanae* NCPPB 4222 and from the Greek strain NCPPB 3491 following the protocol described in Section 2.1 and sequenced using both Oxford Nanopore Technologies (ONT) and Illumina methods. A total of 37k ONT reads (N50 of 17.15k) and 13.7 M Illumina reads were obtained from NCPPB 4222, while 194k reads (N50 of 20.39k) and 15.7 M Illumina reads were obtained by NCPPB 3491. The reads combined in a hybrid de novo assembly allowed us to obtain the complete circular genomes and related plasmids for both strains. The NCPPB 4222 strain resulted in a genome of 6,062,886bp with a GC content of 58.59% and 2 additional plasmids of 110,191bp and 50,625bp, respectively. The Greek strain was arranged in one single circular chromosome of 6,217,249bp (GC content 58.31%) and 3 plasmids of 105,919bp, 87,243bp and 71,760bp, respectively (Figure 1).

TABLE 1 List of the *Pseudomonas* strains used in the comparative genomics and phylogenetic analysis

Species	Pathovar	Strain name	Host	Geographic origin	Accession number
<i>Pseudomonas avellanae</i>	-	CC1416	Epilithon	Montana (USA)	GCA_000452845.2
<i>Pseudomonas avellanae</i>	-	MAFF 212061	<i>Actinidia chinensis</i>	Japan	GCA_002261495.1
<i>Pseudomonas avellanae</i>	-	R2leaf	<i>Prunus avium</i>	United Kingdom	GCA_002905795.2
<i>Pseudomonas avellanae</i>	-	R25255	<i>Prunus avium</i>	United Kingdom	GCA_002905695.1
<i>Pseudomonas avellanae</i>	-	R29095	<i>Prunus avium</i>	United Kingdom	GCA_002917235.1
<i>Pseudomonas avellanae</i>	-	R2sc214	<i>Prunus avium</i>	United Kingdom	GCA_002905955.1
<i>Pseudomonas avellanae</i>	-	R25260	<i>Prunus avium</i>	United Kingdom	GCA_002905775.1
<i>Pseudomonas avellanae</i>	-	ICMP 3690	<i>Prunus avium</i>	New Zealand	GCA_002699825.1
<i>Pseudomonas avellanae</i>	-	BPIC 631	<i>Corylus avellana</i>	Greece	GCA_000302915.1
<i>Pseudomonas avellanae</i>	-	JCM 11937	<i>Corylus avellana</i>	Greece	GCA_014646595.1
<i>Pseudomonas avellanae</i>	-	ICMP 9749	<i>Corylus avellana</i>	Greece	GCA_003702325.1
<i>Pseudomonas avellanae</i>	-	NCPPB 3491	<i>Corylus avellana</i>	Greece	CP091143-CP091146
<i>Pseudomonas avellanae</i>	-	PaVt10	<i>Corylus avellana</i>	Italy	GCA_000935825.1
<i>Pseudomonas avellanae</i>	-	CRAFRUed2	<i>Corylus avellana</i>	Italy	GCA_000978625.2
<i>Pseudomonas avellanae</i>	-	CRAFRUee3	<i>Corylus avellana</i>	Italy	GCA_000978635.2
<i>Pseudomonas avellanae</i>	-	CRAFRUec1	<i>Corylus avellana</i>	Italy	GCA_000441975.1
<i>Pseudomonas avellanae</i>	-	NCPPB 4222	<i>Corylus avellana</i>	Italy	CP091147-CP091149
<i>Pseudomonas syringae</i>	<i>avellanae</i>	ISPaVe013	<i>Corylus avellana</i>	Italy	GCA_000302795.1
<i>Pseudomonas syringae</i>	<i>avellanae</i>	ISPaVe037	<i>Corylus avellana</i>	Italy	GCA_000302815.1
<i>Pseudomonas syringae</i>	<i>coryli</i>	NCPPB 4273	<i>Corylus avellana</i>	Italy	GCA_000972175.1
<i>Pseudomonas syringae</i>	<i>coryli</i>	ICMP 17001	<i>Corylus avellana</i>	Italy	GCA_001400235.1
<i>Pseudomonas syringae</i>	<i>syringae</i>	CRAFRU11	<i>Corylus avellana</i>	Italy	GCA_000972155.1
<i>Pseudomonas syringae</i>	<i>syringae</i>	CRAFRU12	<i>Corylus avellana</i>	Italy	GCA_000972195.1

The annotation performed with Prokka identified 5473 CDS on NCPPB 4222 chromosome, with 63 tRNAs, 16 rRNAs and 1 tmRNA. 118 and 66 CDS were found in the two plasmids, respectively. Instead, 5642 CDS were found on the NCPPB 3491 chromosome, with 59 tRNAs, 16 rRNAs and 1 tmRNA. 114, 98 and 88 CDS were annotated on the plasmids. The genome of the NCPPB 4222 strain and related plasmids has been deposited under the accession numbers CP091147-CP091149, while NCPPB 3491 and its three plasmids can be found under accession numbers CP091143-CP091146.

3.2 | Phylogenetic relationship among pseudomonads

The phylogenetic relationships among the *Pseudomonas avellanae* and *Pseudomonas syringae* strains compared in this study (Table 1) were evaluated based on the alignment of both the core genome identified by Roary (Figure 2) and of seven housekeeping genes (Figure S1). The results are consistent between the two different analyses, showing *Pseudomonas avellanae* clearly separated from *P. syringae* species and pathovars. Four distinct sub-clusters are evident: (i) the Group I, including the *P. avellanae* strains from Italy and Greece, perfectly divided into two subgroups; (ii) the Group

II, including the strains deriving from the United Kingdom and the ICMP 3690 from New Zealand, all isolated from *Prunus avium*, together with the Japanese MAFF 212061 isolated from kiwifruit; (iii) the American CC1416 strain that clustered on its own as Group III; (iv) Group IV, including all the strains belonging to *Pseudomonas syringae* pathovars. The phylogenetic groups identified in this analysis were further used as cluster categories in the pangenome analysis described below.

3.3 | Pangenome analysis

The pangenome analysis performed with Anvi'o identified a total of 127,767 genes arranged in 10,923 gene clusters (Figure 3). Among those, 3944 gene clusters are shared between all the *Pseudomonas* strains and define the 'core' genome, with 2746 of them present in one copy per genome (SCG, Single Core Genome). Besides the 3809 accessory gene clusters, the number of singletons, meaning those gene clusters present only in one strain and thus representing the strain peculiarity, resulted to be 3170 in total. Among those, *Pseudomonas avellanae* strain MAFF 212061, together with BPIC 631, are the ones with the highest number of singletons (938 and 584 gene clusters, respectively), followed by *Pseudomonas avellanae* CC1416 strain with 260 singletons. The two strains sequenced

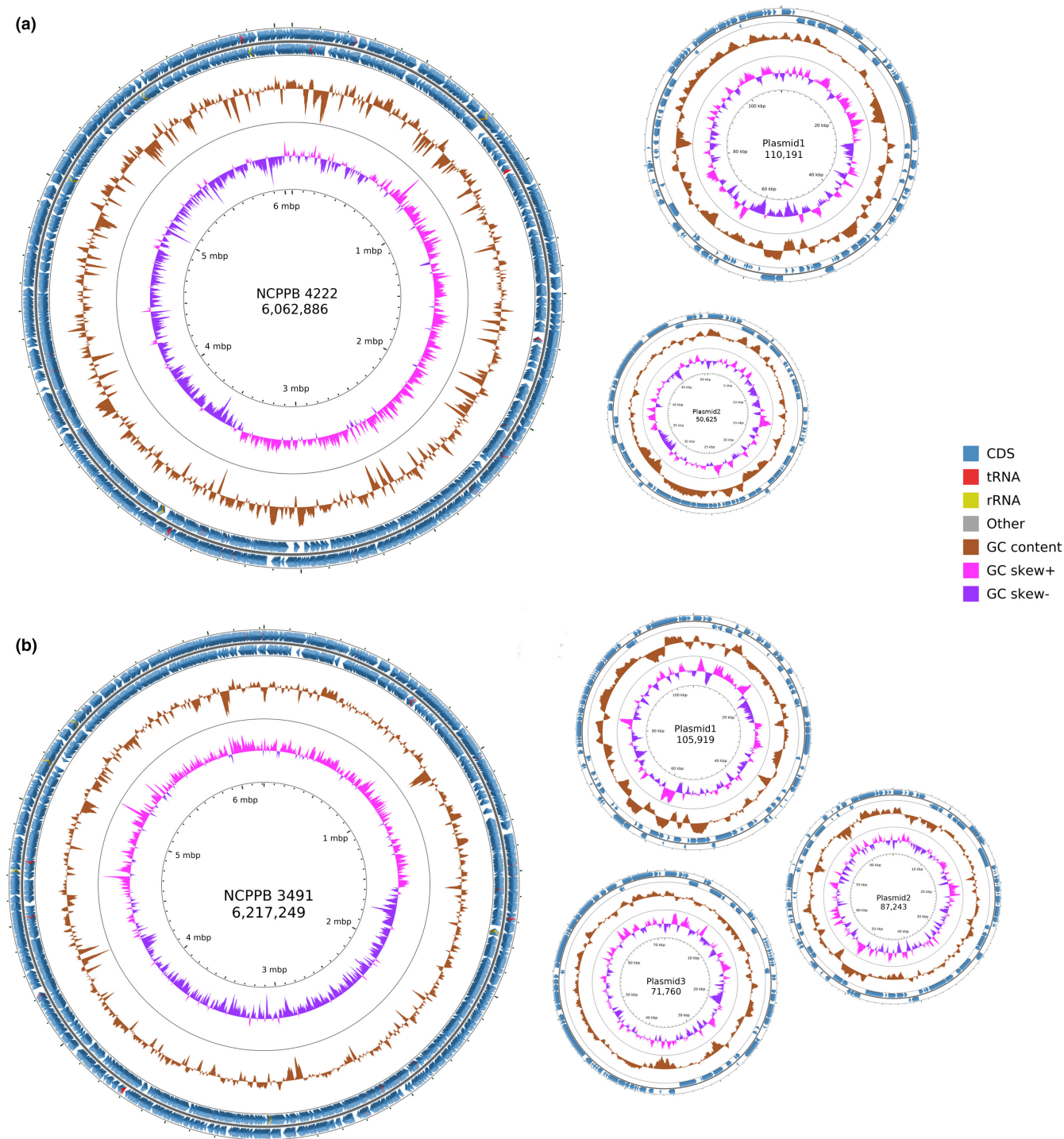


FIGURE 1 Graphical circular maps of the two de novo assembled genomes visualized through CG view. From outside to the centre: ring 1 and ring 2 show the protein-coding sequence (CDS) identified by Prokka in both forward and reverse strand, with the tRNA represented in red (tRNA); ring 3 shows the GC content percentage and ring 4 the GC skew. (a) Graphical representation of the Italian Pav NCPPB 4222 and its two plasmids. (b) Graphical representation of the Greek Pav NCPPB 3491 and its three plasmids

by us, NCPPB 4222 and NCPPB 3491, instead, showed 7 and 13 singletons, respectively. Among the 3170 singleton gene clusters, 2177 could not be assigned to any COG or Pfam category and were annotated by Prokka as a hypothetical protein. Instead, the remaining singletons are spread among different COG categories (Figure 3). Interestingly, MAFF 212061 showed several singletons related to

intracellular trafficking and secretion (U), defence mechanisms (V) and mobilome (X): the type VI protein secretion system component Hcp (secreted cytotoxin) and VasF, the multidrug efflux pump subunit AcrB, the VirB3, VirB8, VirB9 and TrbL components of the Type IV secretory pathway, the GspD/PulD secretin protein of the type II secretory pathway, streptomycin 6-kinase, Virulence-associated

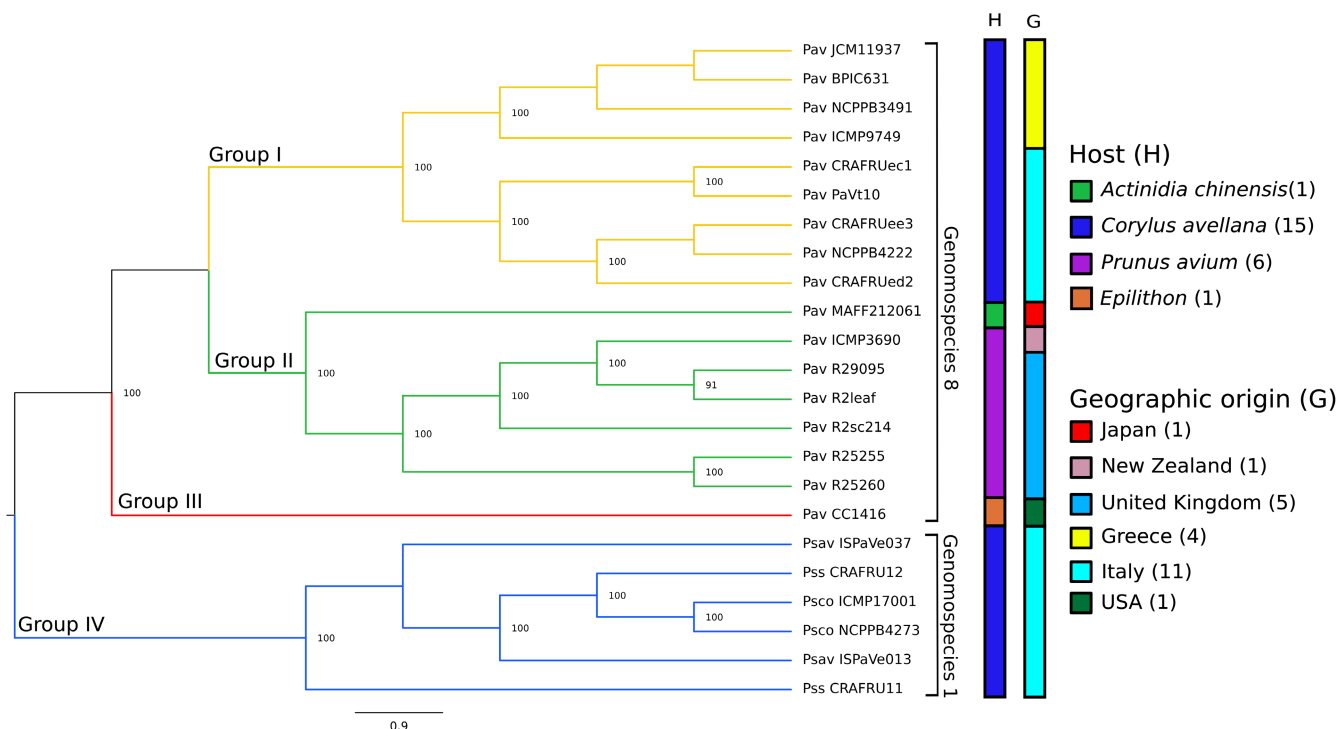


FIGURE 2 Phylogenetic analysis among the *Pseudomonas* strains under study based on the alignment of the core genome identified by Roary. The four different phylogenetic groups, belonging to either genomospecies 8 or genomospecies 1, are shown with different colours: Group I (in yellow) includes the Pav strains from Greece and Italy isolated from *Corylus avellana*; group II (in green) includes the Pav strains from the United Kingdom, New Zealand and Japan isolated from *Prunus avium* and *Actinidia chinensis*; group III is only represented by Pav CC1416 isolated from the river epilithon; group IV includes the *Pseudomonas syringae* strains isolated in Italy from hazelnut as well

protein VagC, EscV protein of the Type III secretion system, transposase and phage-related proteins (Table S1). The American CC1416 strain, instead, besides several singletons clusters associated with the general functional prediction only category R, showed an enrichment in the transcriptional regulator and DNA-binding protein from the COG category K. The majority of the BPIC 631 singletons could not be assigned to any functional annotation, besides 12 singletons associated to mobilome (COG X), signal transduction mechanisms (T), permease component (E), endonuclease (V), possible tellurite resistance (S), transport and transcriptional regulator (R, K, L).

These strains' peculiarities could explain the Average Nucleotide Identity (ANI) between the strains, as shown in the heatmap in Figure 3, where CC1416, MAFF 212061 and BPIC 631 are the most dissimilar. Overall, the colour gradient of the ANIm heatmap is in line with the phylogenetic trees in Figure 2 and Figure S1, where the *Pseudomonas avellanae* and the *Pseudomonas syringae* strains form two different main clades, with the subclades among the *avellanae* strains.

3.4 | Enrichment analysis of Pan groups

In order to find out the features in common within the phylogenetic clusters and the differences between the clusters, an enrichment analysis using Anvi'o *anvi-compute-functional-enrichment* function was performed, with COG functions, Pfam and Prokka annotation (Table S2). The pan groups defined for the analysis were as follows:

(i) the phylogenetic Groups I, II, III and IV; (ii) the host origin; (iii) the geographic origin; (iv) the *Pseudomonas* species. We considered a Pan group enriched for a gene cluster if it was only present in all the isolates present in that group, showing the highest enrichment score and a p -value $< .05$. It is worth mentioning that the Group I of the *Pseudomonas avellanae* strains from Italy and Greece resulted to be enriched, among the others, in gene clusters associated with the HicA and HicB antitoxins, the effector protein HopAB2 and the protein VirC1. Interestingly, Group II with the *Pseudomonas avellanae* strains from the United Kingdom, New Zealand and Japan is enriched in the component involved in the anthranilate catabolism and thus associated with the WHOP region necessary for the woody-host infection (Caballo-Ponce et al., 2017). Note that the enrichment was not run for group III since it is represented only by one component (Shaiber et al. 2020). Group IV including the *Pseudomonas syringae* strains showed gene clusters belonging to several COG categories, for example, methyltransferase, DNA-binding regulator, protease and aldolase. When running the analysis based on the host range, the *Corylus avellana* group resulted enriched only in a potassium binding protein Kpb, while the *Prunus avium* group was enriched in the regulatory protein AsnC, pentapeptide repeat protein MfpA, L-ectoine synthase, a phosphodiesterase PdeB, methionine aminopeptidase 1, ADP-ribosylarginine hydrolase Tri1, Environmental sensor c-di-GMP phosphodiesterase, contains periplasmic CSS-motif sensor and cytoplasmic EAL domain (Table S2). For the same reason as the phylogenetic Group III, the group from *Actinidia chinensis* could not be

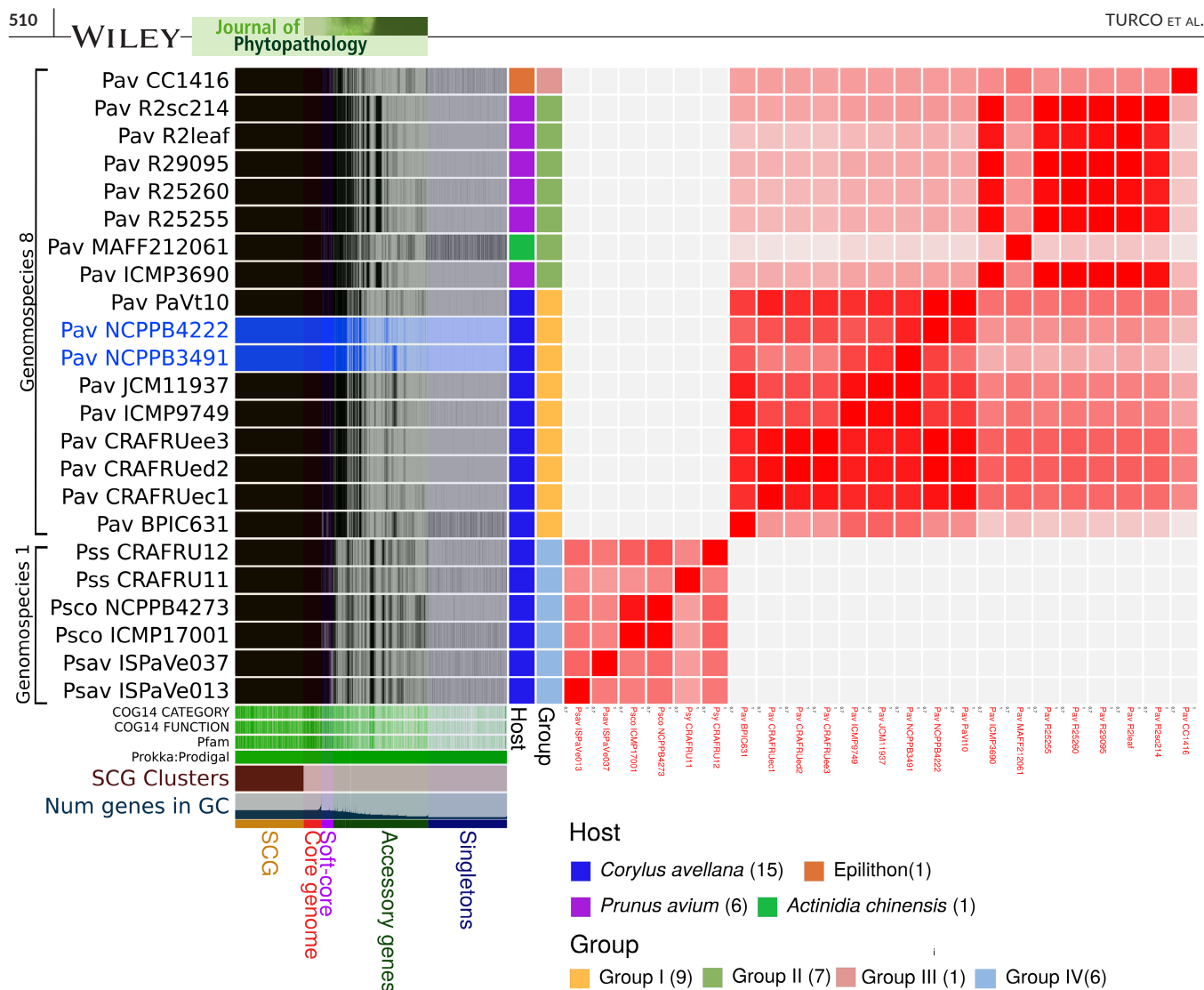


FIGURE 3 Pangenome analysis of the *Pseudomonas* strains performed with Anvi'o. The two strains sequenced and assembled in this work are represented in blue. On the left, the presence/absence of the gene clusters, with the Single Core Genome (SCG) in common among all the strains, the core and soft-core genome, the accessory gene clusters and the singletons. On the right, the heatmap shows the average nucleotide identity (ANI) among the strains. The genomespecies are indicated as well

evaluated. From a geographic point of view, the enrichment was only significant when more groups were combined, such as the United Kingdom, New Zealand and Japan, for example, and the results corroborated the ones from the analysis based on the phylogenetic groups. Instead, the *Pseudomonas avellanae* group was particularly enriched, among the others, in COG clusters related to coenzyme transport and metabolism (H), inorganic ion transport (P), secondary metabolite (Q). The COG enriched categories for *Pseudomonas syringae* were variegated as well, with deaminase from COG category E, to secreted trypsin-like serine protease from COG category O, to more general proteins related to COG category R.

3.5 | Type III secretion effectors

Pathogenic bacteria are known to release effector proteins into the host cell to suppress or escape the host immunity, through the different secretion systems. In particular, the type III secretion effectors

have been deeply studied and characterized by Dillon et al. (2019). We selected and blasted the ones that belong to one of the species/pathovar under study (i.e., *Pseudomonas avellanae*, *Pseudomonas syringae* pv. *avellanae*, *Pseudomonas syringae* pv. *coryli* or *Pseudomonas syringae* pv. *syringae*). The results were recorded as a binary matrix of presence/absence and shown in Figure 4, with a dendrogram clustering comparable to the previous phylogenetic trees. The analysis revealed that six effectors are in common among all the strains (AvrE1, HopAA1, HopAH1, HopAI1, HopB2 and HopM), while others were present in one or less than four isolates and for this reason not shown in Figure 4. Among those, AvrB2 was only found in MAFF 212061, R2leaf and R2sc214, HopBA1 was present only in CRAFRU11, CRAFRU12 and ISPAve013. CC1416 was unique for HopB1 and HopBO1, while CRAFRU11 was unique for HopAX1. HopBP1 was found only in Pav CC1416, MAFF 212061 and CRAFRU11, underlining again that Pss CRAFRU11 is slightly different than CRAFRU12 as reported in Figure S1 and Figure 2. Lastly, HopZ1 was only in R2leaf and R2sc214. Interestingly, HopF3, HopAB2 and HopZ4 are peculiar

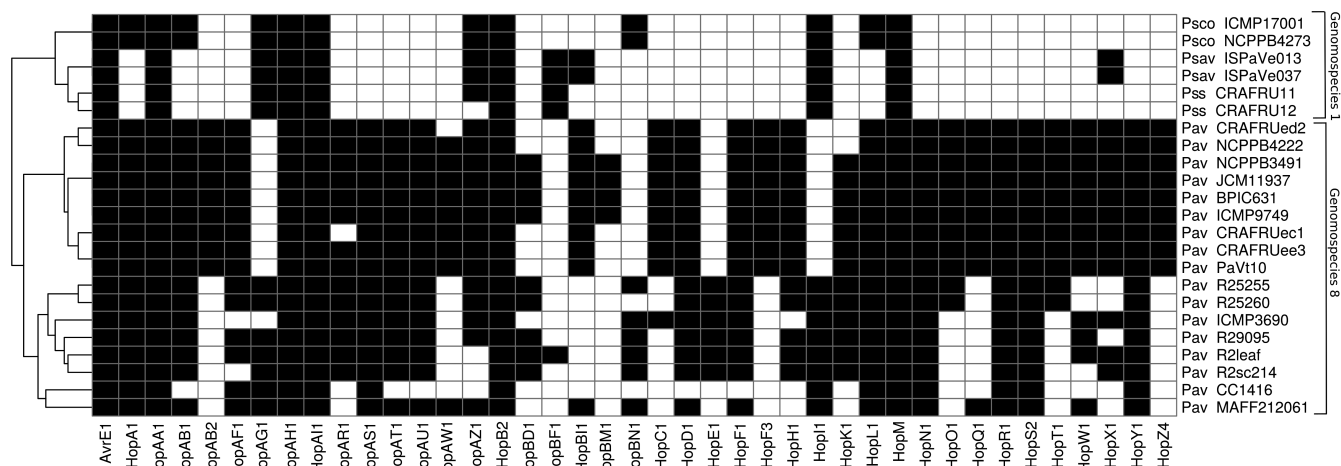


FIGURE 4 Distribution of type III secretion effector genes among the *Pseudomonas* strains used in this study. The black block indicates the presence of the effector retrieved by blast alignment; the white block indicates its absence. The genomospecies are indicated as well

to the *Pseudomonas avellanae* Group I, with the Italian and Greek strains. On the contrary, the same Group I lacks HopAG1 (present in all the others, exception made for Pav ICMP 3690), HopBF1 (present only in the pathovar *avellanae*, *syringae* and in Pav R2leaf), HopBN1 (randomly distributed between pathovar *coryli*, *Prunus* isolates and Pav MAFF 212061) and HopI1 (present in all the others isolates). Furthermore, HopE1 is peculiar to the *Pseudomonas avellanae* strains isolated from *Prunus avium*, while no specific effector was detected for the *Pseudomonas syringae* strains.

3.6 | T3SS secretion systems TPAI and RPAI

In *Pseudomonas*, the ability to infect the host plant and induce a hypersensitive response is guided by the type III secretion system (T3SS), generally having a tripartite mosaic structure composed, with the *hrp/hrc* cluster flanked by an Exchangeable Effector Locus (EEL) and a Conserved Effector Locus (CEL) (Alfano et al., 2000). More recently, five different versions of this T3SS have been identified among different pseudomonads. The canonical tripartite TPAI is present in *Pseudomonas syringae* strains of phylogroup 1, 2, 3, 4, 5, 6 and 10, together with an additional genomic island *Rhizobium* (RPAI) system present in some *P. syringae* of phylogroups 2 and 3. Phylogroups 7 and 11 are carrying the Single Pathogenicity Island (SPA), while the fluorescence (FPAI) system and corrugata (CPAI) island are present only in *Pseudomonas fluorescence* species complex and *Pseudomonas corrugata*, respectively (Dillon et al., 2021). The presence of the tripartite mosaic structure codifying for the type III secretion system was confirmed in all the 23 strains under study, together with an additional genomic island of 20Kb identified as the *Rhizobium* (RPAI) system (Dillon et al., 2021). This region was only of 10Kb in Pav MAFF 212061 and totally absent in Pav CC1416 and Pss CRAFRU11 (Figure 5). The phylogenetic tree based on the alignment of the RPAI region between the 21 remaining strains follows once more the clustering obtained in the previous phylogenetic trees (Figure S2).

3.7 | Other virulence-related genes

Besides the abovementioned effectors, *Pseudomonas* uses other specific tools to promote host infection and colonization. Among those, a genomic region defined as WHOP (Woody-host *Pseudomonas* spp.) is particularly present in *Pseudomonas* species infecting woody plants (Rodriguez-Palenzuela et al., 2010, Caballo-Ponce et al., 2017). Surprisingly, only the *P. avellanae* strains from Group II, isolated from *Prunus* or *Actinidia*, showed the complete WHOP region, while this was not present in the strains isolated from hazelnut, either *P. avellanae* or *P. syringae*. Regarding the gene cluster related to cellulase utilization, all the *wss* genes, except *wssC* and *wssE* were present in all the strains. For the degradation of woody tissues, only *hopAy1* gene was present in the *Pseudomonas avellanae* strains and in the *Pseudomonas syringae* pv. *coryli* (Figure 6). Among the different phytotoxins, the *mgoA*, *mgoB*, *mgoC* and *mgoD* genes of the mangotoxin operon Mgo occurred in all the strains, while only the *Pseudomonas syringae* strains showed the entire mangotoxin operon Mbo as well. All the strains contain the gene clusters to produce tabtoxin, while no phaseolotoxin, syringopeptin nor coronatine genes were detected (Figure 7). Only two *Pseudomonas syringae* pv. *syringae* contain the syringomycin gene cluster, as previously found by Marcelletti and Scortichini (2015). No *Pseudomonas syringae* strains contain *avrD1*, while this feature was detected in all the strains belonging to Group I, as well as is Pav MAFF 212061 (from *Actinidia*) and Pav R25255 (from *Prunus*). All the *P. avellanae* strains contain the indoleacetic acid-lysine gene (*iaaL*) but not the indoleacetic acid phytohormone genes (*iaaM* and *iaaH*). Among the ice nucleation factors, only the *inaA* gene was present in all the strains, while only the *Pseudomonas syringae* pathovars contained the entire ice nucleation cluster. No ethylene-forming enzyme (*efe* gene), isopen-tenyl transferase (*ptz*) and isopen-tenyl-diphosphate delta-isomerase (*idi* gene), both essential for tumorigenesis (Anorga et al., 2020) were detected. All the strains are carrying the HLE for the homoserine lactone. Interestingly, the complete copper resistance gene cluster (*copABCD*) and related regulatory elements (*copSR*) are present only

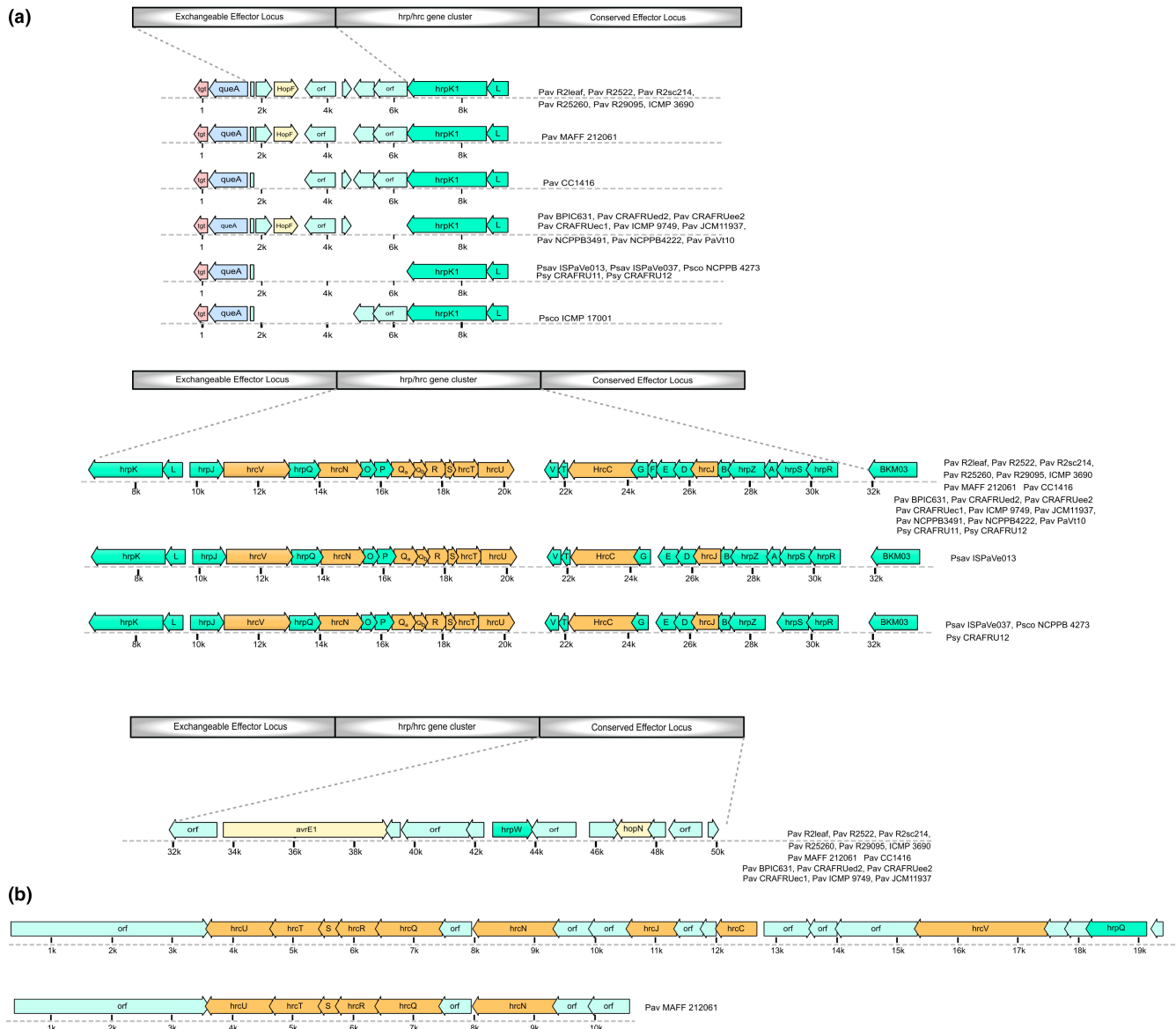


FIGURE 5 Schematic representation of the type III secretion system. (a) The tripartite mosaic structure composed of the variable EEL locus, the *hrp/hrc* gene cluster and the conserved locus. (b) Additional genomic island carrying the type III secretion system as well. Different colours represent different gene categories

in Pss CRAFRU12. Instead, the levansucrase sucrose gene was present in all the strains (Figure 7).

4 | DISCUSSION

Hazelnut is an economically important crop for the Italian market, and during the last decades, it has been threatened by several diseases, either of bacterial or fungal origin (Marcelletti & Scortichini, 2019; Turco et al., 2021). Among those, the bacterial canker of hazelnut has been associated with *Pseudomonas avellanae* (belonging to genomospecies 8), which is able to migrate systemically within the plant and quickly infect the plants within the same hazelnut orchard (Scortichini and Lazzari, 1996; Scortichini et al., 2000). Together with *Pseudomonas syringae* pv. *avellanae*, *Pseudomonas avellanae* results to

be the most aggressive for their aptitude to induce twig wilting, die-back, trunk canker and plant death (Marcelletti & Scortichini, 2019). Two additional *Pseudomonas syringae* pathovars, *coryli* and *syringae*, have been isolated from European hazelnut as well (Scortichini et al., 2002; Loreti et al., 2009).

In the present study, we performed a comparative genomic analysis among *Pseudomonas* strains that converged on the same host, *Corylus avellana*, to assess how they are different from each other and from other *Pseudomonas avellanae* strains infecting different hosts. In literature, there are several examples of *Pseudomonas* belonging to different genomospecies that converged into the same host and that have been analysed through comparative genomics. Already in 2015, Marcelletti and Scortichini performed a comparative genomic analysis of *Pseudomonas* strains that converged into *Corylus avellana*, in particular *Pseudomonas avellanae* strains BPIC

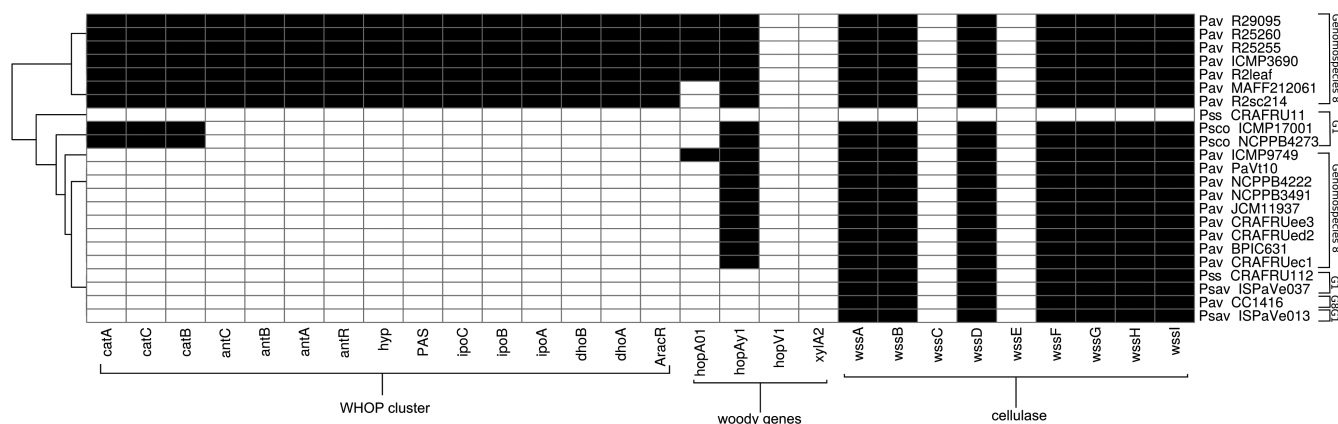


FIGURE 6 Distribution of the WHOP gene cluster, woody genes and cellulase gene cluster among the *Pseudomonas* strains. The black block indicates the presence of the gene retrieved by blast alignment; the white block indicates its absence. Species belonging to genomospecies 8 are indicated in black, and species belonging to genomospecies 1 are indicated in red

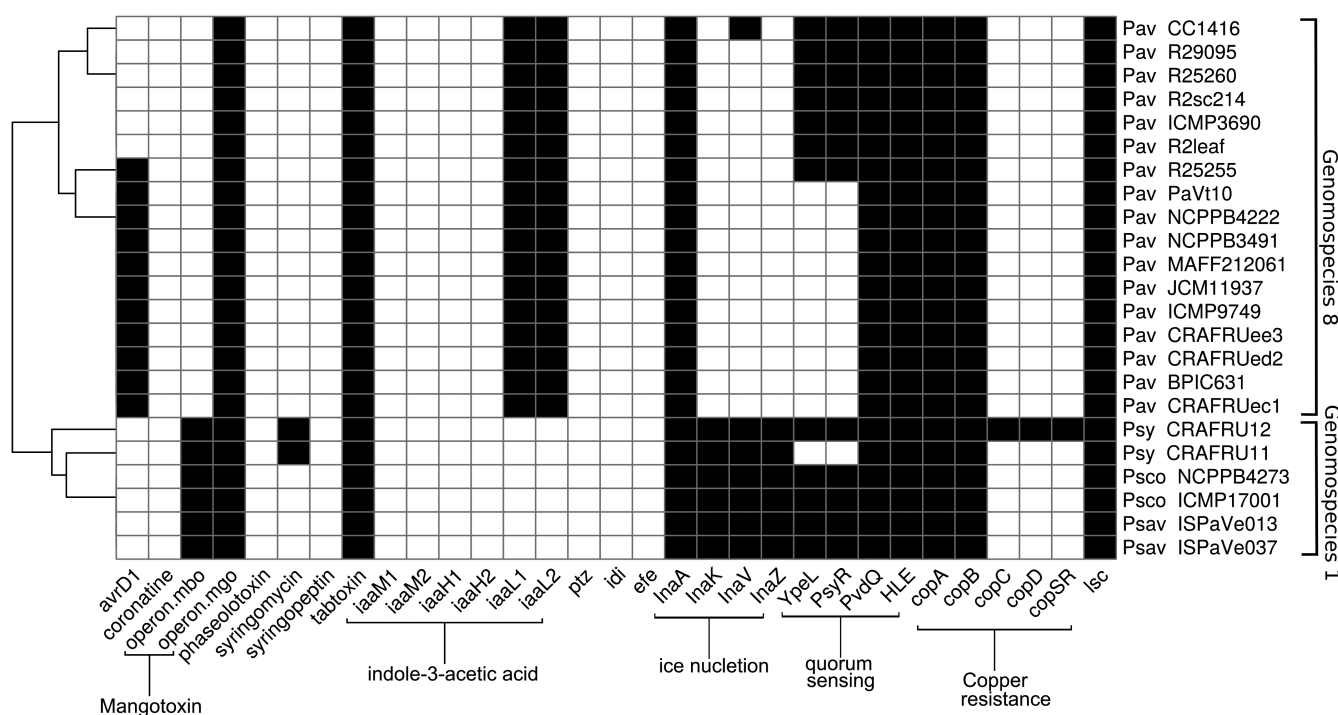


FIGURE 7 Distribution of the virulence-related features among the *Pseudomonas* strains. The black block indicates the presence of the gene retrieved by blast alignment; the white block indicates its absence. The genomospecies are indicated as well

631, CRAFRUec1, CRAFRUed2, CRAFRUee3 (from genomospecies 8), *Pseudomonas syringae* pv. *avellanae* CRAPAV013 and CRAPAV037 (same as ISPAve013 and ISPAve037), *Pseudomonas syringae* pv. *coryli* NCPPB 4273, *Pseudomonas syringae* pv. *syringae* strains CRAFRU11 and CRAFRU12 (all from genomospecies 1). Their findings showed that all the strains contain the same genomic island of 20kb carrying the *hrp/hrc* type III secretion system gene cluster, with an overall higher number of effectors in *Pseudomonas avellanae*, while the *P. syringae* strains carried the complete antimetabolite mangotoxin and ice nucleation activity gene clusters (Marcelletti & Scortichini, 2015). Another study, by Hulin et al. (2018), demonstrated how *Pseudomonas syringae* pv. *morsprunorum* races 1 and

2, *Pseudomonas syringae* pv. *syringae* and *Pseudomonas syringae* pv. *avii* converged into *Prunus avium* and that this host-range pathogenicity was defined by six virulence features: hopBF, hopAB, hopH, hopAR, avrPto and hopBB. In particular, the acquisition of hopAR1 and hopBB1 effectors through phage and the horizontal transfer was depicted as fundamental for cherry pathogenicity (Hulin et al., 2018). In their analysis, *Pseudomonas syringae* pv. *syringae* was the species with less effectors but a higher number of phytotoxins. Later, the taxonomic classification of the strains previously defined as *Pseudomonas syringae* pv. *morsprunorum* race 2 has been changed to *Pseudomonas avellanae* species and thus belonging to genomospecies 8 as well (Hulin et al., 2020). In 2019, Marcelletti and Scortichini

proved once more that *Prunus* can be infected by different strains of *Pseudomonas* using different mechanisms of infection and, since no clear effector pattern distribution could be retrieved, the virulence features could be discerned only at the nucleotide sequence level due to point mutations (Marcelletti & Scortichini, 2019).

Hence, *Pseudomonas avellanae* resulted to have the ability to infect different hosts. Interestingly, pathogenicity tests performed with two strains of *Pseudomonas avellanae* (Pav BPIC 631 and Pav PaVt10) isolated from hazelnut, did not induce any symptoms when tested on cherry fruits (*Prunus avium*) and peach leaves (*Prunus persica*), whilst *Pseudomonas syringae* pv. *avellanae* induced necrotic lesions (Ruinelli et al., 2019). Thus, based on the evidence that at least four different *Pseudomonas* species can infect *Corylus avellana* and since one of them, *Pseudomonas avellanae*, has been isolated from other plant species (*Prunus* and *Actinidia*), as well as from particular natural environments (epilithon), we wanted to further investigate the genetic features related to host-range specificity among these strains.

Hence, in the current work, the complete genome sequence of two additional *Pseudomonas avellanae* strains, NCPPB 4222 strain isolated from hazelnut in Italy and NCPPB 3491 isolated from hazelnut in Greece, was obtained and included in the analysis. By using a combination of Illumina short reads and Oxford Nanopore Technologies (ONT) long reads in a hybrid assembly, it was possible to retrieve the complete bacterial chromosome with the additional plasmids (two for NCPPB 4222 and three for NCPPB 3491, respectively). Indeed, the hybrid assembly combines the advantages of both sequencing technologies, with the ONT long reads able to cover difficult regions like the ones full of tandem repeats or transposase (Goodwin et al. 2016; van Dijk et al., 2018; Turco et al., 2022), while the depth coverage of the Illumina reads allow a solid base calling with higher Phred score (Chen et al., 2020). Having high-quality complete genomes when performing comparative genomics is fundamental for a solid identification of all the features under comparison. In fact, in draft genomes with high contigs number, the total annotated features are underestimated due to contigs interruption and thus, fragmented genes (Smits, 2019).

Besides these two newly sequenced complete genomes, all the other fifteen *Pseudomonas avellanae* strains available on the NCBI database were downloaded and included in the analysis, together with two *Pseudomonas syringae* pv. *avellanae*, two *Pseudomonas syringae* pv. *coryli* and two *Pseudomonas syringae* pv. *syringae*. To reduce the bias due to the genome quality and completeness, all the genomes were reordered and annotated using Mauve and Prokka.

The phylogenetic analysis based either on the seven housekeeping genes or the core genome, as well as the Average Nucleotide Identity (ANI) showed a clear distinction between the *Pseudomonas avellanae* clade and the *Pseudomonas syringae*, in line with previous works (Hulin et al., 2018; Marcelletti & Scortichini, 2019). Furthermore, within the *Pseudomonas avellanae* clade, it is possible to identify three subgroups referable to the geographic origin and/or the source of isolation. The Italian and Greek strains isolated from hazelnut clustered together, with the British strains from *Prunus* and the Japanese strain from kiwifruit in a separated clade. Not surprisingly, the strain isolated from the river epilithon was clearly

divergent from all the others. This first analysis showed that already at the nucleotide level, the *Pseudomonas avellanae* appeared to be different. We then moved to the pangenome analysis and virulence features comparison to appreciate possible differences at the functional level. As expected, all the *Pseudomonas* shared a relevant core genome with conserved features, but also accessory genes and singletons were identified delineating strains diversity. The most divergent strains, MAFF 212061 and CC1416, consistently showed the highest number of singletons.

The enrichment analysis performed within Anvi'o allowed us to identify the features distinctive for a particular strain selection. Interestingly, the Group I of *Pseudomonas avellanae* strains from Italy and Greece resulted to be enriched in a feature similar to VirC1 of *Agrobacterium tumefaciens*, known to be involved in tumour formation (Toro et al., 1989). They are also enriched in the toxin/antitoxin HicAB system, quite spread among Prokaryotes and involved in programmed cell death, stress response, biofilm formation and virulence (Sberro et al., 2013). The type III secretion system effector E3 ubiquitin-protein ligase HopAB2 may also provide host-species-specificity since its complete sequence is not present in any other Groups. On the contrary, the Group II of *Pseudomonas avellanae* from *Prunus* and *Actinidia* contains the complete WHOP region, known to be necessary for woody-host infection (Caballo-Ponce et al., 2017). As found by Hulin et al. (2018), six effectors (AvrE1, HopAA1, HopAH1, HopAI1, HopB2 and HopM) are present in all the strains, while the *syringae* group appears to have less effectors than *avellanae* group. It is noteworthy that HopF3, HopZ4 and a full copy of HopAB2 are characteristic of *Pseudomonas avellanae* Group I and may have contributed to the hazelnut-specificity of this species. On the contrary, HopE1 may be specific for *Prunus* infection. Besides the canonical tripartite pathogenicity island (TPAI), all the strains except for Pav CC1416 and Psav CRAFRU11 contain an additional genomic island *Rhizobium* (RPAI) system, as previously described by Dillon et al. (2021) for other *Pseudomonas*. This island of about 20Kb is most likely the same identified by Marcelletti and Scortichini (2015) in the nine pseudomonads infecting European hazelnut trees. As proved here, this genomic region may not be involved in host-specificity since it is present in all the strains. On the contrary, phytotoxins may play a pivotal role in the evolution of host specificity, since they particularly counteract the immunity response (Dillon et al., 2021). The phytotoxins are more specific for the *syringae* group, which all contain the complete Mbo and Mgo operons for the mangotoxin production, necessary for mango necrosis, as well as a complete ice nucleation cluster (Marcelletti & Scortichini, 2015). Tablatoxin, instead was present in all the strains.

Pathogenicity is a complex mechanism where many actors are involved, and their synergistic effects may be more relevant than the simple presence or absence of a particular effector. For example, the phytotoxins present in the *syringae* group do not necessarily imply heavier symptomatology, while the phytotoxin deficient *Pseudomonas avellanae* have always been reported as virulent.

Hazelnut, despite its acknowledged resilience to adverse environmental conditions, must face the aggression of a plethora of biotic

entities. Either insects, fungi and bacteria can lead to heavy production losses or even to the death of plants. In such a sense, the concurrence of different pathogenic pseudomonads on the same host plant is of particular significance. According to the variability of cultivation conditions, in which the climate change plays a key role, one or more of them are prone to overcome plant defences inciting new and unexpected outbreaks, as already experienced in the recent past.

Therefore, it is of vital importance to deepen as much as possible the knowledge of the genetic traits of these pathogens discovering their specific virulence factors, infection mechanisms or host-specificity features, and finding any potential weaknesses to be used against them for both prevention measures and new control strategies.

AUTHOR CONTRIBUTIONS

Silvia Turco conceptualize the project, performed the sequencing, the data analysis and wrote the manuscript. Leonardo Zuppante contributed to the data analysis. Mounira Inas Draï contributed to the sequencing and experimentation. Angelo Mazzaglia contributed to project conceptualization, supervised the project and contributed to the preparation of the draft manuscript and its review and editing.

ACKNOWLEDGEMENTS

The research was carried out in the frame of the Italian MIUR (Ministry of Education, University and Research) initiative 'Department of Excellence' (Law 232/2016). All the bioinformatics calculations and analyses were performed at the DAFNE HPC scientific computing center of the Università degli Studi della Toscana. Open Access Funding provided by Università degli Studi della Toscana within the CRUI-CARE Agreement.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

PEER REVIEW

The peer review history for this article is available at <https://publons.com/publon/10.1111/jph.13101>.

ORCID

Silvia Turco  <https://orcid.org/0000-0001-9041-6461>

Mounira Inas Draï  <https://orcid.org/0000-0002-9934-5898>

REFERENCES

- Alfano, J. R., Charkowski, A. O., Deng, W.-L., Badel, J. L., Petnicki-Ocwieja, T., van Dijk, K., & Collmer, A. (2000). The *Pseudomonas syringae* Hrp pathogenicity Island has a tripartite mosaic structure composed of a cluster of type III secretion genes bounded by exchangeable effector and conserved effector loci that contribute to parasitic fitness and pathogenicity in pl. *Proceedings of the National Academy of Sciences*, 97, 4856–4861.
- Andrews, S., (2010). FastQC: A quality control tool for high throughput sequence data.
- Añorga, M., Pintado, A., Ramos, C., De Diego, N., Ugena, L., Novák, O., & Murillo, J. (2020). Genes ptz and idi, coding for cytokinin biosynthesis enzymes, are essential for tumorigenesis and in planta growth by *P. syringae* pv. *Savastanoi* NCPPB 3335. *Frontiers in Plant Science*, 11, 1294.
- Caballo-Ponce, E., Van Dillewijn, P., Wittich, R. M., & Ramos, C. (2017). WHOP, a genomic region associated with woody hosts in the *Pseudomonas syringae* complex contributes to the virulence and fitness of *Pseudomonas savastanoi* pv. *Savastanoi* in olive plants. *Molecular Plant-Microbe Interactions*, 30, 113–126.
- Chen, Z., Erickson, D. L., & Meng, J. (2020). Benchmarking hybrid assembly approaches for genomic analyses of bacterial pathogens using Illumina and Oxford nanopore sequencing. *BMC Genomics*, 21, 631.
- Darling, A. C. E., Mau, B., Blattner, F. R., & Perna, N. T. (2004). Mauve: Multiple alignment of conserved genomic sequence with rearrangements. *Genome Research*, 14, 1394–1403.
- 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. *Frontiers in Plant Science*, 10, 418.
- Dillon, M. M., Ruiz-Bedoya, T., Bundalovic-Torma, C., Guttman, K. M., Kwak, H., Middleton, M. A., Wang, P. W., Horuz, S., Aysan, Y., & Guttman, D. S. (2021). Comparative genomic insights into the epidemiology and virulence of plant pathogenic pseudomonads from Turkey. *Microbial Genomics*, 7, 585.
- Edgar, R. C. (2004). MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, 32, 1792–1797.
- El-Gebali, S., Mistry, J., Bateman, A., Eddy, S. R., Luciani, A., Potter, S. C., Qureshi, M., Richardson, L. J., Salazar, G. A., Smart, A., Sonhammer, E. L. L., Hirsh, L., Paladin, L., Piovesan, D., Tosatto, S. C. E., & Finn, R. D. (2019). The Pfam protein families database in 2019. *Nucleic Acids Research*, 47, D427–D432.
- Eren, A. M., Esen, O. C., Quince, C., Vineis, J. H., Morrison, H. G., Sogin, M. L., & Delmont, T. O. (2015). Anvi'o: An advanced analysis and visualization platform for 'omics data'. *PeerJ*, 2015, 1–29.
- Goodwin, S., McPherson, J. D., & McCombie, W. R. (2016). Coming of age: Ten years of next-generation sequencing technologies. *Nature Reviews Genetics*, 17, 333–351.
- Gurevich, A., Saveliev, V., Vyahhi, N., & Tesler, G. (2013). QUAST: Quality assessment tool for genome assemblies. *Bioinformatics*, 29, 1072–1075.
- Hulin, M. T., Armitage, A. D., Vicente, J. G., Holub, E. B., Baxter, L., Bates, H. J., Mansfield, J. W., Jackson, R. W., & Harrison, R. J. (2018). Comparative genomics of *Pseudomonas syringae* reveals convergent gene gain and loss associated with specialization onto cherry (*Prunus avium*). *The New Phytologist*, 219, 672–696.
- Hulin, M. T., Jackson, R. W., Harrison, R. J., & Mansfield, J. W. (2020). Cherry picking by pseudomonads: After a century of research on canker, genomics provides insights into the evolution of pathogenicity towards stone fruits. *Plant Pathology*, 69, 962–978.
- ISTAT. (2021). Istituto Nazionale di Statistica: Coltivazioni legnose fruttifere. <http://dati.istat.it/Index.aspx?QueryId=33705>. Accessed April 21, 2021.
- Janse, J. D., Rossi, P., Angelucci, L., Scortichini, M., Derks, J. H. J., Akkermans, A. D. L., De Vrijer, R., & Psallidas, P. G. (1996). Reclassification of *Pseudomonas syringae* pv. *Avellanae* as *Pseudomonas avellanae* (spec. Nov.), the bacterium causing canker of hazelnut (*Corylus avellana* L.). *Systematic and Applied Microbiology*, 19, 589–595.
- King, E. O., Ward, M. K., & Raney, D. E. (1954). Two simple media for the demonstration of pyocyanin and fluorescein. *The Journal of Laboratory and Clinical Medicine*, 44, 301–307.
- Kolde, R., (2019). Pheatmap: Pretty heatmaps. R Packag. <https://cran.r-project.org/web/packages/pheatmap/pheatmap.pdf>
- Loreti, S., Gallelli, A., de Simone, D., & Bosco, A. (2009). Detection of *Pseudomonas avellanae* and the bacterial microflora of hazelnut

- affected by "Moria" in Central Italy. *Journal of Plant Pathology*, 91, 365–373.
- Marcelletti, S., & Scortichini, M. (2019). Some strains that have converged to infect *prunus* spp. trees are members of distinct *Pseudomonas syringae* genomospecies and ecotypes as revealed by in silico genomic comparison. *Archives of Microbiology*, 201, 67–80.
- Marcelletti, S., & Scortichini, M. (2015). Comparative genomic analyses of multiple *Pseudomonas* strains infecting *Corylus avellana* trees reveal the occurrence of two genetic clusters with both common and distinctive virulence and fitness traits. *PLoS One*, 10(7), e0131112.
- Mayjonade, B., Gouzy, J., Donnadiou, C., Pouilly, N., Marande, W., Callot, C., Langlade, N., & Muñoz, S. (2016). Extraction of high-molecular-weight genomic DNA for long-read sequencing of single molecules. *BioTechniques*, 61, 203–205.
- Morris, C. E., Kinkel, L. L., Kun, X., Prior, P., & Sands, D. C. (2007). A surprising niche for the plant pathogen *Pseudomonas syringae*. *Infection, Genetics and Evolution*, 7, 84–92.
- Page, A. J., Cummins, C. A., Hunt, M., Wong, V. K., Reuter, S., Holden, M. T. G., Fookes, M., Falush, D., Keane, J. A., & Parkhill, J. (2015). Roary: Rapid large-scale prokaryote pan genome analysis. *Bioinformatics*, 31, 3691–3693.
- Petkau, A., Stuart-Edwards, M., Stothard, P., & Van Domselaar, G. (2010). Genome analysis interactive microbial genome visualization with GView. *Bioinformatics*, 26, 3125–3126.
- Psallidas, P. G. (1993). *Pseudomonas syringae* pv. *Avellanae* pathovar nov., the bacterium causing canker disease on *Corylus avellana*. *Plant Pathology*, 42, 358–363.
- Psallidas, P. G. (1984). Bacterial canker of *Corylus avellana*: Taxonomic position of the causal agent. In *Proceedings of the Second Working Group on Pseudomonas syringae Pathovars (International Society for Plant Pathology, Greece, Sounion, 24-28 April 1984)* (pp. 53–55).
- Psallidas, P. G., & Panagopoulos, C. G. (1979). A bacterial canker of *Corylus avellana* in Greece. *Journal of Phytopathology*, 94, 103–111.
- Rodríguez-Palenzuela, P., Matas, I. M., Murillo, J., López-Solanilla, E., Bardaji, L., Pérez-Martínez, I., Rodríguez-Mosquera, M. E., Penyalver, R., López, M. M., Quesada, J. M., Biehl, B. S., Perna, N. T., Glasner, J. D., Cabot, E. L., Neeno-Eckwall, E., & Ramos, C. (2010). Annotation and overview of the *Pseudomonas savastanoi* pv. *Savastanoi* NCPPB 3335 draft genome reveals the virulence gene complement of a tumour-inducing pathogen of woody hosts. *Environmental Microbiology*, 12, 1604–1620.
- 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.
- Sberro, H., Leavitt, A., Kiro, R., Koh, E., Peleg, Y., Qimron, U., & Sorek, R. (2013). Discovery of functional toxin/antitoxin Systems in Bacteria by shotgun cloning. *Molecular Cell*, 50, 136–148.
- Scortichini, M. (1994). Occurrence of *Pseudomonas syringae* pv. *Actinidia* on kiwifruit in Italy. *Plant Pathology*, 43, 1035–1038.
- Scortichini, M., Dettori, M. T., Marchesi, U., Palombi, M. A., & Rossi, M. P. (1998). Differentiation of *Pseudomonas avellanae* strains from Greece and Italy by rep-PCR genomic fingerprinting. *Journal of Phytopathology*, 146, 417–420.
- Scortichini, M., & Lazzari, M. (1996). Systemic migration of *Pseudomonas syringae* pv. *Avellanae* in twigs and Young trees of hazelnut and symptom development. *Journal of Phytopathology*, 144, 215–219.
- Scortichini, M., Marcelletti, S., Ferrante, P., & Firrao, G. (2013). A genomic redefinition of *Pseudomonas avellanae* species. *PLoS One*, 8, 75794.
- Scortichini, M., Marchesi, U., Dettori, M. T., Angelucci, L., Rossi, M. P., & Morone, C. (2000). Genetic and pathogenic diversity of *Pseudomonas avellanae* strains isolated from *Corylus avellana* trees in north-west of Italy, and comparison with strains from other regions. *European Journal of Plant Pathology*, 106, 147–154.
- Scortichini, M., Marchesi, U., Rossi, M. P., & Di Prospero, P. (2002). Bacteria associated with hazelnut (*Corylus avellana* L.) decline are of two groups: *Pseudomonas avellanae* and strains resembling *P. syringae* pv. *Syringae*. *Applied and Environmental Microbiology*, 68, 476–484.
- Scortichini, M., & Morone, C. (1997). Rinvenimento di *Pseudomonas avellanae* in noccioli delle Langhe. *Informatore Fitopatologico*, 47, 25–26.
- Scortichini, M., Rossi, M. P., Loreti, S., Bosco, A., Fiori, M., Jackson, R. W., Stead, D. E., Aspin, A., Marchesi, U., Zini, M., & Janse, J. D. (2005). *Pseudomonas syringae* pv. *Coryli*, the causal agent of bacterial twig dieback of *Corylus avellana*. *Phytopathology*, 95, 1316–1324.
- Scortichini, M., & Tropiano, F. G. (1994). Severe outbreak of *Pseudomonas syringae* pv. *Avellanae* on hazelnut in Italy. *Journal of Phytopathology*, 140, 65–70.
- Seemann, T. (2014). Prokka: Rapid prokaryotic genome annotation. *Bioinformatics*, 30, 2068–2069.
- Shaiber, A., Willis, A. D., Delmont, T. O., Roux, S., Chen, L.-X., Schmid, A. C., Yousef, M., Watson, A. R., Lolans, K., Esen, Ö. C., Lee, S. T. M., Downey, N., Morrison, H. G., Dewhirst, F. E., Mark Welch, J. L., & Eren, A. M. (2020). Functional and genetic markers of niche partitioning among enigmatic members of the human oral microbiome. *Genome Biology*, 21(21), 1–35.
- Smits, T. H. M. (2019). The importance of genome sequence quality to microbial comparative genomics. *BMC Genomics*, 20, 1–4.
- Stamatakis, A. (2014). RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30, 1312–1313.
- Tatusov, R. L., Galperin, M. Y., Natale, D. A., & Koonin, E. V. (2000). The COG database: A tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Research*, 28, 33–36.
- Toro, N., Datta, A., Carmi, O. A., Young, C., Prusti, R. K., & Nester, E. W. (1989). The *agrobacterium tumefaciens* virC1 gene product binds to overdrive, a T-DNA transfer enhancer. *Journal of Bacteriology*, 171, 6845–6849.
- Turco, S., Grottoli, A., Drais, M. I., De Spirito, C., Faino, L., Reverberi, M., Cristofori, V., & Mazzaglia, A. (2021). Draft genome sequence of a new *fusarium* isolate belonging to *fusarium tricinctum* species complex collected from hazelnut in Central Italy. *Frontiers in Plant Science*, 12, 1–11.
- Turco, S., Drais, M. I., Rossini, L., Chaboteaux, E., Rahi, Y. J., Balestra, G. M., Iacobellis, N. S., & Mazzaglia, A. (2022). Complete genome assembly of the levan-positive strain PVFi1 of *Pseudomonas savastanoi* pv. *savastanoi* isolated from olive knots in Central Italy. *Environmental Microbiology Reports*, 14(2), 274–285.
- van Dijk, E. L., Jaszczyszyn, Y., Naquin, D., & Thermes, C. (2018). The third revolution in sequencing technology. *Trends in Genetics*, 34, 666–681.
- 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 Computational Biology*, 13, e1005595.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Turco, S., Zuppante, L., Drais, M. I. & Mazzaglia, A. (2022). Dressing like a pathogen: Comparative analysis of different *Pseudomonas* genomospecies wearing different features to infect *Corylus avellana*. *Journal of Phytopathology*, 170, 504–516. <https://doi.org/10.1111/jph.13101>