The Pseudomonas viridiflava phylogroups in the P. syringae species complex are characterized by genetic variability and phenotypic plasticity of pathogenicity-related traits.

Claudia Bartoli, Odile Berge, Caroline Monteil, Caroline Guilbaud, Giorgio M. Balestra, Leonardo Varvaro, Corbin D. Jones, Jeffery L. Dangl, David A Baltrus, David C. Sands, et al.

The Pseudomonas viridiflava phylogroups in the P. syringae species complex are characterized by genetic variability and phenotypic plasticity of pathogenicity-related traits

Journal:	*Environmental Microbiology and Environmental Microbiology Reports*
Manuscript ID:	Draft
Manuscript Type:	EMI - Research article
Journal:	Environmental Microbiology
Date Submitted by the Author:	n/a
Complete List of Authors:	Bartoli, Claudia; Tuscia University, Departement of Science and Technology for Agriculture, Forestry, Nature and Energy Berge, Odile; INRA, Plant Health and Environment Monteil, Caroline; INRA, Plant Health and Environment Guilbaud, Caroline; INRA, Plant Health and Environment Balestra, Giorgio; Tuscia University, Departement of Science and Technology for Agriculture, Forestry, Nature and Energy Varvaro, Leonardo; Tuscia University, Departement of Science and Technology for Agriculture, Forestry, Nature and Energy Jones, Corbin; University of North Carolina at Chapel Hill, Department of Biology and Carolina Center for Genome Sciences Dangl, Jeff; University of North Carolina at Chapel Hill, Department of Biology and Howard Hughes Medical Institute Baltrus, David; University of Arizona, School of Plant Sciences Sands, David; Montana State University, Dept. of Plant Sciences and Plant Pathology Morris, Cindy; INRA, Plant Health and Environment
Keywords:	microbe:higher organism interactions, microbial ecology, pathogen ecology

The *Pseudomonas viridiflava* phylogroups in the *P. syringae* species complex are characterized by genetic variability and phenotypic plasticity of pathogenicity-related traits

Claudia Bartoli[1,2], Odile Berge[2], Caroline L. Monteil[2], Caroline Guilbaud[2], Giorgio M. Balestra[1], Leonardo Varvaro[1] Corbin Jones[3], Jeffery L Dangl[4], David A. Baltrus[5], David C. Sands[6] and Cindy E. Morris[2,6*]

[1]Departement of Science and Technology for Agriculture, Forestry, Nature and Energy (DAFNE), Tuscia University, Viterbo, Italy

[2]INRA, UR0407 Pathologie Végétale, F-84143 Montfavet cedex, France

[3]Department of Biology and Carolina Center for Genome Sciences, Univ. of North Carolina at Chapel Hill, Chapel Hill, NC 29599, USA

[4]Department of Biology and Howard Hughes Medical Institute, Univ. of North Carolina at Chapel Hill, Chapel Hill, NC 29599, USA

[5]School of Plant Sciences, University of Arizona, P.O. Box 210036, Forbes 303, Tucson, Arizona, 85721, USA

[6]Dept. Plant Sciences and Plant Pathology, Montana State University, Bozeman, MT 59717-3150, USA

*For correspondence. E-mail cindy.morris@avignon.inra.fr; (+33) (0) 432722841;Fax (+33) (0) 432722841

Running title: *Pseudomonas viridiflava* diversity

Summary

As a species complex, *Pseudomonas syringae* exists in both agriculture and natural aquatic habitats. *P. viridiflava*, a member of this complex, has been reported to be phenotypically largely homogenous. We characterized strains from different habitats, selected based on their genetic similarity to previously described *P. viridiflava* strains. We revealed two distinct phylogroups and two different kinds of variability in phenotypic traits and genomic content. The strains exhibited phase variation in phenotypes including pathogenicity and soft rot on potato. We showed that the presence of two configurations of the Type III Secretion System [single (S-PAI) and tripartite (T-PAI) pathogenicity islands] are not correlated with pathogenicity or with the capacity to induce soft rot in contrast to previous reports. The presence/absence of the *avrE* effector gene was the only trait we found to be correlated with pathogenicity of *P. viridiflava*. Other Type III secretion effector genes were not correlated with pathogenicity. A genomic region resembling an exchangeable effector locus (*EEL*) was found in S-PAI strains and a probable recombination between the two PAIs is described. The ensemble of the variability observed in these phylogroups of *P. syringae* likely contributes to their adaptability to alternating opportunities for pathogenicity or saprophytic survival.

Introduction

The plant pathogen *Pseudomonas syringae* is a species complex displaying wide genetic variability and capacity for adaptation to a broad range of habitats, thereby posing a challenge for defining the scope of its diversity. Often referred to as an archetypical plant pathogen and epiphyte (Hirano and Upper, 2000), it is becoming increasingly clear that strains of this species are capable of surviving and diversifying in habitats outside of agriculture. The broad range of ecological niches of *P. syringae* is reflected in the genomic and phenotypic diversity across the whole spectrum of this species complex. Phylogroups within this complex differ dramatically at the genome level (Baltrus *et al.,* 2011; O'Brien *et al.,* 2011, Baltrus *et al*., 2013). This is reflected, at least across pathogenic strains, by variable accumulation of genes encoding Type III secretion systems (TTSS), the Type III effectors (TTEs) that encode substrates for TTSS, and associated phytotoxins that complement and extend TTE virulence functions (Baltrus *et al.,* 2011 Clarke *et al.,* 2010; Araki *et al.,* 2006; Demba Diallo *et al.,* 2012).

In contrast to the well-established heterogeneity within phylogroups of *P. syringae*, *P. viridiflava* has been reported to be relatively homogeneous (Sarris *et al*., 2012). Although designated with a species name, *P. viridiflava* represents one of the multiple phylogroups found within the *P. syringae* species complex (Gardan *et al.,* 1999; Mulet *et al.,* 2010; Parkinson *et al.,* 2011). As described by Billing (Billing, 1970), *P. viridiflava* has pectolytic activity (Liao *et al.,* 1988) and the capacity to induce soft rot of potato slices in laboratory tests and on a range of vegetables during storage (Morris *et al.,* 1991). In the field, *P. viridiflava* has been reported as a pathogen on tomato, on blite goosefoot (*Blitum capitatum*) and eggplant (Goumans and Chatzaki 1998), kiwifruit (Conn and Gubler, 1993), common bean and lettuce (González *et al.,* 2003), basil (Végh *et al.,* 2012), various wild herbaceous species (Goss *et al.,* 2005) and *Arabidopsis thaliana* (Jackson *et al.,* 1999; Goss *et al.,* 2005). Although two different *P. viridiflava* genotypes were detected in populations isolated

from *A. thaliana* (Goss *et al.,* 2005) only differences in virulence were reported (Araki *et al.,* 2006; Jakob *et al.,* 2007). The use of a limited number of specific traits as criteria for identification of strains as *P. viridiflava* that characterized the previously studies, has limited the diversity of the collections that have been studied to date.

Previous studies focused on *P. viridiflava* isolated principally from plants, and little is known about phenotypic diversity outside of agricultural contexts. Questions about the importance of non-agricultural contexts and habitats other than plants in the ecology and evolution of plant pathogens are pertinent in light of evidence that strains from the *P. syringae* complex regularly occur in a range of habitats outside of diseased crop plants (Mohr *et al.,* 2008; Morris *et al.,* 2010). Furthermore, when found in association with plant tissues, *P. viridiflava* is often present in contexts that are favorable for colonization by saprophytes (Balestra and Varvaro, 1997, 1998) such as *P. fluorescens* (Morris *et al.,* 1991; Everett and Henshall, 1994). These observations suggest that habitats fostering saprophytic growth of *P. viridiflava* might be favorable habitats for diversification. We therefore characterized the genetic and phenotypic variability of a set of environmental strains phylogenetically related to phylogroup 7 in which known *P. viridiflava* strains, including the type strain, are found (Parkinson *et al.,* 2011). These strains were further compared to *P. viridiflava* strains from plants capable of causing soft-rot on potato.

Strains related to phylogroup 7 lack a canonical T3SS, but they can be highly aggressive pathogens (Demba Diallo *et al.,* 2012). *P. viridiflava* isolated from wild *Arabidopsis* plants sampled in agricultural sites revealed two mutually exclusive Pathogenicity Islands (PAIs) that each encode a complete T3SS: T-PAI and S-PAI (Jakob *et al.,* 2002). These PAIs are structurally different and situated in two different chromosomal locations. The T-PAI is composed of the *hrp/hrc* gene cluster, the exchangeable effector locus (*EEL*) and the conserved effector locus (*CEL*), and it is organized like the classical tripartite T3SS described in other strains of the *P. syringae* complex (Alfano *et al.,* 2000). By contrast, the S-PAI is composed of only the *hrp/hrc* cluster, with a 10kb-

long insertion in the middle of the *hrp/hrc* containing the *avrE* effector and its chaperone (Araki *et al.,* 2006). Other effector and chaperone genes are present only in the T-PAI. Only one PAI is present in each strain and that the T-PAI and S-PAI strains seemed to differ in their production of pectolytic enzymes and speed of induction of a hypersensitive reaction (HR) on tobacco and *A. thaliana*.

We evaluated the phylogenetic and phenotypic diversity of putative *P. viridiflava* strains, the diversity of their PAIs and the correlation of the PAI profiles with phenotypes including the capacity to induce HR on tobacco, production of pectolytic enzymes and pathogenic host range. Our data confirm the existence of two *P. viridiflava* phylogroups, but with greater phenotypic and genotypic variability than previously appreciated. Importantly, we encountered numerous strains with an atypical LOPAT profile and that also had pronounced phase variation which influenced several phenotypes including pathogenicity. Our results highlight that despite some shared traits across strains, *P. viridiflava* is much more diverse than what was reported previously and our observations provide insights about the balance between the saprophytic and pathogenic life styles of this bacterium.

Results

P. viridiflava *is composed of 2 distinct phylogroups capable to induce potato soft rot*

The strains of *P. viridiflava* used in this study were collected from five different types of natural habitats (stream water, snow, rain, epilithic biofilms and leaf litter from an alpine meadow), from two species of wild plants, from four crop species and from irrigation water (Table 1). We screened the strains for the absence of cytochrome *c* oxidase and we sequenced the *cts* gene of these strains to detect those that were in the *P. syringae* complex. The screening of more than 750 strains of *P. syringae* typed for these traits as well as various other phenotypes demonstrated that the capacity to induce soft rot of potato slices is not found in the *P. syringae* complex outside of the phylogenetyc

groups of strains characterized in this present study (Berge, unpublished data). The delimitation between phylogroups was made by calculating the genomic distances obtained from the concatenated housekeeping genes (*cts*, *gapA*, *gyrB*, *rpoD*) with the Kimura 2-parameter model. A genomic distance <5% was used for delimitation of phylogroups. We used this distance as the maximum value for the delimitation of a phylogroup since it permitted us to obtain the phylogroups already described in the literature (Parkinson *et al*., 2011). Almost all strains belonged to the well known phylogroup 7 (Parkinson *et al.,* 2011) according to their *cts* sequence, or the sequences of all 4 housekeeping genes (Fig. 1). Soft rot was also caused by several strains in the closely-related phylogroup 8 (Fig. 1). These two groups form a monophyletic clade in the *P. syringae* species complex tree, supported by bootstrap values of 83% and 100% (Fig. 1) and were robust to gene or model application (Fig. S1). Following this classification 59 strains from different habitats belonging to phylogroups 7 and 8 were extensively characterized for their phenotypes and genotypes as described below.

Phylogroups 7 and *8 have high phenotypic heterogeneity influenced by phase variation and environmental origin of strains*

We investigated whether the *P. viridiflava* strains isolated from various habitats differed in several phenotypic traits. Only 37% of the strains tested presented the typical *P. viridiflava* LOPAT profile (absence of production of levan exopolysaccharide, induction of HR on tobacco and soft rot of potato). Fifty six percent of the strains produced a yellowish levan capsule after three days on levan sucrose medium, contrary to the classical morphology reported for *P. viridiflava* (levan-negative flat colonies in presence of 5% sucrose). Almost all the levan-producing strains displayed mucoid growth on KB medium. The mucoid yellowish colonies where observed on several rich nutrient media, including KB, independent of the presence of sucrose. Moreover, 55% and 27% of the strains were consistently positive and negative, respectively, for induction of HR on tobacco while

the remaining 18% were variable. All the HR-negative strains were tested in a supplementary experiment where three different clones for each strain were tested on tobacco and none induced HR. Fifty-four of 59 strains (92%) were able to cause soft rot to potato slices. Among the strains incapable of inducing soft rot on potato, two are in phylogroup 7 and three in phylogroup 8. Overall, phenotypic profiles of strains were variable (Fig. S2). The only traits that were identical for all strains were the inability to use sucrose and the capacity to use L–asparagine and D–tartrate as single carbon sources. With the exception of one strain in phylogroup 8, all strains also used arbutin and tween 80 and degraded esculin (Fig. S2). Strains in phylogroup 8 were positive in a bioassay for syringomycin-like toxin production but in PCR they were positive only for the presence of *syrB2* and they lacked *syrB1* and *syrC* genes (data not shown).

As noted, our *P. viridiflava* strains were isolated from a wide range of environmental habitats. Thus we hypothesized that the substrate of origin could influence some phenotypes. We tested the effect of origin of strains on phenotype and genotype, and grouped them into the following categories: i) plant *versus* non-plant origin, (including wild as well as cultivated plants), ii) water *versus* non-water origin (strains collected from water in the planktonic state and in biofilms were included in the water group), and iii) planktonic *versus* biofilm origin among those collected from water. For case "i" (plant *versus* non-plant) the only significant difference was that strains from non-plant habitats had a higher frequency of the lycopene cyclase gene ($P \leq 0.05$). The lycopene cyclase gene was found in the genome of the strains TA0043 and CC1582 and its presence/absence was confirmed by PCR on the total 59 strains. For case "ii" (water vs. non-water), the only significant differences were in pathogenicity tests, with strains from water being more aggressive ($P \leq 0.05$) on cantaloupe seedlings, bean pods, lemon and zucchini fruits than the strains from non-water substrates. This difference in pathogenicity appeared to be due to the biofilm strains from water habitats because for case "iii" (planktonic vs. biofilms), biofilm strains were significantly more aggressive ($P \leq 0.05$) on cantaloupe and zucchini fruit than were planktonic strains. No other

significant differences for case "iii" were observed. These results are in agreement with the previous hypothesis suggesting that water habitats are a reservoir of pathogenic *P. syringae* (Morris *et al.,* 2008, 2007).

In addition to the phenotypic diversity observed among strains, we observed variability among clones of a same strain. The formation of two colony types with different phenotypes is known as phase variation (Hallet *et al.*, 2001). As the definition is based on phenotype and the underlying molecular mechanisms are often unknown, we have referred to the intra-strain variability in *P. viridiflava* as phase variation. Two different colony morphologies were observed in almost all the strains listed in Table 1. Mucoid colonies (M) usually appeared after two days of incubation and large, flat transparent non-mucoid colonies (NM) were visible after four or more days on KB. We obtained stable clones derived from each of these colony types for 11 strains (Table 2) on KB medium, with no detectable reversion. BOX PCR profiles showed that the variants were clonal within a same strain (Fig. 2). Mucoid variants consistently induced soft rot on potato, liquefied gelatin and caused necrotic lesions on bean pods, whereas non-mucoid variants did not (Table 2, Fig. S3). There was no consistent effect of phase variant type on induction of HR. The difference in pathogenicity on cantaloupe seedlings of the M and NM lines was not as distinct as for the lesion test on bean pods. For strains PV841/09, LAB0163 and CC1582, significantly greater severity and incidence of the disease were observed for the M variant clones compared to the NM variants ($P \leq 0.05$). Neither M nor NM variants of strains PVBH nor BS0005 caused marked disease on cantaloupe seedling (Table 3).

The structures of the T3SS of phylogroups 7 and *8 suggest recombination events in genes located in the EEL*

In addition to the thorough phenotypic characterization, variability of traits related to the T3SS was investigated for all 59 *P. viridiflava* strains. Two different pathogenicity islands (T-PAI and S-PAI) were previously reported in *P. viridiflava* (Araki *et al.,* 2006). Only the T-PAI has an *EEL* while the

S-PAI lacks effector genes at this locus. Our analyses of the sequences at this position in T-PAI strains revealed the existence of two open reading frames close to the *hopA1*$_{(T)}$ gene, encoding for a lipoprotein and for a protein with an ABM domain typically present in monooxygenases involved in the biosynthesis of antibiotics. We also found the lipoprotein and the monooxygenase gene in the draft genome of three strains that corresponded to the S-PAI type but that lack *hopA1*$_{(T)}$ and *shcA*$_{(T)}$. This was observed in a genomie analysis of two our *P. viridiflava* strains from non-agricultural habitats (TA0043, Genbank accession AVDV00000000; and CC1582, Genbank accession AVDW00000000) and one other publically available assembly (UASWS0038, Genbank accession number NZ AMQP00000000). We tested the hypothesis that the two open reading frames were widely present in *P. viridiflava* independently of PAIs, by designing specific primers for amplifying both genes. The lipoprotein and the monooxygenese genes were present in almost all the strains analyzed (90%) even in strains lacking *hopA1*$_{(T)}$ and other T-PAI alleles (Fig. 1, Fig. S2). Further genomic analysis of TA0043 and CC1582 revealed that the lipoprotein and the monooxygenase genes are located in a region resembling an *EEL* but lacking effectors. This locus resembles the *EEL* in chromosomal location, bordered by tRNA-$_{leu}$ and the *queA* at the 5' end (Fig. 3), but it does not have an identifiable *hrpK*. In comparison with the *EEL* of strains PsyB728A and *Pto*DC3000 of *P. syringae,* as well as the *EEL* of T-PAI strains, the *P. viridiflava* S-PAI *EEL* lacked the *hrpK* gene, known to delimit the end of the *EEL* and to be a component of the *hrp* PAI with a putative function in translocation (Alfano *et al.,* 2000) and it lacked the *hopA1* and its chaperone *shcA* (Fig. 3). These results suggest that recombination events may have occurred between the PAIs or that the S-PAI strains lost part of the *EEL* during their evolution. The phylogenetic analysis of the concatenated lipoprotein and monoxygenase sequences (Fig. 4) showed evidence of horizontal transfer for these loci compared with the housekeeping phylogeny (Fig. 1), suggestsing that they were potentially acquired with the associated PAIs (Fig. 4). As observed by Araki *et al.,* (2007), strains having *hopA1* were rare: only seven strains were positive for the *hopA1* gene and for the T-PAI and did not have the alleles typical of the S-PAI when tested with PCR (Fig. S2). Forty-nine

strains had the three alleles typical of the S-PAI but had lipoprotein and monoxygenase as well. The remaining six strains had insufficient genes, according to results of PCR, to classify the T3SS according to the criteria of (Araki *et al.,* 2006).

Phylogenies constructed from the *hrcC* gene sequences showed that the T- and S-PAI have different evolutionary histories. As shown in the Bayesian trees of *hrcC* (Fig. 5) and *hrpL* (Fig. S4), the T-PAI strains cluster with the *P. syringae* phylogroups 5, 2 and 3, while the S-PAI strains form a clade more related to a strain of *P. cichorii* than the T-PAI strains and the other *P. syringae* phylogroups. Interestingly, strains in phylogroup 8 which lacked some effectors of T/S-PAI, except *hopA1*, *shcF* and *avrE* (Fig. 1) according to PCR results, are located in the T-PAI clade in the *hrcC* tree (Fig. 5).

Pathogenicity-related traits of P. viridiflava *are not strictly related to T3SS configuration*

The capability to induce disease on cantaloupe as well the ability to cause lesions on bean pods, lemon and zucchini fruits, was tested for all 59 *P. viridiflava* strains to determine the relationship between pathogenicity traits and T3SS configuration.

Results showed that fifteen strains were able to induce disease on cantaloupe and to cause lesions on all the fruits tested. Interestingly, none of these strains carried $hopA1_{(T)}$ or the T-PAI alleles. Eleven strains did not cause disease or lesions on any of the hosts tested, but only one of these had $hopA1_{(T)}$. Among the remaining, ten strains did not cause disease on cantaloupe or lesions on fruits, and among these ten strains, five had the alleles commonly present in the S-PAI and five had apparently incomplete T3SS based on PCR results (Fig. S2). Contrary to the observation of Jakob *et al* (2007), we did not observe clear differences in pathogenicity or potato rot between T- and S-PAI strains. On the other hand, all six strains that lacked the *avrE* gene (having neither the $avrE_{T\text{-}PAI}$ or $avrE_{S\text{-}PAI}$ allele) were not pathogenic on cantaloupe seedlings and did not cause lesions on the fruits

tested. These results suggest that *avrE* has an important role in *P. viridiflava* pathogenicity. Our results suggest that the only correlation between T3SS configuration and pathogenicity in *P. viridiflava* concerns the presence of an *avrE* allele.

Discussion

Our data support a portrait of *P. viridiflava* that differs from previous reports both in terms of phenotypes and importance of the T3SS in pathogenicity. Phenotypes of the LOPAT scheme, have been commonly used to differentiate *P. viridiflava* from other members of the *P. syringae* complex, a practice solidified by reports of homogeneity among strains within *P. viridiflava* (Goss *et al.,* 2005; Sarris *et al.,* 2012). We demonstrated that these characterization schemes are not completely reliable. Although potato soft rot is a phenotype of Pseudomonads unique to the *P. viridiflava* group (phylogroups 7 and 8) (Berge, unpublished data), 8% (5/60) of the *P. viridiflava* characterized here were not able to degrade potato slices. Hence, this trait is also not diagnostic. The phenotypic properties of *P. viridiflava* are further complicated by the almost universal occurrence of phase variation in this group, affecting the expression of previously diagnostic traits such as potato soft rot.

Our results illustrate that the S-PAI TTSS of *P. viridiflava* resembles that of the non-pathogenic strain Psy642 from the 2c clade, where it is correlated with pathogencity (Clarke *et al.,* 2010). However, for clades 7 and 8, our results demonstrate that S-PAI is not predictive of pathogenicity. In contrast, we found that the absence of *avrE* in both T- and S-PAI is correlated with the absence of pathogenicity. Since the presence/absence of other T3SS effectors was not associated with pathogenicity, *avrE* is an attractive target for future studies. Our results reflect findings from the potato soft rot pathogen *Pectobacterium carotovorum* subsp. *carotovorum* in which the only effector secreted during pathogenicity is DspE, a protein similar to AvrE (Hogan *et al.,* 2013). We speculate that the soft rot *P. viridiflava* and *P. carotovorum* strains do not require a wide range of

T3SS genes to suppress host immune responses since AvrE, likely in conjunction with pectolytic enzymes, is likely to be sufficient to induce disease symptoms. The ability to degrade pectin has probably allowed *P. viridiflava* and *P. carotovorum* to simplify their TTE repertories. This hypothesis is also supported by the presence of the S-PAI (a simpler T3SS) in most of the *P. viridiflava* strains isolated from environmental niches. The data we provide for the evolution of the T3SS of *P. viridiflava* may reflect that the T-PAI was acquired later during its evolutionary history. In particular, it seems that strains in phylogroup 7 may have acquired the T-PAI from those in phylogroup 8. This relationship is evident from the position of phylogroup 8 at the root of the *P. viridiflava* tree when the trees made with the housekeeping genes (Fig. 1) are compared to those made from *hrcC*. Since all the strains in this phylogroup, except CST0099, were non-pathogenic (Fig. 1), the tripartite organization of the T3SS did not provide a benefit to the pathogenicity of phylogroup 8. In this light, a more plausible explanation of the evolution of T-PAI in *P. viridiflava* may be its use as an adaptive tool in an environmental context outside of its association with plant habitats. Nevertheless, the presence of two different populations having seemingly the same ecological niche, but different T3SS, still needs to be clarified in *P. viridiflava.*

The results of this study suggest that *P. viridiflava* maintains a high level of adaptability, both as a saprophyte and as a pathogen. The different life-styles of the bacterium are reflected by its ubiquity in the environment. Recently Selezska and co-workers (2012) showed that *P. aeruginosa* is also widely distributed in water habitats. They propose that natural environments, rather than clinical habitats, drive the microevolution of this bacterial species. Phase variation is typically thought to be a means for bacteria to regulate pathogenicity via evasion of host defenses (Dubnau and Losick, 2006). However, phase variation changes phenotypes like motility, production of capsular material and various metabolic capacities and could also contribute to saprophytic survival and multiplication. In *P. viridiflava*, the mucoid variants may have an advantage in plants for two reasons. Firstly, the exopolysaccharide may increase tolerance to plant defense mechanisms. Secondly, the pectolytic ability of the mucoid variants could play an important role in releasing

sugars to support bacterial colonization. It has been demonstrated in *P. syringae* that alginate production confers resistance to toxic compounds and to desiccation, thereby increasing epiphytic fitness (Fett *et al.,* 1989). Furthermore, a correlation between expression of the *algD* gene and induction of HR on tobacco has been noted for *P. syringae* pv. *tomato* DC3000 (Keith *et al.,* 2003).

Mechanisms that regulate phase variation are generally unknown (Hallet *et al*., 2001). In pathogens such as *Escherichia coli*, *Haemophilus influenzae*, and *P. aeruginosa* the formation of antibiotic resistance variants is related to a defective mismatch repair system (MMR) (Matic *et al*., 1997, Watson *et al*., 2004, Ciofu *et al*., 2010). Mutations in MMR genes lead to a non-efficient DNA repair system leading to mutations in loci that influence gene expression. These mutations can be fixed and the re-acquisition of the original phenotype can occur by further mutations in the same genomic loci. However phase variation can also be the result of epigenetic alteration (Hallet *et al*., 2001). The mechanisms that regulate phase variation in *P. viridiflava* are unknown, though it could provide a useful tool for adapting to different habits and modulating bacterial fitness and survival.

The balance between saprophytic and pathogenic modes of *P. viridiflava* has likely also had a role in shaping the nature of its T3SS. As observed for strain Psy642 (Clarke *et al.,* 2010), *P. viridiflava* lacks *hrpK*, encoding a required translocation component of the T3SS (Alfano *et al.,* 2000). Furthermore, a minority of strains of *P. viridiflava* contain *hopA1*, which can have a role in enhancing virulence (Alfano *et al.,* 2000). Two EEL-associated genes, potentially encoding a monooxygenase while the other resembles a lipoprotein, could have toxic functions useful both in pathogenicity and in competition. Although the configuration of the *hrc/hrp* cluster in *P. viridiflava* is similar to that observed in the non pathogenic strain Psy642 (Clarke *et al.,* 2010), *P. viridiflava* clearly has pathogenic potential –albeit unpredictable– whereas strains related to Psy642 (Ps. phylogroup 2c) are not pathogenic (Demba Diallo *et al.,* 2012).

Our results provide new insights into the ecological behaviors of the well-studied *P. syringae* phylogroups 1, 2 and 7. Strains in phylogroups 1 and 2 (except for the 2c clade) have a canonical

T3SS. Group 2 strains are the most widely distributed and most apparently abundant in non-agricultural habitats (Morris *et al.,* 2010; Monteil *et al.,* 2013). They have a reduced number of effectors but carry more genes for production of different toxins than phylogroup 1 strains (Baltrus *et al.,* 2011). Strains from phylogroup 1 have evolved genes for adaptation to woody host plants (Green *et al.,* 2010). The phenotypes of *P. viridiflava* strains seem to reflect their ubiquitous presence in habitats such as biofilms and other aquatic contexts exposed to high light intensity. The lycopene cyclase genes found in *P. viridiflava* were absent from all other *P. syringae* strains with full-sequenced genomes (except strain ES4326 of *P. cannabina* pv. *alisalensis*) and were adjacent to other genes involved in carotenoid biosynthesis such as phytoene synthetase and β-carotene hydroxylase showing an organization that resembled an operon (data not shown). Carotenoids in non-photosynthetic bacteria are known to play an important role in protection against the effect of radicals generated in the presence of light (Armstrong and Hearst, 1996). In *P. viridiflava*, pathways for carotenoid biosynthesis may be crucial either on a leaf surface or in a biofilm ecosystem, providing protection against photo-oxidation. Additionally, carotenoids could modulate some metabolic activities such as motility of *P. viridiflava* under light stress conditions. In other *P. syringae* strains, for example, the photosensory proteins LOV-HK and BphP1 have been reported to influence swarming motility in response to both red and blue light (Wu *et al.,* 2013). The efficiency of these bacteria in degrading cell walls and, in particular of detached plant tissues, illustrates their competence in recycling carbon from primary producers; their phase variation suggests that they are adapted to a rapidly fluctuating availability of such carbon sources. Among the most intriguing question that arises from our results concerns the relative fitness trade-offs of the different modes of saprophytic life styles represented by *P. viridiflava*, *P. fluorescens*, *Pectobacterium carotovora* and the strains of *P. syringae* in phylogroup 2 (in clade 2c) that do not have the canonical T3SS.

The heterogeneity of *P. viridiflava* and the seemingly unpredictable nature of its pathogenicity complicate diagnostics and disease prediction. Based on the results presented here we propose that

detection of the presence of i) the monooxygenase and lipoprotein genes of ii) the allele referred to as *shcF*$_{S\text{-}PAI}$, and of iii) *hopA1* could be very useful in determining if strains that are in the complex are in fact *P. viridiflava*. All strains but one in phylogroup 7 have the monooygenase and lipoprotein couple (Fig. S2) and this pair of genes is not present in the genome sequences of strains of other phylogroups of *P. syringae* that are available. The *shcF*$_{S\text{-}PAI}$ allele is the most regularly present of the T3SS genes in phylogroup 8, and it can co-occur with *hopA1* whereas these genes do not co-occur in phylogroup 7. Therefore, the presence of the monooxygenase and lipoprotein couple or the co-occurrence of *hopA1* and *shcF*$_{S\text{-}PAI}$ would be a strong indication that a strain belongs to phylogroup 7 or 8. Characterization of the pathogenicity of strains suspected to be implicated in disease will require that particular attention is paid to the phase variation of strains during tests. Although phase variation complicates the characterization of *P. viridiflava*, it opens a promising door to disease control. A means to inhibit the emergence of the mucoid variant could be a powerful generic means to inhibit the pathogenicity of *P. viridiflava* independently of its specific relationship with a particular host. This strategy is currently being explored for the control of *P. aeruginosa* in lung infections of patients with cystic fibrosis (Pendersen *et al.,* 1992; Rau *et al.,* 2010; Deziel *et al.,* 2001).

Experimental procedures

Isolation and *selection of bacterial strains*

A total of 59 strains from different substrates collected mainly in the countries of the Mediterranean basin were used in this study (Table 1). The isolation of environmental strains was described previously as indicated in the Table 1. Further information about strain selection is provided in Methods S1.

Biochemical and *pathogenicity tests*

The objective of the phenotypic analysis was to characterize the variability of *P. viridiflava* strains from different substrates and sites. Strains were tested for the characteristics in the LOPAT scheme and for hydrolysis of gelatin, esculin, arbutin and tween80 as described previously (Lelliott *et al.,* 1966). Additional information is presented in Methods S1.

Genetic characterization

The genetic diversity of the strains was characterized in terms of the structure and sequences of the PAIs and the presence of the lycopene cyclase gene that was identified, through comparison of genomes, as being among the genes for pigment production unique to strains of *P. viridiflava*. For genomic analyses and comparisons, we used the draft genome sequences of strains TA0043 and CC1582 (Baltrus *et al.,* 2013).

Characterization of the PAIs is described in Methods S1.

Phylogenetic analyses based on housekeeping genes and *on T3SS genes*

A set of strains were chosen to represent the full diversity of our collection and to avoid clonal strains in the analysis. The criteria of choosing were the phylogeny of the strain according to *cts* and also their phenotypic traits. For this pool , fragments of the housekeeping genes *gapA*, *gyrB*, *rpoD* , in addition to the *cts* gene, were sequenced as described previously (Morris *et al.,* 2008). For phylogenetic analysis, the sequences were trimmed as concatenated with DAMBE version 5.1.1 (Xia 2013). The concatenated sequences (1852 bp) were used to construct a Bayesian phylogeny by using the Mr. Bayes program (http://mrbayes.csit.fsu.edu/) by using 500000 generation. Analysis was concluded when the standard deviation of split frequencies was <0.01 and burned in 100 samples. In addition, maximum likelihood and parsimony phylogenies were created with the Phylip package (http://evolution.genetics.washington. edu/phylip.html). Tree constructed with the different methods had the same topology; these led us to consider that phylogeny was robust. Consensus

trees were created from 100 independent phylogenies for both maximum likelihood and parsimony. Trees for each individual gene were also constructed with the same method.

The open reading frames close to the *hopA1* such as the *hrcC* and *hrpL* genes were sequenced by Macrogen Europe (The Netherlands) with the same primer set used in PCR. The genes *hrcC* and *hrpL* were sequenced to better investigate the evolution of the PAIs. Sequences were deposited on Plant Associated and Environmental Microbes Database (PAMDB) http://genome.ppws.vt.edu/cgi-bin/MLST/home.pl. Un-rooted trees for each gene were constructed as described above for phylogenetic analyses.

Characterization of phase variants

Two different colony types (mucoid and non-mucoid) were re-streaked on KB medium in order to stabilize each variant. After 3 different subcultures obtained by streaking single variants, among the total strains analyzed, 13 strains yielded stable variants. Six clones of each stable colony type per strain were randomly chosen and stored at –20°C in a phosphate buffer solution containing 40% glycerol for further analysis. The genotype of each variant was confirmed with BOX-PCR as described previously (Versalovic *et al.,* 1991). PCR reactions were performed with the Qiagen HotStarTaq®Master kit by using a single pure 48-h-old colony as a template. The PCR products were separated on 2% agarose gel at 4V cm^{-1} for 2 hours. All the stable phase variants with the same BOX profiles were tested for aggressiveness on cantaloupe and bean pods, for soft rot to potato, for gelatin liquefaction, HR on tobacco, utilization of D–tartrate, L–valine and L–alanine, degradation of arbutin and copper resistance as described above. Five of the 6 clones per each variant per each strain were tested.

Statistical analyses

The effect of genotype on the different phenotypes was evaluated with Fisher's exact test. GraphPad software, available on the web site http://graphpad.com/quickcalcs/contingency1.cfm, was used. Values of $P \leq 0.05$ were considered as statistically significant.

Acknowledgements

The salary of C. Bartoli was provided by a scholarship from Tuscia University (Italy). G. M. Balestra received funding from the Italian Minister of Agriculture and Food Polices (MIPAAF), Project ACTINIDIA - OIGA n. 247 that allowed for the isolation of Italian strains. Work on this project at UNC was supported by NIH grant R01-GM066025 (to CDJ, JLD). JLD is an HHMI Investigator and this work was funded in part by the Howard Hughes Medical Institute and the Gordon and Betty Moore Foundation (via Grant GBMF3030). DAB was supported through startup funds from the University of Arizona. We thank Dr. Boris A. Vinatzer, Virginia Tech, for his support in the genomic analysis and for the useful discussion about T3SS in non-pathogenic *P. syringae*. We thank technical staff at INRA for propagation of plant materials used in this study. The authors declare no conflict of interest.

References

Alfano, J.R., Charkowski, A.O., Deng, W.L., Badel, J.L., Petnicki-Ocwieja, T., Van Dijk, K., *et al.* (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 pathogenicit. *Proc Natl Acad Sci USA* **97**: 4856–4861.

Araki, H., Innan, H., Kreitman, M., and Bergelson, J. (2007). Molecular evolution of pathogenicity-island genes in *Pseudomonas viridiflava*. *Genetics* **177**: 1031–1041.

Araki, H., Tian, D., Goss, E.M., Jakob, K., Halldorsdottir, S.S., Kreitman, M., *et al.* (2006). Presence/absence polymorphism for alternative pathogenicity islands in *Pseudomonas viridiflava*, a pathogen of *Arabidopsis*. *Proc Natl Acad Sci USA* **103**: 5887–5892.

Armstrong, G., and Hearst, J. (1996). Carotenoids 2: Genetics and molecular biology of carotenoid pigment biosynthesis. *Fed Am Soc Exp Biol J* **10**: 228–237.

Balestra, G.M., and Varvaro, L. (1997). Epiphytic survival and control of *Pseudomonas viridiflava* on *Actinidia deliciosa*. *Acta Hort* **444**: 745–749.

Balestra, G.M., and Varvaro, L. (1998). Seasonal fluctuations in kiwifruit phyllosphere and ice nucleation activity of *Pseudomonas viriflava*. *J Plant Pathol* **80**: 151–156.

Baltrus, D.A., Nishimura, M.T., Romanchuk, A., Chang, J.H., Mukhtar, M.S., Cherkis, K., *et al.* (2011). Dynamic evolution of pathogenicity revealed by sequencing and comparative genomics of 19 *Pseudomonas syringae* isolates. *PLoS Pathogens* **7**: e1002132.

Baltrus, D.A., Yourstone, S., Lind, A., Guibaud, C., Sands, D.C., Jones, C.D., *et al.,* (2013). Draft genomes for a phylogenetically diverse suite of *Pseudomonas syringae* strains from multiple source populations. *Genome Announc* **2: 01195-13.**

Billing, E. (1970). *Pseudomonas viridiflava* (Burkholder, 1930; Clara 1934). *J Appl Bacteriol* **33**: 492–500.

Caruso, P., and Catara, V. (1996). First report of *Pseudomonas viridiflava* leaf spot of red-leaved chicory. *Plant Dis* **80**: 710.

Ciofu, O., Mandsberg, L.F., Bjarnsholt, T., Wassermann, T., Hoiby, N. (2010). Genetic adaptation of *Pseudomonas aeruginosa* during chronic lung infections of patients with cystic fibrosis: strong and wick mutators with heterogeneous genetic backgrounds emerge in *mucA* and *lasR* mutants. *Microbiology* **156**: 1108–1119.

Clarke, C.R., Cai, R., Studholme, D.J., Guttman, D., and Vinatzer, B.A. (2010). *Pseudomonas syringae* strains naturally lacking the classical *P. syringae hrp/hrc* locus are common leaf colonizers equipped with an atypical type III secretion system. *Mol Plant Microbe Interact* **23**: 198–210.

Conn, K.E., and Gubler, W.D. (1993). Bacteria blight of kiwifruit in California. *Plant Dis* **7**: 228–230.

Demba Diallo, M., Monteil, C.L., Vinatzer, B.A., Clarke, C.R., Glaux, C., Guilbaud, C., *et al.* (2012). *Pseudomonas syringae* naturally lacking the canonical type III secretion system are ubiquitous in nonagricultural habitats, are phylogenetically diverse and can be pathogenic. *Int Soc Microb Ecol* **6**: 1325–1335.

Deziel, E., Comeau, Y., and Villemur, R. (2001). Initiation of biofilm formation by *Pseudomonas aeruginosa* 57RP correlates with emergence of hyperpiliated and highly adherent phenotypic variants deficient in swimming, swarming, and twitching motilities. *J Bacteriol* **183**: 1195–1204.

Dubnau, D., and Losick, R. (2006). Bistability in bacteria. *Mol Microbiol* **61**: 564–572.

Everett, K.R., and Henshall, W.R. (1994). Epidemiology and population ecology of kiwifruit blossom blight. *Plant Pathol* **43**: 824–830.

Fett, W.F., Osman, S.F., and Dunn, M.F. (1989). Characterization of exopolysaccharides produced by plant-associated fluorescent pseudomonads. *Appl Environ Microbiol* **55**: 579–83.

Gardan, L., Shafik, H., Belouin, S., Broch, R., Grimont, F., and Grimont, P.A. (1999). DNA relatedness among the pathovars of *Pseudomonas syringae* and description of *Pseudomonas tremae* sp. nov. and *Pseudomonas cannabina* sp. nov. (ex Sutic and Dowson 1959). *Int J Syst Bacteriol* **49**: 469–478.

González, A.J., Rodicio, M.R., Carmen, M., Gonza, A.J., and Mendoza, M.C. (2003). Identification of an emergent and atypical *Pseudomonas viridiflava* lineage causing bacteriosis in plants of agronomic importance in a spanish region. *Appl Environ Microbiol* **69**: 2936–2941.

Goss, E.M., Kreitman, M., and Bergelson, J. (2005). Genetic diversity, recombination and cryptic clades in *Pseudomonas viridiflava* infecting natural populations of *Arabidopsis thaliana*. *Genetics* **169**: 21–35.

Goss, E.M., Potnis, N., and Jones, J.B. (2013). Grudgingly sharing their secrets: new insight into the evolution of plant pathogenic bacteria. *New Phytol* **199**: 630–632.

Goumans, D.E., and Chatzaki, A.K. (1998). Characterization and host range evaluation of *Pseudomonas viridiflava* from melon , blite , tomato , chrysanthemum and eggplant. *Eur J Plant Pathol* **104**: 181–188.

Green, S., Studholme, D.J., Laue, B.E., Dorati, F., Lovell, H., Arnold, D., *et al.,* (2010). Comparative genome analysis provides insights into the evolution and adaptation of *Pseudomonas syringae* pv. *aesculi* on *Aesculus hippocastanum*. *PloS One* **5**: e10224.

Hallet, B., (2001). Playing Dr Jekyll and Mr Hyde: combined mechanisms of phase variation in bacteria. *Curr Opin Microbiol* **4**: 570–581.

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

Hogan, C.S., Mole, B.M., Grant, S.R., Willis, D.K., and Charkowski, A.O. (2013). The type III secreted effector DspE is required early in *Solanum tuberosum* Leaf infection by *Pectobacterium carotovorum* to cause cell death, and requires Wx(3–6)D/E motifs. *PLoS One* **8**: e65534.

Jackson, R.W., Athanassopoulos, E., Tsiamis, G., Mansfield, J.W., Sesma, A., Arnold, D.L., *et al.* (1999). Identification of a pathogenicity island, which contains genes for virulence and avirulence, on a large native plasmid in the bean pathogen *Pseudomonas syringae* pathovar *phaseolicola*. *Proc Natl Acad Sci USA* **96**: 10875–10880.

Jakob, K., Goss, E.M., Araki, H., Van, T., Kreitman, M., and Bergelson, J. (2002). *Pseudomonas viridiflava* and *P. syringae*-natural pathogens of *Arabidopsis thaliana*. *Mol Plant Microbe Interact* **15**: 1195–1203.

Jakob, K., Kniskern, J.M., and Bergelson, J. (2007). The role of pectate lyase and the jasmonic acid defense response in *Pseudomonas viridiflava* virulence. *Mol Plant Microbe Interact* **20**:146–158.

Keith, R.C., Keith, L.M., Hernàndez-Guzmàn, G., Uppalapati, S.R., and Bender, C.L. (2003). Alginate gene expression by *Pseudomonas syringae* pv. *tomato* DC3000 in host and non-host plants. *Microbiology* **149**: 1127–1138.

Lelliott, R.A, B.illing, E., and Hayward, A.C. (1966). A determinative scheme for the fluorescent plant pathogenic pseudomonads. *J. Appl. Bacteriol* **29**:470–489.

Liao, C-H., Hung, H-Y., and Chatterjee, A.K. (1988). An extracellular pectate lyase in the pathogenicity factor of the soft-rotting bacterium *Pseudomonas viridiflava*. *Mol Plant Microbe Interact* **1**: 199–206.

Lindeberg, M., Stavrinides, J., Chang, J.H., Alfano, J.R., Collmer, A., Dangl, J.L., *et al.* (2005). Proposed guidelines for a unified nomenclature and phylogenetic analysis of type III Hop effector proteins in the plant pathogen *Pseudomonas syringae*. *Mol Plant Microbe Interact* **18**: 275–282.

Matic, I., Radman, M., Taddei, F., Picar, B., Doit, C., Bingen, E., *et al*.(1997). Highly variable mutation rates in commensal and pathogenic *Escherichia coli*. *Science* **227**: 1833–1834.

Mohr, T.J., Liu, H., Yan, S., Morris, C.E., Castillo, J.A., Jelenska, J., *et al.* (2008). Naturally occurring nonpathogenic isolates of the plant pathogen *Pseudomonas syringae* lack a type III secretion system and effector gene orthologues. *J Bacteriol* **190**: 2858–2870.

Monteil, C.L. (2011). Ecologie de *Pseudomonas syringae* dans un basin versant: vers un modèle de transfert dans habitats naturels aux agro-systèmes. *Phd thesis*.

Monteil, C. L., Cai, R., Liu, H., Mechan Llontop, M., Leman, S., Studholme, D., *et al* (2013). Non-agricultural reservoir contribute to emergence and evolution of *Pseudomonas syringae* crop pathogens. *New Phytol* **199**: 800–811.

Morris, C.E., Kinkel, L.L., Xiao, K., Prior, P., and Sands, D.C. (2007). Surprising niche for the plant pathogen *Pseudomonas syringae*. *Infec Genet Evol* **7**: 84–92.

Morris, C.E., Sands, D.C., Vanneste, J.L., Montarry, J., Oakley, B., Guilbaud, C., *et al.* (2010). Inferring the evolutionary history of the plant pathogen *Pseudomonas syringae* from its

biogeography in headwaters of rivers in North America, Europe and New Zealand. *mBio* **1**: 00107–10.

Morris, C.E., Sands, D.C., Vinatzer, B.A., Glaux, C., Guilbaud, C., Buffière, A., *et al.,* (2008). The life history of the plant pathogen *Pseudomonas syringae* is linked to the water cycle. *Int Soc Microb Ecol* **2**: 321–34.

Morris, C.E., Wen, A-M., Xu, X-H., and Di, Y-B. (1991). Ice nucleation-active bacteria on chinese cabbage in Northern China:population dynamics and characteristics and their possible role in storage decay. *Phytopathology* **82**: 739–746.

Mulet, M., Lalucat, J., and García-Valdés, E. (2010). DNA sequence-based analysis of the *Pseudomonas* species. *Environ Microbiol* **12**: 1513–30.

O'Brien, H.E., Thakur, S., and Guttman, D.S. (2011). Evolution of plant pathogenesis in *Pseudomonas syringae*: a genomics perspective. *Annu Rev Phytophatol* **49**: 269–289.

Parkinson, N., Bryant, R., Bew, J., and Elphinstone, J. (2011). Rapid phylogenetic identification of members of the *Pseudomonas syringae* species complex using the *rpoD* locus. *Plant Pathol* **60**: 338–344.

Pendersen, S.S., Hoiby, N., Espersen, F., and Koch, C. (1992). Role of alginate in infection with mucoid *Pseudomonas aeruginosa* in cystic fibrosis. *Thorax* **47**: 6–13.

Rau, M.H., Hansen, S.K., Johansen, H.K., Thomsen, L.E., Workman, C.T., Nielsen, K.F., *et al.* (2010). Early adaptive developments of *Pseudomonas aeruginosa* after the transition from life in the environment to persistent colonization in the airways of human cystic fibrosis hosts. *Environ Microbiol* **12**: 1643–1658.

Sarkar, S.F., and Guttman, D.S. (2004). Evolution of the core genome of *Pseudomonas syringae* , a highly clonal , endemic plant pathogen. *Appl Environ Microbiol* **70**: 1999–2012.

Sarris, P.F., Trantas, E.A., Mpalantinaki, E., Ververidis, F., and Goumas, D.E. (2012). *Pseudomonas viridiflava*, a multi host plant pathogen with significant genetic variation at the molecular level. *PloS One* **7**: e36090.

Selezska, K., Kazmierczack, M., Musken, M., Garbe, J., Shobert, M., Haussler, S. et al. (2012). *Pseudomonas aeruginosa* population structure revisited under environmental focus: impact of water quality and phage pressure. *Environ Microbiol* 14: 1952–1967.

Végh, A., Hevesi, M., Némethy, Z., and Palkovics, L. (2012). First report of bacterial leaf spot of basil caused by *Pseudomonas viridiflava* in Hungary. *Plant Dis* **96**: 141.

Versalovic, J., Koeuth, T., and Lupski, J. (1991). Distribution of repetitive DNA-sequences in eubacteria and application to fingerprinting of bacterial genomes. *Nucleic Acids Res* **19**: 6823–6826.

Watson, M.E., Burns, J.L., Smith, A.L. (2004). Hypermutable *Haemophilus influenzae* with mutation in *mutS* are found in cystic fibrosis sputum. *Microbiology* **150**: 2947-2958.

574 Wu, L., McGrane, R.S., and Beattie, G.A. (2013). Light regulation of swarming motility in
575 *Pseudomonas syringae* integrates signaling pathways mediated by a bacteriophytochrome and a
576 LOV protein. *mBio* **4**:e00334–13.

577 Xia, X. (2013). DAMBE5: A comprehensive software package for data analysis in molecular
578 biology and evolution. *Mol Biol Evol* doi:10.1093/molbev/mst064.

579 Zoina, A., Puopolo, G., Pasini, C., D'Aquila, F., Cozzolino, L., and Raio, A. (2004). *Pseudomonas*
580 *viridiflava* is the casual agent of a bacterial disease of *Ranunculus asiaticus*. *J Plant Pathol* **86**: 339.

581

582

583

584 **Table 1**. List of strains used in this study.

Strain	Year isolated	Substrate	Place of origin	Reference or source
AI0086	2007	stream water	New Zealand	(Morris *et al.,* 2010)
BS0001	2008	*Acitinidia deliciosa*	Italy	This study
BS0002	2008	*Acitinidia deliciosa*	Italy	This study
BS0004	2008	*Acitinidia deliciosa*	Italy	This study
BS0005	2008	*Acitinidia deliciosa*	Italy	This study
CC0657	2004	*Primula officinalis*	France	(Demba Diallo *et al.,* 2012)
CC1486	2006	snowfall	France	(Morris *et al.,* 2008)
CC1492	2006	snowfall	France	(Morris *et al.,* 2008)
CC1582	2006	epilithic biofilm	France	(Morris *et al.,* 2010)
CCE0322	2009	stream water	France	(Monteil 2011)
CCE0328	2009	stream water	France	(Monteil 2011)
CCV0172	2009	stream water	France	(Demba Diallo *et al.,* 2012)
CCV0178	2009	stream water	France	(Demba Diallo *et al.,* 2012)
CCV0180	2009	stream water	France	(Demba Diallo *et al.,* 2012)
CEB0010	2010	epilithic biofilm	France	This study
CEB0022	2010	epilithic biofilm	France	This study
CEB0029	2010	epilithic biofilm	France	This study
CEB0041	2010	epilithic biofilm	France	This study
CEB0085	2010	epilithic biofilm	France	This study
CMA0031	2009	snowpack	Maroc	(Demba Diallo *et al.,* 2012)
CMO0103	2010	rain	France	(Monteil 2011)
CMO0110	2010	rain	France	(Monteil 2011)
CMO0085	2010	rain	France	(Monteil 2011)
CMW0006	2011	river water	France	This study
CMW0028	2011	river water	France	This study
CST0072	2010	rain	France	(Monteil 2011)
CST0079	2010	rain	France	(Monteil 2011)
CST0099	2010	rain	France	(Monteil 2011)
CSZ0285	2009	snowpack	France	(Monteil 2011)
CSZ0297	2009	snowpack	France	(Demba Diallo *et al.,* 2012)
CSZ0341	2009	stream water	France	(Monteil 2011)
CSZ0342	2009	stream water	France	(Monteil 2011)
CSZ0716	2010	snowpack	France	(Monteil 2011)
CSZ0855	2010	leaf litter	France	(Monteil 2011)
FMU0107	1991	*Brassica pekinensis*	Chine	(Monteil 2011)
GAW0092	2011	irrgation water	France	This study
GAW0197	2011	irrgation water	France	This study
GAW0203	2011	irrgation water	France	This study
JT0006	2007	*Actinidia deliciosa*	Italy	This study
LAB0006	2009	epilithic biofilm	France	This study
LAB0023	2009	epilithic biofilm	France	This study
LAB0123	2010	epilithic biofilm	France	This study
LAB0124	2010	epilithic biofilm	France	This study

LAB0126	2010	epilithic biofilm	France	This study
LAB0162	2010	epilithic biofilm	France	This study
LAB0163	2010	epilithic biofilm	France	This study
LYR0041	2011	rain	France	(Monteil 2011)
LYR0042	2011	rain	France	(Monteil 2011)
LYR0045	2011	rain	France	(Monteil 2011)
PV841/09	2004	*Ranunculus acri*	Italy	(Zoina *et al.,* 2004)
PVB-H	2012	*Ocimum basilicum*	Hungary	(Végh *et al.,* 2012)
PVCT26.1.1	1994	*Cichorium intybus*	Italy	(Caruso and Catara 1996)
PVCT26.3.1	1994	*Cichorium intybus*	Italy	(Caruso and Catara 1996)
SZB0012	2009	epilithic biofilm	France	This study
TA0002	2007	stream water	France	(Morris *et al.,* 2010)
TA0020	2007	stream water	France	(Morris *et al.,* 2010)
TA0043	2007	*Primula officinalis*	France	(Morris *et al.,* 2010)
UB0259	2006	stream water	France	(Morris *et al.,* 2010)

585

586

Table 2. Assays conducted to characterize phase variants. For each stable variant, three different clones were used per test and the experiments were repeated twice.

	Utilisation, as sole carbon sources, of								
Stable [a] variants[a]	**Bean pods[b]**	**Potato rot[b]**	**Gelatin hydrolysis[b]**	**HR on tobacco**	**L−Valine**	**D−Tartrate**	**D−Alanine**	**Cu Resistence[c]**	**Arbutin hydrolysis**
CMO0085-M	+	+	+	+	+	+	–	+	+
CMO0085-NM	–	–	–	+	+	+	–	+	+
PVBH-M	+	+	+	+	+	+	+	–	+
PVBH-NM	–	–	–	–	+	+	+	–	+
BS0002-M	+	+	+	+	+	+	–	–	+
BS0002-NM	–	–	–	+	+	+	–	–	+
BS0005-M	+	+	+	+	+	+	+	+	+
BS0005-NM	–	–	–	+	+	+	+	+	+
LAB0163-M	+	+	+	+	+	+	–	–	+
LAB0163-NM	–	–	–	+	+	+	–	–	+
CC1582-M	+	+	+	–	+	+	–	–	+
CC1582-NM	–	–	–	–	+	+	–	–	+
CCV0172-M	+	+	+	+	+	+	–	–	+
CCV0172-NM	–	–	–	–	+	+	–	–	+
JT0006-M	+	+	+	–	+	+	+	–	+
JT0006-NM	–	–	–	–	+	+	+	–	+
PV841/09-M	+	+	+	+	+	+	–	+	+
PV841/09-NM	–	–	–	–	+	+	–	+	+
PVCT26.1.1-M	+	+	+	+	+	+	+	+	+
PVCT26.1.1-NM	–	–	–	+	+	+	+	+	+
TA0043-M	+	+	+	–	+	+	–	–	+
TA0043-NM	–	–	–	–	+	+	–	–	+

[a] Strains in which the two different phases were well separated when re-streaked for a second time each variant on KB medium.

[b] For each strain tested all M variant clones gave positive reactions and no reactions were observed for the NM variant clones. The reactions reported for the variants of each strain were homogeneous.

[c] Two different copper concentrations were tested: 0.64mM and 1.12mM according with the *P. syringae* tolerance curve (Andersen *et al.*, 1991). Results were the same for both copper concentrations is each variant.

597 **Table 3.** Pathogenicity of phase variants on cantaloupe seedlings

Variants[a]	Incidence[b]	Severity[c]
BS0005-M	0.41	0.51
BS0005-NM	0.36	0.45
PVBH-M	0.28	0.25
PVBH-NM	0.11	0.14
PV841/09-M	0.43	0.47
PV841/09-NM	0.00	0.00
LAB0163-M	0.63	0.88
LAB0163-NM	0.1	0.15
CC1582-M	0.75	1.13
CC1582-NM	0.05	0.05

598 [a] Five clones per each variant were inoculated on 12 cantaloupe seedlings
599 [b] Then frequency of cantaloupe seedlings (per 12) showing disease at 7 days after inoculation
600 [c] Severity was evaluated on a scale from 0 to 4 at 7 days after inoculation.

601

602 **Figures Legends**

603 **Figure 1.** Bayesian phylogenetic tree constructed with the concatenated housekeeping genes *cts,*
604 *gapA, gyrB* and *rpoD* (1852bp). The tree was rooted on *P. aeruginosa* (PAO). Posterior
605 probabilities are indicated at each node. Names of the branches correspond to the strains, except for
606 the branches indicated as phylogroups reported in previously work (Parkinson *et al.,* 2011). The
607 substrates are indicated close to the strain names. The black squares indicate a positive reaction or
608 presence of the genes and the white squares stand for a negative reaction or absence of the genes.
609 Grey squares indicate a variable reaction or non-stable phase variants. The two phylogroups, 7 and
610 8 are separated by a mean distance of 5.5% sequence difference. Strains in phylogroup 8 are
611 GAW0197, CST0099, CMO0085, GAW0203 and LYR0041. Abbreviations are: ORF = open
612 reading frames found in the *EEL* of T-PAI strains and encoding for a lipoprotein and a
613 monooxygenase involved in antibiotics biosynthesis. INA = ice nucleation activity. The strains
614 were considered positive when at least 2 of 3 drops containing 10^6 cells per drop froze at
615 temperatures warmer than –9°C. SYR = presence of a syringomycin-like toxin based on inhibition
616 of *Geotricum c*and*idum*.

617 **Figure 2**. Box PCR profiles of mucoid and non-mucoid variants. The first and second lanes of each
618 strain correspond to the mucoid and non-mucoid variants, respectively

619 **Figure 3.** Structure of the locus resembling an *EEL* of *P. viridiflava* TA0043 and CC1582. The *EEL*
620 was identified in draft genomes of strains TA0043 and CC1582. Letters refer to the putative protein
621 function:

622 A= tgt tRNA-guanine transglycosylase, queusine-34-forming, B= Queuine synthetase (queA), C=
623 Laba-A-like N1 domain protein (conserved protein found in different bacteria with unknown
624 function), D= Hypothetical protein found only in *P. viridiflava*, E= Pstpo1411 like protein, F= eel
625 protein found in different *P. syringae* strains but without effector function, G= lipoprotein and

monooxygenase genes, H= Pstpo1371 like protein (conserved effector locus protein), I= hypothetical protein only found in *P. viridiflava*, L=transcriptional factor.

Figure 4. Tree based on the lipoprotein and monooxygenase genes found in the putative *EEL* of the S-PAI strains and in the *EEL* of the T-PAI strains. The Bayesian method was employed to construct the tree. Posterior probabilities are indicated at each node. Sequences for the LU9.1a, PT220.1a, ME210.1b and UASWS0038 strains were extracted from GeneBank. Accession numbers for each strain are: AY859095.1, AY859099.1, AY859100.1 and NZ_AMQP01000083.1, respectively.

Fig 5. Un-rooted Bayesian phylogenetic tree constructed with the *hrcC* gene sequences. T- and S-PAI strains are delimited with black bars. Sequences for LP23.1a, PNA3.3a, RMX23.1 and M3.1b were obtained from GeneBank. Accession numbers are respectively: AY597277.1, AY597278.1, AY597282.1, AY597281.1.

638
639 **Supporting Information**

640 Additional supporting information is available in the online version of the article.

641 **Fig. S1** Neighbor joining trees constructed on the basis of the single housekeeping gene *cts*, *gyrB*,

642 *gapA* and *rpoD.* Posterior probabilities are indicated at each node.

643 **Fig. S2** Neighbor joining tree based on the *cts* sequences was compared with the phenotypic pattern

644 of 59 strains. The genotype of the strains for T3SS genes is also shown.

645 **Fig. S3** Different reactions on bean pods and potato rot between mucoid and non-mucoid variants.

646 **Fig. S4** Bayesian tree based on *hrpL* gene sequences.

647 **Table S1** *hrp/hrc* components found in the *P. viridiflava* genomes (TA0043 and CC1582 strains).

648 **Table S2** *P. syringae* strains used to test the specificity of the primers designed for the type three

649 secretion genes and the lycopene cyclase gene.

650 **Methods S1.** Selection and characterization of the strains. Biochemical tests and genomic typing.

651

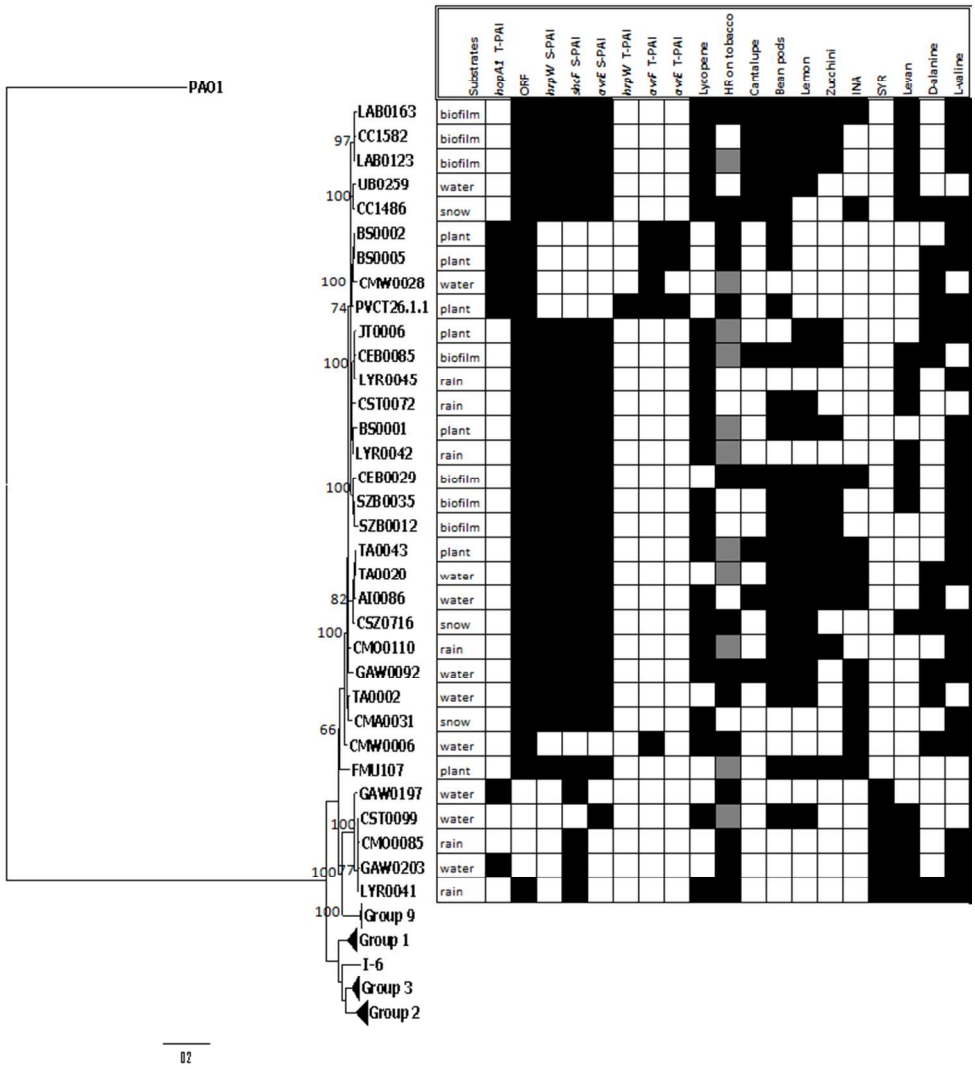


Figure 1. Bayesian phylogenetic tree constructed with the concatenated housekeeping genes cts, gapA, gyrB and rpoD (1852bp). (see manuscript for full legend)
69x76mm (300 x 300 DPI)

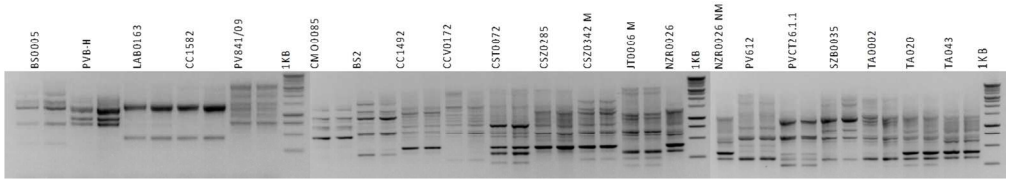


Figure 2. Box PCR profiles of mucoid and non-mucoid variants. The first and second lanes of each strain correspond to the mucoid and non-mucoid variants, respectively
413x190mm (96 x 96 DPI)

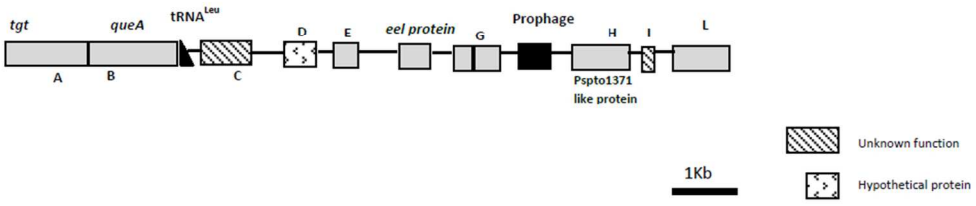


Figure 3. Structure of the locus resembling an EEL of P. viridiflava TA0043 and CC1582. The EEL was identified in draft genomes of strains TA0043 and CC1582. Letters refer to the putative protein function:
(see manuscript for full legend)
257x65mm (96 x 96 DPI)

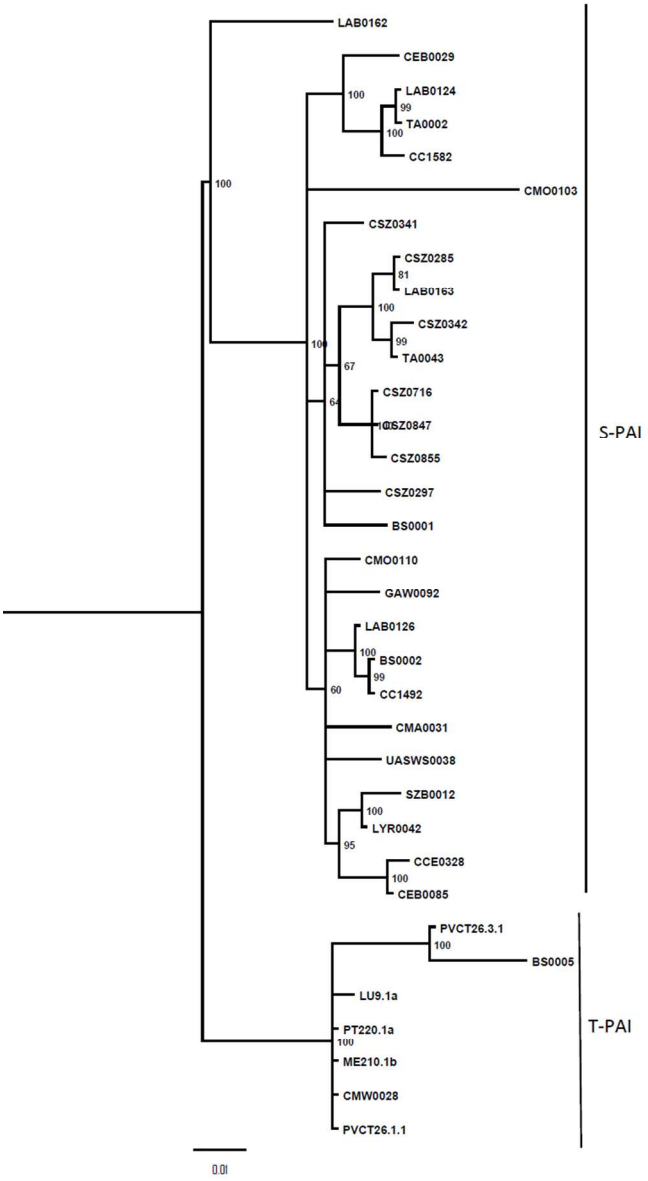


Figure 4. Tree based on the lipoprotein and monooxygenase genes found in the putative EEL of the S-PAI strains and in the EEL of the T-PAI strains. (see manuscript for full legend)
220x375mm (96 x 96 DPI)

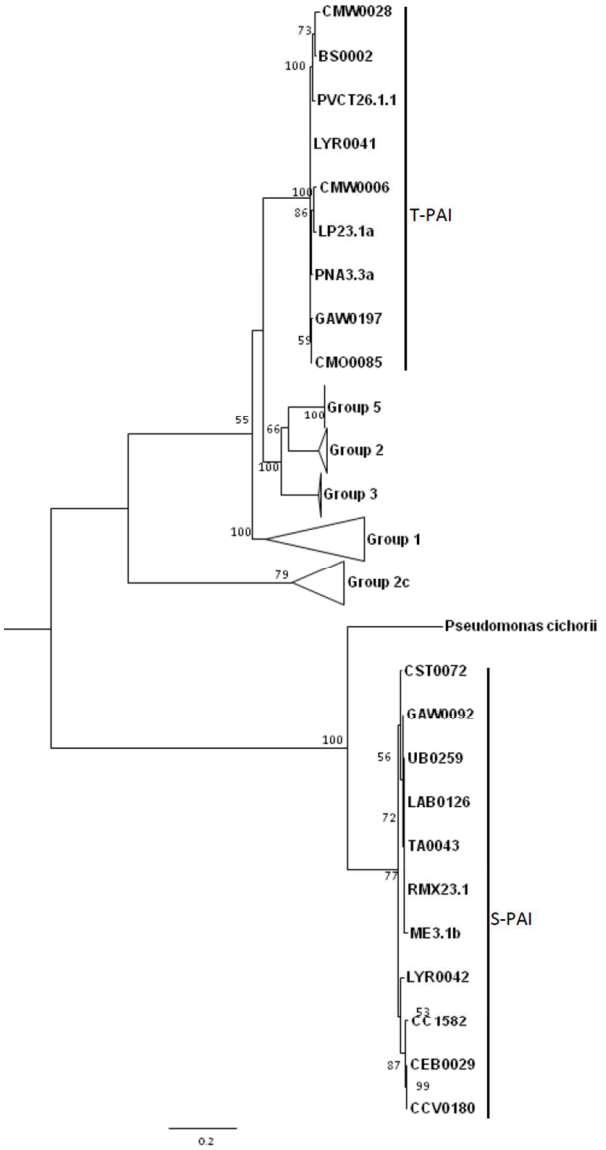
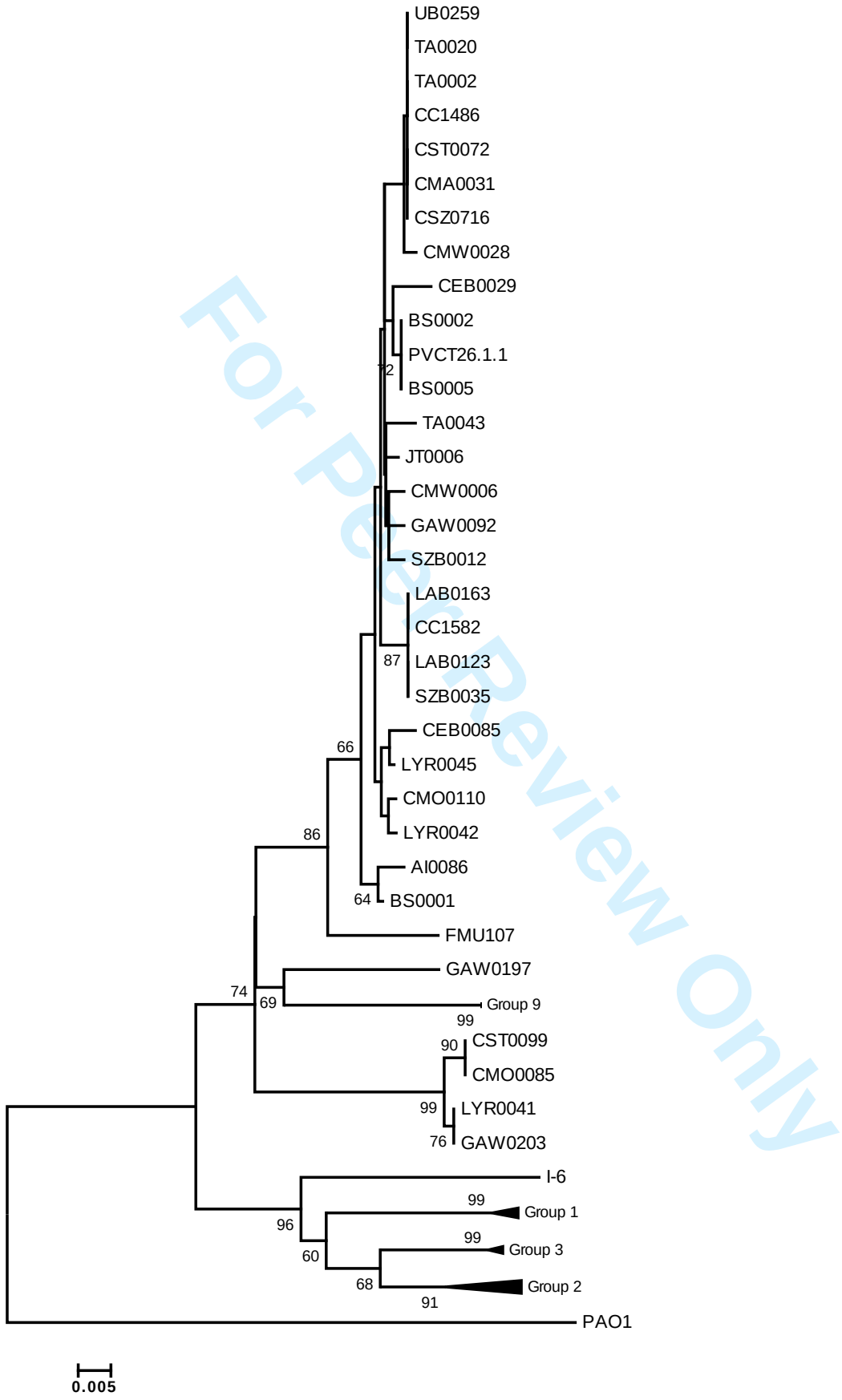
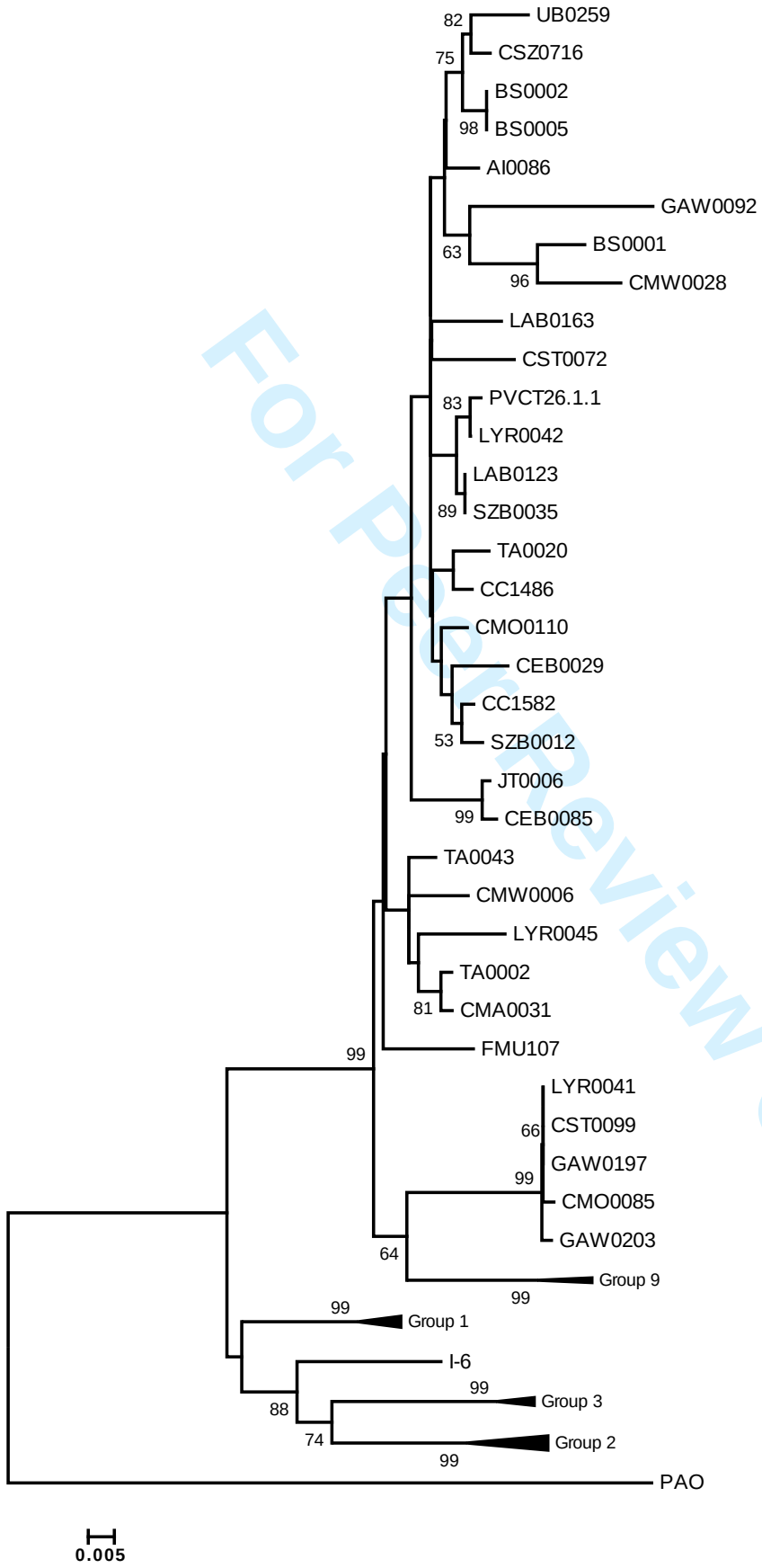


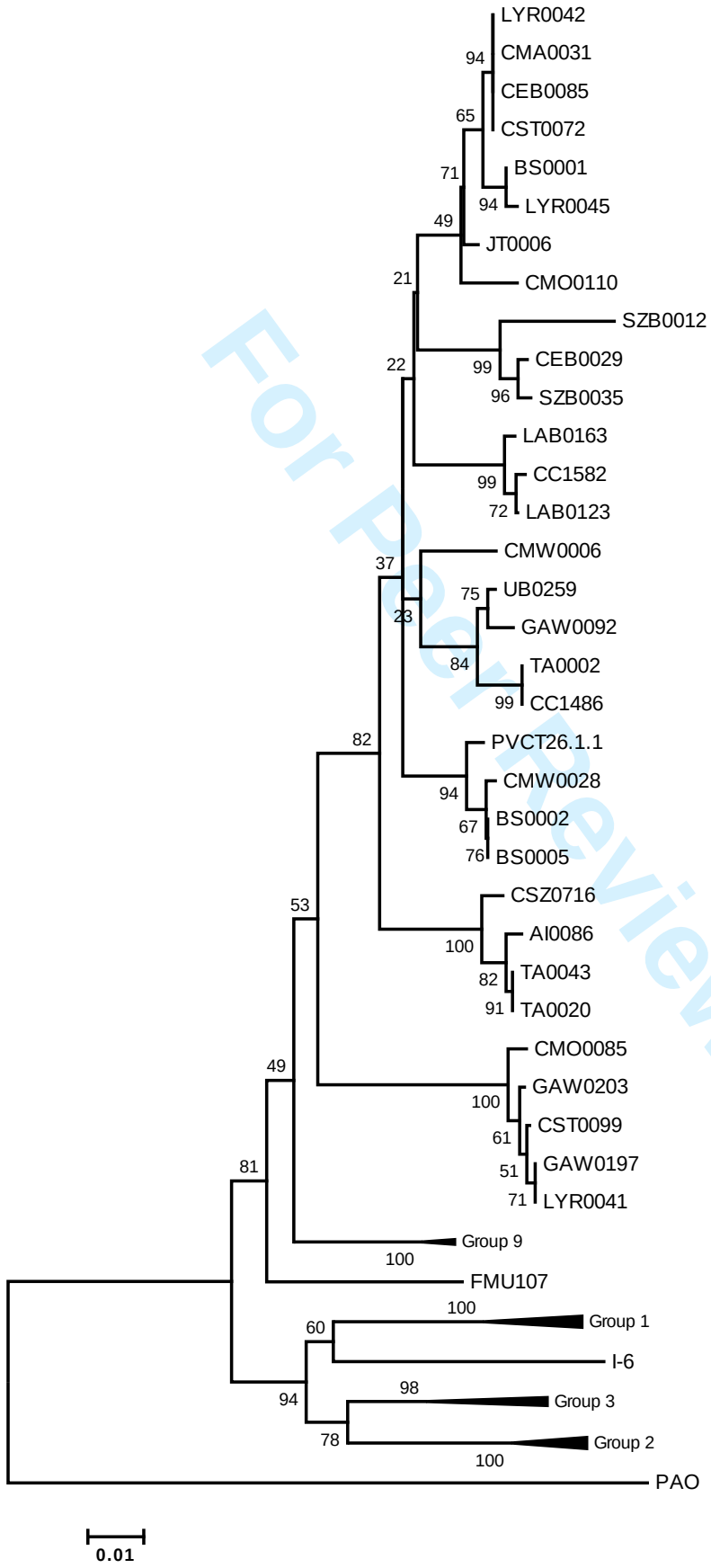
Fig 5. Un-rooted Bayesian phylogenetic tree constructed with the hrcC gene sequences. T- and S-PAI strains are delimited with black bars. Sequences for LP23.1a, PNA3.3a, RMX23.1 and M3.1b were obtained from GeneBank. Accession numbers are respectively: AY597277.1, AY597278.1, AY597282.1, AY597281.1.
210x384mm (96 x 96 DPI)



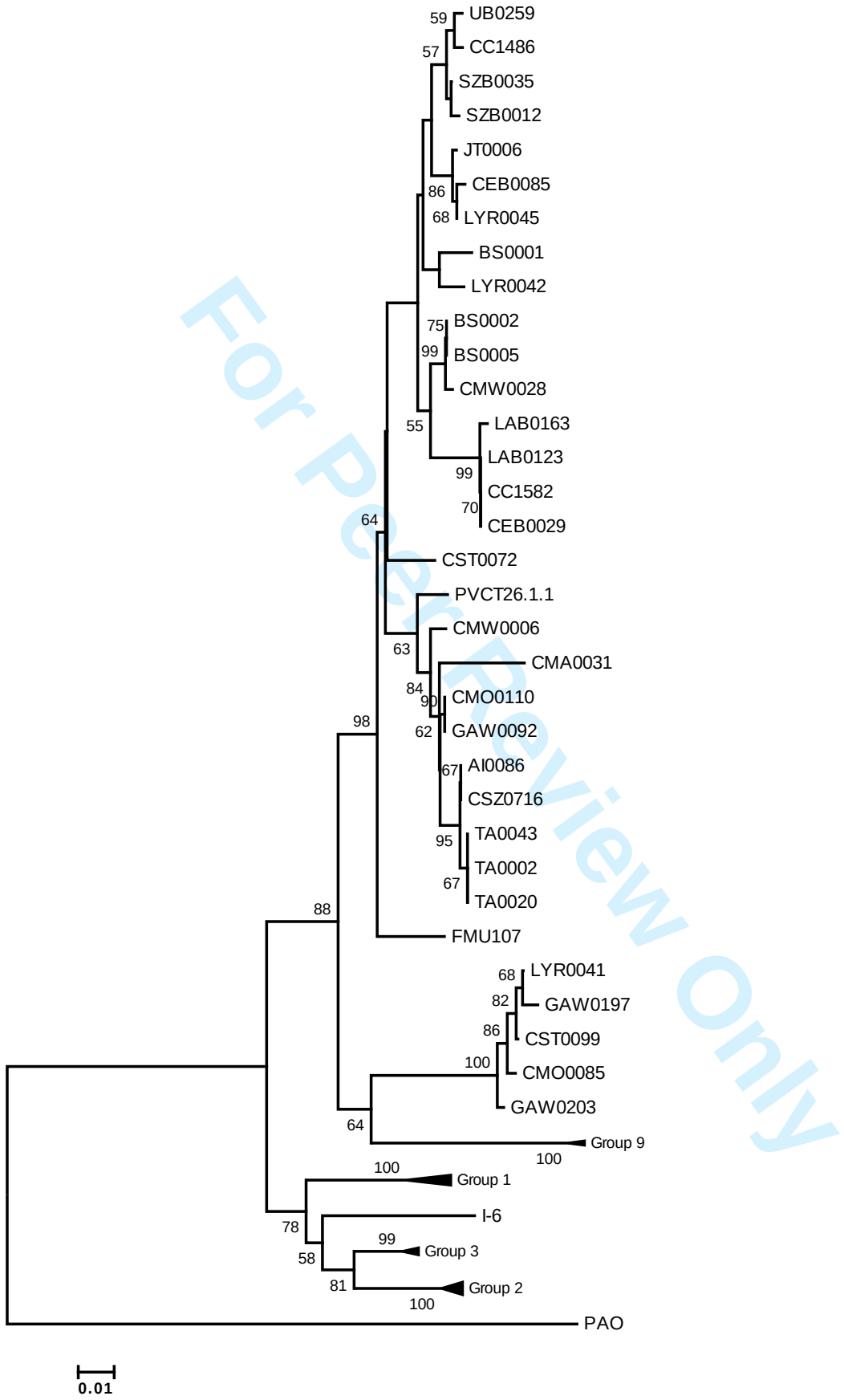
Neighbor joining tree constructed on *cts* gene sequences.



Neighbor joining tree constructed on *gapA* gene sequences

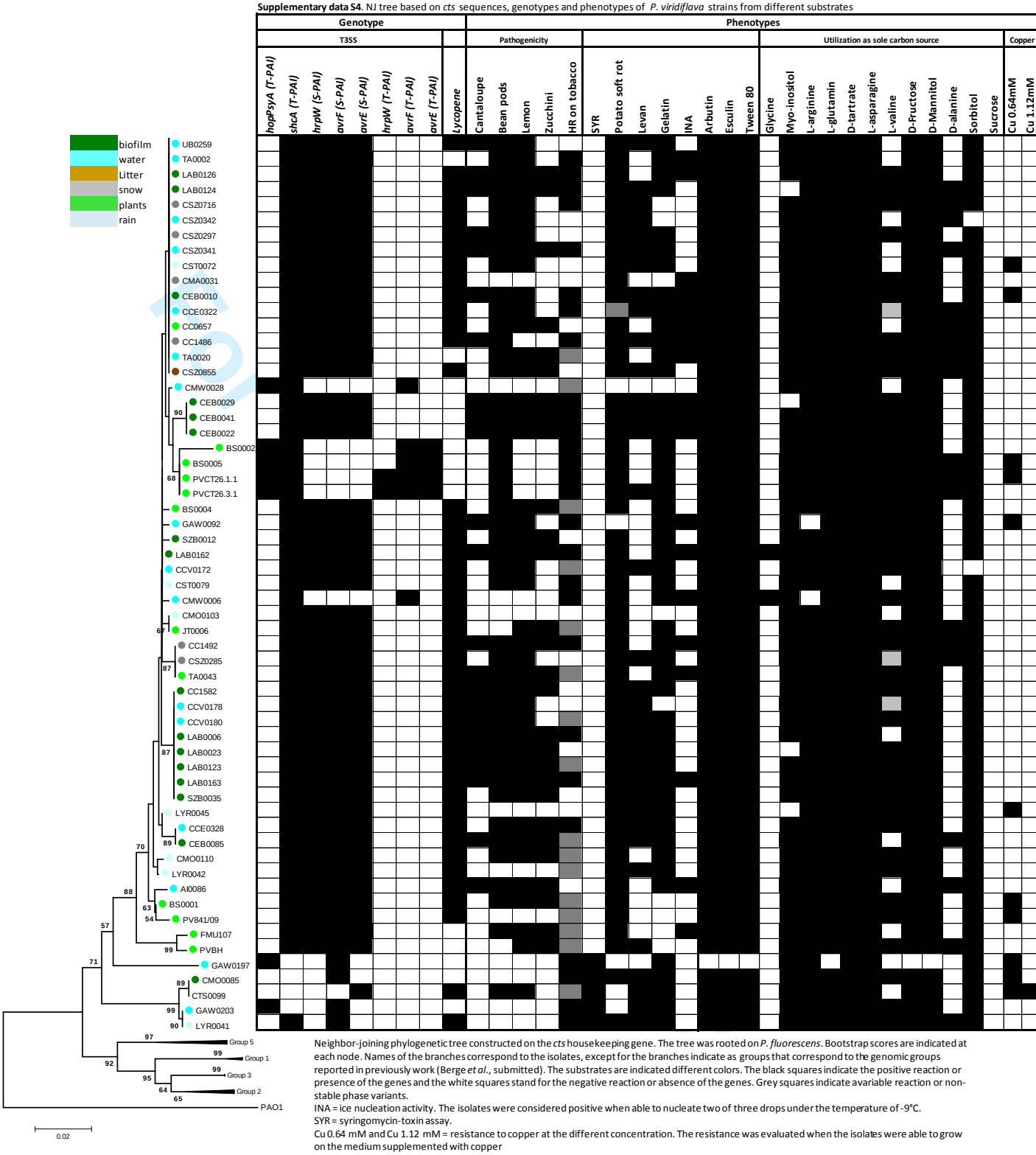


Neighbor joining tree constructed on *gyrB* gene sequences

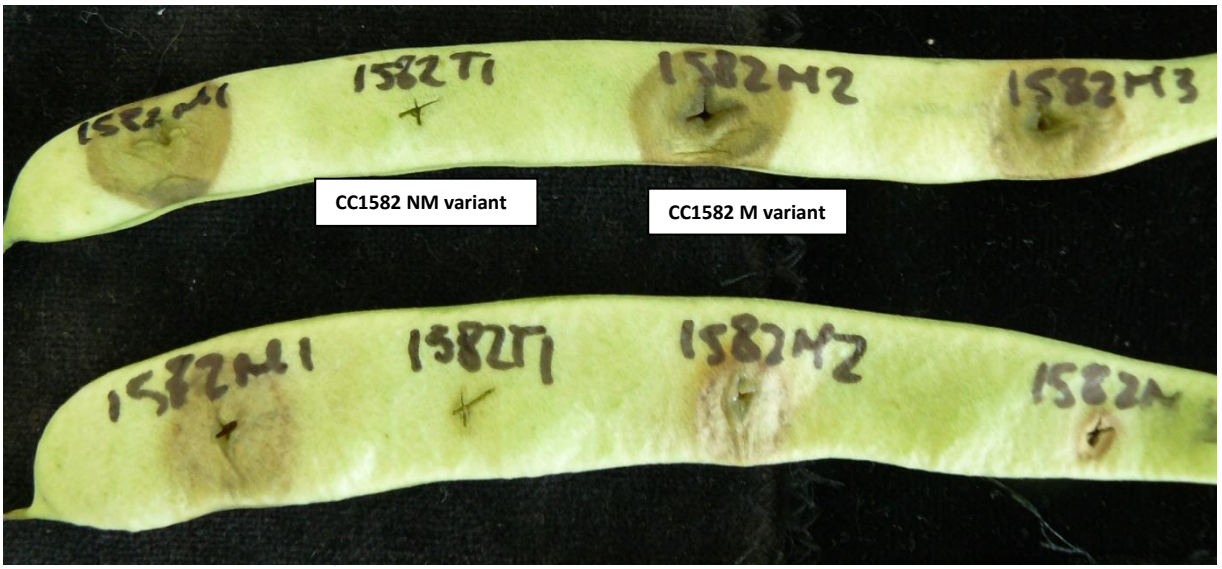


Neighbor joining tree constructed on *rpoD* gene sequences

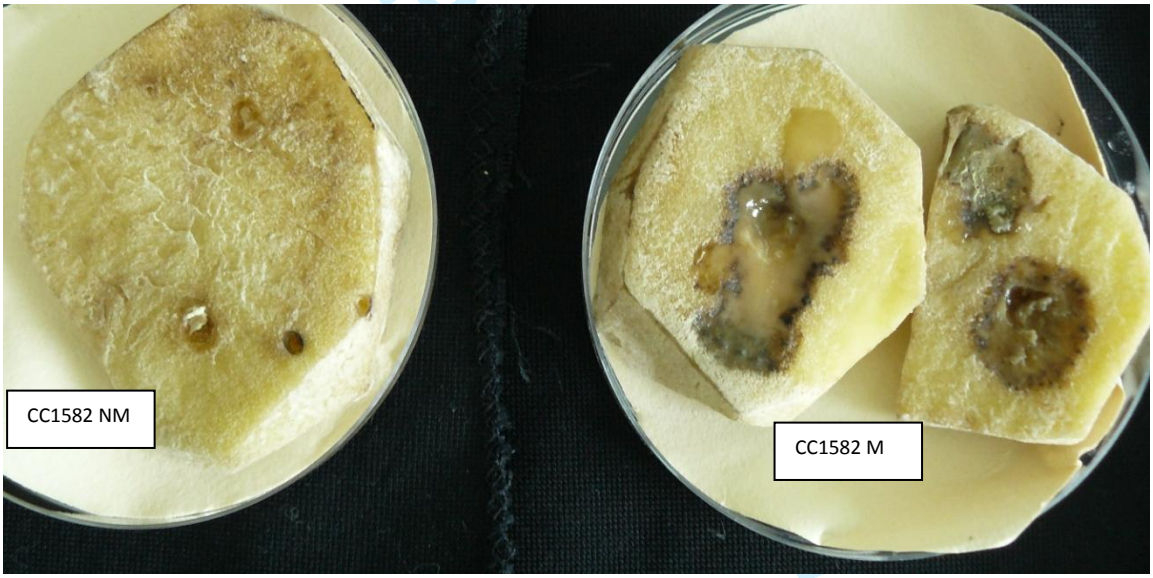
Supplementary data S4. NJ tree based on *cts* sequences, genotypes and phenotypes of *P. viridiflava* strains from different substrates



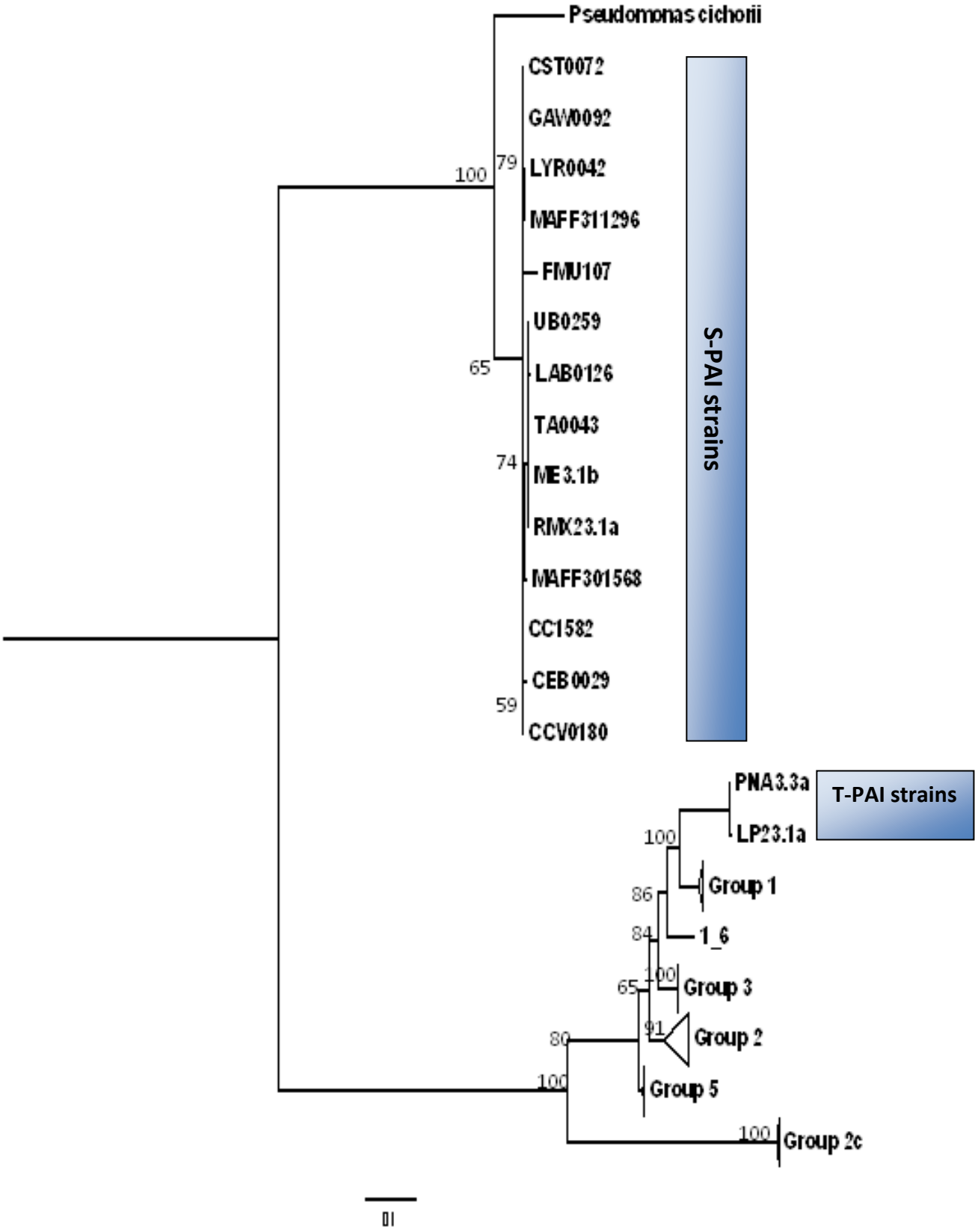
Neighbor-joining phylogenetic tree constructed on the *cts* housekeeping gene. The tree was rooted on *P. fluorescens*. Bootstrap scores are indicated at each node. Names of the branches correspond to the isolates, except for the branches indicate as groups that correspond to the genomic groups reported in previously work (Berge *et al.*, submitted). The substrates are indicated different colors. The black squares indicate the positive reaction or presence of the genes and the white squares stand for the negative reaction or absence of the genes. Grey squares indicate avariable reaction or non-stable phase variants.
INA = ice nucleation activity. The isolates were considered positive when able to nucleate two of three drops under the temperature of -9°C.
SYR = syringomycin-toxin assay.
Cu 0.64 mM and Cu 1.12 mM = resistance to copper at the different concentration. The resistance was evaluated when the isolates were able to grow on the medium supplemented with copper



Lesion test on bean pods with mucoid (M) and non-mucoid (NM) CC1582 variant.



Potato soft rot of CC1582 non mucoid variant (NM) on the left and CC1582 mucoid variant (M) on the right



Supplementary Information Fig. 3. Phylogenies was constructed based on the sequences of the *hrpL* gene locates in the *hrp/hrc* cluster of both T and S-PAI strains. Bayesian method was employed to construct the tree. Posterior probabilities are showed at each node.

Supplementary table. Comparison of the *hrp/hrc* clusters of PsyB728a and of strains TA0043 and CC1582 of *P. viridiflava*

TA0043 and CC1582 *hrp/hrc* protein	Length aligment PsyB728a	Amino acid identity to PsyB728a	E-value	Psy642 component
HrpR	208/276	75%	4.00E-116	+
HrpS	-	-	-	-
HrpL	98/149	66%	3.00E-43	+
HrpJ	166/320	52%	5.00E-32	+
HrpV	64/113	57%	3.00E-19	+
HrcC	435/658	64%	0:00	+
HrcJ	167/204	82%	3.00E-85	+
HrpZ	65/110	59%	3.00E-10	a
HrpA	-	-	-	-
HrpK	-	-	-	-
HrpW	149/198	75%	2.00E-72	-
HrpQ	177/340	52%	2.00E-59	+
HrcN	180/335	54%	7.00E-54	+
HrpO	-	-	-	+
HrpP	-	-	-	+
HrcQ	52/71	73%	3.00E-15	+
HrcR	186/215	87%	1.00E-106	+
HrcS	72/84	86%	1.00E-20	+
HrcT	154/212	73%	2.00E-50	+
HrcU	256/350	73%	3.00E-129	+
HrpT	34/69	49%	2.00E-06	+
HrpG	-	-	-	+
HrpF	-	-	-	+
HrpE	113/192	59%	2.00E-27	+
HrpD	-	-	-	+
HrpB	53/95	56%	2.00E-07	+

a= no significant alignment

Table S1. Strains used to test the specificity of PCR primers for the T3SS *P. viridiflava* components

Strain No.	Genomic group	pathovar or species name	substrate	Country	Year	Reference
CFBP7286	1	*Pseudomonas syringae* pv. *actinidiae*	kiwifruit	Italy	2008	4
DC3000	1	*P. s*. pv. *tomato*	tomato			
NCPPB3487	1	*P. s*. pv. *avellanae*	hazelnut	Greek	1976	6
CFBP 1748	3	*P. s*. pv. *papulans*	apple tree	USA	1975	2
CFBP 1617	2	*P. s*. pv. *aptata*	beet	USA	1959	3
P6	2	P. s. pv. *syringae*	kiwifruit	Portugal	2006	this study
CFBP 1323	1	P. s. pv. *tomato*	tomato	France	1971	
CFBP 1676	3	P. s. pv. *pisi*	pea	New Zealand	1969	7
Pse NE107	3	*P. savastanoi* pv. *savastanoi*	olive	Nepal	2007	1
CC0094	2	Pseudomonas sp.	melon	France	1997	5
36b5		*Pseudomonas fluorescens*				

References:

1) Balestra G. M., Lamichhane J. R., Kshetri M. B., Mazzaglia A.& Varvaro L. (2009). First report of olive knot caused by *Pseudomonas savastanoi* pv. *savastanoi* in Nepal. Plant Pathology 58: 393.
2) Hanvantari B. N. A taxonomic study of *Pseudomonas papulans* Rose 1917 (1977). New Zeal J Agr Res 20: 557-561.
3) Hill C. F (1979). New plant disease records in New Zealand. New Zeal J Agr Res 22: 641-645.
4) Mazzaglia A., Studholme D., J., Taratufolo M. C., Cai R., Almeida N. F., Goodman T., Guttman D. S., Vinatzer B.A.& Balestra G. M (2012). *Pseudomonas syringae* pv. *actinidiae* (PSA) isolated form recent bacterial canker of kiwifruit outbreaks belong to the same genetic lineage. Plos One 7: 36518.
5) Morris C. E., Kinke, L. L., Xiao K., Prior P., Sands D. C (2007). Surprising niches for the plant pathogen *Pseudomonas syringae*. Infect Genet Evol 7: 84-92.
6) O'Brien H.E., Thakur S., Gong Y., Fung P., Zhang J., Yuan L., Wang P.W., Yong C., Scortichini M.& Guttman D.S (2012). Extensive remodeling of the *Pseudomonas syringae* pv. *avellanae* type III secretome associated with two independent host shifts onto hazelnut. BMC Microbiol 12: 141.
7) Young J. M. & Dye D. W (1969). Bacterial blight of peas caused by *Pseudomonas pisi* Sackett, 1916 in New Zealand. New Zeal J Agr Res 13: 315-324.

Supporting Methods S1

Isolation and *selection of bacterial strains*

For strains from crops, the parts of the leaves nearest to the lesions were surface disinfested with 1% NaOH and washed with sterilized distilled water, macerated with a sterilized scalpel and plated on nutrient agar supplemented with 5% sucrose and incubated at 25°C for 3 days. Strains were considered to be in the *P. syringae* complex if they were fluorescent on King's B (KB) medium (Shaad *et al.,* 1980) and tested negative for the presence of cytochrome C oxidase according to the techniques used to identify putative *P. syringae* strains from environmental reservoirs (Mohr *et al.,* 2008; Morris *et al.,* 2010). Among the strains from crops, those that induced soft-rot on potato slices and that did not produce arginine dihydrolase (see methods for these tests below in the biochemical test section) were selected for further characterization to determine if they would be used in this study. But for strains from the environment, soft-rot of potato was not determined prior to selecting the strains for further characterization. All the strains were stored in 0.1M phosphate buffer at 4°C and in 40% of glycerol at –80°C prior to characterization.

To select the definitive group of strains used for this study we determined the phylogenetic position of strains based on the partial sequence of the *cts* (citrate synthase) gene. The strains characterized in this study had *cts* sequences that placed them in or very close to phylogroup 7 (Parkinson *et al.,* 2011), which included clades TA002 and TA020 (Morris *et al.,* 2010). PCR was performed with a fresh 10^8 CFU·ml^{-1} bacterial suspension as template by using *cts* primers (Sarkar and Guttman, 2004). The PCR reagents and conditions used are described by (Morris *et al.,* 2008). Amplified products were loaded on 1% Agarose gels with ethidium bromide and the sequencing analysis was conducted by Eurofins MWG Operon (Ebersberg,

Germany). All the sequences were aligned and cut to the same size with DAMBE version 5.1.1 (Xia 2013). The percent difference in sequence similarity between strains was calculated with Phylip package (http://evolution.genetics.washington. edu/phylip.html).

Biochemical and *pathogenicity tests*

The utilization of D–fructose, D–mannitol, D–sorbitol and sucrose as sole carbon sources was tested in a mineral salts medium composed of 0.05% $NH_4H_2PO_4$, 0.05% K_2HPO_4, 0.02% $MgSO_4 \bullet 7H_2O$, 1.2% agar and supplemented with 1.2 % bromocresol purple (in ethanol) as an indicator of acidification (Gerhardt *et al.,* 1981). Utilization of D–alanine, glycine, myo-inositol, L–arginine, L–glutamine, D–tartrate, L–asparagine and L–valine was detected by adding 0.5% w/v of each filter-sterilized amino acid to a minimal medium containing 0.1% KH_2PO_4, 0.002% $MgSO_4 \bullet 7H_2O$, 0.001% KCl, 0.001% phenol red and 0.8% agar. Resistance to copper was evaluated by using methods described previously (Andersen *et al.,* 1991) and two different $CuSO_4$ concentrations (0.64 mM and 1.12 mM) were assayed. Ice nucleation activity (INA), induction of a HR on tobacco (*Nicotiania tabacum* L. cv Samsum) and production of syringomycin-like toxins (SYR) were evaluated as described previously (Morris *et al.,* 2008). Pathogenicity on several plants was also tested as described below.

To determine the capacity to induce HR on tobacco, three replicates for each strain were tested. For strains that gave negative results, three different clones were tested in a second round of trials. To investigate whether the syringomycin-like toxins influenced results of the HR test, the SYR-positive strains were grown for five days in syringomycin liquid medium (Bender *et al.,* 1999) and an aliquot of the medium filtrate was infiltrated into tobacco leaves.

To test the virulence of the strains on different hosts, bacteria were grown on KB agar for 48 h and bacterial suspensions were prepared in sterilized distilled water at a concentration of 10^8

CFU·ml^{-1}. An aliquot of 700µl of these suspensions was infiltrated into the cotyledons of 7-day-old cantaloupe seedlings (*Cucumis melo* var. *cantalupensis* Naud. cv. Vèdrantis) (Morris *et al.,* 2008). Plants were maintained for one week in a growth chamber at 23°C with an 18 h photoperiod. Symptoms on seedling were scored after 1 week post inoculation and a severity scale from 0 to 4 was associated as 0 (no symptoms), 1 (hypertensive reaction in the point of inoculation), 2 (necrosis on the entire seedling), 3 (two seedlings necrotized), 4 (wilting or death of the entire plant). Bean pods (*Phaseolus vulgaris* L., cv. Corallo), lemon fruits (*Citrus limonum*) and zucchini fruits (*Cucurbita pepo*), purchased at a market, were inoculated by injecting 10µl of bacteria suspension directly in the fruits previously injured with a sterilized scalpel. Before inoculation, all the fruits were partially sterilized with 1% sodium hypochlorite and three replicate inoculations for each strain were made on three separate fruits. Sterilized distilled water was used as a negative control in all the tests. Fruits were kept in a humid chamber at room temperature for three days prior to scoring of the reactions.

Genetic characterization

The comparison of the complete genomes of the strains TA0043 and CC1585 with the PAI regions for strain PNA3.3a (region 1) and strain RMX23.1a (region 2) (Araki *et al.,* 2006) revealed the presence of two open reading frames previously found to be adjacent to *hopA1*$_{(T)}$ and its chaperone, s*hcA*$_{(T)}$, in the *EEL* locus of the T-PAI strains, by Araki *et al.,*, (2006) (Araki *et al.,* 2006) although TA0043 and CC1582 strains carry the alleles typical of the S-PAI. Based on the sequences of strains TA0043, CC1582, PNA3.3a and RMX23.1a we designed primer sets for two open reading frames described above and for partial sequences of *hrpW*$_{(S)}$, *avrE*$_{(S)}$ and *shcF* $_{(S)}$, the chaperone of *avrE* using Primer3 and OligoAnalyzer software available on the website http://frodo.wi.mit.edu/primer3. Primer sets were also designed for the *hrcC* and *hrpL* genes following the same criteria described above. The

specificity of the primers was tested by amplifying reference strains belonging to different phylogroups in the *P. syringae* complex (Table S2). The primers used and their annealing temperatures are reported in Table 1. To amplify the homologous T-PAI genes $hrpW_{(T)}$, $avrE_{(T)}$ and $shcF_{(T)}$, we used the primers described previously (Jakob *et al.,* 2007). The lycopene cyclase gene was found in both TA0043 and CC1582 genomes and in the genome of *Pseudomonas cannabina* pv. *alisalensis* strain ES4326 (also known as *P. syringae* pv. *maculicola* ES4326) but in no other publically available genomes for strains within the *P. syringae* complex. Specific primers for *P. viridiflava* were designed by comparing the three sequences. The presence of the canonical T3SS was checked for some *P. viridiflava* strains used in this study as described previously (Mohr *et al.,* 2008). In addition, for the strains that gave positive results in the syringomicin-like toxin bioassay we performed PCR on the three genes involved in the production of the syringomycin as described before (Bultreys and Gheysen, 1999). All the primers were synthesized by Eurofins MGW Operon (Ebersberg, Germany).

PCR was performed by using GoTaq®Flexi DNA Polymerase Promega with the annealing temperatures listed in Table 4. PCR products were loaded in 1% agarose gels with ethidium bromide to confirm the presence or absence of the amplicons.

Table 1. Primers used in this study and the corresponding temperatures of annealing for PCR.

Gene	Primer name	Sequence (5'-3')	Annealing temp. (°C)	Product size
$hopA1_{(T)}$	HopA-2-Fw	TGT GCG ATC AGA CAC ATC AG	57	500bp
	HopA-2-Rv	AGT ACC TGC GCG ATC TGA TC		
Lipoprotein	ORF1/2-Fw	CGA CCT GCT TTC GAT CA	55	900bp
	ORF1/2-Rv	TCA ATA CTC TGG AGA TCA G		
$hrpW_{(S)}$	HrpW-Fw	TGG AGG TGG AAC ACC TTC	57	480bp
	HrpW-Rv	TGG TCC AGT GGA CGT TAT C		
$shcF_{(S)}$	shcF-Fw	CTA AGT GCC ACT CTC GGT A	57	850bp
	shcF-Rv	ATC CTT GGT CTG CCT GTC		
$avrE_{(S)}$	AvrE-Fw	CAT CCA TCG CGA GGT TGT	57	1000bp

	AvrE-Rv	AGA GGT TGA CCA GCG TAT C		
Lycopene cyclase	Lyco-FW	ACC GGA TGA GTT GCG TC	57	300bp
	Lyco-Rv	CGA CTC GGC TTC GAA GT		
hrcC	hrcC-SPAI-Fw	CGC TGA AAC GGC TTT TCT GGA	60	1500bp
	hrcC-SPAI-Rv	TTG CTC ACC CGA TCC CTT TTC		
hrcC	hrcC-Deg-Fw	ABT TYC AGT GGT TYC TBT AYA ACG	59	1300bp
	hrcC-Deg-Rv	GRT CGA GCT GAT CGC CVA YCA		
hrpL	hrpL-SPAI-Fw	AAG GTT GGT ACG TTC GCT GCT CT	62	600bp
	hrpL-SPAI-Rv	GAA CCT CCT TGG AAT ACA CGC TG		
hrpL	hrpL-Deg-Fw	GAC GTS GAT GAC CTB ATS CAG	59	400bp
	hrpL-Deg-Rv	GCC GYG TCC TGA TAA YTG MC		

References for Supplementary methods.

Andersen, G.L., Menkissoglou, O., and Lindow, S.E. (1991). Occurence and properties of copper-tolerant strains of *Pseudomonas syringae* isolated from fruit in California. *Phytopathology* **8**: 648–656.

Araki, H., Tian, D., Goss, E.M., Jakob, K., Halldorsdottir, S.S., Kreitman, M., *et al.* (2006). Presence/absence polymorphism for alternative pathogenicity islands in *Pseudomonas viridiflava*, a pathogen of *Arabidopsis*. *Proc Natl Acad Sci USA* **103**: 5887–5892.

Bender, C.L., Alarcòn-Chaidez, F., and Gross, D.C. (1999). *Pseudomonas syringae* phytotoxins: mode of action, regulation, and biosynthesis by peptide and polyketide synthetases. *Microbiol Mol Biol Rev* **63**: 266–292.

Bultreys, A., and Gheysen, I. (1999). Biological and molecular detection of toxic lipodepsipeptide-producing *Pseudomonas syringae* strains and PCR identification in plants. *Appl Environ Microbiol* **65**: 1904–1909.

Gerhardt, P., Murray, R.G.E., Costilow, R.N., Nester, E.W., Wood, W.A., Krieg, N.R., *et al.* (1981). Manual of methods for general bacteriology, 1st edn. Pretoria, South Africa: Plant Protection Research Institute.

Jakob, K., Kniskern, J.M., and Bergelson, J. (2007). The role of pectate lyase and the jasmonic acid defense response in *Pseudomonas viridiflava* virulence. *Mol Plant Microbe Interact* **20**:146–158.

Morris, C.E., Sands, D.C., Vinatzer, B.A., Glaux, C., Guilbaud, C., Buffière, A., *et al.,* (2008). The life history of the plant pathogen *Pseudomonas syringae* is linked to the water cycle. *Int Soc Microb Ecol* **2**: 321–34.

Morris, C.E., Sands, D.C., Vanneste, J.L., Montarry, J., Oakley, B., Guilbaud, C., *et al.* (2010). Inferring the evolutionary history of the plant pathogen *Pseudomonas syringae* from its biogeography in headwaters of rivers in North America, Europe and New Zealand. *mBio* **1**: 00107–10.

Parkinson, N., Bryant, R., Bew, J., and Elphinstone, J. (2011). Rapid phylogenetic identification of members of the *Pseudomonas syringae* species complex using the *rpoD* locus. *Plant Pathol* **60**: 338–344.

Sarkar, S.F., and Guttman, D.S. (2004). Evolution of the core genome of *Pseudomonas syringae* , a highly clonal , endemic plant pathogen. *Appl Environ Microbiol* **70**: 1999–2012.

Shaad, N.W., Jones, J. B., and Chun, W. (1980). Laboratory guide for the identification of plant pathogenic bacteria, 2nd edn. St. Paul, Minnesota. The American Phytopathological Society.

Xia, X. (2013). DAMBE5: A comprehensive software package for data analysis in molecular biology and evolution. Mol Biol Evol doi:10.1093/molbev/mst064.