



HAL
open science

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.

► To cite this version:

Claudia Bartoli, Odile Berge, Caroline Monteil, Caroline Guilbaud, Giorgio M. Balestra, et al.. The *Pseudomonas viridiflava* phylogroups in the *P. syringae* species complex are characterized by genetic variability and phenotypic plasticity of pathogenicity-related traits.. *Environmental Microbiology*, 2014, 16 (7), pp.2301-2315. <10.1111/1462-2920.12433>. <hal-02638544>

HAL Id: hal-02638544

<https://hal.inrae.fr/hal-02638544v1>

Submitted on 28 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

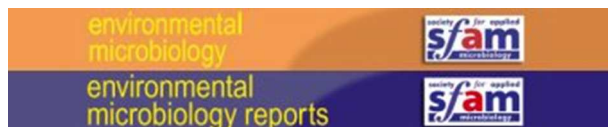
L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Postprint

Version définitive du manuscrit publié dans / Final version of the manuscript published in *Environmental microbiology*, DOI : 10.1111/1462-2920.12433. The original publication is available at Wiley Online library.



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

Journal:	<i>Environmental Microbiology and Environmental Microbiology Reports</i>
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

SCHOLARONE™
Manuscripts

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

3 Claudia Bartoli^{1,2}, Odile Berge², Caroline L. Monteil², Caroline Guilbaud², Giorgio M. Balestra¹,
4 Leonardo Varvaro¹ Corbin Jones³, Jeffery L Dangi⁴, David A. Baltrus⁵, David C. Sands⁶ and
5 Cindy E. Morris^{2,6*}

6

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

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

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

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

14 ⁵School of Plant Sciences, University of Arizona, P.O. Box 210036, Forbes 303, Tucson, Arizona,
15 85721, USA

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

18

19 *For correspondence. E-mail cindy.morris@avignon.inra.fr; (+33) (0) 432722841; Fax (+33) (0)
20 432722841

21

22 Running title: *Pseudomonas viridiflava* diversity

23

24

25 **Summary**

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

41

42

43

44 **Introduction**

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

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

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

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

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

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

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

108 **Results**

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

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

117 groups of strains characterized in this present study (Berge, unpublished data). The delimitation
118 between phylogroups was made by calculating the genomic distances obtained from the
119 concatenated housekeeping genes (*cts*, *gapA*, *gyrB*, *rpoD*) with the Kimura 2-parameter model. A
120 genomic distance <5% was used for delimitation of phylogroups. We used this distance as the
121 maximum value for the delimitation of a phylogroup since it permitted us to obtain the phylogroups
122 already described in the literature (Parkinson *et al.*, 2011). Almost all strains belonged to the well
123 known phylogroup 7 (Parkinson *et al.*, 2011) according to their *cts* sequence, or the sequences of all
124 4 housekeeping genes (Fig. 1). Soft rot was also caused by several strains in the closely-related
125 phylogroup 8 (Fig. 1). These two groups form a monophyletic clade in the *P. syringae* species
126 complex tree, supported by bootstrap values of 83% and 100% (Fig. 1) and were robust to gene or
127 model application (Fig. S1). Following this classification 59 strains from different habitats
128 belonging to phylogroups 7 and 8 were extensively characterized for their phenotypes and
129 genotypes as described below.

130

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

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

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

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

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

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

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

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

192 S-PAI lacks effector genes at this locus. Our analyses of the sequences at this position in T-PAI
193 strains revealed the existence of two open reading frames close to the *hopAI*_(T) gene, encoding for a
194 lipoprotein and for a protein with an ABM domain typically present in monooxygenases involved in
195 the biosynthesis of antibiotics. We also found the lipoprotein and the monooxygenase gene in the
196 draft genome of three strains that corresponded to the S-PAI type but that lack *hopAI*_(T) and *shcA*_(T).
197 This was observed in a genomic analysis of two our *P. viridiflava* strains from non-agricultural
198 habitats (TA0043, Genbank accession AVDV00000000; and CC1582, Genbank accession
199 AVDW00000000) and one other publically available assembly (UASWS0038, Genbank accession
200 number NZ AMQP00000000). We tested the hypothesis that the two open reading frames were
201 widely present in *P. viridiflava* independently of PAIs, by designing specific primers for amplifying
202 both genes. The lipoprotein and the monooxygenase genes were present in almost all the strains
203 analyzed (90%) even in strains lacking *hopAI*_(T) and other T-PAI alleles (Fig. 1, Fig. S2). Further
204 genomic analysis of TA0043 and CC1582 revealed that the lipoprotein and the monooxygenase
205 genes are located in a region resembling an *EEL* but lacking effectors. This locus resembles the
206 *EEL* in chromosomal location, bordered by tRNA_{-leu} and the *queA* at the 5' end (Fig. 3), but it does
207 not have an identifiable *hrpK*. In comparison with the *EEL* of strains PsyB728A and *PtoDC3000* of
208 *P. syringae*, as well as the *EEL* of T-PAI strains, the *P. viridiflava* S-PAI *EEL* lacked the *hrpK*
209 gene, known to delimit the end of the *EEL* and to be a component of the *hrp* PAI with a putative
210 function in translocation (Alfano *et al.*, 2000) and it lacked the *hopAI* and its chaperone *shcA* (Fig.
211 3). These results suggest that recombination events may have occurred between the PAIs or that the
212 S-PAI strains lost part of the *EEL* during their evolution. The phylogenetic analysis of the
213 concatenated lipoprotein and monooxygenase sequences (Fig. 4) showed evidence of horizontal
214 transfer for these loci compared with the housekeeping phylogeny (Fig. 1), suggesting that they
215 were potentially acquired with the associated PAIs (Fig. 4). As observed by Araki *et al.*, (2007),
216 strains having *hopAI* were rare: only seven strains were positive for the *hopAI* gene and for the T-
217 PAI and did not have the alleles typical of the S-PAI when tested with PCR (Fig. S2). Forty-nine

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

221

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

229

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

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

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

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

246

247 Discussion

248 Our data support a portrait of *P. viridiflava* that differs from previous reports both in terms of
249 phenotypes and importance of the T3SS in pathogenicity. Phenotypes of the LOPAT scheme, have
250 been commonly used to differentiate *P. viridiflava* from other members of the *P. syringae* complex,
251 a practice solidified by reports of homogeneity among strains within *P. viridiflava* (Goss *et al.*,
252 2005; Sarris *et al.*, 2012). We demonstrated that these characterization schemes are not completely
253 reliable. Although potato soft rot is a phenotype of Pseudomonads unique to the *P. viridiflava* group
254 (phylogroups 7 and 8) (Berge, unpublished data), 8% (5/60) of the *P. viridiflava* characterized here
255 were not able to degrade potato slices. Hence, this trait is also not diagnostic. The phenotypic
256 properties of *P. viridiflava* are further complicated by the almost universal occurrence of phase
257 variation in this group, affecting the expression of previously diagnostic traits such as potato soft
258 rot.

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

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

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

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

298 Mechanisms that regulate phase variation are generally unknown (Hallet *et al.*, 2001). In
299 pathogens such as *Escherichia coli*, *Haemophilus influenzae*, and *P. aeruginosa* the formation of
300 antibiotic resistance variants is related to a defective mismatch repair system (MMR) (Matic *et al.*,
301 1997, Watson *et al.*, 2004, Ciofu *et al.*, 2010). Mutations in MMR genes lead to a non-efficient
302 DNA repair system leading to mutations in loci that influence gene expression. These mutations can
303 be fixed and the re-acquisition of the original phenotype can occur by further mutations in the same
304 genomic loci. However phase variation can also be the result of epigenetic alteration (Hallet *et al.*,
305 2001). The mechanisms that regulate phase variation in *P. viridiflava* are unknown, though it could
306 provide a useful tool for adapting to different habits and modulating bacterial fitness and survival.

307 The balance between saprophytic and pathogenic modes of *P. viridiflava* has likely also had a
308 role in shaping the nature of its T3SS. As observed for strain Psy642 (Clarke *et al.*, 2010), *P.*
309 *viridiflava* lacks *hrpK*, encoding a required translocation component of the T3SS (Alfano *et al.*,
310 2000). Furthermore, a minority of strains of *P. viridiflava* contain *hopAI*, which can have a role in
311 enhancing virulence (Alfano *et al.*, 2000). Two EEL-associated genes, potentially encoding a
312 monooxygenase while the other resembles a lipoprotein, could have toxic functions useful both in
313 pathogenicity and in competition. Although the configuration of the *hrc/hrp* cluster in *P. viridiflava*
314 is similar to that observed in the non pathogenic strain Psy642 (Clarke *et al.*, 2010), *P. viridiflava*
315 clearly has pathogenic potential –albeit unpredictable– whereas strains related to Psy642 (Ps.
316 phylogroup 2c) are not pathogenic (Demba Diallo *et al.*, 2012).

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

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

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

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

360 **Experimental procedures**

361 *Isolation and selection of bacterial strains*

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

366 *Biochemical and pathogenicity tests*

367 The objective of the phenotypic analysis was to characterize the variability of *P. viridiflava* strains
368 from different substrates and sites. Strains were tested for the characteristics in the LOPAT scheme
369 and for hydrolysis of gelatin, esculin, arbutin and tween80 as described previously (Lelliott *et al.*,
370 1966). Additional information is presented in Methods S1.

371 *Genetic characterization*

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

377 Characterization of the PAIs is described in Methods S1.

378

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

380

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

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

394

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

401 *Characterization of phase variants*

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

414 *Statistical analyses*

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

418

419 **Acknowledgements**

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

430

431

432 **References**

- 433 Alfano, J.R., Charkowski, A.O., Deng, W.L., Badel, J.L., Petnicki-Ocwieja, T., Van Dijk, K., *et al.*
 434 (2000). The *Pseudomonas syringae* Hrp pathogenicity island has a tripartite mosaic structure
 435 composed of a cluster of type III secretion genes bounded by exchangeable effector and conserved
 436 effector loci that contribute to parasitic fitness and pathogenicity. *Proc Natl Acad Sci USA* **97**: 4856–
 437 4861.
- 438 Araki, H., Innan, H., Kreitman, M., and Bergelson, J. (2007). Molecular evolution of pathogenicity-
 439 island genes in *Pseudomonas viridiflava*. *Genetics* **177**: 1031–1041.
- 440 Araki, H., Tian, D., Goss, E.M., Jakob, K., Halldorsdottir, S.S., Kreitman, M., *et al.* (2006).
 441 Presence/absence polymorphism for alternative pathogenicity islands in *Pseudomonas viridiflava*, a
 442 pathogen of *Arabidopsis*. *Proc Natl Acad Sci USA* **103**: 5887–5892.
- 443 Armstrong, G., and Hearst, J. (1996). Carotenoids 2: Genetics and molecular biology of carotenoid
 444 pigment biosynthesis. *Fed Am Soc Exp Biol J* **10**: 228–237.
- 445 Balestra, G.M., and Varvaro, L. (1997). Epiphytic survival and control of *Pseudomonas viridiflava*
 446 on *Actinidia deliciosa*. *Acta Hort* **444**: 745–749.
- 447 Balestra, G.M., and Varvaro, L. (1998). Seasonal fluctuations in kiwifruit phyllosphere and ice
 448 nucleation activity of *Pseudomonas viridiflava*. *J Plant Pathol* **80**: 151–156.
- 449 Baltrus, D.A., Nishimura, M.T., Romanchuk, A., Chang, J.H., Mukhtar, M.S., Cherkis, K., *et al.*
 450 (2011). Dynamic evolution of pathogenicity revealed by sequencing and comparative genomics of
 451 19 *Pseudomonas syringae* isolates. *PLoS Pathogens* **7**: e1002132.
- 452 Baltrus, D.A., Yourstone, S., Lind, A., Guibaud, C., Sands, D.C., Jones, C.D., *et al.*, (2013). Draft
 453 genomes for a phylogenetically diverse suite of *Pseudomonas syringae* strains from multiple source
 454 populations. *Genome Announc* **2**: 01195-13.
- 455 Billing, E. (1970). *Pseudomonas viridiflava* (Burkholder, 1930; Clara 1934). *J Appl Bacteriol* **33**:
 456 492–500.
- 457 Caruso, P., and Catara, V. (1996). First report of *Pseudomonas viridiflava* leaf spot of red-leaved
 458 chicory. *Plant Dis* **80**: 710.
- 459 Ciofu, O., Mandsberg, L.F., Bjarnsholt, T., Wassermann, T., Hoiby, N. (2010). Genetic adaptation
 460 of *Pseudomonas aeruginosa* during chronic lung infections of patients with cystic fibrosis: strong
 461 and wick mutators with heterogeneous genetic backgrounds emerge in *mucA* and *lasR* mutants.
 462 *Microbiology* **156**: 1108–1119.
- 463 Clarke, C.R., Cai, R., Studholme, D.J., Guttman, D., and Vinatzer, B.A. (2010). *Pseudomonas*
 464 *syringae* strains naturally lacking the classical *P. syringae* *hrp/hrc* locus are common leaf colonizers
 465 equipped with an atypical type III secretion system. *Mol Plant Microbe Interact* **23**: 198–210.

- 466 Conn, K.E., and Gubler, W.D. (1993). Bacteria blight of kiwifruit in California. *Plant Dis* **7**: 228–
467 230.
- 468 Demba Diallo, M., Monteil, C.L., Vinatzer, B.A., Clarke, C.R., Glaux, C., Guilbaud, C., *et al.*
469 (2012). *Pseudomonas syringae* naturally lacking the canonical type III secretion system are
470 ubiquitous in nonagricultural habitats, are phylogenetically diverse and can be pathogenic. *Int Soc*
471 *Microb Ecol* **6**: 1325–1335.
- 472 Deziel, E., Comeau, Y., and Villemur, R. (2001). Initiation of biofilm formation by *Pseudomonas*
473 *aeruginosa* 57RP correlates with emergence of hyperpiliated and highly adherent phenotypic
474 variants deficient in swimming, swarming, and twitching motilities. *J Bacteriol* **183**: 1195–1204.
- 475 Dubnau, D., and Losick, R. (2006). Bistability in bacteria. *Mol Microbiol* **61**: 564–572.
- 476 Everett, K.R., and Henshall, W.R. (1994). Epidemiology and population ecology of kiwifruit
477 blossom blight. *Plant Pathol* **43**: 824–830.
- 478 Fett, W.F., Osman, S.F., and Dunn, M.F. (1989). Characterization of exopolysaccharides produced
479 by plant-associated fluorescent pseudomonads. *Appl Environ Microbiol* **55**: 579–83.
- 480 Gardan, L., Shafik, H., Belouin, S., Broch, R., Grimont, F., and Grimont, P.A. (1999). DNA
481 relatedness among the pathovars of *Pseudomonas syringae* and description of *Pseudomonas tremae*
482 sp. nov. and *Pseudomonas cannabina* sp. nov. (ex Sutic and Dowson 1959). *Int J Syst Bacteriol* **49**:
483 469–478.
- 484 González, A.J., Rodicio, M.R., Carmen, M., Gonza, A.J., and Mendoza, M.C. (2003). Identification
485 of an emergent and atypical *Pseudomonas viridiflava* lineage causing bacteriosis in plants of
486 agronomic importance in a spanish region. *Appl Environ Microbiol* **69**: 2936–2941.
- 487 Goss, E.M., Kreitman, M., and Bergelson, J. (2005). Genetic diversity, recombination and cryptic
488 clades in *Pseudomonas viridiflava* infecting natural populations of *Arabidopsis thaliana*. *Genetics*
489 **169**: 21–35.
- 490 Goss, E.M., Potnis, N., and Jones, J.B. (2013). Grudgingly sharing their secrets: new insight into
491 the evolution of plant pathogenic bacteria. *New Phytol* **199**: 630–632.
- 492 Goumans, D.E., and Chatzaki, A.K. (1998). Characterization and host range evaluation of
493 *Pseudomonas viridiflava* from melon , blite , tomato , chrysanthemum and eggplant. *Eur J Plant*
494 *Pathol* **104**: 181–188.
- 495 Green, S., Studholme, D.J., Laue, B.E., Dorati, F., Lovell, H., Arnold, D., *et al.*, (2010).
496 Comparative genome analysis provides insights into the evolution and adaptation of *Pseudomonas*
497 *syringae* pv. *aesculi* on *Aesculus hippocastanum*. *PloS One* **5**: e10224.
- 498 Hallet, B., (2001). Playing Dr Jekyll and Mr Hyde: combined mechanisms of phase variation in
499 bacteria. *Curr Opin Microbiol* **4**: 570–581.
- 500 Hirano, S.S., and Upper, C.D. (2000). Bacteria in the leaf ecosystem with emphasis on
501 *Pseudomonas syringae*-a pathogen, ice nucleus, and epiphyte. *Microbiol Mol Biol Rev* **64**: 624–653.

- 502 Hogan, C.S., Mole, B.M., Grant, S.R., Willis, D.K., and Charkowski, A.O. (2013). The type III
503 secreted effector DspE is required early in *Solanum tuberosum* Leaf infection by *Pectobacterium*
504 *carotovorum* to cause cell death, and requires Wx(3–6)D/E motifs. *PLoS One* **8**: e65534.
- 505 Jackson, R.W., Athanassopoulos, E., Tsiamis, G., Mansfield, J.W., Sesma, A., Arnold, D.L., *et al.*
506 (1999). Identification of a pathogenicity island, which contains genes for virulence and avirulence,
507 on a large native plasmid in the bean pathogen *Pseudomonas syringae* pathovar *phaseolicola*. *Proc*
508 *Natl Acad Sci USA* **96**: 10875–10880.
- 509 Jakob, K., Goss, E.M., Araki, H., Van, T., Kreitman, M., and Bergelson, J. (2002). *Pseudomonas*
510 *viridiflava* and *P. syringae*-natural pathogens of *Arabidopsis thaliana*. *Mol Plant Microbe Interact*
511 **15**: 1195–1203.
- 512 Jakob, K., Kniskern, J.M., and Bergelson, J. (2007). The role of pectate lyase and the jasmonic acid
513 defense response in *Pseudomonas viridiflava* virulence. *Mol Plant Microbe Interact* **20**:146–158.
- 514 Keith, R.C., Keith, L.M., Hernández-Guzmán, G., Uppalapati, S.R., and Bender, C.L. (2003).
515 Alginate gene expression by *Pseudomonas syringae* pv. *tomato* DC3000 in host and non-host
516 plants. *Microbiology* **149**: 1127–1138.
- 517 Lelliott, R.A., Billing, E., and Hayward, A.C. (1966). A determinative scheme for the fluorescent
518 plant pathogenic pseudomonads. *J. Appl. Bacteriol* **29**:470–489.
- 519 Liao, C-H., Hung, H-Y., and Chatterjee, A.K. (1988). An extracellular pectate lyase in the
520 pathogenicity factor of the soft-rotting bacterium *Pseudomonas viridiflava*. *Mol Plant Microbe*
521 *Interact* **1**: 199–206.
- 522 Lindeberg, M., Stavrinides, J., Chang, J.H., Alfano, J.R., Collmer, A., Dangl, J.L., *et al.* (2005).
523 Proposed guidelines for a unified nomenclature and phylogenetic analysis of type III Hop effector
524 proteins in the plant pathogen *Pseudomonas syringae*. *Mol Plant Microbe Interact* **18**: 275–282.
- 525 Matic, I., Radman, M., Taddei, F., Picar, B., Doit, C., Bingen, E., *et al.*(1997). Highly variable
526 mutation rates in commensal and pathogenic *Escherichia coli*. *Science* **227**: 1833–1834.
- 527 Mohr, T.J., Liu, H., Yan, S., Morris, C.E., Castillo, J.A., Jelenska, J., *et al.* (2008). Naturally
528 occurring nonpathogenic isolates of the plant pathogen *Pseudomonas syringae* lack a type III
529 secretion system and effector gene orthologues. *J Bacteriol* **190**: 2858–2870.
- 530 Monteil, C.L. (2011). Ecologie de *Pseudomonas syringae* dans un bassin versant: vers un modèle de
531 transfert dans habitats naturels aux agro-systèmes. *Phd thesis*.
- 532 Monteil, C. L., Cai, R., Liu, H., Mechan Llontop, M., Leman, S., Studholme, D., *et al* (2013). Non-
533 agricultural reservoir contribute to emergence and evolution of *Pseudomonas syringae* crop
534 pathogens. *New Phytol* **199**: 800–811.
- 535 Morris, C.E., Kinkel, L.L., Xiao, K., Prior, P., and Sands, D.C. (2007). Surprising niche for the
536 plant pathogen *Pseudomonas syringae*. *Infect Genet Evol* **7**: 84–92.
- 537 Morris, C.E., Sands, D.C., Vanneste, J.L., Montarry, J., Oakley, B., Guilbaud, C., *et al.* (2010).
538 Inferring the evolutionary history of the plant pathogen *Pseudomonas syringae* from its

- 539 biogeography in headwaters of rivers in North America, Europe and New Zealand. *mBio* **1**: 00107–
540 10.
- 541 Morris, C.E., Sands, D.C., Vinatzer, B.A., Glaux, C., Guilbaud, C., Buffière, A., *et al.*, (2008). The
542 life history of the plant pathogen *Pseudomonas syringae* is linked to the water cycle. *Int Soc Microb*
543 *Ecol* **2**: 321–34.
- 544 Morris, C.E., Wen, A-M., Xu, X-H., and Di, Y-B. (1991). Ice nucleation-active bacteria on chinese
545 cabbage in Northern China: population dynamics and characteristics and their possible role in
546 storage decay. *Phytopathology* **82**: 739–746.
- 547 Mulet, M., Lalucat, J., and García-Valdés, E. (2010). DNA sequence-based analysis of the
548 *Pseudomonas* species. *Environ Microbiol* **12**: 1513–30.
- 549 O'Brien, H.E., Thakur, S., and Guttman, D.S. (2011). Evolution of plant pathogenesis in
550 *Pseudomonas syringae*: a genomics perspective. *Annu Rev Phytophatol* **49**: 269–289.
- 551 Parkinson, N., Bryant, R., Bew, J., and Elphinstone, J. (2011). Rapid phylogenetic identification of
552 members of the *Pseudomonas syringae* species complex using the *rpoD* locus. *Plant Pathol* **60**:
553 338–344.
- 554 Pendersen, S.S., Hoiby, N., Espersen, F., and Koch, C. (1992). Role of alginate in infection with
555 mucoid *Pseudomonas aeruginosa* in cystic fibrosis. *Thorax* **47**: 6–13.
- 556 Rau, M.H., Hansen, S.K., Johansen, H.K., Thomsen, L.E., Workman, C.T., Nielsen, K.F., *et al.*
557 (2010). Early adaptive developments of *Pseudomonas aeruginosa* after the transition from life in
558 the environment to persistent colonization in the airways of human cystic fibrosis hosts. *Environ*
559 *Microbiol* **12**: 1643–1658.
- 560 Sarkar, S.F., and Guttman, D.S. (2004). Evolution of the core genome of *Pseudomonas syringae* , a
561 highly clonal , endemic plant pathogen. *Appl Environ Microbiol* **70**: 1999–2012.
- 562 Sarris, P.F., Trantas, E.A., Mpalantinaki, E., Ververidis, F., and Goumas, D.E. (2012).
563 *Pseudomonas viridiflava*, a multi host plant pathogen with significant genetic variation at the
564 molecular level. *PloS One* **7**: e36090.
- 565 Selezska, K., Kazmierczack, M., Musken, M., Garbe, J., Shobert, M., Haussler, S. *et al.* (2012).
566 *Pseudomonas aeruginosa* population structure revisited under environmental focus: impact of water
567 quality and phage pressure. *Environ Microbiol* **14**: 1952–1967.
- 568 Véghe, A., Hevesi, M., Némethy, Z., and Palkovics, L. (2012). First report of bacterial leaf spot of
569 basil caused by *Pseudomonas viridiflava* in Hungary. *Plant Dis* **96**: 141.
- 570 Versalovic, J., Koeuth, T., and Lupski, J. (1991). Distribution of repetitive DNA-sequences in
571 eubacteria and application to fingerprinting of bacterial genomes. *Nucleic Acids Res* **19**: 6823–6826.
- 572 Watson, M.E., Burns, J.L., Smith, A.L. (2004). Hypermutable *Haemophilus influenzae* with
573 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

For Peer Review Only

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 <i>et al.</i> , 2010)
BS0001	2008	<i>Acitinidia deliciosa</i>	Italy	This study
BS0002	2008	<i>Acitinidia deliciosa</i>	Italy	This study
BS0004	2008	<i>Acitinidia deliciosa</i>	Italy	This study
BS0005	2008	<i>Acitinidia deliciosa</i>	Italy	This study
CC0657	2004	<i>Primula officinalis</i>	France	(Demba Diallo <i>et al.</i> , 2012)
CC1486	2006	snowfall	France	(Morris <i>et al.</i> , 2008)
CC1492	2006	snowfall	France	(Morris <i>et al.</i> , 2008)
CC1582	2006	epilithic biofilm	France	(Morris <i>et al.</i> , 2010)
CCE0322	2009	stream water	France	(Monteil 2011)
CCE0328	2009	stream water	France	(Monteil 2011)
CCV0172	2009	stream water	France	(Demba Diallo <i>et al.</i> , 2012)
CCV0178	2009	stream water	France	(Demba Diallo <i>et al.</i> , 2012)
CCV0180	2009	stream water	France	(Demba Diallo <i>et al.</i> , 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 <i>et al.</i> , 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 <i>et al.</i> , 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	<i>Brassica pekinensis</i>	Chine	(Monteil 2011)
GAW0092	2011	irrigation water	France	This study
GAW0197	2011	irrigation water	France	This study
GAW0203	2011	irrigation water	France	This study
JT0006	2007	<i>Actinidia deliciosa</i>	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

Bartoli et al

Wiley-Blackwell and Society for Applied Microbiology

p 24

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	<i>Ranunculus acris</i>	Italy	(Zoina <i>et al.</i> , 2004)
PVB-H	2012	<i>Ocimum basilicum</i>	Hungary	(Végh <i>et al.</i> , 2012)
PVCT26.1.1	1994	<i>Cichorium intybus</i>	Italy	(Caruso and Catara 1996)
PVCT26.3.1	1994	<i>Cichorium intybus</i>	Italy	(Caruso and Catara 1996)
SZB0012	2009	epilithic biofilm	France	This study
TA0002	2007	stream water	France	(Morris <i>et al.</i> , 2010)
TA0020	2007	stream water	France	(Morris <i>et al.</i> , 2010)
TA0043	2007	<i>Primula officinalis</i>	France	(Morris <i>et al.</i> , 2010)
UB0259	2006	stream water	France	(Morris <i>et al.</i> , 2010)

585

586

587 **Table 2.** Assays conducted to characterize phase variants. For each stable variant, three different clones were used per
 588 test and the experiments were repeated twice.
 589

Stable ^a variants ^a	Utilisation, as sole carbon sources, of								
	Bean pods ^b	Potato rot ^b	Gelatin hydrolysis ^b	HR on tobacco	L-Valine	D-Tartrate	D-Alanine	Cu Resistance ^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	-	-	-	-	+	+	-	-	+

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

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

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

596

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 seedlings599 ^b Then frequency of cantaloupe seedlings (per 12) showing disease at 7 days after inoculation600 ^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 candidum*.

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

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

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

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

637

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

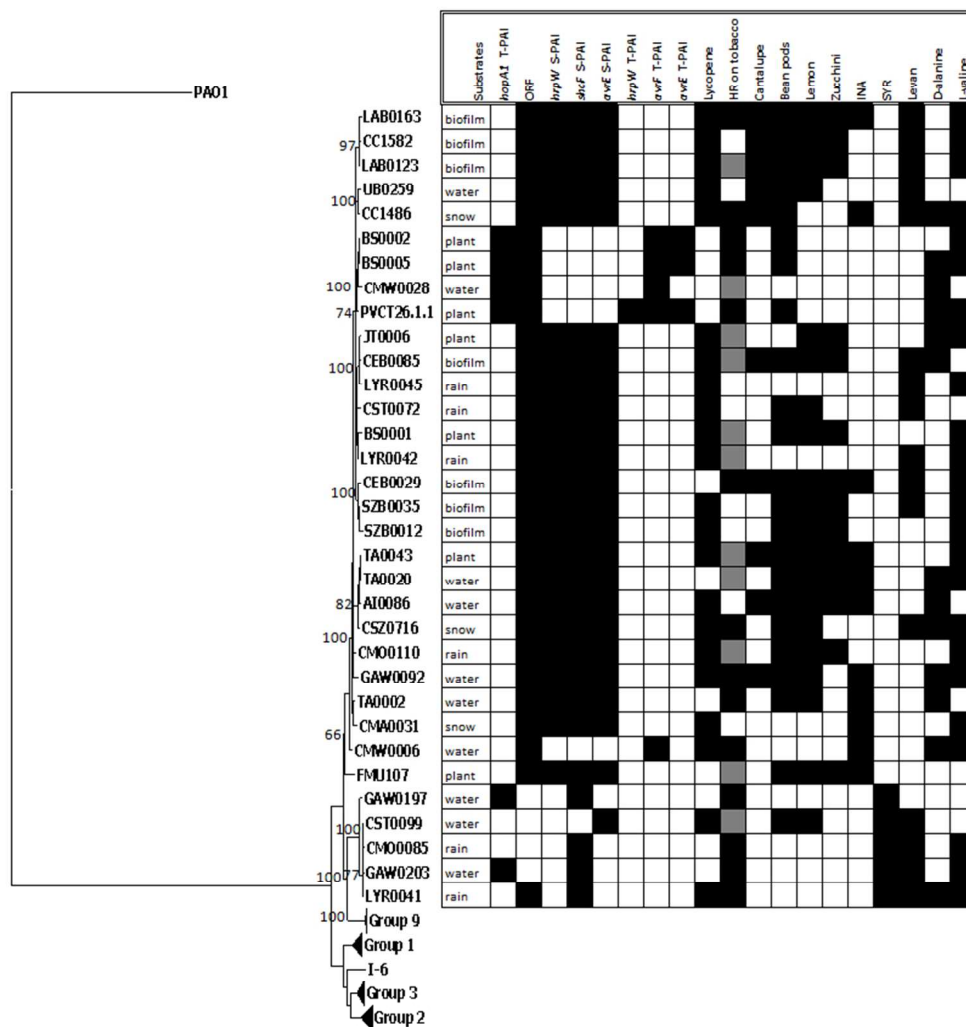
644 of 59 strains. The genotype of the strains for T3SS genes is also shown.

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



02

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)



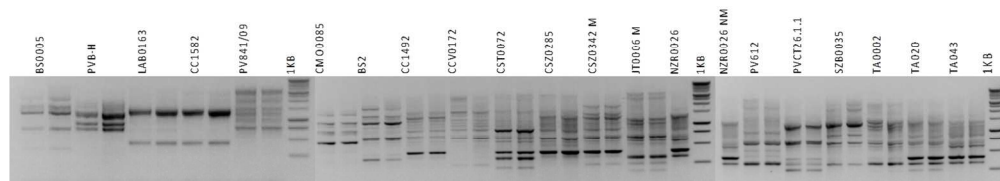


Figure 2. Box PCR profiles of mucoid and non-mucoid variants. The first and second lanes of each strain correspond to the mucoid and non-mucoid variants, respectively
413x190mm (96 x 96 DPI)

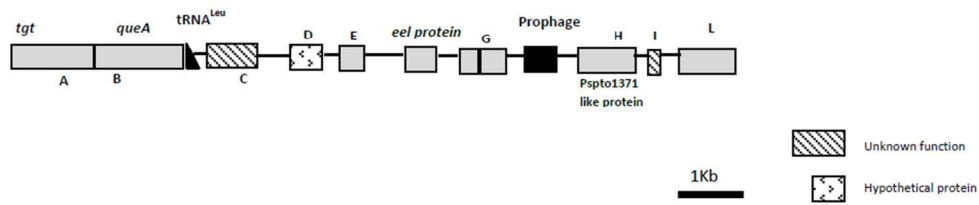


Figure 3. Structure of the locus resembling an EEL of *P. viridiflava* TA0043 and CC1582. The EEL was identified in draft genomes of strains TA0043 and CC1582. Letters refer to the putative protein function: (see manuscript for full legend)
257x65mm (96 x 96 DPI)

Peer Review Only

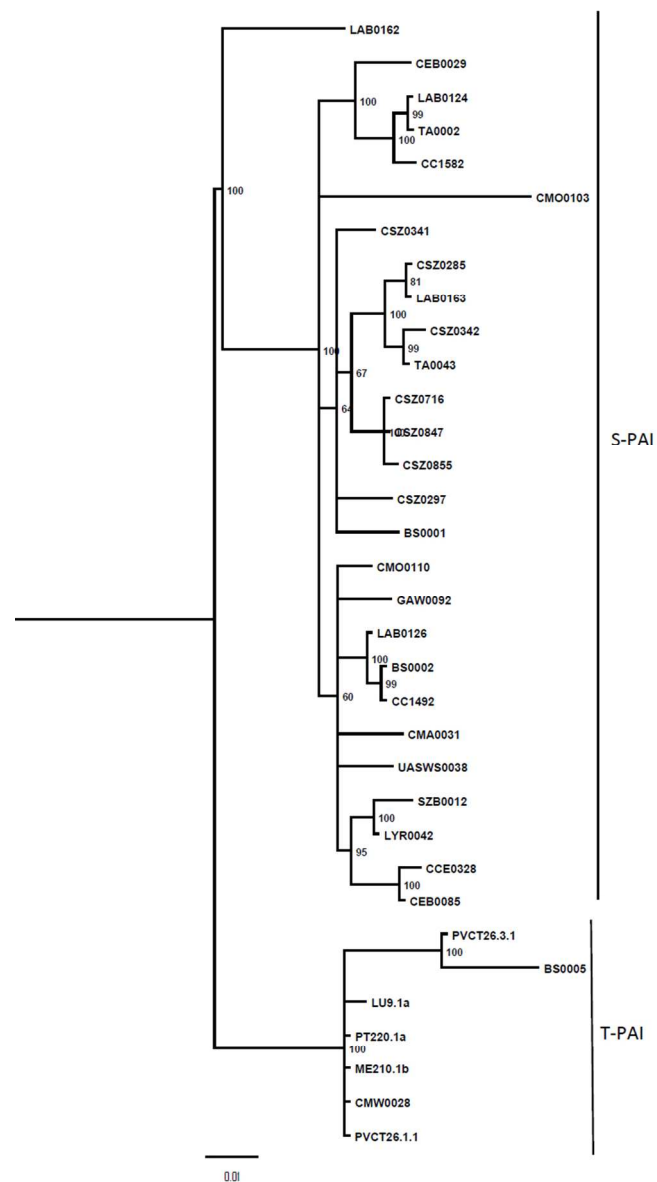


Figure 4. Tree based on the lipoprotein and monoxygenase genes found in the putative EEL of the S-PAI strains and in the EEL of the T-PAI strains. (see manuscript for full legend)
220x375mm (96 x 96 DPI)

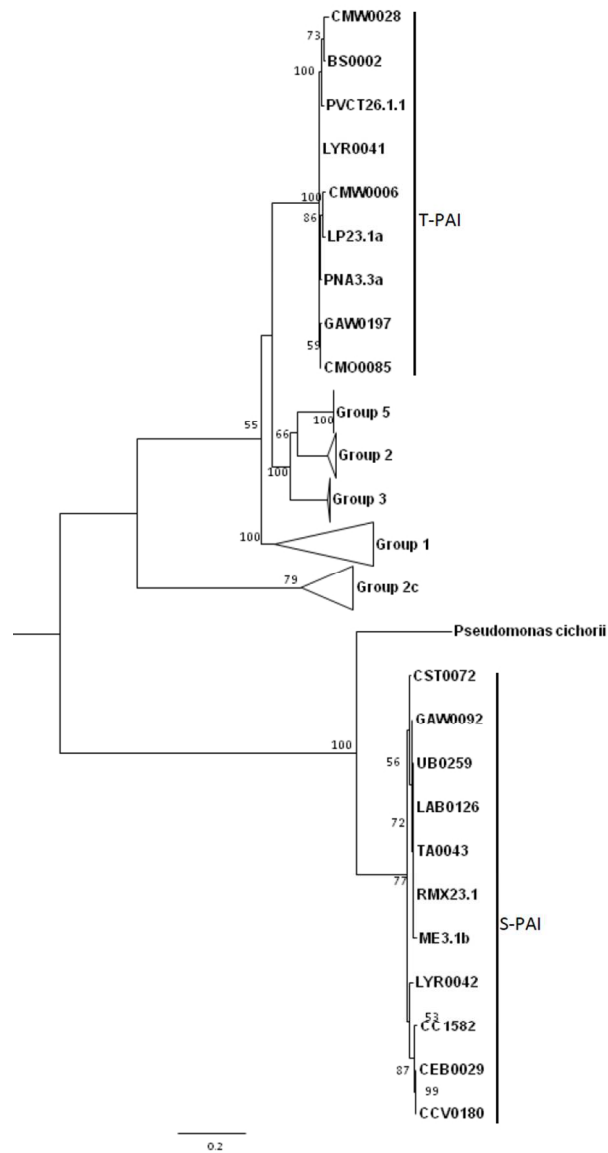
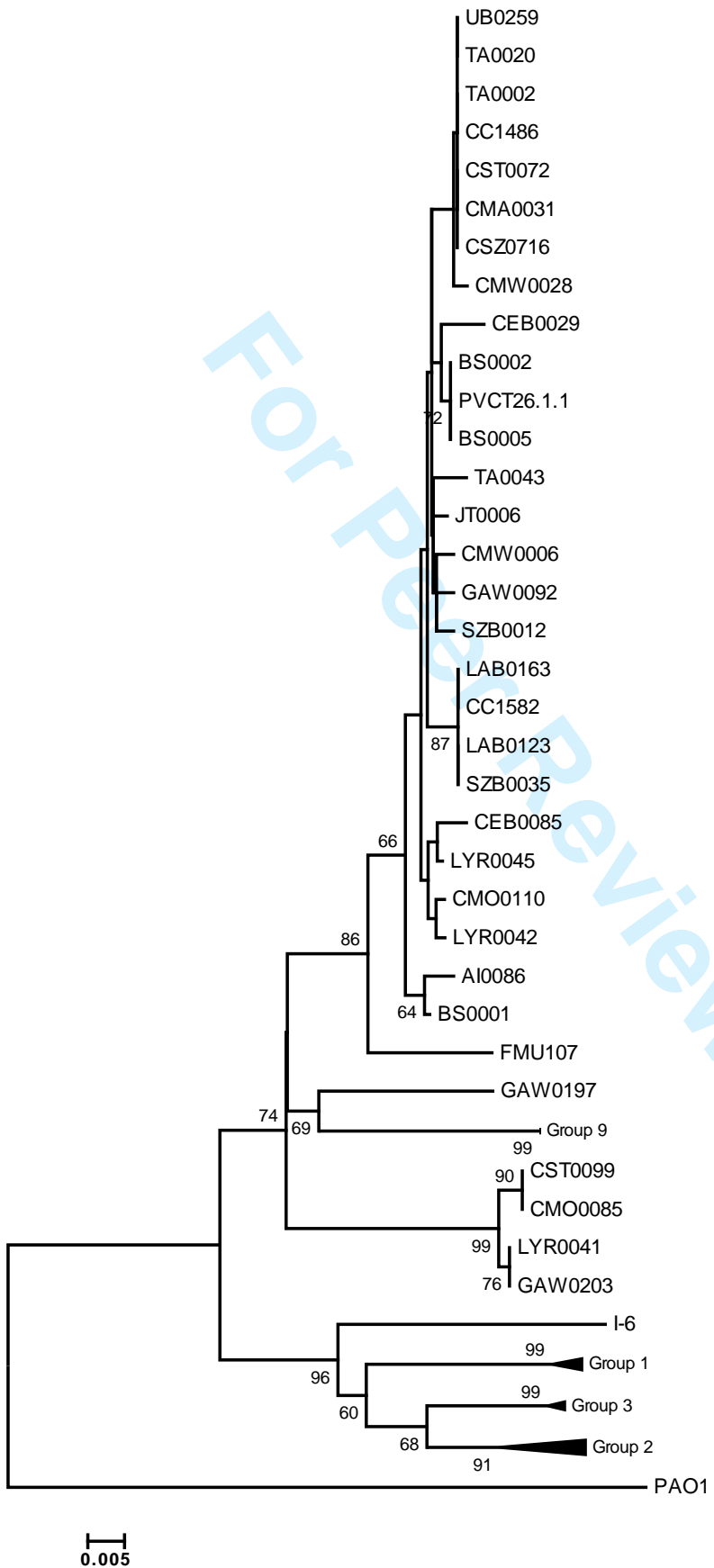
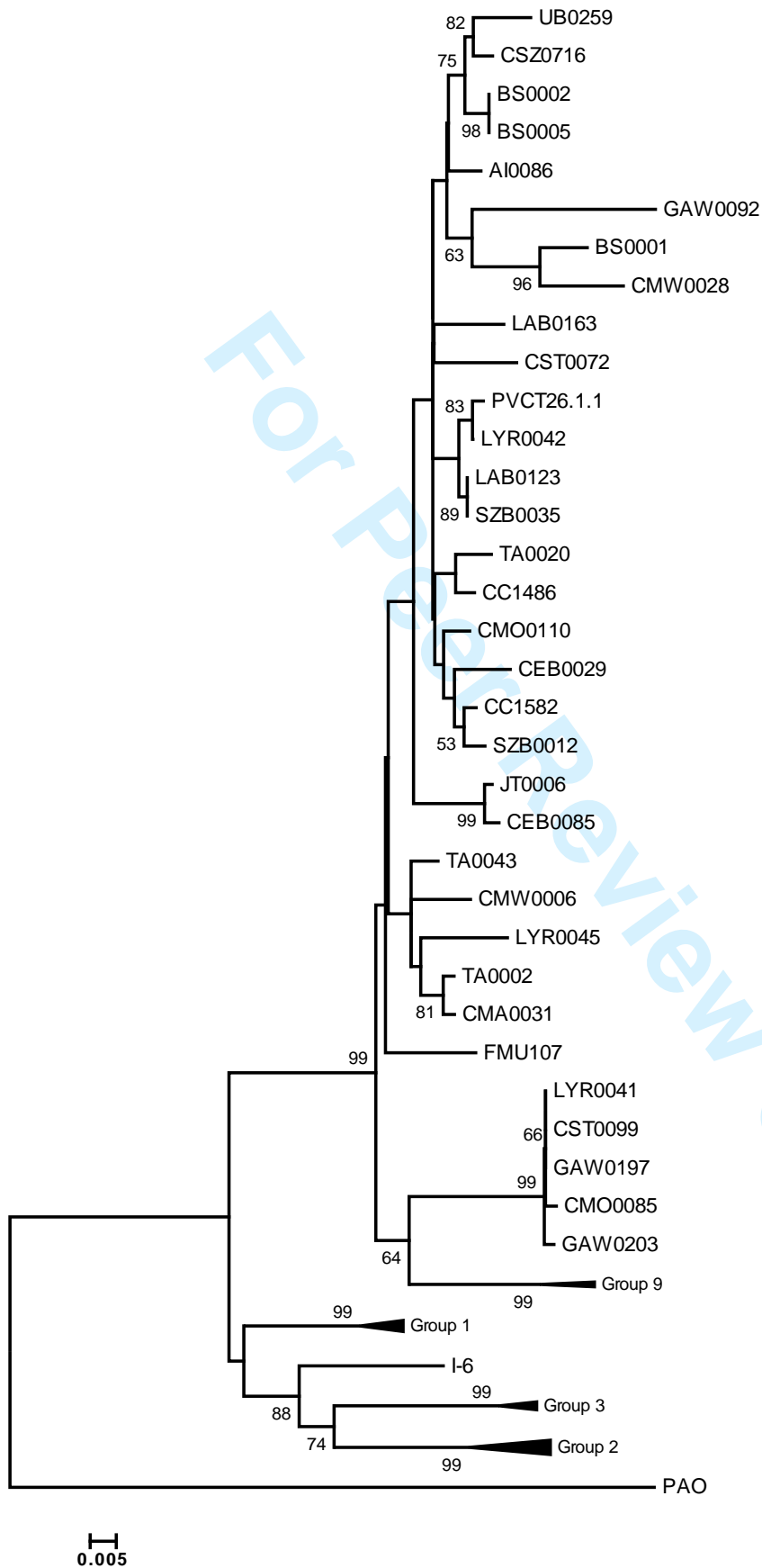


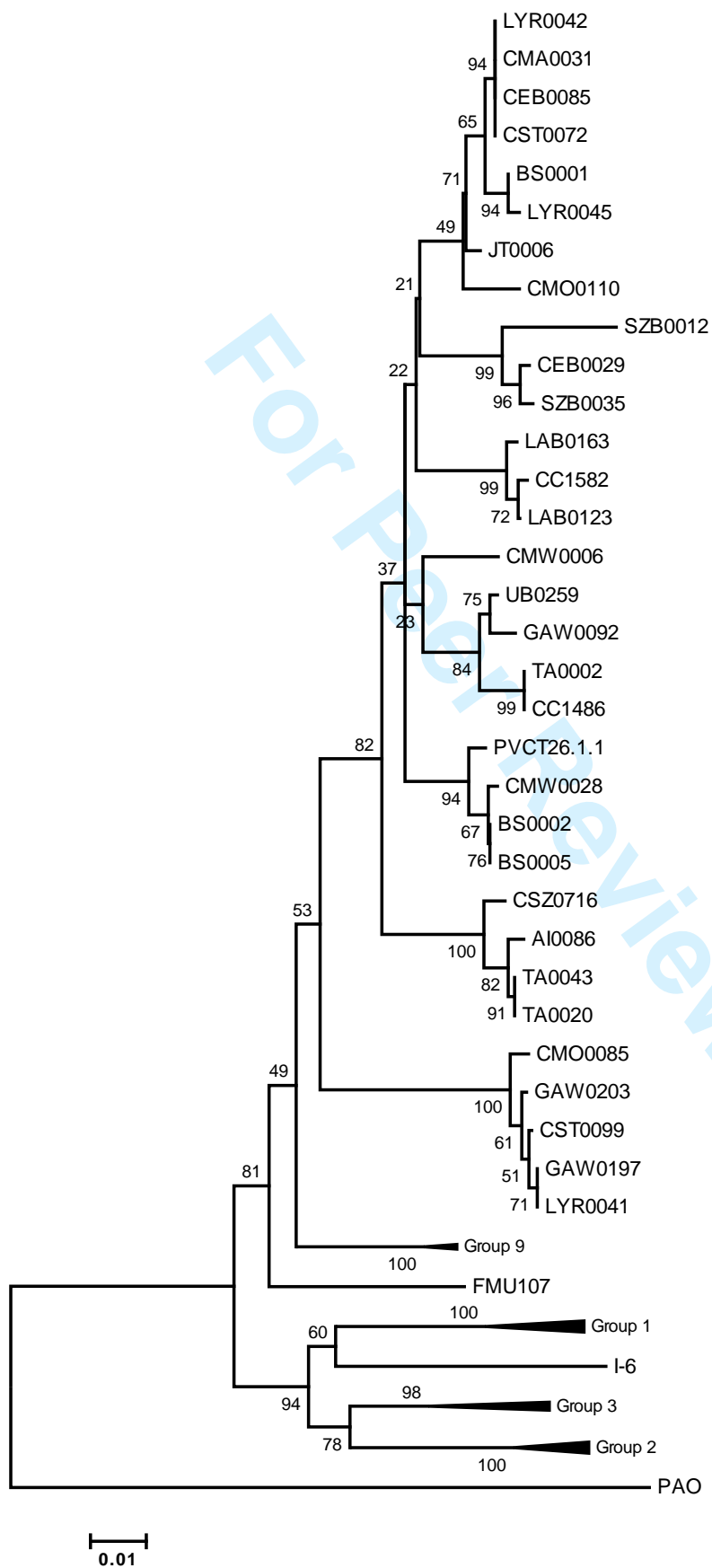
Fig 5. Un-rooted Bayesian phylogenetic tree constructed with the *hrcC* gene sequences. T- and S-PAI strains are delimited with black bars. Sequences for LP23.1a, PNA3.3a, RMX23.1 and M3.1b were obtained from GeneBank. Accession numbers are respectively: AY597277.1, AY597278.1, AY597282.1, AY597281.1.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



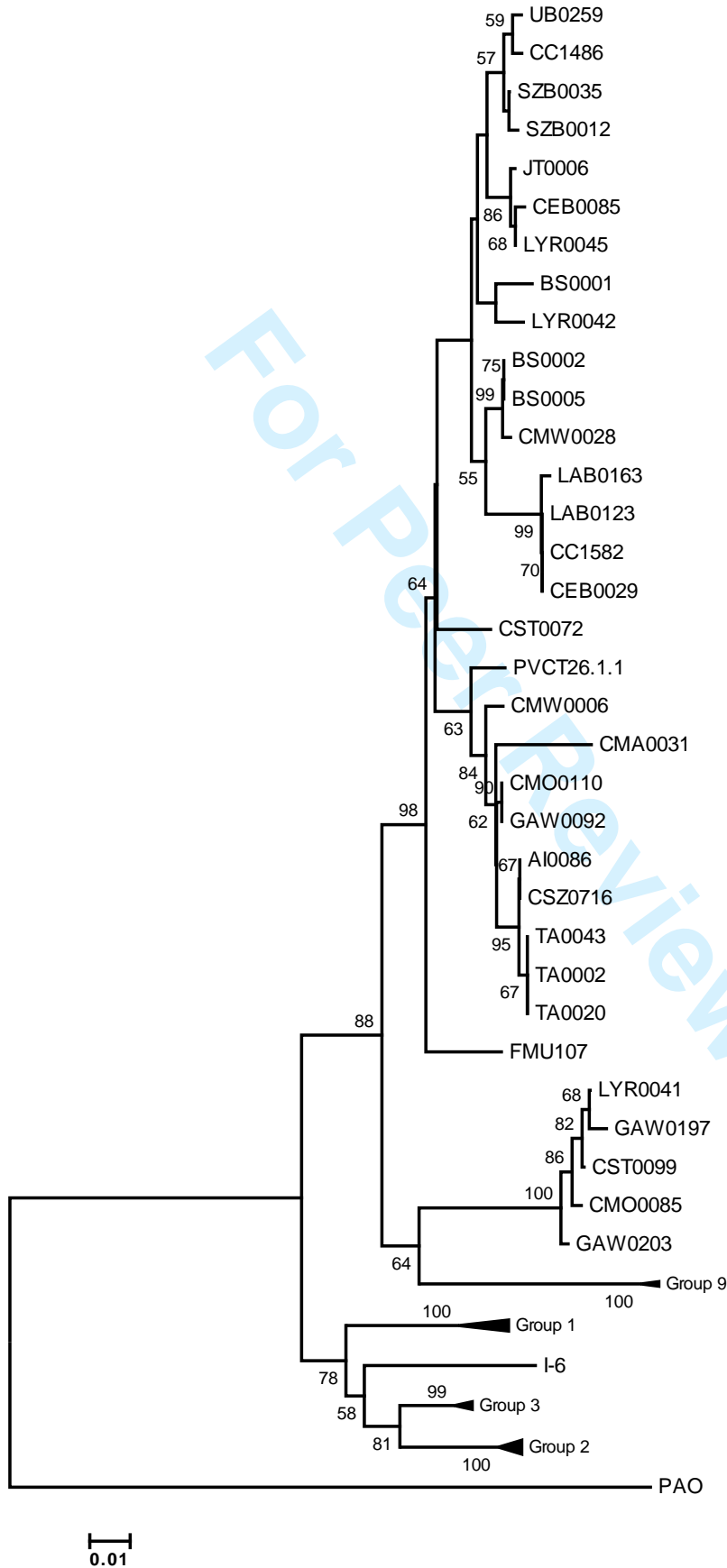
Neighbor joining tree constructed on *cts* gene sequences.



Neighbor joining tree constructed on *gapA* gene sequences

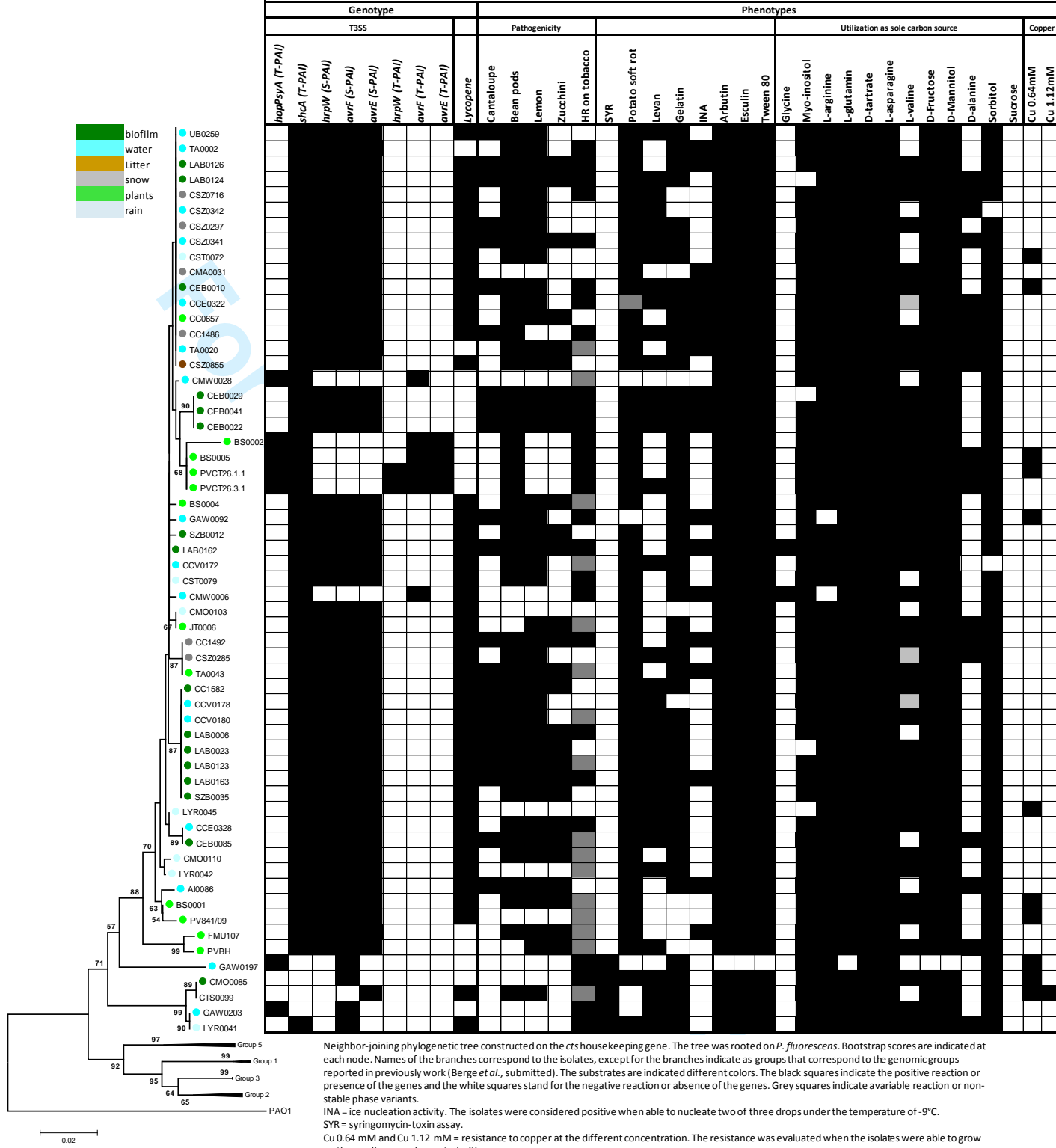


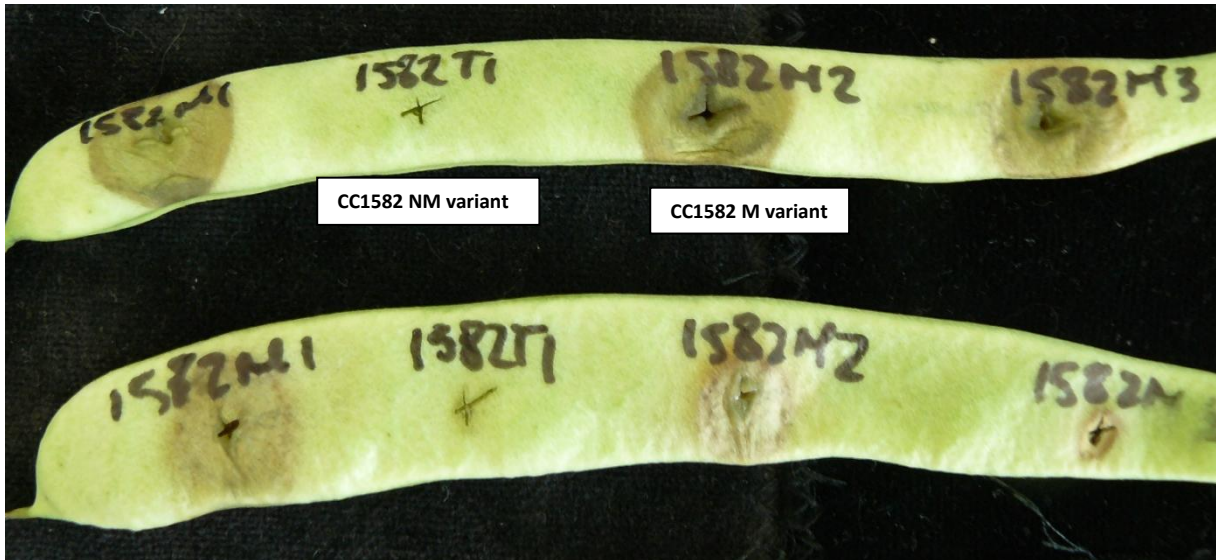
Neighbor joining tree constructed on *gyrB* gene sequences



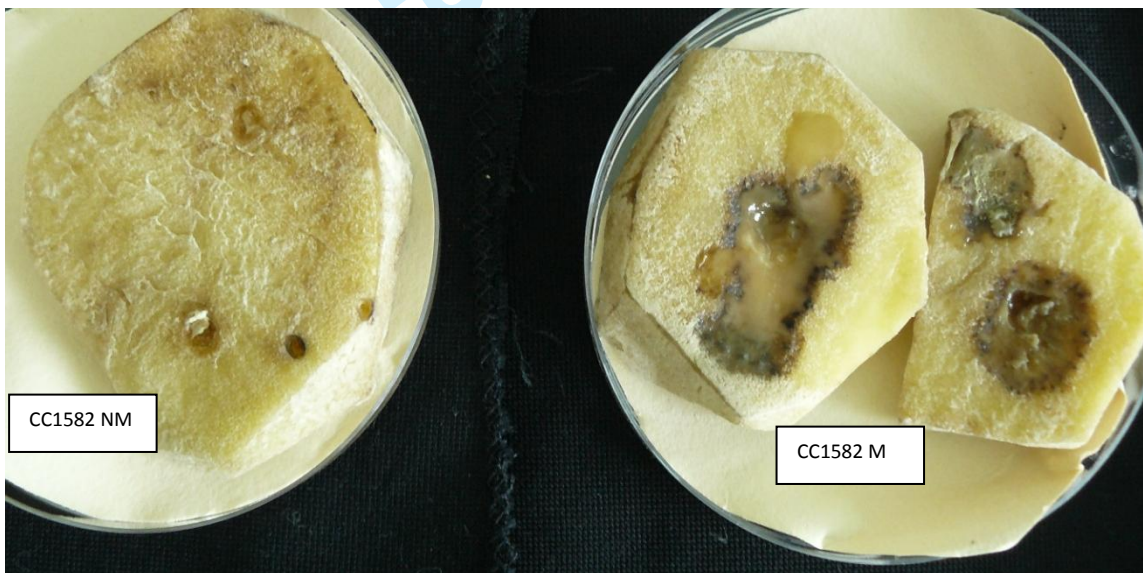
Neighbor joining tree constructed on *rpoD* gene sequences

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

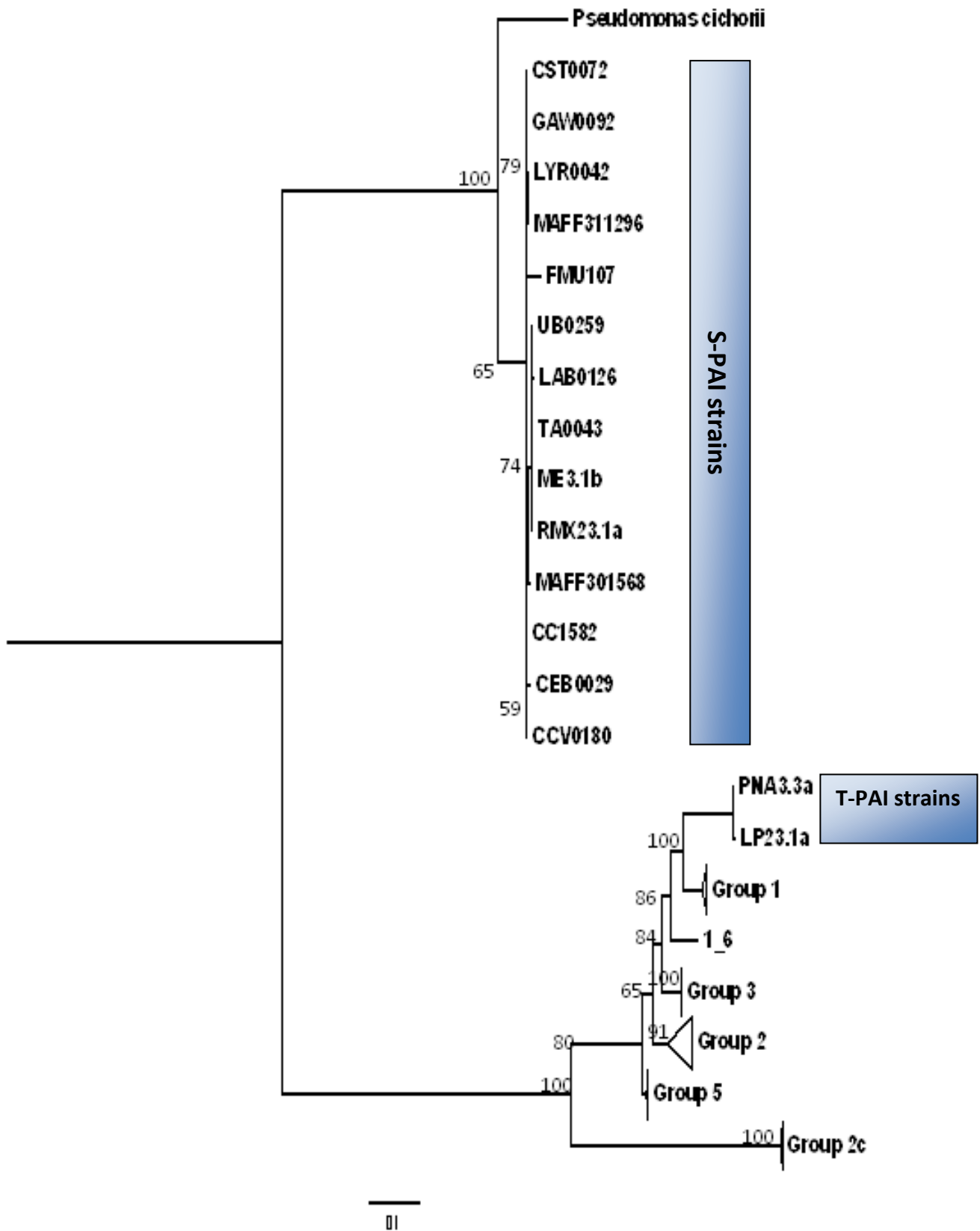




Lesion test on bean pods with mucoid (M) and non-mucoid (NM) CC1582 variant.



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 <i>hrp/hrc</i> protein	Length alignment 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	<i>Pseudomonas syringae</i> pv. <i>actinidiae</i>	kiwifruit	Italy	2008	4
DC3000	1	<i>P. s.</i> pv. <i>tomato</i>	tomato			
NCPPB3487	1	<i>P. s.</i> pv. <i>avellanae</i>	hazelnut	Greek	1976	6
CFBP 1748	3	<i>P. s.</i> pv. <i>papulans</i>	apple tree	USA	1975	2
CFBP 1617	2	<i>P. s.</i> pv. <i>aptata</i>	beet	USA	1959	3
P6	2	<i>P. s.</i> pv. <i>syringae</i>	kiwifruit	Portugal	2006	this study
CFBP 1323	1	<i>P. s.</i> pv. <i>tomato</i>	tomato	France	1971	
CFBP 1676	3	<i>P. s.</i> pv. <i>pisi</i>	pea	New Zealand	1969	7
Pse NE107	3	<i>P. savastanoi</i> pv. <i>savastanoi</i>	olive	Nepal	2007	1
CC0094	2	<i>Pseudomonas</i> sp.	melon	France	1997	5
36b5		<i>Pseudomonas fluorescens</i>				

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 from 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% $\text{NH}_4\text{H}_2\text{PO}_4$, 0.05% K_2HPO_4 , 0.02% $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 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% $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 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 (*Nicotiana 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 *hopAI*_(T) and its chaperone, *shcA*_(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 lycopen 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
<i>hopA1_(T)</i>	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		
<i>hrpW_(S)</i>	HrpW-Fw	TGG AGG TGG AAC ACC TTC	57	480bp
	HrpW-Rv	TGG TCC AGT GGA CGT TAT C		
<i>shcF_(S)</i>	shcF-Fw	CTA AGT GCC ACT CTC GGT A	57	850bp
	shcF-Rv	ATC CTT GGT CTG CCT GTC		
<i>avrE_(S)</i>	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	
<i>hrcC</i>	hrcC-SPAI-Fw	CGC TGA AAC GGC TTT TCT GGA	60 1500bp
	hrcC-SPAI-Rv	TTG CTC ACC CGA TCC CTT TTC	
<i>hrcC</i>	hrcC-Deg-Fw	ABT TYC AGT GGT TYC TBT AYA ACG	59 1300bp
	hrcC-Deg-Rv	GRT CGA GCT GAT CGC CVA YCA	
<i>hrpL</i>	hrpL-SPAI-Fw	AAG GTT GGT ACG TTC GCT GCT CT	62 600bp
	hrpL-SPAI-Rv	GAA CCT CCT TGG AAT ACA CGC TG	
<i>hrpL</i>	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). Occurrence 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., Alarcón-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.

For Peer Review Only