

Tuscia University

**Department of Science and Technology for Agriculture, Forestry, Nature
and Energy (DAFNE)**

XXVI Cycle of Doctorate in Plant Protection

**In collaboration with INRA Institute of Avignon, France, Department of
Plant Pathology**

Thesis:

**Importance of environmental habitats as a reservoir of phytopathogenic
bacteria and their role in the evolution of pathogenicity traits
AGR/12**

By:

Dott. Claudia Bartoli

Coordinator and Supervisor:

Prof. Leonardo Varvaro

Signature

Assistant supervisor

Dott. Cindy E. Morris

Signature

Dott. Giorgio M. Balestra

Signature

**Defended on:
Monday 30 June 2014**

Thesis Committee

Prof. Laura Mugnai

Dr. Elena Di Mattia

Prof. Kubilay Kurtulus Bastas

Referee

Prof. Matthieu Arlat

Prof. Gonçalo Almeida

**Importance of environmental habitats as a reservoir of phytopathogenic
bacteria and their role in the evolution of pathogenicity traits**

Claudia Bartoli

Thesis

Submitted in partial fulfilment of requirements for the degree of Doctor of Philosophy in Plant
Protection at Tuscia University by the authority of the Rector Magnificus

Prof. Alessandro Ruggieri

In the presence of the

Thesis committed appointed by the Doctorate Board

To be defended in public

On Monday 30 June 2014

Ai miei ragazzacci Himal and Jay Ram

*Fidelity is the faithful driver
Microbes depend on good housekeeping
They thrive, survive and flourish
Yet new forms appear
Confounding the scientist
Woe to the status quo*

.

David Sands

Contents

Chapter 1- General Introduction.....	1
Chapter 2 - Methyl-directed mismatch repair: a mechanism underlying emergence of pathogenic bacteria.....	7
Chapter 3 - A user's guide to a data base of the diversity of <i>Pseudomonas syringae</i> and its application to classifying strains in this phylogenetic complex.....	25
Chapter 4 - The <i>Pseudomonas viridiflava</i> phylogroups in the <i>P. syringae</i> species complex are characterized by genetic variability and phenotypic plasticity of pathogenicity-related traits.....	59
Chapter 5 - A framework to gage the epidemic potential of plant pathogens in environmental reservoirs: the example of kiwifruit canker.....	95
Chapter 6 – Hypermotability in <i>Pseudomonas viridiflava</i> : programmed bet-hedging via antibiotic resistance, exopolysaccharide production and pathogenicity.....	144
Chapter 7 - General Conclusions.....	175
Chapter 8 - English Summary.....	181
Italian Summary.....	183
Acknowledgments.....	185

Chapter 1

General Introduction

General Introduction

Pseudomonas syringae is a ubiquitous phytopathogenic bacterium well adapted both to plants and environmental habitats linked with the water cycle (Hirano and Upper, 2000; Morris et al., 2007, 2008). Since only the beginning of this century, 55 reports of disease outbreaks in 25 countries have been associated with *P. syringae* in 25 woody hosts, for example (Lamichhane et al., 2014). All these new disease are caused by new genomic *P. syringae* lineages that after their occurrence rapidly spread worldwide causing plant diseases. Examples of new emerging disease caused by *P. syringae* are the bacterial canker of kiwifruit and the bleeding canker of horse chestnut for what concern the woody plants. The kiwifruit bacterial canker is caused by what is currently called *P. syringae* pv. *actinidiae* (Psa). It was demonstrated that the current kiwifruit epidemic is caused by a clonal Psa population which occurred in all countries where kiwifruit is intensively grown (Chapman et al., 2012; McCann et al., 2013; Vanneste et al., 2013). Likewise, horse chestnut bleeding canker was reported to be caused by a clonal population of *P. syringae* pv. *aesculi* (Green et al., 2010). By contrast, there are examples of *P. syringae* diseases whereby more than one genetic lineage is associated. An example is the bacterial canker of hazelnut caused by two distantly related lineages of *P. syringae* pv. *avellanae* (O'Brien et al., 2012). The authors reported that these lineages convergent the virulence phenotype but dramatically differ in terms of Type Three Secretion System effectors. It is likely that in *P. syringae* more than one evolutionary strategies exist whereby both clonal and convergent populations are able to colonize and cause diseases on plants.

The *P. syringae* genomic lineages responsible of the several diseases on both herbaceous and woody plants, that in the past were attributed to the pathovar name, are all grouped in the *P. syringae* complex. The terminology complex emerged because of the wide diversity that composes *P. syringae*. In particular, genomic analysis based on housekeeping genes of strains isolated from diseased crops demonstrated the existence of 7 phylogenetic groups (phylogroups) which comprise the so-called pathovars (Parkinson et al., 2011). Further studies on the ecology of *P. syringae* in different environmental habitats demonstrated that the diversity estimated in the agricultural context (in particular diseased crops) is only a small fraction of the whole diversity that characterize this bacterial species (Morris et al., 2008) In addition strains from environmental substrates not only form new genomic lineages but they are also characterized by pathogenic strains (Morris et al., 2010).

To date, little is known about the reservoir of phytopathogenic *P. syringae* lineages and the evolutionary mechanisms that drive the emergence of new pathogenic lines. Only recently, it has been demonstrated that *P. syringae* strains isolated from the environment are pathogenic to several plants (Morris et al., 2008, 2010). In addition, a recent study on strains isolated from environmental substrates and pathogenic to tomato, demonstrates that the environment could represent a reservoir of new pathogenic tomato strains (Montail et al., 2013). The authors showed that strains isolated from the environment share some of the effectors with the tomato pathogen, *P. syringae* pv. *tomato*. Because *P. syringae* has been isolated from all habitats related to the water cycle (Morris et al., 2007; 2008) suggests the role of the environment in the evolution of *P. syringae* strains that could represent new emerging pathogens. The frequency of emergence of diseases caused by *P. syringae* and the high diversity of this bacterium, armed with traits for pathogenicity, in the environment at large begs questions about the processes underlying pathogenic diversification. In this light, the objectives of my thesis were: i) to investigate the evolutionary mechanisms (at the molecular level) that can drive the emergence of *P. syringae* pathogenic lineages and, ii) to explore the role of the environment as a reservoir of pathogenic *P. syringae* strains. In particular, I focused on strains isolated from different environmental habitats (rain, snow, fresh water, litter, epilithic biofilms and wild and cultivated plants) that were genetically close to strains isolated from diseased kiwifruit. In **Chapter 2**, I reviewed the molecular mechanisms responsible for the evolution of bacterial pathogens in both human and plant pathogens. From the literature review it became apparent that a defective methyl-directed mismatch repair system (involved in DNA repair) drives the occurrence of horizontal gene transfer, pathoadaptive mutations and prophage integration in bacteria (Meyers and Bull, 2002). These molecular mechanisms are responsible for the evolution of pathogenicity determinants in bacteria. In **Chapter 3**, I present the results of a thorough study on the diversity (both at phenotypic and genetic level) of the *P. syringae* complex, based on a collection of 7000 strains of *P. syringae*. I participated in this research that was initiated in the laboratory at INRA well before my arrival. The collection was established by the MISTRAL team of INRA at Avignon, and includes strains isolated both from agricultural and non-agricultural habitats. This study revealed that the diversity of the *P. syringae* species complex was underestimated by only considering strains isolated from crops. In light of this, we asked how this diversity could influence the evolution of pathogenic strains.

Chapter 1 - General Introduction

After clarifying the classification of strains representing the whole of the diversity of the *P. syringae* complex, my work focused mainly on the kiwifruit pathogens within several *P. syringae* phylogroups. I first performed an in-depth study on *P. viridiflava* (**Chapter 4**), the causal agent of kiwifruit blossom blight (Balestra et al., 2008), which was previously reported to belong to phylogroup 7. However, here we showed that *P. viridiflava* is composed of two different phylogroups. In addition, two different types of Type Three Secretion System are present in *P. viridiflava* but they do not apparently affect the ability of strains to induce disease. Finally, we observed and characterized phase variation in *P. viridiflava* which influences phenotypes related to pathogenicity. The molecular basis of this phase variation is addressed in a subsequent chapter.

In **Chapter 5**, I present work on the investigation of the capacity of *P. syringae* strains from environmental habitats to cause canker of kiwifruit. To date, three different kiwifruit outbreaks associated with different genetic lineages have occurred. Work has been conducted to examine the origin of some of these different outbreaks, but we wondered about the potential for new future outbreaks. Here, we tested the pathogenicity of environmental strains to cause disease to kiwifruit. The first step in addressing the question was to choose candidate strains. This was accomplished by identifying a marker linked to the pathogenicity of *P. syringae* to woody plant species. We found several strains in the collection with this marker and characterized their pathogenic potential for kiwifruit as well as for other woody plants and herbaceous plants as well. In addition to being pathogenic on a range of plant species, environmental strains could co-exist endophytically with epidemic strains thereby suggesting that they could be in favorable conditions to acquire genes for virulence via horizontal gene transfer. This allowed us to gauge the potential for emergence of environmental strains and to contribute to the perspective of strategies for surveying for pre-emergent strains before they cause major epidemics.

Finally in **Chapter 6**, I present work on the role in pathogenicity of hypermutable variants produced by *P. viridiflava* strains. The formation of different phase variants in *P. viridiflava* seems to be a strategy for the regulation of pathogenicity and antibiotic resistance. In human pathogens, however, it was reported that hypermutable variants were due to the mutation of genes related to methyl-directed mismatch repair system. Surprisingly, in *P. viridiflava* the methyl-directed mismatch repair system does not seem to be involved in phase variation. This

Chapter 1 - General Introduction

opens questions about the regulation of this phenomenon and if there are molecular mechanisms that could be targeted as a means of disease control.

References

- Balestra, G. M., Mazzaglia, A., Quattrucci, A., Renzi, M., & Rossetti, A. (2009). Current status of bacterial canker spread on kiwifruit in Italy. *Australasian Plant Disease Notes*, 4, 34–36.
- Balestra, G. M., Mazzaglia, A., & Rossetti, A. (2008). Outbreak of bacterial blossom blight caused by *Pseudomonas viridiflava* on *Actinidia chinensis* kiwifruit plants in Italy. *Plant Disease*, 92, 707.
- Chapman, J. R., Taylor, R. K., Weir, B. S., Romberg, M. K., Vanneste, J. L., Luck, J., & Alexander, B. J. R. (2012). Phylogenetic relationships among global populations of *Pseudomonas syringae* pv. *actinidiae*. *Phytopathology*, 102, 1034–1044.
- Green, S., Studholme, D. J., Laue, B. E., Dorati, F., Lovell, H., Arnold, D., Kamoun, S. (2010). Comparative genome analysis provides insights into the evolution and adaptation of *Pseudomonas syringae* pv. *aesculi* on *Aesculus hippocastanum*. *PloS One*, 5, e10224.
- Hirano, S. S., & 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.
- Koh, Y. J., & Nou, I. S. (2002). DNA markers for identification of *Pseudomonas syringae* pv. *actinidiae*. *Mol Cells*, 13, 309–314.
- Lamichhane, J. R., Varvaro, L., Parisi, L., Audergon, J.-M., & Morris, C. E. (2014). Disease and frost damage of woody plants caused by *Pseudomonas syringae*: seeing the forest for the trees. In *Adv Agron*, 126, 231-291.
- McCann, H. C., Rikkerink, E. H. A., Bertels, F., Fiers, M., & Lu, A. et al. (2013). Genomic analysis of the kiwifruit pathogen *Pseudomonas syringae* pv. *actinidiae* provides insight into the origins of an emergent plant disease. *PLoS Pathogens*, 9, e100.

Chapter 1 - General Introduction

- Meyers, L. A., & Bull, J. J. (2002). Fighting change with change: adaptive variation in an uncertain world. *Trends Ecol Evol*. doi:10.1016/S0169-5347(02)02633-2
- Monteil, C. L., Cai, R., Liu, H., Mechan Llonetop, M. E., Leman, S., Studholme, D. J., Vinatzer, B. A. (2013). Nonagricultural reservoirs 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., & Sands, D. C. (2007). Surprising niche for the plant pathogen *Pseudomonas syringae*. *Infect Gen Evol*, 7, 84–92.
- Morris, C. E., Sands, D. C., Vinatzer, B. A., Glaux, C., Guilbaud, C., Buffière, A., Thompson, B. M. (2008). The life history of the plant pathogen *Pseudomonas syringae* is linked to the water cycle. *The ISME J*, 2, 321–34.
- Morris, C. E., Sands, D. C., Vanneste, J. L., Montarry, J. Oakley, B., Guilbaud, C., Glaux, C (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., & Elphinstone, J. (2011). Rapid phylogenetic identification of members of the *Pseudomonas syringae* species complex using the rpoD locus. *Plant Pathology*, 60, 338–344.
- Scortichini, M. (1994). Occurrence of *Pseudomonas syringe* pv. *actinidiae* on kiwifruit in Italy. *Plant Pathology*, 43, 1035–1038.
- Takikawa, Y., Serizawa Setsuo, I., Takeshi Tsuyumo, S., & Goto, M. (1989). *Pseudomonas syringae* pv. *actinidiae* pv. nov.: the casual bacterium of canker of kiwifruit in japan. *Ann Phytopathol Soc Japan*, 55, 437–444.
- Vanneste, J. L. (2013). Recent progress on detecting , understanding and controlling *Pseudomonas syringae* pv . *actinidiae*: a short review. *New Zeal Plant Prot*, 177, 170–177.

Chapter 1 - General Introduction

Vanneste, J. L., Yu, J., Cornish, D. A., Tanner, D. J., Windner, R., Chapman, J. R., Dowlut, S. (2013). Identification, virulence, and distribution of two biovars of *Pseudomonas syringae* pv. *actinidiae* in New Zealand. *Plant Disease*, 97, 708–719.

Chapter 2

Methyl-directed mismatch repair: a mechanism underlying emergence of pathogenic bacteria

Manuscript in preparation

Methyl-directed mismatch repair: a mechanism underlying emergence of pathogenic bacteria

Claudia Bartoli^{1,2}, Odile Berge², Benoît Moury², Leonardo Varvaro¹ & Cindy E. Morris²

¹Département of Science and Technology for Agriculture, Forestry, Nature and Energy (DAFNE), Tuscia University, 01100 Viterbo, Italy

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

Abstract

The evolution of non-pathogenic bacteria toward a pathogenic form is a process that involves the acquisition, loss or rearrangement of genomic determinants through different molecular mechanisms. A defective methyl-directed mismatch repair system (MMR) leads to increased mutation rate and makes bacteria more competent for horizontal gene transfer, genomic rearrangements and prophage integration. These molecular mechanisms regulate the acquisition of pathogenic determinants leading to emerging diseases. Environmental stresses enhance the number of DNA mismatches and contribute to emergence of new variants including new pathotypes. Understanding the molecular mechanisms responsible of the occurrence of new pathogenic bacteria could lead to better control of disease emergence.

Key words:

Evolution, Horizontal gene transfer, prophage integration, pathogenicity determinants

Emerging pathogens: a threat for animal and plants

Emerging diseases, defined as infections that have newly appeared in a population or have existed but are rapidly increasing in incidence or geographic range [1], are serious threats to the economy of food production systems, medical systems and to human well-being in general. The emergence of new diseases is caused by a combination of ecological or environmental factors and socio-economic factors [2]. For example, several emerging bacterial pathogens come from human-modified environments such as the Shiga-toxin producing *Escherichia coli* O157:H7 that gained access to the food chain because of food production practices. The emergence of *Legionella pneumophila* was facilitated by air-conditioning systems [2]. With the oncoming climate changes, a large number of new infectious diseases have been reported. The occurrence of these diseases rose in correspondence to the climate anomalies that occurred during the 1990s [3]. The latter is especially true for diseases carried by vectors such as mosquitos that are particularly sensitive to meteorological conditions [3]. It was reported that in the last 40 years, 335 new infectious diseases of humans emerged and that the majority of these diseases are caused by multi-drug resistant bacteria [4]. On the other hand, the peak of new infectious diseases caused by bacteria occurred between 1980 and 1990 as a result of the influence of HIV/AIDS on susceptibility to infections [4]. New pathotypes also emerge among plant pathogens. For example, the phytopathogenic bacterium *Pseudomonas syringae* has been reported to be responsible for more than 55 new diseases of fruit, ornamental and forest trees in the last 10 years, most caused by different genetic lines of the bacterium [5]. The seeming continuous emergence of new epidemics raises questions about the origin of new genetic lines of pathogens and specifically about the evolutionary processes involved.

Among the five evolutionary forces, two of them (mutation and recombination or sexual reproduction) lead directly to changes in DNA sequences and provide the substance on which gene flow, genetic drift and natural selection act to permit tangible emergence [6]. Concerning evolution of traits underlying pathogenicity, the origin of the variation in DNA has been attributed to point mutations, horizontal gene transfer (HGT), pathoadaptive mutations and integration of bacteriophages into genomes (Table. 1). Here we summarize the role of these different mechanisms in emergence of bacteria causing diseases to plants, animals and humans and we highlight a point in common for the ensemble of these mechanisms – the need for error in the mechanisms of DNA repair. In particular, a defective methyl-directed mismatch repair system (MMR) that itself produces bacterial variability by facilitating the

occurrence of point mutations is also indirectly responsible for HGT, for pathoadaptive mutations and for prophage integrations. All these molecular mechanisms can take place in the same bacterium and drive the evolution of bacterial pathogens (Fig. 1). Finally we also point out how the environmental context can modulate the occurrence of these molecular mechanisms and in particular how it can influence the expression of genes involved in MMR.

The role of defective methyl-directed mismatch repair in hypermutability and pathoadaptive mutation

Methyl-directed mismatch repair is a conserved system, from prokaryotes to eukaryotes, responsible for post-replicative repair by correcting base-base and insertion/deletion mismatches that escaped the proofreading of DNA polymerases [7] (Box 1). When MMR is defective, mutations are incorporated into the DNA of the replicating cell leading to variability. Hypermutability of bacterial cells, that is associated with the emergence of antibiotic resistant lines, is due to a defective MMR that abolishes their capacity to correct the replication mistakes leading to an increase of the mutation rate [8]. Hence, hypermutable bacteria control poorly DNA repair and they can accumulate mutations that could eventually be favourable or unfavourable for the emergence of new allelic forms of genes associated with pathogenicity. In *E. coli*, it has been demonstrated that mutations in over 20 genes can confer the hypermutable phenotype [9]. Alterations in *mutS*, *mutL*, *mutY* and *mutH* genes in the MMR system are known to be responsible for the hypermutable phenotype in natural, clinical and experimental conditions [10,11]. Mutants in *mut* genes are associated with antibiotic resistance in several bacterial species such as *P. aeruginosa* [11,12], *Staphylococcus aureus* [13], *E. coli* [14], *Neisseria meningitidis* [15], *Salmonella enterica* [16], *Haemophilus influenza* [17] and *Streptococcus pneumoniae* [18]. The resulting antibiotic resistant mutants are often less adapted in normal conditions compared to their non-mutant counterparts. But in presence of antibiotics they present a fitness advantage. Furthermore, resistant strains can acquire secondary mutations that help them to compensate the fitness cost even in antibiotic-free environments [19].

Hypermutable strains have been found more frequently in pathogenic bacterial populations than in commensal populations [14]. This observation could be explained by the fact that the hypermutable phenotype might accelerate the evolution of pathogenic strains by the acquisition of pathogenicity factors as well as antibiotic resistance that is reinforced by natural selection in medical environments [20]. However, Matic et al. [14] demonstrated that in 504

natural isolates, the proportion of commensal strains with a defective MMR was almost the same as in the pathogenic strains, suggesting that further study on the mechanisms of hypermutability is needed.

Because of its role in the rapid occurrence of new mutations, a defective MMR is related to what is called pathoadaptive mutation comprising all the molecular mechanisms that lead to the evolution of virulence traits without horizontal gene transfer (HGT). In general, pathoadaptive mutation is a genetic mechanism that enhances bacterial virulence without HGT but by loss or modification of functions of pre-existing pathogenicity genes toward optimal fitness in new hosts [21]. Several authors proved that pathoadaptive mutation is strongly related to a defective MMR system. For example in *E. coli*, down-regulation of *mut* genes occurred during stationary phase and long-term periods of starvation leading to a greater rate of production of mutants under conditions of starvation [22]. In this light, the MMR apparatus could be down-regulated in stressful conditions thereby favouring the bacterium to adapt to a new environment or host.

Pathoadaptive bacterial lines could therefore show a disadvantage in the ancestral habitat since the mutations they carry might be favourable only in the new host. To overcome this, the pathoadapted pathogens could directly pass to a new host with the same niche characteristics as the former host and progressively become an obligate pathogen [21]. Pathoadaptive mutation is important for obligate pathogens in which the main molecular mechanisms to evolve their virulence determinants are mutation, recombination and transpositions that occur within the host during the infection [23].

An example of pathoadaptive mutation via point mutations is illustrated by *P. aeruginosa*. It was proposed that chronic cystic fibrosis is caused by strains of *P. aeruginosa* that over-express genes for production of alginate. The alginate exopolysaccharide allows *P. aeruginosa* to evade the pulmonary clearance mechanisms thereby fostering lung infections. Knockout mutations in *mucA*, a repressor of alginate biosynthesis, leads to the over-expression of the *algU* gene involved in alginate biosynthesis fostering *P. aeruginosa* lung colonization and biofilm formation [24]. Another recent study conducted over 38 years on cystic fibrosis patients demonstrated that *P. aeruginosa* accumulated mutations in genes involved in antibiotic resistance such as *gyrA/B* and *rpoB*, in which the amino acid changes confer resistance against fluoroquinolones and rifampicin [25]. The same authors found mutations in genes involved in synthesis of the cell envelope and gene regulation, and they also showed that the pathoadaptive mutation of the clonal DK2 strain was correlated with a

defective MMR that led to hypermutable strains with a number of mutations proportional to the infection time.

The occurrence of pathoadaptive mutations and other molecular evolutionary mechanisms such as HGT are not mutually exclusive and both mechanisms might occur in the same bacterial cell (Fig. 1). As discussed below, *Y. pestis* evolved from a less pathogenic *Yersinia* species by acquiring different genomic islands, but the pathogenicity of this bacterium was enhanced also by a point-mutation in the *yopA* gene. The knockout mutation of *yopA*, involved in neutrophil adhesion, led to *Y. pestis* being able to successfully evade host phagocytes. When the mutation was restored the pathogenicity of this bacterium was strongly reduced [26].

Another example of how pathoadaptive mutations, coupled to HGT, can drive evolution comes from the plant pathogen *P. syringae*. This bacterium acquired the Type Three Secretion System (T3SS) *hopZ2* and *hopZ3* genes via HGT while the *hopZ1* effector gene, having a cysteine-protease activity, evolved via deletions and insertions into three functional forms (*hopZ1a*, *hopZ1b* and *hopZc*) and two non-functional forms [27]. The evolution of the three *hopZ1* homologous genes was driven by bacterial-host interactions in which the ancient *hopZ1a* form under selective pressure evolved toward a more effective form [27].

Pathoadaptive mutations can also involve deletion of genes or operons that are unfavourable for host colonization and thus for optimal fitness of the pathogen. A clear example is the complete deletion of the *cadA* gene, encoding the lysine decarboxylase enzyme in shiga-toxin producing *E. coli* strains that favoured the adherence of the bacterium to the human intestinal epithelium [28,29].

A defective MMR favours horizontal gene transfer: the rapid acquisition of new pathogenic determinants

In plants, animal and human bacterial pathogens MMR seems to be a major barrier to chromosomal gene transfer among different bacterial genomic lineages or species [30–33]. In fact, HGT is highly inhibited by *mutS* and *mutL* genes in *E. coli* under experimental conditions and those genes appeared to be acquired and lost different times during the evolution of this bacterial species enhancing the acquisition of foreign DNA [34]. One of the more emblematic examples of how the MMR system can dramatically influence genomic exchange between different species was described for *E. coli* and *S. typhimurium*. These bacterial species can not normally exchange DNA fragments, but mutations in *mutS* and *mutL*

genes allow *E. coli* and *S. typhimurium* to incorporate DNA from each other [31]. In this light, we can suggest that the MMR system is responsible for the “species barrier” in bacteria. The role of *mutS* and *recA* genes in HGT was also recently demonstrated in the plant pathogen *Ralstonia solanacearum* [35] in which the exchange of genetic elements has been demonstrated during tomato infection [36]. It is also important to specify that after the acquisition of new genomic elements via HGT, rearrangements of the novel genes by mutation are an important requisite for aiding the integration of the element into the new host genome [37]. This observation highlights that defective MMR could play an important role not only in the first phases of HGT but also after the acquisition of the new genomic determinants.

Recently the metabolic cost of HGT has been assessed [37] leading to the suggestion that only a clear fitness advantage could compensate the metabolic cost linked to HGT. The rapid acquisition of virulence determinants that lead to higher reproduction rates is the main advantage of HGT in bacterial pathogens. Pathogenicity islands (PAIs) are one of the most impressive examples of HGT of virulence traits both in animals and plant pathogens [38]. The presence of cryptic genes, bacteriophage attachment sites and plasmids confirm that PAIs have spread among bacteria via HGT [39]. For example, the *hrp/hrc* cluster, that encodes for the T3SS in the genera *Pseudomonas*, *Erwinia*, *Xanthomonas* and *Ralstonia* showed the existence of two different *hrp/hrc* groups. The discrepancy between the distributions of these two groups in relationship with the phylogeny of the different bacterial species is strong evidence that the *hrp/hrc* cluster has been horizontally acquired [40].

HGT may lead to the formation of new emerging pathogens in a very short period compared with other mechanisms such as point mutations that could require thousands of years before the emergence of a new pathogen [41]. One impressive example of how a commensal bacterium became a pathogen through HGT comes from *Pantoea agglomerans*. The acquisition of a plasmid harbouring a *hrp/hrc* gene cluster, T3SS effectors and a gene cluster encoding for biosynthesis of indole-3-acetic acid, transformed the epiphytic *P. agglomerans* into a host-specific gall-forming pathogen of two different hosts, gypsophila and sugar beet [42]. Recently, it has been demonstrated that strains in the *P. syringae* phylogroup 10 horizontally acquired pathways for the production of syringomicin-like toxins. Given the few Type Three Secretion effectors present in this phylogroup, it is tempting to speculate that the toxin is involved in the capacity of strains to be pathogenic on *Nicotiana benthamiana* [43].

Human bacterial pathogens can also rapidly shift from a weakly to a highly aggressive form by acquiring virulence determinants. A recent example comes from the epidemic methicillin-resistant *Staphylococcus aureus* in which its emergence coincided with the acquisition, via HGT, of an arginine catabolic mobile element harbouring spermidine genes coding for proteins degrading the toxic polyamines produced in human skin. The early acquisition of this genomic element from *S. epidermis* provided an advantage to *S. aureus* during skin colonization, leading to its emergence and spread [44].

HGT can contribute to the evolution of a non-pathogenic environmental strain into a pathogenic one adapted to a given host. This is what occurred within the *Yersinia* genus in which environmental non-pathogenic strains, by acquiring the pCD1 plasmid carrying T3SS genes, became aggressive human pathogens. Subsequently, acquisition of several pathogenic traits led to the divergence of the three human *Yersinia* pathogens (*Y. enterocolitica*, *Y. pseudotuberculosis* and *Y. pestis*) [45]. *Y. pestis* then further bifurcated from *Y. pseudotuberculosis* by acquiring different plasmids and a genomic island carrying pathogenicity determinants that enhanced its virulence [46].

The role of phages in the evolution of bacterial pathogens and their link with MMR

Bacteriophages are viruses that infect almost all bacterial species including human bacterial pathogens [47] and phytopathogenic bacteria [48]. When a bacteriophage invades a host, two different cycles are possible: the lytic cycle or the lysogenic cycle that leads to the prophage state. The prophage state can cause bacteria can acquire new genomic elements (parts of or all of the prophage) some of which can be pathogenicity determinants [49]. A prophage can also provide a benefit to the bacterial host by increasing its fitness. For example, lambda *E. coli* lysogens are more fit than the non-lysogenic strains when grown under low carbon source conditions [50]. Also P1, P2 and Mu lysogen *E. coli* strains reproduce more rapidly than their non-lysogenic counterparts when they are growth in glucose-limited chemostats [51].

The demonstration that *Corynebacterium diphtheriae* contains bacteriophages encoding diphtheria toxin is the first study in which the role of phages in bacterial pathogenicity was demonstrated [52]. After that, several studies focusing on bacteriophages showed their role in driving the evolution of pathogenic determinants in bacteria.

The role of phages in the evolution of bacterial pathogens may involve acquisition of toxins that enhance the aggressiveness of a given bacterium, passage of PAIs or single T3SS genes and passage of genes different from T3SS effector or toxin genes that play an important role

in bacterial host colonization. Examples where phages have been involved in the acquisition of toxins include the bacteriophage CTX Φ carrying the cholera toxin, a potent A-B-type enterotoxin, that has been shown to increase the aggressiveness of the current O1 and O139 *Vibrio cholerae* strains [53] and also the Shiga-like toxin carried by the 933J phage and responsible for aggressiveness of *E. coli* strains causing food borne toxi-infections [54].

The passage of T3SS genes via temperate phage to pathogenic bacteria was reported in *Salmonella typhimurium*, the casual agent of human and animal salmonellosis. *S. typhimurium* required the Pathogenicity Island 1 SopE effector protein for efficient entry into cell hosts. Miold et al; [55] demonstrated that the SopE effector was transferred in the genome of *S. typhimurium* from a P2 temperate phage after lysogenic conversion of the phage.

Lipopolysaccharides (LPS) are known to be important virulence and colonization factors in bacterial pathogens. In *Shigella flexneri* the temperate bacteriophage Sf6 harbors genes that encode for LPS that are indispensable virulence and colonization factors in *S. flexneri* [56]. The *S. flexneri*/Sf6 model is a clear example of how phages can also harbour genes not related to T3SS or toxin production but are fundamental in the conversion of non-pathogenic into highly virulent strains responsible for global epidemics.

Integration of prophages into host DNA could be regulated by a “permissive” MMR system. *Streptococcus pyogenes* controls the expression of the *mutL* gene by excision and re-integration of a prophage in response to growth. During stationary growth, *S. pyogenes* integrates the prophage that blocks the expression of the *mutL* gene leading to a mutator phenotype [57]. We could consider this as a regulatory system that *S. pyogenes* uses for regulation of *mutL* expression, or we could also consider that *mutL* expression is down-regulated every time that the prophage integrates into the genome of its host. This could occur for several bacteria-prophage systems. However, the role of MMR in prophage integration needs to be further investigated.

Environmental stress exacerbates defective MMR

The emergence of new pathogenic variants can result from processes that do not necessarily involve direct interaction with the host. Abiotic and indirect biotic selection pressures can be important factors in the evolution of new pathogenic lines [58]. Furthermore, the environment is a reservoir of pathogens of both humans and plants [58]. However the environment can also enhance the occurrence of the genetic processes underlying changes in genomic traits. For example, reactive oxygen species (ROS) are a by-product of bacterial metabolism and they

are also produced when bacteria are exposed to high oxidative environments. During the early stages of infection, ROS rapidly increase in the host and they contribute to pathogen clearance and to signalling cascades linked to inflammation and immune responses [59]. In humans it is known that several bacteria can also limit the production of ROS in the host thereby increasing the persistence of infections. The latter has been reported for *Francisella tularensis*, *Anaplasma phagocytophilum* and *Chlamydia trachomatis* [59]. Also during the early phase of plant infections, ROS are produced as a defence response against the microbes suggesting the universality of this defence system in both animals and plants [60]. During host infection reactive oxygen species can influence the expression of MMR genes and the conversion to a hypermutable phenotype, as described in *P. aeruginosa* [61]. In general, the efficiency of MMR can decrease under particular environmental conditions as a result of the decrease in the amount of MMR proteins. In *E. coli* for example, a treatment with DNA-damaging agents can saturate the MMR of the bacterium leading to the incapacity to correct mismatches [62]. In addition, high level nutritional conditions, starvation and nutrient limitations can also induce mutagenesis leading to antibiotic resistance. In *E. coli*, stress-induced mutations have led to resistance to fluoroquinolone ofloxacin and β -lactame [63]. HGT is also induced under particular environmental conditions; in *Bacillus subtilis*, activation of genes for cell competence occurs only in high nutritional conditions while in *Pseudomonas stutzeri* it occurs only in low nutritional environments as an adaptive strategy based on acquisition of genomic determinants triggered by poor food supplies [64]. Recently, it has also been demonstrated that marine α -Proteobacteria isolated in a subtropical ocean are rich in genes that favour HGT in marine environments. High salinity seems to be an important requisite for HGT development in this group of bacteria [65] suggesting that marine habitats may represent a reservoir of genomic element exchanges. This could be of fundamental importance for bacterial pathogens, such as *Vibrio cholerae*, that have part of their cycle in sea waters. If the acquired genetic elements confer traits that are dual use factors for adaptations to general survival and for pathogenicity, then environmental selective pressures can inadvertently contribute to positive selection for pathogens [58].

Concluding remarks

The rapid emergence of new bacterial pathogens raises questions about the molecular mechanisms that regulate this emergence and about the reservoirs in which the new emerging pathogens reside. Here we describe how the molecular mechanisms leading to diversification

of bacteria and to new pathotypes all fundamentally depend on errors in DNA repair and hence on the MMR system. The behaviour of the MMR system can be influenced by rather generic environmental conditions. Therefore, mutations, HGT and prophage integration could occur at different stages of bacterial life history in response to a range of habitats (Fig. 1). Knowing these molecular evolutionary mechanisms and the environmental forces that drive them could help in evaluating the probability of occurrence of new pathogenic bacteria. Furthermore, full characterization of the MMR system and other mechanisms directly involved in DNA repair could open new perspectives for mitigating emergence by inhibiting the molecular mechanisms that underlie diversification.

Acknowledgements

We thank the graphic artist Mr. Louis Wallet for drawing the schematic picture in this review. We also thank Prof. David C. Sands, Montana State University, for the fruitful discussions and for reading and editing the manuscript and Prof. David Baltrus, University of Arizona, for providing crucial references.

Table 1. Evolutionary molecular mechanisms involved in the emergence of new bacterial pathogens

Evolutionary mechanism	Brief description	Consequences	Examples	References
Defective mismatch repair system	The bacterium is not able to repair the mistakes linked to DNA replication. This state is called hypermutable.	• Fitness reduction • Antibiotic resistance • Flexibility in HGT and acquisition of foreign DNA • Rearrangement of new genes	<i>Pseudomonas aeruginosa</i> , <i>Staphylococcus aureus</i> , <i>Escherichia coli</i> , <i>Neisseria meningitidis</i> , <i>Salmonella enterica</i> , <i>Haemophilus influenza</i> , <i>Streptococcus pneumoniae</i> , <i>Ralstonia solanacearum</i>	[11,13,15–18,20,35]
Horizontal gene transfer (HGT)	Transfer of genomic elements from a donor organism to a host organism. HGT is usually intra-specific but it can occur also between more distant taxa.	Acquisition of pathogenicity determinants or entire pathogenicity islands	<i>Pseudomonas</i> spp, <i>Ralstonia</i> spp, <i>Erwinia</i> spp, <i>Xanthomonas</i> spp, <i>Pantoea</i> spp, <i>Staphylococcus</i> spp, <i>Salmonella</i> spp and <i>Streptococcus</i> spp	[33,42,44,46,40,43]
Pathoadaptive mutation	A complex of mutations, loss and modification of pre-existing genes that leads to the emergence of a new pathogen without HGT.	• Fitness increase in a new host • Fitness penalty in the ancestral host/habitat • Enhancement of the functionality of the pathogenicity determinants	<i>Pseudomonas aeruginosa</i> , <i>Yersinia pestis</i> , <i>Pseudomonas syringae</i> , <i>Escherichia coli</i> ,	[24–26,29]
Bacteriophage infection	Infection of phage followed by a lysogenic phase that leads to the formation of the prophage state. The prophage is integrated into the genome of the host leading to the acquisition of new genomic elements by the bacterium.	Acquisition of toxins, Type Three Secretion System effectors or other genes related to pathogenicity	<i>Escherichia coli</i> , <i>Corynebacterium diphtheriae</i> , <i>Vibrio cholerae</i> , <i>Salmonella typhimurium</i> , <i>Shigella flexneri</i>	[51,52,56,66,67]

The methyl-directed mismatch repair system

During DNA replication the proofreading exonuclease associated with the DNA polymerase is able to edit the mistakes made by the DNA polymerase and then to make a second attempt at correct synthesis. However mistakes can escape from the proofreading and they can be corrected just after the DNA strand duplication by the methyl-mismatch repair system (MMR). The latter increase the replication fidelity 50-1000-fold [68]. In *Escherichia coli* the MMR system is composed by MutS, MutL and MutH proteins that can recognize the hemi-methylated DNA. In fact the parent DNA strand can be distinguished from the daughter strand because the methylation of newly synthesized DNA does not occur immediately after the replication [68]. When the methyl directed mismatch is functional, MutS and MutL incise the unmethylated strand at a hemi-methylated GATC site (the signal for the mismatch repair proteins) and then the excision of that portion is made by MutH [68]. Then the DNA is repaired and the DNA polymerase III holoenzyme fills the gap created by the excision. Every time that the MMR system is defective because of spontaneous mutations in the genes that codify for the Mut proteins or because of DNA damages caused by external factors, the replication mistakes are undetected leading to mutations and DNA instability that is also favourable for a high level of DNA recombination.

MMR is considered to be a conserved system; in fact, genes encoding homologous of MutS and MutL have been identified in several eukaryotes, but no homologous of the MutH excision protein have been found in different organisms other than bacteria.

In bacteria, DNA instability caused by a defective MMR is correlated with mutations and recombination, while in humans; there is a direct link between mutant MMR genes and tumour cells. The latter is well demonstrated for colorectal cancer in which an increase in mutation rate was observed in the tumor cells [69]. A defective human MMR gives similar phenotypes as observed for bacteria but with some exceptions in cell behaviour. For example, in humans a defective MMR is related to tolerance to DNA damaging agents as it is for bacteria; however, in human cells tolerance to damage is also related to the apoptotic response of the cell. However, the common phenotype for both prokaryotic and eukaryotic cells carrying a defective MMR is the increased rate of mutation.

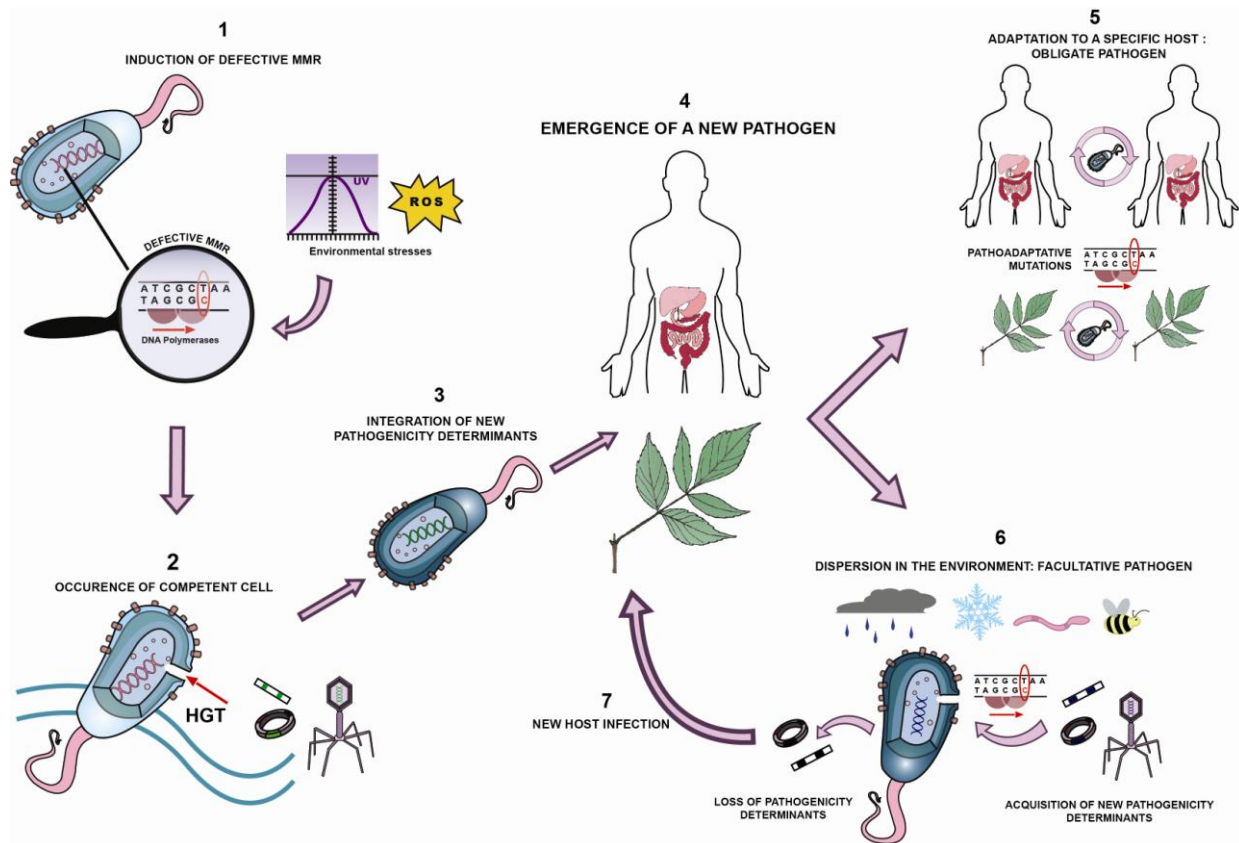


Figure 1. A putative evolutionary scenario regulated by a defective methyl-directed mismatch repair system.

1= A defective methyl-mismatch repair system (MMR), induced by environmental stresses such as reactive oxygen species or UV light, or spontaneously defective, induces mutations in a commensal bacterium, 2= the bacterial cell then become competent and can acquire pathogenicity determinants *via* horizontal gene transfer (HGT) or prophage integration, 3= the bacterial cell carrying the new pathogenicity determinant(s) can infect a new host and become a newly emerging pathogen 4= the new bacterial pathogen can have two different evolutionary destinies: 5= to become an obligate pathogen (an evolutionary dead end) or 6= to maintain a facultative life style, to pass into a new environment and acquire or lose pathogenicity determinants (mechanisms driven by a defective MMR system) and be transformed into a new emerging pathogen or in a more aggressive pathogen that infects the same host species (7). In each step in which the bacterium acquires, rearranges or loses pathogenicity determinants, a defective MMR is the main driver and it can be influenced by environmental stresses. In addition acquisition, or loss of genes related to pathogenicity can be under the selection of environmental pressures.

References

- 1 Morse, S.S. (1995) Factors in the emergence of infectious diseases. *Emerg. Infect. Dis.* 1, 7–15
- 2 Morens, D.M. *et al.* (2004) The challenge of emerging and re-emerging infectious diseases. *Nature* 430, 242–249
- 3 Epstein, P.R. (2001) Climate change and emerging infectious diseases. *Microbes Infect.* 3, 747–754
- 4 Jones, K.E. *et al.* (2008) Global trends in emerging infectious diseases. *Nature* 451, 990–993
- 5 Lamichhane, J.R. *et al.* (2014) Disease and frost damage of woody plants caused by *Pseudomonas syringae*: seeing the forest for the trees. In *Adv Agron* 126, 231–291
- 6 McDonald, B.A. and Linde, C. (2002) Pathogen population genetics, evolutionary potential, and durable resistance. *Annu. Rev. Phytopathol.* 40, 349–379
- 7 Schofield, M.J. and Hsieh, P. (2003) DNA mismatch repair: molecular mechanisms and biological function. *Annu. Rev. Microbiol.* 57, 579–608
- 8 Meyers, L.A. and Bull, J.J. (2002) Fighting change with change: adaptive variation in an uncertain world. *Trends in Ecol Evol*, 17, 551–557
- 9 Horst, J.P. *et al.* (1999) *Escherichia coli* mutator genes. *Trends Microbiol.* 7, 29–36
- 10 Moxon, E.R. *et al.* (1994) Adaptive evolution of highly mutable loci in pathogenic bacteria. *Curr. Biol.* 4, 24–33
- 11 Oliver, A. *et al.* (2000) High frequency of hypermutable *Pseudomonas aeruginosa* in cystic fibrosis lung infection. *Science*. 288, 1251–1253
- 12 Ciofu, O. *et al.* (2010) Genetic adaptation of *Pseudomonas aeruginosa* during chronic lung infection of patients with cystic fibrosis: strong and weak mutators with heterogeneous genetic backgrounds emerge in *mucA* and/or *lasR* mutants. *Microbiology* 156, 1108–1119
- 13 Wang, S. *et al.* (2013) Hypermutable *Staphylococcus aureus* strains present at high frequency in subclinical bovine mastitis isolates are associated with the development of antibiotic resistance. *Vet. Microbiol.* 165, 410–415
- 14 Matic, I. *et al.* (1997) Highly variable mutation rates in commensal and pathogenic *Escherichia coli*. *Science*. 277, 1833–1834
- 15 Richardson, A.R. *et al.* (2002) Mutator clones of *Neisseria meningitidis* in epidemic serogroup A disease. *Proc. Natl. Acad. Sci. U. S. A.* 99, 6103–6107

- 16 LeClerc, J.E. *et al.* (1998) Detection of mutator subpopulations in *Salmonella typhimurium* LT2 by reversion of his alleles. *Mutat. Res.* 400, 89–97
- 17 Watson, M.E. *et al.* (2004) Hypermutable *Haemophilus influenzae* with mutations in mutS are found in cystic fibrosis sputum. *Microbiology* 150, 2947–2958
- 18 Del Campo, R. *et al.* (2005) Population structure, antimicrobial resistance, and mutation frequencies of *Streptococcus pneumoniae* isolates from cystic fibrosis patients. *J. Clin. Microbiol.* 43, 2207–2214
- 19 Lenski, R.E. (1998) Bacterial evolution and the cost of antibiotic resistance. *Int. Microbiol.* 1, 265–270
- 20 Taddei, F. *et al.* (1997) To be a mutator, or how pathogenic and commensal bacteria can evolve rapidly. *Trends Microbiol.* 5, 427–428
- 21 Sokurenko, E. V *et al.* (1999) Pathoadaptive mutations: gene loss and variation in bacterial pathogens. *Trends Microbiol.* 7, 191–195
- 22 Feng, G. *et al.* (1996) Depletion of the cellular amounts of the MutS and MutH methyl-directed mismatch repair proteins in stationary-phase *Escherichia coli* K-12 cells. *J. Bacteriol.* 178, 2388–2396
- 23 Levin, B.R. and Bull, J.J. (1994) Short-sighted evolution and the virulence of pathogenic microorganisms. *Trends Microbiol.* 2, 76–81
- 24 Boucher, J.C. *et al.* (1997) Muroid *Pseudomonas aeruginosa* in cystic fibrosis: characterization of muc mutations in clinical isolates and analysis of clearance in a mouse model of respiratory infection. *Infect. Immun.* 65, 3838–3846
- 25 Marvig, R.L. *et al.* (2013) Genome analysis of a transmissible lineage of *Pseudomonas aeruginosa* reveals pathoadaptive mutations and distinct evolutionary paths of hypermutators. *PLoS Genet.* 9, e1003741
- 26 Rosqvist, R. *et al.* (1988) Increased virulence of *Yersinia pseudotuberculosis* by two independent mutations. *Nature* 334, 522–524
- 27 Ma, W. *et al.* (2006) Type III effector diversification via both pathoadaptation and horizontal transfer in response to a coevolutionary arms race. *PLoS Genet.* 2, e209
- 28 Torres, A.G. *et al.* (2005) Pathoadaptive mutation that mediates adherence of shiga toxin-producing *Escherichia coli* O111. *Infect. Immun.* 73, 4766–4776
- 29 Maurelli, A.T. *et al.* (1998) “Black holes” and bacterial pathogenicity: a large genomic deletion that enhances the virulence of *Shigella* spp. and enteroinvasive *Escherichia coli*. *Proc. Natl. Acad. Sci.* 95, 3943–3948
- 30 Matic, I. *et al.* (1996) Genetic barriers among bacteria. *Trends Microbiol.* 4, 69–72

- 31 Rayssiguier, C. *et al.* (1989) The barrier to recombination between *Escherichia coli* and *Salmonella typhimurium* is disrupted in mismatch-repair mutants. *Nature* 342, 396–401
- 32 Townsend, J.P. *et al.* (2003) Horizontal acquisition of divergent chromosomal DNA in bacteria: effects of mutator phenotypes. *Genetics* 164, 13–21
- 33 Bertolla, F. and Simonet, P. (1999) Horizontal gene transfers in the environment: natural transformation as a putative process for gene transfers between transgenic plants and microorganisms. *Res Microbiol* 150, 375–384
- 34 Denamur, E. *et al.* (2000) Evolutionary implications of the frequent horizontal transfer of mismatch repair genes. *Cell* 103, 711–721
- 35 Fall, S. *et al.* (2007) Horizontal gene transfer regulation in bacteria as a “spandrel” of DNA repair mechanisms. *PLoS One* 2, e1055
- 36 Bertolla, F. *et al.* (1999) During infection of its host, the plant pathogen *Ralstonia solanacearum* naturally develops a state of competence and exchanges genetic material. *Mol. Plant-Microbe Interact.* 12, 467–472
- 37 Baltrus, D. a (2013) Exploring the costs of horizontal gene transfer. *Trends Ecol. Evol.* 28, 489–495
- 38 Schmidt, H. and Hensel, M. (2004) Pathogenicity islands in bacterial pathogenesis. *Clin. Microbiol. Rev.* 17, 14–56
- 39 Hacker, J. *et al.* (1997) Pathogenicity islands of virulent bacteria: structure, function and impact on microbial evolution. *Mol. Microbiol.* 23, 1089–1097
- 40 Alfano, J.R. and Collmer, A. (1997) The type III (Hrp) secretion pathway of plant pathogenic bacteria: trafficking harpins, Avr proteins, and death. *J. Bacteriol.* 179, 5655–5662
- 41 Stukenbrock, E.H. and McDonald, B.A. (2008) The origins of plant pathogens in agro-ecosystems. *Annu. Rev. Phytopathol.* 46, 75–100
- 42 Barash, I. and Manulis-Sasson, S. (2009) Recent evolution of bacterial pathogens: the gall-forming *Pantoea agglomerans* case. *Annu. Rev. Phytopathol.* 47, 133–152
- 43 Hockett, K.L. *et al.* (2013) Interactions between genome architecture and virulence genes in *Pseudomonas syringae*, strain CC1557 as a model. *bioRxiv* DOI: 10.1101/000869
- 44 Planet, P.J. *et al.* (2013) Emergence of the epidemic methicillin-resistant *Staphylococcus aureus* strain USA300 coincides with horizontal transfer of the arginine catabolic mobile element and *speG*-mediated adaptations for survival on skin. *MBio* 4, e00889–13

- 45 Wren, B.W. (2003) The *Yersinia* a model genus to study the rapid evolution of bacterial pathogens. *Nat. Rev. Microbiol.* 1, 55–64
- 46 Zhou, D. *et al.* (2004) DNA microarray analysis of genome dynamics in *Yersinia pestis*: insights into bacterial genome microevolution and niche adaptation. *J. Bacteriol.* 186, 5138–5146
- 47 Wommack, K.E. and Colwell, R.R. (2000) Virioplankton: viruses in aquatic ecosystems. *Microbiol. Mol. Biol. Rev.* 64, 69–114
- 48 Kidambi, S.P. *et al.* (1994) Evidence for phage-mediated gene transfer among *Pseudomonas aeruginosa* strains on the phylloplane. *Appl. Environ. Microbiol.* 60, 496–500
- 49 Brüssow, H. *et al.* (2004) Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. *Microbiol. Mol. Biol. Rev.* 68, 560–602
- 50 Lin, L. *et al.* (1977) Increased reproductive fitness of *Escherichia coli* lambda lysogens. *J. Virol.* 21, 554–559
- 51 Edlin, G. *et al.* (1977) Reproductive fitness of P1, P2, and Mu lysogens of *Escherichia coli*. *J. Virol.* 21, 560–564
- 52 Freeman, V.J. (1951) Studies on the virulence of bacteriophage-infected strains of *Corynebacterium diphtheriae*. *J. Bacteriol* 61, 675–688
- 53 Waldor, M.K. and Mekalanos, J.J. (1996) Lysogenic conversion by a filamentous phage encoding cholera toxin. *Science* 272, 1910–1914
- 54 Newland, J.W. *et al.* (1985) Cloning of Shiga-like toxin structural genes from a toxin converting phage of *Escherichia coli*. *Science* 230, 179–181
- 55 Mirolid, S. *et al.* (1999) Isolation of a temperate bacteriophage encoding the type III effector protein SopE from an epidemic *Salmonella typhimurium* strain. *Proc. Natl. Acad. Sci. U. S. A.* 96, 9845–9850
- 56 Clark, C.A. *et al.* (1991) The *oac* gene encoding a lipopolysaccharide O-antigen acetylase maps adjacent to the integrase-encoding gene on the genome of *Shigella flexneri* bacteriophage Sf6. *Gene* 107, 43–52
- 57 Scott, J. *et al.* (2008) Phage-associated mutator phenotype in group A streptococcus. *J. Bacteriol.* 190, 6290–6301
- 58 Morris, C.E. *et al.* (2009) Expanding the paradigms of plant pathogen life history and evolution of parasitic fitness beyond agricultural boundaries. *PLoS Pathog.* 5, e1000693

- 59 Spooner, R. and Yilmaz, O. (2011) The role of reactive-oxygen-species in microbial persistence and inflammation. *Int. J. Mol. Sci.* 12, 334–352
- 60 Torres, M.A. *et al.* (2006) Reactive oxygen species signaling in response to pathogens. *Plant Physiol.* 141, 373–378
- 61 Torres-barceló, C. *et al.* (2013) A trade-off between oxidative stress resistance and DNA repair plays a role in the evolution of elevated mutation rates in bacteria . *Proc. R. Soc. B* 280, 20130007
- 62 Cupples, C.G. and Miller, J.H. (1989) A set of *lacZ* mutations in *Escherichia coli* that allow rapid detection of each of the six base substitutions. *Proc. Natl. Acad. Sci. U. S. A.* 86, 5345–5349
- 63 Poole, K. (2012) Stress responses as determinants of antimicrobial resistance in Gram-negative bacteria. *Cell* 20, 227–234
- 64 Bertolla, F. and Simonet, P. (1999) Horizontal gene transfers in the environment: natural transformation as a putative process for gene transfers between transgenic plants and microorganisms. *Res Microbiol* 150, 375–384
- 65 McDaniel, L.D. *et al.* (2012) Environmental factors influencing gene transfer agent (GTA) mediated transduction in the Subtropical Ocean. *PLoS One* 7, e43506
- 66 Waldor, M.K. and Mekalanos, J.J. (1996) Lysogenic conversion by a filamentous phage encoding cholera toxin. *Science* 272, 1910–1914
- 67 Miroid, S. *et al.* (1999) Isolation of a temperate bacteriophage encoding the type III effector protein SopE from an epidemic *Salmonella typhimurium* strain. *Proc. Natl. Acad. Sci. U. S. A.* 96, 9845–9850
- 68 Li, G.-M. (2008) Mechanisms and functions of DNA mismatch repair. *Cell Res.* 18, 85–98
- 69 Fishel, R. *et al.* (1993) The human mutator gene homolog MSH2 and its association with hereditary nonpolyposis colon cancer. *Cell* 75, 1027–1038

Chapter 3

**A user's guide to a data base of the
diversity of *Pseudomonas syringae* and its
application to classifying strains in this
phylogenetic complex**

Accepted in PloS ONE in May 2014

Chapter 3- A user's guide to a data base of the diversity of *Pseudomonas syringae* and its application to classifying strains in this phylogenetic complex.

A user's guide to a data base of the diversity of *Pseudomonas syringae* and its application to classifying strains in this phylogenetic complex.

Odile Berge¹, Caroline L. Monteil¹, **Claudia Bartoli**^{1,2}, Charlotte Chandeysson¹, Caroline Guilbaud¹, David C. Sands³. & Cindy E. Morris^{1,3*}

¹INRA, UR0407 Pathologie Végétale, 84143 Montfavet, France

²Department of Science and Technology for Agriculture, Forestry, Nature and Energy (DAFNE), Tuscia University, Viterbo, Italy

³Department of Plant Sciences and Plant Pathology, Montana State University, Bozeman, MT 59717-3150, USA

*For correspondence: E-mail cindy.morris@avignon.inra.fr; (+33) (0) 432722886; Fax (+33) (0) 432722841

Chapter 3- A user's guide to a data base of the diversity of *Pseudomonas syringae* and its application to classifying strains in this phylogenetic complex.

ABSTRACT

The *Pseudomonas syringae* complex is composed of numerous genetic lineages of strains from both agricultural and environmental habitats including habitats closely linked to the water cycle. The new insights from the discovery of this bacterial species in habitats outside of agricultural contexts per se have led to the revelation of a wide diversity of strains in this complex beyond what was known from agricultural contexts. Here, through Multi Locus Sequence Typing (MLST) of over 216 strains, we identified 23 clades within 13 phylogroups among which the seven previously described phylogroups were included. Robustness of phylogroups was shown by using core genome phylogeny on 29 strains representative of nine phylogroups. We show that phenotypic traits almost never provide a satisfactory means for classification of strains. We demonstrate that the citrate synthase (*cts*) housekeeping gene can accurately predict the phylogenetic affiliation for more than 97 % of strains tested and we propose a list of *cts* sequences to be used as a simple tool for quickly and precisely classifying new strains. Finally, our analysis leads to predictions about the diversity of *P. syringae* that is yet to be discovered. Nonetheless, we present here an expandable framework mainly based on *cts* genetic analysis into which more diversity can be integrated.

Chapter 3- A user's guide to a data base of the diversity of *Pseudomonas syringae* and its application to classifying strains in this phylogenetic complex.

INTRODUCTION

Pseudomonas syringae was first reported as a plant pathogen of lilac by van Hall in 1902 [1]. Since its first description, *P. syringae* has become recognized as a phylogenetic complex of strains from terrestrial and aquatic habitats [2]. The classification of strains into the various sub-groups that constitute this complex has mirrored the historical trends in bacterial classification that were initially based on phenotypes (physiological and ecological characteristics) and then progressively were based on genotypes (DNA-DNA hybridization, phylogenetic analysis of housekeeping genes sequences) [3]. Commonly, seven phylogroups based on housekeeping gene phylogeny are recognized in the *P. syringae* complex [4] and some authors also include *Pseudomonas cichorii* a closely related phytopathogenic species [5, 6]. These seven groups are more or less consistent with the species or genomospecies described based on DNA-DNA hybridization [7, 8] such as *P. savastanoi* [7], *P. viridiflava* [9] and *P. avellanae* [10] the latter recently re-defined with more accurate genomic analysis [11]. As for many bacterial pathogens, the allocation of strains into pathovars is very common for the *P. syringae* group. Although the concept of pathovar is not related to phylogeny, these pathovars are frequently used as an analytical framework for classifications based on physiological phenotypes [12, 13], MLST (Multi Locus Sequence Typing) phylogeny [14–16] or DNA-DNA hybridization [7]. More recently, strains of *P. syringae* were isolated from contexts where they were saprophytes in a range of environmental substrates. For these strains, the concept of pathovar had no apparent relevance, especially as they sometimes represented phylogroups not previously described among the strains isolated from diseased plants [2, 17]. These discoveries raise questions about how to classify these strains that have not been resolved in a standardized way.

In light of the growing diversity of what is being called *P. syringae* and of the lack of a guide for homogenous classification and naming of strains, we were led to examine the validity of the biochemical indicators and to attempt to clarify the situation. Here we present the results of genotypic and phenotypic characterization of 764 strains of *P. syringae* collected from a wide range of habitats in which this bacterium has been described up to date. These strains were selected to represent the full breadth of the genetic diversity in a collection of over 1600 strains of *P. syringae* for which some phylogenetic information was available. Through phylogenetic analyses based on 4 housekeeping genes we defined 23 clades within 13 phylogroups. Robustness of phylogroups was shown through core genome phylogeny on 29

Chapter 3- A user's guide to a data base of the diversity of *Pseudomonas syringae* and its application to classifying strains in this phylogenetic complex.

strains representative of 9 of the 13 phylogroups. Phenotypic characterization on 764 strains illustrated that phenotypic traits do not provide a satisfactory means for identification of strains at the clade or phylogroup level. We illustrate that the *cts* housekeeping gene alone can accurately predict the phylogenetic situation for most strains at the phylogroup and clade level. Overall, we describe the diversity of *P. syringae* and the utility of the data-base as a tool for classifying strains. Our analysis permits predictions about the diversity of *P. syringae* beyond what has been discovered and hence it provides a framework for future studies of the ecology of this bacterium.

MATERIALS AND METHODS

Bacterial strains

Most strains used in this study were taken from a collection of over 7000 strains of *P. syringae* maintained at INRA in Montfavet. This collection was initiated in about 1995 and consists of strains collected from crops as well as different environmental habitats (rain, snow, epilithic biofilms, litter and water) via isolation on modified medium B of King (KBC) [18]. Some strains from crops were kindly provided by colleagues or obtained from public collections. All strains were purified before being stored at – 80° C in 40 % glycerol and they all lacked arginine dihydrolase and cytochrome *c* oxidase activity except for the *P. cichorii* reference strains that are oxidase positive. In addition, they varied in their production of fluorescent pigment on King's medium B (KB) [19] and in induction of a hypersensitive reaction (HR) on tobacco. Over the past several years the housekeeping gene encoding for citrate synthase, *cts* (also named *gltA*) was sequenced for a subset of 1630 strains isolated from environmental habitats and crops [2, 20] during the exploratory work for our studies.

Construction of a phylogenetic tree based on the 1630 *cts* sequences allowed us to select strains that represented the range of genetic diversity within each of the phylogroups and clades that could be delimited in the first-approximation analysis (unpublished data). This led us to select the 764 strains isolated from fresh water and epilithic biofilms (56 %), snowpack (16 %), plants (11 %), precipitation (9 %), and litter (8 %) that were subsequently characterized for 12 phenotypic traits typically used in characterization of *P. syringae* (see below) thereby allowing us to evaluate phenotypic diversity within different genetic groups. We selected voluntarily many non-plant derived strains to better describe the unknown

Chapter 3- A user's guide to a data base of the diversity of *Pseudomonas syringae* and its application to classifying strains in this phylogenetic complex.

phylogenetic groups of the *P. syringae* complex. The total of 837 strains used in this study is listed in Table S1 with their origin, alternative name and characteristics. Among them 6 were reference strains chosen outside of the *P. syringae* group of strains mainly for phylogenetic analyses. The 831 remaining strains of *P. syringae* consisted of the 764 phenotyped strains and 67 strains from public data-bases used for their MLST profiles (see below) and not phenotyped here. From the 764 characterized strains, 149 were selected to cover the maximum variability observed for MLST. To this data-base of 149 MLST-typed (4 genes) strains we added the MLST profiles of the 67 strains from public data bases. The pooled set of 216 MLST profiles was used to construct more robust trees and to evaluate the reliability of phylogenetic predictions based on single housekeeping genes vs. combined gene sequences. The affiliation to phylogroups of the remaining 615 strains of *P. syringae* was based on *cts* sequence analysis. Finally, a set of 29 strains (Table S1) that represented the maximum diversity among the *P. syringae* genomes available in GenBank and chosen from the set of 216 MLST strains, was used to compare phylogenetic positioning of strains based on core genome sequences vs. that based on single and multiple housekeeping genes.

Genomic and phylogenetic analysis

MLST analysis was performed by sequencing four housekeeping genes: *cts* (encoding citrate synthase), *gapA* (encoding for glyceraldehyde-3-phosphate dehydrogenase A), *rpoD* (encoding for RNA polymerase sigma⁷⁰ factor) and *gyrB* (encoding for gyrase B), using the Morris MLST schema of the Plant Associated and Environmental Microbes Database (PAMDB, <http://genome.ppws.vt.edu/cgi-bin/MLST/home.pl>) in combination with *gapA* and *gyrB* of the Hwang PAMDB schema [16, 20]. For each locus, sequences were extracted from GenBank and PAMDB and aligned with the *P. syringae* sequences by using DAMBE software version 5 [21] and they were cut to the same size (1859 bp for the concatenated sequences). In order to clarify the phylogenetic position of strain LzW4 isolated from Antarctica and misclassified as *P. syringae* [22], housekeeping gene sequences were obtained from its genome (accession number AOGS000000000). The concatenated sequences were used to construct the phylogeny with maximum likelihood and Bayesian methods by using the PHYLIP package version 3.6 (<http://evolution.genetics.washington.edu/phylip.html>) and Mr. Bayes version 3.1.2, respectively [23]. For maximum likelihood analysis, consensus trees were created from 100 independent phylogenies. Bayesian trees were constructed by using

Chapter 3- A user's guide to a data base of the diversity of *Pseudomonas syringae* and its application to classifying strains in this phylogenetic complex.

500 000 generations with a burn-in period of 250 000. All sequences and critical metadata of strains were deposited on the PAMDB data base [24]. Genetic distances among the strains were determined with the Kimura 2-parameter model, with a gamma correction of 1, by using the PHYLIP package. For delimitation of phylogroups, the distance used as a criterion was chosen to allow delineation of the seven previously-described phylogroups of *P. syringae*. For delineation of clades, we used the threshold value of 2.3 % as well as the tree structure as previously described [2].

For strains for which only the *cts* gene was sequenced, phylogenetic affiliation was determined based solely on this *cts* sequence. We first validated this method on the set of 216 MLST-typed strains. The phylogenetic affiliations of the 216 strains based on the *cts* sequence were compared with affiliations obtained in the MLST analysis. This allowed us to develop *cts* distance thresholds (< 4.0 % for phylogroup and < 2.0 % for clade affiliations) and criteria to avoid erroneous assignment (see Methods S1 for details). Affiliation of new strains was determined based on both the *cts* tree topology and the minimum distance of the *cts* sequence to one of the 216 strains. Phylogenetic analysis of a partial core genome (sequences of 107 genes) was also performed. Core genomes were extracted from 29 *P. syringae* genomes (AVEQ000000000, AEAD000000000, ABZR000000000, AVEP000000000, AVEO000000000, AVEN000000000, AVEK000000000, AVEJ000000000, AVEH000000000, AVEG000000000, AVEF000000000, AVEE000000000, AVED000000000, AVEA000000000, AEAL000000000, AEAE000000000, AEAF000000000, AEAO000000000, AEAK000000000, AEAJ000000000, ADWY000000000, AVDV000000000, AVDY000000000, ADGB000000000, AE016853, ADMI020000000, AABP020000000, AAEZ010000000) as described previously [25]. Alignment of the core genome was made by using DAMBE version 5 as described above and a Bayesian tree was built with Mr. Bayes.

Rarefaction curves

Rarefaction curves were constructed by randomly sampling a set of 831 individuals representing 13 different phylogroups or 23 different clades in the same proportions as delimited by their assignment to clades and phylogroups as described above. Random samples were drawn 831 times from the set of individuals and the average cumulative number of clades or phylogroups observed for each draw were calculated with R software version 2.9.1

Chapter 3- A user's guide to a data base of the diversity of *Pseudomonas syringae* and its application to classifying strains in this phylogenetic complex.

with an in-house program (The R Development Core Team, 2009) after 1000 iterations of the succession of draws.

Phenotypic characterization

Phenotypic tests (Table S1) included production of fluorescent pigments on KB and tests in the LOPAT scheme (levan production, presence of cytochrome *c* oxidase, induction of potato soft rot, presence of arginine dihydrolase and induction of a hypersensitive reaction (HR) on tobacco) were performed as described previously [17]. In addition, tests for esculin degradation, acidification of sucrose, and utilization of D(-) tartrate as a sole carbon source were performed as previously described [26, 27]. Ice nucleation activity (INA) was also tested on 10^7 CFU ml⁻¹ suspensions, obtained from cultures growth on KB for 3 days at 26°C. For each strain, 3 drops of 20 µl were deposited on an aluminum plate floating on a cooling bath as described previously [17]. Freezing was determined at one degree intervals from -2 °C to -8°C. Freezing was scored when at least 2 of the droplets froze. Strains were tested for production of syringomycin-like toxins based on a bioassay with *Geotricum candidum* as described previously [28]. The inhibition zone around the bacterial colony was measured as the distance between the edge of the fungus and the bacterial colony. Bacteria were grown for 8 days before the plate was sprayed with a suspension of *G. candidum* after which it was incubated for 2 days before inhibition zones were measured. *Cucumis melo* var. *cantalupensis* Naud. cv. Vedrantaïs seedlings were used as an indicator plant to estimate the pathogenicity and aggressiveness of strains. These parameters were assessed after injecting 10 µl of an aqueous suspension at 10^8 CFU ml⁻¹ prepared from 48 h bacterial cultures [20]. Seedlings were incubated for 7 days with a photoperiod of 16 h of light at 21°C during the day and 18°C during the dark period. Symptoms on seedlings were scored as follows: 0 (no symptoms), 1 (one cotyledon with necrosis or completed wilted), 2 (necrosis on both cotyledons), 3 (both cotyledons wilted and stem symptoms) and 4 (death of the cotyledons and stem). Pathogenicity was recorded positive when the frequency of seedlings with symptoms (F) was > 50 % and aggressiveness was calculated as the mean score of symptoms (μ) (Table S1). This test was also used as a proxy of the extent of host-range of pathogenic strains as demonstrated previously [17].

Chapter 3- A user's guide to a data base of the diversity of *Pseudomonas syringae* and its application to classifying strains in this phylogenetic complex.

RESULTS AND DISCUSSION

Delimitation of the phylogroups and clades represented in the strain collection

Phylogeny of 216 *P. syringae* strains constructed on *cts*, *gyrB*, *gapA* and *rpoD* sequences with maximum likelihood and Bayesian models revealed 13 phylogenetic groups composed of multiple sub-groups (clades) constituting 23 total clades (Fig. 1). Only seven phylogroups were subdivided into clades (Fig. 1). All strains analyzed had a genetic distance less than 5 % with strains from their own phylogroup and more than 5 % with strains outside their phylogroup with some exceptions (PsyCit7, CCV0213, CCV0567 and FMU107) (Table S3). The mean genetic distances within and outside phylogroups showed that they are relatively homogeneous and distinctly different from each other (Table 1).

The phylogroups were robust independent of the phylogenetic model used to construct the phylogeny. Delimitation of phylogroups was determined by accounting for both tree branches (Fig. S1a) and genetic distances among the strains (Table S3). A genetic distance between concatenated sequences of less than 5 % defined clearly the seven previously described phylogroups: phylogroups 1, 2 and 3 [29], phylogroup 4 [15], phylogroup 5 [16] and phylogroups 6 and 7 [4]. We used the same distance threshold of 5 % for the delimitations of additional phylogroups (Table S3, Fig. S1a). Names previously attributed to well-known phylogroups (1 to 7) or clades, such as 2a, 2c or the recently named phylogroup 8 [30], were maintained to avoid confusion. Correspondences between the names proposed here and those of reference strains named as species, pathovars and genomospecies are indicated in Table S2. Among the reference strains used to construct the tree, *P. graminis* and *P. rhizosphaerae* strains were included to better delimit the *P. syringae* monophyletic group. These species are among the closest species outside the boundary of the *P. syringae* complex [6]. Three reference strains of *P. cichorii* formed a monophyletic clade clearly included in the *P. syringae* group (Fig. S1a). The *P. cichorii* CFBP 4407 strain had a distance higher than 5 % from the two others (Table S3) but we nevertheless compressed all three strains in the *P. cichorii* phylogroup 11 (Fig. 1) since we considered that the diversity and phylogeny of *P. cichorii* is not well characterized and needs further investigation. This group represents to date the oxidase positive lineage of the *P. syringae* group of strains. Finally, our analysis reveals that strain LzW4 isolated from Antarctica and previously named *P. syringae* [22] is

Chapter 3- A user's guide to a data base of the diversity of *Pseudomonas syringae* and its application to classifying strains in this phylogenetic complex.

most closely related to *P. protegens* Pf-05 and it is clearly outside the *P. syringae* complex (Fig. 1).

The phylogeny based on MLST of the 216 *P. syringae* strains is the framework we used to classify the remaining strains. To evaluate its robustness, we compared this phylogeny to that based on sequences of nearly whole core genomes of 29 strains representative of all the phylogroups except phylogroups 6, 8, 11 and 12 since no genomes were available for those phylogroups. The phylogeny based on 107 open reading frames (64,000 bp) illustrated in the unrooted tree in Fig. 2 showed the same phylogroup topology as the tree based on four housekeeping genes (Fig. 1). This result suggests that phylogeny at phylogroup level is robust enough to be represented by MLST analysis and that core genome analysis is not indispensable for studying the diversity of *P. syringae* and classifying strains in phylogroups. The robustness of phylogeny based on MLST was demonstrated for the *P. syringae* phylogroups 1 to 5 with seven genes independently [15]. Here we confirmed the robustness of phylogeny for all *P. syringae* phylogroups 1 to 5 and demonstrated it for phylogroups 7, 9, 10 and 13.

Construction of a data base of 764 phenotyped strains of P. syringae classified into phylogroups and clades

To obtain the broadest range of information about the characteristics of the phylogroups and clades delimited here, we sought to classify the remaining 615 strains by using a reliable method that is simpler than MLST. With this aim, we used genetic distances and a tree constructed with only the *cts* housekeeping gene previously validated on the 216 strains used in the MLST analysis.

The *cts* gene was chosen as a *P. syringae* tool classification because, as previously described [15], it corresponds to one of the most reliable gene sequences among the genes used in MLST because it has the minimum number of recombinations and the most congruence among the trees constructed with housekeeping genes. To reinforce these previous observations, we compared the tree based on the core genomes (Fig. 2) with that built on only *cts* genes and we showed that phylogeny at phylogroup and clade levels among trees was consistent (Fig 3). The classification of the 216 strains based on the *cts* gene sequence analysis validated our *P. syringae* tool classification with few exceptions (4/216) that show some strains needed more than one housekeeping gene to be robustly classified (Methods S1).

Chapter 3- A user's guide to a data base of the diversity of *Pseudomonas syringae* and its application to classifying strains in this phylogenetic complex.

Since we demonstrated that *cts* phylogeny is suitable for classification of strains, such classification was then performed on the 615 non-classified *P. syringae* strains. By comparing *cts* distances between 615 individual strains and the 216 identified strains and by taking into account the minimum distance, we classified the 615 strains without ambiguities except for 16 strains that were equidistant to phylogroup 1a and one strain of phylogroup 1b (Table S4). These strains were finally affiliated to phylogroup 1a based on their placement in the phylogenetic tree (data not shown). On the total set of the 831 strains of *P. syringae*, we calculated that 97.6 % were easily classified via the method proposed here. Finally, a total of 764 strains were classified into phylogroups and clades and characterized for their phenotypes. This constitutes a rather complete database useful for classification and characterization of strains that belong to the *P. syringae* complex (Table S1).

Characteristics of the 13 phylogroups and associated clades

The results we provide for phylogroups clearly showed that phenotypes are variable among and within phylogroups (Table 2). However, some interesting results emerged from phenotypic traits: i) arginine dihydrolase was absent in all strains without exception but since this phenotype was used as criterion for elimination during isolation steps, positive strains could have been missed ii) as expected, the only phylogroup that was oxidase-positive was phylogroup 11 containing the *P. cichorii* strains; the possibility that oxidase positive strains in the *P. syringae* complex from other phylogroups were discarded during isolation could not be excluded, iii) the production of fluorescent pigments and degradation of esculin are the phenotypes that were positive for all strains in all phylogroups except for phylogroup 1, 3 and 5, iv) 65 % of strains were ice nucleation active and the absence of this activity was observed for all strains in phylogroups 8, 11, 12 and 13, v) HR on tobacco was positive for 73 % of the strains, vi) only 28 % of the strains were pathogenic on cantaloupe seedlings, vii) production of toxins inhibiting *Geotricum candidum* was frequent for strains of phylogroup 2, but also for phylogroup 8, 10 and 11. For phylogroup 2 and recently in the phylogroup 10 [31] genes for syringomycin toxins have been described and are likely to be involved in the toxicity observed here. But the mechanisms for the production of this toxin remain to be investigated for phylogroups 1, 4, 8 and 11.

Each phylogroup was characterized based on its phenotypic and genotypic traits as follows:

Chapter 3- A user's guide to a data base of the diversity of *Pseudomonas syringae* and its application to classifying strains in this phylogenetic complex.

***P. syringae* phylogroup 1.** This phylogroup contains many strains from diseased plants but also from numerous environmental habitats and substrates [2] including precipitation, fresh water, epilithic biofilms and leaf litter (Table S1). Phylogroup 1 consists of two clades described by other authors [8,11]. Strains in clade 1a (including *P. s. pv. tomato*) gave variable results for phenotypic tests but all produced fluorescent pigment on KB medium and degraded esculin (Table 2). Clade 1b includes *P. avellanae* and *P. s. pv. actinidiae*, respectively the causal agents of bacterial canker of hazelnut and kiwifruit and also numerous strains isolated from snow, water and leaf litter (Table S1, S2). Strains in this clade, as well as strains in phylogroup 3, contain a catechol operon regrouping genes for degradation of aromatic compounds (unpublished data). Strains in clades 1a and 1b were similar in terms of their phenotypic variability, except that ca. 32 % of the latter did not produce fluorescent pigment on KB and 17 % did not degrade esculin. All the strains that did not degrade esculin carried the genes for degradation of aromatic compounds (unpublished data). Genomic studies have shown that among all phylogroups, strains of phylogroup 1 have the greatest number of Type Three Effector (T3E) genes coding for virulence determinants [25, 32]. More recently, Monteil and coworkers [33] demonstrated that strains closely related to the tomato speck pathogen *P. s. pv. tomato* isolated from snowpack and streams harbor the T3E genes found in epidemic strains. Most T3E gene expressions are driven by the HrpL sigma factor that also regulates non-T3E genes associated with virulence. All the genes regulated by the HrpL sigma factor are called HrpL regulons [34]. Consistent with this observation, the clade 1a strain *P. s. pv. tomato* DC3000 has the greatest number of HrpL regulons described to date [34].

***P. syringae* phylogroup 2.** This is the most ubiquitous phylogroup of *P. syringae* found in all habitats analyzed to date [2]. In this phylogroup, three subgroups had been described previously, 2a, 2b and 2c [35] (Table S2). Phylogroup 2 is in fact composed of five different clades all containing some non-plant derived strains: i) *P. syringae* clade 2a contains strain PsyCit7 isolated from an asymptomatic orange tree [36], a strain from rain and one from an irrigation basin ii) *P. syringae* clade 2b includes the *P. syringae* pv. *syringae* type strain (CFBP 1392^T), *P. s. pv. aptata* (CFBP 1906), *P. s. pv. atrofaciens* (CFBP 2256) and many strains isolated from all environmental substrates , iii) *P. syringae* clade 2c is represented by non-pathogenic *P. syringae* strains (such as Psy642 and SZ0030) isolated from plants and

Chapter 3- A user's guide to a data base of the diversity of *Pseudomonas syringae* and its application to classifying strains in this phylogenetic complex.

environmental substrates having an atypical Type Three Secretion System (T3SS) similar to the T3SS of S-PAI *P. viridiflava* [35]. Many strains in this clade contain identical sequences of a bacteriophage unique to this clade [37], iv) *P. syringae* clade 2d contains strain B728a and is closely related to clade 2b, v) *P. syringae* clade 2e is presently represented by only two strains isolated from fresh water and snow (Table S1). Although phylogroup 2 contains some strains incapable of inducing HR on tobacco due to the presence of an inefficient non canonical T3SS (clade 2c strains), we confirmed that strains in this phylogroup are on average more aggressive on cantaloupe seedlings than strains in all other phylogroups (Table 1) [2]. They are among the most consistently ice nucleation active (85 % of strains) and most of them (90 %) produce a syringomycin-like toxin (Table 2). Up until the recent characterization of strains in phylogroup 10 (described below) [31], the genomes of phylogroup 2 strains had been considered to carry the fewest T3E genes among all phylogroups [25]. In parallel, they have numerous genes for phytotoxins such as syringolin, syringopeptin and syringomycin [25]. Strain B728a in clade 2d was recently reported to carry the fewest HrpL regulons [34].

***P. syringae* phylogroup 3.** The previously reported descriptions of this group that included many pathovars (Table S2) are not greatly influenced by our study because only very few strains in phylogroup 3 were isolated from environmental sources. This result could be partly due to a bias in the isolation method. Strains of phylogroup 3 tend to grow more slowly on KB media than strains of phylogroup 2 for example, and hence they could have been missed. Only two non-plant derived strains in this group, CMO0010 and LYR0002, were isolated from rain in a cultivated area [38] but none from river water, epilithic biofilms, litter or snowpack from which we have extensively sampled particularly in mountainous pastoral areas [27]. This group contains pathogens of woody plants (*P. savastanoi* pvs. *savastanoi*, *aesculi*, and *mori*) that have been found to carry genes for the degradation of aromatic compounds [39, 40], but also contains pathogens of other types of host plants such as soybean and French bean (*P. savastanoi* pvs. *glycinea* and *phaseolicola*). Phylogroup 3 was not divided into clades because clade branches were not robust perhaps due to recombination events. Among the notable phenotypic traits, incapacity to degrade esculin and to produce fluorescent pigment on KB medium were frequent, similar to the properties observed in clade 1b. Phylogroup 3 contained many strains that induce HR on tobacco, but none were

Chapter 3- A user's guide to a data base of the diversity of *Pseudomonas syringae* and its application to classifying strains in this phylogenetic complex.

aggressive on cantaloupe seedlings. In addition, phylogroup 3 strains were rarely ice nucleation active (20 %) and none produced a syringomycin-like toxin (Table 2).

***P. syringae* phylogroup 4.** Few strains have been described in this phylogroup to date but seven pathovars have been reported (Table S2). Strains were isolated from diverse sources including cropped and wild plants (mostly monocotyledonous); rain, snowpack and plant litter (Table S1). As for phylogroup 3, strains from environmental substrates were rare and this could have resulted from a sampling bias. The lack of genetic diversity of the eight strains from this group prevented us from delineating clades. The strains in this phylogroup were remarkably homogenous in their phenotypes, but this might be due to the limited number of strains tested. In contrast to other phylogroups, all strains in phylogroup 4 were ice-nucleation active. Interestingly, although these strains have been rarely detected in the environment, they were nevertheless among the highly ice-nucleation active strains from clouds on the Puy de Dôme in France [41]. Concerning the T3E repertoire within this phylogroup, two new T3SS genes, hopBH1 and hopBI1, were recently described in the strains 1_6 (pathogenic on rice), CC1513 (from healthy wild *Hutchinsia alpine*) and CC1629 (from cropped *Avena sativa*). Strain 1_6 has been reported to have the greatest number of HrpL regulons (T3E and non-T3E) in the *P. syringae* complex [34].

***P. syringae* phylogroup 5.** This phylogroup is represented by only five strains of which four were isolated from crops and one from stream water (Table S1). Phylogroup 5, that includes strains pathogenic on diverse plants like *Cannabis sativa*, *Brassicaceae*, or coriander (Table S2), was not found to be abundant in the environment. In contrast to phylogroup 4, phenotypes within this group were highly variable in spite of the limited number of strains (Table 2). A larger collection of strains from this group is needed to understand its diversity. One of the four strains phenotyped in this study was negative for production of fluorescent pigments and no strain produced a syringomycin-like toxin (Table 2).

***P. syringae* phylogroup 6.** At present, phylogroup 6 contains only strains isolated from diseased crops, including *Asteraceae* (*P. s.* pv. *tagetis*) and papaya (*P. caricapapayae*) in particular (Table S2). Only one strain, *P. s.* pv. *helianthi* (CFBP 2067) was included in this study. Diversity of this phylogroup in the environmental context still needs to be investigated.

Chapter 3- A user's guide to a data base of the diversity of *Pseudomonas syringae* and its application to classifying strains in this phylogenetic complex.

***P. syringae* phylogroup 7.** This phylogroup represents most of the strains called *P. viridiflava* in previous studies, as well as two *P. syringae* pathovars (*P. s. pv. ribicola* and *pv. primulae*) (Table S2) and many strains from a wide range of environmental reservoirs (Table S1). Almost all strains from phylogroup 7 are capable of causing soft rot to potato slices and to display phase variation [30]. This latter behavior has a considerable impact on several phenotypes including soft rot of potato and pathogenicity. Phylogroup 7 consists of two clades, 7a containing most of the strains including those isolated from fresh water, epilithic biofilms, plant litter, snowpack and precipitation and 7b containing only two strains from two different plants (Table S1) thus limiting the analysis of this group. Strains from phylogroup 7 harbor one non-canonical T3SS that resembles the one found in clade 2c [30, 42].

***P. syringae* phylogroup 8.** This phylogroup (Fig. 1, Table 1) also contains strains that could be called *P. viridiflava*. All were isolated from environmental sources (Table S1). These strains share numerous characteristics with those in phylogroup 7, including phase variation [30]. However, they less frequently cause soft rot of potato slices (40 % of strains) than those in phylogroup 7, have no ice nucleating activity and no aggressiveness on cantaloupe seedlings, but all produce a toxin in the bioassays with *G. candidum*. Due to the absence of syringomycin genes (unpublished data), this toxicity could be the result of the production of an antimycotic peptide such as ecomycin identified in *P. viridiflava* [43]. No clades were differentiated in this phylogroup.

***P. syringae* phylogroup 9.** Strains that represent phylogroup 9 have only been reported in aquatic habitats including lakes, epilithic biofilms and irrigation canal water (Table S1). Strains in phylogroup 9 did not produce syringomycin-like toxin, and only 4 % were ice nucleating active. Three clades were delineated; however there had no distinct phenotypic differences. Phylogroup 9 corresponds to the phylogroup previously named CC1524 [2] (Table S2). Analysis of the genome of strain CC1524 revealed that it harbored novel HrpL regulons that were not found in strains in the other phylogroups [34].

***P. syringae* phylogroup 10.** Strains in phylogroup 10 were the second most abundant in the collection analyzed in this study. Strains in this phylogroup were exclusively from

Chapter 3- A user's guide to a data base of the diversity of *Pseudomonas syringae* and its application to classifying strains in this phylogenetic complex.

environmental reservoirs outside of areas cultivated for agriculture (Table S1). To date, no strains of this group have been isolated from diseased plants (Table S2). However, almost all strains in phylogroup 10 (98 %) induce HR on tobacco, 94 % are ice nucleating active and 10 % are pathogenic on cantaloupe seedlings. Genes for syringomycin-like toxins were found in the two full genome sequences available for this group (CC1583 and CC1557) and 51 % of strains produced a syringomycin-like toxin in the bioassay. Among phylogroup 10 strains, seven clades were delimited, three of them corresponding to previously described clades “USA102” (10a), “TA0003” (10b), “USA032” (10e) [2] and three others containing only one strain (10c, 10f and 10g containing respectively CC1586, CCE0153 and CCV0213). The strains in the 10a, 10b, 10d and 10e clades are variable for most traits but show distinct differences in the production of levan and syringomycin-like toxins and in their pathogenicity and aggressiveness on cantaloupe (Table 2). The genomes of strains in phylogroup 10 have been recently reported to have the fewest T3E genes among all the strains in the *P. syringae* complex for which the T3SS has been characterized [31]. As mentioned above, this characteristic had been previously attributed to phylogroup 2 before genomes of strains of phylogroup 10 were available [25]. Finally, phylogroup 10 is quite comparable to phylogroup 2 in terms of its ubiquity, phenotypes, and number of T3E. Since these two phylogroups are phylogenetically distant, convergent evolution could have shaped their behavior through horizontal gene transfer and other evolutionary processes linked to environmental pressures.

***P. syringae* phylogroup 11.** Phylogroup 11 is formed by strains that were classified in the *P. cichorii* species. This species was distinguished originally from *P. syringae* because of its cytochrome *c* oxidase, absent from *P. syringae* [18]. In this study we did not isolate strains of this phylogroup because only oxidase-negative isolates were retained in our basic isolation process. The inclusion of *P. cichorii* lineage in the *P. syringae* complex was already proposed on the basis of phylogeny of housekeeping genes [6]. Strains belonging to phylogroup 11 are reported to be pathogenic on many crops such as lettuce or tomato [44]. Strains of *P. cichorii* have also been isolated from irrigation water [45]. The ecology of this phylogroup in environmental habitats and its diversity in a non-agricultural context remain to be explored. The three strains of phylogroup 11 tested here produced a toxin in the bioassays with *G. candidum*. It is likely that this toxicity is not due to syringomycin-like toxins as already shown by Hu et al. (1998). The main difference between phylogroup 11 and the rest of the *P.*

Chapter 3- A user's guide to a data base of the diversity of *Pseudomonas syringae* and its application to classifying strains in this phylogenetic complex.

syringae phylogroups is its ability to induce a positive reaction in the cytochrome *c* oxidase test. Interestingly the cytochrome *c* oxidase operon was also found in the genome of phylogroup 7 strains (CC1582 and TA043) but in none of the other phylogroups (unpublished data). However, all phylogroup 7 strains are negative for the phenotypic oxidase test. Interestingly, the single component-T3SS (S-PAI) of phylogroup 7 strains is evolutionarily related to the T3SS of phylogroup 11 strains [30] and we hypothesize that the S-PAI configuration represents the most ancient form of T3SS in the *P. syringae* complex.

***P. syringae* phylogroup 12.** This phylogroup is composed of two strains (GAW0112 and GAW0113) isolated from water in an irrigation canal. According to the criteria used to delimit clades, each of these strains represents a distinct clade. These two strains did not induce HR on tobacco, did not cause disease when inoculated on cantaloupe seedlings and gave negative results for many of the other phenotypic tests. Overall, phylogroup 12 strains resemble phylogroup 13 strains (see below) more than those in other phylogroups in terms of the phenotypes characterized here. This finding is consistent with the fact that these two groups are phylogenetically similar and somewhat distant from the other phylogroups (Fig. 1, Table 1).

***P. syringae* phylogroup 13.** Numerous strains isolated from non-plant substrates were found to be affiliated to phylogroup 13 (Table S1), previously called group UB246 [2] (Table S2). At present, phylogroup 13 consists of strains isolated from fresh water; epilithic biofilms, snowpack and plant litter (Table S1). A recent study showed the existence of phylogroup 13 strains in wild alpine plants [46], suggesting that phylogroups such as this one are more widespread than our work indicates. The phenotypes of strains in phylogroup 13 were relatively homogenous as all of them produced a fluorescent pigment on KB medium, were able to degrade esculin, utilize D (-) tartrate as a sole carbon source and none of them were positive for the following tests: HR on tobacco, levan production, sucrose acidification, ice nucleation activity, toxin production and induction of symptoms on cantaloupe seedlings. Two clades were delimited, 13a containing most of the strains and 13b containing only two strains isolated from plant litter. A phylogenetic analysis of the *hrcC* T3SS gene showed that the T3SS of phylogroup 13 is more related to the *P. viridiflava* S-PAI rather than to the canonical T3SS of other *P. syringae* phylogroups [30].

Chapter 3- A user's guide to a data base of the diversity of *Pseudomonas syringae* and its application to classifying strains in this phylogenetic complex.

Tools and guidelines for classifying strains in the P. syringae complex.

The use of phenotypic traits to classify strains in the *P. syringae* complex is sometimes the only option for identification, in particular for small diagnostic laboratories or when resources are limited. However, our results clearly illustrate that the phenotypic traits described here can be readily misleading due to the inherent heterogeneity within phylogroups and within clades. For specific epidemics, clonal lineages of pathogens might have identical phenotypic traits, but those traits could be similar to those of strains from other clades. Furthermore, numerous diseases are caused by a diversity of *P. syringae* strains with different phenotypic profiles and in some cases from multiple phylogroups [34, 47-50]. Phenotypic criteria used during the screening of strains can markedly limit the diversity revealed in ecological or epidemiological studies. The only traits common to all strains used in this study (except phylogroup 11 strains) are the capacity to grow on KBC medium, which contains cephalixin and boric acid as selective agents, and the absence of arginine dihydrolase and cytochrome *c* oxidase. These traits might have limited the diversity uncovered in this study because they were used in the initial selection of strains. However, without a selective medium it would have been impossible to reveal the presence of *P. syringae* in most environmental reservoirs where it can constitute a mere 0.1% or less of the total bacterial population [20, 27]. Isolation of *P. syringae* with classical microbiological methods still remains a technique of choice for studying a bacterium with relatively very low abundance in most of the substrates it inhabits, diseased plants being the principal exception. Production of fluorescent pigments has been very useful to differentiate colonies of *P. syringae* and the occurrence of non-fluorescent strains diminishes the utility of differential media for fluorescent production and seriously complicates comprehensive ecological studies. The presence of the operon for pyoverdinin production in the genomes of non-fluorescent strains (such as all *P. s. pv. actinidiae*) and recent successes to express fluorescence in these strains on different media [51] suggest the possibility to improve differential media for production of fluorescent pigment.

Phylogroup 11, the “oxidase-positive lineage” of *P. syringae*, is an exception that requires another isolation procedure since not all strains are able to grow on KBC medium (data not shown). Moreover, the positive oxidase test could not distinguish strains of phylogroup 11 from other ubiquitous fluorescent pseudomonads related to *P. fluorescens* [52].

Chapter 3- A user's guide to a data base of the diversity of *Pseudomonas syringae* and its application to classifying strains in this phylogenetic complex.

Our results illustrate that, at present, the most precise and efficient means to classify strains in the *P. syringae* complex is to compare the sequences of their *cts* genes to that of the strains used in this study. We propose a list of *cts* reference sequences (Table S5) representing all phylogroups and clades identified in this study. For each clade and each phylogroup the most distant strains from the set of 216 *P. syringae* characterized by MLST were selected for *cts* sequence analysis. These sequences have also been deposited in GenBank (accession numbers indicated in Table S5). The procedure of classification consists of the following steps: 1) alignment of the *cts* gene sequence of the strain to be identified with those in Table S5, 2) analysis of tree branches and of the matrix of pair-wise distances to find the phylogroup and clade with which it is most similar, 3) assignment to the phylogroup or clade if the following criteria are met: < 4 % difference in the sequences for assignment to a phylogroup, and < 2 % difference for assignment to a clade, keeping in mind certain caveats. Due to possible recombination in housekeeping genes, affiliation to the clade level can be uncertain and especially when differences are near or > 4 % as for strain PsyCit7 [15]. Furthermore, as suggested below, a few new phylogroups and many new clades of *P. syringae* are yet to be found. Uncertainties in classification can be addressed by sequencing additional housekeeping genes and performing phylogenetic analyses based on sequences of multiple genes as recommended previously [8]. The data base that we describe here could provide a useful framework for characterization of new biodiversity.

New diversity to anticipate in the P. syringae complex

The diversity we described here is likely to be only a fraction of the entire *P. syringae* diversity. We performed rarefaction analysis on a meta-population structure based on the strains classified here plus an additional 67 strains from public data bases (Fig. 4). This analysis suggests that the number of phylogroups revealed in this study is near its maximum but the number of clades is much smaller than the maximum in the total *P. syringae* meta-population. Hence, descriptions of many new clades and some new phylogroups should be anticipated as ecological studies of the *P. syringae* complex continue. It should be kept in mind that this prevision of new diversity is based on the ecosystems explored in this study. A preponderance of our strain sources are from France. Exploration of additional ecosystems in other geographic locations is likely to increase the probability of discovery of even more genetic diversity of *P. syringae* than we predict here. Populations of endophytic *P. syringae* in

Chapter 3- A user's guide to a data base of the diversity of *Pseudomonas syringae* and its application to classifying strains in this phylogenetic complex.

native bitter cress (*Cardamine cordifolia*) growing in a subalpine context in Colorado at an elevation above 3000 m have recently been characterized [46]. These endophytes were highly diverse belonging to phylogroups 1, 5, 7, 10, and 13 and to three putative new phylogroups. New diversity of *P. syringae* might also be found in association with hosts other than plants (algae, insects, fungi,) or in marine or other more extreme habitats. For example, strain CFII64 isolated from the highly contaminated Clark Fork river in Montana in a study of tolerance to cadmium exposure (http://www.ebi.ac.uk/ena/data/view/GCA_000416235) is closely related to phylogroup 13 (data not shown). This illustrates the diversity that remains to be discovered in this bacterial group and the need for a consistent way to classify strains.

Chapter 3- A user's guide to a data base of the diversity of *Pseudomonas syringae* and its application to classifying strains in this phylogenetic complex.

CONCLUSION

We propose a clear and standard classification of 13 phylogroups and associated clades forming the *P. syringae* complex. This classification considers the ensemble of strains described to date and provides a comprehensive analysis of the phenotypic variation in these phylogroups and clades relative to traits that have commonly been used to identify *P. syringae*. We clearly illustrate that, although phenotypes provide important ecological information, all phenotypic traits tested here other than absence of cytochrome c oxidase activity and arginine dihydrolase can be misleading as a means to identify strains in the *P. syringae* complex or to classify them into a specific phylogroup or clade. In this light we describe a simple method to identify strains of *P. syringae* based on the sequence of a single housekeeping gene and provide the data base needed for this approach. As population genomics emerges [53], it is likely that similar general conclusions will be made about genetic heterogeneity within phylogroups and clades.

By clarifying the classification of strains from a wide range of habitats and describing the genotypic and phenotypic profiles of the different phylogroups, we reveal a fascinating diversity of adaptive strategies deployed within the *P. syringae* complex (Table 3). These phylogroups vary in the nature of their T3SS and its efficiency in inciting plant disease, the balance of effectors, HrpL regulons and toxins in their genomes, and in the production of enzymes to degrade cell walls, for example (Table 3). Yet they all are apparently capable of surviving and multiplying in some environments in sufficient quantities to be detected in isolation schemes. These contrasting profiles raise questions about the fundamental traits of *P. syringae* that are essential for its survival and fitness and which of these are important in the potential of this bacterium to emerge in new epidemics of plant disease.

ACKNOWLEDGEMENTS

We thank Catherine Glaux for her support in preliminary experiments, Benoît Moury for the R code to build the rarefaction curves, David Baltrus for his suggestions for core genome analysis, the technicians in the Plant Growth facilities at the INRA Plant Pathology Research Unit in Montfavet for their support in preparing plant materials, our team mates Luciana Parisi and Jay Ram Lamichhane for reading and improving the manuscript and to the colleagues who have provided us with strains for this work.

Chapter 3- A user's guide to a data base of the diversity of *Pseudomonas syringae* and its application to classifying strains in this phylogenetic complex.

REFERENCES

1. Van Hall CJJ (1902) Bijdrage tot de kennis der bacteriële plantenziekten. Inaug Diss, Amsterdam 8: 372.
2. Morris CE, Sands DC, Vanneste JL, Montarry J, Oakley B, 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: 00107–00110.
3. Young JM (2010) Taxonomy of *Pseudomonas syringae*. J Plant Pathol 92: S1.5–S1.14.
4. Parkinson N, Bryant R, Bew J, Elphinstone J (2011) Rapid phylogenetic identification of members of the *Pseudomonas syringae* species complex using the *rpoD* locus. Plant Pathol 60: 338–344.
5. Bodilis J, Barray S (2006) Molecular evolution of the major outer-membrane protein gene (*oprF*) of *Pseudomonas*. Microbiology 152: 1075–1088.
6. Mulet M, Lalucat J, García-Valdés E (2010) DNA sequence-based analysis of the *Pseudomonas* species. Environ Microbiol 12: 1513–1530.
7. Gardan L, Shafik H, Belouin S, Broch R, Grimont F, et al. (1999) DNA relatedness among the pathovars of *Pseudomonas syringae* and description of *Pseudomonas tremiae* sp. nov. and *Pseudomonas cannabina* sp. nov. (ex Sutic and Dowson 1959). Int J Syst Bacteriol 49: 469–478.
8. Bull CT, Clarke CR, Cai R, Vinatzer BA, Jardini TM, et al. (2011) Multilocus sequence typing of *Pseudomonas syringae sensu lato* confirms previously described genomospecies and permits rapid identification of *P. syringae* pv. *coriandricola* and *P. syringae* pv. *apii* causing bacterial leaf spot on parsley. Phytopathology 101: 847–858.
9. Billing E (1970) *Pseudomonas viridiflava* (Burkholder, 1930; Clara 1934). J Appl Bacteriol: 492–500.
10. Janse JD, Rossi P, Angelucci L, Scortichini M, Derks JHJ, et al. (1996) Reclassification of *Pseudomonas syringae* pv. *avellanae* as *Pseudomonas avellanae* spec. nov, the bacterium causing canker of hazelnut (*Corylus avellana* L.). Syst Appl Microbiol 19: 589–595.
11. Scortichini M, Marcelletti S, Ferrante P, Firrao G (2013) A Genomic redefinition of *Pseudomonas avellanae* species. PLoS One 8: e75794.

Chapter 3- A user's guide to a data base of the diversity of *Pseudomonas syringae* and its application to classifying strains in this phylogenetic complex.

12. Hildebrand D, Schroth M, Sand D (1988) *Pseudomonas*. Laboratory guide for identification of plant pathogenic bacteria. St. Paul, Minnesota, USA. pp. 60–80.
13. Misaghi I, Grogan RG (1969) Nutritional and biochemical comparisons of plant-pathogenic and saprophytic fluorescent pseudomonads. *Phytopathology* 59: 1436–1450.
14. Sawada H, Suzuki F, Matsuda I, Saitou N (1999) Phylogenetic analysis of *Pseudomonas syringae* pathovars suggests the horizontal gene transfer of *argK* and the evolutionary stability of *hrp* gene cluster. *J Mol Evol* 49: 627–644.
15. Sarkar SF, Guttman DS (2004) Evolution of the core genome of *Pseudomonas syringae*, a highly clonal , endemic plant pathogen. *Appl Env Microbiol* 70: 1999–2012.
16. Hwang MSH, Morgan RL, Sarkar SF, Wang PW, Guttman DS (2005) Phylogenetic characterization of virulence and resistance phenotypes of *Pseudomonas syringae*. *Appl Env Microbiol* 71: 5182–5191.
17. Morris CE, Kinkel LL, Xiao K, Prior P, Sands DC (2007) Surprising niche for the plant pathogen *Pseudomonas syringae*. *Infect Genet Evol* 7: 84–92.
18. Schaad NW (1979) Serological identification of plant pathogenic bacteria. *Annu Rev Phytopathol* 17: 23–47.
19. King EO, Ward MK, Raney DE (1954) Two simple media for the demonstration of pyocyanin and fluorescin. *J Lab Clin Med* 44: 301–307.
20. Morris CE, Sands DC, Vinatzer BA, Glaux C, Guilbaud C, et al. (2008) The life history of the plant pathogen *Pseudomonas syringae* is linked to the water cycle. *ISME J* 2: 321–334.
21. Xia X (2013) DAMBE5: A comprehensive software package for data analysis in molecular biology and evolution. *Mol Biol Evol*: mst064.
22. Sulthana S, Rajyaguru PI, Mittal P, Ray MK (2011) *rnr* gene from the antarctic bacterium *Pseudomonas syringae* Lz4W, encoding a psychrophilic RNase R. *Appl Environ Microbiol* 77: 7896–7904.
23. Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19: 1572–1574.
24. Almeida NF, Yan S, Cai R, Clarke CR, Morris CE, et al. (2010) PAMDB, a multilocus sequence typing and analysis database and website for plant-associated microbes. *Phytopathology* 100: 208–215.

Chapter 3- A user's guide to a data base of the diversity of *Pseudomonas syringae* and its application to classifying strains in this phylogenetic complex.

25. Baltrus DA, Nishimura MT, Romanchuk A, Chang JH, Mukhtar MS, et al. (2011) Dynamic evolution of pathogenicity revealed by sequencing and comparative genomics of 19 *Pseudomonas syringae* Isolates. PLoS Pathog 7: e1002132.
26. Gerhardt P, Murray RGE, Costilow RN, Nester EW, Wood WA, et al. (1981) Manual of methods for general bacteriology. American Society for Microbiology 1913, Street, NW, Washington DC, 20006.
27. Monteil CL, Guilbaud C, Glaux C, Lafolie F, Soubeyrand S, et al. (2011) Emigration of the plant pathogen *Pseudomonas syringae* from leaf litter contributes to its population dynamics in alpine snowpack. Environ Microbiol 22: 2099–2112.
28. Xu GW, Gross DC (1988) Evaluation of the role of syringomycin in plant pathogenesis by using Tn5 Mutants of *Pseudomonas syringae* pv. *syringae* defective in syringomycin production. Appl Environ Microbiol 54: 1345–1353.
29. Sawada H, Takeuchi T, Matsuda I (1997) Comparative analysis of *Pseudomonas syringae* pv. *actinidiae* and pv. *phaseolicola* based on phaseolotoxin-resistant ornithine carbamoyltransferase gene (*argK*) and 16S-23S rRNA intergenic spacer sequences. Appl Environ Microbiol 63: 282–288.
30. Bartoli C, Berge O, Monteil CL, Guilbaud C, Balestra GM, et al. (2014) The *Pseudomonas viridiflava* phylogroups in the *P. syringae* species complex are characterized by genetic variability and phenotypic plasticity of pathogenicity-related traits. Environ Microbiol. In press.
31. Hockett KL, Nishimura MT, Karlsrud E, Dougherty K, Baltrus DA (2013) Interactions between genome architecture and virulence genes in *Pseudomonas syringae*, strain CC1557 as a model. bioRxiv. Doi: 10.1101/000869.
32. O'Brien HE, Thakur S, Gong Y, Fung P, Zhang J, et al. (2012) Extensive remodeling of the *Pseudomonas syringae* pv. *avellanae* type III secretome associated with two independent host shifts onto hazelnut. BMC Microbiol 12: 141.
33. Monteil CL, Cai R, Liu H, Mehan Llontop ME, Leman S, et al. (2013) Nonagricultural reservoirs contribute to emergence and evolution of *Pseudomonas syringae* crop pathogens. New Phytol 199: 800–811.
34. Mucyn TS, Yourstone S, Lind AL, Biswas S, Nishimura MT, et al. (2014) Variable suites of non-effector genes are co-regulated in the type III secretion virulence Regulon across the *Pseudomonas syringae* phylogeny. PLoS Pathog 10: e1003807.

Chapter 3- A user's guide to a data base of the diversity of *Pseudomonas syringae* and its application to classifying strains in this phylogenetic complex.

35. Clarke CR, Cai R, Studholme DJ, Guttman D, Vinatzer BA (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.
36. Orser C, Staskawicz BJ, Panopoulos NJ, Dahlbeck D, Lindow SE (1985) Cloning and expression of bacterial ice nucleation genes in *Escherichia coli*. *J Bacteriol* 164: 359–366.
37. Demba Diallo M, Monteil CL, Vinatzer BA, Clarke CR, Glaux 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. *ISME J* 6: 1325–1335.
38. Monteil C, Bardin M, Morris CE (2014) Features of air masses associated with the deposition of *Pseudomonas syringae* and *Botrytis cinerea* by rain and snowfall. *ISME J* in press.
39. Green S, Studholme DJ, Laue BE, Dorati F, Lovell H, 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.
40. Rodríguez-Palenzuela P, Matas IM, Murillo J, López-Solanilla E, Bardaji L, et al. (2010) Annotation and overview of the *Pseudomonas savastanoi* pv. *savastanoi* NCPPB 3335 draft genome reveals the virulence gene complement of a tumour-inducing pathogen of woody hosts. *Environ Microbiol* 12: 1604–1620.
41. Joly M, Attard E, Sancelme M, Deguillaume L, Guilbaud C, et al. (2013) Ice nucleation activity of bacteria isolated from cloud water. *Atmos Environ* 70: 392–400.
42. Jakob K, Goss EM, Araki H, Van T, Kreitman M, et al. (2002) *Pseudomonas viridiflava* and *P. syringae*-natural pathogens of *Arabidopsis thaliana*. *Mol Plant Microbe Interact* 15: 1195–1203.
43. Miller CM, Miller RV, Garton-Kenny D, Redgrave B, Sears J, et al. (1998) Ecomycins, unique antimycotics from *Pseudomonas viridiflava*. *J Appl Microbiol* 84: 937–944.
44. Trantas, Emmanouil , Sarris PF, Mpalantinaki EE, Pentari MG, Ververidis FN, Goumas DE (2013) A new genomovar of *Pseudomonas cichorii*, a causal agent of tomato pith necrosis. *Eur J Plant Pathol*.

Chapter 3- A user's guide to a data base of the diversity of *Pseudomonas syringae* and its application to classifying strains in this phylogenetic complex.

45. Cottyn B, Baeyen S, Pauwelyn E, Verbaendert I, De Vos P, et al. (2011) Development of a real-time PCR assay for *Pseudomonas cichorii*, the causal agent of midrib rot in greenhouse-grown lettuce, and its detection in irrigating water. *Plant Pathol* 60: 453–461.
46. Humphrey PT, Nguyen TT, Villalobos MM, Whiteman NK (2014) Diversity and abundance of phyllosphere bacteria are linked to insect herbivory. *Mol Ecol* 23: 1497–1515.
47. Morris CE, Glaux C, Latour X, Gardan L, Samson R, et al. (2000) The relationship of host range, physiology, and genotype to virulence on cantaloupe in *Pseudomonas syringae* from cantaloupe blight epidemics in France. *Phytopathology* 90: 636–646.
48. Marchi G, Viti C, Giovannetti L, Surico G (2005) Spread of levan-positive populations of *Pseudomonas savastanoi* pv. *savastanoi*, the causal agent of olive knot, in central Italy. *Eur J Plant Pathol* 112: 101–112.
49. Natalini, Rossi MP, Barionovi D, Scortichini M (2006) Genetic and pathogenic diversity of *Pseudomonas syringae* pv. *syringae* isolates associated with bud necrosis and leaf spot of pear in a single orchard. *J Plant Pathol* 88: 219–223.
50. Lamichhane JR, Varvaro L, Parisi L, Audergon J-M, Morris CE (2014) Disease and frost damage of woody plants caused by *Pseudomonas syringae*: seeing the forest for the trees. *Adv Agron*.
51. Lamichhane JR, Varvaro L (2013) A new medium for the detection of fluorescent pigment production by pseudomonads. *Plant Pathol* 62: 624–632.
52. Wilson M, Lindow SE (1994) Coexistence among epiphytic bacterial populations mediated through nutritional resource partitioning. *Appl Environ Microbiol* 60: 4468–4477.
53. Marakeby H, Badr E, Torkey H, Song Y, Leman S, et al. (2014) A system to automatically classify and name any individual genome-sequenced organism independently of current biological classification and nomenclature. *PLoS One* 9: e89142.

Chapter 3- A user's guide to a data base of the diversity of *Pseudomonas syringae* and its application to classifying strains in this phylogenetic complex.

Table 1: Mean genetic distances within (boldface values) and between phylogroups

Number of strains	<i>P. syringae</i> Phylogroup	Mean genetic distance ^a												
		1	2	3	4	5	6	7	8	9	10	11	12	13
30	1	0.02												
84	2	0.10	0.02											
14	3	0.09	0.07	0.02										
5	4	0.10	0.10	0.09	0.00									
2	5	0.10	0.11	0.11	0.11	0.00								
1	6	0.10	0.09	0.08	0.11	0.11	0.00							
12	7	0.11	0.13	0.12	0.12	0.12	0.12	0.03						
4	8	0.13	0.14	0.13	0.14	0.13	0.13	0.07	0.00					
11	9	0.12	0.13	0.13	0.14	0.13	0.13	0.08	0.08	0.02				
38	10	0.11	0.11	0.11	0.11	0.07	0.12	0.12	0.14	0.13	0.03			
3	11	0.17	0.17	0.16	0.17	0.17	0.17	0.17	0.17	0.16	0.17	0.06		
2	12	0.16	0.16	0.16	0.17	0.16	0.16	0.18	0.19	0.17	0.17	0.19	0.05	
10	13	0.17	0.18	0.18	0.19	0.17	0.17	0.18	0.18	0.18	0.17	0.20	0.13	0.03
2	<i>P. graminis</i>	0.21	0.21	0.21	0.22	0.21	0.20	0.21	0.22	0.22	0.21	0.23	0.22	0.22
1	<i>P. rhizosphaerae</i>	0.21	0.21	0.21	0.21	0.22	0.23	0.20	0.19	0.22	0.22	0.23	0.22	0.23
1	<i>P. protegens</i>	0.20	0.19	0.19	0.20	0.19	0.19	0.19	0.18	0.20	0.20	0.20	0.21	0.21
1	<i>P. aeruginosa</i>	0.34	0.22	0.35	0.35	0.35	0.35	0.33	0.34	0.34	0.35	0.35	0.35	0.35

^aDistances based on 1854 bp sequences of four housekeeping genes are those used for constructing the MLST tree of the set of 216 strains in Figure 1.

Chapter 3- A user's guide to a data base of the diversity of *Pseudomonas syringae* and its application to classifying strains in this phylogenetic complex.

Table 2 : Phenotypic characterization of 764 strains representing the genetic diversity in the *P. syringae* complex. Values are the percent of *P. syringae* strains giving positive reactions for the different phenotypes. Arginine dihydrolase production was negative for all strains.

Phylogenetic affiliation	Number of strains	Levan +	Oxidase +	Potato soft rot	HR on tobacco+	Fluorescence on KB medium +	Esculin degradation +	Sucrose utilization+	D(-) Tartrate utilization +	I NA ^b +	Broad host-range toxin ^c +	Pathogenicity ^d +	Avirulent ^e	Mean sdisease severity $\geq 2.0^f$
Total	764	41	0	10	73	98	97	73	32	65	51	28	53	12
PG^a 01	89	39	0	0	92	85	92	91	58	33	2	16	60	1
PG 01a	49	45	0	0	90	100	100	88	78	24	4	14	53	2
PG 01b	40	33	0	0	95	68	83	95	35	43	0	18	68	0
PG 02	323	50	0	0	72	100	100	95	10	85	90	43	41	25
PG 02a	3	0	0	0	100	100	100	100	33	67	67	33	33	0
PG 02b	140	67	0	0	98	100	99	95	4	90	87	44	30	19
PG 02c	91	7	0	0	7	100	100	93	9	68	96	0	89	0
PG 02d	87	70	0	0	100	100	100	98	20	93	91	89	6	62
PG 02e	2	0	0	0	50	100	100	100	0	100	100	0	100	0
PG 03	10	70	0	0	80	70	20	90	0	20	0	0	90	0
PG 04	6	83	0	0	100	100	100	100	0	100	17	0	67	0
PG 05	4	50	0	0	75	75	25	75	25	50	0	25	50	0
PG 07	81	46	0	98	62	100	100	9	96	33	0	47	37	11
PG 07a	79	46	0	97	61	100	100	9	96	33	0	48	37	11
PG 07b	2	50	0	100	100	100	100	0	100	50	0	0	50	0
PG 08	5	80	0	40	100	100	80	0	100	0	100	0	80	0
PG 09	28	11	0	0	68	100	100	14	32	4	0	0	71	0
PG 09a	23	9	0	0	65	100	100	13	26	4	0	0	70	0
PG 09b	4	25	0	0	75	100	100	25	50	0	0	0	75	0
PG 09c	1	0	0	0	100	100	100	0	100	0	0	0	100	0
PG 10	167	34	0	0	98	100	100	84	11	94	51	10	66	3
PG 10a	71	8	0	0	97	100	100	82	23	94	1	7	80	1
PG 10b	77	58	0	0	100	100	100	86	3	94	91	1	62	0
PG 10c	1	0	0	0	100	100	100	0	0	100	0	0	100	0
PG 10d	8	13	0	0	100	100	100	75	0	100	75	50	25	13
PG 10e	8	50	0	0	88	100	100	100	0	100	100	75	13	38
PG 10f	1	100	0	0	100	100	100	100	100	100	100	0	100	0
PG 10g	1	0	0	0	0	100	100	100	0	0	0	0	100	0
PG 11	3	0	100	0	67	100	100	0	33	0	100	67	33	0
PG 12	2	0	0	0	0	100	100	0	50	0	0	0	100	0
PG 12a	1	0	0	0	0	100	100	0	0	0	0	0	100	0
PG 12b	1	0	0	0	0	100	100	0	100	0	0	0	100	0
PG 13	46	0	0	0	0	100	100	0	100	0	0	0	78	0
PG 13a	44	0	0	0	0	100	100	0	100	0	0	0	77	0
PG 13b	2	0	0	0	0	100	100	0	100	0	0	0	100	0

^aPG= *P. syringae* phylogroup.

^bINA = ice nucleation activity of at least 10^6 cells at $\geq 8^\circ$ C.

Chapter 3- A user's guide to a data base of the diversity of *Pseudomonas syringae* and its application to classifying strains in this phylogenetic complex.

^cProduction of a broad host range toxin was evaluated with the test habitually used to reveal syringomycin-like toxins based on the capacity to produce an inhibition zone of growth of *Geotricum candidum*.

^dStrains were considered to be pathogenic on the cantaloupe indicator plant if at least half (6/12) of the seedlings showed compatible reactions.

^eAvirulent strains did not induce any disease reaction on cantaloupe seedlings.

^fFrequency of strains for which the mean disease severity on cantaloupe seedlings was ≥ 2.0

Chapter 3- A user's guide to a data base of the diversity of *Pseudomonas syringae* and its application to classifying strains in this phylogenetic complex.

Table 3. Highlight features of the phylogroups of *Pseudomonas syringae* based on the profiles of strains characterized to date. These features are a general summary of the phenotypic and genotypic profiles garnered in this study as well as in other reports cited throughout the text above.

Phylogroup	Known habitats ^a	T3SS organization ^b	n° of T3SS effectors	n° of HrpL regulons	Broad host range toxins ^d	Soft rot induction	Putative Host range ^e	INA ^f at $\geq -8^{\circ}$ C
1	Ubiquitous	Canonical	+++	+++	+	-	Narrow	++
2	Ubiquitous	Canonical & atypical	+	++	+++	-	Wide for some clades	+++
3	Mostly plants	Canonical	++	++	-	-	Host specific or narrow	++
4	Ubiquitous	Canonical	++	+++	++	-	Host specific or narrow	+++
5	Plants	Canonical	++	++	-	-	Intermediate	++
6	Plants	Canonical	++	+	ND	ND	ND	ND
7	Ubiquitous	Canonical & atypical	++	+	-	+++	Wide	++
8	Ubiquitous	Canonical & atypical	++	ND	++	+++	Narrow	-
9	Environment	ND ^c	ND	+		-	Null or very narrow	+
10	Environment	Canonical	+	+	+++	-	Wide for some clades	+++
11	Ubiquitous	Atypical	ND	ND	++	-	Wide	-
12	Environment	ND	ND	ND	-	-	Null or very narrow	-
13	Environment	Atypical	++	ND	-	-	Null or very narrow	-

^aRepresentatives of some phylogroups have been found associated only with plants, or only in non-agricultural (environmental) habitats or are ubiquitously present in all or nearly all habitats investigated to date.

^bBoth canonical and non-canonical (atypical) T3SS are found in some phylogroups, but they have not been found to co-exist in the same strain.

^cThis property has not been described to date.

^dThe production of broad host range toxins is based on the results of antibiosis tests reported in this work.

^eAggressiveness on cantaloupe seedlings was used as a proxy for host range as described previously [17]. Here, host range concerns the number of plant species on which disease symptoms are caused. The description presented for each phylogroup is relative to the other phylogroups and is based on the results of our analyses here. The range of epiphytic plants that can be colonized asymptotically can be much larger than the host range for disease.

^fIce nucleation activity

Chapter 3- A user's guide to a data base of the diversity of *Pseudomonas syringae* and its application to classifying strains in this phylogenetic complex.

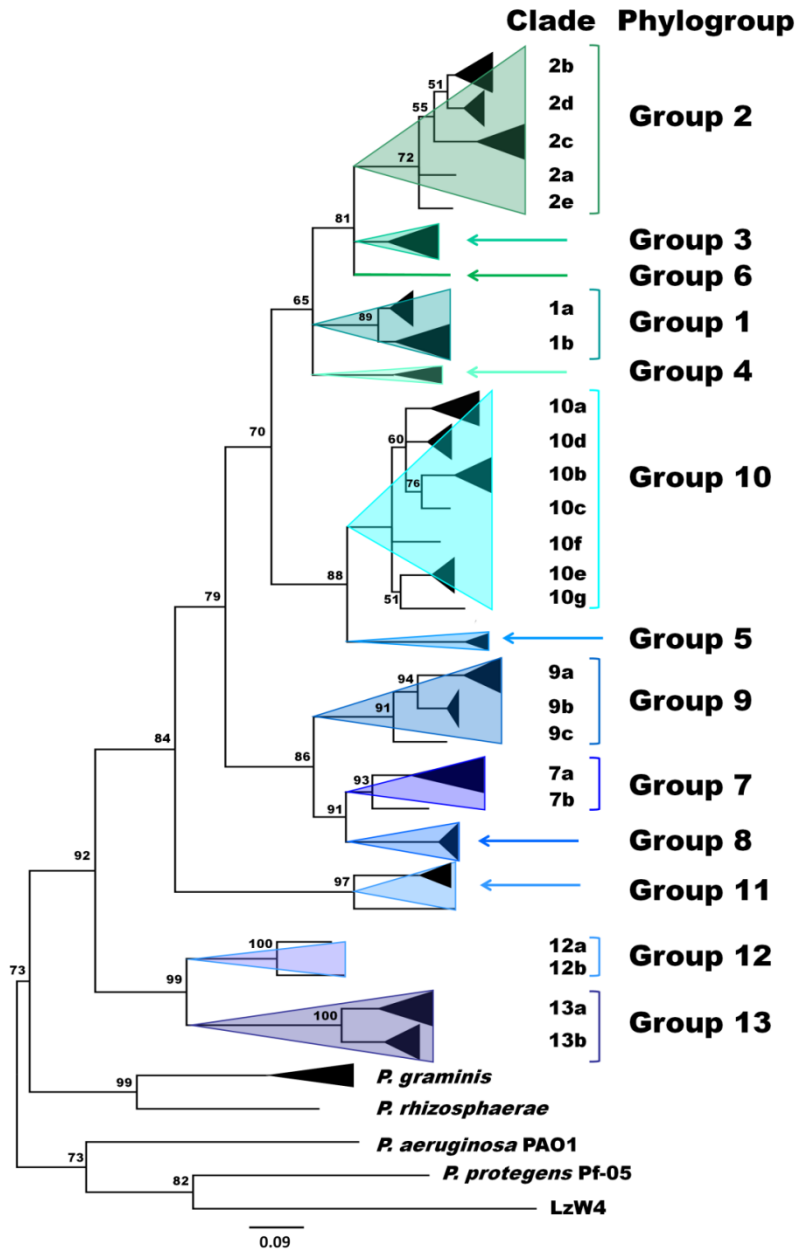


Figure 1. Bayesian tree constructed on the concatenated sequences *cts*, *gyrB*, *gapA* and *rpoD* of 216 *P. syringae* strains.

Bootstrap values are showed at each node. Strain taxa were compressed and clade and phylogroup names are indicated (see the expanded tree with strain names in Fig. S1). Phylogroups from 1 to 7 were already reported in Parkinson et al., [4], phylogroup 8 in [30], phylogroups 9, 10, and 13 were described with other names by Morris and coworkers [2] (see Table S2 for name correspondence), phylogroup 11 corresponds to *P. cichorii* strains and phylogroup 12 was not described previously. The tree was rooted on *P. aeruginosa* PAO1.

Chapter 3- A user’s guide to a data base of the diversity of *Pseudomonas syringae* and its application to classifying strains in this phylogenetic complex.

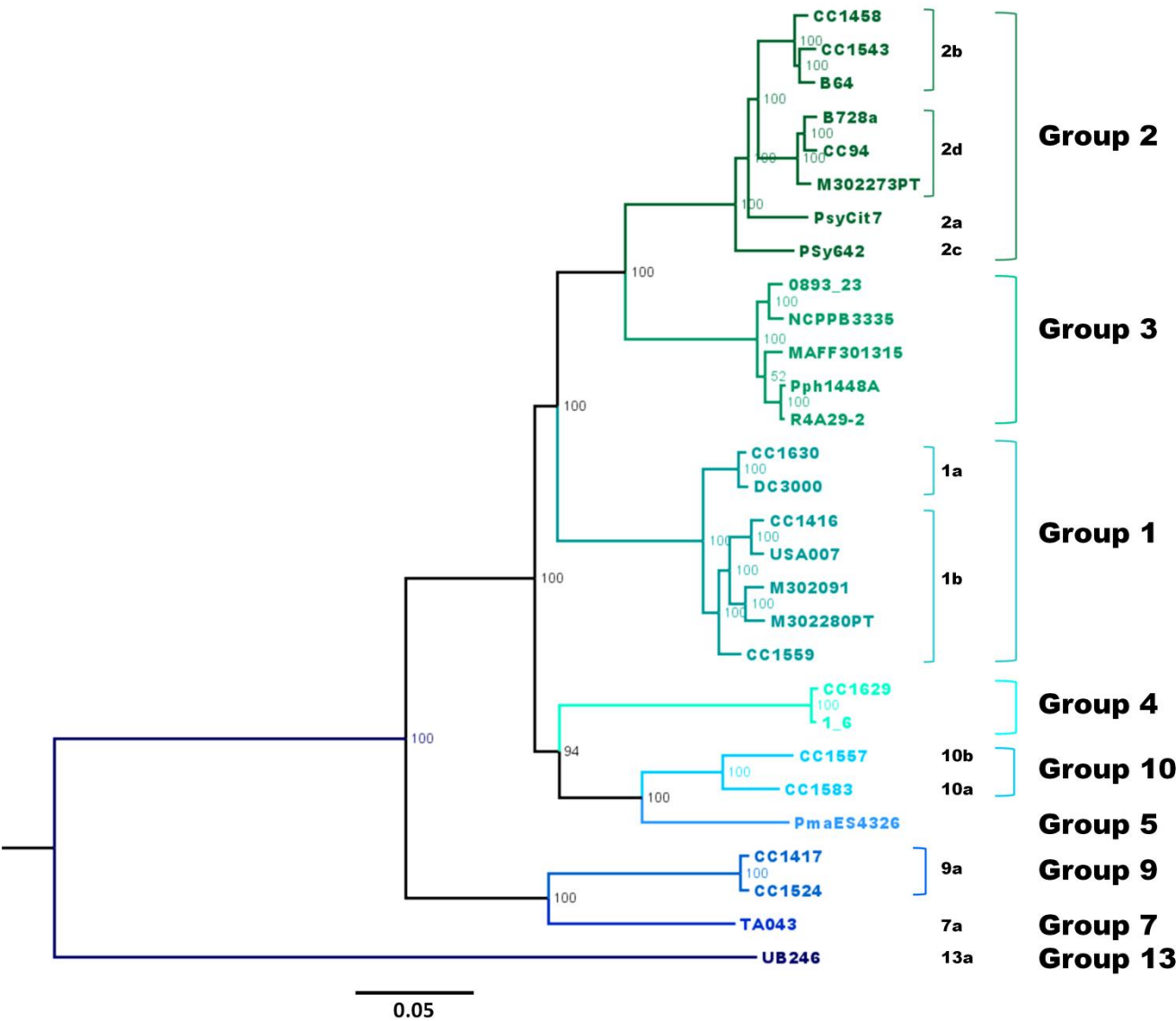


Figure 2. Bayesian phylogeny of the core genome of 29 *P. syringae* strains. An un-rooted tree was constructed on 107 open reading frames (64,000 bp) common to 29 *P. syringae* strains. Bootstrap values are indicated at each node and strain names are indicated at tree branches. Phylogroup and clade names are also indicated in the tree.

Chapter 3- A user's guide to a data base of the diversity of *Pseudomonas syringae* and its application to classifying strains in this phylogenetic complex.

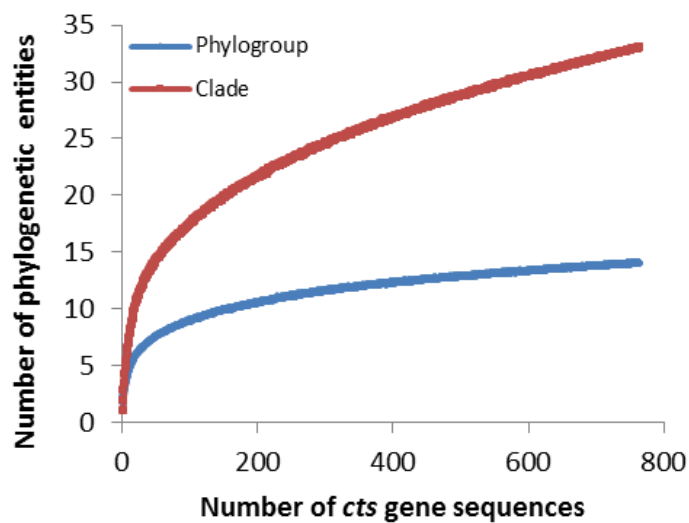


Figure 4. Rarefaction analysis of the *cts* gene sequences from strains of *P. syringae* at both phylogroup (blue curve) and clade (red curve) levels.

Chapter 3- A user's guide to a data base of the diversity of *Pseudomonas syringae* and its application to classifying strains in this phylogenetic complex.

SUPPORTING INFORMATION

Methods S1: Validation of the *cts* gene as *P. syringae* tool classification

The validation of the *cts* gene as *P. syringae* tool classification was made by using the set of 216 strains as follows: i) the *cts* tree was compared with the MLST tree for the same 216 strains, showing that the tree topologies were the same except for some clades in phylogroups 2 and 10 (Fig. S1a, Fig. S1b), ii) genetic distance criteria were established for finding the phylogroups and clades already delineated in the MLST analysis; the maximum of 4 % was used for *cts*-phylogroups (previously 5 % for MLST-phylogroups) and 2 % for *cts*-clades (previously 2.3 % for MLST-clades) (Table S4), iii) for each of the 216 strains, a new phylogroup and clade affiliation was performed based on the *cts* matrix of distances by looking at the minimum distance of each strain with the remaining 215 strains, affiliation of one strain being considered the same as its closest relative; this step allowed us to confirm the affiliation of each strain to its phylogroup and clade by the use of the minimum distance of *cts* sequences and to validate the genomic distances of 4 and 2 % used (Table S4). By applying this method we found some exceptions at clade level, in particular strain K93001 (*P. s. pv theae*) affiliated to MLST-clade 1b was equidistant from strains of *cts*-clade 1a and 1b, and the strains USA0035, CC1586 and CCE0153 that were respectively grouped into 2d, 10b and 10d/10a *cts*-clades when they formed respectively the single-strain clades 2e, 10c and 10f in the MLST analysis (Table S3 and S4). This illustrates that some strains need more than one housekeeping gene to be robustly classified.

Chapter 4

**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) DOI: 10.1111/1462-2920.12433

Chapter 4- The *Pseudomonas viridiflava* phylogroups in the *P. syringae* species complex are characterized by genetic variability and phenotypic plasticity of pathogenicity related traits

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 Varvaro¹ Corbin Jones,³ Jeffery L. Dangl,⁴ David A. Baltrus,⁵ David C. Sands⁶ and Cindy E. Morris^{2,6*}

¹Department of Science and Technology for Agriculture, Forestry, Nature and Energy (DAFNE), Tuscia University, Viterbo, Italy.

²UR0407 Pathologie Végétale, INRA, Montfavet, France.

³Department of Biology, Carolina Center for Genome Sciences and ⁴Department of Biology, Howard Hughes Medical Institute, University of North Carolina, Chapel Hill, NC 29599,

⁵School of Plant Sciences, University of Arizona, Tucson, AZ 85721, and ⁶Department Plant Sciences and Plant Pathology, Montana State University, Bozeman, MT 59717-3150, USA.

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*

Chapter 4- The *Pseudomonas viridiflava* phylogroups in the *P. syringae* species complex are characterized by genetic variability and phenotypic plasticity of pathogenicity related traits

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.

Received 29 September, 2013; accepted 13 February, 2014. *For correspondence. E-mail cindy.morris@avignon.inra.fr; Tel. (+33) (0) 432722841; Fax (+33) (0) 432722841.

Chapter 4- The *Pseudomonas viridiflava* phylogroups in the *P. syringae* species complex are characterized by genetic variability and phenotypic plasticity of pathogenicity related traits

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; 2013; O'Brien *et al.*, 2011). 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 (Araki *et al.*, 2006; Clarke *et al.*, 2010; Baltrus *et al.*, 2011; 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 has 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

Chapter 4- The *Pseudomonas viridiflava* phylogroups in the *P. syringae* species complex are characterized by genetic variability and phenotypic plasticity of pathogenicity related traits

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 with *P. viridiflava* strains from plants capable of causing soft-rot on potato.

Strains related to phylogroup 7 lacks a canonical T3SS, but they can be highly aggressive pathogens (Demba Diallo *et al.*, 2012). *Pseudomonas 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 10 kb-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 the T-PAI and S-PAI

Chapter 4- The *Pseudomonas viridiflava* phylogroups in the *P. syringae* species complex are characterized by genetic variability and phenotypic plasticity of pathogenicity related traits

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

Pseudomonas viridiflava is composed of two distinct phylogroups capable of inducing 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 phylogenetic 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*)

Chapter 4- The *Pseudomonas viridiflava* phylogroups in the *P. syringae* species complex are characterized by genetic variability and phenotypic plasticity of pathogenicity related traits

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 four 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 (Supporting Information 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.

Chapter 4- The *Pseudomonas viridiflava* phylogroups in the *P. syringae* species complex are characterized by genetic variability and phenotypic plasticity of pathogenicity related traits

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
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)

Chapter 4- The *Pseudomonas viridiflava* phylogroups in the *P. syringae* species complex are characterized by genetic variability and phenotypic plasticity of pathogenicity related traits

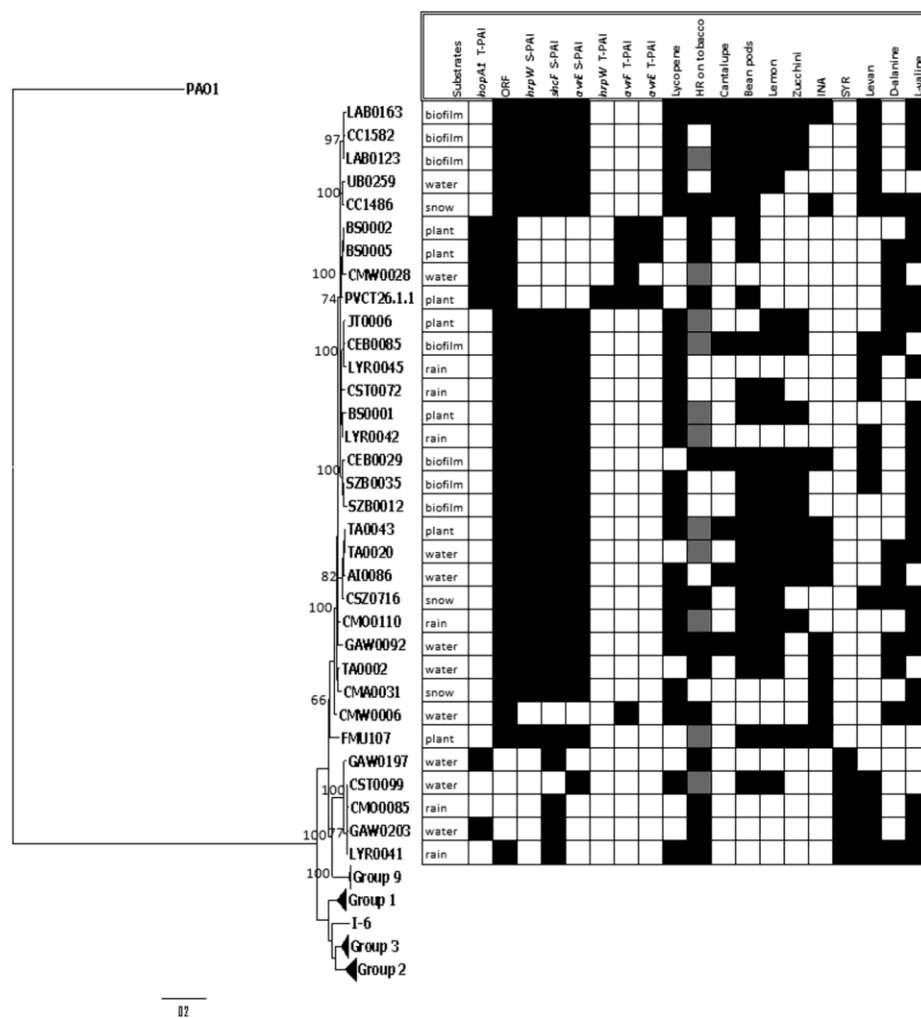


Fig. 1. Bayesian phylogenetic tree constructed with the concatenated housekeeping genes *cts*, *gapA*, *gyrB* and *rpoD* (1852 bp). The tree was rooted on *P. aeruginosa* (PAO). Posterior probabilities are indicated at each node. Names of the branches correspond to the strains, except for the branches indicated as phylogroups reported in previous work (Parkinson *et al.*, 2011). The substrates are indicated close to the strain names. The black squares indicate a positive reaction or presence of the genes and the white squares stand for a negative reaction or absence of the genes. Grey squares indicate a variable reaction or non-stable phase variants. The two phylogroups, 7 and 8 are separated by a mean distance of 5.5% sequence difference. Strains in phylogroup 8 are GAW0197, CST0099, CMO0085, GAW0203 and LYR0041.

Abbreviations: ORF = open reading frames found in the *EEL* of T-PAI strains and encoding for a lipoprotein and a monooxygenase involved in antibiotics biosynthesis. INA = ice nucleation activity. The strains were considered positive when at least two of three drops containing 10^6 cells per drop froze at temperatures warmer than -9°C . SYR = presence of a syringomycin-like toxin based on inhibition of *Geotricum candidum*.

Chapter 4- The *Pseudomonas viridiflava* phylogroups in the *P. syringae* species complex are characterized by genetic variability and phenotypic plasticity of pathogenicity related traits

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 per cent of the strains produced a yellowish levan capsule after 3 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 King's medium B. The mucoid yellowish colonies were 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 (Supporting Information 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 (Supporting Information 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)

Chapter 4- The *Pseudomonas viridiflava* phylogroups in the *P. syringae* species complex are characterized by genetic variability and phenotypic plasticity of pathogenicity related traits

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 versus 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 versus 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.*, 2007; 2008).

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, 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. Muroid colonies (M) usually appeared after 2 days of incubation and large, flat transparent non-muroid colonies (NM) were visible after 4 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). Muroid variants consistently induced soft rot on potato, liquefied gelatin and caused necrotic lesions on bean pods, whereas non-muroid variants did not (Table 2, Supporting Information 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

Chapter 4- The *Pseudomonas viridiflava* phylogroups in the *P. syringae* species complex are characterized by genetic variability and phenotypic plasticity of pathogenicity related traits

strains PV841/09, LAB0163 and CC1582, significantly greater severity and incidence of the disease were observed for the M variant clones compared with the NM variants ($P \leq 0.05$). Neither M nor NM variants of strains PVBH nor BS0005 caused marked disease on cantaloupe seedlings (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 antibiotic biosynthesis monooxygenase (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 genome analysis of two *P. viridiflava* strains from non-agricultural habitats (TA0043, GenBank accession AVDV000000000; and CC1582, GenBank accession AVDW000000000) and one other publically available assembly (UASWS0038, GenBank accession number NZ AMQP000000000). 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 monooxygenase genes were present in almost all the strains analyzed (90%) even in strains lacking *hopA1(T)* and other T-PAI alleles (Fig. 1, Supporting Information 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

Chapter 4- The *Pseudomonas viridiflava* phylogroups in the *P. syringae* species complex are characterized by genetic variability and phenotypic plasticity of pathogenicity related traits

comparison with the *EEL* of strains PsyB728A and PtoDC3000 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 monooxygenase sequences (Fig. 4) showed evidence of horizontal transfer for these loci compared with the housekeeping phylogeny (Fig. 1), suggesting that they were potentially acquired with the associated PAIs (Fig. 4). As observed by Araki and colleagues (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 (Supporting Information Fig. S2). Forty-nine strains had the three alleles typical of the S-PAI but had lipoprotein and monooxygenase as well. The remaining six strains had insufficient genes, according to results of PCR, to classify the T3SS according to the criteria of Araki and colleagues (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* (Supporting Information 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).

Chapter 4- The *Pseudomonas viridiflava* phylogroups in the *P. syringae* species complex are characterized by genetic variability and phenotypic plasticity of pathogenicity related traits

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.

Stable variants ^a	Bean pods ^b	Potato rot ^b	Gelatin hydrolysis ^b	HR on tobacco	Utilisation, as sole carbon sources, of				
					L–Valine	D–Tartrate	D–Alanine	Cu resistanc ^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: 0.64 mM and 1.12 mM according with the *P. syringae* tolerance curve (Andersen *et al.*, 1991). Results were the same for both copper concentrations is each variant.

Chapter 4- The *Pseudomonas viridiflava* phylogroups in the *P. syringae* species complex are characterized by genetic variability and phenotypic plasticity of pathogenicity related traits

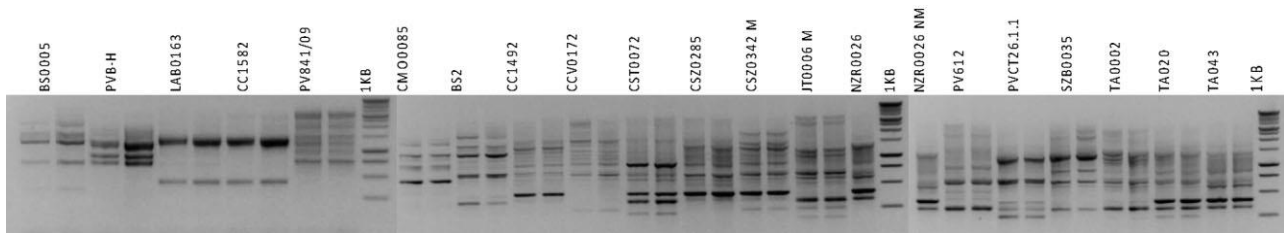


Fig. 2. BOX PCR profiles of mucoid and non-mucoid variants. The first and second lanes of each strain correspond to the mucoid and non-mucoid variants respectively.

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

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

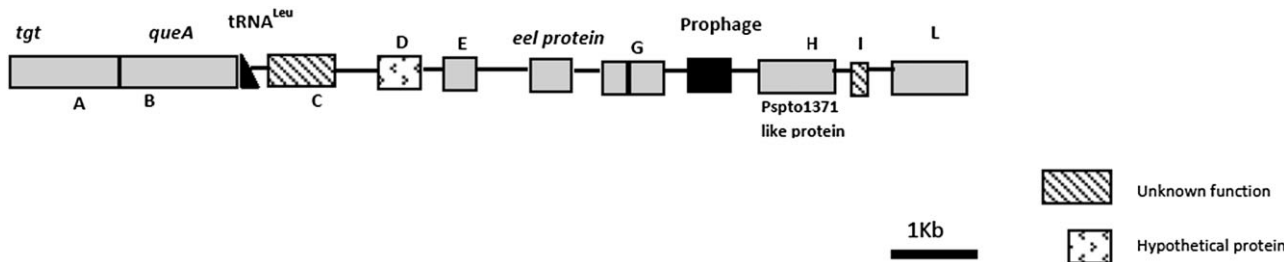


Fig. 3. Structure of the locus resembling an *EEL* of *P. viridiflava* TA0043 and CC1582. The *EEL* was identified in draft genomes of strains TA0043 and CC1582. Letters refer to the putative protein function: A = *tgt* tRNA-guanine transglycosylase, queusine-34-forming, B = Queuine synthetase (*queA*), C = Laba-A-like N1 domain protein (conserved protein found in different bacteria with unknown function), D = Hypothetical protein found only in *P. viridiflava*, E = Pstpo1411 like protein, F = *eel* 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

Chapter 4- The *Pseudomonas viridiflava* phylogroups in the *P. syringae* species complex are characterized by genetic variability and phenotypic plasticity of pathogenicity related traits

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 15 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, 10 strains did not cause disease on cantaloupe or lesions on fruits, and among these 10 strains, five had the alleles commonly present in the S-PAI and five had apparently incomplete T3SS based on PCR results (Supporting Information Fig. S2). Contrary to the observation of Jakob and colleagues (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 *avrET-PAI* or *avrES-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%

Chapter 4- The *Pseudomonas viridiflava* phylogroups in the *P. syringae* species complex are characterized by genetic variability and phenotypic plasticity of pathogenicity related traits

(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 previous 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 (Table S1), where it is correlated with pathogenicity (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 de- grade pectin has probably allowed *P. viridiflava* and *P. carotovorum* to simplify their TTE repertoires. 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 with 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

Chapter 4- The *Pseudomonas viridiflava* phylogroups in the *P. syringae* species complex are characterized by genetic variability and phenotypic plasticity of pathogenicity related traits

presence of two different populations having seemingly the same ecological niche, but different T3SS still need 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 lifestyles of the bacterium are reflected by its ubiquity in the environment. Recently, Selezska and colleagues (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. First, the exopolysaccharide may increase tolerance to plant defence mechanisms. Second, 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, 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, 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.

Chapter 4- The *Pseudomonas viridiflava* phylogroups in the *P. syringae* species complex are characterized by genetic variability and phenotypic plasticity of pathogenicity related traits

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* contains *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 behaviours 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

Chapter 4- The *Pseudomonas viridiflava* phylogroups in the *P. syringae* species complex are characterized by genetic variability and phenotypic plasticity of pathogenicity related traits

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

Chapter 4- The *Pseudomonas viridiflava* phylogroups in the *P. syringae* species complex are characterized by genetic variability and phenotypic plasticity of pathogenicity related traits

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 lifestyles 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 *shcFS-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 monooxygenase and lipoprotein couple (Supporting Information 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 *shcFS-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 *shcFS-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; Deziel *et al.*, 2001; Rau *et al.*, 2010).

Chapter 4- The *Pseudomonas viridiflava* phylogroups in the *P. syringae* species complex are characterized by genetic variability and phenotypic plasticity of pathogenicity related traits

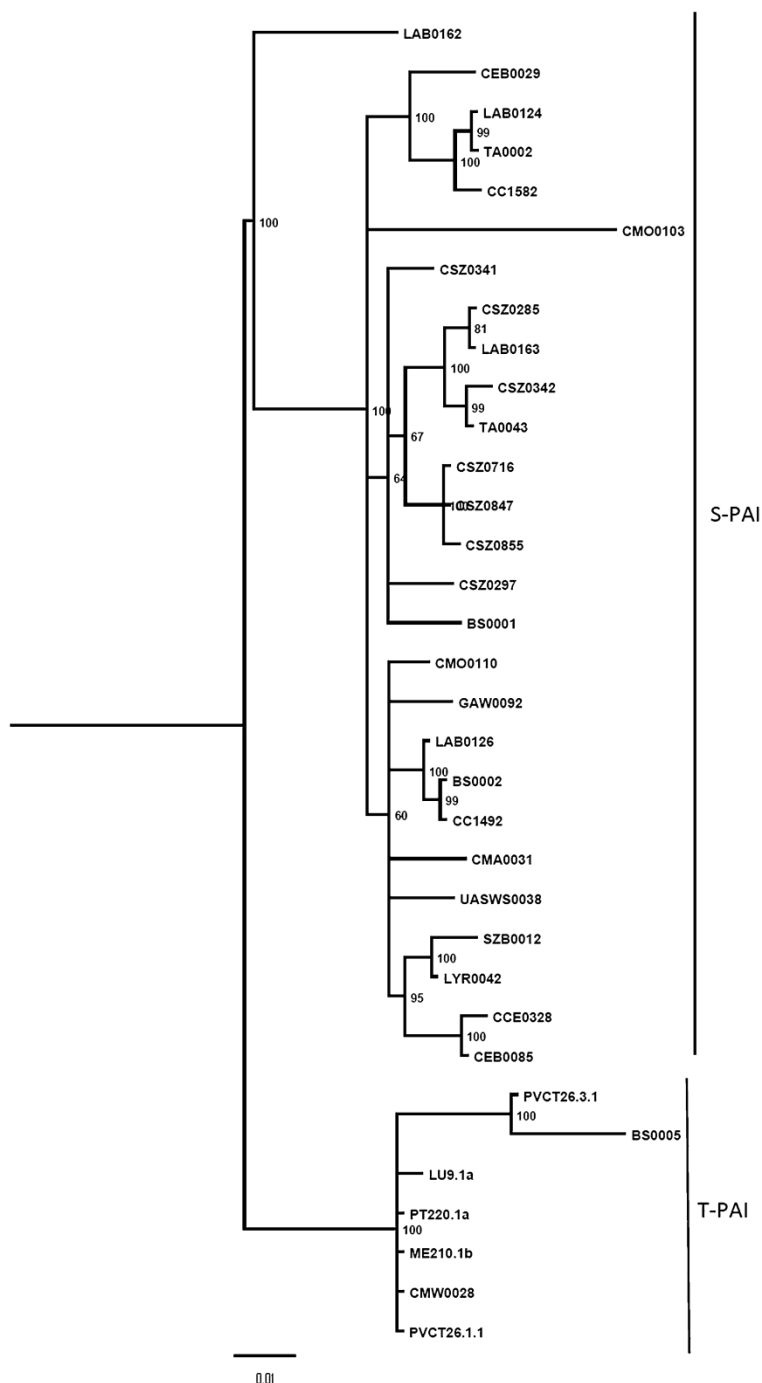


Fig. 4. Tree based on the lipoprotein and monooxygenase genes found in the putative *EEL* of the S-PAI strains and in the *EEL* of the T-PAI strains. The Bayesian method was employed to construct the tree. Posterior probabilities are indicated at each node. Sequences for the LU9.1a, PT220.1a, ME210.1b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

Chapter 4- The *Pseudomonas viridiflava* phylogroups in the *P. syringae* species complex are characterized by genetic variability and phenotypic plasticity of pathogenicity related traits

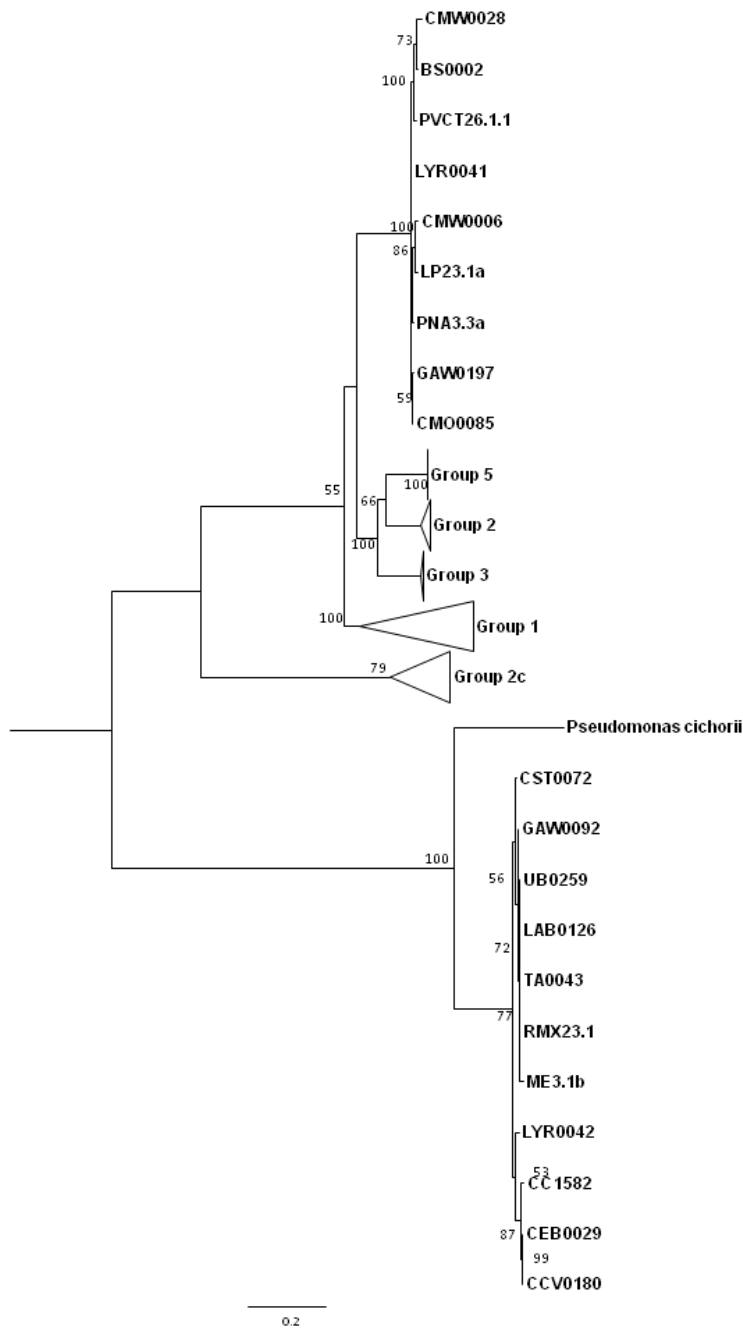


Fig. 5. Un-rooted Bayesian phylogenetic tree constructed with the *hrcC* gene sequences. T- and S-PAI strains are delimited with black bars. Sequences for LP23.1a, PNA3.3a, RMX23.1 and M3.1 b were obtained from GeneBank. Accession numbers are respectively: AY597277.1, AY597278.1, AY597282.1, AY597281.1

Chapter 4- The *Pseudomonas viridiflava* phylogroups in the *P. syringae* species complex are characterized by genetic variability and phenotypic plasticity of pathogenicity related traits

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 Supporting Information 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 Supporting Information 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 Supporting Information Methods S1 and Table S2.

Phylogenetic analyses based on housekeeping genes and on T3SS genes

A set of strains was 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 and 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 500 000 generations. Analysis was concluded when the standard deviation of split

Chapter 4- The *Pseudomonas viridiflava* phylogroups in the *P. syringae* species complex are characterized by genetic variability and phenotypic plasticity of pathogenicity related traits

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>). Trees 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 three 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 h. 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 six 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.

Chapter 4- The *Pseudomonas viridiflava* phylogroups in the *P. syringae* species complex are characterized by genetic variability and phenotypic plasticity of pathogenicity related traits

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 Policies (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 the Institut National de Recherche Agronomique in Avignon 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 tri- partite mosaic structure composed of a cluster of type III secretion genes bounded by exchangeable effector and conserved effector loci that contribute to parasitic fitness and pathogenicity. *Proc Natl Acad Sci USA* **97**: 4856–4861.
- Andersen, G.L., Menkissoglou, O., and Lindow, S.E. (1991) Occurrence and properties of copper-tolerant strains of *Pseudomonas syringae* isolated from fruit trees in California. *Phytopathology* **81**: 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.
- Araki, H., Innan, H., Kreitman, M., and Bergelson, J. (2007) Molecular evolution of pathogenicity-island genes in *Pseudomonas viridiflava*. *Genetics* **177**: 1031–1041.
- 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.

Chapter 4- The *Pseudomonas viridiflava* phylogroups in the *P. syringae* species complex are characterized by genetic variability and phenotypic plasticity of pathogenicity related traits

- 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 Pathog* **7**: e1002132.
- Baltrus, D.A., Yourstone, S., Lind, A., Guibaud, C., Sands, D.C., Jones, C.D., *et al.* (2013) Draft genome sequences for a phylogenetically diverse suite of *Pseudomonas syringae* strains from multiple source populations. *Genome Announc* **2**: e01195–13.
- 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., and 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 hyperpilated and highly adherent phenotypic variants deficient in swimming, swarming, and twitching motilities. *J Bacteriol* **183**: 1195–1204.

Chapter 4- The *Pseudomonas viridiflava* phylogroups in the *P. syringae* species complex are characterized by genetic variability and phenotypic plasticity of pathogenicity related traits

- 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**: 579–583.
- 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.
- 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

Chapter 4- The *Pseudomonas viridiflava* phylogroups in the *P. syringae* species complex are characterized by genetic variability and phenotypic plasticity of pathogenicity related traits

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., Hernández-Guzmán, 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., Billing, 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.

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 bassin versant: vers un modèle de transfert dans habitats naturels aux agro-systèmes. Phd thesis.

Monteil, C.L., Cai, R., Liu, H., Mechan, L., Lontop, 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., 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.

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

Chapter 4- The *Pseudomonas viridiflava* phylogroups in the *P. syringae* species complex are characterized by genetic variability and phenotypic plasticity of pathogenicity related traits

- 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**: 321–334.
- 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**: 107–110.
- Mulet, M., Lalucat, J., and García-Valdés, E. (2010) DNA sequence-based analysis of the *Pseudomonas* species. *Environ Microbiol* **12**: 1513–1530.
- O’Brien, H.E., Thakur, S., and Guttman, D.S. (2011) Evolution of plant pathogenesis in *Pseudomonas syringae*: a genomics perspective. *Annu Rev Phytopathol* **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.
- 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., Kazmierczak, 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.

Chapter 4- The *Pseudomonas viridiflava* phylogroups in the *P. syringae* species complex are characterized by genetic variability and phenotypic plasticity of pathogenicity related traits

- Watson, M.E., Burns, J.L., and Smith, A.L. (2004) Hypermutable *Haemophilus influenzae* with mutation in *mutS* are found in cystic fibrosis sputum. *Microbiology* **150**: 2947–2958.
- Wu, L., McGrane, R.S., and Beattie, G.A. (2013) Light regulation of swarming motility in *Pseudomonas syringae* integrates signaling pathways mediated by a bacterio- phytochrome and a LOV protein. *mBio* **4**: e00334-13.
- Xia, X. (2013) DAMBE5: a comprehensive software package for data analysis in molecular biology and evolution. *Mol Biol Evol* **30**: 1720–1728. doi:10.1093/molbev/mst064.
- Zoina, A., Puopolo, G., Pasini, C., D'Aquila, F., Cozzolino, L., and Raio, A. (2004) *Pseudomonas viridiflava* is the casual agent of a bacterial disease of *Ranunculus asiaticus*. *J Plant Pathol* **86**: 339.

Supporting information

Additional supporting information is available in the online version of the article at the publisher's web-site.

Fig. S1. Neighbour-joining trees constructed on the basis of the single housekeeping gene *cts*, *gyrB*, *gapA* and *rpoD*. Posterior probabilities are indicated at each node.

Fig. S2. Neighbour-joining tree based on the *cts* sequences was compared with the phenotypic pattern of 59 strains. The genotype of the strains for T3SS genes is also shown.

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

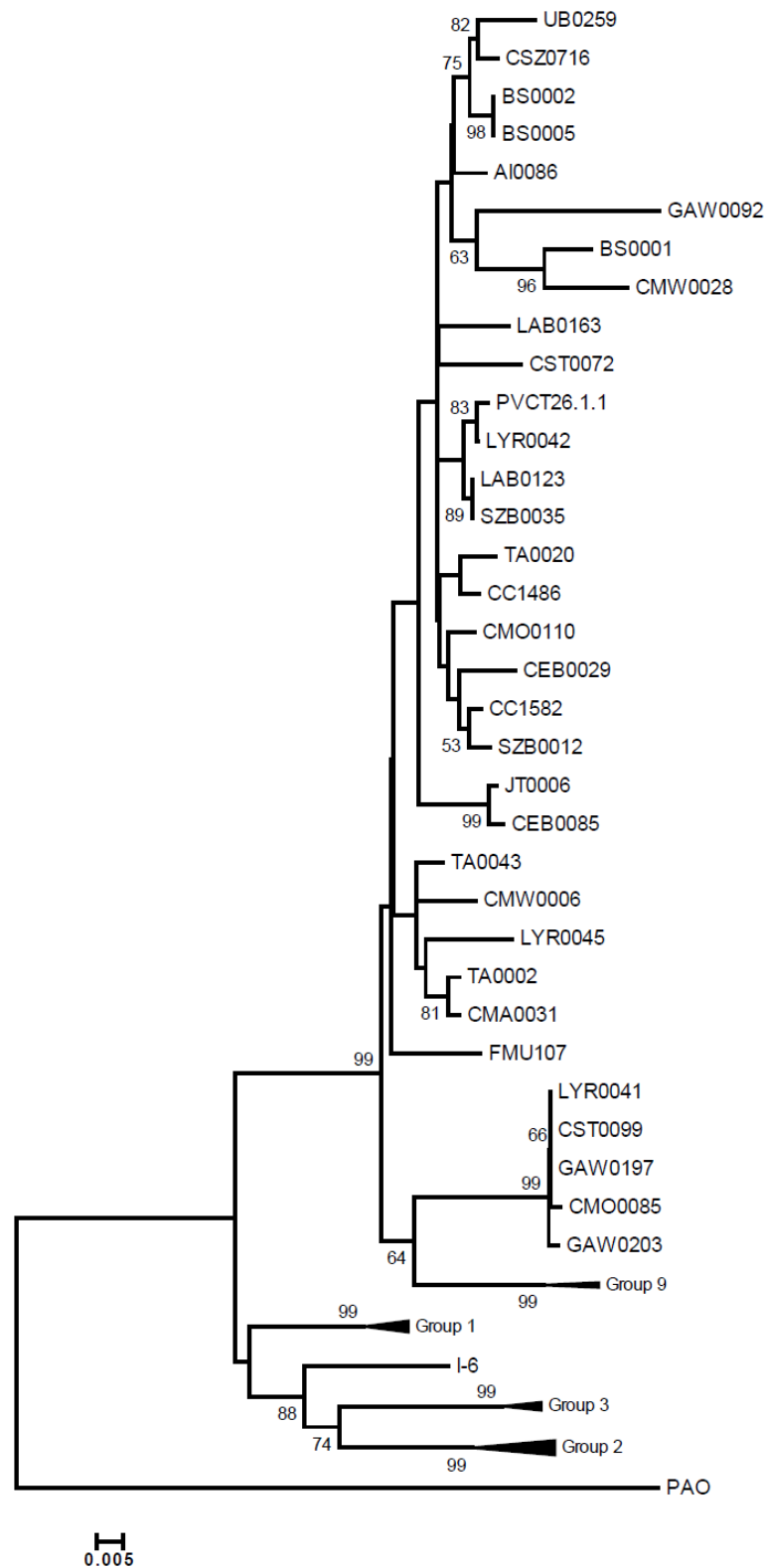
Fig. S4. Bayesian tree based on *hrpL* gene sequences.

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

Table S2. *Pseudomonas syringae* strains used to test the specificity of the primers designed for the type three secretion genes and the lycopene cyclase gene.

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

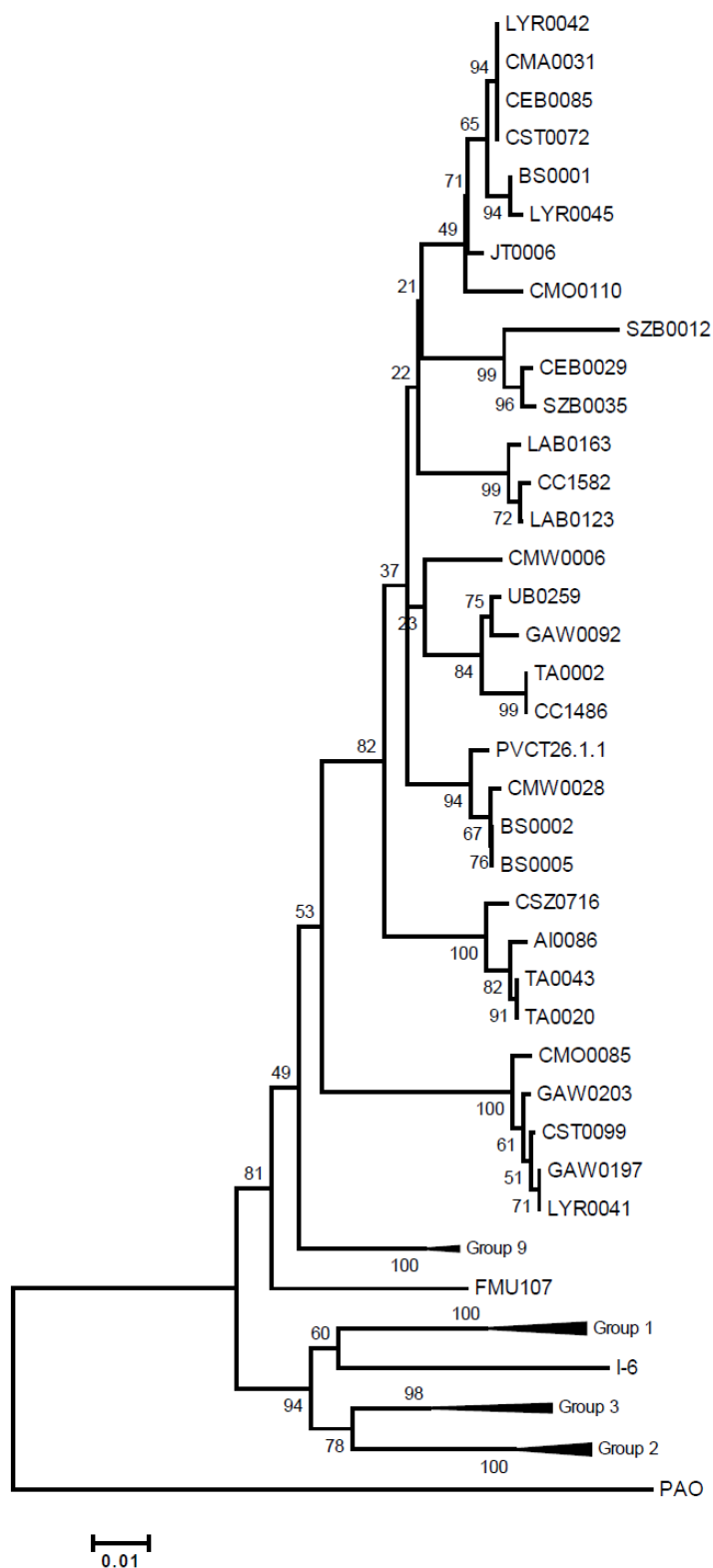
Chapter 4- The *Pseudomonas viridiflava* phylogroups in the *P. syringae* species complex are characterized by genetic variability and phenotypic plasticity of pathogenicity related traits



Neighbor joining tree constructed on *gapA* gene sequences

Supplementary Figure S1

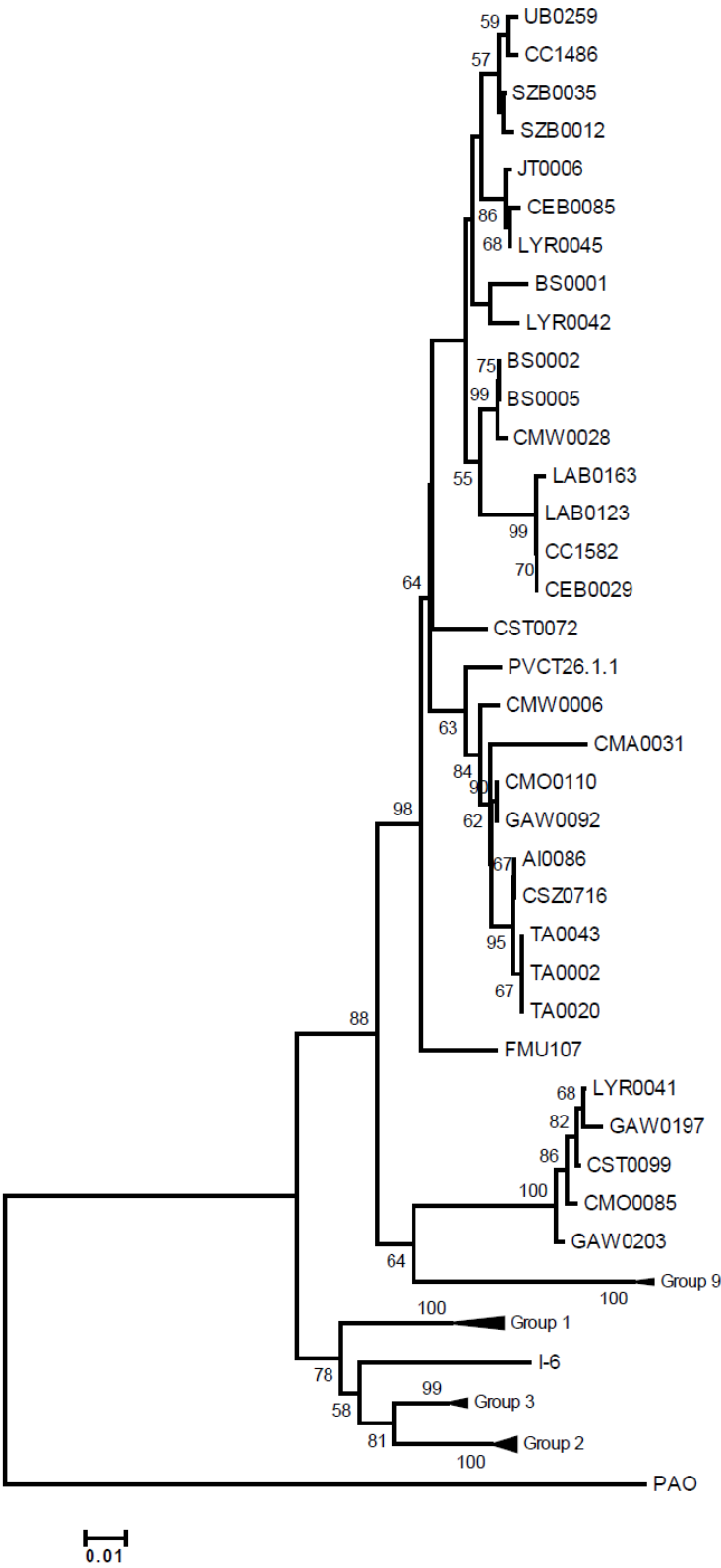
Chapter 4- The *Pseudomonas viridiflava* phylogroups in the *P. syringae* species complex are characterized by genetic variability and phenotypic plasticity of pathogenicity related traits



Neighbor joining tree constructed on *gyrB* gene sequences

Supplementary Figure S1

Chapter 4- The *Pseudomonas viridiflava* phylogroups in the *P. syringae* species complex are characterized by genetic variability and phenotypic plasticity of pathogenicity related traits



Neighbor joining tree constructed on *rpoD* gene sequences

Supplementary Figure S1

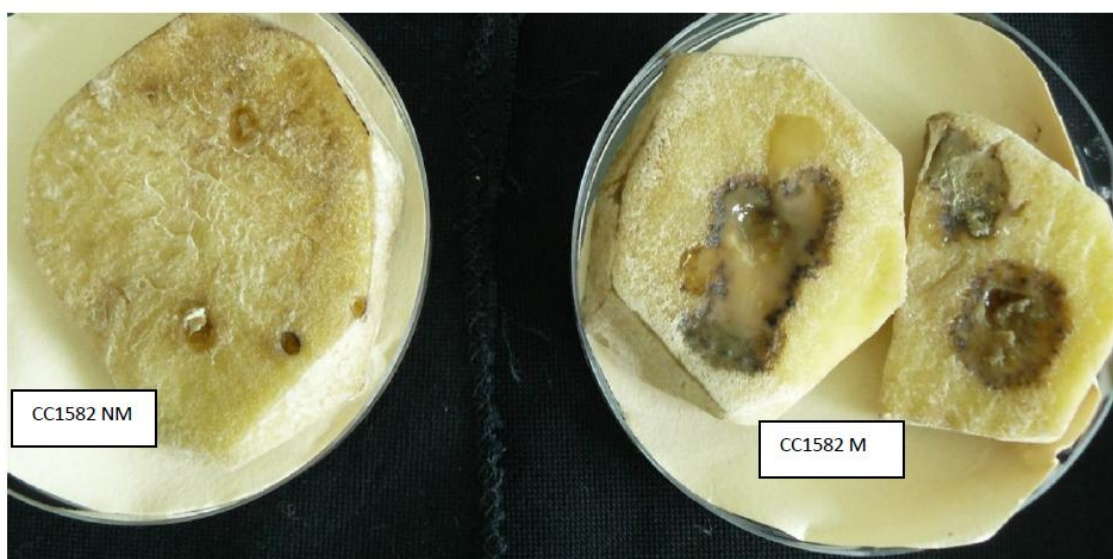
characterized by genetic variability and phenotypic plasticity of pathogenicity related traits



Chapter 4- The *Pseudomonas viridiflava* phylogroups in the *P. syringae* species complex are characterized by genetic variability and phenotypic plasticity of pathogenicity related traits



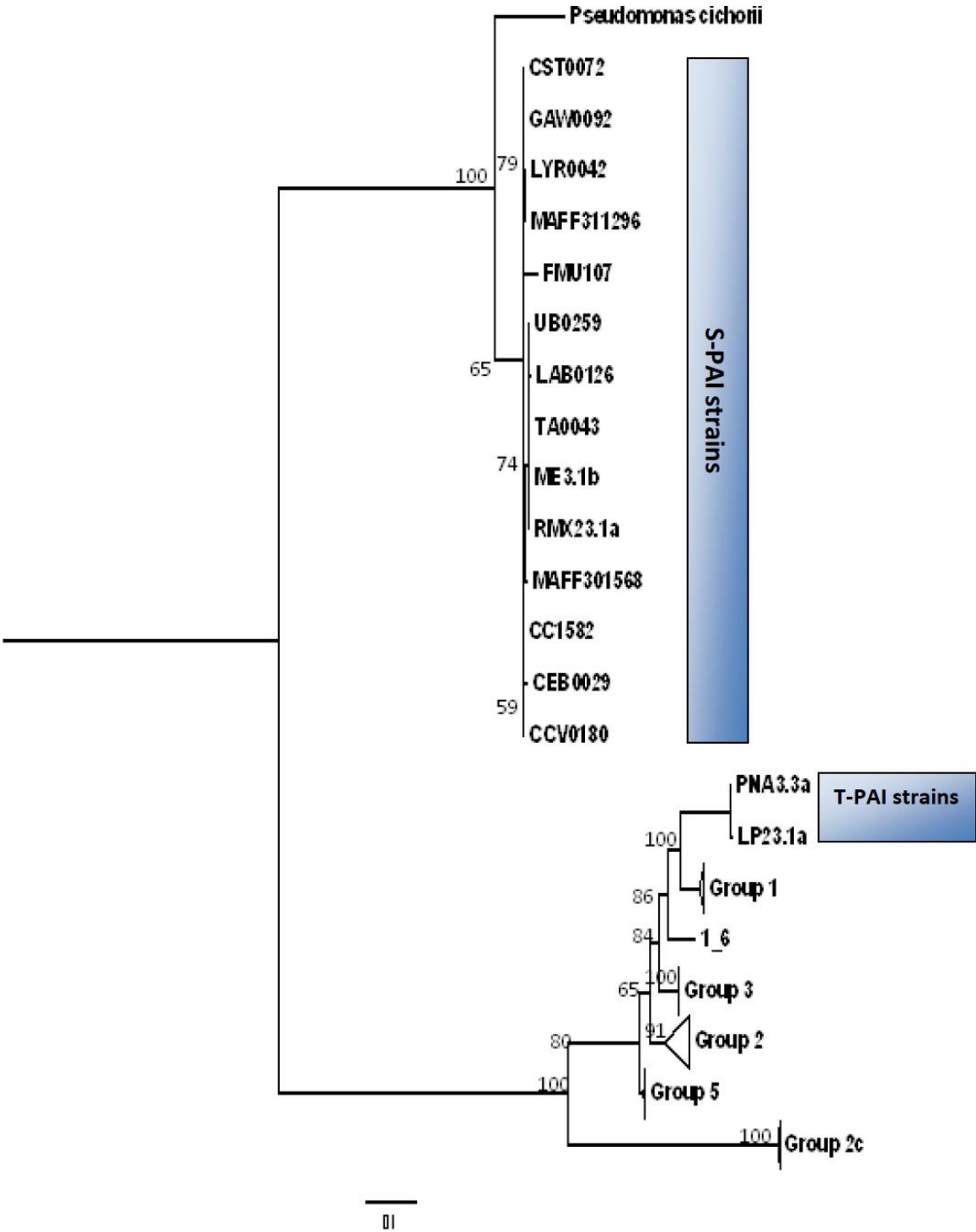
Lesion test on bean pods with mucoid (M) and non-mucoid (NM) CC1582 variant.



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

Supplementary Figure S3

Chapter 4- The *Pseudomonas viridiflava* phylogroups in the *P. syringae* species complex are characterized by genetic variability and phenotypic plasticity of pathogenicity related traits



Supplementary Figure S4

Chapter 4- The *Pseudomonas viridiflava* phylogroups in the *P. syringae* species complex are characterized by genetic variability and phenotypic plasticity of pathogenicity related traits

Supplementary Table S1

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	+

Chapter 4- The *Pseudomonas viridiflava* phylogroups in the *P. syringae* species complex are characterized by genetic variability and phenotypic plasticity of pathogenicity related traits

Supplementary Table S2

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			
NCPBP3487	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.

Chapter 5

**A framework to gage the epidemic
potential of plant pathogens in
environmental reservoirs: the example of
kiwifruit canker**

Submitted to Molecular Plant Pathology in April 2014

Chapter 5 - A framework to gage the epidemic potential of plant pathogens in environmental reservoirs: the example of kiwifruit canker

A framework to gage the epidemic potential of plant pathogens in environmental reservoirs: the example of kiwifruit canker

Claudia Bartoli^{1,2}, Jay Ram Lamichhane^{1,2}, Odile Berge², Caroline Guilbaud², Leonardo Varvaro¹, Giorgio M. Balestra¹, Boris A. Vinatzer³ and Cindy E. Morris^{2*}

¹Departement of Science and Technology for Agriculture, Forestry, Nature and Energy (DAFNE), Tuscia University, 01100 Viterbo, Italy

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

³Departement of Plant Pathology, Physiology, and Weed Science, Virginia Tech, Blacksburg, Virginia, USA

*Corresponding author:

cindy.morris@avignon.inra.fr

Plant Pathology Research Unit 0407, INRA-PACA at Avignon

67 Allée des Chênes, CS60094, 84143 Montfavet cedex, France

Phone 33 (0)4 32 72 28 86

Fax 33 (0)4 32 72 28 42

Chapter 5 - A framework to gage the epidemic potential of plant pathogens in environmental reservoirs: the example of kiwifruit canker

Summary

New, economically important diseases on crops and forest trees emerge recurrently. Understanding where new pathogenic lines come from and how they evolve is fundamental for deploying accurate surveillance methods. We have used kiwifruit bacterial canker as a model for assessing the importance of potential reservoirs of new pathogenic lineages. The current kiwifruit canker epidemic is at least the fourth outbreak of disease on kiwifruit caused by *Pseudomonas syringae* in the mere 50 years that this crop has been cultivated worldwide with each outbreak being caused by different genetic lines of the bacterium. Here we ask if strains in natural environments could cause future epidemics of canker to kiwifruit. To answer this question, we evaluated pathogenicity, endophytic colonization capacity and competitiveness on kiwifruit of *P. syringae* strains chosen from a collection from aquatic and subalpine habitats based on criteria of genetic similarity to epidemic strains. All environmental strains possessing an operon involved in the degradation of aromatic compounds via the catechol pathway grew endophytically and caused symptoms in kiwifruit vascular tissue. Environmental and epidemic strains showed a wide host range revealing their potential as future pathogens of a variety of hosts. Environmental strains readily co-existed endophytically with epidemic strains and shared about 20 virulence genes but were missing six virulence genes found in all epidemic strains. By identifying specific gene content in genetic backgrounds similar to known epidemic strains we developed criteria to assess epidemic potential and to survey for such strains as a means of forecasting and managing disease emergence.

Running title: The epidemic potential of plant pathogen

Key words: emerging pathogens, co-existence, effector repertoire

Introduction

There is growing concern worldwide about emergence of new diseases of crops and forest trees (Elena *et al.*, 2011; Santini *et al.*, 2013). This concern is heightened by changes in climate, land use and commercial networks that could intensify emerging epidemics of plant disease (Bebber *et al.*, 2013; Chakraborty, 2013; Santini *et al.*, 2013). Considerable effort has been invested in elucidating the evolutionary processes that lead to the emergence of new biotypes and pathotypes of pathogens and to their establishment in disease epidemics (Gladieux *et al.*, 2008; Elena *et al.*, 2011; McCann *et al.*, 2013; Ali *et al.*, 2014). An important challenge for managing upcoming epidemics is to deploy the insight gained from these studies to survey reservoirs of pathogens for the genetic or phenotypic lines with the greatest potential for emergence and to monitor crops that are at the greatest risk for new diseases.

Among the plant pathogens, *Pseudomonas syringae* has been responsible for a large number of disease emergences. In this century alone, new diseases caused by *P. syringae* have been reported on over 20 different species of woody plants, for example (Lamichhane *et al.*, 2014). Among the plants that have suffered new diseases caused by this bacterium, kiwifruit has experienced at least four independent emergences of disease caused by different genetic lines of this bacterium over the past 35 years of the mere 50 years that kiwifruit has been cultivated commercially worldwide. Bacterial canker of kiwifruit was first observed in California in 1980 (Opgenorth *et al.*, 1983). Successively, bacterial canker of kiwifruit was reported in Japan (Takikawa *et al.*, 1989) Italy (Scortichini, 1994) and in South Korea (Koh and Nou, 2002). In parallel, multiple emergences occurred in China starting in 1985 and leading to declaration of the disease in 11 different provinces at present. The recent global kiwifruit canker epidemic, responsible for severe economic losses, appeared in 2008-2011 in Italy (Balestra *et al.*, 2009) and in the rest of the world, including Spain (Balestra *et al.*, 2011), Portugal (Balestra *et al.*, 2010), Switzerland (EPPO, 2011a), Turkey (Bastas and Karakaya, 2012), France (Vanneste *et al.*, 2011), Australia (IPPC, 2011), Chile (EPPO, 2011b), Germany (Organization, 2013) and Slovenia (on declaration). The disease has since been reported in New Zealand (Everett *et al.*, 2011) and is jeopardizing its kiwifruit industry which represents a main economic revenue for the country (Vanneste, 2012). More recently, leaf spots on kiwifruit have been reported to be caused by strains in yet another new lineage (Vanneste *et al.*, 2013b). This progression portends future emergences of new lines of *P. syringae* pathogenic to kiwifruit.

Chapter 5 - A framework to gage the epidemic potential of plant pathogens in environmental reservoirs: the example of kiwifruit canker

Recent studies have illustrated the importance of non-agricultural environments as sources of *P. syringae* strains pathogenic for plants (Morris *et al.*, 2007). In particular, it has been demonstrated that environmental strains genetically related to the tomato speck pathogen (*P. syringae* pv. *tomato*) are pathogenic on tomato and have a higher diversity of Type Three Secretion System (T3SS) effectors (Monteil *et al.*, 2013). This led to the hypothesis that highly aggressive clonal *P. syringae* pv. *tomato* populations evolved from a less aggressive environmental strain through a small number of evolutionary events and spread worldwide. In this light, the emergence of new pathotypes of *P. syringae* could be relatively simple in terms of the evolutionary processes involved.

It has been proposed that wild *Actinidia* species could have been the source of the progenitors of the pathogens of cultivated kiwifruit (McCann *et al.*, 2013). However, the ubiquitous nature of *P. syringae* (Morris *et al.*, 2013) and the link that has been demonstrated between environmental and epidemic populations for other pathovars (Monteil *et al.*, 2013) lead us to wonder about the potential of environmental habitats to harbor pathotypes for future disease outbreaks of kiwifruit canker. To evaluate this potential, we characterized the pathogenicity, fitness, host range and effector repertoire of strains from non-agricultural habitats that belong to the *P. syringae* phylogroup 1 and compared them to known epidemic strains of *P. syringae* pv. *actinidiae* (Psa) and to other closely related strains known to attack woody crops (Lamichhane *et al.*, 2014). The strains that were characterized here were selected from our collection of environmental strains based on their genetic similarity with Psa strains for various molecular markers. All strains characterized, including the Psa epidemic strains, had broad host ranges. Several environmental strains induced vascular discoloration, slight external symptoms and were equal in their fitness as endophytes to the epidemic strains. The presence of genes in the catechol pathway operon was particularly accurate in predicting the capacity of strains to colonize the kiwifruit vascular system endophytically and to cause vascular discoloration. Our approach is novel because, by using a marker for a function that we suspected to be useful in colonization of woody tissue, we identified environmental strains that could represent future emerging pathogens for kiwifruit as well as for other woody plant hosts. We propose that the evolution of less aggressive strains toward greater aggressiveness might involve few evolutionary steps including the acquisition of new pathogenic determinants such as T3SS effector genes. Genes in the catechol operon could be targeted in surveillance programs for future epidemic pathogens of kiwifruit and other crops.

Results

Multiple distinct genetic lines of *P. syringae* strains from fresh water habitats are pathogenic to kiwifruit

Comparison of gene content between publicly available genomes of several *P. syringae* pathogens of woody plants with those of herbaceous plants revealed an operon with predicted function in the catechol pathway that was present only in pathogens of woody plants. The screening of several *P. syringae* strains from environmental habitats based on the presence of the catechol operon then allowed us to establish a collection of 10 strains from aquatic habitats and 10 reference strains representing previously described pathovars in addition to 4 strains of Psa bv. 1 (representing strains of the first outbreak in Japan and Italy), Psa bv. 3 (current epidemic strains) and Psa bv.4 (representing the so called low virulent strains) (Vanneste *et al.*, 2013a). Strains in Psa bv. 2 isolated in Korea were not included in the study (Fig. 1). Of the 10 environmental strains, 8 had the genes of the catechol operon. Two strains were phylogenetically close to the reference strains but without the genes in the catechol operon (AF0007, USA0003). All the environmental strains formed different genetic lineages based on MLST (Fig. 1). When inoculated on kiwifruit, the strains caused four types of reactions: i) severe disease with external and internal symptoms (CFBP 7286, CC1676, PA459) (Fig. 2A), ii) marked but less severe disease with mild external and internal symptoms (AF0015, USA0007, CC1544, USA0001, CSZ0350, CSZ0343, CSZ0761 and SZB0070) (Fig. 2B), iii) faint canker and/or swelling on the inoculated point (NCPPB 3335 and PseNe107) and faint internal necrosis (NCPPB 2598, PseNe107 and ICMP 18882) (Fig. 2C); and iv) no external or internal disease symptoms (MAFF 302280PT, DC3000, KN203, MAFF 302273, MAFF 30120, AF0007, USA0003, PaVt10 and 0893_23) (Fig. 2D). All strains causing visible symptoms were successfully re-isolated from plant tissue. In addition, a few other strains that didn't cause any visible symptom (MAFF 302280PT, and 0893-23) were consistently re-isolated. Finally, the remaining strains that didn't cause any symptoms (DC3000, KN203, MAFF 302273, MAFF 30120 and PaVt10) could not be re-isolated. In terms of distance from the point of inoculation, Psa bv.1 and Psa bv. 3 strains were isolated up to 3 cm below the point of inoculation and the environmental strains were isolated up to 1 cm below the point of inoculation. By contrast, strain ICMP 18882 (Psa bv. 4) and MAFF 302280PT (pv. *morsprunorum*) could be isolated

only from the point of inoculation but not beyond. All the bacteria re-isolated from the inoculated plants had the expected BOX-PCR profiles.

Environmental strains attain population sizes in the vascular tissue of kiwifruit similar to epidemic *Psa* bv. 3 strains and can co-exist with epidemic strains without disease occurrence

Endophytic growth on kiwifruit assessed for CFPB 7286, USA0007 and CC1544 showed that, when inoculated independently,: i) all strains attained their maximum population size between 3 and 7 days post inoculation (dpi) (Fig. 3A, 3B), and after that they slightly decrease (Fig. 3A, B), ii) there was no significant difference ($P > 0.1$) between environmental strain population densities and those of the CFPB 7286 at all time points , iii) and strain CFBP 7286 caused severe external and internal necrosis by 21 dpi (Fig. 4A) while environmental strains induced mild external and internal symptoms by 21 dpi.

When strains were co-inoculated we observed that: i) for the couple CFBP 7286-CC1544 population density of strain CFBP 7286 was significantly reduced at 3 and 7 dpi compared with the single inoculation ($P \leq 0.005$), and in contrast, the population density of strain CC1544 was not influenced ($P > 0.05$) by the presence of CFBP 7286 (Fig. 3B), ii) for the couple USA0007-CFBP 7286 the population density of CFBP 7286 was reduced significantly ($P \leq 0.05$) at 3 and 7 dpi compared with the single inoculation , while the strain USA0007 was not influenced ($P > 0.05$) (Fig. 3A), iii) for all treatments (co-and individual inoculations) bacteria attained and maintained almost the same density (10^7 - 10^8). The ability of the strains to move into the kiwifruit vessels was reduced for both co-inoculation experiments, in particular for the co-inoculation of CFBP 7286 and CC1544 where no bacterial cells were re-isolated at any time point below or above the point of inoculation (Fig. S1). In the co-inoculation experiments, no symptoms were recorded (Fig. 4B).

During the *in vitro* test no effects on growth were observed for the environmental strains when they were co-inoculated with the CFBP 7286 strain. In contrast, strain CFBP 7286 was significantly inhibited 28 h after inoculation ($P \leq 0.0005$) by both environmental strains (Fig. S2), more so than in co-inoculations *in planta*.

Endophytic growth in kiwifruit is correlated with the presence of a highly conserved operon for degradation of aromatic compounds that is restricted to phylogroups 1 and 3 of *P. syringae*

We observed that strains that grew endophytically on kiwifruit were also positive for the presence of the catechol operon. These included strains belonging to Psa bv. 3 (CC1676, CFBP 7286), to Psa bv. 1 (PA459), Psa bv. 4 (ICMP 18882), and the environmental strains CC1544, AF0015, SZB0070, CSZ0343, CSZ0761, USA0001, USA0007, CSZ0350 as well as other strains of the *P. syringae* phylogroup 1 (MAFF 302280PT (pv. *morsprunorum*), NCPPB 2598 (pv. *theae*)), and strains of phylogroup 3 (0893_23 (pv. *aesculi*), NCPPB 3335 (pv. *savastanoi*), PseNe107 (pv. *savastanoi*)) (Fig. 1). In contrast, the strains lacking the operon involved in the catechol branch β -ketoadipate pathway (pv. *avellanae* strain PaVt10, the environmental strains AF0007 and USA0003, the tomato pathogen DC3000, the pv. *maculicola* strain KN203 and the pv. *mori* strain MAFF 302273) were not able to grow endophytically on kiwifruit (Fig. 1). Other molecular markers for Psa (*hopZ3*) (Balestra *et al.*, 2013) and the ITS markers (Rees-George *et al.*, 2010) were present in only some of the environmental strains and in the epidemic strains causing severe disease symptoms. Among the strains, the presence of *hopZ3* and ITS markers was not correlated with disease occurrence and endophytic colonization capacity in kiwifruit (Fig. 1).

Analysis of the draft genomes of the two environmental strains (USA0007 and CC1544), several of the Psa genomes available in GenBank (accession numbers are listed in Table S1), and those of *P.s.* pv. *morsprunorum* (strains MAFF 302280PT), *P. s.* pv. *aesculi* (strains 0893_231, NCPPB 3681 and 2250) and, *P. savastanoi*pv. *savastanoi* (strain NCPPB 3335) revealed that the genes involved in the degradation of phenolic and aromatic compounds, via the catechol branch β -ketoadipate pathway, are organized in a unique operon. This operon is composed of i) a PAS domain S-box protein that corresponds to the putative promoter, ii) a protein involved in the meta-pathway phenol degradation, iii) a conserved flavoprotein-oxygenase with a flavin reductase like domain, iv) a FAD-dependent oxido-reductase, v) a short-chain alcohol dehydrogenase involved in oxidation of different aromatic compounds (such as antibiotics and compounds involved in nitrogen metabolism), and vi) a dienelacton hydrolase. The latter is responsible for the degradation of chloroaromatic compounds by three different pathways: 4-chlorocatechol degradation, 3-chlorocatechol degradation II (*ortho*) and 3-chlorocatechol degradation I (*ortho*) (Sridevi *et al.*, 2012).

Using the primers we developed here, a survey conducted on a large collection of strains representing all known phylogroups of *P. syringae* showed that these genes are present only in *P. syringae* phylogroups 1 and 3. The absence of the operon from *P. syringae* strains that are pathogens of herbaceous crops was also confirmed by searching all *P. syringae* genomes in GenBank with the proteins encoded in this operon. The presence of the operon was also investigated in all the other bacterial genomes available in GenBank. Outside of *P. syringae*, the complete operon was only found in the genomes of *P. pseudoalcaligenes* and *P. fuscovagine*.

The phylogenetic tree constructed on the concatenated sequences forming the operon (Fig. 5) showed the same phylogenetic structure as the MLST tree (Fig. S3). To better understand the genetic relationships between the operon and the *cts*, *gapA*, *gyrB*, *rpoD* housekeeping genes, we made strain-by-strain pairwise comparisons of the *p-distances* of the concatenated sequences obtained by the four housekeeping genes with those obtained from the concatenated genes forming the operon. This showed a significant correlation between operon gene sequences and housekeeping genes (Fig. S3). The same tree topology between operon and housekeeping genes, the lack of anomalous nucleotide composition (for example G/C ratio), and absence of phage integrases, transposons and vestiges of other genetic mobile elements flanking the operon led us to conclude that the operon was not transmitted via horizontal gene transfer and probably was acquired before the diversification of the *P. syringae* complex.

Environmental and epidemic strains of *P. syringae* pathogenic to kiwifruit have broad host ranges and partially overlapping effector repertoires

All four strains characterized for host range were pathogenic on 16 to 18 of 22 host plants tested (Fig. 6). For woody plants, we recorded both external and internal disease symptoms. Some hosts did not bear any external disease symptoms while they were affected internally showing clear vascular necrosis (Fig. 6). All of the inoculated bacterial strains were re-isolated from vascular tissue 60 dpi independent of the presence or absence of symptoms except for strain ICMP 18882 which could not be recovered from Mariana plum (Fig. 6).

All woody hosts tested were susceptible to at least one strain (in terms of either internal or external symptoms) whereas some herbaceous plants (spinach and liverwort) did not manifest disease after inoculation with any of the strains. In addition, on hazelnut and oleander all strains caused only internal symptoms (Fig. 6). The environmental strain CC1544 did not cause

external symptoms on woody hosts (except for kiwifruit) but internal symptoms were observed in almost all the plant species tested (Fig. 6).

The presence of type III secretion effector genes in the genomes of two environmental strains (USA0007 and CC1544) and the *Psa* epidemic strains (accession numbers in Table S1) was determined by blasting the sequences of known effector genes that were obtained from <http://www.pseudomonas-syringae.org>. *Psa* bv.3 strains carried 44 full length effectors, while *Psa* bv. 4 and *Psa* bv. 1 strain carried 32 and 40 effectors respectively (Fig. 7). In contrast, strains USA0007 and CC1544 had 29 and 27 full length effector genes, respectively (Fig. 7). Eighteen effector genes were found to be shared among the *Psa* biovars and environmental strains, while four (*hopH1*, *hopZ5*, *hopAA1-2*, *hopAM1-2*) and two (*hopF1* and *hopE1*) effector genes were unique for *Psa* bv. 3 and for *Psa* bv.4, respectively (Fig. 7). Three effectors (*hopAJ1*, *hopV1*, *hopA2*) were carried only in the environmental strains. Finally, *Psa* bv.4 and the environmental strains lacked six effectors (*avrD1*, *avrRpm1*, *hopAM1-1*, *hopD1*, *hopQ1-1*, *hopBB1*) that were present in both *Psa* bv. 1 and *Psa* bv. 3.

The *hopM1* gene, known to be consistently present in the *P. syringae* phylogroup 1 (Baltrus *et al.*, 2011) was observed in all strains (Fig. 7). These authors previously hypothesized that different allelic forms of *hopM1* could influence the host range of the strains from this phylogroup. The same authors showed that phylogenies based on the *hopM1* genes split strains from phylogroup 1 in two divergent groups, the pv. *morspunorum*/pv. *actinidiae* and the pv. *tomato*/pv. *lachrymans* group. In this light, we investigated the *hopM1* phylogenetic relationships among *Psa* epidemic strains, environmental strains and strains from other pathovars of phylogroup 1. The phylogeny constructed with a Bayesian model showed that environmental strains are divided into at least four different genetic lineages distinct from the *Psa* epidemic strains (Fig. S4). Strain CC1544 isolated from France and strain AF0015 from New Zealand had identical *hopM1* alleles whereas strains isolated from the same water sample in the USA (USA0001 and USA0007) had distinctly different *hopM1* alleles. Phylogeny also showed that the *hopM1* alleles of *Psa* bv. 3 and *Psa* bv. 1 are clonal while that of *Psa* bv. 4 is genetically related to that of pv. *morsprunorum* but distinct from the environmental strains.

Discussion

Our results contribute to the emerging concept that non-agricultural environments are reservoirs of plant pathogens with broad host ranges and complex effector repertoires (Monteil *et al.*,

Chapter 5 - A framework to gage the epidemic potential of plant pathogens in environmental reservoirs: the example of kiwifruit canker

2013; Morris *et al.*, 2013; Goss *et al.*, 2013). *P. syringae* strains from aquatic habitats can colonize and cause symptoms in kiwifruit and other woody and herbaceous plants, and share numerous effectors with epidemic strains of the bacterial canker pathogen in addition to possessing some unique effectors. The strains pathogenic to kiwifruit characterized here represent clonal lines that are distinct from the lines of the pathogen that have emerged in epidemics over the past several decades. Therefore, the environmental strains characterized here are unlikely to have contributed to epidemics of bacterial canker seen to date. However, given the numerous processes that can bring environmental strains of *P. syringae* into encounters with agriculture (Morris *et al.*, 2013) these strains have important potential to contribute to future kiwifruit canker epidemics and, in light of their host ranges, to diseases of herbaceous and other woody species. Contact of environmental strains with kiwifruit plants – either wild or cultivated - might explain the history of emergences of canker in China where the diversity of genetic lineages implicated in the disease is greater than that in Europe and New Zealand. Furthermore, the capacity of the environmental strains to colonize vascular tissue of woody plants, and most importantly with low severity of apparent symptoms that would not lead to phytosanitary interventions, suggests that there are opportunities for co-existence with epidemic strains and subsequent opportunity for acquisition of effectors by horizontal gene transfer (as discussed below). Therefore, our work makes an important contribution to identifying the processes that need to be monitored in order to come closer to predicting new disease emergences.

For the bacterial speck pathogen, it has been proposed that the epidemic lineage probably emerged from a less aggressive environmental strain through a small number of evolutionary events, such as by losing effector genes deleterious for disease development (Monteil *et al.*, 2013). Here, we found that environmental strains had a lower number of effectors but almost the same host range as that of epidemic Psa strains. However the environmental strains did not move through the vascular system as efficiently as the epidemic strains (environmental strains were isolated only 1 cm from the point of the inoculation whereas epidemic strains were found 3 cm from the point of inoculation). In light of this, we argue that environmental strains could become more highly aggressive by acquiring some effectors via horizontal gene transfer (HGT). For HGT to occur, nutrient levels should be sufficiently high in order to support an elevated bacterial density (Dröge *et al.*, 1998). Hence, endophytic growth could be a means to attain the needed population density and also to assure proximity between donor and receiver populations (Manceau *et al.*, 1986). Acquisition of one or more of the six effectors absent from Psa bv.4 and

the environmental strains (*avrD1*, *avrRpm1*, *hopAM1-1*, *hopD1*, *hopQ1-1*, *hopBB1*) might enhance aggressiveness. Future studies should focus on the potential HGT between epidemic and environmental strains harbouring the catechol operon to determine the conditions and likelihood of this mechanism contributing to the evolution of environmental strains. To predict the potential for emergence of new kiwifruit pathogens, it will be necessary to elucidate which combination of effectors would enhance pathogenicity to kiwifruit. For example, *avrRpm1* is known to be essential for maximum aggressiveness and for growth of the *Brassica* pathogen *P. syringae* pv. *maculicola* on *Arabidopsis* (Ritter and Dangl, 1995). Likewise, *hopAM1* expressed in *P. syringae* strains led to an enhanced growth on water-stressed *Arabidopsis* plants (Goel *et al.*, 2008). Similar results have been observed on bean, in which the *hopQ1-1* of *P. syringae* pv. *phaseolicola* suppresses plant defences but the same effector expressed in *P. syringae* pv. *tabaci* strains triggers host defences in *Nicotiana* spp. (Ferrante *et al.*, 2009).

For the current epidemic of bacterial canker of kiwifruit, it has been proposed that wild *Actinidia* spp. are likely to have been reservoirs of pathogenic strains (McCann *et al.*, 2013). Here we propose a complementary element to such a scenario whereby aquatic milieu are reservoirs of a vast diversity of strains, many of which are disseminated across various habitats, via the water cycle in particular, where they can survive and thrive (Morris *et al.*, 2010). This dissemination could bring the strains into contact with plants in a range of biomes including those where there are wild relatives of cultivated plants. The small evolutionary steps toward aggressive pathogenicity, such as acquisition of new genes by HGT or deletion of the genomic determinants deleterious for disease occurrence, could occur across the ensemble of the habitats that the bacteria encounter during the dissemination process. Wild plant hosts could then serve as forces of natural selection for strains having gained in fitness in plant tissues.

Niche overlap is one of the first obstacles for co-existence because of the competition for limiting nutritional resources. Therefore, it is not surprising that co-existence of genetically related lines has rarely been described to occur among epiphytic *P. syringae* strains, at least not under experimental conditions (Wilson and Lindow, 1994). However, here we demonstrated the co-existence of genetically close lines of *P. syringae* when co-inoculated in kiwifruit. Co-existence had a cost on fitness and on systemic movement leading to the absence of the disease. This suggests that co-inoculation triggers responses in plants that are not triggered when pure lines are inoculated. Nevertheless, the symptomless nature of this co-existence leads to questions about the frequency of such events in seemingly healthy plants and to the role of co-

existing endophytic microflora in regulating the proliferation of certain genetic lines. A recent study on healthy and diseased kiwifruit plants demonstrated the potential of some endophytes and epiphytes to inhibit *Psa* epidemic strains (Giovanardi *et al.*, 2012). We also have preliminary data (not shown) illustrating that *P. syringae* populations colonizing kiwifruit tissue are composed of mixed genotypes including strains pathogenic to kiwifruit. The importance of multispecies bacterial populations has been well demonstrated in other pathosystems. For example, black band disease of scleractinian coral is caused by a complex bacterial community (Frias-Lopez *et al.*, 2004). The bacterial community of healthy and diseased corals differs dramatically and the bacterial population structure and diversity is a key element in the conditions that trigger the disease.

Our results also suggest that metabolic processes associated with the degradation of aromatic compounds are important for fitness in the kiwifruit vascular system. McCann *et al.*, (2013) recently showed that epidemic strains of *Psa* possess a plasmid with genes involved in aromatic carbon metabolism (genes other than those in the catechol operon) and that these genes are present in other woody pathogens such as *P. syringae* pv. *aesculi* and the xylem pathogen *Xylella fastidiosa*. Our results showed that strains closely related to kiwifruit-competent strains from phylogroup 1 but lacking the genes in the operon involved in phenol degradation via catechol branch β -ketoadipate pathway did not colonize kiwifruit plants. These genes are likely to be involved in processes related to the degradation of some lignin intermediates. Lignin contains many phenolic moieties and its degradation can involve different chemical pathways leading to toxic by-products. Microorganisms commonly use the catechol oxygenase pathway in degrading phenolic compounds (Sridevi *et al.*, 2012) and the genus *Pseudomonas* is known to possess enzymes in this biochemical pathway (Kang and Park, 1997). By metabolizing lignin intermediates, bacteria avoid the toxic effects during colonization of the plant vascular system where lignin is an important structural component. The content of phenolic compounds often increases during the attack of plants by different bacterial pathogens (Oi-Kano, *et al.*, 2008). Hence, the ability of *P. syringae* to degrade phenolic compounds could be a fundamental requisite for endophytic colonization and for pathogenicity. In other bacteria, such as *Pseudomonas putida*, genes for the degradation of aromatic compounds reside on plasmids or transposons (Williams and Sayers, 1994). By contrast, the catechol operon in *P. syringae* is highly conserved and we have shown that its phylogenesis exhibits congruence with housekeeping genes. The maintenance of the operon, both in the *P. syringae* complex and in

Chapter 5 - A framework to gage the epidemic potential of plant pathogens in environmental reservoirs: the example of kiwifruit canker

other *Pseudomonas* spp. abundantly present in the environment, led us to speculate that the environment at large might be an important driver for the evolution of this operon. In the environment, exposure to aromatic compounds might have been a positive selective for the maintenance of this operon. Exposure to plant degradation products might have further reinforced the selection of this operon (Harwood and Parales, 1996). Upon encounter with plant vascular tissues the presence of these genes could then have provided a fitness advantage. Other strains not exposed to the selective forces could have lost the operon. In agreement with this scenario, we observed the presence of this operon only in the *P. syringae* phylogroup 1 and 3 strains known to attack woody plants (pv. *actinidiae*, pv. *savastanoi*, pv. *aesculi*, pv. *morsprunorum* and pv. *theae*). In addition, contact with xenobiotic compounds derived from human activities for more than a century (Díaz, 2004) might have selected bacteria able to use such compounds in their metabolic pathways.

To date, surveillance is one of the most important methods for control of kiwifruit canker as well as for other bacterial plant pathogens (Scortichini *et al.*, 2012). Different detection methods have been proposed for Psa identification. The current molecular markers are based on primers that amplify the ITS region (Rees-George *et al.*, 2010) and the *hopZ3* gene (Balestra *et al.*, 2013) identified in strains from the current disease epidemics. In surveys for the presence of the pathogen, absence of these markers is often a criterion to eliminate a strain from further consideration about its contribution to plant health. As our work illustrates, these detection tools target a narrow range of the ensemble of strains able to grow endophytically in kiwifruit and to cause vascular discoloration. Although such surveillance provides useful information for quarantine measures, it can mask information about synergistic interactions or emergence of new pathotypes. As a means to estimate the potential for emergence of new pathotypes, it would be important to also survey for the presence of other endophytic *P. syringae* able to colonize kiwifruit vascular tissue. Genes in the catechol operon could provide a tool for this goal. In this way, our work contributes to creating a framework of foresight for disease emergence.

Materials and methods

Strain isolation and selection

The reference strains used in this study were sourced from public collections or kindly provided by scientists (Table 1). The strains from different environmental substrates (water, epilithic biofilms) were isolated previously (Morris *et al.*, 2007, 2008) while additional strains from

kiwifruit and hazelnut were isolated during this study. Isolations were made as described previously (Lamichhane *et al.*, 2013). The environmental strains were selected from an in-house data base of about 1600 strains of *P. syringae* based on their phylogenetic co-location in the neighbor-joining tree constructed on the *cts* housekeeping gene. Strains that formed clades close to the reference Psa strains CFBP 7286 and KW30 were selected. Finally, all the isolates were tested for the presence of target genes involved in the degradation of aromatic compounds (see below). Phenotypic tests performed on the isolates are listed in Table S2.

Identification of genes putatively associated with pathogenesis in woody plants

OrthoMCL (54) was used to identify protein families present only in the genomes of the woody pathogens Psa 346, CFBP 7285, CFBP 7286, CFBP 7287, K2, KW41, PA459, V, and VI, *P. syringae* pv. *morsprunorum* MAFF 302280PT, and *P. savastanoi* pv *savastanoi* but absent from the genomes of the herbaceous pathogens *P. cannabina* pv. *alisalensis* BS91, *P. syringae* pv. *maculicola* F1, *P. syringae* pv. *tomato* DC3000 and JL1065, *P. syringae* pv. *pisi* 1704B, *P. syringae* 642, *P. syringae* pv. *syringae* FF5, *P. syringae* pv. *tabaci* ATCC115128, *P. fluorescens* 0-1 and 5 using the OrthoMCL tool on the website <http://pacu.facom.ufms.br/kiwi/orthologsorter/>.

Primers, PCR conditions, sequencing and phylogenetic analysis

Primers and PCR conditions used to amplify the housekeeping genes *gyrB*, *rpoD* and *cts* (Sarkar and Guttman, 2004) and the *gapA* gene (Morris *et al.*, 2008) were as previously described. Information about amplification and sequencing of the marker genes is provided in Text S1.

Phylogenetic trees were constructed on the concatenated sequences of *cts*, *gapA*, *gyrB* and *rpoD* (1852bp). Several Psa sequences available in GenBank and other sequences of strains belonging to the *P. syringae* phylogroup 1 were added into the analysis. Strains and GenBank accession numbers of the sequences are listed in a supplementary data table (Table S1), while sequences for the strains KN.2 and 346 were extracted from the draft genome available on http://pacu.facom.ufms.br/blast_kiwi/blast.cgi. The two environmental strains USA0007 and CC1544 were sequenced previously (Baltrus *et al.*, 2013). Accession numbers for genomes are AVDY01000000 and AVEI01000000, respectively.

Sequences obtained in this study were deposited in PAMBD <http://genome.ppws.vt.edu/cgi-bin/MLST/home.pl>. We also performed phylogenies on the single and concatenated genes

involved in aromatic compound degradation and for *hopM1* sequences. For each phylogenetic analysis, PCR amplicons were aligned and cut to obtain sequences of the same length with DAMBE version 5.1.1 (Xia, 2013). Bayesian trees were constructed with the Mr. Bayes program <http://mrbayes.csit.fsu.edu> by (Ronquist & Huelsenbeck, 2003) using 500 000 generations. Maximum likelihood and parsimony phylogeny were created with the Phylip package (<http://evolution.genetics.washington.edu/phylip.html>).

Pathogenicity test

The strains listed in Table 1 were tested for pathogenicity on kiwifruit plants and for lesion production on various excised fruits. For this test, three-month-old, vegetatively propagated kiwifruit plants (*Actinidia deliciosa* Liang and Ferguson cv. Hayward) were used. For inoculation on kiwifruit plants, the inoculum consisted of aqueous suspensions of 48h bacterial cultures from KB plates adjusted to 1×10^8 CFU ml⁻¹. For all strains inoculated, their fingerprint profiles were determined prior to inoculation by PCR with BOX primers as described previously (Versalovic *et al.*, 1991) with GoTaq Flexi DNA Polymerase (Promega). Each strain had a unique BOX fingerprint. Inoculation was performed only at a single point. At half the height of the main stem, a leaf petiole was manually detached at the base and 10 µl of bacterial suspension were deposited on the leaf scar which had been slightly punctured with a sterile hypodermic needle prior to inoculation. The inoculated point was marked with masking tape for later retrieval. Four plants were inoculated for each strain and for the negative control (sterile distilled water) treatment. The inoculated plants were maintained for 35 days post inoculation (dpi) in the growth chamber at night/day temperatures of 18/23°C and with a light/dark period of 16/8 hours. Plants were watered twice a day, without wetting the leaves, to maintain a constant high humidity. All plants were examined individually for disease symptoms up to 35 dpi and the tests were repeated twice for each strain.

At 35 dpi, disease severity on each plant was scored by considering both the external (necrosis, canker and/or gall) and internal (necrosis) symptoms. Internal disease symptoms on each plant were assessed by excising the sub-cortical tissues of the woody stem and measuring the length of the necrosis developed along the plant vessels. The following scale was used for disease scoring: no symptoms (apparently healthy external and internal tissues), faint symptoms (external symptoms of < 0.20 cm without any internal necrosis), moderate symptoms (both external and internal symptoms of 0.21-1 cm), severe symptoms (both external and internal

Chapter 5 - A framework to gage the epidemic potential of plant pathogens in environmental reservoirs: the example of kiwifruit canker

symptoms of 1.1-3 cm) and highly severe symptoms (both external and internal symptoms of > 3 cm).

Re-isolation of the inoculated strains was performed 5 weeks after inoculation. Re-isolation of the bacteria from the inoculated woody tissues was made by excising bark with a sterile scalpel and a small fragment of sub-cortical tissue was removed and placed in a microtube containing 900 µl of phosphate buffer. The suspensions were incubated at room temperature overnight and then 3 drops of 10 µl were plated onto KB medium supplemented with 50 mg l⁻¹ of cycloheximide. The putative *P. syringae* colonies from each isolation point were streaked for purity and their identity was checked by comparing their BOX-PCR profile to that of the strain used for inoculation.

Endophytic growth in kiwifruit vascular tissue

The capacity for endophytic growth of strains CFPB 7286 (Psa bv.3), and the two environmental strains, USA0007 and CC1544 was evaluated on kiwifruit plants under conditions of single- and mixed-culture inoculation. Experiments were performed in the green house under ambient conditions with an approximate temperature of 27°C during the day (13 h) and 13°C at night (11 h) and repeated twice. More information is provided in Text S1.

Host range determination

Host range of the strains was determined on thirteen herbaceous and nine woody plant species. The plant species and the strains used as positive controls are listed in Table S3.

All of the annual plants tested were seed-grown and transplanted in fresh medium-decomposed white sphagnum peat. *Prunus* spp. plants were sourced from commercial nurseries producing plants from seed, re-potted in the same substrate and maintained in a greenhouse at ambient temperature until inoculation. Annual plants were inoculated one month after sowing while 2-year-old *Prunus* spp. and 3-month-old hazelnut, oleander and poplar plants were used for the inoculation. More information about host range determination is available in Text S1.

Statistical analysis

Data from the endophytic growth assay and competition tests were analyzed with paired Student's *t*-tests on data expressed as log₁₀ of CFU g⁻¹, using Graphpad software

Chapter 5 - A framework to gage the epidemic potential of plant pathogens in environmental reservoirs: the example of kiwifruit canker

<http://www.graphpad.com/quickcalcs/>. Values of $P \leq 0.05$ were considered statistically significant.

Acknowledgements

The salary of Claudia Bartoli was provided by Tuscia University (Italy). We thank INRA's Plant Health and Environment Department (SPE) for supporting this work. We also thank the greenhouse staff of INRA's Plant Pathology Research unit at Avignon for providing and maintaining all plant materials tested in this study.

References

- Ali, S., Gladieux, P., Leconte, M., Gautier, A., Justesen, A.F., Hovmøller, M.S., Enjalbert, J. and de Vallavieille-Pope, C.** (2014) Origin, migration routes and worldwide population genetic structure of the wheat yellow rust pathogen *Puccinia striiformis* f.sp. *tritici*. *PLoS Pathogens* **10**, e1003903.
- Balestra, G.M., Mazzaglia, A., Quattrucci, A., Renzi, M. and Rossetti, A.** (2009) Current status of bacterial canker spread on kiwifruit in Italy. *Australas. Plant Dis. Notes* **4**, 34–36.
- Balestra, G.M., Renzi, M. and Mazzaglia, A.** (2010) First report of bacterial canker of *Actinidia deliciosa* caused by *Pseudomonas syringae* pv. *actinidiae* in Portugal. *New Dis. Reports* **22**, 10.
- Balestra, G.M., Renzi, M. and Mazzaglia, A.** (2011) First report of *Pseudomonas syringae* pv. *actinidiae* on kiwifruit plants in Spain. *New Dis. Reports* **24**, 10.
- Balestra, G.M., Taratufolo, M.C., Vinatzer, B.A. and Mazzaglia, A.** (2013) A multiplex PCR assay for detection of *Pseudomonas syringae* pv. *actinidiae* and differentiation of populations with different geographic origin. *Plant Dis.* **97**, 472–478.
- Baltrus, D.A., Nishimura, M.T., Romanchuk, A., Chang, J.H., Mukhtar, M.S., Cherkis, K., Roach, J., Grant, S.R., Jones, C.D., Dangl, J.L.** (2011) Dynamic evolution of pathogenicity revealed by sequencing and comparative genomics of 19 *Pseudomonas syringae* isolates. *PLoS Pathogens* **7**, e1002132.
- Baltrus, D.A., Yourstone, S., Lind, A., Guibaud, C., Sands, D.C., Jones, C.D., Morris, C.E. and Dangl, J.L.** (2013) Draft genome sequences of phylogenetically diverse suite of *Pseudomonas syringae* strains from multiple source populations. *Genome Announc.* **2**, 01195–13.
- Bastas, K.K., Karakaya, A.** (2012) First Report of bacterial canker of kiwifruit caused by *Pseudomonas syringae* pv. *actinidiae* in Turkey. *Plant Dis.* **96**, 452–452.

Chapter 5 - A framework to gage the epidemic potential of plant pathogens in environmental reservoirs: the example of kiwifruit canker

Bebber, D.P., Ramotowski, M. A. T. and Gurr, S.J. (2013) Crop pests and pathogens move polewards in a warming world. *Nat. Clim. Chang.* **3**, 985–988.

Chakraborty, S. (2013) Migrate or evolve: options for plant pathogens under climate change. *Glob. Chang. Biol.* **19**, 1985–2000.

Cuppels, D.M. (1986) Generation and characterization of Tn5 insertion mutations in *Pseudomonas syringae* pv. *tomato*. *Appl. Environ. Microbiol.* **51**, 323–327.

Díaz, E. (2004) Bacterial degradation of aromatic pollutants: a paradigm of metabolic versatility. *Int. Microbiol. Off. J. Spanish Soc. Microbiol.* **7**, 173–180.

Dröge, M., Pühler, A. and Selbitschka, W. (1998) Horizontal gene transfer as a biosafety issue: a natural phenomenon of public concern. *J. Biotechnol.* **64**, 75–90.

Elena, S.F., Bedhomme, S., Carrasco, P., Cuevas, J.M., de la Iglesia, F., Lafforgue, G., Lalić, J., Pròsper, A., Tromas, N., Zwart and M.P. (2011) The evolutionary genetics of emerging plant RNA viruses. *Molecular plant-microbe interactions Mol Plant Microbe Interact* **24**, 287–93.

European Plant Protection Organization (2011a) First report of *Pseudomonas syringae* pv. *actinidiae* in Switzerland EPPO Reporting Service n°8 2011/168. Available at www.eppo.org [accessed on February 14,2012].

European Plant Protection Organization (2011b) First report of *Pseudomonas syringae* pv. *actinidiae* in Chile EPPO Reporting Service n°3 2011/055. Available at www.eppo.org [accessed on Dicember 14,2013].

European Plant Protection Organization (2013) First report of *Pseudomonas syringae* pv. *actinidiae* in Germany. EPPO Reporting Service n°9, 2013/185. Available at www.eppo.org [accessed on January 17,2013].

Everett, K.R., Taylor, R.K., Romberg, M.K., Rees-George, J., Fullerton, R.A., Vanneste, J.L. and Manning, M.A. (2011) First report of *Pseudomonas syringae* pv. *actinidiae* causing kiwifruit bacterial canker in New Zealand. *Australas. Plant Dis. Notes* **6**, 67–71.

Chapter 5 - A framework to gauge the epidemic potential of plant pathogens in environmental reservoirs: the example of kiwifruit canker

Ferrante, P., Clarke, C.R., Cavanaugh, K.A., Micheltore, R.W., Buonauro, R. and Vinatzer, B.A. (2009) Contributions of the effector gene *hopQ1-1* to differences in host range between *Pseudomonas syringae* pv. *phaseolicola* and *P. syringae* pv. *tabaci*. *Mol. Plant Pathol.* **10**, 837–842.

Frias-Lopez, J., Klaus, J.S., Bonheyo, G.T., Fouke, B.W. (2004) Bacterial community associated with black band disease in corals. *Appl. Environ. Microbiol.* **70**, 5955–62.

Giovanardi, D., Tontou, R., Facchini, C. and Stafani, E. (2012) Endophytes and epiphytes from *Actinidia* spp. as potential agents for the biocontrol of the bacterial canker of kiwifruit. *J Plant Pathol* **94**, 64.

Gladieux, P., Zhang, X.G., Afoufa-Bastien, D., Sanhueza, R.M.V., Sbaghi, M. and Le Cam, B. (2008) On the origin and spread of the scab disease of apple: out of central Asia. *PLoS ONE* **3**, e1455.

Goel, A.K., Lundeberg, D., Torres, M.A., Matthews, R., Akimoto-Tomiyama, C., Farmer, L., Dangel, J.L. and Grant, S.R. (2008) The *Pseudomonas syringae* type III effector HopAM1 enhances virulence on water-stressed plants. *Mol. Plant. Microbe Interact.* **21**, 361–370.

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

Green, S., Studholme, D.J., Laue, B.E, Dorati, F., Lovell, H., Arnold, D., Cottrell, J.E., Bridgett, S., Blaxter, M., Huitema, E., Thwaites, R., Sharp, P.M., Jackson, R.W. and Kamoun, S. (2010) Comparative genome analysis provides insights into the evolution and adaptation of *Pseudomonas syringae* pv. *aesculi* on *Aesculus hippocastanum*. *PloS ONE* **5**, e10224.

Harwood, C.S., Parales, R.E. (1996) The β -ketoadipate pathway and the biology of self-identity. *Annu. Rev. Microbiol.* **50**, 553–590.

International Plant Protection Convention (2011) Detection of the Asian strain of bacterial canker of kiwifruit in Victoria, Australia IPPC AUS-45/1. Available at www.ippc.int [accessed on October 2013]

Chapter 5 - A framework to gauge the epidemic potential of plant pathogens in environmental reservoirs: the example of kiwifruit canker

Kang, M.H., Park, J.M. (1997) Sequential degradation of phenol and cyanide by a commensal interaction between two microorganisms. *J. Chem. Technol. Biotechnol.* **69**, 226–230.

Koh, Y.J., Nou, I.S. (2002) DNA markers for identification of *Pseudomonas syringae* pv. *actinidiae*. *Mol. Cells.* **13**, 309–314.

Lamichhane, J.R., Fabi, A., Ridolfi, R. and Varvaro, L. (2013) Epidemiological study of hazelnut bacterial blight in central Italy by using laboratory analysis and geostatistics. *PloS ONE* **8**, e56298.

Lamichhane, J.R., Varvaro, L. (2013) Epiphytic *Pseudomonas savastanoi* pv. *savastanoi* can infect and cause olive knot disease on *Olea europaea* subsp. *cuspidata*. *Australas. Plant Pathol.* **42**, 219–225.

Lamichhane, J.R., Varvaro, L., Parisi, L., Audergon, J-M. and Morris, C.E. (2014) Disease and frost damage of woody plants caused by *Pseudomonas syringae*: seeing the forest for the trees. *Adv Agron.* In press.

Manceau, C., Gardan, L. and Devaux, M. (1986) Dynamics of RP4 plasmid transfer between *Xanthomonas campestris* pv. *corylina* and *Erwinia herbicola* in hazelnut tissues, in planta. *Can. J. Microbiol.* **32**, 835–841.

Mazzaglia, A., Studholme, D.J., Taratufolo, M.C., Cai, R., Almeida, N.F., Goodman, T., Guttman, D.S., Vinatzer, B.A. and Balestra, G.M. (2012) *Pseudomonas syringae* pv. *actinidiae* (PSA) isolates from recent bacterial canker of kiwifruit outbreaks belong to the same genetic lineage. *PLoS ONE* **7**, e36518.

McCann, H.C., Rikkerink, E.H.A., Bertels, F., Fiers, M., Lu, A., Rees-George, J., Andersen, M.T., Gleave, A.P., Haubold, B., Wohlers, M.W., Guttman, D.S., Wang, P.W., Straub, C., Vanneste, J., Rainey, P.B. and Templeton, M.D. (2013) Genomic analysis of the kiwifruit pathogen *Pseudomonas syringae* pv. *actinidiae* provides insight into the origins of an emergent plant disease. *PLoS Pathogens* **9**, e1003503.

Chapter 5 - A framework to gauge the epidemic potential of plant pathogens in environmental reservoirs: the example of kiwifruit canker

Monteil, C., Cai, R., Liu, H., Mehan, L.M., Leman, S., Studholme, D., Morris, C.E., Vinaterz, B.A. (2013) Non-agricultural reservoir contribute to emergence and evolution of *Pseudomonas syringae* crop pathogens. *New phytologist* **199**, 800-811.

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

Morris, C.E., Monteil, C.L. and Berge, O. (2013) The life history of *Pseudomonas syringae*: linking agriculture to earth system processes. *Ann Rev Phytophatol* **51**, 85–104.

Morris, C.E., Sands, D.C., Vanneste, J.L., Montarry, J., Oakley, B., Guilbaud, C. and Glaux, C (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**, e00107–10.

Morris, C.E., Sands, D.C., Vinatzer, B.A., Glaux, C., Guilbaud, C., Buffière, A., Yan, S., Dominguez, H. and Thompson, B.M. (2008) The life history of the plant pathogen *Pseudomonas syringae* is linked to the water cycle. *The ISME journal* **2**, 321–334.

Oi-Kano, Y., Kawada, T., Watanabe, T., Koyama, F., Watanabe, K., Senbongi, R., and Iwai K. (2008) Oleuropein, a phenolic compound in extra virgin olive oil, increases uncoupling protein 1 content in brown adipose tissue and enhances noradrenaline and adrenaline secretions in rats. *J Nutr Sci Vitaminol* **54**, 363-370.

Opgenorth, D., Lai, M., Sorrell, M., Whithe, J.B. (1983) *Pseudomonas* canker of kiwifruit. *Plant Dis* **67**, 1283–1284.

Rees-George, J., Vanneste, J.L., Cornish, D.A., Pushparajah, I.P.S., Yu, J., Templeton, M.D. and Everett, K.R. (2010) Detection of *Pseudomonas syringae* pv. *actinidiae* using polymerase chain reaction (PCR) primers based on the 16S-23S rDNA intertranscribed spacer region and comparison with PCR primers based on other gene regions. *Plant Pathol.* **59**, 453–464.

Ritter, C., Dangl, J.L. (1995) The *avrRpm1* gene of *Pseudomonas syringae* pv. *maculicola* is required for virulence on Arabidopsis. *Mol Plant Microbe Interact* **3**, 444–453.

Chapter 5 - A framework to gage the epidemic potential of plant pathogens in environmental reservoirs: the example of kiwifruit canker

Rodríguez-Palenzuela, P., Matas, I.M., Murillo, J., López-Solanilla, E., Bardaji, L., Pérez-Martínez, I., Rodríguez-Mosquera, M.E., Penyalver, R., López, M.M., Quesada, J.M., Biehl, B.S., Perna, N.T., Glasner, J.D., Cabot, E.L., Neeno-Eckwall, E. and Ramos, C. (2010) Annotation and overview of the *Pseudomonas savastanoi* pv. *savastanoi* NCPPB 3335 draft genome reveals the virulence gene complement of a tumour-inducing pathogen of woody hosts. *Environ microbiol* **12**, 1604–1620.

Ronquist, F. and Huelsenbeck, J.P. (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**, 1572–1574.

Santini, A., Ghelardini, L., De Pace, C., Desprez-Loustau ML, Capretti P, Chandelier A, Cech T, Chira D, Diamandis S, Gaitniekis T, Hantula, J., Holdenrieder, O., Jankovsky, L., Jung, T., Jurc, D., Kirisits, T., Kunca, A., Lygis, V., Malecka, M., Marcais, B., Schmitz, S., Schumacher, J., Solheim, H., Solla, A., Szabò, I., Tsopelas, P., Vannini, A., Vettraino, A.M., Webber, J., Woodward, S., and Stenlid, J. (2013) Biogeographical patterns and determinants of invasion by forest pathogens in Europe. *The New Phytol.* **197**, 238–50.

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.

Sawada, H., Suzuki, F., Matsuda, I., Saitou, N. (1999) Phylogenetic analysis of *Pseudomonas syringae* pathovars suggests the horizontal gene transfer of *argK* and the evolutionary stability of *hrp* gene cluster. *J. Mol. Evol.* **49**, 627–644.

Scortichini, M. (1994) Occurrence of *Pseudomonas syringae* pv. *actinidiae* on kiwifruit in Italy. *Plant Pathol.* **43**, 1035–1038.

Scortichini, M., Marcelletti, S., Ferrante, P., Petriccione, M. and Firrao, G. (2012) *Pseudomonas syringae* pv. *actinidiae*: a re-emerging, multi-faceted, pandemic pathogen. *Mol Plant Pathol.* **13**, 631–640.

Sridevi, V., Lakshmi, M.V.V.C., Manasa, M. and Sravani, M. (2012) Metabolic pathways for the biodegradation of phenol. *Int. J. Eng. Sci. Adv. Technol.* **2**, 695–705.

Chapter 5 - A framework to gage the epidemic potential of plant pathogens in environmental reservoirs: the example of kiwifruit canker

Takikawa, Y., Serizawa, S.I., Takeshi, Y.S. and Goto, M. (1989) *Pseudomonas syringae* pv. *actinidiae* pv. nov.: the casual bacterium of canker of kiwifruit in japan. *Annu. Phytopathol. Soc. Japan* **55**, 437–444.

Vanneste, J. (2012) *Pseudomonas syringae* pv. *actinidiae* (Psa): a threat to the New Zealand and global kiwifruit industry. *New Zeal. J. Crop Hortic. Sci.* **40**, 265–267.

Vanneste, J.L., Poliakoff, F., Audusseau, C., Cornish, D.A., Paillard, S., Rivoal, C. and Yu, J. (2011) First report of *Pseudomonas syringae* pv. *actinidiae* , the causal agent of bacterial canker of kiwifruit in France. *Plant Dis.***95**, 1311–1311.

Vanneste, J.L., Yu, J., Cornish, D.A., Tanner, D.J., Windner, R., Chapman, J.R., Taylor, R.K., Mackay, J.F. and Dowlut, S. (2013) Identification, virulence and distribution of two biovars of *Pseudomonas syringae* pv. *actinidiae* in New Zealand. *Plant Dis.* **97**, 708-719.

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

Williams, P. A., Sayers, J.R. (1994) The evolution of pathways for aromatic hydrocarbon oxidation in *Pseudomonas*. *Biodegradation* **5**, 195–217.

Wilson, M., Lindow, S.E. (1994) Coexistence among epiphytic bacterial populations mediated through nutritional resource partitioning. *Appl. Environ. Microbiol.* **60**, 4468–77.

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

Chapter 5 - A framework to gage the epidemic potential of plant pathogens in environmental reservoirs: the example of kiwifruit canker

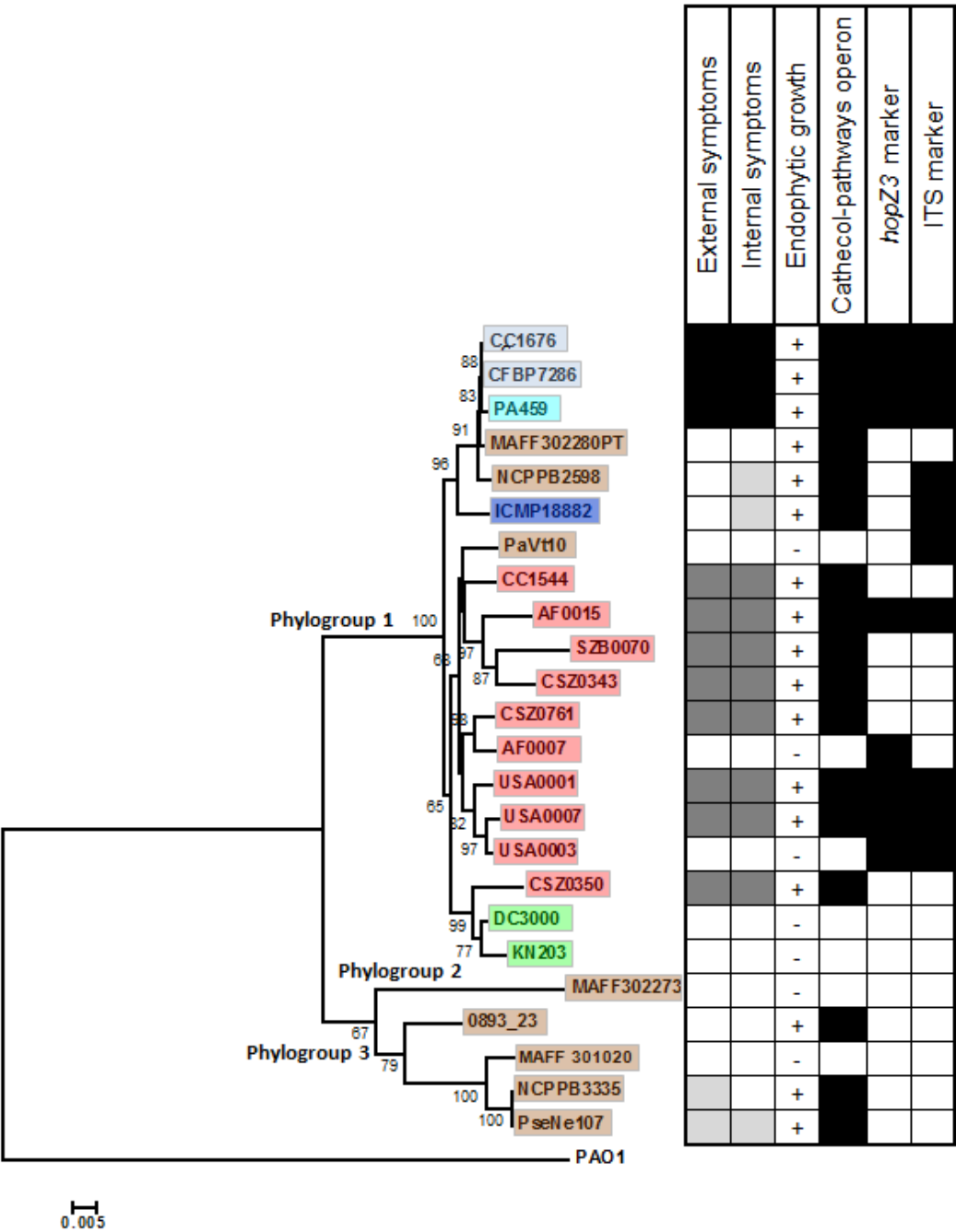


Figure 1. Bayesian tree constructed on the concatenated sequences *cts*, *gyrB*, *rpoD* and *gapA* (1852bp) of the strains tested for pathogenicity, endophytic growth and molecular markers. Bootstrap values are indicated at each node. External and internal symptoms are indicated with black boxes (severe disease), dark grey boxes (mild symptoms), light grey boxes (faint symptoms) and white boxes (no symptoms). The presence of the catechol operon involved in aromatic compound degradation, of *hopZ3* and ITS markers, is indicated in black (present) and white (absent) squares. For the catechol operon all the genes were present for the positive strains. The origin of strains and their biovar affiliation was indicated with a colour legend: light blue for Psa bv.3,

Chapter 5 - A framework to gage the epidemic potential of plant pathogens in environmental reservoirs: the example of kiwifruit canker

turquoise for *Psa* bv.1, dark blue for *Psa* bv.4, brown for strains isolated from woody hosts, green for strains isolated from herbaceous hosts and red for the environmental strains.



Figure 2. Characteristic external symptoms observed on kiwifruit stems at 30 days post inoculation. External necrosis and canker caused by the strain from *Actinidia deliciosa* CC 1676 (A), necrosis caused by the environmental strain AF0015 (B), swelling caused by the strain NCPPB 3335 isolated from *Olea europea* (C) and the absence of symptoms on stems inoculated with strain ICMP 18882 isolated from *Actinidia deliciosa*.

Chapter 5 - A framework to gage the epidemic potential of plant pathogens in environmental reservoirs: the example of kiwifruit canker

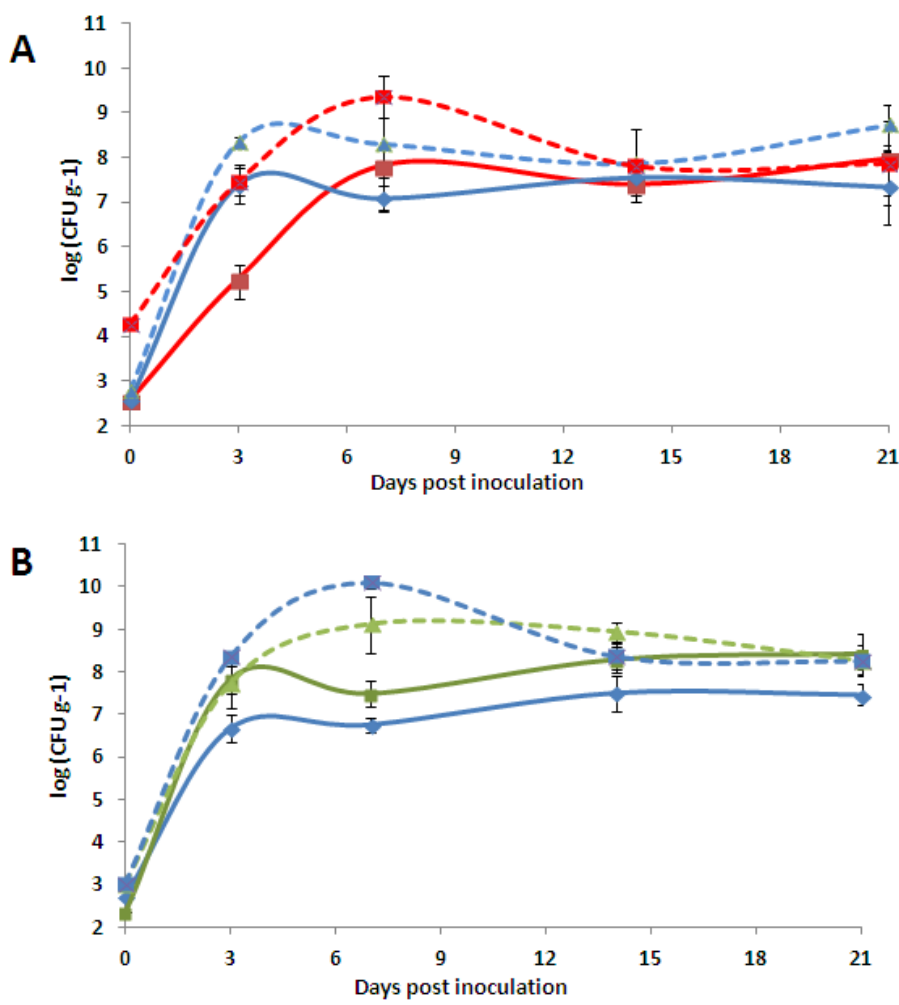


Figure 3. Growth of strains in 3 week-old kiwifruit plants. Population densities were determined at the point of the inoculation for strains CFBP 7286 (blue) and USA007 (red) of *P. syringae* (A) and for strains CFBP 7286 (blue) and CC1544 (green) (B). Dotted lines indicate the growth of each strain independently, and solid lines represent growth for mixed inoculations. Mean $\pm$ standard error at each time point is indicated.

Chapter 5 - A framework to gage the epidemic potential of plant pathogens in environmental reservoirs: the example of kiwifruit canker



Figure 4. External symptoms recorded on kiwifruit plants 3 weeks after inoculation. The same symptoms were observed on all plants (3 per each time point). Necrosis observed with strain CFBP 7286 of *P. syringae* (A). No symptoms were recorded on kiwifruit 3 weeks after co-inoculation with environmental strains (USA0007 and CC1544) and CFBP 7286 (B).

Chapter 5 - A framework to gage the epidemic potential of plant pathogens in environmental reservoirs: the example of kiwifruit canker

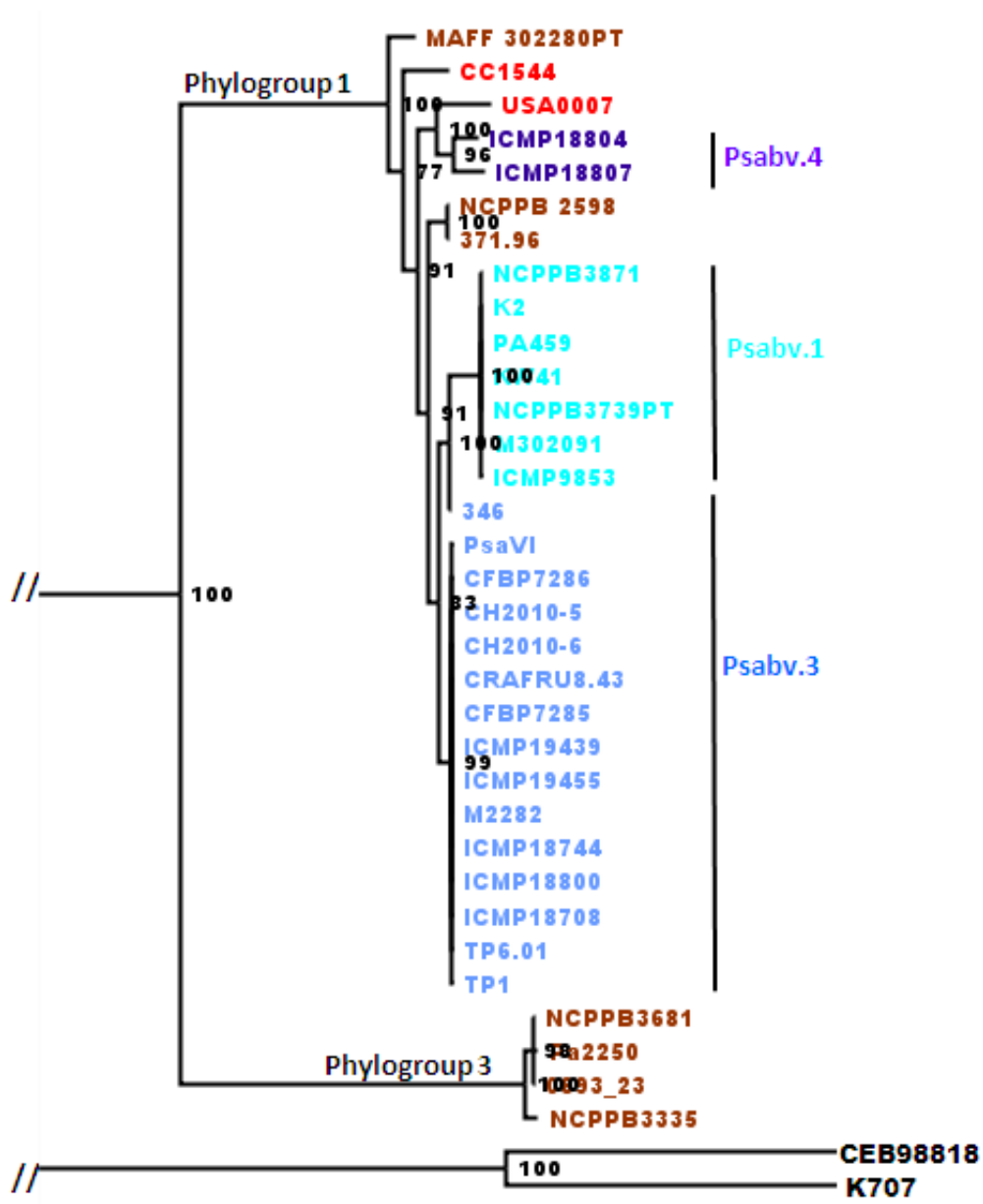


Figure 5. Bayesian tree of the concatenated complete sequences of phenol-MetA, FAD-dependent oxygenase, short-chain alcohol dehydrogenase and diene lacton hydrolase. Posterior probabilities are shown at each node. Phylogroup 1 and 3, Psa biovars and names of the strains are indicated on branches. The tree was rooted with the sequences extracted from the genomes of strain KF707 of *Pseudomonas pseudoalcaligenes* and strain CEB98818 of *Pseudomonas fuscovaginae*. The origin of strains and their biovar affiliation was indicated with a color legend: light blue for Psa bv.3, turquoise for Psa bv.1, dark blue for Psa bv.4, brown for strains isolated from woody hosts, green for strains isolated from herbaceous hosts and red for the environmental strains.

Chapter 5 - A framework to gage the epidemic potential of plant pathogens in environmental reservoirs: the example of kiwifruit canker

		External symptoms				Internal symptoms				Endophytic growth			
Host species		CFBP 7286	ICMP 18882	USA 0007	CC1544	CFBP 7286	ICMP 18882	USA 0007	CC1544	CFBP 7286	ICMP 18882	USA 0007	CC1544
Woody plants	Manicot apricot												
	Pontaleb cherry												
	Myrobolan plum												
	Mariana GF8-1 plum												
	Monclar peach												
	Rubira peach												
	Hazelnut												
	Oleander												
	Poplar												
Herbaceous plants	Soybean*												
	Faba bean*												
	Cowpea												
	Clover*												
	Sorghum												
	Tomato												
	Melon												
	Sunflower *												
	Spinach												
	Geranium*												
	Mustard												
	Ranunculus*												
	Liverwort												

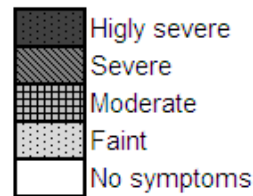


Figure 6. Host range of strains CFBP 7286, ICMP 18882, USA0007 and CC1544. Severity of the disease is expressed with a grey color scale. For endophytic growth, black boxes indicate that strains grow into the host plant and that they were re-isolated even in absence of symptoms.

For strains pathogenic on soybean, faba bean, clover, sunflower, geranium, and ranunculus, disease was observed only on stems (*). Psa bv.3 is indicated in blue, Psa bv.4 in dark blue and environmental strains with red color.

Chapter 5 - A framework to gage the epidemic potential of plant pathogens in environmental reservoirs: the example of kiwifruit canker

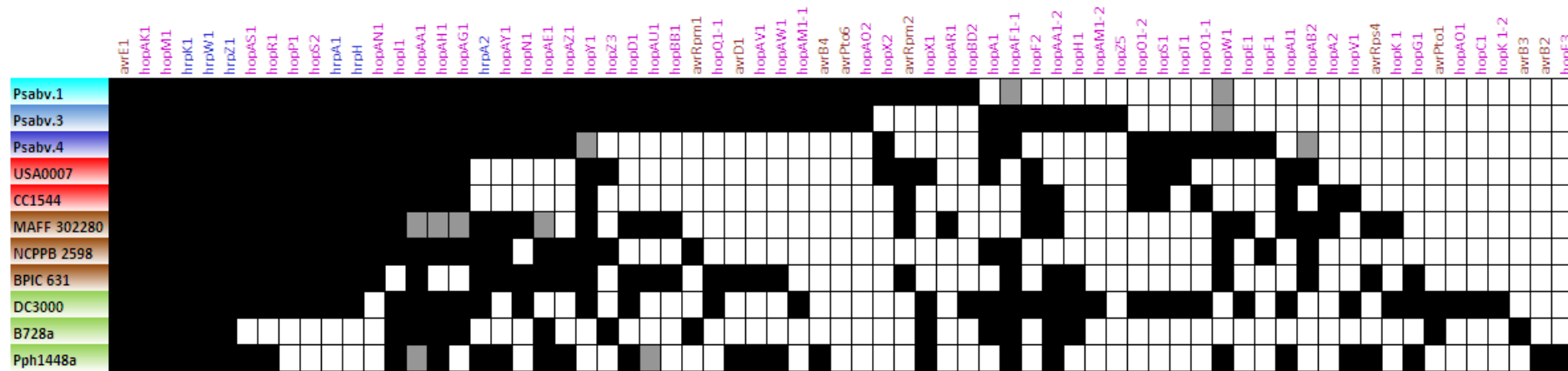


Figure 7. Type III secretion gene repertoires in strains of *P. syringae*. Results for Psabv.1, Psabv.3, Psabv.4 were compressed since the results were identical for all the strains analyzed (all the genomes available on GeneBank). Strains USA0007 and CC1544 were isolated from river headwaters. Other strains indicated include MAFF 302280 (*P. s. pv. morsprunorum*), NCPPB 2598 (*P. s. pv. theae*), BPIC 631 (*P. s. pv. avellanae*), DC3000 (*P. s. pv. tomato*), B728a (*P. s. pv. syringae*) and Pph1448a (*P. s. pv. phaseolicola*). Black boxes indicate that the genes were found in full length, white boxes indicate the absence of the genes and grey boxes indicate that genes were truncated. The origin of strains and their biovar affiliation was indicated with a colour legend: light blue for Psabv.3, turquoise for Psabv.1, dark blue for Psabv.4, brown for strains isolated from woody hosts, green for strains isolated from herbaceous hosts and red for the environmental strains.

Supplementary Information

Supplementary Method S1.

Primers, PCR conditions and sequencing

To amplify and sequence the *hopZ3* gene and the ITS region we used previously described primers and PCR conditions (Rees-George et al., 2010; Balestra et al., 2013) whereas for the *hopM1* gene primers and amplification conditions were developed in this study (Table 1). Four genes involved in the degradation of phenolic and aromatic compounds, *via* the catechol branch β -ketoadipate pathway, were reported in the draft genomes of strains attacking woody plant species including *P. syringae* pv. *aesculi* strain NCPPB 3681 (Green et al., 2010) and *P. savastanoi* pv. *savastanoi* strain NCPPB 3335 (Rodríguez-Palenzuela et al., 2010) and are in the draft genomes of *Psa* genomes available on GeneBank (accession numbers: NZ_AGNO01000000, NZ_AEAL01000000, NZ_AGNQ01000000) and in *P. syringae* pv. *morsprunorum* (MAFF 302280PT). These genes were found to be present in these genomes but absent from *P. syringae* pathogens of herbaceous plants using the ortholog sorter tool at <http://pacu.facom.ufms.br/kiwi/orthologsorter/>. Primer pairs were designed for these genes with Primer 3 software (<http://frodo.wi.mit.edu>). Primers synthesized by Eurofins MWG Operon (Ebersberg, Germany) are listed in Table 1. For all PCR reactions, the mix for each primer set was prepared by using the GoTaq Flexi DNA Polymerase (Promega). Sequencing was conducted by Eurofins MWG Operon and by Macrogen Europe (Amsterdam, the Netherlands).

Chapter 5 - A framework to gage the epidemic potential of plant pathogens in environmental reservoirs: the example of kiwifruit canker

Table 1. Primers used in this study.

Primer name	Gene or protein product	Sequence (5'-3')	Annealing T.	References
Phenol MetA F	Protein involved in meta pathway phenol degradation	cagaattacagcgctgatcg	60°C	This study
Phenol MetA R		ctgttcgaagtacgcgttga		
Put_oxy_sub F	FAD-dependent-oxidoreductases	gtcacttgccgagggcta	60°C	This study
Put_oxy_sub R		caggctactggcgaacgattt		
Dienelactone F	Dienelactone hydrolase	ggaagacagttccaggcctac	60°C	This study
Dienelactone R		gtctcgaactcatggcgaat		
Short_CH F	Short chain-dehydrogenase	cgacactgatcgccatgg	60°C	This study
Short_CH R		gtcacgacgttggccact		
hopM1 R	<i>hopM1</i>	ahgcgvccrgaaagctckgc	58	This study
hopM1 F		gacawacgcacawarcccamgt		
PSA F	<i>hopZ3</i>	cagaggcgctaacaggaaaa	60°C	(Balestra et al., 2013)
PSA R		cgagcatacatcaacagggtca		
PsaF1	ITS	ttttgctttgcacaccgatttt	65°C	(Rees-George et al., 2010)
PsaR2		cacgccccttcaatcaggatg		

Endophytic growth in kiwifruit vascular tissue

For endophytic assay in kiwifruit plants, inoculum consisted of aqueous suspensions of 48 h cultures grown on KB and adjusted to 10^4 CFU ml⁻¹. For the co-inoculation treatments, each environmental strain was combined in the inoculum at equal concentrations with the epidemic strain for a final inoculum concentration of 10^4 CFU ml⁻¹.

Plants were inoculated by injecting 20 µl of bacterial inoculum (pure or mixed-culture suspensions) into a petiole as described above. For each bacterial inoculum, 60 plants were inoculated. Bacterial population sizes in kiwifruit tissue were determined at 3, 7, 14 and 21 dpi by re-isolating bacteria from 3 plants at each date. Plant fragments of ca. 1 cm were cut with a sterilized scalpel at three places on the stem relative to the point of inoculation: at the point of inoculation, 1 cm above and 1 cm below from the point of inoculation. Each portion was separately weighed and chopped with a sterile scalpel in 500 µl of sterile distilled water. The resulting suspensions were serially diluted and plated on KB medium supplemented with 50 mg l⁻¹ of cycloheximide. Plates were incubated at room temperature for 3 days and bacterial colonies were counted. The epidemic strain CFBP 7286 was discriminated from the environmental strains because the former was not fluorescent on KB medium whereas the latter strains were fluorescent. For each treatment, all colonies on dilution plates with about 30 to 300

colonies were transferred to KB and checked for fluorescence. The validity of discrimination based on pigment production was determined by BOX PCR as described above. For the co-inoculation of CFBP 7286 and CC1544, because of the faint fluorescence of strain CC1544, BOX PCR was performed for all colonies at each time point. On the contrary, for the combination of CFBP 7286 and USA0007, fluorescence of USA0007 was a reliable marker and BOX PCR was conducted only on randomly selected colonies. Bacterial populations from each isolated point were expressed as CFU g⁻¹ of fresh plant weight.

To compare competition of strains to their behavior in defined culture medium, 10 % strength tryptic soy broth (DifoTM Detroit, MI, the USA) was inoculated with strains at a final total concentration of 10³ CFU ml⁻¹. For mixed inoculations, each strain constituted half of the total bacterial concentration in the medium. Inoculated media were incubated with agitation at 26° C and population sizes were determined after 14, 19, 24 and 48 h of incubation by dilution-plating on KB medium. Colonies of the different strains were discriminated as described above. Three replicates of each growth curve were realized.

Host range determination

For the host range test, strains of *P. syringae* were streaked on KB agar from glycerol stocks preserved at -80°C. The inoculum was prepared as described above. On annual plants, two different techniques of inoculation were performed. On true leaves, small drops consisting of 10 µl of bacterial suspension were deposited on upper leaf surfaces without causing any lesion. The number of drops varied from 1 to 10 depending on the leaf size. On stems, 10 µl of bacterial suspension were deposited on the lesions (3 in total, 5 cm apart) made by pricking the tissue with a sterile hypodermic needle. On woody plants, inoculation was performed as described above. For each treatment, four plants per species were tested. On each plant species, reference strains reported to cause disease were used as positive controls (Table S3) while sterile distilled water was used as the negative control. All of the inoculated plants were maintained in the greenhouse at ambient temperature and watered twice a day without wetting the leaves.

Annual plants were examined at 5, 10 and 15 dpi. The plants were examined individually for disease symptoms and disease scores at 15 dpi were used for analysis. Leaves and stems on a plant were scored according to the following scale: no symptoms (apparently healthy leaf and stem), faint symptoms (< 5 lesions/leaf with no symptoms on stems), moderate symptoms (5-10 lesions/leaf and/or < 5 mm necrosis, canker and/or gall on stem), severe symptoms (> 10

Chapter 5 - A framework to gage the epidemic potential of plant pathogens in environmental reservoirs: the example of kiwifruit canker

lesions/leaf and/or >5-10 mm necrosis, canker and/or gall on stem) and highly severe symptoms (> 10 lesions/leaf and > 10 mm necrosis, canker and/or gall on stem).

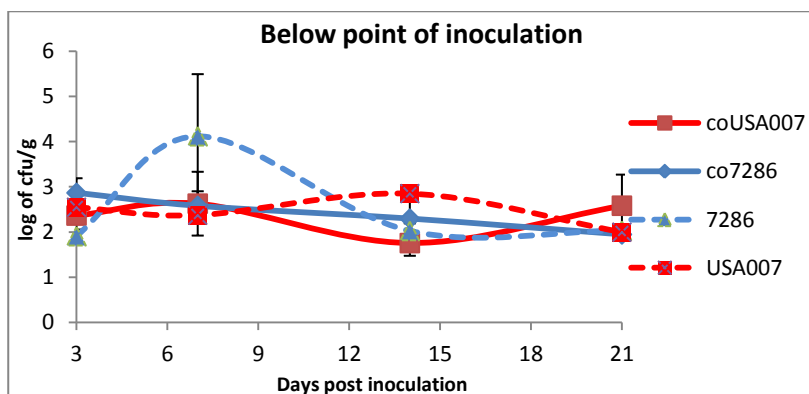
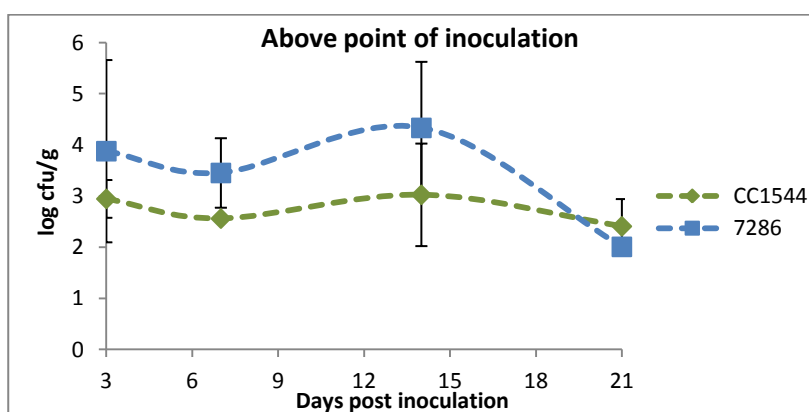
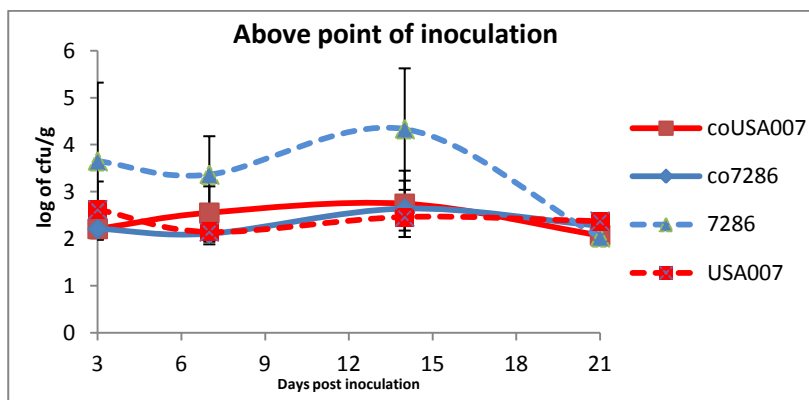
For woody plants, external disease symptoms were assessed at 15, 30, 45 and 60 dpi. The plants were examined individually for external and internal disease symptoms and disease scores at 60 dpi were used for analysis. Disease severity was rated by slightly modifying the disease scale described above. No symptoms (apparently healthy external and internal tissues), faint symptoms (< 0.20 cm), moderate symptoms (0.21-1 cm), severe symptoms (1.1-3 cm) and highly severe symptoms (> 3 cm). For two of the inoculated plants per species and per strain, bacteria were re-isolated from the tissues and their BOX-PCR profiles were compared to those of the strain used for inoculation as described above.

References

1. Balestra GM, Taratufolo MC, Vinatzer BA, Mazzaglia A (2013) A multiplex PCR assay for detection of *Pseudomonas syringae* pv. *actinidiae* and differentiation of populations with different geographic origin. *Plant Dis* 97:472–478.
2. Rees-George J et al. (2010) Detection of *Pseudomonas syringae* pv. *actinidiae* using polymerase chain reaction (PCR) primers based on the 16S-23S rDNA intertranscribed spacer region and comparison with PCR primers based on other gene regions. *Plant Pathol* 59:453–464.
3. Green S 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.
4. Rodríguez-Palenzuela P et al. (2010) Annotation and overview of the *Pseudomonas savastanoi* pv. *savastanoi* NCPPB 3335 draft genome reveals the virulence gene complement of a tumour-inducing pathogen of woody hosts. *Env microbiol* 12:1604–20.
5. Versalovic J, Koeuth T, Lupski J (1991) Distribution of repetitive DNA-sequences in eubacteria and application to fingerprinting of bacterial genomes. *Nucleic Acids Res* 19:6823–6826.

Chapter 5 - A framework to gage the epidemic potential of plant pathogens in environmental reservoirs: the example of kiwifruit canker

6. Scortichini M, Marchesi U, Dettori MT, Rossi MP (2003) Genetic diversity, presence of the *syxB* gene, host preference and virulence of *Pseudomonas syringae* pv. *syringae* strains from woody and herbaceous host plants. *Plant Pathol* 52:277–286.



Chapter 5 - A framework to gage the epidemic potential of plant pathogens in environmental reservoirs: the example of kiwifruit canker

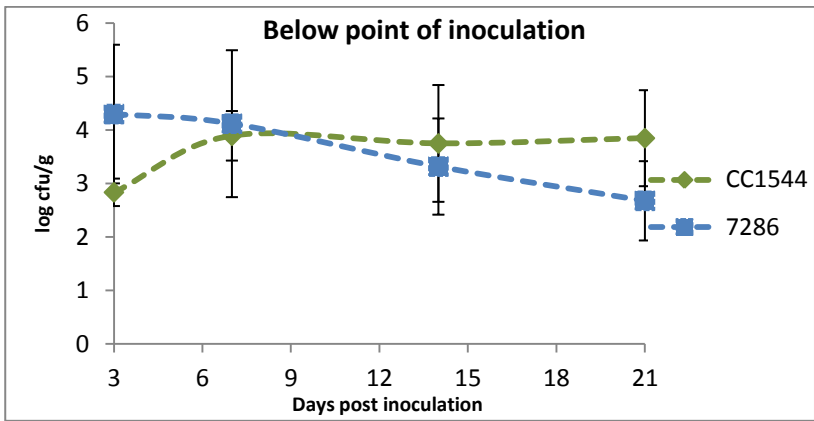


Figure S1. Endophytic population density of CFBP 7286, CC1544 and USA0007 when co-inoculated and alone in kiwifruit stem taken above and below the point of inoculation.

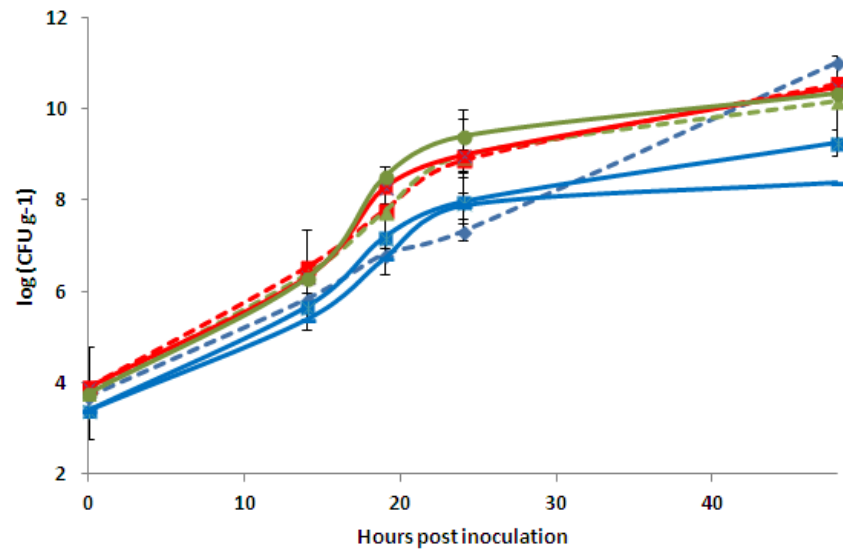


Figure S2. In vitro growth of strains CFBP 7286 (blue), USA007 (red) and CC1544 (green) of *P. syringae* in 10% strength tryptic soy broth. Solid lines indicate the growth of each strain independently, and dotted lines represent growth in mixed culture. Mean $\pm$ standard error at each time point is indicated.

Chapter 5 - A framework to gage the epidemic potential of plant pathogens in environmental reservoirs: the example of kiwifruit canker

Figure S3. Phylogenetic comparison between housekeeping genes and the genes forming the catechol operon.

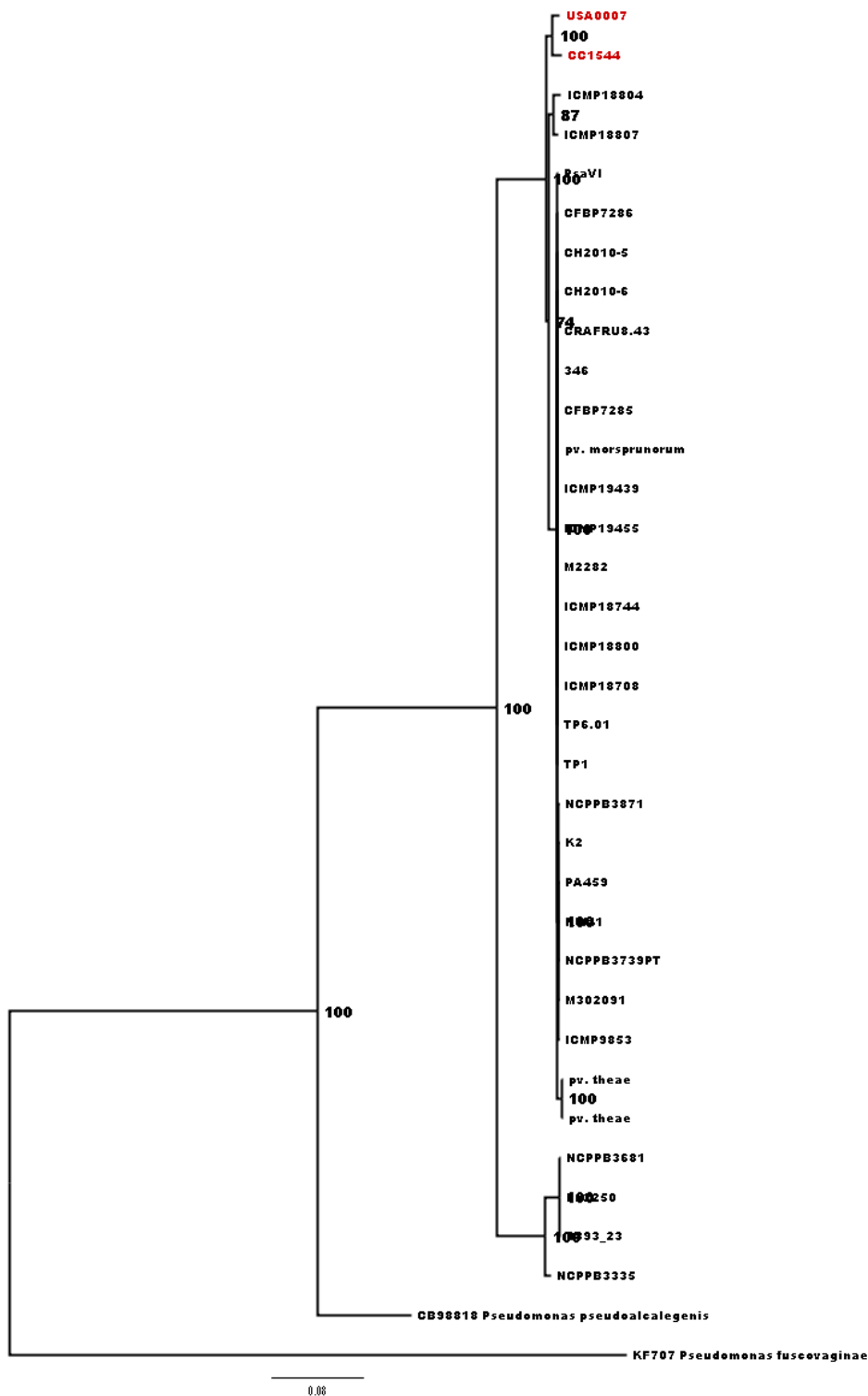


Figure S3 (A). Bayesian tree based on the concatenated *cts*, *gyrB*, *rpoD*, and *gapA* sequences. The tree shows that the phylogeny of the housekeeping genes reflects the phylogeny of the genes involved in the degradation of aromatic compounds.

Chapter 5 - A framework to gage the epidemic potential of plant pathogens in environmental reservoirs: the example of kiwifruit canker

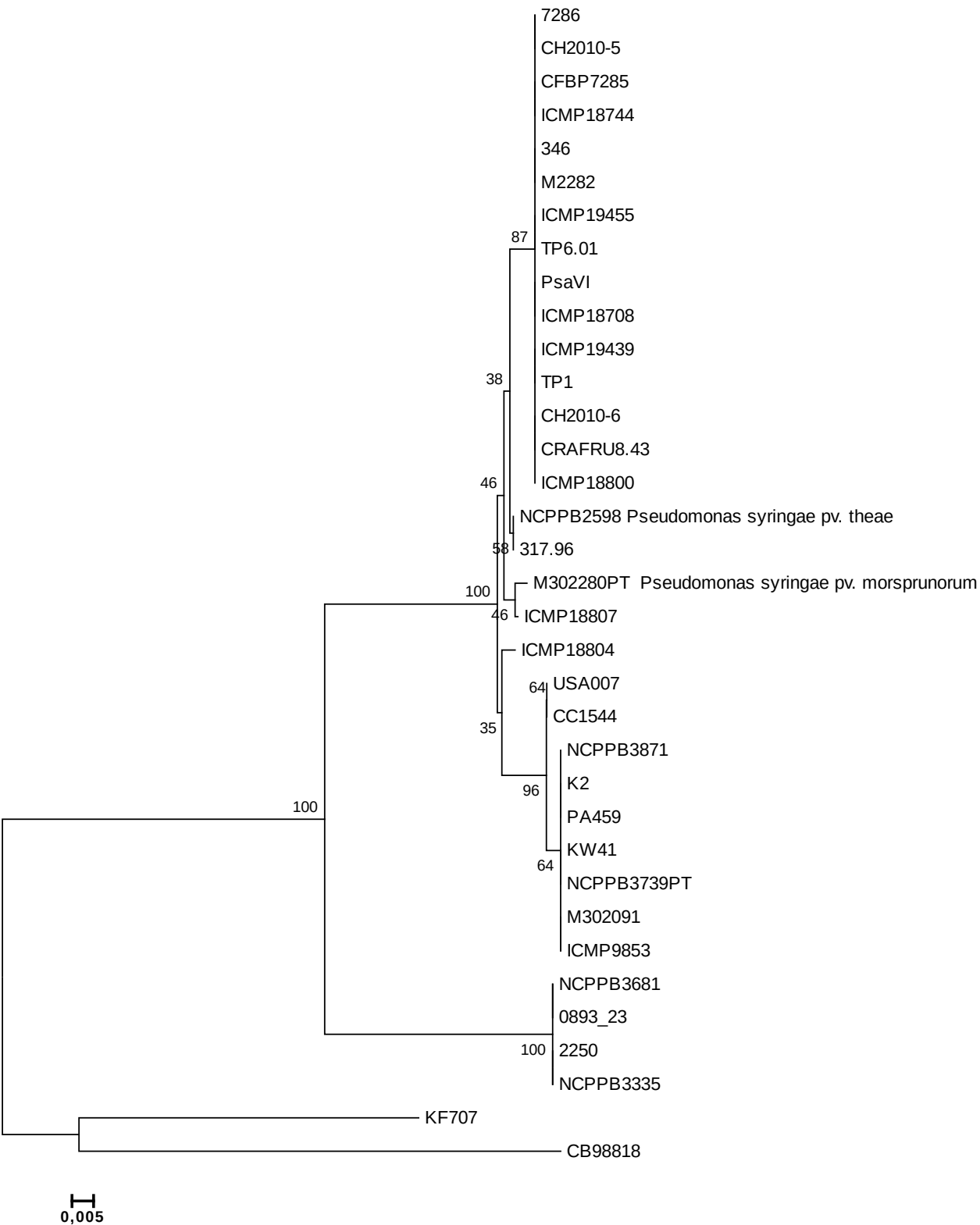


Figure S3 (B). Neighbor-joining tree constructed on the short-chain dehydrogenase gene sequences.

Chapter 5 - A framework to gage the epidemic potential of plant pathogens in environmental reservoirs: the example of kiwifruit canker

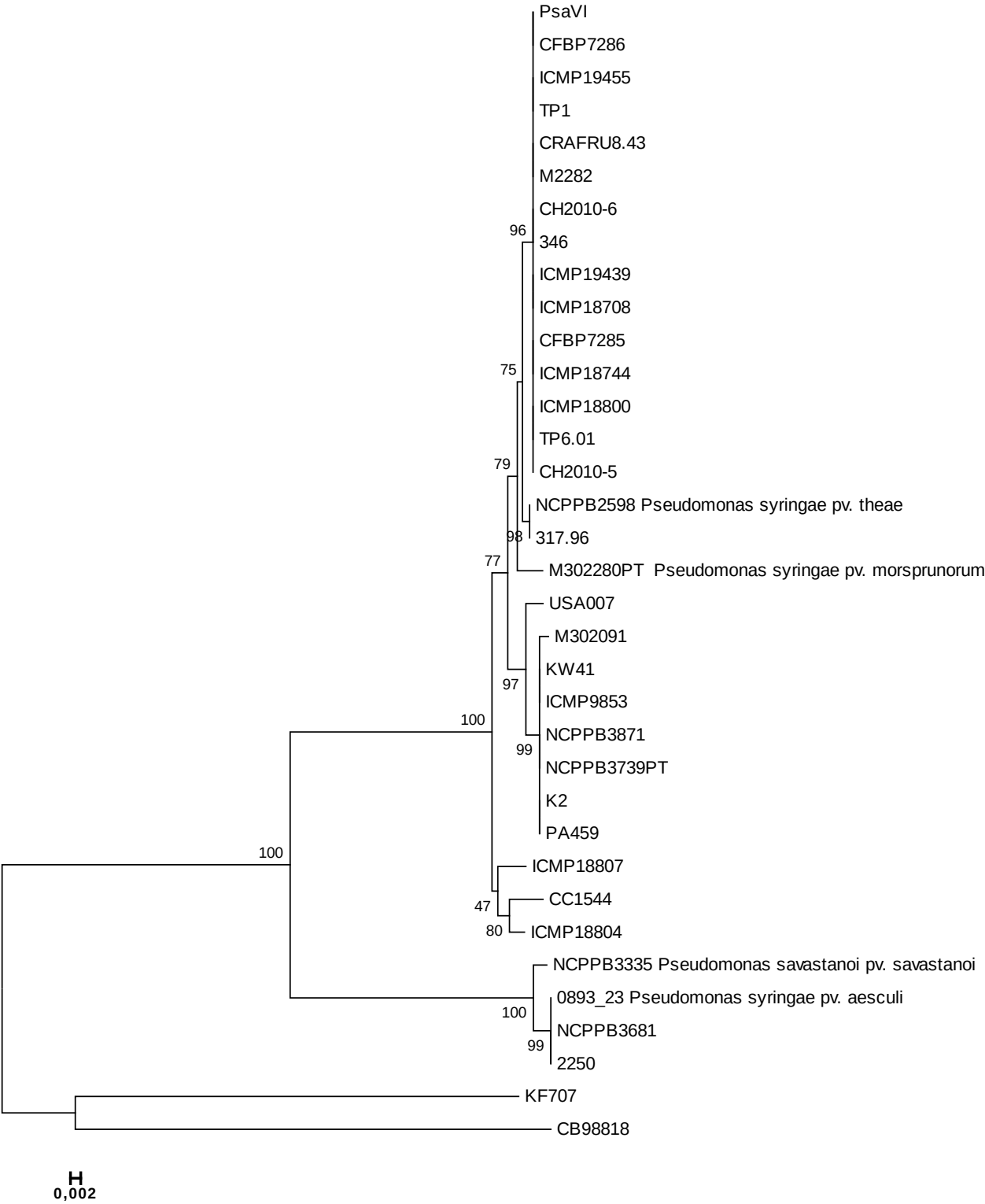


Figure S3 (C). Neighbor-joining tree constructed on the complete gene sequence of FAD-dependent oxidoreductase.

Chapter 5 - A framework to gage the epidemic potential of plant pathogens in environmental reservoirs: the example of kiwifruit canker

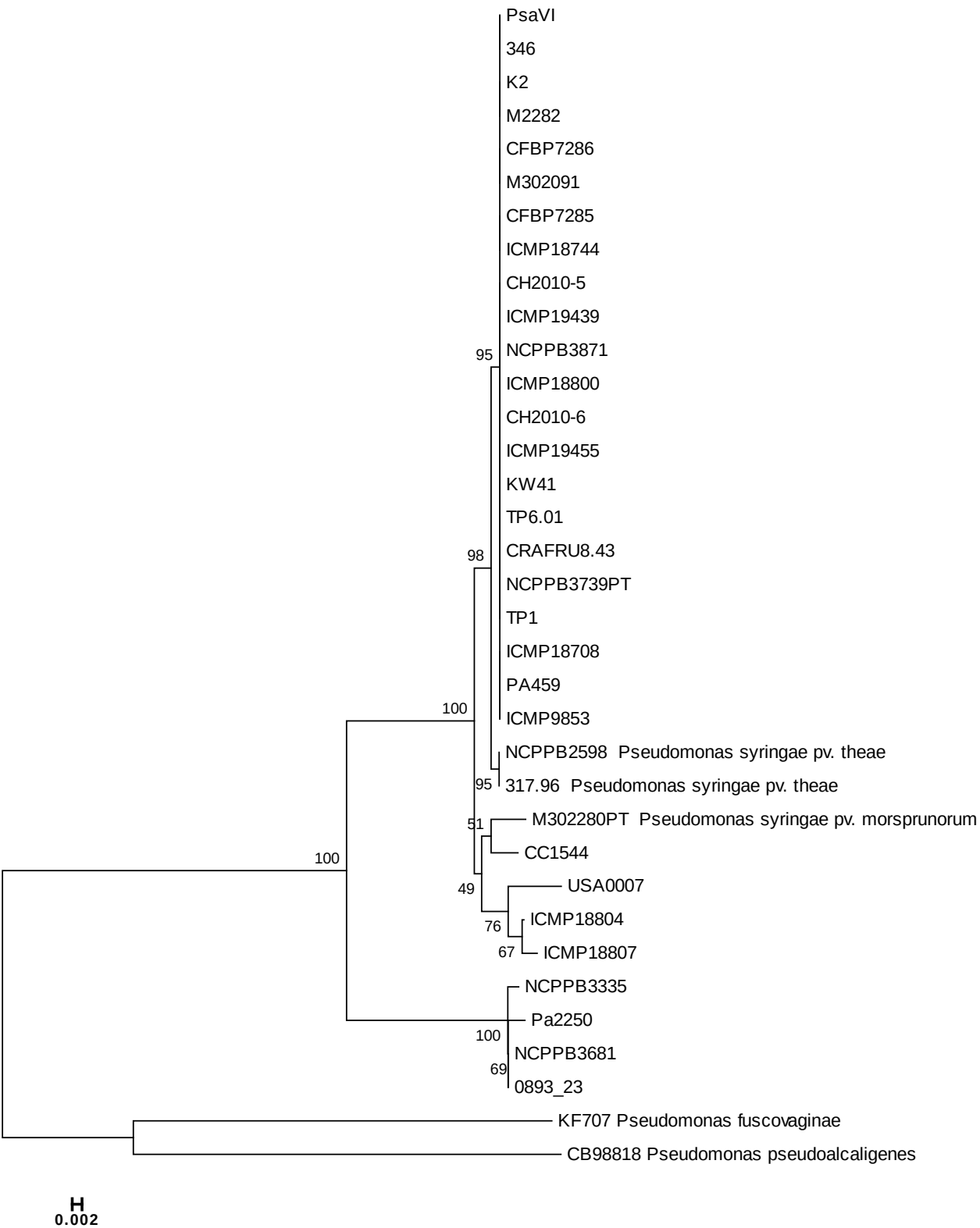


Figure S3 (D). Neighbor-joining tree constructed on the complete gene sequence of meta-pathway phenol degradation.

Chapter 5 - A framework to gage the epidemic potential of plant pathogens in environmental reservoirs: the example of kiwifruit canker

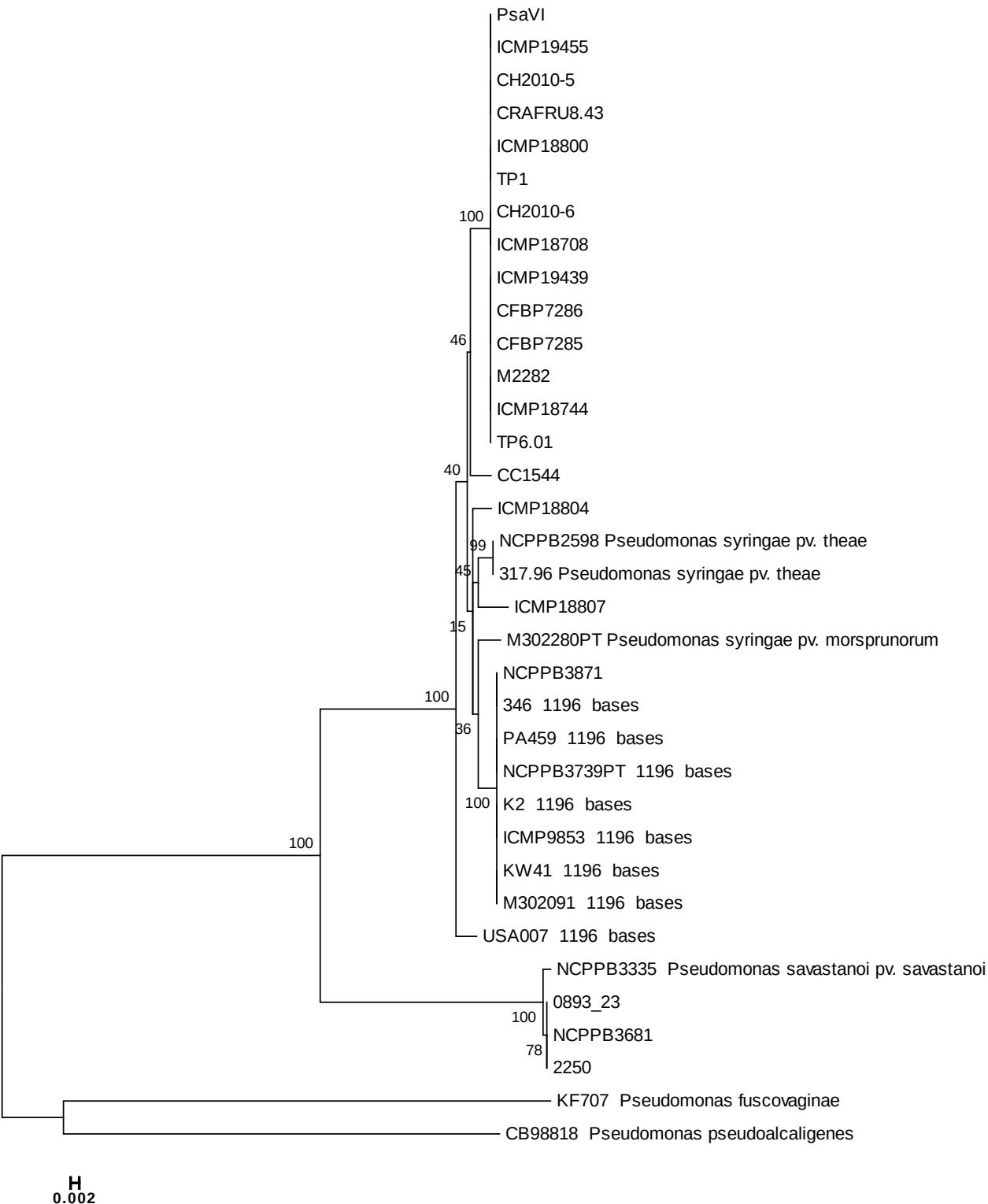


Figure S3 (E). Neighbor-joining constructed on the complete sequence of the dienelacton hydrolase gene.

Chapter 5 - A framework to gage the epidemic potential of plant pathogens in environmental reservoirs: the example of kiwifruit canker

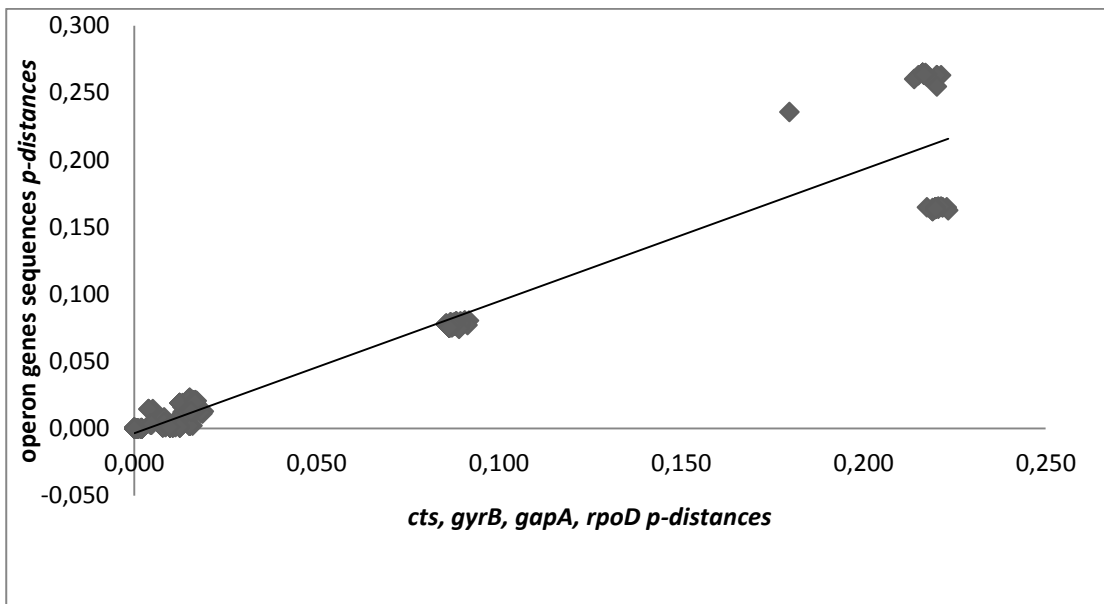


Figure S3 (F). Graphical representation of the relationships between the *p-distances* of the concatenated sequences located in the operon for the degradation of aromatic compounds and the *p-distances* of the concatenated housekeeping genes *cts*, *gyrB*, *gapA* and *rpoD*. The *p-distance* were calculated with MEGA. The points representing the values clearly showed the linear correlation between the gene sequences.

Chapter 5 - A framework to gage the epidemic potential of plant pathogens in environmental reservoirs: the example of kiwifruit canker

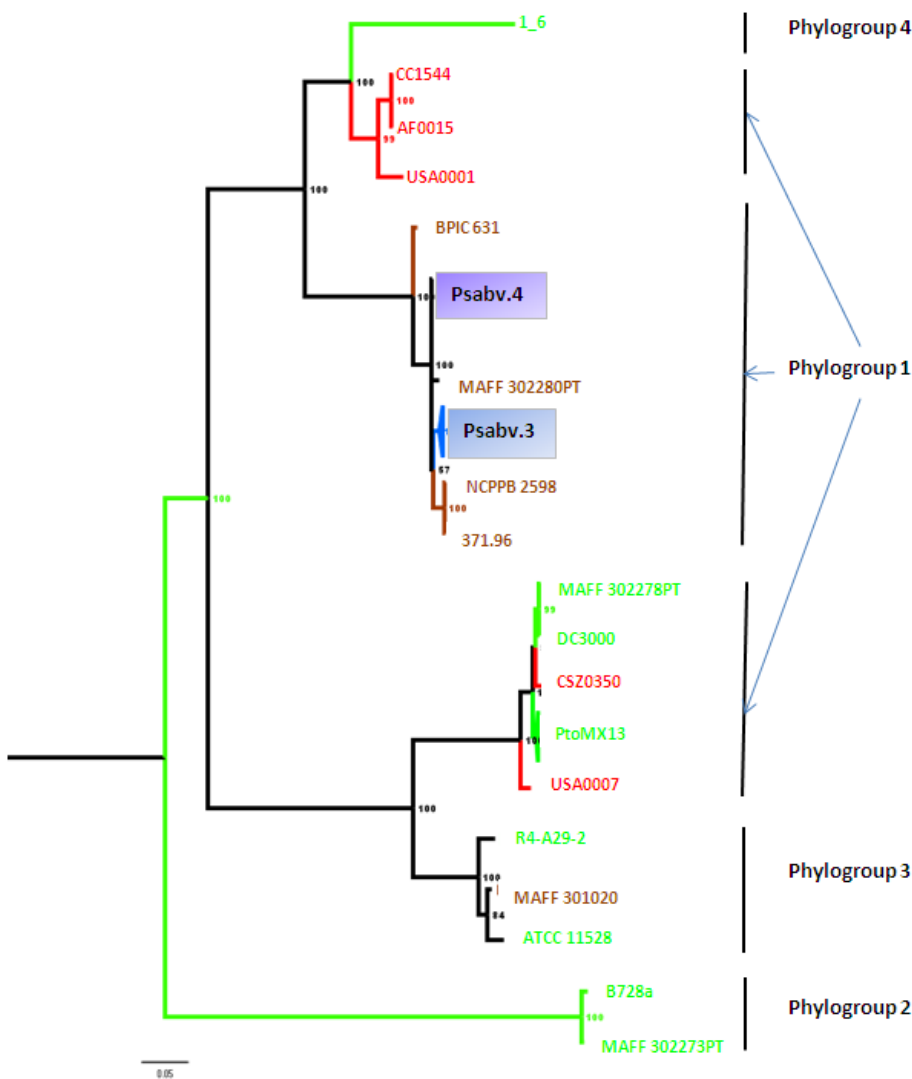


Figure S4. Bayesian phylogeny of the *hopM1* gene. Names of the strains are indicated on branches and posterior probabilities are indicated at each node. Taxon names are coded with a color legend: light blue for *Psa* bv.3, dark blue for *Psa* bv.4, brown for strains isolated from woody hosts, green for strains isolated from herbaceous hosts and red for the environmental strains. Phylogroups are indicated at the right of the tree based on the housekeeping gene phylogeny (Fig. 1). The phylogroup 1 is formed by three polyphyletic clades and the environmental strains are distributed along different clades and a new clade was observed for the strains AF0015, CC1544 and USA0001.

Chapter 5 - A framework to gage the epidemic potential of plant pathogens in environmental reservoirs: the example of kiwifruit canker

Table S1. Accession numbers of strains used in this study.

Strain Name	Cts	GapA	GyrB	RpoD	PhenolMetA	Dienelacton	Short-chain	Oxygenase	Reference or Source
T10_05188	JN683500	JN683481	JN683462	JN683404	–	–	–	–	Chapman <i>et al.</i> , (2012)
346	http://pacu.facom.ufms.br/blast_kiwi/blast.cgi								Mazzaglia <i>et al.</i> , (2012)
28830	JX287994	JX288011	JX288028	JX287911	–	–	–	–	Chapman <i>et al.</i> , (2012)
4649.1	FN651806	FN651852	FN651834	FN651798	–	–	–	–	Ferrante <i>et al.</i> , (2010)
0893_23	AEAD00000000	AEAD00000000	AEAD00000000	AEAD00000000	AEAD00000000	AEAD00000000	AEAD00000000	AEAD00000000	Baltrus <i>et al.</i> , (2011)
2871B5	JX287992	JX288009	JX288026	JX287909	–	–	–	–	Chapman <i>et al.</i> , (2012)
287A63	JX287990	JX288007	JX288024	JX287907	–	–	–	–	Chapman <i>et al.</i> , (2012)
4252A.1	FN652850	FN652929	FN652896	FN652876	–	–	–	–	Ferrante <i>et al.</i> , (2010)
B728a	NC_007005	NC_007005	NC_007005	NC_007005	–	–	–	–	Feil <i>et al.</i> , (2005)
CB98818	ALAQ00000000	ALAQ00000000	ALAQ00000000	ALAQ00000000	ALAQ00000000	ALAQ00000000	ALAQ00000000	ALAQ00000000	Xie <i>al.</i> , (2012)
CFBP 7286	AGNO00000000	AGNO00000000	AGNO00000000	AGNO00000000	AGNO00000000	AGNO00000000	AGNO00000000	AGNO00000000	Mazzaglia <i>et al.</i> , 2012
CH2010-6	AGUH01000004	AGUH01000004	AGUH01000300	AGUH01000248	AGUH00000000	AGUH00000000	AGUH00000000	AGUH00000000	Mazzaglia <i>et al.</i> , (2012)
CRA-FRU 1.2	FN652851	FN652931	FN652891	FN652871	–	–	–	–	Ferrante <i>et al.</i> , (2010)
CRA-FRU 10.10	FN652848	FN652927	FN652888	FN652868	–	–	–	–	Ferrante <i>et al.</i> , (2010)
CRA-FRU 10.14	FN651805	FN651851	FN651833	FN651797	–	–	–	–	Ferrante <i>et al.</i> , (2010)
CRA-FRU 10.15	FN652854	FN652934	FN652894	FN652874	–	–	–	–	Ferrante <i>et al.</i> , (2010)
CRA-FRU 10.16	FN652855	FN652935	FN652895	FN652875	–	–	–	–	Ferrante <i>et al.</i> , (2010)
CRA-FRU 10.22	JX287997	JX288014	FN651832	FN651796	–	–	–	–	Ferrante <i>et al.</i> , (2010)
CRA-FRU 10.22	JX287997	JX288014	JX288031	JX287914	–	–	–	–	Ferrante <i>et al.</i> , (2010)
CRA-FRU 10.24	FN652856	FN652936	FN652897	FN652877	–	–	–	–	Ferrante <i>et al.</i> , (2010)
CRA-FRU 10.25	FN652862	FN652930	FN652890	FN652870	–	–	–	–	Ferrante <i>et al.</i> , (2010)
CRA-FRU 10.26	FN652843	FN652922.1	FN652883	FN652863	–	–	–	–	Ferrante <i>et al.</i> , (2010)
CRA-FRU 10.28	FN652849	FN652928	FN652889	FN652869	–	–	–	–	Ferrante <i>et al.</i> , (2010)
CRA-FRU 10.5	FN652853	FN652933	FN652893	FN652873	–	–	–	–	Ferrante <i>et al.</i> , (2010)
CRA-FRU 10.6	FN652847	FN652926	FN652887	FN652867	–	–	–	–	Ferrante <i>et al.</i> , (2010)
CRA-FRU 11.40	JX287999	JX288016	JX288033	JX287916	–	–	–	–	Chapman <i>et al.</i> , (2012)

Chapter 5 - A framework to gage the epidemic potential of plant pathogens in environmental reservoirs: the example of kiwifruit canker

CRA-FRU 11.41	JX288000	JX288017	JX288034	JX287917	—	—	—	—	Chapman <i>et al.</i> , (2012)
CRA-FRU 11.42	JX288001	JX288018	JX288035	JX287918	—	—	—	—	Chapman <i>et al.</i> , (2012)
CRA-FRU 11.43	JX288002	JX288019	JX288036	JX287919	—	—	—	—	Chapman <i>et al.</i> , 2012
CRA-FRU 3.1	FN652852	FN652932	FN652892	FN652872	—	—	—	—	Ferrante <i>et al.</i> , (2010)
CRA-FRU 5.1	FN652857	FN652937	FN652898	FN652878	—	—	—	—	Ferrante <i>et al.</i> , (2010)
CRA-FRU 8.43	FN651803	FN651849	FN651831	FN651795	AFTG00000000	AFTG00000000	AFTG00000000	AFTG00000000	Ferrante <i>et al.</i> , (2010)
CRA-FRU 8.51	FN652845	FN652924	FN652885	FN652865	—	—	—	—	Ferrante <i>et al.</i> , (2010)
CRA-FRU 8.52	FN652844	FN652923	FN652884	FN652864	—	—	—	—	Ferrante <i>et al.</i> , (2010)
CRA-FRU 8.57	FN652846	FN652925	FN652886	FN652866	—	—	—	—	Ferrante <i>et al.</i> , (2010)
CRA-FRU 8.69	FN652858	FN652938	FN652899	FN652879	—	—	—	—	Ferrante <i>et al.</i> , (2010)
CRA-FRU 8.75	FN652859	FN652939	FN652900	FN652880	—	—	—	—	Ferrante <i>et al.</i> , (2010)
CRA-FRU 8.76	JX287998	JX288015	JX288032	JX287915	—	—	—	—	Chapman <i>et al.</i> , (2012)
CRA-FRU 8.80	FN651802	FN651848	FN651830	FN651794	—	—	—	—	Ferrante <i>et al.</i> , (2010)
DC3000	NC_004578	NC_004578	NC_004578	NC_004578	—	—	—	—	Buel <i>et al.</i> , (2003)
ICMP18708	ANJC00000000	ANJC00000000	ANJC00000000	ANJC00000000	ANJC00000000	ANJC00000000	ANJC00000000	ANJC00000000	Butler <i>et al.</i> , (2013)
ICMP18744	ANGD00000000	ANGD00000000	ANGD00000000	ANGD00000000	ANGD00000000	ANGD00000000	ANGD00000000	ANGD00000000	Butler <i>et al.</i> , (2013)
ICMP18800	ANJD00000000	ANJD00000000	ANJD00000000	ANJD00000000	ANJD00000000	ANJD00000000	ANJD00000000	ANJD00000000	Butler <i>et al.</i> , (2013)
ICMP18804	ANJE00000000	ANJE00000000	ANJE00000000	ANJE00000000	ANJE00000000	ANJE00000000	ANJE00000000	ANJE00000000	Butler <i>et al.</i> , (2013)
ICMP18807	ANJL00000000	ANJL00000000	ANJL00000000	ANJL00000000	ANJL00000000	ANJL00000000	ANJL00000000	ANJL00000000	Butler <i>et al.</i> , (2013)
ICMP19439	ANJM00000000	ANJM00000000	ANJM00000000	ANJM00000000	ANJM00000000	ANJM00000000	ANJM00000000	ANJM00000000	Butler <i>et al.</i> , (2013)
ICMP19455	ANJK00000000	ANJK00000000	ANJK00000000	ANJK00000000	ANJK00000000	ANJK00000000	ANJK00000000	ANJK00000000	Butler <i>et al.</i> , (2013)
ICMP9853 (KW1)	ANJB00000000	ANJB00000000	ANJB00000000	ANJB00000000	ANJB00000000	ANJB00000000	ANJB00000000	ANJB00000000	Butler <i>et al.</i> , (2013)
I-Psa Z	JX288003	JX288020	JX288037	JX288020	—	—	—	—	Chapman <i>et al.</i> , (2012)
I-Psa8	JX288004	JX288021	JX288038	JX287921	—	—	—	—	Chapman <i>et al.</i> , (2012)
KACC 10584	JN683505	JN683486	JN683467	JN683410	—	—	—	—	Chapman <i>et al.</i> , (2012)
KACC10594	JN683506	JN683487	JN683468	JN683410	—	—	—	—	Chapman <i>et al.</i> , (2012)
KACC10754	JN683507	JN683488	JN683469	JN683412	—	—	—	—	Chapman <i>et al.</i> , (2012)
KF707	AJMR01000113	AJMR01000113	AJMR01000113	AJMR01000113	AJMR01000113	AJMR01000113	AJMR01000113	AJMR01000113	Tiscari <i>et al.</i> , (2012)
KN.2	http://pacu.facom.ufms.br/blast_kiwi/blast.cgi				http://pacu.facom.ufms.br/blast_kiwi/blast.cgi				Mazzaglia <i>et al.</i> , (2012)

Chapter 5 - A framework to gage the epidemic potential of plant pathogens in environmental reservoirs: the example of kiwifruit canker

KN203	AY610722	AY610662	AY610782	AY610902	—	—	—	—	Morris <i>et al.</i> , (2010)
KW30	FN652861	FN652941	FN652901	FN652882	—	—	—	—	Chapman <i>et al.</i> , (2012)
KW31	FN652860	FN652940	FN652902	FN652881	—	—	—	—	Chapman <i>et al.</i> , (2012)
KW41	AGNP00000000	AGNP00000000	AGNP00000000	AGNP00000000	AGNP00000000	AGNP00000000	AGNP00000000	AGNP00000000	Chapman <i>et al.</i> , (2012)
M228	ANJI00000000	ANJI00000000	ANJI00000000	ANJI00000000	ANJI00000000	ANJI00000000	ANJI00000000	ANJI00000000	Butler <i>et al.</i> , (2013)
MAFF302091	JN683508	JN683489	JN683470	JN683412	AEAL00000000	AEAL00000000	AEAL00000000	AEAL00000000	Baltrus <i>et al.</i> , (2011)
MAFF302143	JN683509	JN683490	JN683471	JN683413	—	—	—	—	Chapman <i>et al.</i> , (2012)
MAFF302145	JN683510	JN683491	JN683472	JN683414	—	—	—	—	Chapman <i>et al.</i> , (2012)
MAFF302280PT	AEAE00000000	AEAE00000000	AEAE00000000	AEAE00000000	AEAE00000000	AEAE00000000	AEAE00000000	AEAE00000000	Baltrus <i>et al.</i> , 2011 Sarkar and Guttman (2004)
NCPPB2598	AGNN00000000	AGNN00000000	AGNN00000000	AGNN00000000	AGNN00000000	AGNN00000000	AGNN00000000	AGNN00000000	Rodriguez <i>et al.</i> , (2010)
NCPPB3335	ADMI01000151	ADMI01000094	ADMI01000001	ADMI01000048	ADMI00000000	ADMI00000000	ADMI00000000	ADMI00000000	Green <i>et al.</i> , (2010)
NCPPB3681	ACXS00000000	ACXS00000000	ACXS00000000	ACXS00000000	ACXS00000000	ACXS00000000	ACXS00000000	ACXS00000000	Chapman <i>et al.</i> , (2012)
NCPPB3739	AFTH00000000	AFTH00000000	AFTH00000000	AFTH00000000	AFTH00000000	AFTH00000000	AFTH00000000	AFTH00000000	Scortichini <i>et al.</i> , (1994)
NCPPB3871	AFTF00000000	AFTF00000000	AFTF00000000	AFTF00000000	AFTF00000000	AFTF00000000	AFTF00000000	AFTF00000000	Scortichini <i>et al.</i> , (1994)
NCPPB3873	JX287995	JX288012	JX288029	JX287912	—	—	—	—	not published
<i>P. aeruginosa</i> PAO1	NC_002516	NC_002516	NC_002516	NC_002516	—	—	—	—	Stover <i>et al.</i> , (2000)
<i>P. fluorescens</i> Pf0-1	NC_007492	NC_007492	NC_007492	NC_007492	—	—	—	—	Green <i>et al.</i> , (2010)
Pa2250	ACXT00000000	ACXT00000000	ACXT00000000	ACXT00000000	ACXT00000000	ACXT00000000	ACXT00000000	ACXT00000000	Mazzaglia <i>et al.</i> , (2012)
PA459	AGNQ01000413	AGNQ01000595	AGNQ01000070	AGNQ01000464	AGNQ00000000	AGNQ00000000	AGNQ00000000	AGNQ00000000	Chapman <i>et al.</i> , (2012)
PSA 1A	JQ836616	JQ836620	JQ836624	JQ836579	—	—	—	—	Chapman <i>et al.</i> , (2012)
T10_04758	JN683493	JN683474	JN683455	JN683397	—	—	—	—	Chapman <i>et al.</i> , (2012)
T10_04782	JN683494	JN683475	JN683456	JN683398	—	—	—	—	Chapman <i>et al.</i> , (2012)
T10_04782	JN683494	JN683475	JN683456	JN683399	—	—	—	—	Chapman <i>et al.</i> , (2012)
T10_04976	JN683498	JN683479	JN683460	JN683402	—	—	—	—	Chapman <i>et al.</i> , (2012)
T10_05163	JN683499	JN683480	JN683461	JN683403	—	—	—	—	Chapman <i>et al.</i> , (2012)
T10_05195	JN683501	JN683482	JN683463	JN683405	—	—	—	—	Chapman <i>et al.</i> , (2012)
T10_05196	JN683502	JN683483	JN683464	JN683406	—	—	—	—	Chapman <i>et al.</i> , (2012)
T10_05454	JN683495	JN683476	JN683457	JN683400	—	—	—	—	Chapman <i>et al.</i> , (2012)
T10_05815	JX287988	JX288005	JX288022	JX287905	—	—	—	—	Chapman <i>et al.</i> , (2012)

Chapter 5 - A framework to gage the epidemic potential of plant pathogens in environmental reservoirs: the example of kiwifruit canker

T10_06115	JN683496	JN683477	JN683458	JN683401	–	–	–	–	Chapman <i>et al.</i> , (2012)
T11_01369A	JQ836618	JQ836622	JQ836626	JQ836581	–	–	–	–	Chapman <i>et al.</i> , (2012)
T11_01395A	JQ836619	JQ836623	JQ836627	JQ836582	–	–	–	–	Chapman <i>et al.</i> , (2012)
T11_01581B	JX287993	JX288010	JX288027	JX287910	–	–	–	–	Chapman <i>et al.</i> , (2012)
T11_0918	JN683497	JN683478	JN683459	JN683402	–	–	–	–	Chapman <i>et al.</i> , (2012)
TP1	ANJG00000000	ANJG00000000	ANJG00000000	ANJG00000000	ANJG00000000	ANJG00000000	ANJG00000000	ANJG00000000	Butler et al., (2013)
TP6.01	ANJH00000000	ANJH00000000	ANJH00000000	ANJH00000000	ANJH00000000	ANJH00000000	ANJH00000000	ANJH00000000	Butler et al., (2013)

Chapter 5 - A framework to gage the epidemic potential of plant pathogens in environmental reservoirs: the example of kiwifruit canker

Table S3. Plant species used in host range tests and bacterial strains used as positive controls.

Herbaceous plant species				Strain used as control	
Species name	Cultivar	Common name	Family	Strain no.	pv. designation
<i>Glicine max</i> (L) Merr	Paoki	Soybean	<i>Fabaceae</i>	CFBP 3211	<i>glycinea</i>
<i>Vicia faba</i> L.	Agua Dulce	Faba bean	<i>Fabaceae</i>	41A(France)	<i>syringae</i>
<i>Vigna unguiculata</i> (L) Walp.		Cowpea	<i>Fabaceae</i>	CFBP 3637	<i>phaseolicola</i>
<i>Trifolium pretense</i> L.		Clover	<i>Fabaceae</i>	41A	<i>syringae</i>
<i>Sorghum bicolor</i> L.	Solarius	Sorghum	<i>Poaceae</i>	MAFF 311018	<i>oryze</i>
<i>Solanum lycopersicum</i> L.	Monalbo	Tomato	<i>Solanaceae</i>	KN10	<i>tomato</i>
<i>Cuculis melo</i>	Védrantais	Melon	<i>Cucurbitaceae</i>	MAFF 301315	<i>lachrymans</i>
<i>Helianthus annus</i> L.	Solarius	Sunflower	<i>Asteraceae</i>	CFBP 2067	<i>helianthi</i>
<i>Spinacia oleracea</i> L	Giant d'hiver	Spinach	<i>Chenopodiaceae</i>	CFBP 1617	<i>aptata</i>
<i>Pelargonium hortorum</i>		Geranium	<i>Geraniaceae</i>	41A	<i>syringae</i>
<i>Sinapis alba</i> L	Bladue Architect	Mustard	<i>Brassicaceae</i>	CFBP 1657	<i>maculicola</i>
<i>Ranunculus bulbosus</i> L.		Bulbous buttercup	<i>Ranunculaceae</i>	41A	<i>syringae</i>
Non-vascular plants					
<i>Lunularia cruciata</i> (L.) Dumort		Liverwort	<i>Lunulariaceae</i>		
Woody plant species					
<i>Prunus armeniaca</i> L.		Manicot apricot	<i>Rosaceae</i>	41A	<i>syringae</i>
<i>Prunus mahaleb</i> L.		Pontaleb cherry	<i>Rosaceae</i>	41A	<i>syringae</i>
<i>Prunus cerasifera</i> L.		Myrobolan plum	<i>Rosaceae</i>	41A	<i>syringae</i>
<i>Prunus persica</i> (L.) Stokes		Montclar peach	<i>Rosaceae</i>	41A	<i>syringae</i>
<i>Prunus persica</i> L.Batsch		Rubira peach	<i>Rosaceae</i>	41A	<i>syringae</i>
<i>Prunus cerasifera</i> X <i>Prunus munsoniana</i>		Mariana GF8-1 plum	<i>Rosacea</i>		
<i>Corylus avellana</i> L.	Tonda Gentile Romana	hazelnut	<i>Corylaceae</i>	CFBP 4060	<i>avellanae</i>
<i>Nerium oleander</i> L.		Oleander	<i>Apocynaceae</i>	41A	<i>syringae</i>
<i>Populus alba</i> L.		Poplar	<i>Salicaceae</i>	41A	<i>syringae</i>

Chapter 6

**Hypermutableity in *Pseudomonas viridiflava*:
programed bet-hedging via antibiotic
resistance, exopolysaccharide production
and pathogenicity**

Manuscript in preparation

Hypermutability in *Pseudomonas viridiflava*: programed bet-hedging via antibiotic resistance, exopolysaccharide production and pathogenicity

Claudia Bartoli^{1,2}, Jay Ram Lamichanne^{1,2}, Odile Berge², Leonardo Varvaro¹ and Cindy E. Morris²

¹Departement of Science and Technology for Agriculture, Forestry, Nature and Energy (DAFNE), Tuscia University, 01100 Viterbo, Italy

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

Abstract

Adaptation to fluctuating environments depends on the production of offspring with a diversity of traits. Hypermutable bacterial cells are defective in their DNA repair system and they often have a phenotype different than their wild-type counterparts. In human bacterial pathogens, the hypermutable phenotype is often reported to be associated with antibiotic resistance. Here, we quantified the occurrence of phase variants in *Pseudomonas viridiflava*, a phytopathogenic bacterium in the *P. syringae* complex with a broad host range and capacity to live as a saprophyte. Two obvious phenotypic variants (transparent and mucoid) were produced by this bacterium. The transparent one was hypermutable, antibiotic resistant and could not induce disease on the plant species tested whereas the mucoid variant did not display hypermutability and resistance to antibiotics and was capable of inducing disease on a range of plant species. Both the transparent and mucoid variants were less fit when growing *in vitro*, while *in vivo* both of the variants and wild-types attained similar population densities. Given the importance of the methyl-directed mismatch repair system (MMR) in the occurrence of hypermutable cells in human bacterial pathogens, we investigated if mutations in *mut* genes were associated with phase variation in *P. viridiflava*. Our results showed no mutations in MMR genes in any of the variants tested. Studies on the regulation of *mut* genes would be the next step to elucidating whether MMR plays a role in phase variation in *P. viridiflava*.

Introduction

Bacterial adaptation to new environments or hosts is the consequence of both phenotypic and genotypic changes. Evolution of new genomic elements - by *de novo* mutations or other mechanisms that allow the gain and loss of genomic fragments, plasmids or the movement of insertion sequences - is the basis of bacterial variability which also permits adaptation (1). In human bacterial pathogens, bacterial genomic plasticity is often correlated with the ability to colonize new hosts or to become resistant to antibiotic treatments thereby seriously impacting human healthcare. Antibiotic resistant cells could also suffer a fitness cost in the absence of antibiotics. However, bacteria overcome this fitness cost through the evolution of new mutations which jeopardizes the effectiveness of disease control by antibiotic treatments (2). Indeed, the correlation between hypermutable bacterial cells and the development of antibiotic resistance has been described for different human pathogens such as *Escherichia coli*, *Pseudomonas aeruginosa*, *Staphylococcus aureus* and *Haemophilus influenzae* (3–6). For all these pathogens, a defective methyl-directed mismatch repair system (MMR) was the molecular basis that regulates the occurrence of hypermutable antibiotic resistance cells. The MMR system corrects the post-replication mistakes that escape from polymerase proof-reading after DNA replication (7). In *E. coli*, the MutS and MutL proteins bind the mismatches at the hemi-methylated GATC site and MutL further interacts with a third protein (MutH) thereby forming a network with the helicase II (7). Long term experiments on *P. aeruginosa* chronic lung infections demonstrated the formation of a conspicuous number of *mutS*, *mutL* and *mutY* hypermutators associated with antibiotic resistance (4). In addition, the *mutS* *P. aeruginosa* mutants are associated with oxidative-stress resistance because of their ability to produce catalase (8) suggesting that these cells are less susceptible to the oxygen and nitrogen radicals produced during the early stage of the host immune responses. Hence, in addition to antibiotic resistance, hypermutable cells produced by a defective MMR are also linked to resistance to several stress conditions.

Phenotypic plasticity or variation has also been reported to occur in several plant pathogens or plant-associated bacteria in addition to bacteria pathogenic to humans. However, these studies did not demonstrate a correlation between antibiotic resistance and hypermutability phenotypes or a correlation of hypermutability with pathogenicity. In *Pseudomonas brassicacearum*, the ability to resist high cadmium concentrations was correlated with the production of phenotypic variants under both *in vitro* and *in vivo* conditions (9). Since soil and the rhizosphere are

environments with high fluctuating conditions, phase variation can represent a strategy for adaptation to such fluctuating environments. The plant growth-promoting rhizobacterium *Azospirillum lipoferum*, for example, produces phase variants that differ in sugar assimilation, swimming and the production of nitrous oxide (10). Phase variants in *A. lipoferum* are regulated by the *recA* gene and the genomic rearrangements of this bacterium are concomitant with the occurrence of phase variation (10).

Studies concerning the occurrence of hypermutable bacterial cells focused mainly on clinical bacterial strains that are under continuous antibiotic treatments. This led to the assumption that the use of antibiotics is the main stimulating factor for the formation of resistant cells and for the evolution of secondary mutations able to compensate the fitness cost of the antibiotic resistance (2). However, the occurrence of hypermutable bacteria in natural populations is still poorly understood. Matic and coworkers (11) found that commensal *E. coli* strains were characterized by a strong mutation rate associated with a defective MMR. The authors suggested that all bacterial populations experience adaptive mutations leading to antibiotic resistance which in turn leads to adaptation to different environments not strictly related with the susceptible host. In general, the environment could have an important role in selecting mutator phenotypes in ubiquitous bacteria such as *P. aeruginosa* or *Vibrio cholerae* (12, 13). For example, environmental *P. aeruginosa* populations have been observed to carry 10 % of mutator phenotypes related to point mutations in the *mutS* gene (14). As hypothesized by the authors, these early mutator populations could represent future antibiotic resistant (and pathogenic) strains reinforcing the idea that hypermutable *P. aeruginosa* lines are not only associated with clinic strains.

In a recent study we demonstrated that the plant pathogen *Pseudomonas viridiflava* produces phenotypic variants displaying a different behavior in proteolytic activity and exopolysaccharide production (15). Here, we wondered if these variants were affected by phenotypes similar to what has been observed for human pathogens and plant associated bacteria as illustrated above. We also investigated the putative fitness costs of the variants both *in vitro* and *in vivo*. We demonstrated that naturally-produced transparent variants are hypermutable both *in vitro* and *in planta* and they also present antibiotic resistance. In addition, we found a correlation between hypermutability and the absence of pathogenicity in *P. viridiflava*. Interestingly, in contrast to what has been reported for several human pathogens, in

P. viridiflava the mutator phenotype was not related to mutations in genes that code for components of the MMR system. Our study suggests that the formation of antibiotic resistant cells seems to be a common trait of hypermutability but the molecular mechanisms that underline it are not the same in all bacteria. Finally by understanding the mechanisms that regulate the occurrence of non-pathogenic transparent variants we could attempt to control diseases caused by *P. viridiflava* by stimulating these mechanisms.

Materials and methods

Bacterial strains and isolation of variants

Three *P. viridiflava* strains were used in this study. Strain BS0005 was isolated from leaves of kiwifruit (*Actinidia deliciosa*) showing symptoms of bacterial necrosis in Italy in 2008 (15), strain PVB-H was isolated from leaf spots of basil (*Ocimum basilicum*) in Hungary in 2012 (16) and strain CC1582 was isolated from an epilithic biofilm in a creek in France in 2006 (17). Strains were stored at -20°C in phosphate buffer supplemented with 40 % glycerol. Prior to further analysis, strains were streaked on 10% strength Tryptone Soya Agar (TSA/10) (DifoTM Detroit, MI, the USA) and incubated at 26°C for three days.

The presence of phenotypic variation was previously described in all three strains (15). The phase variants were characterized by two different morphologies, a mucoid (M) and a transparent (NM) colony type based on their ability to produce exopolysaccharide. The mucoid variant degrades gelatin indicating a proteolytic activity while the transparent one does not (15). Given the difference in proteolytic activity, TSA/10 medium supplemented with skimmed milk (10 %) (TSAM) was used in all experiment for rapid discrimination of the variants. M variants produced a transparent halo around the colony on the TSA/10 skimmed milk medium while the NM variant did not produce any halo and had a flat and transparent morphology.

To obtain phase variants, the wild type strains were grown in tubes containing 10% Tryptic Soya Broth (TSB/10) for 24 hours at 26°C. Bacterial cultures were then serially diluted and plated on TSAM and plates were incubated at room temperature. After 48 hours, plates were checked and the phase variants were selected based on their morphology on TSAM. Three colonies of each variant were picked with sterilized toothpicks, transferred into a new TSB/10 tube and incubated for 48 hours. The bacterial cultures were serially diluted, plated and checked

for purity for each variant to confirm a fixed phenotype. Ten of these stable variants for each strain were stored at - 20°C in phosphate buffer supplemented with 40% glycerol. To verify stability of the variants, they were grown in TSB/10 for four successive periods of 1 week. At the end of each week of incubation, the cultures were checked for M and NM variants by dilution plating on TSAM, and then 1 ml of the culture was transferred to 9 ml of fresh TSB/10. These three stable variants of each strain were used for the further experiments described below.

Antibiotic resistance assay and determination of mutation frequencies

Antibiotic resistance of the strains was tested on carbanicillin (100 mg l⁻¹), gentamicin (50 mg l⁻¹), cephalotin (50 mg l⁻¹), amoxicillin (50 mg l⁻¹), streptomycin (50 and 100 mg l⁻¹), tetramycin (25 and 50 mg l⁻¹) kanamycin (50 and 100 mg l⁻¹), ampicillin (50 mg l⁻¹) and rifampicin (50 mg l⁻¹). An aliquot (10 µl at 10⁵ CFU ml⁻¹) of bacterial suspension was plated on KB medium (18) supplemented with an antibiotic. In addition bacterial suspensions were also plated on KB without antibiotics. Five plates were used for each of the variants and the wild-type for each. Plates were incubated at 26°C and scored for bacterial growth at 5 days after inoculation.

Mutation frequencies of wild-types as well as M and NM variants were assayed on 100 mg l⁻¹ rifampicin-plates as previously described (19). These frequencies were determined from cultures of stable variants after 24 h periods of growth in 15 replicate tubes of 10 ml of TSB/10. Cultures from each tube were plated on 5 KB plates supplemented with rifampicin. The same bacterial cultures were also dilution plated on KB without rifampicin to quantify the total bacterial population size. Colonies were counted after 3 days of incubation at 26°C. Mutation frequencies were determined by dividing the density of cells resistant to rifampicin by the total population. The experiment was repeated three times and strains with a mutation frequencies > 10⁻⁸ were considered to be hypermutators according to Pruniet *et al.*, (20). Finally, all the rifampicin mutants were transferred into gelatin tubes to determine their proteolytic activity.

In vitro and in planta fitness assays and disease severity test

To investigate putative differences in fitness, growth curve assays for each strain were performed on wild-types, M and NM variants both *in vitro* and *in planta*. For the *in vitro* test, 10² CFU ml⁻¹ bacterial suspensions were inoculated into 4 tubes of TSB/10 and incubated on a rotary shaker at 150 rpm for 48 hours at 26°C. At 14, 17, 20, 24 and 42 hours post inoculation

(p.i.), 100 µl of bacterial cultures were taken and serially diluted and plated on TSAM. Plates were incubated at room temperature and colonies were counted after 48 hours of incubation. The experiment was repeated three times and data were averaged to determine the total bacterial population at each time point. For wild-type lines, the possible occurrence of the NM variants as well as potential reversion of M and NM treatments were monitored.

Fitness *in planta* was evaluated on cantaloupe seedlings (at cotyledon stage), 1 month-old cantaloupe and bean plants. Six cantaloupe seedlings and 4 bean and 4 cantaloupe plants for each time point and for each treatment (wild-type, M and NM variants) were used. Cantaloupe seedlings were inoculated with bacterial suspensions at 10^5 CFU ml⁻¹ as described previously (21) while the stems of cantaloupe and bean plants were inoculated with bacterial suspensions at 10^5 CFU ml⁻¹ as described by Bartoli *et al.*, (22). Plants and seedlings were incubated in the growth chamber at night/day temperatures of 18/23°C and with a light/dark period of 16/8 hours. Re-isolations were made as described by Bartoli *et al.*, (22) for up to 7 days for seedlings and up to 14 days for plants by serially diluting and plating on TSAM. Plates were incubated at room temperature for 48 h and colonies were counted. Bacterial populations were expressed as CFUs per gram of sample for wild-types and the proportions of M and NM variants were calculated for each time point.

Disease severity was assessed on 1 month-old bean plants to evaluate whether the wild-type, M and NM variants had different levels of aggressiveness. Twenty µl of bacterial suspensions at 10^8 CFU ml⁻¹ for each treatment were inoculated into the stems on three different points on each of 4 plants per mutant or wild-type for each strain. Plants were incubated as described above. Disease severity was scored at 4 and 7 days p.i. based on the intensity of necrosis that developed on the inoculated stems. The following scale was used: 0 (no symptoms), <1 cm (moderate symptoms), >1 cm (severe symptoms). Re-isolations were made from all the plants and the percent of M and NM variants was calculated from plants inoculated with wild-types. No symptoms were detected on plants inoculated with sterilized distilled water.

mutS, mutL, mutT, recA and recX gene sequencing and analysis

The entire *mutL* (1965 bp), *mutT* (948 bp), *recA* (1065 bp), and *recX* (423 bp) genes and a part of *mutS* (750 bp) were sequenced to investigate whether the hypermutable NM variants were associated with mutations involved in a defective mismatch repair system (MMR). Primers and

Chapter 6 - Hypermutableity in *Pseudomonas viridiflava*: programed bet-hedging via antibiotic resistance, exopolysaccharide production and pathogenicity

PCR conditions are described in supplementary methods S1. Bayesian trees were constructed with Mr. Bayes (23).

Statistical analysis

Student *T*-tests were performed to compare wild-type, M and NM variants by using GraphPad online software <http://www.graphpad.com/>. For comparison of mutation frequencies between wild-types and transparent variants we used the Mann-Whitney *U* test. A *P* value < 0.05 was considered as statistically significant.

Results

P. viridiflava phase variants do not revert and the temporal dynamics of their wild-types is strain-dependent

The two colony variants (M and NM) of *P. viridiflava* were characterized by a different behavior in proteolytic activity. The mucoid variant was proteolytic while the transparent one lacked that activity. During growth, the all colonies of the wild-type inocula were mucoid but they were able to produce NM during their growth. In contrast, neither M nor NM reverted during the incubation period. The mucoid colonies of the M variants were different from those of the wild-type since they did not produce NM cells. Nevertheless, during the 48 hours *in vitro* growth, the rate of occurrence of detectable levels of NM variants was different from one strain to another. In particular, in both strains BS0005 and PVB-H the NM variants were observed when the population size was 10^6 CFU ml⁻¹ (Table S1). While for strain CC1582, the transparent variants were observed when the bacterium attained a population size of 10^9 CFU ml⁻¹ (Table S1). Likewise, the proportion between transparent and mucoid variants also varied among strains and globally we observed the highest number of transparent variants for strain BS0005 (Table S1). Taken together, these observation led us to conclude that: i) the wild-type state was able to spontaneously produce two phenotypic variants, a mucoid and a transparent and the phenotype of the wild-type before long-term culturing was mucoid; ii) both of the variants were unable to revert and as such they could be considered as stable mutants whereby the original phenotype cannot be restored under normal culturing conditions and; iii) the mucoid variants are different from the wild-types because of their inability to produce transparent variants (or mutants).

Transparent variants are hypermutable and resistant to antibiotics

In several human bacterial pathogens, phase variants were associated with a high mutation ($> 10^{-8}$ mutants per generation) level and the ability to resist several antibiotics. We wondered whether in *P. viridiflava* the mucoid or the transparent variants were associated with the phenotypes previously observed for human bacterial pathogens, viz. with antibiotic resistance. The investigation of mutation frequencies between wild-type, M and NM variants revealed three interesting results. Firstly, M variants did not generate any rifampicin-resistant mutants.

Secondly, the transparent variants showed higher mutation levels compared to the wild-types. The NM variant of strain BS0005 was characterized by a mutation frequency of 1.10×10^{-7} , significantly higher than its wild-type counterpart (4.92×10^{-8}) ($P < 0.05$). In strain PVB-H however, the mutation frequency was 1.06×10^{-7} for NM and 2.50×10^{-8} for the wild-type and these values were not statistically significant ($P > 0.05$). Concerning strain CC1582, the mutation frequency for NM and the wild-type were 1.20×10^{-7} and 5.80×10^{-8} , respectively, and the difference was not statistically significant ($P > 0.05$). Finally, when the mutants obtained from the wild-types were transferred to tubes containing gelatin, mutants were unable to liquefy it confirming that the rifampicin -mutants were characterized by a transparent phenotype. Indeed, when transferred to TSAM medium, the rifampicin -mutants had the same phenotype as that of the NM variants. These results corroborated the observation that for M variants, rifampicin -mutants never occurred and for this reason the mutation frequency in mucoid variants was not detectable.

To investigate whether the hypermutable transparent phenotype of *P. viridiflava* was also associated with other antibiotic resistance, we tested susceptibility to several antibiotics of both M and NM variants for strains BS0005, PVB-H and CC1582. In all the cases, we found that all cells of transparent variants showed resistance to amoxicillin, ampicillin, kanamycin and rifamycin. In contrast, the mucoid variants did not grow in the presence of these antibiotics. Finally, all the strains were inherently resistant to carbenicillin and cephalotin and susceptible to gentamicin, streptomycin and tetracycline and no differences were observed between the N and NM variants.

Mucoid and transparent variants are less fit than the wild-type counterparts in vitro but not in planta

To determine if differences in fitness could exist between the two variants and the wild-type, we assessed growth kinetics both *in vitro* and in bean and cantaloupe plants. Statistically significant differences were observed *in vitro* between the variants and wild-type counterparts. Overall, both the mucoid and transparent variants of each strain tested were significantly less fit when growing *in vitro* compared to their wild-type counterparts. In particular, the fitness difference between wild-types and their variants was highly significant for all the strains at 14 h after

inoculation ($P < 0.005$), 17, 20, 24 and 42 h p.i. ($P < 0.0005$) (Fig. 1A, 1B, 1C). In contrast, no significant differences in fitness were observed between M and NM variants (Fig. 1A, B, C).

In contrast to what we observed *in vitro*, no statistically significant differences ($P > 0.05$) were observed on bean plants in terms of population size between wild-types and their variants for all of the strains tested. The M and NM variants were able to attain the same population density as that of the wild-type counterparts at 3, 10 and 14 days after inoculation (Fig. 2A, B, C) demonstrating that there was no fitness cost of the mucoid and transparent variants. The only exception was for strain BS0005 in which the growth of the NM variant was higher in bean plants compared to the M variant and wild-type counterpart at 7 and 14 days p.i. ($P < 0.005$) (Fig. 1A). When the ratio between mucoid and transparent colonies was monitored in the wild-types of the three strains tested we observed that the BS0005 and PVB-H wild-types produced transparent variants at 3 days after inoculation when their population density was 10^6 CFU g⁻¹ (Table S1). By contrast, the CC1582 wild-type strain did not produce any transparent variant during host infection (Table S1). Concerning strain BS0005, the maximum number of NM variants was observed at 7 days p.i. represented by 15.0 % of the transparent variants in the total bacterial population. Likewise, for strain PVB-H the peak of NM variants was detected at 3 days p.i. and the transparent variants represented 35.5 % of the total bacterial population. An analysis of growth kinetics was also performed on cantaloupe seedlings and plants since we previously demonstrated a correlation between phase variation and aggressiveness on cantaloupe seedlings (15). None of the variants and wild-types of the strains tested grew in cantaloupe seedlings and plants; their population size was constant during the experiment (data not shown). In both cantaloupe seedlings and plants the wild-types did not produce NM variants. In addition, the wild-type strains induced disease only on the seedlings but not on 1 month-old plants (data not shown). These results suggest that the aggressiveness reaction on the cantaloupe seedlings was only a hypersensitive reaction not associated with bacterial multiplication in the host cells.

Mucoid variants and wild-types are pathogenic to bean plants while transparent variants do not induce disease

Although we observed an increase in bacterial population size in bean plants when we inoculated plants with suspensions at 10^5 CFU ml⁻¹, we did not observe any disease symptoms. We hypothesized that the low initial inoculum concentration used in the experiment did not

favor induction of symptoms. To test this hypothesis, we performed another inoculation on bean plants with all the three strains. The wild-types and their variants were separately inoculated at 10^8 CFU ml⁻¹. In this case, all the wild-types and mucoid variants induced disease on bean plants while the transparent variants were unable to cause disease (Fig. 3 A, B, C). In particular, both the wild-type and mucoid variants caused severe (stem necrosis > 1 cm) and moderate (stem necrosis < 1 cm) disease symptoms for strains BS0005 and PVB-H, respectively. However, for strain CC1582, disease severity was high (necrosis > 1) for the wild-type and moderate for the mucoid variant.

For all strains, population size of the NM variants was lower compared to that of the wild-type and mucoid counterparts ($P < 0.0005$) (Fig. 4). Furthermore, strain CC1582 was not able to produce transparent variants during the colonization of bean plants while for strains BS0005 and PVB-H, 9% and 33 %, respectively, of the colonies were NM variants.

Mutations in genes for DNA repair are not involved in the production of hypermutable antibiotic resistant transparent variants

To investigate whether the formation of variants in *P. viridiflava* might be due to a defective MMR or genes related to this system, we sequenced the *mutS*, *mutL*, *mutT*, *recA*, and *recX* genes of mucoid, transparent, wild-types and rifampicin-mutants. We found no differences in gene sequences between the wild-types and variants or mutants. In *E. coli* it has been shown that genes in the MMR are under horizontal gene transfer (HGT) and that they have been lost and acquired different time during the evolution of this bacterial species permitting the genome plasticity of the bacterium (24). We then investigated the putative HGT in the MMR genes of *P. viridiflava*. We built the phylogeny of the MMR genes of the strains BS0005, PVB-H, CC1582, TA0043 and UASW0038 and those of other *Pseudomonas syringae* phylogroups (Fig. S1). Bayesian phylogeny demonstrated that the *P. viridiflava* strains, in the phylogroup 7 of the *P. syringae* complex, did not acquire the MMR genes *via* HGT from other phylogroups (Fig. S1). Indeed, HGT did not occur in any of the strains of the *P. syringae* phylogroups used in our phylogeny. Taken together, these results suggest that the formation of the variants in *P. viridiflava* is not regulated by a mutation or HGT in MMR genes.

Discussion

Phase variation is the genetically-programmed switching from one phenotype to another for a given trait. The occurrence of variants in phase variation is often related to processes that mediate gene expression in response to particular environmental conditions (25). Unlike *bona fide* mutants, phase variants are able to restore their original phenotype ("reversion") when conditions are appropriate. In addition, phase variants occur at frequencies higher than 10^{-5} switches per cell per generation unlike spontaneous mutations which were reported to occur with an average frequency of 10^{-7} mutations per cell per generation (26). Here, we showed that the phase variants spontaneously produced by *P. viridiflava* are unable to revert after one to several days of growth in synthetic media or after being inoculated into a host plant. The rate of occurrence of the transparent and mucoid variants was different among the strains tested with the highest value observed of 10^{-6} variants per generations. These results led us to suggest that the mucoid and transparent variants are probably programmed mutants rather than *bona fide* phase variants. Programmed mutation in the absence of particular stimuli or environmental stresses is a form of bet-hedging that allows generation of variation by increasing the probability that individuals express one or more phenotypes which will be adapted to future environmental conditions (25, 27). In *Pseudomonas fluorescens*, one mutation has been demonstrated as the responsible of bet-hedging (27). In particular, the authors demonstrated that strain SBW25 of *P. fluorescens* produced capsulated and non-capsulated variants and that the non-capsulated phenotype is due to a single non-synonymous nucleotide change in the *carB* gene encoding for the large subunit of carbamoylphosphate synthase. Given the stable behavior of the M and NM variants and their frequency we speculate that in *P. viridiflava* there is a constitutive mechanism similar to bet-hedging that stimulates the production of mutants even without any apparent specific stimuli.

Another possible scenario could be that the production of these mutants is regulated by stressful conditions such as nutritional limitations. For example, *E. coli* mutants are produced under amino acid limitation during stationary phase (28). In biofilms of *Streptococcus pneumoniae*, the development of non-phase variation colony variants was related to high hydrogen peroxide conditions (26). However, our results from *in vitro* experiments support bet-hedging rather than stress induction because no variants were detected in the stationary phase. Another hypothesis

about the basal mechanisms that regulate phase variation suggests a possible role for cell-cell communication. For example, a high mutation rate and phase variation in *E. coli* were related to population density and was regulated by the quorum-sensing regulating gene *luxS* (29). In particular, the authors suggested that a low population density can increase the mutation frequency. Here, since the hypermutable transparent variants could not be detected until bacterial population density was 10^6 CFU ml⁻¹, a similar phenomenon might also be involved in *P. viridiflava* phase variation. Whole genome sequencing of both variants could elucidate on the kind of mutations responsible for the different phenotypes.

In several human pathogens, spontaneous mutants arising during *in vitro* culture or during host infection had a hypermutable phenotype associated with resistance to antibiotics (4–6). Here, we investigated whether the M and NM variants were also characterized by a hypermutable phenotype and resistant to antibiotics. As previously observed for human bacterial pathogens, we found that only the transparent variants were hypermutable and resistant to several antibiotics. In human clinical strains, antibiotic resistance cells were responsible of chronic diseases and it was assumed that these mutants were selected by use of antibiotics (2). However, for a plant pathogen we can wonder about the advantage of resistance to antibiotics. Antibiotic resistance is known to provide fitness advantages in soil (30). Although *P. viridiflava* is a ubiquitous bacterium isolated from different environments including those related with the water cycle (15), it has not been isolated from soil. However, its ability to degrade pectin and the presence of genes for putrescine found in its genome (15) suggest that it might be able to live in the soil. A recent study showed how *P. syringae* is transported through the soil to the water table during snow melt and rain run-off (31). This trajectory could transport *P. viridiflava* to localizations in soil strata that have not yet been explored where it could survive. Water habitats are also rich in antibiotics due to the use of antimicrobial compounds in agriculture, to run-off from medical environments and to the presence of antibiotic-producing organisms (32). Antibiotic resistance can occur in surprising habitats such as in the sources of mineral and spring waters (32). This suggests that the population dynamics of antibiotic resistant strains is not only related to the presence of antibiotics in contaminated waters but also to other mechanisms such as the presence of microbial competitors (antagonists) which could serve as a positive selective force for resistance. In this light, selection of mutations that leads to antibiotic resistance could occur outside of situations where antibiotics are deployed by humans. Given the wide presence

of *P. viridiflava* in fresh-waters (17), these habitats could be reservoirs where transparent hypermutable antibiotic resistant variants can proliferate. In addition, intrinsic antibiotic resistance (comprising the ensemble of non-acquired genes that influence the susceptibility to antibiotics) has also been demonstrated in different bacterial communities (33, 34). This leads us to assume that *P. viridiflava* presents an intrinsic resistance to antibiotic. Furthermore, phase variation can also lead to resistance to toxic metals such as cadmium as in the case of *Pseudomonas brassicacearum* (9).

Our results demonstrate that both mucoid and transparent variants are less fit *in vitro*, but not *in vivo*, compared to their wild-types counterparts suggesting that they are less adapted to a constant environment and are favored in stressful conditions. Hypermutable *P. aeruginosa* cells isolated from cystic fibrosis patients showed a fitness cost both *in vitro* and in the host (35). In contrast, hypermutable carbapenem *P. aeruginosa* mutants had a fitness advantage *in vivo* raising questions about the adaptive advantage of antibiotic resistant bacteria (36). This latter observation is in agreement with our results in that *P. viridiflava* variants showed no loss in fitness in bean plants. Reactive oxygen species (ROS) are produced in plants after pathogen infection as a defense response (37). It is possible that *P. viridiflava* variants survive better than their wild-type counterparts when ROS are produced by the host. However, preliminary results on oxidative stress resistance did not show any apparent differences between the wild-types and variants (data not shown). Long-term experiments in host plants are necessary to better elucidate the population dynamics of *P. viridiflava* variants and their effective role in assuring infection of the host. Transparent variants were defective in causing disease in bean plants compared to the wild-type and mucoid counterparts. Since these NM variants are unable to produce extracellular polysaccharides (EPS) (15) the lack of the disease could be related to the absence of EPS. Indeed, EPS are known to be an important virulent factor in several phytopathogenic bacteria such as *Ralstonia solanacearum*, *Erwinia amylovora* and *Xanthomonas campestris* (38, 39). Understanding the mechanisms that underline the occurrence of these transparent non-pathogenic variants could be an important point in the development of new disease control strategies based on the enhancement of occurrence of the non-pathogenic variants.

Hypermutableity associated with antibiotic resistance, as described in several bacterial pathogens of humans, is linked to a defective MMR system due to mutation in genes for proteins that are involved in the MMR process (4–6, 40). Here we showed that neither the *mut* genes nor the *rec*

genes were mutated in the transparent hypermutable variants. From this we can speculate that modulation of gene expression could be involved in a defective MMR in *P. viridiflava* strains. For example in *Bacillus subtilis* the over-expression of *mutS* gene decreases the mutation rate of this bacterium (41). In the same way the down-regulation of *mutS*, such as its inactivity, is related to hypermutability in *B. subtilis* (41). However, other mechanisms could underline the occurrence of this phenomenon. For example in human cells small RNAs regulate the expression of *mutL* affecting the occurrence of tumor cells (42). Since the MMR system is highly conserved in all organisms (7), from prokaryotes to eukaryotes, the small RNAs may regulate the expression of MMR genes in bacteria and then be related with the hypermutable phenotype. These mechanisms could be revealed by sequencing total mRNA or the whole genomes of M and NM variants.

Chapter 6 - Hypermotability in *Pseudomonas viridiflava*: programmed bet-hedging via antibiotic resistance, exopolysaccharide production and pathogenicity

Acknowledgements

We thank the technicians of the Experimental Installation team at INRA of Avignon, for providing us the plant material used in the pathogenicity tests. The PhD fellowship for Claudia Bartoli was provided by Tuscia Univeristy.

References

1. **Spratt BG, Maiden MC.** 1999. Bacterial population genetics, evolution and epidemiology. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **354**:701–710.
2. **Lenski RE.** 1998. Bacterial evolution and the cost of antibiotic resistance. *Int. Microbiol.* **1**:265–270.
3. **Miller K.** 2002. Response of *Escherichia coli* hypermutators to selection pressure with antimicrobial agents from different classes. *J. Antimicrob. Chemother.* **49**:925–934.
4. **Marvig RL, Johansen HK, Molin S, Jelsbak L.** 2013. Genome analysis of a transmissible lineage of *Pseudomonas aeruginosa* reveals pathoadaptive mutations and distinct evolutionary paths of hypermutators. *PLoS Genet.* **9**:e1003741.
5. **Watson ME, Burns JL, Smith AL.** 2004. Hypermutable *Haemophilus influenzae* with mutations in *mutS* are found in cystic fibrosis sputum. *Microbiology* **150**:2947–2958.
6. **Wang S, Wu C, Shen J, Wu Y, Wang Y.** 2013. Hypermutable *Staphylococcus aureus* strains present at high frequency in subclinical bovine mastitis isolates are associated with the development of antibiotic resistance. *Vet. Microbiol.* **165**:410–415.
7. **Schofield MJ, Hsieh P.** 2003. DNA mismatch repair: molecular mechanisms and biological function. *Annu. Rev. Microbiol.* **57**:579–608.
8. **Torres-barceló C, Cabot G, Oliver A, Buckling A, Maclean RC, Torres-barcelo C.** 2013. A trade-off between oxidative stress resistance and DNA repair plays a role in the evolution of elevated mutation rates in bacteria. *Proc. R. Soc. B* **280**:20130007.
9. **Pagès D, Sanchez L, Conrod S, Gidrol X, Fekete A, Schmitt-Kopplin P, Heulin T, Achouak W.** 2007. Exploration of intracolonial adaptation mechanisms of *Pseudomonas brassicacearum* facing cadmium toxicity. *Env. microbiol* **9**:2820–2835.
10. **Vial L, Lavire C, Mavingui P, Blaha D, Haurat J, Moënne-Loccoz Y, Bally R, Wisniewski-Dyé F.** 2006. Phase variation and genomic architecture changes in *Azospirillum*. *J. Bacteriol.* **188**:5364–5373.
11. **Matic I, Radman M, Taddei F, Picard B, Doit C, Bingen E, Denamur E, Elion J.** 1997. Highly variable mutation rates in commensal and pathogenic *Escherichia coli*. *Science.* **277**:1833–1834.
12. **Faruque SM, Chowdhury N, Kamruzzaman M, Dziejman M, Rahman MH, Sack DA, Nair GB, Mekalanos JJ.** 2004. Genetic diversity and virulence potential of environmental *Vibrio cholerae* population in a cholera-endemic area. *Proc. Natl. Acad. Sci. U. S. A.* **101**:2123–2128.

13. **Rau MH, Hansen SK, Johansen HK, Thomsen LE, Workman CT, Nielsen KF, Jelsbak L, Høiby N, Yang L, Molin S.** 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.
14. **Kenna DT, Doherty CJ, Foweraker J, Macaskill L, Barcus VA, Govan JRW.** 2007. Hypermutable in environmental *Pseudomonas aeruginosa* and in populations causing pulmonary infection in individuals with cystic fibrosis. *Microbiology* **153**:1852–1859.
15. **Bartoli C, Berge O, Monteil C, Guilbaud C, Balestra GM, Varvaro L, Jones C, Dangl JL, Baltrus DA, Sands DC, Morris CE.** 2014. The *Pseudomonas viridiflava* phylogroups in the *P. syringae* species complex are characterized by genetic variability and phenotypic plasticity of pathogenicity-related traits. *Env. microbiol* DOI: 10.1111/1462-2920.12433.
16. **Végh A, Hevesi M, Némethy Z, Palkovics L.** 2012. First report of bacterial leaf spot of basil caused by *Pseudomonas viridiflava* in Hungary. *Plant Dis* **96**:141.
17. **Morris CE, Sands DC, Vanneste JL, Montarry J, Oakley B, Guilbaud C, Glaux C.** 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.
18. **King EO, Ward MK, Raney DE.** 1954. Two simple media for the demonstration of pyocyanin and fluorescin. *J. Lab. Clin. Med.* **44**:301–307.
19. **Besier S, Zander J, Kahl BC, Kraiczky P, Brade V, Wichelhaus TA.** 2008. The thymidine-dependent small-colony-variant phenotype is associated with hypermutability and antibiotic resistance in clinical *Staphylococcus aureus* isolates. *Antimicrob. Agents Chemother.* **52**:2183–2189.
20. **Prunier A-L, Malbruny B, Laurans M, Brouard J, Duhamel J-F, Leclercq R.** 2003. High rate of macrolide resistance in *Staphylococcus aureus* strains from patients with cystic fibrosis reveals high proportions of hypermutable strains. *J. Infect. Dis.* **187**:1709–1716.
21. **Morris CE, Sands DC, Vinatzer BA, Glaux C, Guilbaud C, Buffière A, Yan S, Dominguez H, Thompson BM.** 2008. The life history of the plant pathogen *Pseudomonas syringae* is linked to the water cycle. *ISME J* **2**:321–334.
22. **Bartoli C, Lamichhane JR, Berge O, Guilbaud C, Varvaro L, Balestra GM, Vinatzer BA, Morris CE.** 2014. A framework to gage the epidemic potential of plant pathogens in environmental reservoirs: the example of kiwifruit canker. **Submitted article.**

23. **Ronquist F, Huelsenbeck JP.** 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**:1572–1574.
24. **Denamur E, Lecoindre G, Darlu P, Tenaillon O, Acquaviva C, Sayada C, Sunjevaric I, Rothstein R, Elion J, Taddei F, Radman M, Matic I.** 2000. Evolutionary implications of the frequent horizontal transfer of mismatch repair genes. *Cell* **103**:711–721.
25. **Meyers LA, Bull JJ.** 2002. Fighting change with change: adaptive variation in an uncertain world. *Trends Ecol. Evol.* **17**: 551-557.
26. **Allegrucci M, Sauer K.** 2008. Formation of *Streptococcus pneumoniae* non-phase-variable colony variants is due to increased mutation frequency present under biofilm growth conditions. *J Bacteriol* **190**:6330–6339.
27. **Beaumont HJE, Gallie J, Kost C, Ferguson GC, Rainey PB.** 2009. Experimental evolution of bet hedging. *Nature* **462**:90–93.
28. **Poole K.** 2012. Stress responses as determinants of antimicrobial resistance in Gram-negative bacteria. *Cell* **20**:227–234.
29. **Krasovec R, Belavkin R V., Aston JAD, Channon A, Aston E, Rash BM, Kadirvel M, Forbes M, Knight CG.** 2014. Interactions, mutation rate plasticity in rifampicin resistance depends on *Escherichia coli* cell–cell. *Nature* DOI: **10.10**.
30. **Chait R, Vetsigian K, Kishony R.** 2011. What counters antibiotic resistance in nature? *Nat. Chem. Biol.* **8**: 2-5
31. **Monteil CL, Lafolie F, Laurent J, Clement J-C, Simler R, Travi Y, Morris CE.** 2013. Soil water flow is a source of the plant pathogen *Pseudomonas syringae* in subalpine headwaters. *Environ. Microbiol.* DOI: 10.1111/1462-2920.12296
32. **Vaz-Moreira I, Nunes OC, Manaia CM.** 2014. Bacterial diversity and antibiotic resistance in water habitats: searching the links with the human microbiome. *FEMS Microbiol. Rev.* DOI:10.1111/1574-6976.12062
33. **Baquero F, Tedim AP, Coque TM.** 2013. Antibiotic resistance shaping multi-level population biology of bacteria. *Front. Microbiol.* **4**:15.
34. **Fajardo A, Martínez-Martín N, Mercadillo M, Galán JC, Ghysels B, Matthijs S, Cornelis P, Wiehlmann L, Tümmler B, Baquero F, Martínez JL.** 2008. The neglected intrinsic resistome of bacterial pathogens. *PLoS One* **3**:e1619.
35. **Montanari S, Oliver A, Salerno P, Mena A, Bertoni G, Tümmler B, Cariani L, Conese M, Döring G, Bragonzi A.** 2007. Biological cost of hypermutation in

Pseudomonas aeruginosa strains from patients with cystic fibrosis. Microbiology **153**:1445–1454.

36. **Skurnik D, Roux D, Cattoir V, Danilchanka O, Lu X, Yoder-Himes DR, Han K, Guillard T, Jiang D, Gaultier C, Guerin F, Aschard H, Leclercq R, Mekalanos JJ, Lory S, Pier GB.** 2013. Enhanced in vivo fitness of carbapenem-resistant *oprD* mutants of *Pseudomonas aeruginosa* revealed through high-throughput sequencing. Proc. Natl. Acad. Sci. U. S. A. **110**:20747–20752.
37. **Kotchoni SO, Gachomo EW.** 2006. The reactive oxygen species network pathways: an essential prerequisite for perception of pathogen attack and the acquired disease resistance in plants. J. Biosci. **31**:389–404.
38. **Milling A, Babujee L, Allen C.** 2011. *Ralstonia solanacearum* extracellular polysaccharide is a specific elicitor of defense responses in wilt-resistant tomato plants. PLoS One DOI: 10.1371
39. **Oh CS, Beer S V.** 2005. Molecular genetics of *Erwinia amylovora* involved in the development of fire blight. FEMS Microbiol. **253**: 185-192
40. **Oliver A, Cantòn R, Campo P, Baquero F, Blázquez J.** 2000. High frequency of hypermutable *Pseudomonas aeruginosa* in cystic fibrosis lung infection. Science. **288**:1251–1253.
41. **Pedraza-Reyes M, Yasbin RE.** 2004. Contribution of the mismatch DNA repair system to the generation of stationary-phase-induced mutants of *Bacillus subtilis*. J. Bacteriol. **186**:6485–6491.
42. **Mao G, Lee S, Ortega J, Gu L, Li G-M.** 2012. Modulation of microRNA processing by mismatch repair protein MutL α . Cell Res. **22**: 973-985

Figure Legends

Figure 1. *In vitro* growth of wild-types and mucoid and transparent variants for strains BS0005 (A), PVB-H (B) and CC1582 (C). Wild-types are indicated with solid lines and phase variants with dotted lines (mucoid: ●, transparent: ○). Standard error is indicated for each mean.

Figure 2. Growth of wild-types and mucoid and transparent variants in 1-month old bean plants for strains BS0005 (A), PVB-H (B) and CC1582 (C). Wild-types are indicated with solid lines and phase variants with dotted lines (mucoid: ●, transparent: ○). Standard error is indicated for each mean.

Figure 3. Disease symptoms on bean stems caused by wild-type (A) and mucoid variants (B) of strain BS0005. The transparent BS0005 variant did not cause any symptoms (C). The behavior of wild-type and mucoid and transparent variants was the same for all the strains tested.

Figure 4. Population density of wild type (solid bars) and of mucoid (hashed bars) and transparent (open bars) variants during the disease assay on bean plants for strains BS0005, PVB-H and CC1582. Bacterial population was determined at 7 days after inoculation. Error bars are indicated for each mean. Statistically significant differences of means between wild-type and mucoid or transparent variants (NM) were expressed as * ($P < 0.05$), ** ($P < 0.005$) or *** ($P < 0.0005$).

Supplementary Information

Table S1. Occurrence of transparent variants during growth curve *in vitro* and in bean plants.

Figure S1. Bayesian phylogeny constructed on the concatenated *cts*, *gyrB*, *gapA* and *rpoD* housekeeping genes and *mutS*, *mut*, *mutT* *recA* and *recX* genes. Names of phylogroups are indicated according to Berge *et al.*, 2014

Methods S1. Protocol for *mut* and *rec* gene amplification and sequencing

Chapter 6 - Hypermutableity in *Pseudomonas viridiflava*: programed bet-hedging via antibiotic resistance, exopolysaccharide production and pathogenicity

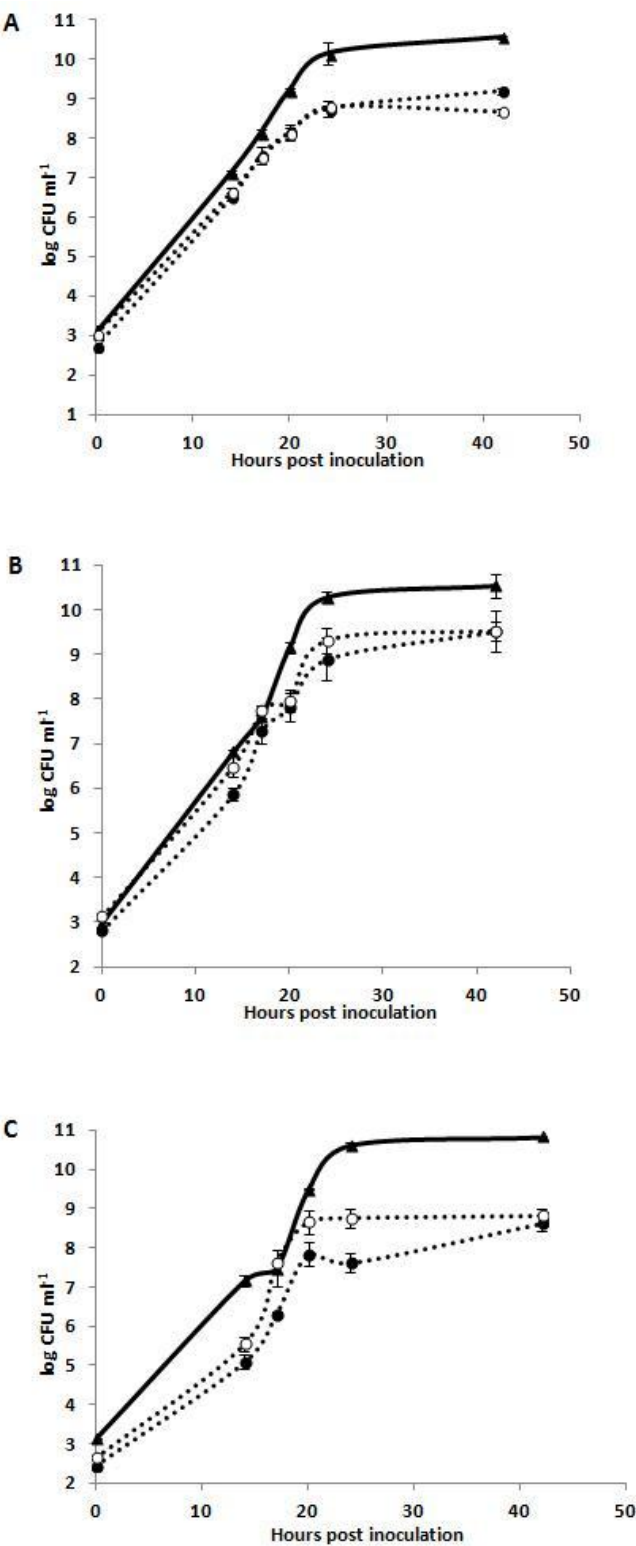


Figure 1

Chapter 6 - Hypermutableity in *Pseudomonas viridiflava*: programed bet-hedging via antibiotic resistance, exopolysaccharide production and pathogenicity

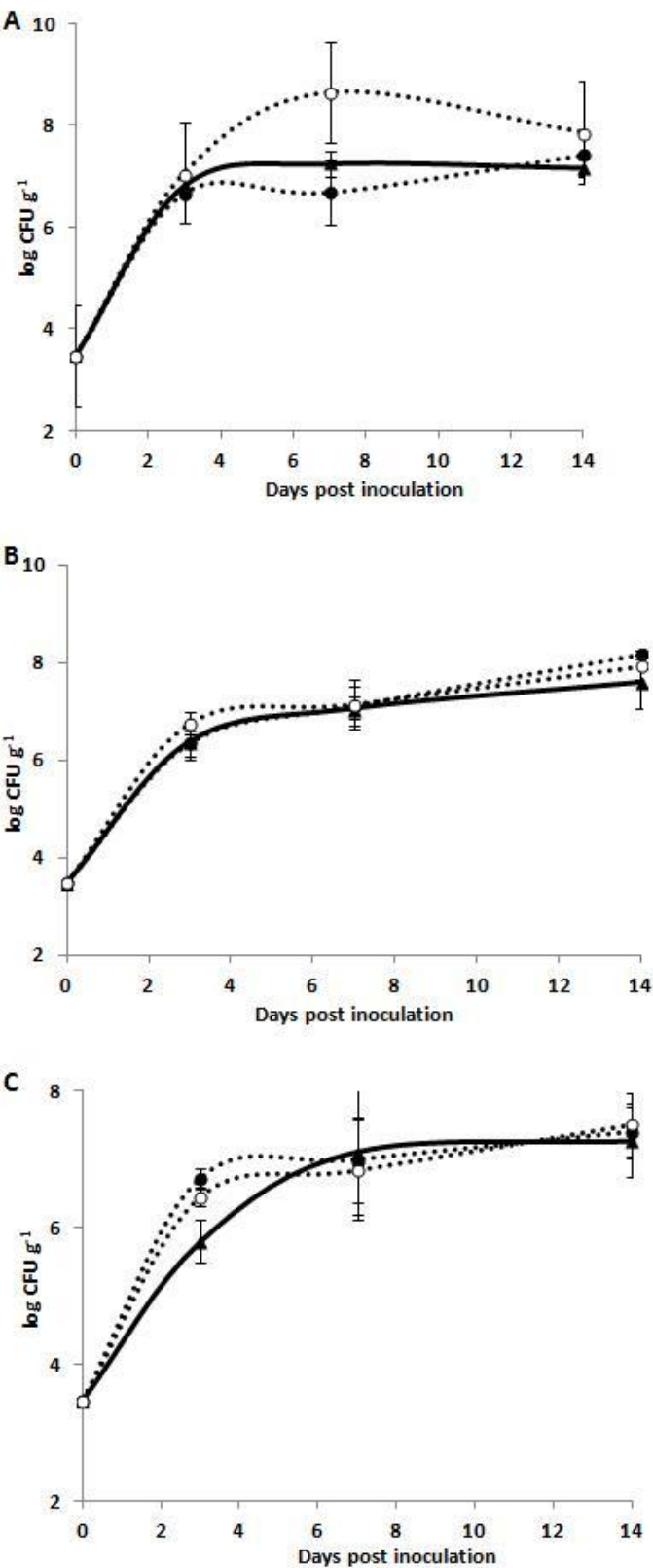


Figure 2

Chapter 6 - Hypermotability in *Pseudomonas viridiflava*: programed bet-hedging via antibiotic resistance, exopolysaccharide production and pathogenicity



Figure 3

Chapter 6 - Hypermutability in *Pseudomonas viridiflava*: programed bet-hedging via antibiotic resistance, exopolysaccharide production and pathogenicity

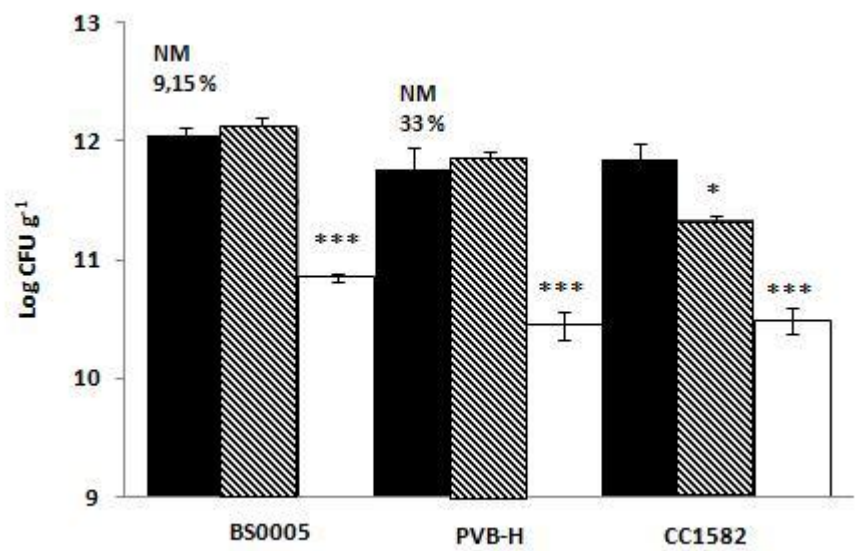


Figure 4

Supplementary Data

Supplementary Method S1.

mutL, *mutT*, *mutS*, *recA* and *recX* gene sequencing

For all PCR reactions GoTaq® Flexi polymerase was used according to manufactures instruction. For *mutL* two primers sets were used for amplifying the entire gene; for the first gene part we used MutL-fw1 (ATGAAGAAGCCGCCAGCAG) and MutL-rv1 (TGGGTGACGGCCATTGAACA), while for the second part MutL-fw2 (CAGTGCGCGACAAACTTGTC) and MutL-rv2 (TCATTGACCGCGCAGGAAT) were used. The *mutT* gene was amplified with MutT-fw (GTGAAACGCATTCACGTGGCC) and MutT-rv (TCAGAGCTCTTCAGGCCAGAAC). The partial *mutS* gene was amplified with the degenerated primers MutSDeg-fw1 (TGATGTTYTAYCGCATGGGCGA) and MutSDeg-rv1 (RTTCAGCCAGCGDGTTCAGCA). Degenerated primers were use since primers designed on the CC1582 strains did not permitted the amplification of the *mutS* gene on BS0005 and PVBH strain. For the *recA* gene the RecA-fw (ATGGACGACAACAAGAAGAAAG) and the degenerated primer RecA-rw (GAACTTGGCCGAGTTGGCYTT) were used. Finally for the *recX* gene we used the primer set RecX-fw (TCGAGCTGACGCGAAAAGT) and ReX-rv (AACAGGCGGCCGATCATGT). For all genes the following PCR program was used: 5 min at 95°C for the initial denaturation, followed by 30 cyclers with 95°C for 30 sec, 60°C for 30 sec and 75°C for 30 sec. Final extension consisted in 10 min at 75°C.

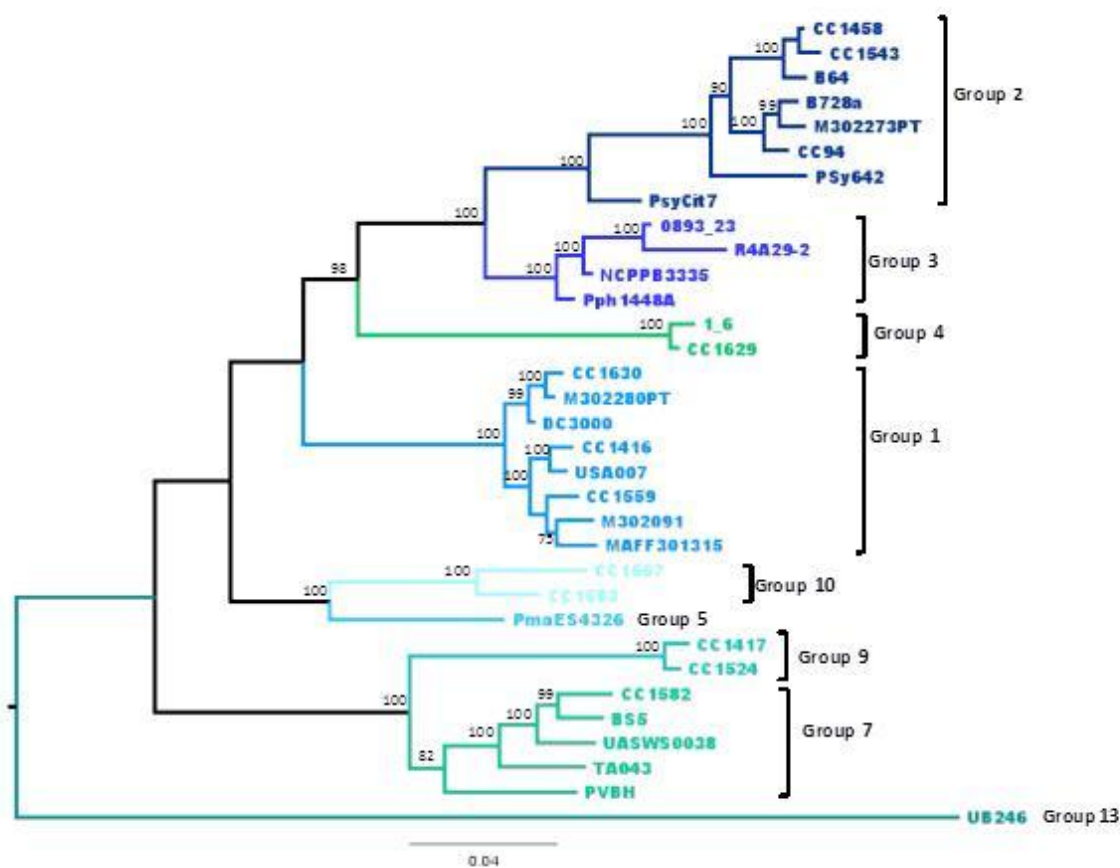
Chapter 6 - Hypermutability in *Pseudomonas viridiflava*: programed bet-hedging via antibiotic resistance, exopolysaccharide production and pathogenicity

Supplementary Table S1

Percent of transparent variants during *in vitro* and *in vivo* growth

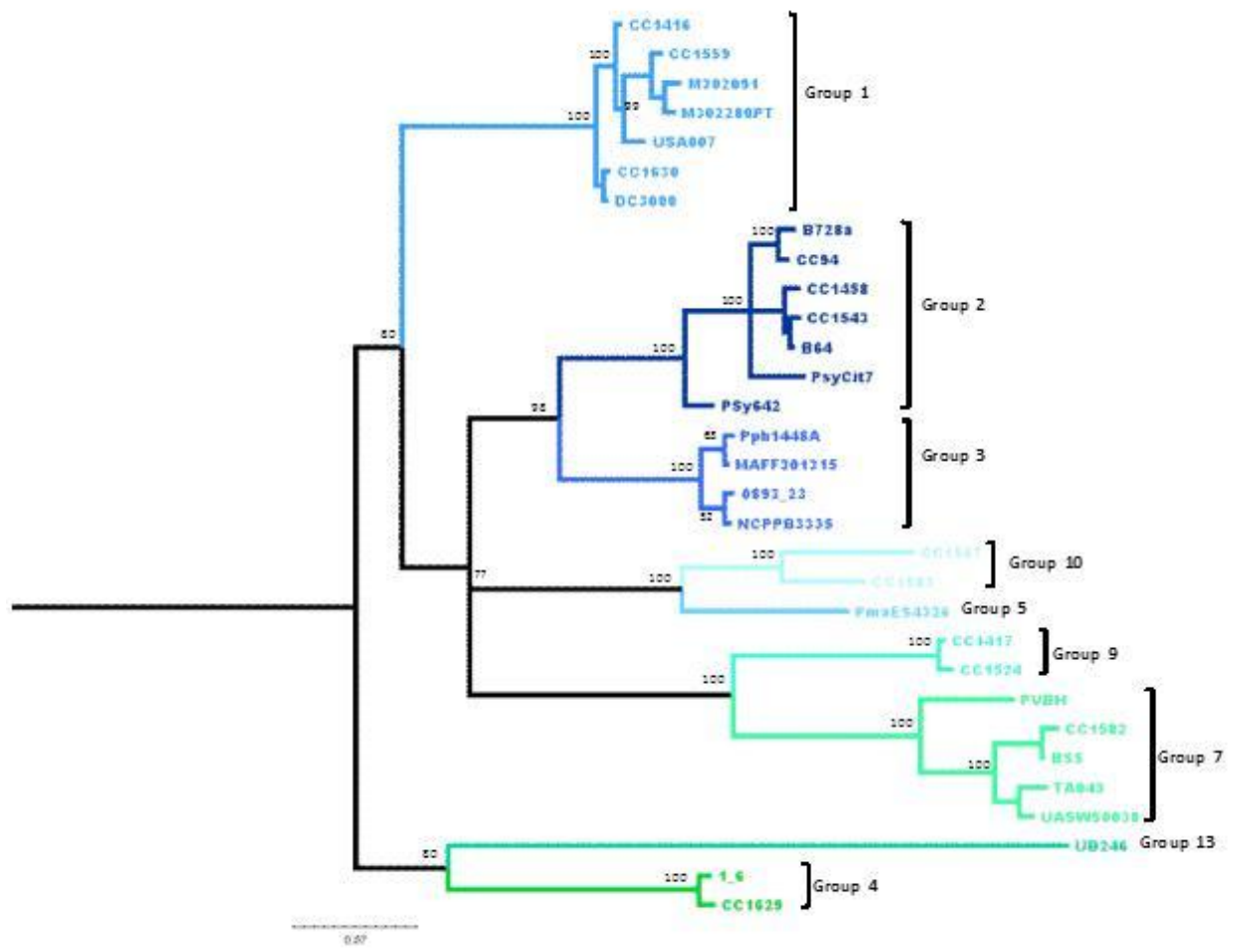
Strain	Population Density	Days Post Inoculation	% Transparent variants
<i>In vitro</i> growth			
BS0005	10^2	0	0
	10^6	14	9
	10^8	17	15.05
	10^9	20	77
	10^{10}	24	24
	10^{10}	42	25
PVB-H	10^2	0	0
	10^6	14	9
	10^8	17	9
	10^9	20	5.05
	10^{10}	24	7.05
	10^{10}	42	8
CC1582	10^2	0	0
	10^6	14	0
	10^8	17	0
	10^9	20	0
	10^{10}	24	1,5
	10^{10}	42	2
<i>In vivo</i> growth (bean plants)			
BS0005	10^3	0	0
	10^6	3	4
	10^7	7	15
	10^7	14	5
PVB-H	10^3	0	0
	10^6	3	38,5
	10^7	7	1
	10^7	14	25,8
CC1582	10^3	0	0
	10^5	3	0
	10^7	7	0
	10^7	14	0

Supplementary Fig. S1



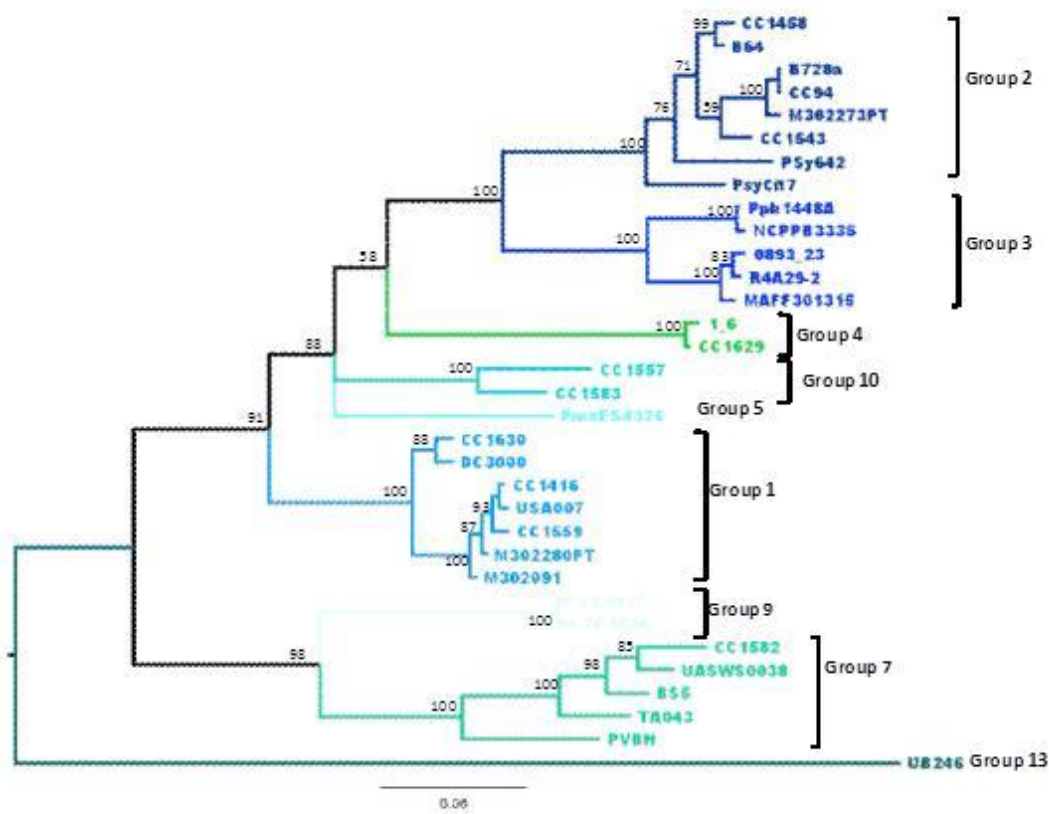
Phylogeny based on *cts*, *gyrB*, *gapA* and *rpoD* housekeeping genes

Chapter 6 - Hypermutableity in *Pseudomonas viridiflava*: programed bet-hedging via antibiotic resistance, exopolysaccharide production and pathogenicity



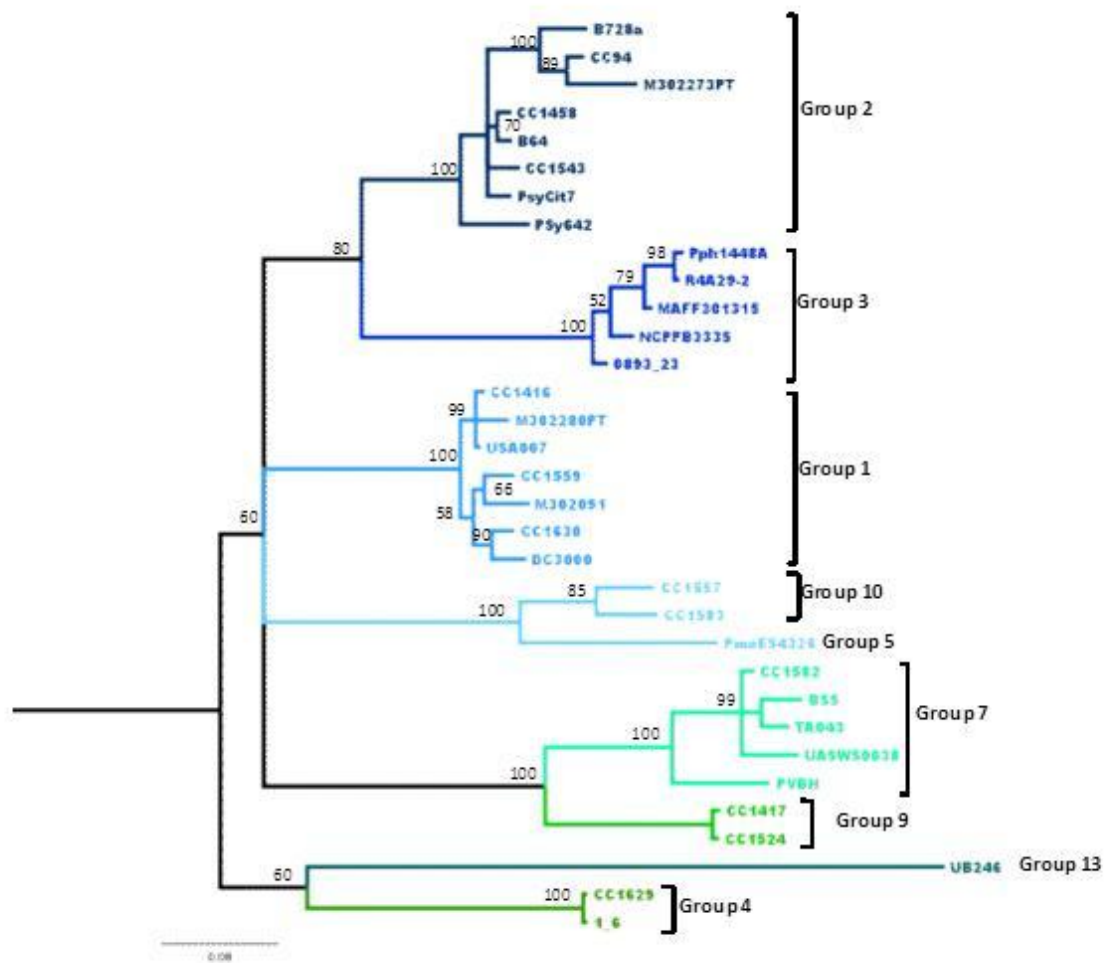
Phylogeny based on *mutS*

Chapter 6 - Hypermutableity in *Pseudomonas viridiflava*: programed bet-hedging via antibiotic resistance, exopolysaccharide production and pathogenicity



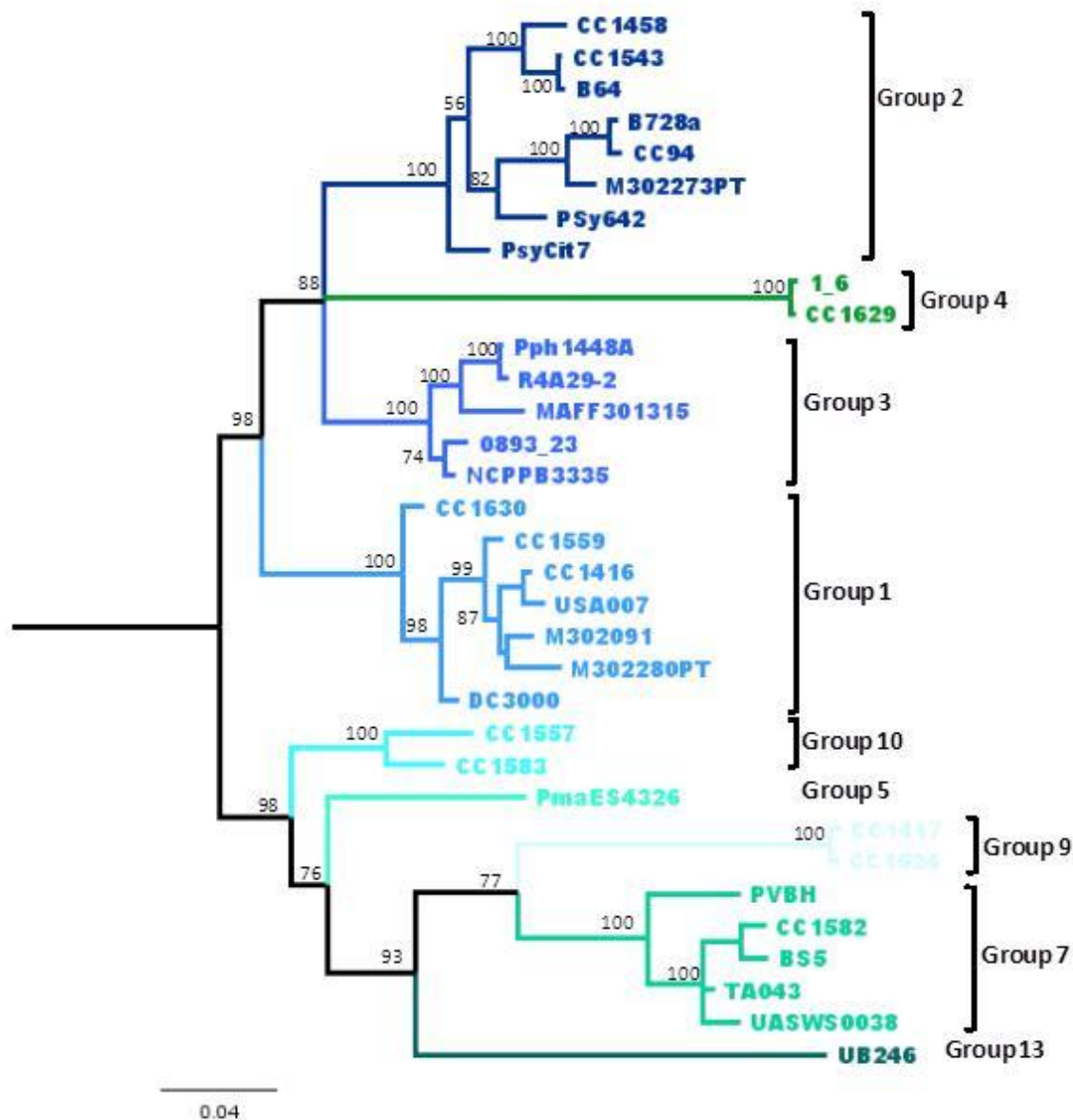
Phylogeny based on *mutL*

Chapter 6 - Hypermutability in *Pseudomonas viridiflava*: programed bet-hedging via antibiotic resistance, exopolysaccharide production and pathogenicity



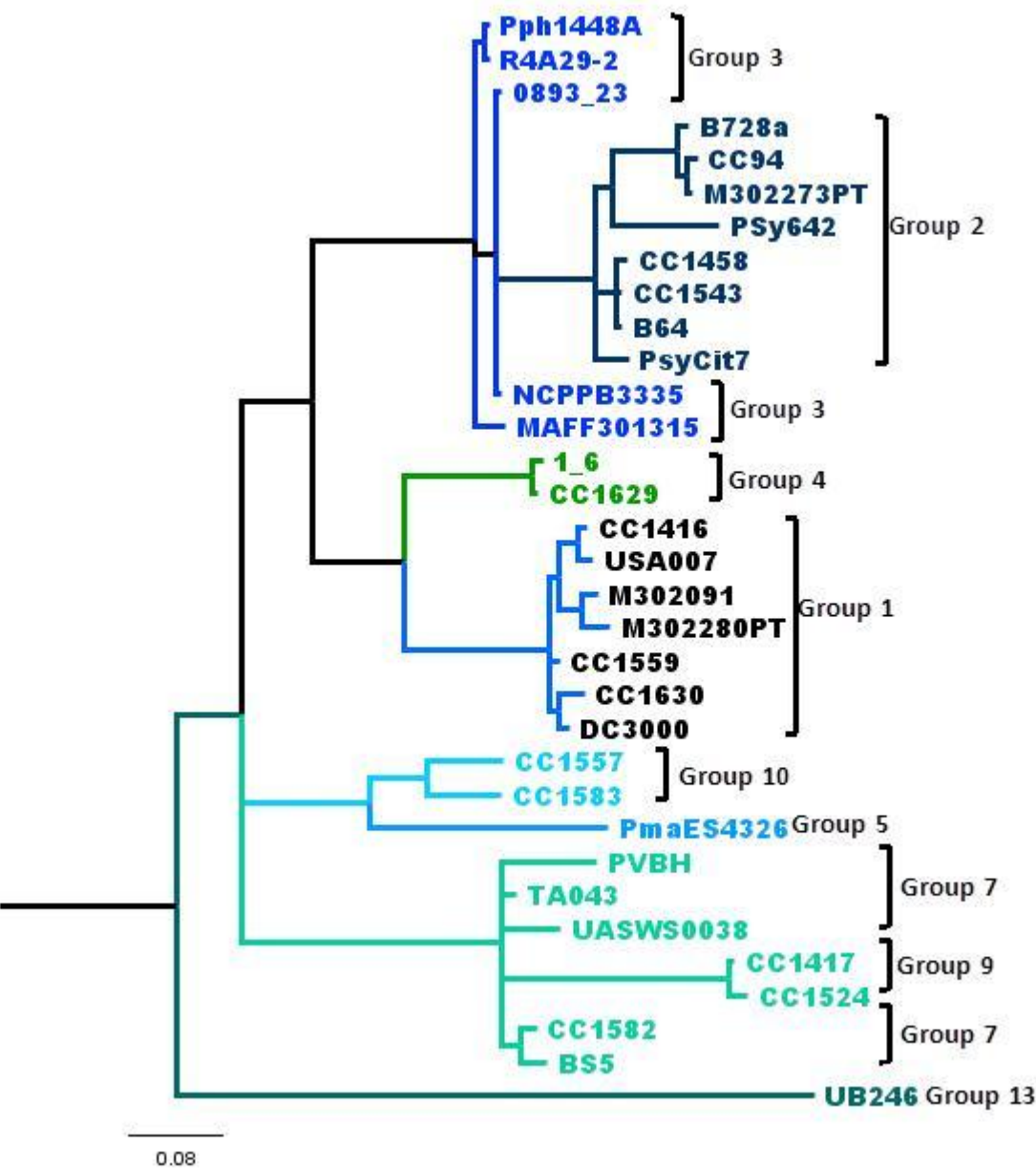
Phylogeny based on *mutT*

Chapter 6 - Hypermutability in *Pseudomonas viridiflava*: programed bet-hedging via antibiotic resistance, exopolysaccharide production and pathogenicity



Phylogeny based on *recA*

Chapter 6 - Hypermutability in *Pseudomonas viridiflava*: programed bet-hedging via antibiotic resistance, exopolysaccharide production and pathogenicity



Phylogeny based on *recX*

Chapter 7

General Discussion

General Conclusions

To date, studies concerning the emergence of new diseases have mainly focused on human pathogens while little is known about the mechanisms and the factors that are involved in the occurrence and evolution of new plant pathogens. The environment is not only the reservoir but also plays an important role in the evolution of pathogenicity traits in several human pathogens such as *Pseudomonas aeruginosa* or *Vibrio cholera*. For plant pathogens, however, the role of the environment has been demonstrated only in terms of being a reservoir of pathogenic strains.

In this thesis, I made an effort to highlight the different strategies that strains in the *P. syringae* complex could have adopted during their evolution that contributed to pathogenicity traits. In the review article I pointed out that although horizontal gene transfer (HGT), pathoadaptive mutations and prophage integration are the main molecular mechanisms that drive the acquisition of pathogenic determinants in bacteria, a defective methyl-directed mismatch repair system (MMR) underlies these mechanisms. In the past, studies about a defective MMR mainly focused on its role in the occurrence of antibiotic resistance cells. However, if we consider that MMR is the main barrier for HGT and for bacterial competence, we also should consider that MMR is an important factor in the occurrence of the molecular mechanisms that regulate the evolution of pathogenicity traits in bacteria. Further studies of the MMR system and its relationship to other mechanisms involved in DNA repair could provide a better understanding of the evolution of bacterial pathogens. It is also worth noting that the environment can modulate the occurrence of defective MMR through oxidative stresses or stresses related to exposure to UV-light. This opens clear hypotheses about the emergence of pathogens in the absence of direct interaction with the host thereby giving credence to paradigms of evolution of pathogens, plant pathogens in particular, beyond the host-pathogen co-evolution paradigm that currently reigns plant pathology (Morris et al, 2009).

Because we found highly diverse *P. syringae* phylogroups from different non-agricultural habitats, we can speculate that cultivated areas also harbour a wider diversity of this pathogen than what is currently appreciated. The bias in our knowledge likely comes from the focus on diagnosing the pathogenic agent responsible for disease. Further studies, that target the broader community of microorganisms associated with plants, independent of their suspected pathogenicity, are needed to elucidate whether the same diversity is present throughout the cultivated areas. Metagenomic analyses of plant-associated microbial communities are one

Chapter 6 - General Discussion

effort in this direction. But isolation of strains still provides the great advantage of allowing complete characterization of traits including fitness, host range and the overall combination of behaviours combined in each individual.

Different evolutionary strategies seem to have been deployed by strains in the *P. syringae* complex according to the substrates or life styles that typify them. We demonstrated that strains in phylogroup 7 and 8, previously named *P. viridiflava*, are widely distributed in the environment and that both environmental and crop strains possess one of either two types of T3SS. However, the evolution of the two T3SS does not seem to be related to the substrate of the origin or with the ability to induce disease. The only correlation with pathogenicity that we could find was that the absence of the conserved effector gene *avrE* was consistently associated with lack of pathogenicity in our tests. Hence, we speculate that the ability to produce pectolytic enzymes capable of inducing soft rot on potato, found only in phylogroup 7 and 8, might have led to the evolution of a less complex T3SS such as the Single-Partite T3SS (S-PAI). The latter was found in 90% of the *P. viridiflava* strains isolated from all the substrates. This is consistent with the prevailing view of *P. viridiflava* as a saprophytic opportunist. In this light, the capacity to produce pectolytic enzymes associated with a potato soft-rot phenotype, in conjunction with the expression of *avrE*, seems to reflect an important evolutionary strategy in phylogroup 7 and 8 that has involved maintaining both saprophytic and pathogenicity life-styles. In-depth future studies about the role of the *avrE* gene during host infection may help understanding and predicting pathogenicity of *P. viridiflava* strains. In fact, although we found a suitable molecular marker for the detection of *P. viridiflava* strains in general, it cannot discriminate between pathogenic and non-pathogenic strains.

Recently, it has been demonstrated that water habitats harbor *P. syringae* strains weakly pathogenic on tomato and that these strains share some of the Type Three Secretion System (T3SS) effectors with the clonal epidemic tomato strain (Monteil et al., 2014). By considering this recent study and the previous studies of the team in which I carried out my thesis, about the existence of *P. syringae* in non-agricultural contexts, I investigated the putative role of environmental habitats as reservoirs of *P. syringae* strains pathogenic to kiwifruit. To understand this role we used genes involved in the degradation of aromatic compounds, found in the genomes of *P. syringae* strains pathogenic for kiwifruit as well as for other woody pathogens, as molecular markers for strains potentially pathogenic for kiwifruit. These genes that are organized in an operon could represent an indispensable tool for some *P. syringae*

Chapter 6 - General Discussion

strains in colonizing woody tissue. By applying this approach, we found a pool of strains with potential to colonizing kiwifruit, such as other woody and herbaceous hosts and we observed that these strains were also able to cause disease symptoms on kiwifruit and on the other hosts. We demonstrated that this operon is conserved and that it was probably acquired by *P. syringae* before its diversification. The acquisition of genes related with the degradation of phenolic and aromatic compounds may represent an evolutionary strategy for adapting to woody hosts. In fact, such genes could degrade aromatic compounds produced during host infection which are toxic for the bacterium. Symptoms caused by environmental strains were less severe when compared with the epidemic strains responsible of the current epidemic of kiwifruit bacterial canker. Since the importance of the T3SS in bacterial plant colonization, we compared the effector repertoire of two environmental strains with the epidemic *P. syringae* pv. *actinidiae* strains, as well as other strains known to cause the first kiwifruit bacterial canker outbreak and those that are known to be weak pathogens. We found that environmental strains share 18 T3SS effectors with the epidemic Psa strains but lack 13 of the effectors present in the epidemic Psa. We hypothesised that the environmental strains that can co-exist with the epidemic Psa could horizontally acquire one or more of the 13 effectors and become more aggressive to kiwifruit. This evolutionary scenario, in which HGT plays a major role, could lead to the evolution of new pathogenic lineages in a “rapid” evolutionary scale. Future studies should focus on the potential occurrence of HGT in kiwifruit between the weak environmental strains and the more aggressive Psa strains to predict new emerging diseases.

The production of offspring with different phenotypes related with pathogenicity is another important strategy that has been described for bacterial pathogens of humans that leads to efficient infection and development of antibiotic resistance cells. Although phase variation has been previously described in some phytopathogenic bacteria such as *Ralstonia solanacearum*, little efforts have been made to describe and understand the mechanisms that characterize the occurrence of these variants, often associated with antibiotic resistance, in bacteria plant pathogens. In this thesis, I demonstrated that the formation of antibiotic resistance variants occur also in *P. viridiflava* strains both *in vivo* and *in vitro*. The evidence that these antibiotic resistance variants were not pathogenic and lacked pectolytic activity led us to speculate that these variants could be less sensitive to oxidative or other stresses during host infection. In human pathogens, the formation of antibiotic resistance variants is associated with a defective MMR system linked to mutations in MMR genes. However, and surprisingly, we did not find

Chapter 6 - General Discussion

mutations in MMR genes in *P. viridiflava*. Studies on the expression of MMR genes in *P. viridiflava* variants might provide insights on the modulation of MMR gene expression which may represent the basis of phase variation in *P. viridiflava*. A better understanding of the molecular mechanisms that regulate the formation of non-pathogenic variants in *P. viridiflava* strains could unveil clues for disease control.

The research activities that I carried out during the thesis allowed me to uncover several thematics which should be investigated through future studies in order to comprehend the mechanisms that regulate the emergence of *P. syringae* form reservoirs and its capacity to cause new diseases on different woody and herbaceous plants. There are several questions raised during the course of my thesis that I could not answer and that I believe deserve attention, especially for predicting new disease events and to improve disease control methods. These questions are:

- i) The *P. viridiflava* work clearly showed that it is difficult to predict pathogenicity of this bacterial species. On the other hand, we demonstrated that the *avrE* gene seems to be linked to the ability to induce disease thereby suggesting the possibility of disease control. The next step should focus on making targeted *avrE* mutants and test them on different host plants such as kiwifruit. In addition, we also described in the phase variation paper that transparent *P. viridiflava* variants are not able to induce disease contrarily to the wild-type and the mucoid variants counterparts although they grow in bean plants. Hence, by understanding the mechanisms that underline the occurrence of transparent variants, we could attempt to induce over-expression of these variants thereby controlling the disease. We could also investigate a way that allows to shift the mixed population of *P. viridiflava* toward stable transparent non-pathogenic populations. To achieve this goal, we should find the signal that regulates the occurrence of phase variation/hypermutableability in *P. viridiflava*.
- ii) In the paper that described the role of the environment in the emergence of kiwifruit bacterial canker we pointed out two interesting points. Firstly, we observed that the presence of the catechol operon involved in the degradation of phenolic compounds was related to endophytic growth in kiwifruit plants. The second step to be taken should be the creation of targeted mutants for this operon and testing them on kiwifruit and other woody plants. This will allow us to understand whether its presence is strictly correlated with the ability of *P. syringae* strains to colonize

woody tissue and to resist secondary lignin compounds that could otherwise be toxic for the bacterium. Secondly, environmental strains share some of the effectors found in the epidemic kiwifruit strains but they lack others that could be important in induction of disease. Since we demonstrated that environmental and epidemic strains can co-exist, the next step should be an experimental evolution assay that could elucidate the possible horizontal gene transfer of the effectors between less aggressive environmental strains and the epidemic one.

- iii) The work about hypermutability variants in *P. viridiflava* strains provided me a lot of aspects for reflection. As I mentioned in the first part of this discussion section, we can work on the mechanisms that drive the occurrence of non-pathogenic variants to control disease emergence. However this work let me think about the fact that a phytopathogenic bacterium has the same “behavior” as clinical strains. We found that hypermutable variants are also antibiotic resistant and they lost phenotypes associated with pathogenicity. The occurrence of these variants, that happens both *in vitro* and *in vivo* conditions, is not driven by the presence of antibiotics. This seems to be a natural mechanism that all bacteria have and that actually it is not related to the stresses to which the bacterium is exposed. I think that this is a common statement that starts to be really “famous” in the field of clinical microbiology, but the question that I ask is: could we use a plant pathogenic bacterium, as model to study the occurrence of antibiotic resistant mutants in human pathogens? I think that we could and I also think that we still need to understand if antibiotic resistant mutants are really pathogenic in some human bacterial pathogens. There are still a lot of inconsistencies in the literature between pathogenic variants (or mutants) and mutants associated with antibiotic resistance. One of the BIG projects I have in my mind is a comparative study of the mechanisms regulating this sort of phase variation in *P. viridiflava* and *P. aeruginosa* to better understand whether the mechanisms are the same and whether the antibiotic resistant mutants are really pathogenic.
- iv) From my point of view, the methyl-directed mismatch repair system is one of the most fascinating subjects of my thesis. In the preliminary results regarding the occurrence of hypermutable variants in *P. viridiflava*, I did not find any correlation with the MMR. However, this mechanism seems to be correlated with all the rest of the mechanisms that drive the evolution of pathogenic determinants at molecular level. We know that DNA polymerases in conjunction with the proofreading

Chapter 6 - General Discussion

exonuclease confer an error rate of $\sim 10^{-7}$ per bp per replication. The MMR corrects the mistakes that escape from this system by elevating fidelity up to 1000-fold. The biological effect of a defective MMR is then the occurrence of mutations that lead to variability. We can assume that the imperfect DNA replication system is made to create diversity that is linked with complexity. This complexity in the bacterial world could represent a new effector allele, or the acquisition of alien DNA that can confer to the bacterium the ability to colonize new substrates or hosts. I think that if we consider the importance of a defective MMR in disease emergence we can attempt to regulate and keep under a certain limit the occurrence of mismatch as disease control. This should universally be applicable for all pathogens during the infection period. In addition, finding the way to control the number of mismatches in eukaryotic cells means controlling diseases linked with DNA damage such as cancer.

Overall, I think that evolutionary experiments that elucidate the genomic exchanges between bacteria co-existing in a given host or substrate and the investigating the molecular mechanisms that regulate the occurrence of a defective MMR should be the next step in predicting new emerging bacterial disease.

References

- Morris, C., Bardin, M., Kinkel, L.L., Moury, B., Nicot, P.C., Sands, D.C. (2009). Expanding the paradigms of plant pathogen life history and evolution of parasitic fitness beyond agricultural boundaries. *PLoS Pathogens* 5: e1000693.
- Monteil, C. L., Cai, R., Liu, H., Mechan Lloncop, M. E., Leman, S., Studholme, D. J., Vinatzer, B. A. (2013). Nonagricultural reservoirs contribute to emergence and evolution of *Pseudomonas syringae* crop pathogens. *New Phytol*, 199, 800–811.

Chapter 8

English Summary

Italian Summary

Acknowledgements

English Summary

Pseudomonas syringae is a phytopathogenic bacterium that attacks both herbaceous and woody plants. In the last 20 years, more than 55 disease outbreaks caused by this bacterium have been reported on woody plants. The recent studies concerning the ecology of *P. syringae* demonstrated that this bacterium is present in several environmental habitats. These studies also showed that the environment is a reservoir of different *P. syringae* genetic lineages. In this thesis, I made an effort to understand what kind of evolutionary mechanisms can underline the emergence of new *P. syringae* pathogenic strains and how the environment can play a role in the evolution of pathogenicity traits of this bacterium. In particular, I focused on *P. syringae* lineages known to be pathogenic on kiwifruit. As a first step, I participated in a study about the diversity of *P. syringae*. This study was carried out in the laboratory of Cindy Morris at INRA of Avignon (France), where I performed almost all the experiments of the thesis. In the study concerning the *P. syringae* diversity, we showed that the genetic diversity (investigated on strains isolated from a wide range of substrates) of this species complex is wider than what was previously reported by considering only strains isolated from diseased crops. In fact, we found that *P. syringae* is composed by 13 phylogroups. In this thesis I showed that the phylogroup 7 and 8, also called as *P. viridiflava*, are formed by two different Type Three Secretion Systems. However, the evolution of these two systems does not seem to be related with the ability of *P. viridiflava* to cause disease. In addition I showed that *avrE* was the only effector correlated with the pathogenicity; it seems to be necessary but not sufficient for strains to be pathogenic.

In the study concerning the role of the environment as reservoirs of *P. syringae* pathogenic to kiwifruit, we demonstrated that strains from aquatic habitats are kiwifruit pathogens since they are able to multiply into the kiwifruit tissues as well as cause the bacterial canker disease. These environmental strains, such as the strains that cause the bacterial canker of kiwifruit, have an operon for the degradation of phenolic compounds. We then hypothesized that this operon is important during the kiwifruit vascularization. In addition, we demonstrated that the environmental strains share 18 effectors with the epidemic strains but they lack 13 of the effectors present in all the epidemic strains. We then hypothesized that less aggressive environmental strains can acquire new effectors by horizontal gene transfer and become new future epidemic strains. This scenario, as I described in the review article, could be regulated by a defective methyl-directed repair system. The understanding of why this system can be

defective could provide insights on the emergence of new diseases caused by *P. syringae* or other pathogens.

Italian Summary

Pseudomonas syringae é un batterio patogeno di piante erbacee e legnose. Negli ultimi 20 anni piantepiù di 55 nuove malattie causate da questo fitopatogeno sono state descritte. Dato che recenti studi sull'ecologia di *P. syringae* hanno messo in evidenza la sua ubiquità in differenti habitat ambientali, domande riguardanti il ruolo dell'ambiente come "reservoir" di linee genetiche patogene di *P. syringae* sono rapidamente emerse negli ultimi cinque anni. In questa tesi é stato fatto uno sforzo per comprendere quali meccanismi evolutivi possono entrare in gioco nell'evoluzione di ceppi patogeni di *P. syringae* e su come l'ambiente in *sensu lato* possa essere coinvolto nell'evoluzione di tratti connessi con la patogenicità. In particolare mi sono focalizzata su linee genetiche di *P. syringae* conosciute essere patogene per l' actinidia. Come primo passo nella mia ricerca ho collaborato ad un lavoro sulla diversità di *P. syringae* nel laboratorio di Cindy Morris, INRA di Avignone-France) in cui ho condotto la maggior parte delle mie ricerche di tesi. Nel lavoro riguardante la diversità di *P. syringae*, abbiamo messo in evidenza che la diversità conosciuta fino ad oggi, e basata su ceppi isolati da piante di origine agraria, é molto sottostimata. Analizzando ceppi provenienti da tutti gli ambienti abbiamo trovato che il complesso *P. syringae* é costituito da 13 gruppi genetici. Nella mia tesi, in particolare, ho messo in evidenza che il gruppo 7 e 8, costituiti da ciò che viene chiamato *P. viridiflava*, sono formati da due diversi Type Three Secretion System. L'evoluzione di questi due sistemi non sembra però essere connessa con l'abilità di *P. viridiflava* in indurre la malattia. Inoltre abbiamo messo in risalto che l'unico effettore connesso con la patogenicità é l'*AvrE*.

Nel lavoro riguardante il ruolo dell'ambiente come "reservoir" di patogeni dell'actinidia abbiamo dimostrato che ceppi isolati da ambienti acquatici rappresentano patogeni deboli ma capaci di crescere e sopravvivere in piante di actinidia allo stesso modo di ceppi che sono attualmente responsabili dell'epidemia del cancro batterico. Questi ceppi ambientali, come tutti i ceppi che causano il cancro batterico, hanno un operone per la degradazione di sostanze fenoliche, e abbiamo ipotizzato che questi geni siano importanti nella vascolarizzazione in tutte le infezioni di piante legnose. Inoltre abbiamo dimostrato che i ceppi ambientali condividono 18 effettori con i ceppi dell'epidemia ma ne mancano 13. Abbiamo quindi proposto che ceppi meno aggressivi provenienti da differenti substrati possano acquisire tramite trasferimento orizzontale nuovi effettori e quindi diventare più aggressivi e essere i rappresentanti di nuove future epidemie. Questo scenario evolutivo, come ho descritto nell'articolo di sintesi riguardante i meccanismi di evoluzione di nuove malattie emergenti, potrebbe essere regolato da un sistema

di riparazione del DNA (methyl-directed mismatch repair system) difettoso. Capire il perché questo sistema a volte non funziona bene ci potrebbe aiutare nel capire e nel controllare l'emergenza di nuove malattie causate da *P. syringae* ed altri patogeni.

Acknowledgements

Thousands of people I should thank in this thesis.....All of you that help me in the hard moments, when I was too tired to continue and when I was waiting for getting my papers accepted or when I was too stupid to understand the mistakes I was making. In all the cases, you kindly helped me to return in the good way.

First of all I definitively say thanks to you **Dr. Cindy Morris!!!!** I know you do not like me either using the word Dr. or calling you by your family name, but here I have to do it ☺. I remember the first time I met you, I was afraid about the BIG Cindy Morris, I mean BIG from the scientific point of view, of course!!!! I was shivering and I was not able even to talk to you...Oh my God.....And one minute after you just told me “please don’t call me Dr. Morris or I feel so old, in my mind I’m still a young student”. At that time, I really understood what kind of person you were. You taught me a lot of stuffs..... but firstly you taught how to be modest and how to have fun in our work. I was in elementary school before to meet you and now I could say that I’m in high school!!! Thanks for all the times you said NO and for all the times you said it’s a GOOD idea or it’s just a STUPID idea. Thanks for having taught me that nobody is right but that it’s just a question of how we look like in the world. I know that I’m not a genius and I don’t have sex appeal!!! but thanks to you that now I know I can do good science just being honest and modest with myself and with my colleagues. Really thanks for all the time you spent with me and for all the nice words you always had for me during the difficult time of my work and of my life.

Now it’s time for you **Dr. Odile Berge...**How I can describe what every day you’ve done for me. I think that your help to me is completely beyond words. You are the best French teacher that I ever had although I’m really sorry for not being a good student and also for what I every day do of your mother tongue ☺. You helped me in the lab with a good scientific way as well as for several daily problems. How I can forget the hours you spent with me for Sécurité Social or other crazy French social stuffs ??? And how I can forget all the conversation we had..... You were and you are so friendly Odileand having worked with you was a great fun and mostly fruitful...

Now that I thanked my two major points of reference I can start with all of you at INRA-Plant Pathology Research Unit. Firstly, many thanks to you ALL. A particular thanks to those who worked with me: **Caroline Monteil...**Yes my dear how many raw food we had together and

how much fun we had. Thanks a lot for all the awesome moments we had in the lab and out of the lab. You are a sweet girl and a fantastic researcher, keep going like this!!!! And what about **Charlotte Chandeysson** and **Caroline Guilbaud**, the two technicians that helped me during my experiments and that always had a nice words for me. You were always ready to help me when the amount of the work was too much, thanks!!! ...I would also say thank to **Dr. Philippe Nicot**, **Dr. Marc Bardin**, **Claire Troulet**, **Magali Duffaud**, **Dr. Chrystel Leyronas**, **Dr. Dominique Andurand** the mycology group of the Mistral team. All of you were too nice and friendly with me. I really enjoyed the time with you. Also a particular thanks to the new comers that joined the Mistral team recently, **Dr. Luciana Parisi**, **Benoit Borschinger** and **Jean-François Bourgeay**. I also like to thank some researchers and technicians of the virology group at INRA PACA, **Dr. Benoit Moury**, **Dr. Karine Berthier**, **Dr. Alexandra Schoeny**, **Vincent Simon**, **Catherine Rys**, **Dr. Eric Verdin** and **Dr. Cécile Desbiez**. Thanks to you all for the useful discussions. Many thanks to all the Green House technicians for your daily help in growing plants!! Thanks to **Claudine Laurent**, **Pascale Gros**, **Magali Eygrier**, **Helene Lopusziak**, (the Big Mama of the Department!!). A particular thanks to **Gregory Girardot** for the useful technical assistance and to **Karine Nozeran** and **Isabelle Bornard** for your friendly and nice words.

And now it's time for the external collaborators. Firstly to YOU dear **Prof. David Sands** from the Montana State University. Thanks Dave for all the ideas you provided to me, for all the times you called me just to say Hi!! and for believing in my research. You are the best, the youngest over 70 year-old scientist and one of the most creative people that I ever known. Also I want to thank **Prof. Boris Vinatzer** from Virginia Tech for all the fruitful discussions and **Prof. David Baltrus** from Arizona University for all nice discussions we had.

And after that, thanks to the Italian part of my thesis. First thank to **Prof. Leonardo Varvaro** the head of the DAFNE Department and the tutor of the thesis. Thanks for leaving me always free to take decision about the experiments and **THANKS** for having given me the possibility to work at the INRA labs until the end of the PhD. You really got me the opportunity to improve my knowledge. Also thanks to you **Dr. Giorgio Mariano Balestra**, you are the first one that gave me the opportunity to start the PhD and you are the first that gave me the chance to go in France for work with cindy. Particular thanks also to the people I met in the first PhD year at the Tuscia University in particular **Dr. Antonio Rossetti** (Heyyy you gave me interesting strains!! Thanks so much Anto), **Dr. Alfredo Fabi** you were always nice and friendly with me!! Thanks

also to **Dr. Marianna Stradaoli** and **Dr. Lorenzo Gallipoli** my PhD colleagues and the other colleagues of the Tuscia Plant Pathology Department.

A particular thanks to **Angélique** and **Jean-Marc Bastien** for taking care of my son Himmi during these two French years of PhD. You really make me the life easy....Sorry for all the times that I picked him up after 6 p.m!!!

And thanks to you that did not directly enjoyed my thesis but that helped me in other ways. Thanks to you my adorable son **Himal Lamichhane (Himmi)** I would write in Italian now...Mi hai accompagnato dall'inizio del mio Master. Stavi nel pancione quando lavoravo 8 ore al giorno in laboratorio per finire la tesi e sei nato poco dopo sono riuscita a discuterla. La tua nascita mi ha dato la motivazione di fare sempre del mio meglio nella vita. Mi hai insegnato tantissime cose ed è solo grazie a te se ora sto prendendo il dottorato. Immagino quanto sia stato difficile per te avere una mamma che lavora tanto, sei sempre stato troppo buono, ho scritto articoli, analizzato dati e fatto tanto altro con te a casa. E soprattutto grazie per l'amore che mi dai ogni giorno tesoro, che mi ha dato e mi darà la forza di andare avanti nel mio lavoro. Himmi grazie per avermi dato le armi per costruire la mia creatività. Ed ora grazie a mio marito, **Jay Ram Lamichhane (Buba)** per aver sempre creduto in me e per avermi dato la possibilità di partire in Francia e lasciare Himmi per 7 mesi solo con te. Siete stati troppo bravi tesori miei. Mi hai lasciato sempre libera di lavorare, mi hai dato la possibilità di andare in laboratorio durante il week-end facendomi sempre trovare la cena pronta al rientro. Mi hai enormemente aiutato con Himmi e senza di te non sarei mai riuscita a fare quello che ho fatto. Scusatemi tanto ragazzacci se in questi tre anni vi ho a volte trascurato...

In fine grazie ai mie genitori **Andrea e Maria** ed ai miei fratelli **Luca e Monica** per avermi aiutato con Himmi durante il mio dottorato soprattutto nei sette mesi in cui sono partita in Francia senza di lui. Grazie per essergli stati vicino in questo momento difficile. Grazie anche per avermi dato la possibilità di studiare ed iniziare l'Università.