The ecological genetics of *Pseudomonas syringae* in the kiwifruit phyllosphere

A thesis submitted in partial fulfilment of the requirements for the degree of

Ph.D.

in

Evolutionary Genetics

New Zealand Institute for Advanced Study

at Massey University, Auckland, New Zealand

Submitted by

Christina Straub

2017

Acknowledgements

An email written four years ago, joking about my taste buds appreciating the taste of the golden kiwifruit, resulted in me moving to the other side of the world.

Here we are, on what feels like a lifetime later and first and foremost, I would like to thank my supervisor Distinguished Professor Paul Rainey for giving me the opportunity to work on this fascinating project, for believing in me for starting from scratch, for the ability to pursue my own research project and for all the helpful discussions and ideas along the way. I am very grateful for all the years of financial support and the opportunities to present at various national and international conferences.

Dr Honour McCann, I would have been lost so many times without you, thank you for all your support and positivity at all hours and for teaching me all I know and so much more about plant pathogenicity experiments and genomics.

My fellow PhD lab mate in crime, Elena Colombi, who alternately makes me laugh or drives me crazy (in a split second), four years was not enough, lets do it for another one!

The past and present members of the bRain(e)y lab, who have been part of my journey: Katrin and Peter (lunch time cakes along with honest academic life advice!), Sylke, Chhavi, Yeserin, Chaitanya (whose passion is

not just game theory, but who is also a big movie buff), Kyo (thanks for showing me your Tokyo), Michael, Joanna and Gayle, Daniel and Philippe.

Dist Prof Paul Rainey, Dr Honour McCann, Dr Philippe Remigi and Dr Gayle Ferguson helped improve this thesis with valuable comments.

This project would not have been possible without a PhD scholarship from the NZIAS and financial support from Plant & Food Research. From Plant & Food Research I want to particularly thank Assoc Prof Matt Templeton and Dr Erik Rikkerink. Vesna Davidovic-Alexander, the secret star of the NZIAS for dealing with all financial related issues.

A special thanks goes to Punchbowl and the hard working kiwifruit growers, who granted me access to their orchards. Without them this study would not have been possible. For me it was a sampling site, but for them it was their livelihood. Keep the hard work going, David French, the Aitkens family and Rex Reed and thanks for sharing your experience.

On a more personal note, my family and friends. Your support has made this journey much easier and more bearable. Florian, words are not going to cut it to express my feeling of gratitude for all these years. Thanks for your understanding!

Krista, John and George, thank you for being my NZ family, all three of you will be forever in my heart. John, it will take me however a couple of years until I have verified all of your stories.

Abstract

The impact of disease-causing bacteria on their hosts is shaped by interactions with co-occurring microbes, but such interactions are rarely studied. *Pseudomonas syringae* is a ubiquitous and significant plant pathogen infecting a wide range of plants, often of agricultural importance. The community context of *P. syringae* in infected plant hosts has been little explored. I determined the population structure and genetic diversity of *Pseudomonas syringae* strains collected from infected and uninfected orchards over the course of a growing season during the current outbreak of bacterial canker of kiwifruit (*P. syringae* pv. *actinidiae, Psa*) in New Zealand. A total of 148 strains comprising Phylogroups 1, 2, 3 and 5 were characterised by Multi Locus Sequence Typing (MLST). The overall population structure was clonal, but with a low level of recombination for single housekeeping genes within phylogroups. More than half of the isolates belonged to a new Phylogroup 3 clade (PG3a) that was also commonly found on kiwifruit leaves in China and previously reported from kiwifruit leaves in Japan. To understand the ecological basis of the co-occurrence of PG3a and PG1 (*Psa*) I looked for evidence of niche specialisation by performing reciprocal invasion from rare assays of a selected representative from each lineage both *in vitro* and *in planta*. *P. syringae* G33C (PG3a) demonstrates antagonistic behaviour towards *Psa* NZ54, whereas *Psa* NZ54 exhibits a beneficial effect on growth of *P. syringae* G33C; an effect that could not be attributed to virulence activity encoded by the Type 3 Secretion System. Given this antagonistic

behaviour, I explored the virulence repertoire in these commensal strains to determine their potential in the emergence of future more virulent types of *Psa*. In addition, I used comparative genomics to unravel the phylogenetic resolution of the novel *P. syringae* clade in context with known representatives of *P. syringae* PG3. Together my data draw attention to the community context of disease and demonstrate the value of incorporating an ecological dimension into the study of the genetic structure of pathogen populations.

Table of Abbreviations

Abbreviation	**Meaning**
AHL	*N*-acyl derivatives of homoserine lactone
Avr	Avirulence
bp	Basepair
CDS	Coding sequence
CEL	Conserved effector locus
cfu	Colony forming unit
CDI	Contact-dependent inhibition
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Dinucleotide triphosphate
dpi	Days post inoculation
EEL	Exchangeable effector locus
EPS	Extracellular polysaccharides
ETI	Effector triggered immunity
HGT	Horizontal Gene Transfer
Hop	Hrp outer protein
HR	Hypersensitive response
IAA	Indole 3-acetic acid
ICE	Integrative conjugative element
Ice+/-	Ice-nucleation active/negative bacteria
INA	Ice-nucleation activity
Kan	Kanamycin

Abbreviation	**Meaning**
KB	King's B medium
LB	Lysogeny broth
M9	M9 medium
MAMP	Microbial associated molecular patterns
Mbp	Megabasepairs
MCMC	Markov chain Monte Carlo
ML	Maximum Likelihood
MLST	Multi Locus Sequence Typing
MST	Minimum Spanning Tree
NRPS	Non-ribosomal peptide synthetases
OD	Optical density
P.	*Pseudomonas*
PAI	Pathogenicity island
PAMP	Pathogen associated molecular patterns
PCR	Polymerase Chain Reaction
PG	Phylogroup
PGPR	Plant growth promoting bacteria
PRR	Pattern-recognition receptors
Psa	*Pseudomonas syringae* pv. *actinidiae*
PTI	Pattern triggered immunity
pv.	Pathovar
R genes	Resistance genes
RK	Receptor kinase
RLP	Receptor- like protein

Abbreviation	**Meaning**
ROS	Reactive Oxygen Species
SE	Standard error
ST	Sequence type
T3SE	Type 3 Secretion Effectors
T3SS	Type 3 Secretion System
WGS	Whole genome sequencing
WT	Wildtype

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Phyllosphere

The term phyllosphere was first coined in the 1950s by F. T. Last and J. Ruinen and originally referred to the plant foliage and solely the external surface of the leaf (Last, 1955; Ruinen, 1956). Nowadays the academic definition of the phyllosphere can vary widely. Morris (2002) considers the phyllosphere as the leaf surface and leaf tissue, because of a lack of clear-cut boundaries owing to natural openings such as stomata, whereas Doan & Leveau (2015) have gone as far as distinguishing between the leaf surface landscape ('phylloplane') and the leaf surface waterscape ('phyllotelma'). Considering the different plant components, the phyllosphere has been further divided up and referred to as the phylloplane (leaf), caulosphere (stem), anthospere (flowers) and carposphere (fruit).

In general, the phyllosphere is home to a diverse and lavish range of microorganisms (bacteria, yeasts, filamentous fungi and algae), but by far the most abundant are bacteria, which include pathogens, mutualists and non-pathogenic (commensal) microbes (Lindow & Brandl, 2003; Vacher *et al.*, 2016; Vorholt, 2012).

The phyllosphere can be considered as a short-lived habitat, e.g. in perennial deciduous plants, a new phyllosphere microbial community has to be recruited every spring. Local reservoirs have been described as the most likely source, with immigration from other part plants, neighbouring plants, soil, air and rainfall all possible origins for bacterial recolonization (Vorholt, 2012).

Microbial communities are defined as a group of different organisms occupying the same space at the same time and they are vital components of a functioning ecosystem. These communities present themselves as complex consortia of diverse microorganisms interacting with their environment and each other. Bacteria occupying the leaf surface are termed epiphytic (as opposed to endophytic bacteria residing in the apoplast) and can reach average population sizes of 10^6 - 10^7 per cm^2 in favourable conditions (Hirano & Upper, 2000). Considering their large population size and the number of plants on the planet, this amounts to a total of 10^{26} for the global plant foliage bacterial population (Morris & Kinkel, 2002). Nonetheless there is considerable variation in population sizes and composition of phyllosphere microbiota based on host genotypes, climatic factors, geographic factors and even at the level of individual plants with

temporal variations (Jackson & Denney, 2011; Lindow & Brandl, 2003; Thompson *et al.*, 1993). In a horticultural context, management practices such as the application of biocontrol agents (Zhang *et al.*, 2008), the use of antibiotics (Balint-Kurti *et al.*, 2011) or the farming method (organic vs. conventional farming; Li *et al.* (2012)) can all influence microbial foliage populations.

1.1.1 A challenging habitat for microbes

The living conditions encountered by epiphytic bacteria on the leaf surface are often hostile. The specific microclimate with rapid fluctuation in UV radiation, water availability and limited nutritional resources on a varying temporal scale make the phylloplane a challenging habitat for microbes. But bacteria are known to cope with adverse conditions and compared to extremophile bacteria which are notorious for surviving in extreme environments from hot springs to Antarctic ice, tolerating high salinity, extreme temperatures or proliferating in metal-rich milieus, the conditions on the leaf surface seem benign and almost inviting (Rampelotto, 2013).

The leaf surface itself is a heterogeneous habitat given its topographic features made up of stomata, trichomes, hydathodes or leaf veins. The first hurdle that bacteria encounter is the cuticle, the outermost layer of the leaf covering the epidermal cells. The cuticle is composed of the cuticle layer, a wax film and wax crystals, whose main function is to regulate water loss/uptake for the plant, limit the damaging effects of UV radiation and to provide a structural barrier against invasion by microorganisms (Müller &

Riederer, 2005; Serrano *et al.*, 2014). Apart from this physical barrier against microbial attacks, there are other mechanisms to defend plants from pathogens. Plant cells are highly dependant on innate immune responses like effector triggered immunity (ETI) and pattern triggered immunity (PTI) (Jones & Dangl, 2006) (described in further detail in section 1.2).

Phytobacteria employ several strategies to increase their epiphytic fitness and to exploit the sparse resources available in the phyllosphere. In adaptation to the habitat, epiphytic bacteria evolved traits like production of biosurfactants, extracellular polysaccharides (EPS), antibiotics, quorum sensing and production of plant hormones (e.g. auxin) (Vorholt, 2012) (Figure 1.1). The excretion of biosurfactants can change the physicochemical properties of the cuticle to their advantage by increasing the wettability of the leaf and water availability (D'aes *et al.*, 2010; Schreiber *et al.*, 2005). Production of biosurfactants, e.g. EPS, improves motility, shields bacteria from osmotic fluctuations and protects from oxidative stress caused by reactive oxygen species (ROS) (Lindow & Brandl, 2003; Quiñones *et al.*, 2005). Quorum sensing signalling molecules (*N*-acyl derivatives of homoserine lactone, AHL) are secreted by many Gram-negative proteobacteria and are used to initiate a concerted gene expression action (motility, virulence and EPS production) upon reaching a certain threshold (or critical concentration) (Quiñones *et al.*, 1999; Whitehead *et al.*, 2001). Apart from these traits, fluorescent microscopy studies revealed that bacteria tend to migrate towards nutrients, where they form aggregates, which usually consist of multiple species with a higher likelihood for aggregation for species of the same phylogenetic group

(Monier & Lindow, 2005; Remus-Emsermann *et al.*, 2014). This subjects clusters of bacteria far less to changing environmental conditions in comparison to solitary cells. An additional observed strategy for epiphytic bacteria is stress avoidance by actively choosing protected sites on the leaf surface (Wilson *et al.*, 1999).

Another well-known characteristic for phyllosphere bacteria is bacterial ice-nucleation activity (INA). The colonisation of bacteria possessing ice nucleation abilities (Ice^+) can cause frost damage in plants, releasing nutrients and water from their cells to be consumed by bacteria. INA is often found, but not restricted to, Gram-negative proteobacteria like *Pseudomonas syringae*, *Xanthomonas* sp., *Pantoea* sp., *Erwinia* sp. and *Pseudomonas fluorescens* (Hill *et al.*, 2014; Lindow *et al.*, 2013; Failor *et al.*, 2017) and allows Ice^+ bacteria to reduce the supercooling ability of plants (usually $-6\,°C$ to $-8\,°C$, reduced to slightly below $-1.5\,°C$), resulting in frost-injury to plants (Lindow, 1983). The INaZ protein, an outer membrane lipoglycoprotein, was found to be involved and ice-nucleation activity highly depended on the cell concentration (Cochet & Widehem, 2000; Green & Warren, 1985). In fact, recombinant Ice^- *P. syringae* bacteria were the first engineered microorganisms to be released into the environment, as competitive exclusion proved efficient in frost damage control (Lindow, 1995).

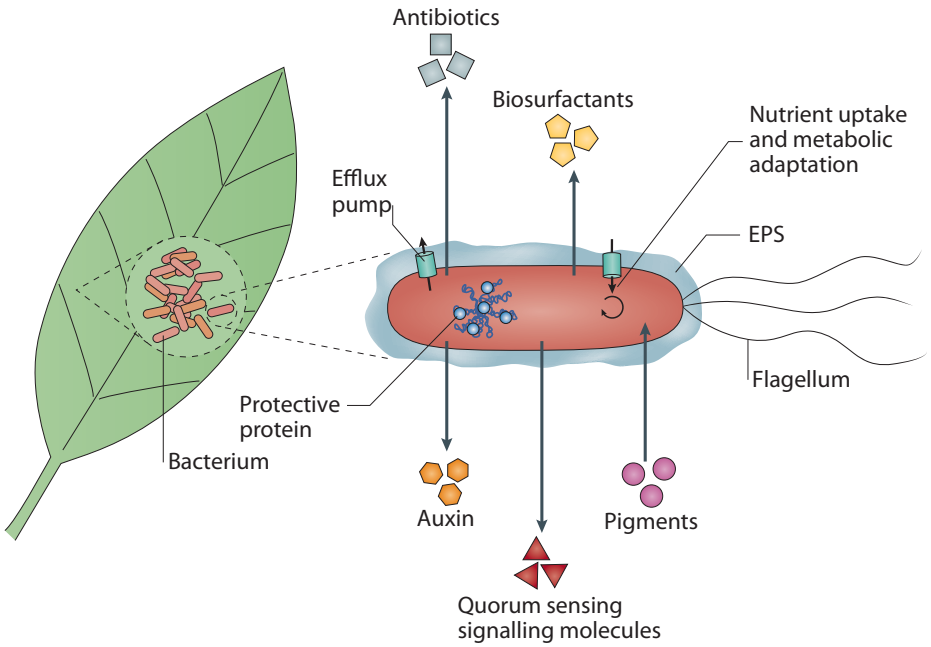

Figure 1.1: Overview of traits evolved by bacteria to adapt to the specific conditions in the phyllosphere. Reprinted with permission from Macmillan Publishers Ltd: Nature Reviews Microbiology, Vorholt (2012), copyright 2012.

1.1.2 Outcome of the plant-bacteria interplay

Bacterial colonisation exerts a profound influence on the host, and hosts evolve because of continual exposure to and interaction with microbes, with microbiota being transmitted between generations (Rosenberg & Zilber-Rosenberg, 2013, 2016). These interactions can result in a parasitic, mutualistic or commensal relationship between host and microbes (Bulgarelli *et al.*, 2013), with the distinction between mutualistic and commensal being hard to pinpoint, as interactions represent a spectrum (Vorholt, 2012).

The majority of epiphytic bacteria are commensal or mutualistic, with mutualists benefiting the plant and by definition commensals providing no harm or benefit to the plant. The presence of either has no negative outcome for the host in terms of virulence or disease. However, commensals are not only a random assortment of bacteria associated with a plant by chance, rather they appear to be regulated by the plant itself (Delmotte *et al.*, 2009; Redford *et al.*, 2010) and are passed on to offspring (Rosenberg & Zilber-Rosenberg, 2013). These commensals can in fact be of beneficial use to the plant and have even been referred to as 'plant probiotics' in regard to potential applications of a mixture of individual commensal species in an agricultural context aiding plant health (Berlec, 2012).

Plants have long been known to benefit from bacterial colonisation without conferring pathogenicity by boosted plant growth, increased plant health or reduction in pathogen infections (Berg, 2009; Lugtenberg & Kamilova, 2009). This area of research became established in the 1980s with a focus on plant growth-promoting bacteria in the rhizosphere (PGPR) (Kloepper *et al.*, 1980a,b). A variety of pseudomonads can have a beneficial effect on plant growth (Cook *et al.*, 1995; Preston, 2004; Rainey, 1999; Schippers *et al.*, 1987). Direct mechanisms involved are e.g. the stimulation of plant hormone production like indole-3-acetic acid (IAA), ethylene, sequestering of iron by production of siderophores, solubilization of posphate for plants, toxins, cytokinins, auxins or gibberellins, although these hormones can be directly synthesised by some beneficial bacteria (Calvo *et al.*, 2014; Compant *et al.*, 2005; Haas & Défago, 2005; Handelsman & Stabb, 1996). Some commensals produce protective

antimicrobial compounds, allow fixation of atmospheric N_2 or contribute indirectly via suppression of plant pathogenic bacteria e.g. through competitive exclusion (Fürnkranz *et al.*, 2008; Lindow & Leveau, 2002; Raaijmakers & Mazzola, 2011). Over the past decades, these beneficial traits have been utilized in agriculture by developing biocontrol agents from plant growth-promoting bacteria like e.g. *Agrobacterium, Arthrobacter, Bacillus, Pantoea, Pseudomonas* or *Stenotrophomonas* to increase productivity, although with limited success (Berg, 2009; Compant *et al.*, 2005; Pérez-García *et al.*, 2011; Raupach & Kloepper, 1998; Weller, 1988).

However, not all plant-bacteria interactions result in neutral or beneficial outcomes for the plant as a host. Plant pathogens as disease causing agents lead to reduced fitness of the host; they can invade the leaf apoplast and penetrate the plant cell wall, resulting in weakened plant growth and reproduction and ultimately plant death. Apart from this obvious outcome, there is also an ongoing evolutionary arms race between pathogens evolving to outwit and manipulate the plant immune defence system and the plant evolving resistance to defend against the pathogens; this arms race is strongly dependent on the environmental and life history attributes of either associated partner (Burdon *et al.*, 2013) (described in more detail in section 1.2 and 1.4.2).

1.2 Plant immune defence

Plants are constantly exposed to microbes and possess innate immunity and resistance to many bacteria. The immune response of a plant triggered by pathogens has been described as a 4-phase zigzag model by Jones & Dangl (2006) (Figure 1.2).

Plants possess pattern-recognition receptors (PRRs), which can be either surface-localized receptor kinases (RKs) or receptor-like proteins (RLPs) (Zipfel, 2014). The first level of immune reaction involves PRRs recognizing specific conserved molecules of microorganisms called MAMPs (microbial associated molecular patterns). Prior classification referred to those molecules as PAMPs (pathogen-associated molecular patterns), but as they are not unique to pathogens, the more generic term MAMP was introduced (Mackey & McFall, 2006). These MAMPs can be general elicitors like lipopolysaccharides, flagellin, elongation factor EF-Tu or cold shock protein (Boller & Felix, 2009; Felix & Boller, 2003; Gómez-Gómez & Boller, 2002). Perceiving MAMPs results in Pattern Triggered Immunity (PTI) as a basal immune defence, which includes production of reactive oxygen species and nitric oxygen, hormonal changes, callose production to reinforce plant cell walls or induction of MAP kinase signalling pathways (Nürnberger *et al.*, 2004; Tsuda & Katagiri, 2010).

Successful pathogens suppressing or interfering with PTI face the second tier of the plant defence system called effector-triggered susceptibility (ETS), which is caused by effector proteins deployed into the

cytoplasm by bacterial pathogens upon successful penetration of the plant cell (Jones & Dangl, 2006).

Third, if a specific effector is recognized directly or indirectly (via nucleotide binding and leucine rich repeat domains, NB-LRR) by plant resistance (*R*) genes, effector triggered immunity (ETI) occurs (Collier & Moffett, 2009).

ETI typically results in disease resistance and programmed cell death (hypersensitive response, HR) and is a faster, prolonged and more stringent response compared to PTI (Chisholm *et al.*, 2006; Dangl *et al.*, 2013; Jones & Dangl, 2006). ETI produces similar physiological reactions as PTI, including callose deposition, phytoalexin synthesis, production of ROS, changes in Ca^{2+} levels and HR (Jeong *et al.*, 2009). Strictly speaking, we cannot differentiate between ETI and PTI, but there is rather a continuum among the immune responses (Thomma *et al.*, 2011).

The fourth tier involves the interplay between bacterial virulence genes and *R* genes, which is driven by a constant and complex co-evolutionary arms race of host defence versus pathogen attack across spatial and temporal scales (Ravensdale *et al.*, 2011). Phytopathogens evolve or alter effectors to suppress ETI, which drives plants to evolve new receptors targeted towards the "novel" effectors to again elicit ETI, with the evolutionary cycle continuously putting selection pressure on each organism.

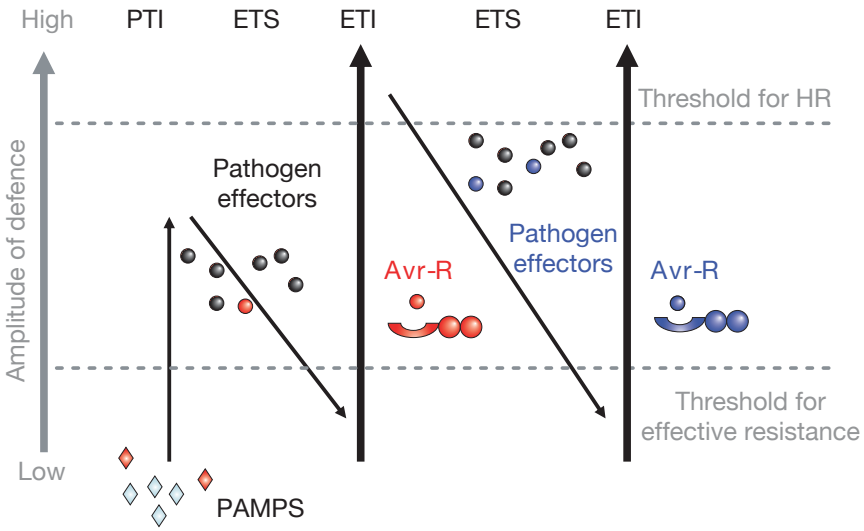

Figure 1.2: Zig-zag model of plant immune response versus plant pathogen infection. Reprinted with permission from Macmillan Publishers Ltd: Nature, Jones & Dangl (2006), copyright 2006.

1.3 Factors driving pathogen emergence

Emerging plant diseases are not only a major threat to global food security and agriculture leading to substantial economic loss, but they also affect wild plants and levels of biodiversity. By definition, according to WHO (World Health Organisation), an emerging disease either occurs for the first time (newly emerging disease) or re-emerges by infecting a new host or by acquiring additional virulence traits towards the original host. Bacteria account for approximately 16% of emerging plant diseases worldwide (Anderson *et al.*, 2004).

There are a variety of pathways for how novel pathogens can arise, via

(i) anthropogenic or zoogenic transmission or dissemination via environmental factors like e.g. rain or climate change; (ii) adaptation via molecular mechanisms (e.g. horizontal gene transfer); (iii) host shifts or host jumps or (iv) arising from a pool of environmental strains (Anderson *et al.*, 2004; Bartoli *et al.*, 2016; Goss *et al.*, 2013; Stukenbrock & McDonald, 2008).

Introduction of bacterial plant pathogens via anthropogenic sources has been identified as the major source for novel outbreaks (Anderson *et al.*, 2004). Although the collected data from this study do not cover the past decade, the trade of plant material (e.g. pollen, fruit, germplasm, grafts, live plants or seeds) between countries is ongoing and thriving. *Pseudomonas syringae* pv. *actinidiae* (*Psa*) is an example of a disseminated plant pathogen, as it was most likely introduced to New Zealand in 2010 with the import of contaminated pollen. Human-mediated transmission of plant diseases can clearly not be denied, and the uniformity of the host populations in an agricultural setup and the environmental homogeneity facilitate the spread of pathogens (McDonald & Stukenbrock, 2016).

Several molecular mechanisms result in the emergence of novel bacterial pathogens, including point mutations, genomic rearrangements or horizontal gene transfer (e.g. acquisition of novel virulence clusters or other ecologically highly relevant factors like, for example, adaptive genes for colonization of specific hosts, and nutrient acquisition) (Bartoli *et al.*, 2016; Polz *et al.*, 2013). Genomic rearrangements include inversions, deletions and duplications, which can arise via recombination through transposons,

plasmids, mobile genetic elements (e.g ICE integrative conjugative elements) or prophages. These variations in genomic structure are observed in many pathogens, from plant-associated *Pseudomonas* sp. (Silby *et al.*, 2011), to the human pathogen *Yersinia pestis*, which showed extreme genomic variability even within members of the same species (Darling *et al.*, 2008). Horizontal gene transfer (HGT) refers to processes where bacteria directly acquire foreign genetic elements: either freely available DNA from the environment (transformation), via plasmids (conjugation) requiring cell-to-cell contact, or via phages (transduction). HGT is a major driver of bacterial evolution which has long been recognised to fuel genomic diversity (Koonin *et al.*, 2001; Ochman *et al.*, 2000) and only recently has been viewed in an ecological context (Polz *et al.*, 2013; Popa & Dagan, 2011). In fact, ecological context has a bigger influence on the frequency and extent of HGT than either phylogeny or geography, with higher levels of HGT occurring among members of the same ecological niche, independent of phylogenetic distance (Smillie *et al.*, 2011). Processes of HGT can greatly influence the evolution of pathogens by acquiring new traits like antibiotic resistance (e.g. Han *et al.*, 2004; Sundin & Bender, 1993) or copper resistance (e.g. Colombi *et al.*, 2017). HGT can result in the ability to colonize and infect a novel host (Smillie *et al.*, 2011), or confer other phenotypic advantages which increase bacterial fitness like lipopolysaccharide synthesis (Patil & Sonti, 2004) or cellulose polymers (encoded on the wss operon) (Gal *et al.*, 2003). A major driver of plant pathogen emergence undoubtedly is the acquisition of virulence genes via HGT resulting in a more virulent phenotype. Genetic material containing

virulence genes transferred horizontally often comes in the form of syntenic chunks, which are also referred to as pathogenicity islands (PAIs) (Hacker *et al.*, 1997) and they comprise mobile DNA elements like integrative conjugative elements (ICEs), conjugative transposons and prophages (Juhas *et al.*, 2009).

Host shifts of a pathogen occur when the newly colonized host is genetically similar, whereas a host jump involves a genetically distinct host. Factors involved in the adaptation of a pathogen to a novel host and disease incidence are the evolvability of the pathogen, the invasibility of the host-environment (community composition and population structure) and associated ecological and environmental dynamics (like temperature or humidity, changes in host phenotype and genotype) (Engering *et al.*, 2013; Woolhouse *et al.*, 2005). Host shifts are largely associated with anthropogenic dissemination, the exchange and introduction of plant material or a change in agricultural practices. The latter was most likely the cause of an epidemic of *P. syringae* pv. *avellanae*, which had been infecting cultivated hazelnuts (*Corylus avellana*) and wild relatives in Greece and Italy in the past (O'Brien *et al.*, 2012). Scenarios like co-cultivation in an agricultural context facilitate host jumps where the original infected host plant serves as the reservoir and co-cultivated healthy plants, or wild plants and weeds adjacent to the field, are exposed to the pathogen until a spill-over occurs, provided that host and pathogen are compatible. Alternatively, we can imagine a scenario of a pathogen established on a wild plant, which evolved to infect a related cultivated crop variety. For example, this was most likely the case for *Xanthomonas*

campestris pv. *musacearum*, which was originally isolated from cultivated and wild enset (*Ensete ventricosum*) in the 1960s, but has since become an epidemic upon discovery of infected cultivated varieties of banana (*Musa* sp.) (Biruma *et al.*, 2007).

The role of environmental and commensal bacteria in the emergence, evolution and outcome of pathogenic interactions is well-appreciated for studies of human pathogens, but remains understudied for plant pathogens. A hypothetical evolutionary scenario begins with a diverse community of frequently recombining bacteria in the environment. By chance, a member of this community encounters a suitable plant host and emerges from the pool, undergoing clonal expansion. Selection for increased fitness on this specific host occurs, and these specialised strains are then dispersed back into the environmental pool (Goss *et al.*, 2013). For example, *Vibrio cholerae*, a widely distributed human pathogen encountered in aquatic environments, causes devastating epidemics of cholera. Environmental populations of *V. cholerae* carry virulence gene clusters, which may have been independently acquired from epidemic clones and serve as a reservoir for future horizontal gene transfer to epidemic lineages (Azarian *et al.*, 2016; Faruque *et al.*, 2004). Such a scenario has also been demonstrated for *Pseudomonas syringae*, a ubiquitous phytopathogen, which is known to persist in different environments including irrigation water, rivers, ice, leaf litter and phloem-feeding insects, irrespective of the presence of a host plant (Monteil *et al.*, 2012; Morris *et al.*, 2008, 2007; Stavrinides *et al.*, 2009). Previous population genetics studies have focused mainly on pathogenic isolates collected from diseased hosts (Cai *et al.*, 2011;

Gutiérrez-Barranquero *et al.*, 2013; Hwang *et al.*, 2005; Kaluzna *et al.*, 2010; Sarkar & Guttman, 2004), but recently interest has shifted to reveal more about the diversity of environmental isolates and their evolutionary potential in the emergence of virulent types (Bartoli *et al.*, 2015; Monteil *et al.*, 2013, 2016). Recently it was shown that epidemic clones of *P. syringae* pv. *tomato* (*Pto*) most likely have emerged from a diverse metapopulation in the environment via a small number of evolutionary steps (Monteil *et al.*, 2013).

1.4 *P. syringae* - a taxonomist's nightmare

Pseudomonas syringae, a Gram-negative gamma proteobacterium, is a ubiquitous plant pathogen and a major component of the phyllosphere bacterial community, by contributing up to 50% (Ercolani, 1991). *P. syringae* is notoriously famous in part because of the breadth of its host range. More than 80 plant species have been listed as hosts of *P. syringae* (Bradbury, 1986). The species name originates from its host of isolation: a diseased lilac (*Syringa vulgaris* L) sampled in 1899 by M. W. Beijerinck and named by C. J. J. van Hall in 1902. Identification as *P. syringae* was originally done according to the LOPAT test (+ for levan sucrose, - for oxidase activity, - for potato rot, + for arginine dihydrolase and + for hypersensitive response on tobacco) (Lelliott *et al.*, 1966). Because of the large number of associated hosts, *P. syringae* strains have been further sub-classified into 60 pathovars (pv.) according to the host of isolation (Bull *et al.*, 2010), but pathovar designation based on phenotypic tests has

important limitations as some isolates are able to infect a variety of different host species and pathovars were defined based on a limited host range test. This heterogeneous nature has its complications for taxonomic classification and assessing the whole host range is difficult, especially if the bacteria were isolated from the environment or from asymptomatic plants.

P. syringae has been further classified using DNA-DNA hybridization studies, where nine 9 genomospecies were identified (Gardan *et al.*, 1999; Young, 2010). Multi Locus Sequence Typing (MLST), where a number of housekeeping genes are compared, led to discovery of 13 different phylogroups (PG) for environmental and agricultural isolates and has been the classification scheme of choice (Berge *et al.*, 2014; Hwang *et al.*, 2005; Sarkar & Guttman, 2004). Some of these phylogroups harbour strains that have been designated with a species name such as *Pseudomonas viridiflava* and *P. cichorii*, which do not possess a canonical Type 3 Secretion System (see section 1.4.2.1), but are nonetheless assigned to *P. syringae* (Bartoli *et al.*, 2014; Berge *et al.*, 2014). In summary, the classification of *P. syringae* has been biased strongly towards phenotypic categorization, disregarding evolutionary principles (Baltrus *et al.*, 2017).

Over the years *P. syringae* has become a model organism for studying plant-microbe interactions (most famously the zigzag model of plant immune response by Jones & Dangl (2006))), the molecular basis of pathogenicity (e.g. Alfano & Collmer, 2004; Collmer *et al.*, 2000; Galán & Collmer, 1999), the epidemiology of disease (McCann *et al.*, 2013, 2017) and unravelling the natural diversity in an agricultural and environmental

context (e.g. Berge *et al.*, 2014; Monteil *et al.*, 2013; Morris *et al.*, 2013). In fact, *P. syringae* has been (in)famously listed as number 1 in the Top 10 plant pathogenic bacteria in plant pathology (Mansfield *et al.*, 2012). *P. syringae* is a major problem in agriculture and is responsible for many disease reports on economically important plants. Plant diseases cause legitimate concern for food production, particularly if the outbreak is severe and widespread (Strange & Scott, 2005). In the last century, diseases caused by a member of the *P. syringae* complex were reported on as many as 40 annual plant species (Lamichhane *et al.*, 2015) and 55 diseases in woody plants (Lamichhane *et al.*, 2014). For example, *P. syringae* pv. *aesculi* has been causing major damage to European horse chestnut trees since 2002 (Green *et al.*, 2010). A frightening example of a re-emerging disease caused by *P. syringae* was *P. syringae* pv. *actinidiae* (*Psa*). In 2008 a devastating outbreak of *Psa* spread in a pandemic fashion across the world.

1.4.1 *P. syringae* causing havoc in the kiwifruit industry

Pseudomonas syringae pv. *actinidiae* (*Psa*) is the causative agent of bacterial canker on kiwifruit (*Actinidia* spp.). It was first reported and described to cause canker, leaf spot lesions and browning of flower buds on green kiwifruit (*Actinidia chinensis* var. *deliciosa*) in the early 1980s in the US (California) and Japan (Opgenorth *et al.*, 1983; Takikawa *et al.*, 1989). Subsequently, *Psa* was detected in other kiwifruit growing regions around

the world, including Italy and Korea in the early 1990s (Koh, 1995; Scortichini, 1994). These disease occurrences were only reported on green kiwifruit, the only commercially available cultivar at that time. After the introduction of a yellow-fleshed cultivar 'Hort16A' (*A. chinensis* var. *chinensis*) in early 2000, it wasn't long until symptoms of *Psa* were first discovered, in Korea in 2006 (Koh *et al.*, 2010). A severe global pandemic soon followed, starting with observations from central Italy in 2007/2008 (Balestra *et al.*, 2008), with subsequent disease reports from the neighbouring European countries Spain, Portugal, France and Turkey (Abelleira *et al.*, 2011; Balestra *et al.*, 2010; Bastas & Karakaya, 2011; Vanneste *et al.*, 2011). *Psa* has since spread globally and has been recovered in most of the kiwifruit growing regions around the world, including South East Asia (Koh *et al.*, 2012; Sawada, 2015; Zhao *et al.*, 2013), Chile (Promedmail archive 20110325.0940), Australia (EPPO, 2011) and New Zealand, where it arrived in 2010 (Everett *et al.*, 2011). Four different clades have been identified for *Psa*, with the earliest isolates from Japan and Italy forming clade 1 (*Psa*-1), Korean isolates from the early 90s forming clade 2 (*Psa*-2), the current pandemic isolates referred to as clade 3, and clade 5 has so far only been identified in a very localised area in Japan (Fujikawa & Sawada, 2016). Considerable effort has gone into identifying the origin of the canker-causing pandemic (*Psa*-3) and it was shown conclusively that the current pandemic caused by *Psa*-3 originated from China, but the native home of *Psa* is most likely in Korea/Japan, which have shown the presence of multiple clades of *Psa* (McCann *et al.*, 2013, 2017). In New Zealand, currently 90% of land dedicated to kiwifruit

production in New Zealand has been infected with *Psa* (March 20, 2017, http://www.kvh.org.nz/maps_stats) with *A. chinensis* var. *chinensis* cv. Hort16A being particularly susceptible to *Psa*. Since the outbreak of *Psa* in New Zealand, 2000 ha of kiwifruit orchards had to be removed (Frampton *et al.*, 2012). Coincidentally, the new gold variety 'G3' was introduced to the market simultaneous to the start of the outbreak and has proven to be more resistant to *Psa* infections, therefore the remaining rootstock of cut down Hort16A orchards had G3 grafted. As a foliar pathogen, *Psa* is initially a member of the phyllosphere bacterial community, as colonisation and proliferation of the leaf surface is an important part of the life cycle. Upon successful establishment of a stable population in the phyllosphere, it enters the apoplast via natural openings in the leaf like stomata, trichomes, hydathodes or wounds, to further proliferate in the intercellular space. Infected plants present typical symptoms like browning of flower buds, shoot dieback, shrivelling of fruit, brown leaf spots surrounded by chlorotic haloes and the typical canker with white or red exudate oozing from wounds in the stem. Induction of disease does not ensure a one directional colonization pathway for a pathogen, in fact ingress and egress of bacteria throughout the infection process has been described for *P. syringae* as well as *Xanthomonas campestris* (Beattie & Lindow, 1999).

1.4.2 Plant pathogen artillery

Many plant pathogens possess the Type 3 Secretion System (T3SS), which is a particular structural machinery that allows infiltration of plant cells and

the delivery of certain effector proteins (T3SE) into the cytoplasm. Effectors are proteins that are delivered directly into the cytoplasm via the needle-like structure of the T3SS and have been termed avr (avirulence) proteins for their ability to induce resistance and Hop (Hrp outer proteins) for those injected via the needle, although some Hop proteins can also induce resistance (Alfano & Collmer, 2004).

The Type 3 Secretion pathway resembles a needle-like structure and is encoded by *hrp* (HR and pathogenicity) and *hrc* (HR and conserved) genes, with some hrp proteins being secreted as extracellular accessory proteins and the conserved hrc proteins directing secretion of effectors across the bacterial envelope (Alfano *et al.*, 2000; Collmer *et al.*, 2000). Hrp proteins are main components of the pilus, acting as translocators and harpins (Jeong *et al.*, 2009).

Once T3SEs are delivered into the host cell they can interfere with PAMP-triggered immunity (PTI) and effector-triggered immunity (ETI) via different strategies: e.g. blocking RNA pathways, interference with vesicle trafficking or organelle function alteration (Block & Alfano, 2011; Deslandes & Rivas, 2012). This results in suppression of plant immunity and consequently allows the pathogen to proliferate in the intercellular spaces and elicit HR response in the host (Guo *et al.*, 2009). The T3SS is prominent in many phytopathogens which colonize the apoplast, including *Xanthomonas, Ralstonia, Erwinia* and *Pseudomonas* and is central to their ability to cause disease (Alfano & Collmer, 2004). The T3SS is best studied

in *P. syringae*, in which the plant-host interactions have been determined at a very fine molecular level:

1.4.2.1 Functional types of T3SS

The *P. syringae* pathogenicity island (PAI) refers to the whole genomic region, which is composed of the *hrp/hrc* gene cluster bordered by the conserved effector locus (CEL) on one side and the exchangeable effector locus (EEL) flanking the opposite side (Alfano *et al.*, 2000). The CEL typically encodes the effector genes *hopM1, avrE1, hopAA* and *hrpW*, with the former two effectors interfering with the plant immune system by suppressing cell-wall based defence (Debroy *et al.*, 2004) and *hrpW* encoding for a harpin targeting the plant cell wall (Charkowski *et al.*, 1998). Mutants of *P. syringae* pv. *tomato* DC3000 ΔCEL with a deficient CEL locus failed to cause disease in tomato and showed drastically reduced growth *in planta* (Badel *et al.*, 2006; Collmer *et al.*, 2000; Debroy *et al.*, 2004). The EEL encodes for a variable repertoire of effectors, which varies among strains and determines the extent of parasitic fitness of the individual strain (Alfano *et al.*, 2000). Additional effectors are encoded in various other parts of the genome. The T3SS structure as described above is generally referred to as the canonical or tripartite pathogenicity island (T-PAI), which is the most common in pathogenic *P. syringae*. Another type of T3SS, the atypical PAI (A-PAI), was identified in non-pathogenic *P. syringae* isolates, which are common leaf colonizers. These strains contain a *hrp/hrc* cluster, which is distantly related to the canonical type, but is

missing *hrpK* and *hrpS*, missing the flanking regions and in a different genomic location (Clarke *et al.*, 2010; Mohr *et al.*, 2008). Additionally, some strains may possess an additional T3SS that appears to resemble the one found in *Rhizobium* (R-PAI), although this system is not a prerequisite to cause disease (Martínez-García *et al.*, 2015). As each *P. syringae* contains a certain type of type III secretion system, Baltrus *et al.* (2017) suggested identification of T3SS type as a basis to start defining ecotypes.

1.4.2.2 Effector repertoires and host range

The pangenome of the *P. syringae* complex encompasses 58 families of effectors with an average of 15-35 effectors delivered into the host cell per strain (Baltrus *et al.*, 2011), but the number of effectors may be as low as 5 (Hockett *et al.*, 2014). There is a significant amount of variation among the amount of effectors secreted by each isolate. For example, isolates from Phylogroup 2 generally possess fewer T3SEs than Phylogroup 1 and 3 (Baltrus *et al.*, 2011; Block & Alfano, 2011; Nowell *et al.*, 2016). When looking at a closer level, even within a pathovar, there is a lot of variability, e.g. *P. syringae* pv. *actinidiae* has been distinguished into four different clades and there is considerable variation in virulence genes among those four clades (Fujikawa & Sawada, 2016). Another example is *P. syringae* pv. *tomato*, as the two isolates *Pto* DC3000 and *Pto* T1 share only 14 effectors, with another 15 unique to DC3000 and 11 unique to T1, despite both being the causative agent of bacterial speck of tomato (Almeida *et al.*, 2009). The T3S effector repertoire can therefore not be used to determine host range,

and in fact this leads to the assumption that *P. syringae* are equipped as generalists and the diversification of the repertoire is driven by the individual interaction with the immune system of the host it encounters. Additionally, horizontal gene transfer plays a major role in the distribution of the T3SE and often effectors are associated with mobile genetic elements (Lindeberg *et al.*, 2012).

1.4.2.3 Phytotoxin production

Apart from the T3SS virulence machinery, *Pseudomonas syringae* produces a variety of non-host specific phytotoxic compounds, which play a role in the pathogen-host interaction by directly influencing disease development and severity (Bender *et al.*, 1999). Typically genes encoding toxin synthesis and secretion are encoded on the chromosome, although some are often found on a plasmid, like coronatine (Alarcón-Chaidez *et al.*, 1999), production of ethylene (Nagahama *et al.*, 1994) and auxin (Glickmann *et al.*, 1998). Ethylene is a plant hormone, which has a significant influence on physiological processes during plant growth and development. The production of ethylene by pathogens affects the strength of plant defence reactions by creating a hormonal imbalance, which in turn has an impact on the severity of symptoms caused by plant pathogens (Weingart *et al.*, 2001). Auxin (indole acetic acid, IAA) is another phytohormone that is produced by pathogenic bacteria. Synthesis occurs via conversion of tryptophan to IAA by the enzymes iaaM (tryptophan monooxygenase) and iaaH (indoleacetamide hydrolase) (Rodríguez-Palenzuela *et al.*, 2010).

Some pathogens possess the *iaaL* gene, which inactivates IAA by converting it to IAA-lysine (Romano *et al.*, 1991). Production of IAA has been linked with various biological roles, including shortening the incubation period of disease, increase of disease severity in *P. savastanoi* pv. *savastanoi* or inhibition of plant defence mechanisms (Gardan *et al.*, 1992; Glickmann *et al.*, 1998). Syringolide is produced via the action of the avirulence gene D (*avrD*) (Keith *et al.*, 1997), which particularly elicits a hypersensitive response in soybean plants carrying the *Rpg4* disease resistance gene (Keen & Buzzell, 1991). The toxin coronatine has a variety of effects on plants. It simulates the plant hormone jasmonic acid, suppressing stomatal plant defence, resulting in stomata opening and clearing the path for *P. syringae* to invade the apoplast (Melotto *et al.*, 2006). It can be required for endophytic bacterial proliferation (Elizabeth & Bender, 2007) and can inhibit the accumulation of salicylic acid (SA), which mediates plant defence responses (Zheng *et al.*, 2012). The virulence factor syringolin A contributes to stomata opening by proteasome inhibition, hence counteracting the innate plant immune defence (Schellenberg *et al.*, 2010). Biosynthesis of phaseolotoxin inhibits ornithine carbamoyl transferase (OCTase), an enzyme of the urea cycle that converts ornithine to citrulline, which results in arginine deficiency (Bender *et al.*, 1999). This conversion is a critical step, hence phaseolotoxin-producing strains possess two OCTases (sensitive OCTase [*argF*] and resistant OCTase [*argK*]), where the resistant version is expressed only in conditions favourable for phaseolotoxin production (Peet & Panopoulos, 1987; Staskawicz & Hoogenraad, 1980). Tabtoxin is the precursor of the monocyclic β-lactam antibiotic, which is

hydrolysed by a peptidase to reach the active state tabtoxinine-β-lactam, which then inhibits glutamine synthetase (Kinscherf *et al.*, 1991; Kinscherf & Willis, 2005). The phytotoxic activity of the cyclic lipopeptide toxins syringopeptin and syringomycin is very similar. Production of either toxin results in plasma membrane pore formation, followed by subsequent cell lysis and leakage of nutrients (Bender *et al.*, 1999; Raaijmakers *et al.*, 2006). Phytotoxins not only affect the plant, but can inhibit the growth of other microorganisms by having an antibiotic effect.

1.4.3 Commensal *P. syringae*

P. syringae is well known for its diversity, being able to cause disease on a broad range of hosts. Given the economic importance of some of its hosts, it comes as no surprise that research initially had focussed on isolates collected from symptomatic plants. In recent years, the environment as a reservoir for a diverse range of *P. syringae* has been investigated, with a focus on both terrestrial (leaf litter, soil) and aquatic (river, snow, precipitation) habitats (Monteil *et al.*, 2012, 2014; Morris *et al.*, 2010, 2007). In an attempt to describe the genetic diversity of *P. syringae* isolated from diseased hosts and environmental sources, four out of 13 phylogroups were purely associated with environmental isolates (Berge *et al.*, 2014), which strongly suggests that only a fraction of the real diversity of *P. syringae* has been recovered and a wide range of diverse isolates is still waiting to be collected.

Not all isolates of *P. syringae* elicit a plant immune response and *P. syringae* can be found in the phyllosphere as a non-pathogenic commensal

(Hirano & Upper, 2000). The most famous example of a non-pathogenic isolate is perhaps *P. syringae* cit7, which was isolated from a leaf of a healthy orange tree and later released as an engineered mutant lacking the Ice nucleation gene for biological control (Lindow *et al.*, 1987). Other non-pathogenic strains used for biological control were *P. syringae* TLP2 (Byrne *et al.*, 2005) to control bacterial speck of tomato (*P. syringae* pv. *tomato*) and *P. syringae* 508 (Burr *et al.*, 1996), which showed strong antifungal activity against apple scab fungus (*Venturia inaequalis*), which were isolated from healthy potato and apple leaves respectively.

Investigations into the inability to incite disease of these non-pathogenic *P. syringae* strains isolated from wild plants revealed that these isolates were simply lacking the canonical Type III secretion system (absence of *hrp/hrc* locus) (Mohr *et al.*, 2008). Though, this was soon rectified by the discovery of an atypical Type III secretion system with a novel *hrp/hrc* locus and putative effectors for these strains, which formed a unique subclade in Phylogroup 2 (2c) based on MLST. These strains did not elicit an HR in any of the tested plants (Clarke *et al.*, 2010) and did not exhibit any significant differences in fitness on *in planta* growth (Kniskern *et al.*, 2011). This atypical T3SS was later also revealed among a variety of environmental isolates originating from epilithic biofilms, snowpack, river and lakes (Demba Diallo *et al.*, 2012).

1.5 Bacterial interactions in the phyllosphere

Interactions between microbes is a substantial field of research spanning from experimental evolution studies (e.g. Barrick & Lenski, 2013; Elena & Lenski, 1997; Griffin *et al.*, 2004; Rainey & Rainey, 2003; Rainey & Travisano, 1998) to field experiments (e.g. Hansen *et al.*, 2007; Huang *et al.*, 2011; Pastar *et al.*, 2013; Tong *et al.*, 2007). Although interactions between microbes, particularly those leading to the maintenance of diversity in itself, is a comprehensive field (Fierer & Lennon, 2011; Hibbing *et al.*, 2010; Kassen & Rainey, 2004; Rainey *et al.*, 2000), bacterial interactions discussed below are limited to those relevant to the phyllosphere, to align with the aim of this thesis.

The phyllosphere is an environment where multipartite interactions occur: interactions between microorganisms and the plant host, among pathogenic bacteria, between pathogenic and commensal microbes and between other microorganisms and bacteria (Vorholt, 2012). The interactions between plant and bacteria are manifold and have been described in section 1.1.2. Microbial interactions can have wide-reaching consequences, from affecting the local tissue to the entire host (Stubbendieck *et al.*, 2016) and influencing the community structure. Interplay among bacteria occurs at the level of individuals, populations and communities and represents an essential driving force to evolution and function. Understanding how, where and between which microorganisms these interactions take place and how environmental factors shape

interactions will support our understanding of what drives the divergence, emergence or extinction of certain bacterial lineages (Hibbing *et al.*, 2010).

A bacterial 'community' constitutes different populations of a number of species, with bacterial 'populations' defined as a subset of bacteria of the same species. The community structure is highly dynamic and evolves over time, and is influenced by the number and types of interactions amongst its members and the influence of abiotic and biotic factors. Given that a suitable microbial habitat is typically densely populated (e.g. 10^6 - 10^7 per cm^2 leaf surface, Hirano & Upper, 2000), the level of interactions must be quite high and associated with a high degree of variability.

Such interactions range from cooperative to antagonistic, whereas neutral interactions are not to the benefit or detriment of either party. Cooperative interactions occur when at least one gains a profit, whereas the other is not affected (commensalism), or both benefit from the interaction (mutualism). Amensalism, on the other hand, involves one member being negatively affected from the interaction, which is a form of competition. Competition among bacteria, as in every living being, is probably the most common type of interaction. It is broadly distinguished between exploitative and interfering competition (Birch, 1957), with nutritional resources typically being a major limitation in a natural environment. Exploitation refers to competition for resources by depleting the environment for e.g. nutrients, water, light, or simply two-dimensional in space. Interfering competition involves the secretion of toxins, enzymes or antibiotics that kill or inhibit growth of the competitor. The ecological

similarity (e.g. identical metabolic needs) of competing species can increase the level of the competition. The leaf surface is home to a comprehensive and diverse number of bacteria occupying a small space. As previously mentioned, the structure and composition of this bacterial community is shaped by a variety of environmental factors like nutrient and water availability, UV radiation, or pollution (Hirano & Upper, 2000; Lindow & Brandl, 2003; Vorholt, 2012). Given the physical and physiological surface properties of the leaf, only restricted areas are available for colonization, hence pathogenic and non-pathogenic microbes must be in close contact with each other. The spatial proximity facilitates HGT events, which is a major influence on genotypic variation (described in section 1.3). In terms of infectious diseases, synergistic interactions have been shown, where a pathogen does not necessarily act independently, but virulence was affected by interaction with co-inhabiting strains (Singer, 2010). For example, this has been demonstrated for olive knot disease, where *P. savastanoi* pv. *savastanoi* is the causative agent, but *Erwinia* sp. and *Pantoea* sp. were discovered as endophytes co-colonizing the cankers and impacting the severity of disease (Marchi *et al.*, 2006; Moretti *et al.*, 2011). Related pathogen species can co-exist on an individual plant host (Fitt *et al.*, 2006). Another interesting example of synergistic interactions displays when bacterial strains lacking virulence factors benefit from the coexisting pathogenic isolate and reap their benefits, e.g. population sizes of *hrp* mutants were smaller when inoculated alone, as compared to co-inoculation with the wildtype *P. syringae* pv. *phaseolicola* B728a strain (Hirano *et al.*, 1999).

Clearly co-residing strains are involved in complex interactions with plant pathogens, but the outcome of these interactions remains to be elucidated. In fact, in diseased hosts the diversity of non-pathogenic isolates has been neglected in the past. There is a need for studies in which a properly designed sample scheme allows discovery of the diversity of the natural populations as a prelude to the investigation of potential interactions. But first, it is necessary to define the level at which interactions are studied. It is important to ask, what constitutes a bacterial population? Do we consider only members of the same species that engage with each other? Within that population, what roles do commensals play, and what is the outcome of competition with the pathogenic strain? What is the evolutionary potential of these commensal bacteria and what ecological factors drive interactions between pathogenic and non-pathogenic isolates?

1.6 Bacterial population genetics

1.6.1 Definition of a bacterial population

A bacterial population can be defined as a collection of cells, and collections of populations make up communities (Rainey, 2005). Hence, a set of closely related individuals (i.e. members of the same species) in a localised area can be referred to as a population. When studying the population structure of a given bacterial species, the focus should ideally lay on a highly localised population, like e.g. *Staphylococcus aureus* in cystic fibrosis

patients or *Vibrio cholerae* in freshwater reservoirs, however this is rarely achieved. Many studies involve collections of spatially widely distributed bacterial isolates, but population structures should be inferred from biologically meaningful datasets (Istock *et al.*, 1992), like it was implemented in e.g. Haubold & Rainey (1996).

Defining a population gets a little difficult when the species boundaries are as unclear as for *P. syringae* (see section 1.4). It is necessary to delineate evolutionarily relevant clusters, meaning an entity in which genotypes at least have the ability to exchange genetic material. There has been much debate as to what actually can be considered as a *P. syringae* population (Baltrus *et al.*, 2017). For this study, the investigated *P. syringae* population refers to all bacterial isolates for which the identity as *P. syringae* was revealed based on four housekeeping genes and which were isolated from the leaf surface of kiwifruit vines.

1.6.2 Population structure

Population genetics is the study of patterns of genetic variation within populations, which can be explained in terms of the evolutionary processes involved (mutation, recombination, migration, selection and genetic drift). Population biologists describe the genetic variation within natural populations, while exploring to what extent each of the microevolutionary processes has contributed. Genetic variations can be introduced via mutation, recombination and migration, with the fate of variants being determined by genetic drift and selection.

Although bacteria reproduce asexually, recombination is a major source of genetic variation. Theoretically bacterial populations consist of essentially clonal lineages, with the only genetic variation arising via *de novo* mutations, which are passed on to progeny. This strictly clonal population structure described above is only hypothetical, as bacteria do engage in recombination (Horizontal Gene Transfer), which was referred to as localized sex by Maynard Smith *et al.* (1991). HGT is the receipt of genetic content from other organisms (closely or distantly related), resulting in either replacement of existing genes (homologous recombination) or introduction of novel genetic content (Mazodier & Davies, 1991). Recombination is a major source of genetic variation in bacteria (Levin & Cornejo, 2009), although the frequency of recombination varies depending on the population. The introduction of genetic variation via recombination results in a disruption of these clonal bacterial structures and this novel gene content can spread throughout a lineage independent of its origin.

So what does the population structure look like for bacterial populations? In fact they exhibit complex genetic structures ranging from strictly clonal (no recombination) to essentially panmictic populations (freely recombining and complete randomization of alleles) (Maynard Smith *et al.*, 1993; Spratt & Maiden, 1999). It is quite rare for bacterial populations to exhibit either extreme, but for example the population structure of *Neisseria gonorrhoeae* was described as panmictic with very high levels of recombination, which was found in early studies based on multi locus enzyme electrophoresis (MLEE) studies and was later confirmed based on whole genome data (Ezewudo *et al.*, 2015; Maynard Smith *et al.*,

1993; O'Rourke & Stevens, 1993). On the contrary, a clonal population structure was shown for many pathogenic bacteria like *Escherichia coli* (Ochman & Selander, 1984; Selander & Levin, 1980; Tenaillon *et al.*, 2010) or *Salmonella enterica* populations (Nelson & Selander, 1994; Selander *et al.*, 1990). Clonality does not suggest that recombination never occurs, more likely it is not frequent enough to break the pattern of clonality (Tibayrenc & Ayala, 2012), or it depends on the level and the intensity (short sequences versus whole genome data) at which population structure is investigated, as recombination was in fact detected for e.g. *Salmonella enterica*, but at the subspecies level (Didelot *et al.*, 2011; Lan *et al.*, 2009).

The population structure of many pathogenic bacteria has been described so far and the characterization of isolates has been used to answer epidemiological questions like where did novel types emerge from and how do they relate to isolates from other geographical regions (Spratt & Maiden, 1999). The genetic structure of a pathogen allows us to assess its evolutionary potential and in turn make informed decisions about the associated risk of disease emergence.

The diversity and genetic structure of pathogenic strains and isolates from the environment have been extensively investigated for *P. syringae* in various MLST studies (e.g. Bartoli *et al.*, 2015; Berge *et al.*, 2014; Bull *et al.*, 2011; Hall *et al.*, 2016; Hwang *et al.*, 2005; Sarkar & Guttman, 2004). The major conclusion for isolates from diseased hosts was that homologous recombination among *P. syringae* is limited, resulting in an overall clonal population structure (Sarkar & Guttman, 2004). This finding was

supported by another study, which found well-defined phylogenetic clades for *P. syringae* (Bull *et al.*, 2011), however all of these studies are purely based on a small number of housekeeping genes, not taking into account genomic data, which potentially could reveal a different picture, as has been described for closely related lineages in pv. *pisi* (Baltrus *et al.*, 2014). Nonetheless, although the cost of whole genome sequencing decreases constantly, MLST studies remain a simple and cost-effective way of determining population structures, especially in regard to comparison with existing databases such as the Plant Associated and Environmental Microbes database (http://www.pamdb.org). Much is known about the genetic structure of pathogenic strains, but there is a lack of knowledge about the population structure and the diverse ecological and evolutionary forces acting on commensal *P. syringae* populations. Selective pressures on commensal strains could drive host range extension. Understanding the distribution of commensal strains, whether random or host specific, will lead to identification of potential resident strains, which would have an even higher likelihood of contributing to the evolution of virulent types and antibiotic resistance.

1.7 Research objectives

This thesis investigates the eco-evolutionary dynamics of *P. syringae* inhabiting the kiwifruit phyllosphere. I am combining three different approaches to gain knowledge about the diversity and evolutionary potential of non-pathogenic *P. syringae* in the presence of a pathogen in an agricultural environment.

Chapter 2 provides insights into the diversity and microevolutionary parameters shaping the genetic structure of *Pseudomonas syringae* residing in the phyllosphere of an agriculturally important crop (kiwifruit, *Actinidia* sp.). This led to the discovery of a novel clade (PG3a) of *P. syringae*, which was recovered from leaves in three different kiwifruit growing regions around the world.

In Chapter 3 I describe the ecological perspective of interactions occurring between a plant pathogen (*Psa*) and a non-pathogenic isolate from the new kiwifruit-associated clade. Mixed inoculations revealed that both bacteria are able to co-exist in a stable manner, however there is a reciprocal effect of the presence of one strain on the other. Given the antagonistic activity of the commensal strain on the pathogen *Psa* from the competition experiments, there is a potential development of these commensal strains as a biocontrol agent.

Taking a comparative genomics approach (Chapter 4), I investigate the genomic fluctuations of *P. syringae* Phylogroup 3 and novel clade 3a,

clarifying phylogenetic resolution at the core genome level, looking into the repertoire of virulence genes (effectors, toxins) to compare the diversity within the new clade of kiwifruit resident strains with the kiwifruit pathogen *Psa*, revealing regions which have recently undergone recombination, and looking at other regions of interest, such as copper resistance genes or genes involved in host range expansion.

Chapter 2

Population genetics of *P. syringae* on kiwifruit leaves

2.1 Introduction

Measures of population structure are key to understanding pathogen evolution and the origin and spread of bacterial diseases (Levin *et al.*, 1999; Maynard Smith *et al.*, 2000; Spratt & Maiden, 1999). The genetic structure of a population reflects its evolutionary history. This provides important information about a pathogen's evolutionary potential, which is indispensable for making informed decisions about disease management. Genetic variation plays a critical role in the evolution of pathogenicity, as populations with large genetic variation increase their potential to adapt to changing environments or evolve resistance mechanisms.

2.1.1 Recombination and its effect on population structure

Genetic variation arises through mutation and recombination, with additional contributions from migration. According to neutral theory, mutations accumulate in a more or less clock-like manner, but recombination in bacteria is neither obligatory, nor absent, but varies considerably among species. For this reason, understanding the extent of recombination in bacterial populations has received considerable attention (Didelot *et al.*, 2011; Feil & Spratt, 2001; Holmes *et al.*, 1999; Levin & Cornejo, 2009; Maynard Smith *et al.*, 1993; Ochman *et al.*, 2000; Polz *et al.*, 2013; Yahara *et al.*, 2016). Virulence genes, antibiotic/heavy metal resistance, or factors involved in host range expansion are features that can be acquired via recombination from the same, closely related, or completely different bacterial species, eukaryotes and Archaea (Juhas *et al.*, 2009; Marraffini & Sontheimer, 2008; Nelson *et al.*, 1999). Transfer of DNA between bacterial cells can occur via three parasexual processes (transduction, transformation, conjugation). Incoming foreign DNA is typically recognised and eliminated by various defence systems including restriction and modification systems (Jeltsch, 2003; Thomas & Nielsen, 2005) and CRISPR (Barrangou *et al.*, 2007; Horvath & Barrangou, 2010). However, as these systems do not operate with perfect fidelity, or are maybe otherwise overcome by attributes of the foreign DNA, foreign DNA may become incorporated into the host genome. This occurs via either homologous, or non-homologous recombination. Homologous recombination

involves replacement of existing parts of the genome with new genetic content of high nucleotide identity, resulting in either identical or new alleles (Didelot & Maiden, 2010). Non-homologous recombination (horizontal gene transfer) involves new fragments of DNA from a donor organism being inserted into the genome of a recipient, typically the accessory genome. In contrast, homologous recombination affects both the accessory and core genome, which consists of genes essential to the survival of the cell. Both types of recombination are accountable for maintaining and shaping genetic diversity in bacterial populations, as well as being an important driving force in the evolution of bacteria.

The frequency of genetic exchange is unknown and can likely occur at different rates. Low/absent levels of recombination essentially result in a population being designated 'clonal', whereas high levels of recombination among individuals generate panmixia and the population is referred to as 'recombining' (Spratt & Maiden, 1999). Moderate levels of recombination introduce genetic variation to avoid the emergence of clonal lineages. For a population to be strictly clonal, genetic isolation at an extreme level would be required. In a clonal population, genetic variation is created by indels, transposon insertions, inversions and *de novo* point mutations, which becomes apparent from linkage disequilibrium (although non-random association of alleles can be observed, even if individuals are recombining) and congruence among phylogenetic trees (Didelot & Maiden, 2010). These extreme scenarios of population structures are however unlikely, with most bacterial populations lying somewhere in between. In 1973 Roger Milkman first claimed that natural *E. coli* bacterial populations were freely

recombining (Milkman, 1973), but this was questioned by Selander & Levin (1980) who isolated identical genotypes from different host organisms. Until the early 1990s the prevailing view was that *E. coli* populations were clonal (Lenski, 1993; Maynard Smith, 1991; Ochman & Selander, 1984). In the seminal work of John Maynard Smith from 1993, a variety of population structures were detected for bacterial populations ranging from strictly clonal (*Salmonella*) to panmictic (*Neisseria gonorrhoea*) (Maynard Smith *et al.*, 1993). This led these authors to question universal clonality and they drew attention to the issue of temporary clonality, where the outbreak and dissemination of a pathogenic isolate can result in the appearance of a clonal population structure, especially when sampling is biased on diseased hosts. For example, even in bacteria where a clonal population structure has long been assumed, such as for *Salmonella enterica*, clonality is in reality rather rare in a natural environment (Didelot *et al.*, 2011; Lan *et al.*, 2009).

For *P. syringae*, recombination rates appeared to occur at a low level, suggesting a clonal population structure (Sarkar & Guttman, 2004), however this resulted from studying isolates from diseased hosts, which ignores a lot of the diversity found in other environments. The narrow focus on disease-causing isolates meant that researchers had only captured a small snapshot of the diversity, implying that we should employ a much broader sampling scheme than has been done so far to unravel genetic structure and capture the full evolutionary potential. Recent *P. syringae* research has focussed on discovering the diversity of strains found in a non-agricultural setting and bacteria were isolated from various

environmental sources, including river water, snow, irrigation water, leaf litter, ice and even phloem-feeding insects (Monteil *et al.*, 2012, 2013, 2014; Morris *et al.*, 2008, 2010; Stavrinides *et al.*, 2009). These strains convey a large and diverse genetic pool and could potentially contribute to the emergence of a new virulent epidemic strain. In fact, Monteil *et al.* (2013) found signals of environmental isolates recombining with the crop pathogen *P. syringae* pv. *tomato* and suggested that eventually an epidemic clone had emerged from a freely recombining pool of *P. syringae* by having adapted to a specific host through a number of evolutionary events.

However, little work has focussed on diversity of non-pathogenic *P. syringae* strains that reside on the same host, either healthy or diseased. The limited space and crowded conditions in a densely populated environment like the leaf surface, suggests that these commensal bacteria might have an even higher probability of exchanging genetic material with pathogens, as they are in geographical proximity (Didelot *et al.*, 2012). The occupation of the same ecological niche has important implications for the evolution of bacteria, where recombination can result in exchange of genetic factors that might confer a selective benefit in a new habitat (Fraser *et al.*, 2009; Toft & Andersson, 2010). This has been observed in different lineages that have adapted to the same ecological niche (Didelot *et al.*, 2007; Luo *et al.*, 2011; Sheppard *et al.*, 2008).

The current outbreak of *Psa* on kiwifruit presents an opportunity to study commensals in an environment where they either encounter a highly virulent pathogen, or reside on a healthy host. I employed Multi Locus

Sequence Typing to infer the micro evolutionary parameters affecting structuring of diversity in the focal population.

2.1.2 Multi Locus Sequence Typing (MLST)

Knowledge of evolutionary forces is central to understanding, and making sense of, the ecological, evolutionary and population dynamics of pathogens (Spratt & Maiden, 1999). Multi Locus Sequence Typing (MLST) involves the unified amplification of short sequences (400-500 bp) of typically up to seven core (housekeeping) genes, which allows portability around the world. It is a commonly used and easily reproducible method to characterize and identify bacteria species and was introduced by Maiden *et al.* (1998). Typically the data are deposited into open-access databases, e.g. http://pubmlst.org, University of Oxford, UK or http://www.pamdb.org, Vinatzer Lab, Virgina Polytechnic Institute, US, making it available for global studies.

Two approaches are used when analysing MLST data: allele- and nucleotide based schemes. For the allele-based method, each unique nucleotide sequence is assigned an allele number and the unique allelic profile of typically 7 genes is then assigned a sequence type (ST). The weakness of this approach is the fact that much of the evolutionary information is actually disregarded. For example, relatedness information gets lost in the translation of the nucleotide sequences into STs, as a strain showing one single nucleotide polymorphism (SNP) is being treated in the same way as one that shows many SNPs, by simply assigning a different

allele (Didelot & Falush, 2007). The second strategy, the nucleotide based approach, makes use of this information and is based on the actual nucleotide sequences.

Applications for MLST are wide-ranging, from determining strain-relatedness (clonal expansion and diversification), to phylogenetic inference, population dynamics and evolutionary history and the amount of genetic diversity; it can also be used for unravelling patterns of historic migration or inference of divergence time (Pérez-Losada *et al.*, 2013). In an era when bacterial genomes are being sequenced at increasing pace, MLST remains a simple, effective and, importantly, cost-efficient method of inferring important population genetics parameters. MLST is still applicable to genome sequencing data and the existing comprehensive databases will keep its worth at least for now. In fact, sequencing data improve the practicality of MLST, as genomes of the target species provide us with a higher resolution and with a reference genome to design markers if the need arises (Pérez-Losada *et al.*, 2013).

2.1.3 The population structure of *P. syringae*

P. syringae encompasses a genetically diverse group of bacteria. From initial efforts of classification based on phenotypic traits (Lelliott *et al.*, 1966; Young, 2010), comprehensive efforts to understand the genetic structure of *P. syringae* were done using DNA-DNA hybridization studies, where 9 genomospecies were identified (Gardan *et al.*, 1999; Young, 2010). Since then the population structure of *P. syringae* has been explored using

MLST based on seven housekeeping genes (*acn*, *cts* (=*gltA*), *gapA*, *gyrB*, *prfk*, *pgi* and *rpoD*) as first proposed by Sarkar & Guttman (2004). Later this was reduced to *gapA, gyrB, gltA* and *rpoD* as these provide nearly as good phylogenetic resolution compared to the set of seven genes (Hwang *et al.*, 2005). The MLST studies mentioned previously focussed on isolates from diseased hosts and found a mostly clonal population structure for *P. syringae*, with little evidence of recombination among housekeeping genes (Sarkar & Guttman, 2004; Hwang *et al.*, 2005). Further studies investigating phylogenetic relationships among different pathovars of *P. syringae* showed well supported phylogenetic lineages, which provided further indication that horizontal gene transfer is limited (Bull *et al.*, 2011; Clarke *et al.*, 2010), but not absent (McCann *et al.*, 2013, 2017). The MLST scheme has been used extensively to classify *P. syringae* into major monophyletic groups. Recently, a comprehensive study investigating the phylogenetic grouping of a large sample of pathogenic and environmental *P. syringae* strains identified 13 phylogroups (PGs) representing 24 clades (Berge *et al.*, 2014). The sample included isolates from snow, plants, leaf litter, precipitation and epilithic biofilms, as well as from plant pathogenic scenarios, and showed that environmental isolates were often found in the same genetic lineage as *P. syringae* isolated from diseased crop plants (Berge *et al.*, 2014). The phylogroups relevant to this study include PG1, PG2, PG3 and PG5. PG1 contains mainly pathogens of members of the family *Brassicaceae* (e.g. cauliflower, cabbage, radish), in addition to pathogens of tomato (pv. *tomato*), tea (pv. *theae*) and kiwifruit (pv. *actinidiae*). PG2 is the most ubiquitous group and encompasses a diverse

range including tomato, bean, lilac and pea pathogens; many pathogens of woody plants (mulberry, olive, horse chestnut) group with PG3, which is also associated with bean (pv.*phaseolicola*), tobacco (pv.*tabaci*) and cucumber (pv.*lachrymans*) pathogens. PG5 isolates are pathogenic on *Brassicaceae* and coriander (pv.*coriandricola*) (Berge *et al.*, 2014; Hwang *et al.*, 2005; Sarkar & Guttman, 2004).

2.2 Aims

For this study *P. syringae* isolates were collected from infected and uninfected orchards in the greater Auckland region. Using Multi Locus Sequence Typing (MLST) the aim was to:

1. Unravel the diversity of *P. syringae* associated with the leaf surface of kiwifruit vines during the current outbreak of *P. syringae* pv. *actinidiae* (*Psa*).

2. Estimate the rates of recombination and whether recombination affects phylogenetic reconstruction.

3. Define any subpopulations existing within *P. syringae* based on ancestry of polymorphisms (sequence type analysis) or nucleotides (phylogenies).

4. Identify whether biogeographical factors like cultivar, infection status or time of isolation influence genotypic diversity.

5. Place *P. syringae* isolated from NZ-grown kiwifruit vines in a global context by making use of the MLST sequences deposited in the PAMDB database and other kiwifruit *P. syringae* isolates available from NCBI.

2.3 Material & Methods

2.3.1 Isolation of putative *P. syringae* isolates from kiwifruit leaf washes

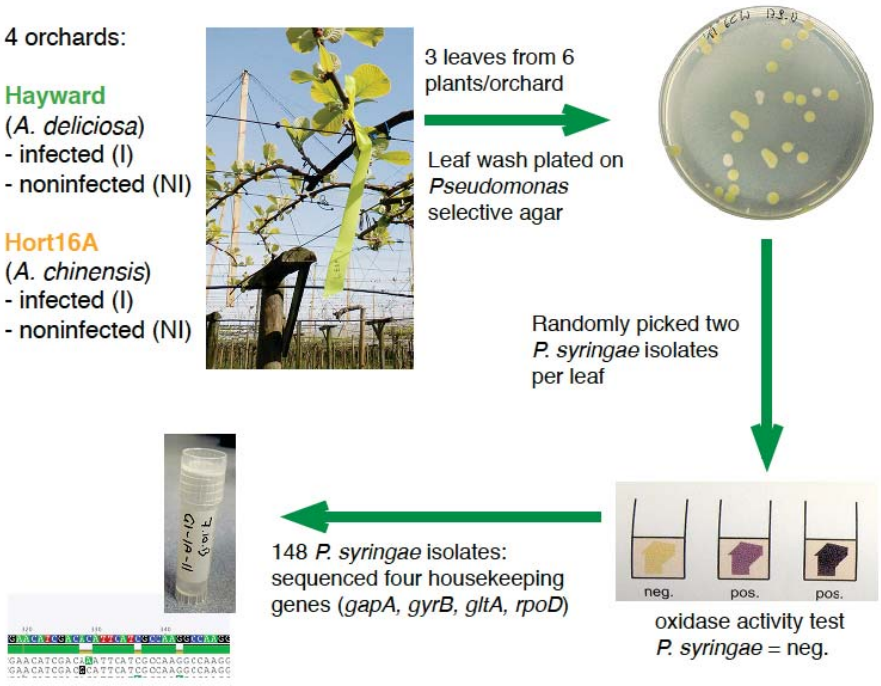

Figure 2.1: Overview of the sampling strategy.

Leaf samples were collected from four separate kiwifruit orchards: two orchards cultivating green *A. chinensis* var. *deliciosa* 'Hayward' and two cultivating gold *A. chinensis* var. *chinensis* 'Hort16A' (sampling procedure described in Figure 2.1). One orchard of each cultivar was already infected with *P. syringae* pv. *actinidiae*, the other two remained uninfected for the duration of the study. Sampling occurred at three intervals during the

growing season: spring (after bud break), summer and autumn (prior to harvest). Three leaves were taken per plant from a total of six kiwifruit vines per orchard, sampling across the orchard in a diagonal transect (approx. 400 m). Some *A. chinensis* var. *chinensis* canes were removed during routine disease management; neighbouring canes on the same vine were then sampled and tagged. All uninfected *A. chinensis* var. *deliciosa* vines were cut down prior to the last sampling day, so the adjoining block of *A. chinensis* var. *deliciosa* was sampled instead. The location of each sampled orchard is listed in Table 2.1. Sampling days were chosen according to weather conditions, so that physical circumstances remained comparable throughout the study.

Table 2.1: Sampling locations. Specification of cultivar and infection status according to KVH (Kiwifruit Vine Health), orchard ID, GPS coordinates, location and month of sampling.

Orchard	**ID**	**GPS coordinates**	**Location**	**Sampling time**
Hort16A uninfected	K	36 46' 39.5" S 174 34' 42.9" E	Kumeu	September 2013 January 2014 May 2014
Hort16A infected	G	37 10' 19.0" S 174 43' 25.9" E	Glenbrook	October 2013 February 2014 May 2014
Hayward uninfected	R	36 47' 25.4" S 174 33' 21.6" E	Kumeu	October 2013 February 2014 May 2014
Hayward infected	H	37 10' 23.4" S 174 43' 08.1" E	Glenbrook	October 2013 February 2014 May 2014

Leaves were placed in sterile 50 mL conical centrifuge tubes (Corning CentriStar, US) and 40 mL of 10 mM $MgSO_4$ buffer with 0.2% Tween (Invitrogen, US) was added. Tubes were alternatively shaken and vortexed at slow speed for 30 s for a total of three minutes. Leaves were removed from tubes with sterile tweezers and photographed to record the leaf area and the presence of any symptoms of disease. Tubes of leaf wash were centrifuged at 4600 rpm for 10 min, followed by removal of the supernatant. The pellet was resuspended in 200 μL 10 mM $MgSO_4$ buffer and dilutions of leaf washes were plated on *Pseudomonas* agar base (Oxoid, UK) supplemented with 10 mg/L cetrimide, 10 mg/L fucidin and 50 mg/L cephalosporin (CFC supplement, Oxoid, UK).

2.3.2 Selection of colonies for genotypic testing

After 48 h incubation at 28 °C, two isolates resembling *P. syringae* (round, creamy white) were randomly picked per leaf wash and restreaked on KB agar plates to confirm colony morphology. The absence of cytochrome C oxidase was verified by using Bactident Oxidase strips (Merck KgaA, Germany). Strains lacking cytochrome C oxidase were grown overnight in liquid KB and then stored in 30% glycerol at −80 °C.

2.3.2.1 DNA preparation for genotyping

Each isolate was streaked on KB agar plates from freezer stock and grown for two days. Using a sterile toothpick one single colony was transferred and

suspended in 100 μL $d\mathrm{H_20}$. Subsequently the mix was lysed at 95 °C for 10 min and stored at −20 °C for use in all PCR reactions.

2.3.3 MLST genotyping

Strains were sequenced using the Hwang *et al.* (2005) MLST scheme for four housekeeping genes, *gapA, gyrB, gltA* and *rpoD*. The *gapA* gene encodes glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and is hence vital for glycolysis; *gyrB* encodes subunit B of DNA gyrase, an enzyme introducing negative supercoils in double-stranded closed-circular DNA in bacteria (Reece & Maxwell, 1991); *gltA* encodes citrate synthase, an essential enzyme in the first step of the Krebs cycle (citric acid cycle) and *rpoD* encodes RNA polymerase sigma factor 70, which is essential for initiating contact between RNA polymerase and a gene promoter. Due to amplification problems, the forward primer for *rpoD* from Sarkar & Guttman (2004) was used.

2.3.4 Polymerase Chain Reaction (PCR)

2.3.4.1 PCR primers

For this study the primers designed by Hwang *et al.* (2005) were used for *gapA, gyrB, gltA* and *rpoD* (reverse), as well as the forward primer for *rpoD* from Sarkar & Guttman (2004). The list of primer sequences can be found in Table 2.2.

2.3.4.2 PCR amplification

PCR amplification was performed with a BIO-RAD T100 Thermal Cycler following an adapted protocol of Hwang *et al.* (2005) using a total reaction volume of 50 μL with a final concentration of 1x PCR buffer (Invitrogen, US), 1 μM for each primer, 0.2 mM dNTP's (Bioline, UK), 1 U Taq Polymerase (Invitrogen, US), 1 μL lysed bacterial cells, 2% DMSO (Sigma-Aldrich, US) and 1.5 mM $MgCl_2$. Initial denaturation was at 94 °C for 2 min, followed by 30 cycles of amplification with denaturation at 94 °C for 30 s, annealing at 63 °C for 30 s and elongation at 72 °C for 1 min. Final elongation was for 3 min at 72 °C. Isolates which didn't amplify successfully for any of the housekeeping genes were sequenced at the 16S rRNA gene using the universal primers 8F (forward) and rp2 (reverse) to confirm identity at least to the genus level (Turner *et al.*, 1999; Weisburg *et al.*, 1991). PCR amplification for 16S was performed with a BIO-RAD T100 Thermal Cycler using a total reaction volume of 50 μL with a final concentration of 1x PCR buffer (Invitrogen, US), 1 μM for each primer, 0.2 mM dNTP's (Bioline), 1 U Taq Polymerase (Invitrogen, US), 1 μL lysed bacterial cells, 2% DMSO and 3 mM $MgCl_2$ (Invitrogen, US). Initial denaturation was at 95 °C for 2 min, followed by 30 cycles of amplification with denaturation at 95 °C for 45 s, annealing at 55 °C for 1 min and elongation at 72 °C for 2 min. Final elongation was for 10 min at 72 °C.

2.3.5 Sample purification

Successful amplification was confirmed on 1% agarose gels at 120 V for 20 min using 5 μL of the PCR product (Figure 2.2). The remaining sample was purified using an enzymatic reaction: Exonuclease I (Exo1) and Calf Intestinal Alkaline Phosphatase (CIP). Per 45 μL PCR product, 4.25 μL dH_2O, 0.25 μL Exo1 (20.000 U/mL, Biolabs) and 0.5 μL CIP (10.000 U/mL, Biolabs) were added and incubated at 37 °C for 60 min, 85 °C for 15 min and then stored at 4 °C.

For samples producing multiple bands a gel extraction of bands of the expected product size was performed using the QIA quick Gel extraction kit (Qiagen, The Netherlands). The purified PCR products were sequenced by Macrogen Inc (South Korea) using an Applied Biosystems 3730xl DNA Analyzer.

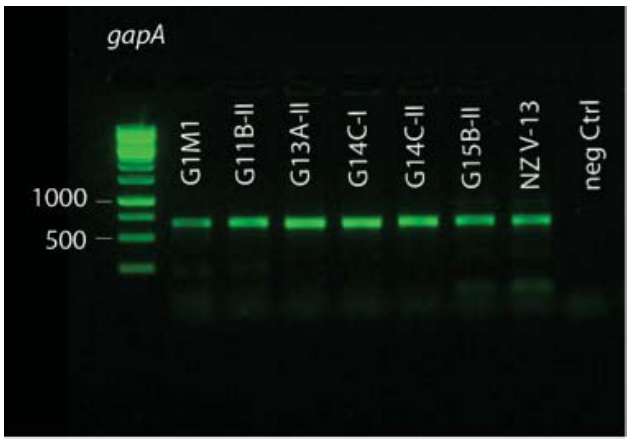

Figure 2.2: Agarose gel. Example of the successful amplification of *gapA* for the first collected samples (G1M1-G15B-II) and the reference strain *Psa* NZ13 (ICMP 19884).

Table 2.2: List of primers. Primers f = forward, r = reverse, P = PCR, S = sequencing, length in basepairs and melting temperature T_m in °C, source 1 = Hwang et al. 2005, source 2 = Sarkar & Guttman 2004

Primer	Gene	Sequence	length	T_m	Source
f P	*gapA*	CCGGCSGARCTGCCSTGG	18	63	1
f S	*gapA*	TCGARTGCACSGGBCTSTTCACC	23	66	1
r P,S	*gapA*	GTGTGRTTGGCRTCGAARATCGA	23	63	1
f P,S	*gyrB*	TCBGCRGCVGARGTSATCATGAC	23	63	1
r P,S	*gyrB*	TTGTCYTTGGTCTGSGAGCTGAA	23	61	1
f P	*gltA*	GCCTCBTGCGAGTCGAAGATCACC	24	64	1
f S	*gltA*	CCTGRTCGCCAAGATGCCGAC	21	62	1
r S	*gltA*	CGAAGATCACGGTGAACATGCTGG	24	62	1
r P	*gltA*	CTTGTAVGGRCYGGAGAGCATTTC	24	63	1
f P	*rpoD*	AAGGCGARATCGAAATCGCCAAGCG	25	63	2
f S	*rpoD*	AAGCGTATCGAAGAAGGCATYCGTG	25	63	2
r P,S	*rpoD*	CCGATGTTGCCTTCCTGGATCAG	23	61	1
F8	16S	AGAGTTTGATCCTGGCTCAG	20		
rp2	16S	ACGGCTACCTTGTTACGACTT	21		

2.3.6 Sequence assembly

Both strands of DNA were sequenced and then forward and reverse sequences of each isolate were assembled and trimmed with Geneious v7.1.4 (Kearse *et al.*, 2012) to obtain the consensus sequence. If the two sequenced strains showed mismatches, the PCR and sequencing was repeated to confirm the true consensus sequence. All sequences were run through the Basic Local Alignment Research Tool (BLAST+) (Camacho *et al.*, 2009) to find the most

similar sequence deposited in the NCBI (National Center for Biotechnology Information) database.

2.3.7 Global *P. syringae* strains

For the MLST analysis 165 non-redundant sequence type profiles of *P. syringae* strains matching in length were downloaded from Plant Associated and Environmental Microbes Database (PAMDB) (Almeida *et al.*, 2010). Sequences from *P. syringae* recently isolated from kiwifruit and air samples in Japan (Tomihama *et al.*, 2016) and kiwifruit isolates from NZ, France and the United States (Visnovsky *et al.*, 2016) were also included. A reduced set of 37 *P. syringae* PAMDB isolates representing the different monophyletic groups of *P. syringae*, as well as the Japanese kiwifruit strains and the US, France and NZ kiwifruit isolates from previous years were used to provide better resolution in the phylogenies. STs were allocated to global sequences starting from 100 to avoid overlap with the STs assigned within this study. A list of all strains used in this study can be found in Table S 2.1 (appendix).

2.4 Analysis

2.4.1 Quality of the sample

Rarefaction analysis was used for assessing the completeness of the sampling by measuring species richness. It uses a random re-sampling without replacement method by calculating the expected number of species of randomly drawn subsamples. Once the number of collected strains stays similar with less observations (in this case Operational Taxonomic Units, OTU's), then the sampling was sufficient to cover the amount of diversity observed in the population. The program MOTHUR v1.34.4 (Schloss *et al.*, 2009) was used to perform rarefaction analysis. Gaps were removed from the alignment prior to running a rarefaction command. Analysis was performed for the whole dataset (n=148), as well as data grouped according to monophyletic *P. syringae* groups. Prior to the rarefaction analysis, a neighbor joining phylogenetic tree was built with reference sequences from various monophyletic *P. syringae* groups. The isolates were grouped according to the cluster they formed in the NJ tree and the rarefaction analysis was then run for each group. Pairwise genetic distances between isolates were calculated and sequences were then assigned to OTU's. This can be done based on various cut-offs, e.g. unique OTU's means that the genetic distance between strains is 0.00, whereas $\leq$0.05 means that strains are assigned into OTU's based on pairwise genetic distances of equal or less than 0.05. MOTHUR implements rarefaction calculations via subsampling and it uses 1,000 iterations to generate the rarefaction curve. The

calculators Shen, Solow and Efron, implemented in MOTHUR v1.34.4, were used to determine whether the sample could have been improved with further sampling. Using an extrapolation formula based on a species accumulation curve, these calculators estimate the number of additional OTU's that would have been found if the sample size had been increased by e.g. 10%.

2.4.2 Indices of diversity

α-diversity was measured by calculating Simpson's index of diversity. Rather than just counting the number of species (or haplotypes, genera) found in a population (species richness), the abundance data are also taken into account. This is important, as in principle a population with high species richness does not necessarily have to be high in diversity. Diversity increases if the relative abundance of each species is distributed more evenly among the population, hence a community with one or two dominant species is considered less diverse. Simpson's index of diversity was converted to the effective number of species (D_c) highlighting the non-linear properties of the Simpson's index (Jost, 2006).

2.4.2.1 Simpson's index of diversity (Gini-Simpson)

$$D = 1 - \sum p_i^2 \tag{2.1}$$

p_i = proportion of the species relative to the total number of species; values range from 0 to 1, with 0 being no diversity and 1 counting as high diversity

The Simpson's index is converted to the effective number of species following Jost (2006):

$$D_c = \frac{1}{1 - D} \tag{2.2}$$

Equitability is the evenness of the distribution of individuals over species and is calculated as follows:

$$E_D = \frac{D}{ln(S)} \tag{2.3}$$

S = total number of species of the site, D = Simpson's index of diversity; equitability values range between 0 and 1, with 1 being complete evenness among the population

For estimating the diversity among sampling sites (β-diversity), the Sorenson-Dice calculator (Dice, 1945; Sorensen, 1947) was used. The β-diversity of pairs of sites is calculated with the number of species shared and not shared among sites treated separately, revealing a dissimilarity index which does not increase with sample size. A Sorensen index of 1 means no shared species among sites and 0 all species are shared. The various α and β diversity indices were calculated in R v3.3.1 (R.Core.Team, 2016) using the vegan package (Oksanen *et al.*, 2016).

2.4.3 Best fitting evolutionary model

The inference of phylogenetic relations relies heavily on the assumption of how evolutionary processes occur (Felsenstein, 1988). The change of one nucleotide to another can be described by a wide variety of models, which need to be tested beforehand to determine the best-fitting model for the

dataset. The correct assumption of the model is tremendously important for constructing phylogenies, as it affects parameters like the branch length and transition/transversion ratio, which could be over- or underestimated (Posada, 2003). Choosing the best-fitting model can be computationally assessed in a variety of ways, either by comparing two models at a time (Likelihood Ratio tests, LRTs) or comparing all competing models at the same time (information criteria). A common strategy for statistical model selection among a finite set is using the Bayesian Information Criterion (Schwarz Information Criterion, BIC) (Schwarz, 1978). The BIC score is an estimation of the performance of a certain model on a dataset, with the lowest BIC score signaling the best model. BIC specifically takes into account over-fitting of a model caused by the introduction of too many parameters by introducing a penalty term. Akaike Information Criterion (AIC) (Akaike, 1974) is a very similar measure to BIC, but is less stringent with the penalty of an increased number of parameters. jModeltest v2.1.7 (Darriba *et al.*, 2012; Guindon & Gascuel, 2003) was used to determine the best fitting evolutionary model (88 models) for each dataset (individual loci and concatenated dataset) using default parameters based on the parameter BIC.

2.4.4 Sequence diversity

The degree of polymorphism is a measure of the amount of nucleotide diversity that can be found at a specific site in the sequence. Translating the nucleotide sequence into a protein sequence allows the determination of

the number of polymorphisms, which result in an amino acid change (non-synonymous mutation) vs. synonymous mutations, which do not change the amino acid. Two other measures of polymorphism were calculated: the nucleotide diversity index π (Nei, 1987), which denotes the average number of polymorphism/nucleotide differences at any site between a pair of sequences; and Watterson's θ (Watterson, 1975) which describes the number of segregating sites. The average genomic GC (guanine + cytosine) content is a feature that is used in bacterial classification, but can also be used to determine regions with atypical GC content which could point at regions having undergone recent horizontal gene transfer. Bacteria show a wide range of diversity in GC contents from as low as 16.5% (*Carsonella ruddii*) to as high as 75% (*Anaeromyxobacter dehalogenans*) (Nakabachi *et al.*, 2006; Thomas *et al.*, 2008), but the average GC content for *P. syringae* ranges from 59-61%.

START2 v0.9.0 beta (Jolley *et al.*, 2001) was employed to calculate parameters of genetic diversity, number of alleles and polymorphic sites and GC content. The number of polymorphic amino acid changes, number of mutations and the nucleotide diversity parameter π were calculated using the program DnaSP v5.10.1 (Rozas & Rozas, 1995). Pairwise genetic variability among and between phylogroups was calculated using MEGA7 (Kumar *et al.*, 2016) using the Maximum Composite Likelihood model with a gamma distribution of 1.

2.4.5 Selection

As housekeeping genes are vital for basic cellular function, horizontal gene transfer is less likely to occur in these parts of the genome. The expected form of selection acting upon those loci is purifying selection. Four different tests were performed to test for selection.

2.4.5.1 Ratio of non-synonymous versus synonymous substitutions

The ratio of non-synonymous versus synonymous substitution (d_N/d_S or K_A/K_S) can be used as an indicator of selective pressure on a protein-coding gene and allows us to infer the direction of natural selection. Non-synonymous substitution means a nucleotide change results in an alteration of the amino acid sequence, whereas a synonymous substitution has no effect on the amino acid sequence. A $d_N/d_S <$ or > 1 indicates negative and positive selection respectively, whereas values close to 0 imply purifying selection. d_N/d_S values were calculated for each locus using START2 v0.9.0 beta (Jolley *et al.*, 2001).

2.4.5.2 Tajima's *D*

Tajima's D statistics (Tajima, 1989) are used to test whether a DNA sequence is evolving randomly (under neutrality) or if some form of selection is acting upon it. It is based on the assumption that under

neutrality the average pairwise difference between sequences (π) is more or less equal to an estimate of Watterson's theta ($\theta = 2\text{Ne}\mu$), calculated based on the number of segregating sites.

$$D = \frac{\pi - S/\alpha_1}{\sqrt{var[\pi - S/\alpha_1]}} \tag{2.4}$$

The difference between π and S/α_1 should be equal to zero if no selection is acting. When the difference is normally distributed, we expect D to lie between -2 and 2. All values outside of this range can be regarded as a guidance that the null hypothesis has to be rejected. A positive value suggests either balancing selection is acting in favour of a few types, but still maintaining distinct types, or that a previously large population suddenly contracted in size. A negative value suggests selection is removing variation within the population (purifying selection), e.g. the population size increases after a bottleneck.

2.4.5.3 Fu and Li's *D** & *F**

Fu and Li's D^* (Fu & Li, 1993) is an expansion of Tajima's D statistics to test neutrality among a set of DNA sequences based on coalescence. The theory is based on the assumption that 'new' mutations happen in the more recent part (external branches) of a genealogy, so mutations in the internal and external branches are compared. Fu and Li's D^* is based on the differences between the number of singletons (unique mutations) and the total number of mutations, whereas Fu and Li's F^* is based on differences between number of

singletons and the average number of nucleotide differences among sequences. The expected value is zero and deviations from zero can be indications of selection. *D* statistics were calculated using the program DnaSP v5.10.1 (Rozas & Rozas, 1995).

2.4.6 Linkage & recombination

2.4.6.1 Testing for linkage

The term linkage equilibrium refers to the probability that all alleles are randomly assorted and there is no linkage between the different loci, meaning that the presence of certain alleles does not allow the prediction of other alleles. By contrast, linkage disequilibrium means that alleles tend to occur in certain combinations across loci and not in a random manner. A way of determining the extent of linkage in a given MLST dataset is using Brown's index of association (I_A) (Brown *et al.*, 1980), which was defined by Maynard Smith *et al.* (1993) as

$$I_A = \frac{V_D}{V_e} - 1 \tag{2.5}$$

V_D = observed mismatch variance V_e = expected mismatch variance

Under the null hypothesis of linkage equilibrium the variance in pairwise differences is expected to be around zero. A population in linkage disequilibrium is of a clonal pattern. However, interpretation of I_A can lead to wrong conclusions about population structure, as it is dependent on the number of loci. As a way to address this issue, a standardized index of

association, I^s_A ,was proposed (Haubold *et al.*, 1998). The program LIAN v3.0 (Haubold & Hudson, 2000) was used to calculate I^s_A using the Monte Carlo test and 10,000 iterations.

2.4.6.2 Recombination

Traditionally, phylogenetic methods assume that the whole length of sequences in a dataset shares a single evolutionary history. Uptake of foreign genetic material is a major force shaping the evolution of bacteria and is mediated via three mechanisms: conjugation via plasmids, transduction by bacteriophages and transformation via uptake of DNA from the environment. These processes result in either homologous or non-homologous recombination. Recombination needs to be taken into account before performing various population genetics analysis, as otherwise the inferred phylogenetic relations will be inaccurate and might not reflect the true evolutionary history. Evolutionary stochasticity will be reduced as a result of recombination obscuring genealogical histories over a genome (McVean *et al.*, 2002). LDHAT v2.2a (Auton & McVean, 2007) was used to test for recombination in population genetics data. The program estimates the rate of recombination using the composite likelihood method of Hudson (Hudson, 2001) with an adaptation to finite-site models. Using a finite-sites model is important because in bacteria patterns originating from recurrent mutations can look like signs of recombination.

Let the mutation rate be

$$\theta = 2N_e\mu \tag{2.6}$$

and recombination rate be

$$\rho = 2N_e r \tag{2.7}$$

with N_e = effective population size, μ = mutation rate, r= recombination rate

The ratio ϵ of recombination rate to mutation rate was calculated as ρ/θ, which gives the likelihood of any single nucleotide polymorphism resulting from recombination rather than mutation. Intragenic recombination rates were calculated for individual loci and intergenic recombination rate was calculated based on the concatenated dataset. In addition, recombination rates were calculated for the concatenated and single gene alignments sorted according to disease status of the orchard (uninfected vs. infected) and host (green vs. gold). Additionally I tested for recombination within PGs, by preparing individual datasets for each gene for each of the different PGs (PG1, PG2, PG3, PG5), as well as including the global sequences from PAMDB. All analyses were performed in LDHAT using the pairwise command including only polymorphic sites with two alleles and the frequency cut-off for missing data was set to 0.2.

2.4.6.3 Detection of breakpoints for recombination events

The analysis was performed with GARD (Genetic Algorithm for Recombination Detection) on http://www.datamonkey.org (Pond & Frost, 2005). The website allows the upload of a given alignment and performs various analyses on an external server. GARD (Kosakovsky Pond *et al.*, 2006) identifies all possible break points in the alignment and then searches

for the most likely location of a break point by reconstructing the phylogeny for each fragment and calculating an information-based criterion, e.g. Bayesian Information Criterion (BIC) or Akaike's Information Criterion (AIC). The GARD analysis was run using the HKY85 nucleotide substitution model with beta-gamma site substitution. After the initial upload of the 148 sequence alignment, 105 duplicate sequences were removed. The inclusion of duplicate sequences would not add any value to the output itself, but only slow down the computation. GARD was run with standard parameters, site-to-site variation set to none and rate classes = 2.

2.4.7 MLST analysis - assignment of sequence types (ST)

Before blasting the nucleotide sequences against the allele profiles deposited in the PAMDB database, the non-redundant sequences were determined using CD-HIT-est (Li & Godzik, 2006). Alleles were assigned to the isolates based on the BLAST hits from the PAMDB database, and novel alleles given numbers starting from 300. Upon completion of the allelic profile, a total of 45 unique sequence types (ST) were assigned.

2.4.8 Group patterns of evolutionary descent

The simple concept of clonality in bacterial evolution predicts that a founding genotype increases in frequency, but diversifies over time, resulting

in a clonal complex. The eBURST algorithm (Feil *et al.*, 2004) uses sequence types (ST's) and their allelic profiles to (1) identify and group related isolates according to a certain group definition and (2) aims at predicting the founding genotype of each group (clonal complex).

eBURST v3 (http://eburst.mlst.net/), which uses an algorithm to identify the founder of a group of related genotypes, was used to group the ST's according to similarity (Feil *et al.*, 2004; Spratt *et al.*, 2004). ST's sharing at least three alleles (SLV - single locus derivant) were allocated to the same group using a bootstrap method (1,000,000 resamplings). Each ST can only be allocated to a single group and the ST with the highest number of SLV in a group is the predicted founder. For a more relaxed group definition, eBURST also identifies and groups according to double locus variant (DLV), which differ at two out of four alleles and triple locus variants (TLV) which differ at three out of four loci. Unlike phylogenetic methods, BURST ignores the relationships between more distantly related groups, which might not be represented correctly anyway due to homologous recombination.

2.4.8.1 Global *P. syringae* evolutionary relationships

PHYLOViZ v2.0 (Francisco *et al.*, 2012), a Java platform, makes use of the goeBURST algorithm to visualize the evolutionary relationships among isolates.

Apart from the eBURST approach, the data can also be presented in a

Minimum Spanning Tree (MST). MST is a spanning tree, whose weight (the weight of the edges combined) is equal to or smaller than the weight of any other spanning tree, which means it is a subset of the data with the minimum possible number of edges. The dataset used included the 45 unique STs from this study, 165 unique STs from the PAMDB database, as well as the global kiwifruit isolates (Tomihama *et al.*, 2016; Visnovsky *et al.*, 2016).

2.4.9 Clustering analysis

Inferring the degree of ancestry among the different samples by assigning them into distinct genetic clusters is a common approach to investigate the genetic structure of a population. This information can be used to correlate to abiotic or biotic factors, e.g. the geographical area where the samples were taken or phenotypic traits.

STRUCTURE v2.3.2 (Pritchard *et al.*, 2000) is a widely used Bayesian clustering approach, which uses a likelihood model to assign individual genotypes to a number of genetic clusters K, where K may be unknown. The program xmfa2struct (developed by Didelot, available from http://www.xavierdidelot.xtreemhost.com/clonalframe.htm) was used to convert the concatenated nucleotide alignment files to xmfa (eXtended MultiFasta), the input file format for STRUCTURE. Xmfa2struct can be used for haploid organisms under linkage equilibrium, where recombination is frequently occurring. Each polymorphic site in the alignment is considered as an allele. No *a priori* population information (sampling sites) were provided, 25,000 iterations were used for burn-in and 100,000 for

MCMC. For consistency of results three replicate runs were performed for each number of K (from 1 to 12). The admixture model, where individuals can have mixed ancestry, was chosen with $\alpha = 1$ as starting value. After the initial run discovered two ancestral groups, the computation was repeated with 500,000 repetitions, 25,000 burn-in, 10 replicate runs for K=1 to K=5.

Structure Harvester v0.6.94 (Earl & VonHoldt, 2012) (http://taylor0.biology. ucla.edu/structureHarvester/) was used to collate the final five STRUCTURE runs. The program has implemented the Evanno method (Evanno *et al.*, 2005), which detects the true number of K by calculating the difference of the log probability of each successive K value (ΔK). CLUMPP v1.1.2 (Jakobsson & Rosenberg, 2007) and distruct v1.1.(Rosenberg, 2004) were used to produce the graphical output.

2.4.10 Phylogenetic reconstruction

Phylogenetics infer the evolutionary descent and hierarchical clustering of a study organism e.g. to classify unknown species, understand the diversity of a sampled population, reveal events that occurred throughout the evolution or simply observe the fate of lineages. There are several methods to construct phylogenetic trees. Distance-based methods like UPGMA (Michener & Sokal, 1957) and Neighbor Joining (NJ) (Saitou & Nei, 1987) are both based on a pairwise genetic distance matrix, however NJ corrects for varying rates of evolution throughout the evolutionary past of an organism and is not based on assumption of a molecular clock. The other three popular methods to infer phylogenetic relationships are based on the

actual nucleotide sequences themselves with the accurate and best-fitting model of evolution chosen for determination of the correct tree: Maximum Parsimony, Maximum Likelihood and Bayesian inference.

Maximum Parsimony (Edwards & Cavalli-Sforza, 1963) is based on the number of substitutions which occurred that minimize the cost. Basically it looks for the tree with the least character conflict. However, the caveat is that approach can lead to inaccurate phylogenies, which result from long branch lengths, high substitution rates, or unequal evolutionary rates across different lineages (Felsenstein, 1978).

Edwards & Cavalli-Sforza (1965) first proposed the idea of implementing likelihood to reconstruct phylogenies, but it was not until Felsenstein (1981) published the first computationally feasible and reasonably fast Maximum Likelihood (ML) algorithm to reconstruct phylogenies, that the ML method was more practicable. Under the chosen model of evolution, the likelihood for each possible tree topology is calculated and the tree with the highest likelihood is regarded as the best-fitting tree.

Bayesian inference (Mau *et al.*, 1999; Yang & Rannala, 1997) is based on the posterior probability (the probability the tree fits the presented data, taking into account the prior probabilities of the tree) and employs the use of Markov Chain Monte Carlo (MCMC). The algorithm returns a number of trees providing the highest probability, from which a consensus tree is built. Bayesian methods allow the use of more complex evolutionary models.

The most popular and commonly used methods nowadays are ML and

Bayesian inference. Both methods were employed to compare the resulting phylogenetic trees of the two different inference methodologies.

2.4.10.1 Maximum Likelihood (ML) Tree

The program TREEPUZZLE v5.3 (Schmidt *et al.*, 2002) was used to construct ML trees for each single gene and the concatenated dataset. The most fitting evolutionary model, as determined with jModeltest, was used and program parameters were set to 100,000 puzzling steps, a neighbor-joining tree for parameter estimation use and quartet puzzling as tree search procedure. All duplicate sequences were removed and the input file consisted of a single representative of each of the 45 unique sequence types.

The reconstruction of trees takes place in three steps: (1) TREEPUZZLE computes the maximum likelihood distances between groups of four sequences and weights them according to posterior probabilities. (2) Stepwise, starting from one quartet tree, sequences are added, using the ML information from the first step, producing a number of intermediate trees. (3) The consensus tree is built with a 50% majority consensus.

The Shimodaira-Hasegawa (SH) test is applied for testing the congruence between single trees (Shimodaira & Hasegawa, 1999). The SH-test is a way of comparing tree topologies and hence determining if recombination has occurred between the housekeeping genes. The program Dnaml, incorporated

in PHYLIP v3.695. (Felsenstein, 1989), was used to perform the SH-test. Dnaml is a DNA Maximum Likelihood program and allows the comparison of user-specified trees with log likelihood values. The SH-test is a statistical test, which makes use of the branch lengths and evaluates differences in log-likelihoods of each tree against each other, without altering branch lengths any further. The difference between the highest log-likelihood and the other tree's values is compared and the output gives an indication of significance, i.e. whether the tree is significantly worse than the best one. It was run using the default parameters, but with user trees as input file and a random number seed of 333.

2.4.10.2 Bayesian Tree

For Bayesian inference of genealogical relations between strains, taking into account point mutation and homologous recombination, the program Clonalframe v1.1 (Didelot & Falush, 2007) was employed. As discussed previously, recombination is a major problem in phylogenetic inference, as horizontal gene transfer can obscure and falsify the resulting tree. Clonalframe works under the assumption that all recombination events introduce an unknown number of novel polymorphisms, however it does not try to determine the origins of stretches of DNA sequence created via homologous recombination.

Duplicate runs of Clonalframe were performed starting from 100,000 to 500,000 iterations, using a burn-in of 50,000 iterations based on the concatenated dataset of 45 STs. Values of various parameters (θ, ρ,..) of

duplicate runs were compared to see if they have converged and the best run was chosen according to converged values. In addition topologies of the trees were compared, with each branch in the tree having a minimum of 50% support based on the posterior distribution.

2.4.10.3 Phylogenetic network

Phylogenetic trees are based on evolutionary models, where point mutations are supposedly the sole source of evolutionary differences between taxa. However, microevolutionary events, such as e.g. gene duplication/loss, recombination, Horizontal Gene Transfer can obscure the true evolutionary history presented by phylogenetic trees. Although there are some reconstruction methods which consider recombination, another method of inferring the evolutionary history of individual taxa are phylogenetic networks. These allow for reticulate events and model the relationships as a network based on a choice of distances, sequences or trees, with the nodes representing taxa and their evolutionary relationships being represented by edges, where multiple edges can be parallel (Huson & Bryant, 2006). Splitstree v4.13.1 (Huson & Bryant, 2006) was used to create a splits network based on the neighbor-net method using the concatenated dataset of the 45 unique STs.

2.4.11 Testing for biogeographic structure - PERMANOVA

A permutational multivariate analysis of variance (PERMANOVA) (Anderson, 2001; McArdle & Anderson, 2001) was used to test whether community composition and genetic diversity varied among biogeographical structures (sampling location, time points of sampling, infection status of orchard and cultivar). Abundance data of each ST and the pairwise genetic distance matrix were used as input for PRIMER v6.1.12 (PRIMER-E Ltd., Plymouth, UK, PERMANOVA+ add-on v1.0.2.) respectively. In cases of too few permutations available to produce a reasonable P-value, the Monte-Carlo random draws were applied to obtain correct P-values.

One-factor PERMANOVA designs were used to compare abundance and genetic diversity among the different factors (fixed) with 9999 permutations using unrestricted permutation of raw data. A two-factor design for cultivar (fixed, 2 levels) + infection (fixed, 2 levels) and orchard (fixed, 4 levels) + time (fixed, 3 levels), as well as a three-factor nested PERMANOVA design with 9999 permutations with cultivar (random factor, 2 levels), infection (random factor, 2 levels) and time (random factor 3 levels) was performed in PRIMER to determine the effect of the different factors.

The non-metric multidimensional scaling maps (MDS) to visualize similarity among each represented sample were run with 100 random restarts to ensure a low stress-level. The MDS plots highlight the effect of the different factors on the differences in abundance and genetic diversity.

2.5 Results

P. syringae was isolated from the leaf surfaces of two different cultivars of *Actinidia chinensis*: *A. chinensis* var. *chinensis* Hort16A (gold) and *A. chinensis* var. *deliciosa* Hayward (green), which vary in their susceptibility to *Psa*: Hort16A is more susceptible than the green Hayward (Cameron & Sarojini, 2014; Ferrante & Scortichini, 2010). One infected and one uninfected orchard of each variety was sampled by collecting leaves from six separate kiwifruit vines along a diagonal path.

A total of 233 *Pseudomonas* strains were isolated from leaf washes, based purely on colony morphology (creamy white, round). Out of the 233 isolates, 148 were determined to belong to the species *P. syringae* (Table 2.3) based on sequencing of four housekeeping genes (*gapA, gyrB, gltA* and *rpoD*). Initially, 71 out of the 148 sequences turned out to be too short at the 5' end for *rpoD* using the forward and reverse primers of Hwang *et al.* (2005). The PCR was therefore repeated with the forward *rpoD* primer from Sarkar & Guttman (2004), so the sequences were overlapping with the reference sequences from PAMDB.org. The final alignment lengths for individual genes were 476 bp, 507 bp, 529 bp and 495 bp for *gapA, gyrB, gltA* and *rpoD* respectively (2010 bp concatenated). The following analysis is based on the dataset consisting of 148 *P. syringae* strains, unless stated otherwise.

Table 2.3: Overview of the number of samples. The number of *P. syringae* isolates sampled from each orchard per time point; numbers in brackets equals to the total number of pseudomonads isolated.

	Spring	Summer	Fall	Total
K (H16A uninfected)	9 (13)	7 (14)	12 (23)	28 (50)
G (H16A infected)	13 (22)	15 (20)	16 (26)	44 (68)
H (Hayward infected)	10 (11)	11 (18)	15 (26)	36 (55)
R (Hayward uninfected)	16 (18)	13 (20)	11 (22)	40 (60)
				148 (233)

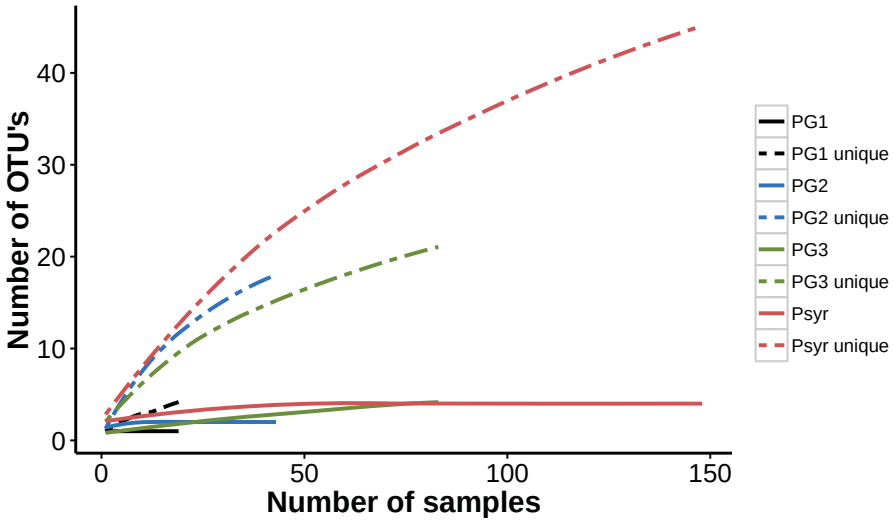

Figure 2.3: Rarefaction curves based on the concatenated sequences (2010 bp). Two curves each are shown for *P. syringae* (n=148) and the sequences grouped according to phylogroups (PG): solid lines represent grouping based on unique STs and dashed lines according to a cut-off equal to the average pairwise genetic distance of the group: PG 1, PG 2 & PG 3 = 0.02, Psyr all = 0.05.

2.5.1 Depth of sampling

The rarefaction curve for unique OTU's was not saturated, however choosing a cut-off of 0.05 resulted in the curve levelling off (Figure 2.3). The cut-off was chosen based on the average pairwise genetic distance (=0.0531, 101 nucleotide changes in the alignment of 2010 bp). This suggests the sampling was sufficient to cover the diversity present in the population. After the collection of 50 samples, 22 unique OTU's were identified; after reaching 100 isolates 30 unique OTU's were identified. A total of 45 unique OTU's and 5 OTU$\leq$0.05 were found in 148 sampled strains.

The Shen calculator revealed that the quality of the sample size would have increased only to a minimal extent with further sampling. If the sample size had been increased by 10%, only two additional unique OTU's would have been identified. The Solow and Efron calculator showed very similar results. This further confirms that sampling was sufficient and the forthcoming analysis is based on a satisfactory dataset.

The rarefaction curves based on datasets for each monophyletic group (PG) showed that based on the cut-off being equal to the mean pairwise genetic distance of each group, all curves are saturated (Figure 2.3). Hence, the sample of 148 isolates covered the diversity of all three phylogroups.

2.5.2 Indices of diversity

Two parameters as a measure of α-diversity, the Simpson's index of diversity and effective number of species, were calculated. The infected green orchard showed the highest amount of diversity with D=0.904, whereas the uninfected gold orchard showed the least diversity (D=0.737). The evenness (or equitability) E_D was low among all sampling sites with values from 0.136 to 0.290, meaning that the species are not equally represented in the population (Table 2.4). Simpson's index of diversity is a highly non-linear measure. Jost (2006) suggests to convert the indices to the true diversities (Dc), which equals the effective number of species, so the values can be directly compared: e.g. the infected green orchard (Dc = 10.42) had three times the number of effective species, compared to the uninfected gold orchard (Dc = 3.8). Comparing these numbers to the actual index of diversity highlights its non-linear properties. Sorensen's index of dissimilarity, as a measure for β-diversity, was 0.847, indicating that the species composition across the different sampling varies heavily.

Table 2.4: Indices of diversity. D = Simpsons index of diversity, E_D = Simpsons evenness, Dc = converted to effective number of species.

	D	**E_D**	**Dc**
Uninfected Hort16A	0.737 ± 0.111	0.136	3.8
Infected Hayward	0.904 ± 0.041	0.290	10.42
Uninfected Hayward	0.866 ± 0.054	0.187	7.46
Infected Hort16A	0.834 ± 0.083	0.137	6.02

2.5.3 Testing models of evolution

For the concatenated alignment, the best fitting model of evolution was the Hasegawa, Kishino and Yano (HKY) model + I + G (Figure 2.4). It showed the lowest BIC value and was therefore chosen for subsequent analysis. The best substitution models for the individual gene alignments were HKY + I (*gapA*), HKY + G (*gltA*), HKY + I + G (*rpoD*) and TrNef + G (*gyrB*) [G = Gamma distribution, gamma distributed rate variation among sites; I = proportion of invariable sites, extent of non-changing sites].

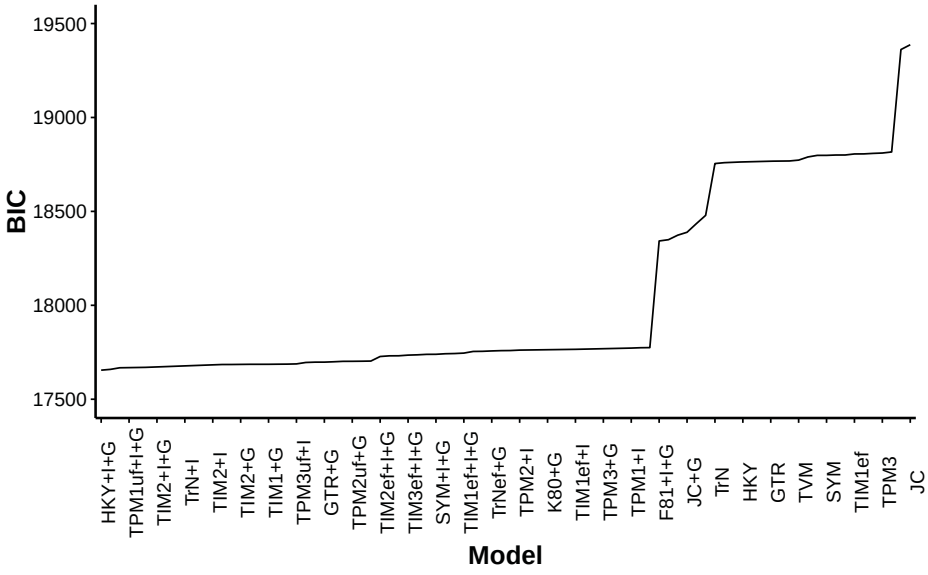

Figure 2.4: Best nucleotide substitution model. The graph shows the BIC values for the different nucleotide substitution models tested for in jModeltest2 for the concatenated dataset.

2.5.4 Sequence diversity

The total concatenated alignment length was 2010 bp, with no insertions or deletions detected for any of the four loci. The number of alleles ranged from 25 (*gapA*) to 35 (*rpoD*) (Table 2.5). There was a total of 412 polymorphic sites, ranging from 80 (16.81%, *gapA*) to 145 (28.6%, *gyrB*). At the level of polymorphic amino acids, *rpoD* was the locus with the highest amount of amino acid changes (19/166, 11.4%), whereas *gltA* displayed the lowest number (2/176, 1.13%). The locus with the highest degree of nucleotide polymorphism (*gyrB*) did not show the highest number of amino acid changes. No nonsense mutations (point mutations resulting in an early stop codon) were detected for any of the loci. The nucleotide diversity index π and Watterson's θ were highly consistent among loci, varying from 0.040 to 0.055 for π and from 0.024 to 0.041 for θ. The average GC content 57.99% is similar to that found in other *P. syringae* studies (59- 61%), however locus *gyrB* displayed an unusually low GC content of 53.15%.

Table 2.5: Nucleotide and amino acid diversity. L = length in bp, AA length = amino acid chain length, GC = average GC content in %, N_A = number of alleles, P = number of polymorphic sites (%), d_N/d_S ratio of non-synonymous versus synonymous substitutions, π = nucleotide diversity indices, θ = Watterson's theta.

Locus	**L (bp)**	**AA length**	**GC**	$\mathbf{N_A}$	**P**	$\mathbf{d_N/d_S}$	π	θ
gapA	476	158	60.81	25	80 (16.81)	3.365	0.055	0.024
gyrB	507	169	53.15	28	145 (28.60)	0.018	0.054	0.041
gltA	529	176	58.45	27	88 (16.64)	0.011	0.040	0.025
rpoD	495	166	59.56	35	99 (20)	2.022	0.042	0.03
Mean	502	167	57.99	29	105 (20.51)	1.354	0.048	0.03

2.5.5 Pairwise genetic distance within and among phylogroups

The pairwise genetic difference within phylogroups was not greater than 2.7%, whereas among phylogroups the variability ranged from 6-11% (Table 2.6). These values are consistent with previously described measures of genetic variability among these phylogroups for *P. syringae* (Berge *et al.*, 2014; Morris *et al.*, 2010; Sarkar & Guttman, 2004).

Table 2.6: Mean pairwise genetic distances within (bold values) and between phylogroups. Analyses were conducted on the concatenated alignment (2006 bp, gaps removed) using the Maximum Composite Likelihood model (gamma distribution with alpha parameter of 1).

Number of strains	Phylogroup	1	2	5	3
19	1	**0.01**			
43	2	0.098	**0.027**		
3	5	0.111	0.106	**0.008**	
83	3	0.099	0.063	0.107	**0.014**

2.5.6 Testing for selection

All four genes are housekeeping genes, so the main form of selection acting on these loci is expected to be purifying selection. None of the loci deviated significantly from neutrality based on Tajima's D values (Table 2.7).

Fu & Li's D^* values were positive for three loci (*gapA, gltA* and *rpoD*), with only *gapA* being significant by the F^* test. Positive values are indicative of a selective sweep affecting these genes, which is undermined by the lack of singletons. In general Fu & Li's D^* is more sensitive than Tajima's D, which could explain the non-significant values for Tajima's D.

The d_N/d_S values for *gyrB* (0.0175) and *gltA* (0.0106) showed values <1 and by contrast, *gapA* (3.3653) and *rpoD* (2.0221) showed values of >1. The typical interpretation of these values would be that >1 indicates positive selection, whereas <1 suggests strong purifying selection acting on these loci. However, originally the theoretical framework of d_N/d_S measures were developed for more distantly related species and not assuming a

population context (Kryazhimskiy & Plotkin, 2008). An analysis by Kryazhimskiy & Plotkin (2008) showed that for population samples and more closely related isolates the interpretation of d_N/d_S changes drastically; strong positive selection produces d_N/d_S values of <1 and it was shown that d_N/d_S close to 1 might even under both strong positive or mederate negative selection. This means that the power for inferring selection pressures by using d_N/d_S for population samples is reduced and the interpretation may not be valid from this measurement.

Table 2.7: Testing for selection. Estimation of Tajima's D and Fu & Li's D^* and F^* in DnaSP v5 and the calculation of d_N/d_S ratio in START. P = level of significance, NS = not significant, ** = P<0.02.

	Tajima's *D*	***P***	**Fu&Li's D***	***P***	**Fu&Li's F***	***P***	d_N/d_S
gapA	1.598	NS	2.568	**	2.560	**	3.365
gyrB	-0.530	NS	0.605	NS	0.097	NS	0.018
gltA	0.525	NS	1.886	**	1.529	NS	0.011
rpoD	-0.050	NS	1.897	**	1.218	NS	2.022

Table 2.8: Index of association. Determining the linkage between alleles using the standardized index of association. I^s_A = index of association, P = level of significance. [a] dataset consisting of individual ST profiles n= 45.

	$\mathbf{I^s_A}$	***P***
All	0.5975	<0.0001
PG1	0.9829	<0.0001
PG2	0.7292	<0.0001
PG3	0.4919	<0.0001
PG5	1.1667	0.111
STs_a	0.1935	<0.0001

2.5.7 Linkage & recombination

2.5.7.1 Index of association

Significant linkage disequilibrium was detected for the whole population 0.5975 (P <0.0001), and at the level of individual phylogroups, except for PG5 (Table 2.8.) Running the analysis with only one representative of each ST (n=45), the value of I^s_A dropped down to 0.1935 (P <0.0001), but remained significant. This suggests association, which is indicative of limited recombination within the data.

2.5.7.2 Analysis of recombination using LDhat

The testing for linkage disequilibirum in the previous section suggested that recombination has occurred. In order to obtain the ratio of mutation versus recombination, LDhat was used to calculate recombination and mutation rates per locus.

NZ isolates

Intragenic recombination rates (ρ) ranged from 0.012 (*rpoD*) to 0.038 (*gyrB*) and a value of 0.006 for the concatenated dataset calculated using the coalescence-based method implemented in LDhat. The ratio ϵ (recombination rate/mutation rate) gives the likelihood that a given genetic variability is due to recombination versus mutation. Ratio ϵ ranged from 0.187 (concatenated) to 0.931 (*gyrB*), suggesting that any single nucleotide

polymorphism found is up to five times more likely to have arisen from a mutation than recombination (Table 2.9).

Clustering sequences by phylogroup revealed there was no recombination in PG1 and PG5 (ρ=0), but this could be due to the small sample size in these phylogroups. There was evidence for recombination in PG3, more specifically for *gltA* (ϵ=1.18) and *rpoD* (ϵ=4.07), whereas in PG2 recombination only appeared to be acting on *rpoD* (ϵ=1.734). These trends are consistent with what has been described earlier (Sarkar & Guttman, 2004), where recombination occurred in PG2 and PG3, whereas PG1 showed almost strictly clonal patterns.

The test for recombination of an influence of the disease status (uninfected vs. infected) or host (Hayward vs. Hort16A), did not detect any outstanding patterns of recombination, in fact recombination rates ρ ranged from 0.003 to 0.09 and the ratio ϵ varied from 0.098 to 0.285 (Table 2.10).

Evidence of recombination at a global scale

A total of 165 unique *P. syringae* MLST profiles across different phylogroups from PAMDB were included to add a global component to the analysis. In addition, various *P. syringae* isolated from kiwifruit and air in Japan (Tomihama *et al.*, 2016) and from kiwifruit in NZ and USA (Visnovsky *et al.*, 2016) were included to demonstrate the population structure of *P. syringae* found in the phyllosphere of kiwifruit.

On a global scale, the overall intergenic recombination rates were low

among phylogroups, ranging from 0.005 to 0.012 for the concatenated dataset. However some signs of recombination were detected for individual genes (Table 2.11). Including the global strains resulted in observing recombination among PG1, however this was not too surprising, since most NZ kiwifruit strains in PG1 were *Psa* isolates and for the global dataset more diverse members of PG1 were included.

Table 2.9: LDhat recombination analysis. Showing the length of the alignment in bp, mutation rate θ (=2Neμ) per site, recombination rate ρ (=2Ner) per site and $\epsilon = \rho/\theta$. First group of three columns showing the results for the whole alignment of 148 *P. syringae* sequences, the remaining columns show the results for each monophyletic group (PG1 1: n=19, PG2: n=43, PG3: n= 83, PG5: n = 3). Signs of recombination highlighted in grey.

	Gene	**Length**	θ	ρ	ϵ
All	**concat**	2010	0.030	0.006	0.187
	gapA	476	0.024	0.021	0.902
	gyrB	507	0.041	0.038	0.931
	gltA	529	0.025	0.012	0.461
	rpoD	498	0.030	0.012	0.416
	Mean		0.030	0.018	0.579
Phylogroup 1	**concat**	2010	0.009	0.000	0.000
	gapA	476	0.005	0.000	0.000
	gyrB	507	0.011	0.000	0.000
	gltA	529	0.010	0.000	0.000
	rpoD	498	0.011	0.000	0.000
	Mean		0.009	0.000	0.000
Phylogroup 2	**concat**	2010	0.017	0.002	0.120
	gapA	476	0.014	0.006	0.457
	gyrB	507	0.029	0.000	0.000
	gltA	529	0.012	0.008	0.654
	rpoD	498	0.013	0.023	1.734
	Mean		0.017	0.008	0.593
Phylogroup 3	**concat**	2010	0.016	0.015	0.937
	gapA	476	0.014	0.013	0.898
	gyrB	507	0.038	0.024	0.636
	gltA	529	0.005	0.006	1.175
	rpoD	498	0.008	0.033	4.073
	Mean		0.016	0.018	1.544
Phylogroup 5	**concat**	2010	0.008	0.000	0.000
	gapA	476	-	-	-
	gyrB	507	0.017	0.000	0.000
	gltA	529	0.004	0.000	0.000
	rpoD	498	0.009	0.000	0.000
	Mean		0.009	0.000	0.000

Table 2.10: LDhat recombination analysis for host and disease status. Based on an alignment of 2010 bp. N = number of sequences, S = number of segregating sites, mutation rate θ per site, recombination rate ρ per site and ratio $\epsilon = \rho/\theta$.

	N	**S**	θ	ρ	ϵ
concat green	76	316	0.032	0.009	0.285
concat gold	72	304	0.031	0.003	0.098
concat uninfected	68	226	0.023	0.004	0.173
concat infected	80	310	0.031	0.004	0.114

Table 2.11: LDhat recombination analysis for global data sorted according to phylogroups (PG). Length of alignment in bp, N = number of sequences, S = number of segregating sites, mutation rate θ, recombination rate ρ and ratio $\epsilon = \rho/\theta$.

Gene	**Length (bp)**	**N**	**S**	θ	ρ	ϵ	**Tajima's D**
Phylogroup 1							
concatenated	2013	30	175	0.022	0.005	0.231	-0.952
gapA	476	30	57	0.030	0.000	0.000	-2.224
gyrB	507	30	53	0.026	0.014	0.534	-0.582
gltA	529	30	25	0.012	0.008	0.647	0.718
rpoD	501	30	40	0.020	0.047	2.324	-0.563
Phylogroup 2							
concatenated	2013	101	330	0.032	0.007	0.225	-0.562
gapA	476	101	63	0.026	0.015	0.588	0.357
gyrB	507	101	98	0.037	0.012	0.324	0.722
gltA	529	101	52	0.019	0.029	1.527	-0.283
rpoD	501	101	117	0.045	0.000	0.000	-2.227
Phylogroup 3							
concatenated	2013	51	220	0.024	0.012	0.501	0.027
gapA	476	51	48	0.022	0.021	0.956	0.59
gyrB	507	51	77	0.034	0.016	0.477	0.769
gltA	529	51	49	0.021	0.010	0.468	0.122
rpoD	501	51	46	0.020	0.039	1.896	-1.874

2.5.7.3 GARD recombination breakpoint

Out of 414 potential breakpoints, evidence for two breakpoints was discovered using GARD, the first one located at 551 bp and the second one at 836 bp. Both of the breaks were located within the *gyrB* gene. The Kishino Hasegawa (KH) test for topological incongruence found that trees segregated at the breakpoints were highly significant ($p<0.01$) at these breakpoints.

2.5.8 Clonal complexes & ancestry based on MLST analysis

A total of 45 unique ST's were identified among the 148 *P. syringae* strains. All ST's except ST904 (*Psa*) were novel when compared to all available *Pseudomonas syringae* ST's deposited in the Plant Associated and Environmental Microbes Database (PAMDB).

Infected orchards (both green and gold) harbored the highest number of unique ST's, sharing only three ST's between them (Figure 2.5). There were no ST's present in all four orchards, but two STs were shared among three orchards (ST1 and ST3). The frequency of ST's found in the different orchards varied accordingly, with all orchards clearly showcasing a few very abundant ST's (Figure 2.6).

Two clonal complexes (CC) (21 isolates), five doubletons (32 isolates) and 28 singletons (95 strains) were identified among the 148 tested isolates (Figure 2.7). CC1 and CC2 are comprised of 11 and 10 strains, respectively.

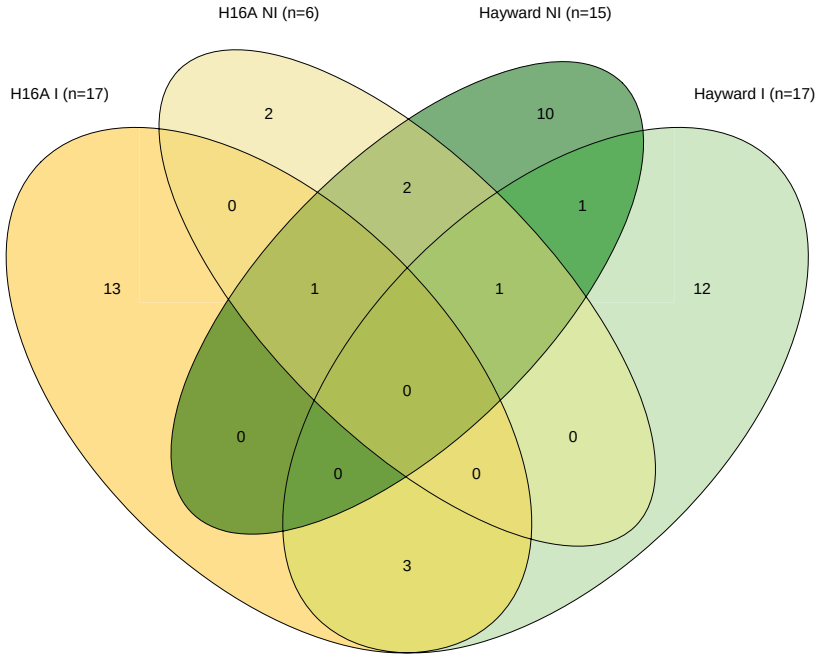

Figure 2.5: Venn diagram displaying the unique and shared ST's among the orchards. I = infected, NI = uninfected.

In both CC's, the predicted founder is the most predominant ST in the complex and the SLV (single locus derivant) was represented by one strain only. The most common ST's in the dataset were ST1 (24/148) and ST904 (15/148). The highest amount of SLV's was found in the uninfected green orchard. ST6 (freq=8) was found in equal frequencies in the two uninfected gold and green orchards, with its SLV ST38 (freq=1) and ST40 (freq=1) present in the uninfected green and ST39 (freq=1) in the infected green orchard. Interestingly, ST6 and ST38 were recovered from the same leaf in autumn, whereas one ST6 isolate was isolated from the same leaf as ST40

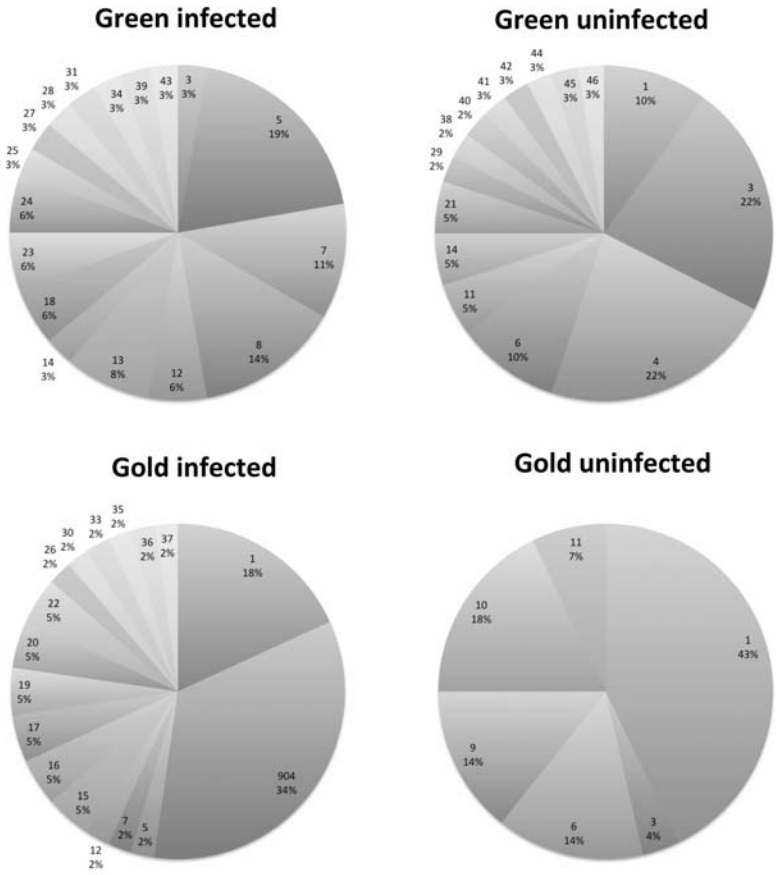

Figure 2.6: Overview of the distribution of ST's per orchard. Each pie graph shows the frequency of observed ST's in %.

in spring. For CC2 (ST5, ST27 and ST28), ST5 was found in two orchards, the infected gold (freq=1) and the infected green (freq=7), whereas the two SLV of ST5 (ST27, ST28) were found in the infected green orchard only. The doubletons were mostly isolated from the same leaf of the same orchard, except for ST20 & ST21, which were isolated from the infected gold and the uninfected green respectively, as well as ST14 (both green orchards) and ST45 (uninfected green only).

In order to display any temporal effect in the detection of STs from the different orchards, the population snapshot is displayed according to the time point of isolation (Figure 2.8).

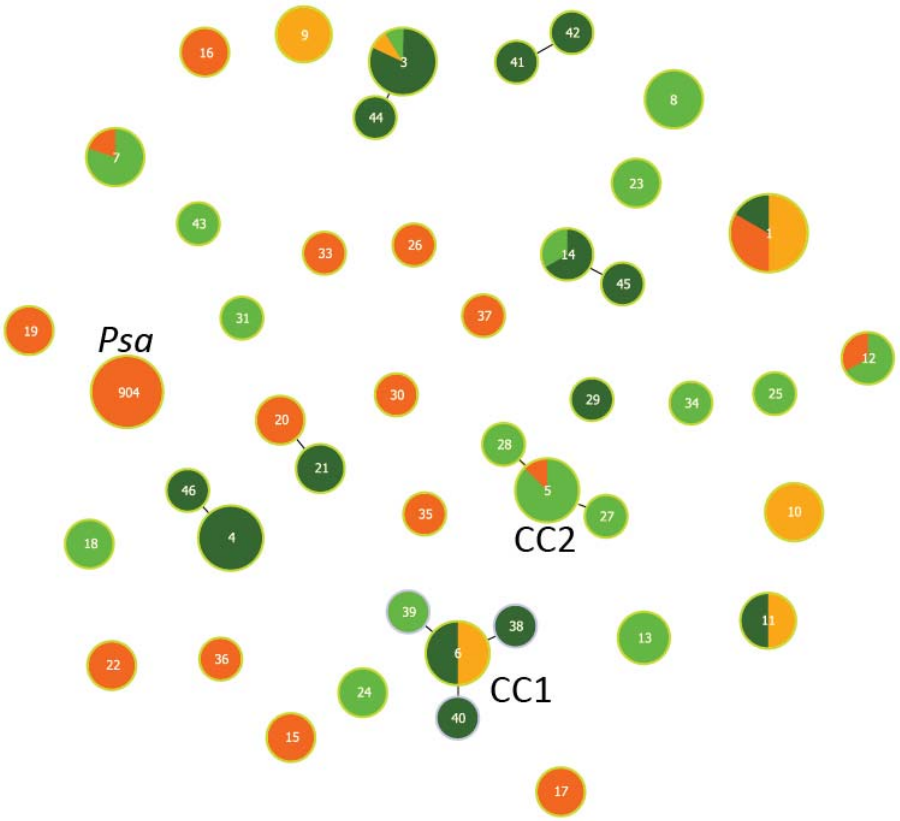

Figure 2.7: Population snapshot. The graph shows a population snapshot of the ST's discovered in 148 isolates (illustrated using PHYLOViZ, single-locus variant level). The size of the circle increases with the frequency of the ST in the dataset. Two clonal complexes (CC) were identified. Colours correspond to the different orchards: orange = infected gold, yellow = uninfected gold, light green = uninfected green, dark green = infected green.

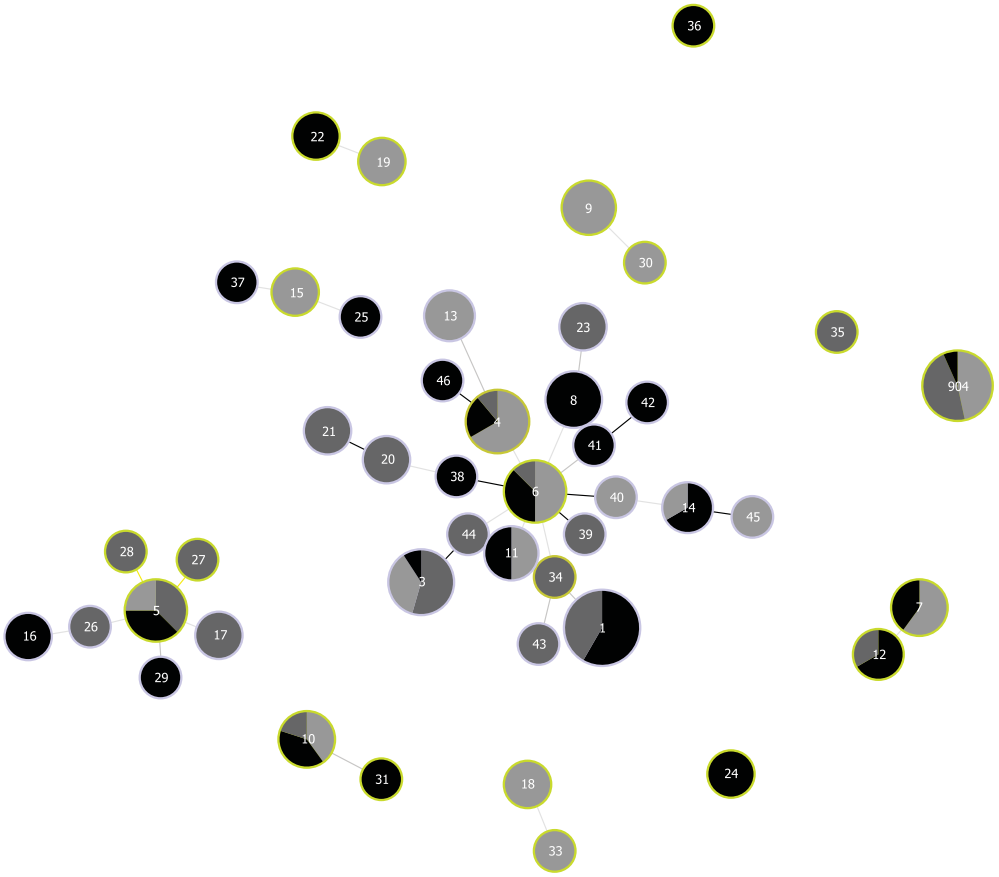

Figure 2.8: Population snapshot (triple-locus variant level) highlighting the time of isolation using PHYLOViZ. Spring (after bud break, n=48, 32.43%) = light grey, summer (n=46, 31.08%) = dark grey, autumn (prior to harvest, n=54, 36.49%) black.

2.5.8.1 Shared alleles and STs among global *P. syringae* isolates

The Minimum Spanning Tree (MST) shows an overview of the global STs with strains isolated from NZ (years 1991, 2010-11, 2013-14), Japan (2015) and the US (Figure 2.9). Surprisingly, we discovered a shared ST (ST 3), which was isolated from non-symptomatic *A. chinensis* var. *deliciosa* in Japan (KW11, 2015) and New Zealand (e.g. R22B-1, H31A-I, 2013, 2014) and non-symptomatic *A. chinensis* var. *chinensis* (LKCH0, 2010; K22A-I, 2014) in NZ; other Japanese kiwifruit ST's cluster closely with *P. syringae* originating from kiwifruit in NZ. In addition, ST16 was recovered from kiwifruit leaves in NZ 1991 (LKEP0) and 2013 (G35B-I, G35B-II).

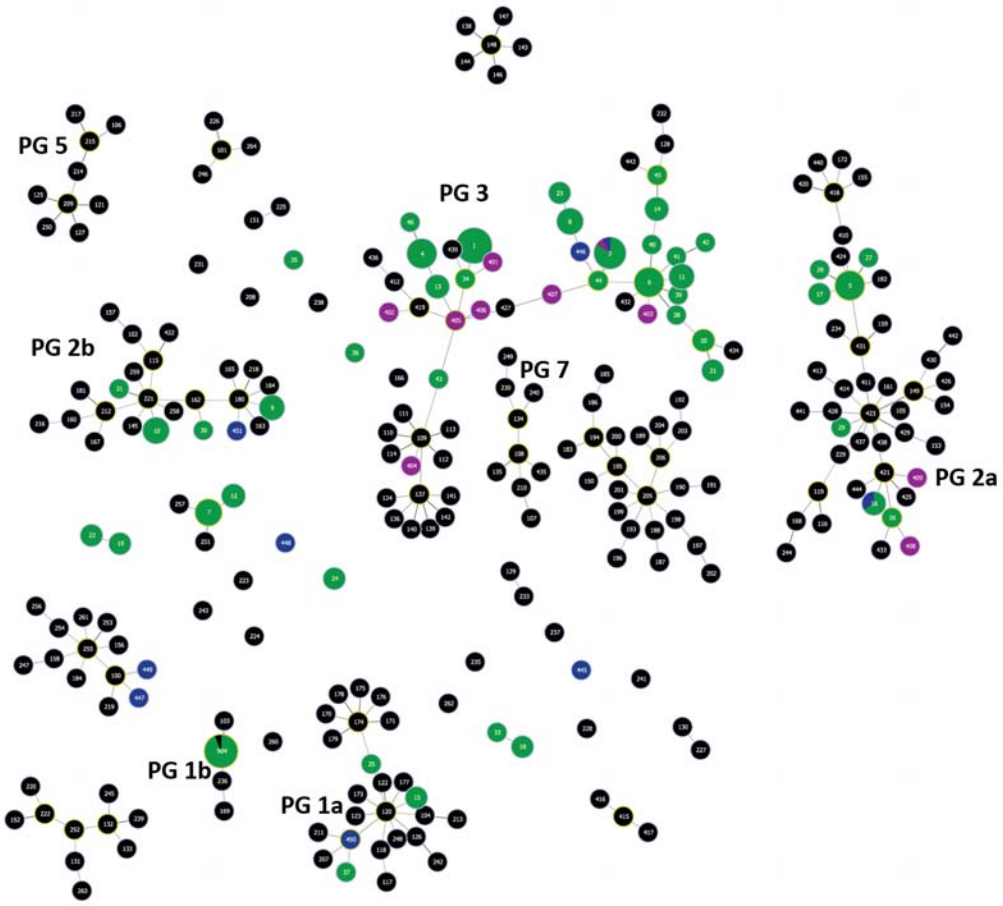

Figure 2.9: Global Minimum Spanning Tree at the triple-locus variant level. Global isolates from PAMDB in black, NZ isolates this study (green), Japanese isolates from kiwifruit (purple), NZ and US isolates (dark blue).

2.5.9 Ancestry of polymorphisms

The admixture model was run in STRUCTURE to assign samples based on their proportion of ancestral subpopulations. Multiple runs were performed from K=1 to K=5 and the most likely number of subpopulations among *P. syringae* was K=3 based on ΔK (Earl & VonHoldt, 2012). Although this number of K biologically makes sense, with the correlation to the three major phylogroups (Figure 2.10), it has to be regarded with some caution. ΔK identifies the highest level of population structure and is known to miss fine scale or hierarchical population structure (e.g. Waples & Gaggiotti, 2006). When looking at the highest probability of $\ln(K)$, the lowest value actually lies at K=5 (Figure 2.11).

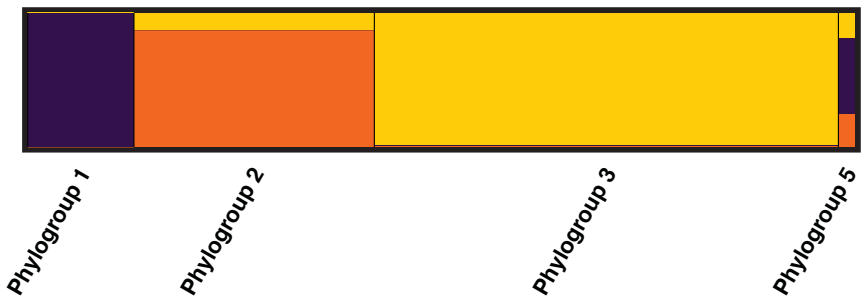

Figure 2.10: Membership coefficients of K=3, grouped according to Phylogroup distinction. Graph is the mean value of 10 replicate runs of K=3 summarised using CLUMPP v1.1.2 (Jakobsson & Rosenberg, 2007) and distruct v1.1 (Rosenberg, 2004).

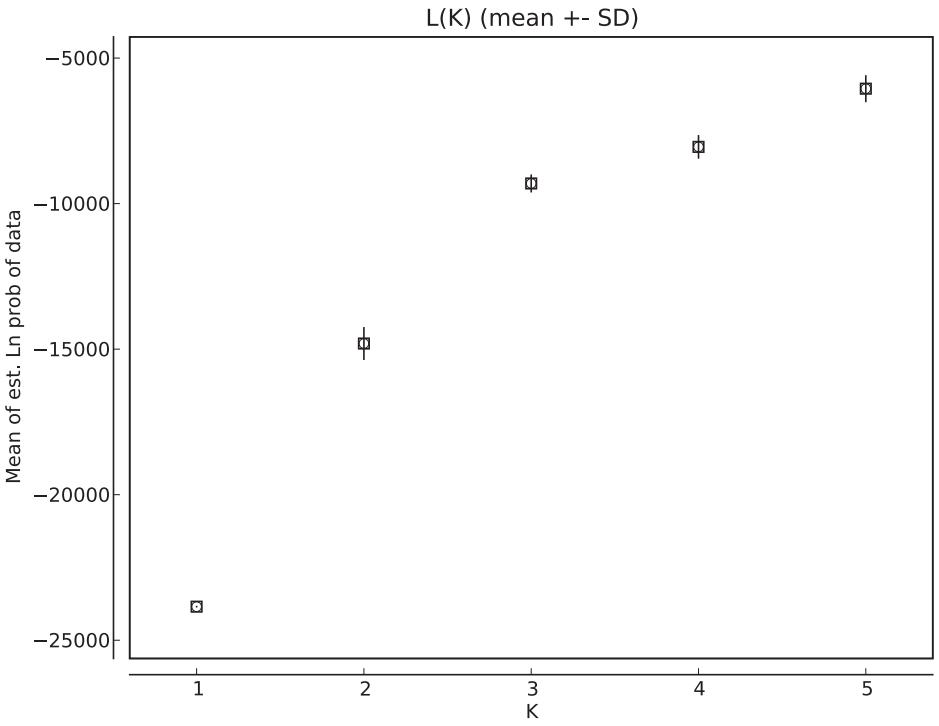

Figure 2.11: Mean of ln (probability of data being assigned to N number of clusters) vs. the most likely number of *K*. SD is presented as error bars.

2.5.10 Ancestry based on nucleotide sequences

Prior analysis (section 2.5.7) showed limited evidence of recombination overall and, on the basis of these results, it is justified to obtain a picture of evolutionary ancestry. However, to ensure recombination does not obscure phylogenetic inference, three different models of phylogenetic construction were employed and the program Clonalframe was used, which removes chunks in the alignment which are potentially the result of a recombination event.

2.5.10.1 Maximum Likelihood Tree

In order to obtain a clearer picture of the ancestry of sequence types based on the actual nucleotide sequences, a Maximum Likelihood Tree was constructed using TREEPUZZLE (Figure 2.12). The analysis found only 1.6% partly resolved and 0.7% unresolved quartets, meaning that the tree is resolved reliably. A high number of partially resolved an unresolved quartets would result in a multifurcating tree. The phylogeny of the ML tree corresponds to the anticipated clustering in *P. syringae* phylogroups: Phylogroup 1,2, 3 and 5.

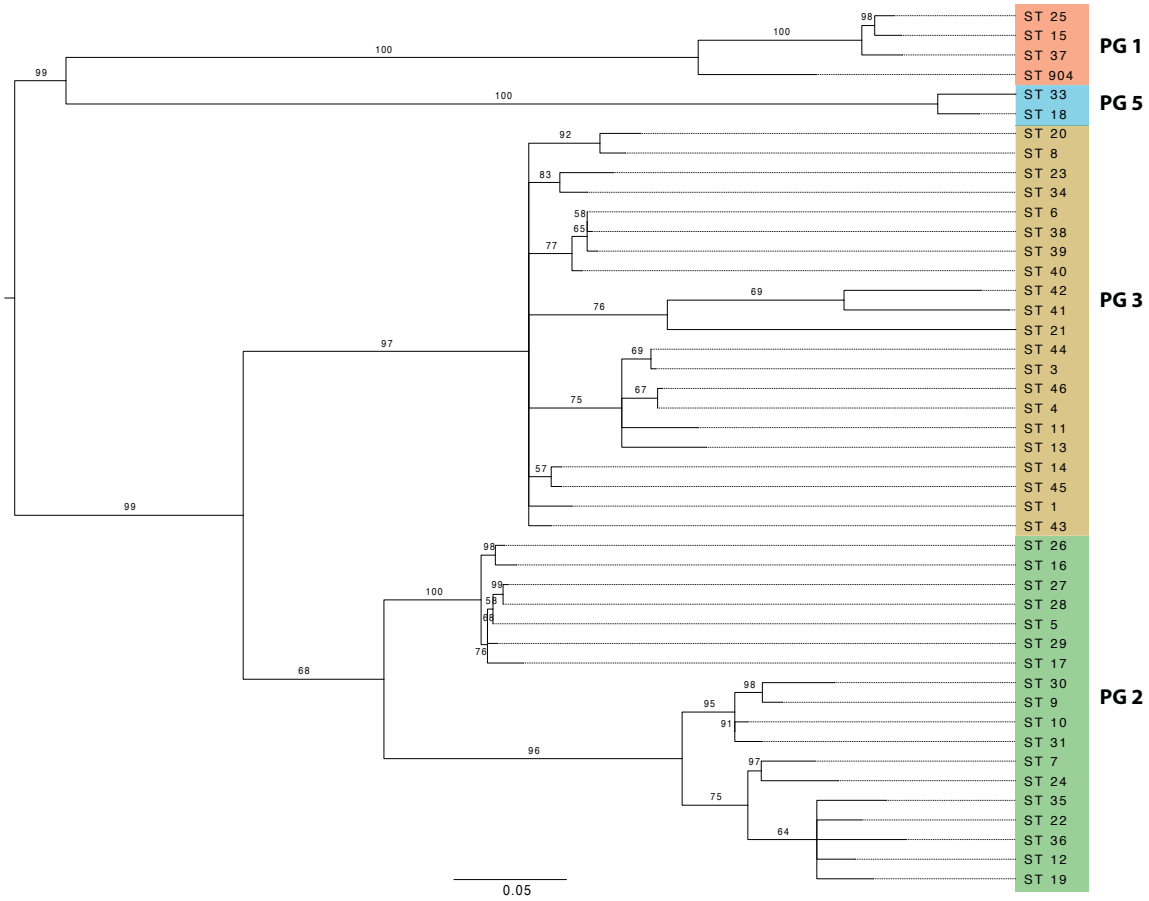

Figure 2.12: Maximum Likelihood phylogenetic reconstruction of all STs. The tree was constructed with TREEPUZZLE based on all ST's and using the HKY+G+I nucleotide substitution model. Bootstrap values are indicated at the nodes. PG = Phylogroup.

2.5.10.2 Bayesian Tree

Clonalframe was used to infer the evolutionary relationship by removing parts of the alignment that have undergone recombination. The Gelman-Rubin test, which assesses convergence among consecutive runs, showed that convergence was reached at 500,000 MCMC. The Bayesian tree was built after removing 170 bp of recombinant regions. The congruence graph between the two different phylogenetic trees based on the Maximum Likelihood and Bayesian approach, revealed that recombination had only a minimal impact on the phylogenetic reconstruction (Figure 2.13).

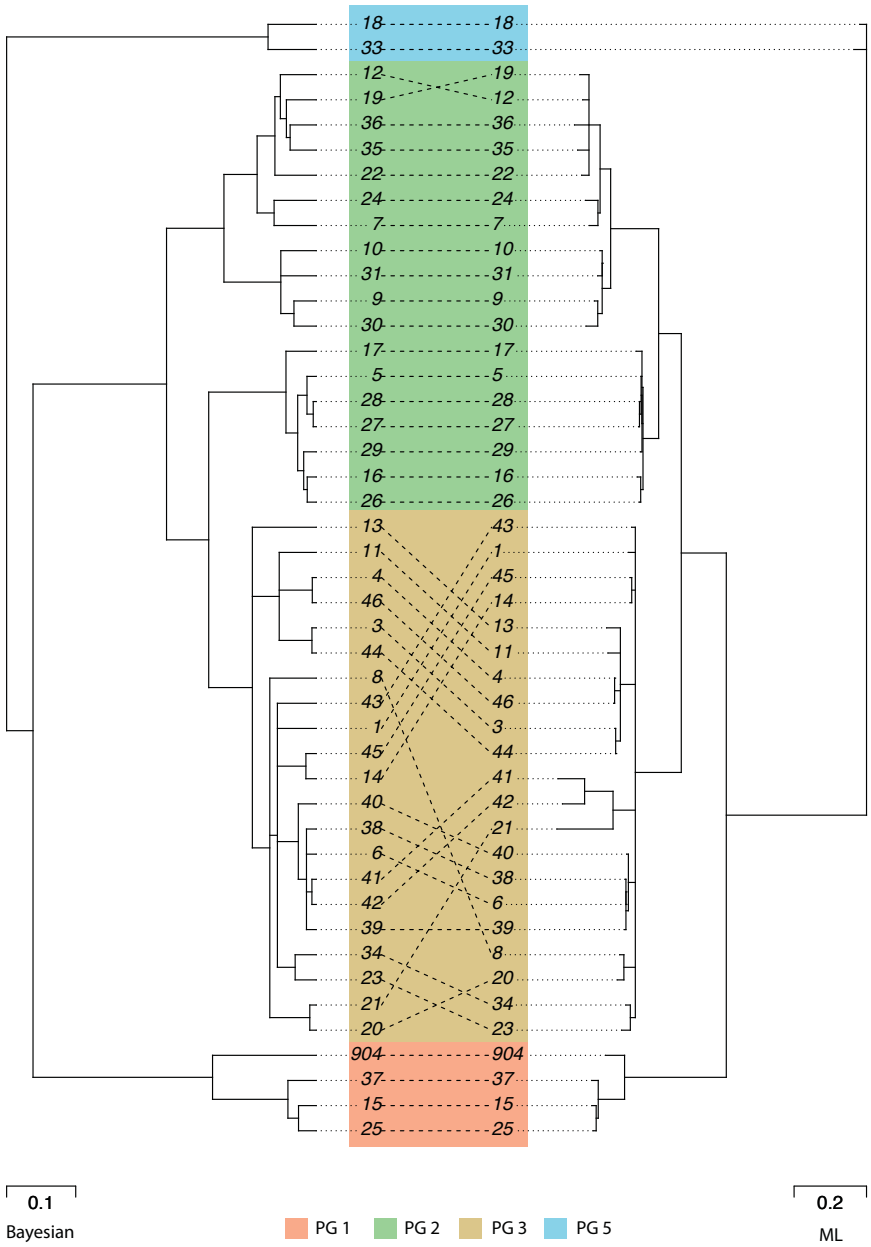

Figure 2.13: Bayesian tree (Clonalframe) vs. Maximum Likelihood tree (Treepuzzle). A comparison of the congruence between the two methods, the dotted lines connect the same node names. Numbers correspond to ST's.

2.5.10.3 Testing for phylogenetic congruence

Maximum likelihood trees were constructed using TREEPUZZLE and the corresponding best fitting nucleotide substitution model. Then the Shimodaira-Hasegawa (SH)-test was used to test for congruence between single gene trees and the concatenated tree. Two loci, *gyrB* and *gltA*, were significantly incongruent with every other gene tree, except for their own (Table 2.12). All loci, except *gyrB*, were congruent with the ML tree based on the concatenated dataset. The SH-test itself does not suggest which taxa might be responsible for the incongruence, but the quartet analysis done using TREEPUZZLE provides some information. All taxa were found to contribute equally to the the low number of partially resolved (1.6 %) and unresolved quartets (0.7%), which suggests the incongruence is not due to a small number of taxa in the dataset, but by an overall contribution from the sample.

Table 2.12: Shimodaira-Hasegawa test for phylogenetic congruence between ML trees. Values in bold indicate significance. Upper and lower triangle represent SH-tests based on different starting trees. The upper triangle shows results for using the tree of each gene in the first row as comparison, whereas the lower triangle the results for using the gene trees listed in the first column.

	concat ML	***gapA***	***gyrB***	***gltA***	***rpoD***
concat ML		0.060	**0.000**	**0.000**	**0.037**
gapA	0.274		**0.000**	**0.000**	**0.019**
gyrB	**0.001**	**0.000**		**0.000**	**0.000**
gltA	0.218	**0.042**	**0.000**		0.055
rpoD	0.148	**0.033**	**0.000**	**0.008**	

2.5.10.4 Phylogenetic network

A splits network is commonly used when reticulate events like recombination or horizontal gene transfer can obscure the phylogenetic relations. The splits network drawn using Splitstree, based on the neighbor-net method, suggests there is recombination occurring in PG2 and PG3 (Figure 2.14), as shown by the multiple edges. Recombination is confirmed for the entire dataset by the PHI test (P=2.656E-8) implemented in Splitstree.

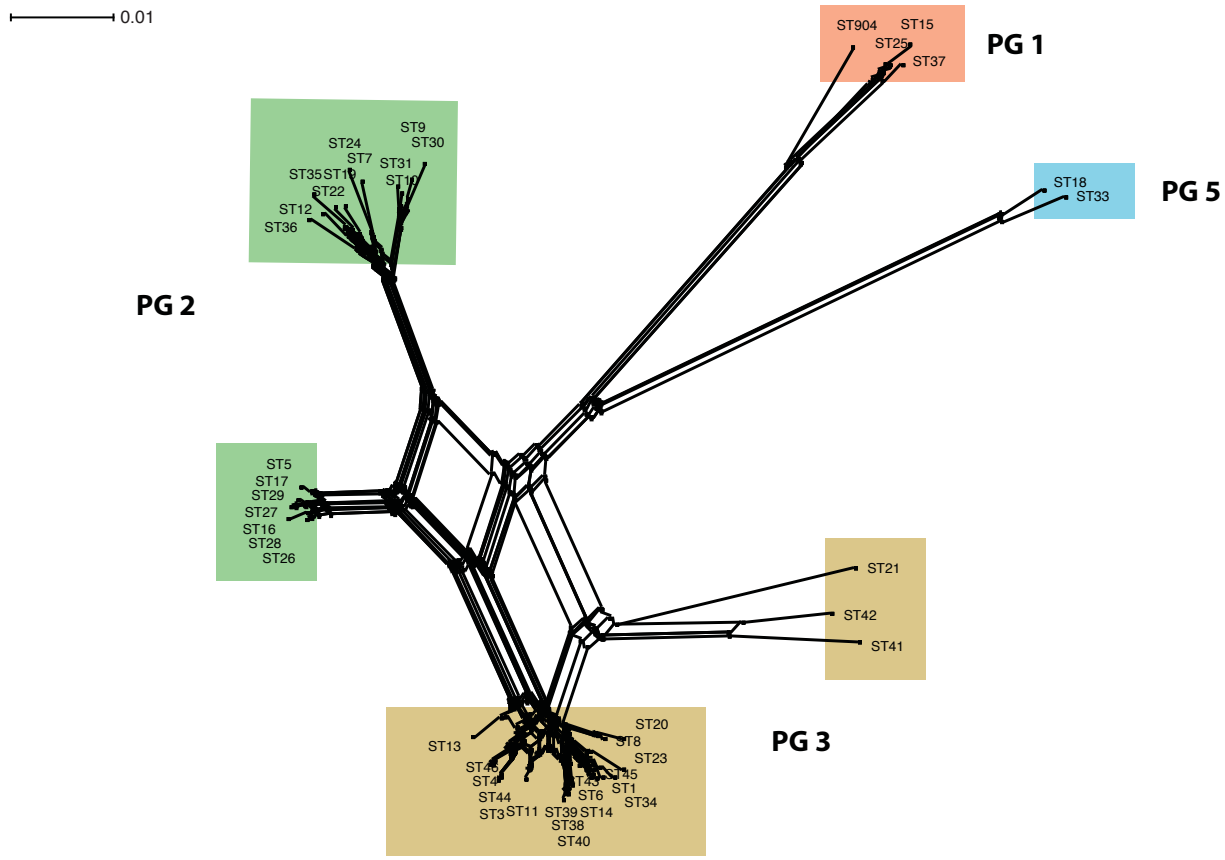

Figure 2.14: Splits network of 45 unique STs. Splits network is based on the neighbour-net method using Splitstree4.

2.5.11 Phylogenetic analysis and population structure of *P. syringae* associated with the phyllosphere in kiwifruit

As a reference, 37 *P. syringae* isolates from PAMDB.org database, representing the different monophyletic groups of *P. syringae*, as well as isolates from kiwifruit from Japan (n=10) (Tomihama *et al.*, 2016) and the US, France and NZ from previous years (n=10), were included to show the phylogenetic diversity among *P. syringae* of the phyllosphere of kiwifruit.

Phylogenetic trees were based on the concatenated alignment of the unique sequence types to improve readability of the tree; an overview of the frequency of each ST and the corresponding strain names can be found in (Table S 2.1, appendix).

A total of 13 different phylogroups were described for *P. syringae* recently based on a collection of agricultural and environmental strains, which allows for classification of strains of unknown pathogenicity (Berge *et al.*, 2014). Overall, most of the *P. syringae* isolates from NZ kiwifruit were grouping with three major monophyletic groups; PG1 (13%), PG2 (29%) and PG3 (56%), with only a few isolates grouping most closely with PG5 (2%) (Figure 2.15).

PG1 is made up of two clades: 1a containing tomato pathogens (pv. *tomato*) and clade 1b including kiwifruit (pv. *actinidiae*) and hazelnut (*P. avellanae*) (Bull *et al.*, 2011; Scortichini *et al.*, 2013). PG2 is the most

ubiquitous group with a wide range of host diversity varying from pea pathogens (pv. *pisi*) to tomato, stonefruit, brown rice and lilacs (pv. *syringae*), but it also includes many non-pathogenic strains from plant and environmental sources. PG3, the most prominent phylogroup in this study, holds many known pathogens of woody plants (e.g. pv. *savastanoi*, pv. *aesculi*, pv. *mori*) and bean pathogens (pv. *phaseolicola*, pv. *glycinea*). A couple of strains grouped with PG5, which contains known pathogens of *Brassicaceae* and coriander (*P. syringae* pv. *coriandricola*).

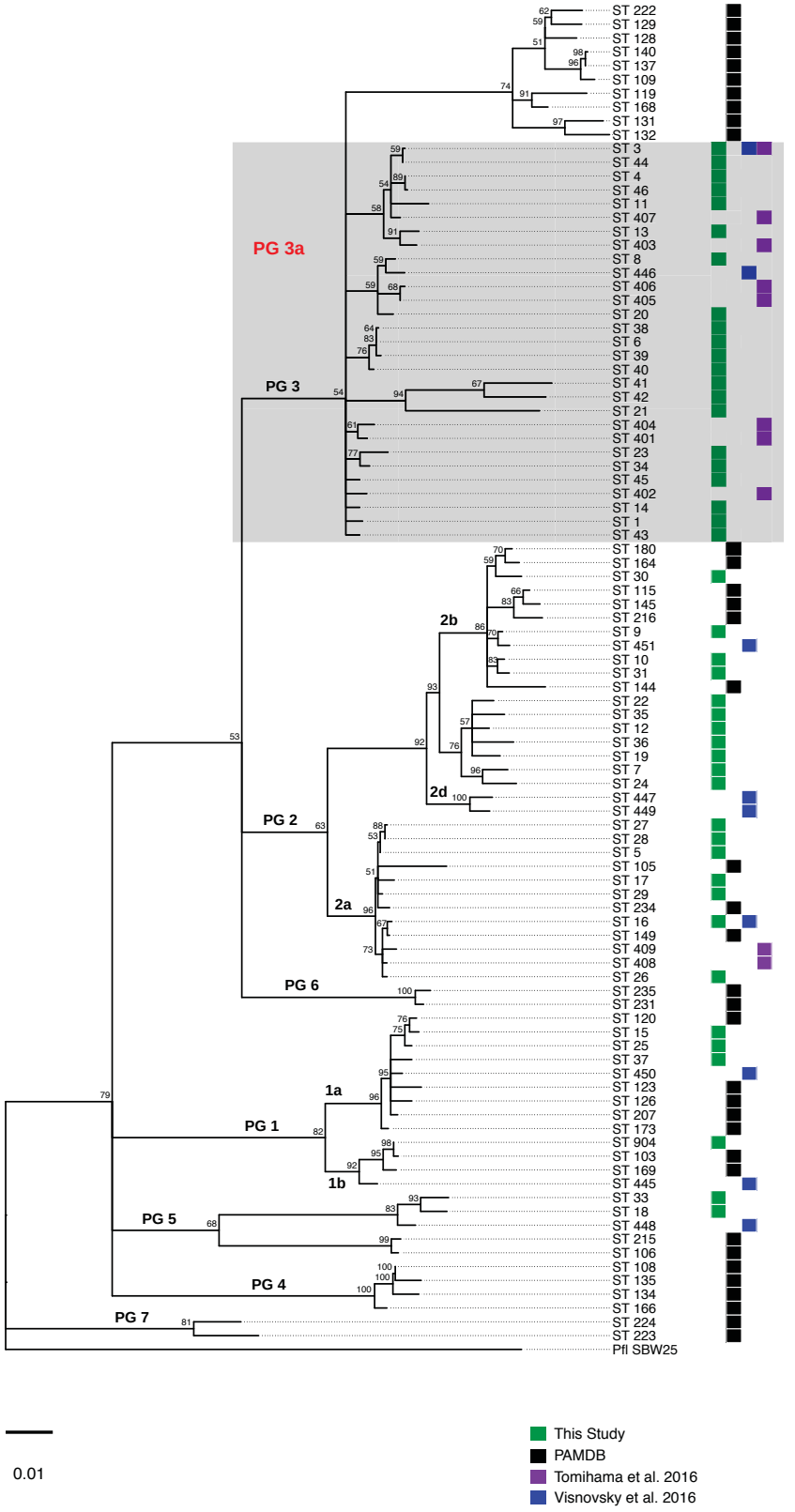

Figure 2.15: (Caption next page.)

Figure 2.15: (Previous page.) **ML tree based on the concatenated alignment (2010 bp) of *gapA, gyrB, gltA* and *rpoD*.** Tree reconstructed using TREEPUZZLE based on the Tamura-Nei model using 100,000 puzzling steps. Single representative sequences for each ST were used to improve readability (frequency of each ST and corresponding strain names listed in Table S 2.1, appendix). Values indicated at nodes are bootstrap values. The corresponding phylogroups (PG) are indicated, eg. PG1 = phylogroup 1 with clades 1a and 1b. Origin of isolates is illustrated in colour coded boxes, green = this study, black = PAMDB, blue = Visnovsky *et al.* (2016), purple = Tomihama *et al.* (2016).

2.5.11.1 Discovery of a new clade in PG3: clade 3a

Curiously, over half of all kiwifruit-associated strains (56%) group within a novel clade (referred to as clade 3a) of PG3. Strains from this new clade were also recovered during a small-scale sampling from Japanese kiwifruit (Tomihama *et al.*, 2016) and were previously isolated in NZ in 2010 and 2011 (Visnovsky *et al.*, 2016).

Interestingly, all strains from our study, which fall into PG3, form this new clade with strains having been isolated from every orchard. The inclusion of *gltA* sequences from a wide range of *P. syringae* isolated from wild and cultivated kiwifruit in China (McCann et al. unpublished) revealed that Chinese isolates are also present in clade 3a (Figure 2.16). No other *P. syringae* strains downloaded from the PAMDB database group with this particular clade, suggesting clade 3a is strongly and persistently associated with kiwifruit in different kiwifruit growing countries. The topologies of the single gene trees were significantly different from each other and the concatenated dataset (P<0.0.5, SH test), however the same classification for phylogroups was obtained for *gltA, rpoD* and *gapA* (with a couple of outliers); only *gyrB* grouped some isolates differently together (Figure 2.17).

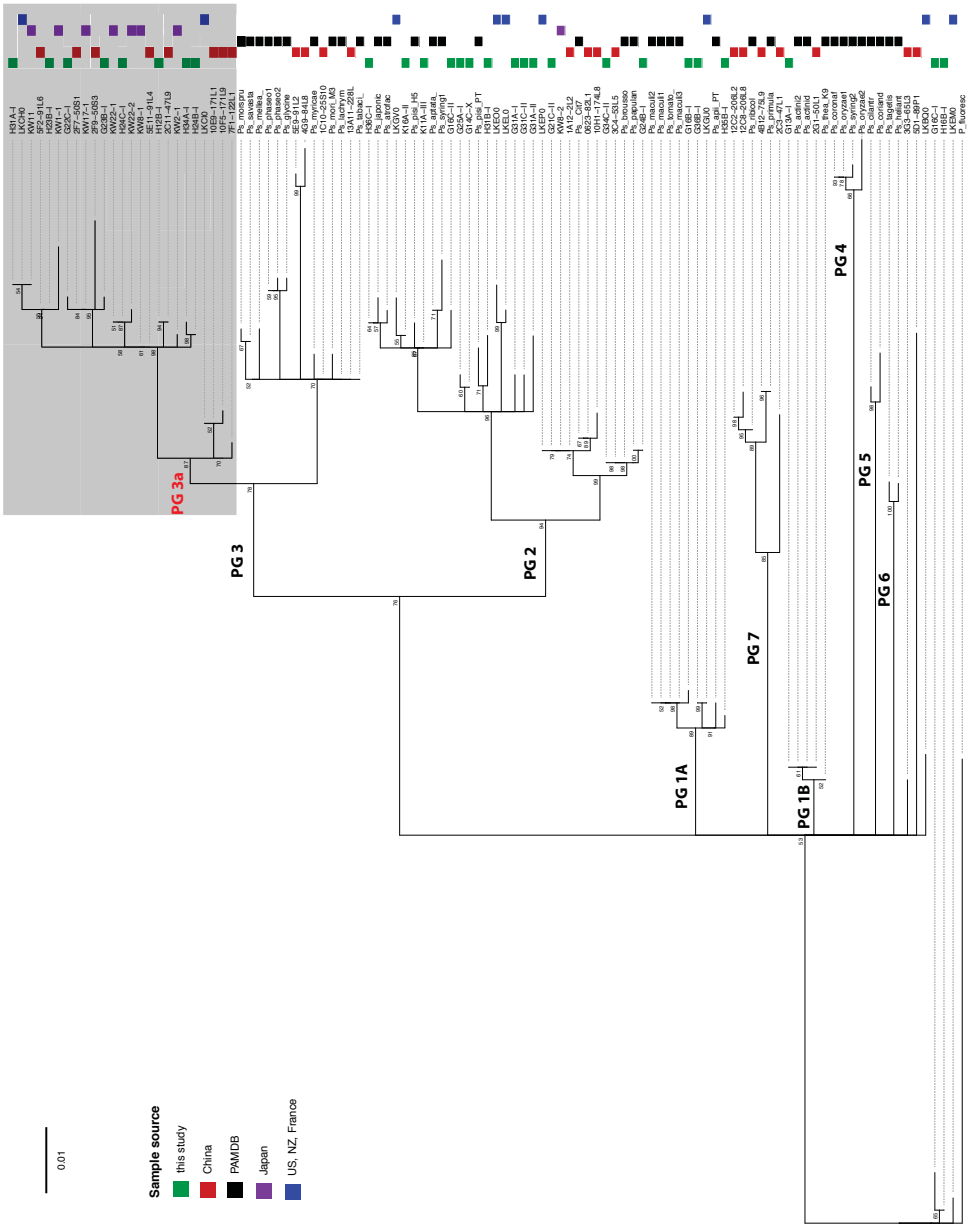

Figure 2.16: (Caption next page.)

Figure 2.16: (Previous page.) **Global ML tree reconstructed from *gltA* sequences highlighting the particularity of clade 3a, which includes kiwifruit isolates from NZ, China, Japan, the US and France.** The tree was built on a 529 bp alignment using TREEPUZZLE (HKY model; 100,000 puzzling steps), with *Pseudomonas fluorescens* SBW25 as outgroup. Values indicated at nodes are bootstrap values. The source of each isolate is highlighted in colour-coded boxes, green = this study, red = China, black = PAMDB, purple = Japan, blue = US, NZ and France.

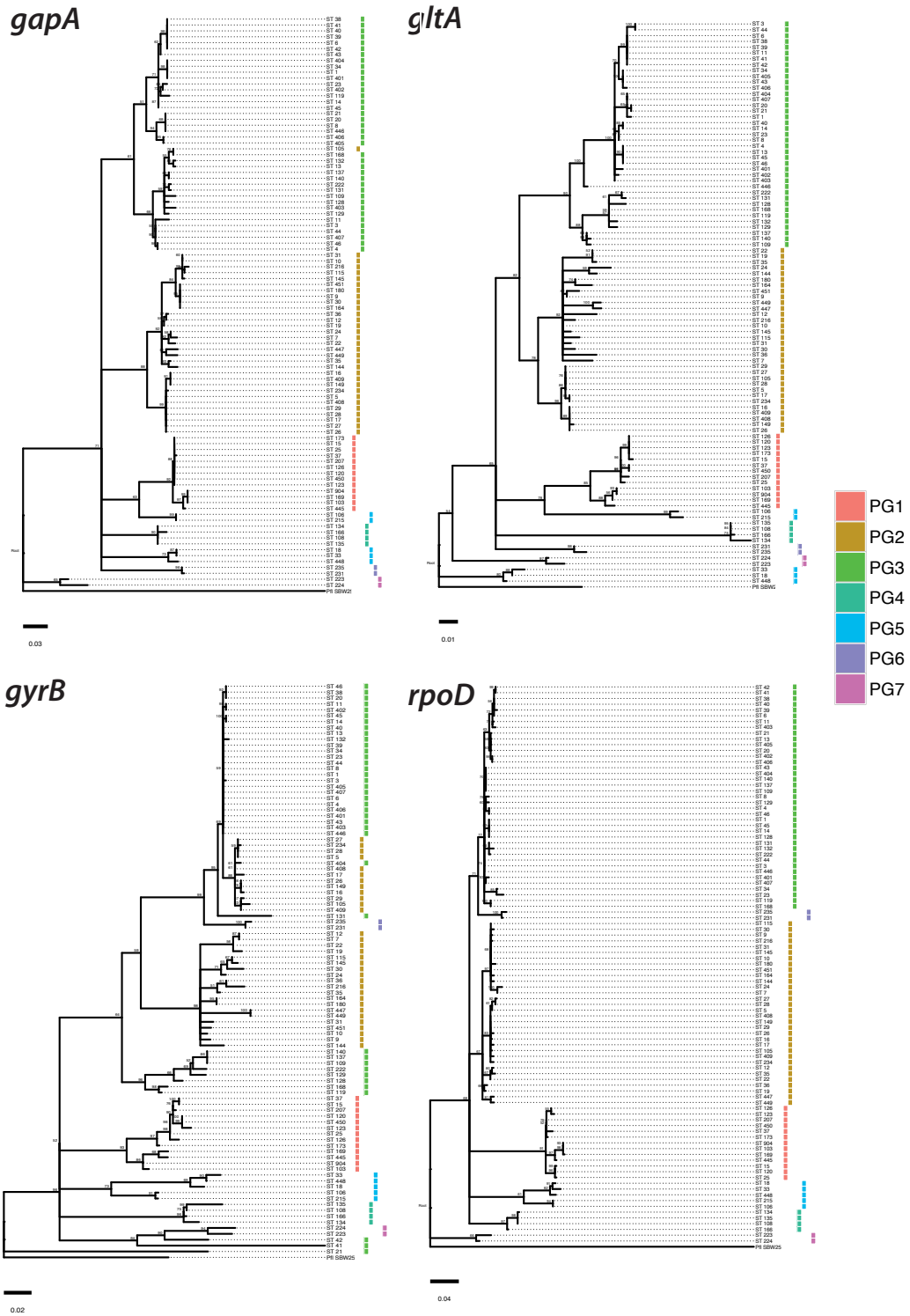

Figure 2.17: Caption next page

Figure 2.17: (Previous page.) **Maximum Likelihood trees based on single genes.** Each Maximum Likelihood tree is rooted on *Pseudomonas fluorescens* SBW25 and was reconstructed using TREEPUZZLE based on the Tamura-Nei model using 100,000 puzzling steps. Trees were built using single representatives of each unique ST to improve readability of the tree. Values indicated at nodes are bootstrap values. The corresponding phylogroup distinctions based on the concatenated ML tree are indicated with the coloured squares.

2.5.12 Biogeographic structure

PERMANOVA analysis was conducted for testing the impact of biogeographic structure on haplotype diversity (based on STs) or genetic diversity (based on pairwise genetic distance matrix). Unequality of sample size was taken into account by using the Type III SS (sum of squares) option.

2.5.12.1 Community composition

PERMANOVA analysis found there was neither influence of the cultivar (Figure 2.18 A, pseudo-F = 1.88, P >0.06) nor a temporal effect (Figure 2.18 C, pseudo-F = 0.74, P >0.5) on the community composition, however the infection status of an orchard did have a highly significant effect (Figure 2.18 B, pseudo-F 3.78, P <0.01). Testing for differences among sampling locations was highly significant (Pseudo-F 3.26, P <0.001) with the pairwise tests among orchards finding significant (P <0.05) or near-significant (P =0.05) differences in diversity between all orchard pairs, except between the infected gold and green orchard (P >0.1) (Figure 2.18 D).

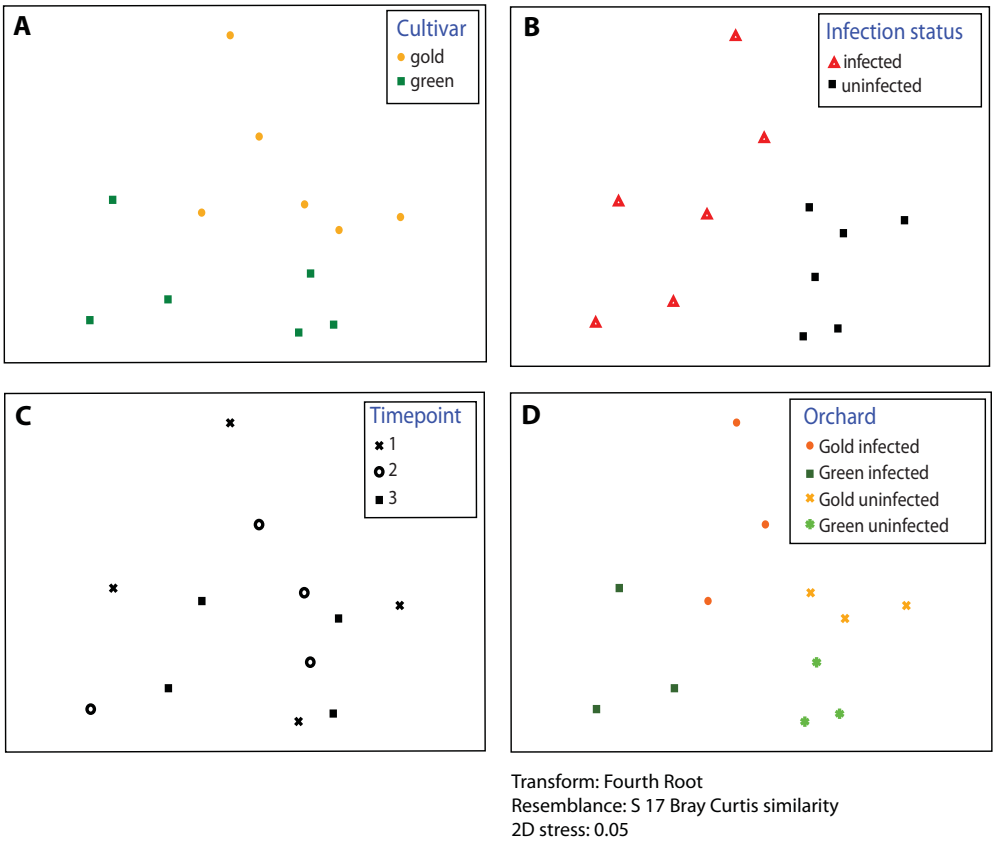

Figure 2.18: Multi Dimensional Scaling (MDS) plot based on sequence types. The low stress level (0.03) was reached after 100 permutations. STs are colour coded according to (A) cultivar, (B) infection status, (C) time point and (D) orchard of isolation.

2.5.12.2 Genetic diversity

Genetic diversity was determined based on the overall sample, hence also duplicate STs were considered. A significant effect of cultivar (Figure 2.19 A, Pseudo-F 11.96, P <0.001), infection status of the orchard (Figure 2.19 B, Pseudo-F 23.18, P <0.001) and time point of isolation (Figure 2.19 C, Pseudo-F 2.54, P <0.02) was also found based on variations in genetic

distance. Differences among overall genetic diversity of sampling locations were highly significant (Figure 2.19 D, Pseudo-F 14.98, P <0.001), with all pairwise tests among orchards being significantly different (P <0.001, except for infected green vs. uninfected gold (P >0.06)). The two-factor analysis found highly significant interaction between the factors orchardxtime (P <0.001) (Table 2.13). The three-factor nested analysis determined a highly significant effect of infection (Pseudo-F 5.52, P <0.002) and time point (pseudo-F 3.30, P <0.001) on the genetic diversity, but no effect of cultivar (pseudo-F 0.62, P >0.6) (Table 2.13).

Measuring differences in the "true" genetic diversity, the analysis was repeated on the dataset containing only the unique STs per location. Overall, the genetic diversity among sampled orchards was highly significant (Pseudo-F 5.99, P <0.0001), with the uninfected green orchard being significantly different from each other orchard (P <0.003). The cultivar (Pseudo-F 5.62, P <0.001) and the infection status of an orchard (Pseudo-F 11.72, P <0.001) also had a significant effect on genetic diversity, whereas time of isolation did not (Pseudo-F 1.10, P >0.34). Testing the effect of three factors, only the infection status has a relevant impact on genetic diversity (Table 2.13). MDS plots did not vary from the analysis based on the dataset containing duplicate STs.

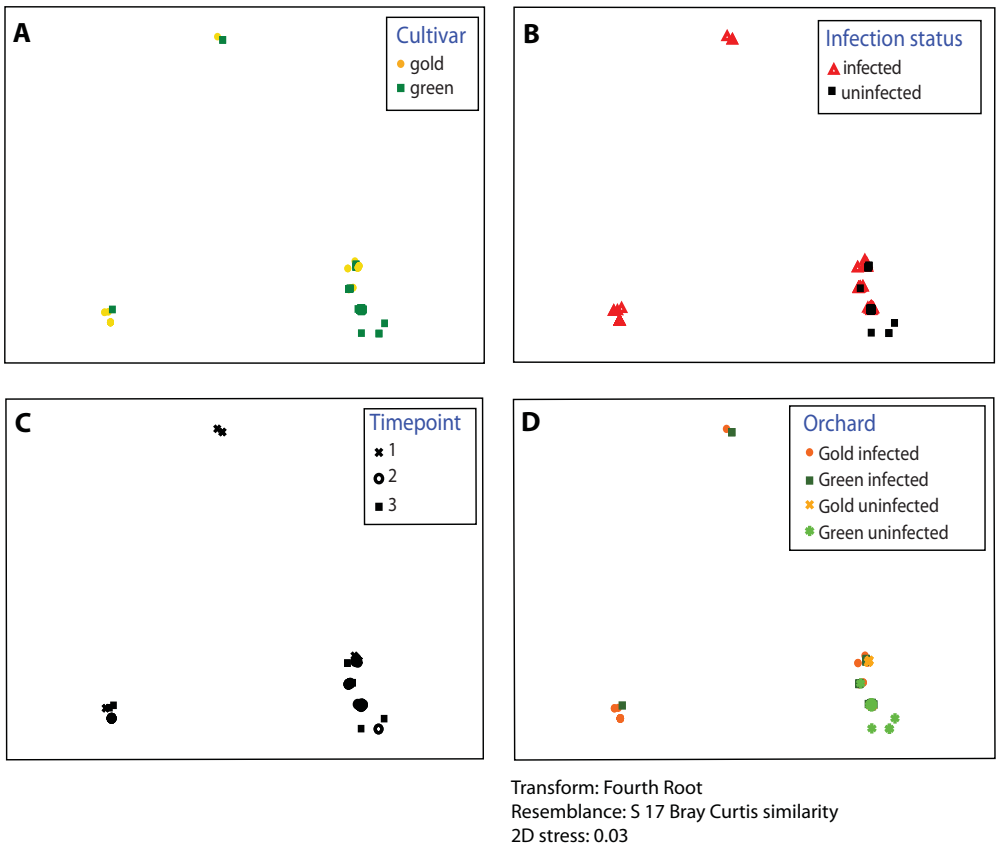

Figure 2.19: Multi Dimensional Scaling (MDS) plot based on the pairwise genetic distance. The low stress level (0.03) was reached after 100 permutations. STs are colour coded according to (A) cultivar, (B) infection status, (C) time point and (D) orchard of isolation.

Table 2.13: PERMANOVA results of 2-factor and 3-factor nested analysis for the variable genetic distance based on pairwise genetic distances.

All pairwise	**df**	**Sum of squares**	**Mean square**	**Pseudo F**	***P*(perm)**	**Unique perms**
orchard	3	7.39E-02	2.46E-02	17.707	0.0001	9934
time	2	1.26E-02	6.32E-03	4.5425	0.0006	9952
orchardxtime	6	2.42E-02	4.03E-03	2.8991	0.0005	9906
Residuals	136	0.18908	1.39E-03			
Total	147	0.29627				
cultivar	1	1.53E-02	1.53E-02	0.62472	0.6649	6
infection(cultivar)	2	4.92E-02	2.46E-02	5.519	0.0013	9951
time(infection(cultivar))	8	3.67E-02	4.59E-03	3.3021	0.0001	9916
Residuals	136	0.18908	1.39E-03			
Total	147	0.29627				
Unique sequences	**df**	**Sum of squares**	**Mean square**	**Pseudo F**	***P*(perm)**	**Unique perms**
Orchard	3	2.94E-02	9.81E-03	5.9737	0.0001	9917
Time	2	4.52E-03	2.26E-03	1.3773	0.2124	9948
orchardxtime	6	6.35E-03	1.06E-03	0.64479	0.8756	9908
Residuals	61	0.1002	1.64E-03			
Total	72	0.14088				
Cultivar	1	6.40E-03	6.40E-03	0.70802	0.6644	6
infection(cultivar)	2	1.88E-02	9.39E-03	6.4248	0.0045	9949
time(infection(cultivar))	8	1.16E-02	1.45E-03	0.88104	0.633	9901
Residuals	61	0.1002	1.64E-03			
Total	72	0.14088				

2.6 Discussion

Traditionally, most of the research looking into the population structure of *Pseudomonas syringae* has been biased towards pathogenic strains isolated from their respective hosts (e.g. Sarkar & Guttman, 2004; Hwang *et al.*, 2005), although the pool of *P. syringae* residing in the environment (aquatic and terrestrial) has been the focus of more recent attention (e.g. Monteil *et al.*, 2012; Morris *et al.*, 2010). This environmental pool of isolates is more diverse than what has been recovered on diseased crops so far (Berge *et al.*, 2014) and evidence was found that some of these isolates are pathogenic on some tested crop plants (e.g. Bartoli *et al.*, 2015).

Only a few recent studies have explored the diversity of non-pathogenic *P. syringae* isolated from infected plants (Beiki *et al.*, 2016; Hall *et al.*, 2016; Humphrey *et al.*, 2014). Knowledge of the ecological and genetic structure of commensal strains will enable understanding of how transitions from a commensal to a pathogenic life style might take place and what factors promote the emergence of virulence and resistance. These commensal strains might present a diverse gene pool, which could potentially provide a source of genetic material like virulence or resistance genes awaiting to be acquired by pathogens, as was previously shown for *P. syringae* pv. *tomato* (Monteil *et al.*, 2013).

My work has provided new perspectives on the diversity and genetic structure of *P. syringae* isolated from infected and non-infected kiwifruit vines during an outbreak of bacterial canker of kiwifruit (*Pseudomonas*

syringae pv. *actinidiae*). These data help to understand the microevolutionary processes shaping the population structure of this ubiquitous plant pathogen.

2.6.1 Higher diversity in infected orchards

The results suggest that infected orchards display a higher amount of haplotype diversity compared to non-infected orchards (2.18, 2.19). However, it is important to keep in mind that the unequal sampling size could easily impact these measures. Interestingly, a study comparing the diversity of *Pseudomonas* sp. in herbivore-damaged bitter cress (*Cardamine cordifolia, Brassicaceae*) leaves, compared to healthy undamaged leaves, found a higher abundance and diversity in the damaged leaves (Humphrey *et al.*, 2014). This could suggest that the innate immune response of a plant to an invader possibly has a cross-talk effect by positively influencing the simultaneous or secondary colonization by microbes e.g. via increased plant susceptibility or nutrient availability by cell leaching.

2.6.2 Low levels of recombination among *P. syringae* in the phyllosphere

Horizontal gene transfer can disrupt the expected patterns of inheritance of organisms, hence it is important to consider for correct inference based on nucleotide sequences. Based on the full dataset (all sequences), neither intragenic nor intergenic recombination was detected, with mutation being up

to 5 times more likely to contribute to genetic variation than recombination. However, some recombination was found at the phylogroup level (PG2 and PG3), but not for PG1, when calculating recombination rates for individual lineages (Table 2.9). However, most strains in PG1 showed 100% nucleotide identity, as these were all *Psa* isolates (ST904), which affects calculating rates of recombination.

Including the global *P. syringae* strains from the PAMDB database revealed that, in addition to PG2 and PG3, recombination was also occurring in PG1 (now including more divergent strains in addition to *Psa*), but this was limited to a single gene (Table 2.11). This low level of recombination confirms what has been found previously for isolates from mostly diseased hosts (Sarkar & Guttman, 2004): a mostly clonal population structure with some recombination within phylogroups, but not across different phylogroups.

Comparing the Maximum Likelihood tree to the Bayesian phylogenetic reconstruction using Clonalframe, which removed the chunk of alignment which has undergone recombination, reveals that only within PG3, is incongruence evident, but this part of the tree is not well resolved (low bootstrap values). As stated in the previous paragraph, recombination occured among PG3, but not among all lineages (Figure 2.13).

2.6.3 Commensal *P. syringae* classified in subpopulations

The symbiotic relationship between PG3a and their associated host can be described as commensal, with one type benefitting and the other being largely unaffected. The population of *P. syringae* in the kiwifruit phyllosphere is apparently commensal, as no evidence of disease was observed in the field and in the laboratory (see section 3.4.3). Hence, the commensal strains are also referred to as non-pathogenic isolates of kiwifruit in the scope of this thesis. However, it is necessary to keep in mind that some of these strains may be able to infect other hosts than kiwifruit, but that this host would need to be identified by performing pathogenicity tests on a broad range of plants.

2.6.3.1 ST-based analysis

From the total of 45 unique ST's, only three had been previously isolated with all others found to be novel. This highlights the fact that not much attention has been paid to non-pathogenic commensals in an agricultural environment and that the diversity of *P. syringae* is still understudied. No clear temporal pattern could be observed for STs and clonal complexes (Figure 2.8), which is also evident in not finding any temporal influence on genetic diversity using PERMANOVA. Determining the ancestry of the polymorphisms revealed evidently the same pattern as the phylogenetic analysis: three clusters were detected, which were identical to the three major phylogroups.

2.6.3.2 Nucleotide-based analysis

The commensal *P. syringae* isolated from kiwifruit can be distinguished into three prominent phylogroups (PG1, PG2, PG3) and three strains which group with PG5. This amount of diversity in a cultivated environment is not surprising, compared to what other studies have found in an agricultural setting. Comparably, two clades of endophytic *P. syringae* pv. *syringae* were recovered from symptomatic grapevines in Australia, with a mix of pathogenic and non-pathogenic isolates clustering together (Hall *et al.*, 2016). A sampling done in citrus orchards suffering from citrus blast caused by *P. syringae* pv. *syringae* found isolates associated with PG2, PG7 and an unknown clade (Beiki *et al.*, 2016). In a non-agricultural environment native bittercress (*Cardamine cordifolia*) leaves revealed *P. syringae* isolates from PG1, PG5, PG7, PG10 and PG13 (Humphrey *et al.*, 2014). A range of *P. syringae* were collected from wild and cultivated *Arabidopsis thaliana*, where resident strains from wild plants showed less diversity and all grouped within PG2c, whereas non-resident strains grouped with other clades within PG2 (Kniskern *et al.*, 2011). These findings suggest there is a vast diversity of commensal isolates waiting to be still discovered, independent of whether a *P. syringae* infection is already taking place, or the plant is seemingly healthy. This also suggests that there is potential for interactions among these co-residing strains and further work is required to unravel the level and intensity of these interactions.

2.6.4 Global kiwifruit *P. syringae* form a new clade in Phylogroup 3

More than 50% of the sampled isolates belonged to a new group (PG3a) of *P. syringae* colonizing kiwifruit plants in New Zealand. However this new group does not appear to be restricted to NZ, as strains in PG3a were also isolated from kiwifruit plants in Japan and France, as well as from kiwifruit sampled in New Zealand in 1991 (Figure 2.15) (Tomihama *et al.*, 2016; Visnovsky *et al.*, 2016). There is further evidence for the presence of this clade of commensals in China, where a number of *P. syringae* samples were collected from wild and cultivated kiwifruit (McCann et al., pers. comm., Figure 2.16). Members of PG3a form the majority of kiwifruit *P. syringae* in this study, with isolates collected from every single orchard. None of the strains recorded in PAMDB group with PG3a, which suggests that this new group of *P. syringae* forms a persistent association with kiwifruit on a national and global level.

2.6.5 Genotypic diversity affected by biogeographical factors

The higher genetic diversity of *P. syringae* among infected orchards might be explained by the beneficial effect of the presence of *Psa*. *Psa*'s suppression of plant immune responses and migration into the apoplast and vascular tissues could allow colonization by other bacteria. Plants experiencing either fungal or bacterial infection display changes in

composition and abundance of bacteria. Such a shift of population dynamics was shown during a *Pectobacterium atrosepticum* infection in potato tubers, where post-infection commensal *Enterobacter* and *Pseudomonas* thrived and did not hinder the pathogenesis process (Kõiv *et al.*, 2015). Cucumber and Japanese spindle plants suffering from a powdery mildew infection had increases in bacterial population size, diversity and species richness (Suda *et al.*, 2009). The same was observed for bitter cress leaves attacked by herbivores (Humphrey *et al.*, 2014). There were cultivar-specific differences in *P. syringae* genetic diversity. Host species identity has a greater effect on bacterial community composition than other factors like age, time or geographical location (Laforest-Lapointe *et al.*, 2016), and the relatedness of trees even corresponded with the phylogeny of the bacterial community (Redford *et al.*, 2010). Others showed a host genotype-dependent colonization in different cultivars of crop species, (e.g. Adams & Kloepper, 2002; Van Overbeek & Van Elsas, 2008; Wagner *et al.*, 2016). However, the observed differences in phyllosphere diversity may also be attributed to variation in environmental factors affecting the different orchards (Whipps *et al.*, 2008), as well as orchard management practices. Different fertilizer and spray regimes (copper, antibiotics, biological agents) are employed by growers to prevent or manage *Psa* infection throughout the growing season (http://www.kvh.org.nz/vdb/document/99346). These practices may have selected for copper and streptomycin resistance in *Psa* in NZ and elsewhere (Colombi *et al.*, 2017; Han *et al.*, 2003).

Chapter 3

Ecological interactions of commensal *P. syringae* with the pathogen *P. syringae* pv. *actinidiae*

3.1 Introduction

3.1.1 Maintenance of variation

Genetic variation can be introduced via mutation, recombination and migration, with the fate of variants being determined by genetic drift and selection. The maintenance of variation in natural populations and its

underlying mechanisms are a major focus of evolutionary ecology (Fierer & Lennon, 2011; Hansen *et al.*, 2007; Huang *et al.*, 2011; Pastar *et al.*, 2013; Tong *et al.*, 2007). From an ecological perspective understanding the forces and mechanisms responsible for maintaining variation is important, but the dynamic and variable conditions in the natural environment make this much more difficult than, for example, in a controlled environment like the laboratory that allows constant surveillance of conditions (Hibbing *et al.*, 2010; Rainey *et al.*, 2000).

3.1.2 Maintenance of genetic polymorphism in natural bacterial populations

Bacteria form complex multispecies communities comprised of numerous populations (Rainey, 2005). A population consists of a collection of individual genotypes; for bacterial populations, these typically constitute individuals from the same species. Interactions occur between individuals of the same genotype, between different genotypes and between different species, with a wide range of consequences (Stubbendieck *et al.*, 2016). An understanding of these interactions is necessary for discerning what drives co-existence and how patterns of diversity are shaped at the species level (Weber & Strauss, 2016).

Bacteria play an essential role in the life history of plants. Bacterial lifestyles range from endophytic to epiphytic, pathogenic to commensal, and bacteria form strong associations with plants. Ecological populations of

bacteria can be considered as groups of closely related coexisting types that require similar resources (Cordero *et al.*, 2012).

The presence of multiple different bacterial types enables interactions. Of interest is whether two or more organisms can occupy the same niche. The competitive-exclusion principle states that two co-existing competing genotypes cannot stably co-inhabit the same ecological niche when only a single limiting resource is available (Gause, 1934; Hardin, 1960). Hence when two types are commonly found in the same environment, either there is competitive interaction or a stable polymorphism by occupation of different ecological niches. Competition between types can occur via the production of allelopathic factors like toxins or siderophores, with one type ultimately being driven extinct (Blanchard *et al.*, 2014; Cornforth & Foster, 2013; Hibbing *et al.*, 2010; Stacy *et al.*, 2012). Maintenance of diversity on the other hand depends on the availability of niches, which is coupled to the structure of the environment. In a spatially structured environment, selection of variants can lead to occupation of different ecological niches (Haubold *et al.*, 1998; Hol *et al.*, 2015; Rainey & Travisano, 1998). For example, in a static liquid culture, *Pseudomonas fluorescens* diversifies into genotypes adapted to the air-liquid interface by producing a biofilm, whereas the ancestor proliferates in the nutrient broth (Rainey & Travisano, 1998).

3.1.3 Bacterial interactions influencing the outcome of plant disease

When bacteria colonize and infect plants, they are typically not the sole residents occupying the leaf surface. The phyllosphere and apoplast is home to a wide diversity of microorganisms and the interaction between the pathogen and co-inhabiting strains can influence the infection process (Burmølle *et al.*, 2006; Lindow & Brandl, 2003; Marchi *et al.*, 2006; Marcon *et al.*, 2002). Depending on the nature of the interaction, the outcome of pathogenicity is altered, which in turn is important for disease epidemiology. Synergistic interactions between co-occurring bacteria can increase disease severity. For example, the severity of olive knot canker caused by *P. savastanoi* pv. *savastanoi* was amplified when non-pathogenic *Pantoea agglomerans*, *Erwinia toletana* or *E. olea* were present (Buonaurio *et al.*, 2015; Marchi *et al.*, 2006). Although these strains are non-pathogenic, they are involved in the production of various compounds (e.g. IAA) which aid *P. savastanoi* pv. *savastanoi* during the infection process. Another example is tomato pith necrosis, where severity is influenced by co-infection of species duos, including *Pseudomonas cichorii, P. corrugata - P. marginalis* and *P. corrugata - P. mediterranea* (Kudela *et al.*, 2010; Moura *et al.*, 2004). The interaction among co-residing strains can also have a positive impact by suppressing disease symptoms or development, as was shown for the non-pathogenic *Ralstonia solanacea*, which suppressed bacterial wilt in eggplant (Nakahara *et al.*, 2016; Ogawa *et al.*, 2012).

3.1.4 Interactions among *P. syringae*

The ecological dynamics and factors shaping the coexistence of *P. syringae* have not received much attention. In the late 1960s and 1970s mixed inoculations of pathogenic, non-pathogenic or avirulent *P. syringae* were conducted to investigate whether intraspecific interactions were beneficial, neutral or antagonistic for the different variants (Klement & Lovrekovich, 1961; Omer & Wood, 1969; Young, 1974). Non-pathogenic strains tend to benefit from the presence of a pathogenic host strain, as they reach larger population sizes compared to when grown in isolation. The increased growth supposedly results from an increase in cell-membrane permeability and subsequent nutrient and water leakage induced by the pathogenic strain (Rufián *et al.*, 2017; Young, 1974). These effects were also dependent on the inoculum ratios (Macho *et al.*, 2007; Young, 1974). Others have shown that avirulent strains of *P. syringae* activate plant defence mechanisms and consequently restrict the growth of the co-inoculated virulent strain (Klement & Lovrekovich, 1961; Omer & Wood, 1969).

In the late 1980s, antagonistic interactions in *P. syringae* resulting in pre-emptive exclusion of one strain were used to drive the development of natural biocontrol agents. This pioneering work by Lindow (1986) was done by constructing ice-nucleation negative (Ice^-) *P. syringae* strains and resulted in the release of the first genetically modified bacterial strains into the wild. These Ice^- bacteria engaged in antagonistic interactions with Ice^+ isolates and reduced the severity of frost damage to plants in an agricultural environment (Lindow, 1991; Lindow & Panopoulos, 1988; Wilson & Lindow, 1994b).

In 2001, the efficiency of various non-pathogenic *P. syringae* pathovars as biocontrol agents to treat *P. syringae* pv. *glycinea* infection in soybean fields was tested; the presence of a non-pathogenic *P. syringae* strain on the leaf successfully reduced the epiphytic population size of the pathogen to 0.12% of its individual size (Völksch & May, 2001). Recently it was shown by Bartoli *et al.* (2015) that *P. syringae* strains from an aquatic environmental origin can grow in kiwifruit. Interestingly, these strains were able to reach almost the same population size as *Psa* in the apoplastic space, but they did cause symptoms of a varying degree in kiwifruit and a variety of other tested woody and herbaceous plant species. Confocal microscopy was used to observe the colonization *in planta* when mixed inoculations were applied; the non-pathogenic isolate was seen to colonize larger areas in the presence of pathogenic bacteria (Rufián *et al.*, 2017). Together with measures of evolutionary parameters, characterisation of these ecological interactions is an important means to gain insight into the emergence of bacterial diseases.

The kiwifruit canker pathogen *P. syringae* pv. *actinidiae* was found to co-inhabit the same geographical space as a variety of non-pathogenic commensal strains on the kiwifruit leaf surface. In fact a diverse range of commensal *P. syringae* were discovered on healthy hosts and even in heavily infected kiwifruit plants. The results from Chapter 2 - investigating the diversity of *P. syringae* inhabiting the kiwifruit phyllosphere using MLST - suggest the presence of a clade of kiwifruit resident strains within Phylogroup 3 (clade 3a) that form a persistent association with kiwifruit. Having investigated the evolutionary forces driving the diversity and

population structure within *P. syringae*, I now focus on the ecological interactions shaping the co-existence within *P. syringae* associated with kiwifruit.

3.2 Aims

In this chapter I investigate the dynamics of *P. syringae* ecological interactions by testing for stable co-existence of two closely related lineages (pathogenic vs. non-pathogenic) of *P. syringae* isolated from the same leaf. Experiments explore interactions both *in vitro* and *in planta*. The aim was to:

1. Determine whether either type can invade from rare (testing stability of an apparent polymorphism)
2. Determine the nature and magnitude of the interactions between lineages (e.g. beneficial, detrimental, neutral) when grown together
3. Test whether early colonization of one type impacts on the ability of the other to establish a viable phyllosphere population

3.3 Material & Methods

3.3.1 Strains and culture conditions

Pseudomonas strains were cultured in King's B medium at 28 °C; *E. coli* was cultured in Luria Bertani medium at 37 °C. Liquid overnight cultures were inoculated from single colonies and shaken at 250 rpm for 16 h. The antibiotics kanamycin (kan) and nitrofurantoin (nf) were used at a concentration of 50 μg/mL. A list of all bacterial strains used in this study can be found in Table 3.1. Kanamycin resistant *Psa* NZ54 and *Psa* NZ13 Δ*hrcC* were employed in all *in vitro* and *in planta* experiments.

3.3.2 Tn*5* transposon mutagenesis

Triparental matings were performed to introduce a kanamycin resistant Tn*5* transposon into *Psa* NZ54 and *Psa* V13 Δ*hrcC*. *E. coli* S17-1 Tn*5hah Sgid1* (donor), *E. coli* (*pRK2013*) (helper plasmid) and *Psa* NZ54/*Psa* NZ13 Δ*hrcC* (recipient) were grown in shaken liquid media overnight. 200 μL of donor and helper and 2 mL of recipient were individually washed, pelleted and combined in 30 μL 10 mM $MgCl_2$. The mixture was plated on a pre-warmed LB agar plate and incubated at 28 °C for 24 h. The cells were scraped off and resuspended in 1 mL 10 mM $MgCl_2$ and plated on KB+Kan+Nf plates. Mutants were screened and a few clones were stocked in the freezer. Growth of four different clones was compared to the wild

type (WT) by measuring OD_{600} over the course of three days in a 96-well plate on a shaker; the experiment was repeated twice. The insertion site was identified using Polymerase Chain Reaction as described in the next section.

Table 3.1: List of bacterial strains. Original ID refers to the ID assigned upon sample collection (as used in Chapter 2); new names were assigned for further experiments to avoid confusion over names.

Strain	**Genotype description**	**Reference**	**Original ID**
P. syringae G33C	*P. syringae* WT strain (PG 3)	This study	G33C-I
Psa NZ54	*P. syringae* pv. *actinidiae* (PG1)	This study	G33C-II
Psa NZ54::Tn5	*P. syringae* pv. *actinidiae* carrying the kanamycin resistance cassette	This study	
Psa NZ13 Δ*hrcC*	*P. syringae* pv. *actinidiae* NZ13 knockout mutant with deletion of hrcC	Colombi (2017)	
Psa NZ13 Δ*hrcC*::Tn5	*P. syringae* pv. *actinidiae* NZ13 hrcC knockout mutant carrying the kanamycin resistance cassette	This study	
E. coli S17-1 Tn*5hah Sgid1*	Donor for triparental conjugation	Zhang *et al.* (2015)	
E. coli pRK2013	Helper plasmid for triparental conjugation	Ditta *et al.* (1980)	

3.3.3 Insertion site identification

The PCR template was prepared by centrifuging 1 mL of overnight culture (2 min, 13,000 g) and re-suspending the cells in 400 μL de-ionised H_2O. PCR amplification (list of primers: Table 3.2) was performed with a BIO-RAD T100 Thermal Cycler as follows: 3 μL of washed cells were used as template in the following PCR in a volume of 20 μL: 2.5 μL 10x Buffer, 0.8 μL 50 mM $MgCl_2$, 1 μL 10 mM dNTPs, 2 μL primer TnphoAII (10 pmol/μL), and primer CEKG 2A:2B:2C (equal mix) (10 pmol/μL), 0.5 μL 5 U Taq. The PCR was run with the following thermal cycle: initial denaturation at 94 °C for 10 min followed by 5 cycles of amplification with denaturation at 94 °C for 30 s, annealing at 42 °C for 30 s (decreasing by 1 °C after every cycle) and elongation at 72 °C for 3 min. This was followed by another 24 cycles at 94 °C for 30 s, 65 °C for 30 s and elongation at 72 °C for 3 min.

For the second PCR, 2 μL of diluted PCR1 product was used as template together with 2.5 μL 10x Buffer, 0.8 μL 50 mM $MgCl_2$, 1 μL 10 mM dNTPs, 2 μL each of primer Hah-1 (10 pmol/μL) and primer CEKG 4 (10 pmol/μL), 0.5 μL 5U Taq and run using the following thermal cycle: initial denaturation at 94 °C for 10 min, followed by 30 cycles of 94 °C for 30 s, 65 °C for 30 s and 72 °C for 3 min. Samples were purified using the Exo-CIP method detailed below and sequenced by Macrogen Inc. (South Korea).

Table 3.2: List of primers

Primer name	Sequence	Target region
Tn*pho*A-II	GTGCAGTAATATCGCCCTGAGCA	IS-Ω-Km/hah
CEKG 2A	GGCCACGCGTCGACTAGTACNNNNNNNNNNAGAG	non-specific
CEKG 2B	GGCCACGCGTCGACTAGTACNNNNNNNNNNACGCC	non-specific
CEKG 2C	GGCCACGCGTCGACTAGTACNNNNNNNNNNGATAT	non-specific
Hah-1	ATCCCCCTGGATGGAAAACGG	IS-Ω-Km/hah
CEKG4	GGCCACGCGTCGACTAGTAC	5' end of CEKG 2A, B & C

3.3.3.1 Exo-CIP purification

The PCR product was purified using an enzymatic reaction by adding 4 units of CIP (Calf Intestinal Phosphatase, NEB) and Exo I (Exonuclease I, NEB). The reaction was incubated at 37 °C for 60 min. The enzymes were inactivated at 85 °C for 15 min and the final purified product was stored at 4 °C.

3.3.4 Competition experiments

The stability of the polymorphism of *Psa* NZ54 (PG1) vs. *P. syringae* G33C (clade 3a, PG3) was tested *in vitro* and *in planta*. As a selective marker, *Psa* NZ54 and *Psa* NZ13 Δ*hrcC* were marked with a kanamycin resistance cassette using Tn*5* transposon mutagenesis. The two strains were distinguished by plating on KB supplemented with kanamycin (50

μg/mL) and M9 agar plates, with *P. syringae* G33C growing significantly faster on M9.

In vitro. Experiments were performed using rich (King's B) and minimal (M9) media in a shaken and static environment. Competition experiments were performed in 1:1, 1:10 and 10:1 ratios for each of the four assay conditions. Liquid overnight cultures of each strain were established from single colony inoculations. 30 mL vials containing 4 mL of the appropriate media were inoculated with each strain, adjusted to a founding density of either OD_{600} 0.006 ($5x10^6$ cfu/mL) or OD_{600} 0.0004 ($4x10^4$ cfu/mL). Control vials were inoculated with a single strain, adjusted to OD_{600} 0.006 ($5x10^6$ cfu/mL). Cultures were incubated at 28 °C and grown over a period of 72 h, either still or shaken at 250 rpm. Bacterial density was calculated after 0, 24, 48 and 72 h, by plating dilutions on KB+Kan and M9 agar plates to distinguish between competing strains. The experiment was performed using three replicates and repeated three times.

In planta. Bacterial isolates were grown on KB agar plates for two days. Inoculum was prepared in a volume of 50 mL 10 mM $MgSO_4$ buffer with the addition of 0.002% Silwet-70 for better adherence. Endo- and epiphytic growth of strains *Psa* NZ54, *Psa* NZ13 Δ*hrcC* and *P. syringae* G33C was evaluated on clonally propagated 4-week old *A. chinensis* var. *chinensis* Hort16A and G3 kiwifruit plantlets using single and mixed-culture inoculation. Cultures were adjusted to OD_{600} 0.1 ($8x10^7$ cfu/mL) [OD_{600} 0.001 ($8x10^5$ cfu/mL) for invasion from rare]. Single

culture inoculations, as well as competition experiments in the mixes 1:1, 1:100 and 100:1, were performed.

Plants were inoculated by dipping them into the inoculum for 5 s and then leaving to air-dry with timepoint 0 being sampled just after the leaves were dry. Plants were maintained in a Conviron CMP6010 growth chamber at 20 °C with a light/dark period of 14/10 h, a humidity of 70% and watered every second day. Bacterial density was assessed either at 0, 2, 4, 7 and 10 days post-inoculation (dpi), or 0, 3, 7 dpi ($\Delta hrcC$ competition experiments).

Epiphytic growth. Sampled leaves were placed in separate sterile plastic bags with 35 mL 10 mM $MgSO_4$ buffer and shaken gently for 3 min. The leaf wash was pipetted into 50 mL centrifuge tubes and the solution was centrifuged at 4600 rpm for 3 min and the supernatant discarded. The bacteria were re-suspended in 200 μL 10 mM $MgSO_4$ buffer and serial dilutions were plated on M9 and KB+Kan agar plates.

Endophytic growth. For each leaf, a 1 cm^2 leaf disk was punched from the midrib and surface sterilized in 70% EtOH for 30 s, dried and then put in a 1.5 mL Eppendorf tube containing 200 μL 10 mM $MgSO_4$ buffer and two metal beads. Each leaf disk was disrupted for 1 min using the TissueLyser II (QIAGEN) and the lysate plated on M9 and KB+Kan agar plates.

Disease symptoms were recorded by taking photographs at each sampling point. All experiments were performed in duplicate, with at least 4 replicates per experiment.

Statistical analysis

A Student's t-test was used to verify the statistical difference where applicable. For non-normally distributed data and non-equal variance, the Mann Whitney U test was performed.

Fitness parameters

Fitness of co-culture consortia is expressed as the Malthusian parameter (Lenski *et al.*, 1991). The Malthusian parameter was calculated as

$$M = ln(N1_f/N1_i) - ln(N2_f/N2_i) \tag{3.1}$$

where $N1_i$ is the initial number of cfu of strain 1 at 0 h and $N1_f$ the cfu after 24 h (*in vitro*) or 2/3 dpi (*in planta*, Hort16A/G3).

3.4 Results

The stability of the polymorphism involving *P. syringae* from two different phylogroups was tested *in vitro* and *in planta*. For the competition experiments one representative of each lineage was chosen at random: *P. syringae* G33C, as a representative member of PG3a and *Psa* NZ54 from PG1. Coincidentally, these two strains were isolated from the same leaf, thus it is possible that they share a history of coexistence in the natural environment. Competition assays were performed *in vitro* and *in planta* on two different kiwifruit gold cultivars (Hort16A and G3) to test whether these two isolates are stably maintained. In addition plant pathogenicity tests were performed on kiwifruit plants to assess the infection potential of some PG3a strains.

3.4.1 Tn*5* mutagenesis

Prior to performing the competition experiments, a kanamycin resistant clone of *Psa* was constructed in order to distinguish the two types by agar plate culture during the experiment. Two different *Psa* strains were used throughout the experiment, *Psa* NZ54 for all 1:1 and invasion from rare experiments and *Psa* NZ13 Δ*hrcC*, to test whether a non-functional T3SS had any effect on bacterial growth of *P. syringae* G33C in competition.

The growth of the different kanamycin resistant clones was compared to wild type (WT) *Psa* NZ54 and *Psa* NZ13 Δ*hrcC* strains in King's B and

M9 media. There was no cost to fitness in carrying the kanamycin resistance cassette (Figure 3.1). *Psa* NZ54::Tn*5*-Clone 1 and *Psa* NZ13 Δ*hrcC*::Tn*5* Clone 3 were chosen to confirm the insertion site of the cassette. For *Psa* NZ54, the Tn*5* inserted in peptidase C69, disrupting the ORF at 1,104 bp, whereas for *Psa* NZ13 Δ*hrcC*, it inserted into the *AraC* family transcriptional regulator disrupting the ORF at 379 bp. Neither insertion site involves a transposable element. *Psa* NZ54::Tn*5*-Clone 1 and *Psa* NZ13 Δ*hrcC*::Tn*5* Clone 3 were subsequently used for all experiments and the names *Psa* NZ54 and *Psa* NZ13 Δ*hrcC* are used for clarity throughout the thesis.

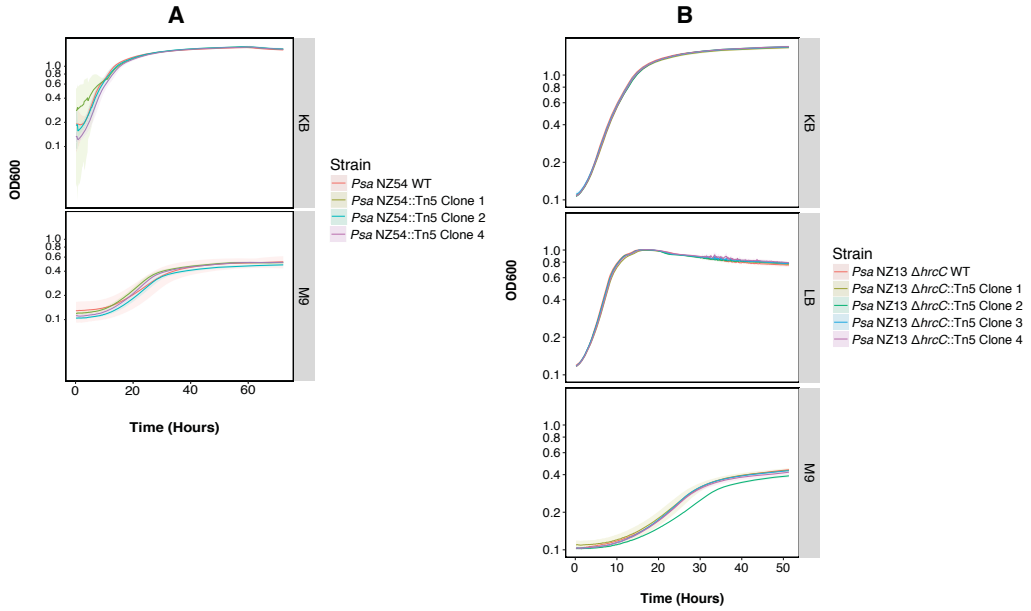

Figure 3.1: Growth curves of Tn*5* mutants. Strains were grown in KB, M9 and LB media under shaken conditions. Curves represent the mean of three individual replicates and the shaded area correlates to the standard deviation. (A) Growth curves of *Psa* NZ54 WT versus various clones of *Psa* NZ54::Tn*5*. (B) Growth curves of *Psa* NZ13 Δ*hrcC* WT versus various clones of *Psa* NZ13 Δ*hrcC*::Tn*5*.

3.4.2 Competition experiments *in vitro*

3.4.2.1 Individual growth of *P. syringae* G33C and *Psa* NZ54 *in vitro*

When grown individually, both strains showed similar growth dynamics and reached roughly the same maximum population size (ranging from $8.67\text{x}10^7 \pm 1.68\text{x}10^7$ to $2.43\text{x}10^9 \pm 4.6\text{x}10^8$ $\log_{10}$ cfu/mL) at 48 h *in vitro* (Figure 3.2).

3.4.2.2 A significant advantage of *P. syringae* G33C over *Psa* NZ54 *in vitro*

In competition, when *Psa* NZ54 and *P. syringae* G33C were co-cultured from an equal starting ratio of each competitor, *P. syringae* G33C grew to the same population size as when inoculated individually, whereas growth of *Psa* NZ54 was significantly reduced up to 100-fold (Figure 3.2). This effect was amplified in shaken M9 media, which mimics the nutrient-poor conditions encountered on the leaf surface: *Psa* NZ54 failed to establish a population in shaken M9 media. Overall, *P. syringae* G33C exhibited a fitness advantage in all environments, which was reflected by the poor fitness of *Psa* NZ54 relative to *P. syringae* G33C (ranging from -1.862 to -12.097, $P < 0.01$, Students t-test, see Table 3.3).

Table 3.3: Parameters of relative fitness of *Psa* NZ54 relative to *P. syringae* G33C for *in vitro* competition. The competition assay was performed for 72 h in shaken and static conditions for KB and M9 liquid media. Fitness parameters were calculated as ln difference *Psa* NZ54 – *P. syringae* G33C using the Malthusian parameters at 24 h from the average of three replicates. *indicating significance at the 1% level (Students *t*-test).

in vitro	**ln difference**	**SE**
KB shaken	-2.076*	±0.015
KB static	-1.862*	±0.053
M9 shaken	-3.876*	±0.002
M9 static	-12.097*	±0.085

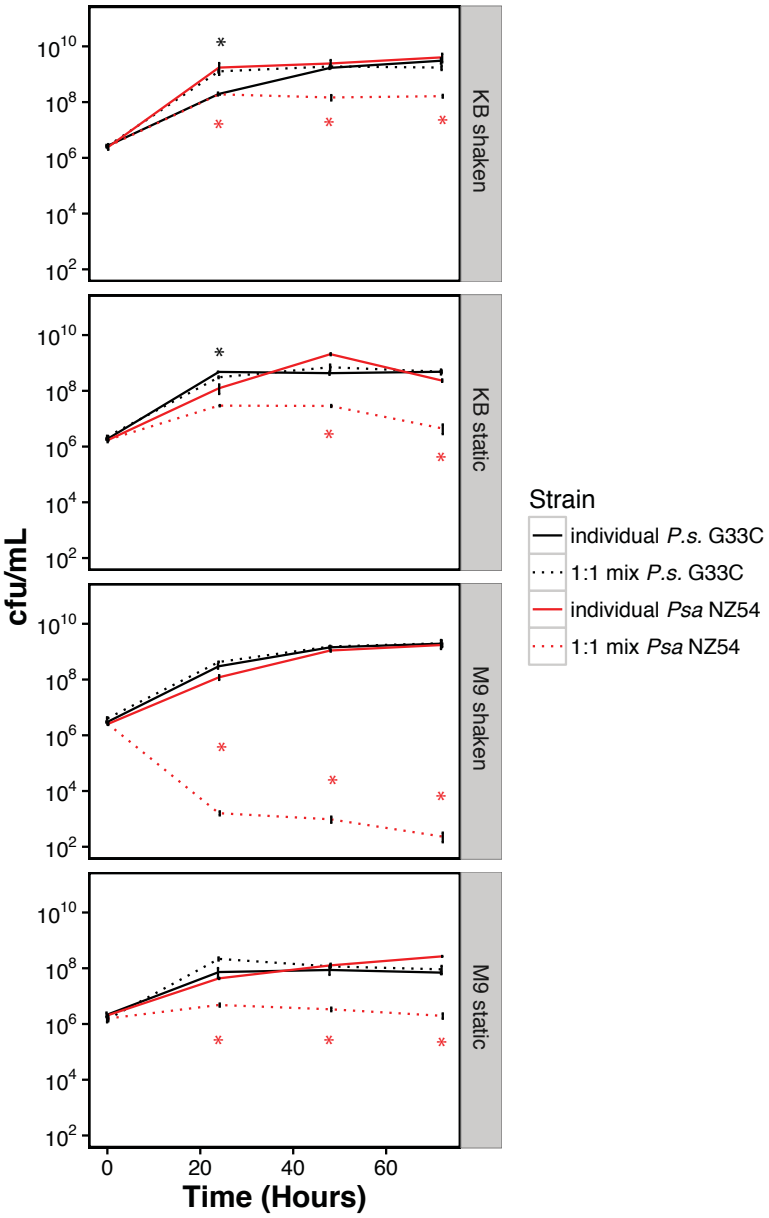

Figure 3.2: Individual growth dynamics compared with co-inoculation (1:1 ratio) *in vitro*. The solid lines represent the individually grown strains and the dotted lines the co-inoculated culture. Three replicate experiments were conducted with three independent measurements. The presented mean and standard error were calculated from a single experiment (three replicates). Asterisks indicate significant differences between individual and co-cultured growth at the 5% level (paired t-tests).

3.4.2.3 Invasion from rare

The results from 1:1 competition experiments suggest that *P. syringae* G33C is the superior competitor and that the polymorphism should therefore be unstable. To further test this apparent instability, experiments were performed to determine whether either type can invade from rare against a numerically superior competitor. Cultures were inoculated with a founding density of OD_{600} 0.006 ($5x10^6$ cfu/mL) and 0.0004 ($4x10^4$ cfu/mL) respectively. These invasion-from-rare curves (Figure 3.3A&B) can be compared to the growth of each strain on its own (solid lines in Figure 3.2) and the difference provides insights into the kind of interaction occurring between the two strains. *P. syringae* G33C successfully invaded *Psa* NZ54 cultures from rare after only 24 h in both KB and M9 media (Figure 3.3A). In fact it reached an equal population size to that achieved when cultured alone in M9.

Conversely, the reciprocal experiment showed that *Psa* NZ54 can invade *P. syringae* G33C from rare, except in the shaken M9 environment, establishing a population size of 10^5 to 10^6 cfu/mL (Figure 3.3B). Comparing these numbers to the individual growth rates of *Psa* NZ54 observed in Figure 3.2 revealed that the population size was 100-fold reduced. This suggests that *P. syringae* G33C exerts a suppressive effect on *Psa* NZ54. In shaken M9, *Psa* NZ54 failed to invade from rare, which confirmed the failure to proliferate in shaken M9 previously observed for the 1:1 mix. The striking suppression of *Psa* NZ54 by *P. syringae* G33C

provoked two additional experiments to check the validity of the findings: there was complete concurrence across the three experiments.

From these data it is clear that both isolates can invade from rare *in vitro*, except for *Psa* NZ54 in shaken M9, which suggests that in a controlled environment the polymorphism is stable. Clearly *P. syringae* G33C is the superior competitor *in vitro*. However, *in planta* there are numerous other factors (nutrient availability, humidity, UV radiation) that come into play compared to the controlled environment *in vitro*, so it is important to see whether the same dynamics hold true on the leaf.

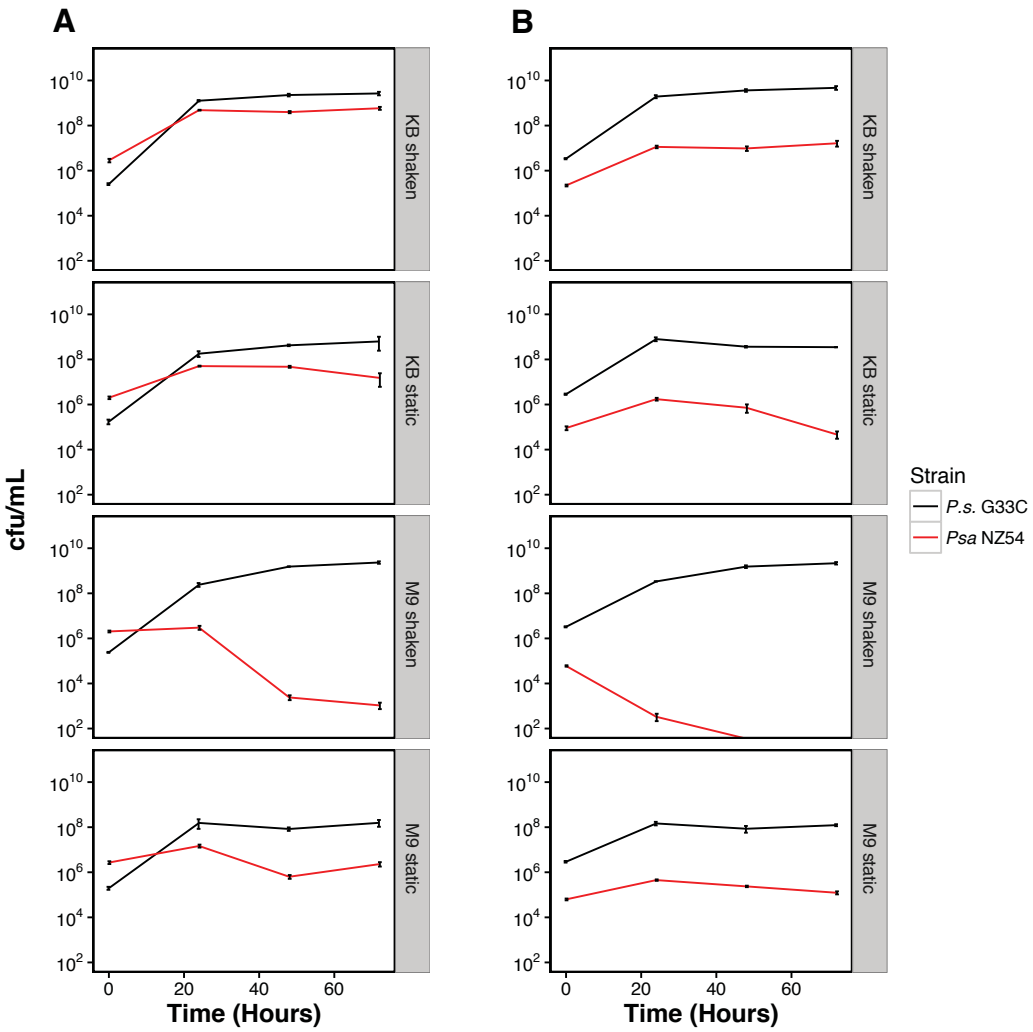

Figure 3.3: *In vitro* growth curves from invasion from rare experiments for *Psa* NZ54 : *P. syringae* G33C and vice versa. (A) shows the curves for *Psa* NZ54 : *P. syringae* G33C 10:1 ratio and (B) *Psa* NZ54 : *P. syringae* G33C 1:10 ratio. The presented mean and standard error were calculated from three individual measurements.

3.4.3 Competition experiments *in planta*

Experiments were performed on four-week-old plantlets from two gold kiwifruit cultivars: *A. chinensis* var. *chinensis* Hort16A and G3, with cultivar Hort16A being very susceptible to *Psa* and cultivar G3 more resistant to *Psa* infection.

3.4.3.1 Individual growth of *Psa* NZ54 and *P. syringae* G33C in kiwifruit

First, the long-term stability of *P. syringae* G33C growth was examined in cultivar Hort16A over the course of 28 days (Figure 3.4) to ensure the isolate was in fact able to colonize kiwifruit and establish a stable population. At 7 days post inoculation (dpi) the population maximum for endophytic growth of $2.22\text{x}10^4$ ($\pm 5.89\text{x}10^3$) was reached and remained stable, whereas the epiphytic growth appeared to be fluctuating by 100-fold over the course of the experiment. These fluctuations in population size on the leaf surface are most likely explained by variations in physical factors like humidity.

P. syringae G33C was able to establish and maintain a stable population size in both cultivars (G3, Hort16A); reaching a bacterial density of $5.3\text{x}10^3$ ($\pm 6.05\text{x}10^3$) cfu/cm^2 for endophytic growth and $1.1\text{x}10^3$ ($\pm 2.17\text{x}10^3$) cfu/cm^2 for epiphytic growth at 7 dpi for Hort16A; however, epiphytic growth was reduced on cultivar G3 with $1.1\text{x}10^1$ ($\pm 1.3\text{x}10^1$) cfu/cm^2 ($P > 0.1$, paired t-test) (Figure 3.7). *Psa* NZ54 attained a population size at least 10,000-fold greater than that of *P. syringae* G33C in both hosts. Comparing the growth

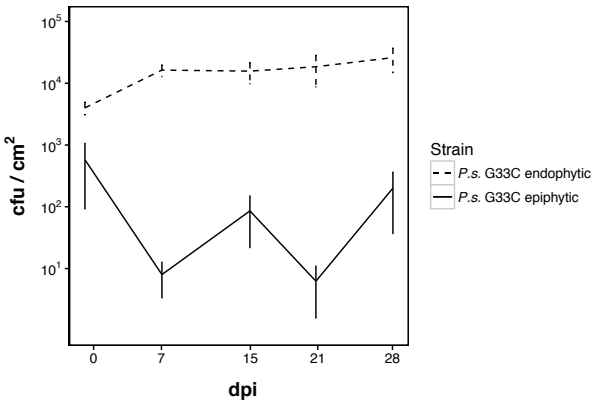

Figure 3.4: Long-term stability of *P. syringae* G33C growth in cultivar Hort16A over the course of 28 days. The dashed line represents endophytic growth, the solid line epiphytic growth. The presented mean and standard error were calculated from the mean of six individual measurements.

of *Psa* NZ54 on the two cultivars revealed that both endophytic and epiphytic growth were significantly reduced (P <0.05, Mann-Whitney U-test) by 10-fold at 3 dpi in cultivar G3 compared to Hort16A (Figure 3.7). There are no long-term data (>10 days) available for *Psa* NZ54 growth *in planta*, because infected 4-week old kiwifruit plantlets do not survive beyond 10 days.

Disease symptoms

P. syringae G33C did not produce any visible symptoms in either cultivar. *Psa* symptoms on leaves typically present as dark brown spots surrounded with a yellow halo. Kiwifruit plantlets inoculated with *Psa* NZ54 showed first leaf spots at 4 dpi and exhibited severe external symptoms at 7 dpi in cultivar Hort16A (Figure 3.5), whereas in cultivar G3 minor symptoms were observed at 7 dpi (Figure 3.6).

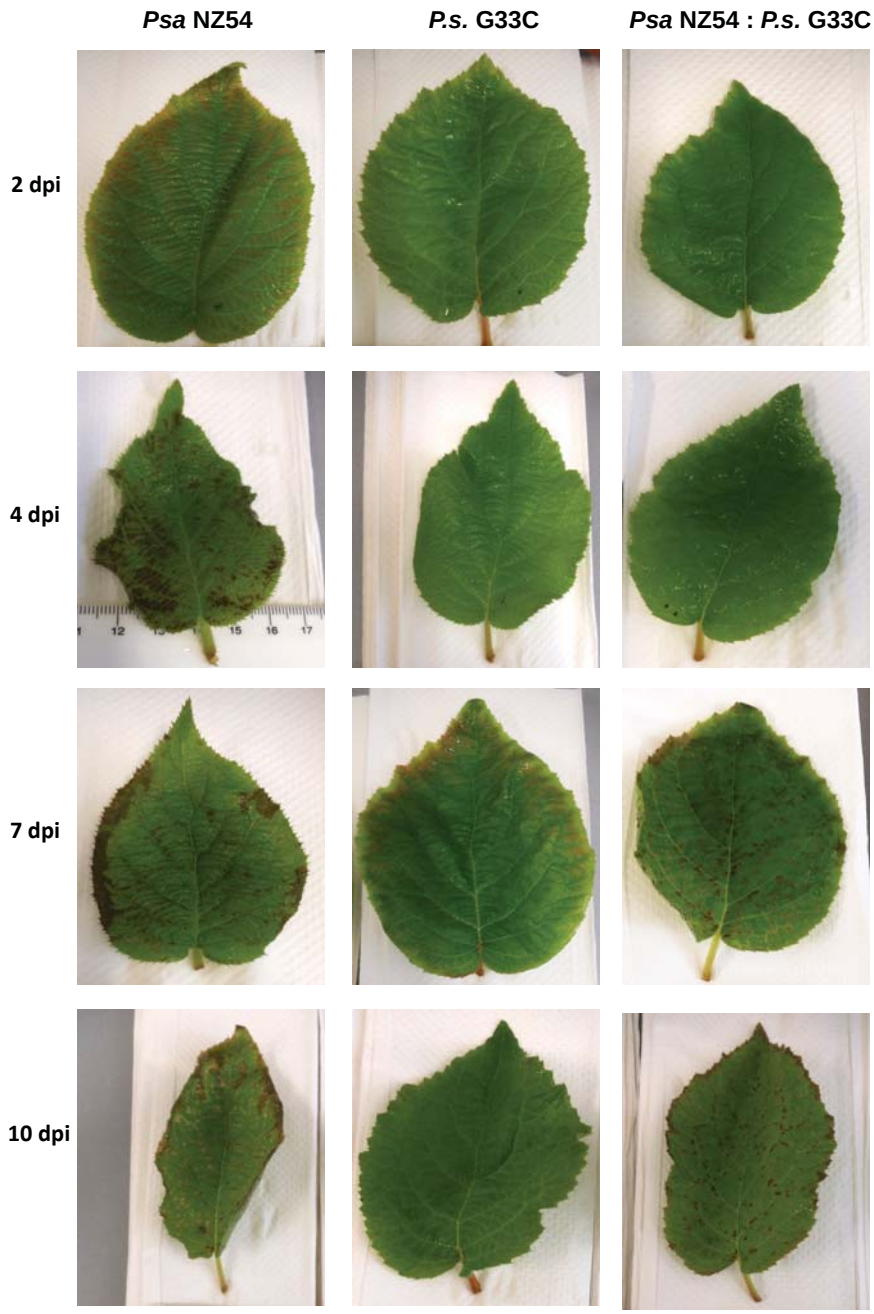

Figure 3.5: Leaves of 4-week old Hort16A plantlets inoculated with *Psa* NZ54, *P. syringae* G33C and 1:1 mix. Leaves sampled at 2, 4, 7 and 10 days post inoculation.

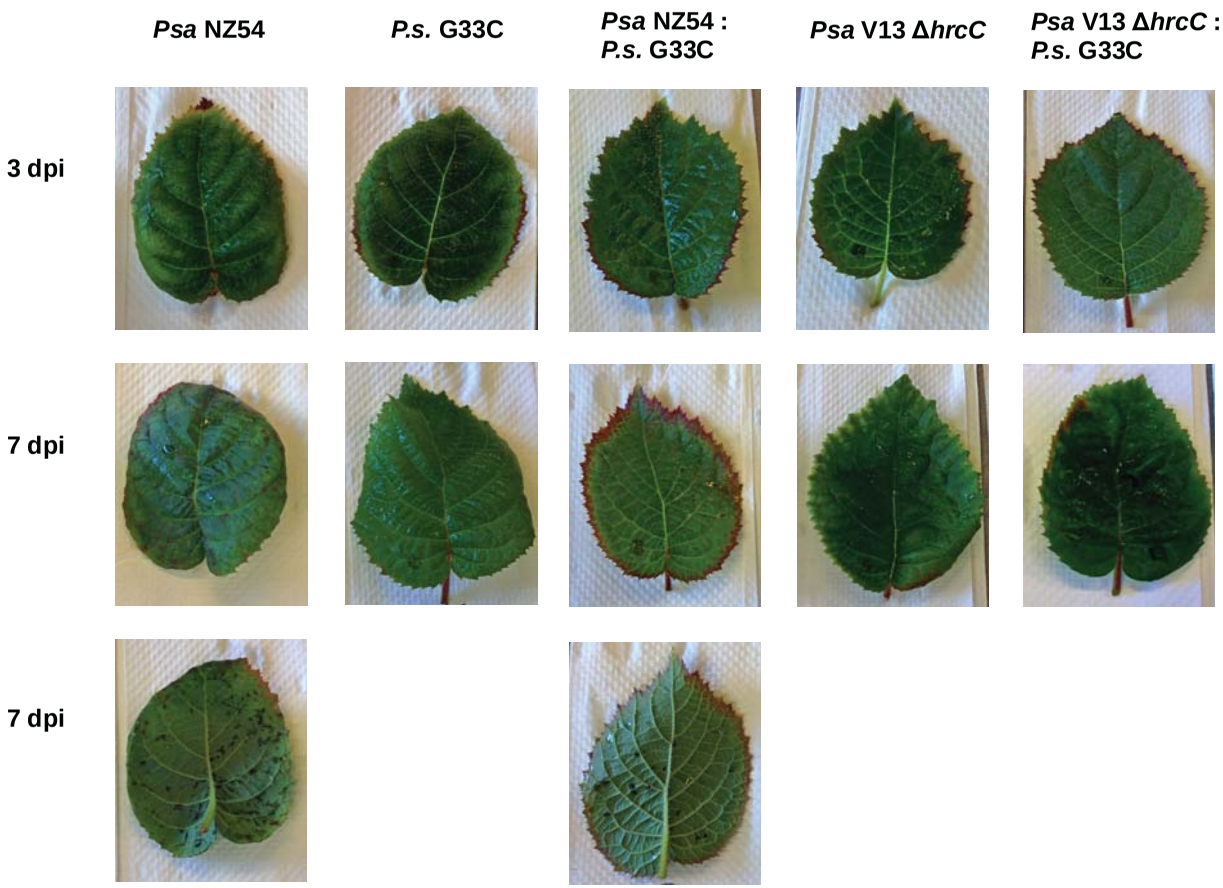

Figure 3.6: Leaves of 4-week old G3 plantlets inoculated with *Psa* NZ54, *P. syringae* G33C, *Psa* NZ13 Δ*hrcC* and 1:1 mixes. Leaves sampled at 3 and 7 days post inoculation. For leaves showing minor leaf spots, the underside of the leaf is also shown.

3.4.3.2 1:1 competition *in planta*

The results from the 1:1 competition experiments *in planta* differ markedly from those observed *in vitro*. Competition experiments using an equal starting ratio of each competitor revealed that both the epiphytic and endophytic population sizes of *Psa* NZ54 were significantly reduced (P <0.05, paired t-test) on cultivar Hort16A when grown in competition with *P. syringae* G33C. However, the growth was reduced only at early stages of the experiment and after 7 dpi no significant differences were detected. The reduction of growth of *Psa* NZ54 in the presence of *P. syringae* G33C was more pronounced in cultivar G3, where endophytic growth of *Psa* NZ54 was decreased 100-fold (7 dpi) when co-inoculated with *P. syringae* G33C (P <0.05, paired t-test, Figure 3.7B).

The non-pathogenic *P. syringae* G33C was able to maintain a stable population size in 1:1 competition experiments, despite the increasing abundance of *Psa* NZ54 in both the apoplast and the phyllosphere. In fact, the presence of *Psa* NZ54 had a highly significant positive effect on the growth of *P. syringae* G33C in both plant hosts. At 7 dpi, *P. syringae* G33C established 1,000-fold higher epiphytic population densities in *A. chinensis* var. *chinensis* Hort16A (P <0.01, paired t-test) and 10-fold higher epiphytic and endophytic population densities in G3 plants. Co-inoculated Hort16A plants exhibited a notable delay in symptom onset compared to singly inoculated plantlets (Figure 3.5), whereas there was no difference in symptoms for G3 plantlets (Figure 3.6).

The increased fitness of *Psa* NZ54 over *P. syringae* G33C in Hort16A competition experiments was also reflected in the relative fitness parameters (0.7 ± 0.1 for epiphytic and 4.9 ± 0.8 for endophytic, P <0.05, t-test), whereas in G3 plants *P. syringae* G33C performed better in the epiphytic phase (Table 3.4).

Table 3.4: Parameters of fitness of *Psa* NZ54 relative to *P. syringae* G33C for *in planta* competition. The parameters were calculated as ln difference *Psa* NZ54 – *P. syringae* G33C using the Malthusian parameters at Day 2/Day 0 and Day 3/Day 0 from the average of four replicates. Asterisks indicate significance at the 5% level (Students t-test).

Cultivar	**Hort16A**	**G3**
epiphytic	0.7 ± 0.1*	-1.6 ± 0.2
endophytic	4.9 ± 0.8*	4.7 ± 0.1*
Time	T_2/T_0	T_3/T_0

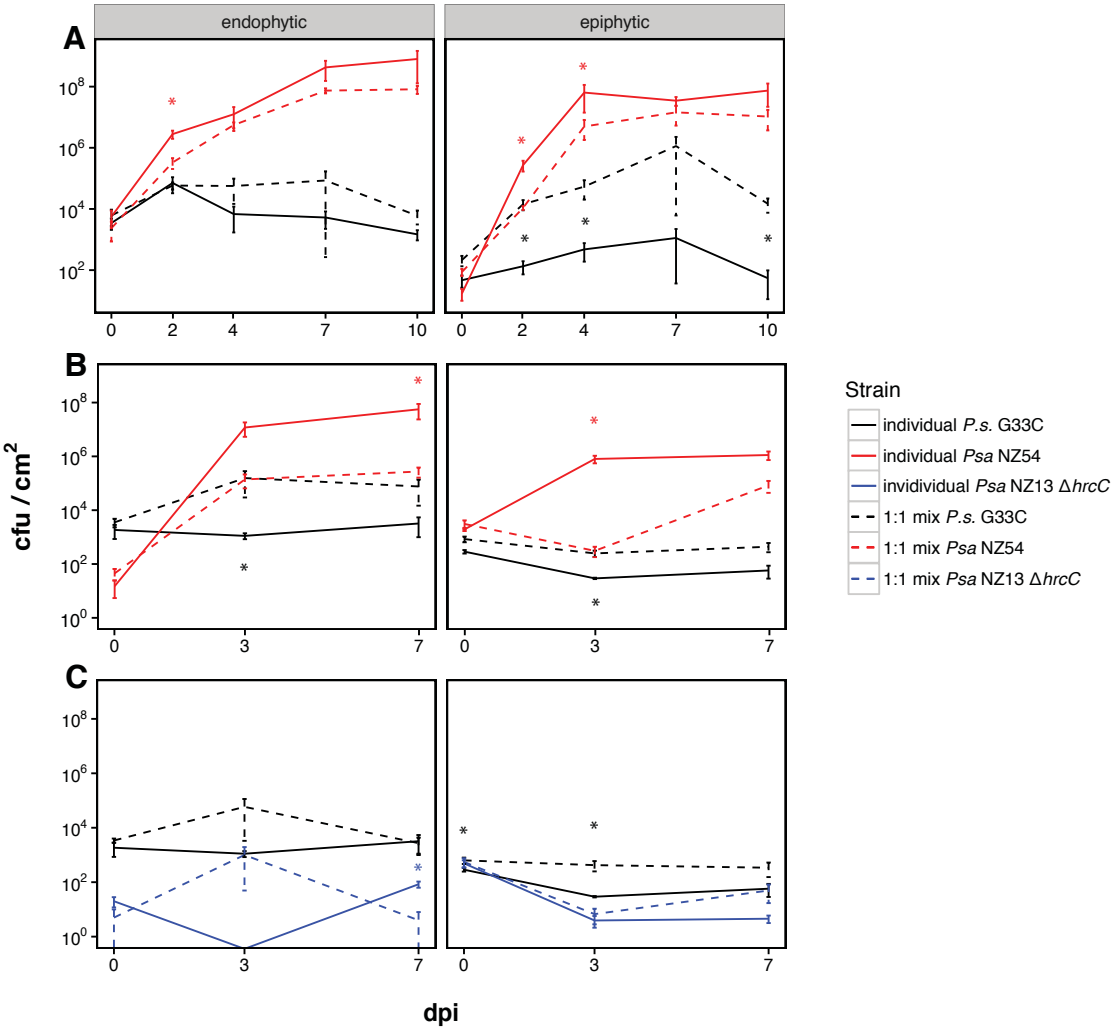

Figure 3.7: 1:1 competition growth assays of *Psa* NZ54 vs. *P. syringae* G33C *in planta*. Solid lines represent the single growth of each strain; dashed curves show the growth of each strain in competition. The presented mean and standard error were calculated from the mean of four (Hort16A) and five (G3) individual measurements. Hort16A plantlets (A) and G3 plantlets (B) were inoculated with a 1:1 mix of *P. syringae* G33C : *Psa* NZ54. (C) G3 plants were inoculated with 1:1 mix of *P. syringae* G33C : *Psa* NZ13 Δ*hrcC*. Asterisks indicate significant differences between individual and co-cultured growth at the 5% level (paired *t*-tests).

3.4.3.3 Invasion from rare

In order to establish the stability of the interaction between *Psa* NZ54 and *P. syringae* G33C, invasion from rare experiments were performed on *A. chinensis* var. *chinensis* Hort16A. Upon co-inoculation in a 100:1 (*Psa* NZ54 : *P. syringae* G33C) ratio, *P. syringae* G33C was able to invade from rare over the first 4 days, but was then excluded by *Psa* NZ54 (Figure 3.8B). An increase in growth of *P. syringae* G33C from 0 dpi to 4 dpi was followed by a population collapse by 7 dpi with no endophytic growth detected and reduced epiphytic growth in environments dominated by *Psa* NZ54. Conversely, *Psa* NZ54 grew to the same population size as when inoculated individually in the presence of *P. syringae* G33C ($P > 0.1$, paired t-tests).

In the reciprocal experimental setup (1:100 *Psa* NZ54 : *P. syringae* G33C ratio), *Psa* NZ54 invaded from rare against *P. syringae* G33C in the endophytic and epiphytic environment, although the invasion was slowed down in the epiphytic phase. However, both epi- and endophytic population sizes were significantly reduced compared to single inoculations ($P < 0.01$, paired t-tests) (Figure 3.8C). Despite the growing population of *Psa* NZ54, *P. syringae* G33C was able to maintain the same epiphytic population size as when inoculated individually ($P > 0.2$, paired t-tests). The endophytic population size of *P. syringae* G33C was increased ($P < 0.01$, 7 dpi, paired t-test), which mirrored the results from the 1:1 mix, where the presence of *Psa* NZ54 also had a positive effect on growth.

Contrary to the *in vitro* experiments, *in planta* the results suggest that

the polymorphism is not stable. This is either because both strains are occupying the same niche and competing for the same resources or because *Psa* is allelopathically active via the production of toxins, antibiotics siderophores or similar bioactive molecules. Clearly, in its native plant host, *Psa* NZ54 is the superior competitor.

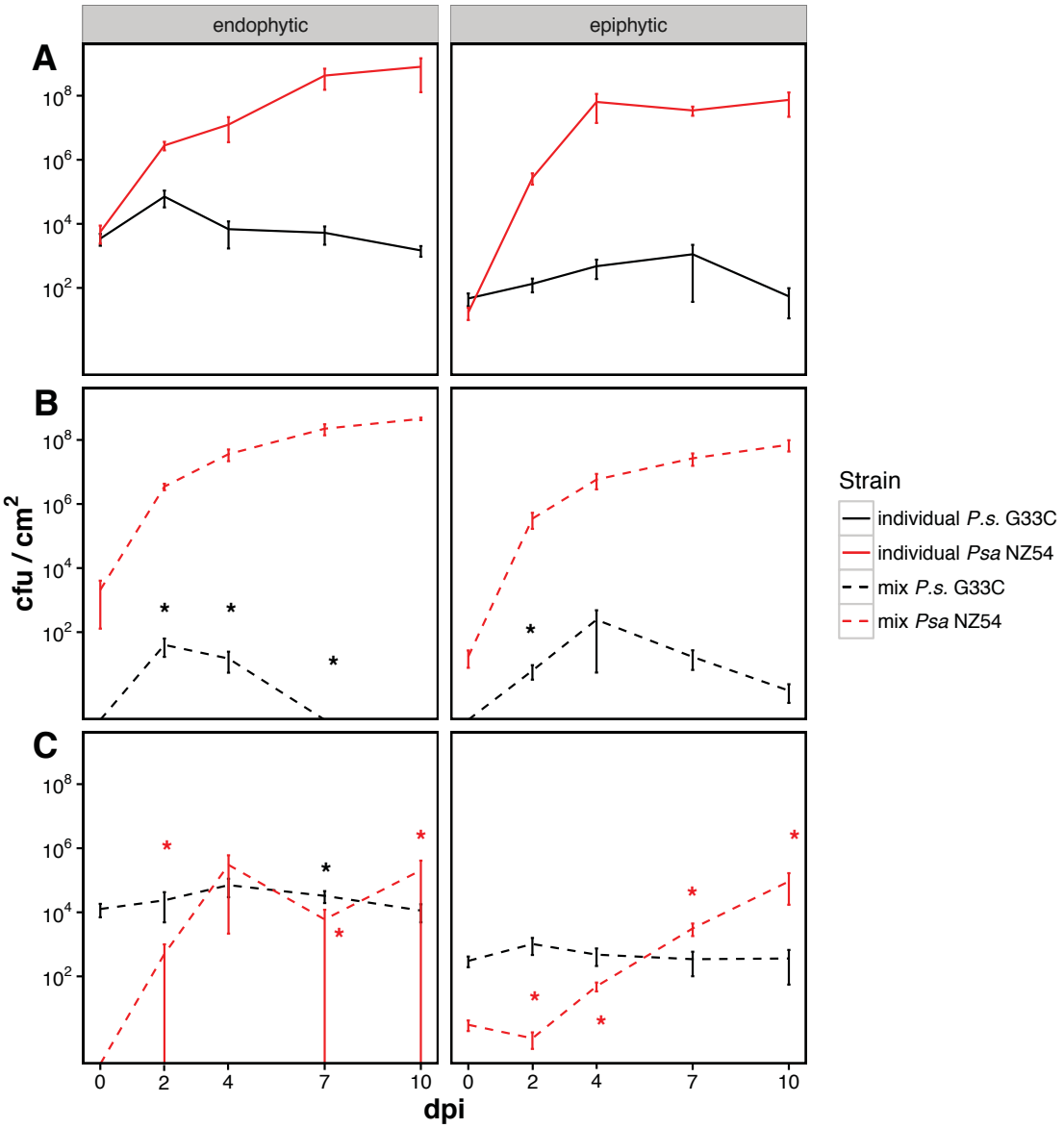

Figure 3.8: *In planta* invasion from rare experiments for *Psa* NZ54 and *P. syringae* G33C. Hort16A plantlets were inoculated with different ratios of strains *Psa* NZ54 and *P. syringae* G33C. Because different inoculation ratios were used for the different treatments the individual growth is plotted separately for comparison. Row A) shows individual growth, B) *P. syringae* G33C invading from rare (100:1) and C) *Psa* NZ54 invading from rare (1:100). The presented mean and standard error were calculated from the mean of four individual measurements. Asterisks indicate significant differences between individual and co-cultured growth at the 5% level (paired *t*-tests).

3.4.3.4 No priority effect was found for either of the strains

In order to pinpoint any advantage of being an early colonist, a time-stagger experiment was performed. Kiwifruit plants were pre-inoculated (8x10^7 cfu/mL) with either of the two strains and followed by a subsequent inoculation (8x10^7 cfu/mL, or 8x10^5 cfu/mL) of the other strain after three days. Being an early colonizer did not provide any advantage for *Psa* NZ54 (Figure 3.9A) as *P. syringae* G33C was still able to maintain and establish a stable population size by 7 dpi and there was no significant difference in growth compared to the individual growth (P >0.3, paired t-tests). Reducing the secondary inoculation density, however, resulted in a reduced initial population size at 3 and 7 dpi for *P. syringae* G33C (P <0.05, paired t-tests), but by 10 dpi it reached the same size as when grown individually (Figure 3.10C).

With *P. syringae* G33C being the early colonist (Figure 3.9B) *Psa* NZ54 (inoculated at 8x10^7 cfu/mL) grew to the same population size by 7 dpi as compared to individually inoculated plants (P >0.5, paired t-tests). Using a reduced inoculum density (8x10^5 cfu/mL), the growth of *Psa* NZ54 was initially lower when compared to individual growth (P <0.05, paired t-tests), but these differences were remedied for the epiphytic population by 10 dpi (Figure 3.9B).

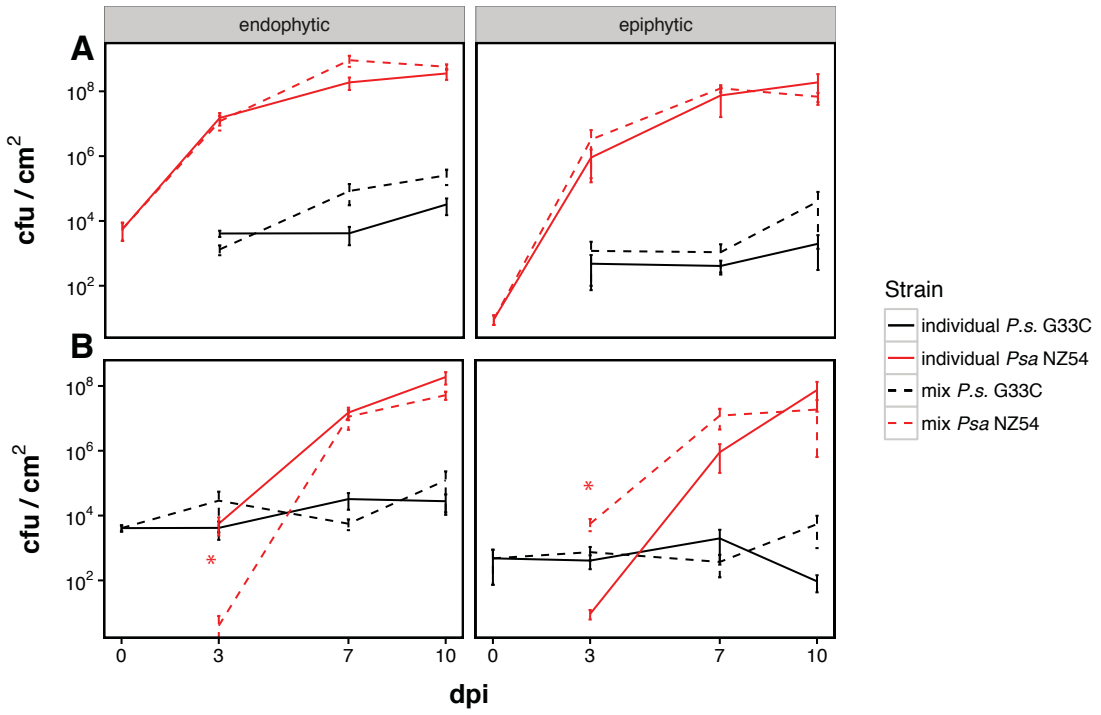

Figure 3.9: *In planta* priority effect of either *Psa* NZ54 or *P. syringae* G33C, with subsequent inoculation of the respective second strain with the same concentration. Solid curves represent the growth of each strain in individually inoculated plants. Dashed curves represent the growth in competition. Data presented is the mean and SE for 5 individual replicates. (A) *In planta* growth assay of *P. syringae* G33C using Hort16A plantlets pre-inoculated for three days with *Psa* NZ54 (8x10^7 cfu/mL). (B) *In planta* growth assay of *Psa* NZ54 using Hort16A plantlets pre-inoculated for three days with *P. syringae* G33C (8x10^7 cfu/mL). Asterisks indicate significant differences between individual and co-cultured at the 5% level (paired *t*-tests).

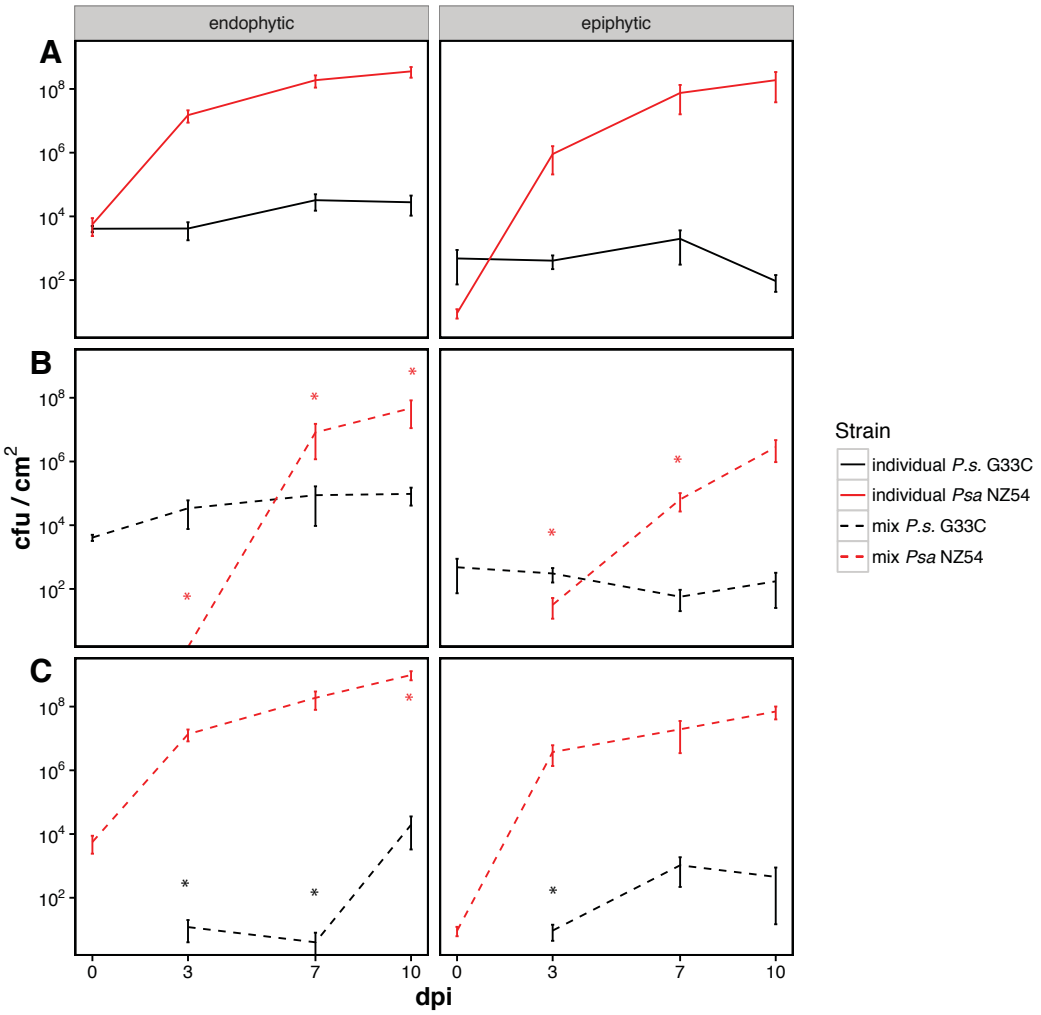

Figure 3.10: *In planta* priority effect of either *Psa* NZ54 or *P. syringae* G33C, with subsequent inoculation of the respective second strain with a 100-fold lower concentration. *In planta* growth assay using Hort16A plantlets pre-inoculated with each respective strain for three days ($8x10^7$ cfu/mL) and subsequent inoculation of the second isolate at $8x10^5$ cfu/mL. Row A) shows individual growth of each strain for comparison, B) inoculation of *Psa* NZ54 at Day 3 and C) inoculation of *P. syringae* G33C at Day 3. Data presented is the mean and SE for 5 individual replicates. Asterisks indicate significant differences between individual and co-cultured growth at the 5% level (paired *t*-tests).

3.4.3.5 T3SS mutant

To assess whether the heightened growth of *P. syringae* G33C in the presence of *Psa* NZ54 was due to the virulence activity of the pathogen elicited by the Type 3 Secretion system (that might have provided additional resources for *P. syringae* G33C), the 1:1 competition experiment was repeated using a Type 3 secretion system (T3SS) deficient mutant (*Psa* NZ13 $\Delta hrcC$) on cultivar G3. The epiphytic growth of *P. syringae* G33C remained elevated when co-inoculated with *Psa* NZ13 $\Delta hrcC$ (Figure 3.7C), indicating that the virulence activity encoded by the T3SS was not responsible for the advantage conferred to the non-pathogenic strain (P <0.05, paired t-test).

3.5 Discussion

Few studies have investigated the co-existence of non-pathogenic and pathogenic *P. syringae* isolates *in planta* by determining the location and survival, the level of co-existence, interactions and their epidemic potential (Bartoli *et al.*, 2015; Rufián *et al.*, 2017; Wilson *et al.*, 1999; Wilson & Lindow, 1994b), although there is early work from the late 60s looking at competitive interactions (e.g. Omer & Wood, 1969; Young, 1974). I explored the nature of the association between PG3a, which appears to be a resident clade of kiwifruit commensals, and the kiwifruit pathogen *Psa*. I was interested in how interactions between plant epiphytes and pathogens might alter the disease outcome by performing experiments using a pair of strains (virulent *Psa* NZ54 and the kiwifruit commensal *P. syringae* G33C), which were originally isolated from the same leaf.

3.5.1 *P. syringae* G33C as superior competitor *in vitro*

The commensal strain *P. syringae* G33C displayed a fitness advantage over *Psa* NZ54 *in vitro*, which was even more prominent in minimal media (M9) than rich media (KB) (Figure 2 & Figure 3). Despite the competitor's presence, *P. syringae* G33C grew to the same population size as when grown individually, irrespective of media or what initial inoculation ratio (1:1, 1:10, 10:1). Surprisingly, *P. syringae* G33C appeared to have an

advantage in minimal media, with *Psa* NZ54 being excluded even more quickly in shaken M9. This is despite M9 being the medium that most closely resembles conditions that bacteria might face *in planta*. M9 media is commonly used for simulating endophytic growth conditions, because it has a limited supply of nutrients compared to KB media. The collapse of *Psa* NZ54 in shaken M9 was also observed at 1:10 and 10:1 inoculation ratios, although in 10:1 (*Psa* NZ54 : *P. syringae* G33C, Figure 3.3A) the collapse of *Psa* NZ54 was sightly delayed, which would be expected if inhibition were due to a build-up of *P. syringae* G33C antimicrobial compounds.

Phytotoxin production is widespread among fluorescent pseudomonads, with some toxins also having antimicrobial activity (Bender *et al.*, 1999). Interestingly, some toxins are only produced in shaken conditions (e.g. phaseolotoxin produced by *P. syringae* pv. *phaseolicola*), whereas *P. syringae* pv. *syringae* produces syringomcyin in unshaken culture (Durbin, 1982). Alternatively, the two isolates could engage in contact-dependant growth inhibition (CDI), where upon cell-to-cell contact toxins are delivered into the opponent cell (Hayes *et al.*, 2010; Ruhe *et al.*, 2013). CDI, expressed as a Type V secretion system, was first described in *E. coli*, but *Yersinia pestis* and *Bacillus pseudomallei* most likely also engage in CDI (Aoki *et al.*, 2005). Additionally, the Type VI secretion system, which is commonly found in Gram-negative bacteria, plays a role in the ecological success of pathogens; expression of *hcp2* by *P. syringae* pv. *tomato* was induced when grown in competition with enterobacteria and yeasts (Haapalainen *et al.*, 2012). Other protein-based systems to kill competitors were shown in *P. syringae* pv. *phaseolicola* B728a, which carries a R-type

syringacin, a bacteriophage-derived bacteriocin (Hockett *et al.*, 2015). The genomic region encoding those bacteriophage fragments was conserved in several other pseudomonads (Hockett *et al.*, 2015).

In principle both isolates were able to coexist in a stable manner (with the exception of growth in shaken M9), but clearly *P. syringae* G33C is the superior competitor *in vitro*. This can possibly be attributed to the fact that as a resident commensal on kiwifruit it is more of a generalist with an enhanced capacity to take up available nutrients, whereas *Psa* NZ54 is a specialist on kiwifruit. *Psa* NZ54 may simply trade off growth against enhanced capacity to take resources from a host. The advantage of *P. syringae* G33C *in vitro* does not ensure overall dominance *in planta*, where other factors (e.g. environmental factors like UV radiation, nutrient availability, humidity) come into play.

3.5.2 A reverse outcome of competition *in planta*

3.5.2.1 Growth of *P. syringae* G33C in kiwifruit

P. syringae G33C was demonstrated to be non-pathogenic on *A. chinensis* var. *chinensis* cultivars Hort16A and G3, which are highly susceptible to infection by *Psa*. *P. syringae* G33C was able to colonize the leaf epidermis and apoplasm, attaining a population size four logs reduced compared to *Psa* NZ54 (Figure 3.7) and produced no visible symptoms of disease (Figure 3.5 and Figure 3.6). Limited growth of 10^3-10^4 cfu/cm^2 for *P. syringae* isolates *in planta* has been described previously. Early work done by Young

(1974) demonstrated that population sizes for the pathogen *P. syringae* pv. *phaseolica* on its respective host (*Phaseolus vulgaris*) were double that of the non-pathogenic strain. Omer & Wood (1969) showed that *P. syringae* pv. *phaseolicola* attains double the population size in susceptible plants when compared to resistant bean plants.

Overall, studies have shown that non-pathogenic isolates reach lower population sizes (up to 10^4 cfu/cm^2) that are similar to the growth of T3SS mutants, whereas highly aggressive pathogens exceed population sizes of 10^6 cfu/cm^2 (Clarke *et al.*, 2010; Demba Diallo *et al.*, 2012; Kniskern *et al.*, 2011; Mohr *et al.*, 2008). Recent work by Bartoli *et al.* (2015) showed that environmental strains have the potential to reach population sizes very similar to *Psa*, however these isolates induced symptoms *in planta*.

It is therefore likely that the ability to infect a plant by employing the use of the T3SS and various virulence effectors to overcome the plant's immune system is a prerequisite for *P. syringae* to reach high densities.

3.5.2.2 Stability of the polymorphism

The inability of *P. syringae* G33C to invade from rare *in planta* (Figure 3.8B) suggests that the polymorphism between *Psa* NZ54 and *P. syringae* G33C is not stable. This is an indication that density-dependent competitive exclusion can occur.

Intriguingly, this effect was not observed when plants were pre-inoculated with *Psa* NZ54 for three days (Figure 3.10C, 100:1 ratio).

With *Psa* NZ54 pre-established on the leaf, *P. syringae* G33C successfully invaded from rare. This might be explained by the fact that *Psa* NZ54 had already successfully colonized the apoplast, hence *P. syringae* G33C was able to occupy freely available space on the leaf surface. In contrast, when strains colonise at the same time they directly compete for resources and space and the unequal inoculum ratio results in the more abundant isolate taking over.

This density-dependent effect was also observed for the pathogenic isolate. Usually, there was no observable difference in growth of *Psa* NZ54 at 0 dpi when inoculated either alone or with the competitor. Yet, when *P. syringae* G33C was allowed to grow on the leaf for three days before co-inoculation of *Psa* NZ54 (Figure 3.9B), the initial population of endophytic *Psa* NZ54 at 0 dpi was drastically reduced. Epiphytic growth of *Psa* NZ54 was also significantly reduced when co-inoculated with *P. syringae* G33C in equal amounts on *A. chinensis* var. *chinensis* G3 (Figure 3.7B) and to a lesser extent on *A. chinesis* var. *chinensis* Hort16A. This indicates that plant epiphytes may suppress pathogen growth, either as direct antagonists or indirectly via resource competition (Wilson & Lindow, 1994a). *Psa* growth suppression has also been demonstrated by Bartoli *et al.* (2015) using environmental *P. syringae* isolates from freshwater habitats. Niche specificity and availability, and the dynamics of aggregate formation may all play important roles in maintaining diversity on the plant leaf (Monier & Lindow, 2005; Wilson *et al.*, 1999). It is also possible that *P. syringae* G33C might be lacking certain Type III effectors

suppressing plant defence responses and essentially triggering the plant defence system when colonizing the leaf prior to arrival of *Psa*.

3.5.2.3 Beneficial effect of *Psa* NZ54 on *P. syringae* G33C *in planta*

P. syringae G33C seemed to benefit from the presence of the pathogen *Psa* NZ54 *in planta*, as its growth was significantly enhanced in co-inoculated plants for both cultivars (Figure 3.7). The interaction of *P. syringae* G33C and *Psa* NZ54 appears to be synergistic in nature, where one isolate gains in fitness by reaping the benefits of the co-residing strain. Studies exploring the dynamics of mixed infections have been very limited, but boosted growth of the non-pathogen has been reported in the presence of pathogenic lineages (Macho *et al.*, 2007; Young, 1974), although this effect appeared to be inoculum density dependant (Macho *et al.*, 2007). Just recently, confocal microscopy revealed increased growth of non-pathogenic *P. syringae* and an ability to colonize wider territories in the presence of a pathogenic strain (Rufián *et al.*, 2017).

Such a fitness benefit for one strain has been previously attributed to pathogenicity genes of the co-residing isolate, e.g. *P. syringae* B728a mutants with a defective *hrpJ* gene showed poor individual growth (10^3-10^4 cfu/sample), but displayed a 1000-fold increase in growth when co-inoculated with a strain possessing the *hrp* cluster (Hirano *et al.*, 1999). I investigated whether enhanced growth of *P. syringae* G33C resulting from *Psa* NZ54 infection could be due to greater nutrient or water availability as

a result of *Psa* virulence activity, but found that the positive effect on growth of *P. syringae* G33C persisted, even in the absence of a functional T3SS for the competing *Psa* strain (Figure 3.7C). It remains possible that virulence activities not encoded by the T3SS may be responsible for this outcome, but hitch-hiking effects demonstrated elsewhere were T3SS dependant (Hirano *et al.*, 1999). Nevertheless, the results rule out a type of synergistic effect where the resident strains play a role in the pathogenicity process, as was for example demonstrated for *P. syringae* pv. *savastanoi*, where the severity of olive knot canker was increased in the presence of other bacteria (Marchi *et al.*, 2006; Rodríguez-Palenzuela *et al.*, 2010). The outcome of the competition experiments suggests there was no additive effect on disease severity due to the presence of *P. syringae* G33C.

In summary, these results clearly show that commensal bacteria engage in complex interactions with pathogenic isolates. What has long been recognised in human and mammalian diseases caused by microorganisms remains mostly neglected in plants, as the view of infections caused by single clonal lineages is engrained in plant pathology (Lamichhane & Venturi, 2015).

A focus of plant pathogenic studies should shift more towards other members of the microbial community and what role they play. Interactions are not influenced only by environmental factors; a spatial component can also have an impact on the intensity of ecological interactions. In the phyllosphere, the usual pathways through which bacteria gain entry into the apoplast are openings like stomata, trichomes, hydathodes or small

wounds in the leaf epidermis. These openings exert spatial bottlenecks, where the geographical space is reduced to a small gap (e.g. the typical stomata size in kiwifruit leaves is 20 μm, Przywara *et al.* (1988)). The close proximity can further facilitate or even force engagement between commensal strains and the kiwifruit pathogen *Psa*.

It has been shown that interactions either between commensals and pathogenic isolates, or between different pathogenic isolates, can have an influence on the outcome of the infection process in plants (Buonaurio *et al.*, 2015; Kudela *et al.*, 2010; Nakahara *et al.*, 2016). This can take place via a variety of mechanisms, although many remain to be explored. Resident isolates can produce secondary metabolites (like IAA) as public goods, which in turn can increase pathogenicity (Buonaurio *et al.*, 2015; Kudela *et al.*, 2010). Some co-colonizing bacteria exhibit antagonism by inhibiting growth via production of toxins, biosurfactants or antibiotics (Raaijmakers & Mazzola, 2011). An alternative tactic is induction of plant immune responses, which in turn restrict the growth of a virulent strain (Klement & Lovrekovich, 1961; Omer & Wood, 1969). In the past these attributes have been used for the development of biocontrol agents, as host-adapted and ecologically more similar species tend to be more effective in controlling disease (Völksch & May, 2001).

The diminishing effect of *P. syringae* G33C on *in planta* growth of the kiwifruit pathogen *Psa* NZ54 was interesting and naturally such a finding leads to the question of pathogenic properties of these commensal isolates. Before driving research towards the development of biocontrol agents, there

is a need to investigate the range of virulence traits. These isolates could provide a pool of readily available DNA sequences to be horizontally transmitted to other bacteria, resulting in the possible emergence of new pathogenic types.

Chapter 4

Comparative genomics of kiwifruit-associated *P. syringae* Phylogroup 3a

4.1 Introduction

4.1.1 The status of *P. syringae* genomics

P. syringae encompasses strains with versatile life styles reflected in a remarkable natural diversity both in an agricultural and environmental setting (Berge *et al.*, 2014; Hwang *et al.*, 2005; Sarkar & Guttman, 2004). This diversity is also mirrored in the genomes of these strains, with ∼3,400

core genes found for *P. syringae* across phylogroups and ∼12,750 strain-specific genes found in the pan genome (Baltrus *et al.*, 2011).

P. syringae genomes range from 5.0 to 6.4 Mbp in size, some PG3 isolates have genomes larger than 7 Mbp (NCBI accession GCA_001293825.1) with a varying GC content of 56-60% (O'Brien *et al.*, 2011; Thakur *et al.*, 2016). For a handful of *P. syringae* isolates a closed genome is available (e.g. *Pto* DC3000, *Psy* B728a, *Pph* 1448a, *Psa* NZ13, CC1557 ...) (Buell *et al.*, 2003; Feil *et al.*, 2005; Joardar *et al.*, 2005; Templeton *et al.*, 2015; Hockett *et al.*, 2014), with the rest being draft genomes presented in a varying number of contigs (e.g. ranging from 86 to 1,143 contigs in Thakur *et al.* (2016)). Around 5,500 predicted coding sequences (CDS) are represented in *P. syringae* genomes (Thakur *et al.*, 2016), but genome variation is extensive within *P. syringae*, not only among phylogroups (e.g. Baltrus *et al.*, 2011; O'Brien *et al.*, 2011), but even within pathovars, like pv. *actinidiae* (McCann *et al.*, 2013, 2017).

Most *P. syringae* genomes include a plasmid (up to 120 kilobases (kb) in size), which can play a role in virulence or resistance factors. For example, determinants of copper resistance were carried on mobile elements, including plasmids (Colombi *et al.*, 2017).

4.1.2 Evolution of pathogenicity

The focus of comparative genomics studies is typically on virulence factors like the Type 3 Secretion System (T3SS) or phytotoxins carried on the

chromosomes of recognised pathogens (Baltrus *et al.*, 2011, 2017; Nowell *et al.*, 2016; O'Brien *et al.*, 2011). Genome analysis of non-pathogenic isolates has received far less attention, despite knowledge that a diverse population of commensal *P. syringae* exist in the environment and on wild plants (Monteil *et al.*, 2012; Morris *et al.*, 2010, 2007). Comparative analysis of the genomes of commensals could offer knowledge about genomic attributes that contribute to more general epiphytic fitness.

An interesting aspect that has been receiving attention from disease epidemiologists lately is the transition from mutualism or commensalism to pathogenicity. *Yersinia* spp. has served as a model example of the evolution of human pathogenic isolates from environmental or commensal bacteria. Evolution of virulence in *Yersinia pestis* was mainly attributed to the acquisition of plasmids (particularly plasmid pYV), mobile genetic elements and bacteriophages (McNally *et al.*, 2016). Such transitions and the potential pool of commensal isolates in contributing to HGT events has also attracted interest in other pathogens, such as for example *Escherichia coli* (Rasko *et al.*, 2008; Tenaillon *et al.*, 2010) or *Candida albicans* (Hellstein *et al.*, 1993; Hube, 2004).

Recent studies have shown that non-pathogenic *P. syringae* isolated from aquatic habitats can evolve the capacity to infect tomato and kiwifruit via a small number of evolutionary changes (Bartoli *et al.*, 2015; Monteil *et al.*, 2013). A significant proportion of *P. syringae* strains identified in various habitats are in fact commensals and most likely non-pathogenic, despite some efforts to identify targeted hosts (e.g. Bartoli *et al.*, 2015;

Mohr *et al.*, 2008). The host range for a large number of *P. syringae* strains is unknown and it is possible that some of these isolates solely exist in association with the environment without displaying any pathogenic qualities. However, it is likely that isolates with virulence abilities are residing on wild plants/non-host crops, but not much is known about the selection pressures that maintain these pathogenic populations in the wild and how this shapes their evolutionary history (Meaden & Koskella, 2017). Taken together this leads to an interesting open question on the evolutionary potential of commensal strains towards specialised/targeted pathogenesis and the reverse scenario of a transition of pathogenic isolates into non-pathogenic saprophytes: what evolutionary steps are involved to acquire/lose pathogenic potential?

Genomic rearrangements, mutations and the acquisition/loss of genomic content via Horizontal Gene Transfer (HGT) play a role in the evolution of pathogenicity and host-pathogen interactions (Arnold & Jackson, 2011; Didelot & Maiden, 2010; Ochman *et al.*, 2000; Yahara *et al.*, 2016). Uptake of genetic material and rearrangements (insertions, deletions) can have a profound impact on the phenotype, including the ability to colonise and infect new hosts (Arnold & Jackson, 2011; Toft & Andersson, 2010). Genes involved in virulence appear to evolve faster compared to the rest of the genome, as was shown in a study of *Ralstonia solanacearum*, the causal agent of bacterial wilt in tomato (Remenant *et al.*, 2010).

4.1.3 Types of Type 3 Secretion Systems and pathogenicity

The T3SS has long been known to confer a key fitness advantage to plant pathogenic bacteria by delivering specific effector proteins (T3SEs) into the plant cytoplasm via a needle-like structure (Alfano & Collmer, 2004; Collmer *et al.*, 2000). The structural and regulatory components of the T3SS are encoded on a 26 kb genomic region, the *hrp/hrc* cluster. This cluster typically consists of the conserved effector locus (CEL), *hrp/hrc* genes and the exchangeable effector locus (EEL) (Alfano *et al.*, 2000) and is referred to as the tripartite pathogenicity island (T-PAI). An atypical type of pathogenicity island (A-PAI) has been observed for naturally occurring *P. syringae* isolates grouping with PG2c, which contains a distantly related *hrp/hrc* cluster missing *hrpK* and *hrpS* (Clarke *et al.*, 2010; Mohr *et al.*, 2008). *P. syringae* carrying the A-PAI do not cause an HR (HR-) in tobacco. The growth of strains with an A-PAI was comparable to the growth of pathogenic isolates in non-host plants (Clarke *et al.*, 2010; Kniskern *et al.*, 2011; Mohr *et al.*, 2008). Environmentally occurring *P. syringae* isolated from wild plants, epilithic biofilms, snow pack and freshwater had a deficient T3SS, by lacking one or more genes from the canonical *hrp/hrc* locus or the CEL (Demba Diallo *et al.*, 2012). However the ability to cause disease is not dependant on a functional canonical *hrp/hrc* locus, as some of the T3SS-deficient isolates also induced symptoms (Demba Diallo *et al.*, 2012). Interestingly, T3SS are also encountered among non-pathogenic *P. fluorescens* isolates (Marchi *et al.*,

2013; Mavrodi *et al.*, 2011; Preston *et al.*, 2001; Rainey, 1999). The T3SS of strain *P. fluorescens* SBW25 is capable of transferring effectors into plants (Preston *et al.*, 2001), but the precise role of the system remains unclear (Jackson *et al.*, 2005).

The number of Type 3 effector (T3E) proteins produced by a strain ranges from about 10 to a total of 40 T3Es (Baltrus *et al.*, 2011; O'Brien *et al.*, 2011), but can be as low as 5 (Hockett *et al.*, 2014). Variability in the amount of effectors produced by strains of even the same pathovar makes determining the host range based on effectors alone very difficult (Almeida *et al.*, 2009; Fujikawa & Sawada, 2016). Interestingly, a deletion mutant of the tomato model pathogen *Pto* DC3000 lacking a single effector (*HopQ1-1*) resulted in a shift in host range by causing disease symptoms in tobacco plants (Wei *et al.*, 2007). There still remains the need to understand how the quantity of T3SE correlates with host adaptation, the ability to cause disease or reach pathogenic population sizes *in planta* (Lindeberg *et al.*, 2009).

Apart from the T3SS as a key virulence factor, *P. syringae* employs several other strategies for infecting their hosts, like the production of antibiotics or phytotoxins. The variations in effector repertoire and the types of phytotoxins produced by *P. syringae* are described in more detail in section 1.4.2.

4.1.4 Factors involved in host range expansion (woody hosts)

The best characterized *P. syringae* strains are pathogenic on herbaceous hosts, like tomato (pv. *tomato*) or beans (pv. *phaseolicola*). Recent studies have focussed on the genomic factors involved in host range expansion of pathogens to woody hosts (Green *et al.*, 2010; Nowell *et al.*, 2016; Rodríguez-Palenzuela *et al.*, 2010). A first report revealed genes specific for adaptation of *P. syringae* pv. *aesculi* (*Pae*) to its woody host: genes involved in the catabolism of aromatic compounds, sucrose utilization genes and a novel gene cluster associated with fatty acid synthesis specific for this woody pathogen (Green *et al.*, 2010).

Aromatic substances like lignin, a major component of wood, are a source of growth substrate for organisms able to degrade aromatic rings (Fuchs *et al.*, 2011). Aromatic compounds are also produced as secondary soluble products (phenols, flavonoid, coumarins, tannins or lignins) by woody plants and these compounds help to protect plants against disease causing organisms (Harwood & Parales, 1996; Pallardy, 2008). Hence, plant-associated bacteria naturally encounter a variety of aromatic compounds produced by the plant during growth, which accumulate due to the plants' inability to make use of these compounds themselves (Fuchs *et al.*, 2011).

In order to cleave and degrade aromatic rings, dioxygenase enzymes are required. These enzymes are predominantly encoded in the chromosome

and part of the metabolic β-ketoadipate pathway. The pathway is ubiquitous in bacteria associated with soil and plays a central role in degrading a wide variety of aromatic compounds to the central metabolites pyruvate, acetaldehyde and β-ketoadipate (Harwood & Parales, 1996). It bears two main branches (Figure 4.1), the protocatechuate branch (degrading phenolic compounds, *pca* genes) and the catechol branch (converting catechol, *cat* genes). Aromatic compounds like benzoate, 4-hydroxybenzoate, vanillate or quinate are hydroxylated to the product protocatechuate, whereas phenols, cinnamate, benzenes and benzoate are hydroxylated to catechol, before being further degraded via the respective pathways (Harwood & Parales, 1996).

The general importance of the catechol pathway has been demonstrated in various bacteria of many genera, including *Pseudomonas* sp., *Ralstonia* sp., *Corynebacterium* sp., *Acinetobacter* sp. (Jiménez *et al.*, 2002; Shen *et al.*, 2005; Tao *et al.*, 2004; Zhan *et al.*, 2008), but it has also been found in fungi including some yeasts (Aranda, 2016; Cook & Cain, 1974; Middelhoven, 1993). *P. syringae* pv. *aesculi*, the causal agent of bleeding canker of horse chestnut, was the first *P. syringae* pathovar shown to possess both pathways, the protocatechuate and the catechol branch (Green *et al.*, 2010), whereas the genome of the bean pathogen encoded only the catechol branch (Joardar *et al.*, 2005). Since then, it was discovered that *P. syringae* pv. *aesculi*, pv. *actinidiae* and pv. *savastanoi*, which fall into PG1 and PG3, appear to possess the set of well-conserved genes involved in the degradation of plant-derived aromatic compounds (protocatechuate branch) (Green *et al.*, 2010; Marcelletti & Scortichini, 2011; Rodríguez-Palenzuela *et al.*, 2010).

After the discovery of a new clade (PG3a) of commensal *P. syringae* globally associated with kiwifruit (Chapter 2) and evidence of their antagonistic behaviour towards *Psa* (Chapter 3), I wanted to determine the evolutionary potential of these strains, especially in regard to evolving from a commensal non-pathogenic resident strain towards being a pathogen. This is especially important in order to avoid the mistake of potentially using these strains as biocontrol agents only to discover that they readily evolve capacity to become pathogenic. The strains characterized in this study were selected from the collection of commensal *P. syringae* PG3 strains isolated from kiwifruit in New Zealand and China.

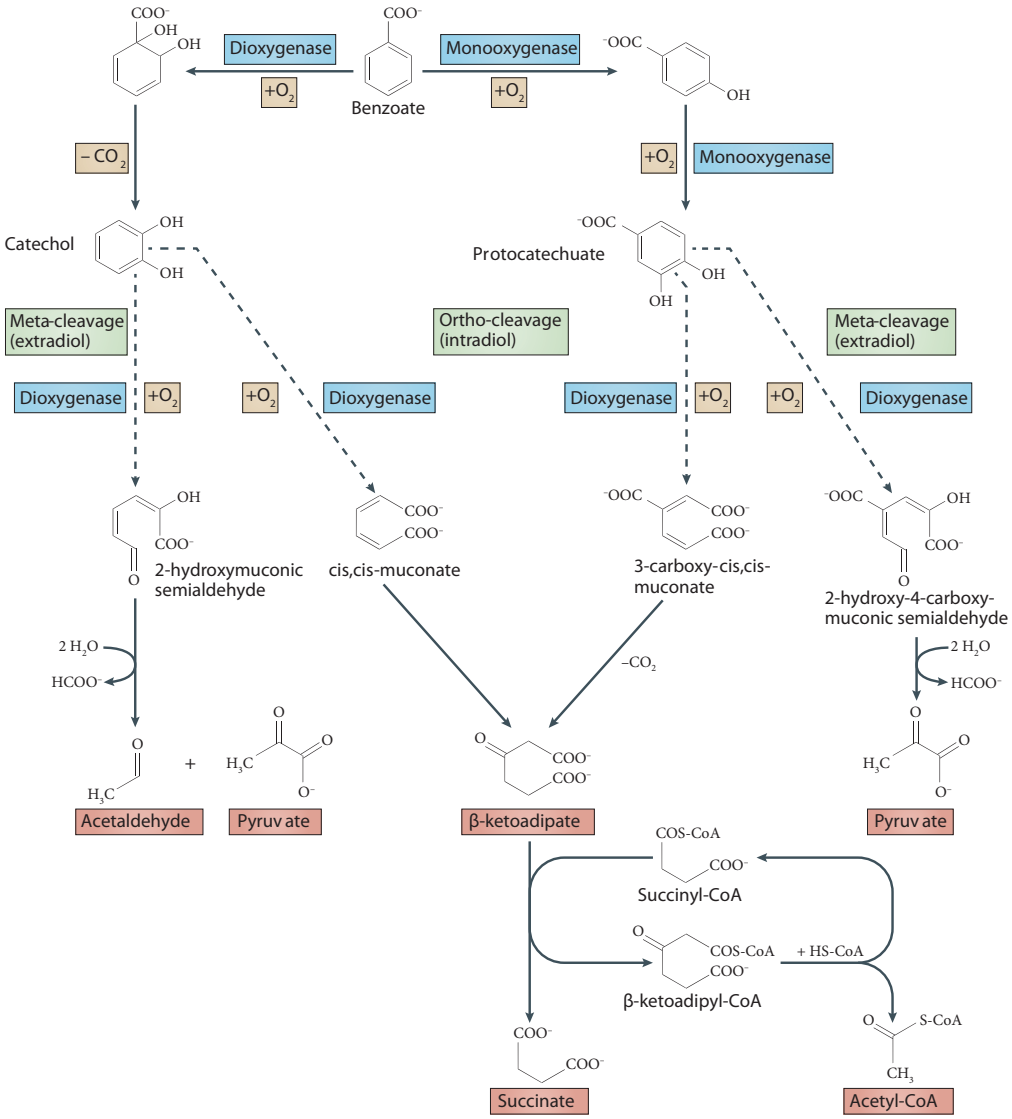

Figure 4.1: Overview of the β-ketoadipate pathway showing the example of degradation of benzoate. The aromatic rings are either cleaved between the two hydroxyl groups (ortho-cleavage, see β-ketoadipate pathway) or ajacent to one of the hydroxyl groups (meta-cleavage) by dioxygenases. Reprinted with permission from Macmillan Publishers Ltd: Nature Reviews Microbiology, Fuchs *et al.* (2011), copyright 2011.

4.2 Aims

The aim of this project was to gauge the evolutionary potential of commensal strains to evolve into pathogens by applying the tools of comparative genomics. The aim was to

1. Determine the fine-scale phylogenetic resolution of *P. syringae* PG3 based on MLST and the core genome
2. Screen T3SS proteins and pathogenicity islands (PAI) and presence/absence to identify host-specific effectors
3. Determine the presence of genes with predicted ability to produce other phytotoxic substances to interfere with the plants' immune system (toxins, plant hormone production)
4. Screen for putative secondary metabolite biosynthesis gene clusters
5. Screen for other genomic factors which might be involved in host range expansion to woody plants (catechol operon) and resistance to certain metallic substances commonly used in an orchard environment.

4.3 Material & Methods

4.3.1 Strain information

A total of 7 NZ strains and 16 Chinese strains of clade 3a (PG3) were chosen for whole genome sequencing (Table S 4.1, appendix). The NZ isolates were recovered from kiwifruit leaf washes, whereas the Chinese strains were isolated from homogenised leaf tissue. The NZ strains chosen for whole genome sequencing are highlighted in grey in Figure 4.2, the ST and corresponding identities can be found in Table S 4.1 (appendix). The Chinese sequences do not have STs listed, as only *gltA* sequences were available.

Isolates were re-streaked on King's B agar and grown for 48 h at 28 °C. A single colony was selected and grown overnight in 5 mL King's B media. Genomic DNA was extracted from 1 mL of overnight culture using the Promega Wizard Genomic DNA Purification Kit following the recommended protocol. Strains were sequenced using the Illumina MiSeq platform (2x300 bp PE; Max-Planck Institute for Evolutionary Biology Plön, Germany and NZGL, Auckland, New Zealand) for NZ isolates and Illumina HiSeq 2500 platform (2x125 bp PE; Novogene, Guangzhou, 451 China) for the Chinese isolates.

4.3.2 Comparison of *in planta* growth of *metA*+ vs. *metA*- strain

The growth assay was performed following the growth assay procedure described in Chapter 3 by individually dip-inoculating 3-week-old *A. chinensis* var. *chinensis* Hort16A plantlets with a concentration of 8x10^7 cfu/mL of the strains *P. syringae* G33C (*metA*$^-$) and NZ50 (*metA*$^+$). The plants were grown over the course of 7 days and the endophytic population size of each strain was counted at 0, 4 and 7 dpi.

Statistical analysis

A Student's *t*-test was used to verify the statistical difference where applicable.

4.3.3 Whole genome assembly

Raw sequence reads were quality trimmed using Trimmomatic v0.36 (Bolger *et al.*, 2014) with the following settings: HEADCROP:10; LEADING:20; TRAILING:20; SLIDINGWINDOW:4:20; MINLEN:50. SPAdes v3.6.2 (Bankevich *et al.*, 2012) was used for error correction and genome assembly using the trimmed PE reads. Assembly quality was improved using Pilon v17 (Walker *et al.*, 2014) and genomes were annotated with Prokka (rapid prokaryotic genome annotation) (Seemann, 2014).

4.3.4 Sequence data

Additional sequencing data were downloaded from the National Center for Biotechnology Information (NCBI) Genbank. A total of 25 PG3 strains from NCBI were included in the analysis, including the closed reference genome of *P. syringae* pv. *phaseolicola* 1448A (Table S4.1, appendix).

4.4 Analysis

4.4.1 Assembly quality assessment

Quality of the draft assemblies was assessed using QUAST v4.3. (Gurevich *et al.*, 2013) and compared to the reference genome of *P. syringae* pv. *phaseolicola* 1448A (*Pph* 1448A) (PG3) and *P. syringae* pv. *actinidiae* NZ54 (*Psa* NZ54) (PG1).

4.4.2 Phylogenetics

4.4.2.1 MLST

Phylogenies were built using the sequences of four MLST housekeeping genes (*gapA, gyrB, gltA* and *rpoD*) typically used for population genetics work. However the full length (7,106 bp) of all four loci was used (compared to sub-fragments totalling 2010 bp used for standard MLST),

which were extracted from the draft genomes. Prior to tree building, the best fitting model of evolution was determined using jmodeltest2 (Darriba *et al.*, 2012). Maximum Likelihood trees were built using TREEPUZZLE v5.3 (Schmidt *et al.*, 2002) with 100,000 puzzling steps, a neighbour-joining tree for parameter estimation and quartet puzzling as tree search procedure. Clonalframe v1.1 (Didelot & Falush, 2007) was used for Bayesian inference using 500,000 MCMC and a burn-in of 50,000 iterations.

4.4.2.2 Variant calling

DNA sequence reads were mapped to the reference genome of *P. syringae* pv. *phaseolicola* 1448A (Joardar *et al.*, 2005) to determine the core genome using the default settings for paired-end read-mapping in bowtie2 v2.2.9 (Langmead *et al.*, 2009). For assembled genomes from NCBI, paired-end Illumina HiSeq 2500 sequencing reads (150 bp read length, 500 bp fragment size, fold coverage 20, standard deviation of insert size 10) were simulated using ART (Huang *et al.*, 2012).

SAMtools v1.3.1 (Li *et al.*, 2009) (mpileup –ugf) was used to calculate likelihood of data given each possible genotype. The genomic variant calling was done using bcftools v1.3.1 (call –consensus-caller –output-type v –ploidy 1), followed by a filtering step using VCFtools v0.1.13 (Danecek *et al.*, 2011) (–minDP10 –non-ref-af 0.95 –minQ 20 –removeindels – recode –recode-INFO-all). BEDTools v2.26.0 (Quinlan & Hall, 2010) was used to filter genomic regions with coverage of less than 10. Variant calling included indels and multiple nucleotide insertions as well as single nucleotide insertions; only

SNPs were retained for downstream phylogenetic analyses. The consensus alignment was created by removing regions in the genome with coverage of less than 10 and a quality score under 20.

Recombination analysis and phylogenetic reconstruction

For removal of recombinant regions in the core genome alignment, ClonalFrameML (Didelot & Wilson, 2015) was run using default settings and using a NJ phylogenetic tree generated by Geneious v7.1.7 as starting tree. A Maximum Likelihood phylogenetic tree of 48 PG3 strains comprising NZ and Chinese kiwifruit isolates and strains reflecting the diversity of PG3 was built using RAxML v7.2.8 (Stamatakis, 2006). A total of 1,511,224 bp recombinant positions were removed from the core genome alignment of 1,644,402 bp, resulting in a non-recombinant core genome alignment of 133,178 bp including 5,886 variant sites. The tree was rooted on the most divergent member: C58. Trees were built with the generalized time-reversible model and gamma distribution of site-specific rate variation (GTR+I) and 100 bootstrap replicates.

4.4.3 Identification of the Pangenome

The program Roary v3.6.1 (Page *et al.*, 2015) was used to determine the core and flexible genome when comparing kiwifruit *P. syringae* isolates of PG3a with (a) all other kiwifruit *P. syringae* of PG3 and the reference genome of *Pph* 1448A and (b) compared to *Psa* NZ54. Core genes are present in

99-100% of isolates, soft core genes in 95-99%, shell genes in 15-95% and the cloud genes in 0-15%.

4.4.4 Type Three Secretion System and Effectors

For determination of absence or presence of T3S effectors among the strains, the protein sequences for each effector were downloaded from www.pseudomonas-syringae.org (March 01, 2017). This resulted in a database of 751 different protein sequences for a total of 85 effectors. A custom BLAST database was created with all PG3 strains used in this study and the effector protein sequences were blasted against this database using tBLASTn (evalue 10^{-5}). Presence was confirmed when the sequence similarity was >80% for >80% of the query. Effectors were reported as truncated when the sequence similarity was >80% for <80% of the query sequence. Pathogenicity islands were determined using the genomic location of key effector genes typically located in the *hrp/hrc* cluster, CEL and EEL.

4.4.5 Phytotoxins

P. syringae strains produce a variety of phytotoxins and the genomes were screened for the presence of pathways of a variety of toxins: phaseolotoxin, coronatine, syringomycin, syringopeptin, tabtoxin. In addition the genomes were searched for syringolide and genes encoding for the plant hormones ethylene and auxin. Protein sequences for genes involved in these pathways (Table 4.1) were downloaded from NCBI and blasted against the genomes

using tBLASTn (cut-off evalue 10^{-5}). Presence of genes was confirmed with a query coverage and identity of >80%.

Table 4.1: Phytotoxin screen.

	Gene		**Encoding for**
Syringolide	*avrD*	avirulence gene D	production of syringolide
Ethylene production	*efe*	ethylene-forming enzyme	ethylene-forming enzyme
Coronatine	*COR*	coronofacic acid synthetase	coronofacic acid synthetase
Auxin inactivation	*iaaL*	indoleacetic acid-lysine	IAA-lysin ligase
Auxin production	*iaaM*	tryptophan monooxygenase	indole-acetic acid biosynthesis
	iaaH	indoleacetamide hydrolase	
Tabtoxin	*tabA*		
Phaseolotoxin	*argK*	phaseolotoxin-insensitive ornithine carbamoyltransferase	resistance to phaseolotoxin
Syringomycin	*syrB1*	Syringomycin synthethase	
	syrB2		synthesis
	syrC	Syringomycin synthethase	
	syrE		
	syrD		secretion
	syrP		regulation
Syringopeptin	*syrD*		secretion
	sypA	Syringopeptin	
Syringolin	*sylA*	syringolin A	
	sylB	syringolin B	
	sylC	syringolin C	synthesis
	sylD	syringolin D	
	sylE	syringolin E	

4.4.6 Secondary metabolite biosynthesis genes

The phytotoxin screen described in the previous section was purely based on BLAST similarity searches and any potential new candidates might have been missed. Hence all genomes were screened for any putative secondary metabolite biosynthesis gene clusters (non-ribosomal peptide synthetases, NRPS such as toxins or siderophores) using antiSMASH v3.0 (Weber *et al.*, 2015). antismash detects known gene clusters by comparing to a database, but also detects novel gene clusters by looking for frequencies of observed PFAM domains in and outside of known BGC (biosynthetic gene cluster) clusters using ClusterFinder. Analysis options included the ClusterFinder using probabilistic BGC detection, a minimum cluster size in CDS and biosynthesis-related PFAM domains set to 5 and the threshold set to 60%. However, NRPS are long proteins containing many repeated units, which potentially could be difficult to find, when the genome assemblies are of inferior quality.

4.4.7 Copper resistance

Genes conferring resistance to copper were first discovered in *Psa* isolates in New Zealand in 2014, whereas previously isolated *Psa* were copper sensitive. Copper resistance in *Psa* was conferred by integrative conjugative elements and plasmids carrying copper resistance genes of both the copper efflux (*cusABC*) and copper sequestration (*copABCD*) systems (Colombi *et al.*, 2017). The genomes of kiwifruit commensals were screened for the

presence of the copper efflux (*cusABC*) and/or copper sequestration (*copABCD*) systems and if present the genomic organisation was illustrated.

4.4.8 Catechol operon

Genomes were searched for five genes previously used to determine the presence/absence of the catechol operon in *Pseudomonas syringae* (Bartoli *et al.*, 2015): Phenol degradation protein (*metA*), Catechol 1,2 dioxygenase, FAD-dependent-oxidoreductase, dienelactone hydrolase and short chain dehydrogenase. tBLASTn searches were done using genes extracted from *Psa* NZ13 (GenBank CP011972.2) and *P. savastanoi* pv. *aseculi* (GenBank: ACXT01000012).

4.5 Results

From the collection of kiwifruit-isolated *P. syringae* strains resulting from the MLST study presented in Chapter 2, several representatives of PG3 were chosen for whole genome sequencing (Figure 4.2), along with a subset of Chinese kiwifruit *P. syringae* isolates identified as PG3. All Chinese strains were isolated from kiwifruit with the exception of strain C46 (PG3a), which was isolated from an orchard, previously growing kiwifruit and thought to have become infected by *Psa* in 1985-86, which is now growing tea (*Camellia* sp.). Few *Actinidia* sp. were observed to grow on the margins of the orchard.

4.5.1 Quality of the draft genomes

The number of contigs for the draft genomes ranged from 51 to 118, with a N50 value of up to 578,625 (Table 4.2). All draft genomes were aligned to the most closely related fully closed reference genome available: *P. syringae* pv. *phaseolicola* 1448A (*Pph* 1448A). The percentage of genome being successfully aligned to *Pph* 1448A ranged from 55.6% to 83.1% (average 79%). Aligning the PG3 draft genomes to the draft genome of *Psa* NZ54 revealed an average of 67.9% aligned to *Psa* NZ54, ranging from 51.6% to 84.5% (including unaligned fragments in otherwise aligned contigs). However, these statistics depend highly on the quality of the genome assembly.

Table 4.2: Assembly statistics for NZ and Chinese PG3a strains. * including partially unaligned contigs

	C25	**C32**	**C37**	**C39**	**C40**	**C41**	**C42**	**C46**	**C49**	**C50**	**C56**	**C57**
# contigs	2003	80	220	71	102	53	64	93	100	51	166	97
Largest contig	555,411	700,579	1,062,581	814,248	1,053,138	1,358,347	750,139	331,911	680,924	1,246,926	719,564	963,063
Total length	7,844,274	6,110,836	6,156,507	5,954,232	6,028,468	5,763,156	5,914,177	5,957,168	6,047,550	6,608,276	5,943,750	5,887,108
N50	178,031	239,967	363,383	264,939	428,979	298,651	330,867	167,111	159,596	578,625	192,143	192,143
GC (%)	57.25	58.37	57.78	58.22	58.06	58.39	58.24	58.42	58.35	58.19	58.42	58.52
% aligned to *Pph* 1448A	61.4	79.2	82.9	85	83.7	88	86	79	80.4	84.4	81.2	82
% aligned to *Psa* NZ54*	53.7	67.7	72	84.5	82.9	80.2	83.6	56.9	53.3	79.2	61.9	66.6

	C58	**C59**	**C64**	**C65**	**NZ50**	**NZ51**	**NZ55**	**NZ56**	**NZ57**	**NZ58**	***P.s.* G33C**
# contigs	59	73	84	89	51	58	71	106	118	62	192
Largest contig	937,503	868,485	586,930	586,822	838,741	1,102,065	553,495	591,022	373,695	1,120,796	496,864
Total length	6,177,965	5,943,754	5,999,344	5,999,538	6,015,443	6,148,182	6,085,667	6,014,717	6,010,405	5,922,456	6,016,813
N50	382,084	188,922	236,581	295,061	228,020	363,770	288,627	180,715	140,612	287,427	235,653
GC (%)	58.35	58.48	58.42	58.4	58.46	58.44	58.45	58.44	58.46	58.49	58.55
% aligned to *Pph* 1448A	55	80.6	80.8	80.8	79.5	77.7	79.4	80.3	79	80.7	79.1
% aligned to *Psa* NZ54*	72.6	58.3	64.3	65.4	62.2	74.4	68.7	58.7	51.6	79.4	63.8

4.5.2 Phylogenetic reconstruction

4.5.2.1 MLST

Phylogenies were built based on the complete gene sequences of the four housekeeping genes. The best fitting evolutionary model identified by jmodeltest was TIM1+I+G. Two different approaches were used, Maximum Likelihood and Bayesian inference. The program Clonalframe, which removes the part of the alignment which has undergone recombination, was used as Bayesian approach and it was found that Maximum Likelihood and Bayesian inference were congruent. A subset of isolates collected from kiwifruit in China and New Zealand form a unique subclade (3a) within PG3, based on the full coding sequences of *gapA, gyrB, gltA* and *rpoD* (Figure 4.3). This confirms the results from the MLST study whilst using more diverse PG3 representatives. The most closely related pathovars to PG3a were pv. *myricae* (isolated from Chinese bayberry), pv. *raphiolepidis* (isolated from Japanese hawthorn) and pv. *eriobotryae* (isolated from loquat). All of these plants belong to the families *Myrica* and *Rosaceae*, which are small trees and shrubs, mostly native to Asia.

4.5.2.2 Core genome tree

In order to obtain greater resolution of the relationship between PG3 isolates and the new clade 3a of kiwifruit strains, the core genome of *P. syringae* PG3 was identified.

Read-mapping and variant calling of the 48 *P. syringae* PG3 genomes resulted in a core genome alignment of 1,644,402 bp including 149,925 variant sites (88,931 within PG3a). Clonalframe ML was run to identify any recombinant sites, to minimize the effect of recombination on phylogenetic reconstruction. A total of 1,511,224 recombinant positions were identified, which resulted in a nonrecombinant core genome of 133,178 bp with 5,886 variant sites (2,875 within PG3a). Membership within PG3 corresponds to a minimum of 98.65% pairwise identity and for clade 3a to 98.98%. The ML tree (Figure 4.4) is based on a 133,178 bp core genome alignment and clearly shows that PG3 is a very diverse phylogroup, with two clades. The kiwifruit isolates from NZ and China form a separate clade (3a), whereas all other strains form the second clade.

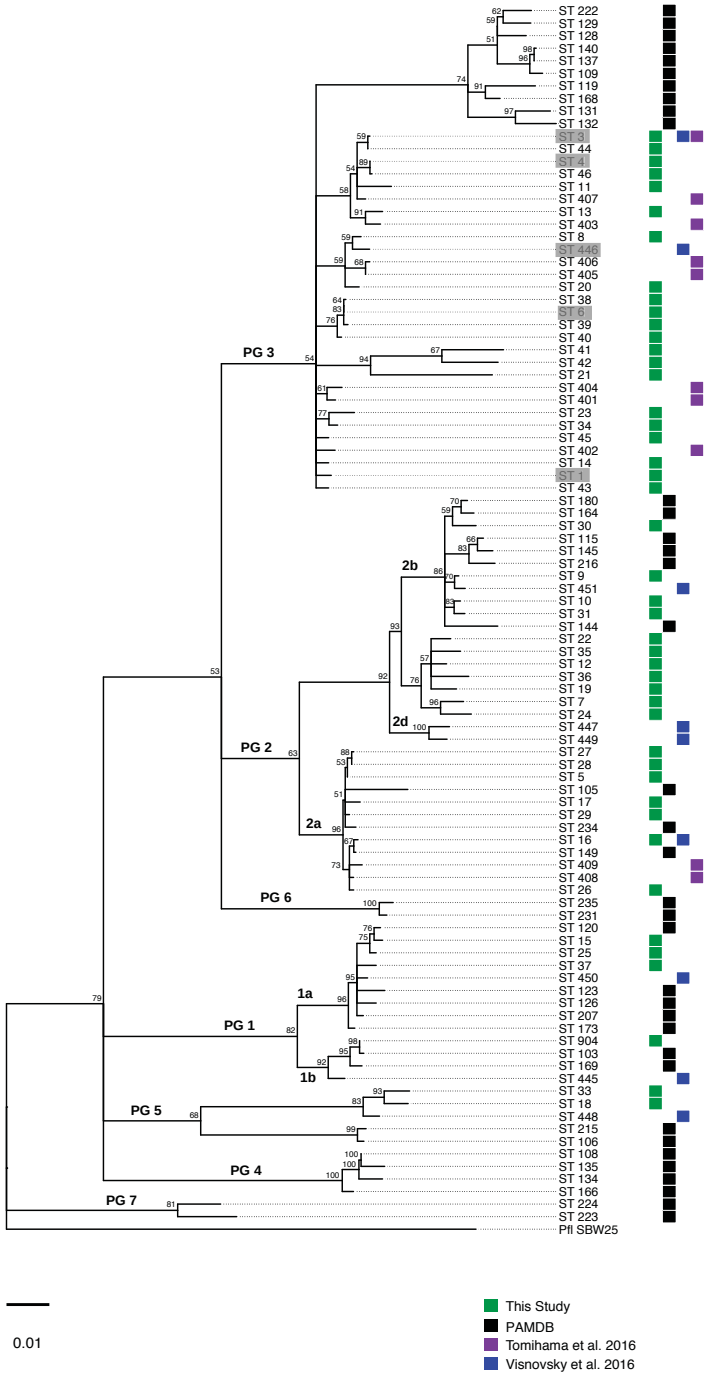

Figure 4.2: ML tree of *P. syringae* as illustrated in Chapter 2 based on the concatenated alignment (2010 bp) of *gapA*, *gyrB*, *gltA* and *rpoD*. Sequence types of NZ isolates in PG3 chosen for whole genome sequencing are shaded in grey, the corresponding names can be found in Table S 2.1 (appendix).

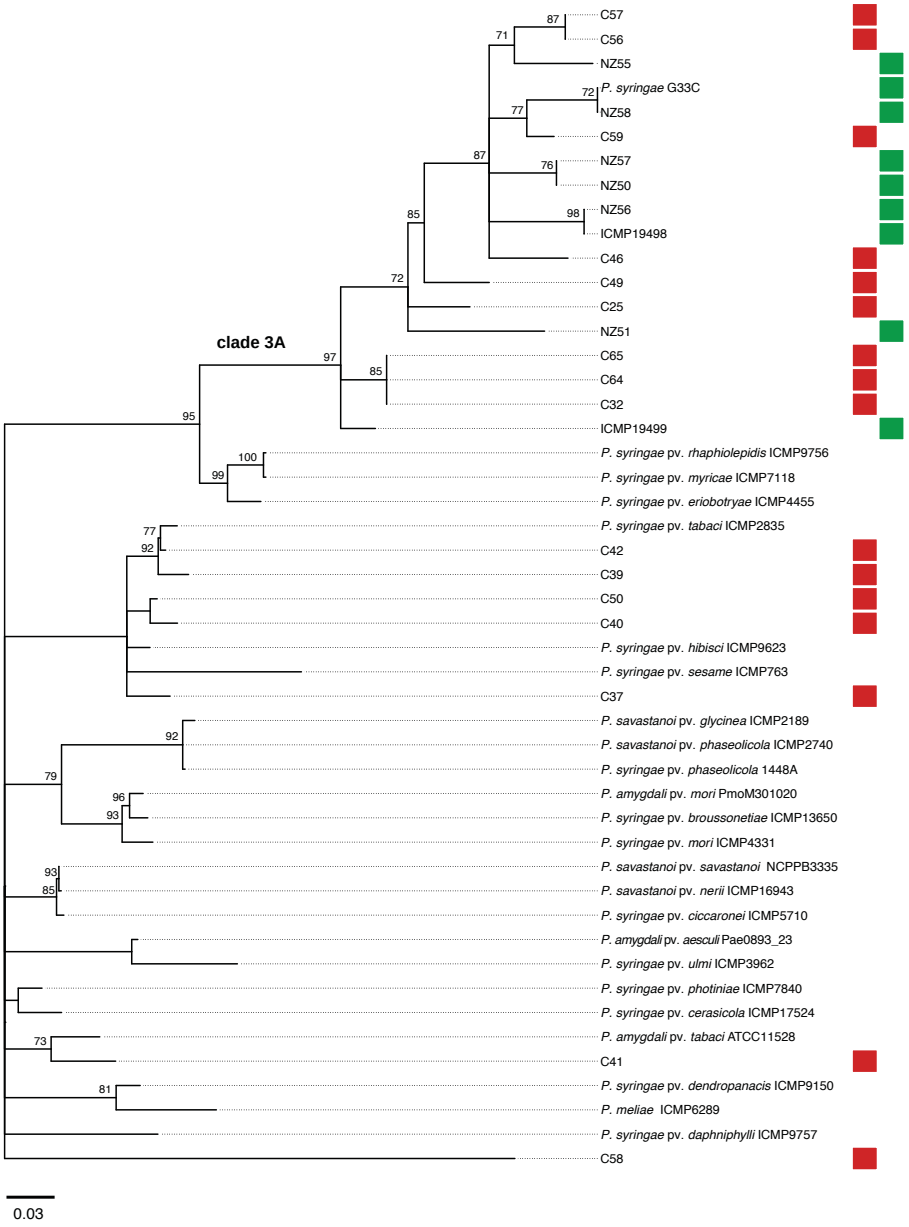

Figure 4.3: Phylogeny of *P. syringae* PG3 based on full coding sequences of four housekeeping genes (MLST). ML tree constructed using TREEPUZZLE (GTR model) based on the concatenated alignment of the full coding sequence for *gapA, gyrB, gltA*and *rpoD* (7,120 bp). The most divergent strain, C58, was used as an outgroup and the kiwifruit isolates are colour coded according to their sampling location (red = China, green = New Zealand).

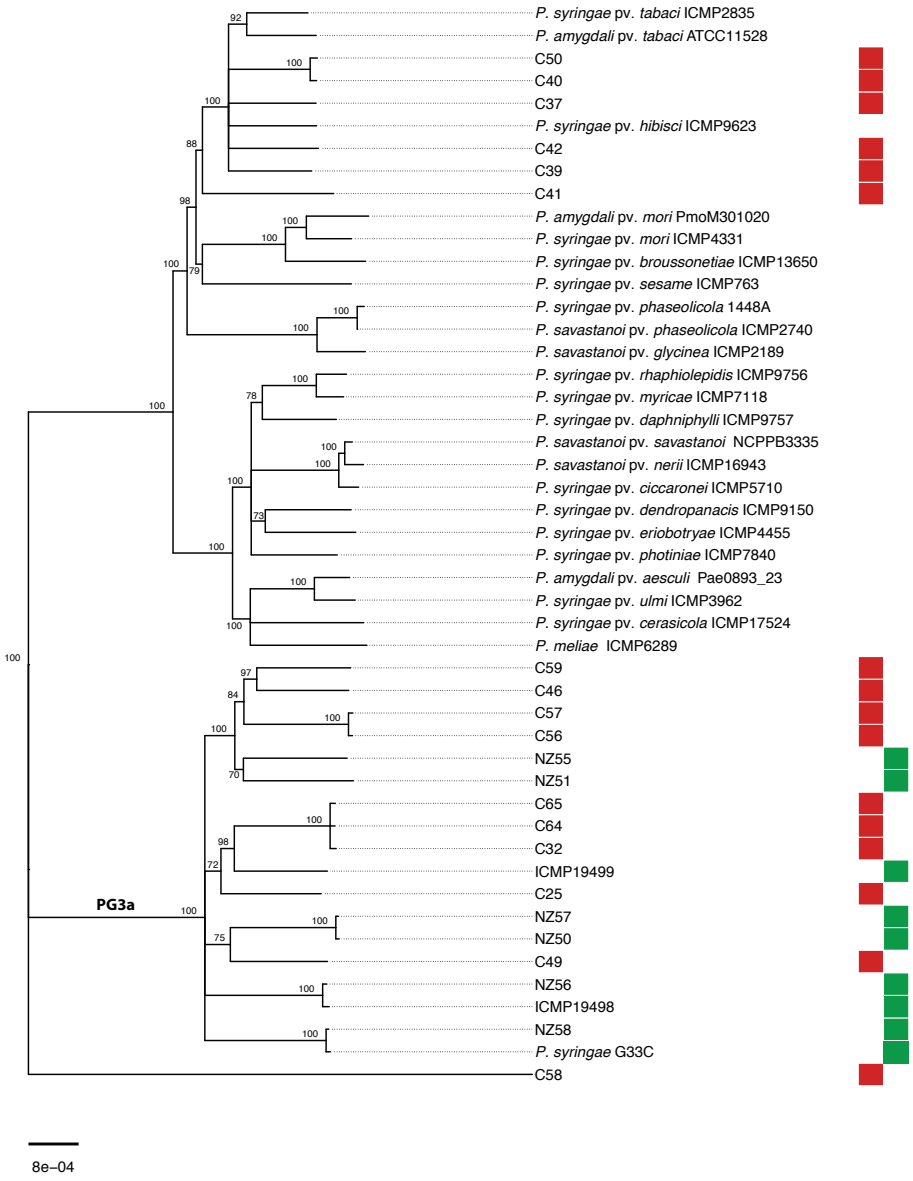

Figure 4.4: Phylogeny of *P. syringae* PG3 based on the core genome. RaxML Maximum Likelihood tree based on a nonrecombinant core genome alignment of 133,178 bp including 5,886 variant sites. Bootstrap values (>70%) are indicated at the nodes. The most divergent strain, C58, was used as an outgroup and the kiwifruit isolates are colour coded according to their sampling location (red = China, green = New Zealand).

4.5.3 Unique and shared genes for Phylogroup 3a

Comparing the genomes of kiwifruit *P. syringae* isolates falling into PG3a (n=18) with other PG3 kiwifruit *P. syringae* strains (n=7) and *Pph* 1448A, revealed 4,167 core genes comprising of 496 soft core (present in 95-99% of strains) and 3,671 core genes (present in 99-100% of strains), which equates to 28.7% of the pan genome (14,540 genes). The same analysis was repeated comparing PG3a to a single reference of PG1 (*Psa* NZ54) and revealed 2,223 core genes (16.2% of the pan genome; 13,769 genes) (Figure 4.5). The shell genome (present in 15-95% of strains) is twice as large for the comparison with *Psa* NZ54, with a noticeable absence of a soft core gene cluster. However, the number of unique (cloud genes), which are only present in 1-15% of isolates, is bigger for PG3 vs. PG3a. Despite a larger amount of core genes within PG3, there is still a considerable amount of variation.

A study investigating genome evolution in *Psa* found that the core genome of 60 *Psa* biovar 3 isolates amounted to 50.5% of the pan genome (McCann *et al.*, 2017), whereas a comparative analysis of *P. syringae* across phylogroups estimated a core genome of 48% in an average *P. syringae* genome (20.6% of the total pangenome) (Nowell *et al.*, 2016). These analyses are, however, biased by sequencing and assembly quality and annotation errors.

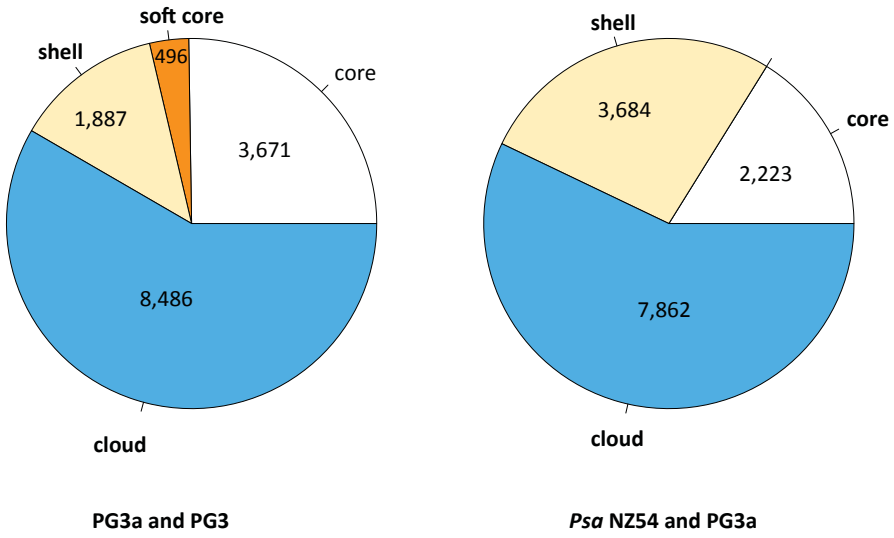

Figure 4.5: The core and flexible genome of *P. syringae* PG3a when compared to *P. syringae* PG3 (*Pph* 1448A) or *Psa* NZ54.

4.5.4 Type 3 Secretion System

4.5.4.1 Effector repertoire

The effector screen revealed differences in the effector repertoire within *P. syringae* PG3. Among the conserved subclade 3a, which is solely found associated with kiwifruit, only a limited number of T3E families is present: *avrE, hopA, hopAA, hopAB, hopAH, hopI, hopM, hopV* and *hopZ3* (Table 4.3), with *avrE, hopI, hopAA* and *hopM* being the core effector families present in each genome sequence (Figure 4.6). Three of these conserved effectors (*avrE, hopM* and *hopAA*) are located in the gene cluster CEL (conserved effector locus) (Debroy *et al.*, 2004), which plays a role in suppressing basal plant immunity and is conserved among *P. syringae*

(Badel *et al.*, 2006; Debroy *et al.*, 2004). Noticeably, the strains that possess a shorter allele of *hopAA1* (186aa) compared to the typical 486 aa, show the presence of additional effectors (*hopA1* and *hopAB3*), but also lack *hopZ3*. *HopZ3* is encoded in the EEL, in the same genomic location as *hopA1*. The remaining isolates of PG3 show a much higher number of effectors, with an average of 20 effectors per strain, which highlights the significant reduction in effectors for clade 3a. Other studies reported the lowest number of effectors previously for *P. syringae* PG2 with an average of 13, whereas PG1 was the most diverse with 35 and 29 for PG3 (Nowell *et al.*, 2016). *P. syringae* G33C, which was used in the competition experiments in Chapter 3, possesses six effectors (*avrE1, hopI1, hopM1, hopZ3, hopAA1, hopAH1*), which can all be found in *Psa* NZ54 as well, but *Psa* NZ54 harbours an additional 28 effectors (34 total).

Table 4.3: Activity, plant target and localisation of effectors found in Phylogroup 3a.

T3E	**Cellular function affected**	**Pathovar**	**Target**	**Subcellular localisation**	**Reference**
avrE1	Promote lesion formation	*Tomato* DC3000	Cya translocator reporter	unknown	Badel *et al.* (2006)
hopA1	Regulation of basal resistance		Enhanced Disease Suceptibility 1 (EDS1)	cytoplasm	Bhattacharjee *et al.* (2011)
hopAA1	Lesion formation	*Tomato* DC3000	unknown	mitochondria	Munkvold *et al.* (2008, 2009)
hopAB3	unknown	unknown	unknown	unknown	
hopAH1	unknown	unknown	unknown	unknown	
hopI1	Suppression of salicylic acid accumulation	*Maculicola* ES4326	Heat shock protein Hsp70	chloroplast	Jelenska *et al.* (2010)
hopM1	Host vesicle trafficking pathway	*Tomato* DC3000	AtMIN7	endomembrane	Nomura *et al.* (2006)
hopV1	unknown	unknown	unknown	unknown	
hopZ3	unknown	*Syringae* B728a	Unknown	Non-discrete	Lewis *et al.* (2008)

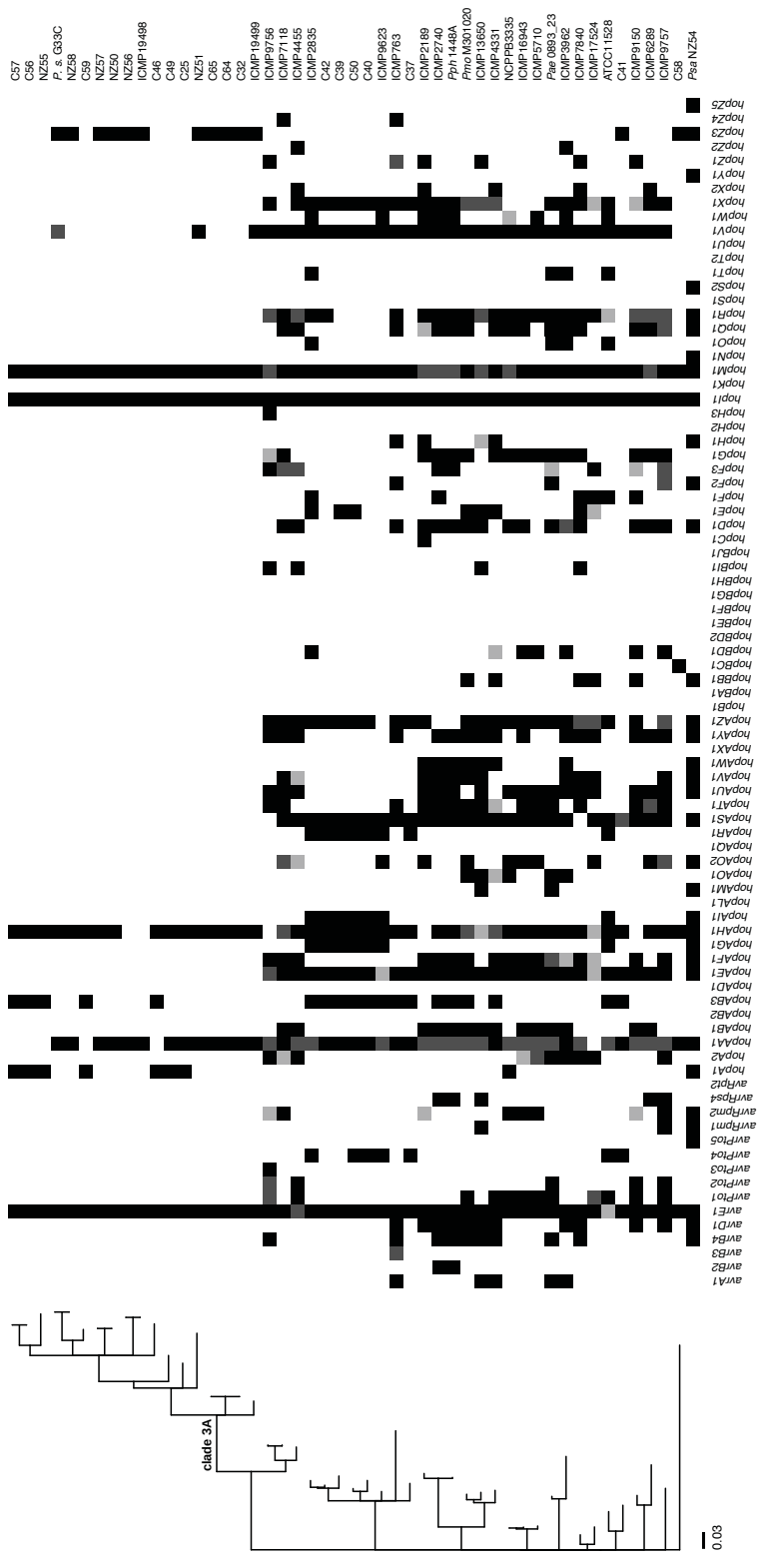

Figure 4.6: T3S effector repertoire. Presence of effectors is indicated with a black box, truncated versions (contig break) with a light grey box, truncated effectors with a dark grey box, and white depicts absence. The phylogenetic tree corresponds to the ML tree based on MLST, as shown in Figure 4.3.

4.5.4.2 Pathogenicity islands

The T3SS pathway of *P. syringae* is encoded on so-called pathogenicity islands (PAI). A few different types of PAI's have been described, with the "traditional" and most common pathogenicity cluster being referred to as the tripartite pathogenicity island (T-PAI), but other types such as the atypical (A-PAI) are present in pathogenicity-destitute isolates from PG2c, or the rhizobium PAI (R-PAI), which can be found as an additional T3SS (Baltrus *et al.*, 2017).

A search for the core effectors *avrE, hopM* and *hopAA* present in all commensal strains revealed a PAI (50 kb) located between *gstA* (glutathione S-transferase) (downstream) and tRNA/*queA* (upstream). It was extracted from all PG3 kiwifruit strains and compared to the T-PAI of *Pto* DC3000 (PG1), *Pph* 1448A (PG3) and *Psa* NZ54 (PG1). The alignment revealed that the PAI's found in commensal kiwifruit *P. syringae* show the typical T-PAI structure of conserved effector locus (CEL), *hrp/hrc* cluster and exchangeable effector locus (EEL). The PAI's in PG3 are more closely related to the *Pph* 1448A-PAI (78% pairwise nucleotide identity) than *Pto* DC3000-PAI (59%) or *Psa* NZ54-PAI (56%). Several clusters of kiwifruit isolates within PG3 share 100% identical PAIs at the nucleotide level: ICMP19498, NZ56, NZ50 and NZ57; C56 and C57; C32, C64 and C65, G33C and NZ58; C39, C40 and C50.

As shown in Figure 4.7, the CEL is a 24.6 kb region downstream of *hrpR* and comprises the conserved effector genes *avrE1, hopM1* and *hopAA1*

(93% identity at the nucleotide level). The *hrp/hrc* cluster stretches 21.9 kb downstream from *hrpR* to *hrpL* and is organized identically to *Pto* DC3000, *Psa* NZ54 and *Pph* 1448A, except for a 1 kb insertion between *hrpV* and *hrcU* for *Psa* NZ54 (copper translocating P-type ATPase) and some kiwifruit strains (hypothetical protein, 100% amino acid identity, E-value 1e-110).

The EEL resides between *hrpK* a $tRNA^{Leu}$, contain various effector genes (*hopZ3* and *hopA1*) and ranges from 6.7 kb to 50.5 kb in size. This large variation in size is explained by the insertion of apparent phage genes (from 43 to 46.9 kb) just downstream of $tRNA^{Leu}$ for six strains (NZ 50-51, NZ56-57, C58 and ICMP19498). A BLAST search using PHASTER (http://www.phaser.ca) (Arndt *et al.*, 2016; Zhou *et al.*, 2011) confirmed the presence of intact prophage regions (head, integrase, lysis, portal, capsid, plate, recombinase, tail) including chromosomal attachment sites (*attR* and *attL*) in four strains.

Expression of T3SS genes is mediated by an alternative sigma factor (HrpL). HrpL requires binding by two DNA binding proteins HrpR and HrpS, which activates transcription of *hrpL* (Hutcheson *et al.*, 2001; Ortiz-Martin *et al.*, 2010; Xiao *et al.*, 1994). The *hrpRS* operon, as well as *hrpL* are present in the *hrp/hrc* cluster in the typical genomic organisation in all of the kiwifruit *P. syringae* isolates.

The highly conserved order of T3SS genes in the PAI suggests that the T3S system in *P. syringae* PG3 strains isolated from kiwifruit could be fully functional, however this needs to be demonstrated experimentally.

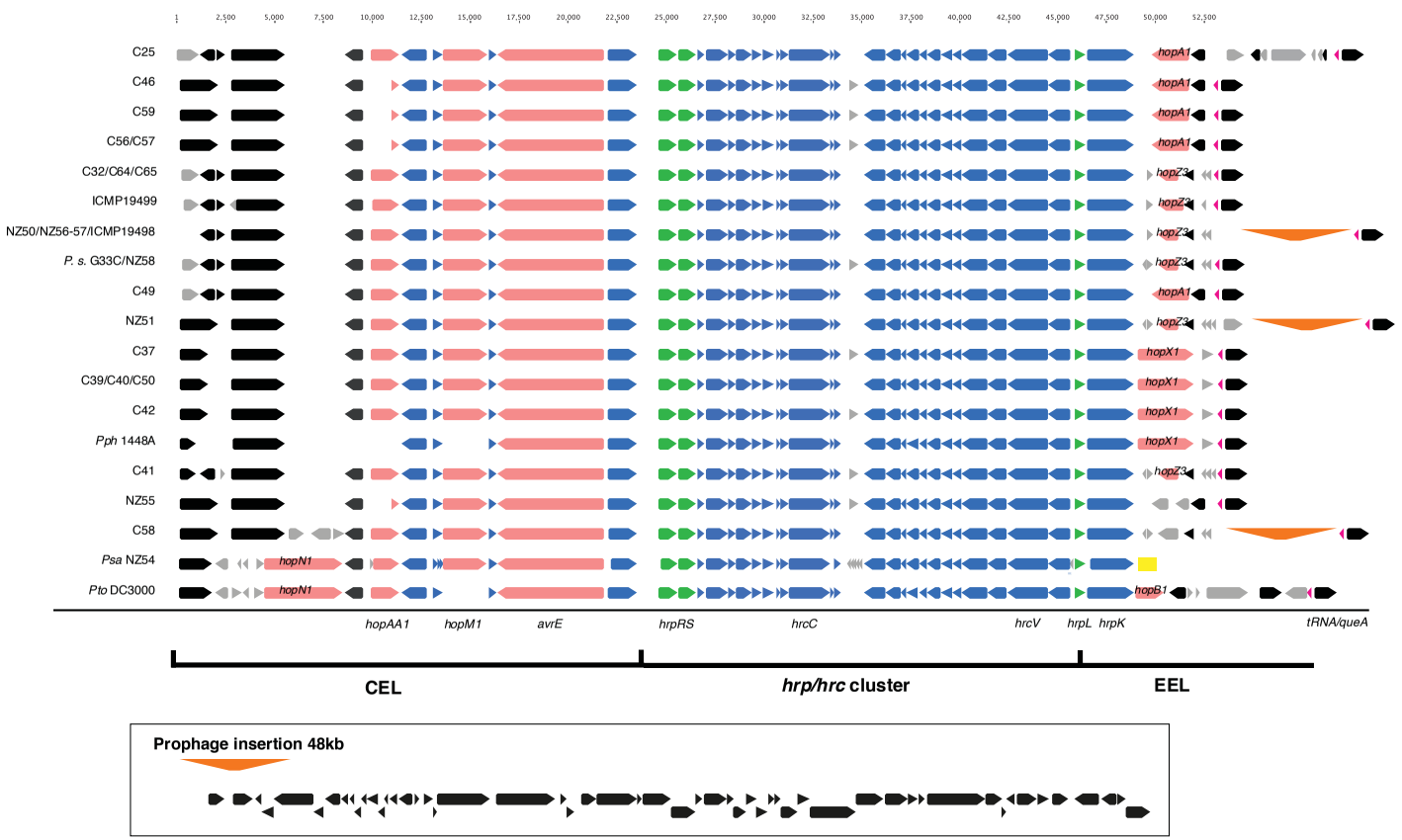

Figure 4.7: Genetic organisation of PAIs found in PG3, compared to *Pph* 1448A, *Pto* DC3000 and *Psa* NZ54. Red boxes depict effector genes, blue boxes are *hrp/hrc* genes, regulatory genes are depicted in green boxes, pink boxes are tRNA, black boxes denote other genes, a yellow box reveals contig breaks and prophage insertions are shown as an orange box. The 48 kb prophage insertion sequence is shown in the box.

4.5.5 Phytotoxins

The toxin screen revealed diversity in the phytotoxin repertoire among PG3 strains. Only one of the tested pathways was conserved across all strains: the ability to produce the plant hormone auxin (*iaaM* gene), which influences bacterial growth and expression of other virulence factors and is conserved across various *Pseudomonas* spp. (Glickmann *et al.*, 1998). Interestingly, all strains isolated from kiwifruit lacked the *iaaL* gene, which inactivates auxin, along with the two representatives of *P. syringae* pv. *phaseolicola*, whereas all other pathovars possess *iaaL* to modify IAA levels. Tabtoxin was not discovered in any strains apart from pv. *tabaci*, for which it was primarily studied (Kinscherf *et al.*, 1991; Kinscherf & Willis, 2005). Phaseolotoxin is produced by *P. syringae* pv. *phaseolicola* and some biovars of *P. syringae* pv. *actinidiae*. *Psa* biovar 1 and 6 produce phaseolotoxin, biovar 2 strains produce only coronatine and biovar 3 strains (current pandemic) produce neither coronatine nor phaseolotoxin (Cunty *et al.*, 2015). *Psa* NZ54 is a biovar 3 strain, hence neither coronatine nor phaseolotoxin is produced by this isolate.

Pathways for ethylene production (ethylene-forming enzyme) were found only in pvs. *sesame* and *glycinea*. Genes involved in coronatine production were discovered in pvs. *ulmi, aesculi* and *glycinea*.

Interestingly, there appears to be a trade-off between the number of T3SE and toxin production, as those strains that showed a limited number of effectors in the effector screen (clade 3a and strain C58, Figure 4.6)

possess the pathways to produce syringomycin and in some cases syringolin. This negative correlation was shown previously for PG2, where syringolin A, syringopeptin and syringomycin were conserved for strains showcasing a limited number of effectors (Baltrus *et al.*, 2011; Hockett *et al.*, 2014).

The protein identity among the tblastN searches for syringomycin was quite high: *syrB1* (91%), *syrB2* (99%), *syrD* (96%), *syrP* (93%). Only *syrC* showed solely 75% identity. The protein identity for *syrE* was 92% for >80% coverage. The syringomycin gene cluster organisation (Figure 4.8) was conserved for all strains possessing the pathway and showed the same genomic organisation as previously described (Scholz-Schroeder *et al.*, 2001). Genes encoding for syringopeptin synthetase (*sypABC*) are usually transcribed downstream of the syringomycin cluster (Scholz-Schroeder *et al.*, 2001). The PG3a strains show some similarities in blast hits of genes adjacent to *syrD*, but none are highly similar, like it was found for syringomycin protein identities.

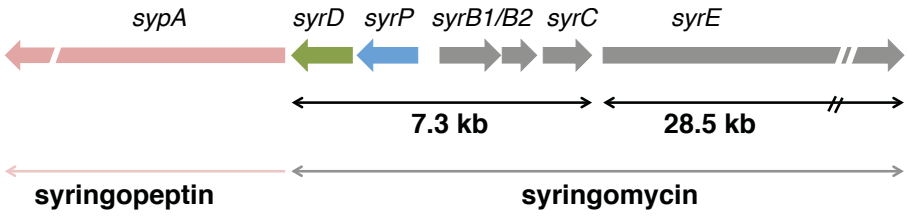

Figure 4.8: Genetic organisation of the syringomycin/syringopeptin gene cluster among PG3 isolates. Syringomycin synthesis genes are coloured grey, regulatory genes in blue and secretion genes in green.

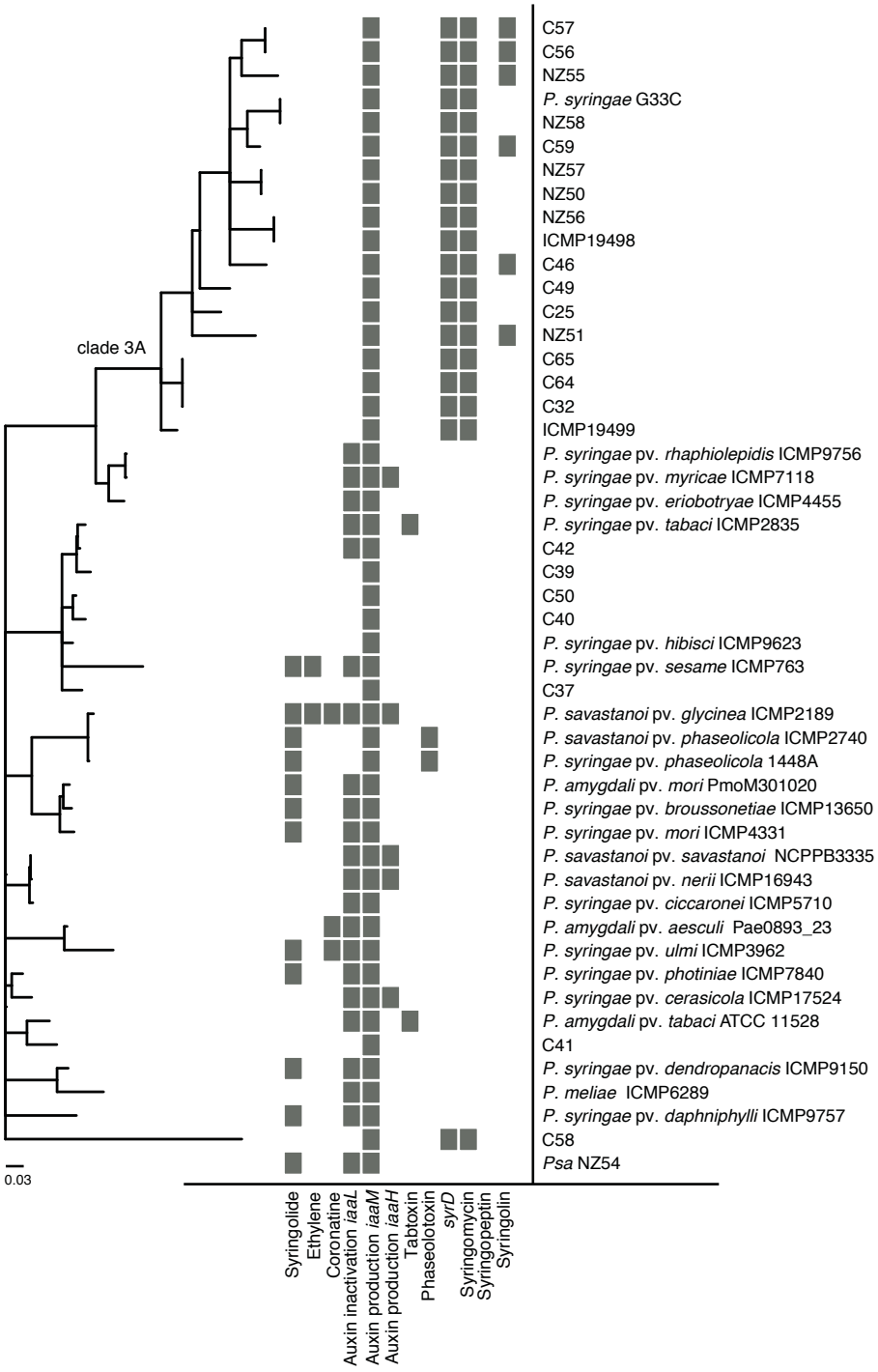

Figure 4.9: Presence/absence of genes involved in phytotoxin production in *P. syringae* PG3. Grey boxes indicate presence of full-length gene, whereas white indicates absence. The phylogenetic tree corresponds to the ML tree based on MLST, as shown in Figure 4.3.

4.5.6 Secondary metabolite cluster mining

The screen using antiSMASH v3.0 for automated identification of putative secondary metabolite biosynthesis genes in kiwifruit isolates revealed a number of clusters with at least 40% similarity. A few clusters were found in all isolates: an arylpolyene (APE) cluster, which encodes for typically yellow pigmentation (Cimermancic *et al.*, 2014); alginate, a common exopolysaccharide and major virulence factor of *Pseudomonas aeruginosa* that has been associated with epiphytic fitness of *P. syringae* (Fakhr *et al.*, 1999; Keith *et al.*, 2003); the biosurfactant syringafactin which increases tolerance to fluctuating water availability on the leaf (Burch *et al.*, 2014); and pseudopyronine B, an antimicrobial compound (Nishanth Kumar *et al.*, 2016). A mangotoxin gene cluster was found in 50% of the isolates. Mangotoxin is known to inhibit ornithine acetyltransferase, an enzyme involved in arginine biosynthesis (Arrebola *et al.*, 2003).

A cluster of genes involved in phaseolotoxin production was found in two strains (C37, C39), which do not group with clade 3a. A NRPS ralsolamysin cluster was found for two strains (C64 and C65), cyanopeptin was found in C56 and cichopeptin, a recently discovered bioactive compound that has biosurfactant and antimicrobial properties (Huang *et al.*, 2015), was more widely distributed and found in most isolates of PG3a, except C46, C64, C65, NZ50, NZ55, NZ58 and *P. syringae* G33C.

Compared to PG3, the following NRPS were found for *Psa* NZ54: pseudopyronine AB, arylpolyene, mangotoxin, alginate and cichopeptin.

4.5.7 Copper resistance

An assay determining the minimal concentration of copper (MIC) which inhibits bacterial growth done by Colombi (2017) revealed that all NZ strains (NZ50, NZ51, NZ55-NZ58, *P. syringae* G33C) had a MIC superior to 0.8 mM $CuSO_4$, which is the threshold upon which they were considered resistant. Hence all kiwifruit *P. syringae* genomes were screened for the presence/absence of the *copABCD* and *cusABC* systems conferring copper resistance. The Chinese kiwifruit isolates lacked both copper resistance systems, whereas a screening of the NZ kiwifruit isolates revealed that six out of seven NZ isolates possess both the *copABCD* and *cusABC* system. In addition, the *copABCD* system was found in two kiwifruit isolates from 2010 and 2011 (ICMP 19498 and ICMP19499). None of the copper resistance clusters were associated with either ICE-elements or mobile elements. The organization of the two systems varies (Figure 4.11). Typically *copABCD* is arranged as a single operon, however this was only found to be the case for NZ55, NZ56 and ICMP19498. In all other isolates *copAB* and *copCD* were organized as two separate operons. The two-component regulatory system (*copRS/cusRS*) (Bondarczuk & Piotrowska-Seget, 2013) was present in all isolates carrying *cop* genes. In addition arsenic resistance genes were detected in six out of eight isolates.

The pairs *P. syringae* G33C/NZ58 and NZ50/NZ57 (except one SNP in *cusA*) show 100% identity at the amino acid level for *copABCD* and *cusABC*. Both isolate pairs were collected in the same orchards respectively. The isolate from Te Puke collected in 2010 (ICMP19498)

shared identical *copABCD* genes with NZ55, whereas ICMP19499 is highly similar to *P.s.* G33C/NZ58. A phylogenetic tree (Figure 4.10) based on the individual protein sequences of *copABCD* genes identified in *Psa* ICE's (Colombi *et al.*, 2017) and commensal strains revealed that protein sequences of *copABCD* of *Psa*NZ64_ICE are more similar (83%-98%) to the commensals NZ50, NZ57 than other *Psa* ICEs (58-72%).

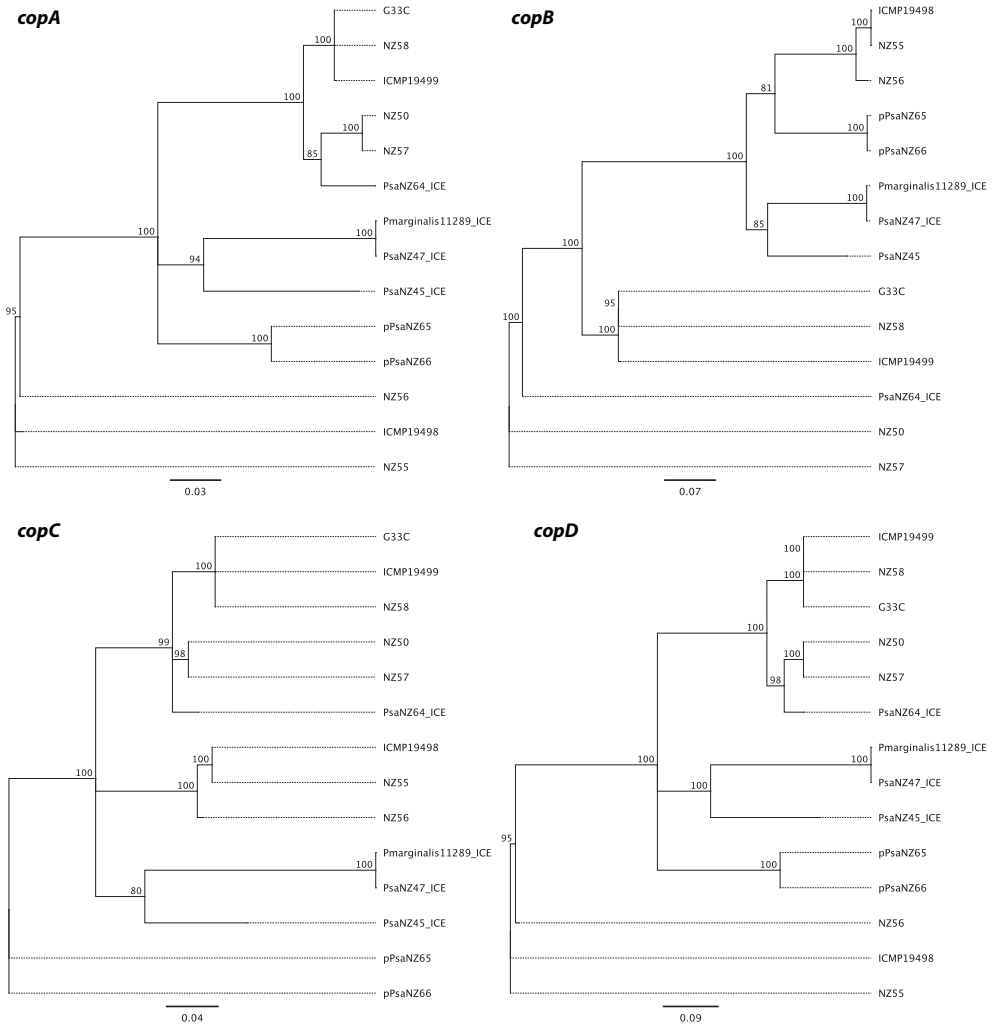

Figure 4.10: NJ tree for *copABCD* genes based on protein sequences. Includes the copper genes extracted from ICE elements (*Pmarginalis*11289_ICE, *Psa*NZ47_ICE, *Psa*NZ45_ICE, *Psa*NZ64_ICE) and plasmids (p*Psa*NZ65, p*Psa*NZ66) carrying copper resistance (Colombi *et al.*, 2017). Bootstrap values are indicated at the nodes.

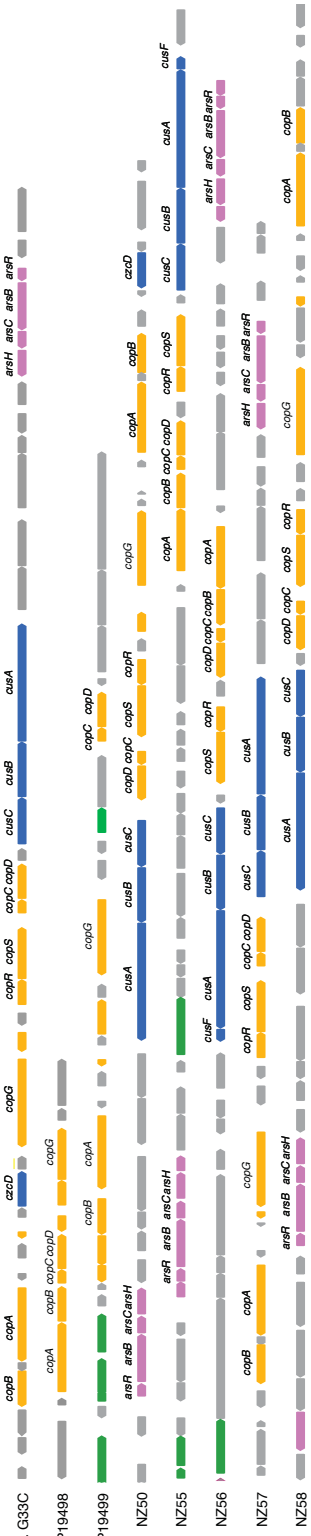

Figure 4.11: Genetic organisation of metal resistance loci in NZ isolates. Orange boxes depict copper resistance genes, green boxes refer to mobile genes (transposases or recombinases), blue boxes are the *czc*/*cus* system, pink boxes are arsenic resistance genes and grey boxes denote other genes.

4.5.8 Catechol operon

The catechol operon was discovered in a third of all isolates, evenly distributed throughout PG3, without any recognisable pattern. Only three strains which were isolated from kiwifruit (NZ50, NZ57 and C58) were shown to possess the phenol degradation pathway, two of which group in PG3a. The arrangement of genes shows the same makeup as the one described in detail in *P. syringae* pv. *aesculi* (Green *et al.*, 2010). Noticeably, most isolates from the kiwifruit-associated clade are not able to degrade phenol, including the strain *P. syringae* G33C, which was used for competition assays in Chapter 3.

Catechol operon is not required for endophytic growth in kiwifruit

The presence of the catechol operon, involved in the microbial degradation of phenolic compounds, was found to be essential for environmental strains, which were able to grow endophytically in kiwifruit (Bartoli *et al.*, 2015).

P. syringae G33C lacked the operon to degrade phenolic compounds, which is present in *Psa* NZ54. The previous *in planta* competition experiments (Chapter 3) have shown extensively that despite lack of the operon, *P. syringae* G33C grows successfully in the apoplast of *A. chinensis* var. *chinensis* Hort16A and G3 to a population size of 10^4-10^5 cfu/cm^2. To demonstrate that the operon is not a constraint for endophytic growth in *A. chinensis* var. *chinensis*, I tested the growth of *metA*$^+$ (*P.*

syringae NZ50) and *metA*$^{-}$ (*P. syringae* G33C) strains on cultivar Hort16A and found no significant differences in growth between those strains at any timepoint (P >0.3, paired t-tests) (Figure 4.12).

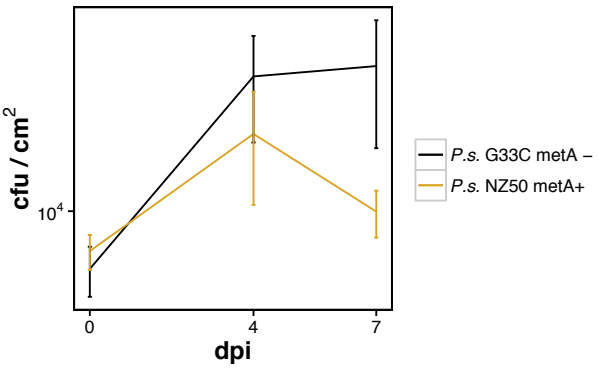

Figure 4.12: Comparison of endophytic growth in Hort16A of a strain possessing (*P. syringae* NZ50) and lacking (*P. syringae* G33C) the catechol operon The presented mean and standard error were calculated from the mean of 5 individual measurements.

4.6 Discussion

This study has unravelled the genetic diversity of PG3 with more detailed insight into the pathogenicity traits of newly described PG3a, a clade associated with kiwifruit isolates from NZ, Japan and China. The analysis was based on 7 NZ and 16 Chinese PG3 strains, as no whole genome sequences were available for any Japanese isolates. PG3 is a very diverse phylogroup with two well-supported clades based on a maximum likelihood phylogenetic analysis (Figure 4.4). Clade 3a contains solely *P. syringae* associated with kiwifruit, with the exception of C46, which was isolated from a tea orchard in Henan (China) where kiwifruit was grown previously. All other known PG3 strains form a cohesive subclade within PG3. Previously very few environmental isolates were found to belong to PG3 based on MLST (Berge *et al.*, 2014). There is a remarkable amount of diversity within PG3a and the paraphyly of NZ and Chinese kiwifruit isolates within PG3a suggests that these isolates have diversified on kiwifruit independent from each other.

Given the fact that isolates from NZ, China and Japan group together in clade 3a, it appears as if these commensal strains have been associated with kiwifruit for a long time. Kiwifruit is grown in a large number of countries worldwide, with China, Italy, New Zealand, Chile and Greece being the major producers. There is ongoing exchange of plant material (pollen, cuttings, fruit) between these countries, enabling the associated microbial kiwifruit community to be potentially translocated together.

Considering that since the outbreak of *Psa* in New Zealand in 2010 (Everett *et al.*, 2011), these commensal strains have been coexisting with *Psa*, there is the possibility of exchange of genetic material among members of the phyllosphere community. Acquisition of mobile elements, like genomic islands, via HGT can alter pathogenicity (Hacker *et al.*, 1997; Juhas *et al.*, 2009) and virulence mechanisms in the species complex *P. syringae* are dynamic across the different phylogroups (Baltrus *et al.*, 2011; Sarkar *et al.*, 2006).

4.6.1 Variable effector suite within PG3 *P. syringae*

There was a great amount of variation in the T3SE suite within PG3, ranging from an average of 20 effectors to a drastically reduced set of only six effectors for clade 3a. Extensive variations in the effector suites have long been recognised, even within a single pathovar, as was shown for example for *Psa* (Bartoli *et al.*, 2015; McCann *et al.*, 2013), where e.g. for 25 *Psa* strains, out of 51 discovered T3SEs only 17 effectors were shared among all isolates (McCann *et al.*, 2017). In previous studies, PG2 was shown to possess the smallest amount of effectors and a negative correlation between number of effectors and presence of additional toxin pathways (Baltrus *et al.*, 2011; Nowell *et al.*, 2016). The same trade-off was described for isolates from PG10, which possess only 5 effectors, but have simultaneously acquired additional phytotoxin genes (Hockett *et al.*, 2014). It is possible that due to the failure of effectors to elicit a defence response in kiwifruit and lack of a co-evolutionary arms race, commensal strains have

simply lost effectors. The few effectors that remain could relate to the minimum amount required to allow sufficient growth on kiwifruit. These effectors are however not enough to allow these isolates to reach pathogenic population sizes and/or elicit disease symptoms.

Only a few core effectors were shared among all kiwifruit *P. syringae* genomes: *avrE, hopI, hopAA* and *hopM*. Effectors *hopM1* and *avrE1* are functionally redundant and among the most widely distributed effectors (Badel *et al.*, 2006; Debroy *et al.*, 2004). The data suggest that the presence of additional effectors (*hopA1* and *hopAB3*) could compensate for possession of a truncated allele of *hopAA1*. In turn they lack also *hopZ3*, as it is located in the same genomic position as *hopA1*. It has been shown previously that T3SE can have similar functions and functional redundancies have been described (Kvitko *et al.*, 2009). Despite finding genes encoding for a small variety of T3SEs, their functionality remains to be elucidated. Point mutations, genomic rearrangements or indels can alter amino acid sequences and result in loss of function, changes in virulence, host range shifts or evasion of the plant defence system (Arnold & Jackson, 2011; Stevens *et al.*, 1998). An example for the diversification of effectors is the *hopZ* family in *P. syringae*. Different alleles of *hopZ* have arisen via mutational changes or via horizontal gene transfer from ecologically similar bacteria in response to the co-evolutionary arms race with the host (Ma *et al.*, 2006).

4.6.2 PG3a strains possess functional pathogenicity islands

The reduced effector repertoire of kiwifruit *P. syringae* led to the search for pathogenicity islands (atypical, tripartite, rhizobial) and the strains were found to harbour the tripartite pathogenicity island (Figure 4.7), which is the most common type of PAI associated with pathogenic *P. syringae* strains (Baltrus *et al.*, 2017). This was somewhat unexpected, given bacterial growth in the apoplast of a subset of these strains (*P. syringae* G33C and NZ50) was only around 10^4 cfu/cm^2 (Figure 4.12). This reduction in bacterial growth for *P. syringae* has been described for non-pathogenic isolates harbouring an atypical pathogenicity island (Clarke *et al.*, 2010; Mohr *et al.*, 2008) or for T3SS mutants [e.g. Δ*hrpJ* (Hirano *et al.*, 1999) or Δ*hrpL* (Ortiz-Martin *et al.*, 2010)]. Despite possession of gene clusters typical for a T-PAI, growth in kiwifruit is restricted. This suggests that other factors rather than the T3SS play a role in reaching pathogenic population sizes in kiwifruit. Despite the presence of the T-PAI, the type strain (*P. syringae* G33C) used for the competition experiments in the previous chapter did not elicit any symptoms of disease in different cultivars of kiwifruit, leaving the question of a suitable host range for these kiwifruit commensals to be determined. There remains the possibility that these commensals are purely associated with kiwifruit and are simply non-pathogenic strains in that environment.

4.6.3 Phytotoxins

Apart from the T3SS being a major contributor to pathogenicity in *P. syringae*, phytotoxins play a major role in terms of virulence (Bender *et al.*, 1999). The ability to produce toxins is highly variable among different pathovars and appears not to be conserved across phylogroups. There is even variability in phytotoxin production among closely related strains, with polymorphisms resulting in different products (Baltrus *et al.*, 2011; Hockett *et al.*, 2014; Hwang *et al.*, 2005). Out of the four primary toxins (coronatine, phaseolotoxin, syringomycin and tabtoxin), typically only one is produced by an individual strain (Hwang *et al.*, 2005). None of the genomes analysed in this study showed the presence of more than one of the non-ribosomal peptide toxin pathways (NRP) (Figure 4.9). The most interesting finding in the toxin screen was the presence of syringomycin genes in clade 3a, which harbour a reduced set of effectors. These results confirm the trade-off between amount of T3SE and NRP pathways which was described previously for *P. syringae* PG2, where the small number of effectors appears to be compensated by the ability to produce syringomycin (Baltrus *et al.*, 2011; Hockett *et al.*, 2014). This suggests in turn that strains with a very large set of effectors do not rely on the production of toxins to be able to colonize and infect their plant host, e.g. the isolates *Psa* NZ54 (34 effectors) or *P. syringae* pv. *daphniphylli* ICMP9757 (30 effectors) do not produce any of the primary toxins. The diverse range of T3SEs evolve under strong selection pressure by the host, as the pathogen

constantly interacts with the host's immune system to overcome resistance (arms-race) (McCann & Guttman, 2007).

Hypothetically, the absence of the host-pathogen co-evolutionary arms race and selection pressures resulted in certain *P. syringae* (some members of PG2, PG3a) isolates losing effector genes and showcasing a smaller set of effectors. In contrast, the production of toxins is less aimed at a particular host, but at targeting a variety of plant species (Bender *et al.*, 1999). Hence these strains could be regarded as generalists. Perhaps this trade-off in virulence mechanisms leads to increased epiphytic fitness, which gives them an advantage resulting in long-term resident strains. However, future work should aim at the discovery and identification of any novel T3SE families which may be present in these commensal *P. syringae* isolates.

4.6.4 Secondary metabolite production

Genome mining for secondary metabolite biosynthesis clusters is a common approach to detect novel gene clusters, feasible because of a high protein sequence similarity for core genes of biosynthetic enzymes. This analysis was done in addition to the screen of well known phytotoxins produced by *P. syringae*. Apart from confirming the presence of syringomycin and syringolin already found using the BLAST approach in some of the isolates, other expected biosynthesis clusters for production of compounds beneficial to the epiphytic phase of *P. syringae* life cycle were found. Additionally, gene clusters encoding for recently discovered NRPS like cichopeptin, ralsolamycin and cyanopeptin were identified. Some of these NRPS have

been only very recently described, like cichopeptin discovered in *Pseudomonas cichorii* (Huang *et al.*, 2015).

4.6.5 Copper resistance

The use of copper compounds to control plant diseases is common practice in agriculture given its low price and high efficiency. However, the effectiveness of copper applications is reduced by the appearance of bacterial resistance mechanisms. Copper sprays have been used extensively in kiwifruit orchards in New Zealand to protect the vines and in 2014 the first copper sulphate-resistant *Psa* strains were detected. Copper resistance mechanisms had evolved via the acquisition of plasmids and integrative conjugative elements (Colombi *et al.*, 2017). Copper-resistant *Psa* has not been found in other countries, which has led to the hypothesis that environmental copper levels have led to selective evolution of copper resistance (Colombi *et al.*, 2017). A screen of the NZ commensal kiwifruit *P. syringae* showed that 8 out of 9 strains possess copper resistance systems (Figure 4.11): the *copABCD* and/or *cusABC* system with the functional regulator system *cop/cusRS*. The two strains isolated in 2010 and 2011 (ICMP19498, ICMP19499) at the same time as the incursion event of *Psa*, only possess the *copABCD* system, whereas all others harboured both. Nonetheless, the presence of copper resistance mechanisms in commensal *P. syringae* (2010-2013) prior to the discovery of copper resistant *Psa* in 2014 supports the hypothesis that copper resistance has been present before the heavier use of copper compounds in a bid to protect vines from *Psa*

(Colombi *et al.*, 2017). The copper systems showed high similarity at the protein level, suggesting that copper resistance has evolved in these isolates some time ago. It is highly likely that copper resistance has indeed spread from commensal isolates to *Psa*. The high similarity of copper genes in *Psa*NZ64_ICE suggests that indeed the copper resistance genes might have been acquired from these commensal isolates (Figure 4.10).

4.6.6 Catechol operon

The catechol operon, a set of 10 genes involved in the β-ketoadipate pathway for degradation of plant-derived aromatic compounds, appears to be well conserved in *P. syringae* PG1 and 3, whereas strains falling into the ubiquitous PG2 seem to be lacking the operon (Bartoli *et al.*, 2015; Nowell *et al.*, 2016). The catechol operon has been suspected of conferring advantage to *P. syringae* in their ability to infect woody hosts (Green *et al.*, 2010; Marcelletti & Scortichini, 2011; Rodríguez-Palenzuela *et al.*, 2010) and it was argued that it contributed to a host range extension across PG3 (Nowell *et al.*, 2016). My results suggest that the catechol operon is in fact not conserved across *P. syringae* PG3 kiwifruit strains, with only one third of strains showing the presence of genes involved in the degradation of aromatic rings. The epiphyte *P. syringae* G33C (PG3a) is also lacking this particular conserved region, but is nonetheless able to grow endophytically. Japanese kiwifruit isolates, which are also associated with PG3a also tested negative for *metA* (Tomihama *et al.*, 2016), a protein involved in phenol degradation and part of the catechol operon, which could suggest that these

strains are not fully adapted to the woody hosts they reside on. Comparing the growth of two strains from PG3a, *P. syringae* G33C (*metA*$^{-}$) to *P. syringae* R14C (*metA*$^{+}$), no significant differences in endophytic growth were discovered. This suggests the possession of the operon alone does not contribute to increased growth within the host. It remains to be elucidated whether the possession of the operon might be advantageous to a phyllosphere resident strain when grown in competition with *Psa*, which possesses the operon most likely for pathogenesis on kiwifruit, and whether the operon actually contributes to virulence.

Taken together, these results provide insights into the diversity and genomic characteristics like pathogenicity, host range and resistance mechanisms of commensal *P. syringae* strains. Although copper resistance has only evolved in NZ strains, virulence factors like the presence/absence of effectors and toxin pathways are well conserved for NZ and Chinese strains of PG3a, suggesting that indeed these isolates have been associated and co-evolved on kiwifruit for quite some time.

Interactions within the bacterial community play an important role in the evolution of pathogenic isolates by acquisition of DNA sequences via Horizontal Gene Transfer (HGT) (Bartoli *et al.*, 2016; Monteil *et al.*, 2013). HGT can result in the transfer of plasmids, Integrative Conjugative Elements (ICE), pathogenicity islands or bacteriophages (Juhas *et al.*, 2009). Bacteria colonizing the same niche, e.g. the leaf, are more prone to engage in HGT (Smillie *et al.*, 2011). This was shown, for example, for *Pectobacterium atrosecpticum*, *Salmonella enterica* and *Pseudomonas*

syringae, where transfer of ICEs was favoured in the natural environment compared to *in vitro* studies (Lovell *et al.*, 2010; Pitman *et al.*, 2005; Quiroz *et al.*, 2011; Vanga *et al.*, 2015).

The *P. syringae* commensals on kiwifruit carry a very limited amount of effectors, which are most likely required to manipulate the host defence system in order to successfully colonise the host. The acquisition of further T3SEs via HGT could have a direct effect on pathogenicity, but could also simply contribute to increased epi- or endophytic fitness. Ideally, the functionality of the T3SEs and pathogenicity islands already present in commensal kiwifruit *P. syringae* should be verified using bioinformatics and wet-lab techniques. In addition construction of mutants would be an option to see whether loss/acquisition of T3SE results in changes in growth, host range or pathogenicity. In addition phylogenetic reconstruction of T3SE families would help identify whether these effectors have been acquired recently or have contributed to their lifestyle in the past.

Chapter 5

Concluding discussion

5.1 Summary of findings

In the past, studies of plant pathology have focused mainly on pathogenic bacteria isolated from respective hosts. Only recently the focus of interest has shifted towards nonpathogenic members and the role they might play in the evolution and emergence of pathogens (Mohr *et al.*, 2008; Monteil *et al.*, 2013; Morris *et al.*, 2010).

In my doctoral research I took yet another approach by focusing on the diversity of *P. syringae* in an agricultural setting, but particularly using population-based sampling in non-infected and *Psa*-infected kiwifruit orchards. I have provided insights into a diverse population of *P. syringae* inhabiting the phyllosphere of kiwifruit, combining different fields by inferring micro-evolutionary parameters using population genetics,

investigating the ecological interactions of commensals with a pathogenic isolate using *in planta* and *in vitro* assays and gauging their pathogenicity potential using a comparative genomics approach.

5.1.1 Diversity of *P. syringae* in the phyllosphere

I performed a detailed study of the diversity and population genetics of *P. syringae* in *Psa* infected and non-infected kiwifruit orchards. Commensal *P. syringae* appeared to have a mostly clonal population structure, with little evidence of recombination within phylogroups, but not between phylogroups. This supports the results of previous studies done using mostly pathogenic isolates (Sarkar & Guttman, 2004). Clonal population structures have been described for other bacterial pathogens such as *E. coli* (Ochman & Selander, 1984; Selander & Levin, 1980; Tenaillon *et al.*, 2010), and *Salmonella enterica* (Nelson & Selander, 1994; Selander *et al.*, 1990). However, this does not mean that recombination does not occur, rather that the level of recombination occurring is most likely not frequent enough to disrupt clonal patterns (Tibayrenc & Ayala, 2012), or that the level of resolution was not sufficient, as partial sequences of housekeeping genes were used. Using whole genome sequencing would provide a greater level of depth. Genetic variation is introduced by mutation, recombination and migration. The fate of these variations is determined by genetic drift and selection and contributes to the evolutionary history of an organism. These population genetics measures are of particular interest for disease epidemiologists in terms of unravelling the source of pathogenic isolates,

how bacterial pathogens emerge or reemerge from environmental reservoirs and routes of dissemination (Vinatzer *et al.*, 2014).

Based on MLST analysis, a new clade (3a) within *P. syringae* Phylogroup 3 was identified, which was solely associated with *P. syringae* isolated from kiwifruit. This clade was briefly mentioned for *P. syringae* isolates from kiwifruit in Japan (Tomihama *et al.*, 2016) and I was able to show that *P. syringae* isolated from kiwifruit in New Zealand, Japan and China form a monophyletic clade. These findings strongly support the hypothesis of a remarkably variable lifestyle of *P. syringae* with the presence of a diverse population of *P. syringae* found in the environment (Monteil *et al.*, 2013, 2014; Morris *et al.*, 2010, 2007), and on plants, in particular in an agricultural setting. It would be interesting to perform sampling with a focus on a potential temporal effect on diversity and whether these seemingly resident strains are consistent epi- and endophytes on kiwifruit over consecutive years. Most likely the diversity found in this study still represents only a fraction of the existent population in the wild, due to sampling constraints. This, taken together with the diverse number of strains isolated from various environmental locations (leaf litter, soil, irrigation water, snow, ice), suggests that the natural diversity and ecology of non-pathogenic *P. syringae* has been neglected in the past and further research should address this lack of knowledge. In order to increase our understanding of natural populations and their involvement in the evolution of pathogens, the sampling strategies need to be more population-based.

5.1.2 Ecological interactions of commensal *P. syringae* and the pathogen *Psa*

The bacterial population residing on a certain plant host in most cases has either no effect (commensals), or can even be beneficial to the host (mutualists), e.g. phyllosphere microbiota can be involved in plant growth and development or provide disease protection (e.g. Berg, 2009; Berlec, 2012; Lindow & Leveau, 2002). The importance of commensals and their role in pathogen evolution and the origin and spread of bacterial diseases has received attention for human pathogens like *E. coli* (Tenaillon *et al.*, 2010). Most research is based on identifying recombinant regions in genomes and virulence clusters that have been lost/gained according to the different lifestyles e.g. for *E. coli* (Didelot *et al.*, 2012; Rasko *et al.*, 2008) or the highly recombinant human *Neisseria* sp. (Maynard Smith *et al.*, 2000). Studies investigating ecological interactions among commensals and pathogens are, however, rare. This chapter provided insight into the ecological interactions of commensal *P. syringae* with a pathogenic *P. syringae* isolate by performing competition experiments using members of a population in a cultivated environment. Assuming closely related bacteria form an ecologically and genetically cohesive unit occupying the same niche, genetic exchange and competition for nutrients and space could be occurring among those strains (Shapiro & Polz, 2014).

The commensal *P. syringae* G33C showed epi- and endophytic growth similar to other non-pathogenic *P. syringae* and pathogenic isolates in non-susceptible hosts (e.g. Clarke *et al.*, 2010; Kniskern *et al.*, 2011; Mohr

et al., 2008), without causing disease symptoms in kiwifruit. The competition experiments *in planta* revealed complex interactions between *Psa* NZ54 and *P. syringae* G33C. Both isolates were able to establish a population when co-inoculated in equal ratios. The failure of *P. syringae* G33C to invade from rare suggests, however, that the polymorphism between the two strains was not stable. Density-dependent effects on the outcome of these ecological interactions remain to be elucidated. The initial inoculum dose could have an effect on the outcome of the interactions, as was shown for growth interference occurring in mixed infections of *P. syringae* at higher inoculation densities (10^8), but not at 10^3-10^4 (Macho *et al.*, 2007).

The commensal isolate *P. syringae* G33C showed antagonistic behavior towards the kiwifruit canker pathogen *Psa* in two different gold kiwifruit cultivars (Hort16A and G3). An interesting aspect would be to test mixtures of commensal *P. syringae* and see whether the diminishing effect on growth of *Psa* NZ54 could be amplified. The external physical factors (humidity, light, leaf surface properties) influencing the complex interactions between pathogens and closely related commensal bacteria remain to be further elucidated, but one could explore the potential of these ecologically similar antagonistic strains as a use of biocontrol agents to curtail *Psa* infection. More broadly, this work highlights the importance of understanding the genotypic diversity and bringing the ecological interactions among pathogens and non-pathogens in a natural population into context.

5.1.3 The pathogenicity potential of commensal *P. syringae*

The third chapter investigated pathogenicity attributes of commensal *P. syringae* using whole genome sequencing. The T3SS is the most important and most thoroughly investigated virulence machinery of *P. syringae* (Collmer *et al.*, 2000; Lindeberg *et al.*, 2009), however relating host range to the T3SE repertoire remains difficult given the redundant functions of many effectors and large variations in presence/absence of effectors observed even within single pathovars. Not many environmental isolates have been subjected to a thorough search of pathogenicity attributes; hence the screen of kiwifruit isolates revealed an interesting insight into potential virulence genes carried by commensals. *P. syringae* of clade 3a, which was solely associated with kiwifruit possess an extremely reduced set of effectors, but in return these isolates all appeared to carry genes required for the production of syringomycin/syringopeptin. This suggested that there has been a trade-off, in which acquiring pathways for phytotoxin production has compensated the loss of effectors. This could point to these strains being generalists, as phytotoxins are aimed at a broad range of targets, compared to T3SE, which are specifically for interaction with a particular plant host. These results confirmed what has been described before for Phylogroup 2c (Baltrus *et al.*, 2011; Hockett *et al.*, 2014), with these strains supposedly excelling in epiphytic fitness compared to other related *P. syringae* (Feil *et al.*, 2005). This suggests that this trade-off could be a common approach for *P. syringae* in response to a less

specified/focused lifestyle, having shown this occurred in two different phylogroups.

Interestingly, the growth of commensal strains tested in kiwifruit showed a 10,000x reduced population size compared to the pathogenic *Psa* isolated. This was despite the commensal strains possessing a typical tripartite pathogenicity island with the presence of all relevant genes. Although the functionality of the T-PAI remains to be confirmed by predicting HrpL-dependant regulation and comparative expression analysis, the question arises as to what is the minimal number of effectors and their role required to reach pathogenic population sizes.

Considering the global distribution of kiwifruit-associated *P. syringae*, it stands to reason that these isolates have been co-existing on kiwifruit for quite some time. Given the occupancy of a particular ecological niche (woody host), presumably adaptations to this particular environment have taken place. The catechol operon supposedly evolved as an adaptation in response to the colonization of woody hosts (Bartoli *et al.*, 2015; Nowell *et al.*, 2016) and was so far identified in various representatives of *P. syringae* PG1 and PG3. Kiwifruit *P. syringae* showed no conclusive possession of the catechol operon, with only 33% of isolates carrying relevant genes. A growth assay comparing commensal kiwifruit *P. syringae* with/without the operon, suggested the presence of the operon was not a prerequisite for colonization of the apoplastic space in kiwifruit, as there were no differences in endophytic growth. It is possible that the degradation of aromatic compounds is beneficial in other stages of the life

cycle, e.g. the epiphytic phase or to ensure long-term stability of bacterial populations.

Taken together, the results of my doctoral research contribute to our understanding of the natural diversity of a plant pathogen with a very comprehensive life style and potential interactions occurring in the natural environment. Insights into the population genetics were given for a representative sample of a natural pathogen population in an agricultural context, which has rarely been done before. The competition experiments revealed the complex ecological interactions occurring between a commensal and a pathogenic isolate, dependent on various environmental factors. Combined with comparative genomics to detect pathogenic attributes present in commensals, my work revealed the potential of these isolates in the use of disease control.

5.2 Future directions

In terms of plant pathogens, the ultimate goal is the prevention of further outbreaks and control of current epidemics. The use of naturally occurring plant-associated microorganisms for biologically controlling plant diseases is of great interest to modern agriculture (Berg, 2009). The discovery of antagonistic behavior of a commensal *P. syringae* towards the kiwifruit canker pathogen *Psa* in two different cultivars of kiwifruit provokes interest in determining the full potential of these kiwifruit resident strains in terms of potential disease control or even prevention. The efficacy of a bacterial

strain as a biocontrol agent is related to its ecological similarity with the target pathogen (Völksch & May, 2001), making the collection of *P. syringae* the repository of perfect candidates. Large-scale plant pathology assays would allow elimination of any strains with pathogenic potential and to determine whether *Psa* infection is positively affected by interaction with other commensal isolates. This is important as multispecies synergistic interactions have been described to increase disease severity for a variety of microbial diseases in humans and animals, but such microbial interactions have remained neglected in the study of plant diseases (Lamichhane & Venturi, 2015).

Using closely related strains to control disease opens up the question of their involvement in the evolution of pathogenicity. Acquisition of genomic content via Horizontal Gene Transfer (HGT) plays a major role in the evolution of pathogenicity and can have a profound impact on the phenotype, including the ability to colonize and infect new hosts. The observed variations in the T3SE repertoire of *P. syringae* among different pathovars, but also even within a single pathovar, have been attributed to relatively frequent HGT events, especially because they are often found on mobile genetic elements (Fujikawa & Sawada, 2016; Lindeberg *et al.*, 2012; O'Brien *et al.*, 2012).

An interesting way of testing how likely HGT processes occur between co-residing isolates would be to use an experimental evolution approach by co-inoculating commensal isolates with the pathogen and serially passaging through kiwifruit for several generations. Ideally also implementing varying

ratios of inoculation to be able to make conclusions about density dependent effects. Whole-genome sequencing would reveal the amount and location of HGT events that occurred.

References

Abelleira, A., López, M. M., Peñalver, J., Aguín, O., Mansilla, J. P., Picoaga, A., & García, M. J. (2011). First report of bacterial canker of kiwifruit Caused by *Pseudomonas syringae* pv. *actinidiae* in Spain. *Plant Disease*, 95(12), 1583.

Adams, P. D. & Kloepper, J. W. (2002). Effect of host genotype on indigenous bacterial endophytes of cotton (*Gossypium hirsutum* L.). *Plant and Soil*, 240(1), 181–189.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), 716–723.

Alarcón-Chaidez, F. J., Peñaloza-Vázquez, a., Ullrich, M., & Bender, C. L. (1999). Characterization of plasmids encoding the phytotoxin coronatine in *Pseudomonas syringae*. *Plasmid*, 42(3), 210–220.

Alfano, J. R., Charkowski, A. O., Deng, W. L., Badel, J. L., Petnicki-Ocwieja, T., van Dijk, K., & Collmer, A. (2000). The *Pseudomonas syringae Hrp* pathogenicity island has a tripartite mosaic structure composed of a cluster of type III secretion genes bounded by exchangeable effector and conserved effector loci that contribute to parasitic fitness and pathogenicity in pl. *Proceedings of the National Academy of Sciences*, 97(9), 4856–4861.

Alfano, J. R. & Collmer, A. (2004). Type III secretion system effector proteins: double agents in bacterial disease and plant defense. *Annual Review of Phytopathology*, 42, 385–414.

Almeida, N. F., Yan, S., Cai, R., Clarke, C. R., Morris, C. E., Schaad, N. W., Schuenzel, E. L., Lacy, G. H., Sun, X., Jones, J. B., Castillo, J. A., Bull, C. T., Leman, S., Guttman, D. S., Setubal, J. C., & Vinatzer, B. A. (2010). PAMDB, a multilocus sequence typing and analysis database and website for plant-associated microbes. *Phytopathology*, 100(3), 208–215.

Almeida, N. F., Yan, S., Lindeberg, M., Studholme, D. J., Schneider, D. J., Condon, B., Liu, H., Viana, C. J., Warren, A., Evans, C., Kemen, E., Maclean, D., Angot, A., Martin, G. B., Jones, J. D., Collmer, A., Setubal, J. C., & Vinatzer, B. A. (2009). A draft genome sequence of *Pseudomonas syringae* pv. *tomato* T1 reveals a type III effector repertoire significantly divergent from that of *Pseudomonas syringae* pv. *tomato* DC3000. *Molecular Plant-Microbe Interactions*, 22(1), 52–62.

Anderson, M. J. (2001). A new method for non parametric multivariate analysis of variance. *Austral Ecology*, 26(2001), 32–46.

Anderson, P. K., Cunningham, A. A., Patel, N. G., Morales, F. J., Epstein, P. R., & Daszak, P. (2004). Emerging infectious diseases of plants: Pathogen pollution, climate change and agrotechnology drivers. *Trends in Ecology and Evolution*, 19(10), 535–544.

Aoki, S. K., Pamma, R., Hernday, A. D., Bickham, J. E., Braaten, B. a., & Low, D. a. (2005). Contact-dependent inhibition of growth in *Escherichia coli*. *Science*, 309(5738), 1245–1248.

Aranda, E. (2016). Promising approaches towards biotransformation of polycyclic aromatic hydrocarbons with Ascomycota fungi. *Current Opinion in Biotechnology*, 38, 1–8.

Arndt, D., Grant, J. R., Marcu, A., Sajed, T., Pon, A., Liang, Y., & Wishart, D. S. (2016). PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Research*, 44(W1), W16–W21.

Arnold, D. L. & Jackson, R. W. (2011). Bacterial genomes: evolution of pathogenicity. *Current Opinion in Plant Biology*, 14(4), 385–391.

Arrebola, E., Cazorla, F. M., Durán, V. E., Rivera, E., Olea, F., Codina, J. C., Pérez-Garcıa, A., & de Vicente, A. (2003). Mangotoxin: a novel antimetabolite toxin produced by *Pseudomonas syringae* inhibiting ornithine/arginine biosynthesis. *Physiological and Molecular Plant Pathology*, 63(3), 117–127.

Auton, A. & McVean, G. (2007). Recombination rate estimation in the presence of hotspots. *Genome Research*, 17(8), 1219–1227.

Azarian, T., Ali, A., Johnson, J. A., Jubair, M., Cella, E., Ciccozzi, M., Nolan, D. J., Farmerie, W., Rashid, M. H., Sinha-Ray, S., Alam, M. T., Morris, J. G., & Salemi, M. (2016). Non-toxigenic environmental *Vibrio cholerae* O1 strain from Haiti provides evidence of pre-pandemic cholera in Hispaniola. *Scientific Reports*, 6, 36115.

Badel, J. L., Shimizu, R., Oh, H.-S., & Collmer, A. (2006). A *Pseudomonas syringae* pv. *tomato avrE1/hopM1* mutant is severely reduced in growth and lesion formation in tomato. *Molecular Plant-Microbe Interactions*, 19(2), 99–111.

Balestra, G. M., Mazzaglia, A., Spinelli, R., Graziani, S., Quattrucci, A., & Rossetti, A. (2008). Cancro batterico su *Actinidia chinensis*. *L'informatore Agrario*, 38, 75–76.

Balestra, G. M., Renzi, M., & Mazzaglia, A. (2010). First report of bacterial canker of *Actinidia deliciosa* caused by *Pseudomonas syringae* pv. *actinidiae* in Portugal. *New Disease Reports*, 22, 10.

Balint-Kurti, P., Pridgen, P., & Stapleton, A. E. (2011). Application of an antibiotic resets the maize leaf phyllosphere community and increases resistance to southern leaf blight. In *Acta Horticulturae*, volume 905 (pp. 57–62).

Baltrus, D. A., Dougherty, K., Beckstrom-Sternberg, S. M., Beckstrom-Sternberg, J. S., & Foster, J. T. (2014). Incongruence between multi-locus sequence analysis (MLSA) and whole-genome-based phylogenies: *Pseudomonas syringae* pathovar *pisi* as a cautionary tale. *Molecular Plant Pathology*, 15(5), 461–465.

Baltrus, D. A., McCann, H. C., & Guttman, D. S. (2017). Evolution, genomics and epidemiology of *Pseudomonas syringae*. *Molecular Plant Pathology*, 18(1), 152–168.

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(7), e1002132.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., Lesin, V. M., Nikolenko, S. I., Pham, S., Prjibelski, A. D., Pyshkin, A. V., Sirotkin, A. V., Vyahhi, N., Tesler, G., Alekseyev, M. a., & Pevzner, P. a. (2012). SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology*, 19(5), 455–477.

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D. A., & Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. *Science*, 315(5819), 1709–1712.

Barrick, J. E. & Lenski, R. E. (2013). Genome dynamics during experimental evolution. *Nature Reviews Genetics*, 14(12), 827–839.

Bartoli, C., Berge, O., Monteil, C. L., Guilbaud, C., Balestra, G. M., Varvaro, L., Jones, C., Dangl, J. L., Baltrus, D. A., Sands, D. C., & Morris, C. E. (2014). The *Pseudomonas viridiflava* phylogroups in the *P. syringae* species complex are characterized by genetic variability and phenotypic plasticity of pathogenicity-related traits. *Environmental Microbiology*, 16(7), 2301–2315.

Bartoli, C., Lamichhane, J. R., Berge, O., Guilbaud, C., Varvaro, L., Balestra, G. M., Vinatzer, B. A., & Morris, C. E. (2015). A framework to gauge the epidemic potential of plant pathogens in environmental reservoirs: the example of kiwifruit canker. *Molecular Plant Pathology*, 16(2), 137–49.

Bartoli, C., Roux, F., & Lamichhane, J. R. (2016). Molecular mechanisms underlying the emergence of bacterial pathogens: An ecological perspective. *Molecular Plant Pathology*, 17(2), 303–310.

Bastas, K. K. & Karakaya, A. (2011). First report of bacterial canker of kiwifruit caused by *Pseudomonas syringae* pv. *actinidiae* in Turkey. *Plant Disease*, 96(3), 452.

Beattie, G. a. & Lindow, S. E. (1999). Bacterial colonization of leaves: a spectrum of strategies. *Phytopathology*, 89(5), 353–359.

Beiki, F., Busquets, A., Gomila, M., Rahimian, H., Lalucat, J., & García-Valdés, E. (2016). New *Pseudomonas* spp. are pathogenic to citrus. *PLoS ONE*, 11(2), e0148796.

Bender, C. L., Alarcon-Chaidez, F., Gross, D. C., Alarcón-Chaidez, F., & Gross, D. C. (1999). *Pseudomonas syringae* phytotoxins: Mode of action, regulation, and biosynthesis by peptide and polyketide synthetases. *Microbiology and Molecular Biology Reviews*, 63(2), 266–292.

Berg, G. (2009). Plant-microbe interactions promoting plant growth and health: perspectives for controlled use of microorganisms in agriculture. *Applied Microbiology and Biotechnology*, 84(1), 11–18.

Berge, O., Monteil, C. L., Bartoli, C., Chandeysson, C., Guilbaud, C., Sands, D. C., & Morris, C. E. (2014). A user's guide to a data base of the diversity of *Pseudomonas syringae* and its application to classifying strains in this phylogenetic complex. *PLoS ONE*, 9(9), e105547.

Berlec, A. (2012). Novel techniques and findings in the study of plant microbiota: Search for plant probiotics. *Plant Science*, 193-194, 96–102.

Bhattacharjee, S., Halane, M. K., Kim, S. H., & Gassmann, W. (2011). Pathogen effectors target

Arabidopsis EDS1 and alter its interactions with immune regulators. *Science*, 334(6061), 1405–1408.

Birch, L. C. (1957). The meanings of competition. *The American Naturalist*, 91(856), 5–18.

Biruma, M., Pillay, M., Tripathi, L., Blomme, G., Abele, S., Mwangi, M., Bandyopadhyay, R., Muchunguzi, P., Kassim, S., Nyine, M., Turyagyenda, L., & Eden-Green, S. (2007). Banana *Xanthomonas* wilt: a review of the disease, management strategies and future research directions. *African Journal of Biotechnology*, 6(8), 953–962.

Blanchard, A. E., Celik, V., & Lu, T. (2014). Extinction, coexistence, and localized patterns of a bacterial population with contact-dependent inhibition. *BMC Systems Biology*, 8(23).

Block, A. & Alfano, J. R. (2011). Plant targets for *Pseudomonas syringae* type III effectors: Virulence targets or guarded decoys? *Current Opinion in Microbiology*, 14(1), 39–46.

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), 2114–2120.

Boller, T. & Felix, G. (2009). A renaissance of eicitors: perception of microbe-associated molecular patterns and danger signals by pattern-recognition receptors. *Annual Review of Plant Biology*, 60, 379–406.

Bondarczuk, K. & Piotrowska-Seget, Z. (2013). Molecular basis of active copper resistance mechanisms in Gram-negative bacteria. *Cell Biology and Toxicology*, 29(6), 397–405.

Bradbury, J. F. (1986). *Guide to Plant Pathogenic Bacteria*. CAB International, Farnham Royal, Great Britain.

Brown, A. H., Feldman, M. W., & Nevo, E. (1980). Multilocus structure of natural populations of *Hordeum spontaneum*. *Genetics*, 96(2), 523–536.

Buell, C. R., Joardar, V., Lindeberg, M., Selengut, J., Paulsen, I. T., Gwinn, M. L., Dodson, R. J., Deboy, R. T., Durkin, A. S., Kolonay, J. F., Madupu, R., Daugherty, S., Brinkac, L., Beanan, M. J., Haft, D. H., Nelson, W. C., Davidsen, T., Zafar, N., Zhou, L., Liu, J., Yuan, Q., Khouri, H., Fedorova, N., Tran, B., Russell, D., Berry, K., Utterback, T., Van Aken, S. E., Feldblyum, T. V., D'Ascenzo, M., Deng, W. L., Ramos, A. R., Alfano, J. R., Cartinhour, S., Chatterjee, A. K., Delaney, T. P., Lazarowitz, S. G., Martin, G. B., Schneider, D. J., Tang, X., Bender, C. L., White, O., Fraser, C. M., & Collmer, A. (2003). The complete genome sequence of the Arabidopsis and tomato pathogen *Pseudomonas syringae* pv. *tomato* DC3000. *Proceedings of the National Academy of Sciences*, 100(18), 10181–10186.

Bulgarelli, D., Schlaeppi, K., Spaepen, S., van Themaat, E. V. L., & Schulze-Lefert, P. (2013). Structure and functions of the bacterial microbiota of plants. *Annual Review of Plant Biology*, 64(1), 807–838.

Bull, C. T., Clarke, C. R., Cai, R., Vinatzer, B. A., Jardini, T. M., & Koike, S. T. (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(7), 847–858.

Bull, C. T., De Boer, S. H., Denny, T. P., Firrao, G., Fischer-Le Saux, M., Saddler, G. S., Scortichini, M., Stead, D. E., & Takikawa, Y. (2010). Comprehensive list of names of plant pathogenic bacteria, 1980-2007. *Journal of Plant Pathology*, 92(3), 551–592.

Buonaurio, R., Moretti, C., da Silva, D. P., Cortese, C., Ramos, C., & Venturi, V. (2015). The olive knot disease as a model to study the role of interspecies bacterial communities in plant disease. *Frontiers in Plant Science*, 6(June), 1–12.

Burch, A. Y., Zeisler, V., Yokota, K., Schreiber, L., & Lindow, S. E. (2014). The hygroscopic biosurfactant syringafactin produced by *Pseudomonas syringae* enhances fitness on leaf surfaces during fluctuating humidity. *Environmental Microbiology*, 16(7), 2086–2098.

Burdon, J. J., Thrall, P. H., & Ericson, L. (2013). Genes, communities & invasive species: understanding the ecological and evolutionary dynamics of host–pathogen interactions. *Current Opinion in Plant Biology*, 16, 1–6.

Burmølle, M., Webb, J. S., Rao, D., Hansen, L. H., Sørensen, S. J., & Kjelleberg, S. (2006). Enhanced biofilm formation and increased resistance to antimicrobial agents and bacterial invasion are caused by synergistic interactions in multispecies biofilms. *Applied and Environmental Microbiology*, 72(6), 3916–3923.

Burr, T., Matteson, M., Smith, C., Corral-Garcia, M., & Huang, T.-C. (1996). Effectiveness of bacteria and yeasts from apple orchards as biological control agents of apple scab. *Biological Control*, 6(2), 151–157.

Byrne, J., Dianese, A., Ji, P., Campbell, H., Cuppels, D., Louws, F., Miller, S., Jones, J., & Wilson, M. (2005). Biological control of bacterial spot of tomato under field conditions at several locations in North America. *Biological Control*, 32(3), 408–418.

Cai, R., Lewis, J., Yan, S., Liu, H., Clarke, C. R., Campanile, F., Almeida, N. F., Studholme, D. J., Lindeberg, M., Schneider, D., Zaccardelli, M., Setubal, J. C., Morales-Lizcano, N. P., Bernal, A., Coaker, G., Baker, C., Bender, C. L., Leman, S., & Vinatzer, B. A. (2011). The plant pathogen *Pseudomonas syringae* pv. *tomato* is genetically monomorphic and under strong selection to evade tomato immunity. *PLoS Pathogens*, 7(8), e1002130.

Calvo, P., Nelson, L., & Kloepper, J. W. (2014). Agricultural uses of plant biostimulants. *Plant and Soil*, 383, 3–41.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009). BLAST plus: architecture and applications. *BMC Bioinformatics*, 10(421), 421.

Cameron, A. & Sarojini, V. (2014). *Pseudomonas syringae* pv. *actinidiae*: chemical control, resistance mechanisms and possible alternatives. *Plant Pathology*, 63(1).

Charkowski, A. O., Alfano, J. R., Preston, G., Yuan, J., He, S. Y., & Collmer, A. (1998). The *Pseudomonas syringae* pv. *tomato* HrpW protein has domains similar to harpins and pectate lyases and can elicit the plant hypersensitive response and bind to pectate. *Journal of Bacteriology*, 180(19), 5211–5217.

Chisholm, S. T., Coaker, G., Day, B., & Staskawicz, B. J. (2006). Host-microbe interactions: Shaping the evolution of the plant immune response. *Cell*, 124(4), 803–814.

Cimermancic, P., Medema, M. H., Claesen, J., Kurita, K., Wieland Brown, L. C., Mavrommatis, K., Pati, A., Godfrey, P. A., Koehrsen, M., Clardy, J., Birren, B. W., Takano, E., Sali, A., Linington, R. G., & Fischbach, M. A. (2014). Insights into secondary metabolism from a global analysis of prokaryotic

biosynthetic gene clusters. *Cell*, 158(2), 412–421.

Clarke, C. R., Cai, R., Studholme, D. J., Guttman, D. S., & 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. *Molecular Plant-Microbe Interactions*, 23(2), 198–210.

Cochet, N. & Widehem, P. (2000). Ice crystallization by *Pseudomonas syringae*. *Applied Microbiology and Biotechnology*, 54(2), 153–161.

Collier, S. M. & Moffett, P. (2009). NB-LRRs work a "bait and switch" on pathogens. *Trends in Plant Science*, 14(10), 521–529.

Collmer, A., Badel, J. L., Charkowski, A. O., Deng, W. L., Fouts, D. E., Ramos, A. R., Rehm, A. H., Anderson, D. M., Schneewind, O., van Dijk, K., & Alfano, J. R. (2000). *Pseudomonas syringae* Hrp type III secretion system and effector proteins. *Proceedings of the National Academy of Sciences*, 97(16), 8770–8777.

Colombi, E. (2017). *The role of integrative conjugative elements in evolution of the kiwifruit pathogen Pseudomonas syringae pv. actinidiae*. PhD thesis, Massey University.

Colombi, E., Straub, C., Künzel, S., Templeton, M. D., McCann, H. C., & Rainey, P. B. (2017). Evolution of copper resistance in the kiwifruit pathogen *Pseudomonas syringae* pv. *actinidiae* through acquisition of integrative conjugative elements and plasmids. *Environmental Microbiology*, 19(2), 819–832.

Compant, S., Duffy, B., Nowak, J., Cle, C., & Barka, E. a. (2005). Use of plant growth-promoting bacteria for biocontrol of plant diseases: principles, mechanisms of action, and future prospects. *Applied and Environmental Microbiology*, 71(9), 4951–4959.

Cook, K. A. & Cain, R. B. (1974). Regulation of aromatic metabolism in the fungi: metabolic control of the 3-oxoadipate pathway in the yeast *Rhodotorula mucilaginosa*. *Journal of General Microbiology*, 85(1), 37–50.

Cook, R. J., Thomashow, L. S., Weller, D. M., Fujimoto, D., Mazzola, M., Bangera, G., & Kim, D. S. (1995). Molecular mechanisms of defense by rhizobacteria against root disease. *Proceedings of the National Academy of Sciences*, 92(10), 4197–4201.

Cordero, O. X., Ventouras, L.-a., DeLong, E. F., & Polz, M. F. (2012). Public good dynamics drive evolution of iron acquisition strategies in natural bacterioplankton populations. *Proceedings of the National Academy of Sciences*, 109(49), 20059–20064.

Cornforth, D. M. & Foster, K. R. (2013). Competition sensing: the social side of bacterial stress responses. *Nature Reviews Microbiology*, 11, 285–293.

Cunty, A., Cesbron, S., Poliakoff, F., Jacques, M.-A., & Manceau, C. (2015). The origin of the outbreak in France of *Pseudomonas syringae* pv. *actinidiae* biovar 3, the causal agent of bacterial canker of Kiwifruit, revealed by a Multilocus Variable-number of Tandem Repeat Analysis. *Applied and Environmental Microbiology*, (pp. AEM.01688–15).

D'aes, J., De Maeyer, K., Pauwelyn, E., & Höfte, M. (2010). Biosurfactants in plant – *Pseudomonas* interactions and their importance to biocontrol. *Environmental Microbiology*, 2, 359–372.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., Depristo, M. A., Handsaker, R. E., Lunter,

G., Marth, G. T., Sherry, S. T., Mcvean, G., Durbin, R., & 1000 Genomes Project Analysis Group (2011). The variant call format and VCFtools. *Bioinformatics*, 27(15), 2156–2158.

Dangl, J. L., Horvath, D. M., & Staskawicz, B. J. (2013). Pivoting the plant immune system from dissection to deployment. *Science*, 341(6147), 746–51.

Darling, A. E., Miklós, I., & Ragan, M. A. (2008). Dynamics of genome rearrangement in bacterial populations. *PLoS Genetics*, 4(7), e1000128.

Darriba, D., Taboada, G. L., Doallo, R., & Posada, D. (2012). jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods*, 9(8), 772–772.

Debroy, S., Thilmony, R., Kwack, Y.-B., Nomura, K., He, S. Y., & Lindow, S. E. (2004). A family of conserved bacterial effectors inhibits salicylic acid-mediated basal immunity and promotes disease necrosis in plants. *Proceedings of the National Academy of Sciences*, 101(26), 9927–9932.

Delmotte, N., Knief, C., Chaffron, S., Innerebner, G., Roschitzki, B., Schlapbach, R., von Mering, C., & Vorholt, J. A. (2009). Community proteogenomics reveals insights into the physiology of phyllosphere bacteria. *Proceedings of the National Academy of Sciences*, 106(38), 16428–16433.

Demba Diallo, M., Monteil, C. L., Vinatzer, B. A., Clarke, C. R., Glaux, C., Guilbaud, C., Desbiez, C., & Morris, C. E. (2012). *Pseudomonas syringae* naturally lacking the canonical type III secretion system are ubiquitous in nonagricultural habitats, are phylogenetically diverse and can be pathogenic. *The ISME Journal*, 6(7), 1325–1335.

Deslandes, L. & Rivas, S. (2012). Catch me if you can: Bacterial effectors and plant targets. *Trends in Plant Science*, 17(11), 644–655.

Dice, L. R. (1945). Measures of the amount of ecologic association between species. *Journal of Chemical Information and Modeling*, 53(26), 297–302.

Didelot, X., Achtman, M., Parkhill, J., Thomson, N. R., & Falush, D. (2007). A bimodal pattern of relatedness between the *Salmonella* Paratyphi A and Typhi genomes: Convergence or divergence by homologous recombination? *Genome Research*, 17, 61–68.

Didelot, X., Bowden, R., Street, T., Golubchik, T., Spencer, C., McVean, G., Sangal, V., Anjum, M. F., Achtman, M., Falush, D., & Donnelly, P. (2011). Recombination and population structure in *Salmonella enterica*. *PLoS Genetics*, 7(7), e1002191.

Didelot, X. & Falush, D. (2007). Inference of bacterial microevolution using multilocus sequence data. *Genetics*, 175(3), 1251–1266.

Didelot, X. & Maiden, M. C. J. (2010). Impact of recombination on bacterial evolution. *Trends in Microbiology*, 18(7), 315–322.

Didelot, X., Meric, G., Falush, D., & Darling, A. E. (2012). Impact of homologous and non-homologous recombination in the genomic evolution of *Escherichia coli*. *BMC Genomics*, 13(256).

Didelot, X. & Wilson, D. J. (2015). ClonalFrameML: Efficient inference of recombination in whole bacterial genomes. *PLoS Computational Biology*, 11(2), e1004041.

Ditta, G., Stanfield, S., Corbin, D., & Helinski, D. R. (1980). Broad host range DNA cloning system for gram-negative bacteria: construction of a gene bank of *Rhizobium meliloti*. *Proceedings of the National*

Academy of Sciences, 77(12), 7347–7351.

Doan, H. K. & Leveau, J. H. J. (2015). Artificial surfaces in phyllosphere microbiology. *Phytopathology*, 105(8), 1036 – 1042.

Durbin, R. D. (1982). Toxins and Pathogenesis. In M. S. Mount & G. H. Lacy (Eds.), *Phytopathogenic Prokaryotes Vol 1* (pp. 423–439). Academic Press.

Earl, D. A. & VonHoldt, B. M. (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, 4(2), 359–361.

Edwards, A. & Cavalli-Sforza, L. L. (1963). The reconstruction of evolution. *Annals of Human Genetics*, 27, 105–106.

Edwards, A. W. F. & Cavalli-Sforza, L. L. (1965). A method for cluster analysis. *Biometrics*, 21(2), 362–375.

Elena, S. F. & Lenski, R. E. (1997). Test of synergistic interactions among deleterious mutations in bacteria. *Nature*, 390(6658), 395–398.

Elizabeth, S. V. & Bender, C. L. (2007). The phytotoxin coronatine from *Pseudomonas syringae* pv. *tomato* DC3000 functions as a virulence factor and influences defence pathways in edible brassicas. *Molecular Plant Pathology*, 8(1), 83–92.

Engering, A., Hogerwerf, L., & Slingenbergh, J. (2013). Pathogen–host–environment interplay and disease emergence. *Emerging Microbes and Infections*, 2(e5).

EPPO (2011). First record of *Pseudomonas syringae* pv. *actinidiae* in Australia. *EPPO Reporting Service*, 6, 130.

Ercolani, G. L. (1991). Distribution of epiphytic bacteria on olive leaves and the influence of leaf age and sampling time. *Microbial Ecology*, 21(1), 35–48.

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology*, 14(8), 2611–2620.

Everett, K. R., Taylor, R. K., Romberg, M. K., Rees-George, J., Fullerton, R. A., Vanneste, J. L., & Manning, M. A. (2011). First report of *Pseudomonas syringae* pv. *actinidiae* causing kiwifruit bacterial canker in New Zealand. *Australasian Plant Disease Notes*, 6(1), 67–71.

Ezewudo, M. N., Joseph, S. J., Castillo-Ramirez, S., Dean, D., Del Rio, C., Didelot, X., Dillon, J.-A., Selden, R. F., Shafer, W. M., Turingan, R. S., Unemo, M., & Read, T. D. (2015). Population structure of *Neisseria gonorrhoeae* based on whole genome data and its relationship with antibiotic resistance. *PeerJ*, 3, e806.

Failor, K. C., Schmale, D. G., Vinatzer, B. A., & Monteil, C. L. (2017). Ice nucleation active bacteria in precipitation are genetically diverse and nucleate ice by employing different mechanisms. *The ISME Journal*, (pp. 1–14, doi:10.1038/ismej.2017.124).

Fakhr, M. K., Peñaloza-Vázquez, A., Chakrabarty, A. M., & Bender, C. L. (1999). Regulation of alginate biosynthesis in *Pseudomonas syringae* pv. *syringae*. *Journal of Bacteriology*, 181(11), 3478–3485.

Faruque, S. M., Chowdhury, N., Kamruzzaman, M., Dziejman, M., Rahman, M. H., Sack, D. A., Nair,

G. B., & Mekalanos, J. J. (2004). Genetic diversity and virulence potential of environmental *Vibrio cholerae* population in a cholera-endemic area. *Proceedings of the National Academy of Sciences*, 101(7), 2123–2128.

Feil, E. J., Li, B. C., Aanensen, D. M., Hanage, W. P., & Spratt, B. G. (2004). eBURST: inferring patterns of evolutionary descent among clusters of related bacterial genotypes from Multilocus Sequence Typing Data. *Journal of Bacteriology*, 186(5), 1518–1530.

Feil, E. J. & Spratt, B. G. (2001). Recombination and the population structures of bacterial pathogens. *Microbiology*, 55, 561–590.

Feil, H., Feil, W. S., Chain, P., Larimer, F., DiBartolo, G., Copeland, A., Lykidis, A., Trong, S., Nolan, M., Goltsman, E., Thiel, J., Malfatti, S., Loper, J. E., Lapidus, A., Detter, J. C., Land, M., Richardson, P. M., Kyrpides, N. C., Ivanova, N., & Lindow, S. E. (2005). Comparison of the complete genome sequences of *Pseudomonas syringae* pv. *syringae* B728a and pv. *tomato* DC3000. *Proceedings of the National Academy of Sciences of the United States of America*, 102(31), 11064–11069.

Felix, G. & Boller, T. (2003). Molecular sensing of bacteria in plants: The highly conserved RNA-binding motif RNP-1 of bacterial cold shock proteins is recognized as an elicitor signal in tobacco. *Journal of Biological Chemistry*, 278(8), 6201–6208.

Felsenstein, J. (1978). Cases in which parsimony or compatibility methods will be positively misleading. *Systematic Zoology*, 27(4), 401–410.

Felsenstein, J. (1981). Evolutionary trees from DNA sequences: A maximum likelihood approach. *Journal of Molecular Evolution*, 17(6), 368–376.

Felsenstein, J. (1988). Phylogenies from molecular sequences: inference and reliability. *Annual Review of Genetics*, 22, 521–565.

Felsenstein, J. (1989). Phylip: phylogeny inference package (version 3.2). *Cladistics*, 5, 164–166.

Ferrante, P. & Scortichini, M. (2010). Molecular and phenotypic features of *Pseudomonas syringae* pv. *actinidiae* isolated during recent epidemics of bacterial canker on yellow kiwifruit (*Actinidia chinensis*) in central Italy. *Plant Pathology*, 59(5), 954–962.

Fierer, N. & Lennon, J. (2011). The generation and maintenance of diversity in microbial communities. *American Journal of Botany*, 98(3), 439–448.

Fitt, B. D. L., Huang, Y.-J., van den Bosch, F., & West, J. S. (2006). Coexistence of related pathogen species on arable crops in space and time. *Annual Review of Phytopathology*, 44, 163–182.

Frampton, R. A., Pitman, A. R., & Fineran, P. C. (2012). Advances in bacteriophage-mediated control of plant pathogens. *International Journal of Microbiology*, (326452), 1–11.

Francisco, A. P., Vaz, C., Monteiro, P. T., Melo-Cristino, J., Ramirez, M., & Carriço, J. A. (2012). PHYLOViZ: phylogenetic inference and data visualization for sequence based typing methods. *BMC Bioinformatics*, 13(1), 87.

Fraser, C., Alm, E. J., Polz, M. F., Spratt, B. G., & Hanage, W. P. (2009). The bacterial species challenge: making sense of genetic and ecological diversity. *Science*, 323(5915), 741–746.

Fu, Y. X. & Li, W. H. (1993). Statistical tests of neutrality of mutations. *Genetics*, 133(3), 693–709.

Fuchs, G., Boll, M., & Heider, J. (2011). Microbial degradation of aromatic compounds — from one strategy to four. *Nature Reviews Microbiology*, 9(11), 803–816.

Fujikawa, T. & Sawada, H. (2016). Genome analysis of the kiwifruit canker pathogen *Pseudomonas syringae* pv. *actinidiae* biovar 5. *Scientific Reports*, 6(21399).

Fürnkranz, M., Wanek, W., Richter, A., Abell, G., Rasche, F., & Sessitsch, A. (2008). Nitrogen fixation by phyllosphere bacteria associated with higher plants and their colonizing epiphytes of a tropical lowland rainforest of Costa Rica. *The ISME Journal*, 2(5), 561–570.

Gal, M., Preston, G. M., Massey, R. C., Spiers, A. J., & Rainey, P. B. (2003). Genes encoding a cellulosic polymer contribute toward the ecological success of *Pseudomonas fluorescens* SBW25 on plant surfaces. *Molecular Ecology*, 12(11), 3109–3121.

Galán, J. E. & Collmer, A. (1999). Type III secretion machines: bacterial devices for protein delivery into host cells. *Science*, 284(5418), 1322–1328.

Gardan, L., David, C., Morel, M., Glickmann, E., Abu-Ghorrah, M., Petit, A., & Dessaux, Y. (1992). Evidence for a correlation between auxin production and host plant species among strains of *Pseudomonas syringae* subsp. *savastanoi*. *Applied and Environmental Microbiology*, 58(5), 1780–1783.

Gardan, L., Shafik, H., Belouin, S., Broch, R., Grimon, F., & Grimont, P. A. D. (1999). DNA relatedness among the pathovars of *Pseudomonas syringae* and the description of *Pseudomonas tremae* sp. nov. and *Pseudomonas cannabina* sp. nov. (ex Sutic and Dowson 1959). *International Journal of Systematic Bacteriology*, 49, 469–478.

Gause, G. F. (1934). *The Struggle for Existence*, volume 7. Williams & Wilkins.

Glickmann, E., Gardan, L., Jacquet, S., Hussain, S., Elasri, M., Petit, A., & Dessaux, Y. (1998). Auxin production is a common feature of most pathovars of *Pseudomonas syringae*. *Molecular Plant-Microbe Interactions*, 11(2), 156–162.

Gómez-Gómez, L. & Boller, T. (2002). Flagellin perception: A paradigm for innate immunity. *Trends in Plant Science*, 7(6), 251–256.

Goss, E. M., Potnis, N., & Jones, J. B. (2013). Grudgingly sharing their secrets: New insight into the evolution of plant pathogenic bacteria. *New Phytologist*, 199(3), 630–632.

Green, R. L. & Warren, G. J. (1985). Physical and functional repetition in a bacterial ice nucleation gene. *Nature*, 317(6038), 645–648.

Green, S., Studholme, D. J., Laue, B. E., Dorati, F., Lovell, H., Arnold, D., Cottrell, J. E., Bridgett, S., Blaxter, M., Huitema, E., Thwaites, R., Sharp, P. M., Jackson, R. W., & Kamoun, S. (2010). Comparative genome analysis provides insights into the evolution and adaptation of *Pseudomonas syringae* pv. *aesculi* on *Aesculus hippocastanum*. *PLoS ONE*, 5(4), e10224.

Griffin, A. S., West, S. A., & Buckling, A. (2004). Cooperation and competition in pathogenic bacteria. *Letters to Nature*, 430(7003), 1024–1027.

Guindon, S. & Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology*, 52(5), 696–704.

Guo, M., Tian, F., Wamboldt, Y., & Alfano, J. R. (2009). The majority of the type III effector inventory of *Pseudomonas syringae* pv. *tomato* DC3000 can suppress plant immunity. *Molecular Plant-Microbe Interactions*, 22(9), 1069–1080.

Gurevich, A., Saveliev, V., Vyahhi, N., & Tesler, G. (2013). QUAST: Quality assessment tool for genome assemblies. *Bioinformatics*, 29(8), 1072–1075.

Gutiérrez-Barranquero, J. a., Carrión, V. J., Murillo, J., Arrebola, E., Arnold, D. L., Cazorla, F. M., & de Vicente, A. (2013). A *Pseudomonas syringae* diversity survey reveals a differentiated phylotype of the pathovar *syringae* associated with the mango host and mangotoxin production. *Phytopathology*, 103(11), 1115–29.

Haapalainen, M., Mosorin, H., Dorati, F., Wu, R. F., Roine, E., Taira, S., Nissinen, R., Mattinen, L., Jackson, R., Pirhonen, M., & Lin, N. C. (2012). Hcp2, a secreted protein of the phytopathogen *Pseudomonas syringae* pv. *tomato* DC3000, is required for fitness for competition against bacteria and yeasts. *Journal of Bacteriology*, 194(18), 4810–4822.

Haas, D. & Défago, G. (2005). Biological control of soil-borne pathogens by fluorescent pseudomonads. *Nature Reviews Microbiology*, 3(4), 307–319.

Hacker, J., Blum-Oehler, G., Mühldorfer, I., & Tschäpe, H. (1997). Pathogenicity islands of virulent bacteria: Structure, function and impact on microbial evolution. *Molecular Microbiology*, 23(6), 1089–1097.

Hall, S. J., Dry, I. B., Blanchard, C. L., & Whitelaw-Weckert, M. A. (2016). Phylogenetic relationships of *Pseudomonas syringae* pv. *syringae* isolates associated with bacterial iInflorescence rot in grapevine. *Plant Disease*, 100(3), 607–616.

Han, H. S., Koh, Y. J., Hur, J.-S., & Jung, J. S. (2004). Occurrence of the strA-strB streptomycin resistance genes in Pseudomonas species isolated from kiwifruit plants. *The Journal of Microbiology*, 42(4), 365–368.

Han, H. S., Nam, H. Y., Koh, Y. J., Hur, J.-S., & Jung, J. S. (2003). Molecular bases of high-level streptomycin resistance in *Pseudomonas marginalis* and *Pseudomonas syringae* pv. *actinidiae*. *The Journal of Microbiology*, 41(1), 16–21.

Handelsman, J. & Stabb, E. V. (1996). Biocontrol of soilborne plant pathogens. *The Plant Cell*, 8(10), 1855–1869.

Hansen, S. K., Rainey, P. B., Haagensen, J. A. J., & Molin, S. (2007). Evolution of species interactions in a biofilm community. *Nature*, 445(7127), 533–536.

Hardin, G. (1960). The competitive exclusion principle. *Science*, 131, 1292–1297.

Harwood, C. S. & Parales, R. E. (1996). The b-ketoadipate pathway and the biology of self-identity. *Annual Review of Microbiology*, 50(1), 553–590.

Haubold, B. & Hudson, R. R. (2000). LIAN 3.0: detecting linkage disequilibrium in multilocus data. *Bioinformatics*, 16(9), 847–848.

Haubold, B. & Rainey, P. B. (1996). Genetic and ecotypic structure of a fluorescent *Pseudomonas* population. *Molecular Ecology*, 5, 747–761.

Haubold, B., Travisano, M., Rainey, P. B., & Hudson, R. R. (1998). Detecting linkage disequilibrium in bacterial populations. *Genetics*, 150(4), 1341–8.

Hayes, C. S., Aoki, S. K., & Low, D. A. (2010). Bacterial contact-dependent delivery systems. *Annual Review of Genetics*, 44(1), 71–90.

Hellstein, J., Vawter-Hugart, H., Fotos, P., Schmid, J., & Soll, D. R. (1993). Genetic similarity and phenotypic diversity of commensal and pathogenic strains of *Candida albicans* isolated from the oral cavity. *Journal of Clinical Microbiology*, 31(12), 3190–3199.

Hibbing, M. E., Fuqua, C., Parsek, M. R., & Peterson, S. B. (2010). Bacterial competition: surviving and thriving in the microbial jungle. *Nature Reviews Microbiology*, 8(1), 15–25.

Hill, T. C. J., Moffett, B. F., DeMott, P. J., Georgakopoulos, D. G., Stump, W. L., & Franc, G. D. (2014). Measurement of ice nucleation-active bacteria on plants and in precipitation by quantitative PCR. *Applied and Environmental Microbiology*, 80(4), 1256–1267.

Hirano, S. S., Charkowski, A. O., Collmer, A., Willis, D. K., & Upper, C. D. (1999). Role of the Hrp type III protein secretion system in growth of *Pseudomonas syringae* pv. *syringae* B728a on host plants in the field. *Proceedings of the National Academy of Sciences*, 96(17), 9851–9856.

Hirano, S. S. & Upper, C. D. (2000). Bacteria in the leaf ecosystem with emphasis on *Pseudomonas syringae* - a pathogen, ice nucleus, and epiphyte. *Microbiology and Molecular Biology Reviews*, 64(3), 624–653.

Hockett, K. L., Nishimura, M. T., Karlsrud, E., Dougherty, K., & Baltrus, D. a. (2014). *Pseudomonas syringae* CC1557: a highly virulent strain with an unusually small Type III effector repertoire that includes a novel effector. *Molecular Plant-Microbe Interactions*, 27(9), 923–32.

Hockett, K. L., Renner, T., & Baltrus, D. A. (2015). Independent co-option of a tailed bacteriophage into a killing complex in *Pseudomonas*. *mBio*, 6(4), 1–11.

Hol, F. J. H., Galajda, P., Woolthuis, R. G., Dekker, C., & Keymer, J. E. (2015). The idiosyncrasy of spatial structure in bacterial competition. *BMC Research Notes*, 8, 245.

Holmes, E. C., Urwin, R., & Maiden, M. C. (1999). The influence of recombination on the population structure and evolution of the human pathogen Neisseria meningitidis. *Molecular Biology and Evolution*, 16(6), 741–9.

Horvath, P. & Barrangou, R. (2010). CRISPR/Cas, the immune system of bacteria and archaea. *Science*, 327(5962), 167–170.

Huang, C.-J., Pauwelyn, E., Ongena, M., Debois, D., Leclère, V., Jacques, P., Bleyaert, P., & Höfte, M. (2015). Characterization of cichopeptins, new phytotoxic cyclic lipodepsipeptides produced by *Pseudomonas cichorii* SF1-54 and their role in bacterial midrib rot disease of lettuce. *Molecular Plant-Microbe Interactions*, 28(9), 1009–1022.

Huang, R., Li, M., & Gregory, R. L. (2011). Bacterial interactions in dental biofilm. *Virulence*, 2(5), 435–444.

Huang, W., Li, L., Myers, J. R., & Marth, G. T. (2012). ART: A next-generation sequencing read simulator. *Bioinformatics*, 28(4), 593–594.

Hube, B. (2004). From commensal to pathogen: Stage- and tissue-specific gene expression of *Candida albicans*. *Current Opinion in Microbiology*, 7(4), 336–341.

Hudson, R. R. (2001). Two-locus sampling distributions and their application. *Genetics*, 159(4), 1805–1817.

Humphrey, P. T., Nguyen, T. T., Villalobos, M. M., & Whiteman, N. K. (2014). Diversity and abundance of phyllosphere bacteria are linked to insect herbivory. *Molecular Ecology*, 23, 1497–1515.

Huson, D. H. & Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution*, 23(2), 254–267.

Hutcheson, S. W., Bretz, J., Sussan, T., Jin, S., & Pak, K. (2001). Enhancer-binding proteins HrpR and HrpS interact to regulate hrp-encoded type III protein secretion in *Pseudomonas syringae* strains. *Journal of Bacteriology*, 183(19), 5589–5598.

Hwang, M. S. H., Morgan, R. L., Sarkar, S. F., Wang, P. W., & Guttman, D. S. (2005). Phylogenetic characterization of virulence and resistance phenotypes of *Pseudomonas syringae*. *Applied and Environmental Microbiology*, 71(9), 5182–5191.

Istock, C. A., Duncan, K. E., Ferguson, N., & Zhou, X. (1992). Sexuality in a natural population of bacteria - *Bacillus subtilis* challenges the clonal paradigm. *Molecular Ecology*, 1, 95–103.

Jackson, C. R. & Denney, W. C. (2011). Annual and Seasonal Variation in the Phyllosphere Bacterial Community Associated with Leaves of the Southern Magnolia (*Magnolia grandiflora*). *Microbial Ecology*, 61(1), 113–122.

Jackson, R. W., Preston, G. M., & Rainey, P. B. (2005). Genetic characterization of *Pseudomonas fluorescens* SBW25 rsp gene expression in the phytosphere and *in vitro*. *Journal of Bacteriology*, 187(24), 8477–8488.

Jakobsson, M. & Rosenberg, N. A. (2007). CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, 23(14), 1801–1806.

Jelenska, J., van Hal, J. A., & Greenberg, J. T. (2010). *Pseudomonas syringae* hijacks plant stress chaperone machinery for virulence. *Proceedings of the National Academy of Sciences*, 107(29), 13177–13182.

Jeltsch, A. (2003). Maintenance of species identity and controlling speciation of bacteria: A new function for restriction/modification systems? *Gene*, 317, 13–16.

Jeong, B.-R., van Dijk, K., & Alfano, J. R. (2009). *Pseudomonas syringae* Type III-secreted proteins and their activities and effects on plant innate immunity. *Annual Plant Reviews*, 34, 48–76.

Jiménez, J. I., Miñambres, B., García, J. L., & Díaz, E. (2002). Genomic analysis of the aromatic catabolic pathways from *Pseudomonas putida* KT2440. *Environmental Microbiology*, 4(12), 824–841.

Joardar, V., Lindeberg, M., Jackson, R. W., Selengut, J., Dodson, R., Brinkac, L. M., Daugherty, S. C., Deboy, R., Durkin, A. S., Giglio, M. G., Madupu, R., Nelson, W. C., Rosovitz, M. J., Sullivan, S., Crabtree, J., Creasy, T., Davidsen, T., Haft, D. H., Zafar, N., Zhou, L., Halpin, R., Holley, T., Khouri, H., Feldblyum, T., White, O., Fraser, C. M., Chatterjee, A. K., Cartinhour, S., Schneider, D. J.,

Mansfield, J., Collmer, A., & Buell, C. R. (2005). Whole-genome sequence analysis of *Pseudomonas syringae* pv. *phaseolicola* 1448A reveals divergence among pathovars in genes involved in virulence and transposition. *Journal of Bacteriology*, 187(18), 6488–6498.

Jolley, K. A., Feil, E. J., Chan, M. S., & Maiden, M. C. (2001). Sequence type analysis and recombinational tests (START). *Bioinformatics*, 17(12), 1230–1231.

Jones, J. & Dangl, J. (2006). The plant immune system. *Nature*, 444(7117), 323–329.

Jost, L. (2006). Entropy and diversity. *Oikos*, 113, 363–375.

Juhas, M., Van Der Meer, J. R., Gaillard, M., Harding, R. M., Hood, D. W., & Crook, D. W. (2009). Genomic islands: Tools of bacterial horizontal gene transfer and evolution. *FEMS Microbiology Reviews*, 33(2), 376–393.

Kaluzna, M., Ferrante, P., Sobiczewski, P., & Scortichini, M. (2010). Characterization and genetic diversity of *Pseudomonas syringae* from stone fruits and hazelnut using Repetitive-PCR and MLST. *Journal of Plant Pathology*, 92(3), 781–787.

Kassen, R. & Rainey, P. B. (2004). The ecology and genetics of microbial diversity. *Annual Review of Microbiology*, 58, 207–231.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., & Drummond, A. (2012). Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, 28(12), 1647–1649.

Keen, N. T. & Buzzell, R. I. (1991). New disease resistance genes in soybean agaisnt *Pseudomonas syringae* pv. *glycinea* - evidence that one of them interacts with a bacterial elicitor. *Theoretical and Applied Genetics*, 81(1), 133–138.

Keith, L. W., Boyd, C., Keen, N. T., & Partridge, J. E. (1997). Comparison of *avrD* alleles from *Pseudomonas syringae* pv. *glycinea*. *Molecular Plant-Microbe Interactions*, 10(3), 416–422.

Keith, R. C., Keith, L. M. W., Hernández-Guzmán, G., Uppalapati, S. R., & Bender, C. L. (2003). Alginate gene expression by *Pseudomonas syringae* pv. *tomato* DC3000 in host and non-host plants. *Microbiology*, 149(5), 1127–1138.

Kinscherf, T. G., Coleman, R. H., Barta, T. M., & Willis, D. K. (1991). Cloning and expression of the tabtoxin biosynthetic region from *Pseudomonas syringae*. *Journal of Bacteriology*, 173(13), 4124–4132.

Kinscherf, T. G. & Willis, D. K. (2005). The biosynthetic gene cluster for the β-lactam antibiotic tabtoxin in *Pseudomonas syringae*. *The Journal of Antibiotics*, 58(12), 817–821.

Klement, Z. & Lovrekovich, L. (1961). Defence reactions induced by phytopathogenic bacteria in bean pods. *Journal of Phytopathology*, 41(3), 217–227.

Kloepper, J. W., Leong, J., Teintze, M., & Schroth, M. N. (1980a). Enhanced plant growth by siderophores produced by plant growth-promoting rhizobacteria. *Nature*, 286(5776), 885–886.

Kloepper, J. W., Schroth, M. N., & Miller, T. D. (1980b). Effects of rhizosphere colonization by plant growth-promoting rhizobacteria on potato plant development and yield. *Phytopathology*, 70(11), 1078–1082.

Kniskern, J. M., Barrett, L. G., & Bergelson, J. (2011). Maladaptation in wild populations of the generalist plant pathogen *Pseudomonas syringae*. *Evolution*, 65(3), 818–830.

Koh, Y. J. (1995). Economically important diseases of kiwifruit. *Plant Disease Agriculture*, 1, 3–13.

Koh, Y. J., Kim, G. H., Jung, J. S., Lee, Y. S., & Hur, J. S. (2010). Outbreak of bacterial canker on Hort16A (*Actinidia chinensis* Planchon) caused by *Pseudomonas syringae* pv. *actinidiae* in Korea. *New Zealand Journal of Crop and Horticultural Science*, 38(4), 275–282.

Koh, Y. J., Kim, G. H., Koh, H. S., Lee, Y. S., Kim, S. C., & Jung, J. S. (2012). Occurrence of a new type of *Pseudomonas syringae* pv. *actinidiae* strain of bacterial canker on kiwifruit in Korea. *Plant Pathology*, 28(4), 423–427.

Kõiv, V., Roosaare, M., Vedler, E., Ann Kivistik, P., Toppi, K., Schryer, D. W., Remm, M., Tenson, T., Mäe, A., Kivistik, P. A., Toppi, K., Schryer, D. W., Remm, M., Tenson, T., & Mäe, A. (2015). Microbial population dynamics in response to Pectobacterium atrosepticum infection in potato tubers. *Scientific Reports*, 5, 11106.

Koonin, E. V., Makarova, K. S., & Aravind, L. (2001). Horizontal gene transfer in prokaryotes: quantification and classification. *Annual Review of Microbiology*, 55(1), 709–742.

Kosakovsky Pond, S. L., Posada, D., Gravenor, M. B., Woelk, C. H., & Frost, S. D. W. (2006). Automated phylogenetic detection of recombination using a genetic algorithm. *Molecular Biology and Evolution*, 23(10), 1891–1901.

Kryazhimskiy, S. & Plotkin, J. B. (2008). The population genetics of dN/dS. *PLoS Genetics*, 4(12).

Kudela, V., Krejzar, V., & Pánková, I. (2010). *Pseudomonas corrugata* and *Pseudomonas marginalis* associated with the collapse of tomato plants in Rockwool slab hydroponic culture. *Plant Protection Science*, 46(1), 1–11.

Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, 33(7), 1870–1874.

Kvitko, B. H., Park, D. H., Velásquez, A. C., Wei, C. F., Russell, A. B., Martin, G. B., Schneider, D. J., & Collmer, A. (2009). Deletions in the repertoire of *Pseudomonas syringae* pv. *tomato* DC3000 type III secretion effector genes reveal functional overlap among effectors. *PLoS Pathogens*, 5(4), e1000388.

Laforest-Lapointe, I., Messier, C., Kembel, S. W., Lindow, S., Brandl, M., Herre, E., Mejía, L., Kyllo, D., Rojas, E., Maynard, Z., Butler, A., Fürnkranz, M., Wanek, W., Richter, A., Abell, G., Rasche, F., Sessitsch, A., Morris, C., Kinkel, L., Lindow, S., Hecht-Poinar, E., Elliott, V., Andrews, J., Harris, R., Lambais, M., Crowley, D., Cury, J., Büll, R., Rodrigues, R., Jumpponen, A., Jones, K., Rodriguez, R., White, J., Arnold, A., Redman, R., Redford, A., Bowers, R., Knight, R., Linhart, Y., Fierer, N., Arnold, A., Mejía, L., Kyllo, D., Rojas, E., Maynard, Z., Robbins, N., Herre, E., Vorholt, J., Müller, T., Ruppel, S., Osono, T., Suda, W., Nagasaki, A., Shishido, M., Gilbert, G., Newton, A., Gravouil, C., Fountaine, J., Kim, M., Singh, D., Lai-Hoe, A., Go, R., Rahim, R., Ainuddin, A., Kembel, S., O'Connor, T., Arnold, H., Hubbell, S., Wright, S., Green, J., Kembel, S., Mueller, R., Mercier, J., Lindow, S., Atamna-Ismaeel, N., Finkel, O., Glaser, F., Mering, C., Vorholt, J., Koblížek, M., Abanda-Nkpwatt, D., Müsch, M., Tschiersch, J., Boettner, M., Schwab, W., Innerebner, G., Knief, C., Vorholt,

J., Shade, A., Handelsman, J., Redford, A., Fierer, N., Burke, C., Thomas, T., Lewis, M., Steinberg, P., Kjelleberg, S., Knief, C., Ramette, A., Frances, L., Alonso-Blanco, C., Vorholt, J., Osono, T., Cordier, T., Robin, C., Capdevielle, X., Desprez-Loustau, M., Vacher, C., Finkel, O., Burch, A., Lindow, S., Post, A., Belkin, S., Rastogi, G., Tech, J., Coaker, G., Leveau, J., Delmotte, N., Knief, C., Chaffron, S., Innerebner, G., Roschitzki, B., Schlapbach, R., Schmidt, T., Rodrigues, J., Mering, C., Bulgarelli, D., Schlaeppi, K., Spaepen, S., Themaat, E., Schulze-Lefert, P., Janssen, P., Kim, M., Heo, E., Kang, H., Adams, J., Normand, P., Lindow, S., Arny, D., Upper, C., Biebl, H., Pfennig, N., Lambais, M., Lucheta, A., Crowley, D., Wright, I., Reich, P., Westoby, M., Ackerly, D., Baruch, Z., Bongers, F., Wright, S., Kitajima, K., Kraft, N., Reich, P., Wright, I., Bunker, D., Gloor, G., Hummelen, R., Macklaim, J., Dickson, R., Fernandes, A., MacPhee, R., Chelius, M., Triplett, E., Zhang, J., Kobert, K., Flouri, T., Stamatakis, A., Caporaso, J., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F., Costello, E., Edgar, R., DeSantis, T., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E., Keller, K., Huber, T., Dalevi, D., Hu, P., Andersen, G., Abrams, M., Kubiske, M., Farrar, J., Shipley, B., Vu, T., Niinemets, Ü., Valladares, F., Chave, J., Coomes, D., Jansen, S., Lewis, S., Swenson, N., Zanne, A., Segata, N., Izard, J., Waldron, L., Gevers, D., Miropolsky, L., Garrett, W., Acinas, S., Sarma-Rupavtarm, R., Klepac-Ceraj, V., Polz, M., Paradis, E., Claude, J., Strimmer, K., Wickham, H., Kembel, S., Cowan, P., Helmus, M., Cornwell, W., Morlon, H., Ackerly, D., Oksanen, J., Kindt, R., Legendre, P., O'Hara, B., Stevens, M., Oksanen, M., Lozupone, C., Hamady, M., Knight, R., Anderson, M., Hochberg, Y., Bland, J., & Altman, D. (2016). Host species identity, site and time drive temperate tree phyllosphere bacterial community structure. *Microbiome*, 4, 27.

Lamichhane, J. R., Messéan, A., & Morris, C. E. (2015). Insights into epidemiology and control of diseases of annual plants caused by the *Pseudomonas syringae* species complex. *Journal of General Plant Pathology*, 81(5), 331–350.

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 D. L. Sparks (Ed.), *Advances in Agronomy*, volume 126 (pp. 235–295). Academic Press.

Lamichhane, J. R. & Venturi, V. (2015). Synergisms between microbial pathogens in plant disease complexes: a growing trend. *Frontiers in Plant Science*, 6, 385.

Lan, R., Reeves, P. R., & Octavia, S. (2009). Population structure, origins and evolution of major *Salmonella enterica* clones. *Infection, Genetics and Evolution*, 9(5), 996–1005.

Langmead, B., Trapnell, C., Pop, M., & Salzberg, S. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome biology*, 10(3), R25.

Last, F. T. (1955). Seasonal incidence of *Sporobolomyces* on cereal leaves. *Transactions of the British Mycological Society*, 38(3), 221–239.

Lelliott, R. A., Billing, E., & Hayward, A. (1966). A determinative scheme for the fluorescent plant pathogenic pseudomonads. *Journal of Applied Bacteriology*, 29(3), 470–489.

Lenski, R. E. (1993). Assessing the genetic structure of microbial populations. *Proceedings of the National Academy of Sciences*, 90(10), 4334–4336.

Lenski, R. E., Rose, M. R., Simpson, S. C., & Tadler, S. C. (1991). Long-term experimental evolution in *Escherichia coli* . I . Adaptation and divergence during 2,000 generations. *The American Naturalist*, 138(6), 1315–1341.

Levin, B. R. & Cornejo, O. E. (2009). The population and evolutionary dynamics of homologous gene recombination in bacteria. *PLoS Genetics*, 5(8), e1000601.

Levin, B. R., Lipsitch, M., & Bonhoeffer, S. (1999). Population biology, evolution, and infectious disease: convergence and synthesis. *Science*, 283(5403), 806–809.

Lewis, J. D., Abada, W., Ma, W., Guttman, D. S., & Desveaux, D. (2008). The HopZ family of *Pseudomonas syringae* type III effectors require myristoylation for virulence and avirulence functions in *Arabidopsis thaliana*. *Journal of Bacteriology*, 190(8), 2880–2891.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment / Map format and SAMtools. *Bioinformatics*, 25(16), 2078–2079.

Li, R., Khafipour, E., Krause, D. O., Entz, M. H., de Kievit, T. R., & Fernando, W. G. D. (2012). Pyrosequencing reveals the influence of organic and conventional farming systems on bacterial communities. *PLoS ONE*, 7(12), e51897.

Li, W. & Godzik, A. (2006). Cd-hit: A fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics*, 22(13), 1658–1659.

Lindeberg, M., Cunnac, S., & Collmer, A. (2009). The evolution of *Pseudomonas syringae* host specificity and type III effector repertoires. *Molecular Plant Pathology*, 10(6), 767–775.

Lindeberg, M., Cunnac, S., & Collmer, A. (2012). *Pseudomonas syringae* type III effector repertoires: Last words in endless arguments. *Trends in Microbiology*, 20(4), 199–208.

Lindow, S. E. (1983). The role of bacterial ice nucleation in frost injury to plants. *Annual Reviews of Phytopathology*, 21(1), 363–384.

Lindow, S. E. (1986). Construction of isogenic Ice strains of *Pseudomonas syringae* for evaluation of specificity of competition on leaf surfaces. *Perspectives in Microbial Ecology. Slovene Society for Microbiology, Ljubljana*, (pp. 509–515).

Lindow, S. E. (1991). Tests of specificity of competition amound *Pseudomonas syringae* strains on plants using recombinant Ice-strains and use of ice nucleation genes asprobes of *in situ* transcriptional activity. In *Advances in Molecular Genetics of Plant-Microbe Interactions Vol. 1* (pp. 457–464). Springer Netherlands.

Lindow, S. E. (1995). Control of epiphytic ice nucleation-active bacteria for management of plant frost injury. *Biological ice nucleation and its applications*, (pp. 239–256).

Lindow, S. E., Arny, D. C., & Upper, C. D. (2013). *Erwinia herbicola*: a bacterial ice nucleus active in increasing frost injury to corn. *Phytopathology*, 68(3), 523–527.

Lindow, S. E. & Brandl, M. T. (2003). Microbiology of the phyllosphere. *Applied and Environmental Microbiology*, 69(4), 1875–1883.

Lindow, S. E., Gies, D. R., Willis, D. K., & Panopoulos, N. J. (1987). *Molecular analysis of the*

Pseudomonas syringae pv. syringae ice gene and construction and testing of Ice- deletion mutants for biological frost control, (pp. 1030). Springer Netherlands.

Lindow, S. E. & Leveau, J. H. J. (2002). Phyllosphere microbiology. *Current Opinion in Biotechnology*, 13(3), 238–243.

Lindow, S. E. & Panopoulos, N. J. (1988). Field tests of recombinant ice–*Pseudomonas syringae* for biological frost control in potato. *Release of genetically engineered microorganisms/M. Sussman...[et al.]*.

Lovell, H. C., Jackson, R. W., Mansfield, J. W., Godfrey, S. A. C., Hancock, J. T., Desikan, R., & Arnold, D. L. (2010). *In planta* conditions induce genomic changes in *Pseudomonas syringae* pv. *phaseolicola*. *Molecular Plant Pathology*, 12(2), 167–176.

Lugtenberg, B. & Kamilova, F. (2009). Plant-Growth-Promoting Rhizobacteria. *Annual Review of Microbiology*, 63(1), 541–556.

Luo, C., Walk, S. T., Gordon, D. M., Feldgarden, M., Tiedje, J. M., & Konstantinidis, K. T. (2011). Genome sequencing of environmental *Escherichia coli* expands understanding of the ecology and speciation of the model bacterial species. *Philosophical Transactions of the Royal Society*, 108(17), 7200–7205.

Ma, W., Dong, F. F. T., Stavrinides, J., & Guttman, D. S. (2006). Type III effector diversification via both pathoadaptation and horizontal transfer in response to a coevolutionary arms race. *PLoS Genetics*, 2(12), e209.

Macho, A. P., Zumaquero, A., Ortiz-Martín, I., & Beuzón, C. R. (2007). Competitive index in mixed infections: A sensitive and accurate assay for the genetic analysis of *Pseudomonas syringae*-plant interactions. *Molecular Plant Pathology*, 8(4), 437–450.

Mackey, D. & McFall, A. J. (2006). MAMPs and MIMPs: Proposed classifications for inducers of innate immunity. *Molecular Microbiology*, 61(6), 1365–1371.

Maiden, M. C., Bygraves, J. a., Feil, E., Morelli, G., Russell, J. E., Urwin, R., Zhang, Q., Zhou, J., Zurth, K., Caugant, D. a., Feavers, I. M., Achtman, M., & Spratt, B. G. (1998). Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. *Proceedings of the National Academy of Sciences*, 95(6), 3140–3145.

Mansfield, J., Genin, S., Magori, S., Citovsky, V., Sriariyanum, M., Ronald, P., Dow, M., Verdier, V., Beer, S. V., Machado, M. A., Toth, I., Salmond, G., & Foster, G. D. (2012). Top 10 plant pathogenic bacteria in molecular plant pathology. *Molecular Plant Pathology*, 13(6), 614–629.

Marcelletti, S. & Scortichini, M. (2011). Clonal outbreaks of bacterial canker caused by *Pseudomonas syringae* pv. *actinidiae* on *Actinidia chinensis* and *A. deliciosa* in Italy. *Journal of Plant Pathology*, 93(2), 479–483.

Marchi, G., Sisto, A., Cimmino, A., Andolfi, A., Cipriani, M. G., Evidente, A., & Surico, G. (2006). Interaction between *Pseudomonas savastanoi* pv. *savastanoi* and *Pantoea agglomerans* in olive knots. *Plant Pathology*, 55(5), 614–624.

Marchi, M., Boutin, M., Gazengel, K., Rispe, C., Gauthier, J., Guillerm-Erckelboudt, A., Lebreton, L.,

Barret, M., Daval, S., & Sarniguet, A. (2013). Genomic analysis of the biocontrol strain *Pseudomonas fluorescens* Pf29Arp with evidence of T3SS and T6SS gene expression on plant roots. *Environmental Microbiology Reports*, 5(3), 393–403.

Marcon, J., Maccheroni, W., Elsas, J. D. V., & Vuurde, J. W. L. V. (2002). Diversity of endophytic bacterial populations and their interaction with *Xylella fastidiosa* in citrus plants. *Applied and Environmental Microbiology*, 68(10), 4906–4914.

Marraffini, L. A. & Sontheimer, E. J. (2008). CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science*, 322(5909), 1843–1845.

Martínez-García, P. M., Rodríguez-Palenzuela, P., Arrebola, E., Carrión, V. J., Gutièrrez-Barranquero, J. A., Pèrez-García, A., Ramos, C., Cazorla, F. M., & De Vicente, A. (2015). Bioinformatics analysis of the complete genome sequence of the mango tree pathogen *Pseudomonas syringae* pv. *syringae* UMAF0158 reveals traits relevant to virulence and epiphytic lifestyle. *PLoS ONE*, 10(8), e0136101.

Mau, B., Newton, M. A., & Larget, B. (1999). Bayesian phylogenetic inference via Markov Chain Monte Carlo methods. *Biometrics*, 55(1), 1–12.

Mavrodi, D. V., Joe, A., Mavrodi, O. V., Hassan, K. A., Weller, D. M., Paulsen, I. T., Loper, J. E., Alfano, J. R., & Thomashow, L. S. (2011). Structural and functional analysis of the type III secretion system from *Pseudomonas fluorescens* Q8r1-96. *Journal of Bacteriology*, 193(1), 177–189.

Maynard Smith, J. (1991). The population genetics of bacteria. *Proceedings of the Royal Society B: Biological Sciences*, 245, 37–41.

Maynard Smith, J., Dowson, C. G., & Spratt, B. G. (1991). Localized sex in bacteria. *Nature*, 349(6304), 29–31.

Maynard Smith, J., Feil, E. J., & Smith, N. H. (2000). Population structure and evolutionary dynamics of pathogenic bacteria. *BioEssays*, 22(12), 1115–22.

Maynard Smith, J., Smith, N. H., O'Rourke, M., & Spratt, B. G. (1993). How clonal are bacteria? *Proceedings of the National Academy of Sciences*, 90(10), 4384–4388.

Mazodier, P. & Davies, J. (1991). Gene transfer between distantly related bacteria. *Annual Review of Genetics*, 25(1), 147–71.

McArdle, B. H. & Anderson, M. J. (2001). Fitting multivariate models to community data: a comment on distance-based redundancy analysis. *Ecology*, 82(1), 290–297.

McCann, H. C. & Guttman, D. S. (2007). Evolution of the type III secretion system and its effectors in plant-microbe interactions. *New Phytologist*, 177(1), 33–47.

McCann, H. C., Li, L. L., Liu, Y., Li, D., Pan, H., Zhong, C., Rikkerink, E. H., Templeton, M. D., Straub, C., Colombi, E., Rainey, P. B., & Huang, H. (2017). Origin and evolution of the kiwifruit canker pandemic. *Genome Biology and Evolution*, 9(4), 932–944.

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., & 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(7), e1003503.

McDonald, B. A. & Stukenbrock, E. H. (2016). Rapid emergence of pathogens in agro-ecosystems: global threats to agricultural sustainability and food security. *Philosophical Transactions of the Royal Society B*, 371(1709), 20160026.

McNally, A., Thomson, N. R., Reuter, S., & Wren, B. W. (2016). 'Add, stir and reduce': *Yersinia* spp. as model bacteria for pathogen evolution. *Nature Reviews Microbiology*, 14(3), 177–190.

McVean, G., Awadalla, P., & Fearnhead, P. (2002). A coalescent-based method for detecting and estimating recombination from gene sequences. *Genetics*, 160(3), 1231–1241.

Meaden, S. & Koskella, B. (2017). Adaptation of the pathogen, *Pseudomonas syringae*, during experimental evolution on a native vs. alternative host plant. *Microbial Ecology*, 26, 1790–1801.

Melotto, M., Underwood, W., Koczan, J., Nomura, K., & He, S. Y. (2006). Plant stomata function in innate immunity against bacterial invasion. *Cell*, 126(5), 969–980.

Michener, C. D. & Sokal, R. R. (1957). A quantitative approach to a problem in classification. *Evolution*, 11(2), 130–162.

Middelhoven, W. J. (1993). Catabolism of benzene compounds by ascomycetous and basidiomycetous yeasts and yeastlike fungi. A literature review and an experimental approach. *Antonie van Leeuwenhoek*, 63(2), 125–144.

Milkman, R. (1973). Electrophoretic variation in *Escherichia coli* from natural sources. *Science*, 182(4116), 1024–1026.

Mohr, T. J., Liu, H., Yan, S., Morris, C. E., Castillo, J. A., Jelenska, J., & Vinatzer, B. A. (2008). Naturally occurring nonpathogenic isolates of the plant pathogen *Pseudomonas syringae* lack a type III secretion system and effector gene orthologues. *Journal of Bacteriology*, 190(8), 2858–2870.

Monier, J. & Lindow, S. E. (2005). Spatial organization of dual-species bacterial aggregates on leaf surfaces. *Applied Environmental Microbiology*, 71(9), 5484–5493.

Monteil, C. L., Cai, R., Liu, H., Mechan Llontop, M. E., Leman, S., Studholme, D. J., Morris, C. E., & Vinatzer, B. A. (2013). Nonagricultural reservoirs contribute to emergence and evolution of *Pseudomonas syringae* crop pathogens. *New Phytologist*, 199(3), 800–811.

Monteil, C. L., Guilbaud, C. C., Glaux, C. C., Lafolie, F. F., Soubeyrand, S. S., & Morris, C. E. C. E. (2012). Emigration of the plant pathogen *Pseudomonas syringae* from leaf litter contributes to its population dynamics in alpine snowpack. *Environmental Microbiology*, 14(8), 2099–2112.

Monteil, C. L., Lafolie, F., Laurent, J., Clement, J.-C. C., Simler, R., Travi, Y., & Morris, C. E. (2014). Soil water flow is a source of the plant pathogen *Pseudomonas syringae* in subalpine headwaters. *Environmental Microbiology*, 16(7), 2038–2052.

Monteil, C. L., Yahara, K., Studholme, D. J., Mageiros, L., Méric, G., Swingle, B., Morris, C. E., Vinatzer, B. A., & Sheppard, S. K. (2016). Population-genomic insights into emergence, crop-adaptation, and dissemination of *Pseudomonas syringae* pathogens. *Microbial Genomics*, 2(10), 1–16.

Moretti, C., Hosni, T., Vandemeulebroecke, K., Brady, C., de Vos, P., Buonaurio, R., & Cleenwerck, I. (2011). *Erwinia oleae* sp. *nov.*, isolated from olive knots caused by *Pseudomonas savastanoi* pv.

savastanoi. *International Journal of Systematic and Evolutionary Microbiology*, 61(11), 2745–2752.

Morris, C. E. (2002). Phyllosphere. *eLife*, (pp. doi: 10.1038/npg.els.0).

Morris, C. E. & Kinkel, L. L. (2002). Fifty years of phyllosphere microbiology: Significant contributions to research in related fields. In S. Lindow, I. Hecht-Poinar, & V. Elliott (Eds.), *Phyllosphere Microbiology.* (pp. 365–375). APS Press.

Morris, C. E., Kinkel, L. L., Xiao, K., Prior, P., & Sands, D. C. (2007). Surprising niche for the plant pathogen *Pseudomonas syringae*. *Infection, Genetics and Evolution*, 7(1), 84–92.

Morris, C. E., Monteil, C. L., & Berge, O. (2013). The life history of *Pseudomonas syringae*: linking agriculture to earth system processes. *Annual Review of Phytopathology*, 51, 6.1–6.20.

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(3), e00107.

Morris, C. E., Sands, D. C., Vinatzer, B. A., Glaux, C., Guilbaud, C., Buffière, A., Yan, S., Dominguez, H., & Thompson, B. M. (2008). The life history of the plant pathogen *Pseudomonas syringae* is linked to the water cycle. *The ISME Journal*, 2(3), 321–334.

Moura, M. L., Jacques, M. A., Brito, L. M., Mourao, I. M., & Duclos, J. (2004). Tomato pith necrosis (TPN) caused by *P. corrugata* and *P. mediterranea*: severity of damages and crop loss assessment. In *I International Symposium on Tomato Diseases 695* (pp. 365–372).

Müller, C. & Riederer, M. (2005). Plant surface properties in chemical ecology. *Journal of Chemical Ecology*, 31(11), 2621–2651.

Munkvold, K. R., Martin, M. E., Bronstein, P. A., & Collmer, A. (2008). A survey of the *Pseudomonas syringae* pv. tomato DC3000 type III secretion system effector repertoire reveals several effectors that are deleterious when expressed in *Saccharomyces cerevisiae*. *Molecular Plant-Microbe Interactions*, 21(4), 490–502.

Munkvold, K. R., Russell, A. B., Kvitko, B. H., & Collmer, A. (2009). *Pseudomonas syringae* pv. *tomato* DC3000 type III effector HopAA1-1 functions redundantly with chlorosis-promoting factor PSPTO4723 to produce bacterial speck lesions in host tomato. *Molecular Plant-Microbe Interactions*, 22(11), 1341–1355.

Nagahama, K., Yoshino, K., Matsuoka, M., Sato, M., Tanase, S., Ogawa, T., & Fukuda, H. (1994). Ethylene production by strains of the plant-pathogenic bacterium *Pseudomonas syringae* depends upon the presence of indigenous plasmids carrying homologous genes for the ethylene-forming enzyme. *Microbiology*, 140(9), 2309–2313.

Nakabachi, A., Yamashita, A., Toh, H., Ishikawa, H., Dunbar, H. E., Moran, N. A., & Hattori, M. (2006). The 160-kilobase genome of the bacterial endosymbiont *Carsonella*. *Science*, 314(5797), 267.

Nakahara, H., Mori, T., Sadakari, N., Matsusaki, H., & Matsuzoe, N. (2016). Selection of effective non-pathogenic *Ralstonia solanacearum* as biocontrol agents against bacterial wilt in eggplant. *Journal of Plant Diseases and Protection*, 123(3), 119–124.

Nei, M. (1987). *Molecular Evolutionary Genetics*. Columbia Univ. Press, New York.

Nelson, K. & Selander, R. K. (1994). Intergeneric transfer and recombination of the 6-phosphogluconate dehydrogenase gene (gnd) in enteric bacteria. *Proceedings of the National Academy of Sciences*, 91(21), 10227–10231.

Nelson, K. E., Clayton, R. a., Gill, S. R., Gwinn, M. L., Dodson, R. J., Haft, D. H., Hickey, E. K., Peterson, J. D., Nelson, W. C., Ketchum, K. a., McDonald, L., Utterback, T. R., Malek, J. a., Linher, K. D., Garrett, M. M., Stewart, a. M., Cotton, M. D., Pratt, M. S., Phillips, C. a., Richardson, D., Heidelberg, J., Sutton, G. G., Fleischmann, R. D., Eisen, J. a., White, O., Salzberg, S. L., Smith, H. O., Venter, J. C., & Fraser, C. M. (1999). Evidence for lateral gene transfer between Archaea and bacteria from genome sequence of *Thermotoga maritima*. *Nature*, 399(6734), 323–329.

Nishanth Kumar, S., Aravind, S. R., Jacob, J., Gopinath, G., Lankalapalli, R. S., Sreelekha, T. T., & Dileep Kumar, B. S. (2016). Pseudopyronine B: A potent antimicrobial and anticancer molecule isolated from a *Pseudomonas mosselii*. *Frontiers in Microbiology*, 7(1307), 1–15.

Nomura, K., DebRoy, S., Lee, Y. H., Pumplin, N., Jones, J., & He, S. Y. (2006). A Bacterial Virulence Protein Suppresses Host Innate Immunity to Cause Plant Disease. *Science*, 313(5784), 220–223.

Nowell, R. W., Laue, B. E., Sharp, P. M., & Green, S. (2016). Comparative genomics reveals genes significantly associated with woody hosts in the plant pathogen *Pseudomonas syringae*. *Molecular Plant Pathology*, 17(9), 1409–1424.

Nürnberger, T., Brunner, F., Kemmerling, B., Piater, L., Nurnberger, T., Brunner, F., Kemmerling, B., & Piater, L. (2004). Innate immunity in plants and animals: strinking similarities and obvious differences. *Immunological Reviews*, 198(1), 249–266.

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 Microbiology*, 12(1), 141.

O'Brien, H. E., Thakur, S., & Guttman, D. S. (2011). Evolution of plant pathogenesis in *Pseudomonas syringae*: a genomics perspective. *Annual Reviews of Phytopathology*, 49, 269–289.

Ochman, H., Lawrence, J. G., & Groisman, E. a. (2000). Lateral gene transfer and the nature of bacterial innovation. *Nature*, 405(6784), 299–304.

Ochman, H. & Selander, R. K. (1984). Evidence for clonal population structure in *Escherichia coli*. *Proceedings of the National Academy of Sciences*, 81(1), 198–201.

Ogawa, K., Mori, T., Matsusaki, H., & Matsuzoe, N. (2012). Effect of the Colonization of Non-pathogenic *Ralstonia solanacearum* on the Suppression of Bacterial Wilt in Eggplant. *Horticultural Research (Japan)*, 11(3), 399–403.

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., Dan McGlinn, Peter R. Minchin, R. B. O., Simpson, G. L., Solymos, P., Stevens, M. H. H., Szoecs, E., & Wagner, H. (2016). vegan: Community Ecology Package.

Omer, M. E. H. & Wood, R. K. S. (1969). Growth of *Pseudomonas phaseolicola* in susceptible and in resistant bean plants. *Annals of Applied Biology*, 63, 103–116.

Opgenorth, D. C., Lai, M., Sorrell, M., & White, J. B. (1983). *Pseudomonas* canker of kiwifruit. *Plant Disease*, 67(11), 1283–1284.

O'Rourke, M. & Stevens, E. (1993). Genetic structure of *Neisseria gonorrhoeae* populations: a non-clonal pathogen. *Journal of General Microbiology*, 139(11), 2603–26011.

Ortiz-Martin, I., Thwaites, R., Macho, A. P., Mansfield, J. W., & Beuzon, C. R. (2010). Positive regulation of the Hrp type III secretion system in *Pseudomonas syringae* pv. *phaseolicola*. *Molecular Plant-Microbe Interactions*, 23(5), 665–681.

Page, A. J., Cummins, C. A., Hunt, M., Wong, V. K., Reuter, S., Holden, M. T. G., Fookes, M., Falush, D., Keane, J. A., & Parkhill, J. (2015). Roary: Rapid large-scale prokaryote pan genome analysis. *Bioinformatics*, 31(22), 3691–3693.

Pallardy, S. G. (2008). *Physiology of woody plants*. Elsevier Ltd, third edition.

Pastar, I., Nusbaum, A. G., Gil, J., Patel, S. B., Chen, J., Valdes, J., Stojadinovic, O., Plano, L. R., Tomic-Canic, M., & Davis, S. C. (2013). Interactions of methicillin resistant *Staphylococcus aureus* USA300 and *Pseudomonas aeruginosa* in polymicrobial wound infection. *PLoS ONE*, 8(2), e56846.

Patil, P. B. & Sonti, R. V. (2004). Variation suggestive of horizontal gene transfer at a lipopolysaccharide (lps) biosynthetic locus in *Xanthomonas oryzae* pv. *oryzae*, the bacterial leaf blight pathogen of rice. *BMC Microbiology*, 4(1), 40.

Peet, R. C. & Panopoulos, N. J. (1987). Ornithine carbamoyltransferase genes and phaseolotoxin immunity in *Pseudomonas syringae* pv. *phaseolicola*. *The EMBO Journal*, 6(12), 3585–3591.

Pérez-García, A., Romero, D., & De Vicente, A. (2011). Plant protection and growth stimulation by microorganisms: biotechnological applications of Bacilli in agriculture. *Current Opinion in Biotechnology*, 22(2), 187–193.

Pérez-Losada, M., Cabezas, P., Castro-Nallar, E., & Crandall, K. A. (2013). Pathogen typing in the genomics era: MLST and the future of molecular epidemiology. *Infection, Genetics and Evolution*, 16, 38–53.

Pitman, A. R., Jackson, R. W., Mansfield, J. W., Kaitell, V., Thwaites, R., & Arnold, D. L. (2005). Exposure to host resistance mechanisms drives evolution of bacterial virulence in pants. *Current Biology*, 15(24), 2230–2235.

Polz, M. F., Alm, E. J., & Hanage, W. P. (2013). Horizontal gene transfer and the evolution of bacterial and archaeal population structure. *Trends in Genetics*, 29(3), 170–175.

Pond, S. L. K. & Frost, S. D. W. (2005). Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. *Bioinformatics*, 21(10), 2531–2533.

Popa, O. & Dagan, T. (2011). Trends and barriers to lateral gene transfer in prokaryotes. *Current Opinion in Microbiology*, 14(5), 615–623.

Posada, D. (2003). Selecting models of evolution. In A. M. Vandemme & M. Salemi (Eds.), *The Phylogenetic Handbook* (pp. 256–282). Cambridge University Press.

Preston, G. M. (2004). Plant perceptions of plant growth-promoting *Pseudomonas*. *Philosophical Transactions of the Royal Society*, 359, 907–918.

Preston, G. M., Bertrand, N., & Rainey, P. B. (2001). Type III secretion in plant growth-promoting *Pseudomonas fluorescens* SBW25. *Molecular Microbiology*, 41(5), 999–1014.

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155(2), 945–959.

Przywara, L., Pandey, K. K., & Sanders, P. M. (1988). Length of stomata as an indicator of ploidy level in *Actinidia deliciosa*. *New Zealand Journal of Botany*, 26(2), 179–182.

Quinlan, A. R. & Hall, I. M. (2010). BEDTools : a flexible suite of utilities for comparing genomic features. *Bioinformatics*, 26(6), 841–842.

Quiñones, B., Dulla, G., & Lindow, S. E. (1999). Quorum sensing regulates exopolysaccharide production, motility, and virulence in *Pseudomonas syringae*. *Molecular Plant-Microbe Interactions*, 18(7), 682–693.

Quiñones, B., Dulla, G., & Lindow, S. E. (2005). Exopolysaccharide production, motility, and virulence in *Pseudomonas syringae*. *Molecular Plant-Microbe Interactions*, 18(7), 682–693.

Quiroz, T. S., Nieto, P. A., Tobar, H. E., Salazar-Echegarai, F. J., Lizana, R. J., Quezada, C. P., Santiviago, C. A., Araya, D. V., Riedel, C. A., & Kalergis, A. M. (2011). Excision of an unstable pathogenicity island in *Salmonella enterica* serovar *enteritidis* is induced during infection of phagocytic cells. *PloS one*, 6(10), e26031.

Raaijmakers, J. M., de Bruijn, I., & de Kock, M. J. D. (2006). Cyclic lipopeptide production by plant-associated *Pseudomonas* spp.: diversity, activity, biosynthesis, and regulation. *Molecular Plant-Microbe Interactions*, 19(7), 699–710.

Raaijmakers, J. M. & Mazzola, M. (2011). Diversity and natural functions of antibiotics produced by beneficial and plant pathogenic bacteria. *Annual Review of Phytopathology*, 50, 403–424.

Rainey, P. B. (1999). Adaptation of *Pseudomonas fluorescens* to the plant rhizosphere. *Environmental Microbiology*, 1(3), 243–257.

Rainey, P. B. (2005). Bacteria evolve and function within communities : Observations from experimental *Pseudomonas* populations. In M. J. McFall-Ngai, B. Henderson, & E. G. Ruby (Eds.), *The Influence of Cooperative Bacteria on Animal Host Biology* (pp. 83–100). Cambridge University Press.

Rainey, P. B., Buckling, A., Kassen, R., & Travisano, M. (2000). The emergence and maintenance of diversity: Insights from experimental bacterial populations. *Trends in Ecology and Evolution*, 15(6), 243–247.

Rainey, P. B. & Rainey, K. (2003). Evolution of cooperation and conflict in experimental bacterial populations. *Nature*, 425(6593), 72–74.

Rainey, P. B. & Travisano, M. (1998). Adaptive radiation in a heterogeneous environment. *Nature*, 394(6688), 69–72.

Rampelotto, P. H. (2013). Extremophiles and extreme environments. *Life*, 3, 482–485.

Rasko, D. A., Rosovitz, M. J., Myers, G. S. A., Mongodin, E. F., Fricke, W. F., Gajer, P., Crabtree, J., Sebaihia, M., Thomson, N. R., Chaudhuri, R., Henderson, I. R., Sperandio, V., & Ravel, J. (2008). The pangenome structure of *Escherichia coli*: Comparative genomic analysis of *E. coli* commensal and

pathogenic isolates. *Journal of Bacteriology*, 190(20), 6881–6893.

Raupach, G. S. & Kloepper, J. W. (1998). Mixtures of Plant Growth-Promoting Rhizobacteria Enhance Biological Control of Multiple Cucumber Pathogens. *Phytopathology*, 88(11), 1158–1164.

Ravensdale, M., Nemri, A., Thrall, P. H., Ellis, J. G., & Dodds, P. N. (2011). Co-evolutionary interactions between host resistance and pathogen effector genes in flax rust disease. *Molecular Plant Pathology*, 12(1), 93–102.

R.Core.Team (2016). R: A language and environment for statistical computing.

Redford, A. J., Bowers, R. M., Knight, R., Linhart, Y., & Fierer, N. (2010). The ecology of the phyllosphere: Geographic and phylogenetic variability in the distribution of bacteria on tree leaves. *Environmental Microbiology*, 12(11), 2885–2893.

Reece, R. J. & Maxwell, A. (1991). DNA gyrase: structure and function. *Critical Reviews in Biochemistry and Molecular Biology*, 26(3/4), 335–375.

Remenant, B., Coupat-Goutaland, B., Guidot, A., Cellier, G., Wicker, E., Allen, C., Fegan, M., Pruvost, O., Elbaz, M., Calteau, A., Salvignol, G., Mornico, D., Mangenot, S., Barbe, V., Médigue, C., & Prior, P. (2010). Genomes of three tomato pathogens within the *Ralstonia solanacearum* species complex reveal significant evolutionary divergence. *BMC Genomics*, 11(379).

Remus-Emsermann, M. N. P., Lücker, S., Müller, D. B., Potthoff, E., Daims, H., & Vorholt, J. A. (2014). Spatial distribution analyses of natural phyllosphere-colonizing bacteria on *Arabidopsis thaliana* revealed by fluorescence *in situ* hybridization. *Environmental Microbiology*, 16(7), 2329–2340.

Rodríguez-Palenzuela, P., Matas, I. M., Murillo, J., López-Solanilla, E., Bardaji, L., Pérez-Martínez, I., Rodríguez-Moskera, M. E., Penyalver, R., López, M. M., Quesada, J. M., Biehl, B. S., Perna, N. T., Glasner, J. D., Cabot, E. L., Neeno-Eckwall, E., & Ramos, C. (2010). Annotation and overview of the *Pseudomonas savastanoi* pv. *savastanoi* NCPPB 3335 draft genome reveals the virulence gene complement of a tumour-inducing pathogen of woody hosts. *Environmental Microbiology*, 12(6), 1604–1620.

Romano, C. P., Hein, M. B., & Klee, H. J. (1991). Inactivation of auxin in tobacco transformed with the indoleacetic acid-lysine synthetase gene of *Pseudomonas savastanoi*. *Genes and Development*, 5(3), 438–446.

Rosenberg, E. & Zilber-Rosenberg, I. (2013). Role of microorganisms in adaptation, development, and evolution of animals and plants: The hologenome concept. In E. Rosenberg, E. DeLong, E. Stackebrandt, S. Lory, & F. Thompson (Eds.), *The Prokaryotes: Prokaryotic Biology and Symbiotic Associations* (pp. 347–358). Berlin Heidelberg: Springer.

Rosenberg, E. & Zilber-Rosenberg, I. (2016). Microbes drive evolution of animals and plants: The hologenome concept. *mBio*, 7(2), e01395–15.

Rosenberg, N. A. (2004). DISTRUCT: A program for the graphical display of population structure. *Molecular Ecology Notes*, 4(1), 137–138.

Rozas, J. & Rozas, R. (1995). DnaSP, DNA sequence polymorphism: an interactive program for estimating population genetics parameters from DNA sequence data. *CABIOS*, 11(6), 621–625.

Rufián, J. S., Macho, A. P., Corry, D. S., Mansfield, J., Arnold, D., & Beuzón, C. R. (2017). Confocal microscopy reveals *in planta* dynamic interactions between pathogenic, avirulent and non-pathogenic *Pseudomonas syringae* strains. *Molecular Plant Pathology*, (pp. DOI: 10.1111/mpp.12539).

Ruhe, Z. C., Low, D. A., & Hayes, C. S. (2013). Bacterial contact-dependent growth inhibition. *Trends in Microbiology*, 21(5), 230–237.

Ruinen, J. (1956). Occurrence of *Beijerinckia* species in the phyllosphere. *Nature*, 177(4501), 220–221.

Saitou, N. & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, 4(4), 406–425.

Sarkar, S. F., Gordon, J. S., Martin, G. B., & Guttman, D. S. (2006). Comparative genomics of host-specific virulence in *Pseudomonas syringae*. *Genetics*, 174(2), 1041–1056.

Sarkar, S. F. & Guttman, D. S. (2004). Evolution of the core genome of *Pseudomonas syringae*, a highly clonal, endemic plant pathogen. *Applied and Environmental Microbiology*, 70(4), 1999–2012.

Sawada, H. (2015). Characterization of biovar 3 strains of *Pseudomonas syringae* pv. *actinidiae* isolated in Japan. *Japanese Journal of Phyopathology*, 81, 111–126.

Schellenberg, B., Ramel, C., & Dudler, R. (2010). *Pseudomonas syringae* virulence factor syringolin A counteracts stomatal immunity by proteasome inhibition. *Molecular Plant-Microbe Interactions*, 23(10), 1287–1293.

Schippers, B., Bakker, A. W., & Bakker, P. A. H. M. (1987). Interactions of deleterious and beneficial rhizosphere microorganisms and the effect of cropping practices. *Annual Review of Phytopathology*, 25(1), 339–358.

Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., Lesniewski, R. A., Oakley, B. B., Parks, D. H., Robinson, C. J., Sahl, J. W., Stres, B., Thallinger, G. G., Van Horn, D. J., & Weber, C. F. (2009). Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology*, 75(23), 7537–7541.

Schmidt, H. A., Strimmer, K., Vingron, M., & von Haeseler, A. (2002). TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. *Bioinformatics*, 18(3), 502–504.

Scholz-Schroeder, B. K., Hutchison, M. L., Grgurina, I., & Gross, D. C. (2001). The contribution of syringopeptin and syringomycin to virulence of *Pseudomonas syringae* pv . syringae strain B301D on the basis of *sypA* and *syrB1* biosynthesis mutant analysis. *Molecular Plant-Microbe Interactions*, 14(3), 336–348.

Schreiber, L., Krimm, U., Knoll, D., Sayed, M., Auling, G., & Kroppenstedt, R. M. (2005). Plant-microbe interactions: Identification of epiphytic bacteria and their ability to alter leaf surface permeability. *New Phytologist*, 166(2), 589–594.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6(2), 461–464.

Scortichini, M. (1994). Occurrence of *Pseudomonas syringae* pv. *actinidiae* on kiwifruit in Italy. *Plant Pathology*, 43(6), 1035–1038.

Scortichini, M., Marcelletti, S., Ferrante, P., & Firrao, G. (2013). A henomic redefinition of *Pseudomonas avellanae* species. *PLoS ONE*, 8(9), e75794.

Seemann, T. (2014). Prokka: Rapid prokaryotic genome annotation. *Bioinformatics*, 30(14), 2068–2069.

Selander, R. K., Beltran, P., Smith, N. H., Barker, R. M., Crichton, P. B., Old, D. C., Musser, J. M., & Whittam, T. S. (1990). Genetic population structure, clonal phylogeny, and pathogenicity of *Salmonella* parathyphi B. *Infection and Immunity*, 58(6), 1891–1901.

Selander, R. K. & Levin, B. R. (1980). Genetic diversity and structure in *Escherichia coli* populations. *Science*, 210(4469), 545–7.

Serrano, M., Coluccia, F., Torres, M., L'Haridon, F., & Métraux, J.-P. (2014). The cuticle and plant defense to pathogens. *Frontiers in Plant Science*, 5(274), 1–8.

Shapiro, B. J. & Polz, M. F. (2014). Ordering microbial diversity into ecologically and genetically cohesive units. *Trends in Microbiology*, 22(5), 235–247.

Shen, X.-H., Huang, Y., & Liu, S.-J. (2005). Genomic analysis and identification of catabolic pathways for aromatic compounds in *Corynebacterium glutamicum*. *Microbes and Environments*, 20(3), 160–167.

Sheppard, S. K., Mccarthy, N. D., Falush, D., & Maiden, M. C. J. (2008). Convergence of *Campylobacter* species: implications for bacterial evolution. *Science*, 5873(237-239).

Shimodaira, H. & Hasegawa, M. (1999). Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Molecular Biology and Evolution*, 16(8), 1114–1116.

Silby, M. W., Winstanley, C., Godfrey, S. A. C., Levy, S. B., & Jackson, R. W. (2011). *Pseudomonas* genomes: diverse and adaptable. *FEMS Microbiology Reviews*, 35, 652–680.

Singer, M. (2010). Pathogen-pathogen interaction. *Virulence*, 1(1), 10–18.

Smillie, C. S., Smith, M. B., Friedman, J., Cordero, O. X., David, L. A., & Alm, E. J. (2011). Ecology drives a global network of gene exchange connecting the human microbiome. *Nature*, 480(7376), 241–244.

Sorensen, T. (1947). A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on Danish commons. *Kongelige Danske Videnskabernes Selskab*, 5(4), 1–34.

Spratt, B. G., Hanage, W. P., Li, B., Aanensen, D. M., & Feil, E. J. (2004). Displaying the relatedness among isolates of bacterial species – the eBURST approach. *FEMS Microbiology Letters*, 241(2), 129–134.

Spratt, B. G. & Maiden, M. C. J. (1999). Bacterial population genetics, evolution and epidemiology. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 354(1384), 701–710.

Stacy, A. R., Diggle, S. P., & Whiteley, M. (2012). Rules of engagement: Defining bacterial communication. *Current Opinion in Microbiology*, 15, 155–161.

Stamatakis, A. (2006). RAxML-VI-HPC : maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*, 22(21), 2688–2690.

Staskawicz, B. J. & Hoogenraad, N. J. (1980). Phaseolotoxin-insensitive ornithine carbamoyltransferase of *Pseudomonas syringae* pv. *phaseolicola*: basis for immunity to phaseolotoxin. *Journal of Bacteriology*,

142(2), 720–723.

Stavrinides, J., McCloskey, J. K., & Ochman, H. (2009). Pea aphid as both host and vector for the phytopathogenic bacterium *Pseudomonas syringae*. *Applied and Environmental Microbiology*, 75(7), 2230–2235.

Stevens, C., Bennett, M. A., Athanassopoulos, E., Tsiamis, G., Taylor, J. D., & Mansfield, J. W. (1998). Sequence variations in alleles of the avirulence gene *avrPphE*. R2 from *Pseudomonas syringae* pv. *phaseolicola* lead to loss of recognition of the AvrPphE protein within bean cells and a gain in cultivar-specific virulence. *Molecular Microbiology*, 29(1), 165–177.

Strange, R. N. & Scott, P. R. (2005). Plant disease: a threat to global food security. *Phytopathology*, 43, 83–116.

Stubbendieck, R. M., Vargas-Bautista, C., & Straight, P. D. (2016). Bacterial communities: Interactions to scale. *Frontiers in Microbiology*, 7, 1234.

Stukenbrock, E. H. & McDonald, B. A. (2008). The origins of plant pathogens in agro-ecosystems. *Annual Review of Phytopathology*, 46, 75–100.

Suda, W., Nagasaki, A., & Shishido, M. (2009). Powdery mildew-infection changes bacterial community composition in the phyllosphere. *Microbes and Environments*, 24(3), 217–223.

Sundin, G. W. & Bender, C. L. (1993). Ecological and genetic analysis of copper and streptomycin resistance in *Pseudomonas syringae* pv. *syringae*. *Applied and Environmental Microbiology*, 59(4), 1018–1024.

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, 123(3), 585–595.

Takikawa, Y., Serizawa, S., Ichikawa, T., Tsuyumu, S., & Goto, M. (1989). *Pseudomonas syringae* pv. *actinidiae* pv. *nov.*: the causal bacterium of canker of kiwifruit in Japan. *Annals of the Phytopathological Society of Japan*, 55(4), 437–444.

Tao, Y., Fishman, A., Bentley, W. E., & Wood, T. K. (2004). Oxidation of benzene to phenol, vatechol, and 1,2,3-trihydroxybenzene by toluene 4-monooxygenase of *Pseudomonas mendocina* KR1 and toluene 3-monooxygenase of *Ralstonia pickettii* PKO1. *Applied and Environmental Microbiology*, 70(7), 3814–3820.

Templeton, M. D., Warren, B. A., Andersen, M. T., Rikkerink, E. H. A., & Fineran, P. C. (2015). Complete DNA sequence of *Pseudomonas syringae* pv. *actinidiae*, the causal agent of kiwifruit canker disease. *Genome Announcements*, 3(5), e01054–15.

Tenaillon, O., Skurnik, D., Picard, B., & Denamur, E. (2010). The population genetics of commensal *Escherichia coli*. *Nature Reviews Microbiology*, 8(3), 207–217.

Thakur, S., Weir, B. S., & Guttman, D. (2016). Phytopathogen genome announcement: draft genome sequences of 62 *Pseudomonas syringae* type and pathotype strains. *Molecular Plant-Microbe Interactions*, 29(4), 243–246.

Thomas, C. M. & Nielsen, K. M. (2005). Mechanisms of, and barriers to, horizontal gene transfer between bacteria. *Nature Reviews Microbiology*, 3(9), 711–721.

Thomas, S. H., Wagner, R. D., Arakaki, A. K., Skolnick, J., Kirby, J. R., Shimkets, L. J., Sanford, R. A., & Löffler, F. E. (2008). The mosaic genome of *Anaeromyxobacter dehalogenans* strain 2CP-C suggests an aerobic common ancestor to the delta-proteobacteria. *PLoS ONE*, 3(5), e2103.

Thomma, B. P. H. J., Nürnberger, T., & Joosten, M. H. A. J. (2011). Of PAMPs and effectors: the blurred PTI-ETI dichotomy. *The Plant Cell*, 23(1), 4–15.

Thompson, I. P., Bailey, M. J., Fenlon, J. S., Fermor, T. R., Lilley, A. K., Lynch, J. M., McCormack, P. J., McQuilken, M. P., Purdy, K. J., Rainey, P. B., & Whipps, J. M. (1993). Quantitative and qualitative seasonal changes in the mobial community from the phyllosphere of sugar beet (*Beta vulgaris*). *Plant and Soil*, 150, 177–191.

Tibayrenc, M. & Ayala, F. J. (2012). Reproductive clonality of pathogens: a perspective on pathogenic viruses, bacteria, fungi, and parasitic protozoa. *Proceedings of the National Academy of Sciences*, 109(48), 3305–3313.

Toft, C. & Andersson, S. G. E. (2010). Evolutionary microbial genomics: insights into bacterial host adaptation. *Nature Reviews Genetics*, 11(7), 465–475.

Tomihama, T., Sekita, T., & Takikawa, Y. (2016). Wide distribution of airborne ice-nucleation active *Pseudomonas syringae* in agricultural environments. *Phytopathology*, 5159, 1–49.

Tong, H., Chen, W., Merritt, J., Qi, F., Shi, W., & Dong, X. (2007). *Streptococcus oligofermentans* inhibits *Streptococcus mutans* through conversion of lactic acid into inhibitory H2O2: A possible counteroffensive strategy for interspecies competition. *Molecular Microbiology*, 63(3), 872–880.

Tsuda, K. & Katagiri, F. (2010). Comparing signaling mechanisms engaged in pattern-triggered and effector-triggered immunity. *Current Opinion in Plant Biology*, 13(4), 459–465.

Turner, S., Pryer, K. M., Miao, V. P., & Palmer, J. D. (1999). Investigating deep phylogenetic relationships among cyanobacteria and plastids by small subunit rRNA sequence analysis. *The Journal of Eukaryotic Microbiology*, 46(4), 327–338.

Vacher, C., Hampe, A., Porté, A., Sauer, U., Compant, S., & Morris, C. E. (2016). The phyllosphere: microbial jungle at the plant-climate interface. *Annual Review of Ecology, Evolution, and Systematics*, 47, 1–24.

Van Overbeek, L. & Van Elsas, J. D. (2008). Effects of plant genotype and growth stage on the structure of bacterial communities associated with potato (*Solanum tuberosum* L.). *FEMS Microbiology Ecology*, 64(2), 283–296.

Vanga, B. R., Ramakrishnan, P., Butler, R. C., Toth, I. K., Ronson, C. W., Jacobs, J. M. E., & Pitman, A. R. (2015). Mobilization of horizontally acquired island 2 is induced *in planta* in the phytopathogen *Pectobacterium atrosepticum* SCRI1043 and involves the putative relaxase ECA0613 and quorum sensing. *Environmental Microbiology*, 17(11), 4730–4744.

Vanneste, J. L., Poliakoff, F., Audusseau, C., Cornish, D. A., Paillard, S., Rivoal, C., & Yu, J. (2011). First report of *Pseudomonas syringae* pv. *actinidiae*, the causal agent of bacterial canker of kiwifruit in France. *Plant Disease*, 95(10), 1311.

Vinatzer, B. A., Monteil, C. L., & Clarke, C. R. (2014). Harnessing population genomics to understand how

bacterial pathogens emerge, adapt to crop hosts, and disseminate. *Annual Review of Phytopathology*, 52(1), 19–43.

Visnovsky, S. B., Fiers, M., Lu, A., Panda, P., Taylor, R., & Pitman, A. R. (2016). Draft genome sequences of 18 strains of *Pseudomonas* isolated from kiwifruit plants in New Zealand and overseas. *Genome Announcements*, 4(2), e00061–16.

Völksch, B. & May, R. (2001). Biological control of *Pseudomonas syringae* pv. *glycinea* by epiphytic bacteria under field conditions. *Microbial Ecology*, 41(2), 132–139.

Vorholt, J. A. (2012). Microbial life in the phyllosphere. *Nature Reviews Microbiology*, 10(12), 828–840.

Wagner, M. R., Lundberg, D. S., del Rio, T. G., Tringe, S. G., Dangl, J. L., & Mitchell-Olds, T. (2016). Host genotype and age shape the leaf and root microbiomes of a wild perennial plant. *Nature Communications*, 7, 12151.

Walker, B. J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., Cuomo, C. A., Zeng, Q., Wortman, J., Young, S. K., & Earl, A. M. (2014). Pilon: An integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS ONE*, 9(11), e112963.

Waples, R. S. & Gaggiotti, O. (2006). What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Molecular Ecology*, 15(6), 1419–1439.

Watterson, G. A. (1975). On the number of segregating sites in genetical models without recombination. *Theoretical Population Biology*, 7(2), 256–276.

Weber, M. G. & Strauss, S. Y. (2016). Coexistence in close relatives: beyond competition and reproductive isolation in sister taxa. *Annual Review of Ecology, Evolution, and Systematics*, 47, 359–381.

Weber, T., Blin, K., Duddela, S., Krug, D., Kim, H. U., Bruccoleri, R., Lee, S. Y., Fischbach, M. A., Müller, R., Wohlleben, W., Breitling, R., Takano, E., & Medema, M. H. (2015). AntiSMASH 3.0-A comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic Acids Research*, 43, W237–W243.

Wei, C.-F., Kvitko, B. H., Shimizu, R., Crabill, E., Alfano, J. R., Lin, N.-C., Martin, G. B., Huang, H.-C., & Collmer, A. (2007). A *Pseudomonas syringae* pv. *tomato* DC3000 mutant lacking the type III effector *HopQ1-1* is able to cause disease in the model plant *Nicotiana benthamiana*. *The Plant Journal*, 51(1), 32–46.

Weingart, H., Ullrich, H., Geider, K., & Völksch, B. (2001). The role of ethylene production in virulence of *Pseudomonas syringae* pvs. *glycinea* and *phaseolicola*. *Phytopathology*, 91(5), 511–518.

Weisburg, W. G., Barns, S. M., Pelletier, D. A., Lane, D. J., Pelletie, D. a., & Lane, D. J. (1991). 16S ribosomal DNA amplification for phylogenetic study. *Journal of Bacteriology*, 173(2), 697–703.

Weller, D. M. (1988). Biological control of soilborne plant pathogens in the rhizospehere with bacteria. *Annual Review of Phytopathology*, 26(1), 379–407.

Whipps, J. M., Hand, P., Pink, D., & Bending, G. D. (2008). Phyllosphere microbiology with special reference to diversity and plant genotype. *Journal of Applied Microbiology*, 105(6), 1744–1755.

Whitehead, N. A., Barnard, A. M. L., Slater, H., Simpson, N. J. L., & Salmond, G. P. C. (2001). Quorum-

sensing in Gram-negative bacteria. *FEMS Microbiology Reviews*, 25, 365–404.

Wilson, M., Hirano, S. S., & Lindow, S. E. (1999). Location and survival of leaf-associated bacteria in relation to pathogenicity and potential for growth within the leaf. *Applied and Environmental Microbiology*, 65(4), 1435–1443.

Wilson, M. & Lindow, S. E. (1994a). Coexistence among epiphytic bacterial Populations mediated through nutritional resource partitioning. *Applied and Environmental Microbiology*, 60(12), 4468–4477.

Wilson, M. & Lindow, S. E. (1994b). Ecological similarity and coexistence of epiphytic Ice-nucleating (Ice+) *Pseudomonas syringae* strains and a non-Ice-nucleating (Ice-) biological control agent. *Applied and Environmental Microbiology*, 60(9), 3128–3137.

Woolhouse, M. E. J., Haydon, D. T., & Antia, R. (2005). Emerging pathogens: The epidemiology and evolution of species jumps. *Trends in Ecology and Evolution*, 20(5), 238–44.

Xiao, Y., Heu, S., Yi, J., Lu, Y., & Hutcheson, S. W. (1994). Identification of a putative alternate sigma factor and characterization of a multicomponent regulatory cascade controlling the expression of *Pseudomonas syringae* pv. *syringae* Pss61 *hrp* and *hrmA* genes. *Journal of Bacteriology*, 176(4), 1025–1036.

Yahara, K., Didelot, X., Jolley, K. A., Kobayashi, I., Maiden, M. C. J., Sheppard, S. K., & Falush, D. (2016). The landscape of realized homologous recombination in pathogenic bacteria. *Molecular Biology and Evolution*, 33(2), 456–471.

Yang, Z. & Rannala, B. (1997). Bayesian phylogenetic inference using DNA sequences: a Markov Chain Monte Carlo Method. *Molecular Biology and Evolution*, 14(7), 717–724.

Young, J. M. (1974). Development of bacterial populations *in vivo* in relation to plant pathogenicity. *New Zealand Journal of Agricultural Research*, 17(1), 105–113.

Young, J. M. (2010). Taxonomy of *Pseudomonas syringae*. *Journal of Plant Pathology*, 92(1, Supplement), S1.5–S1.14.

Zhan, Y., Yu, H., Yan, Y., Chen, M., Lu, W., Li, S., Peng, Z., Zhang, W., Ping, S., Wang, J., & Lin, M. (2008). Genes involved in the benzoate catabolic pathway in *Acinetobacter calcoaceticus* PHEA-2. *Current Microbiology*, 57(6), 609–614.

Zhang, B., Bai, Z., Hoefel, D., Tang, L., Yang, Z., Zhuang, G., Yang, J., & Zhang, H. (2008). Assessing the impact of the biological control agent *Bacillus thuringiensis* on the indigenous microbial community within the pepper plant phyllosphere. *FEMS Microbiology Letters*, 284(1), 102–108.

Zhang, X. X., Gauntlett, J. C., Oldenburg, D. G., Cook, G. M., & Rainey, P. B. (2015). Role of the transporter-like sensor kinase CbrA in histidine uptake and signal transduction. *Journal of Bacteriology*, 197(17), 2867–2878.

Zhao, Z. B., Gao, X. N., Huang, Q. L., Huang, L. L., Qin, H. Q., & Kang, Z. S. (2013). Identification and characerization of the causal agent of bacterial canker of kiwifruit in the Shaanxi province of China. *Journal of Plant Pathology*, 95(1), 155–162.

Zheng, X. Y., Spivey, N. W., Zeng, W., Liu, P. P., Fu, Z. Q., Klessig, D. F., He, S. Y., & Dong, X. (2012). Coronatine promotes *Pseudomonas syringae* virulence in plants by activating a signaling cascade that

inhibits salicylic acid accumulation. *Cell Host and Microbe*, 11(6), 587–596.

Zhou, J., Liu, Y., & Huang, H. (2011). Characterization of 15 novel single nucleotide polymorphisms (SNPs) in the *Actinidia chinensis* species complex (*Actinidiaceae*). *American Journal of Botany*, 98(5), e100–e102.

Zipfel, C. (2014). Plant pattern-recognition receptors. *Trends in Immunology*, 35(7), 345–351.

Appendix

Table S 2.1: List of strains used in the MLST study (Chapter 2).

Strain information includes unique ID, ST = sequence type, source, location, host (gold kiwifruit = *A. chinesis* var. *chinensis* Hort16A, green kiwifruit = *A. chinesis* var. *deliciosa* Hayward), year = year of isolation, PG = phylogroups. All strains listed here were used for the MLST analysis based on STs, the strains included in the phylogenetic analysis are highlighted in grey.

	ID	ST	Source	Location	Host	Year	PG
1	G13A-I	904	this study	Pukekohe, NZ	Gold kiwifruit	2013	1B
2	G13A-III	904	this study	Pukekohe, NZ	Gold kiwifruit	2013	1B
3	G13C-II	904	this study	Pukekohe, NZ	Gold kiwifruit	2013	1B
4	G14C-X	19	this study	Pukekohe, NZ	Gold kiwifruit	2013	2
5	G14C-XI	19	this study	Pukekohe, NZ	Gold kiwifruit	2013	2
6	G15B-I	904	this study	Pukekohe, NZ	Gold kiwifruit	2013	1B
7	G15B-II	904	this study	Pukekohe, NZ	Gold kiwifruit	2013	1B
8	G15C-I	904	this study	Pukekohe, NZ	Gold kiwifruit	2013	1B
9	G15C-II	904	this study	Pukekohe, NZ	Gold kiwifruit	2013	1B
10	G16B-I	15	this study	Pukekohe, NZ	Gold kiwifruit	2013	1A
11	G16B-II	15	this study	Pukekohe, NZ	Gold kiwifruit	2013	1A
12	G16C-I	33	this study	Pukekohe, NZ	Gold kiwifruit	2013	5
13	G16C-II	30	this study	Pukekohe, NZ	Gold kiwifruit	2013	2B
14	G21A-I	904	this study	Pukekohe, NZ	Gold kiwifruit	2014	1B
15	G21A-II	904	this study	Pukekohe, NZ	Gold kiwifruit	2014	1B
16	G21C-II	26	this study	Pukekohe, NZ	Gold kiwifruit	2014	2A
17	G22C-I	20	this study	Pukekohe, NZ	Gold kiwifruit	2014	3A
18	G22C-II	20	this study	Pukekohe, NZ	Gold kiwifruit	2014	3A
19	G23B-I	1	this study	Pukekohe, NZ	Gold kiwifruit	2014	3A
20	G23B-II	1	this study	Pukekohe, NZ	Gold kiwifruit	2014	3A
21	G23C-I	904	this study	Pukekohe, NZ	Gold kiwifruit	2014	1B
22	G23C-II	904	this study	Pukekohe, NZ	Gold kiwifruit	2014	1B
23	G24A-I	904	this study	Pukekohe, NZ	Gold kiwifruit	2014	1B
24	G24A-II	904	this study	Pukekohe, NZ	Gold kiwifruit	2014	1B
25	G24B-I	17	this study	Pukekohe, NZ	Gold kiwifruit	2014	2A
26	G24B-II	17	this study	Pukekohe, NZ	Gold kiwifruit	2014	2A
27	G25A-I	35	this study	Pukekohe, NZ	Gold kiwifruit	2014	2
28	G25A-III	904	this study	Pukekohe, NZ	Gold kiwifruit	2014	1B
29	G31A-I	36	this study	Pukekohe, NZ	Gold kiwifruit	2014	2
30	G31A-II	7	this study	Pukekohe, NZ	Gold kiwifruit	2014	2
31	G31C-II	12	this study	Pukekohe, NZ	Gold kiwifruit	2014	2
32	G32A-I	1	this study	Pukekohe, NZ	Gold kiwifruit	2014	3A

	ID	ST	Source	Location	Host	Year	PG
33	G32B-I	22	this study	Pukekohe, NZ	Gold kiwifruit	2014	2
34	G32B-II	22	this study	Pukekohe, NZ	Gold kiwifruit	2014	2
35	G33A-I	1	this study	Pukekohe, NZ	Gold kiwifruit	2014	3A
36	G33A-II	1	this study	Pukekohe, NZ	Gold kiwifruit	2014	3A
37	G33C-I	1	this study	Pukekohe, NZ	Gold kiwifruit	2014	3A
38	G33C-II	904	this study	Pukekohe, NZ	Gold kiwifruit	2014	1B
39	G34C-II	5	this study	Pukekohe, NZ	Gold kiwifruit	2014	2A
40	G35B-I	16	this study	Pukekohe, NZ	Gold kiwifruit	2014	2A
41	G35B-II	16	this study	Pukekohe, NZ	Gold kiwifruit	2014	2A
42	G36A-I	1	this study	Pukekohe, NZ	Gold kiwifruit	2014	3A
43	G36A-II	1	this study	Pukekohe, NZ	Gold kiwifruit	2014	3A
44	G36B-I	37	this study	Pukekohe, NZ	Gold kiwifruit	2014	1A
45	H12B-I	13	this study	Pukekohe, NZ	Green kiwifruit	2013	3A
46	H12B-II	13	this study	Pukekohe, NZ	Green kiwifruit	2013	3A
47	H12B-III	13	this study	Pukekohe, NZ	Green kiwifruit	2013	3A
48	H13B-I	7	this study	Pukekohe, NZ	Green kiwifruit	2013	2
49	H13B-II	7	this study	Pukekohe, NZ	Green kiwifruit	2013	2
50	H13B-III	7	this study	Pukekohe, NZ	Green kiwifruit	2013	2
51	H16A-I	5	this study	Pukekohe, NZ	Green kiwifruit	2013	2A
52	H16A-II	5	this study	Pukekohe, NZ	Green kiwifruit	2013	2A
53	H16B-I	18	this study	Pukekohe, NZ	Green kiwifruit	2013	5
54	H16B-II	18	this study	Pukekohe, NZ	Green kiwifruit	2013	5
55	H22A-I	5	this study	Pukekohe, NZ	Green kiwifruit	2014	2A
56	H22A-II	5	this study	Pukekohe, NZ	Green kiwifruit	2014	2A
57	H23B-I	39	this study	Pukekohe, NZ	Green kiwifruit	2014	3A
58	H24B-I	23	this study	Pukekohe, NZ	Green kiwifruit	2014	3A
59	H24B-II	23	this study	Pukekohe, NZ	Green kiwifruit	2014	3A
60	H24C-I	43	this study	Pukekohe, NZ	Green kiwifruit	2014	3A
61	H25A-I	5	this study	Pukekohe, NZ	Green kiwifruit	2014	2A
62	H25A-II	12	this study	Pukekohe, NZ	Green kiwifruit	2014	2
63	H25B-I	27	this study	Pukekohe, NZ	Green kiwifruit	2014	2A
64	H25B-II	28	this study	Pukekohe, NZ	Green kiwifruit	2014	2A
65	H25C-II	34	this study	Pukekohe, NZ	Green kiwifruit	2014	3A
66	H31A-I	3	this study	Pukekohe, NZ	Green kiwifruit	2014	3A
67	H31B-I	24	this study	Pukekohe, NZ	Green kiwifruit	2014	2
68	H31B-II	24	this study	Pukekohe, NZ	Green kiwifruit	2014	2
69	H32A-II	7	this study	Pukekohe, NZ	Green kiwifruit	2014	2

	ID	ST	Source	Location	Host	Year	PG
70	H33A-I	8	this study	Pukekohe, NZ	Green kiwifruit	2014	3A
71	H33B-I	8	this study	Pukekohe, NZ	Green kiwifruit	2014	3A
72	H33B-II	8	this study	Pukekohe, NZ	Green kiwifruit	2014	3A
73	H33C-I	8	this study	Pukekohe, NZ	Green kiwifruit	2014	3A
74	H33C-II	8	this study	Pukekohe, NZ	Green kiwifruit	2014	3A
75	H34A-I	14	this study	Pukekohe, NZ	Green kiwifruit	2014	3A
76	H34B-II	5	this study	Pukekohe, NZ	Green kiwifruit	2014	2A
77	H35A-II	5	this study	Pukekohe, NZ	Green kiwifruit	2014	2A
78	H35B-I	25	this study	Pukekohe, NZ	Green kiwifruit	2014	1A
79	H36C-I	31	this study	Pukekohe, NZ	Green kiwifruit	2014	2B
80	H36C-II	12	this study	Pukekohe, NZ	Green kiwifruit	2014	2
81	K11A-II	10	this study	Kumeu, NZ	Gold kiwifruit	2013	2B
82	K11A-III	10	this study	Kumeu, NZ	Gold kiwifruit	2013	2B
83	K14C-I	6	this study	Kumeu, NZ	Gold kiwifruit	2013	3A
84	K14C-II	6	this study	Kumeu, NZ	Gold kiwifruit	2013	3A
85	K14C-IV	6	this study	Kumeu, NZ	Gold kiwifruit	2013	3A
86	K16A-II	9	this study	Kumeu, NZ	Gold kiwifruit	2013	2B
87	K16B-III	9	this study	Kumeu, NZ	Gold kiwifruit	2013	2B
88	K16C-III	9	this study	Kumeu, NZ	Gold kiwifruit	2013	2B
89	K16C-IV	9	this study	Kumeu, NZ	Gold kiwifruit	2013	2B
90	K22A-I	3	this study	Kumeu, NZ	Gold kiwifruit	2014	3A
91	K24C-I	10	this study	Kumeu, NZ	Gold kiwifruit	2014	2B
92	K25A-I	1	this study	Kumeu, NZ	Gold kiwifruit	2014	3A
93	K25A-II	1	this study	Kumeu, NZ	Gold kiwifruit	2014	3A
94	K25A-III	1	this study	Kumeu, NZ	Gold kiwifruit	2014	3A
95	K25A-IV	1	this study	Kumeu, NZ	Gold kiwifruit	2014	3A
96	K25C-II	1	this study	Kumeu, NZ	Gold kiwifruit	2014	3A
97	K31A-I	10	this study	Kumeu, NZ	Gold kiwifruit	2014	2B
98	K31B-II	10	this study	Kumeu, NZ	Gold kiwifruit	2014	2B
99	K34C-II	6	this study	Kumeu, NZ	Gold kiwifruit	2014	3A
100	K35C-I	11	this study	Kumeu, NZ	Gold kiwifruit	2014	3A
101	K35C-IV	11	this study	Kumeu, NZ	Gold kiwifruit	2014	3A
102	K36A-I	1	this study	Kumeu, NZ	Gold kiwifruit	2014	3A
103	K36A-II	1	this study	Kumeu, NZ	Gold kiwifruit	2014	3A
104	K36A-III	1	this study	Kumeu, NZ	Gold kiwifruit	2014	3A
105	K36B-I	1	this study	Kumeu, NZ	Gold kiwifruit	2014	3A
106	K36B-II	1	this study	Kumeu, NZ	Gold kiwifruit	2014	3A

	ID	ST	Source	Location	Host	Year	PG
107	K36B-III	1	this study	Kumeu, NZ	Gold kiwifruit	2014	3A
108	K36B-IV	1	this study	Kumeu, NZ	Gold kiwifruit	2014	3A
109	R11B-III	4	this study	Kumeu, NZ	Green kiwifruit	2013	3A
110	R12B-I	3	this study	Kumeu, NZ	Green kiwifruit	2013	3A
111	R12B-II	3	this study	Kumeu, NZ	Green kiwifruit	2013	3A
112	R13A-I	3	this study	Kumeu, NZ	Green kiwifruit	2013	3A
113	R13A-II	3	this study	Kumeu, NZ	Green kiwifruit	2013	3A
114	R14B-I	4	this study	Kumeu, NZ	Green kiwifruit	2013	3A
115	R14B-II	4	this study	Kumeu, NZ	Green kiwifruit	2013	3A
116	R14B-III	4	this study	Kumeu, NZ	Green kiwifruit	2013	3A
117	R14C-I	4	this study	Kumeu, NZ	Green kiwifruit	2013	3A
118	R14C-II	4	this study	Kumeu, NZ	Green kiwifruit	2013	3A
119	R15B-I	45	this study	Kumeu, NZ	Green kiwifruit	2013	3A
120	R15B-II	14	this study	Kumeu, NZ	Green kiwifruit	2013	3A
121	R15C-I	11	this study	Kumeu, NZ	Green kiwifruit	2013	3A
122	R15C-II	11	this study	Kumeu, NZ	Green kiwifruit	2013	3A
123	R16C-I	40	this study	Kumeu, NZ	Green kiwifruit	2013	3A
124	R16C-II	6	this study	Kumeu, NZ	Green kiwifruit	2013	3A
125	R22A-I	1	this study	Kumeu, NZ	Green kiwifruit	2014	3A
126	R22A-II	1	this study	Kumeu, NZ	Green kiwifruit	2014	3A
127	R22B-I	3	this study	Kumeu, NZ	Green kiwifruit	2014	3A
128	R23A-I	3	this study	Kumeu, NZ	Green kiwifruit	2014	3A
129	R23A-II	3	this study	Kumeu, NZ	Green kiwifruit	2014	3A
130	R23C-II	6	this study	Kumeu, NZ	Green kiwifruit	2014	3A
131	R24A-II	1	this study	Kumeu, NZ	Green kiwifruit	2014	3A
132	R24B-I	3	this study	Kumeu, NZ	Green kiwifruit	2014	3A
133	R24B-II	3	this study	Kumeu, NZ	Green kiwifruit	2014	3A
134	R24C-I	4	this study	Kumeu, NZ	Green kiwifruit	2014	3A
135	R25B-II	44	this study	Kumeu, NZ	Green kiwifruit	2014	3A
136	R26A-I	21	this study	Kumeu, NZ	Green kiwifruit	2014	3A
137	R26A-II	21	this study	Kumeu, NZ	Green kiwifruit	2014	3A
138	R31A-II	29	this study	Kumeu, NZ	Green kiwifruit	2014	2A
139	R31B-I	6	this study	Kumeu, NZ	Green kiwifruit	2014	3A
140	R31C-I	1	this study	Kumeu, NZ	Green kiwifruit	2014	3A
141	R33B-I	41	this study	Kumeu, NZ	Green kiwifruit	2014	3A
142	R33B-II	42	this study	Kumeu, NZ	Green kiwifruit	2014	3A
143	R33C-I	4	this study	Kumeu, NZ	Green kiwifruit	2014	3A

	ID	ST	Source	Location	Host	Year	PG
144	R34A-I	38	this study	Kumeu, NZ	Green kiwifruit	2014	3A
145	R34A-II	6	this study	Kumeu, NZ	Green kiwifruit	2014	3A
146	R34B-I	14	this study	Kumeu, NZ	Green kiwifruit	2014	3A
147	R35A-I	46	this study	Kumeu, NZ	Green kiwifruit	2014	3A
148	R35A-II	4	this study	Kumeu, NZ	Green kiwifruit	2014	3A
149	*P. syringae* pv. *aceris*. A10853	100	PAMDB		maple		
150	*P. syringae* pv. *aesculi* 0893_23	101	PAMDB	US	Horse chestnut		
151	*P. syringae* pv. *atrofaciens* DSM5025	102	PAMDB		wheat		
152	*P. syringae* pv. *actinidiae* FTRS_L1	103	PAMDB	Japan	kiwifruit	1984	1B
153	*P. syringae* pv. *apii* 1089_5	104	PAMDB	US	celery		
154	*P. syringae* pv. *broussonetiae* KOZ8101	105	PAMDB	Japan	paper mulberry	1980	2A
155	*P. syringae* pv. *cilantro* 0788_9	106	PAMDB	US	Cilantro		5
156	*P. syringae* pv. *coronafaciens* 3113	107	PAMDB	UK	Oats	1958	
157	*P. syringae* pv. *coronafaciens* KN221	108	PAMDB		Oats	1984	4
158	*P. syringae* pv. *glycinea* BR1	109	PAMDB		Soybean	1989	3
159	*P. syringae* pv. *glycinea* KN127	110	PAMDB		Soybean	1982	
160	*P. syringae* pv. *glycinea* KN44	111	PAMDB	Japan	Soybean	1981	
161	*P. syringae* pv. *glycinea* LN10	112	PAMDB		Soybean	1989	
162	*P. syringae* pv. *glycinea* M301765	113	PAMDB		Soybean	1982	
163	*P. syringae* pv. *glycinea* MOC601	114	PAMDB		Soybean	1994	
164	*P. syringae* pv. *japonica* M301072	115	PAMDB	Japan	Barley	1951	2B
165	*P. syringae* pv. *lachrymans* 107	116	PAMDB		Cucumber		
166	*P. syringae* pv. *lachrymans* 1188_1	117	PAMDB	US	Zucchini		
167	*P. syringae* pv. *lachrymans* A7386	118	PAMDB		Cucumber		
168	*P. syringae* pv. *lachrymans* N7512	119	PAMDB	Japan	Cucumber	1975	3
169	*P. syringae* pv. *maculicola* 4981	120	PAMDB	Zimbabwe	Cauliflower		1A
170	*P. syringae* pv. *maculicola* ES4326	121	PAMDB		Radish	1965	
171	*P. syringae* pv. *maculicola* H7608	122	PAMDB		Chinese Cabbage	1983	
172	*P. syringae* pv. *maculicola* KN84	123	PAMDB		Radish	1982	1A
173	*P. syringae* pv. *maculicola* KN91	124	PAMDB		Radish	1982	
174	*P. syringae* pv. *maculicola* M4	125	PAMDB		Radish		
175	*P. syringae* pv. *maculicola* M2	126	PAMDB				1A
176	*P. syringae* pv. *maculicola* YM7930	127	PAMDB		Radish	1979	
177	*P. syringae* pv. *mellea* N6801	128	PAMDB		Tobacco	1968	3
178	*P. syringae* pv. *mori* M301020	129	PAMDB	Japan	Mulberry	1966	3
179	*P. syringae* pv. *morsprunorum* 19322	130	PAMDB		European plum		
180	*P. syringae* pv. *morsprunorum* FTRS_U7	131	PAMDB	Japan	Japanese apricot	1978	3
181	*P. syringae* pv. *myricae* AZ84488	132	PAMDB		Bayberry	1984	3
182	*P. syringae* pv. *myricae* M302941	133	PAMDB		Bayberry	1989	
183	*P. syringae* pv. *oryzae* 36_1	134	PAMDB		Rice	1983	4

	ID	ST	Source	Location	Host	Year	PG
184	*P. syringae* pv. *oryzae* I_6	135	PAMDB		Rice	1991	4
185	*P. syringae* pv. *phaseolicola* 1302A	136	PAMDB	Ethiopia	Kidney bean	1984	
186	*P. syringae* pv. *phaseolicola* 1448A	137	PAMDB				3
187	*P. syringae* pv. *phaseolicola* 1449B	138	PAMDB	Ehtiopia	Hyacinth bean	1985	
188	*P. syringae* pv. *phaseolicola* HB10Y	139	PAMDB		Snap bean		
189	*P. syringae* pv. *phaseolicola* KN86	140	PAMDB	Japan	Kidney bean	1982	3
190	*P. syringae* pv. *phaseolicola* NPS3121	141	PAMDB		Kidney bean		
191	*P. syringae* pv. *phaseolicola* R6a	142	PAMDB		Kidney bean		
192	*P. syringae* pv. *pisi* 895A	143	PAMDB		Pea		
193	*P. syringae* pv. *pisi* PT NCPPB 2585 BS2753	144	PAMDB				2B
194	*P. syringae* pv. *pisi* H5E3	145	PAMDB		Pea	1993	2B
195	*P. syringae* pv. *pisi* H6E5	146	PAMDB		Pea	1994	
196	*P. syringae* pv. *pisi* H7E7	147	PAMDB		Pea	1995	
197	*P. syringae* pv. *pisi* R6a	148	PAMDB		Pea		
198	*P. syringae* Cit7	149	PAMDB		Navel Orange		2A
199	*P. syringae* TLP2	150	PAMDB		Potato		
200	*P. syringae* pv. *sesami* HC_1	151	PAMDB		Sesame		
201	*P. syringae* pv. *savastanoi* PT 4352	152	PAMDB	Yugoslavia	Olive		
202	*P. syringae* pv. *syringae* 61	153	PAMDB				
203	*P. syringae* pv. *syringae* 1212R	154	PAMDB		Pea		
204	*P. syringae* pv. *syringae* A2	155	PAMDB		Ornamental pear		
205	*P. syringae* pv. *syringae* B48	156	PAMDB	US	Peach		
206	*P. syringae* pv. *syringae* B64	157	PAMDB	US	Wheat		
207	*P. syringae* pv. *syringae* B728A	158	PAMDB	US	Snap bean		
208	*P. syringae* pv. *syringae* B76	159	PAMDB	US	Tomato		
209	*P. syringae* pv. *syringae* FF5	160	PAMDB	US	Ornamental pear	1998	
210	*P. syringae* pv. *syringae* FTRS_W6	161	PAMDB		Japanese apricot	1966	
211	*P. syringae* pv. *syringae* FTRS_W7	162	PAMDB		Japanese apricot	1978	
212	*P. syringae* pv. *syringae* L177	163	PAMDB		Lilac	1983	
213	*P. syringae* pv. syringae LOB2_1	164	PAMDB	Japan	Lilac	1986	2B
214	*P. syringae* pv. *syringae* NCPPB28	165	PAMDB	UK	Lilac		
215	*P. syringae* pv. *syringae* Ps9220	166	PAMDB		Spring onion	1992	4
216	*P. syringae* pv. *syringae* PSC1B	167	PAMDB	US	corn		
217	*P. syringae* pv. *tabaci* 6606	168	PAMDB	Japan	Tobacco	1967	3
218	*P. syringae* pv. *thea* K93001	169	PAMDB		Tea	1993	1B
219	*P. syringae* pv. *tomato* 487	170	PAMDB	Greece	Tomato		
220	*P. syringae* pv. *tomato* 1318	171	PAMDB	Switzerland	Tomato		
221	*P. syringae* pv. *tomato* 2170	172	PAMDB		Tomato	1984	
222	*P. syringae* pv. tomato DC3000	173	PAMDB	UK	Tomato		1A
223	*P. syringae* pv. *tomato* DC84_1	174	PAMDB	Canada	Tomato		
224	*P. syringae* pv. *tomato* DC89_4H	175	PAMDB	Canada	Tomato		
225	*P. syringae* pv. *tomato* DCT6D1	176	PAMDB	Canada	Tomato		

	ID	ST	Source	Location	Host	Year	PG
226	*P. syringae* pv. *tomato* KN10	177	PAMDB		Tomato	1981	
227	*P. syringae* pv. *tomato* PT23	178	PAMDB	US	Tomato	1986	
228	*P. syringae* pv. *tomato* TF1	179	PAMDB	US	Tomato	1997	
229	*P. syringae* pv. *aptata* 601	180	PAMDB		Sugar beet	1966	2B
230	*P. syringae* pv. *aptata* DSM5022	181	PAMDB		wheat		
231	*P. syringae* pv. *aptata* G733	182	PAMDB		Brown rice	1976	
232	*P. syringae* pv. *syringae* 508	183	PAMDB				
233	*P. syringae* pv. *syringae* 915 29	184	PAMDB				
234	*P. syringae* pv. *syringae* 908 20-1	185	PAMDB				
235	*P. syringae* pv. *syringae* 907 7-1	186	PAMDB				
236	*P. syringae* pv. *syringae* 893	187	PAMDB				
237	*P. syringae* pv. *syringae* 909 35-1	188	PAMDB				
238	*P. syringae* pv. *syringae* cc1502	189	PAMDB				
239	*P. syringae* pv. *syringae* cc667	190	PAMDB				
240	*P. syringae* pv. *syringae* 914 19-1	191	PAMDB				
241	*P. syringae* pv. *syringae* 892	192	PAMDB				
242	*P. syringae* pv. *syringae* 895	193	PAMDB				
243	*P. syringae* pv. *syringae* 872	194	PAMDB				
244	*P. syringae* pv. *syringae* 871	195	PAMDB				
245	*P. syringae* pv. *syringae* 645 B1-11-10	196	PAMDB				
246	*P. syringae* pv. *syringae* 912 23	197	PAMDB				
247	*P. syringae* pv. *syringae* 889	198	PAMDB				
248	*P. syringae* pv. *syringae* 898	199	PAMDB				
249	*P. syringae* pv. *syringae* 863	200	PAMDB				
250	*P. syringae* pv. *syringae* 884	201	PAMDB				
251	*P. syringae* pv. *syringae* 879	202	PAMDB				
252	*P. syringae* pv. *syringae* 642 B2-1-1	203	PAMDB				
253	*P. syringae* pv. *syringae* 891	204	PAMDB				
254	*P. syringae* pv. *syringae* 869	205	PAMDB				
255	*P. syringae* pv. *syringae* cc1503	206	PAMDB				
256	*P. syringae* pv. *apii* PT LMG 2132 BS2456	207	PAMDB				1A
257	*P. syringae* BS2725	208	PAMDB				
258	*P. cannabina alisalensis* ICMP 4326	209	PAMDB				
259	*P. syringae* pv. *coronafaciens* PT NCPPB 600 BS165	210	PAMDB				
260	*P. syringae* BS289	211	PAMDB				
261	*P. syringae* pv. *aptata* PT CFBP 1617 BS291	212	PAMDB				
262	*P. syringae* pv. *apii* BS2715	213	PAMDB				
263	*P. syringae* pv. *coriandricola* BS2059	214	PAMDB				
264	*P. syringae* pv. *coriandricola* BS2705	215	PAMDB	US	Parsley	2009	5
265	*P. syringae* pv. *atrofaciens* PT LMG 5095 BS2457	216	PAMDB				2B
266	*P. syringae* pv. *coriandricola* PT ICMP 12471 BS2673	217	PAMDB				
267	*P. syringae* pv. *syringae* PT LMG 1247 BS2733	218	PAMDB				
268	*P. syringae* pv. *aceris* PT LMG	219	PAMDB				

	ID	ST	Source	Location	Host	Year	PG
	2106 BS2735						
269	*P. syringae* pv. *lapsa* PT LMG 2206 BS2737	221	PAMDB				
270	*P. syringae* pv. *savastanoi* PT LMG 2209 BS2738	222	PAMDB				3
271	*P. syringae* pv. *primulae* LMG 2252 BS2740	223	PAMDB				7
272	*P. syringae* pv. *ribicola* PT LMG 2276 BS2741	224	PAMDB				7
273	*P. syringae* pv. *sesami* PT LMG 2289 BS2742	225	PAMDB				
274	*P. syringae* pv. *ulmi* PT LMG 2349 BS2743	226	PAMDB				
275	*P. syringae* pv. *viburni* PT LMG 2351 BS2744	227	PAMDB				
276	*P. syringae* pv. *atropurpurea* PT LMG 5030 BS2918	228	PAMDB				
277	*P. syringae* pv. *dysoxyli* PT LMG 5062 BS2746	229	PAMDB				
278	*P. syringae* pv. *garcae* LMG 5064 BS2747	230	PAMDB				
279	*P. syringae* pv. *helianthi* PT LMG 5067 BS2749	231	PAMDB				6
280	*P. syringae* pv. *mellea* PT LMG 5072 BS2750	232	PAMDB				
281	*P. syringae* pv. *mori* PT LMG 5074 BS2751	233	PAMDB				
282	*P. syringae* pv. *papulans* PT LMG 5076 BS2752	234	PAMDB				2A
283	*P. syringae* pv. *tagetis* PT LMG 5090 BS2754	235	PAMDB				6
284	*P. syringae* pv. *theae* PT LMG 5092 BS2755	236	PAMDB				
285	*P. syringae* pv. *passiflorae* PT LMG 5185 BS2756	237	PAMDB				
286	*P. syringae* pv. *delphinii* PT LMG 5381 BS2757	238	PAMDB				
287	*P. syringae* pv. *ciccaronei* PT LMG 5541 BS2758	239	PAMDB				
288	*P. syringae* pv. *oryzae* PT LMG 10912 BS2760	240	PAMDB				
289	*P. syringae* pv. *avellanae* type NCPPB 3487 BS2931	241	PAMDB				
290	*P. syringae* pv. *maculicola* PT CFBP 1657 BS286	242	PAMDB				
291	*P. syringae* pv. *tabaci* PT NCPPB 1427 BS2917	244	PAMDB				
292	*P. syringae* pv. *zizaniae* PT NCPPB 3690 BS2928	249	PAMDB				
293	*P. syringae* BS2779	250	PAMDB				
294	*P. syringae* pv. *daphniphylli* PT NCPPB 3617 BS2923	252	PAMDB				
295	*P. syringae* pv. *syringae* A21	253	PAMDB				
296	*P. syringae* pv. *syringae* A43	254	PAMDB				
297	*P. syringae* pv. *syringae* AL17	255	PAMDB				
298	*P. syringae* pv. *syringae* C1	256	PAMDB				
299	*P. syringae* pv. *syringae* C47	257	PAMDB				
300	*P. syringae* pv. *syringae* N2	258	PAMDB				
301	*P. syringae* pv. *syringae* P11	259	PAMDB				
302	*P. syringae* pv. *syringae* P18	260	PAMDB				
303	*P. syringae* pv. *syringae* PL35	261	PAMDB				
304	*P. syringae* pv. *syringae* PL38	262	PAMDB				
305	*P. syringae* pv. *syringae* AL12	263	PAMDB				
306	*P. syringae* pv. morsprunorum A65	264	PAMDB				

	ID	ST	Source	Location	Host	Year	PG
307	*P. syringae* pv. actinidiae PT NCPPB 3739 BS2929	904	PAMDB				
308	KW1-1	401	Tomihama et al. 2016	Japan	Green kiwifruit	2015	3A
309	KW2-1	402	Tomihama et al. 2016	Japan	Green kiwifruit	2015	3A
310	KW8-1	403	Tomihama et al. 2016	Japan	Green kiwifruit	2015	3A
311	KW11	3	Tomihama et al. 2016	Japan	Green kiwifruit	2015	3A
312	KW17-1	404	Tomihama et al. 2016	Japan	Green kiwifruit	2015	3A
313	KW22-1	405	Tomihama et al. 2016	Japan	Green kiwifruit	2015	3A
314	KW22-2	406	Tomihama et al. 2016	Japan	Green kiwifruit	2015	3A
315	KW23	407	Tomihama et al. 2016	Japan	Green kiwifruit	2015	3A
316	KW8-2	408	Tomihama et al. 2016	Japan	Green kiwifruit	2015	2A
317	KW22-3	409	Tomihama et al. 2016	Japan	Green kiwifruit	2015	2A
318	OK1	410	Tomihama et al. 2016	Japan	Okra		
319	A1	411	Tomihama et al. 2016	Japan	Tea		
320	KID0001	412	Tomihama et al. 2016	Japan	Air		
321	KID0033	413	Tomihama et al. 2016	Japan	Air		
322	KID0034	414	Tomihama et al. 2016	Japan	Air		
323	KID0190	415	Tomihama et al. 2016	Japan	Air		
324	KID0206	416	Tomihama et al. 2016	Japan	Air		
325	KID0217	417	Tomihama et al. 2016	Japan	Air		
326	KID0218	418	Tomihama et al. 2016	Japan	Air		
327	KID0220	419	Tomihama et al. 2016	Japan	Air		
328	KID0222	420	Tomihama et al. 2016	Japan	Air		
329	KID0231	421	Tomihama et al. 2016	Japan	Air		
330	KID0234	422	Tomihama et al. 2016	Japan	Air		
331	KID0236	423	Tomihama et al. 2016	Japan	Air		
332	KID0237	424	Tomihama et al. 2016	Japan	Air		
333	KID0244	425	Tomihama et al. 2016	Japan	Air		
334	KID0245	426	Tomihama et al. 2016	Japan	Air		
335	KID0246	427	Tomihama et al. 2016	Japan	Air		
336	KID0247	428	Tomihama et al. 2016	Japan	Air		
337	KID0258	429	Tomihama et al. 2016	Japan	Air		
338	KID0260	430	Tomihama et al. 2016	Japan	Air		
339	KID0261	431	Tomihama et al. 2016	Japan	Air		
340	KID0262	432	Tomihama et al. 2016	Japan	Air		
341	KID0269	433	Tomihama et al. 2016	Japan	Air		
342	KID0284	434	Tomihama et al. 2016	Japan	Air		
343	KID0293	435	Tomihama et al. 2016	Japan	Air		

	ID	ST	Source	Location	Host	Year	PG
344	KID0295	436	Tomihama et al. 2016	Japan	Air		
345	KID0297	437	Tomihama et al. 2016	Japan	Air		
346	KID0299	438	Tomihama et al. 2016	Japan	Air		
347	KID305	439	Tomihama et al. 2016	Japan	Air		
348	KID307	440	Tomihama et al. 2016	Japan	Air		
349	KID309	441	Tomihama et al. 2016	Japan	Air		
350	KID313	442	Tomihama et al. 2016	Japan	Air		
351	KID315	443	Tomihama et al. 2016	Japan	Air		
352	KID316	444	Tomihama et al. 2016	Japan	Air		
353	LKBQ0	445	Visnovsky et al. 2016	TePuke, NZ	Gold kiwifruit		1B
354	LKCH0	3	Visnovsky et al. 2016	Te Puke, NZ	Gold kiwifruit	2010	3A
355	LKCI0	446	Visnovsky et al. 2016	Te Puke, NZ	Gold kiwifruit	2011	3A
356	LKEL0	447	Visnovsky et al. 2016	USA	Gold kiwifruit		2D
357	LKEM0	448	Visnovsky et al. 2016	USA	Gold kiwifruit		5
358	LKEO0	449	Visnovsky et al. 2016	France, Livron sur Drome	Green kiwifruit	1985	2D
359	LKEP0	16	Visnovsky et al. 2016	Kumeu, NZ	Green kiwifruit	1991	2A
360	LKGU0	450	Visnovsky et al. 2016	NZ	Green kiwifruit	1991	1A
361	LKGV0	451	Visnovsky et al. 2016	Katikati, NZ	Green kiwifruit	1991	2B
361	*P. syringae* pv. *actinidiae* ICMP 1884	904	McCann et al. 2013	New Zealand	Green kiwifruit	2010	1B
361	*P. fluorescens* SBW25		Rainey & Bailey 1996	UK	sugar beet		outgroup

Table S 4.1: List of strains used in Chapter 4. Host plant *A. c.* = *Actinidia deliciosa*, PG = Phylogroup, collection alias = ICMP numbers for NCBI strains, isolate names / sequence types (ST) for strains used in the MLST study (Chapter 2).

	Isolate ID	Host plant	PG	Country	Year	Reference	Collection alias
1	C25	*A.c.* var. *chinensis*	PG3	China, Chongqing	2014	This study	NA
2	C32	*A.c.* var. *deliciosa*	PG3	China, Sichuan	2014	This study	NA
3	C37	*A.c.* var. *deliciosa*	PG3	China, Shaanxi	2014	This study	NA
4	C39	*A.c.* var. *deliciosa*	PG3	China, Shaanxi	2014	This study	NA
5	C40	*A.c.* var. *deliciosa*	PG3	China, Shaanxi	2014	This study	NA
6	C41	*A.c.* var. *deliciosa*	PG3	China, Henan	2014	This study	NA
7	C42	*A.c.* var. *deliciosa*	PG3	China, Henan	2014	This study	NA
8	C46	*Camellia sp.*	PG3	China, Henan	2014	This study	NA
9	C49	*A.c.* var. *chinensis*	PG3	China, Henan	2014	This study	NA
10	C50	*A.c.* var. *chinensis*	PG3	China, Henan	2014	This study	NA
11	C56	*A.c.* var. *chinensis*	PG3	China, Guizhou	2014	This study	NA
12	C57	*A.c.* var. *chinensis*	PG3	China, Guizhou	2014	This study	NA
13	C58	*A.c.* var. *chinensis*	PG3	China, Guizhou	2014	This study	NA
14	C59	*A.c.* var. *chinensis*	PG3	China, Guizhou	2014	This study	NA
15	C64		PG3	China	2012	This study	NA
16	C65		PG3	China	2012	This study	NA
17	NZ50	*A.c.* var. *deliciosa*	PG3	NZ, Kumeu	2013	This study	R14C-I / ST4
18	NZ51	*A.c.* var. *chinensis*	PG3	NZ, Pukekohe	2014	This study	G33A-I / ST1
19	NZ55	*A.c.* var. *chinensis*	PG3	NZ, Kumeu	2013	This study	K14C-I / ST6
20	NZ56	*A.c.* var. *deliciosa*	PG3	NZ, Kumeu	2014	This study	R23A-I / ST3
21	NZ57	*A.c.* var. *deliciosa*	PG3	NZ, Kumeu	2014	This study	R33C-I / ST4

	Isolate ID	Host plant	PG	Country	Year	Reference	Collection alias
22	NZ58	*A.c.* var. *chinensis*	PG3	NZ, Pukekohe	2014	This study	G23B-II / ST1
23	*P. syringae* G33C	*A.c.* var. *chinensis*	PG3	NZ, Pukekohe	2014	This study	G33C-I / ST1
24	*P. syringae* ICMP19498	*A.c.* var. *chinensis*	PG3	NZ, Te Puke	2010	Visnovsky et al. (2016)	ICMP19498 / ST3
25	*P. syringae* ICMP19499	*A.c.* var. *chinensis*	PG3	NZ, Te Puke	2011	Visnovsky et al. (2016)	ICMP19499 / ST446
26	*P. syringae* pv. *sesame*		PG3			Thakur et al. (2016)	ICMP763
27	*P. savastanoi* pv. *glycinea*		PG3			Thakur et al. (2016)	ICMP2189
28	*P. savastanoi* pv. *phaseolicola*		PG3			Thakur et al. (2016)	ICMP2740
29	*P. syringae* pv. *tabaci*		PG3			Thakur et al. (2016)	ICMP2835
30	*P. syringae* pv. *ulmi*		PG3			Thakur et al. (2016)	ICMP3962
31	*P. syringae* pv. *mori*		PG3			Thakur et al. (2016)	ICMP4331
32	*P. syringae* pv. *eriobotryae*		PG3			Thakur et al. (2016)	ICMP4455
33	*P. syringae* pv. *ciccaronei*		PG3			Thakur et al. (2016)	ICMP5710
34	*P. meliae*		PG3			Thakur et al. (2016)	ICMP6289
35	*P. syringae* pv. *myricae*		PG3			Thakur et al. (2016)	ICMP7118
36	*P. syringae* pv. *photiniae*		PG3			Thakur et al. (2016)	ICMP7840
37	*P. syringae* pv. *dendropanacis*		PG3			Thakur et al. (2016)	ICMP9150
38	*P. syringae* pv. *hibisci*		PG3			Thakur et al. (2016)	ICMP9623
39	*P. syringae* pv. *rhaphiolepidis*		PG3			Thakur et al. (2016)	ICMP9756
40	*P. syringae* pv. *daphniphylli*		PG3			Thakur et al. (2016)	ICMP9757
41	*P. syringae* pv. *broussonetiae*		PG3			Thakur et al. (2016)	ICMP13650
42	*P. syringae* pv. *nerii*		PG3			Thakur et al. (2016)	ICMP16943
43	*P. syringae* pv. *cerasicola*		PG3			Thakur et al. (2016)	ICMP17524

	Isolate ID	Host plant	PG	Country	Year	Reference	Collection alias
44	*P. amygdali* pv. *tabaci* ATCC11528	*Nicotiana benthamiana*	PG3	USA	1905		ICMP10043
45	*P. amygdali* pv. *aesculi* Pae0893_23	*Aesculus hippocastanum*	PG3	India	1969		NCPPB 3681
46	*P. savastanoi* pv. *savastanoi* NCPPB3335	*Olea europaea*	PG3	France	1984		NCPPB3335
47	*P. amygdali* pv. *mori* PmoM301020	*Morus alba*	PG3	Japan (Nagano)	1966		MAFF301020
48	*P. savastanoi* pv. *phaseolicola* 1448A	*Phaseolus vulgaris*	PG3	Ethiopia	1985	Joardar et al. (2005)	
49	*P. syringae* pv. *actinidiae* NZ54	*A.c.* var. *chinensis*	PG1	NZ, Pukekohe	2014	This study	G33C-II / ST904