The role of integrative conjugative elements in evolution of the kiwifruit pathogen *Pseudomonas syringae* pv. *actinidiae*

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

Doctor of Philosophy

in

Genetics

at Massey University, Albany Campus, New Zealand.

Elena Colombi

2017

Abstract

Horizontal gene transfer (HGT) is a major force driving evolution in prokaryotes. Among the different contributions to bacterial fitness, HGT underpins the evolution of pathogenicity and virulence in several bacterial pathogens. *Pseudomonas syringae* pv. *actinidiae* (*Psa*) first emerged as a pathogen of kiwifruit in Asia during the 1980's. In 2008 an outbreak occurred in Italy that rapidly spread to major kiwifruit growing areas of the world. During its global journey the outbreak lineage independently acquired divergent Integrative and Conjugative Elements (ICEs), harbouring an identical set of cargo genes that are hypothesized to be involved in the plant-pathogen interaction. Here I show that the three ICEs acquired by *Psa* belong to a diverse family of ICEs present only in plant-associated *Pseudomonas* ssp. (the PsICEs). The evolution of PsICEs is characterized by extensive inter-ICE recombination events that are frequent enough to mask evolutionary history, producing chimeras with variable patterns of similarity to each other, yet maintaining a syntenic backbone where cargo genes are integrated in conserved positions. Although there are different classes of PsICE cargo genes, one set was frequently recovered: those contained on a Tn*6212* element. Tn*6212* confers a selective benefit when *Psa* is grown on succinate, fumarate, or malate as the only carbon source, but no phenotype was detected *in planta.* Members of the PsICE family also confer copper resistance to *Psa* strains isolated in New Zealand. I analyzed a number of these and showed transfer *in vitro* and *in planta*. I also measured the fitness consequences of ICE carriage, captured the *de novo* formation of novel recombinant ICEs, and explored ICE host-range.

Together my work, which began with observations from genome sequences before moving to experimental studies in the laboratory, has provided new insights into the role that horizontal gene transfer plays in the evolution of virulence.

Acknowledgements

First and foremost, I would like to thank my supervisor, Distinguished Professor Paul Rainey for his continuous support and curiosity in my work. I am really grateful for the opportunities and ideas that were developed during these years. I am also grateful to my co-supervisor Dr Honour McCann for the infinite help and time she dedicated to me. I would also like to thank Zespri, the Te Puke Growers Association and the New Zealand Institute of Advanced Study for funding this project.

Thanks to all the past and present members of the Rainey Lab who have helped and stood by me during my studies. I am particularly grateful to the other member of the *Psa* team Christina Straub, but also to Dr Daniel Rexin and to Dr Philippe Remigi for helpful discussions, advice and sometimes labour. Thanks to Associate Professor Matt Templeton for the kiwifruit plants, plasmids and genome sequencing. Thanks to Professor Kee Sohn and the members of his lab (especially Jay, Cecile, Sera and Maxim) for hosting and helping run experiments in beautiful Palmerston North.

I am profoundly grateful to “la mamma e i’ babbo” who have always supported my life decisions, both great and small. And Marco, otherwise he gets touchy. Last but not least, a very special thank you for Difa, because he is always here, wherever ‘here’ is.

Table of Contents

List of Illustrations

List of Tables

Chapter 1 – Introduction

1.1 The plant-pathogen interaction

Plants face perpetual challenges from the microbes that experience them as sources of nutrients. Through evolutionary time this has generated arms races between microbes and host plants that has driven the evolution of mechanisms by which microbes cause disease, and corresponding defense responses in plants. From time to time, microbes win the race and cause disease.

During the course of human evolution plants have been domesticated and grown as crops. Very often crop plants are genetically uniform and densely planted, which offers particular opportunity for the evolution of virulence leading to the emergence of pathogens. In some instances, outbreaks of disease have caused devastating losses with major implications for human populations dependent on the crop plant. For example the Irish potato famine in the 19th century was caused by *Phytophthora infestans* (Ristaino, 2002), the Panama disease caused by *Fusarium oxysporum* could cause the end of banana industries (Ploetz, 2015), the citrus canker caused by *Xanthomonas axonopodis* has deeply impacted the citrus production worldwide with million of trees destroyed in the American continent (Graham *et al.*, 2004), and the same risk faces olive production

with the introduction of *Xylella fastidiosa* in south Italy (Bosso *et al.*, 2016). In New Zealand and in the other countries that produce kiwifruit, bleeding canker disease caused by *Pseudomonas syringae* pv. *actnidiae* (*Psa*) has led to almost complete eradication of the cultivated susceptible variety of golden kiwifruit (Dwiartama, 2017).

Humans have brought an additional dimension to the interaction between plants and microbes through the use of chemical agents to control disease. Fungicides and bactericides can prevent disease but also select for resistance as is evident from studies of the spread of antibiotic resistance in bacteria (Alonso *et al.*, 2001; Martínez, 2008; Davies & Davies 2010; Toprak *et al.*, 2012; von Wintersdorff *et al.*, 2016). The most common antimicrobial agent used in agriculture is not an antibiotic but copper, also used in biological agriculture (Lindow, 2017).

The work described in this thesis concerns the kiwifruit pathogen *Psa* with focus on its evolution through horizontal gene transfer (HGT), in particular by the acquisition of Integrative and Conjugative Elements (ICEs). Following a brief introduction to mechanisms by which microbes cause disease in plants, I describe the *P. syringae* species complex with particular attention on *Psa*. A general introduction to HGT and ICE biology follows, and ICEs found in the *P. syringae* complex will be described. I will then touch upon the importance of ICEs and their recently discovered role in mediating evolution of copper resistance in *Psa*. The final part of the Introduction concerns copper resistance.

1.1.1 Pathogenicity factors

The prerequisite for disease development in plants is summarized by the concept of the disease triangle introduced by Stevens in 1960. In the disease triangle an epidemic is caused by the overlapping of an infectious (or abiotic) agent, a susceptible host and an environment favourable for disease development (Figure 1.1). Later new variants of this concept called disease pyramid and tetrahedron were introduced incorporating time and human-mediated variables (Franci, 2001).

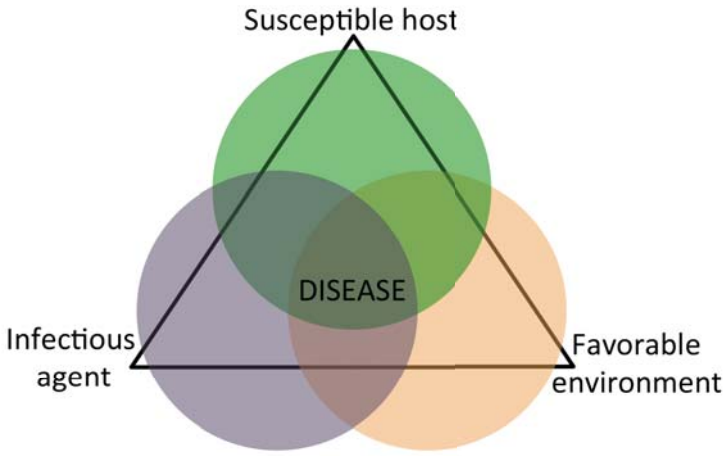

Figure 1.1.The disease triangle. The three necessary causal factors of disease (a susceptible host, an infectious agent and a favorable environment) are positioned at the vertices. Disease is caused by the overlapping of these factors.

Plants are constantly exposed to microbes, some of which have evolved systems to overcome plant defences and access plant tissues. Plant tissues provide plant pathogens with a protected niche and nutrients. To be pathogenic, bacteria must access the plant interior, either by actively penetrating plant surfaces or by entering through wounds or natural openings, such as stomata. Once inside the plant, microbes face a further obstacle as the plant can detect the intruding pathogen by recognizing microbial molecules, known as Pathogen-Associated Molecular Patterns (PAMPs) (Boller & Felix, 2009) and actively defend itself. PAMPs are structurally conserved across a wide range of microbes and normally

they are not present in the host (Nurnberger *et al.*, 2004, Segonzac & Zipfel, 2011). For example, plants recognize multiple cell-surface components of Gram-negative bacteria, including lipopolysaccharides, a major constituent of the outer membrane, and flagellin, the protein subunit of the flagellum (Chisholm *et al.*, 2006; Boller & Felix, 2009). Perception of a microorganism by the plant initiates PAMP-Triggered Immunity (PTI), which usually halts infection before the microbe gains a hold in the plant (Nicaise *et al.*, 2009). PTI may be the plants' first active response to microbial perception. However, pathogenic bacteria have evolved mechanisms to suppress PTI by interfering with recognition or by secreting effector proteins into the plant cell to alter resistance signaling or responses through the Type 3 Secretion System (T3SS). The T3SS is a multi-molecular syringe device, essential for the pathogenicity of Gram-negative bacteria, which use this system for the one-step delivery of T3SS effector proteins (T3SEs) from the bacterial cytoplasm into that of the eukaryotic host cells. Interestingly, ability to deliver pathogen proteins directly into plant host cells to alter plant defense is a unifying theme among phytopathogens (Grant *et al.*, 2006; Zhou & Chai, 2008; Cui *et al.,* 2009).

Once pathogens have suppressed primary defences (PTI), plants have another chance to stop microbial infection through a more specialized mechanism called Effector-Triggered Immunity (ETI) (Dodds & Rathjen, 2010). ETI involves direct or indirect recognition of T3SEs by plant Resistance (R) proteins. Activation of R protein-mediated resistance also suppresses microbial growth, but not before the invader has had an opportunity for limited proliferation. To be a successful pathogen the bacteria needs to overcome the ETI. Not surprisingly, pathogens have

evolved other effectors to suppress ETI (Chisholm *et al.*, 2006; Grant *et al.*, 2006; Zhou & Chai, 2008; Cui *et al.,* 2009).

1.1.2 Virulence factors

The definition of what constitutes a virulence factor is varied and controversial: many authorities consider pathogenicity and virulence as synonymous. In this thesis I refer to these terms as defined by Casadevall & Pirofski (1999). Pathogenicity is the capacity of a microbe to cause damage in a host and virulence is the relative capacity of a microbe to cause damage in a host. This general description of host–microbe interactions proposed by Casadevall & Pirofski in 1999 has been developed in subsequent papers into the "damage–response framework" of microbial pathogenesis (Casadevall & Pirofski 2000, 2001, 2003; Pirofski & Casadevall, 2002). The concept of virulence factors cannot be separated from that of microbial virulence. The problem in defining virulence arises because virulence is a microbial property that can only be expressed in a susceptible host. Hence, virulence is not an independent microbial property, because it cannot be defined independently of a host. In the damage–response framework, a pathogen is a microbe capable of causing host damage, virulence is the relative capacity of a microbe to cause damage in a susceptible host and a virulence factor is a microbial component that can damage a susceptible host. The microbial attributes that confer the potential for virulence fall primarily within several categories including; the ability to enter a host, to evade host defences, to grow in a host environment, to counteract host defence responses, to acquire iron and nutrients from the environment, and to sense environmental change

(Casadevall & Pirofski, 2009). A microbial attribute that confers resistance to antimicrobial compounds can be considered a virulence factor because it contributes to host damage. Indeed antimicrobial resistance often underlies the emergence of new or re-emergence of old bacterial and fungal diseases both in humans and plants (Brent & Hollomon, 2007; Garris *et al.,* 2009; Wozniak *et al.,* 2009; Mutreja *et al*., 2011; Hajiri *et al*., 2010; Cesbron *et al*., 2015; Wong *at al*., 2015; Njamkepo *et al*., 2016).

1.2 Pseudomonas syringae

Pseudomonas syringae is a Gram negative bacteria that causes diseases on monocots and herbaceous and woody dicots worldwide (Fatmi *et al*., 2008). As such it is also one of the model organisms for plant-pathogen interaction (Dangl & Jones, 2001; Jones & Dangl, 2006). Since the beginning of this century *P. syringae* has caused 55 outbreaks of disease on 25 different woody hosts and 72 outbreaks on 40 different annual plant species (Lamichhane *et al*., 2014, 2015). The *P. syringae* species complex currently comprises over 60 pathovars (pv.) divided into 13 phylogroups based on multi-locus sequence typing of 216 strains (Berge *et al*., 2014).

Beyond its prominent role in plant pathogenicity, *P. syringae* is also a model system for microbial evolution in nature (Baltrus *et al*., 2016). Its niche is not limited to cultivated and wild plants (as well as leaf litter), but it is also an ubiquitous environmental bacterium isolated from environments linked to the water cycle such as fresh snow and rainfall, lakes, rivers, snowpack, and even

clouds (Amato *et al.*, 2007; Monteil *et al.*, 2012; Morris *et al.*, 2007, 2008, 2013). It has been proposed that the existence of this highly diverse and recombinogenic background population that lives primarily in environmental habitats serves as reservoir from which pathogenic lineages emerge through clonal expansion (Monteil *et al.*, 2013).

1.2.1 The emergence of *Pseudomonas syringae* pv. *actinidiae*

Pseudomonas syringae pv. *actinidiae* (*Psa*) is a newly re-emerging pathogen causing the kiwifruit bleeding bacterial canker, the most destructive disease affecting kiwifruit (*Actinidia* spp.) cultivation worldwide (Balestra *et al.*, 2010). Symptoms of the disease include shoot die back, blight on young canes; presence of rusty-red or white exudate on canes or trunks; and angular necrotic leaf spots surrounded with yellow chlorotic halos (Figure 1.2), which eventually culminate in the death of the plant.

Figure 1.2. Symptoms produced by *Psa* in a kiwifruit orchard. (KVH).

Kiwifruit canker disease was first reported and identified as *Psa* in Japan in 1984; the disease affected the green kiwifruit cultivar *Actinidia chinensis* var. *deliciosa* Hayward (Serizawa *et al.*, 1989; Takikawa *et al.*, 1989). Although an outbreak of disease with similar symptoms to those produced by *Psa* was described in 1983-1984 in China (Fang *et al.*, 1990), no isolates were collected, and

therefore no positive identification could be made. *Psa* was then reported in 1988 in Korea on Hayward (Koh *et al*., 1994). While only confined to Korea and Japan, the disease was severe and many orchards were destroyed during that time. In 1992 the kiwifruit canker was also recorded in central Italy (Scortichini, 1994), but did not have such severe impact as in Asia. Kiwifruit canker disease only received worldwide attention when a massive outbreak in 2008 on the golden kiwifruit cultivar *A. chinensis* var. *chinensis* Hort16A quickly spread to all major kiwifruit-growing countries. This pandemic has been devastating for the kiwifruit industry. At the beginning the disease was reported in Italy (Balestra *et al*., 2010), and over the following three years *Psa* spread rapidly to neighbouring European countries, reaching Turkey in 2009/2010 (Bastas & Karakaya, 2017), Portugal and France in 2010 (Mazzaglia *et al*., 2012; Vanneste *et al*., 2011), and Spain and Switzerland in 2011 (Abelleira *et al*., 2011; EPPO, 2011). In 2010 the outbreak reached two other major countries cultivating kiwifruit: Chile and New Zealand (EPPO, 2011; Everett *et al*., 2011). Semi-mechanistic and correlative models based on the *Psa* climate suitability showed that *Psa* can potentially spread to wherever kiwifruit is cultivated (Khandan *et al*., 2013).

The structure of the global *Psa* population is represented in Figure 1.3. The *Psa*-1 clade includes strains of *Psa* isolated and described during the first recorded epidemic of bleeding canker disease in Japan (1984-1988). *Psa*-2 includes isolates from an epidemic in South Korea (1997-1998), and *Psa*-3 includes isolates from the latest global pandemic of kiwifruit canker disease (pandemic lineage) as well as more divergent Chinese strains (2008-present). *Psa*-5 is the second most recently reported clade of *Psa*, and is found only in a localized area of Japan at the present (Fujikawa & Sawada, 2016). A novel biovar (clade) 6 has been identified in

Japan; this also has a limited distribution (Sawada *et al.*, 2016). *Psa*-4, which causes only mild foliar symptoms and is more distantly related to canker causing *Psa* than *P. syringae* pv. *theae*, has been reassigned as pathovar *actinidifoliorum* (Cunty *et al.*, 2015).

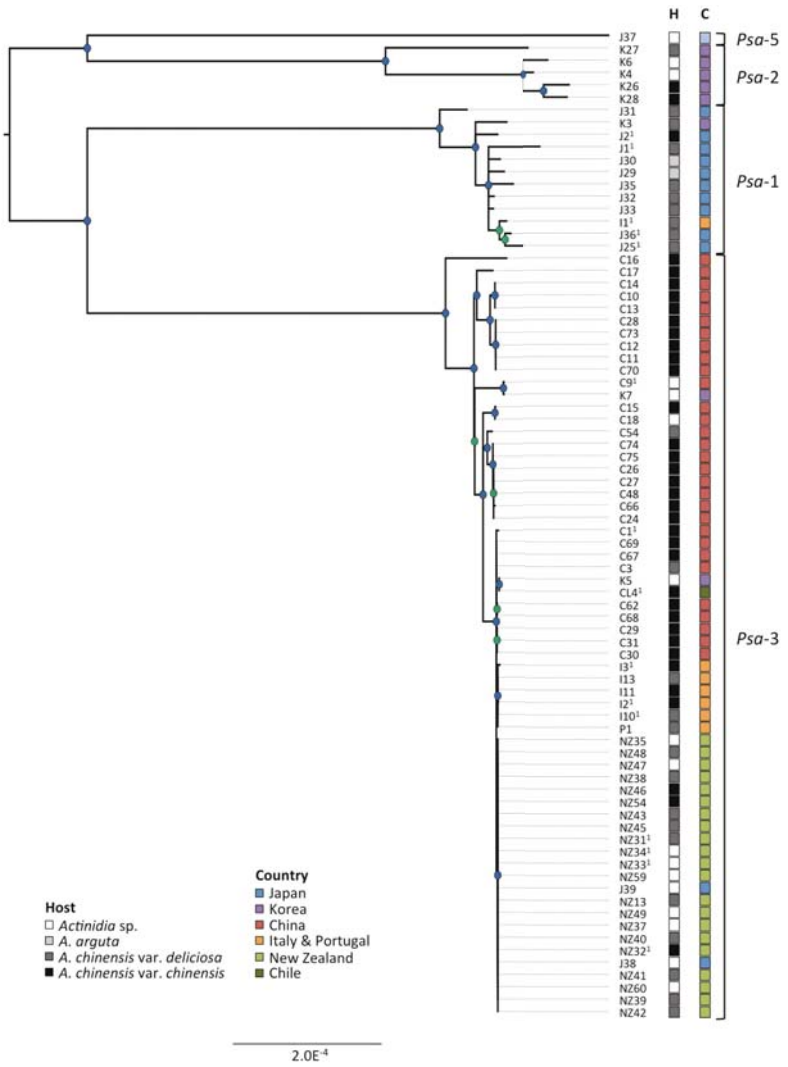

Figure 1.3. Phylogeny of *Psa*. RaxML Maximum likelihood tree based on 1,059,722bp non-recombinant core genome alignment including 2,953 variant sites (reproduced from McCann *et al.*, 2017).

1.2.2 *Pseudomonas syringae* pv. *actinidiae* in New Zealand

New Zealand was infected with a specific lineage of *Psa*-3 via one single transmission event between 2009 and 2010 (possibly via a batch of infected

pollen) (McCann *et al.*, 2017). The introduction of *Psa* in New Zealand was dramatic for the kiwifruit industry, with 90% of the kiwifruit orchards infected (KVH Psa Statistics, 2016) and losses predicted to be over $800 million across a fifteen year period (Greer & Saunders, 2012). The disease has led to the replacement of the golden cultivar Hort16A with the more tolerant golden cultivar G3 (Greer & Saunders, 2012). Although Hort16A is more susceptible to *Psa*-3 pandemic lineage than Hayward (McCann *et al.*, 2013), the *Psa*-3 pandemic lineage evolved to be particularly virulent on the cultivar, being able to spread more quickly in comparison to *Psa*-1 and 2. Genomic analysis carried out during the early pandemic revealed that the only significant genomic change that occurred in the *Psa*-3 pandemic lineage involved the acquisition through HGT of an ICE predicted to manipulate the host metabolism and increase *Psa* virulence (McCann *et al.*, 2013).

1.3 Horizontal gene transfer

HGT is "*the non-genealogical transmission of genetic material from one organism to another*" (Goldenfeld & Woese, 2007). In the microbial world, HGT is a major force shaping bacterial evolution (Mirkin *et al.,* 2003; Kunin *et al.,* 2005; Dagan & Martin 2007; Halary *et al.*, 2010; Kloesges *et al.*, 2011). Such is the prevalence of HGT in prokaryotes that a concept of species in bacteria is fraught with difficulties (Ward, 1998; Rossello-Mora & Amann, 2001; Cohan, 2002; Stackebrandt *et al.*, 2002; Doolittle, 2012) as the classical interpretation of an evolutionary 'tree' of vertical descent is confounded.

In a study of 329 proteobacterial genomes, it has been estimated that at least 75% of the protein families have been affected by HGT during evolution and that 21±9% of the genes among these genomes are of recent acquisition (Kloesges *et al.*, 2011). A study quantifying both insertion and deletion rate of genes relative to mutations estimated that in *Bacillus cereus*, genes were gained and lost around 4.4 times the rate of single nucleotide substitution per site (Hao & Golding, 2006). Similar rates have been estimated for *Streptococcus* and *Corynebacterium* (Marri *et al.*, 2006, 2007). In *P. syringae* this estimate is four orders of magnitude higher than that for *B. cereus*, with individual lineages acquiring thousands of genes in the same period of time in which a 1% amino acid divergence in the core genome occurred (Nowell *et al.,* 2014).

Despite the extent and the predominant role of HGT in bacterial evolution, it has been suggested that the majority of HGT events are short-lived (Hoa & Golding, 2006; Kuo & Ochman, 2009; Vos *et al.*, 2015). Studies have shown a decrease in the rate of gene content change relative to the rate of nucleotide substitution in the deeper branches of genome phylogenies compared to the tips (Marri *et al.*, 2007; Hao & Golding 2010; Didelot *et al.*, 2012) and that the vast majority of individual gene gains are mapped to a single strain (Touchon *et al.*, 2009; Nowell *et al.*, 2014). The reason why HGT events can be transient can be explained by: genetic drift if the HGT changes are neutral; purifying selection if the HGT events are costly and thus eliminated; or the benefits of the gene gain itself being transient (Vos *et al.*, 2015). On the other hand, even if on average gene transfers are more likely to decrease fitness than mutations, it is likely that the proportion of changes with strong beneficial fitness effects is higher for HGT than it is for mutation; indeed HGT can result in the uptake of whole genes, operons,

plasmids or ICEs. Moreover, while mutations are purely random, in the case of HGT via conjugation, the cargo genes can be refined by selection in the environment of the donor strain.

Potential costs of HGT can be generated for different reasons. For instance the absence of coevolution between the HGT element and recipient could represent a cost either at the nucleotide or protein level. Acquisition of a new element could result in a metabolic burden that will lower the fitness of the recipient (Diaz-Ricci & Hernández, 2000; Subbiah *et al*., 2011; Harrison & Brockhurst, 2012; van Rensburg *et al*., 2012) or in alteration of the normal phenotype of the cell (Sato & Kuramitsu, 1998; Gaillard *et al*., 2008). A poorly adapted codon usage that is not optimal for the recipient's tRNA pool could prevent HGT (Medrano-Soto *et al*., 2004; Tuller *at al*., 2011), however compensatory mutations could soften this effect (Kudla *et al.,* 2009; Amorós-Moya *et al*., 2010).

1.3.1 Contribution of HGT to bacterial virulence evolution

HGT allows the rapid evolution of bacteria as, especially with conjugation, it is possible for functional cassettes of genes to move between different strains, species, genera and phyla (Caro-Quintero & Konstantinidis, 2015; Fondi *et al*., 2016). The acquisition of new traits via HGT may allow recipient organisms to exploit novel biochemical functions already refined by selection in the donor organism, thereby facilitating the efficient and effective invasion of new ecological niches (Lawrence, 2002). The colonization of new niches in turn may lead to diversification resulting in bacterial speciation (Polz *et al*., 2013). If the new

horizontally transmitted trait confers more benefits than costs it will be amplified by positive selection, if vice versa (costs>benefits) the trait will be lost by negative selection. If the trait does not influence the host phenotype it could be either lost or even fixed in the population by genetic drift. HGT events tend to be adaptive in prokaryotes, although positive selection on genome dynamics is on average likely to be weak (Sela *et al.*, 2016).

HGT can have different impacts depending on the niche and the type of interaction the recipients have with it. HGT can either evolve harmful traits such as pathogenicity, virulence, antimicrobial resistance but also beneficial traits such as symbiosis. *Pantoea agglomerans* has been shown to be transformed from epiphytic bacterium to pathogen of gypsophila and sugar beet by the acquisition of a plasmid encoding the T3SS, T3SE and a gene cluster responsible for the biosynthesis of indole-3-acetic acid (Barash & Manulis-Sasson, 2009). Also in *P. syringae* the T3SS is thought to have been horizontally acquired, as it is located on a genomic island (Alfano *et al,.* 2000) and its effectors evolve and are dispersed through HGT (McCann & Guttman, 2008). In *P. syringae*, a new type of genomic island called GInts has been demonstrated to excise, circularize and integrate to specific sites; they are thought to be horizontally transmitted and to carry genes predicted to be involved in virulence and resistance to antimicrobial compounds (Bardaji *et al.*, 2017). *Pectobacterium atrosepticum* and *P. carotovorum* acquired an integrative conjugative element (ICE) (HAI2) encoding the phytotoxin coronafacic acid, which was demonstrated to increase its virulence (Bell *et al.*, 2004; Vanga *et al.*, 2012; Panda *et al.*, 2016). *Xanthomonas arboricola* pv. *junglandis* strains responsible for a recent pandemic in France have been proposed to have undergone clonal

expansion with the acquisition of copper resistance genes encoded by the Xaj-ICE (Hajiri *et al.*, 2010; Cesbron *et al.*, 2015).

1.4 Integrative Conjugative Elements

Conjugation is a mechanism of HGT in prokaryotes by which DNA (conjugative transmissible elements) is transferred via pili or adhesins between two cells. Conjugative transmissible elements, such as plasmids and ICEs are thus able to move functional operons over broad phylogenetic distances and mediate abrupt changes in the phenotype of the host (Sullivan & Ronson, 1998; Ochman *et al.*, 2000; Ochman *et al.*, 2005; Guglielmini *et al.*, 2011). ICEs are plasmid-like entities with attributes of temperate phages usually integrated into the genomes (and thus passively replicating as a part of it) but also capable of horizontal transmission by their encoded mechanism for transfer. DNA sequencing analyses suggest that ICEs are the most abundant type of conjugative elements (Guglielmini *et al.*, 2011). Although a large number of ICEs are present in bacterial genomes, very few have been well characterized: the ICE*Bs1* in *Bacillus subtilis*, Tn*916* in *Eenterococcus faecalis*, ICE*SXT* in *Vibrio cholerae*, CTn*DOT* in *Bacteroides thetaiotaomicron*, ICE*St3* in *Streptococcus thermophilus*, ICE*clc* in *P. kamckmussii* and ICE*MlSym*^{R7A} in *Mesorhizobium loti* (Delavat *et al.*, 2017).

1.4.1 ICE life cycle

ICEs start their journey integrated into the genomes at specific attachment (*att*) sites (Figure 1.4a). ICE integration depends on a recombinase, most

commonly a tyrosine recombinase, but also serine. The recombinase catalyzes the reaction between a specific site on the ICE (*attP*) and a target sequence site in the chromosome (*attB*) generating two ICE-chromosome junction sequences (*attR* and *attL*). The recombinase dictates the integration site of the ICE, some recognize *attB* sites that are highly variable while others recognize specific *attB* sites only present one or two times across the genome (Hochhut *et al.*, 1993; Scott *et al.*, 1994; Klockgether *et al.*, 2007). A slight variation to this method of transfer is represented by the ICE*Mc*Sym^{1271} in *M. ciceri*. This is a tripartite island whose three parts are separately integrated across three *attB* sites. These three parts assemble in a unique ICE before the horizontal transfer (Haskett *et al.*, 2016).

ICE transfer can be activated by different signals, such as induction of the cellular response to DNA damage (recA-dependent SOS response) (Beaber *et al.*, 2004; Auchtung *et al.*, 2005; Bellanger *et al.*, 2007), cell-cell signaling from potential recipients (Ramsay *et al.*, 2006; 2009; Auchtung *et al.*, 2007), the growth phase of the host (Carraro *et al.*, 2011; Miyazaki *et al.*, 2012), mechanisms tied to selective advantages conferred by the cargo genes in an ICE (Doucet-Populaire *et al.*, 1991; Wang *et al.*, 2004, 2005), and host colonization (Lovell *et al.*, 2009; Quiroz *et al.*, 2011; Vanga *et al.*, 2015). The signal ICEs respond to are ICE-specific and some are known to respond to more than one stimulus. Moreover studies carried out by van der Meer's group revealed that only a few cells in the population bearing ICEs are active for the horizontal transfer, showing the process is stochastic (Reinhard *et al.*, 2013; Reinhard & van der Meer, 2014; Delavat *et al.*, 2016).

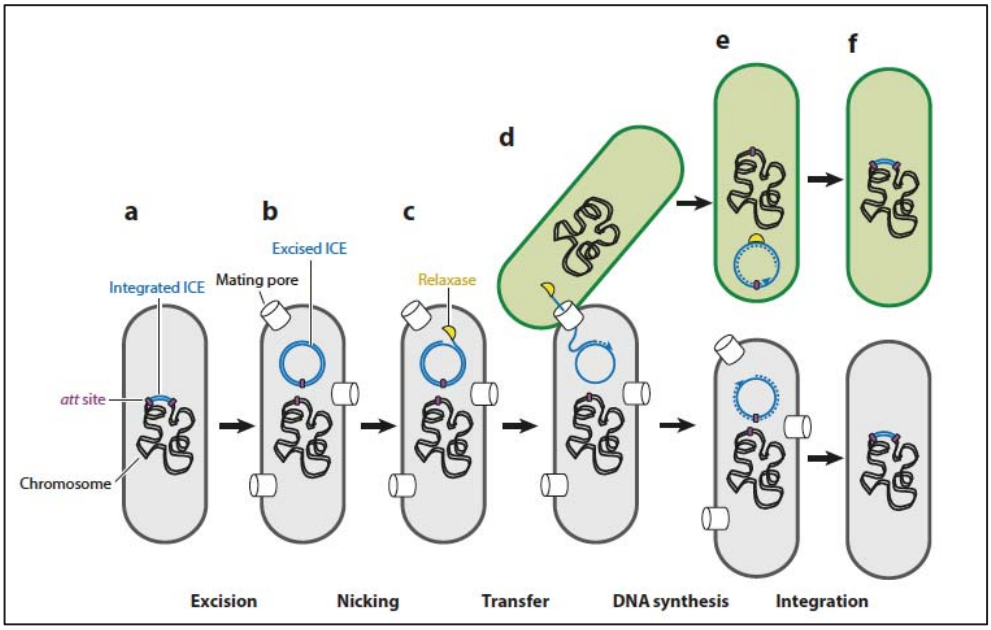

Figure 1.4. The ICE life cycle. ICE model of transfer via conjugation (reproduced from Johnson & Grossman, 2015).

When ICE gene expression is induced, the ICE excises from the host chromosome and circularizes forming a plasmid-like entity. Some ICE-encoded proteins assemble the mating machinery (cylinders spanning the donor cell envelope) (Figure 1.4b). The ICE-encoded relaxase nicks and attaches the 5' of one strand of the ICE (Figure 1.4c). If an appropriate recipient is available, the conjugation machinery transfers the single strand into the recipient cell (Figure 1.4d). For example the PAPI-1 island in *P. aeruginosa* is able to recognize specific lipopolysaccharides on the recipient outer membrane that are required to initiate the transfer (Hong *et al.*, 2017). In the recipient cell, the relaxase ligates the 5' and 3' ends of the DNA then the other strand is synthesized and integrates into the genome. In the donor, the remaining DNA strand is used as a template to generate a dsDNA circle that can then reintegrate into the host chromosome (Figure 1.4e and 1.4f) (Wozniak & Waldor, 2010; Johnson & Grossman, 2015).

Exclusion mechanisms, by which one ICE prevents the establishment of another in the same genome, have been also described. For PAPI-1 in *P. aeruginosa*

this mechanism is based on lipopolysaccharides on the recipient outer membrane (Hong *et al.*, 2017), for the SXT-R391 family in *Vibrio cholerae* element-specific exclusion activity is mediated by two genes encoded in the ICE (*eex* and *traG*) (Marrero & Waldor, 2007).

1.4.2 ICE structure

The genes required for ICEs life cycles are encoded in modules. These essential genetic modules can be functionally divided into: integration and excision, conjugation, and regulation of conjugative activity (Mohd-Zain *et al.*, 2004; Juhas *et al.*, 2007; Roberts & Mullany, 2009). In addition to a set of essential genes, ICEs often harbour accessory genes of adaptive significance to their hosts. These include genes affecting biofilm formation, pathogenicity, antibiotic and heavy metal resistance, symbiosis, bacteriocin synthesis and iron acquisition (Peters *et al.*, 1991; Rauch *et al.*, 1992; Ravatn *et al.*, 1998; Beaber *et al.*, 2002; Drenkard *et al.*, 2002; Burrus *et al.*, 2006; Ramsay *et al.*, 2006; Dimopoulou *et al.*, 2007; Heather *et al.*, 2008; Kung *et al.*, 2010). Historically the study of cargo genes and their transfer in *Enterococcus faecalis* (Franke & Clewell, 1981), *Bacteroides* species (Mays *et al.*, 1982; Rashtchian *et al.*, 1982), *Haemophilus influenza*e (Roberts & Smith, 1980; Stuy, 1980), *Streptococcus pneumoniae* (Shoemaker *et al.*, 1980), *Proteus rettgeri* (Nugent, 1981), and *Clostridium* species (Smith *et al.*, 1981; Magot, 1983) identified antibiotic and heavy metal resistance determinants that were transferred via conjugation. Subsequently, these genes were discovered to be located on the chromosome and not on plasmids as was assumed at the time. However, the first to be described as a genomic island capable of horizontal

transfer was the (now called) ICE*MlSym*^{R7A} in *M. loti* (Sullivan *et al.,* 1995). The term ICE was introduced in 2002 by Burrus *et al*..

The structure of essential functional modules that incorporate accessory and variable genes was illustrated for the ICE*SXT* in *V. cholerae* (Beaber *et al*., 2002; Wozniak *et al*., 2009). The comparison of various members of the SXT-R391 family revealed that the accessory genes are located in defined positions. These ICEs structure can be divided into backbone, hotspots (HS) and variable regions (VR). The backbone genes are the conserved core genes that mediate the ICE life cycles while the HS and VR represent variable regions across the different ICEs. The HS are regions where the accessory genes preferentially integrate while VR are regions of variability not considered true hotspots. For reasons not yet explained, the HS and the VR are located in specific positions in all the ICE*SXT* analyzed.

1.4.3 ICE evolution

ICEs are characterized by mosaicism - described as a puzzle of plasmid, transposons and phages (Wozniak & Waldor, 2010). ICEs are capable of promoting their own recombination and diversification via several mechanisms (Figure 1.5). For example, an ICE can insert in an *att* site nearby a mobile element (*Cis*-mobilizable element (CIME)), generating a construct with a total of three *att* sites that can be moved as a solo segment (Figure 1.5a). Two ICEs can insert into an *att* site next in tandem. This tandem can then recombine at regions of homology, removing the intervening sequence and forming a single, chimeric ICE (Figure 1.5b). Mobile elements can transpose or recombine into an ICE by either site-

specific recombination or homologous recombination, thus becoming part of, and so being transferred with, the ICE (Figure 1.5c). Additionally T-DNA from an external element can recombine into the oriT site on the ICE (Figure 1.5d) (Johnson & Grossman, 2015).

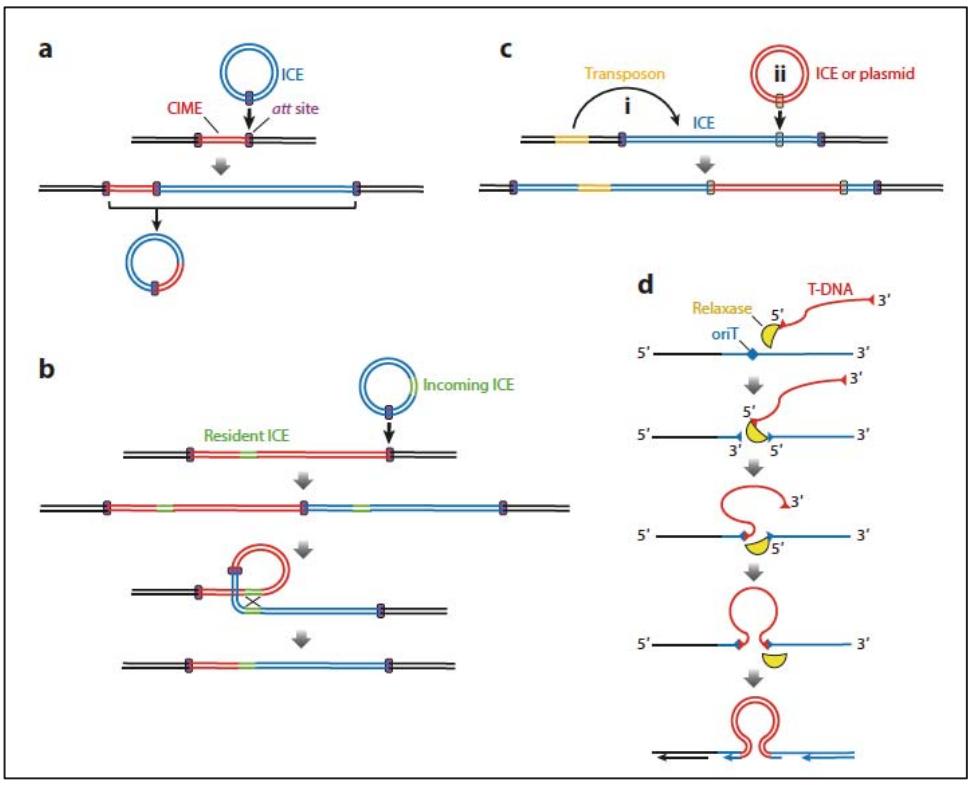

Figure 1.5. Mechanisms that generate ICE diversity. Diversity can be generated via recombination between ICEs or integration of mobile elements (reproduced from Johnson & Grossman, 2015).

To the best of my knowledge chimerism produced by the in-tandem integration of two ICEs has been described only for the SXT-R391 family (Hochhut *et al.*, 2001; Burrus & Waldor, 2004) and for ICEclc (Ravatn *et al.*, 1998). In the SXT-R391 family this inter-ICE recombination is RecA-dependent and promoted by homologous genes (*bet* and *exo*) of the phage lambda Red (Garris *et al.*, 2009).

1.4.4 ICEs in *Pseudomonas syringae*

P. syringae is a model organism for the study of plant-microbe interactions and microbial evolution due to its ubiquity in both agricultural environments, where different lineages are responsible for frequent outbreaks of disease in a variety of hosts; and in non-agricultural habitats, where it is found in asymptomatic association with wild plants, as well as in leaf litter, rivers, snowpack and even clouds (Morris *et al.*, 2007, 2008, 2010). The first ICE described in *P. syringae* was PPHGI-1 in *P. syringae* pv. *phaseolicola* (*Pph*) 1320A (Pitman *et al.*, 2005). In this case the discovery of this element was linked to the phenotypic change, given by the presence in the ICE of the effector gene *hopAR1*, when inoculated in bean cultivar Tendergreen (Pitman *et al.*, 2005). Comparison of the first two sequenced *P. syringae* genomes revealed the presence of PsyrGI-6 in *P. syringae* pv. *syringae* (*Psy*) B728a (Feil *et al.*, 2005). The pandemic of kiwifruit canker disease caused by *Psa* led to the identification of three additional ICEs called the Pacific ICE (Pac_ICE1), Mediterranean ICE (Pac_ICE2) and Andean ICE (Pac_ICE3) (McCann *et al.*, 2013; Butler *et al.*, 2013).

Experimental work has been carried out only for PPHGI-1 by Dawn Arnold's group (Pitman *et al.*, 2005; Lovell *et al.*, 2009; Godfrey *et al.*, 2010; Lovell *et al.*, 2011; Godfrey *et al.*, 2011; Neale *et al.*, 2016). PPHGI-1 (here called Pph_{1302A}ICE) harbors the HopAR1 effector that triggers the hypersensitive response (HR) on the Tendergreen (TG) variety of bean. Evolution experiments showed that repetitive passages on TG lead to an increase in virulence due to the down-regulation of the expression of *hopAR1* when Pph_{1302A}ICE excises from the genome. However the ICE is not completely lost from the population but persists in a small fraction until a favorable host is encountered upon which it increases in the bacterial population. They have also shown that the Pph_{1302A}ICE integrates into two specific sites into

the genome and that its transfer via conjugation and transformation increases *in planta*.

1.4.5 ICEs in *Pseusomonas syringae* pv. *actinidiae*

The first *Psa* strains to be sequenced revealed that the outbreak lineage independently acquired three different ICEs (Mazzaglia *et al.*, 2012; Butler *et al.*, 2013; McCann *et al.*, 2013). These ICEs share synteny but not nucleotide identity, apart from a transposon, Tn*6212*, which is identical. These three PsICEs were called Pacific ICE or Pac_ICE1, Mediterranean ICE or Pac_ICE2, Andean ICE or Pac_ICE3 by Butler *et al.* (2013) and McCann *et al.* (2013) and these authors identified the ~16 kb transposon Tn*6212* (Figure1.6).

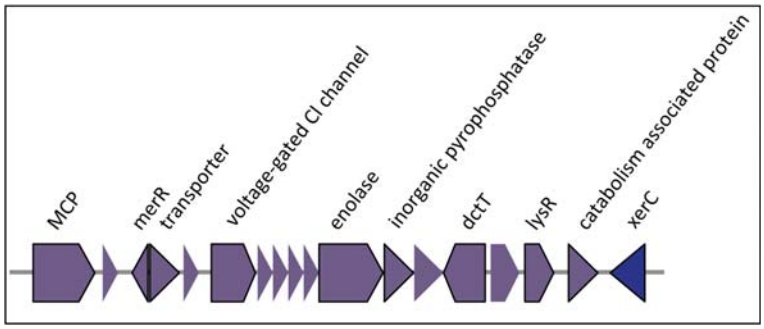

Figure 1.6. Genetic organization of Tn*6212*. Purple genes represent the "enolase region", blue box depicts a recombinase.

Tn*6212* consists of, firstly, a MCP (methyl-accepting chemotaxis) protein involved in taxis toward malate (McCann *et al.*, 2013). The following hypothetical protein (not annotated in Figure 1.6) belongs to a family of proteins containing a number of bacterial and eukaryotic proteins of unknown function that are approximately 200 residues long. Some of these family members' are annotated as putative lipoproteins. The MerR family of transcription regulators have been shown to mediate responses to stress including exposure to heavy metals, drugs,

or oxygen radicals in eubacterial and some archaeal species. They regulate transcription of multidrug/metal ion transporter genes and oxidative stress regulons by reconfiguring the spacer between the -35 and -10 promoter elements. A typical MerR regulator is comprised of two distinct domains that harbor the regulatory (effector-binding) site and the active (DNA-binding) site. Their N-terminal domains are homologous and contain a DNA-binding winged HTH motif, while the C-terminal domains are often dissimilar and bind specific coactivator molecules such as metal ions, drugs, and organic substrates (Brown *et al*., 2003). After the Mer transcriptional regulator there is a transporter with homologs in both prokaryotes and eukaryotes; this family is related to the human bile acid:sodium symporters but the gene also shows homologies with the arsenite efflux pump ACR3. In yeast, overexpression of the ACR3 gene confers an arsenite but not an arsenate resistance phenotype (Wysocki & Bobrowicz, 1997; Mukhopadhyay *et al*., 2002). The voltage-gated Cl channel belongs to a family of proteins that catalyse the selective flow of Cl^- ions across cell membranes and Cl^-/H^+ exchange (Iyer *et al*., 2000). Mutation of this gene (*sycA*) in *Rhizobium tropici* CIAT899 causes deficiencies in nodule development, nodulation competitiveness, and N^2 fixation on *Phaseolus vulgaris* plants, due to its reduced ability for acid resistance, but not for acid tolerance or growth under free-living conditions (Rojas-Jiménez *et al*., 2005). The inorganic phosphatase catalyzes the highly exergonic reaction that hydrolyzes inorganic pyrophosphate (PPi) to two molecules of orthophosphates (Pi) (Harold, 1996). The enolase is the enzyme responsible for the catalysis of the conversion of 2-phosphoglycerate (2-PG) to phosphoenolpyruvate (PEP) and is the ninth and penultimate step of glycolysis. Enolase is an enzyme involved in carbon metabolism, but also in RNA processing

and regulation (Taghbalout & Rothfield, 2007), in bacterial response to oxidative stress (Weng *et al*., 2016), and enolase-like proteins can bind human plasminogens in the outer membrane of *P. aeruginosa* (Ceremuga *et al*., 2004). DctT is a putative di-carboxylic acid transporter, with a sequence at its N-terminus predicted by SecretomeP version 2.0 to be used to target proteins for secretion (SecP score = 0.64). Analysis using EffectiveT3 to predict proteins targeted to the T3SS returned a highly significant probability score (0.97), strongly suggestive of type 3-targeting (McCann *et al*., 2013). In accordance with this 14% of the first 50 amino acids of the protein are either Ser or Pro (Schechter *et al*., 2004; Arnold *et al*., 2009). No other transporter in this region of accessory genes carries a similar signal (McCann *et al*., 2013). LysR belongs to a family of transcriptional regulators represents the most abundant type of transcriptional regulator in the prokaryotic kingdom. LysR-type transcriptional regulators regulate a diverse set of genes, including those involved in virulence, metabolism, quorum sensing and motility (Maddocks & Oyston, 2008). The catabolism-associated protein is a PQQ-dependent catabolism-associated beta-propeller protein. XerC is tyrosine recombinase. These recombinases are involved in the site-specific integration and excision of lysogenic bacteriophage genomes, transposition of conjugative transposons, termination of chromosomal replication, and stable plasmid inheritance (Doungudomdacha *et al*., 2007).

Butler *et al*. (2013) designated this region as an autonomously mobile XerC transposon (Tn*6212*) surrounded by an (imperfect) 14bp direct repeat (AAATACGTTATCAC). The authors also refer to the transporter gene as a mercury resistance gene. In McCann *et al*. (2013) it was named “enolase region” and it was predicted to be involved in the virulence of *Psa*, specifically with a role in the

manipulation of the host cell metabolism. The DctT transporter was thought to be exported via T3SS due to the presence at its N-terminus of a region predicted to be used to target proteins for exportation (by SecretomeP version 2.0) and to be targeted by the T3SS (by EffectiveT3). The putative T3SS-targeted DctT, was hypothesized to enter the plant cell and to be incorporated into host cell membranes to facilitate export of this sugar acid in a manner analogous to that suggested for α-ketoglutarate in *Xanthomonas oryzae* pv. *oryzae* (Guo *et al.*, 2012). The enolase was thought to be involved in the conversion of dicarboxylic acids to glucose or to enhance activity of glycolysis in plant cells (McCann *et al.*, 2013).

1.5. Copper resistance

1.5.1 Resistance

Resistance is the inherited ability to grow at high concentration of a drug, independent of the duration of the treatment, and it is quantified by the minimal inhibitory concentration (MIC) (Brauner *et al.*, 2016). Resistance can be gained via mutation or HGT. The most common mutations lead to the alteration of the drug target or to the increase of its efflux (Nikaido, 2009), but resistance can also be associated with gene amplification (Andersson & Hughes, 2009; Sandegren & Andersson, 2009), reduction in expression of the target (Ince & Hooper, 2003), or alteration of drug modification enzymes (Robicsek *et al.*, 2006). HGT, however, is more likely associated with mechanisms involved in drug modification, target protection, replacement of the susceptible target or acquisition of efflux pumps (Li & Nikaido, 2009).

The acquisition of genes or mutation allowing a bacterium to survive a killing agent is of strong selective benefit and thus in the presence of the drug the acquisition and maintenance of resistance will be strongly favoured. Even a low quantity of the drug can select for resistant populations - it has been shown that the concentrations necessary to maintain resistance plasmids within a population are well below the MIC of the non-plasmid containing susceptible strain (Gullberg *et al*. 2014). Resistance determinants can also be co-selected via the close linkage between two or more different resistance genes. When different resistance genes are spatially closely linked (for example their association in the same mobile element), the presence of one drug can drive the selection of the other resistance gene (Gullberg *et al*., 2014). For example the occurrence of both antibiotic and metal resistance has been noted in almost half of the 5,436 bacterial genomes analyzed by Li *et al*. (2017).

Resistance, however, can be associated with reduced bacterial fitness, and this can lead to its loss. The time required to reduce the abundance of resistant bacteria is inversely related to the cost of resistance (Spratt, 1996; Levin, 2001), this phenomenon is called reversibility. It has been proposed that a reduction in antibiotic use (and therefore a reduction in the selective pressure to acquire resistance) would benefit the fitter susceptible bacterial population, enabling them to outcompete resistant bacteria over time (Levin *et al*., 1997; Andersson & Levin, 1999). This concept is supported by experimental studies and theoretical modelling (Levin *et al*., 1997; Levin, 2002), but other processes such us compensatory evolution and co-selection complicate this picture and make reversibility less probable in a real-life setting. Reversibility, if it occurs at all, is a

process so slow that in most cases it is unlikely to be of practical importance (Andersson & Hughes, 2011).

1.5.2 Copper use in agriculture

Copper is an essential trace element for all organisms living under aerobic conditions, fundamental for cellular processes such as oxidative phosphorylation, photosynthesis, and free radical control (Lee *et al.*, 2002). Despite its essential role in biological processes, it is highly toxic in excess concentration: it degrades iron-sulphur clusters of dehydratases via iron displacement, inactivating these enzymes. The iron released may subsequently initiate the Fenton reaction, which leads to oxidative damage by the generation of reactive oxygen species (ROS) (Macomber & Imlay, 2009). Indiscriminate copper binding may also lead to damaging consequences to protein structure, disrupting their biological functions (Letelier *et al.*, 2005).

For these properties copper has been utilized in medicine as well as in agriculture. Humans have employed copper for tools, disinfection, plumbing, manufacturing, animal husbandry, crop protection and preservation of perishable commodities (Hobman & Crossman, 2015). Copper compounds have their most extensive employment in agriculture where the first recorded use was in 1761, when it was discovered that seed grains soaked in a weak solution of copper sulphate inhibited seed-borne fungi. More than 100 years later it was discovered that copper sulphate also prevented foliage diseases (Millardet, 1885). In 2014, 18,500 thousands of tons of copper was produced worldwide ("pg. 55 – Copper" USGS. 2016). Today probably about 200,000 tons of copper sulphate is used

globally every year in agriculture (www.copper.org). From the beginning of its production ~7,000 years ago total emissions of Cu in the atmosphere were ~3.4 million tons (Hong *et al*., 1996).

1.5.3 Copper resistance in bacteria

Copper homeostasis requires a delicate balance between providing essential nutrients and preventing lethal excess (Puig *et al*., 2002). In bacteria, a number of processes are used to tune this balance, such as intra and extra-cellular sequestration, enzymatic detoxification, and active efflux (Bondarczuk & Piotrowska-Seget, 2013). The widespread use of copper by humans has led to the diversification and selection of copper resistance islands (CHASRI). Current CHASRI diversity is predicted to have originated in a relative of *Enterobacter cloacae* around 6,000 years ago, which corresponds to current estimates of the beginning of Bronze Age, when anthropogenic copper became a factor of environmental stress. Furthermore, three sharp accelerations observed in CHASRI diversification could correspond to three major peaks of copper production (Roman Empire, Chinese Sung Dynasty and post Industrial Revolution). Diversification among livestock-associated CHASRIs is reconstructed to be within the last century, while human-associated diversifications are both ancient and modern (Staehlin *et al*., 2016). Copper resistance genes have been isolated in a variety of environments such as soils, oceans, the phyllosphere, humans and animal guts (Bender & Cooksey, 1986; Sabry *et al*., 1997; Hassen *et al*., 1998). The earliest isolated CHARSIs date to 1890, in *E. cloacae* isolated from the spinal fluid of a patient affected by meningitis (Staehlin *et al*., 2016).

Copper resistance is typically conferred by operons encoding either copper efflux (*cusCFBA* or the *cue* system) and / or sequestration (*copABCD* or *pcoABCDSE*) systems. *cus* and *cop* systems can be under the regulation of the *copRS* / *cusRS* two-component regulatory system (Bondarczuk & Piotrowska-Seget, 2013). CusCFBA was found in *E. coli* and is a tripartite system conferring resistance to copper and silver ions (Grass & Rensing 2001; Franke *et al*., 2001, 2003). CusCBA is an efflux system constituted by a cytoplasmic membrane secondary transporter enabling protein-driven substrate translocation from the cytoplasm (CusA), a periplasmatic membrane fusion protein (CusB) and an outer membrane channels exporting substrate outside the cell (CusC) (Kim *et al*., 2010). CusF is a chaperone required for full resistance (Figure 1.7).

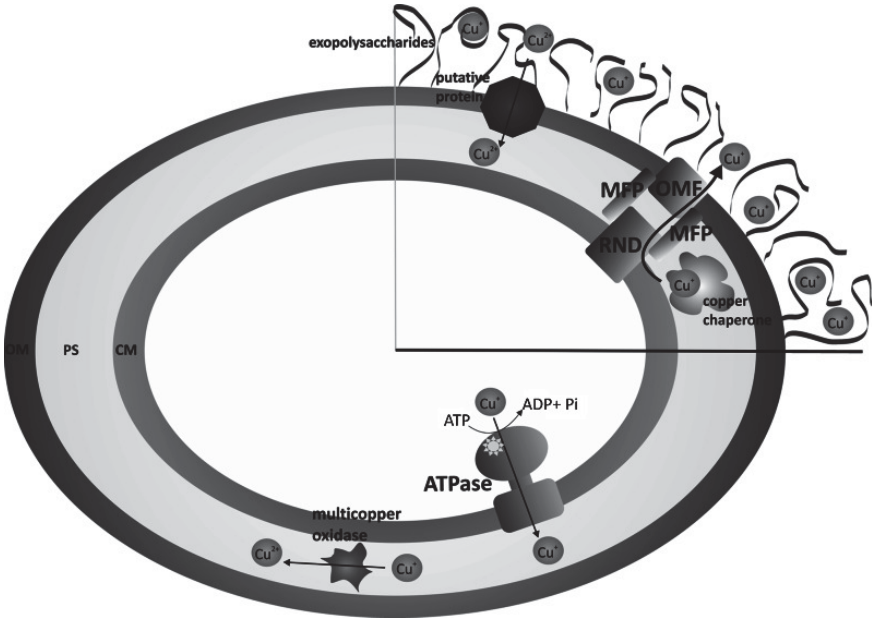

Figure 1.7. Cus System. CM cytoplasmic membrane, PS periplasmic space, OM outer membrane. Assignment of particular proteins: multicopper oxidase (CueO), RND (CusA), MFP (CusB), OMF (CusC), copper chaperone (CusF) (reproduced from Bondarczuk and Piotrowska-Seget, 2013).

CopA is a multicopper oxidase with high copper-binding capacity. CopB is an outer membrane protein containing many methionine residues predicted to bind copper, but this ability hasn't been proven so far (Arnesano *et al*., 2002; Puig *et al*., 2002; Zhang *et al*., 2006). CopC is a soluble periplasmatic chaperone with

two independent binding sites with specific affinity for Cu(I) and Cu(II). It has been proposed that CopC can interact with CopA, CopB, CopD and CopS. CopC can then deliver the copper ions accumulated in the periplasmatic space by CopA to CopD, which is a transmembrane protein that is responsible (with CopC) for copper delivery and accumulation in the cytoplasm through the inner membrane (Figure 1.8) (Cooksey *et al.*, 1993; Arnesano *et al.*, 2002; Puig *et al.*, 2002; Bondarczuk & Piotrowska-Seget, 2013).

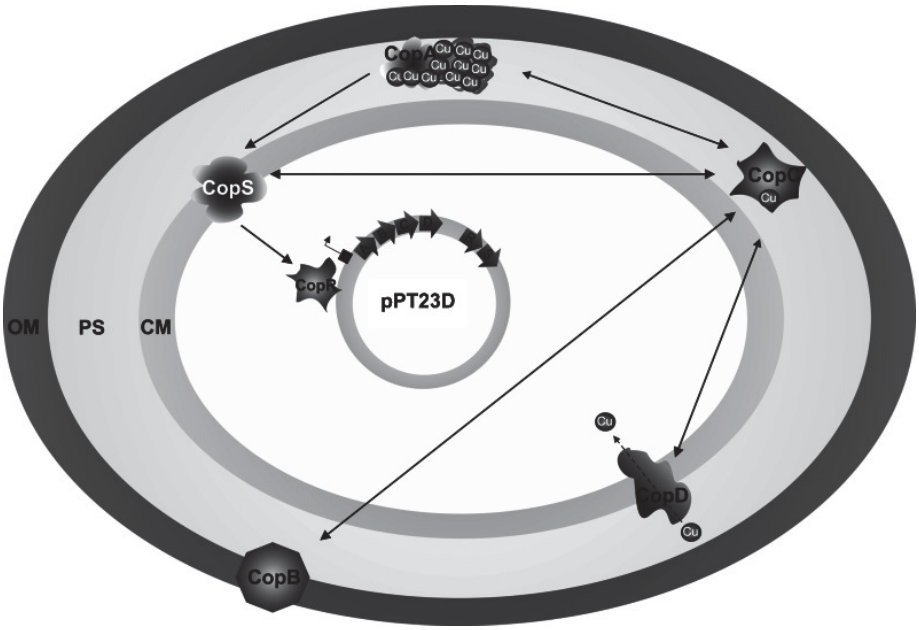

Figure 1.8. A model for cop system in *Pseudomonas syringae* pv. *tomato.* Abbreviations: CM cytoplasmic membrane, PS periplasmic space, OM outer membrane. Arrows indicate presumed interactions between proteins (reproduced from Bondarczuk & Piotrowska-Seget, 2013).

1.5.4 Copper resistance in *Pseudomonas syringae*

Copper resistance has been detected for different pathovars of *P. syringae* and is frequently encoded by plasmids (Cazorla *et al.*, 2002; Hwang *et al.*, 2005; Gutiérrez-Barranquero *et al.*, 2013). However, additional studies have shown that *P. syringae* may harbor other variants of copper resistance (CuR) determinants

(Sundin & Bender, 1987; Rogers *et al.*, 1994; Cazorla *et al.*, 2002). Two *Psy* strains isolated from mango and cherry trees carrying the UMAF0081 or the 6-9 plasmid respectively have been shown to have a MIC for $CuSO_4$ on MGY plates of 1.8 mM, while *Psy* B728a shows a MIC of 1.6 mM (Gutiérrez-Barranquero *et al.*, 2013). In *Psy* B728a copper resistance genes are located in PsyrGI-6. In a survey conducted on *Psy* isolated from mango trees in Sicily, the authors divided the degrees of resistance assessed as MIC for $CuSO_4$ in MGY plates as follows: 0.4-0.8 mM as not resistance, 1-1.6 mM as low resistant, 1.8-2.4 mM as resistance and 2.6-3.2 mM as highly resistant. In this study, among the 71 strains tested, 44 were highly resistant, 16 resistant, and 11 low resistant (no sensitive strains were detected) (Aiello *et al.*, 2015).

In *Psa*, copper resistance has been described only for some Japanese *Psa*-1 lineage isolates from 1980s. CuR determinants were gained via the acquisition of one of both of the plasmids pPaCu1 and pPaCu2, conferring a MIC on PDA plates between 1.75 and 3 mM (Nakajima *et al.*, 2002).

1.6 Research objectives

Psa emerged as kiwifruit pathogen in the 80s in Asia, although the disease was severe it remained confined to this region. From 2008 a particularly virulent strains spread worldwide. First genome sequences revealed that the disease was caused by a single *Psa* clade. *Psa* strains isolated in Italy, New Zealand and Chile were almost identical apart from the independent acquisition of three ICEs. These ICEs shared less than 70% nucleotide identity but carried an identical set of genes (Tn*6212*) that was hypothesized to be involved in virulence. In 2015 I discovered that two *Psa* strains isolated in New Zealand independently acquired two ICEs encoding for copper resistance.

The aims of my thesis are:

- analyze the distribution of the ICEs orthologous to the ones previously isolated in *Psa* and in other *P. syringae* (Chapter 3);

- determine the structure of the PsICEs and infer their evolution history (Chapter 3);

- determine the contribution of Tn*6212* to *Psa* virulence (Chapter 4);

- analyze the evolution of copper resistance in *Psa* in New Zealand, determine the cost and benefits of the acquisition of Psa_{NZ45}ICE_Cu and analyze the dynamics of transfer of Psa_{NZ45}ICE_Cu, member of the PsICE family (Chapter 5).

Chapter 2 - Materials and Methods

2.1 Materials

2.1.1 Strain list

All strains used in this study are listed in Table 2.1.

Strain ID	Genotype or description	Reference
Escherichia coli		
E. coli DH5α	*sup*E, ΔlacU169 (Φ80 lacZΔM15), *rec*A, *end*A, *hsd*R, *gyr*A, *thi*, *rel*A, (oriR6K replication)	Invitrogen
E. coli TOP10	F', mcrA, Δ (mrr-*hsd* RMS-*mcr*BC), Φ80 lacZΔM15, Δ*lac* X74, *deo*R, *ara*D139, Δ(*ara-leu*) 7697, *gal*U, *gal*K, *rps*L, StrR, *end*A1, *nup*G	Invitrogen
Pseudomonas		
Psa NZ13	Wildtype *Psa*-3 isolated in New Zealand	McCann *et al.*, 2013
Psa NZ13^{KmR}	Km resistant *Psa* NZ13 featuring a chromosomical integration of Tn*5*	This study
Psa NZ13 Δ*enoR*	*Psa* NZ13 featuring the enolase region in-frame deletion	This study
Psa NZ13 Δ*hrcC*	*Psa* NZ13 featuring the *hrc* gene in-frame deletion	This study
Psa NZ13 Δ*enoG*	*Psa* NZ13 featuring the *enolase* gene in-frame deletion	This study
Psa NZ13 Δ*dctT*	*Psa* NZ13 featuring the *dctT* gene in-frame deletion	This study
Psa NZ13 Δ*3805*	*Psa* NZ13 featuring the IYO_003805 gene in-frame deletion	This study
Psa NZ13 Δ*3790*	*Psa* NZ13 featuring the IYO_003790 gene in-frame deletion	This study
Psa NZ13 Δ*3805-3790*	*Psa* NZ13 featuring the IYO_003805 and IYO_003790 genes double in-frame deletion	This study
Psa NZ45	Copper-resistant wildtype *Psa*-3 isolated in New Zealand	This study
Psa NZ47	Copper-resistant wildtype *Psa*-3 isolated in New Zealand	This study
Psa NZ62	Copper-resistant wildtype *Psa*-3 isolated in New Zealand	This study
Psa NZ63	Copper-resistant wildtype *Psa*-3 isolated in New Zealand	This study
Psa NZ64	Copper-resistant wildtype *Psa*-3 isolated in New Zealand	This study
Psa NZ65	Copper-resistant wildtype *Psa*-3 isolated in New Zealand	This study
Psa NZ66	Copper-resistant wildtype *Psa*-3 isolated in New Zealand	This study
Psa K28	Wildtype *Psa*-2 isolated in South Korea	McCann *et al.*, 2013
Psa J31	Wildtype *Psa*-1 isolated in Japan	McCann *et al.*, 2013

Psa NZ13 avrRpt-2	*Psa* NZ13 featuring the pavrRpt-2 plasmid	This study
Psa NZ13 pMT1	*Psa* NZ13 featuring the pMT-1 plasmid	This study
Psa NZ13 pMT2	*Psa* NZ13 featuring the pMT-2 plasmid	This study
Psa NZ13 Δ*hrcC* avrRpt2	*Psa* NZ13 Δ*hrcC* featuring the pavrRpt-2 plasmid	This study
Psa NZ13 Δ*hrcC* pMT1	*Psa* NZ13 Δ*hrcC* featuring the pMT-1 plasmid	This study
Psa NZ13 Δ*hrcC* pMT2	*Psa* NZ13 Δ*hrcC* featuring the pMT-2 plasmid	This study
Pfm NZ5	Wildtype *P. syringae pv. actinidifoliorum* isolated in New Zealand	McCann *et al.*, 2013
Pfm NZ9	Wildtype *P. syringae pv. actinidifoliorum* isolated in New Zealand	McCann *et al.*, 2013
Pfm NZ9^{KmR}	*Pfm* NZ9 featuring a chromosomic integration of Tn5	This study
Pto DC3000	Wildtype *Pseudomonas syringae* pv. *tomato* strain DC3000	Buell *et al.*, 2003
Pto DC3000 ^{KmR}	*Pseudomonas syringae* pv. *tomato* DC3000 featuring a chromosomic integration of Tn*5*	This study
Pph 1448a	Wildtype *Pseudomonas syringae* pv. *phaseolicola* strain 1448a	Teverson, 1991
Pph 1448a ^{KmR}	*Pseudomonas syringae* pv. *phaseolicola* 1448a featuring a chromosomic integration of Tn*5*	This study
Ps H24	*P. syiringae* strain H24 strain isolated from kiwifruit in New Zealand	Straub, 2017
Ps H24 ^{KmR}	*P. syiringae* strain H24 strain featuring a chromosomic integration of Tn*5*	This study
Ps H33	*P. syiringae* strain H33 strain isolated from kiwifruit in New Zealand	Straub, 2017
Ps H33 ^{KmR}	*P. syiringae* strain H33 strain featuring a chromosomic integration of Tn*5*	This study
P. fluorescens SBW25	*Pseudomonas fluorescens* strain SBW25	Rainey & Bailey (1996)
P. fluorescens SBW25^{KmR}	Wildtype *Pseudomonas fluorescens* strain SBW25 featuring a chromosomic integration of Tn*5*	This study
P. aeruginosa PAO1	Wildtype *Pseudomonas aeruginosa* strain PAO1	Holloway, 1955
Agrobacterium		
A. tumefaciens AGL1	AGL0 (C58 pTiBo542), recA::bla, T-region deleted ,Mop(+) Cb(R)	Lazo *et al.*, 1991

Table 2.1 Strains used in this study.

2.1.2 Plasmids and transposons

All plasmids and transposons used in this study are listed in Table 2.2.

Name	Characteristic	Reference
Transposon		
IS-Ω-Km/hah	ColE1, *nptII* promoter, IE, OE, *Lox*P, KmR	Giddens *et al.*, 2007
Plasmids		
pGEM®-T	high copy number cloning vector: AmpR, pBR322 ori	Promega
pME6010	Shuttle vector, TcR	Heeb *at al.*, 2000
pKR2013	Helper plasmis for tri-parental mating. KmR, incP4, tra, mob.	Ditta *et al.*, 1980
pK18mobsacB	Mobilizable vector sacB, lacZa, KmR, mcs	Schäfer *et al.*, 1994
pMT1	pUCP22 (GmR), promoter and the putative exportation site (1-52 aa) of dctT fused to avrRpt2Δ1-79aa	GenScript®
pMT2	pUCP22 (GmR), promoter and full leght dctT fused to avrRpt2Δ1-79aa	GenScript®
pavrRpt2	pVSP61 with full length avrRpt2	Kunkel *et al.*, 1993

pICH86988	binary vector,35S, KanR	Kee Sohn Lab
pICSL30005	Plasmid carryin N-terminal epitope tag N-3xFLAG, SpecR	Kee Sohn Lab
pICH86988_RIN4::Flag	Plasmid with RIN4 N-terminal epitope tag N-3xFLAG, KmR	Kee Sohn Lab
pICH86988_RPS2::Myc	Plasmid carryin RPS2 N-terminal epitope tag N-4xMyc, KanR	Kee Sohn Lab

Table 2.2. Plasmids used in this study.

2.1.3 Primers

All primers (Table 2.3) were manufactured by Integrated DNA technologies. Primers were resuspended in deionized water to a stock solution of 100 µM, and diluted to a working concentration of 10 µM. Primers were stored at -20°C prior to use and thawed on ice.

Primer name	Primer sequence (5'→ 3')	Target region
Primers used for transposon mutagenesis		
TnphoA-II	GTGCAGTAATATCGCCCTGAGCA	IS-Ω-Km/hah
CEKG 2A	GGCCACGCGTCGACTAGTACNNNNNNNNNNAGAG	Non-specific
CEKG 2B	GGCCACGCGTCGACTAGTACNNNNNNNNNNACGCC	Non-specific
CEKG 2C	GGCCACGCGTCGACTAGTACNNNNNNNNNNGATAT	Non-specific
Hah-1	ATCCCCCTGGATGGAAAACGG	5' end of CEKG 2A, B, & C
CEKG 4	GGCCACGCGTCGACTAGTAC	IS-Ω-Km/hah
Primers used for cloning		
M13 reverse	CAGGAAACAGCTATGAC	pGEM®-T, pK18mobsacB
M13(-21) forward	TGTAAAACGACGGCCAGT	pGEM®-T, pK18mobsacB
Primers used in the construction of *Psa* in-frame deletion mutants		
hrcC_cross_for	AGTTTAAGTTTAGTGAGCTCGGATTAAGGCCTTGCGCATTT	*hrcC*
hrcC_cross_rev	CCGAGCTCACTAAACTTAAACTAAAATGGCCAAAGAAAGCAG	*hrcC*
hrcC_for_XbaI	TTTTCTAGAACTTGGCTTATGCGGTCAGT	*hrcC*
hrcC_rev_BamHI	TTTGGATCCCTCACACCAACATGGACCAG	*hrcC*
Ext_hrcC_for	TATCAGGGAGGGGTGTATCG	*hrcC*
Ext_hrcC_rev	GTTTGAACCCTTCTGCTGGA	*hrcC*
enoR_XbaI_for	TTTTCTAGACAGGAATCTCTCCCACCAAG	Tn*6212*
enoR_cross_rev	CCGATATCACTAAACTTAAACTTTTTGGTCGCTTTTTCTGG	Tn*6212*
enoR_cross_for	GTTTAAGTTTAGTGATATCGGACGTATACCGGTAGCGACGA	Tn*6212*
enoR_EcoRI_rev	TTTGAATTCTAACGGCCATAACCGACAAT	Tn*6212*
Ext_enoR_for	GGTGTCGGCATGTAGAAAGG	Tn*6212*
Ext_enoR_rev	CCTTGTGCCAGAATCCATCT	Tn*6212*
dctT_cross_for	GGTTTAAGTTTAGTAAGCTTGGGCCCATTATAGCTGCGAGTC	*dctT*
dctT_cross_rev	CCAAGCTTACTAAACTTAAACCCCTGAGCCTTTGGCTACTTG	*dctT*
dctT_for_EcoRI	TTGAATTCGATCTTCGTTGATCGCTTCC	*dctT*

dctT_rev_BamHI	TTGGATCCAGCTACTCCTGGCTCACGAA	*dctT*
Ext_dctT_for	AATCAAATCGGAACGCTGAC	*dctT*
Ext_dctT_rev	TTGTGAATATCGCCGTCAAA	*dctT*
enoG_cross_for	GGTTTAAGTTTAGTAAGCTTGCCGGTAAACCTGTTGGACTG	*enolase*
enoG_cross_rev	CAAGCTTACTAAACTTAAACTCCTGGATTTGCGTTTTCATT	*enolase*
enoG_EcoRI_for	AAGAATTCGCGGTACAGAGGGGCTAGAT	*enolase*
enoG_XmaI_rev	AACCCGGGCAGGGAAAAGCGTCAGTAGC	*enolase*
Ext_enoG_for	ACGGCTTTCAACTGACGTTT	*enolase*
Ext_enoG_rev	TGACGCAGTGTGGACAAAAT	*enolase*
3805_BamHI_for	TTGGATCCTGCTGGACTCGTCAGTTTTG	*Cluster 3*
3805_cross_for	GGTTTAAGTTTAGTCCCGGGGCATGGATTGCAGACTTCGTG	*Cluster 3*
3805_cross_rev	CCCCGGGACTAAACTTAAACTTGAACCGATACCCATTCGAT	*Cluster 3*
3805_SphI_rev	AAGCATGCCGTGCATGATTCAGATCACC	*Cluster 3*
Ext_3805_for	CATCATGGCTAGGCAAAACA	*Cluster 3*
Ext_3805_rev	TTTTGCAAGGTGCTCAACTG	*Cluster 3*
3790_BamHI_for	TTGGATCCCGTATGTACGGCCTTCGTTT	*Cluster 3*
3790_cross_for	GGTTTAAGTTTAGTCCCGGGGAGGCCAACACCATTCTCAAG	*Cluster 3*
3790_cross_rev	CCCCGGGACTAAACTTAAACTACTGGCGGATGAACAATGAT	*Cluster 3*
3790_SphI_rev	AAGCATGCCAAAACTGACGAGTCCAGCA	*Cluster 3*
Ext_3790_for	CCACATGGTGGTCAGATCAA	*Cluster 3*
Ext_3790_rev	CAAGATCGAAACCACGGAGT	*Cluster 3*
Primers used in the identification and location of Psa_{NZ13}ICE and Psa_{NZ45}ICE_Cu		
att-1	TGTAGAATAGCGCGCCTCAG	*att-1*
att-2	AGCCGTAATCCTGCTGTCC	*att-2*
copA for	ATCCGCGGTGACTCGATAAC	Psa_{NZ45}ICE_Cu
copA rev	CAGTCGATGGACCGTACTGG	Psa_{NZ45}ICE_Cu
enolase for	GAGCTGACGTCCGACATAGAG	Psa_{NZ13}ICE
enolase rev	CCAGTCCAACAGGTTTACCG	Psa_{NZ13}ICE
IntPsaNZ13	GTCAGGCTGATCACTTACGTTG	Psa_{NZ13}ICE
IntPsaNZ45	GTCAGGCTGATCACTAGCGTTA	Psa_{NZ45}ICE_Cu
Primers used in the construction of pICH86988_MT1 and MT2		
pMT_BsaI_for	GGTCTCgAATGGGCGGCAGCCATCCACTGG	pMT-1 and pMT-2
pMT_BsaI_rev	GGTCTCaCGAAGCGGTAGAGCATTGCGTGTGGAACCTG	pMT-1 and pMT-2

Table 2.3. Primers used in this study.

2.1.4 Media and growth conditions

Media used for *Psa* and other *Pseudomonas* strains are: King's medium B (KB) (King *et al.*, 1954), M9 minimal medium (M9) (Harwood & Cutting, 1990), mannitol-glutamate yeast extract medium (MGY) (Bender & Cooksey, 1986) and

Lysogeny Broth (LB) (Bertani, 1951). LB was also used for *Escherichia coli* and *Agrobacterium tumefaciens*. KB broth contained: 10 mL glycerol, 20 g Proteosone Peptone No.3 (for liquid media) or tryptone (for solid media), 1.5 g $MgSO_4.7H_2O$, 1.5 g $K_2PO_4.3H_2O$ per L. LB broth contained: 10 g tryptone, 5 g yeast extract, 10 g NaCl per L. M9 media contained: 12.8 g $NaHPO_4$, 3 g KH_2PO_4, 0.5 NaCl, 1 g NH_4Cl, 0.1 mM $CaCl_2.6H_2O$, 2mM $MgSO_4.7H_2O$ per L. MGY media contained: 10 g D-mannitol, 2 g L-glutamic acid monosodium salt, 0.5 g KH_2PO_4, 0.2 NaCl, 0.2 $MgSO_4.7H_2O$, 0.25 yeast extract per L. To prepare solid media, bacteriological agar was added to the media 15 g/L. Media were sterilized in an autoclave at 120°C for 20 minutes in 500 mL Pyrex glass Durans. Sterilized medium with agar was cooled to 55°C before pouring into sterile Petri plates. Plates were allowed to dry on the bench for 24 hours.

Single bacterial colonies were obtained by streaking on agar plates from frozen glycerol stocks, after incubation at 28°C for 24-48 hours (*Psa* and other *Pseudomonas*) or 37°C for 24 hours (*E. coli*). To set up broth cultures, a single colony was picked from a fresh agar culture with a sterile loop and aseptically mixed into the broth prior to incubation. Inoculation plates for *in planta* experiments were set up by heavily streaking onto KB agar plates from frozen stocks. *Pseudomonas* and *A. tumefaciens* strains were grown at 28°C, *E. coli* strains at 37°C.

Glycerol stocks were prepared by mixing 750 µL of a liquid overnight (ON) culture with an equal volume of sterile 50% glycerol solution and storing at -80°C.

2.1.5 Antibiotics and reagents

Antibiotics were purchased from Melford Laboratories and used at the following concentrations: ampicillin (Amp) 100 μg/mL; gentamycin (Gm) 10 μg/mL; kanamycin (Km) 50 μg/mL, spectinomycin (Spe) 50 μg/mL, tetracycline (Tc) 15 μg/mL, nitrofurantoin (Nf) 50 μg/mL dissolved in dimethyl sulfoxide (DMSO). X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) was used at 60 μg/mL in agar plates as indicator of β-galactosidase activity.

2.1.6 Gel electrophoresis

DNA was separated in a 1% agarose gel in a horizontal gel tank containing 1X TAE or TBE. SYBR® Safe DNA (Invitrogen) was added to the gels at a final concentration of 0.1 mg/mL, and loading dye was added to PCR or restriction digests. Gels were run at a constant voltage of 80-120 mV for an appropriate time in SUB DNA cell (BIO-RAD), and bands were visualized under UV transilluminator or a Safe Imager™ 2.0 Blue-Light Transilluminator (Invitrogen). GeneRuler 1 kb DNA Ladder (Thermo Scientific) was used as a size marker.

2.2 Methods

2.2.1 Standard polymerase chain reactions (PCRs)

PCRs were performed using a T100™ Thermal Cycler (Biorad). Standard PCRs were performed using Taq DNA polymerase (Invitrogen). All standard reactions for electrophoresis contained 1 unit of Taq polymerase, 1.25 μL of 10 μM primers, 2.5 μL of 10x PCR buffer, 0.75 μL of 50 mM $MgCl_2$, 0.5 μL of 10 nM dNTP

mix, and were then adjusted to 25 µL with RNAse and DNAse free water. Cells were taken directly from colonies to be used as templates for the reaction. The reaction conditions were an initial denaturation step at 95°C for 2 minutes, followed by 30 seconds denaturation at 95°C, annealing between 55-62°C (depending on primer sequence) for 30 seconds, and extension at 72°C for 60 seconds per kilobase (kb) of target DNA amplified. This was repeated 35 times, followed by a final extension at 72°C for 10 minutes. A list of primers used in this study is shown in Table 2.3.

2.2.2 Overlap extension PCR

Overlap extension PCR (Ho *et al.*, 1989) was used to generate all the in-frame deletion mutants produced in this study. This technique involved the production of a "destruction box" consisting of the up and down stream DNA sequence of the target site to be deleted. The "destruction box" was built by PCR amplifying and PCR joining the sequence upstream and downstream of the target region. In order to minimize mutations in these products, Phusion® High-Fidelity DNA polymerase (New England Biolabs (NEB)) was used in PCR reactions containing 10 µL of 5x reaction buffer, 1.25 µL of 10nM dNTP, 5 µL of 10mM primers, 0.5 µL Phusion® High-Fidelity DNA polymerase and RNAse and DNAse free water to react the final volume of 50 µL. The two regions were amplified using primers (Table 2.3) containing two different features: the sequence for the restriction sites present in the multicloning site of pK18mobsacB was added to the external primers to insert the PCR amplicon into the vector and 21 bp were added to the 5' end of the central primers. Cycling conditions were standard for Phusion® High-Fidelity DNA polymerase: an initial denaturation was performed at

98°C for 30 seconds, with 35 cycles of amplification involving annealing between 60-62°C (depending on primer sequence) for 30 seconds, and extension at 72°C for 30 seconds per kb of target DNA amplified. This was followed by a final elongation step at 72°C for 5 minutes. The resulting PCR product was then verified for size using gel electrophoresis. A second round of PCR was then performed: the downstream and upstream amplicons were then used as templates for the overlap extension PCR experiments, using the external primers. The overlap extension PCR is divided into two steps: the first step is the annealing of the two amplicons, and the second step is the production of the construct used for the deletion. 20 ng/µL of the two PCR products of the previous PCR are used as templates. The concentrations of the reagents were maintained, except for the primers that are added only in the second step of the overlap extension PCR. This first step consists of a series of 10 cycles, which enable the annealing of the two amplicons via the complementarity of the 21 bp added to the primers. This was performed at the temperature standard for the Phusion® High-Fidelity DNA polymerase. In the second step the external primers are added to the PCR mix and the PCR carried out with a standard protocol for 25 cycles. At the end of this step the "destruction box", consisting of the fragment of the target region without the central part, is obtained. The resulting PCR product was then verified for size using gel electrophoresis. The products were gel-extracted and an extra round of elongation was performed to ensure that the "destruction box" contained an adenine 3' overhang required to the TOPO®. To perform the elongation, 17.4 µL of PCR product, 0.4 µL of 10 nM dATP, 0.2 µL of standard Taq polymerase and 2 µL of 10x reaction buffer were mixed and the reaction incubated at 72°C for 15 minutes. The resulting PCR product was stored at 4°C prior to further cloning.

2.2.3 Plasmid and PCR product purification

Plasmid DNA was extracted from overnight cultures of *E. coli* using the E.Z.N.A.® Plasmid Mini Kit 1 (Omega Bio-Tek) according to manufacture's instruction and stored in RNAse and DNAse free water at -20°C. PCR products were either purified via gel extraction using the QIAquick® Gel Extraction Kit (Qiagen) according to manufacture's instruction or via enzymatic reaction, and stored in RNAse and DNAse free water at -20°C. For the enzymatic purification, 25 µL PCR products were treated by adding 2 units of Exonuclease I (NEB) and 2 units of Calf Intestinal Phosphatase (NEB) and incubated at 37°C for 30 minutes followed by 85°C for 15 minutes.

2.2.4 Sanger sequencing

PCR products and plasmids were purified as per Section 2.2.3. Purified products were sequenced by Macrogen Inc (Seoul, South Korea) and the resulting chromatogram sequence aligned to the wildtype sequence of the *Psa* NZ13 genome using Genious version 7.

2.2.5 Genome extraction and sequencing

Genome extractions were performed using the Promega Wizard® Genomic DNA Purification Kit following the recommended protocol. *Psa* strains were sequenced either using the PacBio platform (Sydney, Australia) or the Illumina HiSeq platform (Plön, Germany).

2.2.6 Restriction digestions and ligations

All restriction enzymes and enzymatic buffers were purchased from Thermo Scientific. PCR products were purified from gels and plasmids extracted from *E. coli* prior to enzymatic digestion. DNA was restricted accordingly to manufacturer's instructions. Unless otherwise specified, the digested DNA was directly purified from the reaction or gel-purified using the SureClean Plus (Bioline) kit or the QIAquick Gel Extraction Kit (Qiagen), respectively. Ligations were performed using T4 ligase (Invitrogen) with an excess of insert to vector (from 3:1 to 10:1) (Table 2.4). Reactions were left overnight at 4°C. The morning after 0.5 µL of T4 ligase was added to the mix and left at room temperature for 1 hour. The composition of the ligation mix is reported in Table 2.4.

Reagents	Stock solution	Final concentration
Ligation buffer	5X	1X
Vector	X	50 ng/reaction
Insert	X	ng=(10·ng vector·insert bp) / vector bp
T4 DNA ligase	5 u/µL	5 u/reaction
H_2O		Up to 20 µL

Table 2.4. Ligation reaction.

Salts present in ligation reactions interfere with electroporations. Ligation reactions were purified with either SureClean Plus (Bioline), Sepharose column or with butanol. For the butanol extraction, 30 µL H_2O and 450 µL butanol were added to the ligation (ratio 1:10), after vigorously vortexing, the solution was centrifuged at 13,000 g for 10 minutes. The supernatant was discarded and the Eppendorf tube left to dry at 50°C. After the solution was completely evaporated the pellet was resuspended in 11 µL of pure H_2O. For the purification using Sepharose column, Sepharose solution was prepared mixing 2/3 of Sepharose and 1/3 of H_20. The solution was sterilized in the autoclave and kept at 4°C. To prepare

the columns a small needle was used to make a hole in a 0.5 mL tube, 150 µL of Sepharose solution was added to the 0.5 mL tube and placed inside a 1.5 mL tube. The column was centrifuged at 200 rpm for 2 minutes to dry the column. The 1.5 mL tube was replaced with a new one. The reaction to be purified was placed into a 0.5 mL tube and centrifuged at 200 rpm for 2 minutes.

2.2.7 One-pot restriction and ligation

This technique was used for inserting *dctT:avrRpt2* from the pMT-1 and pMT-2 plasmids into the pUC19B vector. *dctT:avrRpt2* was PCR-amplified with pMT_BsaI_for and pMT_BsaI_rev pair of primers, purified from gel and cloned into the vector pUC19B as follows: PCR product at 240 ng/µL , 60 ng/µL of pUC19B, 1.5 µL of 10x T4 ligase buffer, 0.5 µL of SmaI, 0.4 µL of T4 ligase and DNAse and RNAse free water to the final concentration of 15 µL. The reaction was left at room temperature for at least 2 hours and then purified using Sepharose columns.

2.2.8 Production of electrocompetent cells and electroporation

Replicative plasmids used and produced in this study (Table 2.2) were inserted into *E. coli* DH5α, into *A. tumefaciens* AGL1 and into *Psa* strains using electroporation; all freshly ligated plasmids were inserted into *E. coli* DH5α. Electrocompetent *E. coli* and *A. tumefaciens* AGL1 cells were obtained as follows: 250 mL of fresh LB was inoculated with a 50 mL overnight culture and incubated until the cells reached an optical density (OD_{600}) between 0.6 and 1.0. The 300 mL culture was transferred into 6 centrifuge tubes (50 mL) and centrifuged at 5,000 g for 10 minutes at 4°C. The pellet was carefully resuspended and washed three

times with 50 mL of sterile ultrapure iced water (without salts). Pellets were resuspended in 2 mL of iced 80% glycerol. 1 mL aliquots of each resuspension was transferred into fresh 1.5 mL Eppendorf tubes, centrifuged and resuspended in 250 µL of iced 10% glycerol. 85 µL aliquots were frozen and stored at -80°C.

Electrocompetent *Psa* cells were produced by inoculating 10 mL KB media with 1 mL of a *Psa* overnight culture and incubating until the cells reached OD_{600} between 0.6 and 1.0. 3 mL of the culture was centrifuged at 5,000 g for 10 minutes at 4°C. The pellet was resuspended and washed three times with 750 µL of sterile iced 0.3 M sucrose. The pellet was finally resuspended in 40 µL of sterile iced 0.3 M sucrose. After adding the vector and incubating for 30 minutes on ice, cells were electroporated. Electrocompetent *Psa* cells were solely transformed with plasmids extracted from *E. coli*.

Electroporation was performed with the Electroporator 2510 (Eppendorf) using the following parameters: 2500 V for *Psa* or 1800 V for *E. coli* and *A. tumefaciens*, 25 µFaraday, resistance range 200-500 Ω (Dower *et al.*, 1988). After electroporation, all cells were suspended in 1 mL LB (for *E. coli* and *A. tumefaciens*) or KB (for *Psa*). *E. coli* cells were incubated for 1 hour on a shaker at 250 rpm and 37°C; *Psa* and *A. tumefaciens* cells were incubated on a shaker at 250 rpm and 28°C for 2 hours. The cells were then plated in a dilution series on appropriate selective plates.

2.2.9 Triparental mating

E. coli DH5α was routinely used as a donor. This strain does not express the transfer functions so a helper *E. coli* DH5α strain containing the pRK2013 plasmid

expressing the required transfer proteins was used to ensure successful conjugation. *E. coli* donor, *E. coli* pRK2013 and *Psa* or other recipient *Pseudomonas* strains were grown overnight. 300 µL of helper and donor cells and 3 mL of the recipient were centrifuged separately for 1 minute at 4,000 rpm and washed one time with sterile physiological solution. The cells were finally resuspended in 250 µL of sterile 10mM $MgCl_2$, mixed together and centrifuged for 1 minute. The supernatant was discarded and the pellet resuspended in 30 µL of sterile water, then plated onto a pre-warmed LB agar and incubated for 24 hours at 28°C. Cells were then harvested and resuspended in 1 mL of sterile physiological solution and plated on KB Nf (50 µg/mL) and the selective antibiotic.

2.2.10 Generation of KmR *Pseudomonas* strains

IS-Ω-Km/hah (Tn*5*) transposon (Giddens *et al.*, 2007, who modified IS-*phoA*/hah (Bailey & Manoil, 2002)) was used to generate kanamycin resistant (KmR) mutants of *Psa* NZ13, *Psa* K28, *Psa* J31, *Pseudomonas syringae* pv. *actinidifoliorum* (*Pfm*) NZ9, *Pseudomonas syringae* pv. *tomato* (*Pto*) DC3000, *Pseudomonas syringae* pv *phaseolicola* (*Pph*) 1448a, *Pseudomonas syringae* H24 and H33 and *Pseudomonas fluorescens* SBW25. *E. coli* DH5α Tn*5* was used as a donor in the triparental mating. To amplify the chromosome-transposon junction of Tn*5* insertion in *Psa* NZ13, arbitrary-primed PCR (AP-PCR (Manoil, 2000)) was used to generate amplicons to the junction to allow sequencing and identification of the insertion site. This involved a round of PCR involving the same concentration of reagents as in the standard PCR; however, to ensure consistent levels of templates, 3 µL of template cells was added to the reaction from a 1 mL

overnight culture, pelleted (13,000 g for 1 minute) and resuspended in 400 µL of deionized water. Cycling conditions were as follows: an initial denaturation was performed at 95°C for 10 min, followed by five cycles of amplification involving denaturation at 95°C for 30 seconds, annealing at 42°C for 30 seconds, extension at 72°C for 3 minutes. With each of these five amplifications, the annealing temperature was decreased by 1°C per cycle. This was followed by 25 more cycles with an increased annealing temperature of 65°C. A second round of PCR (25 µL) was performed using nested primers CEKG-4 and Hah-1. Reagents for this PCR round were at the standard concentration, using 2 µL of the PCR product from the first round diluted 10 times using deionized water and 2 µL of each 10 µM primer stock. The initial denaturation was performed at 95°C for 10 min, with 30 cycles of amplification, each involving 95°C for 30 seconds, 65°C for 30 seconds and 72°C for 3 minutes. The sample was then stored at 4°C prior to purification and sequencing.

2.2.11 Generation of *Pseudomonas syringae* pv. *actinidiae* in-frame deletion mutants

The suicide plasmid vector pK18mobsacB (Schäfer *et al*., 1994) was used to create in-frame deletion mutants. One of the characteristics of the pK18mobsacB vector is its incapacity to replicate in *Psa*: it can however integrate into the chromosomal DNA of the bacteria via homologous recombination with any cloned *Psa* DNA within the vector (the "destruction box"). "Destruction boxes" obtained with overlap extension PCR were subsequently cloned into pGEM®-T using the pGEM®-T Easy Vector Systems (Promega), in order to save the construct and to

check the success of the following restriction step. A double digestion was carried out on pGEM®-T carrying the "destruction box" to produce sticky ends. The vector pK18mobsacB was digested with the same enzymes and linearized at the multi-cloning site where both the restriction sequences are present. These two double digestions were loaded and run on a 1% agarose gel. The digested pK18mobsacB vector and the destruction box were extracted, ligated, purified and electroporated into *E. coli* DH5α according to the methods described in this Chapter. Electroporated cells were plated in agar Petri dishes containing LB with Km and incubated at 37°C for 24 hours. Colonies recovered on Km were presumed to carry the plasmid and PCR screened to confirm the presence of the insert. This PCR was carried out with the M13 primers that amplify the multicloning site on the plasmid as well as the insert. The recombinant pK18mobsacB vector was then transferred into *Psa* NZ13 using the triparental mating approach. PCR screening was carried out on the colonies to confirm the first recombination event. To induce the second recombination event, the selected colonies were spread onto a Petri plate of KB containing 5% sucrose. The addition of sucrose to KB is essential to select the recombinant colonies: the pK18mobsacB vector harbors the negative-selection gene *sacB*, which is activated by the presence of sucrose. Since the expression of *sacB* is toxic to the bacterium, strong selection is imposed for the loss of this gene (in practical terms, loss of the vector) for survival. During this process the second homologous recombination occurs, leaving in the bacterial chromosome either the wildtype or the mutated copy of the target region. To select the colonies harboring the recombinant "destruction box", a PCR was carried out using primers that anneal in the chromosome upstream and downstream of the target region, to check for the presence of the deleted copy and to localize the mutation. The PCR

product was also sequenced to confirm the excision of the wildtype copy from the genome of *Psa*. The selected colonies were plated on KB with Km (100 μg/mL) to confirm the excision of the vector and their subsequent sensitivity to this antibiotic. The in-frame deletion was then confirmed via Sanger sequencing.

2.2.12 *In vitro* growth assays

In vitro growth assays were performed either by recording the absorbance or counting cfu/mL. When recording the absorbance, the Synergy 2 plate reader (BioTek) was used to record growth. Bacterial cultures were grown under continuous shaking at 28°C in 200 μL of media. Absorbance was recorded at 600nm every 5 minutes. When cfu were recorded bacteria, were grown in 30 mL vials with 10 mL of media at 28°C under continuous shaking (250 rpm). Serial dilutions were then plated on agar plates.

2.2.13 *Pseudomonas syringae* pv. *actinidiae* isolation from kiwifruit orchards

One cm^2 kiwifruit leaf disks were homogenized by mechanical disruption in 200 μL 10mM $MgCl_2$. The macerate was plated on *Pseudomonas* selective media amended with cetrimide, fucidin and cephalosporin (Oxoid) and incubated at 28˚C for 3 days. *Psa* was identified using either diagnostic PCR or LAMP assays as described in Rees-George *et al*. (2010) and Ruinelli *et al*. (2016).

2.2.14 Copper resistance assays

Copper resistance was evaluated by determining the minimal concentration of copper that inhibited growth (minimal inhibitory concentration, MIC) on MGY plates supplemented with $CuSO_4 \cdot 5H_2O$ (Bender & Cooksey, 1986, Cha & Cooksey 1991). *Psa* NZ13 is inhibited at 0.8 mM $CuSO_4$, *Psa* strains were considered resistant when their MIC exceeded 0.8 mM $CuSO_4$.

2.2.15 Psa_{NZ45}ICE_Cu mobilization assay

Psa NZ45 was used as the ICE donor. Strains tested in the mobilization assays are listed in order of divergence relative to the donor: *Psa* K28, *Psa* J31, *Pfm* NZ9, *Pto* DC3000, *Pph* 1448a, *Pseudomonas syringae* H24 and H33, *Pseudomonas fluorescens* SBW25 and *Pseudomonas aeroginosa* PAO1. The copper sulphate MIC was determined for all tested recipient strains, which were all tagged with kanamycin Tn*5*. A biparental mating was performed using 2 mL and 200 µl of washed recipient and *Psa* NZ45 cells, respectively. The cells were mixed, centrifuged briefly and resuspended in 30 µL of 10 mM $MgCl_2$ alone, 10 mM $MgCl_2$ with 0.5mM $CuSO_4$ or 30 µL of 1 cm^2 kiwifruit plantlet macerate in 200 µL of 10 mM $MgCl_2$ if requested. The cell mixture was plated onto solid media (M9 plates) and incubated at 28˚C for 48 hours. Cells were then harvested and resuspended in 1 mL of sterile 10 mM $MgCl_2$. Serial dilutions were plated on KB kanamycin to count the total number of recipients and on MGY amended with kanamycin and copper sulphate at recipient MIC to count transconjugants.

2.2.16 Hypersensitive response assay

Two weeks after germination, seeds of *Nicotiana benthamiana* were grown in Dalton's Premium Seed raising mix (Fruitfed, NZ), mixed with coarse grain vermiculite for six weeks. Plants were watered every two days in a growth room at 22°C with a photoperiod of 12 hours. *Psa* strains were grown on KB agar plates for 48h at 28°C. The cultures were adjusted at an OD_{600} of 0.2 in 10 mM $MgCl_2$ and were inoculated by infiltrating into the mesophyll of the leaves. Infiltration was achieved by pressing a 1 mL needleless syringe against the leaf underside. The plants were checked after 24 hours post inoculation (hpi). Three plants were inoculated in three or two leaves and the hypersensitive response (HR) checked after 24 hours.

2.2.17 *Pseudomonas syringae* pv. *actinidiae* pathogenicity test on model plants

Two weeks after germination, seeds of *N. benthamiana* were grown in Dalton's Premium Seed raising mix (Fruitfed, NZ), mixed with coarse grain vermiculite for six weeks, seeds of *Arabidopsis thaliana* for four-five weeks, and *Solanum lycopersicum* Money Maker for two-three weeks. Plants were watered every two days in a growth room at 22°C with a photoperiod of 12 hours. The cultures (OD_{600}=0.01 in 10 mM $MgCl_2$) were inoculated by pressure infiltration. Infiltration was achieved by pressing a 1 mL needleless syringe against the leaf underside. Growth was determined by counting cfu/cm^2. Leaf disks were removed with a sterile cork borer from infected tissue and placed into 1 mL sterile 10 mM $MgCl_2$ solution, and homogenized by mechanical disruption. After the mechanical disruption, 10 µL serial dilutions of the resulting bacterial suspensions were plated

on KB agar with Nf and incubated at 28°C until the colonies were visible for counting.

2.2.18 *Pseudomonas syringae* pv. *actinidiae* pathogenicity test on kiwifruit

Clonally propagated *Actinidia chinensis* var. *chinensis* Hort16A were provided by ©Multiflora Laboratories Ltd. 2012 (Auckland). Plantlets were planted on a rockwool support purchased from Hyalite NZ (previously Switched On Gardening). Plantlets were maintained at 20˚C with a light/dark period of 14/10 hours, 70% constant humidity. Plantlets were watered three times per week, two times with water and one time with Max Feed (Miracle-Gro®) fertilizer (3 small scoops in 4 L of H_2O).

Psa strains were grown on KB agar plates for 48 hour at 28˚C. Inoculum with a final OD_{600} of 0.2 of each strain was prepared in 50 mL of 10 mM $MgCl_2$ with 0.002% of Silwet. Three to four week old plantlets were inoculated by dipping the aerial parts in the inoculum solution for 5 seconds. Five separate plantlets were dip-inoculated for each treatment per each time point. When necessary Nordox75 was used at the recommended dosage of 0.375 g/L (www.kvh.org.nz/spray_products). Dip-inoculated plantlets were allowed to dry, then sprayed adaxially and abaxially with Nordox75 until runoff to ensure complete coverage. Bacterial growth was monitored 0, 3 and 7 days post inoculation (dpi).

For sampling the epiphytic growth, sampled leaves were placed in sterile plastic bags with 35 mL 10 mM $MgSO_4$ buffer and gently shaken for 3 minutes. The leaf wash was pipetted into 50 mL centrifuge tube and the solution was

centrifuged at 4,100 g for 3 minutes and the supernatant discarded. The bacteria were resuspended in 200 µL buffer and serial dilutions were plated on agar plates. To assess endophytic bacterial growth one cm^2 disk leaves were cut using a sterile cork borer, surface sterilized in 70% ethanol and ground in 200 µL 10 mM $MgCl_2$. Serial dilutions of the homogenate were plated on agar plates.

For the growth of *Psa* deletion mutants, KB agar plates amended with Nf were used, KB agar plates amended with Km were used to count *Psa* NZ13^{kmR} and MGY agar plates amended with 0.8 mM $CuSO_4$ to count *Psa* NZ45. Each experiment was repeated three times.

2.2.19 *In vitro* and *in planta* competition assays

In vitro and *in planta* competition assays were conducted as described earlier for single strains, except that *Psa* NZ45 and *Psa* NZ13^{kmR} were co-inoculated in a 1:1 mix. Bacterial growth was monitored by plating serial dilutions on KB Km (*Psa* NZ13^{kmR}), MGY 0.8mM $CuSO_4$ (*Psa* NZ45) and on MGY Km 0.8mM $CuSO_4$ (*Psa*NZ13^{kmR} that acquired copper resistance). *In vitro* assays had three replicates per strain; *in planta* assays were conducted using five replicates, each experiment was repeated three times. Fitness was calculated as difference between their Malthusian Parameters: $\ln[N_{PsaNZ45}(t_1)/N_{PsaNZ45}(t_0)] - \ln[N_{PsaNZ13}(t_1)/N_{PsaNZ13}(t_0)]$ (Lenski *et al*., 1991).

2.2.20 Spot inoculation assay

Fruits and vegetables used for this study were purchased from Pak'nSave (Albany, NZ) one day before the experiment. *Psa* inocula were prepared either for

stab or drop inoculation. For stab inoculation *Psa* NZ13 and the Cluster 3 mutants were streaked from glycerol stock onto KB plates to obtain single colonies that were stabbed into the host using a sterile tip. For drop inoculation the inocula were set to a final optical density (OD_{600}) of 0.1 in 10 mM $MgCl_2$, 10 µL of solution was then dropped on a stab obtained with a sterile tip. Inoculated fruits and vegetables were then left at room temperature for 2 days.

2.2.21 Secretion Assay

The plasmids pMT-2 and pMT-1 were constructed using as vector pUCP22 by GenScript®. The promoter and the full-length *dctT* of *Psa* NZ13 were fused to *avrRpt2*$_{\Delta 1\text{-}79aa}$ from *Pto* DC3000 to obtain pMT-2; the promoter and the putative exportation site (1-52 aa) of *dctT* of *Psa* NZ13 were fused to *avrRpt2*$_{\Delta 1\text{-}79aa}$ from *Pto* DC3000 to obtain pMT-1 (plasmids were provided by GenScript®). In the positive control plasmid the full-length *avrRpt2* and its promoter were inserted in pVSP61 (Kunkel *et al*., 1993) (provided by Kee Sohn Lab). Each plasmid was inserted either into *Psa* NZ13 and PsaNZ13Δ*hrcC* via electroporation as described above.

The secretion assay was carried out in *A. thaliana* Col-0*.* Cultures set at an OD_{600} of 0.2; infiltration was achieved by pressing a 1 mL syringe without a needle against the leaf underside. Pictures were taken after 24 hpi. The experiment was repeated two times inoculating two leaves of three plants per strain used.

2.2.22 Ion leakage

Bacterial strains were grown for 2 days at 28˚C on KB plates. The inoculum was set at an optical density (OD_{600}) of 0.2 in 10 mM $MgCl_2$. Four leaves per *A.*

thaliana Col-0 plants were fully pressure-infiltrated into the leaf underside using a syringe without a needle. After 15 minutes two leaf discs per leaf were harvested using a sterile cork borer and transferred in a Petri dish containing 25 mL of sterile milliQ water and placed on gentle shake for 30 minutes. Four leaf discs per strain were transferred into a well of a 12-well plate containing 3 mL of sterile milliQ water. Ion leakage was determined by measuring the conductivity of 60 µL of the solution in the Horiba B-173 Twin Cond Conductivity Meter.

2.2.23 Validation of the Exportation Assay

To demonstrate that the fusion protein was able to induce RIN4-RPS2-dependent cell death (HR), agroinfiltration was used to transiently express DctT:AvrRpt2 in *N. benthamiana.*

A. tumefaciens is capable of transferring DNAs into plants. In nature, *A. tumefaciens* delivers T-DNA (situated in the T-plasmid) into the host cell and it integrates into the plant genome and encodes for the gene responsible for gall tumor formation from which bacteria uptake nutrients. From this plasmid, binary vectors have been created to insert desired genes or sequence fragments under the control of a eukaryotic promoter, most commonly the 35S promoter of the Cauliflower mosaic virus, plus a terminator sequence. The binary vector is compatible with both *Agrobacterium* and *E. coli* and therefore can be manipulated. The vector is delivered by the *A. tumefaciens* Type Four Secretion System while the segment to be integrated into the host cell is recognized by the presence of two imperfect 25 bp direct repeats at the left and right border.

2.2.23a Construct preparation

While *A. tumefaciens* AGL1 carrying RIN4 and RPS2 into binary vectors were provided by the Kee Sohn Lab, the binary vectors carrying the DctT:AvrRpt2 constructs were created during this study.

The binary vector used was pICH86988. The *dctT:avrRpt2* construct was PCR-amplified from the pMT-1 and pMT-2 plasmids with pMT_BsaI_for and pMT_BsaI_rev pair of primers. The PCR product was then cloned into the vector pUC19B as described in Section 2.2.10. The ligation was purified with Sepharose and electroporated into *E. coli* DH5α. Cells were plated on LB Amp with IPTG/X-Gal, to confirm the integration of the MT constructs into the vector, white colonies were PCR screened using the pMT_BsaI_for and pMT_BsaI_rev pair of primers. Positive plasmids were then digested with BsaI to confirm the presence of the restriction site and then sent for sequencing. The final construct pICH86988 carrying *dctT:avrRpt2* fused to the C-terminal epitope tag 3xFlag was constructed using the Golden Gate cloning.

Golden Gate cloning allows the simultaneous and directional assemble of multiple DNA fragments using only a one-step procedure using a PCR machine. The Golden Gate cloning was performed in a final volume of 20 µL as follow: 1 µL of NEB T4 DNA ligase, 1 µL of NEB BsaI 10x, 2 µL of NEB T4 DNA Buffer, equimolar vector and inserts (~50 ng) (pICH86988, 3xFlag pICSL, pUC19B_MT-1 (or MT-2)). The reaction was then loaded on a PCR machine with the following program: a cycle of 3 minutes at 37˚C followed by 2 minutes at 16˚C reaped 25 times, 5 minutes at 50˚C and 5 minutes at 80˚C.

The ligation product was purified with Sepharose columns, electroporated into *E. coli* DH5α and cells plated on LB Km with IPTG/X-Gal. White colonies were PCR screened with the pMT_BsaI_for and pMT_BsaI_rev pair of primers. The

plasmids were extracted and 1 µL was electroporated into *A. tumefacies* AGL1 as described in Section 2.28. Cells were plated on LB Km and Amp, colonies were PCR screened with pMT_BsaI_for and pMT_BsaI_rev pair of primers.

2.2.23b Agroinfiltration

Agroinfiltration is used as a phenotypic marker of the HR-inducing phenotype associated with RPS2. With this method, *A. tumefaciens* carrying RPS2 (AGLRPS2) and RIN4 (AGLRIN4) are infiltrated into *N. benthamiana* leaves, which transfer these two genes into the plant genome. In *A. thaliana*, RIN4 acts as a negative regulator of PTI and ETI. RIN4 is monitored by the NB-LRR immune receptor, RPS2 (and RPM1). The *P. syringae* effector AvrRpt2 is a protease that directly targets RIN4. The AvrRpt2-dependent cleavage of RIN4 is sensed by RPS2 that then activates the HR (Grant *et al*., 2006).

Seeds of *N. benthamiana* were germinated in soil mix for six weeks. Plants were watered every two days in a growth room at 22°C with a photoperiod of 12 hours. The inoculum solutions were prepared in 10 mM $MgCl_2$, each strain used was set at an optical density (OD_{600}) of 0.1 with the exception of AtRIN4 which was adjusted at a final OD_{600} of 0.4.

As negative controls, AGLRIN4, AGLRPS2 and AGLGFP were used. AGLGFP was added to the bacterial mix to balance the final optical density. In the positive control AGLRIN4, AGLRPS2 and AGLAvrRpt2 were co-infiltrated. The additional negative control is represented by the mix of AGLRIN4, AGLRPS2, AGLAvrRpt2-C123A. In this case, AvrRpt2 carries a mutation in its active site and it is not able to cleave RIN4, therefore the plant immune response is not activated. The aim of this experiment was to test the functionality of DctT:AvrRpt2 constructs of the pMT-1

and pMT-2 plasmids. In the fusion of the full-length or putative exported sequence, *dctT* didn't interfere with the functionality or RIN4 the HR should be observed.

Infiltration was achieved by pressing a 1 mL syringe without a needle against the leaf underside and HR was observed 24 hpi. Each experiment was repeated three times.

2.3 ICE search

ICEs were search in the GenBank nucleotide collection with the Blastn algorithm (https://blast.ncbi.nlm.nih.gov/Blast.cgi). Psa_{CL1}ICE, Psa_{I10}ICE, Psa_{NZ13}ICE, Psy_{B728a}ICE and Pph_{1302A}ICE were used as query. Homologous sequences were found only in *Pseudomonas syringae* strains. To expand the search the Blastn algorithm was used in the whole-genome shotgun contig database, restricting the search on *Pseudomonadales*. Psa_{CL1}ICE, Psa_{I10}ICE, Psa_{NZ13}ICE, Psy_{B728a}ICE and Pph_{1302A}ICE were used as query. All the homologous sequences were downloaded and when necessary entire ICEs were reconstructed with Geneious. Contigs unable to be concatenated into a single ICE were discarded from further analysis.

2.4 Software

2.4.1 REALPHY

REALPHY (Bertels *et al.*, 2014) is a free online pipeline that infers phylogenetic trees. All provided sequences are mapped to each of the references selected by the user using bowtie2. Individual reference alignments were merged

to create alignments on which the trees in Figure 3.1 and 4.1 were built. The PsICEs used as reference for Figure 3.1 were: Psa_{C15I}CE, Pfm_{19497}ICE, Psa_{I10}ICE, Psa_{NZ13}ICE, Psy_{B728a}ICE, Pph_{1302A}ICE, Psa_{C3}ICE, Psa_{C6}ICE, Psa_{C11}ICE, Psa_{C18}ICE, and Psa_{C2}ICE. The reference Tn*6212* elements used for Figure 4.1 belonged to *Ps* 47E2, *Paf* ICMP4394, *Ps* Pengzhou8, *Psa* C3, *Psa* C6, Psa C11, and *Psa* NZ13.

2.4.2 Progressive MAUVE

Progressive Mauve (Darling *et al.*, 2010) is a software package that aligns orthologous and xenologous DNA sequences for the detection of large-scale rearrangements among genomes, such as gene loss, duplication, rearrangement, and horizontal transfer.

Alignments with Progressive MAUVE were performed using default settings but with seeds families set at seed weight 11.

2.4.3 Alfy

Alfy (Domazet-Lošo & Haubold, 2011) is an alignment-free program for comparing one or more query DNA sequences. Alfy describes the chimeric structure of a set of sequences based on the lengths of exact matches between pairs of sequences. The program was run selecting a default sliding window of 300 bp, selecting only strong homologies (regions where the average shortest unique substring length is greater than the maximum shortest unique substring length occurring by chance alone).

2.4.4 ClonalFrameML

ClonalFrameML (Didelot & Wilson, 2015) is a computer package that uses the maximum likelihood inference to detect recombination in bacterial genomes, and accounts for it in phylogenetic reconstruction.

ClonalFrameML was run using default settings. The input files were a MAFFT alignment and a PhyML tree of the backbone genes of the non-redundant PsICEs created using Geneious version 7 (Kearse *et al*., 2012).

Chapter 3 - Integrative and Conjugative Elements (ICEs) found in *Pseudomonas syringae* define a new ICE family

3.1 Introduction

Horizontal gene transfer (HGT) plays an important role in evolution. Kloesges *et al*. (2011) estimated that across 329 proteobacterial genomes, at least 75% of the protein families have been affected by HGT. The acquisition of new traits via HGT allows recipient organisms to exploit novel biochemical functions already refined by selection in the donor organism, thereby facilitating the efficient and effective invasion of new ecological niches (Lawrence, 2002). The colonization of new niches in turn may lead to diversification resulting in bacterial speciation (Polz *et al*., 2013).

Conjugative transmissible elements, such as plasmids and Integrative Conjugative Elements (ICEs), are able to move functional genetic units over broad phylogenetic distances (Médigue *et al*., 1991; Lan & Reeves, 1996; Sullivan & Ronson, 1998; Ochman *et al*., 2000; Ochman *et al*., 2005; Guglielmini *et al*., 2011; Polz *et al*., 2013). Sequence analyses suggest that ICEs are the most abundant type of conjugative elements in bacteria (Guglielmini *et al*., 2011). ICEs are plasmid-like

entities with attributes of temperate phages that disseminate vertically as part of the bacterial chromosome and horizontally by virtue of endogenously encoded machinery for conjugative transfer (Wozniak & Waldor, 2010; Guglielmini *et al.*, 2011). During the process of conjugation ICEs circularize and transfer to new hosts, leaving a copy in the original host genome (Wozniak & Waldor, 2010; Johnson & Grossman, 2015). Genes encoding processes responsible for the ICE life cycle (integration, excision, conjugation and regulation) are encoded within recognisable modules (Mohd-Zain *et al.*, 2004; Juhas *et al.*, 2007; Roberts & Mullany, 2009). In addition to essential genes, ICEs often harbour accessory genes of potential adaptive significance for their hosts. These include genes affecting biofilm formation, pathogenicity, antibiotic and heavy metal resistance, symbiosis and bacteriocin synthesis (Peters *et al.*, 1991; Rauch *et al.*, 1992; Ravatn *et al.*, 1998; Beaber *et al.*, 2002; Drenkard *et al.*, 2002; Burrus *et al.*, 2006; Ramsay *et al.*, 2006; Dimopoulou *et al.*, 2007; Kung *et al.*, 2010). It was the observation of transferable antibiotic resistance and the capacity to cause nodulation on leguminous plants and fix nitrogen that first led to the discovery of ICEs (Franke & Clewell, 1981; Mays *et al.*, 1982; Rashtchian *et al.*, 1982; Roberts & Smith, 1980; Stuy, 1980; Shoemaker *et al.*, 1980; Nugent, 1981; Magot, 1983; Smith *et al.*, 1981; Sullivan *et al.,* 1995). However, not all ICEs possess cargo genes that confer an obvious fitness advantages to their host genomes under standard laboratory conditions. Such ICEs are likely to be discovered via sequence analysis rather than phenotypic assay (Johnson & Grossman, 2015).

Pseudomonas syringae is a model organism for the study of plant-microbe interactions and microbial evolution due to its ubiquity in both agricultural environments, where different lineages are responsible for frequent outbreaks of

disease in a variety of hosts, and in non-agricultural habitats, where it is found in asymptomatic association with wild plants, as well as in leaf litter, rivers, snowpack and even clouds (Morris *et al.*, 2007, 2008, 2010). HGT occurs at high frequency in *P. syringae,* and has a disproportionate impact on adaptation in nature (Baltrus *et al.*, 2011, 2016; Nowell *et al.*, 2016). Individual *P. syringae* lineages are thought to have acquired thousands of genes in the same period of time that a 1% amino acid divergence has occurred in the core genome (Nowell *et al.*, 2014).

The first ICE described in *P. syringae* was the PPHGI-1 in *P. syringae* pv. *phaseolicola* (*Pph*) 1302A (Pitman *et al.*, 2005). In this case the discovery was linked to phenotypic changes given by the presence/absence of PPHGI-1 carrying the effector gene *hopAR1* when inoculated in bean, cultivar Tendergreen (Pitman *et al.*, 2005). Comparison of the first two sequenced *P. syringae* genomes revealed the presence of PsyrGI-6 in *P. syringae* pv. *syringae* (*Psy*) B728a (Feil *et al.*, 2005). The pandemic of kiwifruit canker disease caused by *P. syringae* pv. *actinidiae* (*Psa*) led to the identification other three additional ICEs called Pacific ICE (Pac_ICE1), Mediterranean ICE (Pac_ICE2) and Andean ICE (Pac_ICE3) (McCann *et al.*, 2013; Butler *et al.*, 2013). More recently, *Psa* ICEs were found to encode copper resistance, a phenotype of potential adaptive significance in orchards, where copper is frequently applied as a foliar bactericide (Chapter 5; Colombi *et al.*, 2017).

3.1.1 Aims

- Analyze the distribution of the ICEs orthologous to the ones previously isolated in *Psa* and in other *P. syringae*.
- Determine the structure of the PsICEs and also infer their evolutionary history

3.2 Results

3.2.1 Definition of a new family of ICE: PsICE

3,189 *Pseudomonas* whole genome sequences in GenBank (updated to December 2016) were interrogated for the presence of ICEs homologous to the known ICEs in *P. syringae* (Pph_{1302A}ICE, Psy_{B728a}ICE, Psa_{I10}ICE, Psa_{NZ13}ICE, Psa_{CL1}ICE) based on the sequence of the entire ICE. 91 ICEs were identified and these came from bacteria isolated from the 1920s to present from both cultivated and wild plants and from different continents (with the exception of Antarctica). Four ICEs were excluded from this group of 87 (see next Section) because of extensive deletions that have likely destroyed ICE function. The 87 ICEs used in this study are listed in Table 3.1. ICEs ranged in length from 88 kb to 139 kb and were located at one or both of the known chromosomal integration sites (*att-1* and *att-2*) as defined by Lovell *et al*. (2009) and McCann *et al*. (2013). The *att-1* site is within a tRNA upstream of *clpB*, while the *att-2* site is within a tRNA upstream of *queC*. Only four genomes harboured more than one ICE. Two different ICEs are integrated in tandem into the *att-2* site of *Psa* C9. The second ICE in *Psa* C9 could not be reconstructed, therefore it was excluded from subsequent analysis. *Psa* NZ45, NZ47, and NZ64 have two ICEs in each *att* site: they harbour both the PsaNZ13-like ICE (present in *Psa* strains isolated in New Zealand) and strain-specific ICEs conferring copper resistance (Chapter 5). An expanded search against the GenBank nucleotide collection (*P. syringae* spp. excluded) for homologues of the known ICEs in *P. syringae* based on their entire sequences showed the absence of homologous ICEs in any other species.

Here, I define a new family of ICEs: PsICE (as *Pseudomonas syringae* ICEs). ICEs belonging to the PsICE family share the same backbone gene structure (refer to Section 3.2.3), and integrate into one of two loci (*att-1* and *att-2*) near *clpB* and *queC*. I propose the following nomenclature in preparation for the inevitable identification of new ICEs in the future: abbreviation for the species of origin, with the strain in subscript followed by "ICE". If the same strain harbors two ICEs the proposed nomenclature uses an underscore using up to three letters to distinguish them. For example, *Psa* NZ45 harbours the Psa_{NZ45}ICE, which is common to the *Psa* strains isolated in New Zealand; additionally, *Psa* NZ45 acquired a second ICE conferring copper resistance (Chapter 5; Colombi *et al.*, 2017) named Psa_{NZ45}ICE_Cu. This type of nomenclature is a modification of Burrus *et al*. (2002).

ICE ID	Strain	Country	Year	Isolated from	ICE Integration locus	Tn*6212*	GenBank accession number
Paf$_{4394}$ICE *	*P. syringae* pv. *atrofaciens* ICMP 4394	New Zealand	1968	wheat	*att-1*	✓	LJPO01000188.1; LJPO01000111.1
Par$_{4457}$ICE	*P. coronafaciens* pv. *atropurpurea* ICMP4457	Japan	1967	ryegrass	*att-2*	✓	LJPS01000162.1, LJPS01000123.1, LJPS01000111.1, LJPS01000149.1, LJPS01000047.1
Pav$_{013}$ICE	*P. syrinage* pv.*avellanae* ISPaVe013	Italy	1991	hazelnut	N/A	✓	AKCJ01000063.1, AKCJ01000059.1, AKCJ01000060.1, AKCJ01000061.1, AKCJ01000062.1
Pco$_{19117}$ICE *	*P. congelans* ICMP 19117	Germany	1994	grass	*att-2*	✓	LJQB01000076.1, LJQB01000035.1
Pdp$_{529}$ICE	*P. syringae* pv. *delphinii* ICMP529	New Zealand	1957	larkspur	*att-1*	✓	LJQH01000365.1, LJQH01000290.1, LJQH01000355.1
Pfm$_{19497}$ICE	*P. syringae* pv. *actinidifoliorum* ICMP 19497	New Zealand	2010	kiwifruit	*att-1*	✗	LKBQ01000112.1
Pmaginalis$_{11289}$ICE	*P. marginalis* ICMP 11289	New Zealand	1991	kiwifruit	*att-1*	✗	LKGX01000080.1
Ppa$_{2367}$ICE *	*P. syringae* pv. *panici* LMG2367	USA	1921-22	proso millet	*att-2*	✓	ALAC01000062.1, ALAC01000019.1
Pph$_{1302A}$ICE *	*P. syringae pv. phaseolicola* 1302A	Ethiopia	1994	bean	*att-2*	✗	AJ870974.1
Ps$_{11168}$ICE *	*P. syringae* ICMP 11168	New Zealand	1991	kiwifruit	*att-2*	✓	LKGV01000001.1
Ps$_{12500}$ICE	*P. syringae* KCTC12500				*att-1*	✗	AYTM02000002.1
Ps$_{243L2}$ICE	*P. syringae* 243L2	China	2014	kiwifruit	*att-1*	✓	Rainey Lab
Ps$_{247E2}$ICE	*P. syringae* 247E2	China	2014	kiwifruit	*att-1*	✓	Rainey Lab
Ps$_{26L6}$ICE	*P. syringae* 26L6	China	2014	kiwifruit	*att-1*	✓	Rainey Lab
Ps$_{2L4}$ICE *	*P. syringae* 2L4	China	2014	kiwifruit	*att-2*	✓	Rainey Lab
Ps$_{34876}$ICE *	*P. syringae* BRIP34876	Australia	1971	barley	*att-2*	✓	AMXK01000055.1, AMXK01000099.1, AMXK01000014.1
Ps$_{34881}$ICE	*P. syringae* BRIP34881	Australia	1971	barley	*att-2*	✓	AMXL01000033.1, AMXL01000031.1
Ps$_{39023}$ICE *	*P. syringae* BRIP39023	Australia	1988	wheat	*att-1*	✗	AMZX01000047.1, AMZX01000001.1
Ps$_{82L1}$ICE *	*P. syringae* 82L1	China	2014	kiwifruit	*att-1*	✓	Rainey Lab
Ps$_{91L8}$ICE *	*P. syringae* 91L8	China	2014	kiwifruit	*att-2*	✓	Rainey Lab
Psa$_{C11}$ICE *	*P. syringae* pv. *actinidiae* C11	China	2012	kiwifruit	*att-2*	✓	MTHK01000001.1
Psa$_{C15}$ICE *	*P. syringae* pv. *actinidiae* C15	China	2012	kiwifruit	*att-2*	✓	MTHO01000001.1, MTHO01000028.1, MTHO01000264.1
Psa$_{C18}$ICE *	*P. syringae* pv. *actinidiae* C18	China	2012	kiwifruit	*att-1*	✓	MTHR01000006.1
Psa$_{C1}$ICE	*P. syringae* pv. *actinidiae* C1	China	2010	kiwifruit	*att-1*	✓	KC148185.1
Psa$_{C24}$ICE	*P. syringae* pv. *actinidiae* C24	China	2014	kiwifruit	*att-2*	✓	MTHS01000003.1

Psa$_{C26}$ICE	*P. syringae* pv. *actinidiae* C26	China	2014	kiwifruit	*att-2*	✓	MTHT01000001.1
Psa$_{C27}$ICE *	*P. syringae* pv. *actinidiae* C27	China	2014	kiwifruit	*att-2*	✓	MTHU01000001.1
Psa$_{C28}$ICE	*P. syringae* pv. *actinidiae* C28	China	2014	kiwifruit	*att-2*	✓	MTHV01000001.1
Psa$_{C29}$ICE	*P. syringae* pv. *actinidiae* C29	China	2014	kiwifruit	*att-2*	✓	MTHW01000005.1, MTHW01000045.1, MTHW01000257.1
Psa$_{C2}$ICE *	*P. syringae* pv. *actinidiae* C2	China	2012	kiwifruit	*att-2*	✓	Rainey Lab
Psa$_{C30}$ICE	*P. syringae* pv. *actinidiae* C30	China	2014	kiwifruit	*att-2*	✓	MTHX01000007.1, MTHX01000050.1, MTHX01000252.1
Psa$_{C31}$ICE	*P. syringae* pv. *actinidiae* C31	China	2014	kiwifruit	*att-2*	✓	MTHY01000007.1, MTHY01000050.1, MTHY01000250.1
Psa$_{C3}$ICE *	*P. syringae* pv. *actinidiae* C3	China	2012	kiwifruit	N/A	✓	MTCQ01000008.1
Psa$_{C4}$ICE	*P. syringae* pv. *actinidiae* C4	China	2012	kiwifruit	*att-2*	✓	Rainey Lab
Psa$_{C5}$ICE	*P. syringae* pv. *actinidiae* C5	China	2012	kiwifruit	*att-2*	✓	Rainey Lab
Psa$_{C62}$ICE	*P. syringae* pv. *actinidiae* C62	China	2014	kiwifruit	*att-1*	✓	MTIB01000003.1
Psa$_{C66}$ICE *	*P. syringae* pv. *actinidiae* C66	China	2012	kiwifruit	*att-1*	✗	MTYK01000003.1
Psa$_{C67}$ICE	*P. syringae* pv. *actinidiae* C67	China	2012	kiwifruit	*att-2*	✓	MTYL01000003.1
Psa$_{C68}$ICE	*P. syringae* pv. *actinidiae* C68	China	2014	kiwifruit	*att-1*	✓	MTYM01000003.1
Psa$_{C69}$ICE	*P. syringae* pv. *actinidiae* C69	China	2014	kiwifruit	*att-1*	✓	MTYO01000003.1
Psa$_{C6}$ICE *	*P. syringae* pv. *actinidiae* C6	China	2012	kiwifruit	*att-2*	✓	Rainey Lab
Psa$_{C70}$ICE	*P. syringae* pv. *actinidiae* C70	China	2014	kiwifruit	*att-2*	✓	MTYN01000001.1
Psa$_{C73}$ICE	*P. syringae* pv. *actinidiae* C73	China	2014	kiwifruit	*att-2*	✓	MTYP01000001.1
Psa$_{C9}$ICE	*P. syringae* pv. *actinidiae* C9	China	2010	kiwifruit	*att-2*	✓	ANJI02000003.1
Psa$_{CL1}$ICE	*P. syringae* pv. *actinidiae* CL1	Chile	2010	kiwifruit	*att-2*	✓	KC148188.1
Psa$_{CL2}$ICE	*P. syringae* pv. *actinidiae* CL2	Chile	2010	kiwifruit	*att-2*	✓	Rainey Lab
Psa$_{CL3}$ICE	*P. syringae* pv. *actinidiae* CL3	Chile	2010	kiwifruit	*att-2*	✓	Rainey Lab
Psa$_{CL4}$ICE	*P. syringae* pv. *actinidiae* CL4	Chile	2010	kiwifruit	*att-2*	✓	Rainey Lab
Psa$_{I10}$ICE *	*P. syringae* pv. *actinidiae* I10	Italy	2008	kiwifruit	*att-2*	✓	ANGD01000014.1
Psa$_{I13}$ICE	*P. syringae* pv. *actinidiae* I13	Italy	2008	kiwifruit	*att-2*	✓	Rainey Lab
Psa$_{I4}$ICE	*P. syringae* pv. *actinidiae* I4	Italy	2008	kiwifruit	*att-2*	✓	Rainey Lab
Psa$_{I5}$ICE	*P. syringae* pv. *actinidiae* I5	Italy	2009	kiwifruit	*att-2*	✓	Rainey Lab

Psa_{I6}ICE	*P. syringae* pv. *actinidiae* I6	Italy	2009	kiwifruit	*att-2*	✓	Rainey Lab
Psa_{I7}ICE	*P. syringae* pv. *actinidiae* I7	Italy	2009	kiwifruit	*att-1*	✓	Rainey Lab
Psa_{I8}ICE	*P. syringae* pv. *actinidiae* I8	Italy	2010	kiwifruit	*att-2*	✓	Rainey Lab
Psa_{I9}ICE	*P. syringae* pv. *actinidiae* I9	Italy	2010	kiwifruit	*att-2*	✓	Rainey Lab
Psa_{J38}ICE	*P. syringae* pv. *actinidiae* J38	Japan	2014	kiwifruit	*att-2*	✓	MTYS01000003.1
Psa_{K7}ICE	*P. syringae* pv. *actinidiae* K7	South Korea	2014	kiwifruit	*att-2*	✓	MTYW01000054.1, MTYW01000001.1
Psa_{NZ13}ICE *	*P. syringae* pv. *actinidiae* NZ13	New Zealand	2010	kiwifruit	*att-1*	✓	CP011972.2
Psa_{NZ31}ICE	*P. syringae* pv. *actinidiae* NZ31	New Zealand	2010	kiwifruit	N/A	✓	ANJD01000023.1, ANJD01000118.1, ANJD01000111.1
Psa_{NZ32}ICE	*P. syringae* pv. *actinidiae* NZ32	New Zealand	2010	kiwifruit	*att-1*	✓	ANJC01000006.1
Psa_{NZ33}ICE	*P. syringae* pv. *actinidiae* NZ33	New Zealand	2011	kiwifruit	*att-1*	✓	ANJG01000046.1
Psa_{NZ34}ICE	*P. syringae* pv. *actinidiae* NZ34	New Zealand	2011	kiwifruit	*att-1*	✓	ANJH01000079.1, ANJH01000040.1, ANJH01000162.1
Psa_{NZ35}ICE	*P. syringae* pv. *actinidiae* NZ35	New Zealand	2010	kiwifruit	*att-1*	✓	MTYX01000007.1
Psa_{NZ36}ICE	*P. syringae* pv. *actinidiae* NZ36	New Zealand	2012	kiwifruit	*att-2*	✓	Rainey Lab
Psa_{NZ37}ICE	*P. syringae* pv. *actinidiae* NZ37	New Zealand	2010	kiwifruit	*att-1*	✓	MTYY01000003.1
Psa_{NZ38}ICE	*P. syringae* pv. *actinidiae* NZ38	New Zealand	2014	kiwifruit	*att-2*	✓	MTYZ01000003.1
Psa_{NZ39}ICE	*P. syringae* pv. *actinidiae* NZ39	New Zealand	2014	kiwifruit	*att-1*	✓	MTZA01000003.1
Psa_{NZ40}ICE	*P. syringae* pv. *actinidiae* NZ40	New Zealand	2014	kiwifruit	*att-2*	✓	MTZB01000003.1
Psa_{NZ41}ICE	*P. syringae* pv. *actinidiae* NZ41	New Zealand	2014	kiwifruit	*att-2*	✓	MTZC01000003.1
Psa_{NZ42}ICE	*P. syringae* pv. *actinidiae* NZ42	New Zealand	2014	kiwifruit	*att-1*	✓	MTZD01000004.1
Psa_{NZ43}ICE	*P. syringae* pv. *actinidiae* NZ43	New Zealand	2014	kiwifruit	*att-1*	✓	MTZE01000003.1
Psa_{NZ44}ICE	*P. syringae* pv. *actinidiae* NZ44	New Zealand	2014	kiwifruit	N/A	✓	Rainey Lab
Psa_{NZ45}ICE_Cu *	*P. syringae* pv. *actinidiae* NZ45	New Zealand	2014	kiwifruit	*att-1*	✗	MTZF01000001.1
Psa_{NZ46}ICE	*P. syringae* pv. *actinidiae* NZ46	New Zealand	2012	kiwifruit	*att-1*	✓	Rainey Lab
Psa_{NZ47}ICE_Cu *	*P. syringae* pv. *actinidiae* NZ47	New Zealand	2014	kiwifruit	*att-1*	✗	CP017009.1
Psa_{NZ48}ICE	*P. syringae* pv. *actinidiae* NZ48	New Zealand	2013	kiwifruit	*att-2*	✓	MTZH01000003.1
Psa_{NZ49}ICE	*P. syringae* pv. *actinidiae* NZ49	New Zealand	2011	kiwifruit	*att-1*	✓	MTZJ01000003.1
Psa_{NZ62}ICE	*P. syringae* pv. *actinidiae* NZ62	New Zealand	2015	kiwifruit	*att-2*	✗	MOMK01000003.1
Psa_{NZ63}ICE	*P. syringae* pv. *actinidiae* NZ63	New Zealand	2015	kiwifruit	*att-1*	✗	MOML01000002.1

Psa_{NZ64}ICE_Cu *	*P. syringae* pv. *actinidiae* NZ64	New Zealand	2016	kiwifruit	*att-1*	✗	MOMM01000021.1, MOMM01000065.1 MOMM01000141.1 MOMM01000150.1 MOMM01000185.1 MOMM01000221.1 MOMM01000225.1, MOMM01000243.1
Psa_{P1}ICE	*P. syringae* pv. *actinidiae* P1	Portugal	2010	kiwifruit	*att-2*	✓	Rainey Lab
$Ps_{Pengzhou8}$ICE *	*P. syringae* Pengzhou8	China	2014	kiwifruit	*att-1*	✓	Rainey Lab
Psy_{3023}ICE *	*P. syringae* pv. *syringae* ICMP3023	UK	NA	bean	*att-1*	✗	LJRK01000049.1
Psy_{41A}ICE *	*P. syringae* pv. *syringae* 41A	France	2011	apricot	*att-2*	✓	JYHJ01000001.1
Psy_{B728a}ICE *	*P. syringae* pv. *syringae* B728a	USA	1985	bean	*att-2*	✓	CP000075.1
Psy_{SM}ICE *	*P. syringae* pv. *syringae* SM	USA	N/A	wheat	*att-2*	✗	APWT01000010.1; APWT01000011.1

Table 3.1. ICEs used in this study. The strains harbouring the ICE, its country, year and host of isolation; the integration locus, presence of Tn*6212* and Genbank accession of each ICE are described in the Table. * marks the non-redundant PsICEs.

3.2.2 PsICE distribution

The diversity of the 87 PsICEs is represented in Figure 3.1.

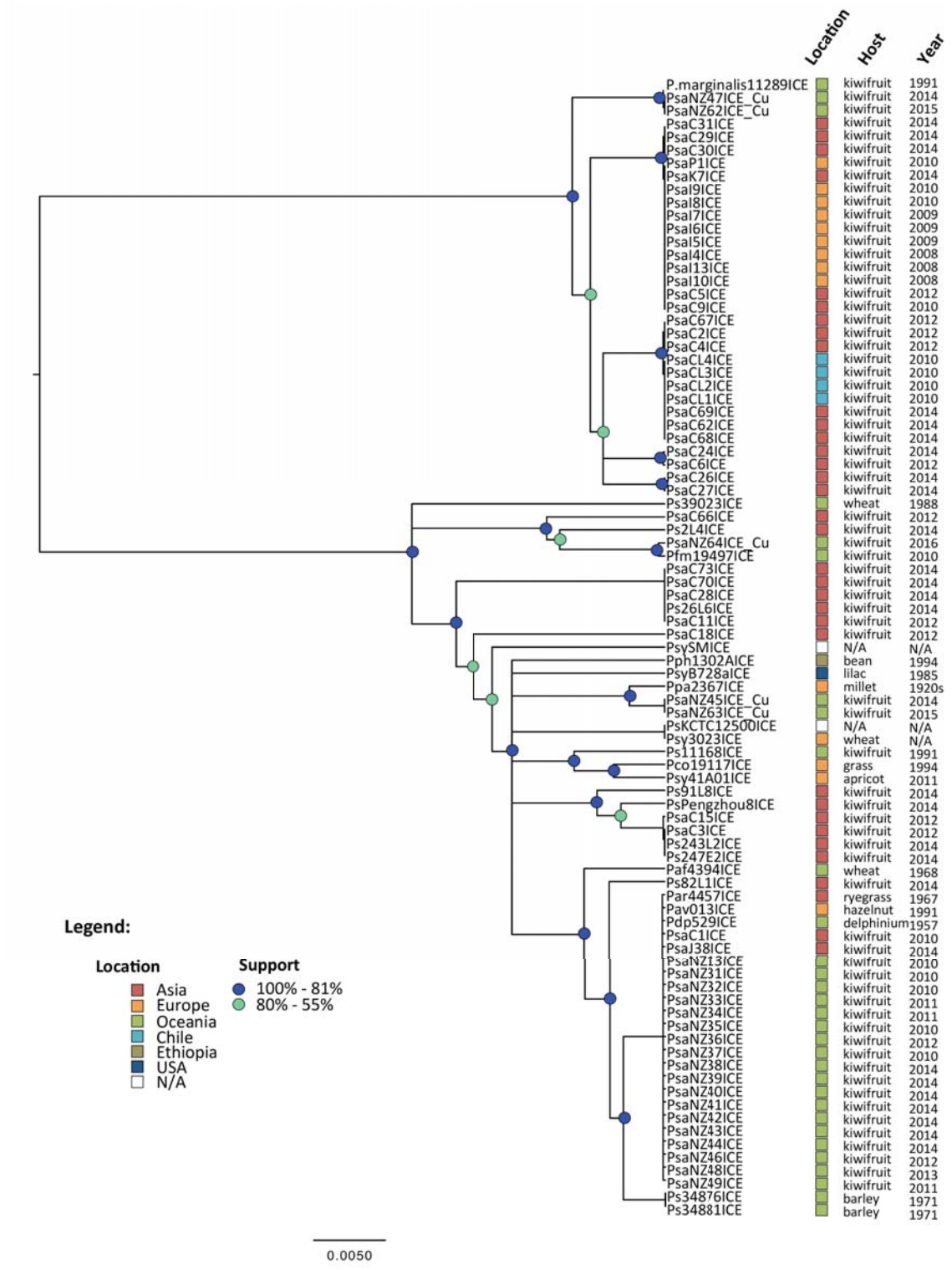

Figure 3.1. Patterns of similarity of the PsICEs. UPGMA tree based on the REALPHY alignment of the PsICEs used in this study. The entire PsICEs sequences were used in REALPHY which produced a ~4.7 kb alignment used to build the tree. Scale bar represents substitutions per site.

The tree is based on patterns of conserved SNPs extracted from whole genome alignments using REALPHY (Bertels *et al.*, 2014). The alignment from which the tree was generated was 4.7 kb, because few sites are conserved across the set of ICEs. The purpose of the tree is simply to cluster together similar ICEs – the tree itself conveys limited phylogenetic information.

PsICEs do not cluster based on year, plant, or location of isolation (Figure 3.1). For example, strains isolated from kiwifruit in China encompass a set of diverse ICEs. Additionally, ICEs from older strains in the data set (1920-1991) cluster with contemporary ICEs. It is important to note that the availability of many *Psa* and *P. syringae* genome sequences isolated from China in our laboratory constitutes a source of sampling-generated bias.

Given the known capacity for horizontal transfer of ICEs, ICE distribution is unlikely to reflect host phylogeny. To explore the relationship between ICE and host phylogeny, the phylogeny of strains harbouring PsICEs was constructed using a concatenation of the alignments of *gapA*, *gltA*, *gyrB* and *rpoD* (Hwang *et al.*, 2005) and is shown in Figure 3.2. Comparison of Figure 3.1 and Figure 3.2 shows that PsICE distribution does not reflect the underlying phylogeny of the strains harbouring them. Lineage 3 of *Psa* (*Psa*-3) harbours a diverse panel of PsICEs. The PsaI10-like PsICE is present both in the *Psa* outbreak strains present in Europe and China and in three *Psa* strains belonging to the more divergent *Psa*-3 strains isolated in China (*Psa* C5 and C9) and Korea (*Psa* K7) (McCann *et al.*, 2017). At the same time, this same ICE is found in otherwise distantly related strains. Psa_{C3}ICE is also present in two suspected non-pathogenic kiwifruit isolates (*P. syringae* (*Ps*) 243L2 and *Ps* 247E2) that belong to phylogroup 3 (PG3, *Psa* is PG1). *Ps* 26L6 is a close relative of *Ps* 243L2, but *Ps* 26L6 harbours a PsaC11-like PsICE and *Ps* 243L2

harbours a PsaC3-like ICE; both these strains were isolated from kiwifruit in China.

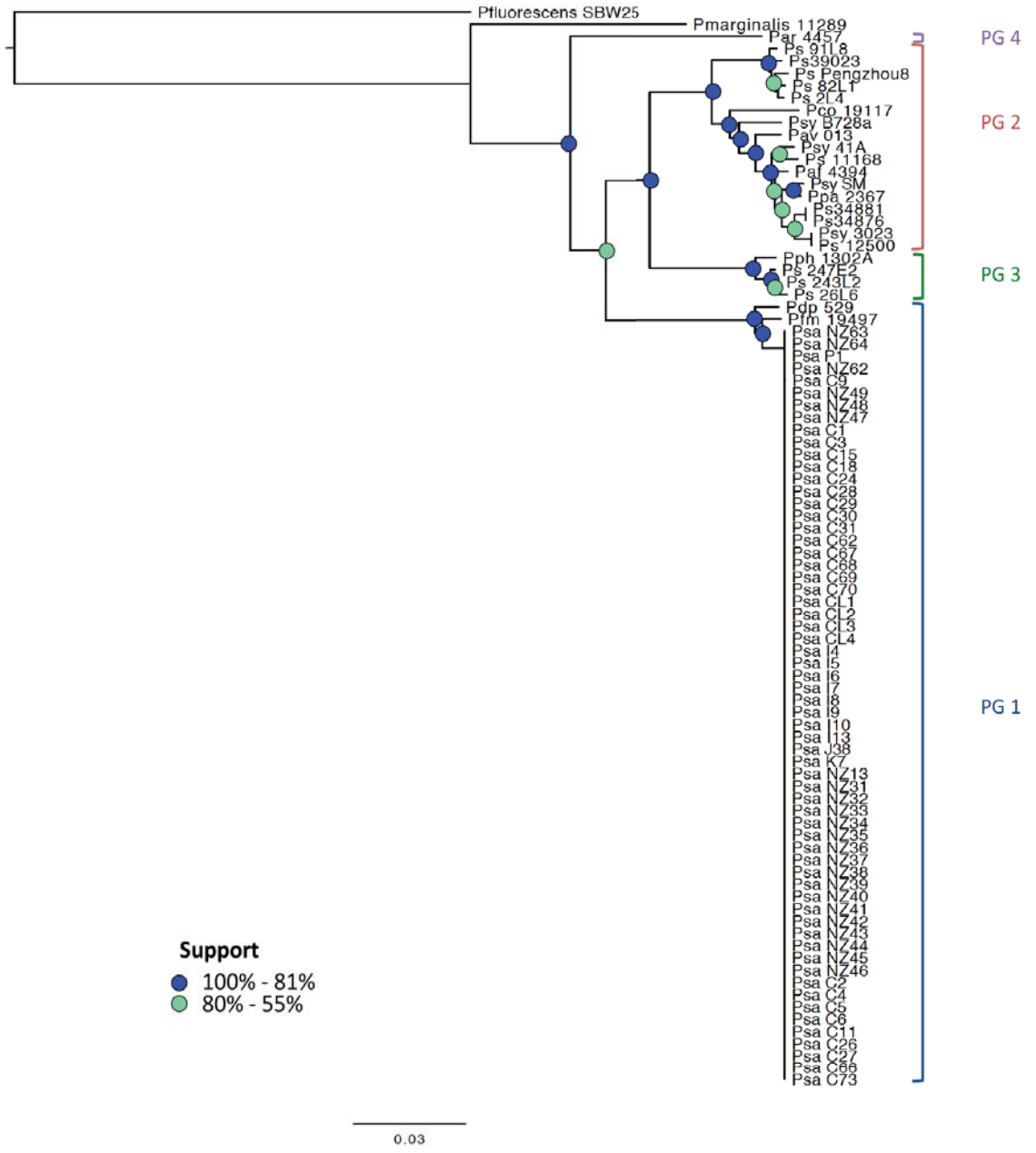

Figure 3.2. Phylogeny of the strains harbouring the PsICEs. PhyML tree based on the concatenation of *gapA*, *gltA*, *gyrB* and *rpoD*. *P. fluorescens* SBW25 was used as outgroup. PG stands for phylogroup.

The same copper resistance-encoding (CuR) ICE is present both in the canker-causing *Psa* NZ64 and foliar *P. syringae* pv. *actinidifoliorum* ICMP 19497 (*Pfm* 19497) (both isolated in New Zealand on kiwifruit 6 years apart). One PsICE has crossed the species boundary: the CuR ICE in *Psa* NZ47 is present in *P. marginalis* ICMP 11289, both isolated from kiwifruit in New Zealand, but 23 years apart.

During these 23 years the PsICE accumulated 6 SNPs and an 18 bp indel. While in these cases transmission is between strains isolated from kiwifruit, in others the distance between strains is more evident. The PsaNZ13-like PsICE is present in the outbreak *Psa* lineage infecting NZ (including two strains from Japan and one from China) (PG1), in *P. syringae* pv. *avellanae* ISPaVe013 (*Pav* 013) (PG2, Italy, 1991), *P. syringae* pv. *delphinii* ICMP529 (*Pdp* 529) (PG1, New Zealand, 1957) and in the more distantly related *P. coronafaciens* pv. *atropurpurea* (*Par* 4457) (PG4, Japan, 1967). These strains were isolated from kiwifruit, hazelnut, larkspur and ryegrass, respectively. The PsICEs in *Pdp* 529, *Par* 4457 and *Pav* 013 isolated 53, 43 and 19 years before *Psa* NZ13 differ from the Psa_{NZ13}ICE by 6, 4 and 8 non recombinant SNPs respectively.

From the initial 87 PsICEs, 28 non-redundant PsICEs were identified (Table 3.1). Redundancy is based on gene content rather than SNPs accumulation. Most of the non-redundant PsICEs are present in only one strain, but others are found in more than one. PsaI10-like ICEs are found in 15 *Psa* strains and Psa_{I10}ICE picked as representative; Psa_{NZ13}ICE was chosen as the non-redundant representative for 18 *Psa* outbreak strains, *Pdp* 529, *Par* 4457 and *Pav* 013. Psa_{C2}ICE, Psa_{C3}ICE, Psa_{C11}ICE, Psa_{C6}ICE, Psa_{C26}ICE, Psa_{NZ45}ICE, Psa_{NZ47}ICE, Psy_{3023}ICE and Ps_{34876}ICE were selected as representative of their respective clusters. Pfm_{19497}ICE differs from Psa_{NZ46}ICE by 458 SNPs, 426 of which concentrated in less than 7 kb over a ~127 kb ICE, and thus considered redundant. Psa_{C15}ICE was considered non-redundant from Psa_{C3}ICE as the difference is in 335 SNPs, a 13 kb transposon insertion coding for copper and arsenic resistance and a 2 kb deletion in Psa_{C15}ICE.

3.2.3 Delineation of PsICE structure

To make sense of diversity it is necessary to find ways of identifying similarities among different types. Of the 28 non-redundant PsICEs it was possible to align the complete sequence of 26 PsICEs using the genome alignment tool MAUVE (Darling *et al.*, 2010) (Figure 3.3). Ppa_{2367}ICE and Psy_{3023}ICE are not represented in Figure 3.3 because their inclusion caused the MAUVE software to malfunction, however these two ICEs are included in subsequent analyses.

Figure 3.3 shows that these ICEs share apparently conserved regions of homology (coloured blocks) but also non-conserved regions (outside the blocks). The regions of homology are either conserved in all the ICEs (i.e. the first green block in all the ICEs) or in few (i.e. dark blue block in the third ICE).

Regions conserved in all ICEs are likely to be those necessary for the ICE life cycle: ICE integrity, excision, conjugation, integration and replication. The core genes appear to be located in blocks coloured in light green, light purple, pink, dark green and bluebell in Figure 3.3. To identify the core shared genes of the PsICEs more precisely, the ".backbone" file originated by MAUVE was used selecting regions that were present in each sequence. These segments were mapped back to the original ICE sequences to discover which genes MAUVE defined as "backbone". Although the algorithm used by MAUVE to create blocks of sequence similarity relies on identification of anchors rejecting those that did not meet "minimum weight" criteria, it is nonetheless desirable to make explicit these criteria so that they can be applied to each open reading frame thus justifying assignment of genes to core or accessory classes. Unfortunately, it is not easy to define such criteria because of the low level of similarity among ICE-encoded proteins predicted to perform the same function (e.g. DSBA in Psa_{I10}ICE and Psa_{NZ13}ICE share ~62% pairwise identity). However, genes identified as conserved do encode the same

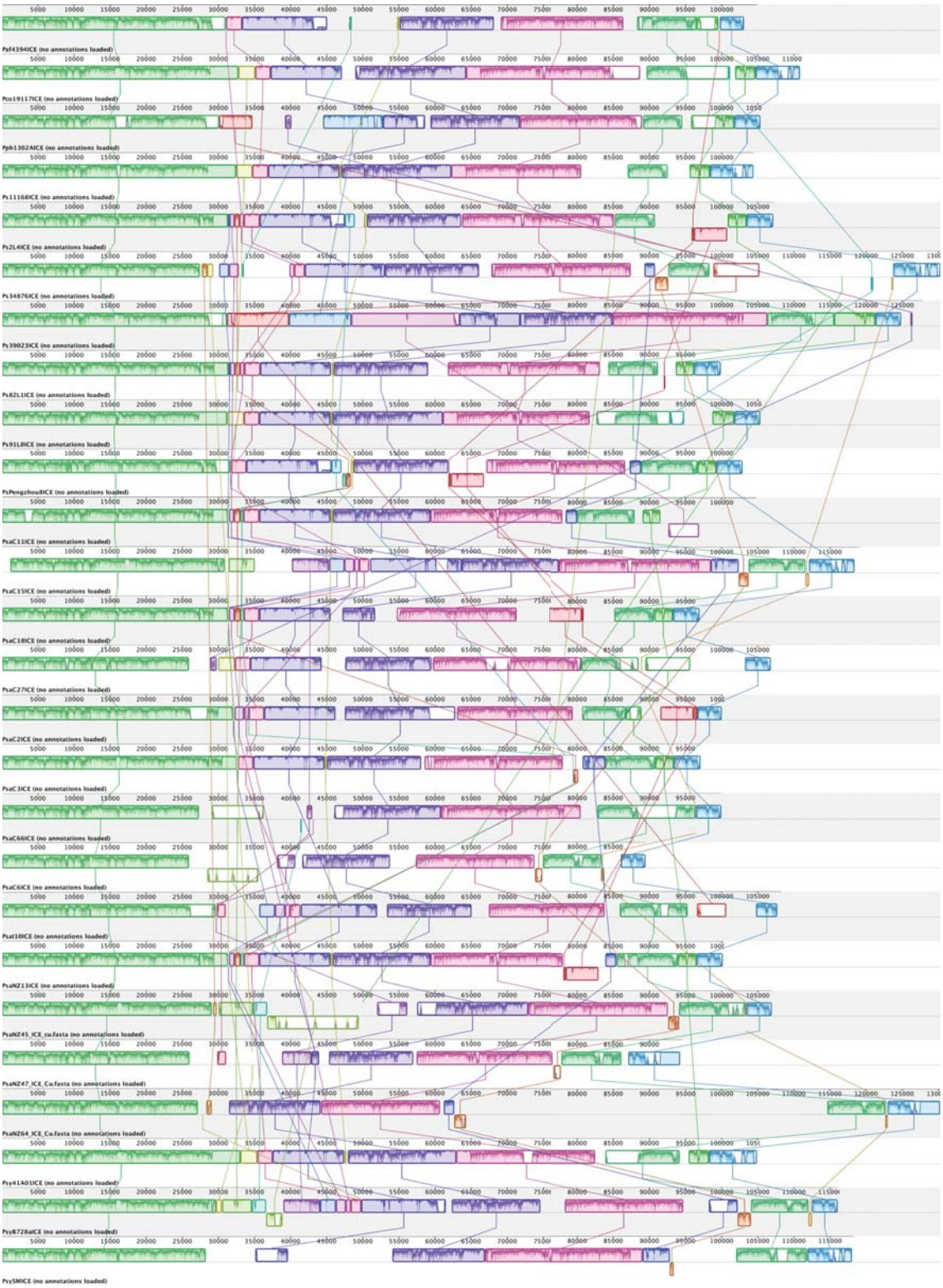

Figure 3.3. Non-redundant PsICE comparison. Alignment of 26 non-redundant ICEs produced by MAUVE using seeds family and seed weight 11. Colour-coded blocks depict regions of homology that are connected through lines; blocks below the center-line indicate inverse orientation. Inside each block a similarity profile of the sequence is drawn, the height of the similarity profile corresponds to the average level of conservation in that region of the sequence. Regions outside blocks were not aligned and probably contain sequence elements specific to a particular PsICE.

protein across ICEs, and these proteins are likely to be involved in the ICE life cycle (refer to Table 3.2).

Genes coding for the conjugation machinery (*pil* and *virB*), transfer (*tra*), partitioning (*par*), integration (*int*), excision (*exc,* a topoisomerase) have been identified (also experimentally) as backbone in other ICEs (Sullivan *et al.*, 2002; Sutanto *et al.*, 2002; Gaillard *et al.*, 2006; Wozniak *et al.*, 2009). MAUVE was then used as guide for the identification of core genes. Genes identified by MAUVE to be conserved in 90% of ICEs were considered backbone genes. This relaxed criteria takes into account sporadic deletion events and misassembly. Genes that are not part of the backbone are considered cargo or accessory genes. These regions, consistently observed in specific positions across the ICEs, are called cargo gene (CG) regions and are described in detail below (Section 3.2.4).

The backbone consists of 65 genes predicted to be involved in ICE maintenance, regulation and movement, as well as a number of conserved hypothetical proteins (Table 3.2).

Gene number	Gene identity
1	*parA* - chromosome partitioning ATPase
2	hypothetical protein
3	hypothetical protein
4	*dnaB* - Replicative DNA helicase
5	hypothetical protein
6	hypothetical protein
7	hypothetical protein
8	*parB* - chromosome partitioning ATPase
9	hypothetical protein
10	hypothetical protein
	CG1
11	*ICE protein*
12	*Integrase regulator R*
13	*ssb* - single-stranded DNA-binding protein
14	hypothetical protein
	CG2
15	*DNA/RNA helicase*
	CG3
16	*pilL* - Conjugative transfer protein
17	*pilN* - Conjugative transfer protein

18	*pilO* - Conjugative transfer protein
19	*pilP* - Conjugative transfer protein
20	*pilQ* - Conjugative transfer protein
21	*pilR* - Conjugative transfer protein
22	*pilS* - Conjugative transfer protein
23	*pilU* - Conjugative transfer protein
24	*pilV* - Conjugative transfer protein
25	*pilM* - Conjugative transfer protein
	CG4
26	*topoisomeraseIII*
27	hypothetical protein
28	hypothetical protein
29	*tonB* - Periplasmic protein
30	hypothetical protein
31	hypothetical protein
32	hypothetical protein
33	hypothetical protein
34	hypothetical protein
35	adenylate methylase - Plasmid related protein
36	*DEAD box helicase* - Superfamily II DNA/RNA helicases
	CG5
37	Methyl-accepting chemotaxis protein
38	*ICE* - probable exported protein
39	Soluble lytic murein transglycosylase and related regulatory proteins
40	hypothetical protein
41	*traD* - Conjugative coupling factor
42	putative membrane protein
43	*helicase*
	CG6
44	*aconitase B*
45	hypothetical conjugal transfer protein
46	hypothetical protein
47	hypothetical protein
48	hypothetical protein
49	possible exported protein
50	hypothetical protein
51	putative exported protein
52	putative lipoprotein
53	*VirB4* - Type IV secretory pathway
54	hypothetical protein
55	*DSBA* - Protein-disulfide isomerase
	CG7
56	hypothetical protein
57	*traU* – ICE protein
58	hypothetical protein

59	hypothetical protein
60	*traG* – ICE protein
	CG8
61	*rulB* - Error-prone, lesion bypass DNA polymerase V
62	*rulA* - Error-prone repair protein
63	Protein unknown function
	CG9
64	*traI* - Pyruvate/2-oxoglutarate dehydrogenase complex
65	*xerC* - Integrase

Table 3.2. Backbone genes of the PsICE family. List of genes identified as backbone. The first column assigns a number to each gene, the second column identifies the gene and its function. CGs (cargo genes regions) are located in conserved intragenic loci, their gene content is variable.

The backbone genes of the PsICE family conserve predicted gene function and synteny, functioning as a scaffold for variability introduced at CG regions. All the genes of the backbone are orientated in the same direction 5' → 3' with the exception of *rulAB*. The relaxed criteria (presence in 90% of ICEs) used to define the core gene set means that there exist exceptions to the core set of backbone genes.

Deletions were identified in four PsICEs: Psa_{C18}ICE has a ~5.8 kb deletion encompassing hypothetical gene #32 (Table 3.1) and *superfamily II DNA/RNA helicase* gene #36. Ppa_{2367}ICE lacks ~9 kb of the backbone, possibly due to misassembly given that the missing region is present in the genome of *P. syringae* pv. *panici* (*Ppa*) LGM2367, albeit in a different contig. A deletion event in the Psa_{NZ45}ICE_Cu and Paf_{4394}ICE resulted in the loss of *rulA* (gene #62) and truncation of the genes up and downstream (*rulB* gene #61 and the hypothetical gene #63). *rulAB* and hypothetical gene #63 are also absent from Ps_{34876}ICE.

Four ICEs were not included in the 87 PsICEs analyzed in this study because of deletions affecting backbone genes that are likely to have crippled each ICE. For example, ICEs Por_{9088}ICE (LJQX01000005.1, LJQX01000019.1), Pzi_{8921}ICE (LJRT01000137.1) and the ICE at the *att-2* site in *P. cannabina* (FNKU01000001.1)

have lost the *pil* cluster (from gene #16 to #25). More contigs (which couldn't be reconstructed into an entire ICE) with the same backbone loss are found in: *P. syringae* pv. *maculicola* (*Pma*) 7930 (LGLD01000027.1), *Pma* M4a (LGLE01000020.1), *P. savastanoi* pv. *glycinea* (*Pgy*) BR1 (LGLO01000042.1), *P. amygdali* pv. *myricae* 2897 (LIHY01000179.1), *Psa* C9 (ANJI02000003.1), *Psa* C54 and C78 (Rainey Lab) *P. cannabina* has also the *att-1* site occupied by an ICE (FNKU01000001.1) which carries deletion encompassing the genes from #1 (*parA*) to #15 (*DNA/RNA helicase*) and from gene #17 (*pilN*) to #25(*pilM*). An uncharacterized phage-encoded protein is present in the same location (between gene #12 and #13) in all but one of the ICEs: Por_{9088}ICE.

3.2.4 Identification of regions harbouring cargo genes

Genes in PsICEs that are not part of the backbone are considered cargo or accessory genes. Comparison of the 28 non-redundant PsICEs revealed cargo genes in eight specific positions across the ICEs (Figure 3.4, Table 3.2). These positions were defined as regions harbouring cargo genes (CGs): loci of variability where DNA appears to preferentially integrate. Some of these variable loci are apparently ancient (e.g. CG1 and 2) with divergence having occurred through time, and probably reflecting that carriage is (or has been) useful to the ICEs. In CG1, 13 of the non-redundant PsICEs harbour a predicted haemin uptake outer membrane receptor, while the other 15 carry a hypothetical protein. In CG2, 23 PsICEs harbour a hypothetical protein predicted to belong to ICEs clusters, while the remaining 5 ICEs carry a shorter version of the same hypothetical gene. Other CGs

(e.g. CG4) appear to be of recent (and repeated) acquisitions of the same genes (e.g. the transposon Tn*6212*, see Chapter 4).

Tn*6212* (Butler *et al*., 2013), also referred to as the enolase region (McCann *et al*., 2013), is the most common element found in the CGs, and is integrated exclusively into CG4 (in purple in Figure 3.4). Only 8 of the 28 non-redundant PsICEs lack Tn*6212* (Psa_{NZ45}ICE, Psa_{NZ47}ICE, Psa_{NZ64}ICE, Psa_{C66}ICE, Pph_{1302A}ICE, Psy_{SM}ICE, Psy_{3023}ICE, Ps_{34876}ICE and Ps_{39023}ICE).

Genes encoding resistance to heavy metals are recurrent in different ICEs. Recovery of CuR in *Psa* ICEs could be due to the nature of the sampling of those strains (Chapter 5, Colombi *et al*., 2017). A transposon coding for a methyl-accepting chemotaxis protein (Tn*6211*, Butler *et al*., 2013) (in red in Figure 3.4) is present in 7 of 28 non-redundant PsICEs, though integrated in different positions. Paf_{4394}ICE and Ps_{34876}ICE harbour ~7 kb transposons encoding mercury resistance integrated at the end of the CG4 and inside Tn*6212*, respectively. These transposons are not identical, and both display 96% pairwise nucleotide identity with Tn*5042* in *P. fluorescens* ED94-62a (AJ563380.2). Tn*5042* is a novel type of mercury resistance (HgR) transposon identified in both permafrost strain collections and present-day bacteria (Mindlin *et al*., 2005). Toxin and antitoxin systems and genes involved in metabolism are present in the CGs. Pph_{1302A}ICE is the only PsICE carrying an effector gene (HopAR1).

For an overall panorama of the 87 PsICEs structure, distribution, and diversity, refer to Figure 3.5.

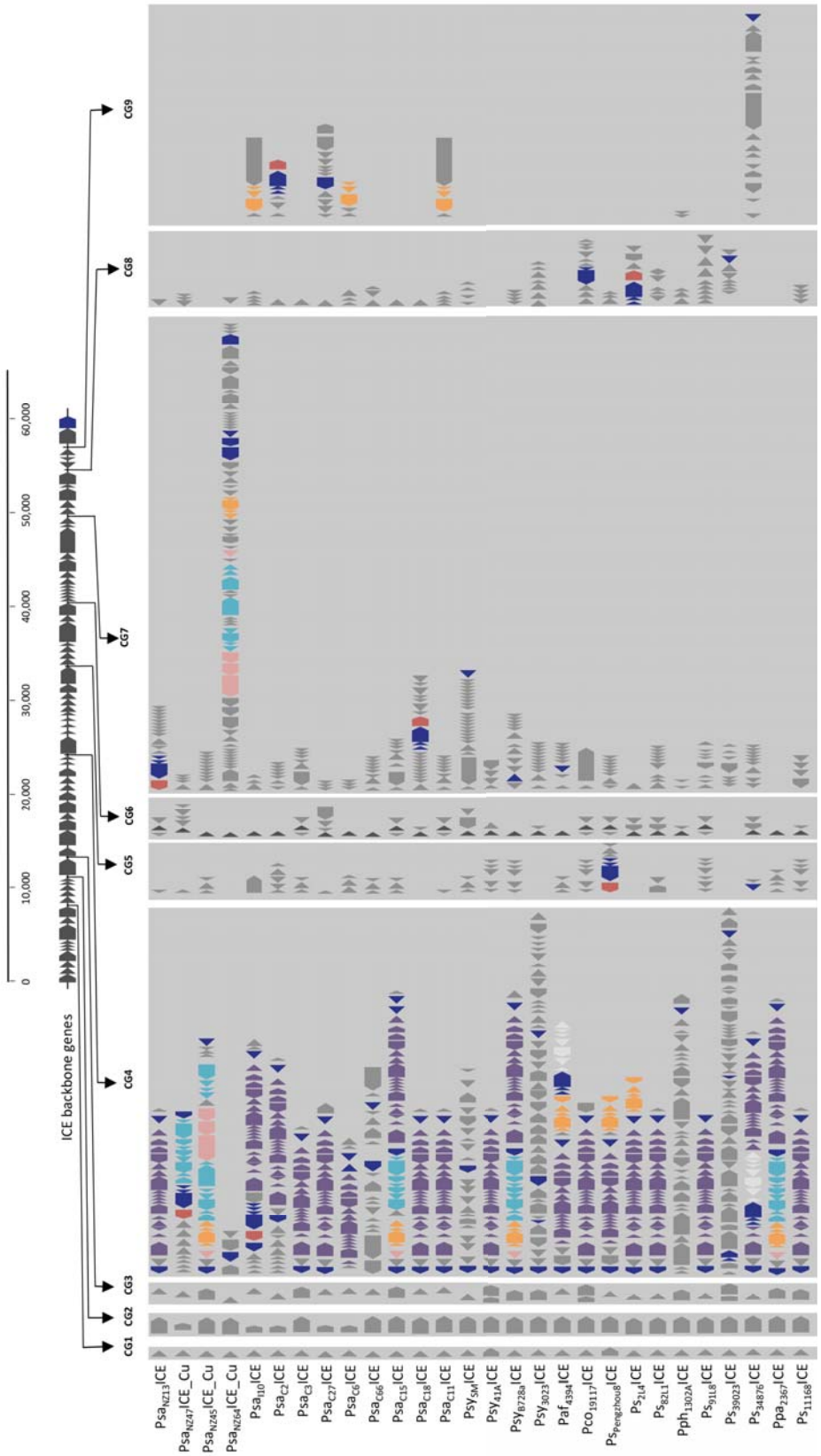

Figure 3.4. Genetic organization of the CG of the non-redundant PsICEs. The gene content of the CG is highlighted with the grey background; arrows indicate their position on the backbone. Backbone genes are depicted in black, blue boxes are mobile genes (transposases or recombinases), purple boxes define the 'enolase region', azure boxes depict copper resistance genes, orange boxes are arsenic resistance genes, pink boxes are genes belonging to the *czc*/*cus* system, white boxes denote genes encoding mercury resistance, red boxes are MCPs, grey boxes depict other genes.

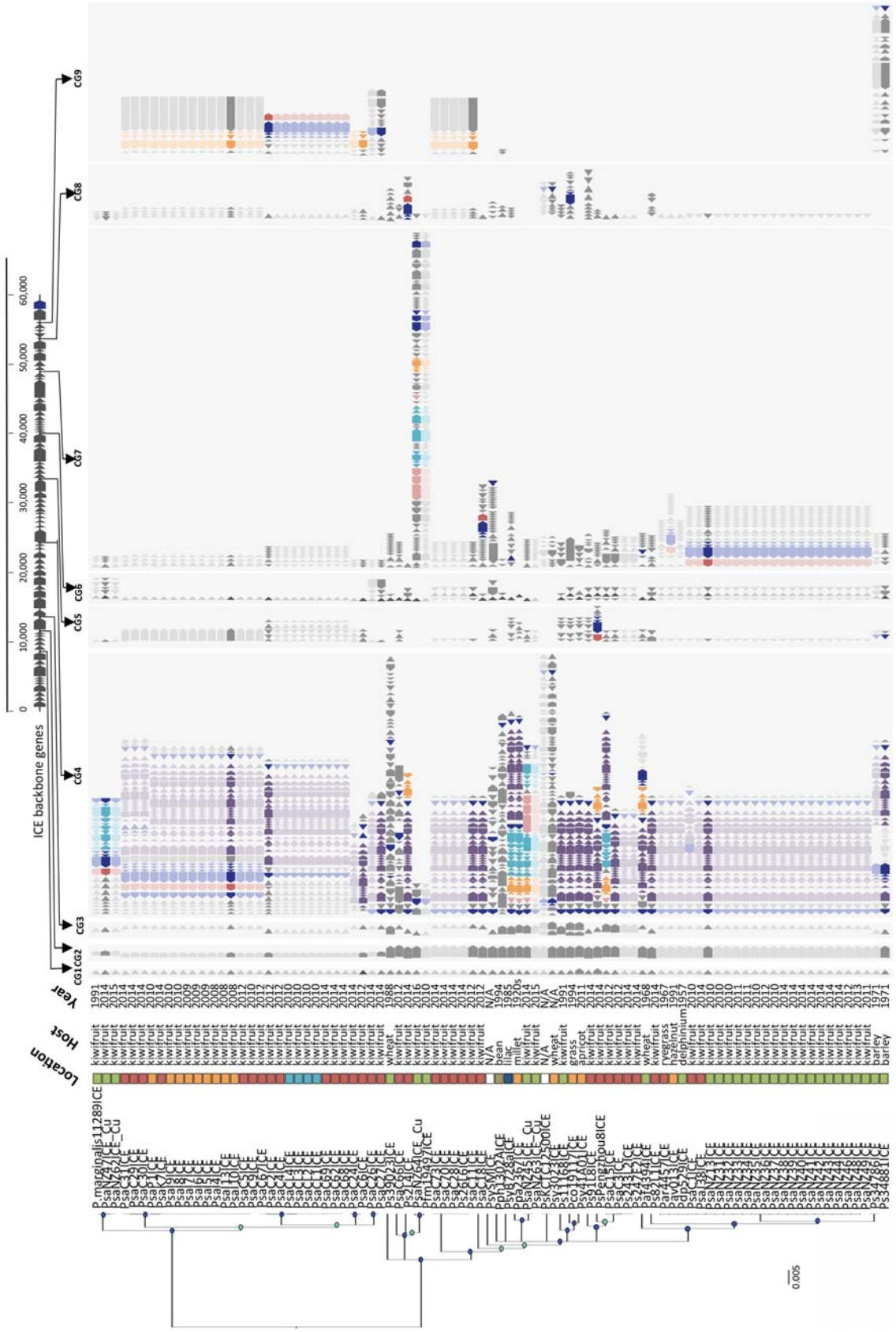

Figure 3.5. The PsICEs. UPGMA tree based on the REALPHY alignment of the PsICEs as in Figure 3.1 and their genetic organization as in Figure 3.4.

3.2.5 Recombination clouds the evolutionary history of PsICEs

Even though the PsICEs code for the same backbone genes, saturation at the third codon position between the PsaI10 and PsaNZ13-like ICEs indicates the backbone genes have an ancient evolutionary past (McCann *et al.*, 2013). For example, their *DEAD box helicases* (a gene of 2,265 bp) differ in 159 SNPs (110 synonymous), of which 114 are in the third codon position. The average variation in nucleotide identity for each backbone gene of the unique PsICEs is displayed in Figure 3.6A. Genes that comprise the backbone share between 95.4% (*DEAD box helicase*) and 70.6 % (*xerC*) pairwise nucleotide identity, with a mean between genes of 88.57%. Between the distantly related Psa_{NZ13}ICE and Psa_{I10}ICE the pairwise identity ranges between 44.4% (backbone gene #30) and 92.8% (*DEAD box helicase*) with an average of 76.4% (Figure 3.6B). The GC% for the concatenation of the backbone genes for all the PsICEs varies between 55.9% and 56.2%.

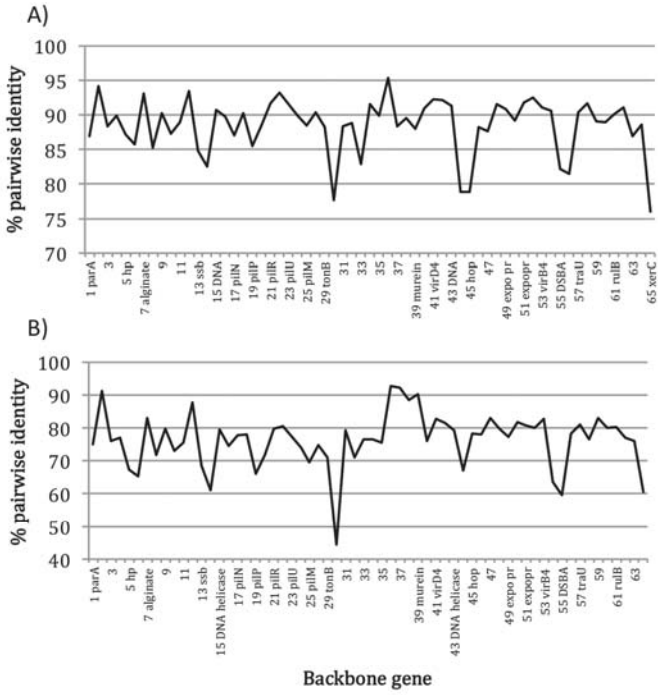

Figure 3.6. Pairwise identity of the backbone genes. In A) each backbone gene of each non-redundant PsICEs was aligned in Geneious and the % pairwise identity plotted in the graph, in B) the % pairwise identity between each backbone gene of Psa_{I10}ICE and Psa_{NZ13}ICE only was considered.

To explore the extent to which conserved genes within ICEs show evidence of vertical descent, the phylogeny of *parA*, *topoisomerase III* and *xerC* (located at the beginning, middle and end of the backbone respectively) was analysed (Figure 3.7).

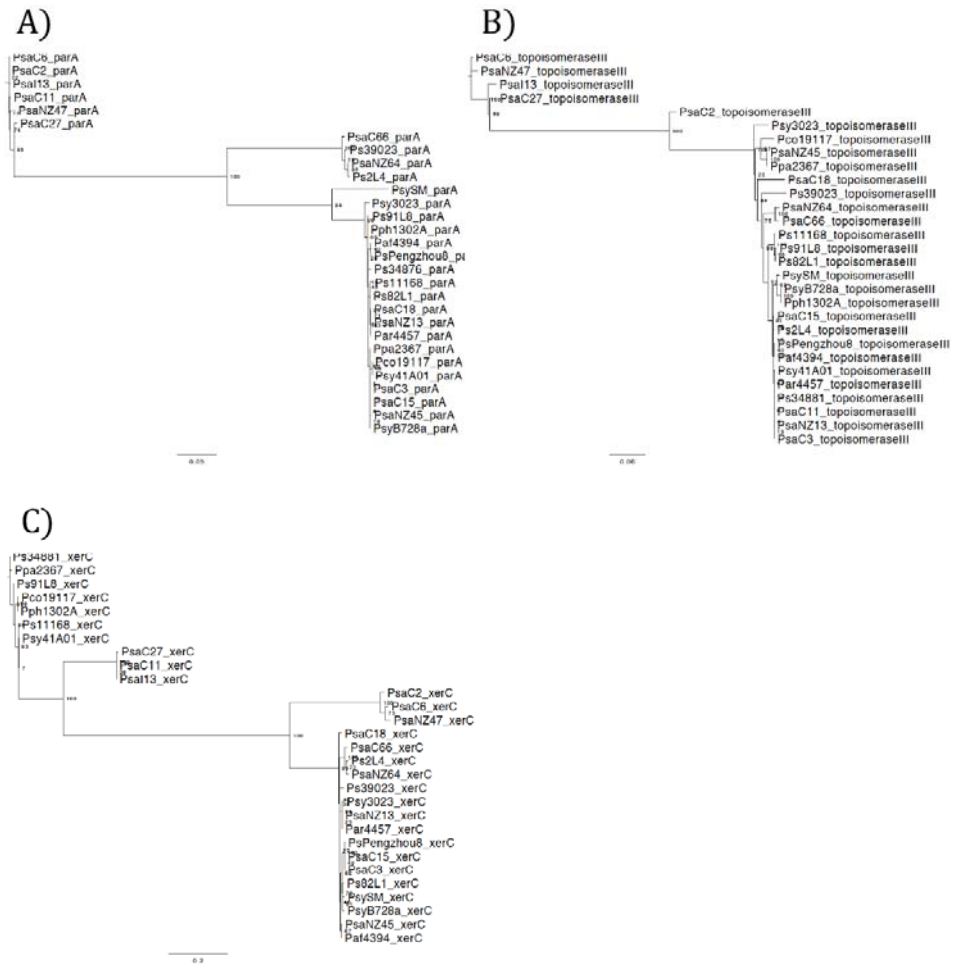

Figure 3.7. Phylogenetic incongruence of the backbone. PhyML trees of A) *parA,* B) *topoisomeraseIII* and C) *xerC* of the non- redundant PsICEs show incongruence.

PhyML trees of A) *parA*, B) *topoisomeraseIII* and C) *xerC* show incongruence, indicative of recombination. Additionally, the incongruence length difference (ILD) test (Farris *et al*., 1994) was carried out on 13 segments of the backbone identified by MAUVE using PAUP* (Swofford, 2003). PAUP* performs a

partitioning homogeneity test (i.e., ILD test) to evaluate phylogenetic congruence between each tree of the 13 backbone segments identified by MAUVE; each pair was found to be incongruent, with probability value (*P*) of type I error (falsely rejecting the true null hypothesis of congruence) of 0.01.

The neighbor-net network was used to portray the relationship among the backbone genes of the 28 non-redundant PsICEs (Figure 3.8). The extent of reticulation of the concatenation of backbone genes indicates that recombination has shaped PsICE diversity. The Phi test for recombination (Bruen *et al.*, 2006) was statistically significant ($P<0.0001$).

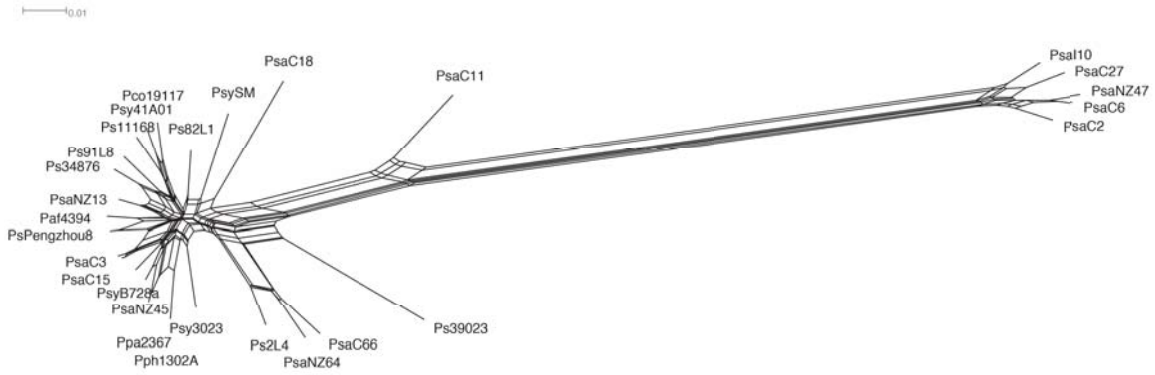

Figure 3.8. Recombination in the PsICEs. Neighbor-net network of the backbone genes of the 28 non-redundant PsICEs. The alignment on which the tree is based on is a MAFFT alignment of the backbone genes created with Geneious.

To further understand the extent and the pattern of recombination, concatenation of the backbone genes of the 28 non-redundant PsICE was analysed with Alfy (Domazet-Lošo & Haubold, 2011). Alfy describes the chimeric structure of a set of sequences, not relying on alignments, but on the lengths of exact matches between pairs of sequences. The program was run selecting a default sliding window (300 bp), selecting only strong homologies (regions where the average shortest unique substring length is greater than the maximum shortest unique substring length occurring by chance alone). Visualization of the recombination patterns reveals the PsICEs are chimeric (Table 3.3). The minimal length of homology for the

inclusion in Table 3.3 was 10% (inferior values are shown when the 10% length cut off is not reached.) Each PsICE displays minimal stretches of sequence displaying no homology with any other PsICE.

PsySM		Psy3023		Ps39023		PsaNZ13		Pco19117		Ppa2367		PsaC3	
ICE_id	% of homology	ICE_id	% of homology	ICE_id	% of homology	ICE_id	% of homology	ICE_id	% of homology	ICE_id	% of homology	ICE_id	% of homology
Ps11168	9.3	Ps82L1	8.7	Pph1302A	15	PsaC18	32.5	Psy41A01	38.1	PsaNZ45	85.4	PsaC15	28.9
nh	2.3	nh	1.2	nh	6.9	Ps34876	22.8	Ps11168	11.2	nh	0.5	PsPengzhou8	14.4

Ps2L4		Ps82L1		Ps91L8		PsaC11		PsaNZ45		PsaC66		Ps34876	
ICE_id	% of homology	ICE_id	% of homology	ICE_id	% of homology	ICE_id	% of homology	ICE_id	% of homology	ICE_id	% of homology	ICE_id	% of homology
PsaNZ64	28.8	Ps11168	15.4	82L1	22	PsaI10	13.9	Ppa2367	63.1	PsaNZ64	46.3	PsaNZ13	28.2
PsaC66	27.7	91L8	11.1	PsaNZ13	11.2	PsaNZ13	10.3	Paf4394	11.8	2L4	27.9	Ps82L1	21.4
nh	1.2	nh	0.5	nh	0.3	nh	1.3	nh	0.3	nh	1.2	nh	0.4

PsaNZ64		Pph1302A		Paf4394		PsPengzhou8		Psy41A0i		PsaC18		PsyB728a	
ICE_id	% of homology	ICE_id	% of homology	ICE_id	% of homology	ICE_id	% of homology	ICE_id	% of homology	ICE_id	% of homology	ICE_id	% of homology
PsaC66	41.6	Ps39023	10.8	PsPengzhou8	51.1	Paf4394	52	Pco19117	29.5	PsaNZ13	27.3	PsaC15	18.1
Ps2L4	28.7	PsyB728a	10.2	PsaNZ45	10.2	PsaC3	14.7	Ps11168	22.4	Paf4394	11.3	PsaNZ45	15.8
nh	0.6	nh	1.3	nh	0.2	PsaC15	11.3	Ps2L4	13.8	PsySM	10.5	Paf4394	10.5
								nh	0.7	nh	10.6	nh	0.5

Ps11168		PsaC2		PsaC6		PsaNZ47		PsaC15		PsaC27		PsaI10	
ICE_id	% of homology	ICE_id	% of homology	ICE_id	% of homology	ICE_id	% of homology	ICE_id	% of homology	ICE_id	% of homology	ICE_id	% of homology
Ps82L1	23.3	PsaC6	32	PsaNZ47	43.9	PsaC6	48.7	PsaC3	27.5	PsaI10	25.7	PsaC27	19.8
Psy41A01	21.2	PsaI10	23.5	PsaC2	22	PsaI10	19.2	PsPengzhou8	13.8	PsaC6	22.2	PsaNZ47	19.6
Pco19117	10.1	PsaC27	16.6	PsaC27	15.5	PsaC27	13.8	Ppa2367	11.5	PsaC2	20.2	PsaC2	19.3
nh	1.8	PsaNZ47	15.6	PsaI10	15.1	PsaC2	12.5	PsyB728a	11.2	PsaNZ47	16.5	PsaC6	17.1
		nh	2	nh	1	nh	0.7	PsaNZ45	11.1	nh	1.1	PsaC11	13.9
												nh	1.7

Table 3.3. Recombination pattern in the PsICEs. Alfy was used to visualise the % of recombination length for each backbone of non-redundant PsICE. nh stands for no homology.

For example, Psy_{B728a}ICE displays homology in 18.1% of its length with Psa_{C15}ICE, in 15.8% with Psa_{NZ45}ICE_Cu and in 10.5% with Paf_{4394}ICE; 0.5% of Psy_{B728a}ICE displays no homology to any other PsICE (Table 3.4). The alignment of the backbone of Psy_{B728a}ICE with the homologous backbones identified by Alfy is showed in Figure 3.9.

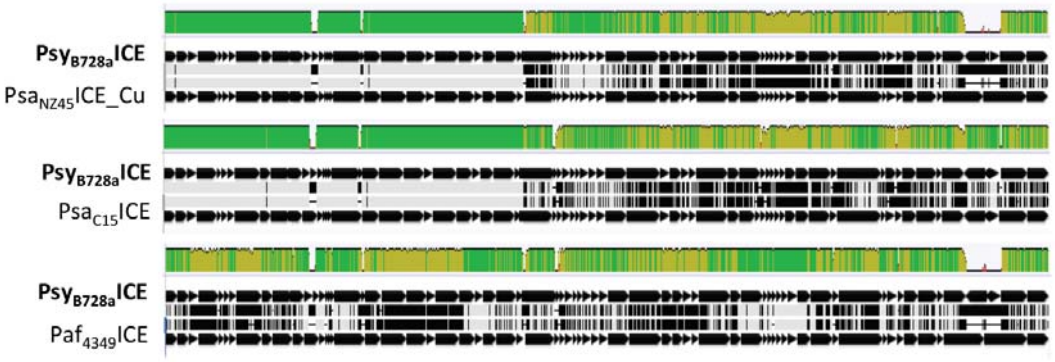

Figure 3.9. Chimerism in the Psy_{B728a}ICEs backbone. MAFFT alignment in Geneious of backbone genes only (depicts as black boxes). Identity between sequences is represented in green. Alfy shows Psy_{B728a}ICE to display homology in 18.1% of its length with Psa_{C15}ICE, in 15.8% with Psa_{NZ45}ICE_Cu and in 10.5% with Paf_{4394}ICE; 0.5% of Psy_{B728a}ICE displays no homology to any other PsICE (see Table 3.3 for details).

The extent of recombination was quantified using ClonalFrameML (Didelot & Wilson, 2015). ClonalFrameML identified recombination events in most of the backbone sequences (Figure 3.10). The removal of the sequences affected by recombination from the alignment of concatenated backbone genes (~65 kb) resulted in an extremely short alignment (~7.5 kb).

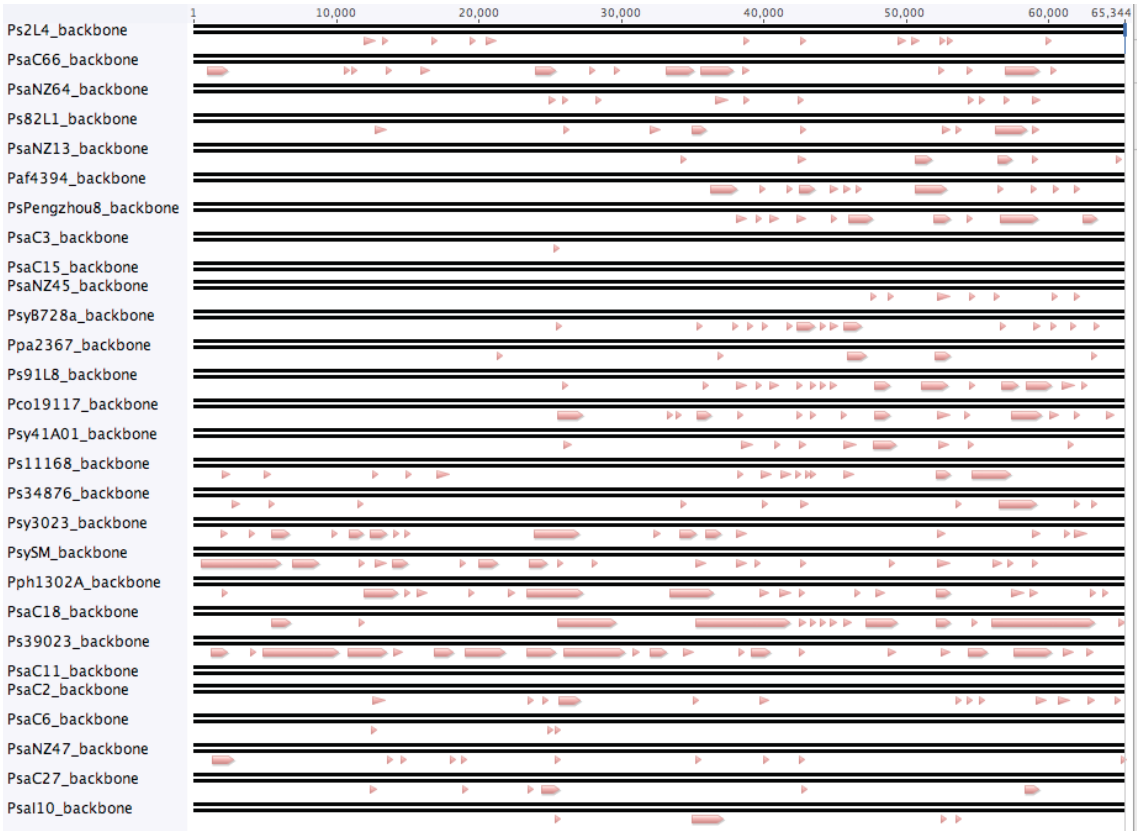

Figure 3.10. Recombination events identified by ClonalFrameML. The alignment of concatenation of the backbone sequences of the non-redundant PsICEs was analysed using ClonalFrameML to identify recombination events. The black doubled line represents the sequence of the alignment of the backbone genes of each non-redundant PsICEs, pink boxes depict recombination events identified by ClonalFrameML. When the recombination events are manually removed from every sequence, from ~65 kb, the resulting alignment is ~7.5 kb.

3.2.6 Distribution and Comparison of the PsICEs

The distribution of a single PsICE between different strains spans across time and space and more interestingly each PsICE itself is a chimera of other PsICEs that can be present in strains from vastly different origins. In this Section I

will use a comparison of three PsICEs as examples of what has been described in the previous Section (Figure 3.11).

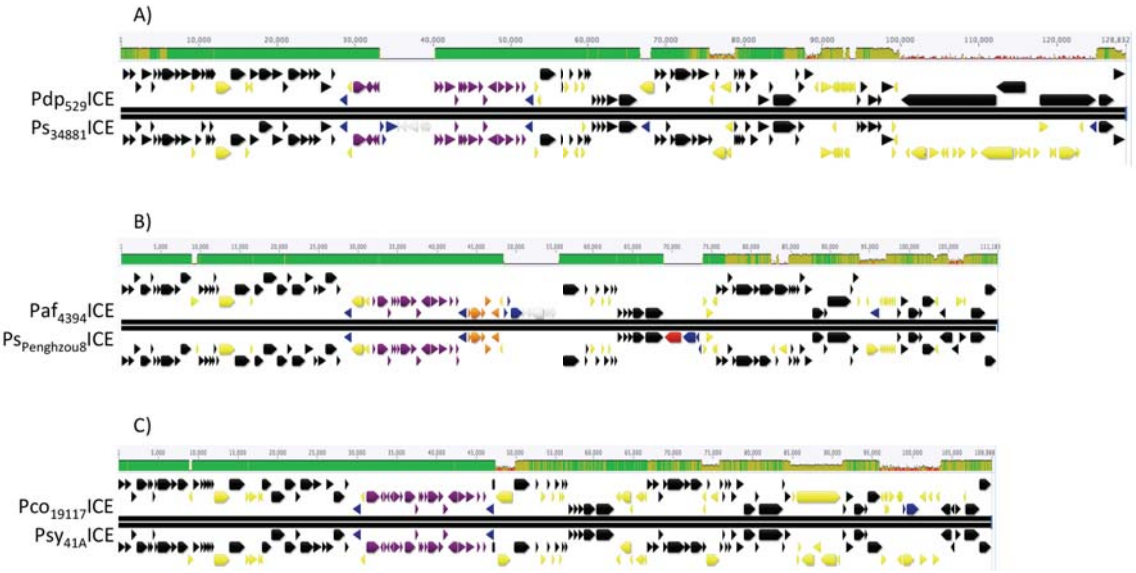

Figure 3.11. Chimerism in the PsICEs. MAFFT alignment in Geneious. Black depicts backbone genes, blue are mobile genes (transposases or recombinases), purple defines the 'enolase region', orange boxes are arsenic resistance genes, white boxes denote genes encoding mercury resistance, red box is a MCP, yellow depicts other genes. Identity between sequences is represented in green.

Pdp 529 was isolated in NZ in 1957 from larkspur while *Ps* BRIP34881 was isolated in Australia in 1971 from barley. Their PsICEs share 98.2% pairwise identity (103 SNPs) across the first 5.6 kb, then near perfect identity across the following 84 kb, but for 4 SNPs and two insertion events (a transposase and HuR transposon) in Ps_{34876}ICE. The ICEs share only 47.3% pairwise identity in the final section of the mosaic (Figure 3.11A). *P. syringae* pv. *atrofaciens* IMCP4394 (*Paf* 4394) was isolated in 1968 from wheat in NZ while *Ps* Penghzou8 was isolated from grassland in China (2012). Their PsICEs differ by 10 SNPs in the first 77 kb. Paf_{4394}ICE also has a 633 bp deletion in the CG1 and HuR transposon insertion, while Tn*6211* is present in $Ps_{Penghzou8}$ICE. Paf_{4394}ICE and $Ps_{Penghzou8}$ICE exhibit very low conservation across the next 35 kb, with upwards of 1,000 SNPs, yet $Ps_{Penghzou8}$ICE shares high similarity with Psa_{C3}ICE in this same region (18 SNPs in ~16 kb of backbone sequence) (Figure 3.11B). *Psy* 41A was isolated from apricot

in France (2011) and *P. congelans* ICMP19117 (*Pco* 19117) from grass in Germany (1994), yet their PsICEs differ by only 3 SNPs and a 288 bp indel across 47 kb. Homology ends before the last gene of CG4 after Tn*6212* (Figure 3.11C). A similar breakpoint is present between Psa_{C15}ICE and Psy_{B728a}ICE, and during experiments involving inter-strain ICE transfer, where PsICE recombination was directly observed (Chapter 5; Colombi *et al*., 2017).

3.2.7 Common origin of PsICE and PAPI-1

The extent of recombination masks phylogeny among members of the PsICE family. However, synteny in the backbone indicates a common origin that can be investigated. Comparison of each backbone gene of the PsICEs revealed *DEAD box helicases* to be the most conserved with 95.4% nucleotide identity. To search for more distantly related ICEs, the protein sequence of this gene was compared against all microbial genomes available as of April 2017 (*P. syringae* spp. excluded) using tBLASTn, with a cut off of 81% of identity. With this selection criterion 45 DEAD box helicase with at least 81% of identity with the DEAD box helicase of Psa_{NZ13}ICE were found. These 45 proteins are distributed in 41 genomes, 34 of which belonging to *P. aeruginosa* (Table 3.4).

Name	Strain ID	Isolation	Gene Bank accession	ICE family	*att* site	% of homology
P agarici NCPPB2472	*P. agarici* strain NCPPB 2472	mushroom	CP014135.1	NA	*queC*	81%
P chlororaphis PA23	*P. chlororaphis* strain PA23	Soy bean root	CP008696.1	PAPI-1†	*queC*	86%
P mosselii SJ10	*P. mosselii* SJ10	wastewater	CP009365.1	PAPI-1†	*queC*	81%
Pa B10W	*P. aeruginosa* isolate B10W	wastewater	CP017969.1	PAPI-1	*queC*	82%
Pa B136-33	*P. aeruginosa* B136-33	human	CP004061.1	PAPI-1	*clpB*	83%
pKLC102	*P. aeruginosa* strain C	human	AY257538.1	PAPI-1	NA	82%
Pa FA-HZ1	*P. aeruginosa* strain	wastewater	CP017353	PAPI-1	*queC*	83%

	FA-HZ1					
Pa MTB-1	*P. aeruginosa* MTB-1	Soil	CP006853.1	PAPI-1	*clpB*	82%
Pa NCGM2.S1	*P. aeruginosa* NCGM2.S1	human	AP012280.1	NA	*queC*	82%
Pa PA1	*P. aeruginosa* PA1	human	CP004054.2	PAPI-1	*queC*	82%
PAPI-1	*P. aeruginosa* PA14	human	AY273869.1	PAPI-1	*clpB*	81%
Pa PA14Or	*P. aeruginosa* isolate PA14Or	NA	LT608330.1	PAPI-1	*clpB*	81%
Pa PA1R	*P. aeruginosa* PA1R	NA	CP004055.1	PAPI-1	*queC*	82%
Pa PA38182	*P. aeruginosa* PA38182	human	HG530068.1	NA	NA	83%
Pa PA7	*P. aeruginosa* PA7	human	CP000744.1	PAPI-1	*queC*	83%
Pa PA96	*P. aeruginosa* PA96	human	CP007224.1	PAPI-1	*queC*	82%
PAGI-5	*P. aeruginosa* PSE9	NA	EF611301.1	PAPI-1	NA	83%
Pa SGVI ST111	*P. aeruginosa* SGVI ST111	human	KT887560.1	PAPI-1	NA	82%
Pa 8380	*P. aeruginosa* strain 8380	human	AP014839.2	PAPI-1	*clpB*	82%
Pa BAMCPA07-48	*P. aeruginosa* strain BAMCPA07-48	human	CP015377.1	PAPI-1	*clpB*	82%
Pa Carb01 63 I	*P. aeruginosa* strain Carb01 63	human	CP011317.1	PAPI-1	*clpB*	83%
Pa Carb01 63 II	*P. aeruginosa* strain Carb01 63	human	CP011317.1	PAPI-1	*clpB*	83%
Pa DN1	*P. aeruginosa* strain DN1	soil	CP017099.1	NA	*queC*	81%
Pa F23197	*P. aeruginosa* strain F23197	human	CP008856.1	PAPI-1	*queC*	82%
Pa F30658 II	*P. aeruginosa* strain F30658	human	CP008857.1	PAPI-1	*queC**	82%
Pa F30658 I	*P. aeruginosa* strain F30658	human	CP008857.1	NA	*clpB*	82%
Pa F63912	*P. aeruginosa* strain F63912	human	CP008858.1	PAPI-1	*clpB*	82%
Pa H27930	*P. aeruginosa* strain H27930	human	CP008860.1	PAPI-1	*queC**	82%
Pa NCGM 1900	*P. aeruginosa* strain NCGM 1990	human	AP014622.1	NA	*queC*	82%
Pa NCGM 1984	*P. aeruginosa* strain NCGM 1984	human	AP014646.1	NA	*queC*	82%
Pa NCGM257	*P. aeruginosa* strain NCGM257	human	AP014651	PAPI-1	*clpB*	82%
Pa PA1RG	*P. aeruginosa* strain PA1RG	hospital sewage	CP012679.1	PAPI-1	*queC*	82%
Pa PSE305	*P. aeruginosa* strain PSE305	sheep	HG974234.1	PAPI-1	*clpB*	82%
Pa S04 90 I	*P. aeruginosa* strain S04 90	human	CP011369.1	PAPI-1	*clpB*	82%
Pa S04 90 II	*P. aeruginosa* strain S04 90	human	CP011369.1	NA	*queC*	82%
Pa S86968 II	*P. aeruginosa* strain S86968	human	CP008865.1	PAPI-1	*queC*	82%
Pa S86968 I	*P. aeruginosa* strain S86968	human	CP008865.1	PAPI-1	*clpB*	83%
Pa T63266	*P. aeruginosa* strain T63266	human	CP008868.1	PAPI-1	*queC*	82%
Pa W16407	*P. aeruginosa* strain W16407	human	CP008869.1	PAPI-1	*clpB*	82%
Pa UCBPP-PA14	*P. aeruginosa* UCBPP-PA14	human	CP000438.1	PAPI-1	*clpB*	81%
Pa VRFPA04	*P. aeruginosa* VRFPA04	human	CP008739.2	NA	*queC*	83%
PA W45909	*P. aeruginosa* strain W45909	human	CP008871.1	PAPI-1	*queC*	83%
PA YL84	*P. aeruginosa* YL84	compost	CP007147	PAPI-1	*queC*	82%
Pf-5 PAPI	*P. protegens* Pf-5	rizhosphere	CP000076.1	PAPI-1	*queC*	83%
P. sp. WCS374	*P.* sp. WCS374	rhizosphere	CP007638.1	PAPI-1†	*clpB*	83%

Table 3.4. List of elements harbouring the *DEAD box helicase* similar to Psa$_{NZ13}$ICE. (* tandem integration; †not canonical).

With the exception of one sequence (Pa PA38182), all the *DEAD box helicases* belong to mobile elements integrated in one of the two *att*-site nearby *clpB* and *queC* described for the PsICEs. Although pKLC102 was deposited in Genbank as a plasmid, it can integrate into one of the two *att*-sites (Klockgether *et al.*, 2007). 33 of these are ICEs belonging to the PAPI-1 family of ICEs (Table 3.4).

PAGI-5 and PsaNZ13 were picked as a representative of each family of ICE to perform an alignment with MAUVE (Figure 3.12). The alignment shows that two ICEs show ICE specific regions but also regions of homology where synteny is conserved with the exception of the yellow block (corresponding in Psa$_{NZ13}$ICE to *rulAB* and gene #63).

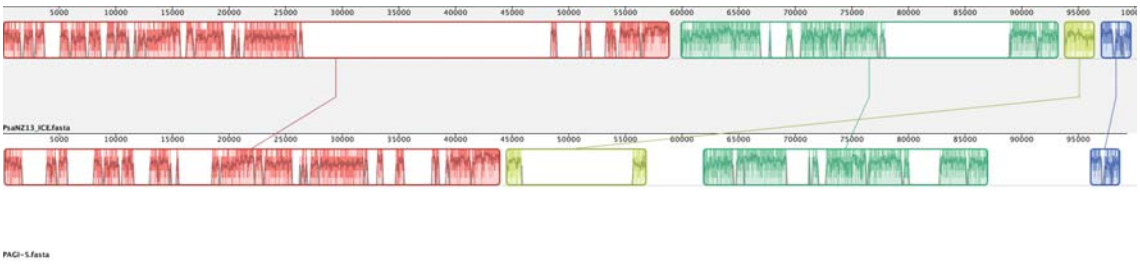

Figure 3.12. Psa$_{NZ13}$ICE and PAGI-5 comparison. Alignment of Psa$_{NZ13}$ICE and PAGI-5 produced by MAUVE using seeds family and seed weight 11. Colour-coded blocks depict regions of homology that are connected through lines. Inside each block a similarity profile of the sequence is drawn, the height of the similarity profile corresponds to the average level of conservation in that region of the sequence. Regions outside blocks were not aligned and probably contain sequence elements specific to each ICE.

The regions of homology identified by MAUVE map in the Psa$_{NZ13}$ICE exclusively in backbone genes (although not in all of them) and in the CG1. *topoisomeraseIII* wasn't recognized as a region of homology although *topoisomeraseIII* is present in PAGI-5 upstream the *pil* cluster and *DEAD box helicase*.

To infer the phylogenetic relationship between the PsICEs and PAPI-1, a concatenation of the protein alignment of ParA, TopoisomeraseIII and XerC of the PsICEs and all ICEs listed in Table 3.4 was used. The resulting phylogenetic tree is

shown in Figure 3.13. Two separated clusters are evident when these families of ICEs are compared, suggesting a common ancestor.

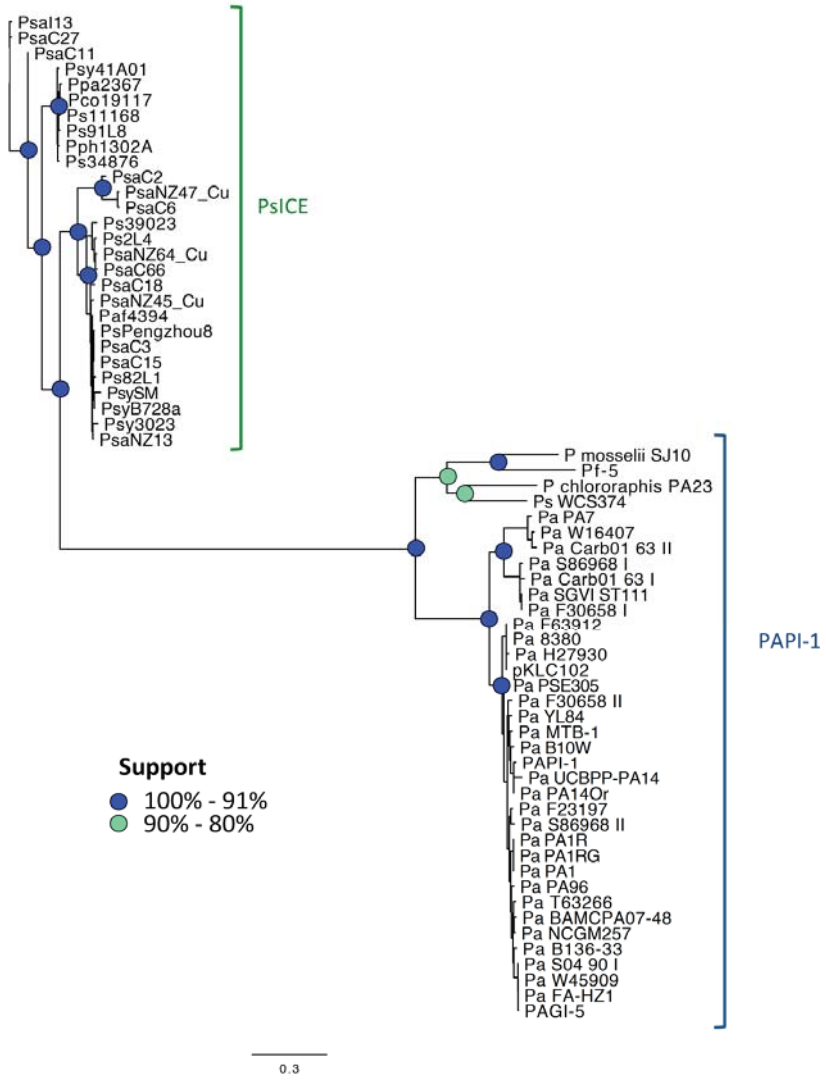

Figure 3.13. Evolutionary relationship between the PsICEs and PAPI-1. PhyML tree based on the concatenation of the alignment of ParA, TopoisomeraseIII and XerC.

3.3 Discussion

Mobile elements play an important role in the horizontal dissemination of new traits across bacteria. The initial discovery of ICEs was related to the phenotypic change caused by cargo genes (Stuy, 1980; Franke & Clewell, 1981; Rashtchian, 1982; Sullivan & Ronson, 1998), and they became of increasing interest due to their role in disseminating antibiotic resistance. But this is just the tip of the ICEberg (Burrus *et al.*, 2002). With the increasing number of strains being sequenced it is possible to find new ICEs independently of their encoded phenotypes, simply by searching for homology in DNA sequence databases.

ICEs are thought to derive from plasmids that eventually became part of the host genome (Novick & Ram, 2016). Novick & Ram (2016) suggest that the basic genome of ICEs is derived primarily from plasmids because ICEs transfer via conjugation as plasmids, but they differ from the latter as they start their travel as part of the chromosome, and because some ICEs may replicate autonomously as plasmids.

Despite not limiting searches on GenBank and using permissive BLASTn search parameters, only ICEs present in plant-associated *Pseudomonas* spp. were found. It appears that ICE presence is more strongly affected by ecology than physical and temporal distance. PsICEs are not distributed according to year or geographical region of origin, yet they share the same ecological niche: plants. This has also been observed over larger scales (Fondi *et al.*, 2016). Moreover, being horizontally transmitted elements, PsICE distribution does not reflect the underlying phylogeny of the strains harbouring them. This emphasizes the

importance of not drawing conclusions on strain phylogeny based on ICE content, as for example, as was done by Butler *et al*. (2013).

The diversity of the PsICEs can be represented using only the backbone genes or the entire sequence of the ICE. In Section 3.2.5 I used a concatenation of the backbone to produce a neighbor-net network using Splitstree. However, the use of only selected backbone genes is not a straightforward process. Backbone genes need first to be identified, extracted, aligned and concatenated. REALPHY instead uses the entire sequence of PsICEs with the additional advantage of allowing lists of ICEs to be updated with new sequences with minimal difficulty. Independently by the two methodologies, two main groups of ICEs are evident: on one side there are the PsaI10, PsaC2, PsaC6 and PsaC27-like ICEs and the CuR ICE in *Psa* NZ47. On the other side, but connected through recombination, are the remaining ICEs.

The diversity of PsICEs is not only driven by the sequence divergence of the backbone, but also by the presence of cargo genes. Cargo genes often reflect the environment where the bacterial host lives. For instance, the Xaj-ICE harboured by *Xanthomonas arboricola* strains (pathogenic or non-pathogenic strains isolated from cultivated walnut trees) has CuR genes (Cesbron *et al*., 2015). Copper is a very common bacteriocide used in agriculture (see Chapter 5). Interestingly, ICEs of the STX-R39 family in *Vibrio cholerae* harbour antibiotic resistance genes, while cargo genes in ICEs of the same family isolated from the free living marine bacteria *Alteromonas* mainly code for restriction modification or metal resistance systems (as Cu and Hg) (López-Pérez *et al*., 2017). This could reflect differences in host/environment ecology: a human body exposed to antibiotics versus polluted marine environment.

Cargo genes are found in fixed intergenic loci across the ICE backbone. The existence of fixed positions that harbour variable loci suggests that other positions across the backbone are 'prohibited'. Also, the maintenance of synteny in PsICEs suggests that different arrangements in the backbone are 'rejected'. In the Selfish Operon Model elaborated by Lawrence & Roth (1996), genes subjected to weak or no selection escape elimination by physical proximity, which facilitates the invasion of new genomes (HGT). Consequently, HGT selects for genes performing similar or a single function to be clustered together, i.e. in operons. The backbone genes don't provide a selective advantage to the host, but just to the ICE itself. If these sets of genes are together performing a single function (e.g. conjugation or replication), an integration or recombination event that destroys their functionality will be selected against. The 'rejection' of breaking these sets of genes suggests that they work as a unit. If the ICE is crippled because one of these units is broken, it is likely that it will be eliminated over time.

Some CG regions seem to be more prone to acquiring foreign genes than others. CG4, for example, has repeatedly acquired Tn*6212*, e.g., as in the distantly related PsaI10 and PsaNZ13-like ICEs, but also in the same cluster of PsICEs as in the PsaNZ13-like ICEs (refer to Chapter 4). At the same time CG4 can also harbour a diverse set of genes, e.g. in Psy_{39023}ICE and Pph_{1302A}ICE. The presence of the same element (Tn*6212*) in the same position (CG4) suggests a sequence based targeting of the integration locus, although the backbone sequence of the ICEs is quite divergent. On the contrary, Tn*6211* shows more flexibility in terms of integration target but still is found in one of the CGs, integrated either in CG3, 4, 5, 7, 8 or 9.

Tn*6212*, also known as the enolase region (McCann *et al.*, 2013) is present in 20 of 28 non-redundant PsICEs. Tn*6212* is a ~16 kb tyrosine recombinase

transposon which comprises a set of 17 open reading frames including a predicted enolase and a C4-dicarboxylate transporter (*dctT*). Together the two genes are predicted to manipulate the plant-host metabolism (McCann *et al.*, 2013). Tn*6212* is found in almost identical forms in Psa_{I10}ICE, Psa_{NZ13}ICE and Psa_{CL1}ICE (Butler *et al.*, 2013; McCann *et al.*, 2013) and in each PsICE, it is found in the same location (CG4). The fact that the locus is essentially identical across diverse ICEs indicates that it provides a selective advantage or that it is of recent acquisition. Tn*6212* is further analysed and discussed in Chapter 4.

PsICEs share some cargo genes with plasmids. For example, Tn*6212*, CuR and HgR transposons present in the PsICEs can be found in plasmids. The movement of these transposons from plasmids to ICEs can facilitate their spread. Plasmid carriage is often costly (Björkman & Andersson, 2000), but studies suggest that there isn't any cost in carrying the PsICEs (Chapter 5; Colombi *et al.*, 2017). Moreover, even under a strong selection to lose it, Pph_{13A02}ICE is not completely lost from the population but persists in a small fraction until a favourable host is encountered upon which its frequency increases in the bacterial population (Neale *et al.*, 2016).

The phylogenetic history of PsICEs is difficult to infer due to the impact of recombination. PsICE evolution is characterized by extensive recombination events as described also for the SXT/R39 family of ICEs (Spagnoletti *et al.*, 2014, Johnson & Grossman, 2015). For example, both ClonalFrameML and Alfy recognized many recombination events and the chimeric nature of the PsICE backbones, respectively. This is evident with comparison of only backbone genes (e.g. the backbone of Psy_{B728a}ICE displays regions identical to Psa_{C15}ICE, Psa_{NZ45}ICE_Cu and Paf_{4394}ICE), but also with the comparison of the entire PsICEs

sequences (e.g. Pdp_{529}ICE and Ps_{34881}ICE alternate regions of complete identity with region of divergence). Inter-ICE recombination is frequent enough to mask evolutionary history, producing chimeras with variable patterns of similarity to each other. Although inter-ICE recombination events have been frequent, PsICEs have maintained their overall synteny. This again suggests that a different shuffled arrangement of the backbone genes is selected against, maybe because it would prevent these genes from working as a unit (Selfish Operon Model (Lawrence & Roth (1996)).

Patterns of inter-ICE recombination have been observed in other ICEs (Wozniak & Waldor, 2010; Spagnoletti *et al.*, 2014). The extent of recombination and sequence identity can seem surprising, considering the vast spatial and temporal distance separating the strains. However, the extent of recombination and the global distribution of the PsICEs could be explained by factors such as i) the dispersal potential of *P. syringae* via the water cycle (Monteil *et al.*, 2013); ii) the high abundance of *P. syringae* (10^8 cfu/cm^2) in wild plants of alpine ecosystems (Monteil *et al.*, 2012); and iii) high transmission rates of the PsICEs, via transformation and conjugation, to different pathovars, potentially to a maximum of 10^{-2} transconjugants per recipient cell *in planta* (Lovell *et al.*, 2009; Colombi *et al.*, 2017; Chapter 5). This pattern of distribution is compatible with the existence of a highly diverse and recombinogenic population that lives primarily in environmental habitats from which lineages emerge due to clonal expansion (Monteil *et al.*, 2013).

Although ICEs recombine frequently *in vitro* (Garris *et al.*, 2009; Colombi *et al.*, 2017; Chapter 5), some PsICEs remain highly conserved over several decades. Psa_{NZ13}ICE has accumulated 6 SNPs in 53 years in comparison to Pdp_{529}ICE (in

~100 kb) and 4 SNPs in comparison with Par$_{4457}$ICE (in ~90 kb, Tn*6212* excluded, see Chapter 4). The accumulation of ~10^{-6} substitution per site per year is in line for what described for other bacterial species (Duchêne *et al*., 2016). Given this mutation rate it is actually not that surprising that some PsICEs have conserved nucleotide identity across decades.

The comparison between PAPI-1 and PsICEs indicated that they share a common ancestral ICE. PAPI-1 and PsICEs share most of the genes of their backbones and also integrate at *att-1* and *att-2* sites. In addition, most of the shared backbone genes are syntenous. Although *P. aeruginosa* is an important human pathogen, it is also found in the environment and can be isolated from different sources such as soil, water and plants (Feltman *et al*., 2001; Kidd *et al*., 2012), they can thus share the same ecological niche. However, the PAPI-1 like ICEs are not exclusively present in *P. aeruginosa* but also in *P. fluorescens* Pf-5, *P. chlororaphis* PA23, *P. putida* WCS374 and *P. mosselii* SJ10 isolated from wastewater soil and plants. The conserved synteny in the PsICEs suggests a common ancestor for this family, as for the PAPI-1 family. The conserved synteny for most of the backbone genes between the PsICE and PAPI-1 families of ICEs and the formation of two separated clusters when these families are phylogenetically compared suggest that there is also a common ancestor between the two families of ICEs. Given a clear separation of these families, I hypothesize a rare transmission event (either direct or indirect) between *P. syringae* and *P. aeruginosa*. If transmission events were frequent, conserved synteny and clear separation would be unlikely to be observed.

The first ICEs described in *P. syringae* were Psy$_{B728a}$ICE and Pph$_{13A02}$ICE (Feil *et al.,* 2005; Pitman *et al*., 2005). Feil *et al.* (2005) merely note the presence of

Psy_{B728a}ICE, while Pph_{13A02}ICE has been the subject of extensive study (Pitman *et al*., 2005; Lovell *et al*., 2009; Godfrey *et al*., 2010; Lovell *et al*., 2011; Godfrey *et al*., 2011; Neale *et al*., 2016). Three additional ICEs were described during the *Psa* outbreak (Butler *et al*., 2013; McCann *et al.,* 2013). This Chapter demonstrates that these are not unrelated solo elements but they are part of a family of ICEs that spread into *P. syringae* species, whose evolutionary history has been clouded by recombination.

Chapter 4 – Horizontally transmitted elements and contributions to *Pseudomonas syringae* pv. *actinidiae* virulence

4.1 Introduction

The kiwifruit bleeding canker disease, caused by *P. syringae* pv. *actinidiae* (*Psa*) (Balestra *et al*., 2010), devastated the kiwifruit industry worldwide, leading to the almost complete eradication of the golden kiwifruit variety *Actinidiae chinensis* var. *chinensis* Hort16A (Greer & Saunders, 2012). Initially the disease was reported in Italy (Balestra *et al*., 2010), and in the following three years *Psa* was detected in neighboring European countries (Abelleira *et al*., 2011; EPPO, 2011; Vanneste *et al*., 2011; Mazzaglia *et al*., 2012; Bastas & Karakaya, 2017). In 2010, the outbreak reached two other major countries cultivating kiwifruit: Chile and New Zealand (EPPO, 2011; Everett *et al*., 2011).

Genomic analysis showed that although the pandemic was derived from a single clone, it acquired a set of distinctive Integrative Conjugative Elements (ICEs) during the course of its global journey (Mazzaglia *et al*., 2012; Butler *et al*., 2013; McCann *et al*., 2013). These ICEs belong to the PsICE family that is common to plant-associated *Pseudomonas* (refer to Chapter 3). Analysis of the PsICE family

revealed that the Tn*6212* transposon is a common (and often highly conserved) element that is integrated within the larger element, suggesting a possible role in the interaction between *P. syringae* and its various hosts.

Transposable elements are specific DNA segments capable of moving from one position in a given genome to another (Campbell *et al.*, 1979; Muñoz-López & García-Pérez, 2010). Transposable elements are widespread in prokaryotes, with variation in genetic organization, mode of insertion/excision and in the accessory genes they carry. With the exception of conjugative transposons (i.e. ICEs), transposons are not capable of self-transmission between hosts, but hitchhike with other genetic elements such as ICEs, plasmids or phages (Roberts *et al.*, 2008; Leclerq *et al.*, 2012). However, even transposons incapable of horizontal transfer play a role in bacterial evolution given their capacity to recruit and move genes cassettes that confer a selective advantage to their hosts (Liebert *et al.*, 1999; Holmes *et al.*, 2003; Ferreira *et al.*, 2015; Reid *et al.*, 2015; Furi *et al.*, 2016). Both the transposons and the 'shuttles' (such as plasmids and ICEs) benefit from this association. Transposons without 'shuttles' (i.e. incapable of horizontal transfer) have a restricted transmission potential while 'shuttles' without the recruitment of transposons conferring a selective advantage to the host bacterium have a lower chance to be maintained over time. The synergic association between transposons and their 'shuttles' was indeed crucial for the dissemination of antibiotic resistance among prokaryotes (Stoke & Gillings, 2007; Roberts & Mullany, 2009).

Tyrosine recombinases (such as the XerC in Tn*6212*) promote the site-specific integration/excision of mobile elements into the genome (Doungudomdacha *et al.*, 2007). Many types of transposable elements rely on tyrosine (or serine) recombinases for their movement, such as ICEs, genomic

islands and unit transposons (Murphy, 1983; Scott *et al*., 1994; Buchrieser *et al*., 1998; Hochhut & Waldor, 1999). Unit transposons encode an enzyme involved in excision and integration, often a site-specific recombinase or resolvase, and one or several accessory genes in one genetic unit (Roberts *et al*., 2008).

Tn*6212* is a ~16 kb tyrosine recombinase unit transposon (Butler *et al*., 2013) encoding several accessory open reading frames. Predicted genes on Tn*6212* include a methyl-accepting chemotaxis protein (MCP), a transporter, a chloride ion channel, an enolase, an inorganic pyrophosphatase, a C4-dicarboxylic acid transporter (DctT), a putative alcohol catabolism-associated protein and two transcriptional regulators. The MCP is predicted to be involved in taxis toward malate (McCann *et al*., 2013). The transporter is related to human bile acid:sodium symporters but also shows homology to the arsenite efflux pump ACR3 (Mansour *et al*., 2007). The voltage-gated Cl^- channel belongs to a family of proteins that catalyzes the selective flow of Cl^- ions across cell membranes (Iyer *et al*., 2002). The inorganic phosphatase catalyzes the highly exergonic reaction that hydrolyzes inorganic pyrophosphate (PPi) to two molecules of orthophosphates (Pi) (Harold, 1966). Enolase is an enzyme involved in carbon metabolism, but also in RNA processing and regulation (Taghbalout & Rothfield, 2007), in bacterial response to oxidative stress (Weng *et al*., 2016); and enolase-like proteins can bind human plasminogens in the outer membrane of *P. aeruginosa* (Ceremuga *et al*., 2014). DctT is a putative C4-dicarboxylic acid transporter: a particularity of this protein is the sequence at the N-terminus that is predicted to be a signal for export via the Type 3 Secretion System (T3SS) (McCann *et al*., 2013). The T3SS is a molecular syringe device through which proteins with roles in pathogenicity and manipulation of host defenses (the so-named T3SS effector proteins (T3SEs)) are

delivered into the host cell (Coburn *et al*., 2007).

McCann *et al*. (2013) named Tn*6212* "the enolase region" and predicted that it was involved in the virulence of *Psa*, specifically with a role in manipulation of host cell metabolism. *Pseudomonas* has a preference for growth on TCA cycle intermediates and the dicarboxylate transporter *dctA1* was demonstrated to play a role in the virulence of *P. syringae* pv. *tomato* (*Pto*) DC3000 (Mellgren *et al*., 2009). The putative T3SS-targeted DctT, is hypothesized to enter the plant cell and to be incorporated into host cell membranes to facilitate export of C4 sugar acids in a manner analogous to that suggested for α-ketoglutarate in *Xanthomonas oryzae* pv. *oryzae* (Guo *et al*., 2012). The enolase , being part of the RNA degradosome, was predicted to be involved in the deprivation of plant sugars. It was hypothesized to be exported and to enhance activity of glycolysis in plant cells and thus to increase the concentration of C4 acid sugars to be exported by DctT (McCann *et al*., 2013).

In this Chapter, after a survey of the diversity of Tn*6212* elements, I report the results of investigations into the possible function of genes carried on the transposon and their role in the plant-*Psa* interaction. However, at the start of this project *Psa* was an unknown quantity in terms of genetics. It was therefore necessary to firstly establish *Psa* genetics and protocols for assaying phenotypes on plants before tackling questions relating to the function of specific genes on Tn*6212*.

4.1.1 Aims

- Describe the diversity and organization of Tn*6212* elements harboured by the PsICE family of ICEs.
- Establish protocols for *Psa* NZ13:
    - explore the possibility to insert foreign DNA,
    - create a T3SS mutant,
    - set up protocols for *in planta* assays using different hosts.
- With genetics established it is possible to test the hypothesis of the role of the enolase region:
    - test the contribution of Tn*6212* toward the colonization of host plants;
    - test the secretion of DctT into the host plant cell;
    - test the contribution of Tn*6212* in *Psa* fitness *in vitro*.

4.2 Results

4.2.1 Variability of Tn*6212*

Tn*6212* is the most common element found in the PsICEs and it is always integrated in CG4 (a conserved hotspot locus for insertions in the PsICEs family (refer to Chapter 3)).

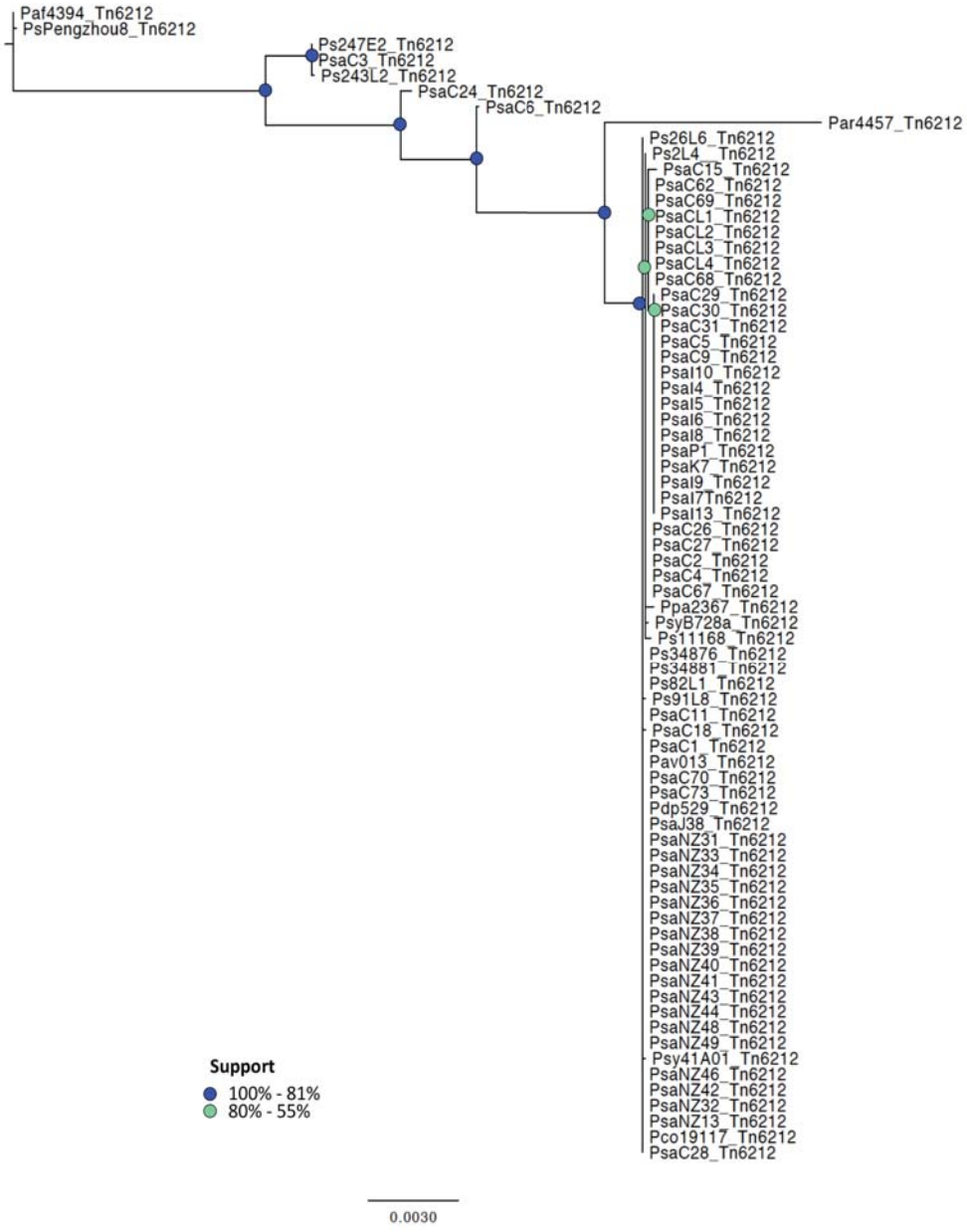

Figure 4.1. Phylogeny of Tn*6212*. PhyML maximum likelihood tree based on the REALPHY (Bertels *et al.*, 2014) alignment (~11 kb) of Tn*6212* (~16 kb) present in the PsICE family.

The phylogeny of Tn*6212* harboured by the PsICE family is represented in Figure 4.1. Most of the transposons form a single cluster, with very little difference within the cluster. Tn*6212* elements inside the main cluster differ only by 8 SNPs in ~16 kb. The main cluster groups together almost identical Tn*6212* elements that are however harboured by very divergent PsICEs (refer to Chapter 3). At the same time, two almost identical PsICEs (Par_{4457}ICE and Psa_{NZ13}ICE) that share only 6 SNPs in their ~60 kb backbone (refer to Chapter 3), harbour Tn*6212* elements that differ by 129 SNPs.

Due to the chimeric nature of the PsICEs, it is impossible to build a meaningful phylogenetic tree of the PsICEs and thus compare the evolutionary history of the Tn*6212* to that of the PsICEs. However, there is an indication that some Tn*6212* elements have co-evolved for a short time with the PsICEs they are integrated in. For instance, in the main cluster in Figure 4.1, the Tn*6212* elements harboured by the same type of PsICE (the PsaI10-like and the PsaC2-like PsICE) form separate clades, although the support is low.

Tn*6212* is conserved among the PsICEs, although there are some modifications compared with the most common variant harboured also by *Psa* NZ13 (Figure 4.2). In Psa_{C3}ICE a 2 kb region encompassing the MerR regulator and the transporter is absent. In Paf_{4394}ICE and $Ps_{Penghzhou8}$ICE the first 4 kb have been substituted with a Na^+/H^+ antiporter (*nhaP*) and a sensor histidine kinase then Tn*6212* continues with the gene upstream of the chloride channel. In Psa_{C6}ICE and Psa_{C24}ICE, Tn*6212* starts with one hypothetical gene upstream of the chloride channel; Tn*6212* in Psa_{C6}ICE harbours also an additional transposase and differs to Tn*6212* in Psa_{C24}ICE by 42 SNPs, 27 of which are in the predicted enolase gene. Most of these substitutions are synonymous (19 over 27), indicative of purifying

selection for the preservation of protein function. Psa$_{C1}$ICE has an IS*3* family transposase insertion upstream of the predicted enolase; Psa$_{C29}$ICE (along with Psa$_{C30}$ICE and Psa$_{C31}$ICE, identical PsICEs) has two IS*3* family transposases inserted downstream of *merR*; while Ps$_{34876}$ICE (and Ps$_{34881}$ICE, identical PsICEs) acquired a mercury resistant (HgR) transposon upstream of the transporter.

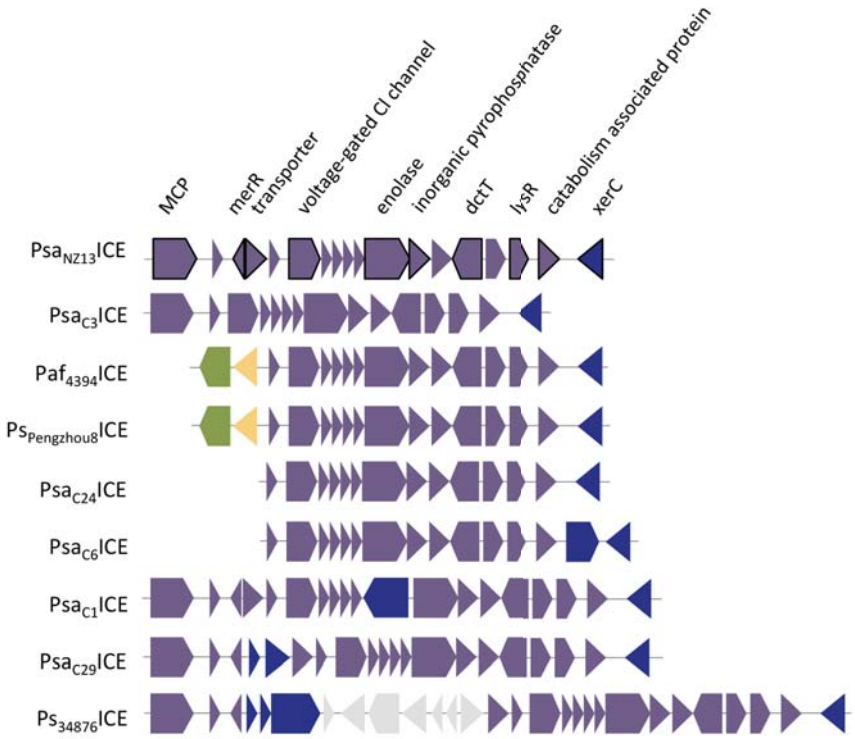

Figure 4.2. Genetic organization of Tn*6212* in the PsICEs. Purple boxes define the "enolase region", blue boxes are mobile genes (transposases or recombinases), light grey boxes denote genes encoding mercury resistance, green boxes are *nhaP,* yellow boxes are sensor histidine kinases.

Tn*6212* is not always associated with ICEs: it is also present in plasmids and chromosomes with no association with ICEs (Table 4.1). The genomes harbouring Tn*6212* are all plant-associated *Pseudomonas*.

Strain ID	Host of isolation	Tn*6212* location	GenBank Accession
P. syringae pv. *syringae* 642	unidentified weed	genome	ADGB01000033.1
P. syringae pv. *aptata* ICMP459	beet	plasmid	LJRP01000009.1
P. syringae pv. *philadelphi* ICMP8903	English dogwood	plasmid	LJQY01000214.1
P. syringae pv. *solidagae* ICMP16925	Canada goldenrod	plasmid	LJRH01000181 .1
P. graminis DSM11363	N/A	genome	FOHW01000016.1

P. syringae pv. *tagetis* ICMP4091	African marigold	plasmid	LJRM01000076.1
P. amygdali pv. *morsprunorum* HRI-W 5269	tart cherry	plasmid	LIHZ01000021.1
P. syringae pv. *avii* strain CFBP 3846	wild cherry	plasmid	LIIJ01000245.1
P. syringae pv. *primulae* ICMP3956	primula	genome	LJRC01000145.1
P. trivialis LMG 21464	grass	genome	MDFJ01000003.1
Pseudomonas sp. Leaf129	*A. thaliana*	genome	LMOC01000006.1
P. syringae C77	kiwifruit	plasmid	Rainey Lab

Table 4.1. List of strains harboring Tn*6212* not integrated in a PsICE.

4.2.2 Optimization of mutagenesis techniques

Having described the structure and diversity of Tn*6212*, I now turn to questions relating to its function and contribution to plant-microbe interactions. To this end it was first necessary to establish genetics for *Psa*. I began by asking whether it was possible for foreign DNA to be inserted into *Psa*. For this I chose to focus on the capacity of *Psa* to incorporate a transposon into its genome. Incorporation would indicate capacity for DNA to be taken up by *Psa* and for it to be recombined into the genome without prior degradation. To this end transposon IS-Ωkm/hah (Tn*5*) (Giddens *et al.*, 2007) was chosen because it has been widely used to mutagenize *Pseudomonas*.

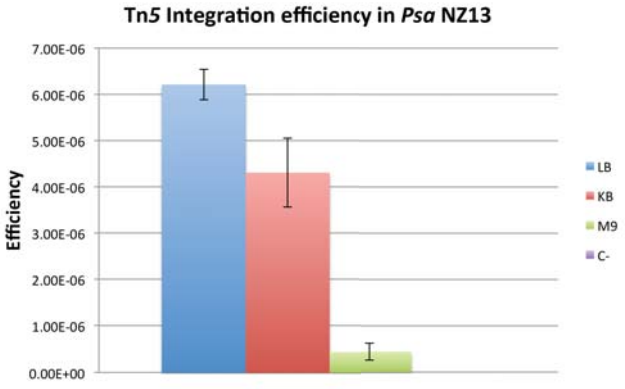

Figure 4.3. Tn*5* transformation efficiency of *Psa* NZ13. Efficiency is calculated as number of *Psa* NZ13^{kmR}/*Psa* NZ13. The means of three independent experiments of three replicas each are plotted with the standard error. The negative control is plotted as C-.

E. coli DH5 α harbouring Tn*5* was used as the donor, *Psa* as recipient and *E. coli* pRK2013 as helper. The triparental mating was carried out on KB, LB and M9 agar plates to identify the most efficient conditions (Figure 4.3). For this evaluation *Psa* NZ13 was used as representative strain of the outbreak clade, while *Psa* J35 and *Psa* K28 were used as representatives of the *Psa*-1 and *Psa*-2 lineages. Transconjugants were obtained for *Psa* NZ13, but for *Psa* J35 and *Psa* K28 no mutants could be detected. For *Psa* NZ13 the efficiency increased when the triparental mating was performed on LB rather than KB. When the triparental mating was carried out on M9 plates the there was a one-log reduction in efficiency.

Having established that DNA can be moved into *Psa* by conjugation, I next turned to transformation via electroporation. This approach was tested with plasmids pME6010 and pK18mobsacBΔ*hrcC.* pME6010 is able to replicate autonomously and it does not need to integrate into the host genome to confer antibiotic resistance (Heeb *et al.*, 2000). pK18mobsacB instead cannot replicate autonomously and has to integrate into the genome via homologous recombination to confer resistance to the host (Schäfer *et al.*, 1994). Homologous recombination is driven by one of the flanking sequences of *hrcC* (refer to next Section and Section 2.2.2).

Electroporation was successful only for the replicative vector, thus conjugation was chosen for the creation of mutants.

4.2.3 Creation of the mutant *Psa* NZ13 Δ*hrcC*

Having shown that performing genetic manipulation in *Psa* NZ13 is possible, I aimed to construct a deletion mutant of *hrcC*. HrcC is a structural protein of the T3SS (Deng & Huang, 1999). T3SS mutants are perfect candidates to study effectiveness of the mutation process and to test possible host plants as they show a clear phenotype. Typically, T3SS mutants are unable to invade the host and cause disease because they are unable to deliver effectors (responsible for the suppression of plants defences) into the host cell (Deng & Huang, 1999; Hauck *et al*., 2003). Moreover, a T3SS mutant could be further used to test the predicted exportation of DctT via T3SS (see below).

The in-frame deletion of *hrcC* was carried out using pK18mobcsacB. pK18mobsacBΔ*hrcC* was inserted in *Psa* NZ13 via triparental mating (Section 2.2.9). The first homologous recombination is driven by the presence of the flanking sequences of *hrcC*. To induce the second recombination event, PsaNZ13::pk18mobsacBΔ*hrcC* colonies were spread onto a Petri plate of KB containing 5% sucrose. pK18mobsacB vector harbors the negative-selection gene *sacB*, which is activated by the presence of sucrose. Since the expression of *sacB* is toxic to the bacterium, strong selection is imposed for the loss of this gene (in practical terms, loss of the vector) for survival. During this process the second homologous recombination event occurs, leaving in the bacterial chromosome either the wildtype or the mutated copy of the target region. To check for the absence of the wildtype copy and to localize the mutation, a PCR was carried out using primers that anneal in the chromosome upstream and downstream of the target region (*hrcC*) (Figure 4.4) (Section 2.2.11). The in-frame deletion was confirmed via Sanger sequencing.

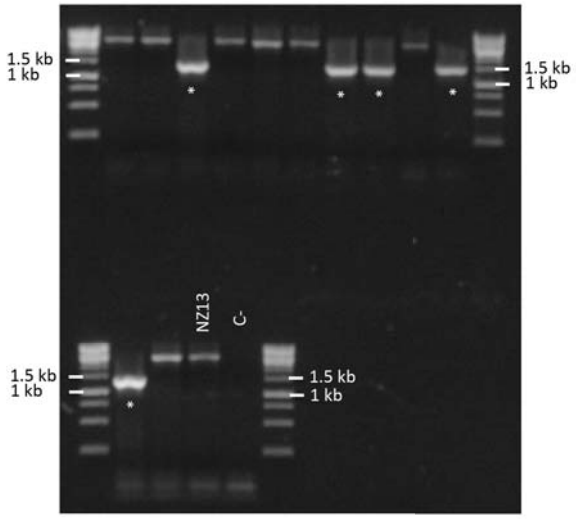

Figure 4.4. Analysis of colonies after the second recombination event. PCR carried out to assess the presence of the Δ*hrcC* copy in 12 *Psa* NZ13 recombinant colonies. PCR was carried out with the primers Ext_hrcC_For and Ext_hrcC_Rev. Control of *Psa* NZ13 (NZ13) shows one bands of ~3.4 kb, positive colonies (*) with the *hrcC* deletion showed one band of ~1.4 kb, the negative control (C-) shows no amplification. The ladder used is O'Gene Ruler 1kb DNA Ladder (Fisher Scientific).

4.2.4 Phenotypic characterisation of *Psa* NZ13 Δ*hrcC*

4.2.4a Hypersensitive Response

The first phenotypic assay carried out to test the effect of the *hrcC* deletion was the Hypersensitive Response (HR). The HR (rapid and localized cell death at the site of bacterial inoculation) is a plant response to the recognition of T3SEs (Greenberg & Yao, 2004; Coll *et al*., 2011). T3SS mutants are unable to deliver the T3SE into the plant cell and thus do not produce HR (Alfano *et al*., 1996; Deng *et al*., 1998). The absence of HR will thus confirm the successful deletion of *hrcC*.

HR was assayed by infiltrating *Psa* NZ13 and *Psa* NZ13 Δ*hrcC* (and *Pfm* NZ5 as positive control) into the mesophyll of *Nicotiana benthamiana* leaves (refer to Section 2.2.16). Results were observed 24 hours post inoculation (hpi) (Figure 4.5). *Psa* NZ13 and *Pfm* NZ5 produced a clear HR, while the *hrcC* deletion mutant did not, confirming the malfunction of the T3SS.

Figure 4.5. HR response in *Nicotiana benthamiana*. The HR is characterized by cell death visible in the dark and translucent infiltrated areas. Picture was taken 24 hpi. The experiment was repeated three times.

4.2.4b Growth and symptoms on kiwifruit

Having shown the destruction of the functionality of the T3SS, I aimed to set up protocols for kiwifruit colonization assays. *Psa* colonization assays were performed on the susceptible kiwifruit cultivar *A. chinensis* var. *chinensis* Hort16A. 3-4 week old Hort16A plantlets were dip-inoculated in a solution containing *Psa* and the surfactant Silwet (refer to Section 2.2.18). The addition of Silwet was necessary to establish a detectable population for day 0 inside the leaves; previous experiments carried out without it produced no detectable endophytic *Psa* immediately after inoculation.

Growth was recorded at 0, 3 and 7 days post inoculation (dpi), symptoms at 7 dpi. While the wildtype was able to grow *in planta*, reaching ~10^9-10^{10} cfu/cm^2 in 7 days of infection, *Psa* NZ13 Δ*hrcC* did not grow: from an initial population of 10^5 cfu/cm^2 at day 0, it decreases to 10^3 and 10^4 cfu/cm^2 at day 3 and 7 respectively (Figure 4.6).

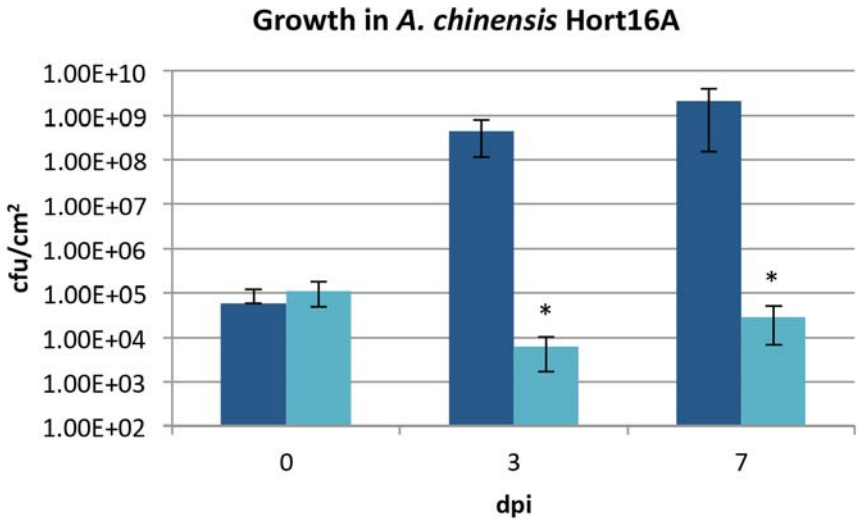

Figure 4.6. Growth on *Actinidiae chinensis* Hort16A leaves. Growth of *Psa* NZ13 (blue bars) and *Psa* NZ13 Δ*hrcC* (azure bars) was assessed endophytically on leaves of the kiwifruit cultivar Hort16A. Data are means and standard deviation of five replicates. Two tailed *t*-test showed statistical difference ($P<0.05$) (*) in growth between of Psa NZ13 and *Psa* NZ13 Δ*hrcC*.

Accordingly, *Psa* NZ13 Δ*hrcC* does not produce any symptoms while *Psa* NZ13 causes spots on leaves (Figure 4.7).

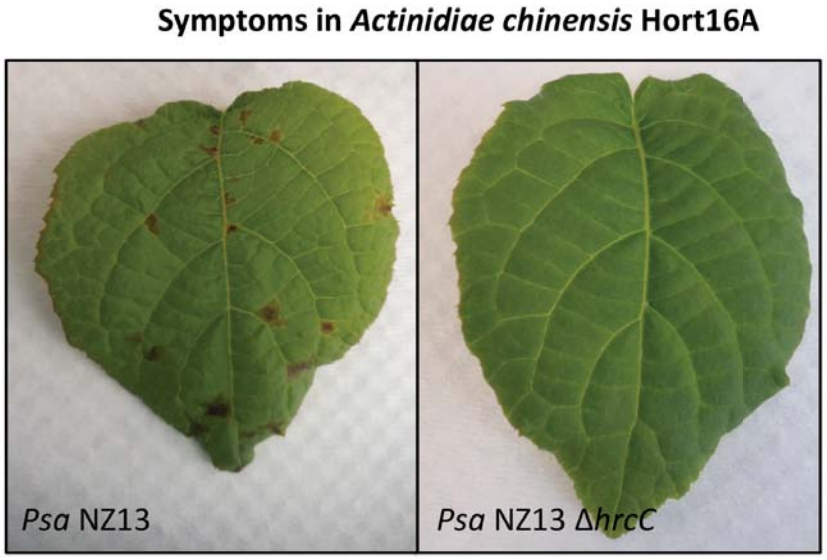

Figure 4.7. Symptoms on Hort16A leaves. Symptoms produced by *Psa* NZ13 and *Psa* NZ13 Δ*hrcC* during the growth assay showed in Figure 4.3. Pictures were taken at 7dpi.

4.2.4c Growth and symptoms on model plants

Kiwifruit is the ideal plant to study the virulence of *Psa* because it is its natural host. However, I also explored the possibility of using alternative host

plants that are more easily grown, cheaper and not prone to contamination from the supplier's facility as kiwifruit. The model plants used were: *N. benthamiana*, *Arabidopsis thaliana* Col-0 and *Solanum lycopersicum* Money Maker (refer to Section 2.2.17).

Psa NZ13 is able to rapidly grow in *N. benthamiana* and from an initial population of 10^3 cfu/cm^2 it reaches a population density of 10^6-10^7 cfu/cm^2 already at 2 dpi. The population then stabilizes at day 3. Also, the *hrcC* deletion mutant is able to grow but it does not reach the same population size as the wildtype with a maximum of 10^5-10^6 cfu/cm^2 at day 3 (Figure 4.8).

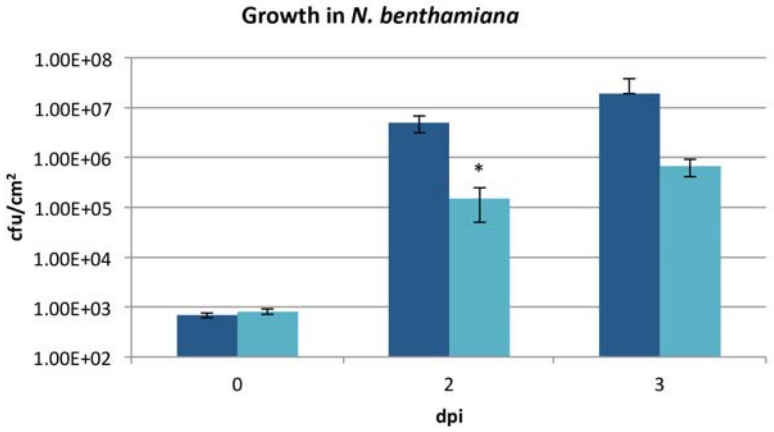

Figure 4.8. Growth on *Nicotiana benthamiana*. Growth of *Psa* NZ13 (blue bars) and *Psa* NZ13 Δ*hrcC* (azure bars) was assessed endophytically on leaves *N. benthamiana*. Data are means and standard deviation of three replicates. Two tailed *t*-test showed statistical difference ($P<0.05$) (*) in growth between of *Psa* NZ13 and *Psa* NZ13 Δ*hrcC*.

Symptoms are shown in Figure 4.9. The wildtype does not produce the classical leaf spots caused on kiwifruit but only a very faint discoloration of the infiltrated area, while the HrcC mutant does not induce any visual change in the leaf.

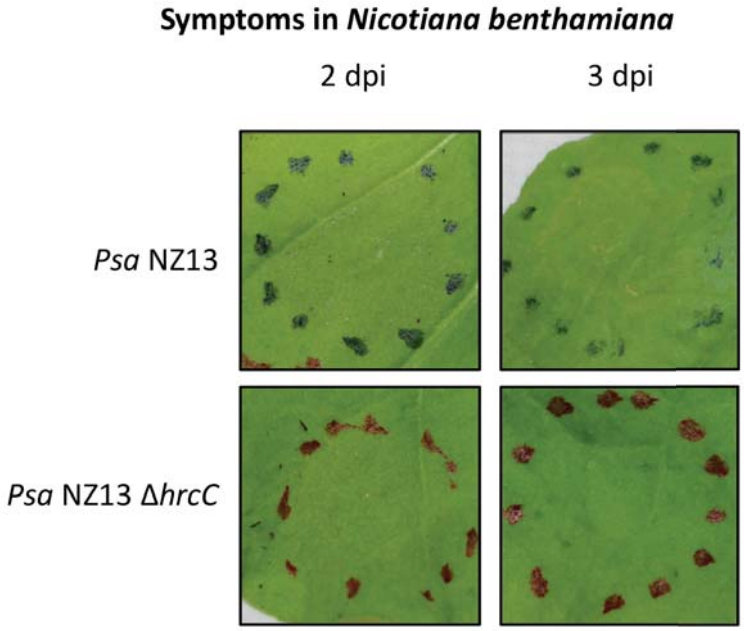

Figure 4.9. Symptoms development on *Nicotiana benthamiana* leaves. Leaves were infiltrated with a needleless syringe and pictures taken before the assessment of bacterial population.

Growth and symptoms on *S. lycopersicum* Money Maker leaves are shown in Figures 4.10 and 4.11.

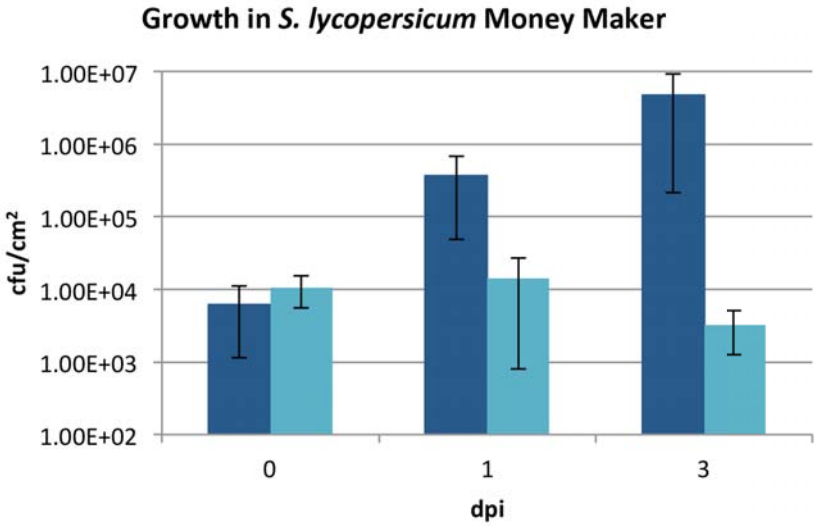

Figure 4.10. Growth on *Solanum lycopersicum* Money Maker leaves. Growth of *Psa* NZ13 (blue bars) and *Psa* NZ13 Δ*hrcC* (azure bars) was assessed endophytically on leaves *S. lycopersicum*. Data are means and standard deviation of three replicates. Two tailed *t*-test showed no statistical difference ($P>0.05$) in growth between of *Psa* NZ13 and *Psa* NZ13 Δ*hrcC*.

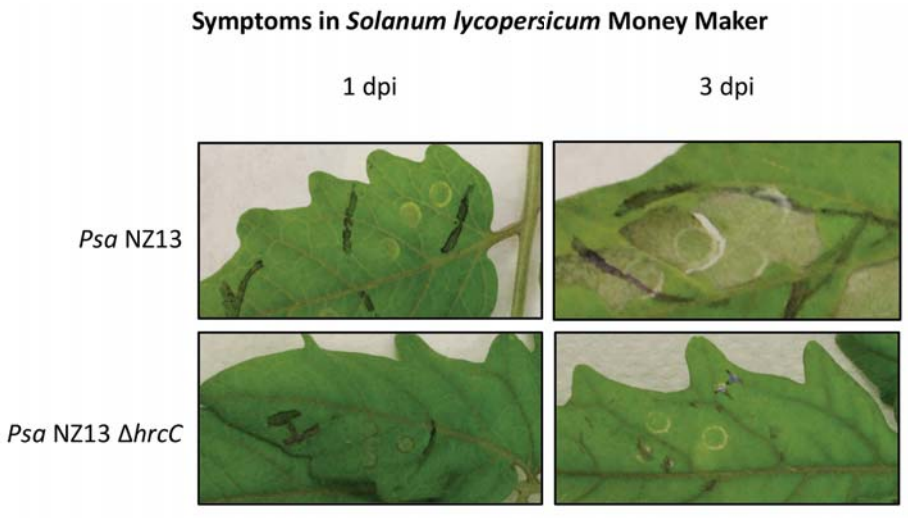

Figure 4.11. Symptoms development on *Solanum lycopersicum* leaves. Leaves were infiltrated with a needleless syringe and picture taken before the assessment of bacterial population showed in Figure 4.10.

Psa NZ13 is able to colonize Money Maker tomato plants. From an initial population size of 10^4 cfu/cm^2 at day 0, it reaches 10^5 – 10^6 cfu/cm^2 and 10^6 – 10^7 cfu/cm^2 at 1 and 3 dpi, respectively. The HrcC mutant remains stable at its inoculation density at 1 dpi and slightly decreases at 3 dpi. Also in tomato the symptoms produced by *Psa* are not the typical leaf spots but a discoloration of the infiltrated area. However, in addition to the discoloration, the infiltrated area quickly dries. The HcrC mutant did not produce any visual change in the infiltrated area.

Growth on *A. thaliana* is shown in Figure 4.12, *Pto* DC3000 was used as a positive control. Both the wildtype and the HrcC mutant were not able to grow, remaining stable at their inoculation density of ~10^2 cfu/cm^2. The positive control *Pto* DC3000 reached instead 10^6 cfu/cm^2 at 2 dpi. Both the strains did not produce any visual change in the leaves (data not shown).

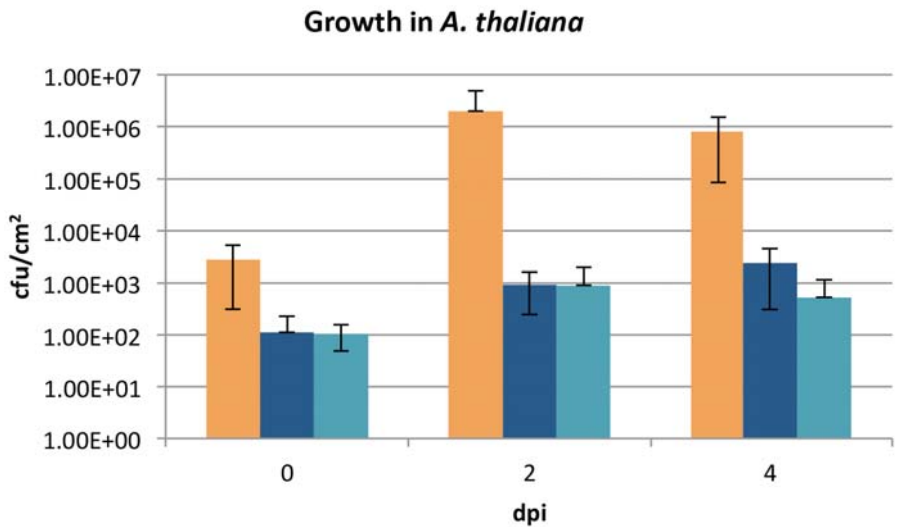

Figure 4.12. Growth on *Arabidopsis thaliana* leaves. Growth of *Psa* NZ13 (blue bars), *Psa* NZ13 Δ*hrcC* (azure bars) and *Pto* DC3000 (orange bars) was assessed endophytically in leaves *A. thaliana*. Data are means and standard deviation of four replicates. Two tailed *t*-test showed no statistical ($P>0.05$) difference in growth between of *Psa* NZ13 and *Psa* NZ13 Δ*hrcC*.

4.2.5 *In planta* phenotypical characterization of Tn*6212*

Having set up protocols for *Psa* NZ13 mutagenesis and *in planta* assays, I finally produced Tn*6212* mutants for their experimental analysis. The deletion mutants created were: *Psa* NZ13 Δ*enoR* ('enolase region' in-frame deletion mutant), *Psa* NZ13 Δ*enoG* (enolase in-frame deletion mutant) and *Psa* NZ13 Δ*dctT* (*dctT* in-frame deletion mutant).

To test the hypothesized secretion of DctT via T3SS, the pMT-1 and pMT-2 plasmids (refer to Section 2.2.21) were inserted in *Psa* NZ13 and *Psa* NZ13 Δ*hrcC*.

4.2.5a Colonization assays

Because Tn*6212* was predicted to be involved in the plant-pathogen interaction, pathogenicity tests were conducted on kiwifruit and also on *N. benthamiana* and *S. lycopersicum* Money Maker (refer to Sections 2.2.17 and

2.2.18). The latter were included because *Psa* showed the ability to colonize these plants even if it does not produce clear symptoms.

Hort16A plantlets were dip-inoculated as described in Material and Methods and growth recorded at 0, 3 and 7 dpi (Figure 4.13). The population growth shows that the wildtype and the Tn*6212* mutants had a similar behaviour, *Psa* NZ13 Δ*hrcC* was used as negative control. The Tn*6212* mutants did not show any impaired growth on kiwifruit and they were able to reach ~10^9 cfu/cm^2 as the wildtype. Consequently, the production of symptoms of the mutants was comparable to *Psa* NZ13 (Figure 4.14).

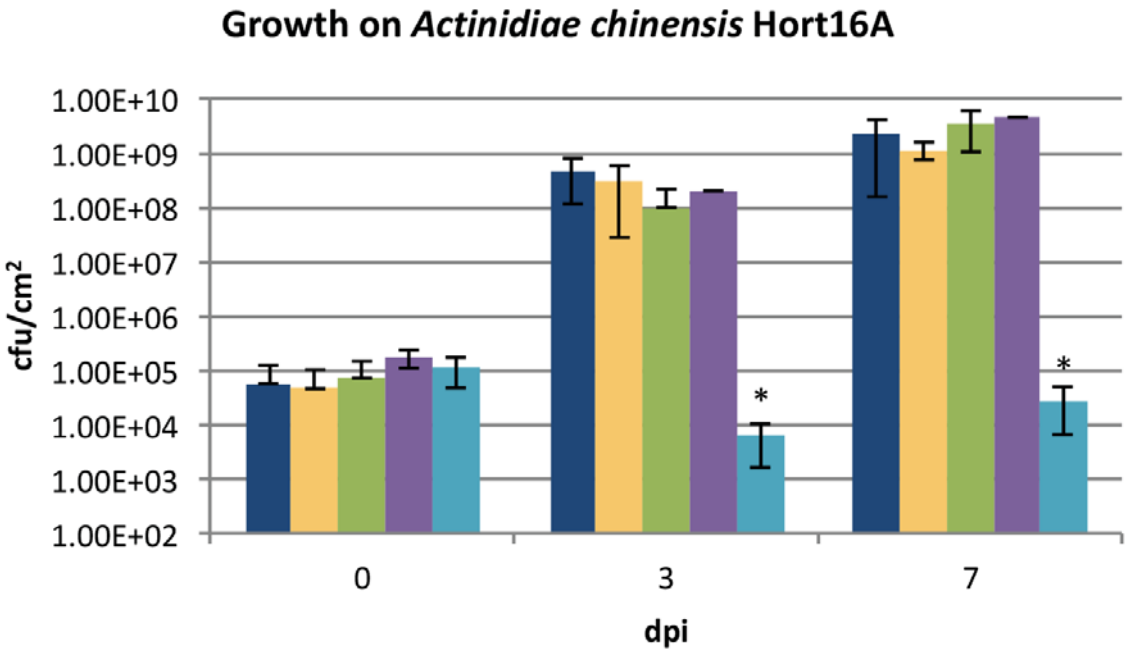

Figure 4.13. Growth on kiwifruit. Growth of *Psa* NZ13 (blue bars), *Psa* NZ13 Δ*enoR* (yellow bars), *Psa* NZ13 Δ*enoG* (green bars), *Psa* NZ13 Δ*dctT* (purple bars) and *Psa* NZ13 Δ*hrcC* (azure bars) was assessed endophytically on leaves of the kiwifruit cultivar Hort16A. Data are means and standard deviation of five replicates. Two tailed *t*- test revealed no statistical difference ($P>0.05$) between *Psa* NZ13 and the Tn*6212* mutants. (*) indicates $P<0.05$.

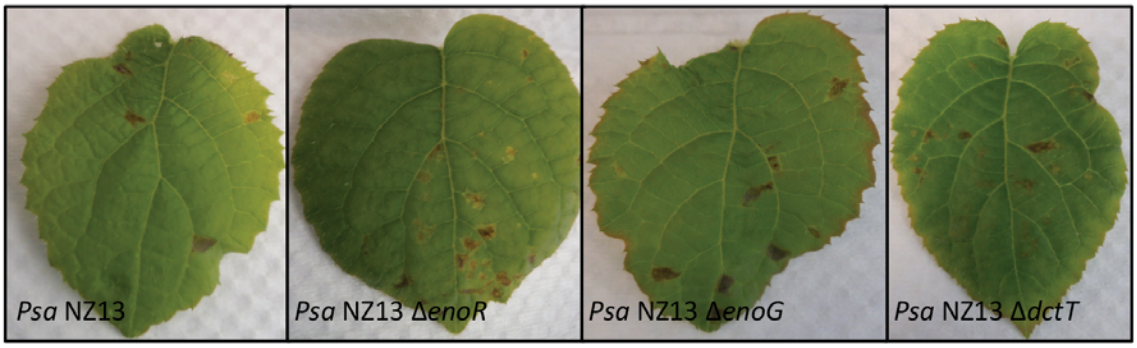

Figure 4.14. Symptom production of *Pseudomonas syringae* pv. actinidiae NZ13 and the Tn*6212* mutants on kiwifruit cultivar Hort16A. Kiwifruit plants were dip-inoculated and pictures were taken at 7 dpi.

Growth and symptom production on *N. benthamiana* are shown in Figure 4.15 and 4.16. Similarly to Hort16A, the Tn*6212* mutants were able to grow and to produce the same decolourization of the inoculated leaf area as the wildtype.

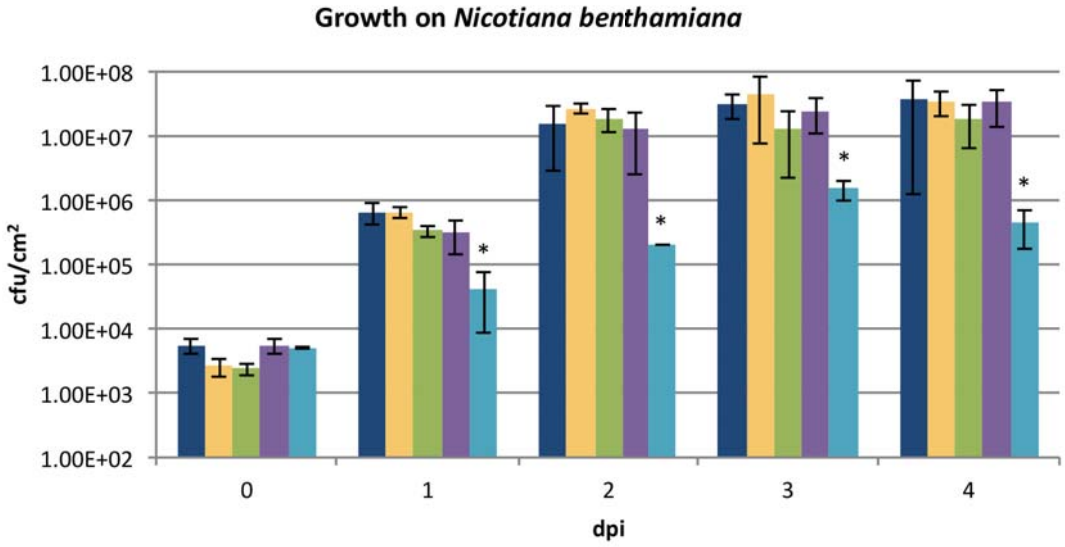

Figure 4.15. Growth on *Nicotiana benthamiana*. Growth of *Psa* NZ13 (blue bars), *Psa* NZ13 Δ*enoR* (yellow bars), *Psa* NZ13 Δ*enoG* (green bars), *Psa* NZ13 Δ*dctT* (purple bars) and *Psa* NZ13 Δ*hrcC* (azure bars) was assessed on pressure-inoculated leaves of *N. benthamiana*. Data are means and standard deviation of three replicates. Two tailed *t*- test revealed no statistical difference ($P>0.05$) between *Psa* NZ13 and the Tn*6212* mutants. (*) indicates $P<0.05$.

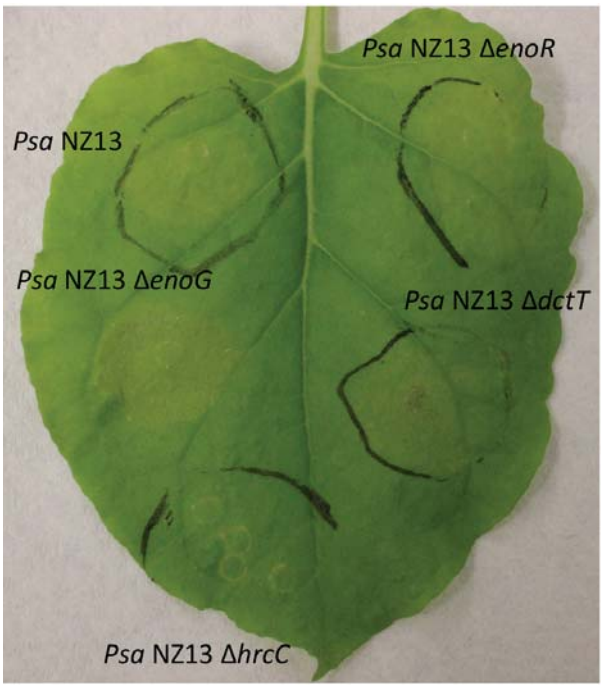

Figure 4.16. Symptoms production on *Nicotiana benthamiana*. Picture was taken at 4 dpi.

No differences between the mutants and the wildtype for colonization and symptom development was assesed on *S. lycopersicum* Money Maker (Figure 4.17 and 4.18).

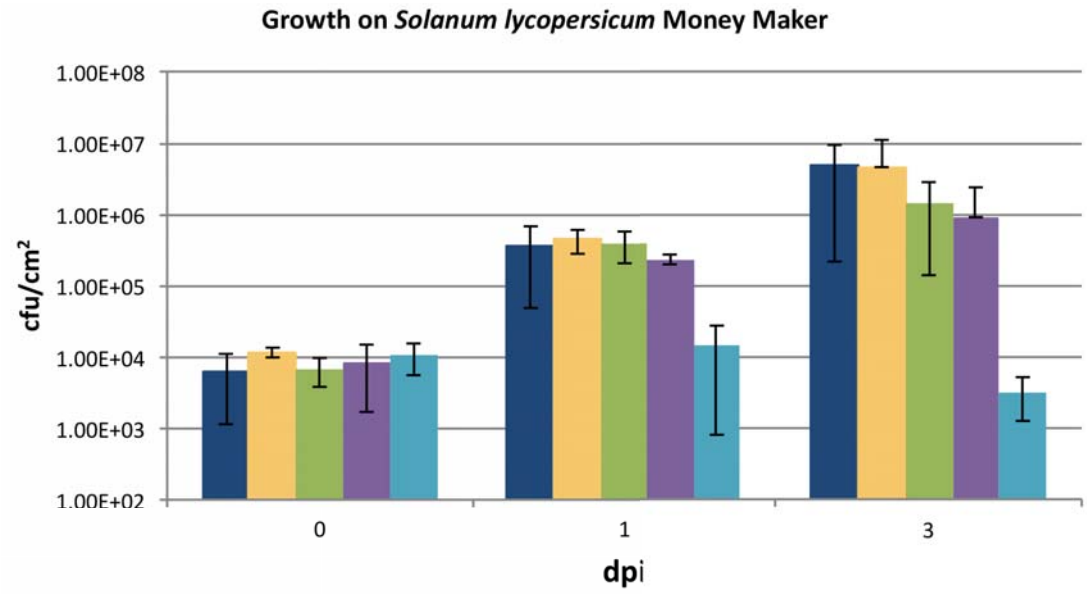

Figure 4.17. Growth on tomato. Growth of *Psa* NZ13 (blue bars), *Psa* NZ13 Δ*enoR* (yellow bars), *Psa* NZ13 Δ*enoG* (green bars), *Psa* NZ13 Δ*dctT* (purple bars) and *Psa* NZ13 Δ*hrcC* (azure bars) was assessed on pressure-inoculated leaves of *S. lycopersicum* variety Money Maker. Data are

means and standard deviation of three replicates. Two-tailed *t*- test revealed no statistical difference ($P > 0.05$) between *Psa* NZ13 and the mutants

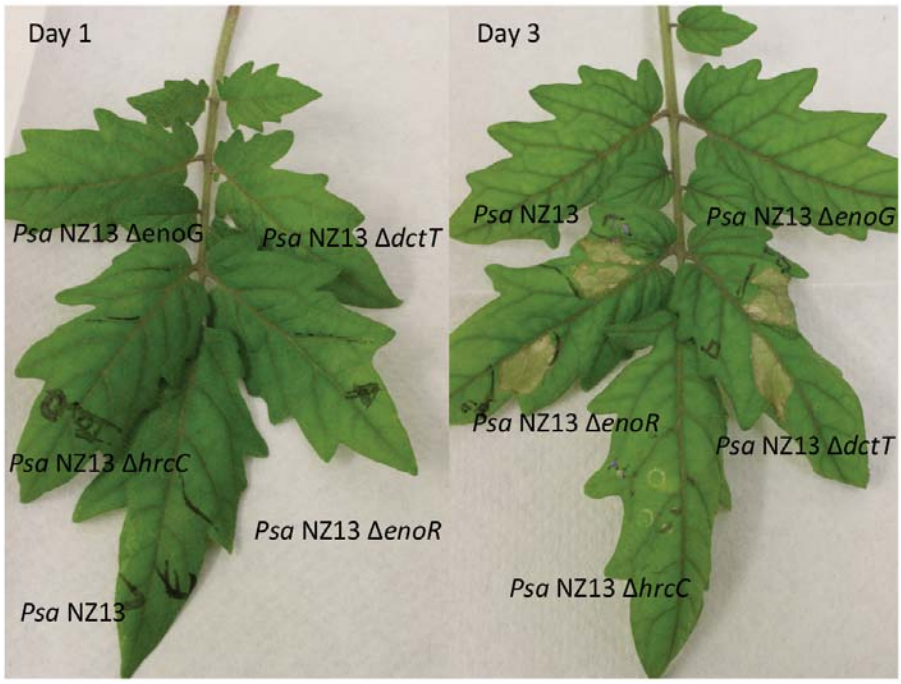

Figure 4.18. Symptoms development on *Solanum lycopersicum* Money Maker leaves.

4.2.5b Secretion assay

DctT was hypothesized to be exported into the host cell via T3SS (McCann *et al*., 2013), secretion assays with AvrRpt2 were carried out to analyse this hypothesis (refer to Section 2.2.21).

AvrRpt2 is a T3SE, when it is delivered into host cells, autocleavage of the N-terminal sequence by host cyclophilin occurs, followed by AvrRpt2-induced cleavage of the plant protein RIN4. Cleavage of RIN4 in turn activates RPS2 that mediates disease resistance, producing the HR (Grant *et al*., 2006). To assay whether DctT is indeed exported, the N-terminal T3SS-targeting sequence of *avrRpt2* (1-79aa) was replaced with the corresponding region in *dctT*: successful secretion results in a visible HR response. Two plasmids were synthesized by GenScript® for this purpose: in pMT-2 the promoter and the full-length *dctT* of *Psa* NZ13 were fused to *avrRpt2*$_{\Delta 1\text{-}79aa}$; in pMT-1 the promoter and the putative

exportation site (1-52 aa) of *dctT* of *Psa* NZ13 were fused to *avrRpt2*$_{\Delta 1\text{-}79aa}$. To confirm T3SS-mediated translocation is responsible, plasmids were also inserted in *Psa* NZ13 Δ*hrcC*. *Psa* NZ13 harbouring the full-length AvrRpt2 controlled by its native promoter (provided by Kee Sohn Lab) represents the positive control.

Leaves of *A. thaliana* were pressure-inoculated using a needleless syringe and HR was checked at 24 hpi and 48 hpi (Figure 4.19). *Psa* NZ13 + AvrRpt2 produced an HR on the leaves, while PsaNZ13Δ*hrcC* carrying the same plasmid did not, confirming the exportation via T3SS of the effector. The negative controls *Psa* NZ13 and *Psa* NZ13 Δ*hrcC* did not induce a HR. *Psa* NZ13 and *Psa* NZ13 Δ*hrcC* carrying either pMT-1 and pMT-2 plasmids did not induce an HR, indicating *dctT* is likely not secreted into the host cell.

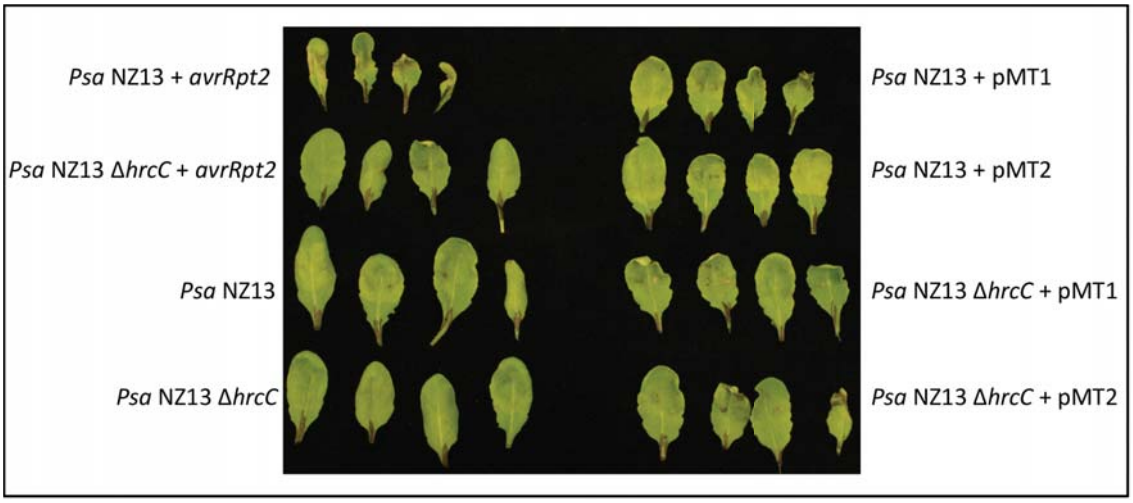

Figure 4.19. Secretion assay. Leaves of *A. thaliana* Col-0 were pressure infiltrate with a needless syringe and HR development recorded at 48 hpi. The experiment was repeated two times inoculating two leaves of three plants per strain used.

4.2.5c Ion leakage

Ion leakage from plant tissue, measured in conductivity, is an early indicator of cell death (Tamagnone *et al.*, 1998) and was thus used here to confirm the absence of HR (refer to Section 2.2.22). The experiment showed no difference

in conductivity between strains with or without pMT-1 and pMT-2 plasmids, confirming the absence of HR (Figure 4.20).

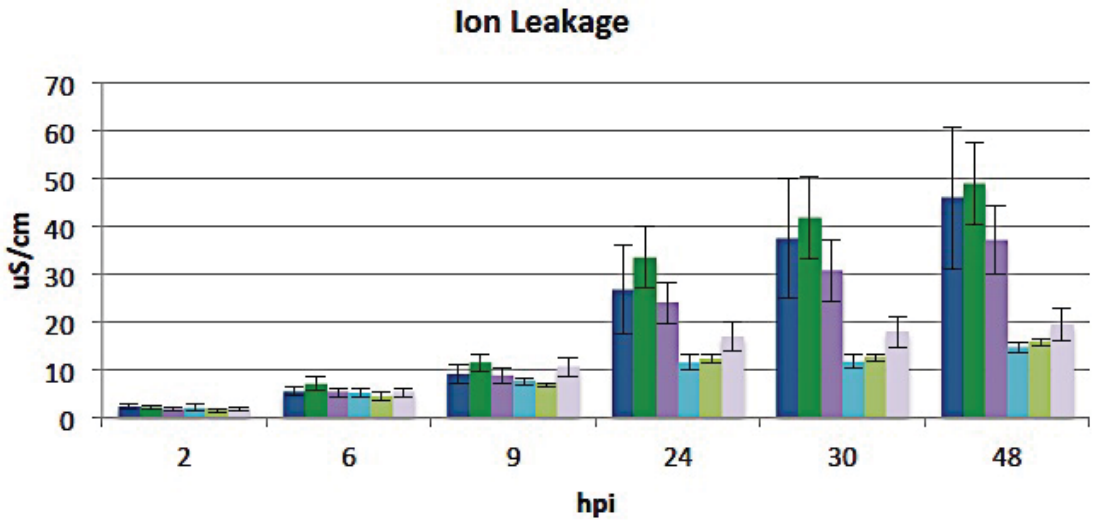

Figure 4.20. Ion leakage in *Arabidopsis thaliana*. Conductivity (μS/cm) of solution containing leaf discs inoculated with *Psa* N13 (blue bars), *Psa* NZ13 + pMT-1 (dark green bars), *Psa* NZ13 + pMT-2 (purple bars), *Psa* NZ13 Δ*hrcC* (azure bars), *Psa* NZ13 Δ*hrcC* + pMT-1 (light green bars) and *Psa* NZ13 Δ*hrcC* + pMT-2 (lilac bars). Data are means and standard deviation of four replicates.

4.2.5d Validation of the exportation assay

The absence of a HR on *A. thaliana* can be due to the unsuccessful delivery of the effector into the host cell; there is however the possibility that the fusion with the full length or the putative exportation signal of DctT compromises the conformation and therefore the functionality of AvrRpt2. To demonstrate that the fusion protein is able to induce the RIN4-RPS2-dependent cell death (HR), DctT:AvrRpt2 constructs of pMT-1 and pMT-2 (MT-1 and MT-2, respectively) were transiently expressed in *N. benthamiana* via infiltration with *Agrobacterium tumefaciens* (agroinfiltration) along with RIN4 and RPS2 (refer to Section 2.2.23).

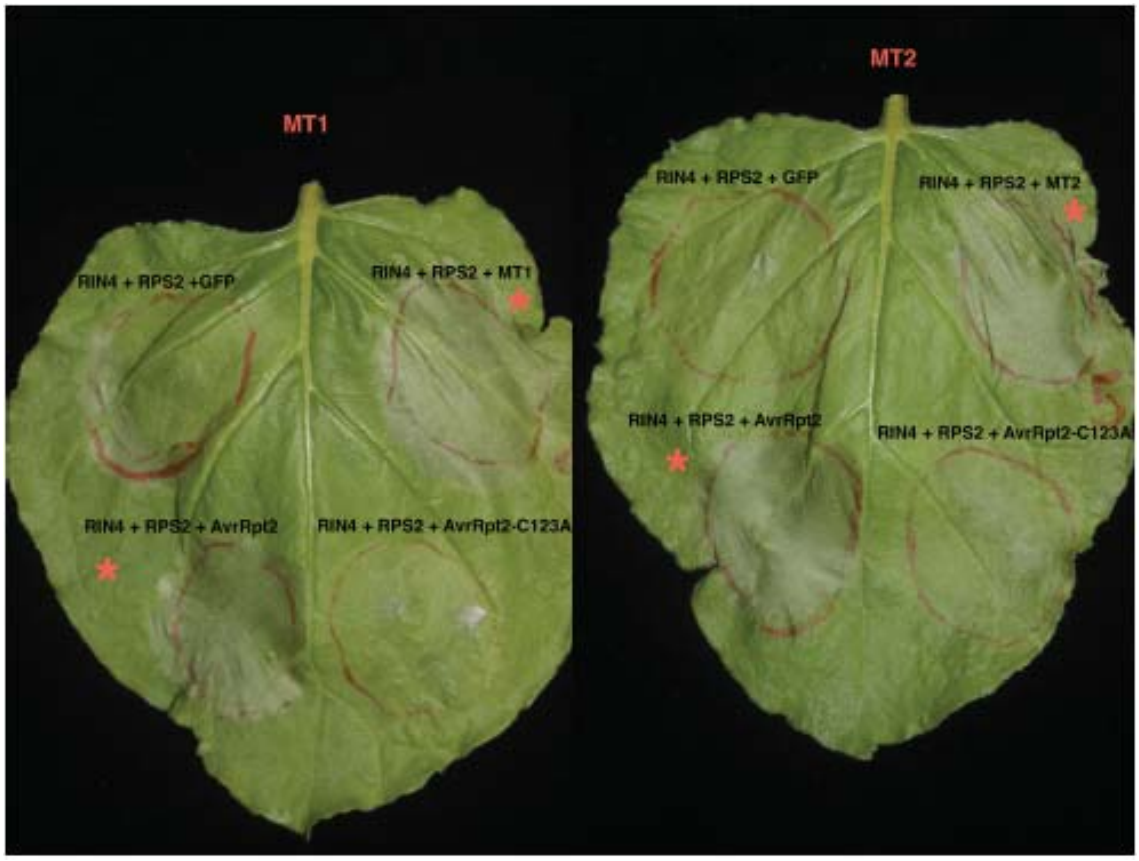

Figure 4.21. Agroinfiltration in *Nicotiana bethamiana* leaves. Pictures were taken 24 hpi. * indicates HR development.

N. benthamiana was pressure-inoculated with a) AGLRIN4, AGLRPS2, AGLGFP (negative control) b) AGLRIN4, AGLRPS2, AGLAvrRpt2 (positive control), c) AGLRIN4, AGLRPS2, AGLMT-1 (or AGLMT-2) and d) AGLRIN4, AGLRPS2, AGLAvrRpt2-C123A (negative control) (Figure 4.21). In the positive control AvrRpt2 interacts with RIN4, this modification is recognized by RPS2, which activates the HR. In the negative controls AvrRpt2 is (a) not present or (b) it is mutated in its active site and the HR is not triggered. When *N. benthamiana* is infiltrated with AGLRIN4, AGLRPS2, AGLMT-1 and AGLRIN4, AGLRPS2, AGLMT-2, the HR develops. This result confirms that the fusion of the full-length or of the putative exportation site does not alter the functionality of AvrRpt2, conclusively demonstrating the absence of DctT secretion into the host cell.

4.2.6 *In vitro* phenotypical characterization of Tn*6212*

The *in planta* experiments did not demonstrate the involvement of Tn*6212* in the plant-pathogen interaction; I thus continued the Tn*6212* characterization *in vitro*. I explored the possibility that Tn*6212* is involved in catabolism.

4.2.6a *In vitro* growth assay

The presence of genes involved in energy production and sugars utilization (enolase, inorganic pyrophosphatase, DctT – C4 sugar transporter, a catabolism-associated protein) suggests that Tn*6212* might be involved in catabolism. Therefore, growth of *Psa* NZ13 and the deletion mutant of the entire "enolase region" (*Psa* NZ13 Δ*enoR*) was evaluated in rich media (KB and LB), and in minimal media (M9 supplemented with glucose, succinate, malate or fumarate as only carbon source) (Figure 4.22) (refer to Section 2.2.12). Growth of the wildtype and the mutant was similar in the rich media and in M9 glucose, but in M9 succinate, malate and fumarate *Psa* NZ13 Δ*enoR* showed impaired growth. Tn*6212* confers a selective advantage when *Psa* NZ13 grows in the TCA cycle intermediates succinate, malate and fumarate as only carbon sources.

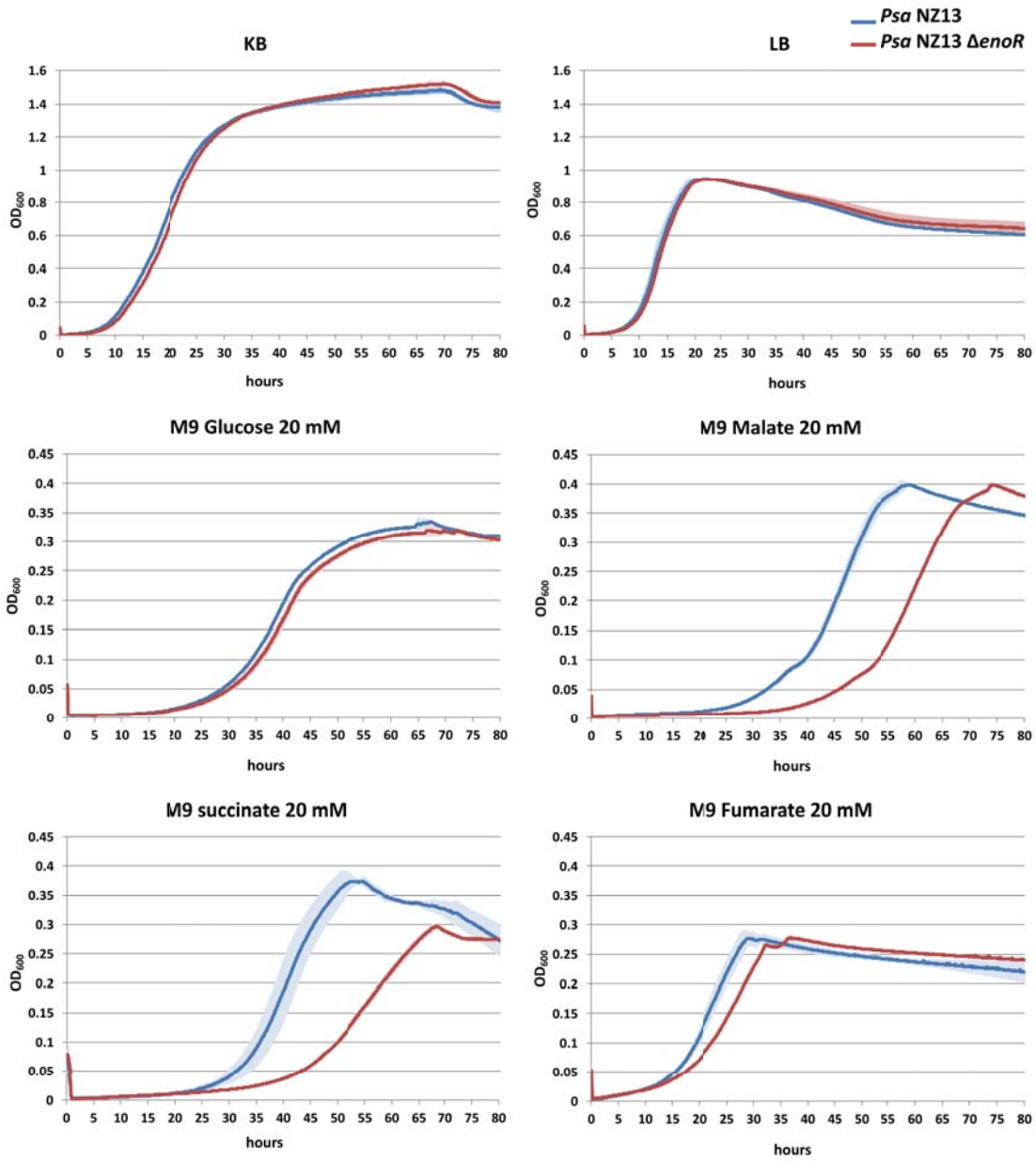

Figure 4.22. Growth of *Psa* NZ13 Δ*enoR* and the wildtype in different media. *Psa* NZ13 and *Psa* NZ13 Δ*enoR* were grown in a 96 weel plate and growth was recorded as optical density (OD_{600}) using a plate reader. Data are average growth of five replicates, error bars represent standard deviation. The experiment was repeated three times.

4.3 Discussion

The emergence of novel pathogens requires the establishment of effective techniques and assays to study their virulence. I have developed molecular techniques for mutagenesis and established protocols for the study of the plant-*Psa* interaction. *hrcC* is the ideal candidate gene to explore the effectiveness of the mutagenesis and to produce a negative control for pathogenicity assays. This mutant has proven to be of use to the community, it was also given to two different laboratories for their specific research projects. I also established a protocol for pathogenicity tests on plants that conformed to strict biosecurity regulations around working with *Psa* in New Zealand.

The absence of growth of both the wildtype and the *hrcC* mutant in *A. thaliana* suggested that this plant is a non-host for *Psa*. This finding was later confirmed in the Kee Sohn Lab. *Psa* is able to rapidly colonize tomato, but the symptoms induced dry papery lesions that resemble a HR. *N. bentamiana* was the best model plant tested with a very short experimental time (maximum density was attained in 2 days), though symptom production was faint and HrcC mutants also proliferated *in planta*. Both growth and symptom development can be tracked when using the natural host, kiwifruit. The two model plants were chosen for further experiments to cover the possibility that Tn*6212* could be involved in the colonization of plants that are not the principal host of the pathogen.

Tn*6212* is present exclusively in plant-associated *Pseduomonas* spp. either on conjugative elements (ICEs and plasmids) or integrated in the genome. Initially the Tn*6212* transposon was described exclusively in the outbreak strains of *Psa*, but the analysis conducted in Chapter 3 revealed it is the most common cargo

element in the PsICEs. Of the 87 PsICEs isolated in Chapter 3, 73 harbour Tn*6212*, of these 73, 65 have an almost identical Tn*6212*. The absence of variation in this element implies Tn*6212* has only recently disseminated across PsICEs or there are strong purifying selective pressures for the conservation of gene function. Even if it is a new element, this initial spread (Tn*6212* is present in 23 of the ~3,000 genomes interrogated in Chapter 3) implies the presence of a selective pressure for its acquisition and retention.

The recurrent presence of Tn*6212* in the PsICE family, harboured by plant-associated *Pseudomonas* spp., suggested it contributes to plant pathogen's fitness. The nature of the genes in the transposon and the presence of a hypothesized signal for the T3SS exportation in *dctT* suggested a role in virulence (McCann *et al.*, 2013). This led to the creation of mutants to explore their contribution to the plant pathogen interaction. Pathogenicity assays were carried out on three different plants (kiwifruit, tomato and *N. benthamiana*), yet Tn*6212* and the single enolase and *dctT* mutations did not show any significant contribution to plant colonization in *Psa*. The T3SS export assay and ion leakage experiments on *A. thaliana* indicate that DctT is not translocated by the T3SS into the host cell. Further experiments established that the cause for the absence of HR was failure of DctT to be translocated and not some malfunction of AvrRpt2.

In vitro growth assays demonstrated the involvement of the enolase region in the catabolism of succinate, fumarate and malate. *P. syringae* strains harbouring Tn*6212* may be favoured in the plant environment where TCA cycle intermediates are involved in energy production and also in other metabolic pathways (Sweetlove *et al.*, 2010). Moreover the nowadays elevated atmospheric concentrations of CO_2 (e[CO_2]) could enhance this advantage. Even if the response

to e[CO_2] varies between different species of plants, common responses are the increase of leaf photosynthetic rate, carbohydrates and biomass production (Rajkumar *et al*., 2013). These changes in the plant metabolism could lead to an increased concentration of the TCA cycle intermediates and make this selective advantage even greater.

Plant pathogenicity assays in laboratory settings might not reflect the natural environment *Psa* lives in. Cultivated kiwifruit vines are productive up to 20 years in an orchard, and do not produce fruit until they are 3 to 5 years old, yet I employed 3-4 week old plants grown on a soil-free substrate. The different stage of plant host development may influence the outcome of the experiments (Kus *et al*., 2002; Panter & Jones, 2002). My *in planta* experiments would also be unable to reveal fine scale impacts, which might cumulatively enhance the fitness of *Psa* over the long term, or at different stages in the *Psa* life cycle. The absence of a phenotypic change in the Tn*6212* mutants during plant colonization can thus not conclusevely be explained by the non-involvement of the transposon in the pathogenicity of *Psa* under a naïve model of pathogenicity (leaf assay in young seedlings).

It is also possible that Tn*6212* might not be involved in the plant-pathogen interaction. The initial hypothesis was formulated from a plant-pathogen interaction perspective, primarily because of the virulence of the outbreak *Psa* clade (McCann *et al*., 2013) and secondly because of the presence of this region in plant pathogens (Chapter 3). This distribution, however, needs to be considered with caution. Fondi *et al*. (2016) demonstrated that gene distribution is strongly affected by ecology: bacteria living in the same niche will have more similar accessory genes than more closely related species living in different environments.

Therefore, is not surprising that plant-associated bacteria share the same mobilome. The prevalence of genome sequencing of plant pathogenic *P. syringae* strains that influence humans from an economical perspective might be misleading and certainly does not reflect the diversity of the *Pseudomonas* spp. associated with plants. Only recently, with the drop of sequencing costs, have also non-pathogenic strains of *P. syringae* begun to be sequenced (Morris *et al.*, 2010; Monteil *et al.*, 2012; 2013). I have looked at Tn*6212* only from a plant-pathogen interaction point of view. Even if the region is present only in pathogenic *P. syringae,* these strains probably spend much of their time surviving in the environment *ex planta* or persisting outside of disease outbreaks (Hirano & Upper, 2000; Morris *et al.*, 2013) and this region may play an important role in this relatively understudied part of the bacterium's life cycle.

On the basis of different annotation deposited in GenBank, DctT could be a tellurite resistance protein, the transporter could be an arsenite efflux pump, the voltage-gated Cl channel might be involved in the maintenance of pH homeostasis in a similar way of *yjcE* (Águia-Clares *et al.*, 2017). In *Erwinia amylovora* the Na^+/H^+ exchanger *yjcE* is activated in response to Cu stress and maybe involved in the maintenance of pH homeostasis in cells, which could reduce the cycling of Cu^+ and Cu^{2+} (Águia-Clares *et al.*, 2017). Butler *et al.* (2013) recognized the presence of a HgR gene in Tn*6212*. Indeed, when considering the whole picture of the accessory genes of the PsICE family, they harbour resistance genes to the most common heavy metal pollutants: chrome, cadmium, mercury, copper and arsenic, with the notable exception of aluminium and lead. Considering this, I speculate that Tn*6212* might be implicated in heavy metal tolerance, though this awaits further experimental analysis.

Chapter 5 – Evolution of copper resistance in *Pseudomonas syringae* pv. *actinidiae*

This Chapter was published as: Colombi, E., Straub, C., Kunzel, 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. Environ Microbiol, 19, 819–832.

The paper was accompanied by a Research Highlight in the same issue: Lindow, S. E. (2017). Horizontal gene transfer gone wild: promiscuity in a kiwifruit pathogen leads to resistance to chemical control. Environ Microbiol, 19, 1363-1365.

5.1 Introduction

Horizontal gene transfer (HGT) is a potent evolutionary process that significantly shapes patterns of diversity in bacterial populations. Horizontally transmissible elements, including plasmids, phages and Integrative Conjugative Elements (ICEs) move genes over broad phylogenetic distances and mediate abrupt changes in niche preferences that may even fuel speciation (Médigue *et al.*, 1991; Lan & Reeves, 1996; Sullivan & Ronson, 1998; Ochman *et al.*, 2000; Ochman *et al.*, 2005; Retchless & Lawrence, 2007; Guglielmini *et al.*, 2011; Popa & Dagan, 2011; Polz *et a*l., 2013).

ICEs are plasmid-like entities with attributes of temperate phages that disseminate vertically as part of the bacterial chromosome and horizontally by virtue of endogenously encoded machinery for conjugative transfer (Wozniak &

Waldor, 2010; Guglielmini *et al.*, 2011). Essential genetic modules include those mediating integration, excision, conjugation and regulation of conjugative activity (Mohd-Zain *et al.*, 2004; Juhas *et al.*, 2007; Roberts & Mullany, 2009). During the process of conjugation ICEs circularize and transfer to new hosts, leaving a copy in the original host genome (Wozniak & Waldor, 2010; Johnson & Grossman, 2015). Conjugation during pathogenesis is often regulated by environmental signals (Lovell *et al.*, 2009; Quiroz *et al.*, 2011; Vanga *et al.*, 2015).

In addition to a set of essential genes, ICEs often harbour 'cargo' genes of adaptive significance to their hosts. These include genes affecting biofilm formation, pathogenicity, antibiotic and heavy metal resistance, symbiosis and bacteriocin synthesis (Peters *et al.*, 1991; Rauch *et al.*, 1992; Ravatn *et al.*, 1998; Beaber *et al.*, 2002; Drenkard *et al.*, 2002; Burrus *et al.*, 2006; Ramsay *et al.*, 2006; Dimopoulou *et al.*, 2007; Kung *et al.*, 2010). The genetic information stored in cargo genes varies considerably resulting in ICEs that range in size from 20 kb to 500 kb (Johnson & Grossman, 2015).

In 2008 a distinct and particularly virulent form of the kiwifruit pathogen *Pseudomonas syringae* pv. *actinidiae* (*Psa*) was identified in Italy. It was subsequently disseminated throughout kiwifruit growing regions of the world causing a global pandemic that reached New Zealand (NZ) in 2010 (Balestra *et al.*, 2010; Abelleira *et al.*, 2011; Everett *et al.*, 2011; Vanneste *et al.*, 2011). Genomic analysis showed that although the pandemic was derived from a single clone it acquired a set of distinctive ICEs during the course of its global journey (Mazzaglia *et al.*, 2012; Butler *et al.*, 2013; McCann *et al.*, 2013). The NZ lineage carries Psa_{NZ13}ICE which harbours a 20 kb "enolase region" that is also found in otherwise divergent *Psa* ICEs (McCann *et al.*, 2013; McCann *et al.*, 2017).

Copper sprays have long been used in NZ to protect plants from a range of diseases. Since the arrival of *Psa* in NZ, kiwifruit orchardists have employed copper-based products to protect vines. From 2011 an ongoing industry-based programme has been in place to monitor copper resistance. In 2014, evidence was first obtained of *Psa* isolates resistant to copper sulphate. Given that the clone of *Psa* originally introduced into NZ was sensitive to copper and lacked genes encoding copper resistance (McCann *et al.*, 2013), detection of copper resistance raised the possibility that the evolution of copper resistance in *Psa* is an evolutionary response to the use of copper-based sprays.

5.1.1 Aims

- Investigate the transfer dynamics of Psa_{NZ45}ICE_Cu both *in vitro* and *in planta*,
- investigate the consequences in term of fitness cost and benefit of the acquisition of Psa_{NZ45}ICE_Cu,
- investigate the host range of Psa_{NZ45}ICE_Cu.

5.2 Results

5.2.1 Occurrence of copper resistance in *Pseudomonas syringae* pv. *actinidiae*

Psa NZ13, isolated in 2010 and representative of the clone introduced in NZ, lacks genes encoding copper resistance (McCann *et al.*, 2013) and is unable to grow at copper concentrations in excess of 0.8 mM $CuSO_4$ on MGY. Prior to 2014 no copper-resistant or tolerant *Psa* strains had been reported in NZ. Two strains isolated from two different kiwifruit orchards in 2014, *Psa* NZ45 and *Psa* NZ47, harboured two distinct PsICEs encoding copper resistance. These strains displayed copper resistance with a MIC of 1.2 mM $CuSO_4$ on MGY plates. This finding

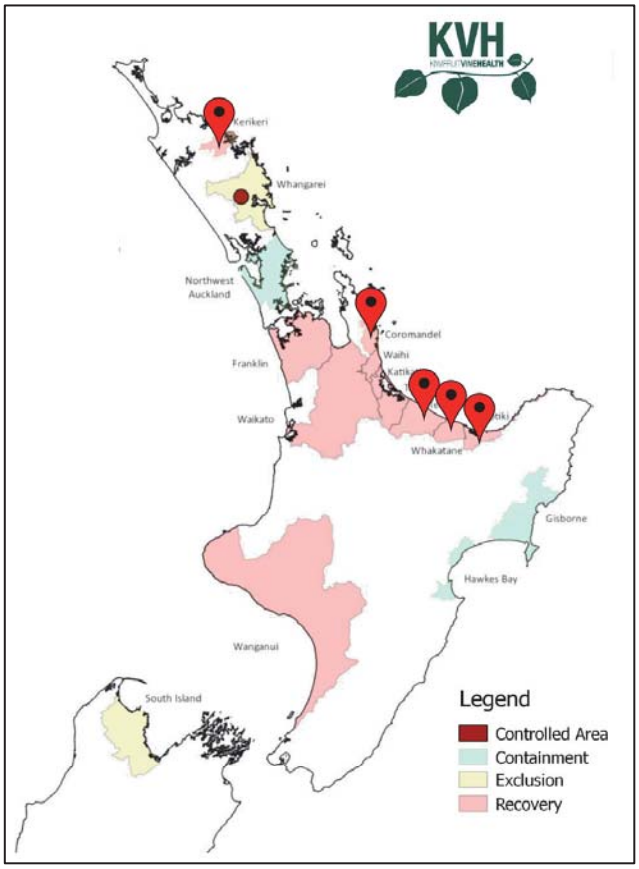

Figure 5.1. New Zealand kiwifruit growing regions with isolation sites of copper-resistant *Pseudomonas syringae* pv. *actinidiae*. Map was modified from regional classification map of June 2016 (KVH).

prompted a small-scale sampling of both copper-treated and untreated orchards in 2015/2016 encompassing the area where resistance was first identified. From a sample of 213 strains isolated from seven orchards 59 were found to be copper-resistant. Copper-resistant isolates were recovered from both copper-treated and untreated orchards. Additional copper-resistant strains were procured from other kiwifruit-growing regions of NZ by KVH (Kiwifruit Wine Health) (Figure 5.1).

5.2.2 ICE and plasmid-mediated acquisition of copper resistance in *Pseudomonas syringae* pv. *actinidiae*

The genome of the focal copper-resistant isolate, *Psa* NZ45, is a direct clonal descendant of the isolate originally introduced into NZ (*Psa* NZ13) with 2 SNPs in the 4,853,155 bp non-recombinant core genome alignment (McCann *et al*., 2017), but differs in two significant regards. Firstly, the 'native' ICE (Psa_{NZ13}ICE) at *att*-1 (immediately upstream of *clpB*), is located at the second *att* site (*att*-2) immediately downstream of *queC* (Figure 5.2). Secondly, the genome harbours a new 107 kb ICE (Psa_{NZ45}ICE_Cu) integrated at the *att*-1 site: Psa_{NZ45}ICE_Cu carries genes encoding copper resistance (Figures 5.2 and 5.3A).

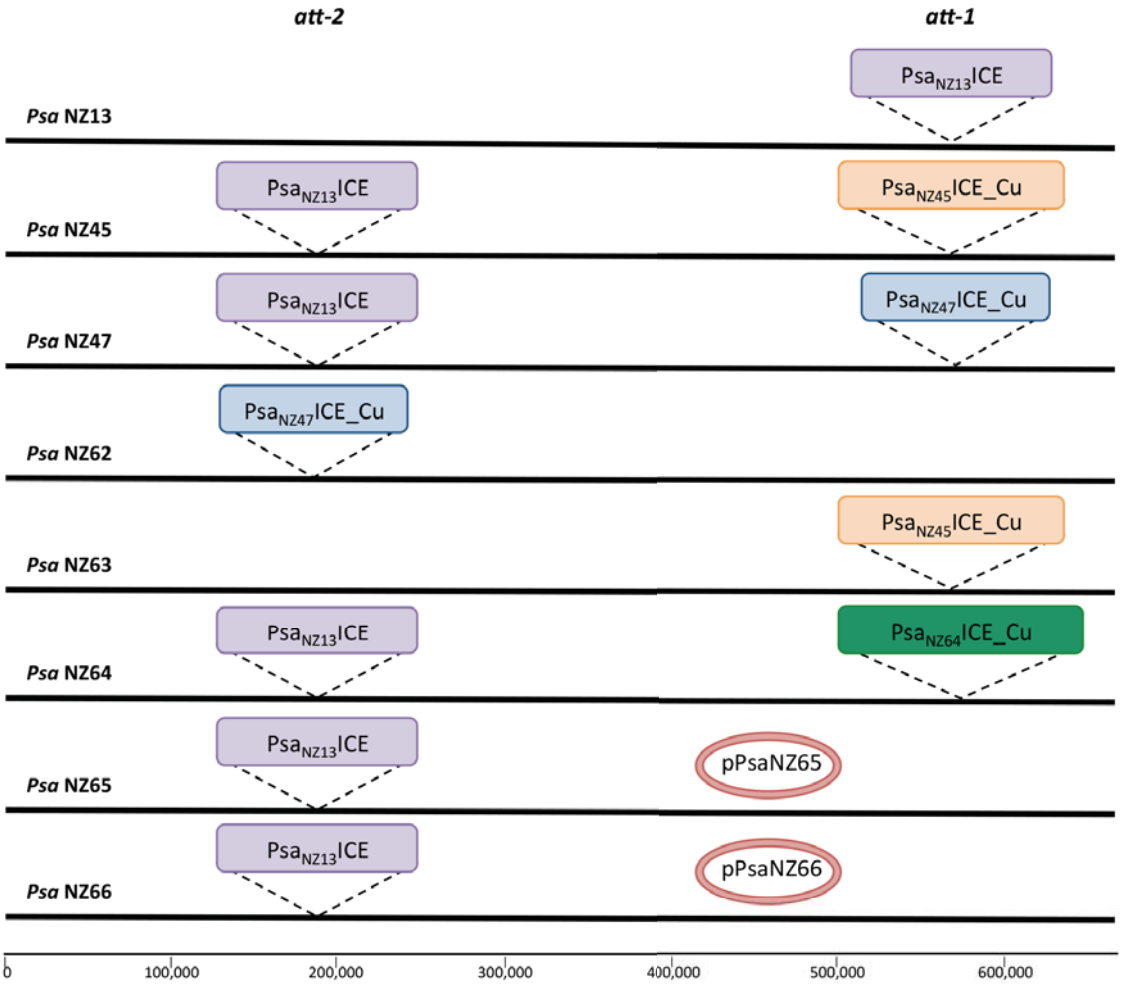

Figure 5.2. Genomic location of PsICEs in *Pseudomonas syringae* pv. *actinidiae*. In purple the Psa$_{NZ13}$ICE (100 kb), in orange Psa$_{NZ45}$ICE_Cu (107 kb), in blue the Psa$_{NZ47}$ICE_Cu (90 kb), in green the Psa$_{NZ64}$ICE_Cu (130 kb), pPsaNZ65 and pPsaNZ66 plasmids are 111 kb. Each island is bounded by 52 bp *att* sequences overlapping tRNALys. In *Psa* NZ13 the *att-1* site is located at 5,534,632 bp, *att-2* at 1,733,972 bp. The figure is not to scale (the entire genome of 6.7 Mbp is indicated a single black line). Both Psa$_{NZ13}$ICE and Psa$_{NZ47}$ICE_Cu ICEs were detected in *Psa* NZ47 by sequencing, but analysis of independent colonies from the freezer stock show that Psa$_{NZ13}$ICE is prone to loss.

The genomes of six additional copper-resistant *Psa* isolates were also sequenced (Table 5.1) and as with *Psa* NZ45, reads were aligned against the *Psa* NZ13 reference genome (McCann *et al.*, 2013; Templeton *et al.,* 2015). All six harbour mobile elements carrying genes encoding copper resistance. The diversity of these elements and genomic location is shown in Figure 5.2 and their structure is represented in Figure 5.3A.

Isolate ID	Place of isolation	Year of isolation	Orchard's copper programme	MIC to $CuSO_4$	GenBank accession number
Psa NZ13	Te Puke, NZ	2010	NA	0.8 mM	CP011972-3
Psa NZ45	Te Puke, NZ	2014	Full spray	1.2 mM	CP017007-8
Psa NZ47	Te Puke, NZ	2014	Spray free	1.2 mM	CP017009-11
Psa NZ62	Te Puke, NZ	2015	Organic	1.2 mM	MOMK00000000
Psa NZ63	Te Puke, NZ	2015	To minimum	1.2 mM	MOML00000000
Psa NZ64	Te Puke, NZ	2016	Spray free	1.2 mM	MOMM00000000
Psa NZ65	Te Puke, NZ	2016	Full spray	1.5 mM	MOMN00000000
Psa NZ66	Coromandel, NZ	2016	Full spray	1.5 mM	MOMJ00000000

Table 5.1. List of the *Pseudomonas syringae* pv. *actinidiae* genomes used in this study

All isolates are direct clonal descendants of *Psa* NZ13 and thus share an almost identical genome with the exception of the determinants of copper resistance. In *Psa* NZ47 the genes encoding copper resistance are located on the 90 kb ICE (Psa_{NZ47}ICE_Cu) integrated at the *att-1* site: the native Psa_{NZ13}ICE is located at the *att-2* site. *Psa* NZ62 carries an ICE identical to that found in Psa_{NZ47}ICE_Cu (Psa_{NZ62}ICE_Cu), but is integrated at the *att-2* site; the native ICE (Psa_{NZ13}ICE) is absent leaving the *att-1* site unoccupied. Isolate *Psa* NZ63 carries Psa_{NZ45}ICE_Cu integrated at the *att-1* site, but, the native Psa_{NZ13}ICE has been lost. Copper resistance genes in isolate *Psa* NZ64 are also ICE-encoded, but the NZ64 ICE (Psa_{NZ64}ICE_Cu) is genetically distinct from both Psa_{NZ47}ICE_Cu and Psa_{NZ45}ICE_Cu – at 130 kb, it is also the largest. In NZ64, Psa_{NZ64}ICE_Cu is located at the *att-1* site and the *att-2* site contains the native (Psa_{NZ13}ICE) ICE. Isolates *Psa* NZ65 and NZ66 both harbour copper resistance genes on a near identical, 120 kb previously undescribed plasmid (pPsaNZ65 and pPsaNZ66, respectively). The only significant difference among the plasmids is the location of a strepomycin resistance-encoding transposon (see below): both isolates have the original Psa_{NZ13}ICE integrated at the *att-2* site (Figure 5.2). *Psa* harbouring copper resistance-

encoding ICEs have a MIC $CuSO_4$ of 1.2 mM while the MIC of plasmid-carrying *Psa* 1.5 mM (Table 5.1).

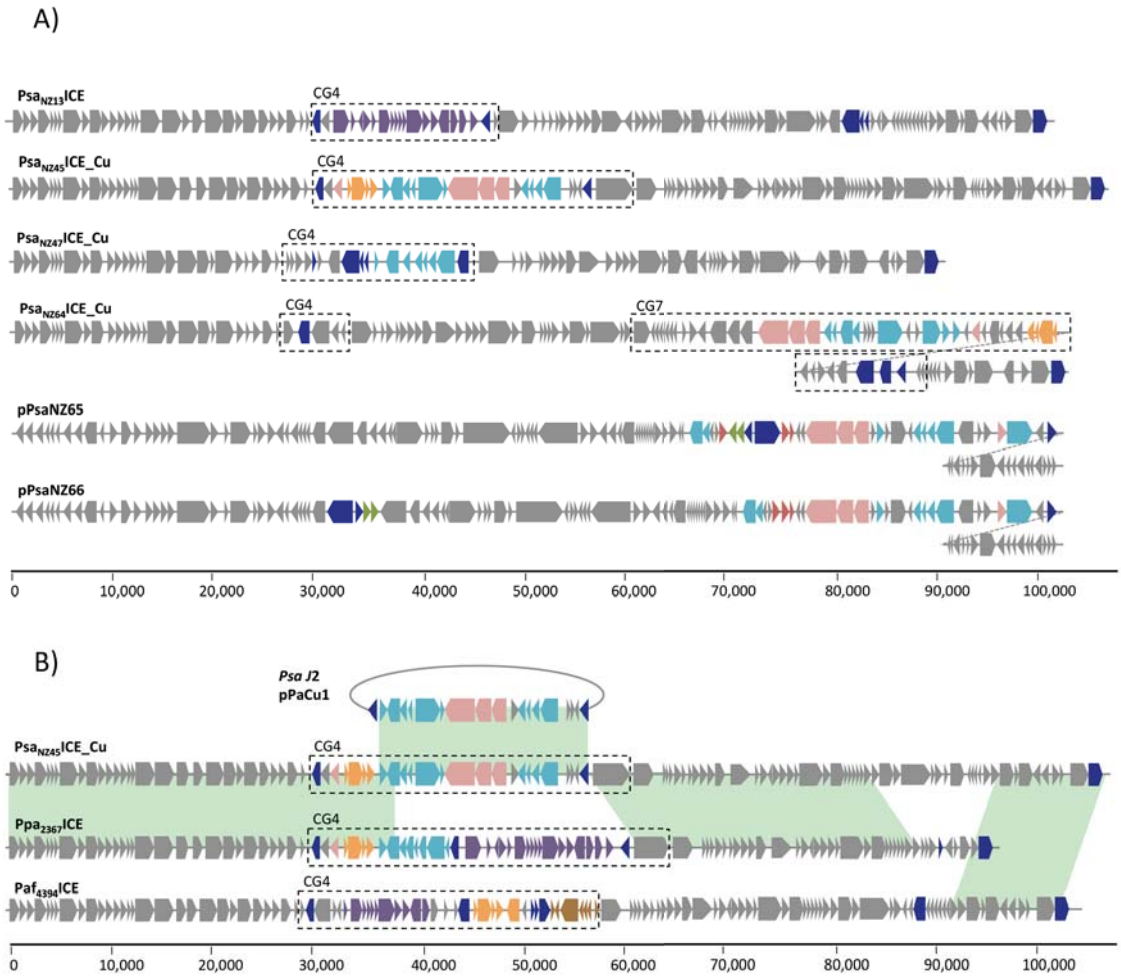

Figure 5.3. Genetic organization of ICEs and plasmids acquired by *Pseudomonas syringae* pv. *actinidiae* and mosaicism of Psa_{NZ45}ICE_Cu.

A) Blue boxes are mobile genes (transposases or recombinases), purple boxes define the "enolase region", azure boxes depict copper resistance genes, orange boxes are arsenic resistance genes, pink boxes are genes belonging to the *czc*/*cus* system, green boxes are streptomycin resistance genes, red boxes are cation transporter ATPases, brown boxes denote genes encoding mercury resistance. Core backbone and other cargo genes are depicted as grey boxes. Dotted diagonal lines indicate continuation of the element.

B) Areas in green show more than 99% pairwise nucelotide identity. Psa_{NZ45}ICE_Cu and Ppa_{2367}ICE share identity both in the first 38 kb and 20 kb downstream of CG4. The remaining 20 kb of the Psa_{NZ45}ICE_Cu CG4 is almost identical to pPaCu1 (it differs by just 2 SNPs). The last 12.5 kb of Psa_{NZ45}ICE_Cu is identical to Paf_{4394}ICE.

That such a small sample of isolates is each unique with regard to the copper resistance-encoding element points to highly dynamic processes shaping their evolution. These samples came from a relatively small geographical location.

Two ICEs, Psa_{NZ64}ICE_Cu and Psa_{NZ47}ICE_Cu were found in different isolates sampled from the same orchard (one year apart), although the two isolates containing near identical plasmids were isolated from orchards located 100 km apart. Two isolates sampled one year apart from the same location (neighbouring orchards in Te Puke) carry the same ICE (Psa_{NZ45}ICE_Cu=Psa_{NZ63}ICE_Cu; Psa_{NZ47}ICE_Cu=Psa_{NZ62}ICE_Cu) (Table 5.1).

The dynamics of ICE evolution becomes especially evident when placed in the broader context possible by comparisons to ICEs recorded in DNA databases. Psa_{NZ45}ICE_Cu is a mosaic of DNA from two known ICEs and a plasmid. It shares regions of near perfect identity (over 66 kb) with ICEs present in the otherwise divergent host genomes of *P. syringae pv. panici* (*Ppa*, LGM2367) isolated from proso millet in Madison (USA) in the 1920s (over the first 38 kb it differs by just 12 SNPs, and one 144 bp indel), *P. syringae* pv. *atrofasciens* (*Paf*, ICMP4394) isolated in NZ in 1968 from wheat, and a 70.5 kb plasmid present in a non-pandemic *Psa* strain (J2), isolated in Japan in 1988 (Figure 5.3B).

Two of the ICEs described here are present in non-*Psa Pseudomonas* isolated from kiwifruit leaves. Psa_{NZ47}ICE_Cu shows 99.7% pairwise nucleotide identity with an ICE found in *P. marginalis* ICMP 11289 isolated in 1991 from *Actinidiae chinensis* var. *deliciosa* in Katikati (NZ). Psa_{NZ64}ICE_Cu is almost identical (99.5% nucleotide pairwise identity) to an ICE from *P. syringae* pv. *actinidifoliorum* (*Pfm*) ICMP19497, isolated from kiwifruit in 2010 in Te Puke (NZ) (Visnovsky *et al.*, 2016) (Table 5.1). Additionally, a 48 kb segment of coding copper resistance genes Psa_{NZ64}ICE_Cu shares 99.3% nucleotide pairwise identity with a locus found in *P. azotoformans* strain S4 (Fang *et al.*, 2016), which was

isolated from soil in 2014 in Lijiang (China). However, the locus from *P. azotoformans* is not associated with an ICE.

5.2.3. Genetic determinants of copper resistance

ICEs identified in *Psa* isolates harbour operons encoding examples of both resistance mechanisms (and regulators), plus genetic determinants of resistance to other metal ions. In each instance the resistance genes are located within cargo gene regions (CG, refer to Chapter 3) (Figure 5.3A). Overall there are notable similarities and differences in the organization of the CGs.

As shown in Figure 5.3B, the first 38 kb of Psa_{NZ45}ICE_Cu is almost identical (99.7% identical at the nucleotide level) to Ppa_{2367}ICE. This region spans the core genes, but extends ~8.2 kb into the variable cargo genes with just two SNPs distinguishing the two PsICEs (across the 8.2 kb CG4). Encoded within this region is an integrase, arsenic resistance genes (*arsRBCH*), a gene implicated in cadmium and cobalt resistance (*czcD*) and the *copRS* regulatory system. Partway through *copS* the two ICEs diverge at a recombination breakpoint with the downstream CG4 from Psa_{NZ45}ICE_Cu being homologous to a set of copper resistance genes found on plasmid pPaCu1 from the divergent (non-pandemic) Japanese isolate of *Psa* J2 (Nakajima *et al.*, 2002). This region comprises a putative copper transporting ATPase encoded by *copG* (Gutiérrez-Barranquero *et al.*, 2013), *cusABC* genes involved in the detoxification of monovalent cations, including copper and silver (Mergeay *et al.*, 2003; Rensing & Grass, 2003) and *copABCD* (Figure 3B). The last 4 kb of the CG4 of Psa_{NZ45}ICE_Cu shares almost complete identity with Ppa_{2367}ICE (Figure 3B).

Detail of the diversity of copper resistance (and related metal resistance) genes is shown in Figure 5.4.

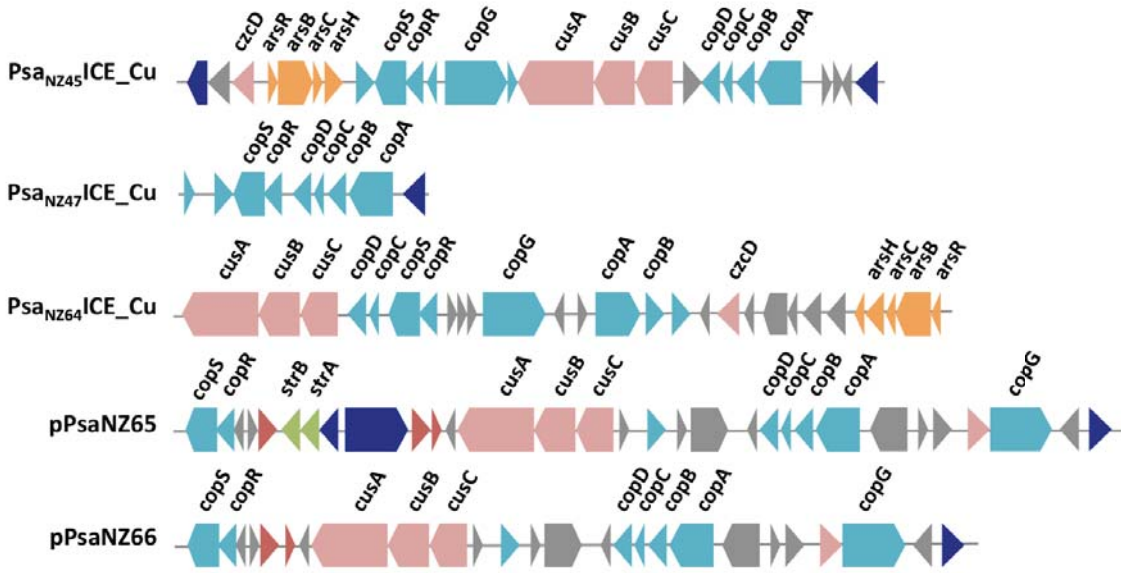

Figure 5.4. Genetic organization of metal resistance loci. Blue boxes are mobile genes (transposases or recombinases), azure boxes depict copper resistance genes, orange boxes are arsenic resistance genes, pink boxes are genes belonging to the *czc/cus* system, green boxes are streptomycin resistance genes and other genes are depicted as grey boxes.

All elements (ICEs and plasmids) harbour the *copRS* regulatory system and, with the exception of Psa$_{NZ47}$ICE_Cu, all carry both *cusABC* and *copABCD*, although their organization varies. For example, while *copABCD* is typical, in Psa$_{NZ64}$ICE_Cu *copAB* and *copCD* are organized as two separate operons (Figure 5.4). The putative copper ABC transport system encoded by *copG* is a common feature, and determinants of arsenic resistance are present in both Psa$_{NZ45}$ICE_Cu and Psa$_{NZ64}$ICE_Cu. The putative cadmium and related metal resistance gene, *czcD* is also present on these two PsICEs. As noted above, a transposon carrying determinants of streptomycin resistance (*strAB*) is present on plasmids pPsaNZ65 and pPsaNZ66. The transposon is of the Tn*3* family and the cassette bears identity to streptomycin resistance carrying transposons found in *P. syringae* pv. *syringae* B728a (Feil *et al.*, 2005), but also on plasmid pMRVIM0713 from *Pseudomonas aeruginosa* strain MRSN17623 (GenBank: KP975076.1), plasmid pPMK1-C from

Klebsiella pneumoniae strain PMK1 (Stoesser *et al.*, 2014), and plasmid pTi carried by *Agrobacterium tumefaciens* LBA4213 (Ach5) (GenBank: CP007228.1).

At the level of the operons determining copper resistance there is marked genetic diversity, however, with the exception of CopR, there is relatively little evidence of within operon recombination. The CusABC system is carried on pPsaNZ65 and pPsaNZ66 (but these are identical) and the ICEs Psa_{NZ45}ICE_Cu and Psa_{NZ64}ICE_Cu: CusA, CusB and CusC show 75.8%, 50.0% and 44.8% pairwise amino acid identity, respectively; phylogenetic trees based on protein sequences show congruence (Figure 5.5A). The CopABCD system is present on Psa_{NZ45}ICE_Cu, Psa_{NZ47}ICE_Cu, Psa_{NZ64}ICE_Cu (but CopAB and CopCD are in different locations (Figure 5.4)) and plasmid pPsaNZ65 (and pPsaNZ66): CopA, CopB, CopC and CopD show 76.4%, 63.1%, 79.1% and 60.8% pairwise amino acid identity, respectively. With the exception of CopC (where bootstrap support is low) phylogenetic trees for each protein show the same overall arrangement (Figure 5.5B). The two-component regulatory system *copRS* is also present on each of the elements with the amino acid sequences of CopR showing 84.3% and those of CopS 63.0% pairwise amino acid identity. Phylogenetic trees show CopS from Psa_{NZ64}ICE_Cu to be the most divergent, and those from Psa_{NZ45}ICE_Cu and Psa_{NZ47}ICE_Cu being most similar: CopR shows the same phylogenetic arrangement, however, CopR sequences from Psa_{NZ45}ICE_Cu and Psa_{NZ47}ICE_Cu are identical at the protein level suggesting a recent recombination event (Figure 5.5C).

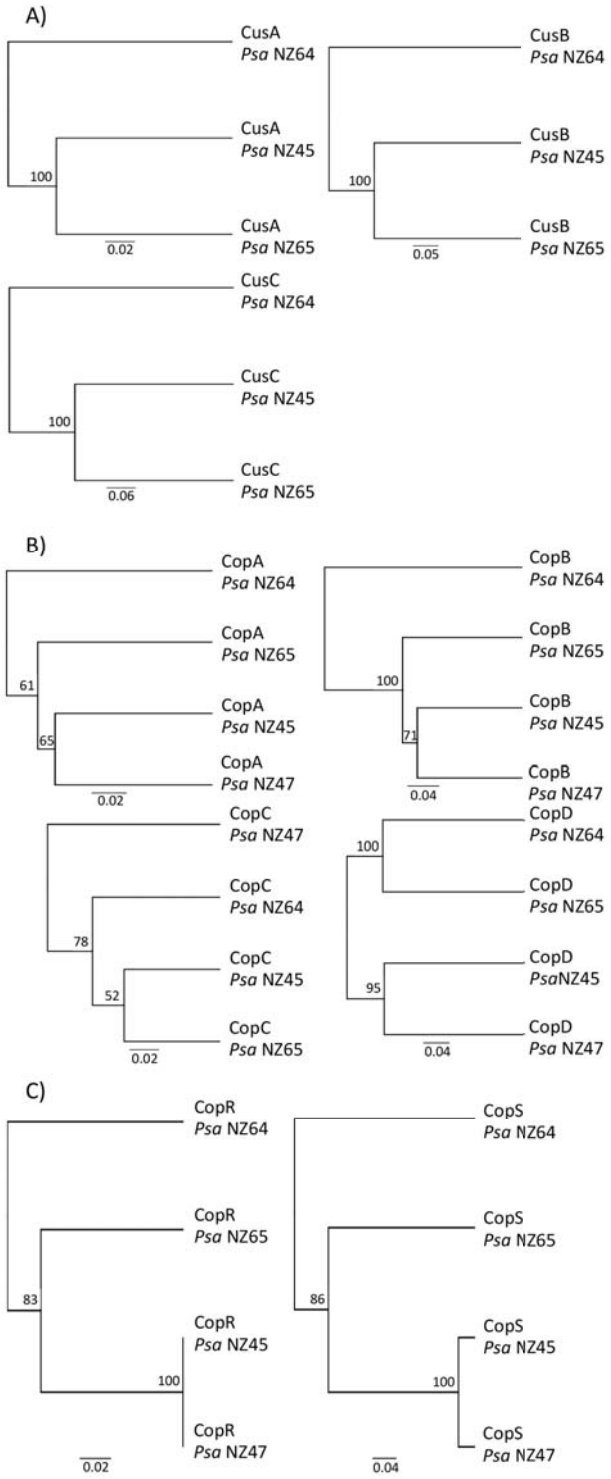

Figure 5.5. UPMGA trees of A) Cus and B) Cop system, C) CopR and CopS proteins in *Psa* NZ. Bootstrap values are shown at each node.

5.2.4 Psa$_{NZ45}$ICE_Cu imposes no detectable fitness cost and confers a selective advantage *in vitro* in the presence of copper

To determine whether ICE carriage confers a fitness cost, I took advantage of the fact that *Psa* NZ13 and *Psa* NZ45 are essentially isogenic, with the exception of the additional ICE in *Psa* NZ45 (Psa_{NZ45}ICE_Cu). Each strain was grown alone and density of cells monitored over a 72 hour period with samples taken every 24 hours. In the absence of copper sulphate, no difference in cell density was detected; however, in the presence of 0.5 and 0.8 mM $CuSO_4$ the density of *Psa* NZ13 was reduced (Figure 5.6). There is thus no apparent fitness cost associated with carriage of Psa_{NZ45}ICE_Cu in the absence of copper sulphate, but there is a fitness advantage in copper-containing environments.

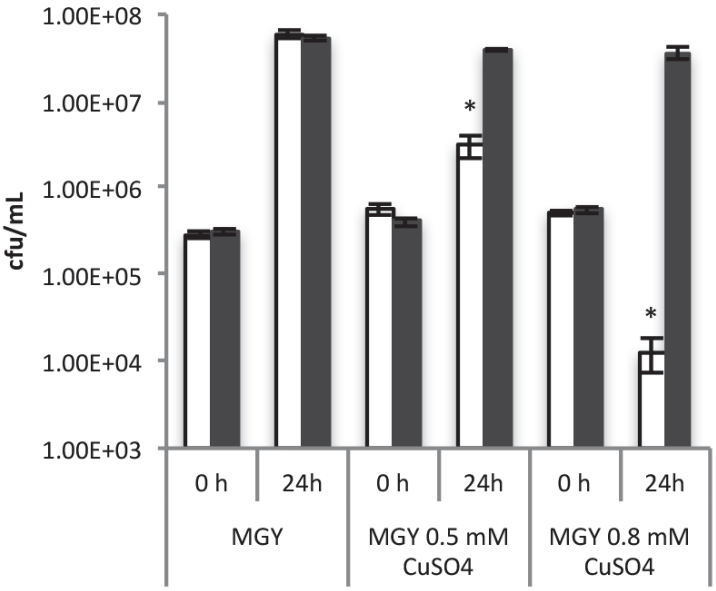

Figure 5.6. Effect of copper ions on growth of *Pseudomonas syringae* pv. *actinidiae* NZ13 and *P. syringae* pv. *actinidiae* NZ45. *Psa* NZ13 (white bars) and *Psa* NZ45 (grey bars) were grown for 24 h in shaken MGY culture and MGY supplemented with 0.5mM and 0.8 mM $CuSO_4$. Data are means and standard deviation of three independent cultures. *indicates significant difference $P<0.05$ (two tailed *t*-test)).

Although carriage of Psa_{NZ45}ICE_Cu appeared not to affect the growth of *Psa* NZ45 in the absence of copper, a more precise measure of fitness was sought by performing competition experiments in which *Psa* NZ13 and *Psa* NZ45 were co-cultured. For this experiment *Psa* NZ13 was marked with a kanamycin resistance

cassette so that it could be distinguished from kanamycin-sensitive, copper-resistant *Psa* NZ45. Over a 24 hour period where the two strains (founded at equal density) competed for the same resources (in shaken MGY medium without copper sulphate), the fitness of *Psa* NZ45 was not significantly different to *Psa* NZ13 (0.29 $\pm$ 0.2; mean and SEM from 3 independent experiments, each comprised of 3 replicates), indicating no significant detectable cost of carriage of Psa_{NZ45}ICE_Cu.

Given that the mechanism of copper resistance in *Psa* NZ45 – based upon *copABCD* – likely involves sequestration of copper ions I considered the possibility that this isolate might confer cross protection to non-copper-resistant isolates, such as *Psa* NZ13. To this end I performed co-culture experiments at sub-inhibitory (0.5 mM) and inhibitory (0.8 mM) copper sulphate concentrations. Growth of *Psa* NZ13 at sub-inhibitory concentrations of copper sulphate was significantly impaired by the presence of *Psa* NZ45 and this was especially evident at 48 and 72 hours (Figure 5.7). At the inhibitory copper sulphate concentration, *Psa* NZ13 appeared to benefit from the presence of *Psa* NZ45 (Figure 5.7).

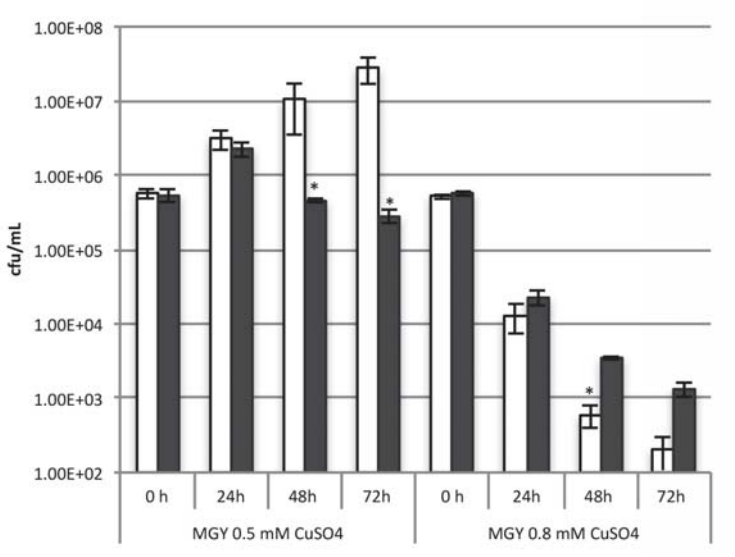

Figure 5.7. Density of single and co-cultured *Pseudomonas syringae* pv. *actinidiae* in liquid MGY supplemented with 0.5 m and 0.8 mM $CuSO_4$. *Psa* NZ13 was cultured alone

(white bars) or co-cultured with *Psa* NZ45 (grey bars). Data are means and standard deviation of three independent cultures. *indicates significance at 5% level by two-tailed *t*-test.

5.2.4. Psa$_{NZ45}$ICE_Cu imposes no detectable fitness cost but confers a minor selective advantage *in planta*

Cost and benefit of carrying Psa$_{NZ45}$ICE_Cu was also evaluated during endophytic colonization of kiwifruit leaves. No significant difference was observed in growth of singly-inoculated *Psa* NZ13 and NZ45 (dip inoculation was used to found colonization) independently of the application of a commonly used commercial copper-based product (Nordox75 (0.375 g/L)) subsequent to dip inoculation (Figure 5.8).

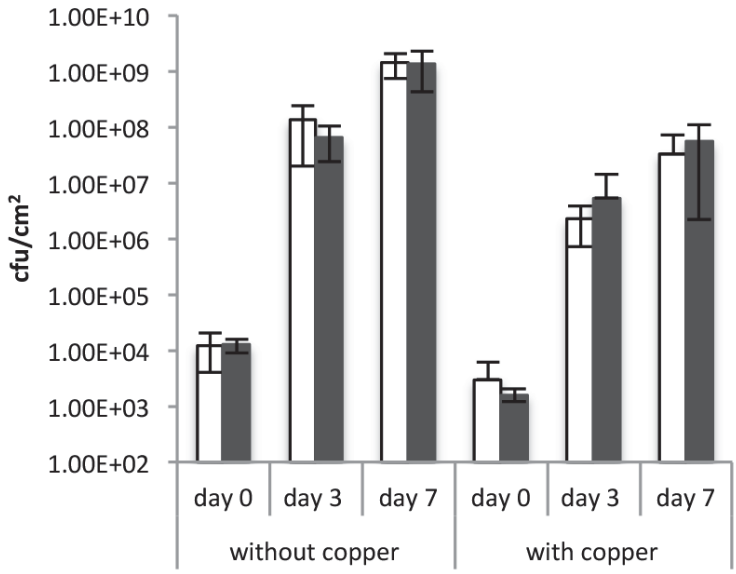

Figure 5.8. *In planta* growth of *Pseudomonas syringae* pv. *actinidiae* NZ13 and *P. syringae* pv. *actinidiae* NZ45. The single growth of *Psa* NZ13 (white bars) and *Psa* NZ45 (grey bars) was assessed endophytically on leaves of the kiwifruit cultivar Hort16A. The copper product Nordox75 (0.375g L^{-1}) was sprayed adaxially and abaxially until run off. Data are means and standard deviation of 5 replicates. two tailed *t*-test showed no statistical difference in growth between of *Psa* NZ13 and NZ45 in absence or presence of copper.

Co-cultivation competition assays in the presence or absence of Nordox75 confirmed carriage of Psa$_{NZ45}$ICE_Cu imposes no significant fitness cost or

advantage during endophytic growth (Figure 5.9). However, there was an advantage to epiphytic cells carrying Psa_{NZ45}ICE_Cu that was observed at day 3, but not at day 7 (by day 7 tissue damage was severe making it impossible to distinguished epiphytic from endophytic colonists) (Figure 5.9).

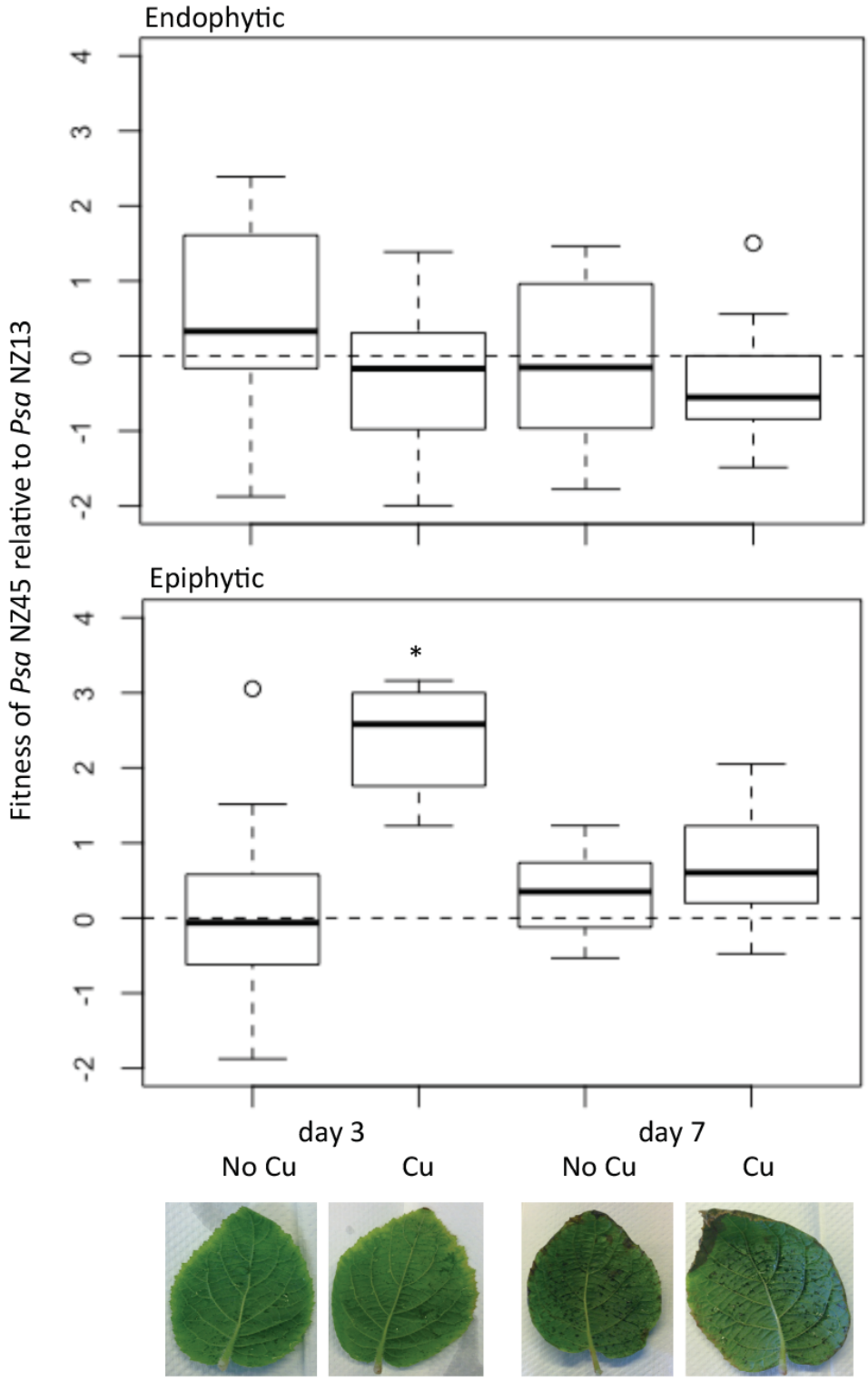

Figure 5.9. Competition assay of *Pseudomonas syringae* pv. *actinidiae* NZ45 and *P. syringae* pv. *actinidiae* NZ13 *in* and *super planta*. The difference in fitness between *Psa* NZ45 and *Psa* NZ13 was assessed endophytically and epiphytically on leaves of the kiwifruit cultivar Hort16A. Data represent the difference in the Malthusian parameters of three experiments of five replicates. The copper product Nordox75 (0.375g L^{-1}) was sprayed adaxially and abaxially until run off (Cu). *indicates significant difference $P<0.05$ from 0 (one sample *t*-test)).

5.2.5 Psa_{NZ45}ICE_Cu transfer dynamics *in vitro* and *in planta*

To determine whether Psa_{NZ45}ICE_Cu is active and capable of self-transmission samples from the *Psa* NZ45 and *Psa* NZ13 mixtures from the co-cultivation experiments were also plated on MGY medium containing both kanamycin and copper sulphate. Copper-resistant, kanamycin-resistant transconjugants were detected both *in vitro* and *in planta.* This means that a fraction of *Psa* NZ13 strains acquired Psa_{NZ45}ICE_Cu. These transconjugants marginally inflate the counts of *Psa* NZ45, however, the number of transconjugants (see below) was several orders of magnitude less that *Psa* NZ13, thus having no appreciable effect on the measures of relative fitness*.*

At 24 hours in shaken MGY broth transconjugants were present at a frequency of $5.04 \pm 2.25 \times 10^{-3}$ per recipient cell (mean and SEM from 3 independent experiments, each comprised of 3 replicates). Analysis of samples from *in planta* experiments showed that at 3 days, transconjugants were present at a frequency of $2.05 \pm 0.63 \times 10^{-2}$ per recipient cell (mean and SEM from 3 independent experiments, each comprised of 5 replicates). On plants in the presence of Nordox (0. 375 g/L) the frequency of transconjugants was $9.37 \pm 1.56 \times 10^{-2}$ per recipient three days after inoculation (mean and SEM from 3 independent experiments, each comprised of 5 replicates). Transfer was also observed in M9 agar, on M9 agar supplemented with 0.5mM $CuSO_4$, on M9 agar supplemented with a macerate of Hort16A fruit with transconjugants present (at 48 hours) at a frequency of $2.16 \pm 0.9 \times 10^{-5}$, $1.11 \pm 0.4 \times 10^{-5}$ and $1.98 \pm 0.8 \times 10^{-5}$ per recipient cell, respectively.

To explore the dynamics of transfer *in vitro*, samples from shaken MGY cultures were taken hourly, for six hours, and then at 24 hours. The data, presented in Figure 5.10, show acquisition of Psa_{NZ45}ICE_Cu by *Psa* NZ13 within one hour of the mating mix being established (approximately 1 recipient per 10^5 recipient cells). The frequency was relatively invariant over the subsequent six hour period, but rose to approximately 1 recipient in 10^3 cells at 24 hours.

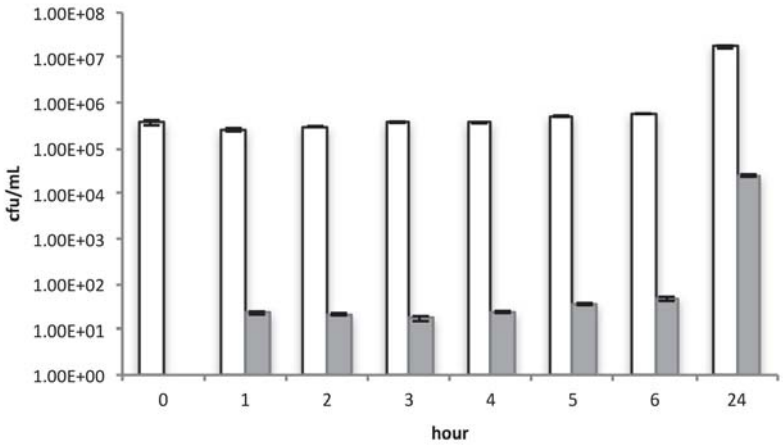

Figure 5.10. *In vitro* transfer of Psa_{NZ45}ICE_Cu from *Pseudomonas syringae* pv. *actinidiae* NZ45 to *P. syringae* pv. *actinidiae* NZ13. Colony forming units of the recipient *Psa* NZ13 (white bars) and *Psa* NZ13 carrying Psa_{NZ45}ICE_Cu (transconjugants, grey bars) was monitored during co-cultivation. Data are means and standard deviation of 3 independent cultures.

Detection of ICE transfer just one hour after mixing donor and recipient cells promoted a further experiment in which transconjugants were assayed at 10 minute intervals. From three independent experiments, each with five replicates, transconjugants were detected at 30 mins (approximately 4×10^{-7} transconjugants per recipient cell).

Analysis of co-cultivation experiments from kiwifruit leaves showed evidence that Psa_{NZ45}ICE_Cu also transferred *in planta.* The frequency of transconjugants at day 3 and day 7 was approximately 1 per 50 recipient cells and the frequency of transconjugants was not affected by changes in the initial founding ratios of donor and recipient cells (Table 5.2). Overall, the frequency of

transconjugants was approximately three orders of magnitude greater *in planta* than *in vitro*.

	Frequency of Psa_{NZ45}ICE_Cu transconjugants	
***Psa* NZ13 : *Psa* NZ45**	**day 3**	**day 7**
1 : 1	$(2.05 \pm 0.63)^{-2}$	$(2.28 \pm 0.7)^{-2}$
1 : 0.1	$(1.34 \pm 0.52)^{-2}$	$(3.14 \pm 2.3)^{-2}$
0.1 : 1	$(1.68 \pm 0.68)^{-2}$	$(1.88 \pm 0.81)^{-2}$

Table 5.2. *In planta* transfer of Psa_{NZ45}ICE_Cu from *Pseudomonas syringae* pv. *actinidiae* NZ45 to *P. syringae* pv. *actinidiae* NZ13 at different founding ratios of donor and recipient. Donor and recipient strains were dip-inoculated onto Hort16A leaves at different founding ratios and frequency of recipients determined at days 3 and 7.

5.2.6. ICE displacement and recombination

To check the genetic composition of transconjugants and to investigate whether Psa_{NZ45}ICE_Cu integration in recipient cells occurred at the *att-1* or *att-2* site, a set of primers were designed to identify the location of ICE integration in the *Psa* NZ13 genome (Table 2.3). 11 independently generated transconjugants from shaken MGY culture were screened. As expected, successful amplification of primers annealing to *copA* in Psa_{NZ45}ICE_Cu was observed in all transconjugants, while amplification of the *enolase* gene primers (indicative of the presence of the native Psa_{NZ13}ICE) occurred only in *Psa* NZ13 and *Psa* NZ45 (Figure 5.11).

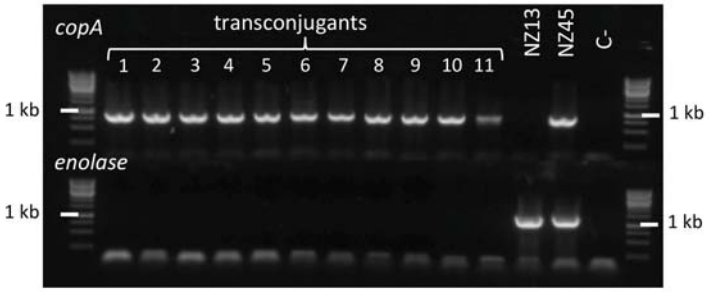

Figure 5.11. Analysis of the presence of the CG4 of Psa_{NZ45}ICE_Cu and Psa_{NZ13}ICE in 11 *Pseudomonas syringae* pv. *actinidiae* NZ13 transconjugants. PCRs were carried out to detect *copA* (CG4 of Psa_{NZ45}ICE_Cu) or *enolase* genes (CG4 of Psa_{NZ13}ICE). Controls of *Psa* NZ13 and *Psa* NZ45 show one and two bands, indicative of Psa_{NZ13}ICE in *Psa* NZ13 and both Psa_{NZ13}ICE Psa_{NZ45}ICE_Cu and in *Psa* NZ45, respectively. The negative control (C-) shows no amplification. All transconjugants, lanes 1-11 have acquired Psa_{NZ45}ICE_Cu.

However, in two transconjugants only the IntPsaNZ13-att-1 primer pair resulted in amplification, suggesting that recombination between Psa$_{NZ45}$ICE_Cu and Psa$_{NZ13}$ICE had occurred (Figure 5.12).

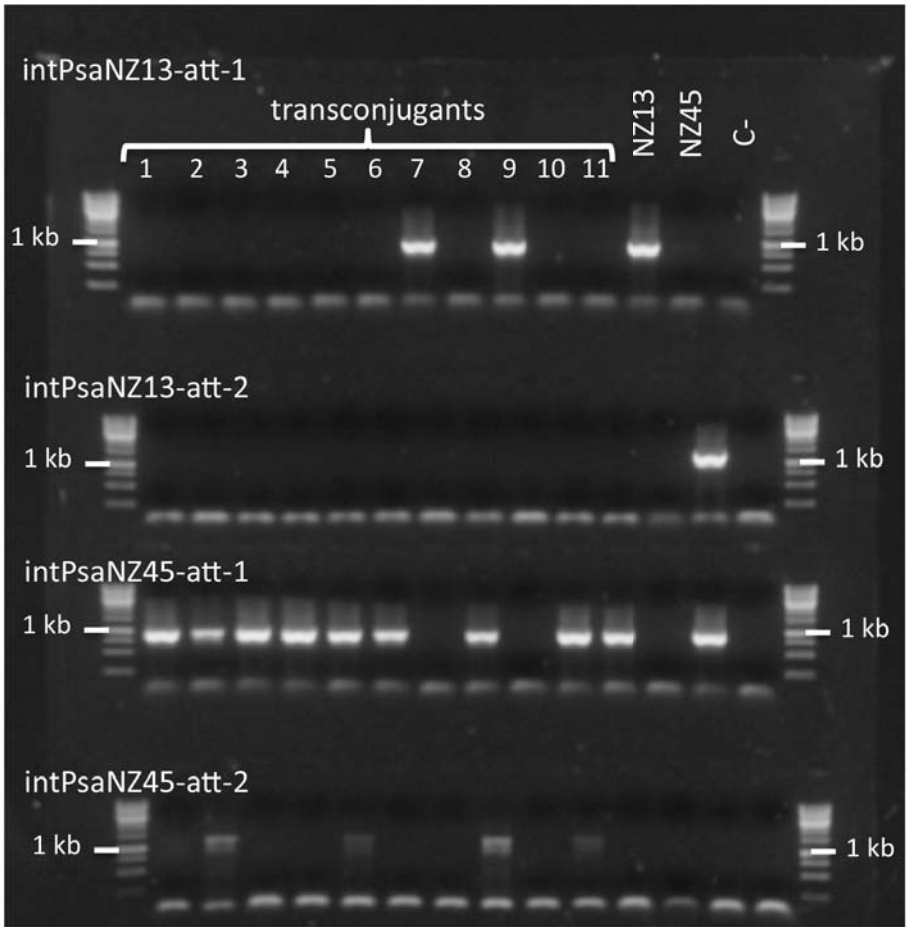

Figure 5.12. Analysis of the insertion site of the Psa$_{NZ13}$ICE and Psa$_{NZ45}$ICE_Cu in 11 *Pseudomonas syringae* pv. *actinidiae* NZ13 transconjugants. PCRs were to detect the integration of Psa$_{NZ13}$ICE in the *att-1* or *att-2* sites (intPsaNZ13-att-1 and intPsaNZ13-att-2) and the integration of Psa$_{NZ45}$ICE_Cu in the *att-1* or *att-2* sites (intPsaNZ45-att-1 and intPsaNZ45-att-2). Controls of *Psa* NZ13 and *Psa* NZ45 show that in *Psa* NZ13 the Psa$_{NZ13}$ICE is integrated in the *att-1* site and in *Psa* NZ45 the Psa$_{NZ13}$ICE is integrated in the *att-2* and the Psa$_{NZ45}$ICE_Cu in the *att-1* site. The negative control (C-) shows no amplification.

Genome sequencing of one of these transconjugants revealed a recombination event inside the variable region of the ICE that produced a chimeric ICE identical to Psa$_{NZ45}$ICE_Cu up to and including the CuR operon, with the remainder identical to the downstream segment of Psa$_{NZ13}$ICE (Figure 5.13).

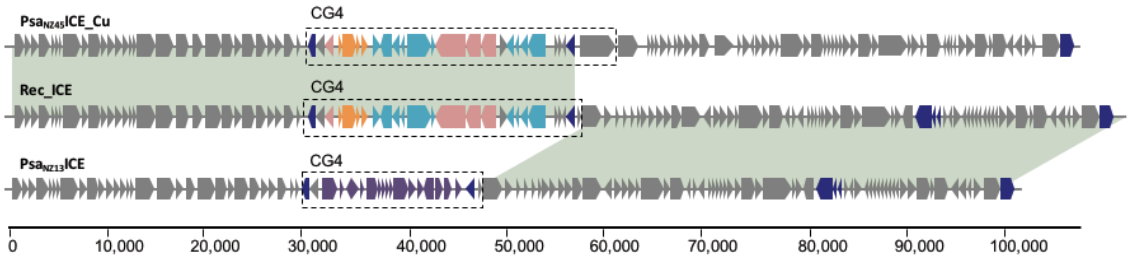

Figure 5.13. Structure and chimerism of the recombinant ICE (Rec_ICE) in transconjugant *Pseudomonas syringae* pv. *actinidiae* NZ13. Areas highlighted in green show 100% pairwise identity. The recombination break point is inside the CG4. Blue boxes are mobile genes (transposases or recombinases), purple boxes define the "enolase region" (McCann *et al*. 2013), azure boxes depict copper resistance genes, orange boxes are arsenic resistance genes and pink boxes are genes belonging to the *czc*/*cus* system. Core backbone and other cargo genes are depicted as grey boxes.

5.2.7 Psa$_{NZ45}$ICE_Cu can be transferred to a range of *P. syringae* strains

The host range of the Psa$_{NZ45}$ICE_Cu was characterised using a panel of nine different *Pseudomonas* strains as recipients, representing the diversity of *P. syringae* and the genus more broadly. Transfer of Psa$_{NZ45}$ICE_Cu to *Psa* J31, *Pfm* NZ9 and *P. syringae* pv. *phaseolicola* (*Pph*) 1448a (on M9 agar plates) was observed with the frequency of transconjugants per recipient cell being 7.64 $\pm$ 1.7 x 10^{-6}, 7.74. $\pm$ 2.5 x 10^{-7} and 1.23 $\pm$ 0.2 x 10^{-4}, respectively. No transconjugants were detected for *P. aeruginosa* PAO1, *P. fluorescens* SBW25, *P. syringae* pv. *tomato* DC3000 or *Psa* K28, despite the fact that these three strains have both *att* sites.

5.3 Discussion

The importance and impact of HGT on the evolution of microbial populations has long been recognized (Sullivan *et al.*, 1995; Lilley & Bailey, 1997; Ochman *et al.*, 2000; Ochman *et al.*, 2005; Wozniak & Waldor, 2010; Polz *et al.*, 2013). Here I have captured the real-time evolution of copper resistance in a plant pathogen, in an agricultural setting, and shown that movement of copper resistance genes occurs via ICEs. Of the seven copper-resistant *Psa* isolates analyzed, five contain copper resistance-encoding ICEs – three unique ICEs in total – with variable placement within the *Psa* genome, including movement and instability of the native ICE (Psa_{NZ13}ICE). Further evidence of dynamism comes from *in vitro* and *in planta* studies, which show not only transfer to isogenic *Psa* and unrelated *P. syringae* strains, but also the ready formation of chimeras between Psa_{NZ45}ICE_Cu and Psa_{NZ13}ICE. Mosaicism of ICEs has been reported elsewhere and is often promoted by the presence of tandem copies (Garriss *et al.*, 2009; Wozniak & Waldor, 2010). The ease with which ICEs move between strains and capacity for intra-ICE recombination emphasizes the futility of drawing conclusions on strain phylogeny based on ICE phylogeny (McCann *et al.*, 2013), but also the impossibility of understanding ICE evolution based on the phylogeny of ICEs themselves.

Evidence of the formation of chimeric ICEs extends beyond the ICEs studied here. Psa_{NZ45}ICE_Cu is a recombinant of two previously reported ICEs and a plasmid: most surprising is the fact that the recombinant components are derived from elements isolated from three geographic regions (USA, Japan and New Zealand) from three different plants (millet, kiwifruit and wheat) and spanning

almost 100 years. Additionally, two of the copper resistance-encoding ICEs found in *Psa* (Psa_{NZ47}ICE_Cu and Psa_{NZ64}ICE_Cu) have been reported in other kiwifruit leaf colonizing organisms emphasizing the ease by which self-transmissible elements can move between members of a single community. Clearly the potency of evolution fuelled by ICEs within the *P. syringae* complex is remarkable, with impacts likely extending well beyond that inferred from the analysis of genome sequences (Fondi *et al.*, 2016).

Evidence of the spectrum and dynamic of transfer inferred from the genomic analysis of natural isolates is bolstered by demonstration of the *in vitro* and *in planta* transfer of Psa_{NZ45}ICE_Cu. The fact that Psa_{NZ45}ICE_Cu can be detected in a recipient strain just 30 minutes after mixing with a donor strain (in shaken broth culture) points to an as yet undetermined proficiency for transfer. At the same time, the frequency of transconjugants *in planta* are several orders of magnitude greater than *in vitro* suggesting even greater potential for transfer in the natural environment.

The selective causes underpinning the evolution of copper resistance in *Psa* is uncertain and to date copper resistance in *Psa* is not known to have evolved outside of NZ. While it is tempting to blame use of copper sprays by orchardists, it is possible that the evolution of copper-resistant *Psa* is a more general response to copper levels in New Zealand soils, copper emission to the atmosphere from anthropomorphic activities (Hong *et al.*, 1996; Staehlin *et al.*, 2016), combined with the use of copper-based sprays in NZ agriculture (Morgan & Taylor, 2004; Dean, 2016). Support for this stems from the fact that Psa_{NZ47}ICE_Cu shows almost perfect identify with an ICE found in *P. marginalis* (ICMP 11289) from kiwifruit

isolated in 1991 (in NZ). In addition, copper resistance-encoding ICEs were found in both copper-treated, and untreated orchards.

In previous studies of copper resistance in *Pseudomonas syringae* beneficial effects due to carriage of resistance genes have been detected during assays of epiphytic growth (Sundin *et al*., 1989; Menkissoglu & Lindow, 1991). In our work, the impact of the copper resistance-encoding ICEs on fitness *in planta* – in the presence of copper sprays – appears to be minimal. The selective advantage conferred by Psa_{NZ45}ICE_Cu was statistically significant only for epiphytic growth at day 3 in presence of Nordox75. The lack of observable benefit at day 7 likely reflects tissue breakdown and escape of endophytic bacteria through stomata, causing inflation of the number of epiphytic bacteria (Figure 5.9).

From the perspective of pathogen control using copper-based sprays there are at least three reasons to treat our detection of a minimal advantage to carriage of genetic determinants of copper resistance with caution. Firstly, for reasons mentioned above and encountered elsewhere (Beattie & Lindow, 1999; Hirano & Upper, 2000) it is difficult to accurately assess fitness of microbes on plants and it is possible that our measures underestimate the contribution of copper resistance to growth in the presence of copper: even a 1% increase in fitness over 24 hours, which is beyond experimental capacity to detect, can have significant long-term consequences. Secondly, the presence of copper resistance genes means opportunity for levels of resistance to increase through, for example, promoter mutations that increase levels of transcription of resistance determinants, or through acquisition of additional copper resistance-encoding genes. Thirdly, and perhaps most significantly, is the fact that the copper resistance-encoding ICEs confer no measurable fitness cost even in the absence of copper. This suggests that

these elements will not be readily lost from *Psa* populations even if copper-based sprays were eliminated (Andersson & Hughes, 2010; Neale *et al.*, 2016). That some strains of the global pandemic now contain two ICEs gives reason to suspect elevated evolutionary potential among these isolates.

While the focus of our investigation has been copper resistance, the ICEs reported here carry a cargo of additional genes, some of which are implicated in resistance to other metals. In some instances the cargo genes have no similarity to genes of known function (some of the grey boxes in Figure 5.3A). ICEs and similar laterally transferred elements provide opportunity for genes unrelated to copper resistance, for example gene connected to virulence, to hitchhike and rapidly spread. In this regard the two plasmids characterized here are of interest: both carry determinants of streptomycin resistance – an antibiotic that is also sprayed on NZ kiwifruit orchards in order to control *Psa* (KVH, *personal communication*). The potential for hitchhiking has been previously noted in the context of antibiotic resistance-encoding plasmids (Gullberg *et al.*, 2014).

Recognition of ICEs along with their potential to change the course of microbial evolution extends less than twenty years (Wozniak & Waldor, 2010). While it might be argued that this potential is no different from that long realized via conjugative plasmids, or phage (Ochman *et al.*, 2000), ICEs, being a composite of both, seem to have an edge. Unlike conjugative plasmids that rarely integrate into the host genome, ICEs integrate as a matter of course and are largely immune to segregational loss; additionally, fitness consequences as a result of carriage are likely to be minimal. Unlike temperate phages, ICEs do not kill the host upon transfer, but they can nonetheless mediate transfer upon encountering transfer proficient conditions. Having control over both vertical and horizontal modes of

transmission, while minimizing costs for host cells, marks these elements as especially potent vehicles of microbial evolution.

MASSEY UNIVERSITY
GRADUATE RESEARCH SCHOOL

STATEMENT OF CONTRIBUTION TO DOCTORAL THESIS CONTAINING PUBLICATIONS

(To appear at the end of each thesis chapter/section/appendix submitted as an article/paper or collected as an appendix at the end of the thesis)

We, the candidate and the candidate's Principal Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidate's contribution as indicated below in the *Statement of Originality*.

Name of Candidate: Elena Colombi

Name/Title of Principal Supervisor: Paul Rainey / Distinguished Professor

Name of Published Research Output and full reference:

Colombi et al (2017). Evolution of copper resistance in the kiwifruit pathogen Pseudomonas syringae pv. actinidiae through acquisition of integrative conjugative elements and plasmids. Environmental Microbiology 19, 819-832.

In which Chapter is the Published Work: Chapter 5

Please indicate either:

- The percentage of the Published Work that was contributed by the candidate:

  and / or

- Describe the contribution that the candidate has made to the Published Work:

  All the experimental work, data analysis, preparation of figures etc was performed by the candidate. Elena also drafted the manuscript and handled the submission and revision process. Elena did the vast majority of the work.

Candidate's Signature

2 June 2017

Date

Principal Supervisor's signature

31 May 2017

Date

Chapter 6 - Concluding remarks

6.1 Background

Horizontal gene transfer (HGT) is a major force driving evolution in prokaryotes (Mirkin *et al.,* 2003; Kunin *et al.,* 2005; Dagan & Martin, 2007; Halary *et al*., 2010; Kloesges *et al*., 2011). Depending on the nature of the transfer, HGT can result in the acquisition of either random segments of DNA or novel clusters, thus allowing a major switch in phenotype. Kloesges *et al.* (2011) have shown that most bacterial protein families have been affected by HGT and that much of it is of recent acquisition. HGT has played a direct role in the evolution of pathogenicity and virulence in several bacterial pathogens (Wren, 2003; Zhou *et al.,* 2004; Barash & Manulis-Sasson, 2009; Alfano *et al.*, 2010; Shapiro, 2016).

Pseudomonas syringae pv. *actinidiae* (*Psa*) first emerged as a kiwifruit pathogen in the 1980's in Asia, with regional outbreaks in both Japan and Korea. Reduced host susceptibility, decreased transmission or suboptimal environmental conditions (or any combination thereof) may have contributed to the limited spread of disease at that time. The current pandemic lineage emerged prior to 2010; this global spread may be attributed to a highly

susceptible host (Hort16A), and insufficient control over the extensive commercial trade of plant material (including plants and pollen). Earlier genomic analysis of pandemic *Psa* revealed that the only significant genomic change among the Psa-3 lineage involved three ICEs, which despite having highly diverse sets of core genes and variable cargo genes, contained a single accessory element (Tn*6212*) identical in the three otherwise divergent ICEs. This element was predicted to encode genes involved in manipulation of host metabolism (McCann *et al.*, 2013). Thus HGT was implicated in the evolution *Psa* virulence on Hort16A.

6.2 Findings

The three ICEs acquired by Psa-3 infecting Chile, Italy and New Zealand belong to a diverse family of ICEs present exclusively in plant-associated *Pseudomonas* spp.. In Chapter 3 I described this novel ICE family, referred to as PsICEs, with members of this family having the same integration loci and backbone genes. A comparison of these PsICEs with a family of ICEs in *P. aeruginosa* (PAPI-1) suggests that these two families share a common ancestor. PsICE evolution is characterized by extensive recombination events: inter-ICE recombination is frequent enough to mask their evolutionary history, producing chimeras with variable patterns of similarity to each other, yet maintaining a syntenic backbone. The diversity of PsICEs is driven by both the sequence divergence and recombination of the backbone, as well as the presence of cargo genes. In the PsICEs the cargo genes are located in conserved intergenic loci across the backbone. Although there are different classes of PsICE cargo genes,

one set was frequently recovered: the set carried on the sub-element Tn*6212*. The recurrence of Tn*6212* in PsICEs isolated exclusively from plant-associated *Pseudomonas* spp. suggested that this region contributes toward the fitness of *Psa* on kiwifruit.

I tested the hypothesis that Tn*6212* contributes to virulence in Chapter 4. As *Psa* was an understudied pathogen, I first set up experimental protocols for mutagenesis and pathogenicity assays. Tn*6212* contributes to enhanced growth on succinate, fumarate and malate. This fitness advantage was not observed during plant colonization, neither on the natural host Hort16A nor the model plants tomato and *N. benthamiana.* No evidence was found to support the hypothesis that DctT is translocated into the plant cell via the TTSS.

The success of an agricultural pathogen is not solely dependent on its capacity to evade host defenses, but also its colonization ability. In Chapter 5, I described my discovery of the evolution of copper resistance in *Psa*. Copper is among the most frequently applied bactericides in agriculture, thus the acquisition of copper resistance in *Psa* is as alarming as the acquisition of antibiotic resistance in human pathogens. The horizontal acquisition of resistance was mediated both by plasmids and (with a stroke of luck for this thesis) by PsICEs harboring different copper resistance determinants. *Psa*$_{NZ45}$ICE_Cu imposed no cost and conferred a selective advantage *in vitro* at both lethal and sub-lethal copper sulphate concentrations. Resistance was also evaluated *in planta* with the most commonly used copper-based spray in New Zealand. The absence of any cost of carriage was confirmed and a selective advantage was identified, though this was reduced relative to *in vitro* experiments. The transfer of the *Psa*$_{NZ45}$ICE_Cu between *Psa* strains was highly

dynamic, with transconjugants detected *in vitro* within 30 minutes of co-inoculation and up to 1/50 recipient cells *in planta* both three and seven days post-inoculation. I also identified the creation of a chimeric ICE formed by the recombination between the *Psa*$_{NZ45}$ICE_Cu and the native *Psa*$_{NZ13}$ICE. Transfer of *Psa*$_{NZ45}$ICE_Cu was also detected to different *P. syringae* pathovars.

6.3 Final Comments

The importance of agents of horizontal transfer for the spread of virulence traits among pathogens is known, but only little has been investigated on ICEs among the *P. syringae* species complex (Baltrus *et al.*, 2017). In 2005, two articles describing Psy$_{B728a}$ICE and Pph$_{13A02}$ICE were published (Feil *et al.,* 2005; Pitman *et al.*, 2005). Feil *et al.* (2005) merely note the presence of Psy$_{B728a}$ICE, while Pph$_{13A02}$ICE has been the subject of extensive study (Pitman *et al.*, 2005; Lovell *et al.*, 2009; Godfrey *et al.*, 2010; Lovell *et al.*, 2011; Godfrey *et al.*, 2011; Neale *et al.*, 2016). These studies were driven by the presence of the effector gene *hopAR1*. Interestingly, of all the PsICEs isolated in this study Pph$_{13A02}$ICE is the only one carrying an effector gene. Three additional ICEs were described during the *Psa* outbreak (Butler *et al.*, 2013; McCann *et al.,* 2013). This work demonstrates that these are not unrelated solo elements but that they are part of a more extensive family of ICEs that spread into the *P. syringae* species a long time ago, possibly distantly related to the PAPI-1 family of ICEs in *P. aeruginosa*. As described for other ICEs, this family is characterized by mosaicisms between plasmids and transposons and a high degree of chimerism between ICEs. Though PsICEs may recombine a lot, those isolated

from unrelated strains in different decades may also be nearly identical. The extent of recombination and the global distribution of the PsICEs may be explained by *P. syringae*'s potential dispersal in the water cycle, the abundance of cells residing in the phyllosphere (10^8 cfu/cm^2), and their transmission efficiency (Monteil *et al*., 2012, 2013; Neale *et al.*, 2016; Colombi *et al*., 2017). However it is not surprising to find the same gene in different locations of the world if they share the same ecological niche (Fondi *et al*., 2016). It remains unclear whether the transfer success/rate is a consequence of some feature of the recipient cell or the PsICE. A full investigation into the regulation and mechanisms of ICE transfer and exclusion is required to fully understand the life of the PsICEs. Mechanisms driving the hotspots and the contribution of cargo genes to the fitness of the host (both inside and outside the plant-pathogen interaction) need to be fully understood.

Tn*6212* was demonstrated to play a role in the utilization of succinate, fumarate and malate, potentially contributing to *P. syringae* fitness *in planta*. I could not demonstrate its involvement in plant colonization, however plant-based assays are not suitable to detect small effects on pathogen fitness, as fitness is a complex phenotype in which multiple traits may contribute. Competitive or fluorescent protein labelled assays *in planta* and the use of older plants might reveal more subtle effects of Tn*6212* in host colonization. As the recruitment of cargo genes and the ease in which ICEs have spread in the microbial world has likely been caused by anthropomorphic activities (Gillings, 2016), the presence of Tn*6212* as a response of the changing environment is not to be excluded, i.e., global warming or pollution.

This study has directly contributed to the New Zealand kiwifruit industry, with concern for the effect of copper resistance in the orchards. I have demonstrated that HGT has "*gone wild*" (Lindow, 2017) in *Psa*, representing a real risk for orchard management. Though the effects of the spread of copper resistance in the orchards have still not been assessed. A recent article suggests that resistance is horizontally transferred when selective pressures are absent (Stevenson *et al.*, 2017), and in an agricultural setting this scenario is common. When copper is sprayed in the orchards it does not reach all the aerial parts of the crop with equal distribution, or inside of the plant, thus niches will be present where the selective pressure for copper resistance is low (or even absent). If "*HGT is most likely to drive the spread of resistance genes in environments where resistance is useless*" (Stevenson *et al.*, 2017), the presence of many CuR *Psa* in copper-free orchards would be explained.

This work contributed by giving an insight into the unstudied world of ICEs in the *P. syringae* species complex, with a particular focus on the evolution of *Psa* through the acquisition of PsICEs.

Reference List

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 Dis, 95, 1583.

Águila-Clares, B., Castiblanco, L. F., Quesada, J. M., Penyalver, R., Carbonell, J., López, M. M., *et al.* (2016). Transcriptional response of *Erwinia amylovora* upon copper shock: in vivo role of the *copA* gene. Mol Plant Pathol.

Aiello, D., Ferrante, P., Vitale, A., Polizzi, G., Scortichini, M., & Cirvilleri, G. (2015). Characterization of *Pseudomonas syringae* pv. *syringae* isolated from mango in Sicily and occurrence of copper-resistant strains. J Plant Pathol, 97(2), 273-282.

Alfano, J. R., Bauer, D. W., Milos, T. M., & Collmer, A. (1996). Analysis of the role of the *Pseudomonas syringae* pv. *syringae* HrpZ harpin in elicitation of the hypersensitive response in tobacco using functionally non-polar *hrpZ* deletion mutations, truncated HrpZ fragments, and *hrmA* mutations. Mol Microbiol, 19(4), 715-728.

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 plants. Proc Natl Acad Sci, 97, 4856–4861.

Alonso, A., Sanchez, P., & Martinez, J. L. (2001). Environmental selection of antibiotic resistance genes. Environ Microbiol, 3(1), 1-9.

Amato, P., Parazols, M., Sancelme, M., Laj, P., Mailhot, G., & Delort, A. M. (2007). Microorganisms isolated from the water phase of tropospheric clouds at the Puy de Dôme: major groups and growth abilities at low temperatures. FEMS Microbiol Ecol, 59(2), 242-254.

Amorós-Moya, D., Bedhomme, S., Hermann, M., & Bravo, I. G. (2010). Evolution in regulatory regions rapidly compensates the cost of nonoptimal codon usage. Mol Biol Evol, 27(9), 2141-2151.

Andersson, D. I., & Hughes, D. (2009). Gene amplification and adaptive evolution in bacteria. Annu Rev Genet, 43, 167-195.

Andersson, D. I., & Hughes, D. (2011). Persistence of antibiotic resistance in bacterial populations. FEMS Microbiol Rev, 35(5), 901-911.

Andersson, D. I., & Levin, B. R. (1999). The biological cost of antibiotic resistance. Curr Opin Microbiol, 2(5), 489-493.

Andersson, D.I., & Hughes, D. (2010). Antibiotic resistance and its cost: is it possible to reverse resistance? Nat Rev Microbiol, 8, 260–271.

Arnold, D. L., Godfrey, S. A., & Jackson, R. W. (2009). *Pseudomonas syringae*

genomics provides important insights to secretion systems, effector genes and the evolution of virulence. In Plant Pathogenic Bacteria: Genomics and Molecular Biology, 202-226. Caister Academic Press Wymondham, UK.

Arrebola, E., Cazorla, F. M., Romero, D., Pérez-García, A., & de Vicente, A. (2007). A nonribosomal peptide synthetase gene (*mgoA*) of *Pseudomonas syringae* pv. *syringae* is involved in mangotoxin biosynthesis and is required for full virulence. Mol Plant Microbe Interact, 20(5), 500-509.

Auchtung, J. M., Lee, C. A., Garrison, K. L., & Grossman, A. D. (2007). Identification and characterization of the immunity repressor (ImmR) that controls the mobile genetic element ICEBs1 of *Bacillus subtilis*. Mol Microbiol, 64(6), 1515-1528.

Auchtung, J. M., Lee, C. A., Monson, R. E., Lehman, A. P., & Grossman, A. D. (2005). Regulation of a Bacillus subtilis mobile genetic element by intercellular signaling and the global DNA damage response. Proc Natl Acad Sci, 102(35), 12554-12559.

Aziz, R. K., Bartels, D., Best, A. A., DeJongh, M., Disz, T., Edwards, R. A., *et al*. (2008). The RAST Server: rapid annotations using subsystems technology. BMC Genomics, 9, 75-10.1186/1471-2164-9-75.

Bailey, J., & Manoil, C. (2002). Genome-wide internal tagging of bacterial exported proteins. Nat Biotechnol, 20,839–842.

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 Dis Rep, 22, 10.

Baltrus, D. A., McCann, H. C., & Guttman, D. S. (2017). Evolution, genomics and epidemiology of *Pseudomonas syringae*: challenges in bacterial molecular plant pathology. Mol Plant Pathol, 18(1), 152-168.

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

Barash, I., & Manulis-Sasson, S. (2009). Recent evolution of bacterial pathogens: the gall-forming *Pantoea agglomerans* case. Annu Rev Phytopathol, 47, 133-152.

Bardaji, L., Añorga, M., Ruiz-Masó, J. A., Del Solar, G., & Murillo, J. (2017). Plasmid replicons from *Pseudomonas* are natural chimeras of functional, exchangeable modules. Front Microbiol, 8.

Bastas, K. K., & Karakaya, A. (2017). First report of bacterial canker of kiwifruit caused by *Pseudomonas syringae* pv. *actinidiae* in Turkey. Phytopathology, 107(2), 184-191.

Beaber, J. W., Hochhut, B., & Waldor, M. K. (2004). SOS response promotes horizontal dissemination of antibiotic resistance genes. Nature, 427(6969), 72-74.

Beaber, J.W., Hochhut, B., & Waldor, M.K. (2002) Genomic and functional analyses of SXT, an integrating antibiotic resistance gene transfer element derived from *Vibrio cholerae*. J Bacteriol, 184, 4259–4269.

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

Bell, K. S., Sebaihia, M., Pritchard, L., Holden, M. T. G., Hyman, L. J., Holeva, M. C., *et al.* (2004). Genome sequence of the enterobacterial phytopathogen

Erwinia carotovora subsp. *atroseptica* and characterization of virulence factors. Proc Natl Acad Sci, 101(30), 11105-11110.

Bellanger, X., Morel, C., Decaris, B., & Guédon, G. (2007). Derepression of excision of integrative and potentially conjugative elements from *Streptococcus thermophilus* by DNA damage response: implication of a cI-related repressor. J Bacteriol, 189(4), 1478-1481.

Bender, C. L., Alarcón-Chaidez, F., & Gross, D. C. (1999). *Pseudomonas syringae* phytotoxins: mode of action, regulation, and biosynthesis by peptide and polyketide synthetases. Microbiol Mol Biol Rev, 63(2), 266-292.

Bender, C. L., Stone, H. E., Sims, J. J., & Cooksey, D. A. (1987). Reduced pathogen fitness of *Pseudomonas syringae* pv. *tomato* Tn*5* mutants defective in coronatine production. Physiol Mol Plant Pathol, 30(2), 273-283.

Bender, C.L., & Cooksey, D.A. (1986) Indigenous plasmids in *Pseudomonas syringae* pv. *tomato*: conjugative transfer and role in copper resistance. J. Bacteriol, 165, 534-541.

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, e105547.

Bertani, G. (1951). STUDIES ON LYSOGENESIS I.: The Mode of Phage Liberation by Lysogenic *Escherichia coli*. J Bacteriol, 62(3), 293-300.

Bertels, F., Silander, O. K., Pachkov, M., Rainey, P. B., & van Nimwegen, E. (2014). Automated reconstruction of whole-genome phylogenies from short-sequence reads. Mol Biol Evol, 31(5), 1077-1088.

Björkman, J., & Andersson, D. I. (2000). The cost of antibiotic resistance from a bacterial perspective. Drug Resistance Updates, 3(4), 237-245.

Blin, K., Medema, M. H., Kazempour, D., Fischbach, M. A., Breitling, R., Takano, E., & Weber, T. (2013). antiSMASH 2.0—a versatile platform for genome mining of secondary metabolite producers. Nucleic Acids Res, W204-W212.

Boller, T., & Felix, G. (2009). A renaissance of elicitors: perception of microbe-associated molecular patterns and danger signals by pattern-recognition receptors. Annu Rev Plant Biol, 60, 379-406.

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

Bosso, L., Russo, D., Di Febbraro, M., Cristinzio, G., & Zoina, A. (2016). Potential distribution of *Xylella fastidiosa* in Italy: a maximum entropy model. Phytopathol Mediterr, 55(1), 62-72.

Brauner, A., Fridman, O., Gefen, O., & Balaban, N. Q. (2016). Distinguishing between resistance, tolerance and persistence to antibiotic treatment. Nat Rev Microbiol, 14(5), 320-330.

Brent, K. J., & Hollomon, D. W. (2007). Fungicide resistance in crop protection, how can it be managed. FRAC Monograph, 1, 56.

Brown, N. L., Stoyanov, J. V., Kidd, S. P., & Hobman, J. L. (2003). The MerR family of transcriptional regulators. FEMS Microbiol Rev, 27(2-3), 145-163.

Bruen, T. C., Philippe, H., & Bryant, D. (2006). A simple and robust statistical test for detecting the presence of recombination. Genetics, 172(4), 2665-2681.

Buell, C. R., Joardar, V., Lindeberg, M., Selengut, J., Paulsen, I. T., Gwinn, M. L., *et al*. (2003) The complete genome sequence of the *Arabidopsis* and tomato

pathogen *Pseudomonas syringae* pv. *tomato* DC3000. Proc Nat Acad Sci, 100, 10181–10186.

Burrus, V., & Waldor, M. K. (2004). Formation of SXT tandem arrays and SXT-R391 hybrids. J Bacteriol, 186(9), 2636-2645.

Burrus, V., Marrero, J., & Waldor, M.K. (2006). The current ICE age: biology and evolution of SXT-related integrating conjugative elements. Plasmid, 55, 173–183.

Burrus, V., Pavlovic, G., Decaris, B., & Guédon, G. (2002). Conjugative transposons: the tip of the iceberg. Mol Microbiol, 46(3), 601-610.

Butler, M. I., Stockwell, P. A., Black, M. A., Day, R. C., Lamont, I. L., & Poulter, R. T. M. (2013). *Pseudomonas syringae* pv. *actinidiae* from recent outbreaks of kiwifruit bacterial canker belong to different clones that originated in China. PLoS ONE, 8, e57464.

Campbell, A., Berg, D., Botstein, D., Lederberg, E. M., Novick, R. P., Starlinger, P., & Szybalski, W. (1979). Nomenclature of transposable elements in prokaryotes. Gene, 5(3), 197-206.

Caro-Quintero, A., & Konstantinidis, K. T. (2015). Inter-phylum HGT has shaped the metabolism of many mesophilic and anaerobic bacteria. ISME J, 9(4), 958-967.

Carraro, N., Durand, R., Rivard, N., Anquetil, C., Barrette, C., Humbert, M., & Burrus, V. (2017). *Salmonella* genomic island 1 (SGI1) reshapes the mating apparatus of IncC conjugative plasmids to promote self-propagation. PLoS Genet, 13(3), e1006705.

Carraro, N., Libante, V., Morel, C., Decaris, B., Charron-Bourgoin, F., Leblond, P., & Guédon, G. (2011). Differential regulation of two closely related integrative and conjugative elements from *Streptococcus thermophilus.* BMC Microbiol, 11(1), 238.

Carrión, V. J., Arrebola, E., Cazorla, F. M., Murillo, J., & de Vicente, A. (2012). The *mbo* operon is specific and essential for biosynthesis of mangotoxin in *Pseudomonas syringae*. PLoS One, 7(5), e36709.

Casadevall, A., & Pirofski, L. A. (1999). Host-pathogen interactions: redefining the basic concepts of virulence and pathogenicity. Infect Immun, 67(8), 3703-3713.

Casadevall, A., & Pirofski, L. A. (2000). Host-pathogen interactions: basic concepts of microbial commensalism, colonization, infection, and disease. Infect Immun, 68(12), 6511-6518.

Casadevall, A., & Pirofski, L. A. (2001). Host-pathogen interactions: the attributes of virulence. J Infec Dis, 184(3), 337-344.

Casadevall, A., & Pirofski, L. A. (2009). Virulence factors and their mechanisms of action: the view from a damage–response framework. J Water Health, 7(S1), S2-S18.

Cazorla, F. M., Arrebola, E., Sesma, A., Pérez-García, A., Codina, J. C., Murillo, J., & de Vicente, A. (2002). Copper resistance in *Pseudomonas syringae* strains isolated from mango is encoded mainly by plasmids. Phytopathology, 92(8), 909-916.

Ceremuga, I., Seweryn, E., Bednarz-Misa, I., Pietkiewicz, J., Jermakow, K., Banaś, T., & Gamian, A. (2014). Enolase-like protein present on the outer membrane of Pseudomonas aeruginosa binds plasminogen. Folia Microbiol, 59(5), 391-397.

Cesbron, S., Briand, M., Essakhi, S., Gironde, S., Boureau, T., Manceau, C., *et al*. (2015). Comparative genomics of pathogenic and nonpathogenic strains of *Xanthomonas arboricola* unveil molecular and evolutionary events linked to pathoadaptation. Front Plant Sci, 6.

Cha, J. S., & Cooksey, D. A. (1993). Copper hypersensitivity and uptake in *Pseudomonas syringae* containing cloned components of the copper resistance operon. Appl Environ Microbiol, 59, 1671–1674.

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.

Cohan, F.M. (2002). What are bacterial species?. Annu Rev Microbiol, 56, 457-487.

Coll, N. S., Epple, P., & Dangl, J. L. (2011). Programmed cell death in the plant immune system. Cell Death Differ, 18(8), 1247-1256.

Colombi, E., Straub, C., Kunzel, 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. Environ Microbiol, 19, 819–832.

Cooksey, D. A. (1993). Copper uptake and resistance in bacteria. Mol Microbiol, 7(1), 1-5.

Cornelis, P., & Matthijs, S. (2002). Diversity of siderophore-mediated iron uptake systems in fluorescent pseudomonads: not only pyoverdines. Environ Microbiol, 4(12), 787-798.

Cui, H., Xiang, T., & Zhou, J. M. (2009). Plant immunity: a lesson from pathogenic bacterial effector proteins. Cellular microbiol, 11(10), 1453-1461.

Cunty, A., Poliakoff, F., Rivoal, C., Cesbron, S., Saux, F. L., Lemaire, C., *et al*. (2015). Characterization of *Pseudomonas syringae* pv. *actinidiae* (*Psa*) isolated from France and assignment of Psa biovar 4 to a de novo pathovar: *Pseudomonas syringae* pv. *actinidifoliorum* pv. nov. Plant Pathol, 64(3), 582-596.

Dagan T., & Martin W. (2007). Ancestral genome sizes specify the minimum rate of lateral gene transfer during prokaryote evolution. Proc Natl Acad Sci, 104(3), 870-875.

Dangl, J. L., & Jones, J. D. (2001). Plant pathogens and integrated defence responses to infection. Nature, 411(6839), 826-833.

Darling, A. E., Mau, B., & Perna, N. T. (2010). progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. PloS one, 5(6), e11147.

Davies, J., & Davies, D. (2010). Origins and evolution of antibiotic resistance. Microbiol Mol Biol Rev, 74(3), 417-433.

Dean, F.P. (2016). Effects of copper sprays on microbial communities in kiwifruit orchard soils. (Doctoral dissertation, University of Waikato).

Delavat, F., Miyazaki, R., Carraro, N., Pradervand, N., & van der Meer, J. R. (2017). The hidden life of integrative and conjugative elements. FEMS Microbiol Rev.

Deng, W. L., & Huang, H. C. (1999). Cellular locations of *Pseudomonas syringae* pv. *syringae* HrcC and HrcJ proteins, required for harpin secretion via the type III pathway. J Bacteriol, 181(7), 2298-2301.

Deng, W. L., Preston, G., Collmer, A., Chang, C. J., & Huang, H. C. (1998).

Characterization of the hrpC and hrpRS operons of *Pseudomonas syringae* pathovars syringae, tomato, and glycinea and analysis of the ability of *hrpF*, *hrpG*, *hrcC*, *hrpT*, and *hrpV* mutants to elicit the hypersensitive response and disease in plants. J Bacteriol, 180(17), 4523-4531.

Diaz-Ricci, J. C., & Hernández, M. E. (2000). Plasmid effects on Escherichia coli metabolism. Crit Rev Biotechnol, 20(2), 79-108.

Didelot, X., & Wilson, D. J. (2015). ClonalFrameML: efficient inference of recombination in whole bacterial genomes. PLoS Comput Biol, 11(2), e1004041.

Didelot, X., Méric, G., Falush, D., & Darling, A. E. (2012). Impact of homologous and non-homologous recombination in the genomic evolution of Escherichia coli. BMC genomics, 13(1), 256.

Dimopoulou, I.D., Kartali, S.I., Harding, R.M., Peto, T.E.A., & Crook, D.W. (2007) Diversity of antibiotic resistance integrative and conjugative elements among haemophili. J Med Microbiol, 56, 838–846.

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*. Proc Natl Acad Sci USA, 77, 7347–7351.

Dodds, P. N., & Rathjen, J. P. (2010). Plant immunity: towards an integrated view of plant–pathogen interactions. Nature Rev Genet, 11(8), 539-548.

Domazet-Lošo, M., & Haubold, B. (2011). Alignment-free detection of local similarity among viral and bacterial genomes. Bioinformatics, 27(11), 1466-1472.

Doolitle, W.F. (2012). Population genomics: how bacteria species form and why they don't exist. Curr Biol, 5,22(11),R451-3

Doucet-Populaire, F., Trieu-Cuot, P., Dosbaa, I., Andremont, A., & Courvalin, P. (1991). Inducible transfer of conjugative transposon Tn*1545* from *Enterococcus faecalis* to *Listeria monocytogenes* in the digestive tracts of gnotobiotic mice. Antimicrob Agents Chemother, 35(1), 185-187.

Doungudomdacha, S., Volgina, A., & Di Rienzo, J. M. (2007). Evidence that the cytolethal distending toxin locus was once part of a genomic island in the periodontal pathogen *Aggregatibacter* (*Actinobacillus*) *actinomycetemcomitans* strain Y4. J Med Microbiol, 56(11), 1519-1527.

Dower, W. J., Miller, J. F., & Ragsdale, C. W. (1988). High efficiency transformation of *E. coli* by high voltage electroporation. Nucleic Acids Research, 16(13), 6127-6145.

Drenkard, E., & Ausubel, F.M. (2002) *Pseudomonas* biofilm formation and antibiotic resistance are linked to phenotypic variation. Nature, 416, 740–743.

Duchêne, S., Holt, K. E., Weill, F. X., Le Hello, S., Hawkey, J., Edwards, D. J., *et al*. (2016). Genome-scale rates of evolutionary change in bacteria. Microb Genom, 2(11).

Dwiartama, A. (2017). Resilience and transformation of the New Zealand kiwifruit industry in the face of *Psa*-V disease. Journal of Rural Studies.

European and Mediterranean Plant Protection Organization. 2011. EPPO Reporting Service - Pests & Diseases, 3,4

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 Dis Note, 6, 67–71.

Fang, Y., Xiaoxiang, Z., Tao, W. Y. (1990). Preliminary studies on kiwifruit disease in Hunan province. Sichuan Fruit Sci Technol, 18,28–29.

Fang, Y., Wu, L., Chen, G., Feng, G. (2016). Complete genome sequence of *Pseudomonas azotoformans* S4, a potential biocontrol bacterium. J Biotechnol, 227, 25-26.

Farris, J. S., Källersjö, M., Kluge, A. G., & Bult, C. (1994). Testing significance of incongruence. Cladistics, 10(3), 315-319.

Fatmi, M. B., Collmer, A., Sante Iacobellis, N., Mansfield, J. W., Murillo, J., Schaad, N. W., & Ullrich, M. (2008). *Pseudomonas syringae* pathovars and related pathogens—identification, epidemiology, and genomics (Vol. 433). Dordrecht, Springer.

Feil, H., Feil, W.S., Chain, P., Larimer, F., Di Bartolo, G., Copeland, A., *et al.* (2005) Comparison of the complete genome sequences of *Pseudomonas syringae* pv. *syringae* B728a and pv. *tomato* DC3000. Proc Natl Acad Sci USA, 102, 11064-11069.

Feltman, H., Schulert, G., Khan, S., Jain, M., Peterson, L., & Hauser, A. R. (2001). Prevalence of type III secretion genes in clinical and environmental isolates of *Pseudomonas aeruginosa*. Microbiology, 147(10), 2659-2669.

Ferreira, R. M., de Oliveira, A. C. P., Moreira, L. M., Belasque, J., Gourbeyre, E., Siguier, P., *et al*. (2015). A TALE of transposition: Tn*3*-like transposons play a major role in the spread of pathogenicity determinants of *Xanthomonas citri* and other xanthomonads. MBio, 6(1), e02505-14.

Fondi, M., Karkman, A., Tamminen, M. V., Bosi, E., Virta, M., Fani, R., *et al.* (2016). “Every gene is everywhere but the environment selects”: global geolocalization of gene sharing in environmental samples through network analysis. Genome Biol Evol, 8, 1388–1400.

Francl, L. J. (2001). The disease triangle: a plant pathological paradigm revisited. Plant Health Instructor DOI, 10.

Franke, A. E., & Clewell, D. B. (1981). Evidence for a chromosome-borne resistance transposon (Tn916) in *Streptococcus faecalis* that is capable of "conjugal" transfer in the absence of a conjugative plasmid. J Bacteriol, 145(1), 494-502.

Franke, S., Grass, G., & Nies, D. H. (2001). The product of the *ybdE* gene of the *Escherichia coli* chromosome is involved in detoxification of silver ions. Microbiology, 147(4), 965-972.

Franke, S., Grass, G., Rensing, C., & Nies, D. H. (2003). Molecular analysis of the copper-transporting efflux system CusCFBA of *Escherichia coli.* J Bacteriol, 185(13), 3804-3812.

Fujikawa, T., & Sawada, H. (2016). Genome analysis of the kiwifruit canker pathogen *Pseudomonas syringae* pv. *actinidiae* biovar 5. Sci Rep, 6,21399–11.

Furi, L., Haigh, R., Al Jabri, Z. J., Morrissey, I., Ou, H. Y., León-Sampedro, R., *et al*. (2016). Dissemination of novel antimicrobial resistance mechanisms through the insertion sequence mediated spread of metabolic genes. Front Microbiol, 7.

Gaillard, M., Pernet, N., Vogne, C., Hagenbüchle, O., & van der Meer, J. R. (2008). Host and invader impact of transfer of the clc genomic island into *Pseudomonas aeruginosa* PAO1. Proc Natl Acad Sci, 105(19), 7058-7063.

Gaillard, M., Vallaeys, T., Vorhölter, F. J., Minoia, M., Werlen, C., Sentchilo, V., *et al*. (2006). The clc element of *Pseudomonas* sp. strain B13, a genomic island with various catabolic properties. J Bacteriol, 188(5), 1999-2013.

Garriss, G., Waldor, M. K., & Burrus, V. (2009). Mobile antibiotic resistance encoding elements promote their own diversity. PLoS Genetic, 5, e1000775.

Giddens, S. R., Jackson, R. W., Moon, C. D., Jacobs, M. A., Zhang, X. X., Gehrig, S. M., & Rainey, P. B. (2007). Mutational activation of niche-specific genes provides insight into regulatory networks and bacterial function in a complex environment. Proc Natl Acad Sci, 104(46), 18247-18252.

Gillings, M. R. (2016). Lateral gene transfer, bacterial genome evolution, and the Anthropocene. Ann N Y Acad Sci.

Godfrey, S. A. C., Mansfield, J. W., Corry, D. S., Lovell, H. C., Jackson, R. W., & Arnold, D. L. (2010). Confocal imaging of *Pseudomonas syringae* pv. *phaseolicola* colony development in bean reveals reduced multiplication of strains containing the genomic island PPHGI-1. Mol Plant Microbe Interact, 23(10), 1294-1302.

Godfrey, S. A., Lovell, H. C., Mansfield, J. W., Corry, D. S., Jackson, R. W., & Arnold, D. L. (2011). The stealth episome: suppression of gene expression on the excised genomic island PPHGI-1 from *Pseudomonas syringae* pv. *phaseolicola*. PLoS Pathog, 7(3), e1002010.

Goldenfeld N., & Woese C. (2007). Biology's next revolution. Nature, 445(7126), 369-369.

Graham, J. H., Gottwald, T. R., Cubero, J., & Achor, D. S. (2004). *Xanthomonas axonopodis* pv. *citri*: factors affecting successful eradication of citrus canker. Mol Plant Pathol, 5(1), 1-15.

Grant, S. R., Fisher, E. J., Chang, J. H., Mole, B. M., & Dangl, J. L. (2006). Subterfuge and manipulation: type III effector proteins of phytopathogenic bacteria. Annu Rev Microbiol, 60, 425-449.

Grass, G., & Rensing, C. (2001). CueO is a multi-copper oxidase that confers copper tolerance in *Escherichia coli*. Biochem Biophys Res Commun, 286(5), 902-908.

Greenberg, J. T., & Yao, N. (2004). The role and regulation of programmed cell death in plant–pathogen interactions. Cell Microbiol, 6(3), 201-211.

Greer, G., & Saunders, C. (2012). The costs of *Psa*-V to the New Zealand kiwifruit industry and the wider community.

Groll, M., Schellenberg, B., Bachmann, A. S., Archer, C. R., Huber, R., Powell, T. K., *et al*. (2008). A plant pathogen virulence factor inhibits the eukaryotic proteasome by a novel mechanism. Nature, 452(7188), 755-758.

Guglielmini, J., Quintais, L., Garcillán-Barcia, M. P., de la Cruz, F., Rocha, E.P. (2011). The repertoire of ICE in prokaryotes underscores the unity, diversity, and ubiquity of conjugation. PLoS Genet, 7: e1002222.

Gullberg, E., Albrecht, L. M., Karlsson, C., Sandegren, L., & Andersson, D. I. (2014). Selection of a multidrug resistance plasmid by sublethal levels of antibiotics and heavy metals. MBio, 5: e01918–14.

Guo, W., Cai, L. L., Zou, H. S., Ma, W. X., Liu, X. L., *et al*. (2012). Ketoglutarate transport protein KgtP is secreted through the type III secretion system and contributes to virulence in *Xanthomonas oryzae* pv. *oryzae*. Appl Environ Microbiol, 78, 5672–5681.

Gutiérrez-Barranquero, J. A., de Vicente, A., Carrión, V. J., Sundin, G. W., &

Cazorla, F. M. (2013). Recruitment and rearrangement of three different genetic determinants into a conjugative plasmid increase copper resistance in *Pseudomonas syringae*. Appl Environ Microbiol, 79, 1028–1033.

Hajri, A., Meyer, D., Delort, F., Guillaumès, J., Brin, C., & Manceau, C. (2010). Identification of a genetic lineage within *Xanthomonas arboricola* pv. *juglandis* as the causal agent of vertical oozing canker of Persian (English) walnut in France. Plant Pathol, 59(6), 1014-1022.

Halary, S., Leigh, J. W., Cheaib, B., Lopez, P., & Bapteste, E. (2010). Network analyses structure genetic diversity in independent genetic worlds. Proc Natl Acad Sci, 107(1), 127-132.

Hao, W., & Golding, G. B. (2010). Inferring bacterial genome flux while considering truncated genes. Genetics, 186, 411–426.

Harold, F. M. (1966). Inorganic polyphosphates in biology: structure, metabolism, and function. Bacteriol Rev, 30(4), 772.

Harrison, E., & Brockhurst, M. A. (2012). Plasmid-mediated horizontal gene transfer is a coevolutionary process. Trends Microbiol, 20(6), 262-267.

Harwood, C. R., & Cutting, S. M. (1990). Chemically defined growth media and supplements, p. 548. In C. R. Harwood and S. M. Cutting (ed.), Molecular biological methods for *Bacillus*. Wiley, Chichester, United Kingdom.

Haskett, T. L., Terpolilli, J. J., Bekuma, A., O'Hara, G. W., Sullivan, J. T., Wang, P., *et al*. (2016). Assembly and transfer of tripartite integrative and conjugative genetic elements. Proc Natl Acad Sc, 201613358.

Hassen, A., Saidi, N., Cherif, M., & Boudabous, A. (1998). Resistance of environmental bacteria to heavy metals. Bioresour Technol, 64(1), 7-15.

Hauck, P., Thilmony, R., & He, S. Y. (2003). A *Pseudomonas syringae* type III effector suppresses cell wall-based extracellular defense in susceptible Arabidopsis plants. Proc Natl Acad Sci, 100(14), 8577-8582.

Heather, Z., Holden, M. T., Steward, K. F., Parkhill, J., Song, L., Challis, G. L., *et al*. (2008). A novel streptococcal integrative conjugative element involved in iron acquisition. Mol Microbiol, 70(5), 1274-1292.

Heeb, S., Itoh, Y., Nishijyo, T., Schnider, U., Keel, C., Wade, J., *et al*. (2000). Small, stable shuttle vectors based on the minimal pVS1 replicon for use in gram-negative, plant-associated bacteria. Mol Plant Microbe Interact, 13(2), 232-237.

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

Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K., & Pease, L. R. (1989). Site-directed mutagenesis by overlap extension using the polymerase chain reaction. Gene, 77(1), 51-59.

Hobman, J. L., & Crossman, L. C. (2015). Bacterial antimicrobial metal ion resistance. J Med Microbiol, 64(5), 471-497.

Hochhut, B., & Waldor, M. K. (1999). Site-specific integration of the conjugal *Vibrio cholerae* SXT element into prfC. Mol Microbiol, 32(1), 99-110.

Hochhut, B., Beaber, J. W., Woodgate, R., & Waldor, M. K. (2001). Formation of chromosomal tandem arrays of the SXT element and R391, two conjugative chromosomally integrating elements that share an attachment site. J Bacteriol, 183(4), 1124-1132.

Holloway, B.W. (1955). Genetic recombination in *Pseudomonas aeruginosa*. J

Gen Microbiol, 13, 572–581.

Holmes, A. J., Gillings, M. R., Nield, B. S., Mabbutt, B. C., Nevalainen, K. M., & Stokes, H. W. (2003). The gene cassette metagenome is a basic resource for bacterial genome evolution. Environ microbiol, 5(5), 383-394.

Hong, S., Candelone, J. P., Soutif, M., & Boutron, C. F. (1996). A reconstruction of changes in copper production and copper emissions to the atmosphere during the past 7000 years. Sci Total Environ, 188, 183-193.

Hong, T. P., Carter, M. Q., Struffi, P., Casonato, S., Hao, Y., Lam, J. S., *et al*. (2017). Conjugative type IVb pilus recognizes lipopolysaccharide of recipient cells to initiate PAPI-1 pathogenicity island transfer in *Pseudomonas aeruginosa*. BMC Microbiol, 17(1), 31.

Hwang, M. S., Morgan, R. L., Sarkar, S. F., Wang, P. W., & Guttman, D. S. (2005). Phylogenetic characterization of virulence and resistance phenotypes of *Pseudomonas syringae*. Appl Environ Microbiol, 71(9), 5182-5191.

Iyer, R., Iverson, T. M., Accardi, A., & Miller, C. (2002). A biological role for prokaryotic ClC chloride channels. Nature, 419(6908), 715-718.

Johnson, C. M., & Grossman, A. D. (2015). Integrative and conjugative elements (ICEs): what they do and how they work. Annu Rev Genet, 49, 577–601.

Jones, J. D., & Dangl, J. L. (2006). The plant immune system. Nature, 444(7117), 323-329.

Juhas, M., Power, P. M., Harding, R. M., Ferguson, D. J., Dimopoulou, I. D., Elamin, A.R., *et al*. (2007). Sequence and functional analyses of *Haemophilus* spp. genomic islands. Genome Biol, 8, R237.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., *et al*. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics, 28, 1647-1649.

Khandan, H. N., Worner, S. P., Jones, E. E., Villjanen-Rollinson, S. L. H., Gallipoli, L., Mazzaglia, A., & Balestra, G. M. (2013). Predicting the potential global distribution of *Pseudomonas syringae* pv. *actinidiae* (*Psa*). New Zealand Plant Protection conference, 12-15.

Kidd, T. J., Ritchie, S. R., Ramsay, K. A., Grimwood, K., Bell, S. C., & Rainey, P. B. (2012). *Pseudomonas aeruginosa* exhibits frequent recombination, but only a limited association between genotype and ecological setting. PLoS One, 7(9), e44199.

Kim, H. M., Xu, Y., Lee, M., Piao, S., Sim, S. H., Ha, N. C., & Lee, K. (2010). Functional relationships between the AcrA hairpin tip region and the TolC aperture tip region for the formation of the bacterial tripartite efflux pump AcrAB-TolC. J Bacteriol, 192(17), 4498-4503.

King, E. O., Ward, M. K., & Raney D. E. (1954). Two simple media for the determination of pyocianine and fluorescein. J Lab Clin Med, 44,301–307.

Klockgether, J., Würdemann, D., Reva, O., Wiehlmann, L., & Tümmler, B. (2007). Diversity of the abundant pKLC102/PAGI-2 family of genomic islands in *Pseudomonas aeruginosa*. J Bacteriol, 189(6), 2443-2459.

Kloesges, T., Popa, O., Martin, W., & Dagan, T. (2011). Networks of gene sharing among 329 proteobacterial genomes reveal differences in lateral gene transfer frequency at different phylogenetic depths. Mol Biol Evol, 28(2), 1057-1074.

Koh YJ, Cha JB, Chung JH, Lee HD. (1994). Outbreak and spread of bacterial

canker in kiwifruit. Korean J Plant Pathol, 10, 68–72.
Kudla, G., Murray, A. W., Tollervey, D., & Plotkin, J. B. (2009). Coding-sequence determinants of gene expression in *Escherichia coli.* Science, 324(5924), 255-258.
Kung, V. L., Ozer, E. A., & Hauser, A. R. (2010). The accessory genome of *Pseudomonas aeruginosa.* Microbiol Mol Biol Rev, 74, 621–641.
Kunin, V., Goldovsky, L., Darzentas, N., & Ouzounis, C. A. (2005). The net of life: reconstructing the microbial phylogenetic network. Genome Res, 15(7), 954-959.
Kunkel, B. N., Bent, A. F., Dahlbeck, D., Innes, R. W., & Staskawicz, B. J. (1993). RPS2, an *Arabidopsis* disease resistance locus specifying recognition of *Pseudomonas syringae* strains expressing the avirulence gene *avrRpt2*. Plant Cell, 5(8), 865-875.
Kuo, C. H., & Ochman, H. (2009). Deletional bias across the three domains of life. Genome Biol Evol, 1, 145-152.
Kus, J. V., Zaton, K., Sarkar, R., & Cameron, R. K. (2002). Age-related resistance in *Arabidopsis* is a developmentally regulated defense response to Pseudomonas syringae. Plant Cell, 14(2), 479-490.
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. J Gen Plant Pathol, 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. Adv Agron, 126, 235-295.
Lan, R., & Reeves, P. R. (1996). Gene transfer is a major factor in bacterial evolution. Mol Biol Evol, 13, 47-55.
Lawrence, J. G. (2002). Gene transfer in bacteria: speciation without species?. Theor Popul Biol, 61(4), 449-460.
Lawrence, J. G., & Roth, J. R. (1996). Selfish operons: horizontal transfer may drive the evolution of gene clusters. Genetics, 143(4), 1843-1860.
Lazo, G. R., Stein, P. A., & Ludwig, R. A. (1991). A DNA transformation–competent *Arabidopsis* genomic library in *Agrobacterium*. Nat Biotechnol, 9(10), 963-967.
Leclercq, S., Gilbert, C., & Cordaux, R. (2012). Cargo capacity of phages and plasmids and other factors influencing horizontal transfers of prokaryote transposable elements. Mol Genet Elements, 2(2), 115-118.
Lee, S. M., Grass, G., Rensing, C., Barrett, S. R., Yates, C. J., Stoyanov, J. V., & Brown, N. L. (2002). The Pco proteins are involved in periplasmic copper handling in *Escherichia coli*. Biochem Bioph Res Co, 295(3), 616-620.
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. Am Nat, 138, 1315−1341.
Letelier, M. E., Lepe, A. M., Faúndez, M., Salazar, J., Marín, R., Aracena, P., & Speisky, H. (2005). Possible mechanisms underlying copper-induced damage in biological membranes leading to cellular toxicity. Chem Biol Interact, 151(2), 71-82.
Levin, B. R. (2001). Minimizing potential resistance: a population dynamics view. Clin Infect Dis, 33(Supplement 3), S161-S169.
Levin, B. R. (2002). Models for the spread of resistant pathogens. Neth J Med, 60

(suppl), 58–64; discussion 64–66.

Levin, B. R., Lipsitch, M., Perrot, V., Schrag, S., Antia, R., Simonsen, L., *et al*. (1997). The population genetics of antibiotic resistance. Clin Infect Dis, 24(Supplement_1), S9-S16.

Li, L. G., Xia, Y., & Zhang, T. (2017). Co-occurrence of antibiotic and metal resistance genes revealed in complete genome collection. ISME J, 11(3), 651-662.

Li, X. Z., & Nikaido, H. (2009). Efflux-mediated drug resistance in bacteria. Drugs, 69(12), 1555-1623.

Liebert, C. A., Hall, R. M., & Summers, A. O. (1999). Transposon Tn*21,* flagship of the floating genome. Microbiol Mol Biol Rev, 63(3), 507-522.

Lilley, A. K., Bailey, M. J. (1997). Impact of plasmid pQBR103 acquisition and carriage on the phytosphere fitness of *Pseudomonas fluorescens* SBW25: burden and benefit. Appl Environ Microbiol, 63, 1584–1587.

Lindow, S. E. (2017). Horizontal gene transfer gone wild: promiscuity in a kiwifruit pathogen leads to resistance to chemical control. Environ Microbiol, 19(4), 1363-1365.

López-Pérez, M., Ramon-Marco, N., & Rodriguez-Valera, F. (2017). Networking in microbes: conjugative elements and plasmids in the genus Alteromonas. BMC genomics, 18(1), 36.

Lovell, H. C., Jackson, R. W., Mansfield, J. W., Godfrey, S. A., Hancock, J. T., Desikan, R., & Arnold, D. L. (2011). *In planta* conditions induce genomic changes in *Pseudomonas syringae* pv. *phaseolicola*. Mol Plant Pathol, 12(2), 167-176.

Lovell, H. C., Mansfield, J. W., Godfrey, S. A. C., Jackson, R. W., Hancock, J. T. & Arnold, D. L. (2009). Bacterial evolution by genomic island transfer occurs via DNA transformation *in planta*. Curr Biol, 19, 1586–1590.

Macomber, L., & Imlay, J. A. (2009). The iron-sulfur clusters of dehydratases are primary intracellular targets of copper toxicity. Proc Natl Acad Sci, 106(20), 8344-8349.

Maddocks, S. E., & Oyston, P. C. (2008). Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins. Microbiology, 154(12), 3609-3623.

Magot, M. (1983). Transfer of antibiotic resistances from *Clostridium innocuum* to *Clostridium perfringens* in the absence of detectable plasmid DNA. FEMS Microbiol Lett, 18(1-2), 149-151.

Manoil, C. (2000). Tagging exported proteins using *Escherichia coli* alkaline phosphatase gene fusions. Adv Bot Res, 326, 35-47.

Mansour, N. M., Sawhney, M., Tamang, D. G., Vogl, C., & Saier, M. H. (2007). The bile/arsenite/riboflavin transporter (BART) superfamily. FEBS J, 274(3), 612-629.

Marrero, J., & Waldor, M. K. (2007). The SXT/R391 family of integrative conjugative elements is composed of two exclusion groups. J Bacteriol, 189(8), 3302-3305.

Marri, P. R., Hao W., & Golding, G. B. (2007). The role of laterally transferred genes in adaptive evolution. BMC Evol Biol, 7(1), S8.

Marri, P. R., Hao, W., & Golding, G. B. (2006). Gene gain and gene loss in streptococcus: is it driven by habitat?. Mol Biol Evol, 23(12), 2379-2391.

Martínez, J. L. (2008). Antibiotics and antibiotic resistance genes in natural environments. Science, 321(5887), 365-367.

Mays, T. D., Smith, C. J., Welch, R. A., Delfini, C., & Macrina, F. L. (1982). Novel antibiotic resistance transfer in *Bacteroides*. Antimicrob Agents Chemother, 21(1), 110-118.

Mazzaglia, A., Studholme, D. J., Taratufolo, M. C., Cai, R., Almeida, N. F., Goodman, T., *et al*. (2012) *Pseudomonas syringae* pv. *actinidiae* (PSA) isolates from recent bacterial canker of kiwifruit outbreaks belong to the same genetic lineage. PLoS One, 7, e36518.

McCann, H. C., & Guttman, D. S. (2008). 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., Liu, Y., Templeton, M. D., Colombi, E., Straub, C., *et al.* (2017). Origin and evolution of a pandemic lineage of the kiwifruit pathogen *Pseudomonas syringae* pv. *actinidiae*. Genome Biol Evol, 9(4), 932–944.

McCann, H. C., Rikkerink, E. H. A., Bertels, F., Fiers, M., Lu, A., Rees-George, J., *et al*. (2013). Genomic analysis of the kiwifruit pathogen *Pseudomonas syringae* pv. *actinidiae* provides insight into the origins of an emergent plant disease. PLoS Pathog, 9, e1003503.

Médigue, C., Rouxel, T., Vigier, P., Hénaut, A., Danchin, A. (1991). Evidence for horizontal gene transfer in *Escherichia coli* speciation. J Mol Biol, 222, 851-6.

Medrano-Soto, A., Moreno-Hagelsieb, G., Vinuesa, P., Christen, J. A., & Collado-Vides, J. (2004). Successful lateral transfer requires codon usage compatibility between foreign genes and recipient genomes. Mol Biol Evol, 21(10), 1884-1894.

Meier, J. L., & Burkart, M. D. (2011). Proteomic analysis of polyketide and nonribosomal peptide biosynthesis. Curr Opin Chem Biol, 15(1), 48-56.

Mellgren, E. M., Kloek, A. P., & Kunkel, B. N. (2009). Mqo, a tricarboxylic acid cycle enzyme, is required for virulence of *Pseudomonas syringae* pv. *tomato* strain DC3000 on *Arabidopsis thaliana*. J Bacteriol, 191(9), 3132-3141.

Menkissoglu, O. & Lindow, S. E. (1991). Chemical forms of copper on leaves in relation to the bactericidal activity of cupric hydroxide deposits on plants. Phytopathology, 81, 1263-1270.

Mergeay, M., Monchy, S., Vallaeys, T., Auquier, V., Benotmane, A., Bertin, P., *et al*. (2003). *Ralstonia metallidurans*, a bacterium specifically adapted to toxic metals: towards a catalogue of metal-responsive genes. FEMS Microbiol Rev, 27, 385–410.

Midonet, C., & Barre, F. X. (2014). Xer site-specific recombination: promoting vertical and horizontal transmission of genetic information. Microbiol Spectrum, 2(6).

Millardet, P. (1885). Sur l'histoire du traitement du mildiou par le sulfate de cuivre.

Miller, J. "II. 1972. Experiments in molecular genetics."

Mindlin, S., Minakhin, L., Petrova, M., Kholodii, G., Minakhina, S., Gorlenko, Z., & Nikiforov, V. (2005). Present-day mercury resistance transposons are common in bacteria preserved in permafrost grounds since the Upper Pleistocene. Res Microbiol, 156(10), 994-1004.

Mirkin, B. G., Fenner, T. I., Galperin, M. Y., & Koonin, E. V. (2003). Algorithms for computing parsimonious evolutionary scenarios for genome evolution, the

last universal common ancestor and dominance of horizontal gene transfer in the evolution of prokaryotes. BMC Evol, 3(1), 2.

Miyazaki, R., Minoia, M., Pradervand, N., Sulser, S., Reinhard, F., & van der Meer, J. R. (2012). Cellular variability of RpoS expression underlies subpopulation activation of an integrative and conjugative element. PLoS Genet, 8(7), e1002818.

Mohd-Zain, Z., Turner, S.L., Cerdeno-Tarraga, A.M., Lilley, A.K., Inzana, T.J., Duncan, A.J., *et al*. (2004). Transferable antibiotic resistance elements in *Haemophilus influenzae* share a common evolutionary origin with a diverse family of syntenic genomic islands. J Bacteriol, 186, 8114–8122.

Monteil, C. L., Cai, R., Liu, H., Mechan Llontop, M. E., Studholme, D. J., Morris, C. E., & Vinatzer, B. A. (2013). Nonagricultural reservoirs contribute to emergence and evolution of *Pseudomonas syringae* crop pathogens. New Phytol, 199(3), 800-811.

Monteil, C. L., Guilbaud, C., Glaux, C., Lafolie, F., Soubeyrand, S., & Morris, C. E. (2012). Emigration of the plant pathogen *Pseudomonas syringae* from leaf litter contributes to its population dynamics in alpine snowpack. Environ Microbiol, 14(8), 2099-2112.

Morens, D. M., & Fauci, A. S. (2013). Emerging infectious diseases: threats to human health and global stability. PLoS Pathog, 9(7), e1003467.

Morgan, R. K., & Taylor, E. (2004). Copper accumulation in vineyard soils in New Zealand. Environ Sci, 1,2, 139-167.

Morris, C. E., Kinkel, L. L., Xiao, K., Prior, P., & Sands, D. C. (2007). Surprising niche for the plant pathogen *Pseudomonas syringae*. Infect Genet Evol, 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. Annu Rev Phytopathol, 51, 85-104.

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), e00107-10.

Morris, C. E., Sands, D. C., Vinatzer, B. A., Glaux, C., Guilbaud, C., Buffiere, A., *et al*. (2008). The life history of the plant pathogen *Pseudomonas syringae* is linked to the water cycle. ISME J, 2(3), 321-334.

Mukhopadhyay, R., Rosen, B. P., Phung, L. T., & Silver, S. (2002). Microbial arsenic: from geocycles to genes and enzymes. FEMS Microbiol Rev, 26(3), 311-325.

Muñoz-López, M., & García-Pérez, J. L. (2010). DNA transposons: nature and applications in genomics. Curr Genomics, 11(2), 115-128.

Murphy, E. (1983). Inhibition of Tn*554* transposition: deletion analysis. Plasmid 10, 260–269.

Mutreja, A., Kim, D. W., Thomson, N. R., Connor, T. R., Lee, J. H., Kariuki, S., *et al.* (2011). Evidence for several waves of global transmission in the seventh cholera pandemic. Nature, 477(7365), 462-465.

Nakajima, M., Goto, M., & Hibi, T. (2002). Similarity between copper resistance genes from *Pseudomonas syringae* pv. *actinidiae* and *P. syringae* pv. *tomato*, J Gen Plant Pathol, 68, 68–74.

Neale, H. C., Laister, R., Payne, J., Preston, G., Jackson, R. W. & Arnold, D. L.

(2016). A low frequency persistent reservoir of a genomic island in a pathogen population ensures island survival and improves pathogen fitness in a susceptible host. Environ Microbiol, 18(11), 4144-4152.

Nicaise, V., Roux, M., & Zipfel, C. (2009). Recent advances in PAMP-triggered immunity against bacteria: pattern recognition receptors watch over and raise the alarm. Plant Physiol, 150(4), 1638-1647.

Nikaido, H. (2009). Multidrug resistance in bacteria. Annu Rev Biochem, 78, 119-146.

Njamkepo, E., Fawal, N., Tran-Dien, A., Hawkey, J., Strockbine, N., Jenkins, C., *et al*. (2016). Global phylogeography and evolutionary history of Shigella dysenteriae type 1. Nat Microbiol, 1, 16027.

Novick, R. P., & Ram, G. (2016). The floating (pathogenicity) island: a genomic dessert. Trends Genet, 32(2), 114-126.

Nowell, R. W., Green, S., Laue, B. E., & Sharp, P. M. (2014). The extent of genome flux and its role in the differentiation of bacterial lineages. Genome Biol Evol, 6(6), 1514-1529.

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*. Mol Plant Pathol, 17(9), 1409-1424.

Nürnberger, T., Brunner, F., Kemmerling, B., & Piater, L. (2004). Innate immunity in plants and animals: striking similarities and obvious differences. Immunol Rev, 198(1), 249-266.

Ochman, H., Lawrence, J.G. and Groisman, E.A. (2000). Lateral gene transfer and the nature of bacterial innovation. Nature, 405, 299–304.

Ochman, H., Lerat, E., Daubin, V. (2005). Examining bacterial species under the specter of gene transfer and exchange. Proc Natl Acad Sci, 102 Suppl 1, 6595-6599.

Panda, P., Vanga, B. R., Lu, A., Fiers, M., Fineran, P. C., Butler, R., *et al*. (2016). *Pectobacterium atrosepticum* and *Pectobacterium carotovorum* harbor distinct, independently acquired integrative and conjugative elements encoding coronafacic acid that enhance virulence on potato stems. Front Microbiol, 7.

Panter, S. N., & Jones, D. A. (2002). Age-related resistance to plant pathogens. Adv Botanical Res, 38, 251-280.

Patil, S. S., Hayward, A. C., & Emmons, R. (1974). An ultraviolet-induced nontoxigenic mutant of *Pseudomonas phaseolicola* of altered pathogenicity. Phytopathology, 64,590–595.

Peters, S. E., Hobman, J. L., Strike, P., & Ritchie, D. A. (1991). Novel mercury resistance determinants carried by IncJ plasmids pMERPH and R391. Mol Gen Genet, 228, 294—299.

Pirofski, L. A., & Casadevall, A. (2002). The meaning of microbial exposure, infection, colonisation, and disease in clinical practice. Lancet Infect Dis, 2(10), 628-635.

Pitman, A. R., Jackson, R. W., Mansfield, J. W., Kaitell, V., Thwaites, R., *et al*. (2005). Exposure to host resistance mechanisms drives evolution of bacterial virulence in plants. Curr Biol, 15, 2230–2235.

Ploetz, R. C. (2015). Fusarium wilt of banana. Phytopathology, 105(12), 1512-1521.

Polz, M. F., Alm, E. J., & Hanage, W. P. (2013). Horizontal gene transfer and the

evolution of bacterial and archaeal population structure. Trends Genet, 29, 170–175.

Popa, O., & Dagan, T. (2011). Trends and barriers to lateral gene transfer in prokaryotes. Curr Opin Microbiol, 14, 615-623.

Puig, S., Lee, J., Lau, M., & Thiele, D. J. (2002). Biochemical and genetic analyses of yeast and human high affinity copper transporters suggest a conserved mechanism for copper uptake. J Biol Chem, 277(29), 26021-26030.

Quiroz, T. S., Nieto, P. A., Tobar, H. E., Salazar-Echegarai, F. J., Lizana, R. J., Quezada, C. P., *et al*. (2011). Excision of an unstable pathogenicity island in *Salmonella enterica* serovar *enteritidis* is induced during infection of phagocytic cells. PLoS ONE, 6, e26031.

Rainey, P. B., & Bailey, M. J. (1996). Physical and genetic map of the *Pseudomonas fluorescens* SBW25 chromosome. Mol Microbiol, 19(3), 521-533.

Rajkumar, M., Prasad, M. N. V., Swaminathan, S., & Freitas, H. (2013). Climate change driven plant–metal–microbe interactions. Environ Int, 53, 74-86.

Ramsay, J. P., Sullivan, J. T., Jambari, N., Ortori, C. A., Heeb, S., Williams, P., *et al*. (2009). A LuxRI-family regulatory system controls excision and transfer of the *Mesorhizobium loti* strain R7A symbiosis island by activating expression of two conserved hypothetical genes. Mol Microbiol, 73(6), 1141-1155.

Ramsay, J. P., Sullivan, J. T., Stuart, G. S., Lamont, I. L., & Ronson, C. W. (2006). Excision and transfer of the *Mesorhizobium loti* R7A symbiosis island requires an integrase IntS, a novel recombination directionality factor RdfS, and a putative relaxase RlxS. Mol Microbiol, 62, 723–734.

Rashtchian, A. Y. O. U. B., Dubes, G. R., & Booth, S. J. (1982). Tetracycline-inducible transfer of tetracycline resistance in *Bacteroides* fragilis in the absence of detectable plasmid DNA. J Bacteriol, 150(1), 141-147.

Rauch, P. J. G., & De Vos, W. M. (1992). Characterization of the novel nisin-sucrose conjugative transposon Tn*5276* and its insertion in *Lactococcus lactis*. J Bacteriol, 174, 1280-1287.

Ravatn, R., Studer, S., Springael, D., Zehnder, A.J.B., & Van Der Meer, J.R. (1998) Chromosomal integration, tandem amplification, and deamplification in *Pseudomonas putida* F1 of a 105-kilobase genetic element containing the chlorocathecol degradative genes from *Pseudomonas* sp. strain B13. J Bacteriol, 180, 4360–4369.

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

Reid, C. J., Chowdhury, P. R., & Djordjevic, S. P. (2015). Tn*6026* and Tn*6029* are found in complex resistance regions mobilised by diverse plasmids and chromosomal islands in multiple antibiotic resistant *Enterobacteriaceae*. Plasmid, 80, 127-137.

Reinhard, F., & van der Meer, J. R. (2014). Life history analysis of integrative and conjugative element activation in growing microcolonies of *Pseudomonas*. J Bacteriol, 196(7),1425–1434.

Reinhard, F., Miyazaki, R., Pradervand, N., & van der Meer, J. R. (2013). Cell

differentiation to “mating bodies” induced by an integrating and conjugative element in free-living bacteria. Curr Biol, 23(3), 255-259.

Rensing, C., & Grass, G. (2003). *Escherichia coli* mechanisms of copper homeostasis in a changing environment. FEMS Microbiol Rev, 27, 197–213.

Retchless, A. C., & Lawrence, J. G. (2007). Temporal fragmentation of speciation in bacteria. Science, 317, 1093-1096.

Ristaino, J. B. (2002). Tracking historic migrations of the Irish potato famine pathogen, *Phytophthora infestans*. Microbes infect, 4(13), 1369-1377.

Roberts, A. P., Chandler, M., Courvalin, P., Guédon, G., Mullany, P., Pembroke, T., *et al*. (2008). Revised nomenclature for transposable genetic elements. Plasmid, 60(3), 167-173.

Roberts, A. P., & Mullany, P. (2009). A modular master on the move: the Tn*916* family of mobile genetic elements. Trends Microbiol, 17, 251–258.

Roberts, M. C., & Smith, A. L. (1980). Molecular characterization of "plasmid-free" antibiotic-resistant *Haemophilus influenzae*. J Bacteriol, 144(1), 476-479.

Rojas-Jiménez, K., Sohlenkamp, C., Geiger, O., Martínez-Romero, E., Werner, D., & Vinuesa, P. (2005). A ClC chloride channel homolog and ornithine-containing membrane lipids of *Rhizobium tropici* CIAT899 are involved in symbiotic efficiency and acid tolerance. Mol Plant Microbe Interact, 18(11), 1175-1185.

Rosselló-Mora, R., & Amann, R. (2001). The species concept for prokaryotes. FEMS Microbiol Ecol, 25(1), 39-67.

Ruinelli, M., Schneeberger, P. H. H., Ferrante, P., Bühlmann, A., Scortichini, M., Vanneste, J. L., *et al*. (2016). Comparative genomics-informed design of two LAMP assays for detection of the kiwifruit pathogen *Pseudomonas syringae* pv. *actinidiae* and discrimination of isolates belonging to the pandemic biovar 3. Plant Pathol doi:10.1111/ppa.12551.

Sabry, S. A., Ghozlan, H. A., & Abou-Zeid, D. M. (1997). Metal tolerance and antibiotic resistance patterns of a bacterial population isolated from sea water. J Appl Microbiol, 82(2), 245-252.

Sato, T., & Kuramitsu, H. (1998). Plasmid maintenance renders bacteria more susceptible to heat stress. Microbiol Immunol, 42(6), 467-469.

Sawada, H., Kondo, K., & Nakaune, R. (2016). Novel biovar (biovar 6) of *Pseudomonas syringae* pv. *actinidiae* causing bacterial canker of kiwifruit (*Actinidia deliciosa*) in Japan. Jpn J Phytopathol, 82(2), 101-115.

Schäfer, A., Tauch, A., Jäger, W., Kalinowski, J., Thierbach, G., & Pühler, A. (1994). Small mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: selection of defined deletions in the chromosome of C*orynebacterium glutamicum*. Gene, 145(1), 69-73.

Schechter, L. M., Roberts, K. A., Jamir, Y., Alfano, J. R., & Collmer, A. (2004). *Pseudomonas syringae* type III secretion system targeting signals and novel effectors studied with a Cya translocation reporter. J Bacteriol, 186(2), 543-555.

Scortichini, M. (1994). Occurrence of *Pseudomonas syringae* pv. *actinidiae* on kiwifruit in Italy. Plant Pathol, 43(6), 1035-1038.

Scott, J. R., Bringel, F., Marra, D., Alstine, G., & Rudy, C. K. (1994). Conjugative transposition of Tn*916*: preferred targets and evidence for conjugative transfer of a single strand and for a double-stranded circular

intermediate. Mol Microbiol, 11(6), 1099-1108.
Segonzac, C., & Zipfel, C. (2011). Activation of plant pattern-recognition receptors by bacteria. Curr Opin Microbiol, 14(1), 54-61.
Sela, I., Wolf, Y. I., & Koonin, E. V. (2016). Theory of prokaryotic genome evolution. Proc Natl Acad Sci, 201614083.
Serizawa S., Ichikawa, T., Takikawa, Y., Tsuyumu S., & Goto M. (1989). Occurrence of bacterial canker of kiwifruit in Japan: description of symptoms, isolation of the pathogen and screening of bactericides. Ann Phytopathol Soc Japan, 55, 427–436
Shapiro, B. J. (2016). How clonal are bacteria over time?. Curr Opin Microbiol, 31, 116-123.
Shoemaker, N. B., Smith, M. D., & Guild, W. R. (1980). DNase-resistant transfer of chromosomal cat and tet insertions by filter mating in *Pneumococcus*. Plasmid, 3(1), 80-87.
Smith, C. J., Markowitz, S. M., & Macrina, F. L. (1981). Transferable tetracycline resistance in *Clostridium difficile*. Antimicrob Agents Chemother, 19(6), 997-1003.
Spagnoletti, M., Ceccarelli, D., Rieux, A., Fondi, M., Taviani, E., Fani, R., *et al*. (2014). Acquisition and evolution of SXT-R391 integrative conjugative elements in the seventh-pandemic *Vibrio cholerae* lineage. MBio, 5(4), e01356-14.
Spratt, B. G. (1996). Antibiotic resistance: counting the cost. Curr Biol, 6(10), 1219-1221.
Stackebrandt, E., Frederiksen, W., Garrity, G. M., Grimont, P. A., Kämpfer, P., Maiden, M. C., *et al.* (2002). Report of the ad hoc committee for the re-evaluation of the species definition in bacteriology. Inte J Syst Evol Microbiol, 52(3), 1043-1047.
Staehlin, B. M., Gibbons, J. G., Rokas, A., O'Halloran, T. V., & Slot, J. C. (2016). Evolution of a heavy metal homeostasis/resistance island reflects increasing copper stress in *Enterobacteria*. Genome Biol Evol, 8, 811-826.
Stevens, R.B. (1960). Plant Pathology, an Advanced Treatise, Vol. 3. J.G. Horsfall & A.E. Dimond, eds. Academic Press, NY.
Stevenson, C., Hall, J. P., Harrison, E., Wood, A. J., & Brockhurst, M. A. (2017). Gene mobility promotes the spread of resistance in bacterial populations. ISME J.
Stoesser, N., Giess, A., Batty, E. M., Sheppard, A. E., Walker, A. S., Wilson, D. J., *et al*. (2014). Genome sequencing of an extended series of NDM-producing *Klebsiella pneumoniae* isolates from neonatal infections in a Nepali hospital characterizes the extent of community- versus hospital-associated transmission in an endemic setting. Antimicrob Agents Chemother, 58, 7347–7357.
Stokes, H. W., & Gillings, M. R. (2011). Gene flow, mobile genetic elements and the recruitment of antibiotic resistance genes into Gram-negative pathogens. FEMS Microbiol Rev, 35(5), 790-819.
Straub, C. (2017). The ecological genetics of *P. syringae* in the kiwifruit phyllosphere. (Doctoral dissertation, Massey University, *in prep*).
Strieker, M., Tanović, A., & Marahiel, M. A. (2010). Nonribosomal peptide synthetases: structures and dynamics. Curr Opin Struct Biol, 20(2), 234-240.

Stuy, J. H. (1980). Chromosomally integrated conjugative plasmids are common in antibiotic-resistant *Haemophilus influenzae*. J Bacteriol, 142(3), 925-930.

Subbiah, M., Top, E. M., Shah, D. H., & Call, D. R. (2011). Selection pressure required for long-term persistence of blaCMY-2-positive IncA/C plasmids. Appl Environ Microbiol, 77(13), 4486-4493.

Sullivan, J., & Ronson, C. (1998). Evolution of rhizobia by acquisition of a 500-kb symbiosis island that integrates into a phe-tRNA gene. Proc Natl Acad Sci USA, 95, 5145–5149

Sullivan, J. T., Trzebiatowski, J. R., Cruickshank, R. W., Gouzy, J., Brown, S. D., Elliot, R. M., *et al*. (2002). Comparative sequence analysis of the symbiosis island of *Mesorhizobium loti* strain R7A. J Bacteriol, 184(11), 3086-3095.

Sullivan, J. T., & Ronson, C. W. (1998). Evolution of rhizobia by acquisition of a 500-kb symbiosis island that integrates into a phe-tRNA gene. Proc Natl Acad Sci USA, 9, 5145–514.

Sullivan, J. T., Patrick, H. N., Lowther, W. L., Scott, D. B., & Ronson, C. W. (1995). Nodulating strains of *Rhizobium loti* arise through chromosomal symbiotic gene transfer in the environment. Proc Natl Acad Sci USA, 92, 8985–8989.

Sundin, G. W., Jones, A. L., & Fulbright, D. W. (1989). Copper resistance in *Pseudomonas syringae* pv. *syringae* from cherry orchards and its associated transfer *in vitro* and *in planta* with a plasmid. Phytopathology, 79, 861-865.

Sutanto, Y., Shoemaker, N. B., Gardner, J. F., & Salyers, A. A. (2002). Characterization of Exc, a novel protein required for the excision of *Bacteroides* conjugative transposon. Mol Microbiol, 46(5), 1239-1246.

Sweetlove, L. J., Beard, K. F., Nunes-Nesi, A., Fernie, A. R., & Ratcliffe, R. G. (2010). Not just a circle: flux modes in the plant TCA cycle. Trends Plant Sci, 15(8), 462-470.

Swofford, D. L. (2003). PAUP*. Phylogenetic analysis using parsimony (* and other methods). Version 4.

Taghbalout, A., & Rothfield, L. (2007). RNaseE and the other constituents of the RNA degradosome are components of the bacterial cytoskeleton. Proc Natl Acad Sci, 104(5), 1667-1672.

Takikawa, Y., Serizawa, S., Ichikawa, T., & Tsuyumu, S. (1989). *Pseudomonas syringae* pv. *actinidiae* pv. nov.: The causal bacterium of canker of kiwifruit in Japan. Jpn J Phytopathol, 55, 437–444.

Tamagnone, L., Merida, A., Stacey, N., Plaskitt, K., Parr, A., Chang, C. F., *et al*. (1998). Inhibition of phenolic acid metabolism results in precocious cell death and altered cell morphology in leaves of transgenic tobacco plants. Plant Cell, 10(11), 1801-1816.

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 Announc, 3, e01054-15.

Teverson, D. M. (1991). Genetics of pathogenicity and resistance in the halo-blight disease of beans in Africa. Ph.D. thesis. University of Birmingham, Birmingham, United Kingdom.

Toprak, E., Veres, A., Michel, J. B., Chait, R., Hartl, D. L., & Kishony, R. (2012). Evolutionary paths to antibiotic resistance under dynamically sustained drug selection. Nat Genet, 44(1), 101-105.

Touchon, M., Hoede, C., Tenaillon, O., Barbe, V., Baeriswyl, S., Bidet, P., *et al.*

(2009). Organised genome dynamics in the *Escherichia coli* species results in highly diverse adaptive paths. PLoS Genet, 5(1), e1000344.

Tuller, T., Girshovich, Y., Sella, Y., Kreimer, A., Freilich, S., Kupiec, M., ..., & Ruppin, E. (2011). Association between translation efficiency and horizontal gene transfer within microbial communities. Nucleic Acids Res, 39(11):4743-55.

van Rensburg, E., Den Haan, R., Smith, J., van Zyl, W. H., & Görgens, J. F. (2012). The metabolic burden of cellulase expression by recombinant *Saccharomyces cerevisiae* Y294 in aerobic batch culture. Appl Microbiol Biotechnol, 96(1), 197-209.

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. Environ Microbiol, 17, 4730–4744.

Vanneste, J. L., Giovanardi, D., Yu, J., Cornish, D. A., Kay, C., Spinelli, F., & Stefani, E. (2011). Detection of *Pseudomonas syringae* pv. *actinidiae* in kiwifruit pollen samples. NZ Plant Protect, 64, 246-251.

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 Announc, 4, e00061–16.

Völksch, B., & Weingart, H. (1998). Toxin production by pathovars of *Pseudomonas syringae* and their antagonistic activities against epiphytic microorganisms. J Basic Microbiol, 38(2), 135-145.

von Wintersdorff, C. J., Penders, J., van Niekerk, J. M., Mills, N. D., Majumder, S., van Alphen, L. B., *et al*. (2016). Dissemination of antimicrobial resistance in microbial ecosystems through horizontal gene transfer. Front Microbiol, 7.

Vos, M., Hesselman, M. C., Te Beek, T. A., van Passel, M. W., & Eyre-Walker, A. (2015). Rates of lateral gene transfer in prokaryotes: high but why?. Trends Microbiol, 23(10), 598-605.

Wang, Y., Rotman, E. R., Shoemaker, N. B., & Salyers, A. A. (2005). Translational control of tetracycline resistance and conjugation in the *Bacteroides* conjugative transposon CTnDOT. J Bacteriol, 187(8), 2673-2680.

Wang, Y., Shoemaker, N. B., & Salyers, A. A. (2004). Regulation of a *Bacteroides* operon that controls excision and transfer of the conjugative transposon CTnDOT. J Bacteriol, 186(9), 2548-2557.

Ward, D. M. (1998). A natural species concept for prokaryotes. Curr Opin Microbiol, 1(3), 271-277.

Weng, Y., Chen, F., Liu, Y., Zhao, Q., Chen, R., Pan, X., *et al*. (2016). *Pseudomonas aeruginosa* enolase influences bacterial tolerance to oxidative stresses and virulence. Front Microbiol, 7.

Wong, V. K., Baker, S., Pickard, D. J., Parkhill, J., Page, A. J., Feasey, N. A., *et al*. (2015). Phylogeographical analysis of the dominant multidrug-resistant H58 clade of *Salmonella Typhi* identifies inter-and intracontinental transmission events. Nat Genet, 47(6), 632-639.

Wozniak, R. A., Fouts, D. E., Spagnoletti, M., Colombo, M. M., Ceccarelli, D., Garriss, G., *et al*. (2009). Comparative ICE genomics: insights into the evolution of the SXT/R391 family of ICEs. PLoS Genet, 5(12), e1000786.

Wozniak, R. A. F., & Waldor, M. K. (2010). Integrative and conjugative elements:

mosaic mobile genetic elements enabling dynamic lateral gene flow. Nat Rev Microbiol, 8, 552–563.

Wren, B. W. (2003). The *yersiniae*—a model genus to study the rapid evolution of bacterial pathogens. Nat Rev Microbiol, 1(1), 55-64.

Wysocki, R., & Bobrowicz, P. (1997). The *Saccharomyces cerevisiae* ACR3 gene encodes a putative membrane protein involved in arsenite transport. J Biol Chem, 272(48), 30061-30066.

Xu, G.-W., & Gross, D.C. (1988). Evaluation of the role of syringomycin in plant pathogenesis by using Tn*5* mutants of *Pseudomonas syringae* pv. *syringae* defective in syringomycin production. Appl Environ Microbiol, 54,1345–1353

Zhang, L., Koay, M., Maher, M. J., Xiao, Z., & Wedd, A. G. (2006). Intermolecular transfer of copper ions from the CopC protein of *Pseudomonas syringae*. Crystal structures of fully loaded CuICuII forms. J Am Chem Soc, 128(17), 5834-5850.

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. J Bacteriol, 197, 2867–2878.

Zhang, X.X., George, A., Bailey, M.J., & Rainey, P.B. (2006) The histidine utilization (*hut*) genes of *Pseudomonas fluorescens* SBW25 are active on plant surfaces, but are not required for competitive colonization of sugar beet seedlings. Microbiol, 152, 1867–1875.

Zhou, D., Han, Y., Song, Y., Huang, P., & Yang, R. (2004). Comparative and evolutionary genomics of Yersinia pestis. Microbes Infect, 6(13), 1226-1234.

Zhou, J. M., & Chai, J. (2008). Plant pathogenic bacterial type III effectors subdue host responses. Curr Opin Microbiol, 11(2), 179-185.

Appendix 1 - Cluster 3

A1.1 Introduction

P. syringae spp. produces a variety of non-ribosomial peptides that can be involved in virulence. Non-ribosomal peptides are a diverse group of peptides that are synthesised on large protein templates and not by ribosomes. They encompass a diverse range of compounds including siderophores, toxins, virulence factors and antibiotics (Cornelis & Matthijs, 2002, Strieker *et al*., 2010; Meier & Burkart, 2011). In *P. syringae,* phytotoxins such as syringolin A (Groll *et al*., 2008) and mangotoxin (Arrebola *et al*., 2007; Carrión *et al*., 2012) are non-ribosomal peptides with a non-host-specific role in virulence (Bender *et al*., 1999). Phytotoxins can be involved in the systemic movement of bacteria *in planta* (Patil, 1974), lesion size (Xu & Gross, 1988; Arrebola *et al*., 2006), and multiplication inside the host (Bender *et al*., 1987). At the same time, non-ribosomal encoded siderophores do not influence the production of symptoms in *P. syringae* pv. *phaseolicola* (Cornelis & Matthijs, 2002).

The presence of genes encoding secondary metabolites in *Psa* NZ13 was assessed using antiSMASH (Blin *et al*., 2013). This program identified clusters predicted to produce mangotoxin, ectoine, pyoverdine, achromobactin,

yersinibactin and one predicted to produce an unknown metabolite. This unknown cluster, termed Cluster 3, is shown in Figure A1.1.

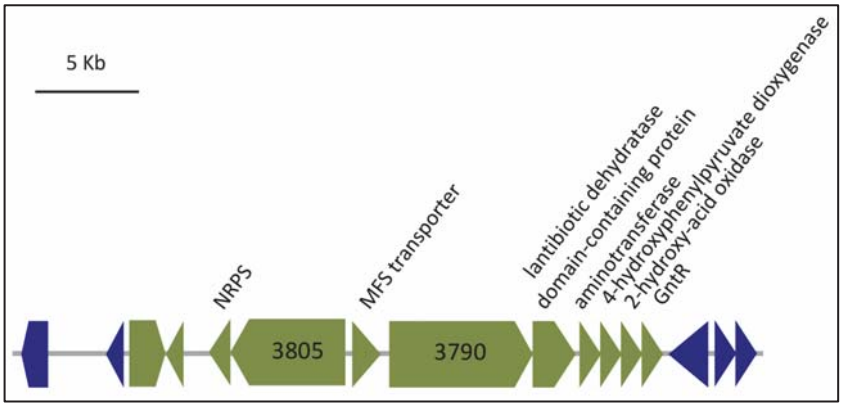

Figure A1.1. Genetic organization of Cluster 3. Green boxes depict genes predicted to be involved in toxin production, and the blue boxes mobile genes (recombinases or transposases).

Cluster 3 is located between nucleotides 771,255 and 805,662 on the *Psa* NZ13 genome, spans 34 kb, 11 open reading frames, and is hypothesized to produce a peptide of unknown structure. The cluster includes two non-ribosomal peptide synthetases (NRPS) (IYO_003805, termed *3805* and IYO_003790, termed *3790*) and a lantibiotic type dehydratase (IYO_003810). NRPS are enzymes that produce peptides without mRNA, and the lantibiotic type dehydratase is involved in post-translational modification of ribosomal peptides. The presence of transposase genes upstream and downstream of Cluster 3 suggests that it has been transferred as one functional unit (Li Bo, *personal communication*).

To explore the function of this novel cluster, the genes termed *3805* and *3790* (Figure A1.1) that both code for NRPSs were used as the target of the mutagenesis. Single and double mutants were constructed by a two-step allelic exchange strategy (refer to Section 2.2.11) and confirmed by PCR and sequencing. The mutants *Psa* NZ13 Δ*3805*, *Psa* NZ13 Δ*3790* and the double mutant *Psa* NZ13 Δ*3805-3790* were tested *in planta* to assess the activity of the putative toxin, while

more sensitive biochemical assays will be carried out by the group of our collaborator Prof. Li Bo at the University of North Carolina.

A1.2 Results

A1.2.1 Spot inoculation

To assess the activity of the putative toxin produced by Cluster 3, a panel of different vegetables, fruits and mushrooms were used for spot inoculation. On bok choy, plum, apple, kale, white and brown bottom mushrooms, parsnip and tomato, *Psa* NZ13 produced no phenotype. On nectarine, roman lettuce and green bean, the wildtype, *Psa* NZ13 Δ*3805*, *Psa* NZ13 Δ*3790* and the double mutant *Psa* NZ13 Δ*3805-3790* produced the same blackening of the inoculated area (Figure A1.2).

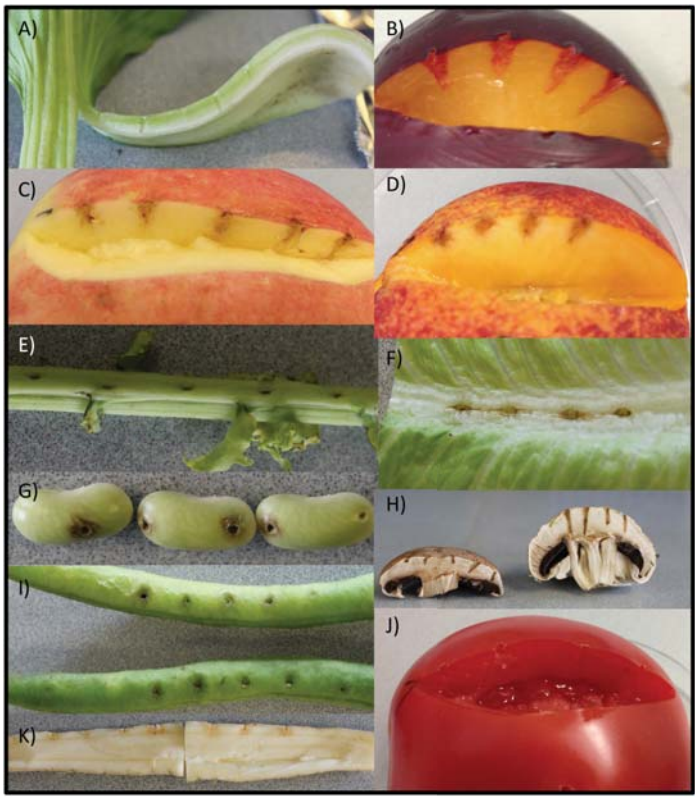

Figure A1.2. Cluster 3 mutants stab inoculations. For each host used the strains were inoculated from left to right in the following order: *Psa* NZ13, *Psa* NZ13 Δ*3805*, *Psa* NZ13 Δ*3790*

and *Psa* NZ13 Δ*3805-3790*. The hosts used are: A) bok choy, B) plum, C) apple, D) nectarine, E) kale, F) roman lettuce, G) and I) green beans, H) brown and white bottom mushrooms, K) parsnip, and J) tomato.

A1.2.2 Growth on kiwifruit

Cluster 3 mutants were tested for their colonization ability on kiwifruit Hort16A and symptoms production (Figure A1.3 and A1.4).

This preliminary experiment showed that the Cluster 3 mutants were able to colonize and produce leaf spots on *A. chinensis* Hort16A plants as the wildtype *Psa* NZ13.

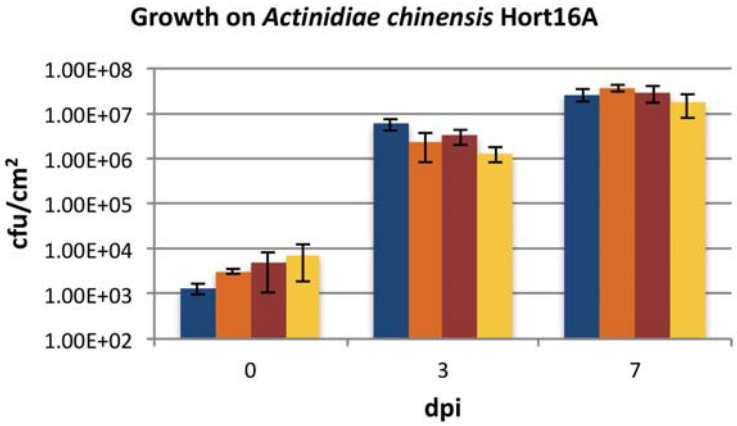

Figure A1.3. Cluster 3 growth on Hort16A. Growth of *Psa* NZ13 (blue bars), *Psa* NZ13 Δ*3805* (orange bars), *Psa* NZ13 Δ*3790* (red bars), and *Psa* NZ13 Δ*3805-3790* (yellow bars) was assessed endophytically on leaves of the kiwifruit cultivar Hort16A. Data are mean and standard deviation of four replicates. Two-tailed t-test showed no statistical difference between the growth of *Psa* NZ13 and the mutants.

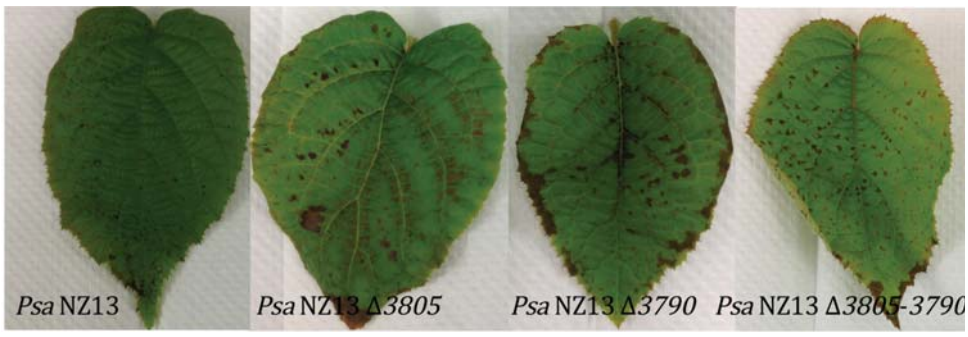

Figure A1.4. Symptoms on Hort16A leaves. Symptoms production of *Psa* NZ13 and Cluster 3 mutants during the growth assay showed in Figure A1.4. Pictures were taken at 7dpi.

A1.3 Discussion

Toxins are important weapons for *P. syringae*, both for host colonization and for antagonistic activity versus endophytic bacteria (Völksch & Weingart, 1998; Bender *et al.*, 1999). Here I have carried out initial work to evaluate the involvement of the novel Cluster 3 in the plant-pathogen interaction. Both the spot inoculations on a diverse panel of hosts and pathogenicity assay on kiwifruit but thus far no phytotoxic activity has been detected.