Quantum pathogen evolution by integron-mediated effector capture

Joseph D. Payne

A thesis submitted in partial fulfilment of the requirements of the University of the West of England, Bristol for the degree of Doctor of Philosophy.

Faculty of Health and Applied Sciences, University of the West of England, Coldharbour Lane, Bristol, BS16 1QY.

October 2017

This copy has been supplied with the understanding that it is copyright material and that no quotation from this thesis may be published without proper acknowledgement.

Abstract

Plant pathogenic Pseudomonads are responsible for the loss of millions of pounds in crop revenue each year. They export effector molecules via the type three secretion system into the plants' cells in order to elicit disease. If the plant has the corresponding resistance genes to detect the type three effector molecule then the plant will mount an immune response called the plant hypersensitive response (HR). Type three effector molecules can also supress the plants' immune response including pathogen associated molecular pattern triggered immunity and effector triggered immunity.

Pseudomonads can evade HR by potentially gaining different effector molecules using mobile DNA elements. Integrons are one such type of element. Integrons are elements that allow bacteria to acquire and store genes from the environment particularly during times of stress. They also allow differential expression of the captured genes dependent on the environmental conditions.

Integron-like elements (ILEs) within *Pseudomonas syringae* pathovars and other Pseudomonads can be identified by using conserved genes such as the *xerC* integrase and the UV damage repair gene *rulB. RulB* encodes a DNA polymerase V which appears to be a hotspot for ILE insertion. Using the *rulAB* operon, the *xerC* gene and the ILE insertion junction, *rulB-xerC,* it was possible to identify a number of ILEs. The screening of 164 plant pathogenic *Pseudomonas* strains revealed new ILEs from 21 strains all containing at least one type three effector molecule. The screening also revealed that the *xerC* integrase was conserved across multiple ILEs within plant pathogens.

Expression studies of the ILE integrase genes, type three effector genes and the disrupted *rulB* gene showed that the genes on both ILEs present in *P. syringae* pv. *pisi* 203 and pv. *syringae* 3023 are upregulated in times of cellular stress and DNA damage. This led to the conclusion that ILEs may be more active when the bacteria was in need of exogenous genes to overcome the cellular stress. The ILE may also be excised following DNA damage to restore full *rulB* functionality.

It was identified that *rulB* was a hotspot for ILE insertion but it was not known why the ILEs choose this site or if any other genes were required for ILE insertion. Cloned versions of the *rulAB* operon from the pWW0 plasmid found in *Pseudomonas putida* PaW340 showed that only *rulAB* was required for *P. fluorescens* ILE insertion but *rulAB* must be intact. *P. syringae* ILEs were also tested but did not show any insertion.

Due to ILEs inserting into and disrupting *rulB* their effect on UV tolerance was tested. A range of strains containing an intact *rulB* gene were tested alongside the ILE containing strains with increasing amounts of UVB irradiation applied. The results showed very minor differences in growth rates between the two groups with only 60 seconds UVB exposure causing a significant difference in growth rate at the 95% confidence interval between the two groups of strains.

This research has contributed to the understanding of ILEs in phytopathogenic bacteria. It has also increased our understanding of the mechanisms of ILE gene expression, the mechanism surrounding ILE excision and insertion and the effect of ILEs on bacterial growth in high UV environments.

Acknowledgements

The past four years at UWE have encompassed some of the best times in my life so far along with some of the hardest. I have thoroughly enjoyed my time here and have learnt a lot about my subject area and myself. I have many friends, colleagues and family members to thank for making this a great experience.

Firstly I would like to thank my amazing family for always being there for me. In particular my mother and father for all their love, support and hard work during my studies. I would also like to thank my sister, Victoria, for always being there for a moan, gossip or just a simple chat. Finally a special thank you to my grandparents and great uncle for all their help.

The most important person throughout my PhD was Professor Dawn Arnold. Without Dawn all of this work would not have been achieved. Dawn was a fantastic supervisor and was always pushing me forward and developing me as a researcher. Her years of experience are invaluable and helped me to generate new ideas for my research. Dawn was not only my supervisor but she also became a great friend, I hope to continue working with her and remain in contact for many years.

I also need to thank Dr Helen Neale and Dr Carrie Brady for all of their help and support in the lab which really aided me in my work. Without Helen and Carrie lab work would have been much slower and nowhere near as fun!

My other supervisors, Professor John Hancock and Professor Robert Jackson were a real help in terms of new ideas and developing my final thesis. Their insight and experience was really appreciated and without them the process would not have been as smooth. I also want to thank all of the technical team at UWE especially Alison Halliday and Paul Deane. I also need to thank Professor Roger Pickup from Lancaster University and Dr Glenn Rhodes from the Centre for Ecology and Hydrology for their continued help and support.

A final thank you needs to go to all of my great friends who have really helped me forget about work at stressful times and helped me relax. Two friends deserve a really special thank you. Becky and Kieran without your help and friendship I would not have been able to complete my PhD.

Contents **Page**

List of Tables

List of Figures

Abbreviations and units

µF	microfarad(s)
µg	microgram(s)
µL	microliter(s)
µm	micrometre(s)
µM	micromole(s)
ADP	adenosine diphosphate
Ap.	appendix
ATP	adenosine triphosphate
AWF	apoplastic wash fluid
bp	base pair(s)
CFU	colony forming unit(s)
cm	centimetre(s)
cv	cultivar
CW	Canadian Wonder (bean)
DNA	deoxyribonucleic acid
DNase I	deoxyribonuclease I
dNTP	deoxynucleotide triphosphate
EDTA	ethylenediaminetetraacetic acid
eg.	exempli gratia (for example)
et al.	et alia (and others)
ETI	effector triggered immunity
g	gram(s)
GI	genomic island
GInts	genomic island with three integrases
Gm	gentamycin
HR	hypersensitive response
hr(s)	hour(s)
hrc	hypersensitive response and conserved
hrp	hypersensitive response and pathogenicity
HSB	high stringency buffer (see Ap. II)
IC	Indigo carmine
IPTG	Isopropyl β-D-1-thiogalactopyranoside
kb	kilo base(s)
KB	Kings B media (see Ap. II)
kg	kilogram(s)
Km	kanamycin (see Ap. II)
KW	Kelvedon Wonder (Pea)
L	litre(s)
LB	Luria-Bertani media (see Ap. II)
LSB	low stringency buffer (see Ap. II)
M	Mole(s)
MDH	malate dehydrogenase
mg	milligram(s)
MGE(s)	mobile genetic element

min(s)	minute(s)
mL	millilitre(s)
mM	millimolar
mm	millimetre(s)
MM	M9 minimal media (see Ap. II)
NA	nutrient agar (see Ap. II)
NB-LRR	nucleotide binding leucine rich repeat
ng	nanogram(s)
^{o}C	degrees Celsius
OD	optical density
ORF(s)	open reading frame(s)
PCR	polymerase chain reaction
PAMP(s)	pathogen associated molecular pattern(s)
PTI	PAMP triggered immunity
PRR	pattern recognition receptor
pv.	pathovar(s)
RE	restriction enzyme
RM	Red Mexican (bean)
RNase	ribonuclease
rpm	revolutions per minute
s/sec	second(s)
SD	standard deviation
SDS	sodium dodecyl sulphate
SEM	standard error of the mean
SSC	saline-sodium citrate buffer (see Ap. II)
Stm	streptomycin (see Ap. II)
TAE	Tris base, acetic acid and EDTA (see Ap. II)
TE	Tris base and EDTA (see Ap. II)
TG	Tendergreen (bean)
U	Units
UV	ultraviolet
V	volt(s)
w/v	weight per volume
w/w	weight per weight
xg	times gravity
x-Gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside
Ω	ohms

Chapter 1. Introduction

1.1: Plant pathogenic bacteria.

Plant pathogens have been responsible for severe crop losses throughout history and often result in human loss. The most famous human loss due to crop failure was the mass starvation in Ireland during the potato famine of 1845-49 caused by *Phytophthora infestans* (Maloy, 2005).

Plant pathogenic bacteria have a tremendous impact on a wide variety of agricultural crops ranging from fruit trees to garden ornamentals (Stavrinides, 2009) and are responsible for the loss of millions of pounds in crop revenue due to crop damage each year. Understanding how these pathogenic bacteria interact and evolve with their host plant species is of vital importance, not only to the global agricultural market but also to smaller individual farmers, suppliers and processers. Understanding this relationship between pathogenic bacteria and their host is paramount to establishing effective and efficient disease management programs around the world (Fry, 1982; Maloy, 2005).

Bacterial plant pathogens are largely confined to the Gram-negative Proteobacteria. The most observed pathogens are represented in either the α, β or γ subclasses (Table 1.1). *Agrobacterium tumefaciens* belongs to the α-subclass and is the causative agent of crown gall in over 140 plant species (Escobar and Dandekar, 2003). Within the β-subclass is *Ralstonia solanacearum* which is responsible for bacteria wilts in plants, by colonising the xylem. Erwin Frink Smith (1905) proved that bacterial wilts of tomato, pepper, eggplant and Irish potatoes were caused by *Ralstonia solanacearum* (Li *et al.,* 2005). Finally the γ-subclass contains the largest number of plant pathogenic bacteria which includes 28 families (Williams *et al.,* 2010), including the *Pseudomonas* genus

(Stavrinides, 2009). *Pseudomonas syringae* pathovars (pv.) cause many plant diseases including bacterial speck on tomatoes and halo blight on beans (*P. syringae* will be discussed further in Section 1.3). In order to better prevent plant disease and crop damage it is essential to understand how plant pathogenic bacteria behave and interact on a molecular basis.

Table 1.1: Representatives of plant pathogenic bacteria from the three subclasses of Proteobacteria, alpha (α), beta (β) and gamma (γ). (adapted from Garnier *et al.* 2000; Coenye and Vanadamme, 2003; Stavrinides, 2009; Madigan *et al.,* 2012).

Subclass of Proteobacteria	Example of Bacterial Species	Example Disease	Example Host
α	*Agrobacterium tumefaciens*	Crown Gall	Rhubarb, Walnuts
	Rhizobium leguminosarum	Root Nodules	Legumes
β	*Burkholderia cenocepacia*	Sour skin	Onions
	Ralstonia solanacearum	Bacterial Wilt	Tomato, Peppers
γ	*Pseudomonas syringae*	Bacterial Speck	Broad Range
	Xanthomonas campestris	Leaf Spot	Cabbage

1.2: Pseudomonads.

The Pseudomonads are a diverse genera containing over one hundred species, many of which are pathogenic species on either plants or animals (Özen and Ussery, 2012). Pseudomonads have remarkable metabolic and physiologic variability enabling them to colonise a diverse range of habitats and hosts ranging from soil and plant environments to human and aquatic environments (Palleroni, 1992; Silby *et al.,* 2011). Due to their variability Pseudomonads are responsible for many plant and human diseases across the globe (Table 1.2). *Pseudomonas* infections in humans can be fatal if not treated quickly and with targeted antimicrobial treatment. For example, *P. aeruginosa* can quickly colonise patients

who are in hospital for prolonged periods and can lead to meningitis, pneumonia, and septicaemia (Pai *et al.,* 2016; Bodey *et al.,* 1983).

Table 1.2: Host and habitats of four different *Pseudomonas* species (adapted from Silby *et al.,* 2011).

Pseudomonad	**Host**	**Habitat**
P. syringae	Plants (tomato, bean, olive, tobacco, chestnut, soybean)	Soil
P. fluorescens	Mostly non-pathogenic, Plants	Soil, Water
P. putida	Mostly non-pathogenic, Plants	Soil, Plants
P. aeruginosa	Animals, Humans, Plants (Arabidopsis thaliana [Thale cress] and Lactuca sativa [lettuce])	Soil, Water, Skin Flora and most man-made environments

Pseudomonas bacteria are Gram negative and are generally aerobic with a few exceptions being denitrifying (Palleroni, 1984). They can be either straight or curved rods that are between 0.5-1 x 1.5-4 μm in size and are non-spore forming with one or multiple polar flagella that assist in the bacteria's motility (Madigan *et al.,* 2012).

Phytopathogenic Pseudomonads have the ability to cause many different plant diseases with various symptoms. These symptoms include cankers, blossom, kernel, leaf or twig blight, dieback, leaf spots, soft or brown rot, galls and mushroom blights (Schaad *et al.,* 2001). Many of these bacterial species are either foliar epiphytes or rhizosphere inhabitants. The most economically important phytopathogenic *Pseudomonas* is *P. syringae* with over 50 pathovars (Berge *et al.,* 2014).

1.3: *Pseudomonas syringae.*

P. syringae was first isolated from a diseased lilac in 1902 (Hirano and Upper, 2000). *P. syringae* is of interest due to its importance as a plant pathogenic bacterium and was ranked the number one plant pathogenic bacteria by Mansfield *et al.* (2012) based on agricultural and economic impact. *P. syringae* is particularly important to plant pathogenesis studies due to the species having a wide host range that includes many commercially important crops (eg. tobacco and soybean) and its ability to cause damage. *P. syringae* typically attacks plant foliage causing the onset of chlorosis and eventually necrotic lesions on leaves (Madigan *et al.,* 2012). *P. syringae* causes a wide variety of plant symptoms ranging from bacterial speck, fleck, cankers, halo blight, galls and brown spot (Figure 1.1). For example, *P. syringae* pv. *syringae* strain B728A causes brown spot disease of bean leaves (Silby *et al.,* 2011), whereas *P. syringae* pv. *pisi* (*Ppi*) causes water-soaked lesions on pea plants (Suzuki and Takikawa, 2004).

The *P. syringae* species contains many different pathovars. The term pathovar relates to a group of bacterial strains which exhibit distinctive pathogenicity towards one or more hosts and can include different symptoms. The term pathovar is very useful when referring to relatedness between strains due to their phenotypic properties, but cannot reveal how related two strains are based on their genetic properties (Denny *et al.,* 1988). Pathovars can be further divided into different races (Young, 2008). Races are identified through the use of plant host differentials.

The occurrence of *P. syringae* pathovars is on the increase with a resurgence of old diseases such as bacterial speck on tomatoes caused by *P. syringae* pv. *tomato* (*Pto*) (Shenge *et al.,* 2007) and also the emergence of new

infections such as bleeding canker of horse-chestnut caused by *P. syringae* pv. *aesculi* (Green *et al.,* 2010).

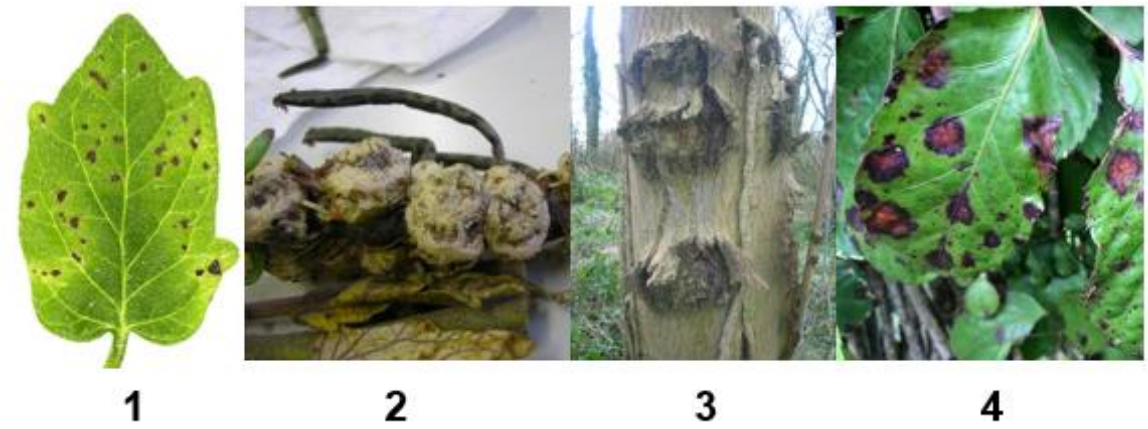


Figure 1.1 Symptoms caused by *Pseudomonas syringae.* 1) Bacterial speck of tomato leaves (*Solanum lycopersicum*) caused by *P. syringae* pv. *tomato*. **2**) Crown galls caused by *P. syringae* pathovars. **3**) Bleeding canker of horse chestnut tree (*Aesculus hippocastanum*) caused by *P. syringae* pv. *aesculi*. **4**) Bacterial blight caused by multiple *P. syringae* pathovars. (Images adapted from commons.wikimedia.org and used under the creative commons licence 3.0; https://creativecommons.org/licenses/by-sa/3.0/ or from flickr.com and used under the creative commons licence 2.0;

It is the many different pathovars of *P. syringae* that allow the species to infect a variety of host organisms. Strains within most of the pathovars have narrow host ranges with the exception being *P. syringae* pv. *syringae* which has more than 80 plant hosts listed (Bradbury, 1986). There has recently been genomic (Multi Locus Sequence Typing (MLST)) and phenotypic analysis of 216 strains of *P. syringae* which identified more than 50 different pathovars of *P. syringae* (Berge *et al.,* 2014). Each pathovar may only have a narrow host range, but due to the high diversity of pathovars multiple plant species can be infected by *P. syringae* leading to its high pathogenicity status (Hirano and Upper, 2000).

1.4: Pathogenicity factors.

P. syringae produce a range of proteins to facilitate plant colonisation including enzymes to degrade the plant cell walls, proteins to allow adherence to the plant and other enzymes such as proteases, lipases and haemolysins (Preston *et al.,* 2005). Many of these secreted molecules are commonly secreted by the type one (proteases, nucleases), two (virulence factors and toxins such as phospholipases) or three (effector molecules) secretion systems and these molecules promote virulence and enhanced fitness during colonisation. There are also type four, five and six secretion systems which facilitate translocation and autotransportation of proteins and single-stranded DNA (Figure 1.2) (Jackson, 2009).

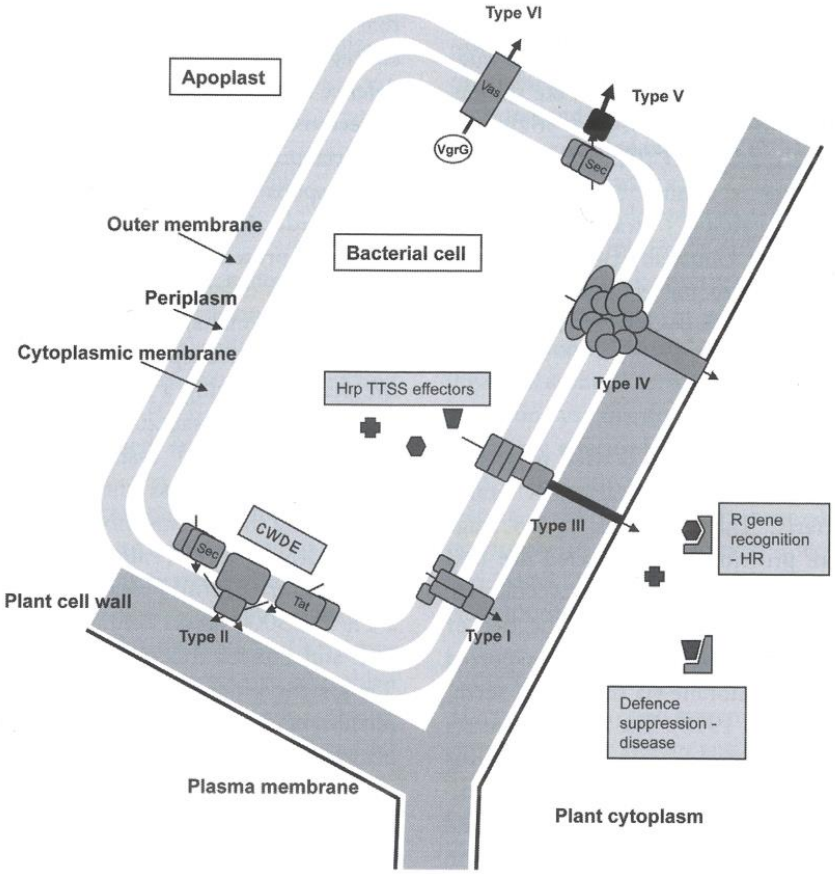


Figure 1.2: Diagram showing the various secretion systems used by *P. syringae.* Types one to three are used for secreting factors that promote virulence and increase fitness, such as type three effectors. CWDE - Cell wall degrading enzymes. (Image from Jackson (2009) and used with permission from Caister Academic Press; http://www.caister.com/copyright).

P. syringae produces four primary toxins: coronatine; phaseolotoxin; syringomycin; and tabtoxin. It is probable that many toxins cause a change in plant metabolism as all of these toxins promote chlorosis (Bender *et al.,* 1999). Coronatine is a molecule that mimics methyl-jasmonate, a key host signalling molecule within the plant, therefore disrupting normal signalling patterns (Bender *et al.,* 1999). Phaseolotoxin has been shown to promote bacterial growth and bacterial spread once inside the plant by altering the membrane permeability to allow sugars and other organic compounds to move more freely (Hutchison *et al.,* 1995). Phaseolotoxin has also been shown to disrupt the urea cycle causing arginine deficiencies within the plant (Hwang *et al.,* 2005). Phaseolotoxin is secreted via the type one secretion system via an oligopeptide permease (Staskawicz and Panopoulos, 1980). Syringomycin is also secreted by the type one secretion system. Syringomycin is a pore-forming toxin which causes electrolytes to leave the host cell and alters the permeability of the membrane allowing sugars and organic compounds to 'leak' out. These compounds favour bacterial growth (Rico *et al.,* 2009). Tabtoxin inhibits glutamine synthetase which leads to ammonia accumulation and visible chlorosis (Turner, 1989).

Not all *P. syringae* strains are pathogenic. When the link between frost injury in plants and ice nucleating bacteria was discovered *P. syringae* was the most frequently found bacteria (Lindow *et al.,* 1978), but the plants did not always have disease symptoms depending on the pathovar and the host plant. Non-pathogenic isolates have also been identified in the field. *Pseudomonas syringae* pv. *syringae* (*Psy*) 508 was isolated from a fallen apple tree leaf and was unable to cause disease on any of the plants tested (Mohr *et al.,* 2008). It is thought that this isolate belongs to a monophyletic group that evolved from a pathogenic *P. syringae* strain by losing its ability to cause disease (Mohr *et al.,* 2008).

1.5: Plant immunity and pathogen recognition.

Plants react to a bacterial infection using a two pronged immune approach. The first approach is responsible for detecting common molecules that many bacterial species express (Jones and Dangl, 2006). This includes both pathogens and non-pathogens. For pathogens this recognition occurs on the external surface of the plant cells before the bacteria have entered the cell. These common microbial molecules, called Pathogen Associated Molecular Patterns (PAMPs), are recognised on the cells surfaces by Pattern Recognition Receptors (PRRs) and elicit PAMP-Triggered Immunity (PTI) (Jones and Dangl, 2006). This causes the deposition of callose to the cell wall to act as a physical barrier (Nicaise *et al.,* 2009), stomatal closure to prevent entry, restriction of nutrient transfer from the cytosol to the apoplast to limit bacterial growth and the production of antimicrobials (Bigeard *et al.,* 2015).

The second immune response the plant has is Effector Triggered Immunity (ETI). This occurs in response to the pathogen exuding virulence proteins, called effectors, into the plant cell. ETI is commonly induced by nucleotide-binding leucine-rich repeat (NB-LRR) receptors recognising effector molecules either directly or indirectly through their effects on host targets once they are inside the plant cell (Dodds and Rathjen, 2010), (Figure 1.3). ETI results in the induction of the hypersensitive response (HR). The HR is characterised by localised cell death at the infection site. The cell death is triggered by gene-for-gene resistance in plants caused by invading pathogens carrying effector proteins encoded by certain avirulence (*avr*) genes. The Avr proteins are recognised by corresponding R proteins in the plant (Erbs and Newman, 2009). If the plant lacks the corresponding R protein no HR will occur, facilitating bacterial proliferation.

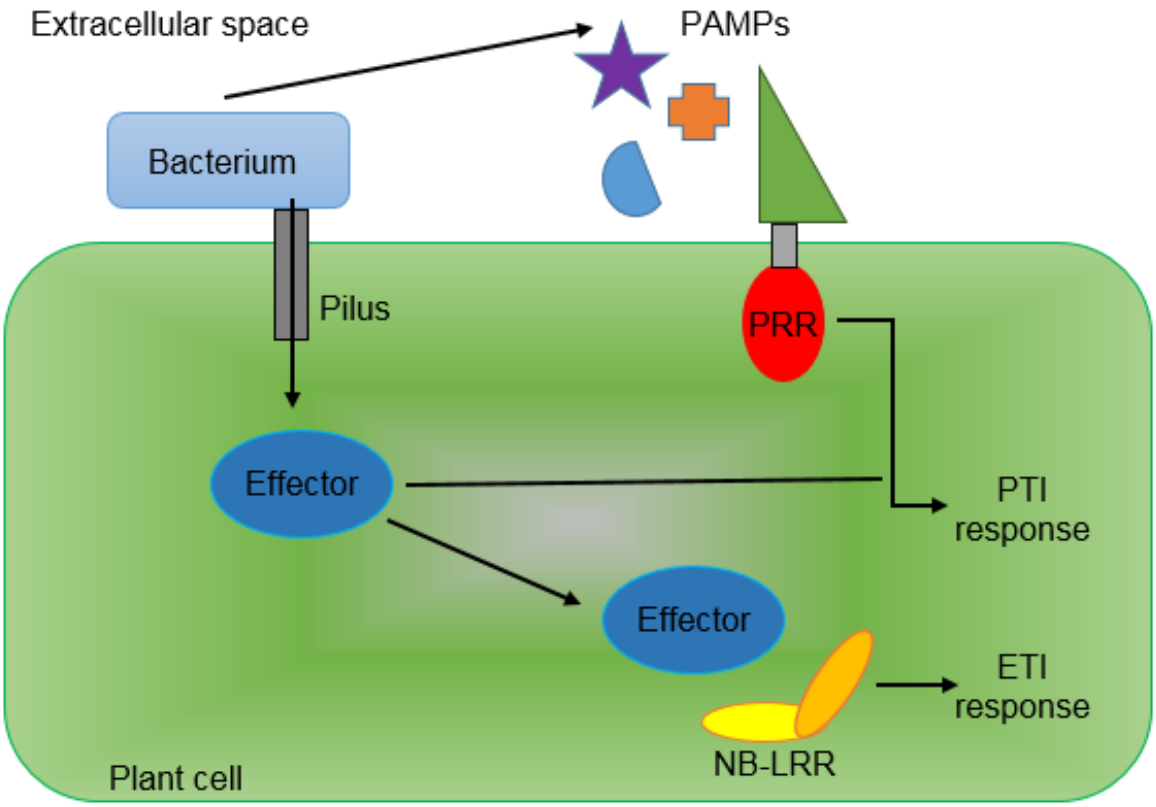


Figure 1.3 Outline of Plant Immune System. Bacterial plant pathogens propagate exclusively in the extracellular spaces of plant tissues. Molecules released from the pathogens into the extracellular spaces, such as lipopolysaccharides, flagellin and chitin (PAMPs) are recognized by cell surface Pattern-recognition receptors (PRRs) and elicit PAMP triggered immunity (PTI). PRRs generally consist of an extracellular leucine-rich repeat (LRR) domain (green), and an intracellular kinase domain (red). When a PAMP is recognised the PTI signalling pathway is triggered. Bacterial pathogens deliver effector proteins into the host cell by a type three secretion pilus. These intracellular effectors often act to suppress PTI. However, many are recognized by intracellular nucleotide-binding (NB)-LRR receptors, which induces effector triggered immunity (ETI).

Many plant-pathogenic bacteria secrete a large number of different effector proteins into host cells to increase the chances of successful infection and evasion of ETI (Zhou and Chai, 2008). It is the constant battle between plant and pathogen that leads to the evolution of new effector molecules and new receptors to recognise them, the problem is that bacteria can evolve much quicker than the plant. In 2006 Jones and Dangl theorised the ‘zig-zag’ model for plant-pathogen interaction and co-evolution (Figure 1.4). Effector DNA can be re-

shuffled which leads to the NB-LRR receptor no longer being able to recognise the effector preventing ETI.

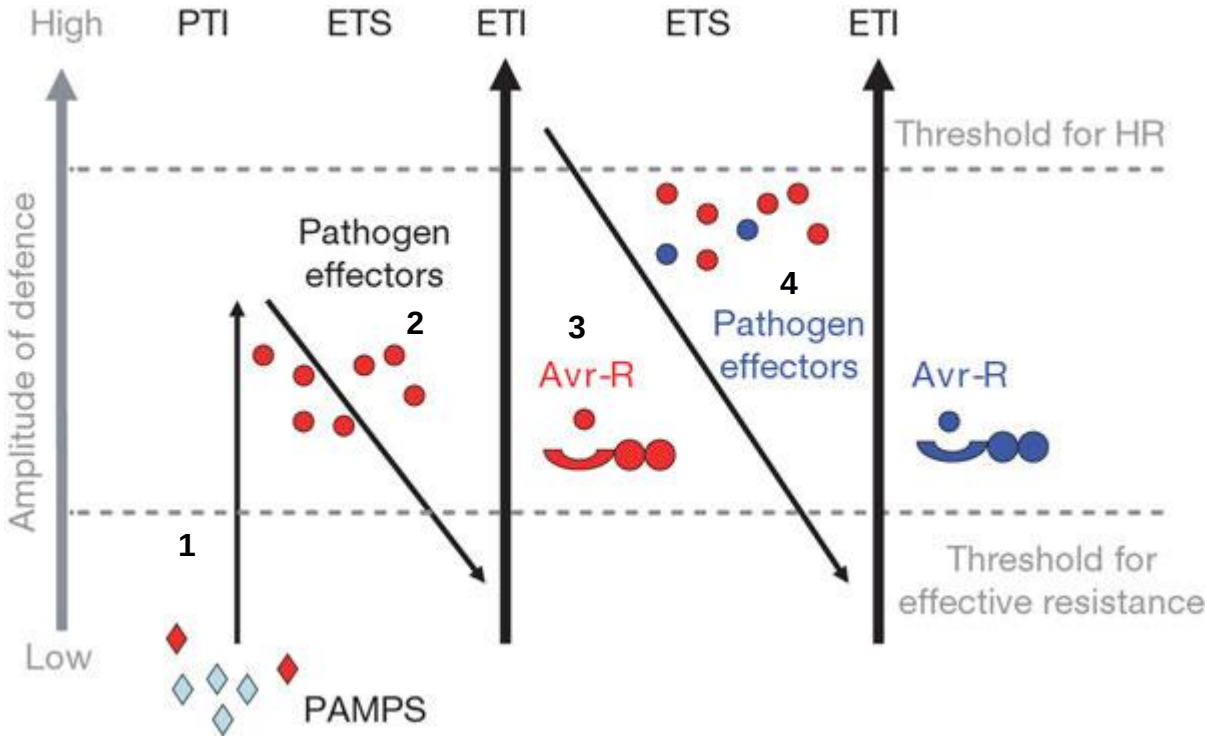


Figure 1.4 Outline Model of Plant Immune Response. In section **1**, plants detect PAMPs (red diamonds) via PRRs to trigger PTI. In section **2**, successful pathogens that evaded PTI deliver effectors that interfere with PTI, or otherwise enable pathogen nutrition and dispersal, resulting in effector-triggered susceptibility (ETS). In section **3**, one effector (indicated in red) is recognized by an NB-LRR protein, activating ETI, an amplified version of PTI that often passes a threshold for induction of hypersensitive cell death (HR). In section **4**, pathogen isolates are selected that have lost the red effector, and perhaps gained new effectors through horizontal gene flow (in blue). These effectors can help pathogens to suppress ETI. Natural selection favours both new pathogen effectors and new plant NB-LRR alleles that can recognize one of the newly acquired effectors, resulting again in ETI. The cycle continues with new effectors and new NB-LRR receptors co-evolving. (Image adapted from Jones and Dangl (2006) and used under Agreement with Nature Publishing Group no. 4052991072647).

1.5.1: Pathogen Associated Molecular Patterns.

Innate immunity consists of molecules designed to prevent pathogen growth and facilitate removal of the pathogen from the tissue. Plants only have innate immunity as a line of defence against infection, unlike animals (van Baarlen *et al.,* 2007). Animals also have innate immunity which is split into two parts. The first, humoral innate immunity, involves a variety of substances found in bodily fluids which interfere with the growth of pathogens. The second innate response is called cellular innate immunity which is carried out by cells called phagocytes that degrade pathogens. The difference between plants and animals is that animals (vertebrates) also have adaptive immunity that can recognise certain substances via antigens on the pathogen and remove the target. This response also allows the host to remember the pathogen and deliver a quicker response if the pathogen infects the host again. Innate immunity is nonspecific; it is not directed against specific invaders, but against any pathogens. Whereas animal adaptive immunity can recognize and destroy specific substances (Nürnberger *et al.,* 2004).

PAMPs trigger the innate immune response and are molecules that are presented on the surface of the pathogen. The plant recognises these molecules before the pathogen has invaded the cell. The four most studied bacterial elicitors that act as PAMPs are flagellin (Flg), elongation factor Tu (EF-Tu), lipopolysaccharides (LPS) and cold shock protein (CSP) (Jackson, 2009). These molecules all initiate the plant's immune system (Erbs and Newman, 2009). Recognition of PAMPs induces plant defence systems that may include: oxidative burst - the rapid release of reactive oxygen species which degrade bacteria, nitric oxide generation – secreted as free radicals in an immune response and is toxic to bacteria (Hausladen and Stamler, 1998), cell wall strengthening – acts as

physical barrier to prevent effector molecules entering the cell, pathogenesis-related (PR) protein accumulation – some are antimicrobial and attack the bacteria whereas others send signals to other areas of the plant (Ebrahim *et al.,* 2011).

The most defined PAMP is flagellin. Flagellin is a globular protein that arranges itself as a hollow cylinder to form the filament in bacterial flagellum. The role of flagellin is essential to pathogens for mobility and increased adhesion of the bacterial cell to the cell wall of the plant. Flagellin is a major activator of PTI as the majority of plant pathogens have flagella containing flagellin which makes it the ideal molecule for the immune system to detect (Erbs and Newman, 2009). The PTI response is not triggered by the entire flagellin molecule but rather a highly conserved domain at the N-terminal, a 22 amino acid peptide named flg22 (Felix *et al.,* 1999). Flagellin has its own unique PRR that allows the plant to trigger an immune response. The PRR was discovered by Gómez-Gómez and Boller (2002) and named flagellin sensing 2 (FLS2) due to it recognising the conserved N-terminal domain of flagellin.

1.6: Effectors.

P. syringae uses many virulence associated systems to infect its host. The most studied is the type three secretion system (TTSS) which exudes an array of effector proteins into the host primarily to suppress plant immune systems (Arnold *et al.,* 2009). Effector proteins are delivered into the plant's cytoplasm by the TTSS (see section 1.7). Effectors disrupt the plants' cellular and signalling pathways to prevent defence mechanisms being triggered such as PTI, ETI and HR (Arnold *et al.,* 2009). Effectors used to be split into two categories, avirulent and virulent depending on their ability to cause disease on a selected plant. This

terminology was confusing though as the same protein can function as virulent or avirulent depending on host, therefore the term effector is now used. The effectors that are specifically recognized by 'matching' resistance proteins (termed R proteins) are termed avirulence (AVR) proteins (Rouxel and Balesdent, 2010). This formed the basis of the 'gene-for-gene' concept which shows that a plant encoding an R gene specific to an effector is resistant to the pathogen that produces it (Keen, 1990; Crute, 1994).

The first effector protein to be discovered was avirulence gene A (*avrA*) in *P. syringae* pv. *glycinea* by Staskawicz *et al.* (1984). In 2000 a review by Vivian and Arnold (2000) estimated that 30 effector genes had been identified in different *P. syringae* pathovars using 'gain-of-function' assays (Cunnac *et al.,* 2009). Effector genes can also be identified via PCR as the DNA sequences flanking effector genes in *Ppi* show high degrees of similarity and primers could be designed from these conserved regions (Arnold *et al.,* 2001).

Effector proteins have many functions, one function is to suppress host defences allowing the pathogen to infect and spread inside the plant. The first type three secreted effector (TTSE) shown to suppress basal defence was AvrPto from *Pto* DC3000 (Nomura *et al.,* 2005). AvrPto prevents the deposition of callose to the cell wall in response to the TTSE. The prevention of callose deposition means that the bacteria can invade the plant more easily. This is because callose-containing cell-wall appositions act as effective physical barriers to prevent bacterial invasion. These appositions are induced at the sites of attack during the relatively early stages of pathogen invasion (Luna *et al.,* 2010). Effector proteins can also suppress advanced plant defences including gene-for-gene and HR resistance. Effectors can also be enzymes, AvrRpt2 is a cysteine protease that can cleave RIN4, a molecule needed for HR signal transduction. RIN4 is broken

down so cannot activate the HR (Luna *et al.,* 2010). These enzymes contain active sites of mono-ADP ribosyltranferases that cause RNA-binding proteins to change, therefore altering RNA metabolism and reducing the amount of immunity related mRNA's available favouring pathogen establishment (Fu *et al.,* 2007). Effectors can also block the recognition of other effectors by suppressing host genes responsible for receptor expression.

Finally effector proteins can alter host pathways. For example, de Torres-Zabala *et al.,* (2007) demonstrated that *Pto* effectors can 'hijack' the abscisic acid (ABA) pathway in *Arabidopsis thaliana* leading to disease.

1.7: The type three secretion system (TTSS).

The phytopathogenic bacteria never enter the plant cells' so a delivery mechanism is required to get the effector molecules into the cell. An essential part of bacterial pathogens is their ability to secrete proteins that facilitate infection and bacterial proliferation and survival. There are a total of six secretion systems that have either been identified or predicted to be utilised by *P. syringae* with the most emphasis being on the TTSS and the effector proteins it delivers (Arnold *et al.,* 2009). The TTSS allows infection to occur by directly 'injecting' effector proteins into the cytoplasm that can suppress plant immune responses and block signalling pathways (Preston *et al.,* 2005).

Once the bacteria have entered the extracellular space of their host environmental changes lead to the activation of a specific gene cluster *hrp-hrc,* responsible for encoding the TTSS (Jin *et al.,* 2003; Kvitko *et al.,* 2007). There is strict regulation of TTSS expression and it is only induced in plant tissue, apoplastic fluid and *hrp* inducing media. This results in the transcription and translation of multi-protein complexes which form a complex, supramolecular

structure with the distinctive syringe and needle-like TTSS (Figure 1.6) which has the ability to 'pierce' the host's cell wall exporting effector proteins (Gerlach and Hensal, 2007). Molecular chaperones, called harpins, are required to aid the transport of effector molecules from the pathogen into the host cell via the TTSS. Harpins are a subset of TTSS substrates found in all phytopathogenic bacteria that utilize a TTSS (Kvitko *et al.,* 2007).

In *P. syringae*, hypersensitive response and pathogenicity (*hrp*) and hypersensitive response conserved (*hrc*) genes encode the TTSS pathway (Alfano and Collmer, 1997), and avirulence (*avr*) and Hrp-dependent outer protein (*hop*) genes encode effector proteins (Schechter *et al.,* 2004). The *hrp/ hrc* genes are required for the development of the HR in non-host and resistant hosts and the onset of pathogenesis in susceptible plants. The *hrp/ hrc* gene cluster for the TTSS is on a pathogenicity island and is bounded by two effector loci, an exchangeable effector locus and a conversed effector locus (Alfano *et al.,* 2000). Typically in *P. syringae* strains the *hrp/ hrc* genes are clustered in a 25kb region organised into seven operons which can encode either regulatory, secretory or effector proteins. (Figure 1.5).

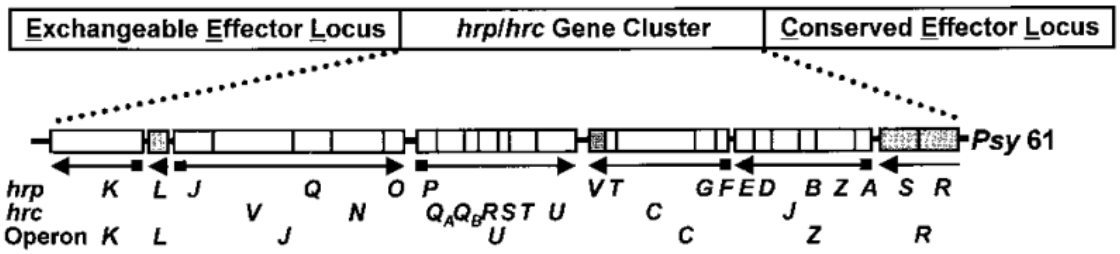


Figure 1.5: The arrangement of the conserved *hrp/ hrc* genes within the pathogenicity island bounded by the exchangeable effector locus and the conserved effector locus. The letters denote the gene names, eg. P is *hrp.* The arrows indicate the direction of transcription and the boxes indicate the Hrp box. (Image from Alfano *et al.,* 2000 with permission from PNAS, Copyright (2000) National Academy of Sciences).

Within the *hrp/ hrc* locus there are at least three classes of *hrp* genes. One class encodes both positive, HrpL, and negative, HrpV, regulatory proteins in *P. syringae,* these are responsible for the regulation of the TTSS associated genes (Ortiz-Martin *et al.,* 2010). Another *hrp* class encodes core structural components of the TTSS including genes that have a high similarity to flagellum assembly genes and proteins that are involved in the breakdown of peptidoglycan in the cell wall allowing the TTSS apparatus to form into the plant cell (Alfano and Collmer, 1997). The final class encodes secreted proteins.

This suggests that the PAI containing the TTSS in *P. syringae* was inherited via horizontal gene transfer (HGT) as it is also present in distantly related bacteria. It has also been suggested that acquisition of the PAI led in part to *P. syringae* becoming a phytopathogen (Mohr *et al.,* 2008).

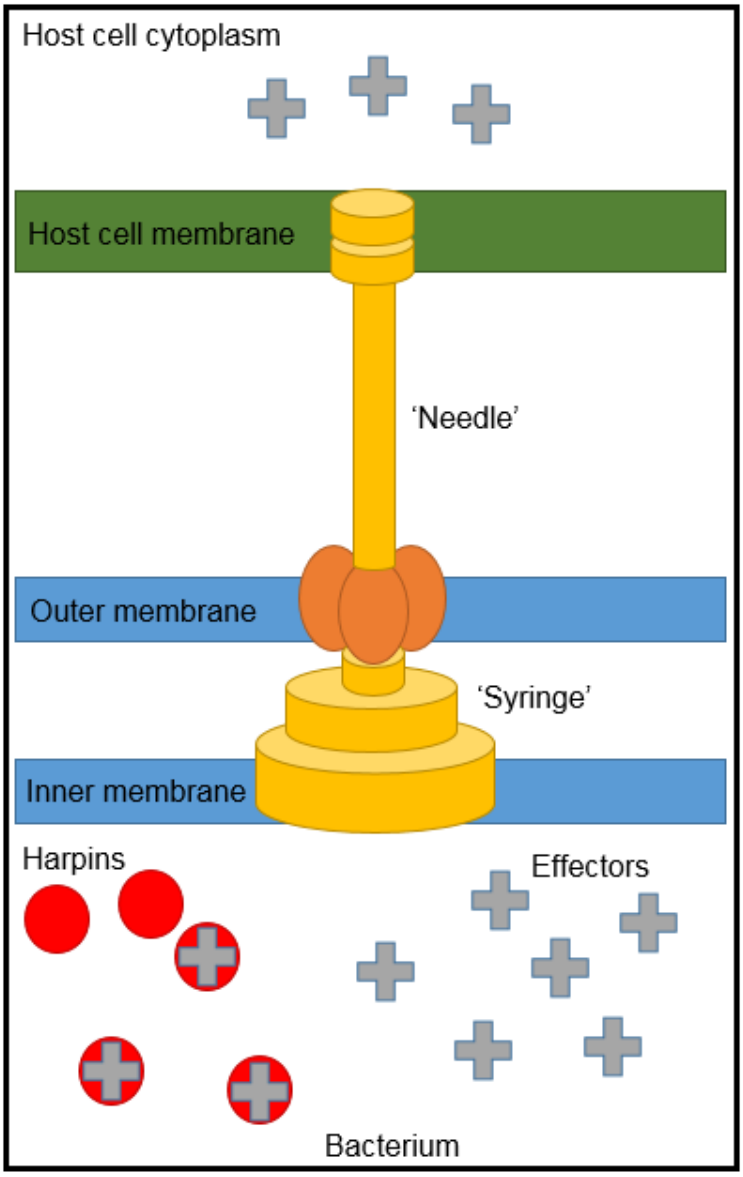


Figure 1.6: The type three secretion system in action, how effector proteins are exported into host cytoplasm. The TTSS forms a syringe like base which spans the inner and outer membranes of the bacteria. This base also contains a pore structure through which effectors enter the secretion system aided by harpin molecules. The effectors then travel along the needle which forms a physical connection between bacteria and host. The end of the needle (pilus) 'pierces' the host's cell wall allowing the effectors to enter straight into the cytoplasm.

1.8: Effector evolution.

Mutations in effector genes can result in evasion of host recognition. The mutation can be in a gene itself or the loss/gain of larger pieces of DNA. The antimicrobial environment created during the HR presents a strong selection pressure for mutants that can avoid triggering the HR. The effector could be lost by mutational insertions, deletions or rearrangements, even a single base pair

change can lead to a non-functional effector (Arnold *et al.,* 2009) (Figure 1.7). A classic example of this is the case of *avrPphE*. All 9 races of *P. syringae* pv. *phaseolicola (Pph)* have *avrPphE* but only races 2,4,5 and 7 have an active version. The remaining races have an inactivated version of *avrPphE* due to a single base pair change which confers an amino acid change (Stevens *et al.,* 1998). The function of *avrPphE* is not yet fully understood but it is a modular protein that acts as a virulence determinant (Nimchuk *et al.,* 2007).

Effector genes can also be disrupted by the insertion of mobile genetic elements (MGEs). The disruption can be caused by insertion sequences, transposons and integrons. As more genome sequences are being analysed more disrupted genes are being identified. The effector genes are disrupted in some strains but not others leading to a variation of functional effectors across different strains of the same species. This variation has been shown in the complete genome of *Pto* DC3000 which revealed a total of 31 effectors secreted by the TTSS (Buell *et al.,* 2003). When compared to the *Psy* B728A genome four effectors appeared to be disrupted by MGE insertion. This variation prompted Greenburg and Vinatzer (2003) to propose that different effector profiles are the reason why highly related phytopathogenic strains have different host range and disease characteristics depending on their effector repertoire.

Certain phytopathogenic strains, such as *Pto* DC3000, contain a lot more transposons than other strains. The high level of transposons could be driving a high level of DNA shuffling in these strains resulting in novel effectors that the plant cannot recognise. The DNA can also be shuffled so there is a TTSS signal included, allowing the protein to be transported into the plant whereas before it could not (Greenburg and Vinatzer, 2003) (Figure 1.7).

Effector genes can also be lost. The entire coding region could be lost during cell replication if the effector gene is carried on a MGE (Arnold *et al.,* 2007). This was shown when the genomic island (GI), PPHGI-1, was lost from *Pph* race 4 strain 1302A. The loss of the GI resulted in the loss of the effector gene, *avrPphB* (Pitman *et al*., 2005). The loss of the GI and subsequent effector causes the bacterial strain to become virulent on beans carrying the R3 resistance gene (Figure 1.7).

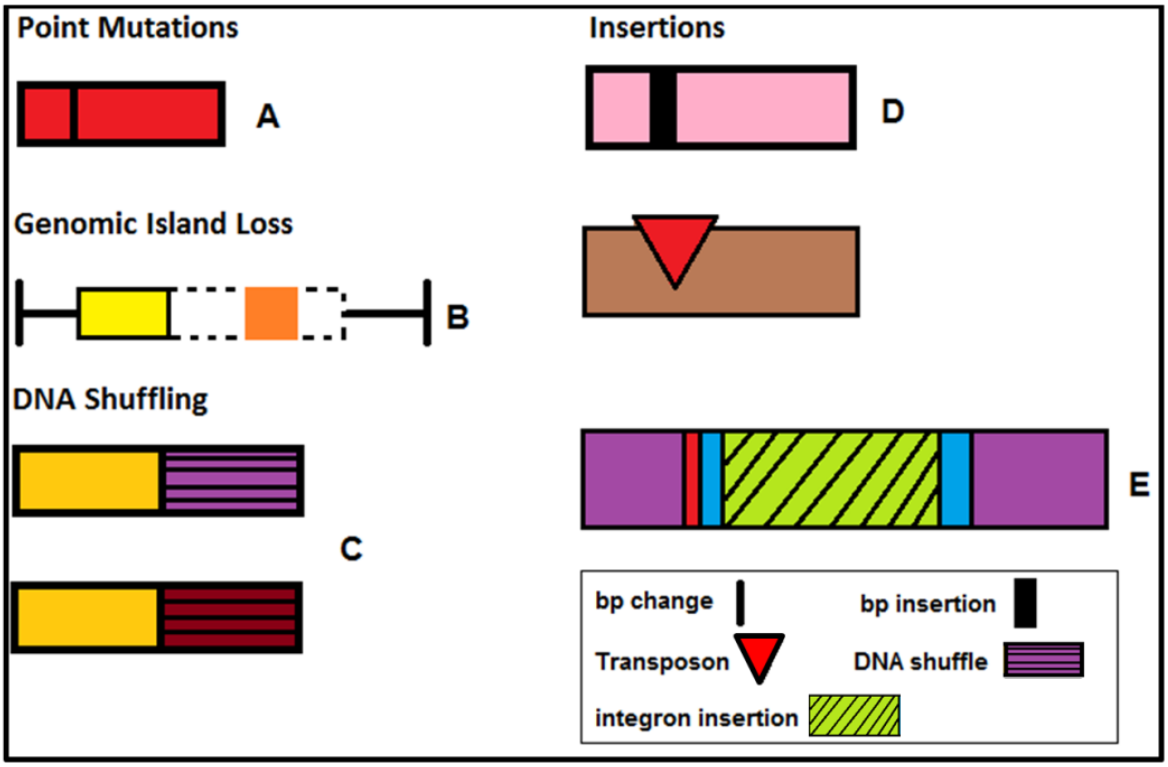


Figure 1.7 The different ways in which effector genes' functions can be changed, lost or suppressed. A) A simple base pair (bp) mutation can cause an effector to lose its avirulence (*avrPphE*). **B)** The loss of a genomic island (hashed lines) which contained an effector (orange box) the plant could detect, strain is now pathogenic on host. **C)** The shuffling of DNA regions, in this case the different ends of an effector gene leading to acquisition of TTSS signals. **D)** The insertion of base pairs (1) or transposons (2) leads to the gene no longer forming a correct and functional effector protein, meaning there is no effector for the plant to recognise. **E)** Integron insertion can also disrupt genes leading to dysfunctional proteins, but integrons can also pick-up and harbour different effector genes from totally different bacterial species, meaning the plant will not recognise the effector.

1.9: Integrons.

Integrons are sections of mobile DNA characterized by their ability to capture and incorporate gene cassettes by site-specific recombination. Integrons were first discovered in the late 1980's (Stoke and Hall, 1989) and were primarily thought to be unique to human clinical settings. However, integrons have now been isolated from bacteria in non-clinical environments including soil and water (Domingues *et al.,* 2012). Integrons have been identified as major determinants in a bacteria's ability to become resistant to multiple antibiotics by 'sharing' resistance genes between bacterial populations and species (Gillings *et al.,* 2014). This 'sharing' of genes is not limited to antibiotic resistance and includes genes that encode effector proteins responsible for plant disease (Arnold *et al.,* 2001). This sudden uptake of new genes could result in a quantum evolutionary jump for the bacteria. This evolution could result in the bacteria being able to potentially survive and infect new host organisms, and if successful the new strain will thrive.

Integrons have been identified in approximately 17% of bacterial genome sequences stored on the NCBI database (Gillings, 2014). Integrons are commonly linked to MGEs to allow mobility between chromosomes and plasmids (Domingues *et al.,* 2012). Integrons can move independent of MGEs but only as a result of captured genes providing and facilitating self-mobility mechanisms. The heritability of an integron via horizontal gene transfer (HGT), and therefore its spread throughout the bacterial population is reliant on two aspects, the first being the genetic stability of the integron and the second is the fitness cost to the new host.

Integrons fundamentally contain three main elements that work together to capture and express exogenous genes cassettes (Jackson *et al.,* 2011). The main elements are a tyrosine recombinase (*intI*) gene with its own promoter (P_{int}), an adjacent recombination site (*attI*) and a promoter (P_c) (Gillings, 2014) (Figure 1.8). The integrase protein encoded by *intI* belongs to the tyrosine recombinase family and is responsible for catalysing site-specific integration and excision of specific gene cassettes into the *attI* site. The *attI* site is a recombination site recognised by the integrase and the site where the gene cassette is inserted. The final element of an integron is the outward facing promoter (P_c) which is required for expressing captured gene cassettes (Gillings, 2014; Jackson *et al.,* 2011; Hall and Collis, 1995).

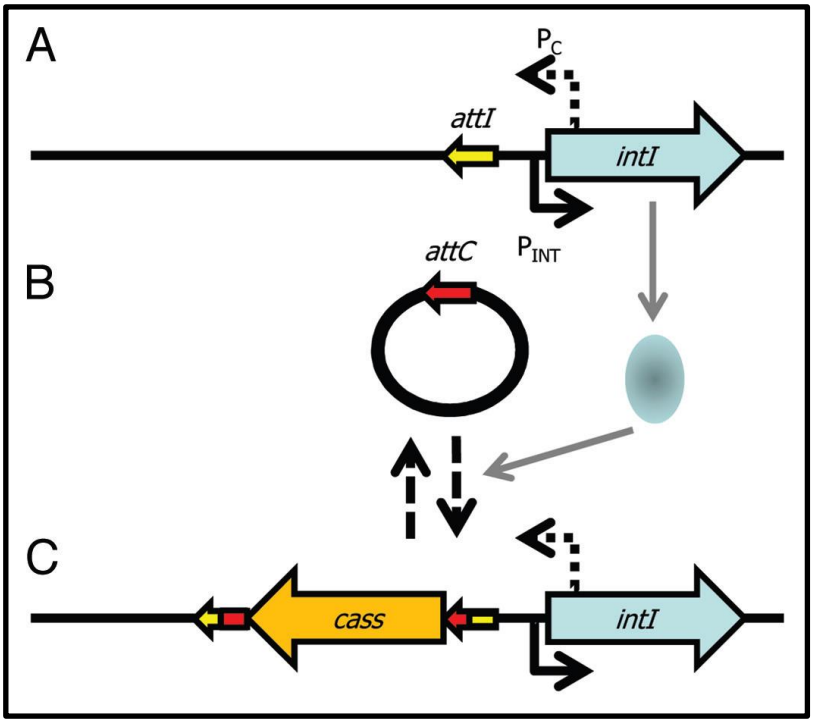


Figure 1.8 Capture of a gene cassette by an integron. A) The core integron assembly comprising of the integrase gene (*intI*), the integrase promoter (P_{int}), the gene cassette promoter (P_c) and the integration site (*attI*). **B)** The expression of integrase occurs during SOS response and catalyses site-specific recombination of circularised gene cassettes between the *attC* site on the cassette and the *attI* site of the integron. **C)** Cassettes (*cass*) are incorporated into the integron complex. More cassettes can be added and they can also be excised. (Image adapted from Jackson *et al.* (2011) used under the creative commons license 3.0; https://creativecommons.org/licenses/by-nc/3.0/).

Integron activity is regulated by the transcriptional repressor LexA and the SOS response upstream of the tyrosine recombinase gene (Guerin *et al.,* 2009). The integrase promoter (P_{int}) contains LexA binding sites that cause the expression of integrase to be down-regulated when LexA is bound. The SOS response causes the release and degradation of LexA from the promoter and therefore activates integrase expression allowing gene cassettes to be captured (Jackson *et al.,* 2011). Plasmid conjugation activates the SOS response meaning integrase is at its highest levels when new gene cassettes are potentially more common. Environmental stress and DNA damage causes the formation of ssDNA molecules which bind non-specifically to the universal recombination protein RecA which in turn promotes LexA inactivation by autoproteolytic cleavage, the inactivation of LexA induces the SOS response and activates the integron integrase (Figure 1.9) (Cambray et al., 2011). It is cell stress and damage that activates the SOS response and allows capture and expression of new genes that may benefit the bacterium in stressful environments.

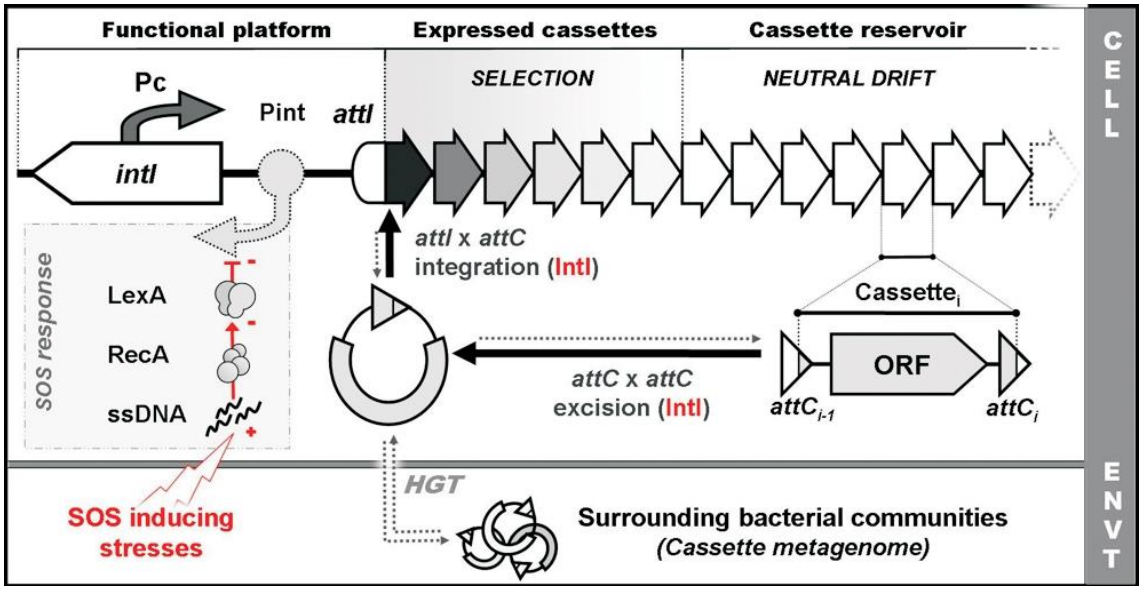


Figure 1.9: How integrons use integrase genes to capture novel gene cassettes and how they are regulated. The diagram shows how the integron integrase, *intI*, is regulated by LexA and how a SOS inducing stress causes the LexA repressor to be inactivated allowing cassette capture. (Image from Cambray *et al.,* (2011) used under creative commons license 2.0; creativecommons.org/licenses/by/2.0/).

1.9.1: Integron-like elements present in the plant pathogen *Pseudomonas syringae*.

Previous work had shown integrons to be abundant in bacterial species that had antibiotic resistance such as *Vibrio* species (Clark *et al.,* 2000). Integrons however are not exclusive to clinical isolates showing antibiotic resistance and integron-like elements have been identified in environmental isolates (Gillings *et al.,* 2008). Integron-like elements (ILEs) may be like integrons and be able to capture, express and help spread genes within bacterial populations. However the term integron-like element is used because the integrase genes are forward facing with respect to the ILE (unlike true integrons) and because in true integrons the integrase gene has its own forward facing promoter and a reverse facing promoter for expression of cassette genes in the variable end (Gillings *et al.,* 2008). ILEs in *P. syringae* and *P. fluorescens* have neither of these and may be under the control of LexA found upstream of *rulA* (Rhodes *et al.,* 2014). ILEs may also be more independently mobile that other integrons, such as class one integrons which require genetic linkage to other MGEs (Domingues *et al.,* 2012).

Prof. Dawn Arnold and her research group identified an integron-like element within the chromosome of the pea pathogen, *Ppi* race 2 strain 203 (Arnold *et al.,* 1999; 2000). An 8.5kb region of DNA was identified that was present in strain *Ppi* 203 but not in any other *Ppi* strains tested. The DNA was flanked by direct repeat sequences. The 8.5kb fragment contained a *rulAB* operon (responsible for UV tolerance and DNA repair) and the *rulB* gene was disrupted by a 4.3kb insertion of DNA which contained the avirulence gene, *avrPpiA1,* and genes with high similarity to transposase genes which may confer mobility (Figure 1.10) (Arnold *et al.,* 2000). The *rulAB* encodes the error-prone polymerase V that can synthesise DNA across lesions caused by UV irradiation

and genotoxic compounds, resulting in bacterial survival but at the cost of genetic mutations in the DNA (Stockwell *et al.,* 2013).

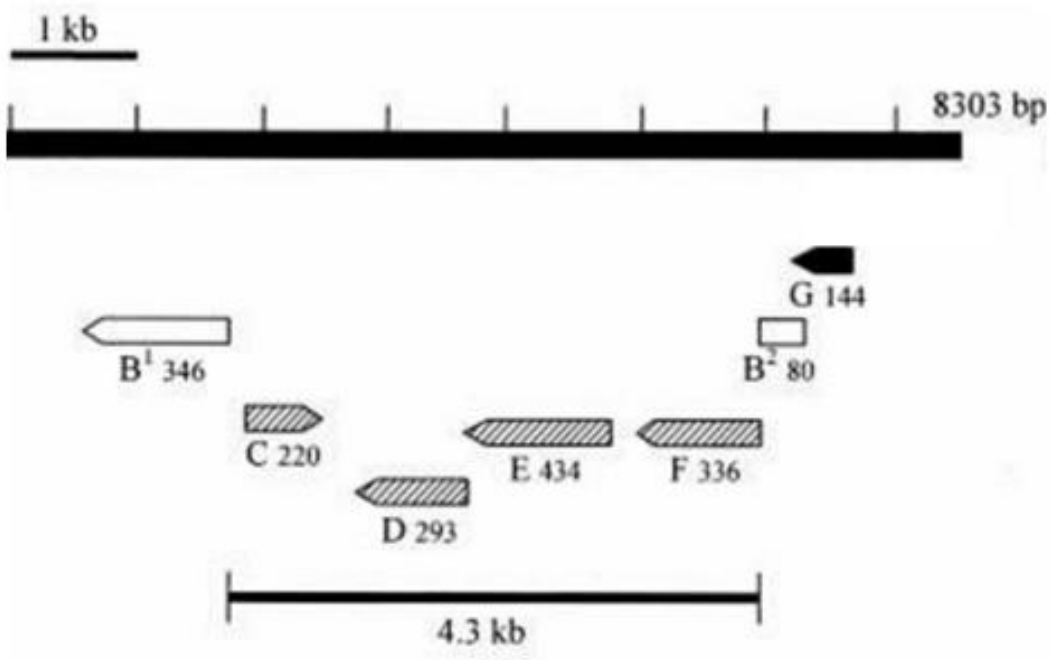


Figure 1.10: Identified ILE within *Pseudomonas syringae* pv. *pisi* 203. The 4.3kb ILE is present with a larger 8.3kb insertion, the ILE is inserted with the *rulB* gene (**B**^{1,2}). The *rulA* gene is also present (**G**) and the ILE contains three integrase genes (**D,E,F**) along with the type three effector gene, *avrPpiA1* (**C**). Image adapted from Arnold *et al.,* 2000. (Used under agreement from John Wiley and Sons, no. 4138340003860).

Inverted repeats also flanked the 4.3kb region meaning there was a high possibility further insertions had occurred. Following this work it was proposed that this potentially mobile region of DNA (ILE) carrying the avirulence gene, *avrPpiA1* can move between plasmids. This was suggested as the disrupted *rulB* gene is plasmid-borne and it also explains the distribution of *avrPpiA1* gene to plasmids in *Ppi* races 5 and 7 (Arnold *et al.,* 2000). Although it is hypothesised that the inserted region carrying the *avrPpiA1* gene is mobile between chromosomes and plasmids of different races and possibly species, the movement of the region has never been observed in the case of *Ppi.* However

Rhodes *et al.,* (2014) have observed a region similar to the one seen in strain 203 move out of the chromosome and into a plasmid in a *P. fluorescens* strain.

Other similar MGEs have been identified in different species of bacteria that also insert into *rulAB* orthologues such as *umuDC.* One recent example of this is in *Pantoea ananatis* which has an integrative and conjugative element (ICE) that contains stress response genes and antibiotic resistance genes was identified and may play a major role in diversification of *Pantoea ananatis.* The really interesting finding was that the ICE was inserted into the *umuC* gene of the *umuDC* operon which is an orthologue of *rulAB* (Maayer *et al.,* 2015). So the same process as the ILE insertion appears to be happening in different species with different MGEs.

1.10: *Pseudomonas fluorescens.*

Pseudomonas fluorescens is a common Gram-negative, rod-shaped bacterium with multiple flagella and can be found in soil and water. It is an obligate aerobe, but certain strains are capable of using nitrate instead of oxygen as a final electron acceptor during cellular respiration (Palleroni, 1984). *P. fluorescens* does not cause plant disease and certain strains can actually suppress plant fungal diseases. The biocontrol abilities depend on root colonization, induction of systemic resistance in the plant and the production of diffusible antifungal antibiotics (Haas and Keel 2003).

1.10.1: An integron-like element moving from *P. fluorescens* chromosomes to plasmid pWW0 and *rulAB* as a hotspot for insertion.

The ILE in *P. fluorescens* strain *FH1* was first discovered in 1989 during an investigation into plasmid-encoded copper resistance in environmental isolates from a copper mine in the Lake District, Cumbria, UK (Pickup, 1989). It was whilst Pickup was attempting to cure the strains via incompatibility experiments using plasmid pWW0, that it was noticed that the plasmid had acquired an extra region of DNA (ILE) and this was repeatable. The region was 10kb in length, and as Arnold *et al.* (2000) had observed, the insertion disrupted the *rulB* gene. Sequencing revealed the region had left the chromosome and inserted into pWW0. This ILE however contained genes responsible for efflux pumping of toluene allowing the bacteria to withstand high toluene levels, although open reading frames (ORFs) 1-3 were conserved and similar to the integron found in *Ppi* race 2 strain 203 and are integrases and recombinases found in most integrons (Rhodes *et al.,* 2014).

It has therefore been established that ILEs can move between the chromosome of a bacterium and into a plasmid. The plasmid is then potentially transmissible to other strains and species transferring the genes the ILE has captured. What has also been deduced is that *rulB* is a hotspot for ILE insertion. It is not fully understood why the insertion in *rulB* has been selected over evolutionary time but one possible reason is that on many plasmids *rulB* is within the core backbone of the plasmid close to the origin of replication. This ensures its passage into another bacterium which may or may not be related depending on the host range of the plasmid (Rhodes *et al.,* 2014).

1.11: Aims.

1-To determine the frequency of integron-like elements (ILEs) and *rulAB* genes in different plant pathogenic *Pseudomonas* bacteria.

An investigation of how widespread ILEs (*rulAB*) are in a bank of >200 plant pathogenic bacteria currently maintained at UWE. Followed by investigating how many of these ILE's are disrupted (whether their *rulB* has been disrupted).

2-To determine whether ILEs and ILE captured genes are mobile.

An examination of whether ILEs and their captured genes are mobile between different integrase systems and different *rulAB* genes. These investigations will include the potential for retrotransfer of ILEs by plasmid pWW0.

3- To identify the conditions in which integron integrase and type three effectors (TTE) are expressed.

The conditions for expression of the integron integrase (*intI*) will be identified. The focus will initially be on the sequenced ILE in *P. syringae* pathovar *pisi* 203 (integron harbours effector *avrPpiA1*). Further to this any newly identified ILEs will also have their expression profiles analysed to assess any differences between them.

4-To characterise the ILE co-localisation with *rulB.*

A previous bioinformatics scan of ~25 genomes (Jackson *et al*., 2011) found evidence that *rulB* (or orthologues) disruption is widespread. Why does the integron insertion occur within *rulB* and does this affect the fitness of the bacteria? The *rulAB* gene may play a role in stabilising the insertion and ensure that it is conserved and always transferred. An investigation will look at the relationship

between integron insertion and the *rulB* hotspot to assess why *rulB* is favoured over other genes.

5-To investigate if the disrupted RulB protein still conveys UV resistance.

Using a bank of strains, it will be possible to compare the relative UV resistance phenotype of each strain. This will thus determine whether the rulB system remains functional even after disruption or whether the capture and integration of the integron represents a trade-off in bacterial fitness.

Chapter 2. Materials and Methods

2.1: Bacterial Strains.

A summary list of bacterial strains and plasmids used in this project is shown in Table 2.1. A complete detailed table is available in Appendix I.

Table 2.1: Bacterial strains used in this study. Km- Kanamycin; Stm- Streptomycin

Strain	Antibiotic Resistance	Reference
Pseudomonas syringae pv. *pisi*		
Race 1 (all available numbered strains from stock)	None (not screened)	HRI culture collection
Race 2 (all available numbered strains from stock)	None (not screened)	HRI culture collection
Race 3 (all available numbered strains from stock)	None (not screened)	HRI culture collection
Race 4 (all available numbered strains from stock)	None (not screened)	HRI culture collection
Race 5 (all available numbered strains from stock)	None (not screened)	HRI culture collection
Race 6 (all available numbered strains from stock)	None (not screened)	HRI culture collection
Race 7 (all available numbered strains from stock)	None (not screened)	HRI culture collection
Pseudomonas putida		
PaW340 (-)	Stm	Rhodes *et al.,* 2014
PaW340 (pWW0::km^r)	Stm, Km	Rhodes *et al.,* 2014
PaW340 (pWW0::km^{r}Δ*rulAB*)	Stm, Km	Rhodes *et al.,* 2014
Pseudomonas fluorescens		
FH1	None (not screened)	Rhodes *et al.,* 2014
FH4	None (not screened)	Rhodes *et al.,* 2014
FH1 (pWW0::km^r::ILE$_{FH1}$)	Km	Rhodes *et al.,* 2014
FH4 (pWW0::km^r::ILE$_{FH4}$)	Km	Rhodes *et al.,* 2014
Pseudomonas syringae pv. *phaseolicola*		
Pph 1302A	None	HRI culture collection
Pph races 1, 3-9 (one representative from each)	None (not screened)	Tsiamis *et al.,* 2000
Pseudomonas syringae pv. *syringae*		
Psy. A selection of strains	None (not screened)	HRI culture collection
Plasmids		
pWW0::km^r (large self-transmissible plasmid with heavy metal resistance)	Km	Rhodes *et al.,* 2014

pWW0::km^r::ILE$_{FH1}$ (large self-transmissible plasmid with heavy metal resistance, also has the integron inserted into *rulAB* from *Pf. FH1*)	Km	Rhodes *et al.,* 2014
pBBR1MCS-2 broad host range vector	Km	Invitrogen
pCR2.1 PCR cloning vector	Km	Invitrogen

2.2: Media and culture conditions.

All *Pseudomonas* strains were plate cultured on either King's B (KB) media (Fluka Analytical, UK) or nutrient agar (NA) (Oxoid, UK) for 48 hours at 25^{o}C. Overnight liquid cultures of *Pseudomonas* strains were grown in Luria Bertani (LB) broth (Difco, UK) for 16 hours at 25^{o}C with shaking at two times g-force (xg). Selection of bacterial matings of *Pseudomonas* strains were carried out on M9 minimal media (MM) (Fluka Analytical, UK) or KB. Agar plates were supplemented with antibiotics (Sigma, UK) at concentrations of 5 mg/ mL kanamycin (Km) and 10 mg/ mL streptomycin (Stm) and LB broths were supplemented with 2.5 mg/ mL Km and 5 mg/ mL Stm. For long term storage bacterial cultures were grown overnight and mixed with 40% glycerol (1:1) then stored in a -80^{o}C freezer (U57085, New Brunswick Scientific, UK). The broths were incubated in a shaking incubator (Innova 4230, New Brunswick Scientific, UK) and agarose plates were incubated in a static incubator (LT2J, LEEC, UK).

2.2.1: Making rifampicin mutants.

The bacterial strain required to be rifampicin resistance was grown in LB broth overnight. One mL of liquid culture was centrifuged for one minute at 15,000 xg, the pelleted cells were resuspended in 100 µL of LB broth and spread onto a KB plate containing 20 mg/ mL rifampicin. The plates were incubated at 25^{o}C for approximately four days until resistant colonies appeared.

2.3: DNA extraction.

2.3.1: Genomic DNA extraction.

Overnight liquid cultures of *Pseudomonas* strains were grown in 10 mL LB broth at 25^{o}C. Genomic DNA was extracted from 1.5 mL of liquid culture using the Puregene Yeast/Bact. Kit B (QIAGEN, UK) as per the manufacturer's instructions.

2.3.2: Plasmid DNA extraction.

Overnight liquid cultures of *Pseudomonas* strains were grown in 10 mL LB broth at 25^{o}C. Plasmid DNA was extracted from 1.5 mL of liquid culture using the QIAprep spin mini-prep kit (QIAGEN, UK) as per the manufacturer's instructions. Larger plasmids were extracted from 50 mL of liquid culture using the QIAGEN plasmid midi kit as per the manufacturer's instructions.

2.4: Polymerase chain reaction (PCR).

All of the PCR reactions carried out were a total volume of 25 µL that consisted of 2 µL of DNA or overnight liquid culture, 1 µL of forward primer and 1 µL of reverse primer (Table 2.2), 12.5 µL of mastermix (Taq PCR mastermix, QIAGEN, UK) and 8.5 µL of sterile deionised water. PCR was performed using a Flexigene thermal cycler (TECHNE, UK). Standard PCR cycling was used which consisted of 94^{o}C for 10 minutes then 35 cycles of 94^{o}C for 30 secs, Tm (-5^{o}C) of primers for 30 seconds (Tm was calculated from the composition of the primers), 72^{o}C for 1 min and a final extension at 72^{o}C for 10 minutes, the samples were then held at 4^{o}C until needed. PCR products were visualised following the method given in Section 2.5.

2.4.1: Semi-degenerate primer PCR.

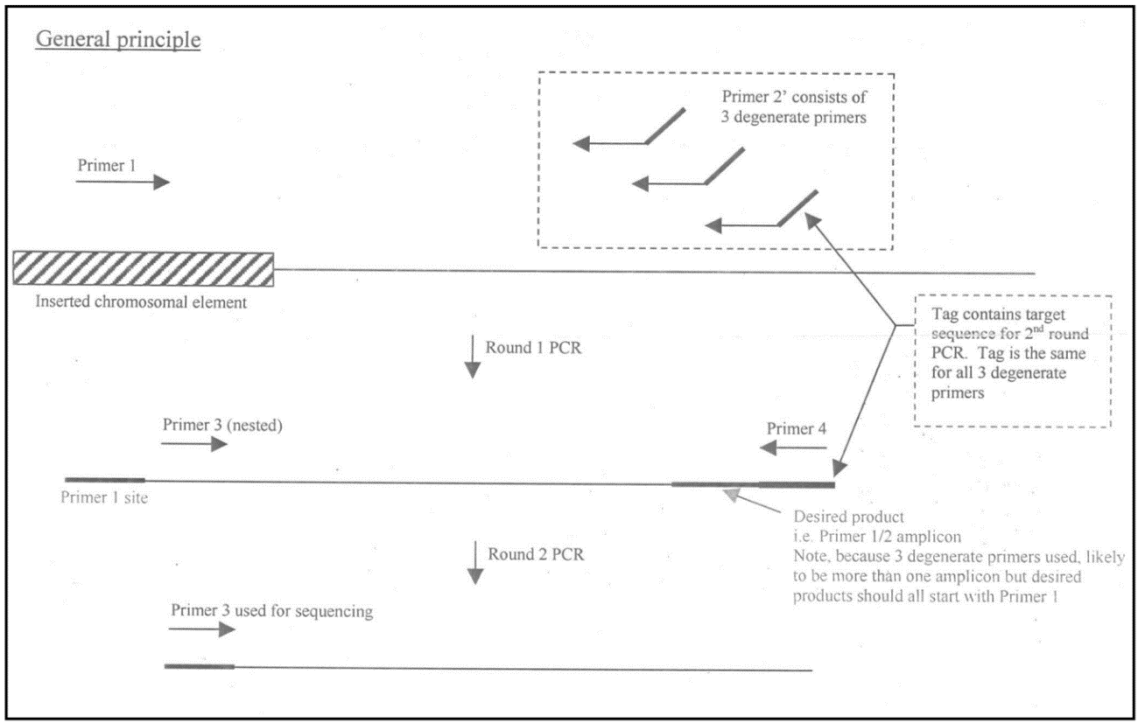


Figure 2.1: Outline of semi-degenerate amplification of chromosomally inserted targets. (Image generated and used with permission from S. Godfrey).

The degenerate primers (Table 2.2; CEKG) bound to multiple regions in the variable end of the ILE. Primer 1 was the primer to bind to *rulB"* (variRP1). The first round of PCR then amplified the region between Primer 1 and a mix of the degenerate primers (CEKG 2A, B and C). The second round of PCR uses Primer 1 again but Primer 4 (CEKG 4) that binds to tags on the degenerate primers (CEKG 2A, B and C). A sequencing primer was then used for sequencing. The amplicon of interest should start with the known sequence (*rulB"*).

2.4.2: Primer design.

Primers were designed using Oligoanalyzer software found at: http://eu.idtdna.com/analyzer/applications/oligoanalyzer. The sequences were first obtained from the NCBI database and copied into Bioedit software where primer sequences were selected manually. These sequences were checked for

secondary structures and their melting temperatures (Tm) using the Oligoanalyzer software. Primers were synthesised by Eurofins (Germany).

Table 2.2: Primers used for PCR and hybridisation probes.

Primer Name	Sequence 5' - 3'	Target Gene(s)
ILE region primers		
GRrulABFP (pWW0) GRrulABRP (pWW0)	TGGCGTATGTCGATAACCAG CAATTCCCCGTACAAGGTGT	*rulAB*
GRxerCFP (pWW0) GRxerCRP (pWW0)	AGCAGCGCAACCTGATAACT GCCTGCCTTCATTAGTCAGC	*xerC*
GRrulAB-xerCFP (pWW0) GRrulAB-xerCRP (pWW0)	TGGCGTATGTCGATAACCAG GTACAGACGCCGTCCATAGG	*rulAB-xerC* flank
2015rulABFP 2015rulABRP	CCATCATGAAGAGCCGCTTGCAGAT TCATTGAAAAGACGGCTCGTGAGTT	*rulAB'*
2015rulAB-xerCFP 2015rulAB-xerCRP	CCATCATGAAGAGCCGCTTGCAGAT TGGTACGACCACCTGGTGTTTCAG	*rulB'-xerC*
2015xerCFP 2015xerCRP	CTGAAACACCAGGTGGTCGTACCA ACCTTGATTTATGTGCACCTGTCGG	*xerC*
ILE variable primers		
CEKG 2A	GGCCACGCGTCGACTAGTACNNNN NNNNNNAGAG	Degenerate primer
CEKG 2B	GGCCACGCGTCGACTAGTACNNNN NNNNNNACGCC	Degenerate primer
CEKG 2C	GGCCACGCGTCGACTAGTACNNNN NNNNNNGATAT	Degenerate primer
CEKG 4	GGCCACGCGTCGACTAGTAC	Binds to degenerate primers
variFP variRP	AGCCAGGAGACGCTTTGCTG TACACTCTCCTCGCATTGGG	ILE variable end and rulB''

2.5: Agarose gel electrophoresis.

Agarose gels of 0.7% were prepared by dissolving agarose powder (Bioline, UK) in 1x TAE (Appendix II) via microwaving (Sanyo, UK). Agarose gels were supplemented with 0.01% v/v (µL) of Sybr Safe (Invitrogen, USA) with a concentration of 1000x. The DNA samples had 20% v/v of loading dye added (Appendix II). The gels were loaded with 5 µL of size ladder (Bioline, UK) and 10

µL of DNA sample per well. The gels were then placed in tanks containing 1x TAE and electrophoresed for 80 minutes at 100V for 150 mL gels and 40 minutes at 80V for 30 mL gels. The gels were visualised and photographed using a UV gel doc system (U:Genius, Syngene, UK).

2.6: Clean-up of PCR products or PCR products from an agarose gel.

PCR products or PCR products from an agarose gel were cleaned-up using the Wizard SV Gel and PCR clean-up system (Promega, UK) as per the manufacturer's instructions.

2.7: Quantifying DNA purity with a Nanodrop 1000.

The purity of DNA samples were measured using a Nanodrop 1000 (Thermo Scientific, UK). The Nanodrop head was cleaned twice with 2 µL of sterile water and lint free tissue. The Nanodrop was then blanked with 2 µL of the solution that the DNA was rehydrated in. The head was then cleaned again and 2 µL of the DNA sample was added and measured.

2.8: DNA hybridisation via vacuum dot blotting.

All hybaid oven glass tubes (Thermo Scientific, UK) were washed with liquid detergent and hot water, followed by a rinse in cold water, followed by a vigourous rinse in deionised water and left to dry in a drying cabinet (LEEC) before use. The Bio-rad dot blot apparatus (Bio-rad, UK) was also washed in the same way but followed by autoclave treatment.

2.8.1: Producing chemiluminescent labelled probes.

The plasmid pWW0::km^r was used to produce the probes for the *rulAB* regions and the plasmid pWW0::km^r::ILE$_{FH1}$ was used as a probe primarily for *xerC* regions but also *rulAB-xerC* crossover regions. *Pseudomonas putida* strain PaW340 (pWW0::km^r) and *Pseudomonas fluorescens* strain *FH1* (pWW0::km^r::ILE$_{FH1}$) were cultured in 10 mL of LB broth supplemented with 25 µg/ mL Km for 20 hours with shaking at 25^{o}C. The plasmid DNA (pDNA) was extracted using a mini prep extraction kit (Section 2.3.2). Following the extraction the purity of the pDNA was checked using a Nanodrop 1000 (Thermo Scientific, UK)**.** The pDNA from the two strains was then used as a template in a PCR reaction (1:100 dilution of pDNA) (primers shown in Table 2.2). The PCR products were cleaned up using a PCR clean-up kit (Promega) (Section 2.6).

The amount of DNA from the PCR was calculated using a Nanodrop and 400 ng of DNA was required for the labelling process (volumes in Table 3.1). The DNA was added to autoclaved, double distilled (dd.) water to make a final volume of 16 µL in a 1.5 mL Eppendorf tube. The DNA was then denatured by heating in a water bath to 100^{o}C (W14, Grant) for 10 minutes then quickly chilled in an ice bath. The denatured DNA then had 4 µL of well mixed DIG-high prime (Roche, UK) added. The mixture was centrifuged (Prism, Labnet) at 15,000 xg for 10 seconds and incubated in a static 37^{o}C incubator (LEEC, UK) overnight. The labelling reaction was stopped by heating the sample to 65^{o}C for 10 minutes.

2.8.2: Transferring DNA onto a nylon membrane.

A small amount of bacterial cells from a plate culture were added to 200 µL of 0.4M NaOH-10mM EDTA in the wells of a microtitre plate. The plate was sealed with autoclave tape and incubated for 15 minutes at 60^{o}C in a Hybaid mini

oven (Thermo Scientific, UK). Once the incubation was complete the microtitre plate was chilled on ice for five minutes.

The positively charged nylon membrane (Biotrans) was wetted with 2x saline-sodium citrate (SSC) buffer (Appendix II) and the dot blotter was assembled by placing the nylon membrane between the rubber gasket and the upper part of the blotter; alternate screws were tightened. A vacuum was applied and the screws tightened further. The vacuum was then held and 180 µL of each sample was loaded. The vacuum was reapplied to draw the samples through. Once the samples were through, the blotter was disassembled and the nylon membrane was washed briefly in 2xSSC and left to air dry for 30 minutes. Once the nylon membrane was dry it was wrapped in Saran wrap and exposed to ultra-violet (UV) light (302nm) for 2 minutes.

2.9: Hybridisation of DIG-labelled probe to DNA.

2.9.1: Prehybridisation of blot with DIG easy hyb.

During the prehybridisation of the blot 10 mL of DIG Easy Hyb (Roche, UK) was used for an 8 cm by 12 cm sized blot. The blot was placed into a 110 mL glass Hybaid oven tube along with 10 mL of DIG Easy Hyb. The blot was incubated at 42^{o}C for 1 hour inside the Hybaid oven with rotation.

2.9.2: Hybridisation of DIG labelled probe to DNA on the blot.

At least 50 ng of labelled probe was used in the hybridisation buffer, the final volume of labelled probe from Section 2.81 used was 2 µL for all of the 3 probes. The labelled probe was added to 50 µL of dd. water in a 1.5 mL Eppendorf tube and denatured in boiling water for 5 minutes. The probe was removed and chilled in an ice bath. The probe was immediately added to 5 mL of prewarmed

DIG Easy Hyb (42^{o}C). The prehybridisation buffer was removed from the Hybaid tube and the probe buffer was added. The hybridisation reaction was left overnight at 42^{o}C with rotation.

Once the hybridisation had finished the blot was removed from the Hybaid tube and the buffer was kept (it can be reused five times). The blot was then washed in 200 mL low stringency buffer (LSB) (Appendix II) twice for 5 minutes. During this step high stringency buffer (HSB) (Appendix II) was prewarmed to 65^{o}C. The blot was then transferred to a clean Hybaid tube where 75 mL of prewarmed H.S.B. was added and left rotating in the Hybaid oven for 20 minutes at 65^{o}C. This was repeated once more.

2.10: Detection of hybridised probes.

Following hybridisation the blot was transferred to a clean plastic tray where 200 mL of washing buffer (Appendix II) was added. The blot was gently shaken for 3 minutes at room temperature; the buffer was then discarded. Using the same tray 170 mL of blocking solution (Appendix II) was added to the blot and left shaking for 1 hour; the solution was then discarded. Following this 30 mL of antibody solution (Appendix II) was added and left shaking for 30 minutes. The blot was then washed in 200 mL of washing buffer for 20 minutes; this was repeated with fresh buffer. Finally the blot was equilibrated for 5 minutes in 30 mL of detection buffer (Appendix II).

Once all the washes were complete 1 mL of chloro-5-substituted adamantyl-1,2-dioxetane phosphate (CSPD) (Roche, UK) was evenly distributed to the DNA side of the blot and placed inside a hybridisation bag (Roche, UK). The blot was left for 5 minutes at room temperature and then excess CSPD was drained off. The bag was resealed and incubated at 37^{o}C for 10 minutes to

activate the CSPD. Blots were then exposed to chemiluminescent X-ray film (GE Healthcare) for 10 minutes and developed using Carestream developer and fixer (Kodak, USA) inside a dark room.

2.11: DNA sequencing.

2.11.1: Sequencing of PCR products.

PCR products were cleaned up using the ExoSAP protocol before sequencing. The ExoSAP mastermix was made up fresh and comprised of 2.5 µL of Exonuclease 1, 25 µL of SAP and 972.5 µL of Milli Q water (Thermo-Scientific, UK). Each PCR sample had 10 µL of ExoSAP mastermix added and was incubated at 37^{o}C for 30 minutes and then at 95^{o}C for 5 minutes in a thermal cycler. This was stored at -20^{o}C until needed. The forward primer was diluted, 1.8 µL of primer to 10.2 µL of dd. water. The diluted primer (3 µL) was then mixed with 12 µL of the ExoSAP cleaned PCR product. The PCR product and primer was sent to Eurofins, Germany for sequencing. Sequence results were used to confirm the PCR results using NCBI BLAST, CLUSTALW and TCOFFEE (ebi.ac.uk) multiple alignments. Single alignments were carried out using EMBOSS Water (http://www.ebi.ac.uk/Tools/psa/emboss_water/).

2.11.2: Preparation for whole bacterial genome sequencing.

Bacterial genome sequencing was performed by MicrobesNG, UK. DNA was extracted using the Puregene Yeast/Bact. Kit B (QIAGEN, UK). The DNA integrity and quantity was checked using a Nanodrop 1000 (Thermo-Scientific, UK). The DNA was eluted in elution buffer (Qiagen, UK). The sequencing was performed using the Illumina MiSeq and HiSeq 2500 platforms and genome sequence quality was verified using the Burrows-Wheeler Aligner by MicrobesNG, UK. The genomes were assembled and partially annotated using

SPAdes (Bankevich *et al.,* 2012), available here; http://cab.spbu.ru/software/spades/. Subsequent analysis and annotation was performed using Prokka (Seemann, 2014) to annotate contigs with genes, blast homologs, identify domains and motifs, find tRNAs and rRNAs and CRISPR-Cas if present; http://vicbioinformatics.com/software.prokka.shtml. The Artemis: sequence visualization and annotation software (Rutherford *et al.,* 2000); http://www.sanger.ac.uk/science/tools/artemis was also used for gene annotation.

2.12: Bacterial mating.

All cultures used for matings were grown in 10 mL LB broth overnight.

2.12.1: Filter matings.

Liquid cultures of both *P. putida* PaW340 (pWW0::km^r) (donor) and *P. fluorescens FH1* (recipient) were used in a control filter mating. A sterile membrane filter (Supor 200, PALL, UK) was placed onto a KB plate, 10 µL of each of the liquid cultures was spotted onto the membrane and in the final spot 1 culture was placed on top of the other to facilitate the mating. These were left to grow over 3 days at 25^{o}C. Following this the bacterial cells from the mating were picked up using an inoculation loop and immersed in 1.5 mL 1/4 Ringers solution (Appendix II). This solution was then streaked onto M9 MM plates containing just Km which would select for *P. fluorescens FH1* (pWW0::km^r) and incubated for four days at 25^{o}C.

2.12.2: Eppendorf tube matings.

Liquid cultures of both *P. putida* PaW340 (pWW0::km^r) (donor) and *P. fluorescens FH1* (recipient) were used in a control Eppendorf mating. Using a

centrifuge 0.5 mL of the donor culture and 1 mL of the recipient culture was spun down for one min at 15,000 xg. The supernatant was discarded and the cells resuspended in 0.5 mL and 1 mL respectively of 1/4 Ringers solution; the centrifugation and resuspension was then repeated. Following this 500 µL of each cell suspension was mixed together in a clean Eppendorf tube and centrifuged for 1 min at 15,000 xg. The supernatant was discarded and the pellet (~10 µL) was drawn up using a pipette. The bacterial cells were placed in the middle of a NA plate and left to grow for 48 hours at 25^{o}C. Bacterial cells were then streaked out onto M9 MM plates containing 5 mg/ mL Km which would select for *P. fluorescens FH1* (pWW0::km^{r}) and incubated for 4 days at 25^{o}C.

2.12.3: Electroporation of plasmid DNA into recipient cells.

2.12.3a: Electroporation of plasmid DNA into *Pseudomonas.*

Overnight liquid cultures of both donor and recipient cells were grown and plasmid DNA (pWW0) was extracted from the donor following method 2.3.2. One mL of the recipient culture was harvested and washed three times in 750 µL ice cold sucrose (0.5M). 100 µL of the recipient cells were incubated on ice with 10 µL of pDNA for 30 minutes. The cell solution was then transferred to an ice cold cuvette and electroporated at 200Ω, 2000V and 25µF with a time constant of ~4.4. Following this 1 mL of LB broth was immediately added and mixed via pipetting. The cells were centrifuged for 1 minute at 13,000 rpm and resuspended in 200 µL of LB broth. The suspension was plated out onto KB + Km plates and incubated at 25^{o}C for 3-4 four days.

2.12.3b: Electroporation of plasmid DNA into *E. coli.*

The recipient strain was grown overnight at 37^{o}C. 100 µL was used to inoculate 10 mL of fresh LB media and grown at 37^{o}C until an OD_{600} of 0.5-1 was

reached. The culture was then chilled on ice for 30 minutes before being harvested and resuspended in 10 mL ice cold sterile water, this was repeated twice more. The cells were then washed twice in 200 µL 10% (v/v) glycerol and a final wash with 2.3 mL 10% (v/v) glycerol. 100 µL of the recipient cells were incubated on ice with 10 µL of pDNA for 30 minutes. The cell solution was then transferred to an ice cold cuvette and electroporated at 200Ω, 2000V and 25µF with a time constant of ~4.4. Following this 1 mL of LB broth was immediately added and mixed via pipetting. The cells were centrifuged for 1 minute at 13,000 rpm and resuspended in 200 µL of LB broth. The suspension was plated out onto KB + Km plates and incubated at 25^{o}C for 3-4 days.

2.13: Apoplastic fluid extraction.

Bean plants were grown for 14 days at 23^{o}C with 80% humidity and 16 hours of light. Following this the leaves were harvested by cutting at the base of the leaf. The leaf was then folded into a quarter of the size taking care not to break or damage the leaf. The folded leaf was placed into a large 100 mL syringe and filled with distilled water, any excess water and air was removed. Water was then forced into the leaf by placing a finger over the end of the syringe and pushing the syringe. Once the leaf was dark green all over it was removed from the syringe and placed into a smaller 20 mL syringe suspended in a Falcon tube. This assembly was then centrifuged at 1500 xg for 15 minutes to collect the expelled apoplastic fluid which was stored at -80^{o}C.

Apoplastic fluid concentration was measured following the method developed by O'Leary *et al.,* (2014). Two volumes of distilled water (100 mL) were prepared with 1 containing 50 µM indigo carmine. The absorbance of both solutions was measured at 610 nm with the OD_{610} of the solution containing no indigo carmine subtracted from the OD_{610} of the indigo carmine containing

solution. This gave the $OD_{610infiltrate}$ value. Using the method above leaf infiltrations with both solutions was performed. The average OD_{610} of at least 3 replicates was measured with the apoplastic fluid containing no indigo carmine subtracted from the indigo carmine containing apoplastic fluid, this gave the OD_{610AWF} value. These values were then used in the equation; Apoplast dilution factor = $OD_{610infiltrate}/(OD_{610infiltrate}-OD_{610AWF})$. The fluid was then diluted and normalised to the same starting dilution factor for all tests (dilution factor of 1 in 5 mL).

A check was also performed to ensure no cell lysate was present in the apoplastic fluid preparation. This was performed using a malate dehydrogenase assay (Toyobo Co., Ltd, 2010). The assay was performed on all apoplastic fluid extractions from the 4 different plants, TG, CW, RM and Pea. Cell lysate samples from the plants were also included as positive MDH samples along with distilled water as the negative MDH control. The enzyme activity was calculated using;

$$\text{Volume activity (U/ml)} = \frac{\Delta OD/min\ (\Delta OD\ test - \Delta OD\ blank) \times Vt \times df}{6.22 \times 1.0 \times Vs} = \Delta OD/min \times 9.807 \times df$$

Weight activity (U/mg)=(U/ml)×1/C

Vt : Total volume (3.05ml)

Vs : Sample volume (0.05ml)

6.22 : Millimolar extinction coefficient of NADH under the assay condition(㎠/micromole)

1.0 : Light path length (cm)

df : Dilution factor

C : Enzyme concentration in dissolution (c mg/ml)

2.14: Quantitative reverse transcription PCR (RT-qPCR).

2.14.1: Preparation of bacteria for gene expression studies in plant apoplastic fluid.

Overnight liquid cultures of *Ppi* 203 and *Psy* 3023 were grown in LB broth at 25ºC. Following overnight growth the cells were washed in 1/4 Ringers solution and 100 µL of *Ppi* 203 or *Psy* 3023 was placed into 1 mL of each of the apoplastic fluid preparations; Bean cv. Tendergreen, Bean cv. Canadian Wonder, Bean cv. Red Mexican or Pea cv. Kelvedon Wonder along with minimal media as a control. These were left shaking at 25ºC for 6 hours. The cells were then harvested. The apoplastic fluid was diluted 50:50 with minimal media.

2.14.2: Preparation of bacteria for gene expression studies *in planta.*

Overnight liquid cultures of *Ppi* 203 and *Psy* 3023 were grown in LB broth at 25ºC. Following overnight growth the cells were washed in 1/4 Ringers solution. These preparations were then injected into Bean cv. Tendergreen, Bean cv. Canadian Wonder, Bean cv. Red Mexican or Pea cv. Kelvedon Wonder plants and left for 6 hours before being extracted via grinding of the leaf. Minimal media was used as a control.

2.14.3: Preparation of bacteria for gene expression studies following bacterial conjugation.

Overnight liquid cultures of *E.coli* DH5α and *E.coli* DH5α (pRK2013) and either *Ppi* race two strain 203 or *Psy* 3023 were grown. The conjugation mix was set up by mixing 60 µL of *Ppi* 203 or *Psy* 3023, 20 µL of *E.coli* DH5α and 20 µL of *E.coli* DH5α (pRK2013). This mix was then placed in the middle of an LB agar (Appendix II) plate and incubated at 25ºC for 6 hours. Following the incubation

the conjugated cells were scraped off the plate and resuspended in 500 µL LB broth. No conjugation of *Ppi* 203 and *Psy* 3023 were used as controls.

2.14.4: Preparation of bacteria for gene expression studies following exposure to sub-optimal temperatures.

Overnight liquid cultures of *Ppi* 203 and *Psy* 3023 were grown in LB broth at 25ºC. Following overnight growth the cells were washed in 1/4 Ringers solution. The preparations were then incubated for 6 hours at -80ºC, -20ºC, 4ºC, 25ºC and 37ºC. The cells incubated at 25ºC were used as a control.

2.14.5: Preparation of bacteria for gene expression studies following exposure to mitomycin C.

Overnight liquid cultures of *Ppi* 203 and *Psy* 3023 were grown in LB broth at 25ºC. MM containing 0.05, 0.1, 0.5 and 1 µg/mL MMC was prepared and 200 µL of the overnight cultures were incubated in the MMC preparations for 6 hours. This method was also used for bacterial growth studies following MMC exposure. The only difference was that the cells were plated onto KB plates and left to grow at 25ºC for 48 hours before being counted. CFU/ mL values were also calculated using the following equation; CFU/ mL = (number of colonies*dilution factor) / volume spread on plate.

2.14.6: RNA protect.

Following the RT-qPCR bacterial preparation steps 500 µL of the samples were added to 1 mL of RNA protect reagent (QIAGEN, UK) in a 1.5 mL Eppendorf tube. The samples were then immediately mixed by vortexing and incubated at room temperature for 5 minutes followed by centrifugation at 7000 xg for 10

minutes. The supernatant was then decanted and the tube was drained on paper to dry. The pellet can be stored for up to two weeks at -20^{o}C.

2.14.7: RNA purification.

The RNA pellet was thawed and 100 µL of TE buffer containing 1 mg/ mL lysozyme was added and incubated at 25^{o}C with shaking for 10 minutes; the samples were vortexed every 2 minutes during the incubation. Following this a RNA purification kit was used, either the RNeasy kit (QIAGEN, USA) or the SV Total RNA Isolation System (Promega, USA) as per the manufacturer's instructions. Following the purification the samples were checked on a Nanodrop 1000 (Thermo Scientific, UK) to ensure purity and to get a RNA concentration to normalise the samples.

2.14.8: Complementary DNA (cDNA) synthesis.

The TaqMan reverse transcription reagents were used to generate the cDNA. For each sample and replicate 2 µL of 10x RT buffer, 1.4 µL of 25mM $MgCl_2$, 4 µL of 10mM dNTP mix, 1 µL of RNase inhibitor and 1 µL of MultiScribe reverse transcriptase was used. The master mix was made up to 20 µL with the addition of purified DNA/RNAase free water. Each sample had 20 µL of the mastermix added along with 1 µL of random hexamers and less than one µg of template RNA. The reactions were then incubated in a thermal cycler (TECHNE, UK) with the following settings: 25^{o}C for 10 minutes, 37^{o}C for 30 minutes, 95^{o}C for 5 minutes and 4^{o}C indefinitely. The cDNA was stored at -20^{o}C or used immediately for real-time quantification.

2.14.9: Reverse transcription quantitative PCR (RT-qPCR).

Primers and probes used for real-time PCR quantification are in Table 2.3. Each well for the real-time quantification contained 12.5 µL of RT-qPCR mastermix, 2 µL of forward and reverse primers, 2 µL of probe, 3 µL of cDNA and 5.5 µL of DNase, RNase free water. The samples were then subjected to 50^{o}C for 2 minutes, 95^{o}C for 10 minutes and then 40 cycles consisting of 95^{o}C for 15 seconds and 60^{o}C for 1 min. DNA gyrase sub-unit B gene was used for the base level gene expression to normalise all other values.

Table 2.3: Primers and fluorescent probes used for RT-qPCR.

Primer/Probe name	Sequence 5' - 3'	Target gene
GyrB-QF	GATGATGGAATCGGTGTCGAA	DNA gyrase subunit B
GyrB-QR	TTGGTGAAGCACAACAGGTTCT	
GyrB-QP	CCCTGCAGTGGAACGACAGCTTCA	
Ppi 203		
ORFA-QF	CGTCCAGGCGCAAACC	Recombinase
ORFA-QR	ACAGCACCGCACCGAGAT	
ORFA-QP	TAGGCAATGATCTGTCG	
203ILExerCF	CTGCGTCGGCCTTTGG	*xerC* on ILE
203ILExerCR	AATGAGTCGATGGGCGAGAT	
203ILExerCP	CGTTTCTGCGCGCACT	
203ORFDF	CATCAGCTCCATTAGCGACATG	ORF D *int*
203ORFDR	GCCTCGGTCGTTACTGCATT	
203ORFDP	CTGCGGAACACTCG	
203ORFEF	CGTACGAAAACCTCCCATGTG	ORF E *int*
203ORFER	TGGGTCGCGCCCTTT	
203ORFEP	TGCGACCGCTGCC	
AvrPpiA1-QF	GTGCAACCGAGGGATCTAGAAC	TTE on ILE
AvrPpiA1-QR	TCTAGCATTTCTGACGAGCATGTC	
AvrPpiA1-QP	TCGAGCTCAGCCCCGA	
Psy 3023		
3023ILExerCF	TCCGTCGTGCCGAACTG	*xerC* on ILE
3023ILExerCR	TCGATGGGCGAGATCCA	
3023ILExerCP	TGGCCCTTAGAATCG	
3023ILEHopH1F	CAGGAGGCAGCTGGAAAAGT	TTE on ILE
3023ILEHopH1R	GCAATGTGCCGATTCTCGTT	
3023ILEHopH1P	CGAACCAGAAAACTT	
3023ILEHopAP1F	TGTGAAATGGATCCGCAGAA	TTE on ILE
3023ILEHopAP1R	TCCGAGCCCTGATCTGTCTT	
3023ILEHopAP1P	CGGGTGGTCATCTC	

2.14.10: Statistics and heat map analysis of gene expression.

A students T-test and a Tukey statistical test was performed on the raw qPCR data following ANOVA to obtain the level of significance between the mean values for each condition within a set experiment. JMP statistical package was used (https://www.jmp.com/). Gene expression heat maps were generated using the x-fold gene expression values and matrix2png software provided at http://www.chibi.ubc.ca/matrix2png/bin/matrix2png.cgi was used to generate the heat maps.

2.15: Ultra-violet tolerance tests.

The strains to be tested for UV tolerance in relation to their *rulB* profile were grown in liquid culture overnight at 25^{o}C with gentle shaking at 2 xg with appropriate antibiotics. Following overnight growth the cells were harvested by centrifugation at 1500 xg for 15 minutes (MSE Mistral 2000) and the cells were then washed in 10 mL 10 mM magnesium chloride and were normalised to an optical density of 0.6 at 600nm. The washed cells were transferred to a 80mm Ø glass petri dish and exposed to UVB light (302nm) on a UV transilluminator (UVP, UK) for either 0, 30, 60, 120 or 300 seconds (UV source was allowed to warm up for 15 minutes prior to use). Following the UV exposure 1 mL was taken from the petri dish and added to the 10 mL LB broth and grown at 25^{o}C with shaking under red light. The growth (optical density) was measured every hour for 8 hours and at 24 hours at 600nm in a Jenway 7315 spectrophotometer; growth curves were produced from the data. If the OD_{600} was above 0.6 the samples were diluted prior to measuring.

2.16: Creation of artificial *rulAB* containing vector.

2.16.1: Ligation of gene into pCR2.1 vector.

The pWW0 plasmid was extracted from *P. putida* PaW340 (pWW0::km^r) using the Qiagen midi kit (Section 2.3.2). DNA was also extracted from *Psy* B728a and *Pph* 1302A (PPHGI-1). The *rulAB* genes and promoter regions from the 3 strains were amplified via standard PCR primers shown in Table 2.4.

Table 2.4: Primers used to clone various versions of *rulAB.*

Primer name	Sequence 5' - 3'
pWW0rulAB cloning forward	TTGGGGATTCAGCCTTTTACG
pWW0rulAB cloning reverse	TAGCCGTTTTTGGTGAACAGG
pWW0rulB cloning forward	TTGGGGATTCAGCCTTTTACG
pWW0rulB cloning reverse	CGACGTATTTGGCGTGGTCA
pWW0rulAB'-60bp forward	TGATCTTGTCCATCGTGGCC
pWW0rulAB'-60bp reverse	TAGCCGTTTTTGGTGAACAGG
pWW0rulAB'-IP forward	CTCAACTCAGGCCGGAAGAT
pWW0rulAB'-IP reverse	TAGCCGTTTTTGGTGAACAGG
PphGI-1rulAB forward	TGCGCTCCCGTGGCTTGGTG
PphGI-1rulAB reverse	TCGCCATCATGCGGCTCAAGC

Following PCR, 10 ng of PCR product was added to 2 µL of ligation buffer, 1 µL pCR2.1 vector and 1 µL of T4 DNA ligase (TOPO® TA cloning kit, Life Technologies). The mix was made up to 10 µL with water and incubated at 14^{o}C overnight.

2.16.2: Transformation.

The Top10F' competent cells (TOPO® TA cloning kit, Life Technologies) were defrosted and 2 µL of ligation mixtures was added. This was incubated on ice for 30 minutes. Following the incubation the cells were heat shocked at 42^{o}C

for 30 seconds and placed back on ice for 2 minutes. The cells were then incubated with 250 µL of SOC medium for 1 hour at 37ºC with shaking. The cells were then spread onto LB + Km + isopropyl β-D-1-thiogalactopyranoside (IPTG) (0.04 mg/ mL) + 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) (0.04 mg/ mL) (Appendix II) plates and incubated overnight at 37ºC. White colonies were picked and spread on LB + Km plates or placed in LB + Km broth and incubated overnight at 37ºC.

2.16.3: DNA digest of cloned fragments.

Plasmid DNA extraction was performed using the Qiagen mini kit (Section 2.3.2) from the overnight cultures. The DNA was digested using either *EcoR1* or *Xba1* and *Spe1*, 1 µL of restriction enzyme, 3 µL of buffer and 7 µL of water were added to 20 µL of DNA and incubated at 37ºC for 1 hour. The digest was analysed on an agarose gel (Section 2.5) and the cloned DNA was cut out and purified (Section 2.6).

2.16.4: Digesting and dephosphorylating the broad host range vectors.

An overnight culture of the broad host range vector was prepared at 37ºC and the plasmid DNA was extracted (Section 2.3.2). The plasmid DNA was digested as in Section 2.16.3 with both *EcoR1* and *Xba1* and *Spe1*, the digest was also analysed on a gel and purified as in Sections 2.16.3 and 2.16.5. The vector was then dephosphorylated by adding 1 µg of vector to 1 µL FastAP alkaline phosphatase, 2 µL reaction buffer (Thermo Scientific, UK), made up to 20 µL with nuclease-free water and incubated at 37ºC for 10 minutes.

2.16.5: Purification of broad host range vectors.

Following on from the restriction digest and dephosphorylation of the broad host range vector (Section 2.16.4) the vector was purified to remove

contaminants and residual enzymes. One volume of phenol: chloroform: isoamyl alcohol (25:24:1) was added to the vector and vortexed for approximately 20 seconds, this was then centrifuged at 15,000 xg for 5 minutes at room temperature. Following this the upper aqueous phase was carefully removed into a fresh Eppendorf tube. Into the fresh tube 0.5 x sample volume of 4M NaAc was added along with 2x sample volume of 100% EtOH. The sample was then incubated either at -20^{o}C overnight or -80^{o}C for 30 minutes to precipitate the DNA. The sample was then centrifuged at 15,000 xg for 20 minutes at room temperature to pellet DNA and the supernatant was carefully removed. The pellet was washed in 150 µL 70% EtOH and centrifuged again at 15,000 xg for 5 minutes at room temperature. The pellet was air dried for 10 minutes and resuspended in 50 µL TE buffer (Appendix II). The vector was then ready for ligation with the desired insert following the same method as in Section 2.16.1.

2.16.6: Electroporation of construct into desired strains.

The cloned construct containing the intact *rulAB* gene was inserted into a strain that was to be tested via electroporation following the same method outlined in Section 2.12.3. ILE movement was detected using PCR with primers listed in Table 2.5.

Table 2.5: Primers used to confirm ILE movement into cloned *rulAB* versions in pBBR1MCS-2.

Primer name	Sequence 5' - 3'
pWW0rulB F	TGGCGTATGTCGATAACCAG
FH1ILE xerC R	GCCTGCCTTCATTAGTCAGC
2015xerC R	ACCTTGATTTATGTGCACCTGTCGG
ILEFH1SulP F	ACAAAGGCGATCACCACGGC
ILE3023HopAP1 F	TTAATGGCAGGCTGGCCGC
pWW0rulB'' R	ATGCCGACACCAACCGGAAG

2.16.7: Producing chemically competent *E. coli* DH5α cells.

An overnight culture of DH5α was grown and 100 µL of this was removed, placed into 10 mL of fresh LB media and grown at 37°C for 90 minutes. 1.5 mL was removed and centrifuged for 1 min at 15,000 xg. The pellet was washed twice in 750 µL 50mM $CaCl_2$ (ice cold) and left on ice for 30 minutes after which the cells were centrifuged at 15,000 xg for 1 minute and resuspended in 100 µL of 50mM $CaCl_2$ (ice cold). Cells were left on ice until required.

2.16.8: Heat shock transformation of *E. coli* cells.

Following the ligation of the insert into the broad host range vector (pBBR1MCS-2 etc) the total ligation mix (10 µL) was added to the chemically competent DH5α cells and left on ice for 30 minutes. Following this the cells were heat shocked by placing them at 42°C for 2 minutes and then rapidly cooling them in ice for 2 minutes. One mL LB media added and incubated at 37°C for 45 minutes. The cells were centrifuged at 15,000 xg for 1 minute and resuspended in 200 µL LB and plated onto LB +X-gal + IPTG + antibiotic plates and incubated overnight at 37°C.

2.17: ILE circular intermediate tests

ILE strains were tested for ILE circularisation following stress to the cells. These stresses were conjugation with *E. coli* DH5α and *E. coli* DH5α (pRK2013) for 24 hours, cold stress for three hours, UV irradiation for 2 minutes and growth in minimal media broth overnight. A PCR product would only form if circularisation had occurred. The primer sets (Table 2.6) were developed to ascertain whether the ILEs circularise during their movement between *rulB* genes. These primers were used in standard PCR as set out in Section 2.4 and visualised as in Section 2.5.

Table 2.6: Primers used to identify any ILEs that had formed a circular intermediate during movement between *rulB* genes.

Primer name	Sequence 5' - 3'
FH1circ.inter F	TGGCTTCCAGAAAGGCAAAG
FH1circ.inter R	AGCGATGGCTTCAGGGATC
203circ.inter F	TGTGATGGATTCAGCCTCCAG
203circ.inter R	ATACACTCTCCTCGCATTGGG
3023circ.inter F	ATCCAAAGTAGCCGGCGC
3023circ.inter R	ATTGCGAATTACCGTCCGAC

2.18: Frequency of ILE movement into *rulB.*

Two strains were tested to determine the frequency of ILE movement into two forms of *rulB.* The two strains were *P. fluorescens FH1* (pWW0::km^r::ILE$_{FH1}$) and *P. fluorescens FH1* (pBBR1MCS-2::km^r::pWW0*rulAB*). Four colonies from both strains were grown in liquid LB media for 16 hours at 25^{o}C. The cultures were diluted to an OD_{600} of 1.0 using fresh liquid LB media and 1 mL from each colony was mixed together for 1 hour. These mixtures were then serial diluted and spread onto LB + Km agar plates and incubated for 48 hours at 25^{o}C. Colony PCR (Section 2.4) was performed on 94 colonies using the GRrulAB-xerC primers (Table 2.2). The PCR products were analysed using gel electrophoresis (Section 2.5) and converted into a percentage of ILE movement for both strains.

Chapter 3. Screening for integron-like elements and associated regions in *Pseudomonas* bacteria.

3.1: Introduction

As stated in the Introduction the first objective of this research project was to screen a number of *Pseudomonas syringae* strains and other *Pseudomonas* species to establish the frequency of integron-like elements (ILEs) and the ILE associated gene *rulB* within the *Pseudomonas* genus. The outcome of the screen via DNA hybridisation and polymerase chain reaction (PCR) are described here and the DNA sequence analysis in Chapter 4.

When integrons were first discovered in the late 1980's (Stoke and Hall, 1989) it was thought that they were exclusive to clinical isolates of bacteria showing antibiotic resistance, however this was not the case. In 1999 an ILE was identified in the chromosome of the pea plant pathogen, *Pseudomonas syringae* pv. *pisi* (*Ppi*) race 2 strain 203 (Arnold *et al.,* 1999). An insertion into the chromosome of *Ppi* 203 was identified to be an 8.5kb region of DNA that was not present in any other *Ppi* strains tested. The insert was flanked by direct repeat sequences. Within the 8.5kb insert was the resistance to ultra-violet light (*rulAB*) operon which is responsible for UV tolerance and DNA damage repair. It was hypothesised that the *rulB* gene in the operon is a hotspot for ILE insertion as it is disrupted by a 4.3kb insert of DNA which contains the effector gene, *avrPpiA1*, and genes with high similarity to integrase genes which may confer mobility (Arnold *et al*., 2000). The *rulB* gene is plasmid-borne and the ILE insert may explain the distribution of *avrPpiA1* to *Ppi* races 5 and 7 (Arnold *et al*., 2000).

A further ILE was observed in *Pseudomonas fluorescens* strain FH1 where a fragment of DNA excised from the chromosome and inserted into *rulB* on the

plasmid pWW0 during a heavy metal tolerance test (Rhodes *et al.,* 2014). The ILE in *Pseudomonas fluorescens* strain FH1 harboured heavy metal resistance genes such as *sulP* (encodes heavy metal transport protein), regulatory genes including *dksA* and *sdiA*. A *tetR* gene was also present.

The association of ILEs with *rulAB* is broad as similar disruptions have been observed *in silico* in orthologues of *rulAB* such as *ruvAB, rumAB, umuDC, impAB, samAB* and *mucAB* in a wide range of bacterial species (Rhodes *et al.,* 2014; Hochhut *et al.,* 2001; Smith and Walker, 1998). Two examples are; the SXT conjugative element that inserts into *rumB* within *Vibrio cholera,* this conjugative element confers pathogenicity to *Vibrio cholera* (Hochhut *et al.,* 2001) and the insertion of a non-conserved insert into the *umuC* section of *umuDC in* the bacterial species *Pantoea ananatis* (Maayer *et al.,* 2015). The *umuC* insertion is an Integrative and Conjugative Element (ICE) and is 13.7kb. The ICE contains cargo genes associated with oxidative stress and acid stress responses (universal stress protein, UspA), chemotaxis genes and enzymes that play a role in energy production and conversion (Maayer *et al.,* 2015). The ICE insertion event also fits the proposed model of ILE insertion as there are conserved *xerC* orthologues adjacent to the small 5' fragment of *umuC* (*rulB'*).

Due to the apparent association of *rulAB* with ILEs *rulAB* was included in the current screening process and proved valuable in identifying possible ILEs within the *Pseudomonas* species. The hybridisation screening focused on three distinct regions linked to ILEs (Figure 3.1; A1, B and C). The first region (A1) was the associated *rulAB* operon in its intact form with no ILE insertion; this indicated possible hotspots for ILE insertion. The *xerC* gene (B) and the ILE insertion region (C) were also screened during the hybridisation tests.

PCR screening focused on a different *rulAB* section (A1) which would detect a *rulAB'* gene in either its intact or disrupted form and indicate the presence of the *rulAB* operon. The next two regions, *xerC* and *rulAB-xerC* (B and C) were the same regions as used during the hybridisation tests. XerC is an integrase that is conserved in the 5' end of the ILE. This screen gave an indication if an ILE was present, but not definitively as *xerC* integrases can be present in other areas of the genome due to them being broad lambda phage integrases (Blakely *et al.,* 1991). The insertion region between the disrupted *rulB* gene and the *xerC* gene at the start of the ILE would indicate that an ILE had inserted into *rulB.*

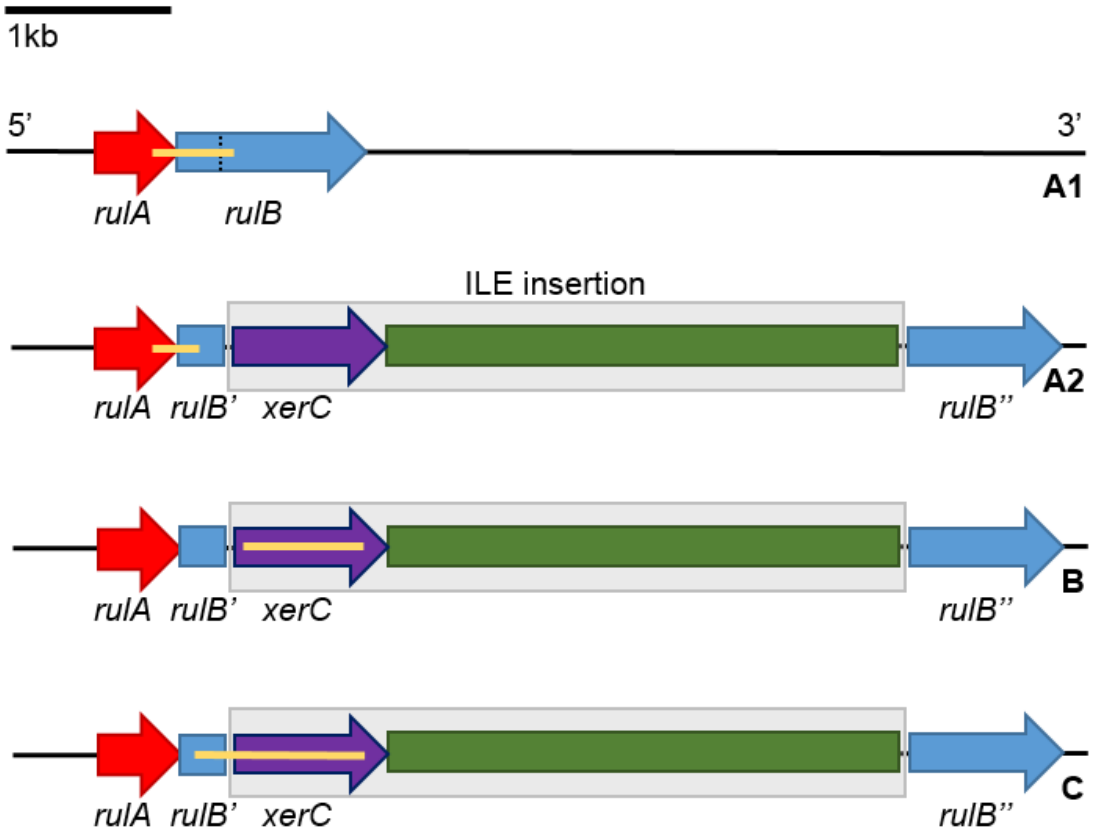


Figure 3.1: Identification of the four ILE related regions. The ILE insert is represented by the grey boxes. The dotted line represents the ILE insertion point and the yellow lines are the amplification targets. **A1)** Intact *rulAB* amplification; **A2)** Presence of *rulAB* but no information on its form; **B)** ILE *xerC* amplification; **C)** *rulAB-xerC* ILE insertion junction, presence of absence of ILE.

The initial screen was performed by DNA hybridisation using fluorescently tagged probes to target the three regions of interest (Figure 3.1; A1, B and C). Positive signals were more likely with hybridisation and would give a preliminary indication of ILE presence. The strains were also screened via PCR with primers derived from *Ppi* race 2 strain 203 (Table 2.2). The advantage of using PCR to screen for ILEs was the ability to sequence the amplifications and confirm the similarity between strains via sequence alignments.

In total 164 plant pathogenic *Pseudomonas* strains were tested with 141 belonging to *Pseudomonas syringae* pathovars, 82 of which were *Ppi* (Table 3.1). Following the screening 22 potential ILE containing strains were identified.

Table 3.1: List of strains, including isolation information, used throughout the screening tests. The table includes isolation location, host organism and year of isolation for the majority of the strains.

No.	Strain	Year isolated	Location	Isolated from
1	Ppi R1 299A	1970	New Zealand	Pea (*Pisum sativum*) cv. *Rondo*
2	Ppi R1 4461			Unknown
3	Ppi R1 862A	1975	USA	Pea (*Pisum sativum*) cv. *Martus*
4	Ppi R1 461	1971	New Zealand	Pea (*Pisum sativum*) cv. *Blue Prussian*
5	Ppi R1 379	1969	Italy	Pea (*Pisum sativum*)
6	Ppi R1 456A	1971	New Zealand	Pea (*Pisum sativum*)
7	Ppi R1 2491B			Unknown
8	Ppi R1 277			Unknown
9	Ppi R2 390			Unknown
10	Ppi R2 1577	1986	France	Pea (*Pisum sativum*) cv. *Hergolt*
11	Ppi R2 1759	1978	UK	Pea (*Pisum sativum*)
12	Ppi R2 2889B			Unknown
13	Ppi R2 202	1944	USA	Pea (*Pisum sativum*)
14	Ppi R2 288			Unknown
15	Ppi R2 223	1968	New Zealand	Pea (*Pisum sativum*) cv. *Partridge*
16	Ppi R2 285	1958	Canada	Pea (*Pisum sativum*)
17	Ppi R2 1124B	1982	USA	Pea (*Pisum sativum*) cv. *Scout*
18	Ppi R2 278	1945	USA	Pea (*Pisum sativum*)
19	Ppi R2 1939	1987	UK	Pea (*Pisum sativum*)
20	Ppi R2 374A	1970	New Zealand	Pea (*Pisum sativum*) cv. *Partridge*
21	Ppi R2 1924			Unknown
22	Ppi R2 1842A			Unknown
23	Ppi R2 1517C	1986	UK	Pea (*Pisum sativum*) cv. *Belinda*
24	Ppi R2 1576A			Unknown
25	Ppi R2 203	1969	New Zealand	Pea (*Pisum sativum*) cv. Small sieve freezer
26	Ppi R3 222			Unknown
27	Ppi R3 283	1970	New Zealand	Unknown
28	Ppi R3 870A	1975	USA	Pea (*Pisum sativum*) cv. *Martus*
29	Ppi R3 895A	1975	USA	Pea (*Pisum sativum*) cv. *Martus*
30	Ppi R3 1125			Unknown
31	Ppi R3 1214			Unknown
32	Ppi R3 1216			Unknown
33	Ppi R3 1380A	1985	USA	Pea (*Pisum sativum*) (seed)
34	Ppi R3 1441	1985	UK	Pea (*Pisum sativum*) cv. *Belinda*
35	Ppi R3 1554A	1987	UK	Pea (*Pisum sativum*) cv. Sprite
36	Ppi R3 1892	1987	USA	Pea (*Pisum sativum*) cv. Avola
37	Ppi R3 2183A			Unknown
38	Ppi R3 2186A			Unknown
39	Ppi R3 2191A	1988	Canada	Pea (*Pisum sativum*) cv. *Rondo*
40	Ppi R3 2817A	1991	Spain	Unknown
41	Ppi R3 4411			Unknown
42	Ppi R3 4574			Unknown
43	Ppi R4 1452	1985	UK	Pea (*Pisum sativum*) cv. *Belinda*
44	Ppi R4 1456A	1985	UK	Pea (*Pisum sativum*) cv. *Belinda*
45	Ppi R4 1456B	1985	UK	Pea (*Pisum sativum*) cv. *Belinda*
46	Ppi R4 1456C	1985	UK	Pea (*Pisum sativum*) cv. *Belinda*
47	Ppi R4 1456D	1985	UK	Pea (*Pisum sativum*) cv. *Belinda*
48	Ppi R4 1456E	1985	UK	Pea (*Pisum sativum*) cv. *Belinda*
49	Ppi R4 1456F	1985	UK	Pea (*Pisum sativum*) cv. *Belinda*
50	Ppi R4 1525			Unknown

51	Ppi R4 1528			Unknown
52	Ppi R4 1758B			Unknown
53	Ppi R4 1811	1987	UK	Pea (*Pisum sativum*) cv. *Green pearl*
54	Ppi R4 1812A	1987	USA	Pea (*Pisum sativum*) cv. *Spain*
55	Ppi R4 2171A			Unknown
56	Ppi R4 5143			Unknown
57	Ppi R5 974B	1978	USA	Pea (*Pisum sativum*) cv. *Puget*
58	Ppi R5 2301C			Unknown
59	Ppi R5 2532A			Unknown
60	Ppi R6 1683	1956	Hungary	Pea (*Pisum sativum*)
61	Ppi R6 1688	1973	France	Pea (*Pisum sativum*)
62	Ppi R6 1704B	1986	France	Pea (*Pisum sativum*) cv. *Stehgolt*
63	Ppi R6 1759	Spontaneous mutant from isolate 870A (race 3) no. 28		
64	Ppi R6 1785A	1987	UK	Pea (*Pisum sativum*) cv. *Countess*
65	Ppi R6 1796A	1987	France	Pea (*Pisum sativum*) cv. *Stehgolt*
66	Ppi R6 1797A	1987	Netherlands	Pea (*Pisum sativum*) cv. *Solara*
67	Ppi R6 1804B	Spontaneous mutant from isolate 1577 (race 2) no. 10		
68	Ppi R6 1807A			Unknown
69	Ppi R6 1842B			Unknown
70	Ppi R6 1842C			Unknown
71	Ppi R6 1745A			Unknown
72	Ppi R6 1746A			Unknown
73	Ppi R6 1755A			Unknown
74	Ppi R6 1842D			Unknown
75	Ppi R6 1942			Unknown
76	Ppi R7 1691	1976	Australia	Pea (*Pisum sativum*)
77	Ppi R7 2491A	1976	Australia	Pea (*Pisum sativum*)
78	Ppi R7 4298			Unknown
79	Ppi R7 4300	1991	UK	Pea (*Pisum sativum*) cv. *Allround*
80	Ppi R7 4409	1991	UK	Pea (*Pisum sativum*) cv. *Bikini*
81	Ppi R7 4426	1991	UK	Pea (*Pisum sativum*) cv. *Allround*
82	Ppi R7 4466			Unknown
83	Psy B728A		Wisconsin, USA	Green Bean (*Phaseolus vulgaris*)
84	Psy 100	1962	Kenya	Butter Bean (*Phaseolus lunatus*)
85	Psy 1142			Unknown
86	Psy 1150			Unknown
87	Psy 1212			Pea (*Pisum sativum*)
88	Psy 1282-8			Unknown
89	Psy 1338A			Unknown
90	Psy 2242A	1988	Zaire	Green Bean (*Phaseolus vulgaris*)
91	Psy 2673C	1990	Lesotho	Green Bean (*Phaseolus vulgaris*)
92	Psy 2675C	1965	Kenya	Okra (*Hibiscus esculentus*)
93	Psy 2676C	1990	Lesotho	Green Bean (*Phaseolus vulgaris*)
94	Psy 2677C	1990	Lesotho	Green Bean (*Phaseolus vulgaris*)
95	Psy 2682C	1990	Lesotho	Green Bean (*Phaseolus vulgaris*)
96	Psy 2692C	1990	Zimbabwe	Green Bean (*Phaseolus vulgaris*)
97	Psy 2703C	1990	Zimbabwe	Green Bean (*Phaseolus vulgaris*)
98	Psy 2732A	1990	Colombia	Green Bean (*Phaseolus vulgaris*)
99	Psy 3023	1950	UK	Lilac (*Syringa vulgaris*)
100	Pph 103			Unknown
101	Pph R1 1281A	1984	UK	Runner Bean (*Phaseolus coccineus*)
102	Pph R6 1299A	1984	Tanzania	Green Bean (*Phaseolus vulgaris*)
103	Pph R3 1301A	1984	Tanzania	Green Bean (*Phaseolus vulgaris*)
104	Pph R4 1302A	1984	Ethiopia	Green Bean (*Phaseolus vulgaris*)
105	Pph R5 1375A	1985	Kenya	Hyacinth bean (*Lablab purpureus*)
106	Pph R7 1449B	1985	Ethiopia	Hyacinth bean (*Lablab purpureus*)
107	Pph R6 1448A	1985	Ethiopia	Green Bean (*Phaseolus vulgaris*)
108	Pph R7 1449A	1985	Ethiopia	Hyacinth bean (*Lablab purpureus*)
109	Pph R8 2656A	1990	Lesotho	Green Bean (*Phaseolus vulgaris*)
110	Pph R9 2709A	1990	Malawi	Green Bean (*Phaseolus vulgaris*)

111	Pma M4	1965	UK	Radish (*Raphanus sativus*)
112	Pma 65			Unknown
113	Pma 1809A			Unknown
114	Pma 1813	1967	UK	Cauliflower (*Brassica oleracea*) cv. *Danish perfection*
115	Pma 1820	1966	USA	Radish (*Raphanus sativus*)
116	Pma 1821A	1967	UK	Cauliflower (*Brassica oleracea*) cv. *Danish perfection*
117	Pma 1838A	1987	UK	Cauliflower (*Brassica oleracea*) cv. *Danish perfection*
118	Pma 1846A	1987	UK	Cauliflower (*Brassica oleracea*) cv. *King*
119	Pma 1848B	1987	UK	Cauliflower (*Brassica oleracea*) cv. *White rock*
120	Pma 1852A	1987	UK	Brussel Sprout (*Brassica oleracea*) cv. *Oliver*
121	Pma 1855A	1987	UK	Cauliflower (*Brassica oleracea*) cv. *Danish perfection*
122	Pma 5422	1995	UK	Cauliflower (*Brassica oleracea*) cv. *Danish perfection*
123	Pma 5429		UK	Cauliflower (*Brassica oleracea*)
124	Pma 6201	1987	Portugal	Kale (*Brassica oleracea*)
125	Pma 6319A/1			Unknown
126	Pma 6328A/1			Unknown
127	Pto DC3000		UK	Tomato (*Solanum lycopersicum*)
128	Pto 19	1961	UK	Tomato (*Solanum lycopersicum*)
129	Pto 119			Unknown
130	Pto 138			Unknown
131	Pto 1108	1960	UK	Tomato (*Solanum lycopersicum*)
132	Pto 2944	1961	UK	Tomato (*Solanum lycopersicum*)
133	Pto 2945		France	Tomato (*Solanum lycopersicum*)
134	Pto 6034			Unknown
135	Pat 152E	1968	UK	Snapdragon (*Antirrhinum majus*)
136	Pat 4303	1965	UK	Snapdragon (*Antirrhinum majus*)
137	Pla 789			Unknown
138	Pla 3988	1935	USA	Cucumber (*Cucumis sativus*)
139	Pgy 1139	1962	Zimbabwe	Soybean (*Glycine javanica*)
140	Pgy 2411	1971	New Zealand	Soybean (*Glycine max*)
141	Pgy 3318	1984	Italy	Soybean (*Glycine max*)
142	P. asplenii 959	1961		Bird's nest fern (*Asplenium nidus*)
143	P. asplenii 1947	1966		Bird's nest fern (*Asplenium nidus*)
144	P. caricapapayae 1873	1966	Brazil	Papaya (*Carica papaya*)
145	P. caricapapayae 3080	1979	Brazil	Papaya (*Carica papaya*)
146	P. caricapapayae 3439	1985	Brazil	Papaya (*Carica papaya*)
147	P. savastanoi 639	1959	Yugoslavia	Olive (*Olea europaea*)
148	P. savastanoi 2716	1975	UK	Ash (*Fraxinus excelsior*)
149	P. savastanoi 3334	1984	France	Oleander (*Nerium oleander*)
150	P. corrugata 2445	1972	UK	Tomato (*Solanum lycopersicum*)
151	P. corrugata 3056	1978	USA	Alfafa (*Medicago sativa*)
152	P. corrugata 3316	1984	S. Africa	Tomato (*Solanum lycopersicum*)
153	P. cichorii 907	1961	USA	Florist's daisy (*Chrysanthemum morifolium*)
154	P. cichorii 943	1961	Germany	Endive (*Cichorium endivia*)
155	P. cichorii 3109A	1979	Brazil	Coffee (*Coffea arabica*)
156	P. cichorii 3109B	1979	Brazil	Coffee (*Coffea arabica*)
157	P. cichorii 3283	1983	UK	Florist's daisy (*Chrysanthemum morifolium*)
158	P. marginalis 247	1924	USA	Lettuce (*Lactuca sativa*)
159	P. marginalis 949	1961	USA	Parsnip (*Pastinaca sativa*)
160	P. marginalis 2380	1971	UK	Lettuce (*Lactuca sativa*)
161	P. marginalis 2644	1974	USA	Alfafa (*Medicago sativa*)
162	P. marginalis 2645	1974	USA	Alfafa (*Medicago sativa*)
163	P. marginalis 2646	1974	USA	Alfafa (*Medicago sativa*)
164	P. marginalis 3210	1982	Canada	Broad bean (*Vicia faba*)

Ppi **= pv. p*isi*** ***Psy*** **= pv. s*yringae*** ***Pph*** **= pv. p*haseolicola*** ***Pma*** **= pv. m*aculicola***
Pto **= pv. t*omato*** ***Pat*** **= pv. a*ntirrhini*** ***Pla*** **= pv. l*achrymans*** ***Pgy*** **= pv. g*lycinea***

3.2: Results

3.2.1: Confirming that *P.putida* PaW340 has no *rulAB* gene.

Firstly, a suitable negative control against *rulAB* was established. *P. putida* PaW340 was used as the negative control as it is a wild type strain that contains no *rulAB* operon. This was demonstrated via PCR (Figure 3.2) to confirm the absence of *rulAB.* The PCR was performed using the GR primers in Table 2.2.

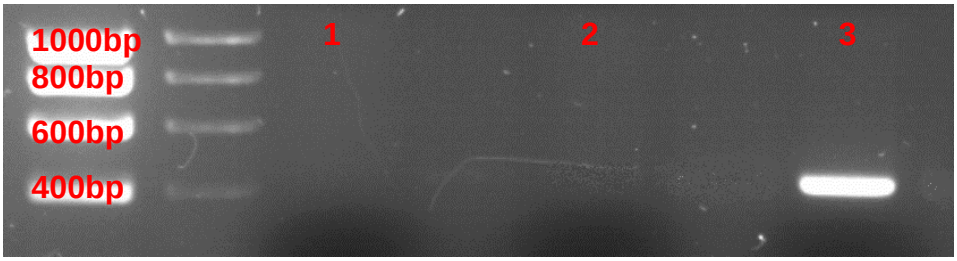


Figure 3.2: PCR confirmation that *P. putida* PaW340 contains no *rulAB* gene. The PCR results confirm that *P. putida* PaW340 (-) contains no *rulAB* gene as no bands were visible on the gel, the only band seen was the positive control which was the correct size of 423bp. Hyperladder 1kb (Bioline, UK) was used a size marker.

No.	PCR Product
1	*PaW 340* with the GR*rulAB* primers. (Cell culture)
2	*PaW 340* with the GR*rulAB* primers. (DNA)
3	*PaW 340* (pWW0::km^r) with the GR*rulAB* primers. Positive control.

3.2.2: Strain confirmation for DNA hybridisation probe construction.

The strains used to produce the chemiluminescent probes had their ILE and *rulAB* properties checked via PCR to ensure the strains contained the regions needed for the screening (Figure 3.3). The strains were also checked via antibiotic resistance screening (Table 3.2) with appropriate antibiotics and media. The control strains were supplied by Dr Glenn Rhodes (Centre for Ecology and Hydrology, Lancaster).

Table 3.2: Differentiation of strains used to make hybridisation probes on selective media.

Bacterial Strain	Probe region	Antibiotic Resistance	Media conditions
***P. putida* PaW340 (pWW0::km^r)**	-Intact *rulAB*	Stm, Km	requires tryptophan to grow
***P. fluorescens FH1* (pWW0::km^r::ILE$_{FH1}$)**	*-rulB-xerC* ILE junction *-xerC* within ILE	Km	N/A

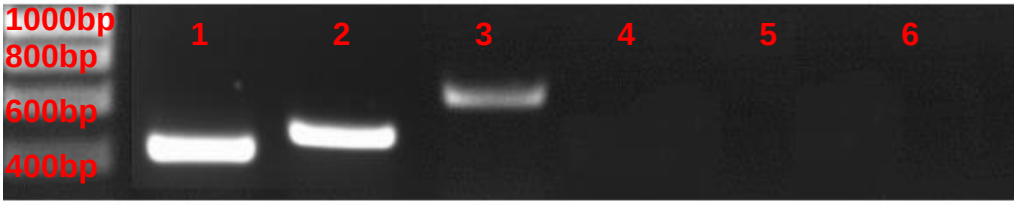


Figure 3.3: PCR amplification of *rulAB, rulAB-xerC and xerC* regions from *P. putida* PaW340 (pWW0::km^r) and *P. fluorescens* FH1 (pWW0::km^r::ILE$_{FH1}$). The PCR amplifications confirm that the strains used for hybridisation probes and as controls contain the correct ILE associated genes. Primers used were the GR primers from Table 2.2. Hyperladder 1kb (Bioline, UK) was used a size marker

1- *P. putida* PaW340 (pWW0::km^r) [*rulAB*];
2- *P. fluorescens* FH1 (pWW0::km^r::ILE$_{FH1}$) [*xerC*];
3- *P. fluorescens* FH1 (pWW0::km^r::ILE$_{FH1}$) [*rulAB-xerC*];
4- *P. putida* PaW340 (pWW0::km^r) [*xerC*];
5- *P. putida* PaW340 (pWW0::km^r) [*rulAB-xerC*];
6- *P. fluorescens* FH1 (pWW0::km^r::ILE$_{FH1}$) [*rulAB*].

The results from the confirmation PCR (Figure 3.3) confirmed that the strains show the correct *rulAB* (423bp), *xerC* (501bp) and *rulAB-xerC* (590bp) profiles for the strains as expected. *P. putida* PaW340 (pWW0::km^r) shows amplification of the *rulAB* region due to the strain containing an intact *rulAB* on pWW0 (Greated *et al.,* 2002). No band was seen for the *rulAB-xerC* or the *xerC* primers as they are not present due to the lack of an ILE. The *rulAB-xerC* junction and *xerC* integrase primers show that *P. fluorescens* FH1 (pWW0::km^r::ILE$_{FH1}$)

contains an ILE as expected. No band was present for the intact *rulAB* as it was disrupted by the ILE.

3.2.3: DNA hybridisation screening for ILEs and associated regions.

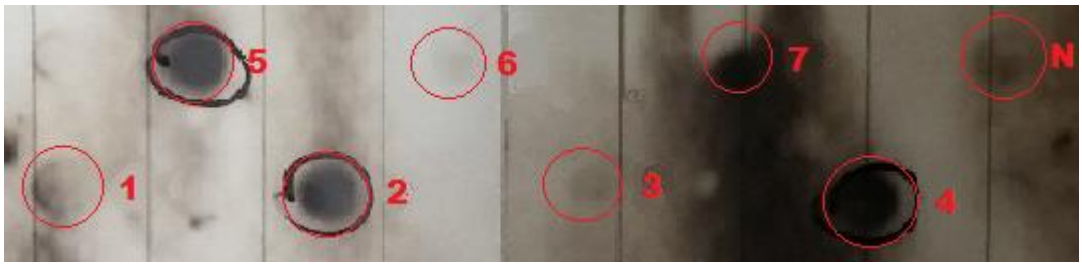
The first objective was to identify how widespread ILEs are amongst plant pathogenic *Pseudomonas* strains. Initial screening were carried out using a colony dot blot hybridisation method that used a chemiluminescent probe for detection.

3.2.4: Producing DIG labelled probes.

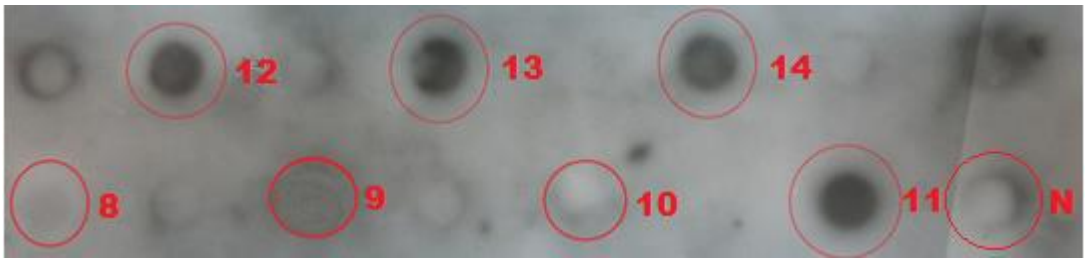
Firstly the relevant DNA regions were amplified to create the hybridisation probes. The *rulAB* gene, the *xerC* gene and *rulAB-xerC* junction region (Figure 3.3; lanes1, 2 and 3 respectively) were amplified via standard PCR. The amplified products were cleaned up and their purity and quantity measured via a Nanodrop 1000 to ensure effective labelling as described in Section 2.8.

3.2.5: Control hybridisations to ensure *rulAB*, *xerC* and *rulAB-xerC* probes hybridise.

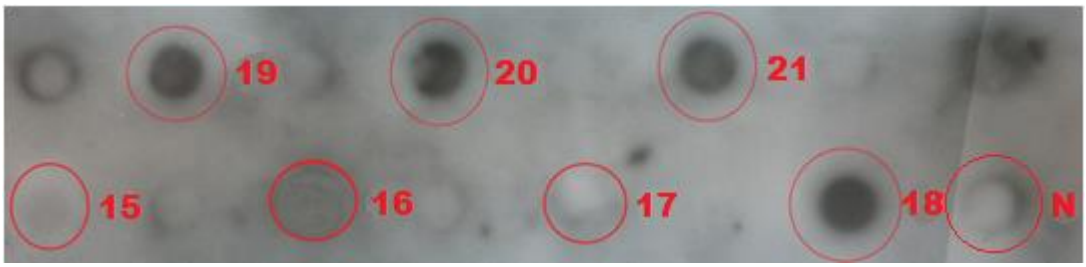
The *rulAB, xerC* and *rulAB-xerC* probes were first tested on control strains of known ILE content and the strains used to produce the hybridisation probes, *P. putida* PaW340 (pWW0::km^r) and *P. fluorescens* FH1 pWW0::km^r::ILE$_{FH1}$) respectively (Figure 3.3). The control tests comprised of one blot per probe (Figure 3.4). It was essential to carry out these control hybridisations to check that the labelling reaction had been successful and to ensure that the hybridisation method was a viable approach to screening for ILEs amongst the *Pseudomonas* strains.



The *rulAB* control dot blot:
1- *P. putida* PaW340 (no *rulAB,* no signal);
2- *P. putida* PaW340 (pWW0::km^r) (positive control, *rulAB* intact on plasmid);
3- *P. putida* PaW340 (pWW0::km^{r}Δ*rulAB*) (*rulAB* has been deleted*,* no signal);
4- *P. fluorescens* FH1 (pWW0::km^r::ILE$_{FH1}$) (disrupted *rulAB* present, signal present);
5- *P. fluorescens* FH4 (pWW0::km^r::ILE$_{FH4}$) (disrupted *rulAB* present, signal present);
6- *P. fluorescens* FH1 (disrupted different form of *rulAB* on chromosome, no signal);
7- *P. fluorescens* FH4 (disrupted different form of *rulAB* on chromosome, no signal).



The *xerC* control dot blot:
8- *P. putida* PaW340 (no ILE *xerC,* no signal);
9- *P. putida* PaW340 (pWW0::km^r) (no ILE *xerC,* no signal);
10- *P. putida* PaW340 (pWW0::km^{r}Δ*rulAB*) (no ILE *xerC,* no signal);
11- *P. fluorescens* FH1 (pWW0::km^r::ILE$_{FH1}$) (positive control, *xerC* present in plasmid ILE);
12- *P. fluorescens* FH4 (pWW0::km^r::ILE$_{FH4}$) (positive control, *xerC* present in plasmid ILE);
13- *P. fluorescens* FH1 (positive control, *xerC* present in chromosomal ILE);
14- *P. fluorescens* FH4 (positive control, *xerC* present in chromosomal ILE).



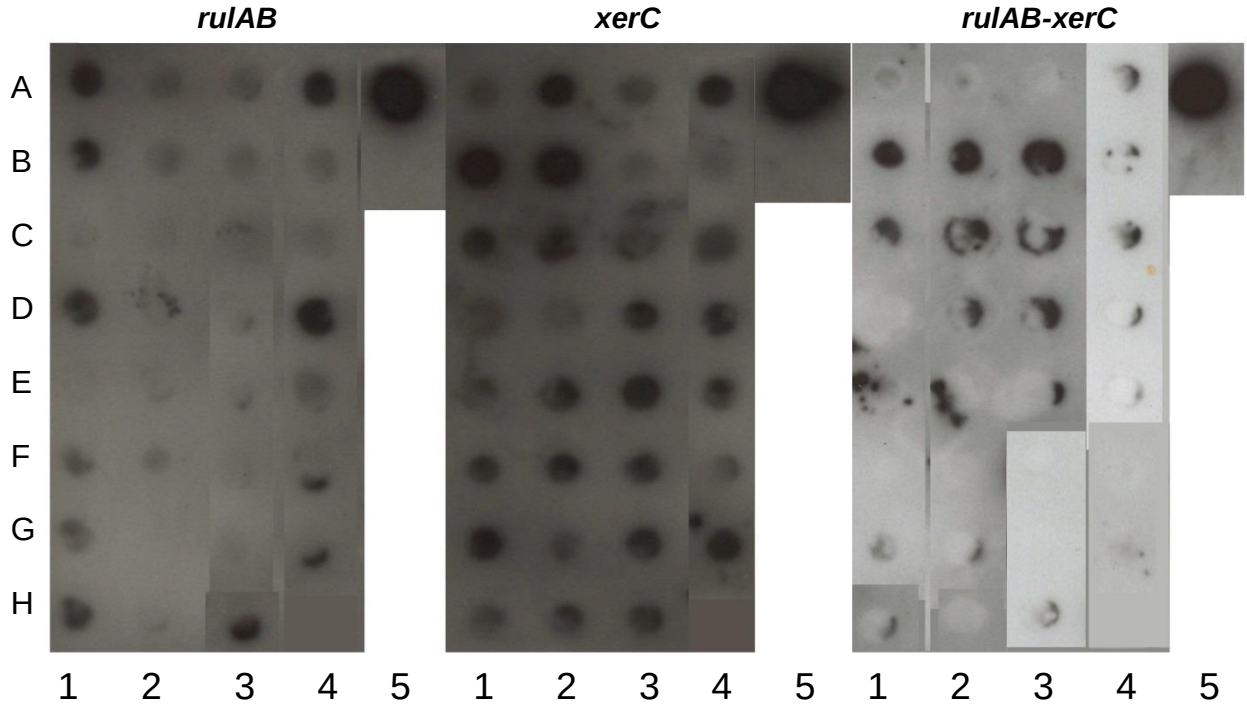
The *rulAB-xerC* ILE insert junction control dot blot:
15- *P. putida* PaW340 (no ILE*,* no signal);
16- *P. putida* PaW340 (pWW0::km^r) (no ILE*,* no signal);
17- *P. putida* PaW340 (pWW0::km^{r}Δ*rulAB*) (no ILE*,* no signal);
18- *P. fluorescens* FH1 (pWW0::km^r::ILE$_{FH1}$) (positive control, ILE present);
19- *P. fluorescens* FH4 (pWW0::km^r::ILE$_{FH4}$) (positive control, ILE present);
20- *P. fluorescens* FH1 (positive control, ILE present);
21- *P. fluorescens* FH4 (positive control, ILE present).

Figure 3.4: Hybridisation of the three ILE probes on control *P. putida* and *P. fluorescens* strains. N- Negative control.

In the control hybridisations (Figure 3.4) the negative controls (position N) showed no hybridisation signals. The *rulAB* probe successfully hybridised to the strain it was constructed from (position 2). *P. fluorescens* FH1 (pWW0::km^r::ILE$_{FH1}$) (position 4) and *P. fluorescens* FH4 (pWW0::km^r::ILE$_{FH4}$) (position 5) showed some signal from the *rulAB* probe. The *xerC* probe successfully hybridised to strains that were known to contain *xerC* regions (positions 12, 13 and 14) and the strain that the probe was constructed from (position 11). The *P. putida* strains lacking a *xerC* site showed no hybridisation as expected. The final probe, *rulAB-xerC* showed hybridisation signals from the strains with an ILE present either in the plasmid (positions 18 and 19) or on the chromosome (positions 20 and 21). No hybridisation occurred in strains that did not contain an ILE (positions 15, 16 and 17). The control hybridisations confirmed the probes had been labelled correctly and that the probes were hybridising as expected.

3.2.6: Hybridisation of *rulAB*, *xerC* and *rulAB-xerC* probes to *Pseudomonas syringae* pv. *pisi* strains.

A total of 82 *Ppi* strains were screened with a selection of representatives from all seven races. All of the *Ppi* races provided hybridisation signals with varying intensity across the three probes. Races three and four provided some variation in signal strength (Figure 3.5). The example dot blots (Figure 3.5) showed that multiple *Ppi* strains from both races three and four share some degree of identity with either the *rulAB* region on pWW0::km^r, the *P. fluorescens* FH1 ILE *xerC* region or the *P. fluorescens* FH1 (pWW0::km^r::ILE$_{FH1}$) *rulAB-xerC* junction region. Some signals appear darker due to the probes sharing more identity to the strain being screened which results in more of the probe hybridising.



	1	2	3	4	5
A	*Ppi* race 3 strain 222	*Ppi* race 3 strain 1441	*Ppi* race 3 strain 4574	*Ppi* race 4 strain 1525	Positive control
B	*Ppi* race 3 strain 283	*Ppi* race 3 strain 1554A	*Ppi* race 4 strain 1452	*Ppi* race 4 strain 1528	Negative control
C	*Ppi* race 3 strain 870A	*Ppi* race 3 strain 1892	*Ppi* race 4 strain 1456A	*Ppi* race 4 strain 1758B	
D	*Ppi* race 3 strain 895A	*Ppi* race 3 strain 2183A	*Ppi* race 4 strain 1456B	*Ppi* race 4 strain 1811A	
E	*Ppi* race 3 strain 1125	*Ppi* race 3 strain 2186A	*Ppi* race 4 strain 1456C	*Ppi* race 4 strain 1812A	
F	*Ppi* race 3 strain 1214	*Ppi* race 3 strain 2191A	*Ppi* race 4 strain 1456D	*Ppi* race 4 strain 2171A	
G	*Ppi* race 3 strain 1216	*Ppi* race 3 strain 2817A	*Ppi* race 4 strain 1456E	*Ppi* race 4 strain 5143	
H	*Ppi* race 3 strain 1380A	*Ppi* race 3 strain 4411	*Ppi* race 4 strain 1456F		

Figure 3.5: Dot blot hybridisations of the *rulAB, xerC* and *rulAB-xerC* probes to *Ppi* strains belonging to races three and four. The positive controls, *P. putida* PaW340 (pWW0::km^r) [*rulAB*] and *P. fluorescens FH1* (pWW0::km^r::ILE$_{FH1}$) [*xerC* and *rulAB-xerC*] show strong hybridisation for their respective probes. The negative controls remained clear. The *rulAB-xerC* blot has been comprised from two different images.

Following the hybridisations the signal intensity for all of the *pisi* strains was visually scored (Table 3.3). The scoring was zero for no hybridisation to four which indicated strong hybridisation in line with the positive control.

Table 3.3: Hybridisation result with *rulAB, xerC* and *rulAB-xerC probes* from *P. syringae* pv. *pisi.* strains. Visual scoring 0 to 4 with 0 representing no signal and 4 representing strong hybridisation equal to positive control. Red numbers are considered as showing strong hybridisation.

Region		intact *rulAB* region			*xerC* region			disrupted *rulB* region		
Probe		*rulAB*			*xerC*			*rulAB-xerC*		
No.	Strain	Rep. 1	Rep. 2	Overall	Rep. 1	Rep. 2	Overall	Rep. 1	Rep. 2	Overall
1	Ppi R1 299A	1	2	**1.5**	1	2	**1.5**	0	1	**0.5**
2	Ppi R1 4461	0	1	**0.5**	1	1	**1**	1	2	**1.5**
3	Ppi R1 862A	0	0	**0**	1	1	**1**	1	1	**1**
4	Ppi R1 461	1	1	**1**	1	1	**1**	1	3	**2**
5	Ppi R1 379	1	1	**1**	2	0	**1**	0	0	**0**
6	Ppi R1 456A	0	1	**0.5**	1	1	**1**	0	4	**2**
7	Ppi R1 2491B	0	0	**0**	1	1	**1**	0	0	**0**
8	Ppi R1 277	0	1	**0.5**	2	2	**2**	1	0	**0.5**
9	Ppi R2 390	3	3	**3**	0	0	**0**	2	0	**1**
10	Ppi R2 1577	2	2	**2**	0	0	**0**	0	1	**0.5**
11	Ppi R2 1759	2	2	**2**	0	0	**0**	0	0	**0**
12	Ppi R2 2889B	1	1	**1**	1	1	**1**	0	1	**0.5**
13	Ppi R2 202	3	2	**2.5**	3	1	**2**	1	0	**0.5**
14	Ppi R2 288	1	1	**1**	1	1	**1**	0	1	**0.5**
15	Ppi R2 223	1	2	**1.5**	1	1	**1**	0	3	**1.5**
16	Ppi R2 285	0	0	**0**	1	1	**1**	1	2	**1.5**
17	Ppi R2 1142B	3	2	**2.5**	0	0	**0**	2	0	**1**
18	Ppi R2 278	1	1	**1**	0	0	**0**	0	0	**0**
19	Ppi R2 1939	0	2	**1**	2	2	**2**	0	2	**1**
20	Ppi R2 374A	2	2	**2**	0	1	**0.5**	0	2	**1**
21	Ppi R2 1924	2	2	**2**	1	0	**0.5**	0	0	**0**
22	Ppi R2 1842A	2	0	**1**	0	0	**0**	0	4	**2**
23	Ppi R2 1517C	1	0	**0.5**	0	0	**0**	0	0	**0**
24	Ppi R2 1576A	0	1	**0.5**	2	2	**2**	1	0	**0.5**
25	Ppi R2 203	3	2	**2.5**	3	3	**3**	1	2	**1.5**
26	Ppi R3 222	2	2	**2**	1	1	**1**	2	1	**1.5**
27	Ppi R3 283	2	2	**2**	4	4	**4**	0	4	**2**
28	Ppi R3 870A	0	0	**0**	4	3	**3.5**	1	3	**2**
29	Ppi R3 895A	2	0	**1**	1	1	**1**	0	0	**0**
30	Ppi R3 1125	0	0	**0**	1	0	**0.5**	0	1	**0.5**
31	Ppi R3 1214	1	1	**1**	2	2	**2**	0	1	**0.5**
32	Ppi R3 1216	1	1	**1**	3	1	**2**	1	1	**1**
33	Ppi R3 1380A	2	2	**2**	1	1	**1**	1	1	**1**
34	Ppi R3 1441	2	2	**2**	4	4	**4**	0	1	**0.5**
35	Ppi R3 1554A	1	1	**1**	4	4	**4**	0	3	**1.5**
36	Ppi R3 1892	0	0	**0**	3	1	**2**	1	2	**1.5**
37	Ppi R3 2183A	0	0	**0**	1	1	**1**	1	2	**1.5**
38	Ppi R3 2186A	1	3	**2**	2	2	**2**	0	0	**0**
39	Ppi R3 2191A	1	1	**1**	2	2	**2**	0	0	**0**
40	Ppi R3 2817A	0	0	**0**	1	1	**1**	0	0	**0**
41	Ppi R3 4411	0	1	**0.5**	2	2	**2**	0	0	**0**
42	Ppi R3 4574	1	1	**1**	1	1	**1**	2	0	**1**
43	Ppi R4 1452	0	0	**0**	1	1	**1**	0	3	**1.5**
44	Ppi R4 1456A	1	1	**1**	1	0	**0.5**	0	1	**0.5**
45	Ppi R4 1456B	0	0	**0**	3	1	**2**	0	2	**1**

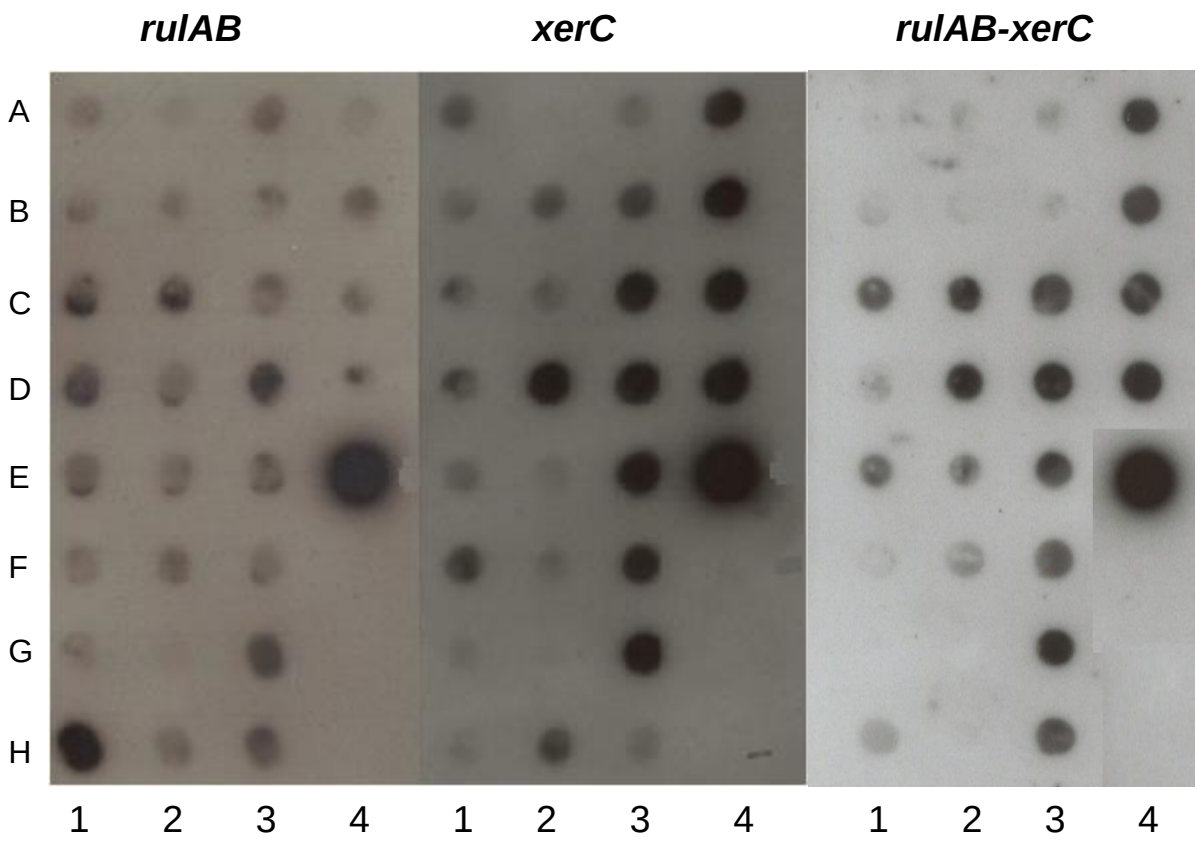
46	Ppi R4 1456C	0	0	**0**		**3**	**3**	**3**		0	1	**0.5**
47	Ppi R4 1456D	0	0	**0**		**3**	**3**	**3**		0	2	**1**
48	Ppi R4 1456E	0	0	**0**		**4**	2	**3**		0	2	**1**
49	Ppi R4 1456F	**4**	2	**3**		2	2	**2**		1	2	**1.5**
50	Ppi R4 1525	**4**	2	**3**		**4**	2	**3**		1	2	**1.5**
51	Ppi R4 1528	1	1	**1**		0	0	**0**		0	2	**1**
52	Ppi R4 1758B	1	0	**0.5**		**3**	1	**2**		1	0	**0.5**
53	Ppi R4 1811	**4**	2	**3**		**3**	2	**2.5**		0	0	**0**
54	Ppi R4 1812A	0	0	**0**		2	2	**2**		0	1	**0.5**
55	Ppi R4 2171A	0	0	**0**		1	1	**1**		0	1	**0.5**
56	Ppi R4 5143	1	1	**1**		**3**	**3**	**3**		0	0	**0**
57	Ppi R5 974B	**4**	2	**3**		2	2	**2**		**4**	2	**3**
58	Ppi R5 2301C	**4**	2	**3**		1	1	**1**		**3**	**3**	**3**
59	Ppi R5 2532A	**3**	1	**2**		2	1	**1.5**		0	2	**1**
60	Ppi R6 1683	2	1	**1.5**		**3**	2	**2.5**		**4**	0	**2**
61	Ppi R6 1688	**3**	1	**2**		1	1	**1**		1	0	**0.5**
62	Ppi R6 1704B	1	1	**1**		2	0	**1**		2	1	**1.5**
63	Ppi R6 1759	2	2	**2**		0	0	**0**		0	1	**0.5**
64	Ppi R6 1785A	**3**	1	**2**		1	0	**0.5**		0	1	**0.5**
65	Ppi R6 1796A	**4**	**4**	**4**		1	1	**1**		2	2	**2**
66	Ppi R6 1797A	1	0	**0.5**		0	0	**0**		1	1	**1**
67	Ppi R6 1804B	2	1	**1.5**		1	1	**1**		1	**4**	**2.5**
68	Ppi R6 1807A	1	1	**1**		1	1	**1**		0	2	**1**
69	Ppi R6 1842B	**4**	0	**2**		**3**	**3**	**3**		0	0	**0**
70	Ppi R6 1842C	1	1	**1**		0	0	**0**		0	1	**0.5**
71	Ppi R6 1745A	2	1	**1.5**		0	0	**0**		0	1	**0.5**
72	Ppi R6 1746A	**3**	**3**	**3**		0	0	**0**		0	**3**	**1.5**
73	Ppi R6 1755A	1	1	**1**		2	1	**1.5**		**3**	**3**	**3**
74	Ppi R6 1842D	0	0	**0**		1	1	**1**		2	**3**	**2.5**
75	Ppi R6 1942	0	0	**0**		1	1	**1**		1	2	**1.5**
76	Ppi R7 1691	**3**	1	**2**		**3**	2	**2.5**		1	**3**	**2**
77	Ppi R7 2491A	**3**	1	**2**		2	2	**2**		0	1	**0.5**
78	Ppi R7 4298	2	1	**1.5**		2	2	**2**		0	0	**0**
79	Ppi R7 4300	**3**	1	**2**		1	0	**0.5**		0	2	**1**
80	Ppi R7 4409	2	1	**1.5**		1	0	**0.5**		0	**3**	**1.5**
81	Ppi R7 4426	**4**	**4**	**4**		1	0	**0.5**		1	**3**	**2**
82	Ppi R7 4466	**3**	1	**2**		0	1	**0.5**		2	2	**2**

The *Ppi* results showed that there is a wide variety of *rulAB, xerC* and *rulAB-xerC* profiles among the strains. The minority of the screened strains showed no hybridisation signals for the probes, meaning the target gene was not identified. This occurred for 20.7% (17) of the strains screened with the *rulAB* probe; 15.9% (13) of the strains showed no hybridisation with the *xerC* probe and 18.3% (15) had no hybridisation with the *rulAB-xerC* probe. However, the majority of strains did display hybridisation. Strong hybridisation was determined as any strain that had a hybridisation score of three or four on replicate blots (indicated by red numbering in Table 3.4). Of the 82 *Ppi* strains screened 14.6% (12) presented strong *rulAB* hybridisation, 17.1% (14) of strains presented strong *xerC* hybridisation and 4.9% (4) presented strong *rulAB-xerC* hybridisation. The

remaining strains did show hybridisation, but with weaker signal levels. The most interesting strains were ones that had strong signals for the *rulAB-xerC* probe as this indicated the potential presence of an ILE.

3.2.7: Hybridisation of *rulAB* and *xerC* probes to *P. syringae* pv. *syringae* and *P. syringae* pv. *phaseolicola* strains.

P. syringae pv. s*yringae* (*Psy*) and *P. syringae* pv. *phaseolicola* (*Pph*) strains were screened for the same three regions as above. Hybridisation of the probes was observed in 17 *Psy* strains and 11 *Pph* strains (Figure 3.6). The signals were scored in the same manner as the *Ppi* strains (Table 3.5).



	1	2	3	4
A	*Psy* strain B728a	*Psy* strain 2673C	*Psy* strain 3023	*Pph* race 6 strain 1448A
B	*Psy* strain 100	*Psy* strain 2675C	*Pph* strain 103	*Pph* race 7 strain 1449A
C	*Psy* strain 1142	*Psy* strain 2676C	*Pph* race 1 strain 1281A	*Pph* race 8 strain 2656A
D	*Psy* strain 1150	*Psy* strain 2677C	*Pph* race 6 strain 1299A	*Pph* race 9 strain 2709A
E	*Psy* strain 1212	*Psy* strain 2682C	*Pph* race 3 strain 1301A	Positive control
F	*Psy* strain 1282-8	*Psy* strain 2692C	*Pph* race 4 strain 1302A	Negative control
G	*Psy* strain 1338A	*Psy* strain 2703C	*Pph* race 5 strain 1375A	
H	*Psy* strain 2242A	*Psy* strain 2732A	*Pph* race 7 strain 1449B	

Figure 3.6: Hybridisations of the *rulAB, xerC* and *rulAB-xerC* probes to *Psy and Pph* strains. The positive controls, *P. putida* PaW340 (pWW0::km^r) [*rulAB*] and *P. fluorescens FH1* (pWW0::km^r::ILE$_{FH1}$) [*xerC* and *rulAB-xerC*] show strong hybridisation for their respective probes. The negative no DNA spots remained clear. The *rulAB-xerC* blot has been comprised from two different images.

Table 3.4: Hybridisation results with *rulAB, xerC* and *rulAB-xerC probes* from *P. syringae* pv. *syringae* and *phaseolicola* strains. Visual scoring 0 to 4 with 0 representing no signal and 4 representing strong hybridisation equal to positive control. Red numbers are considered as showing strong hybridisation.

Region		intact *rulAB* region				*xerC* region				disrupted *rulB* region		
Probe		*rulAB*				*xerC*				*rulAB-xerC*		
83	Psy B728A	1	1	**1**		2	2	**2**		0	1	**0.5**
84	Psy 100	1	1	**1**		1	1	**1**		1	1	**1**
85	Psy 1142	**3**	1	**2**		1	1	**1**		2	1	**1.5**
86	Psy 1150	**3**	**3**	**3**		2	2	**2**		1	0	**0.5**
87	Psy 1212	2	1	**1.5**		1	1	**1**		2	0	**1**
88	Psy 1282-8	1	1	**1**		**3**	**3**	**3**		0	1	**0.5**
89	Psy 1338A	1	0	**0.5**		0	0	**0**		0	0	**0**
90	Psy 2242A	**4**	**4**	**4**		0	0	**0**		2	2	**2**
91	Psy 2673C	0	0	**0**		0	0	**0**		1	0	**0.5**
92	Psy 2675C	1	1	**1**		2	2	**2**		1	1	**1**
93	Psy 2676C	**3**	1	**2**		1	1	**1**		**3**	1	**2**
94	Psy 2677C	1	1	**1**		**4**	**4**	**4**		**3**	2	**2.5**
95	Psy 2682C	1	1	**1**		0	0	**0**		2	1	**1.5**
96	Psy 2692C	1	1	**1**		2	**3**	**2.5**		2	1	**1.5**
97	Psy 2703C	0	0	**0**		0	0	**0**		0	0	**0**
98	Psy 2732A	1	1	**1**		2	2	**2**		0	0	**0**
99	Psy 3023	2	2	**2**		1	1	**1**		1	1	**1**
100	Pph 103	1	0	**0.5**		**3**	2	**2.5**		0	0	**0**
101	Pph R1 1281A	1	0	**0.5**		**4**	**4**	**4**		**3**	2	**2.5**
102	Pph R6 1299A	**4**	**4**	**4**		**4**	**4**	**4**		**3**	2	**2.5**
103	Pph R3 1301A	2	1	**1.5**		**4**	**4**	**4**		2	2	**2**
104	Pph R4 1302A	**3**	**3**	**3**		**4**	**4**	**4**		2	2	**2**
105	Pph R5 1375A	3	1	**2**		**4**	**4**	**4**		**3**	**4**	**3.5**
106	Pph R7 1449B	2	2	**2**		1	2	**1.5**		2	0	**1**
107	Pph R6 1448A	1	1	**1**		**4**	**4**	**4**		**3**	2	**2.5**
108	Pph R7 1449A	2	2	**2**		**4**	**4**	**4**		**3**	2	**2.5**
109	Pph R8 2656A	1	1	**1**		**4**	**4**	**4**		**3**	1	**2**
110	Pph R9 2709A	1	2	**1.5**		**4**	**4**	**4**		**4**	2	**3**

The hybridisation results from both *Psy* and *Pph* (Table 3.4) show clear differences between the strains from both pathovars in regard to their ILE and *rulAB* profiles. In the *Psy* strains 88% (15) showed hybridisation of the *rulAB* probe with 12% (2) showing strong hybridisation. The *xerC* probe hybridised to 71% (12) of the *Psy* strains with 18% (3) showing strong hybridisation of the *xerC* probe. The final probe, *rulAB-xerC,* hybridised to 82% (14) of the *Psy* strains but only 6% (1) showed strong hybridisation.

Of the *Pph* strains, 100% (11) of the strains showed hybridisation of the *rulAB* probe with 18% (2) showing strong hybridisation. The same was also true for the *xerC* probe as 100% (11) of strains showed hybridisation to the probe, but unlike the *rulAB* probe 91% (10) showed strong hybridisation of the *xerC* probe.

The *rulAB-xerC* probe also hybridised to the *Pph* strains with 91% (10) showing hybridisation and 55% (6) showing strong hybridisation.

3.2.8: Hybridisations of *rulAB, xerC* and *rulAB-xerC* probes to *P. syringae* pathovars *maculicola, tomato, antirrhini, lachrymans* and *glycinea.*

The results from the remaining *Pseudomonas syringae* pathovars (Table 3.5) show varying degrees of hybridisation between the different pathovars. *P. syringae* pv. *maculicola* (*Pma*) showed 63% (10) hybridisation of the *rulAB* probe, 19% (3) of which was strong hybridisation. The *xerC* probe successfully hybridised to 88% (14) of the *Pma* strains with 31% (5) being stronger hybridisations. Finally 81% (13) of the *Pma* strains showed hybridisation of the *rulAB-xerC* probe with 19% (3) being strong hybridisations.

The four remaining pathovars, *tomato* (*Pto*), *antirrhini* (*Pat*), *lachrymans* (*Pla*) and *glycinea* (*Pgy*) all showed some hybridisation to the probes with between 33% and 100% hybridisation of the *rulAB* probe, between 75% and 100% hybridisation of the *xerC* probe with the exception of the *Pat* strains which showed 0% hybridisation of the *xerC* probe. The majority of the remaining strains also showed hybridisation of the *rulAB-xerC* probe with values ranging between 66% and 100%, again with the exception of the *Pat* strains which showed 0% hybridisation of the *rulAB-xerC* probe. Although there was some degree of hybridisation of the three probes across the four remaining pathovars, none of the results provided strong hybridisation.

Table 3.5: Hybridisation results with *rulAB*, *xerC* and *rulAB-xerC* probes from various *P.syringae* pathovars. Visual scoring 0 to 4 with 0 representing no signal and 4 representing strong hybridisation equal to positive control. Red numbers are considered as showing strong hybridisation.

Region		intact *rulAB* region				*xerC* region				disrupted *rulB* region		
Probe		*rulAB*				*xerC*				*rulAB-xerC*		
111	Pma M4	1	1	**1**		1	1	**1**		0	0	**0**
112	Pma 65	1	1	**1**		1	2	**1.5**		0	1	**0.5**
113	Pma 1809A	**4**	**4**	**4**		**3**	**3**	**3**		1	0	**0.5**
114	Pma 1813	1	1	**1**		**3**	**3**	**3**		1	0	**0.5**
115	Pma 1820	0	0	**0**		1	1	**1**		1	1	**1**
116	Pma 1821A	0	0	**0**		2	1	**1.5**		0	0	**0**
117	Pma 1838A	1	0	**0.5**		1	1	**1**		1	0	**0.5**
118	Pma 1846A	1	0	**0.5**		0	1	**0.5**		2	1	**1.5**
119	Pma 1848B	2	1	**1.5**		**4**	**4**	**4**		**4**	**4**	**4**
120	Pma 1852A	**3**	2	**2.5**		**4**	**4**	**4**		**4**	**4**	**4**
121	Pma 1855A	0	0	**0**		0	2	**1**		0	0	**0**
122	Pma 5422	**4**	**4**	**4**		1	2	**1.5**		0	1	**0.5**
123	Pma 5429	0	0	**0**		1	1	**1**		0	1	**0.5**
124	Pma 6201	2	1	**1.5**		**4**	**3**	**3.5**		**3**	**4**	**3.5**
125	Pma 6319A/1	0	0	**0**		0	0	**0**		1	2	**1.5**
126	Pma 6328A/1	0	0	**0**		0	0	**0**		0	**4**	**2**
127	Pto DC3000	1	1	**1**		2	1	**1.5**		0	**4**	**2**
128	Pto 19	0	1	**0.5**		0	1	**0.5**		0	3	**1.5**
129	Pto 119	1	1	**1**		2	1	**1.5**		1	2	**1.5**
130	Pto 138	1	0	**0.5**		2	1	**1.5**		1	1	**1**
131	Pto 1108	1	0	**0.5**		0	0	**0**		1	0	**0.5**
132	Pto 2944	0	1	**0.5**		1	0	**0.5**		0	1	**0.5**
133	Pto 2945	1	0	**0.5**		0	0	**0**		0	0	**0**
134	Pto 6034	1	0	**0.5**		2	1	**1.5**		0	1	**0.5**
135	Pat 152E	0	1	**0.5**		0	0	**0**		0	0	**0**
136	Pat 4303	1	0	**0.5**		0	0	**0**		0	0	**0**
137	Pla 789	1	1	**1**		0	1	**0.5**		1	1	**1**
138	Pla 3988	2	1	**1.5**		1	2	**1.5**		0	2	**1**
139	Pgy 1139	1	0	**0.5**		1	1	**1**		1	2	**1.5**
140	Pgy 2411	0	0	**0**		1	1	**1**		0	1	**0.5**
141	Pgy 3318	0	0	**0**		1	1	**1**		0	0	**0**

3.2.9: Hybridisation of *rulAB*, *xerC* and *rulAB-xerC* probes to different *Pseudomonas* species.

The final set of hybridisation screening consisted of screening strains belonging to six different plant pathogenic *Pseudomonas* species (Table 3.6). The six species were *Pseudomonas asperii, P. caricapapayae, P. savastanoi, P. corrugata, P. cichorii* and *P. marginalis.* These species of *Pseudomonas* showed weak hybridisation, no strong hybridisation was observed with any of the probes. The total number of hybridisations of the *rulAB* probe across the six different

species was 65% (15), the *xerC* probe hybridised to 61% (14) of the strains and 39% (9) showed hybridisation of the *rulAB-xerC* probe.

Table 3.6: Hybridisation results with *rulAB, xerC* and *rulAB-xerC* probes from various *Pseudomonas* strains. Visual scoring 0 to 4 with 0 representing no signal and 4 representing strong hybridisation equal to positive control.

Region		intact *rulAB* region				*xerC* region				disrupted *rulB* region		
Probe		*rulAB*				*xerC*				*rulAB-xerC*		
142	P. asplenii 959	1	1	**1**		1	0	**0.5**		1	0	**0.5**
143	P. asplenii 1947	0	0	**0**		0	0	**0**		0	0	**0**
144	P. caricapapayae 1873	0	0	**0**		1	1	**1**		0	0	**0**
145	P. caricapapayae 3080	1	0	**0.5**		1	1	**1**		0	2	**1**
146	P. caricapapayae 3439	0	0	**0**		1	2	**1.5**		0	2	**1**
147	P. savastanoi 639	0	0	**0**		0	0	**0**		0	0	**0**
148	P. savastanoi 2716	1	0	**0.5**		1	1	**1**		1	1	**1**
149	P. savastanoi 3334	0	1	**0.5**		1	1	**1**		0	0	**0**
150	P. corrugata 2445	1	1	**1**		1	1	**1**		0	0	**0**
151	P. corrugata 3056	2	1	**1.5**		1	0	**0.5**		0	0	**0**
152	P. corrugata 3316	1	1	**1**		0	0	**0**		0	0	**0**
153	P. cichorii 907	1	1	**1**		0	0	**0**		0	0	**0**
154	P. cichorii 943	1	1	**1**		1	1	**1**		0	1	**0.5**
155	P. cichorii 3109A	1	1	**1**		1	1	**1**		1	2	**1.5**
156	P. cichorii 3109B	0	1	**0.5**		0	0	**0**		0	0	**0**
157	P. cichorii 3283	0	0	**0**		0	1	**0.5**		0	0	**0**
158	P. marginalis 247	1	1	**1**		0	1	**0.5**		0	0	**0**
159	P. marginalis 949	0	0	**0**		1	0	**0.5**		0	0	**0**
160	P. marginalis 2380	0	1	**0.5**		0	0	**0**		0	0	**0**
161	P. marginalis 2644	0	0	**0**		0	0	**0**		0	0	**0**
162	P. marginalis 2645	1	1	**1**		1	0	**0.5**		0	1	**0.5**
163	P. marginalis 2646	0	0	**0**		0	0	**0**		0	1	**0.5**
164	P. marginalis 3210	1	1	**1**		0	0	**0**		1	1	**1**

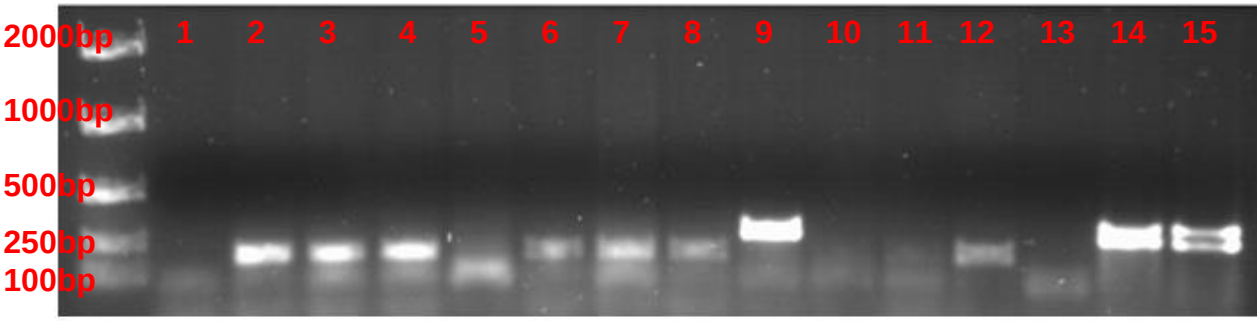
3.2.10: PCR screening for ILEs and associated regions in *Pseudomonas syringae* pathovars and *Pseudomonas* species.

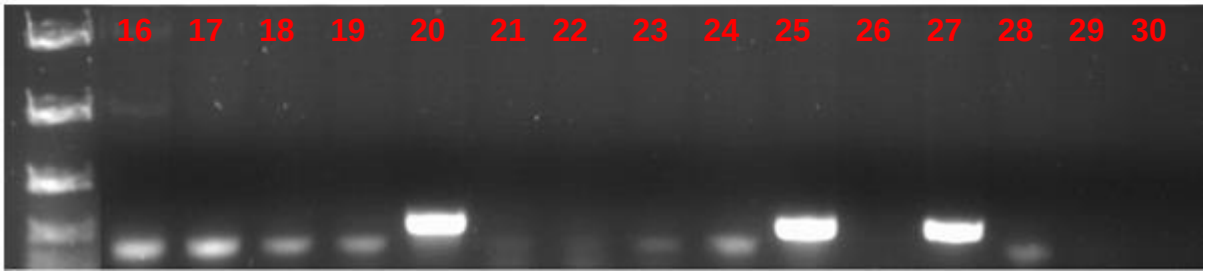
The same *Pseudomonas* strains that were screened via hybridisation were also screened for potential ILEs and *rulAB,* via PCR. PCR was used alongside DNA hybridisation due to PCR being more specific and having the ability to sequence the results downstream. The regions screened for via PCR were similar to the regions screened in the hybridisation screens, *rulAB, xerC* and *rulAB-xerC,* the only difference from the hybridisation being that the *rulAB'* primer only

indicated partial *rulAB* presence. The primers used were derived from *Ppi* race 2 strain 203. *Ppi* race 2 strain 203 is known to harbour an ILE inserted into *rulAB* (Arnold *et al.,* 2000). The PCR screening was carried out for two reasons, firstly to double check the hybridisation results and secondly to allow downstream sequence analysis.

3.2.11: PCR screening of *rulAB* genes in *Pseudomonas syringae* pathovars.

PCR screening of the *rulAB'* region produced varying results across the different pathovars and species. Figure 3.7 shows an example of the PCR results from *Ppi* races one and two. The majority of the positive amplifications came from the *pisi* pathovar with 12% (10) showing amplification. A further 6% (1) of the *Psy* strains showed amplification and 9% (1) of the *Pph* strains showed amplification (Table 3.7). The remaining *P. syringae* pathovars showed no amplification of the *rulAB'* region when using the *pisi* derived primers, this was also true for the different *Pseudomonas* species.





No.	Screened strain	No.	Screened strain
1	*Ppi* race 1 strain 299A	16	*Ppi* race 2 strain 285
2	*Ppi* race 1 strain 4461	17	*Ppi* race 2 strain 1142B
3	*Ppi* race 1 strain 862A	18	*Ppi* race 2 strain 278
4	*Ppi* race 1 strain 461	19	*Ppi* race 2 strain 1939
5	*Ppi* race 1 strain 379	20	*Ppi* race 2 strain 374A
6	*Ppi* race 1 strain 456A	21	*Ppi* race 2 strain 1924
7	*Ppi* race 1 strain 2491B	22	*Ppi* race 2 strain 1842A
8	*Ppi* race 1 strain 277	23	*Ppi* race 2 strain 1517C
9	*Ppi* race 2 strain 390	24	*Ppi* race 2 strain 1576A
10	*Ppi* race 2 strain 1577	25	*Ppi* race 2 strain 203 (+)
11	*Ppi* race 2 strain 1759	26	*PaW* 340 (-)
12	*Ppi* race 2 strain 2889B	27	*Ppi* race 3 strain 283
13	*Ppi* race 2 strain 202	28	*Ppi* race 3 strain 370A

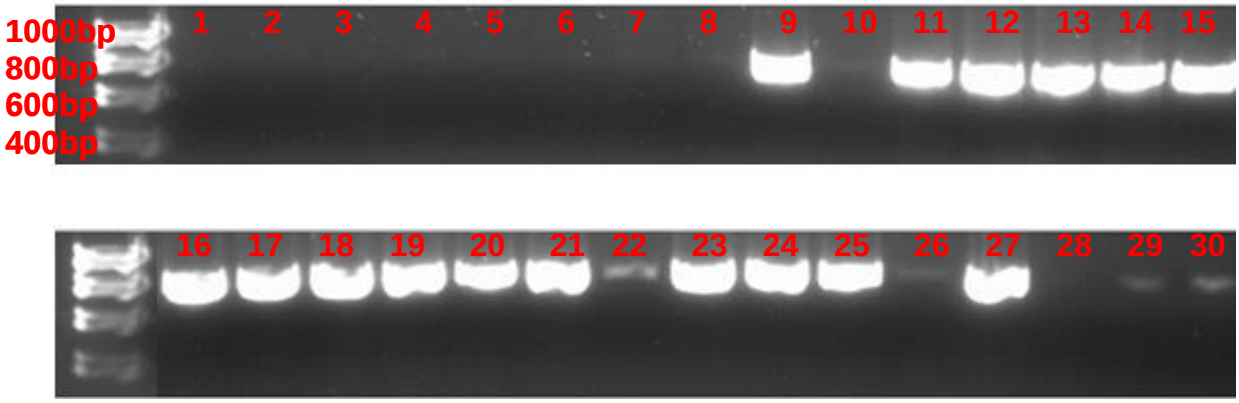
Figure 3.7: PCR Amplification of the *rulAB'* region before the ILE insertion point from *P. syringae* pv. *pisi* races 1, 2 and 3. The primers used produce an amplified fragment of 273bp in length. Amplification using these primers indicate the presence of a *rulAB* operon in the bacteria but not whether it is intact or disrupted and therefore no indication if an ILE is present. The positive control was *Ppi* race 2 strain 203 containing a disrupted *rulB* gene and the negative control, *PaW* 340(-), contains no *rulAB.* Easyladder 1 (Bioline, UK) was used a size marker.

Table 3.7: PCR screening of the *rulAB'* region for *P. syringae* pathovars. The remaining 74 strains screened have been omitted due to zero amplification observed.

No.	Strain	PCR amp.	No.	Strain	PCR amp.	No.	Strain	PCR amp.
1	Ppi R1 299A	-	38	Ppi R3 2186A	-	75	Ppi R6 1942	-
2	Ppi R1 4461	-	39	Ppi R3 2191A	-	76	Ppi R7 1691	-
3	Ppi R1 862A	-	40	Ppi R3 2817A	-	77	Ppi R7 2491A	-
4	Ppi R1 461	-	41	Ppi R3 4411	-	78	Ppi R7 4298	-
5	Ppi R1 379	-	42	Ppi R3 4574	-	79	Ppi R7 4300	-
6	Ppi R1 456A	-	43	Ppi R4 1452	-	80	Ppi R7 4409	-
7	Ppi R1 2491B	-	44	Ppi R4 1456A	-	81	Ppi R7 4426	-
8	Ppi R1 277	-	45	Ppi R4 1456B	-	82	Ppi R7 4466	-
9	Ppi R2 390	+	46	Ppi R4 1456C	-	83	Psy B728A	-
10	Ppi R2 1577	-	47	Ppi R4 1456D	-	84	Psy 100	-
11	Ppi R2 1759	-	48	Ppi R4 1456E	-	85	Psy 1142	-
12	Ppi R2 2889B	-	49	Ppi R4 1456F	-	86	Psy 1150	-
13	Ppi R2 202	-	50	Ppi R4 1525	-	87	Psy 1212	-
14	Ppi R2 288	+	51	Ppi R4 1528	-	88	Psy 1282-8	-
15	Ppi R2 223	+	52	Ppi R4 1758B	+	89	Psy 1338A	-
16	Ppi R2 285	-	53	Ppi R4 1811	-	90	Psy 2242A	-
17	Ppi R2 1124B	-	54	Ppi R4 1812A	-	91	Psy 2673C	-
18	Ppi R2 278	-	55	Ppi R4 2171A	-	92	Psy 2675C	-
19	Ppi R2 1939	-	56	Ppi R4 5143	-	93	Psy 2676C	-
20	Ppi R2 374A	+	57	Ppi R5 974B	-	94	Psy 2677C	-
21	Ppi R2 1924	-	58	Ppi R5 2301C	-	95	Psy 2682C	-
22	Ppi R2 1842A	-	59	Ppi R5 2532A	-	96	Psy 2692C	-
23	Ppi R2 1517C	-	60	Ppi R6 1683	-	97	Psy 2703C	-
24	Ppi R2 1576A	-	61	Ppi R6 1688	-	98	Psy 2732A	-
25	Ppi R2 203	+	62	Ppi R6 1704B	-	99	Psy 3023	+
26	Ppi R3 222	-	63	Ppi R6 1759	-	100	Pph 103	-
27	Ppi R3 283	+	64	Ppi R6 1785A	+	101	Pph R1 1281A	-
28	Ppi R3 370A	-	65	Ppi R6 1796A	-	102	Pph R6 1299A	-
29	Ppi R3 895A	-	66	Ppi R6 1797A	-	103	Pph R3 1301A	-
30	Ppi R3 1125	-	67	Ppi R6 1804B	-	104	Pph R4 1302A	-
31	Ppi R3 1214	-	68	Ppi R6 1807A	+	105	Pph R5 1375A	+
32	Ppi R3 1216	-	69	Ppi R6 1842B	-	106	Pph R7 1449B	-
33	Ppi R3 1380A	-	70	Ppi R6 1842C	-	107	Pph R6 1448A	-
34	Ppi R3 1441	-	71	Ppi R6 1745A	-	108	Pph R7 1449A	-
35	Ppi R3 1554A	-	72	Ppi R6 1746A	+	109	Pph R8 2656A	-
36	Ppi R3 1892	-	73	Ppi R6 1755A	-	110	Pph R9 2709A	-
37	Ppi R3 2183A	-	74	Ppi R6 1842D	-			

3.2.12: PCR screening of *xerC* genes in *Pseudomonas syringae* pathovars.

Following on from the screening of the *rulAB'* region in the *Pseudomonas* strains the same strains were subjected to PCR screening for the *xerC* gene. The *xerC* gene gives a possible indication to ILE presence within the strain, but does not confirm it. An example of the PCR results is shown in Figure 3.8. The results were different to the *rulAB'* PCR screening as more strains were positive for *xerC* genes. In this screen both pathovars *pisi* and *maculicola* had positive amplification of the *xerC* gene across 50% (41 and 8 respectively) of the screened strains. This was not the highest percentage as *Pto* showed 63% (5) amplification of the gene. *Pgy* 2411 also showed amplification. The final pathovar to show amplification was the *syringae* strains which had 12% (2) amplification of the *xerC* gene. Outside of the *P. syringae* pathovars two different *Pseudomonas* species had amplification of the *xerC* gene with one *P. cichorii* strain and one *P. marginalis* showing amplification. Full *xerC* results can be found in Table 3.8.



No.	Screened strain	No.	Screened strain
1	*Ppi* race 1 strain 299A	16	*Ppi* race 2 strain 285
2	*Ppi* race 1 strain 4461	17	*Ppi* race 2 strain 1142B
3	*Ppi* race 1 strain 862A	18	*Ppi* race 2 strain 278
4	*Ppi* race 1 strain 461	19	*Ppi* race 2 strain 1939
5	*Ppi* race 1 strain 379	20	*Ppi* race 2 strain 374A
6	*Ppi* race 1 strain 456A	21	*Ppi* race 2 strain 1924
7	*Ppi* race 1 strain 2491B	22	*Ppi* race 2 strain 1842A
8	*Ppi* race 1 strain 277	23	*Ppi* race 2 strain 1517C
9	*Ppi* race 2 strain 390	24	*Ppi* race 2 strain 1576A
10	*Ppi* race 2 strain 1577	25	*Ppi* race 2 strain 203 (+)
11	*Ppi* race 2 strain 1759	26	*PaW* 340 (-)
12	*Ppi* race 2 strain 2889B	27	*Ppi* race 3 strain 283
13	*Ppi* race 2 strain 202	28	*Ppi* race 3 strain 370A

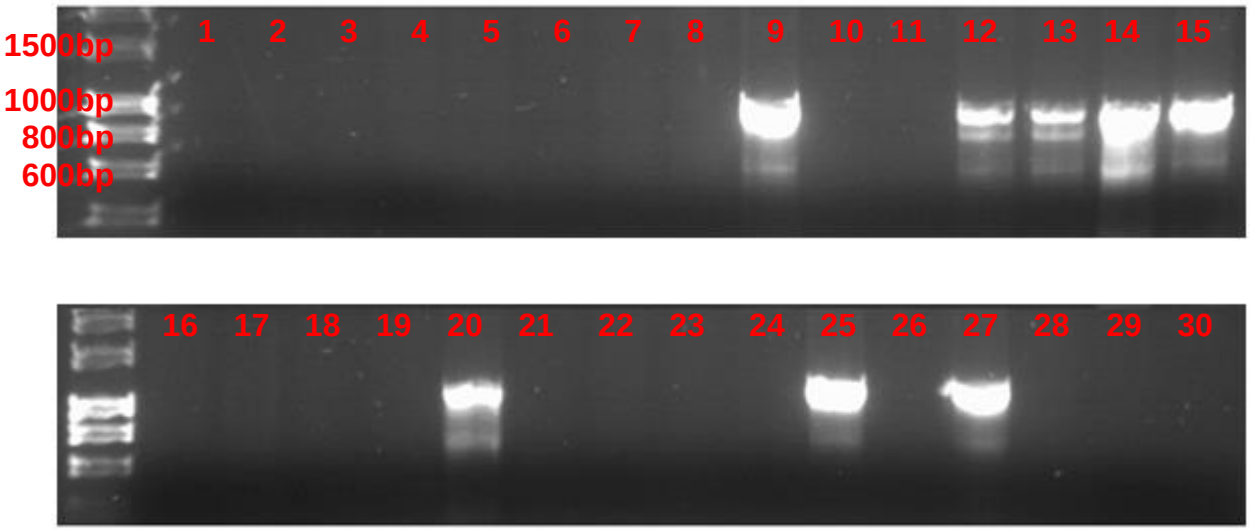
Figure 3.8: PCR Amplification of the *xerC* gene from *P. syringae* pv. *pisi* races 1, 2 and 3. The primers used produce an amplified fragment of 804bp in length. Amplification of the *xerC* may indicate the presence of an ILE. The positive control was *Ppi* race 2 strain 203 which contains a *xerC* gene. The negative control, *PaW* 340(-), contains no *xerC.* Hyperladder 1kb (Bioline, UK) was used a size marker.

Table 3.8: PCR screening of the *xerC* gene in *P. syringae* pathovars and *Pseudomonas* species. The remaining 18 strains have been omitted due to zero amplification.

No.	Strain	PCR amp.	No.	Strain	PCR amp.	No.	Strain	PCR amp.
1	Ppi R1 299A	-	50	Ppi R4 1525	-	99	Psy 3023	-
2	Ppi R1 4461	-	51	Ppi R4 1528	-	100	Pph 103	-
3	Ppi R1 862A	-	52	Ppi R4 1758B	-	101	Pph R1 1281A	-
4	Ppi R1 461	-	53	Ppi R4 1811	-	102	Pph R6 1299A	-
5	Ppi R1 379	-	54	Ppi R4 1812A	-	103	Pph R3 1301A	-
6	Ppi R1 456A	-	55	Ppi R4 2171A	-	104	Pph R4 1302A	-
7	Ppi R1 2491B	-	56	Ppi R4 5143	-	105	Pph R5 1375A	-
8	Ppi R1 277	-	57	Ppi R5 974B	-	106	Pph R7 1449B	-
9	Ppi R2 390	+	58	Ppi R5 2301C	+	107	Pph R6 1448A	-
10	Ppi R2 1577	-	59	Ppi R5 2532A	+	108	Pph R7 1449A	-
11	Ppi R2 1759	+	60	Ppi R6 1683	-	109	Pph R8 2656A	-
12	Ppi R2 2889B	+	61	Ppi R6 1688	-	110	Pph R9 2709A	-
13	Ppi R2 202	+	62	Ppi R6 1704B	-	111	Pma M4	-
14	Ppi R2 288	+	63	Ppi R6 1759	-	112	Pma 65	+
15	Ppi R2 223	+	64	Ppi R6 1785A	+	113	Pma 1809A	-
16	Ppi R2 285	+	65	Ppi R6 1796A	-	114	Pma 1813	-
17	Ppi R2 1142B	+	66	Ppi R6 1797A	-	115	Pma 1820	+
18	Ppi R2 278	+	67	Ppi R6 1804B	-	116	Pma 1821A	-
19	Ppi R2 1939	+	68	Ppi R6 1807A	-	117	Pma 1838A	-
20	Ppi R2 374A	+	69	Ppi R6 1842B	-	118	Pma 1846A	-
21	Ppi R2 1924	+	70	Ppi R6 1842C	-	119	Pma 1848B	+
22	Ppi R2 1842A	-	71	Ppi R6 1745A	-	120	Pma 1852A	+
23	Ppi R2 1517C	+	72	Ppi R6 1746A	-	121	Pma 1855A	+
24	Ppi R2 1576A	+	73	Ppi R6 1755A	-	122	Pma 5422	+
25	Ppi R2 203	+	74	Ppi R6 1842D	-	123	Pma 5429	+
26	Ppi R3 222	-	75	Ppi R6 1942	-	124	Pma 6201	-
27	Ppi R3 283	+	76	Ppi R7 1691	+	125	Pma 6319A/1	+
28	Ppi R3 870A	-	77	Ppi R7 2491A	-	126	Pma 6328A/1	-
29	Ppi R3 895A	-	78	Ppi R7 4298	+	127	Pto DC3000	+
30	Ppi R3 1125	-	79	Ppi R7 4300	+	128	Pto 19	+
31	Ppi R3 1214	-	80	Ppi R7 4409	+	129	Pto 119	-
32	Ppi R3 1216	+	81	Ppi R7 4426	+	130	Pto 138	+
33	Ppi R3 1380A	+	82	Ppi R7 4466	+	131	Pto 1108	-
34	Ppi R3 1441	+	83	Psy B728A	-	132	Pto 2944	+
35	Ppi R3 1554A	+	84	Psy 100	-	133	Pto 2945	-
36	Ppi R3 1892	-	85	Psy 1142	-	134	Pto 6034	+
37	Ppi R3 2183A	+	86	Psy 1150	-	135	Pgy 1139	-
38	Ppi R3 2186A	+	87	Psy 1212	+	136	Pgy 2411	+
39	Ppi R3 2191A	+	88	Psy 1282-8	-	137	Pgy 3318	-
40	Ppi R3 2817A	+	89	Psy 1338A	-	138	P. cichorii 943	+
41	Ppi R3 4411	+	90	Psy 2242A	-	139	P. cichorii 3109A	-
42	Ppi R3 4574	-	91	Psy 2673C	-	140	P. cichorii 3109B	-
43	Ppi R4 1452	+	92	Psy 2675C	-	141	P. cichorii 3283	-
44	Ppi R4 1456A	+	93	Psy 2676C	-	142	P. marginalis 247	-
45	Ppi R4 1456B	+	94	Psy 2677C	-	143	P. marginalis 949	-
46	Ppi R4 1456C	+	95	Psy 2682C	-	144	P. marginalis 2380	-
47	Ppi R4 1456D	+	96	Psy 2692C	+	145	P. marginalis 2644	-
48	Ppi R4 1456E	+	97	Psy 2703C	-	146	P. marginalis 2645	+
49	Ppi R4 1456F	+	98	Psy 2732A	-	147	P. marginalis 2646	-
						148	P. marginalis 3210	-

3.2.13: PCR screening of the *rulAB-xerC* junction region in *Pseudomonas syringae* pathovars and other *Pseudomonas* species.

The last set of PCR amplifications focused on the ILE insertion junction into *rulB*. Any amplification would indicate ILE insertion into *rulB.* Following the PCR tests the products were visualised on an agarose gel to identify positives and confirm the fragment size, an example is shown in Figure 3.9. This screening identified 22 potential ILEs. Again *Ppi* strains provided the highest percentage of positive results with 18.3% (15) of the strains showing amplification, the *Psy* pathovar only had one strain that showed amplification. Out of the 16 *Pma* strains tested 18.8% (3) showed amplification and *Pgy* which showed one strain with amplification of the *rulAB-xerC* region. Full results for these pathovars are shown in Table 3.9. These results were also confirmed using genome analysis on some strains that were previously sequenced including, *Pph* 1448A, *Pph* 1302A and *Pto* DC3000.



No.	Screened strain	No.	Screened strain
1	*Ppi* race 1 strain 299A	16	*Ppi* race 2 strain 285
2	*Ppi* race 1 strain 4461	17	*Ppi* race 2 strain 1142B
3	*Ppi* race 1 strain 862A	18	*Ppi* race 2 strain 278
4	*Ppi* race 1 strain 461	19	*Ppi* race 2 strain 1939
5	*Ppi* race 1 strain 379	20	*Ppi* race 2 strain 374A
6	*Ppi* race 1 strain 456A	21	*Ppi* race 2 strain 1924
7	*Ppi* race 1 strain 2491B	22	*Ppi* race 2 strain 1842A
8	*Ppi* race 1 strain 277	23	*Ppi* race 2 strain 1517C
9	*Ppi* race 2 strain 390	24	*Ppi* race 2 strain 1576A
10	*Ppi* race 2 strain 1577	25	*Ppi* race 2 strain 203 (+)
11	*Ppi* race 2 strain 1759	26	*PaW* 340 (-)
12	*Ppi* race 2 strain 2889B	27	*Ppi* race 3 strain 283
13	*Ppi* race 2 strain 202	28	*Ppi* race 3 strain 370A

Figure 3.9: PCR amplification of the *rulAB-xerC* junction region from *P. syringae* pv. *pisi* races 1, 2 and 3. The primers used produce an amplified fragment 1048bp in length. Any amplification of this region would indicate the possible presence of an ILE inserted into the *rulB* gene. The positive control was *Ppi* race 2 strain 203 which contains a *rulAB-xerC* junction. The negative control, *PaW* 340(-), contains no *rulAB* and no ILE.

Table 3.9: PCR screening of the *rulAB-xerC* junction region in *P. syringae* pathovars. The remaining 46 strains have been omitted due to zero amplification.

No.	Strain	PCR amp.	No.	Strain	PCR amp.	No.	Strain	PCR amp.
1	Ppi R1 299A	-	40	Ppi R3 2817A	-	79	Ppi R7 4300	-
2	Ppi R1 4461	-	41	Ppi R3 4411	-	80	Ppi R7 4409	-
3	Ppi R1 862A	-	42	Ppi R3 4574	-	81	Ppi R7 4426	-
4	Ppi R1 461	-	43	Ppi R4 1452	+	82	Ppi R7 4466	-
5	Ppi R1 379	-	44	Ppi R4 1456A	+	83	Psy B728A	+
6	Ppi R1 456A	-	45	Ppi R4 1456B	+	84	Psy 100	-
7	Ppi R1 2491B	-	46	Ppi R4 1456C	+	85	Psy 1142	-
8	Ppi R1 277	-	47	Ppi R4 1456D	+	86	Psy 1150	-
9	Ppi R2 390	+	48	Ppi R4 1456E	+	87	Psy 1212	-
10	Ppi R2 1577	-	49	Ppi R4 1456F	+	88	Psy 1282-8	-
11	Ppi R2 1759	-	50	Ppi R4 1525	-	89	Psy 1338A	-
12	Ppi R2 2889B	+	51	Ppi R4 1528	-	90	Psy 2242A	-
13	Ppi R2 202	+	52	Ppi R4 1758B	-	91	Psy 2673C	-
14	Ppi R2 288	+	53	Ppi R4 1811	-	92	Psy 2675C	-
15	Ppi R2 223	+	54	Ppi R4 1812A	-	93	Psy 2676C	-
16	Ppi R2 285	-	55	Ppi R4 2171A	-	94	Psy 2677C	-
17	Ppi R2 1124B	-	56	Ppi R4 5143	-	95	Psy 2682C	-
18	Ppi R2 278	-	57	Ppi R5 974B	-	96	Psy 2692C	-
19	Ppi R2 1939	+	58	Ppi R5 2301C	-	97	Psy 2703C	-
20	Ppi R2 374A	+	59	Ppi R5 2532A	-	98	Psy 2732A	-
21	Ppi R2 1924	-	60	Ppi R6 1683	-	99	Psy 3023	+
22	Ppi R2 1842A	-	61	Ppi R6 1688	-	100	Pma M4	-
23	Ppi R2 1517C	-	62	Ppi R6 1704B	-	101	Pma 65	-
24	Ppi R2 1576A	-	63	Ppi R6 1759	-	102	Pma 1809A	-
25	Ppi R2 203	+	64	Ppi R6 1785A	-	103	Pma 1813	-
26	Ppi R3 222	-	65	Ppi R6 1796A	-	104	Pma 1820	-
27	Ppi R3 283	+	66	Ppi R6 1797A	-	105	Pma 1821A	-
28	Ppi R3 870A	-	67	Ppi R6 1804B	-	106	Pma 1838A	-
29	Ppi R3 895A	-	68	Ppi R6 1807A	-	107	Pma 1846A	-
30	Ppi R3 1125	-	69	Ppi R6 1842B	-	108	Pma 1848B	-
31	Ppi R3 1214	-	70	Ppi R6 1842C	-	109	Pma 1852A	+
32	Ppi R3 1216	-	71	Ppi R6 1745A	-	110	Pma 1855A	-
33	Ppi R3 1380A	-	72	Ppi R6 1746A	-	111	Pma 5422	+
34	Ppi R3 1441	-	73	Ppi R6 1755A	-	112	Pma 5429	-
35	Ppi R3 1554A	-	74	Ppi R6 1842D	-	113	Pma 6201	-
36	Ppi R3 1892	-	75	Ppi R6 1942	-	114	Pma 6319A/1	-
37	Ppi R3 2183A	-	76	Ppi R7 1691	-	115	Pma 6328A/1	+
38	Ppi R3 2186A	-	77	Ppi R7 2491A	-	116	Pgy 1139	-
39	Ppi R3 2191A	-	78	Ppi R7 4298	-	117	Pgy 2411	+
						118	Pgy 3318	-

3.2.14: Comparison of hybridisation versus PCR results.

Both hybridisation and PCR screening methods produced different results with the hybridisation screen producing many more positive signals than the PCR screen. This is mainly due to specificity and stringency of the two methods with PCR being a lot more specific. A comparative analysis (Table 3.10) was performed on both sets of screening data to observe if the hybridisation was representative of the PCR and vice versa. The results showed that the majority of the positive PCR amplifications also produced a signal following the hybridisation screen for the three regions, *rulAB, xerC* and *rulAB-xerC.* The only exceptions were during the *xerC* screens where six strains, five *Ppi* and one *Pma,* produced no hybridisation signal but produced a PCR product.

Table 3.10: Comparison of hybridisation and PCR amplification screening results from *P. syringae* strains. Red numbering represents a strong hybridisation signal.

Region		*rulAB*		*xerC*		*rulAB'-xerC*	
No.	Strain	Hybridisation	PCR	Hybridisation	PCR	Hybridisation	PCR
1	Ppi R1 299A	2	-	2	-	1	-
2	Ppi R1 4461	1	-	1	-	2	-
3	Ppi R1 862A	0	-	1	-	1	-
4	Ppi R1 461	1	-	1	-	2	-
5	Ppi R1 379	1	-	1	-	0	-
6	Ppi R1 456A	1	-	1	-	2	-
7	Ppi R1 2491B	0	-	1	-	0	-
8	Ppi R1 277	1	-	2	-	1	-
9	Ppi R2 390	3	+	0	+	1	+
10	Ppi R2 1577	2	-	0	-	1	-
11	Ppi R2 1759	2	-	0	+	0	-
12	Ppi R2 2889B	1	-	1	+	1	+
13	Ppi R2 202	3	-	2	+	1	+
14	Ppi R2 288	1	+	1	+	1	+
15	Ppi R2 223	2	+	1	+	2	+
16	Ppi R2 285	0	-	1	+	2	-

17	Ppi R2 1124B	3	-	0	+	1	-
18	Ppi R2 278	1	-	0	+	0	-
19	Ppi R2 1939	1	-	2	+	1	+
20	Ppi R2 374A	2	+	1	+	1	+
21	Ppi R2 1924	2	-	1	+	0	-
22	Ppi R2 1842A	1	-	0	-	2	-
23	Ppi R2 1517C	1	-	0	+	0	-
24	Ppi R2 1576A	1	-	2	+	1	-
25	Ppi R2 203	3	+	3	+	2	+
26	Ppi R3 222	2	-	1	-	2	-
27	Ppi R3 283	2	+	4	+	2	+
28	Ppi R3 870A	0	-	3	-	2	-
29	Ppi R3 895A	1	-	1	-	0	-
30	Ppi R3 1125	0	-	1	-	1	-
31	Ppi R3 1214	1	-	2	-	1	-
32	Ppi R3 1216	1	-	2	+	1	-
33	Ppi R3 1380A	2	-	1	+	1	-
34	Ppi R3 1441	2	-	4	+	1	-
35	Ppi R3 1554A	1	-	4	+	2	-
36	Ppi R3 1892	0	-	2	-	2	-
37	Ppi R3 2183A	0	-	1	+	2	-
38	Ppi R3 2186A	2	-	2	+	0	-
39	Ppi R3 2191A	1	-	2	+	0	-
40	Ppi R3 2817A	0	-	1	+	0	-
41	Ppi R3 4411	1	-	2	+	0	-
42	Ppi R3 4574	1	-	1	-	1	-
43	Ppi R4 1452	0	-	1	+	2	+
44	Ppi R4 1456A	1	-	1	+	1	+
45	Ppi R4 1456B	0	-	2	+	1	+
46	Ppi R4 1456C	0	-	3	+	1	+
47	Ppi R4 1456D	0	-	3	+	1	+
48	Ppi R4 1456E	0	-	3	+	1	+
49	Ppi R4 1456F	3	-	2	+	2	+
50	Ppi R4 1525	3	-	3	-	2	-
51	Ppi R4 1528	1	-	0	-	1	-
52	Ppi R4 1758B	1	+	2	-	1	-
53	Ppi R4 1811	3	-	3	-	0	-
54	Ppi R4 1812A	0	-	2	-	1	-
55	Ppi R4 2171A	0	-	1	-	1	-
56	Ppi R4 5143	1	-	3	-	0	-
57	Ppi R5 974B	3	-	2	-	3	-
58	Ppi R5 2301C	3	-	1	+	3	-
59	Ppi R5 2532A	2	-	2	+	1	-
60	Ppi R6 1683	2	-	3	-	2	-

61	Ppi R6 1688	2	-	1	-	1	-
62	Ppi R6 1704B	1	-	1	-	2	-
63	Ppi R6 1759	2	-	0	-	1	-
64	Ppi R6 1785A	2	+	1	+	1	-
65	Ppi R6 1796A	4	-	1	-	2	-
66	Ppi R6 1797A	1	-	0	-	1	-
67	Ppi R6 1804B	2	-	1	-	3	-
68	Ppi R6 1807A	1	+	1	-	1	-
69	Ppi R6 1842B	2	-	3	-	0	-
70	Ppi R6 1842C	1	-	0	-	1	-
71	Ppi R6 1745A	2	-	0	-	1	-
72	Ppi R6 1746A	3	+	0	-	2	-
73	Ppi R6 1755A	1	-	2	-	3	-
74	Ppi R6 1842D	0	-	1	-	3	-
75	Ppi R6 1942	0	-	1	-	2	-
76	Ppi R7 1691	2	-	3	+	2	-
77	Ppi R7 2491A	2	-	2	-	1	-
78	Ppi R7 4298	2	-	2	+	0	-
79	Ppi R7 4300	2	-	1	+	1	-
80	Ppi R7 4409	2	-	1	+	2	-
81	Ppi R7 4426	4	-	1	+	2	-
82	Ppi R7 4466	2	-	1	+	2	-
83	Psy B728A	1	-	2	-	1	+
84	Psy 100	1	-	1	-	1	-
85	Psy 1142	2	-	1	-	2	-
86	Psy 1150	3	-	2	-	1	-
87	Psy 1212	1	-	1	+	1	-
88	Psy 1282-8	1	-	3	-	1	-
89	Psy 1338A	1	-	0	-	0	-
90	Psy 2242A	4	-	0	-	2	-
91	Psy 2673C	0	-	0	-	1	-
92	Psy 2675C	1	-	2	-	1	-
93	Psy 2676C	2	-	1	-	2	-
94	Psy 2677C	1	-	4	-	3	-
95	Psy 2682C	1	-	0	-	2	-
96	Psy 2692C	1	-	3	+	2	-
97	Psy 2703C	0	-	0	-	0	-
98	Psy 2732A	1	-	2	-	0	-
99	Psy 3023	2	+	1	-	1	+
100	Pph 103	1	-	3	-	0	-
101	Pph R1 1281A	1	-	4	-	3	-
102	Pph R6 1299A	4	-	4	-	3	-
103	Pph R3 1301A	2	-	4	-	2	-
104	Pph R4 1302A	3	-	4	-	2	-

105	Pph R5 1375A	**2**	**+**	**4**	**-**	**3**	**-**
106	Pph R7 1449B	**2**	**-**	**2**	**-**	**1**	**-**
107	Pph R6 1448A	**1**	**-**	**4**	**-**	**3**	**-**
108	Pph R7 1449A	**2**	**-**	**4**	**-**	**3**	**-**
109	Pph R8 2656A	**1**	**-**	**4**	**-**	**2**	**-**
110	Pph R9 2709A	**2**	**-**	**4**	**-**	**3**	**-**
111	Pma M4	**1**	**-**	**1**	**-**	**0**	**-**
112	Pma 65	**1**	**-**	**2**	**+**	**1**	**-**
113	Pma 1809A	**4**	**-**	**3**	**-**	**1**	**-**
114	Pma 1813	**1**	**-**	**3**	**-**	**1**	**-**
115	Pma 1820	**0**	**-**	**1**	**+**	**1**	**-**
116	Pma 1821A	**0**	**-**	**2**	**-**	**0**	**-**
117	Pma 1838A	**1**	**-**	**1**	**-**	**1**	**-**
118	Pma 1846A	**1**	**-**	**1**	**-**	**2**	**-**
119	Pma 1848B	**2**	**-**	**4**	**+**	**4**	**-**
120	Pma 1852A	**3**	**-**	**4**	**+**	**4**	**+**
121	Pma 1855A	**0**	**-**	**1**	**+**	**0**	**-**
122	Pma 5422	**4**	**-**	**2**	**+**	**1**	**+**
123	Pma 5429	**0**	**-**	**1**	**+**	**1**	**-**
124	Pma 6201	**2**	**-**	**3**	**-**	**3**	**-**
125	Pma 6319A/1	**0**	**-**	**0**	**+**	**2**	**-**
126	Pma 6328A/1	**0**	**-**	**0**	**-**	**2**	**+**
127	Pto DC3000	**1**	**-**	**2**	**+**	**2**	**-**
128	Pto 19	**1**	**-**	**1**	**+**	**2**	**-**
129	Pto 119	**1**	**-**	**2**	**-**	**2**	**-**
130	Pto 138	**1**	**-**	**2**	**+**	**1**	**-**
131	Pto 1108	**1**	**-**	**0**	**-**	**1**	**-**
132	Pto 2944	**1**	**-**	**1**	**+**	**1**	**-**
133	Pto 2945	**1**	**-**	**0**	**-**	**0**	**-**
134	Pto 6034	**1**	**-**	**2**	**+**	**1**	**-**
135	Pat 152E	**1**	**-**	**0**	**-**	**0**	**-**
136	Pat 4303	**1**	**-**	**0**	**-**	**0**	**-**
137	Pla 789	**1**	**-**	**1**	**-**	**1**	**-**
138	Pla 3988	**2**	**-**	**2**	**-**	**1**	**-**
139	Pgy 1139	**1**	**-**	**1**	**-**	**2**	**-**
140	Pgy 2411	**0**	**-**	**1**	**-**	**1**	**+**
141	Pgy 3318	**0**	**-**	**1**	**-**	**0**	**-**
142	P. asplenii 959	**1**	**-**	**1**	**-**	**1**	**-**
143	P. asplenii 1947	**0**	**-**	**0**	**-**	**0**	**-**
144	P. caricapapayae 1873	**0**	**-**	**1**	**-**	**0**	**-**
145	P. caricapapayae 3080	**1**	**-**	**1**	**-**	**1**	**-**
146	P. caricapapayae 3439	**0**	**-**	**2**	**-**	**1**	**-**
147	P. savastanoi 639	**0**	**-**	**0**	**-**	**0**	**-**
148	P. savastanoi 2716	**1**	**-**	**1**	**-**	**1**	**-**

149	P. savastanoi 3334	**1**	**-**	**1**	**-**	**0**	**-**
150	P. corrugata 2445	**1**	**-**	**1**	**-**	**0**	**-**
151	P. corrugata 3056	**2**	**-**	**1**	**-**	**0**	**-**
152	P. corrugata 3316	**1**	**-**	**0**	**-**	**0**	**-**
153	P. cichorii 907	**1**	**-**	**0**	**-**	**0**	**-**
154	P. cichorii 943	**1**	**-**	**1**	**+**	**1**	**-**
155	P. cichorii 3109A	**1**	**-**	**1**	**-**	**2**	**-**
156	P. cichorii 3109B	**1**	**-**	**0**	**-**	**0**	**-**
157	P. cichorii 3283	**0**	**-**	**1**	**-**	**0**	**-**
158	P. marginalis 247	**1**	**-**	**1**	**-**	**0**	**-**
159	P. marginalis 949	**0**	**-**	**1**	**-**	**0**	**-**
160	P. marginalis 2380	**1**	**-**	**0**	**-**	**0**	**-**
161	P. marginalis 2644	**0**	**-**	**0**	**-**	**0**	**-**
162	P. marginalis 2645	**1**	**-**	**1**	**+**	**1**	**-**
163	P. marginalis 2646	**0**	**-**	**0**	**-**	**1**	**-**
164	P. marginalis 3210	**1**	**-**	**0**	**-**	**1**	**-**

3.3: Discussion

Integron-like elements (ILEs) may provide a rapid route for plant pathogenic bacteria to overcome plant defence and surveillance systems because *Ppi* ILEs carry a TTE gene and this may be mobile between different bacterial strains. During the screening phase of this study 164 plant pathogenic *Pseudomonads* were screened for ILEs by looking at the disruption of the resistance to ultraviolet light A and B genes (*rulAB*). The *rulB* gene has been identified as a possible hotspot for ILE insertion (Arnold *et al.,* 2000; Rhodes *et al.,* 2014).

To test the ILE profiles of different *Pseudomonas* strains chemiluminescent probes were made and used to probe for three regions that would give an estimation of whether an ILE was present and if the ILE was inserted into the *rulB* gene. The probe for the ILE positive strains were made from amplified gene products. *Pseudomonas fluorescens* strain FH1

(pWW0::km^r::ILE$_{FH1}$) was used to make the ILE junction probe (*rulAB-xerC*) and the conserved ILE tyrosine recombinase probe (*xerC*). The probe used to identify strains without an ILE inserted into the *rulAB* operon (intact *rulAB*) was made from *Pseudomonas putida* strain PaW340 (pWW0::km^r) which has no ILE present within *rulAB*.

Although the probes were made from species that were different from the tested strains this was not as important for the DNA hybridisation screens compared to the PCR screens, depending on stringency washes and their component chemicals (Roche, 2014). There are advantages and disadvantages to the decreased identity required for DNA hybridisation tests. The advantage is that an overview of the potential number of ILE containing strains is given in a short time period, up to 94 strains could be screened on one blot per probe. The disadvantage is that the identity between the probes and the strains being tested can be relatively low and false positives can occur due to low stringency matches. This was the case for some of the strains that actually contained an ILE and disrupted *rulB* but presented as having an intact *rulB* gene due to the probe hybridising to either the partial *rulAB* sequence or another *rulAB* elsewhere in the genome. An example of this was in *Ppi* race 3 strain 283 (Figure 3.5, position B1) which provided a signal for both disrupted *rulB* and intact *rulB*, sequencing later showed an ILE was present and *rulB* was disrupted. Another example was the hybridisation results for the *Pph* strains when using the *rulAB-xerC* probe. When using the probe all of the strains produced a hybridisation signal (Figure 3.6); however, when the same strains were screened via PCR no amplification of the *rulAB-xerC* region was seen. It was these types of discrepancies along with the potential for sequencing that led onto PCR tests being used. A comparison of the two screening methods is shown in Table 3.10.

The first PCR screening tests used primers designed from the same strains used to make the hybridisation probes. These primers also targeted the same regions of the genome, *rulAB, xerC* and *rulAB-xerC.* However, no amplification was seen apart from the positive controls. This was due to the sequences sharing very little identity between the *Pseudomonas syringae* pathovars and the *Pseudomonas fluorescens* and *putida* strains. The difference between DNA hybridisation and PCR is evident when the two are compared. With hybridisation probes only, >40%+ identity is required with very low stringency, whereas in PCR primers >85%+ identity is required for amplification where the first ~5 nucleotides must be homologous to the target gene (Sommer and Tautz, 1989). However the hybridisations performed in the current work were highly stringent. Therefore, new primers were designed from the ILE harbouring *Ppi* 203 strain (Arnold *et al.,* 2000). As the primers were designed from a different strain to the hybridisation probes, the PCR screen was repeated on all 164 strains. The PCR screening provided results that could be used in later downstream applications. The high identity required for PCR amplification also explains why no amplification was seen from the more distantly related *Pseudomonas* species such as *P. asplenii, P. corrugata* and *P. marginalis.*

The *rulAB* primers that amplified the intact *rulAB* operon provided positive results for only 22 out of 164 strains screened. This number may not represent the true nature of UV resistance genes within these strains as there are many orthologues of *rulAB* that share similar function, but have different sequences (Kim and Sundin, 2000). The amount of identity between the orthologues determined whether the primers would bind and if amplification would occur. The orthologues that have the most identity to *rulAB* and are present in various *Pseudomonas* species include *recA*, *uvrD*, *polB*, *recN*, *sulA*, *uvrA*, *uvrB*, and

umuDC (Smith and Walker, 1998). The orthologue that most closely resembles *rulAB* the most is *umuDC* as this is also two separate genes that have a brief overlap to form a functional UV damage repair operon. ILEs have also been observed to integrate into *umuDC* forming *umuDC'* and *umuC''* in the same manner as *rulAB* (Maayer *et al.,* 2015). The diverse nature of UV damage repair genes and the abundance of *rulAB* orthologues may explain why so few *rulAB* positives were identified.

The *xerC* amplifications provided the highest amount of positive results for all three experimental regions; 78 positives from 164 strains tested. Although *xerC* genes are found in the conserved end of all the ILEs identified so far this does not make them markers for ILEs. This is because *xer* genes are not specific to ILEs and are common recombinase genes belonging to the bacteriophage lambda integrase family found across the bacterial domain (Blakely *et al.,* 1991). The fact that they are found in numerous bacterial species means that the number of positive amplifications will be higher than the number of ILEs present; this was the case as 22 potential ILEs were identified but 58 *xerC* genes were identified. In the case of ILEs *xerC* is conserved. This may be because the gene is required for ILE excision in the same manner as the genomic island, PPHGI-1. Excision of PPHGI-1 relies on *xerC* on the island being present and active (Pitman *et al.,* 2005; Lovell *et al.,* 2011). When the *xerC* gene is knocked out excision halts.

There are also other integrases inside and close to the ILE that contribute to gene cassette capture as in *Vibrio* species. The integrase genes within the ILE are regulated by the SOS pathway and are activated by the formation of ssDNA molecules and damage to the DNA (Cambray *et al.,* 2011). There are *rci* recombinase genes found just outside of *rulB'' (*Arnold *et al.,* 2000) that may be linked to ILE excision. The *rci* genes require far more accessory molecules and

signalling to become active compared to the ILE integrase genes. The *rci* genes require accessory genome markers such as shufflons (DNA binding sites) that are flanked by 19-bp repeat sequences (Gyohda *et al.,* 1997; Farrugia *et al.,* 2015).

The final region was the *rulAB-xerC* region which encompasses the disrupted *rulAB* operon and the start of the ILE which is conserved and contains a *xerC* gene. This region was of most importance as it indicates not only the presence of an ILE, but also that the ILE has inserted and disrupted the *rulAB* gene. When screening with the *rulAB-xerC* primers 22 positive amplifications were seen indicating ILE insertion.

In conclusion, this screening procedure has shown that it is highly likely that ILEs are abundant in *P. syringae* pathovars and that they share a common structure with a conserved end harbouring integrase genes and a more variable end possibly harbouring TTE genes and other virulence associated genes. Further work to look into the ILE sequences was carried out to obtain more information about these regions of DNA (Chapter 4).

Chapter 4. Sequence analysis of previously unidentified integron-like elements from *Pseudomonas syringae* pathovars.

4.1: Introduction

Once a number of potential integron-like elements (ILEs) were identified from a range of *Pseudomonas syringae* pathovars (Chapter 3) the next step was to obtain the sequences of these ILEs to assess if they were similar to other ILEs previous identified (Rhodes *et al.,* 2014) and to also assess where they were inserted into *rulB* and whether they encoded any type three effectors (TTE). There were 22 ILEs identified from four pathovars of *P. syringae; pisi* (Pea pathogen)*, syringae* (broad host pathogen)*, maculicola* (Brassica pathogen) and *glycinea* (Soybean pathogen) (Table 4.1), indicating that the ILEs are not restricted to only one pathovar within *P. syringae.*

The sequencing of these newly identified ILEs not only allowed their genetic content to be analysed, but also if the ILEs were more similar to true integrons in terms of their gene orientation and regulation. Integrons typically contain an integrase gene with its own promoter followed by a second promoter (P_c) to regulate the expression of captured genes in the variable end of the integron (Figure 4.1) (Gillings, 2014). However, if the genes are incorporated in the opposite orientation the P_c will not regulate the genes and another promoter is required for gene expression. This is the case for the previously analysed ILEs which have the variable genes in the opposite orientation to integrons (Figure 4.2). The other difference between ILEs and integrons is that integrons are typically found in non-coding regions of the genome, however the ILEs in *P. syringae* and other *Pseudomonas* are present within the DNA polymerase V gene, *rulB*.

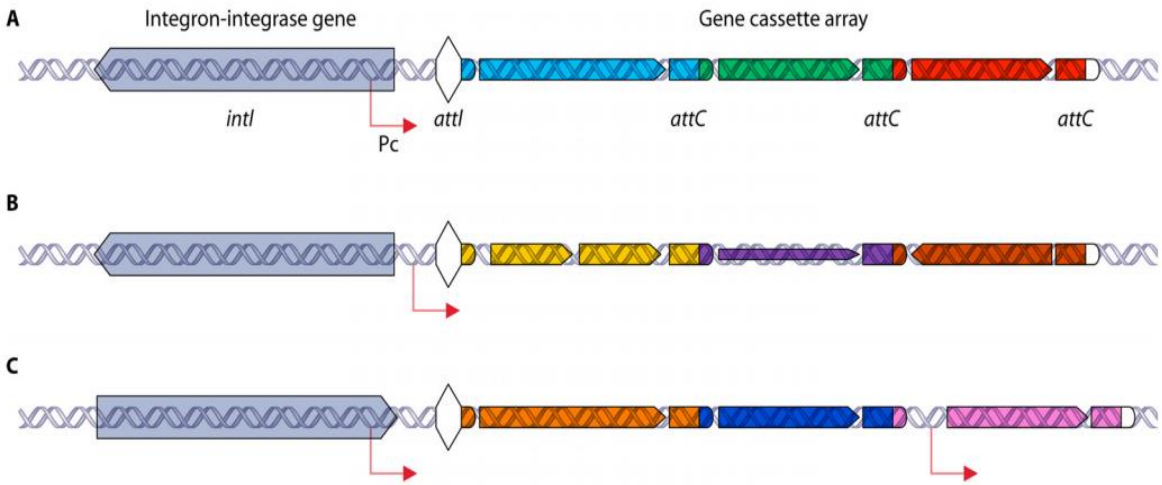


Figure 4.1: The genetic structure of typical integrons. These are three examples of the typical form of an integron. **A)** The P_c is within the integron integrase and regulates the captured genes; **B)** P_c is in the intergenic space between the integrase and the captured genes; **C)** P_c is within the integrase although the integrase is reversed in orientation, similar to ILEs, also there is a second promoter present to regulate genes further downstream. The cassette array is also termed the variable end and the *att* sites are the locations that genes cassettes are incorporated into the integron. (Image from Gillings, 2014. Used with permission from American Society for Microbiology).

Even though there are stark differences between integrons and ILEs they do share similar regions and functions. There were four main regions related to these ILEs that were of interest (Figure 4.2). The first was the region (Figure 4.2; 1) from the disrupted *rulB'* into the *xerC* integrase gene on the 5' end of the ILE. This area was thought to be conserved and that all ILEs contain a conserved integrase gene at this end. The second region (Figure 4.2; 2) falls within the first region and is the intergenic region between the end of the *rulB'* gene and the start of the ILE *xerC,* this was thought to be conserved and the insertion point of the ILEs (Rhodes *et al.,* 2014). The third region (Figure 4.2; 3) was the 3' end of the ILE and assessing its variability and whether or not all ILEs contain the same genes. Finally the whole ILE (Figure 4.2; 4) was studied to again look for variability, but also ILE length and to identify genes harboured within its entirety.

Table 4.1: The 22 ILE containing strains from *Pseudomonas syringae.* The strain name along with original isolation location, isolation date and plant host. *Ppi - pisi, Psy - syringae, Pma - maculicola, Pgy - glycinea.*

Potential ILE containing strains	Isolation location	Isolated from	Isolation year
Ppi R2 202	USA	Pea	1944
Ppi R2 203	New Zealand	Pea cv. *Small sieve freezer*	1969
Ppi R2 223	New Zealand	Pea cv. *Partridge*	1968
Ppi R3 283	New Zealand	Pea	1970
Ppi R2 288	Unknown	Pea	Unknown
Ppi R2 374A	New Zealand	Pea cv. *Partridge*	1970
Ppi R2 390	Unknown	Pea	Unknown
Ppi R2 1939	UK	Pea	1987
Ppi R2 2889B	Unknown	Pea	Unknown
Ppi R4 1452A	UK	Pea cv. *Belinda*	1985
Ppi R4 1456A	UK	Pea cv. *Belinda*	1985
Ppi R4 1456B	UK	Pea cv. *Belinda*	1985
Ppi R4 1456C	UK	Pea cv. *Belinda*	1985
Ppi R4 1456D	UK	Pea cv. *Belinda*	1985
Ppi R4 1456E	UK	Pea cv. *Belinda*	1985
Ppi R4 1456F	UK	Pea cv. *Belinda*	1985
Psy B728A	USA	Green bean	1987
Psy 3023	UK	Unknown	1950
Pma 1852A	UK	Brussel sprout cv. *Oliver*	1987
Pma 5422	UK	Cauliflower cv. *Danish per*	1995
Pma 6328A	Unknown	Unknown	Unknown
Pgy 2411	New Zealand	Soybean	1971

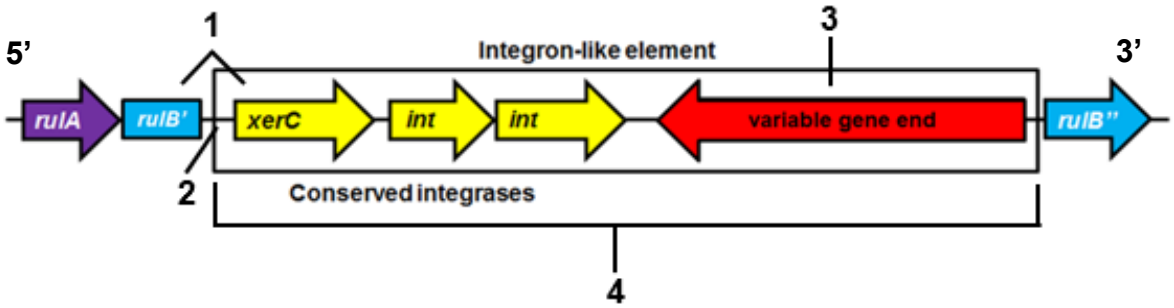


Figure 4.2: Basic outline of an ILE along with the disrupted *rulAB* operon. The figure shows the four main regions of the ILE that were studied and outlines the basic structure of the ILEs found in *P. syringae* pathovars. **1)** disrupted *rulB* flank into the ILE *xerC*; **2)** the intergenic region between *rulB'* and *xerC*; **3)** the variable 3' end of the ILE; **4)** the entire length of the ILE.

The first three regions, the conserved integrase and intergenic region, along with the more variable 3' end region, had their sequences obtained via standard PCR and chain termination sequencing with individual reads of up to ~1000bp. This was enough coverage to establish patterns and trends amongst the regions using multiple alignment and homology software. However, to obtain sequences that were partially unknown, such as differences in the variable end, semi-degenerate primers were used to amplify these regions, however this provide no useful sequence data. To obtain unknown sequences and the entire sequence of the ILEs 10 ILE containing strains were sent for whole genome sequencing from whole DNA preparations (MicrobesNG, Birmingham, UK).

Whole genome sequencing also allowed analysis of the genome surrounding the ILEs to look for any genes that may confer advantageous benefits to the ILE, such as additional mobility genes and integration genes. These may have had an impact on how the ILE moves between systems and how exogenous genes are incorporated. There was an established example of an *rci* integrase gene being located just outside the ILE found in *Ppi* 203 (Arnold *et al.,* 2000) and if other ILEs shared this trait then it indicated other genes are crucial to ILE mechanics.

Knowing the sequences of the newly identified ILEs and the surrounding genome allowed their genetic content to be known which may lead to possible uses of these ILEs in preventing plant diseases due to them all containing TTE genes.

4.2: Results

4.2.1: Sequence analysis of *rulAB-xerC* region indicative of ILE insertion.

PCR screening identified 22 possible ILEs, which included strains from four different *P. syringae* pathovars. All of the 22 positive strains had the conserved 5' ILE *xerC* end sequenced. The PCR product was 804bp and the 2015xerC primer set was used (Table 2.2). Alignment software and analysis confirmed that the region amplified is conserved as the sequences share a high degree of identity (Appendix III). Further analysis revealed that *Pma* 1852 is inserted into a *rulAB* homologue, *umuDC,* with 99% identity to the *umuDC* operon found in *P. syringae* pv. *actinidiae.*

A phylogenetic tree (Figure 4.3) was constructed from the *xerC* sequences from all of the 22 potential ILEs. The phylogenetic tree showed that there are two main clusters present in regard to the *xerC* gene; the first contains 12 strains from *Ppi*, *Pma* and *Psy* pathovars whereas the second cluster contains eight strains from *Ppi*. These clusters are stable as 99% and 86% of the time these clusters would form as in Figure 4.3 if the analysis was repeated 500 times. There are two strains that fall outside of the clusters. The first, *Pgy* 2411 is in its own group and represents approximately a three base pair substitution in the sequence compared to the *Ppi* cluster. Although the identity between *Ppi* 203 and *Pgy* 2411 remains high at 97.0%. The second strain that falls in its own group is *Psy* 3023. Although the phylogenetic tree shows this strain to form its own group *Psy* 3023 shares 95.5% identity with *Ppi* 203. The long branches of the tree relate to the nature of the tree being constructed from sequences within the same species. The phylogenetic tree is stable due to the high numbers on the connecting branches and the main branch with all strains except *Psy* 3023 will cluster as

shown 99% of the time. All *xerC* genes from the 22 ILEs show very high identity that is greater than 95%.

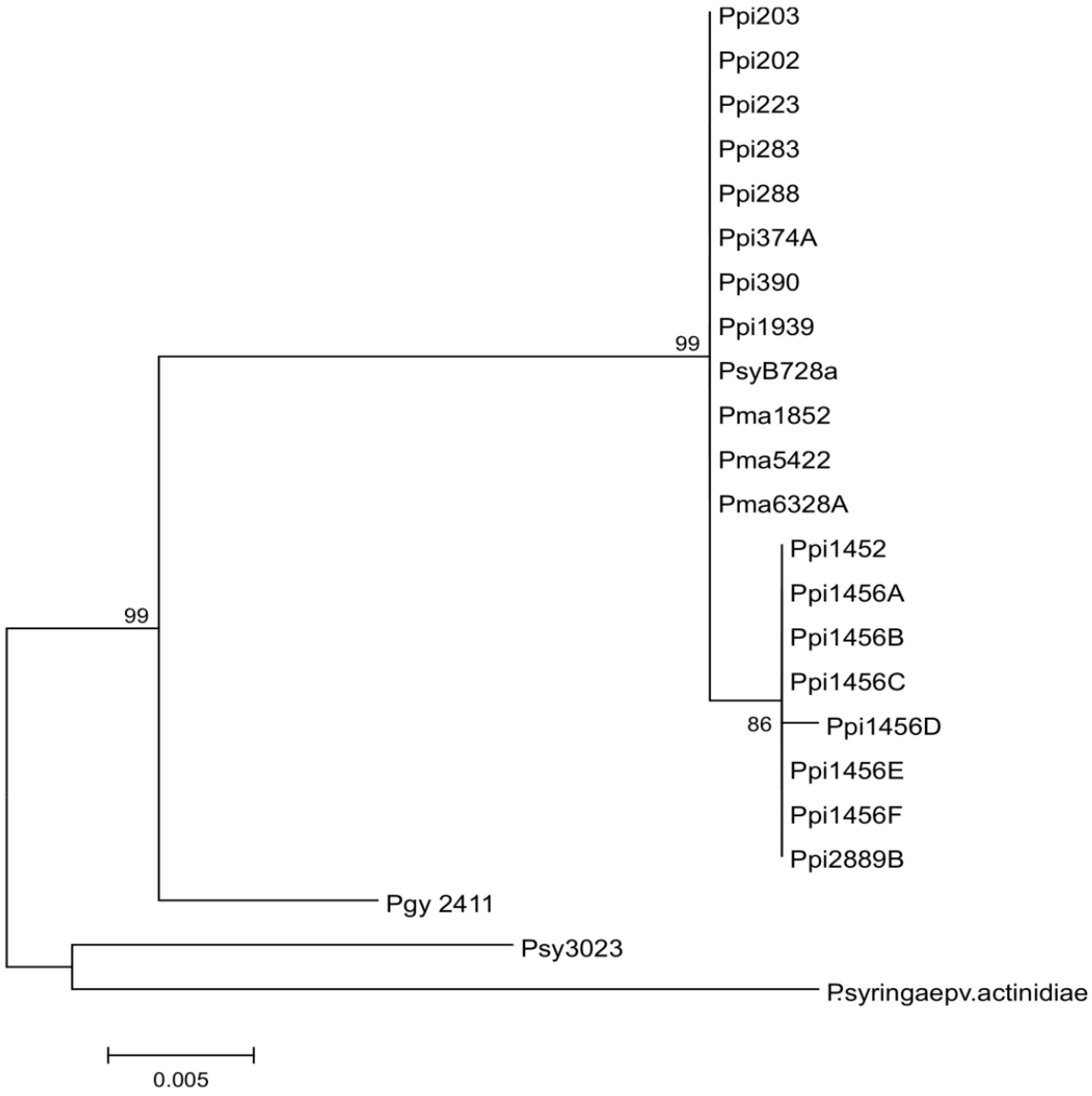


Figure 4.3: Molecular Phylogenetic analysis by Maximum Likelihood method looking at the ILE *xerC.* Phylogenetic tree shows the identity between the multiple potential ILE containing strains in regard to their *xerC* content. The percentage of branches in which the associated strains clustered together is shown next to the branches. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site and a bootstrap support of 500 replicates, using *P. syringae* pv. *actinidiae* as an outgroup. MEGA software was used to generate maximum likelihood tree (http://www.megasoftware.net/).

4.2.2: Sequence analysis of the intergenic region between the end of *rulB'* and the ILE start.

There is preliminary evidence that there is a conserved intergenic region of DNA found in environmental ILEs (*P. fluorescens* FH1, FH4 and FH5) that is always ~119bp long (Rhodes *et al.,* 2014). In the case of *P. syringae* it appears to be ~124bp, but is conserved across all the observed ILEs. However, the sequence is different to the *P. fluorescens* strains, with only 40.8% identity (Rhodes *et al.,* 2014). The region is present between the disruption point of *rulB'* and the start of the ILE. In the current study sequence analysis was carried out on the same area in the strains found to have ILEs. The PCR product used for sequencing was 1048bp and the 2015rulAB-xerCF primer set was used (Table 2.2). The analysis showed that the conserved region is present in *P. syringae* pathovars and is the same sequence and size across all the strains. A possible promoter region was identified with predicted -10 and -35 boxes (Figure 4.4).

```
Ppi203          GATCAAAGAGGTGCTGCGCCGCAATGGCATCAAGGTGTTCAGCAGCAACT
Ppi202          GATCAAAGAGATGCTGCGCCGCAATGGGATCAAGGTGTTCAGCAGCAACT
Ppi223          GATCAAAGAGGTGCTGCGCCGCAATGGCATCAAGGTGTTCAGCAGCAACT
Ppi283          GATCAAAGAGATGCTGCGCCGCAATGGGATCAAGGTGTTCAGCAGCAACT
Ppi288          GATCAAAGAGATGCTGCGCCGCAATGGGATCAAGGTGTTCAGCAGCAACT
Ppi374A         GATCAAAGAGGTGCTGCGCCGCAATGGCATCAAGGTGTTCAGCAGCAACT
Ppi390          GATGCGTTAAGT-CTGCGCCGCAATGGCATCAAGGTGTTCAGCAGCAACT
Ppi1452         GATCAGAGAGGTGTTGCGCCGCAACGGCATCAAGGTGTTCAGCAGCAACT
Ppi1456A        GATCAGAGAGGTGTTGCGCCGCAACGGCATCAAGGTGTTCAGCAGCAACT
Ppi1456B        GATCAGAGAGGTGTTGCGCCGCAACGGCATCAAGGTGTTCAGCAGCAACT
Ppi1456C        GATCAGAGAGGTGTTGCGCCGCAACGGCATCAAGGTGTTCAGCAGCAACT
Ppi1456D        GATCAGAGAGGTGTTGCGCCGCAACGGCATCAAGGTGTTCAGCAGCAACT
Ppi1456E        GATCAGAGAGGTGTTGCGCCGCAACGGCATCAAGGTGTTCAGCAGCAACT
Ppi1456F        GATCAGAGAGGTGTTGCGCCGCAACGGCATCAAGGTGTTCAGCAGCAACT
Ppi1939         GATCAAAGAGGTGCTGCGCCGCAATGGGATCAAGGTGTTCAGCAGCAACT
Ppi2889B        GATCAGAGAGGTGTTGCGCCGCAACGGCATCAAGGTGTTCAGCAGCAACT
PsyB728a        GATCAAAGAGGTGCTGCGCCGCAATGGCATCAAGGTGTTCAGCAGCAACT
Psy3023         GATCAAAGAGGTGTTGCGCCGCCACGGCATAAAGGTGTTCAGCAGCAACT
Pma1852         GATTAAGGATGTACTGAAGCGAAACGGTATCAAGGTTTTCAGCAGCAACT
Pma5422         GATTAAGGATGTACTGAAGCGAAACGGTATCAAGGTTTTCAGCAGCAACT
Pma6328A        GATCAGAGAGGTGTTGCGCCGCAACGGCATCAAGGTGTTCAGCAGCAACT
Pgy2411         GATCAGAGAGGTGTTGCGCCGCAACGGCATCAAGGTGTTCAGCAGCAACT
                ***     *  *  **   **  * ** ** ***** *************

Ppi203          ACGCGCTTTAGTAAGTTAGGTGGAATGCTTCTGGGCTACGCTGATTCTGT
Ppi202          ACGCGCTTTAGTAAGTTAGGTGGAATGCTTCTGGGCTACGCTGATTCTGT
Ppi223          ACGCGCTTTAGTAAGTTAGGTGGAATGCTTCTGGGCTACGCTGATTCTGT
Ppi283          ACGCGCTTTAGTAAGTTAGGTGGAATGCTTCTGGGCTACGCTGATTCTGT
Ppi288          ACGCGCTTTAGTAAGTTAGGTGGAATGCTTCTGGGCTACGCTGATTCTGT
Ppi374A         ACGCGCTTTAGTAAGTTAGGTGGAATGCTTCTGGGCTACGCTGATTCTGT
Ppi390          ACGCGCTTTAGTAAGTTAGGTGGAATGCTTCTGGGCTACGCTGATTCTGT
Ppi1452         ATGCACTCTAGTAAGTTAGGTGGAATGCTTCTGGGCTACGCTGATTCTGT
Ppi1456A        ATGCACTCTAGTAAGTTAGGTGGAATGCTTCTGGGCTACGCTGATTCTGT
Ppi1456B        ATGCACTCTAGTAAGTTAGGTGGAATGCTTCTGGGCTACGCTGATTCTGT
Ppi1456C        ATGCACTCTAGTAAGTTAGGTGGAATGCTTCTGGGCTACGCTGATTCTGT
Ppi1456D        ATGCACTCTAGTAAGTTAGGTGGAATGCTTCTGGGCTACGCTGATTCTGT
Ppi1456E        ATGCACTCTAGTAAGTTAGGTGGAATGCTTCTGGGCTACGCTGATTCTGT
Ppi1456F        ATGCACTCTAGTAAGTTAGGTGGAATGCTTCTGGGCTACGCTGATTCTGT
Ppi1939         ACGCGCTTTAGTAAGTTAGGTGGAATGCTTCTGGGCTACGCTGATTCTGT
Ppi2889B        ATGCACTCTAGTAAGTTAGGTGGAATGCTTCTGGGCTACGCTGATTCTGT
PsyB728a        ACGCGCTTTAGTAAGTTAGGTGGAATGCTTCTGGGCTACGCTGATTCTGT
Psy3023         ACGCGCTCTAGTAAGTTAGGTGGAATGCTTTTGGACTACGCTGATTCTGT
Pma1852         ATGCGCTTTAGTAAGTTAGGTGGAATGCTTCTGGGCTACGCTGATTCTGT
Pma5422         ATGCGCTTTAGTAAGTTAGGTGGAATGCTTCTGGGCTACGCTGATTCTGT
Pma6328A        ATGCACTCTAGTAAGTTAGGTGGAATGCTTCTGGGCTACGCTGATTCTGT
Pgy2411         ATGCACTCTAGTAAGTTAGGTGGAATGCTTCTGGGCTACGCTGATTCTGT
                * ** ** ********************** *** ***************

Ppi203          CGACGTTCTGGGGGATCGCCCATG
Ppi202          CGACGTTCTGGGGGATCGCCCATG
Ppi223          CGACGTTCTGGGGGATCGCCCATG
Ppi283          CGACGTTCTGGGGGATCGCCCATG
Ppi288          CGACGTTCTGGGGGATCGCCCATG
Ppi374A         CGACGTTCTGGGGGATCGCCCATG
Ppi390          CGACGTTCTGGGGGATCGCCCATG
Ppi1452         CGACGTTCTGGGGGATCGCCCATG
Ppi1456A        CGACGTTCTGGGGGATCGCCCATG
Ppi1456B        CGACGTTCTGGGGGATCGCCCATG
Ppi1456C        CGACGTTCTGGGGGATCGCCCATG
Ppi1456D        CGACGTTCTGGGGGATCGCCCATG
Ppi1456E        CGACGTTCTGGGGGATCGCCCATG
Ppi1456F        CGACGTTCTGGGGGATCGCCCATG
Ppi1939         CGACGTTCTGGGGGATCGCCCATG
Ppi2889B        CGACGTTCTGGGGGATCGCCCATG
PsyB728a        CGACGTTCTGGGGGATCGCCCATG
Psy3023         CGACGTTCTGGGGGATTGCCCATG
Pma1852         CGACGTTCTGGGGGATCGCCCATG
Pma5422         CGACGTTCTGGGGGATCGCCCATG
Pma6328A        CGACGTTCTGGGGGATCGCCCATG
Pgy2411         CGACGTTCTGGGGGATCGCCCATG
                *************** *******
```

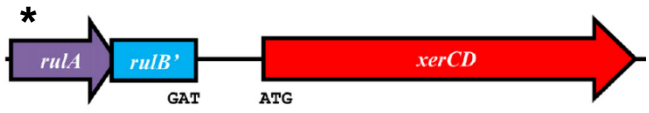


Figure 4.4: Sequence alignment of the intergenic region between rulB' and the ILE and its location between *rulB'* and *xerC*. The alignment shows that the intergenic region of the 22 ILE strains share high identity. Analysis performed using T-Coffee software available at http://www.ebi.ac.uk. -35 box and -10 box predicted using BProm available at http://www.softberry.com. *Image adapted from Rhodes *et al.,* (2014) under the creative commons license 4.0; creativecommons.org/licenses/by/4.0/.

4.2.3: PCR tests of the ILE variable end to observe if variation occurs.

Although the results showed that the conserved region of the *P. syringae* ILEs are conserved across the tested pathovars, little was known about the 3' end of the ILE, i.e. the variable region which in *Ppi* 203 contains the *avrPpiA1* effector gene. The first stage was to identify if the variable region is indeed variable. This was initially performed by using *Ppi* 203 as a template to design the primers, variF and variR (Table 2.2) to produce a variable end amplicon (Figure 4.5). This was to see if any of the ILE containing strains showed amplification of the region indicating that the region was the same as *Ppi* 203. The ILE variable end tests (Figure 4.6) showed that some of the strains did have the same region, same size band and sequence, as *Ppi* 203. However, other strains did not show amplification indicating possible variation or unsuccessful amplification due to experimental error.

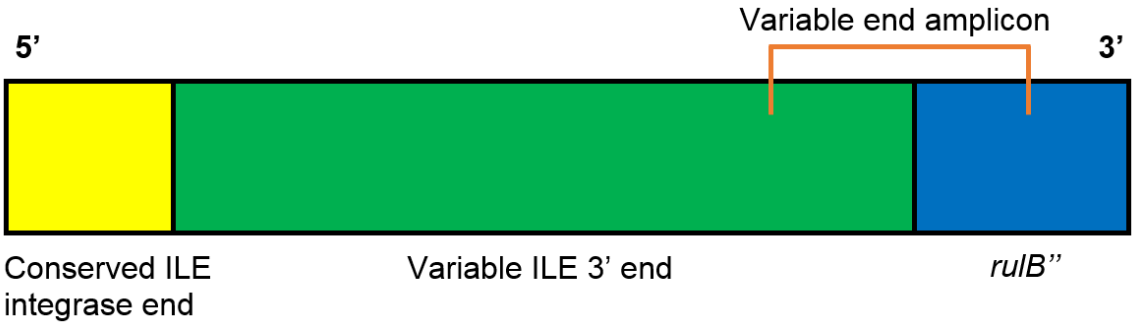
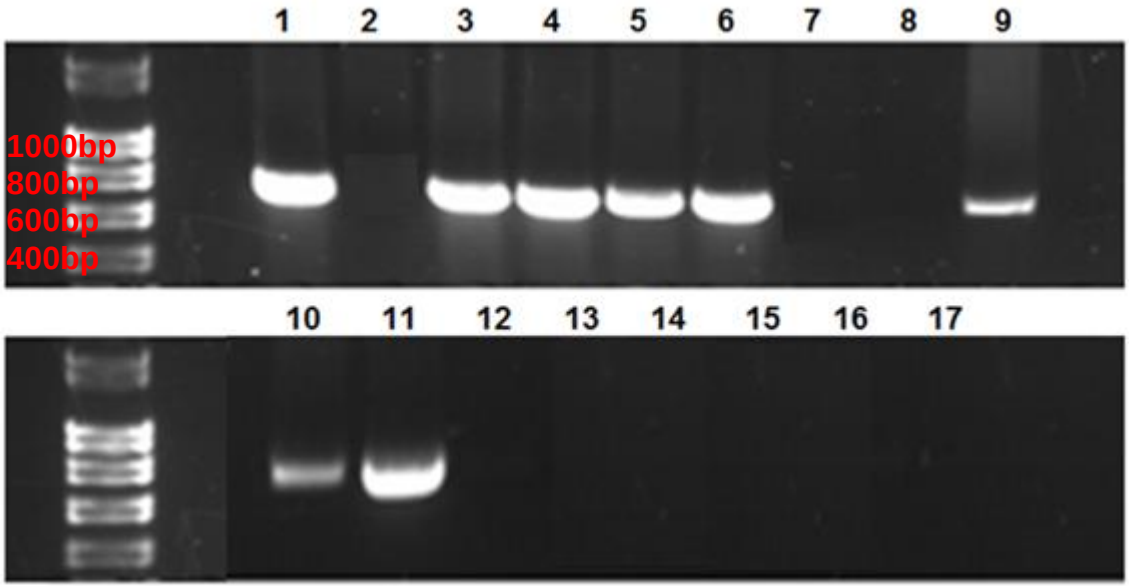


Figure 4.5: Position of variable end primers and the amplicon produced. The variable end primers designed from *Ppi* 203 will only amplify ILE variable ends that are the same as in *Ppi* 203, thus allowing observations into the variable end of multiple ILEs.



No.	Strain
1	*Ppi* 203 (+)
2	*Ppi* 202
3	*Ppi* 223
4	*Ppi* 283
5	*Ppi* 288
6	*Ppi* 374A
7	*Ppi* 390
8	*Ppi* 1456A
9	*Ppi* 1456E
10	*Ppi* 1939
11	*Ppi* 2889B
12	*Psy* B728a
13	*Psy* 3023
14	*Pma* 1852A
15	*Pma* 5422
16	*Pma* 6328A
17	Negative control

Figure 4.6: Identifying if all of the potential ILEs have the same variable end via PCR. Primers designed from *Ppi* 203 would indicate that the variable ILE end was the same if a PCR product was seen (~800bp). Only seven strains, plus the positive control, produced products indicating variability. This was repeated twice with the same results. Hyperladder 1kb (Bioline, UK) was used. (Green coloured box represent positive amplification).

The PCR results (Figure 4.6) showed that *Ppi* strains 223, 283, 288, 374A, 1456E, 1939 and 2889B had the same variable end as *Ppi* 203 as they produced a fragment of the same length: sequencing of the PCR fragment was used to confirm this. The strong amplifications from the variable end PCR tests were sequenced including the fragments that were larger than expected. The sequences obtained were analysed and were identical to each other and to *Ppi* 203, both in the *rulB'* region and the variable end of the ILE. The only exception being that *Pma* 1852 had a variation in the conserved *rulB''* region that confirmed the ILE was inserted into an *umuDC* operon rather than *rulAB*.

This was not unexpected as these results are in agreement with the conserved end tests. The next step was to determine why the other strains did not produce amplification products and how their ILE variable end was different from *Ppi* 203. This was achieved via whole genome sequencing.

4.2.4: Analysis of bacterial genomes to assess ILE content and variability.

Following the identification and analysis of 22 potential ILEs, 13 were found to contain the same genes as the ILE in *Ppi* 203 following PCR fragment sequencing. These strains included *Ppi* 223, 283, 288, 374A, 390, 1456A-F, 1939 and 2889B. The remaining eight potential ILEs were analysed following whole genome sequencing, except for *Psy* 3023 which had been previously sequenced by Thakur *et al.* (2016).

The bioinformatics procedure following genome sequencing by MicrobesNG, (UK) produced numerical outputs (Table 4.2) in relation to the quality and read length of the sequencing. The figures were produced using the QUAST: Quality Assessment Tool for Genome Assemblies software (http://bioinf.spbau.ru/quast, Gurevich *et al.,* 2013). This output from QUAST

provides the number of contigs that are above 1000bp, the total genome length, the number of contigs, the GC ratio and statistics associated with the genome assembly quality, N50 and L50. N50 is similar to the mean of the contig lengths. The two *Pma* strains that were re-sequenced had both genome reads assembled together to reduce the contig number and increase the length of the contigs. The same bioinformatic analysis was performed on this data (Table 4.3).

Bioinformatics of the strains revealed a significant variation across six of the ILEs including different TTE genes (Figure 4.7). This included *hopH1* and *hopAP1* in the pathovar *syringae* strains and *hopC1* in the *glycinea* pathovar. The genome sequences from two of the three *Pma* strains thought to contain an ILE showed that a similar pattern is present in regard to the disrupted *rulB*, or homologue, however the genome sequencing method resulted in the apparent ILE being split between multiple contigs. This meant that only a partial ILE map (Figure 4.8) for these two *Pma* strains could be deduced. One of the strains, *Pma* 6328A, did not contain an ILE as previously thought.

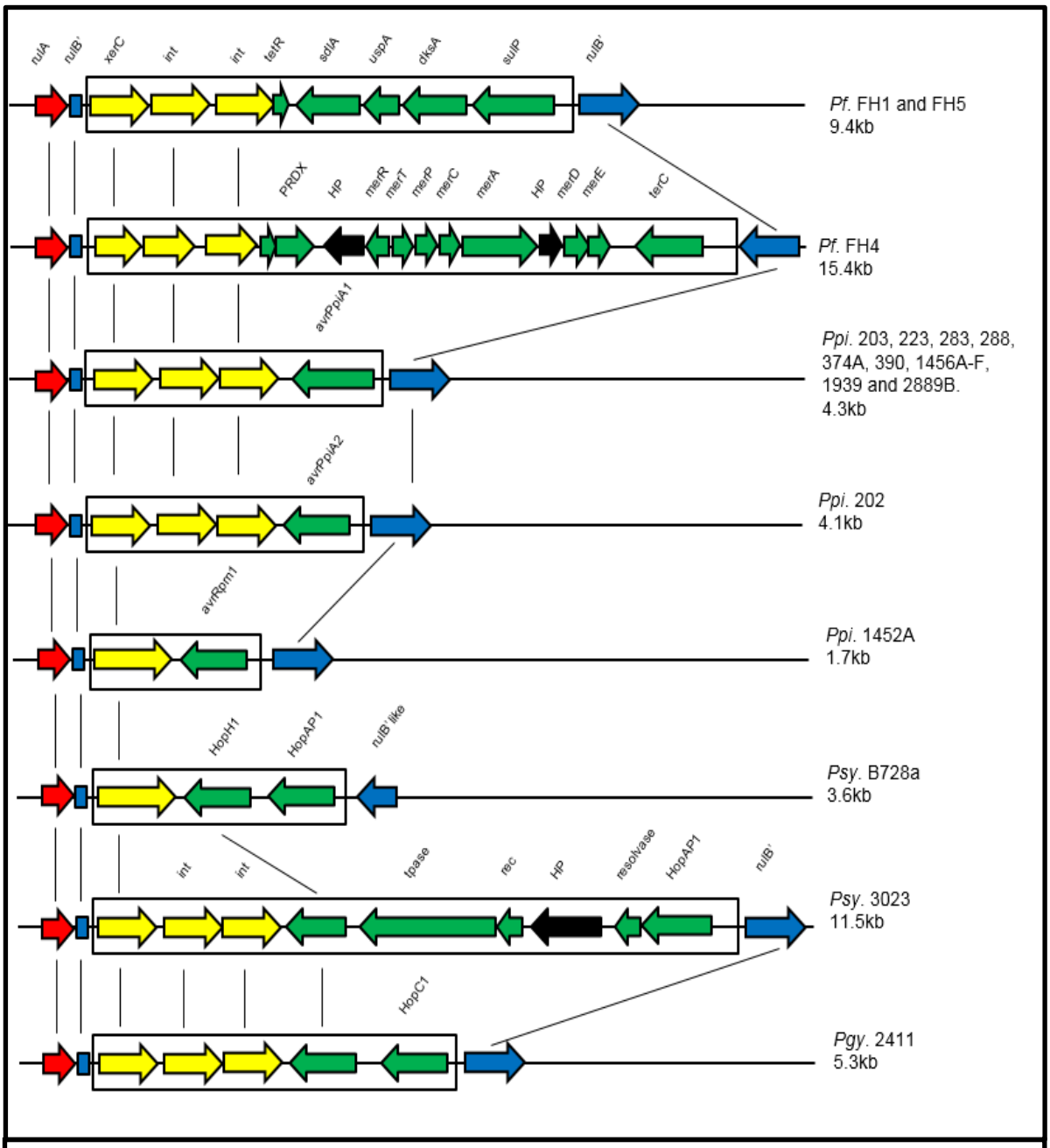


Figure 4.7: Sequence analysis of ILEs from *P. syringae* pathovars. ILE content maps from ILEs identified in *P. syringae* pv. *pisi,* pv. *syringae* and pv. *glycinea* which all contain at least one type three effector molecule. Also included is two previously identified ILEs from *P. fluorescens* which both contain genes conferring heavy metal resistance to the host bacterium (Rhodes *et al.,* 2014).

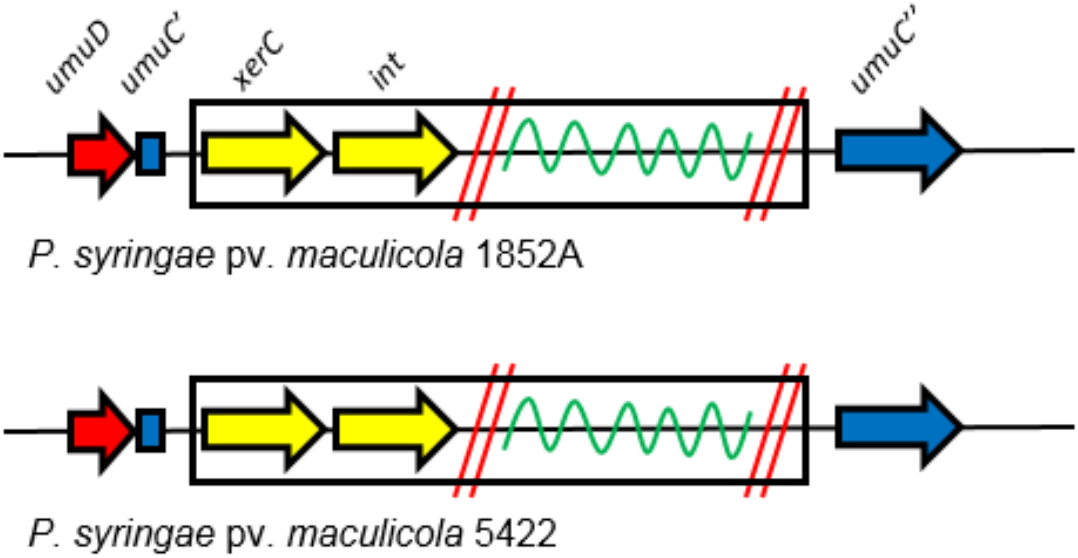


Figure 4.8: Partial ILE interpretation of genome sequence from *Pma* 1852A and 5422. Both of the *Pma* strains appear to have ILEs that resemble the others identified. The *rulB* homologue, *umuC* is disrupted by an insertion containing a *xerC* gene at the 5' end. However due to contig breaks the entire ILE is not viewable in its entirety. The end of the *umuC* gene was identified. The variable end is not known and is represented by the green section and is based on other ILEs identified. The red slashes represent contig breaks.

Table 4.2: QUAST output following *Pseudomonas syringae* pathovar genome sequencing. The QUAST output contains all of the information about the sequencing run, including contig analysis and numbers, genome size, GC ratio and genome quality stats, N50 and L50. *Ppi* 1452A shows a larger genome length due to cross contamination and the sequencing of both genomes, however the true *Ppi* 1452A sequence was still viable for analysis.

Sample ID	# contigs	# contigs >=1kb	Total length (bp)	GC%	N50	L50
***Ppi* 203**	304	148	6431341	58.23	129943	14
***Ppi* 202**	232	123	6431966	58.19	185852	10
***Ppi* 390**	116	103	6261399	58.54	144939	13
***Ppi* 1452A**	674	275	12079518	61.99	115084	31
***Ppi* 1456A**	267	133	6342444	58.34	167434	10
***Psy* B728a**	151	57	6420817	58.96	338366	6
***Pma* 1852A**	315	207	6593317	58.08	67764	27
***Pma* 5422**	363	199	6518403	58.11	70574	25
***Pma* 6328A**	442	213	6208691	58.37	93514	20
***Pgy* 2411**	422	256	6280079	57.72	58825	31

Table 4.3: QUAST output following assembly of two read libraries of two *Pma* strains, 1852A and 5422, into one assembled genome sequence. The QUAST output contains all of the information about the first sequencing run along with the DUAL output which refers to the assembly with two read libraries, including contig analysis and numbers, genome size, GC ratio and genome quality stats, N50 and L50.

Sample ID	# contigs	# contigs >=1kb	Total length (bp)	GC%	N50	L50
***Pma* 1852A**	556	207	6593317	58.08	67784	27
Pma 1852A dual	370	193	6535909	58.19	68921	27
***Pma* 5422**	762	199	6518403	58.11	70574	25
***Pma* 5422 dual**	452	174	6419842	58.26	67763	25

4.3: Discussion

In order to confirm the PCR tests that were carried out to identify potential ILEs (Chapter 3) the 22 potential ILEs had the conserved 5' end amplified and sequenced. This showed that the ILEs were present within *rulAB* or a homologue and that all of the potential ILEs contain a conserved integrase at the 5' end of the ILE (Figure 4.3). An interesting result emerged from sequencing the *rulB'* to *xerC* region. Sequence analysis of *P. syringae* pv. *maculicola* 1852 returned with a sequence where the conserved 5' end of the ILE was inserted into a region with high identity to the *umuDC* operon of *E. coli* rather than a *rulAB* operon. This was confirmed when the genome sequencing revealed the end of the disrupted *umuC* continued after the ILE insertion (Figure 4.7). This result was not unexpected as *rulAB* is a homologue of *umuDC.* It appears this result is not an anomaly as other ILEs have been identified within *umuDC* operons in *Pantoea ananatis* bacteria (Maayer *et al.,* 2015).

In addition, another piece of conserved sequence was identified that had been first identified by Rhodes *et al.,* (2014). This piece of conserved sequence appears to always be 118/119bp in length in *P. syringae* pathovars and is in the intergenic region between the end of the disrupted *rulB'* and the start of the ILE (*xerC*). This region was also present in all of the sequenced samples, but in this case is 124/125bp in length (Figure 4.4). It is not yet clear why this small region is conserved or its significance to ILE mobility and/or fitness, although it may be acting as an accessory genome region for site-specific recombination or as a promoter regulation site due to both a predicted -10 and -35 box being present. If this was the case it would explain the sequence conservation to facilitate promoter binding and gene expression. Once the abundance of ILEs was established and the question of whether they all have the same conserved end

was answered, the next step was to look at the variable end of the ILEs. This region of the ILE can contain a variety of genes including virulence genes. In the case of *Ppi* 203 it is *avrPpiA1* (Arnold *et al.,* 2000), antibiotic resistance in many bacteria including *Vibrio* species (Ploy *et al.,* 2000) and heavy metal resistance genes such as *sdiA, uspA* and *sulP* which were all present on an ILE within *P. fluorescens* FH1 (Rhodes *et al.,* 2014). The first step was to ensure the proposed variable end was variable. This was achieved by designing primers specific to the *Ppi* 203 sequence and the primers spanned a section from the end of the disrupted *rulB, rulB"* to inside the variable end of the ILE (~650bp).

All of the strains that were positive for ILE insertion and had the conserved end sequenced were subjected to PCR amplifications, using the *Ppi* 203 variable end derived primers. Any amplification would suggest that the variable end is shared amongst those strains. Eight strains produced positive amplifications (Figure 4.6). These eight were sequenced and multiple alignment analysis showed that the variable ends shared high identity. The sequence analysis also showed that *Pma* 1852 had an *umuDC* operon rather than a *rulAB* operon, as discussed above.

Many of the strains showed no amplification with the *Ppi* 203 derived primers, this indicated that some variation is present in the ILEs. The next step was to identify what was present in the variable end of these ILEs. This was not possible via normal PCR as the sequences were unknown so primers could not be designed, however the variable end was flanked by the *rulB"* region of *rulAB* and its sequence was known. This meant that a semi-degenerate amplification method (Manoil, 2000) could be used to try and amplify the variable end. The method works by using a two-step PCR and degenerate primers with target tags on (Section 4.2.1).

Once the semi-degenerate primer amplification was complete sequence analysis was performed to determine what genes are present in the variable end of the ILE. This method could be repeated to 'walk' further into the variable end if required. The semi-degenerate amplification method was ineffective at producing results for this study. All of the returned sequences were very short and not specific to any known ILEs.

Another option to determine the variable end of an ILE was to sequence the entire bacterial genome and retrieve the entire ILE sequence from the data. Nine potential ILE containing strains, plus *Ppi* 203, had their genome sequenced using the services provided by MicrobesNG, (UK).

Following ILE sequencing all 22 potential ILEs have been analysed. One strain, *Pma* 6328A, that was thought to contain an ILE did not contain one when its genome was analysed. The two *Pma* strain, 1852A and 5422 were re-sequenced and the two read libraries were assembled into one genome sequence in the hope of reducing the number of contigs and increasing the length of the contigs so that the ILE sequence was not interrupted by the start or end of a contig. The assembly of the two read libraries resulted in a halving of the total number of contigs returned by SPAdes (Bankevich *et al.,* 2012), (available here; http://cab.spbu.ru/software/spades/), but did not increase the length of the largest contigs which suggested that there were repetitive regions that prevented extension using paired-end short reads as used by MicrobesNG for sequencing.

A way to overcome this would be to use mate paired end libraries or Pacific Biosciences sequencing which gives much longer reads of up to an average of 10,000bp and a maximum of 60,000bp (PacBio;http://pacb.com/).

Following this the potential ILEs in both *Pma* 1852A and 5422 were identified and it appears the ILE disrupts the *rulB* homologue, *umuC*. Due to the contig breaks

in the sequence caused by short read lengths it can only be hypothesised that the ILEs in *Pma* 1852A and 5422 follow the same patterns as the other ILEs (Figure 4.7; 4.8) and they may contain a TTE. It has been concluded that the identification of a potential ILE in *Pma* 6328A earlier in the research was a mistake due to cross contamination with another *Pma* strain. This was confirmed with a repeat PCR test that was negative for ILE insertion.

The 21 available ILEs (Figure 4.7; 4.8) have been fully mapped via bioinformatics identifying both similarities and differences between the ILE content. All of the identified ILEs appear to be inserted into a *rulB* gene or a close homologue such as *umuC;* however, ILE insertions into other genes was not considered in this research. Insertion into *rulB* and its homologues is not unprecedented as multiple previous studies have shown mobile genetic elements to insert into these sites (Maayer, *et al.,* 2015; Rhodes *et al.,* 2014; Arnold *et al.,* 2000). This backs up the theory that the UV DNA damage repair polymerase V *rulB* gene is a hotspot for ILE insertion. At this moment it is not known why this site is chosen over a non-coding region, as there are no obvious genetic markers for an insertion event (Rhodes *et al.,* 2014).

Another feature that is shared across the ILEs is a conserved integrase gene at the start of the ILE. These integrase genes share high identity, >99%, to the phage integrase *xerC*. The integrase protein may have a role in ILE mobility or it may be involved in capturing and excising gene cassettes as they are required by the bacterium. The majority of the newly identified ILEs also have two further phage integrase genes following the initial *xerC* gene. The only exceptions are *Ppi* 1452A and *Psy* B728a which only contain the *xerC* integrase. It may be possible that these extra integrases have become redundant over time and have therefore been lost through evolution. The other possibility is that these

integrases play a role in ILE mobility and gene capture resulting in the ILE becoming fixed if these integrases are not present. This has been identified in other MGEs, GInts, which also have multiple integrase genes. When one integrase gene was non-functional the MGE was unable to integrate or excise itself (Bardaji *et al.,* 2017).

The final common feature between all of the ILEs currently analysed is that there is at least one TTE encoding gene present. These proteins play an important role in plant - microbe interactions and their presence on a potential mobile genetic element may indicate how bacterial plant pathogens are able to overcome host resistance by varying their TTE repertoire to avoid host detection (Jones and Dangl, 2006). For instance, if the effector on the ILE causes a plant immune response to be triggered then selective pressure will select for any bacterium that has lost the ILE and therefore the effector, resulting in a bypass of the plants immune response systems. This is a very similar mechanism to the one seen in *Pph* 1302A where a chromosomally located genomic island, PPHGI-1, is excised and lost from the cell during cell division causing *Pph* 1302A to develop a disease phenotype where cells with PPHGI-1 present cause a HR (Neale *et al.,* 2016). A bacterium can also become virulent by gaining TTE genes, such as *avrRpt2* which promotes pathogen virulence on host plants lacking a functional RPS2 gene leading to disease proliferation in plants such as *Arabidopsis thaliana* (Chen *et al.,* 2007). There are six different TTE genes on the analysed ILEs that encode AvrPpiA1, AvrPpiA2, AvrRpm1, HopH1, HopAP1 and HopC1 (Figure 4.7). It is known that these proteins are TTEs but it is not known what protein family they belong to, although they are most likely enzymatic in nature. However AvrPpiA1, AvrPpiA2 and AvrRpm1 are homologues to each

other with only 10 amino acid exchanges across the 220 amino acids in the putative protein (Fritig and Legrand, 1993).

The ILEs do not only vary in genetic composition, but also in length with the largest being 11.5kb and the smallest only 1.7kb. The smallest, *Ppi* 1452A, contains only one integrase gene and one TTE gene. However, the largest, *Psy* 3023 is interesting as it is very similar to the ILE in *Psy* B728a and contains the same TTE genes. Although ILE 3023 is bigger due to the addition of genes related to transposons.

The sequences surrounding the ILEs were also analysed to look for any genes that may benefit the ILE in terms of mobility and/or gene incorporation. However, no genes have been identified and also the nature of the short read length sequencing made it difficult to follow the entire stretch of sequence for some of the ILE strains. However a possible promoter region was identified within the intergenic region in all of the *P. syringae* strains (Figure 4.4). All of the strains showed identity with a -10 box and -35 box linked to a promoter region, TAGTAA-N19-GGCTACGCT. This promoter region may drive ILE integrase expression.

Following analysis the genetic makeup of the 21 newly identified ILEs is known. It will be possible to research them further and identify how and if they have the ability to move from chromosome to plasmid via *rulAB* sites, if there is a preference over which *rulAB* site insertion occurs at and whether a functioning *rulAB* operon is required. Expression profiles of the ILEs can also begin to assess what may cause the genes on the ILEs to be expressed and under what conditions.

Chapter 5. Expression studies of integrase and type three effector genes from two integron-like elements identified in *Pseudomonas syringae.*

5.1: Introduction

Following the identification of 21 ILEs from *P. syringae* (Chapter 4) research began into understanding when the ILE genes are expressed and how different conditions influence gene expression. The gene expression study focused on the integrase gene present in all of the identified ILEs, *xerC,* and also different TTE genes, *avrPpiA1, hopH1 and hopAP1.* All of the identified ILEs have at least one TTE gene present.

Although integrons and gene cassettes function as a combined genetic system, they have a two pronged approached when it comes to gene expression. Firstly, the integrase genes responsible for integron gene cassette excision and integration are controlled by the LexA mediated SOS pathway (Cambray *et al.,* 2011). Secondly, the expression of captured gene cassettes in the variable end of the integron are controlled by one of two promoters; Pc1 is located within the integrase gene facing the gene cassettes and Pc2 is located in the *attI1* integration site (Gillings, 2014). Promoters vary across different integrons and many vary in strength leading to differences in gene cassette expression (Hall and Collis, 1995; Lévesque *et al.,* 1994) and integrons with weaker promoters often have higher excision and gene cassette turnover rates (Jové *et al.,* 2010) (Figure 5.1).

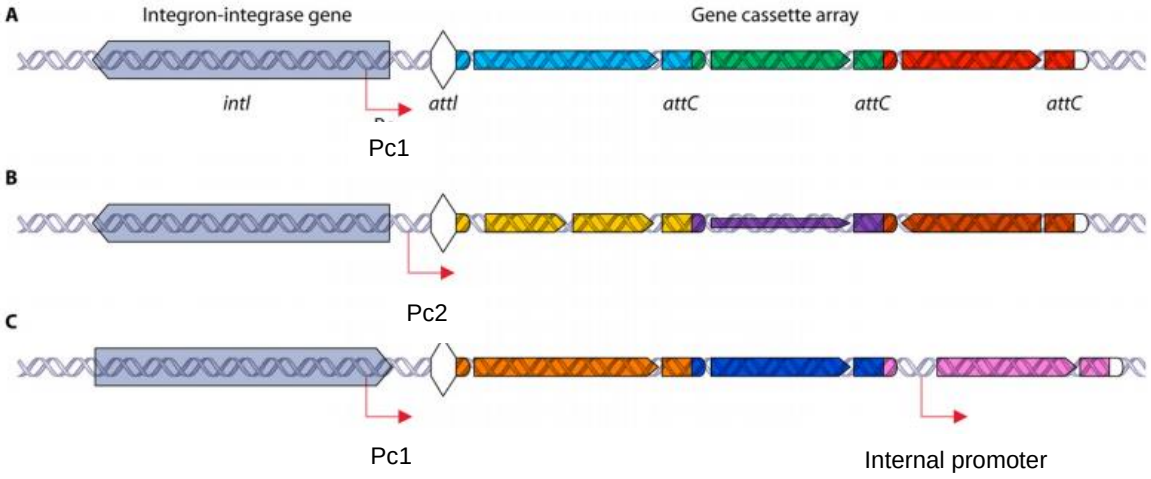


Figure 5.1: Promoter regions in integrons belonging to classes 1, 2 and 3. A) Promoter, Pc1, located within the integrase gene to regulate gene cassettes, maximum expression of ~6 cassettes. **B)** Pc2, located by the *attI1* integration site and acts the same as Pc1. **C)** Promoter Pc1 still present but internal gene cassette promoter also present to facilitate expression of genes further away from the Pc1 and Pc2 promoters. (Image from Gillings, 2014. Used with permission from American Society for Microbiology).

There has been previous studies that show a link between stress, the activation of the SOS response and the expression of integron integrases (Cambray *et al.,* 2011). This link may also be present in the newly identified ILEs in this study. An upregulation of *xerC* is beneficial in times of stress and DNA damage, as new gene cassettes may be available for the ILE to capture via *xerC* mediated integration to overcome the current environmental stress (Cambray *et al.,* 2011). The integron integrase is under the control of the LexA repressor protein that has a binding domain (CTGTatatatatACAG) in the integrase promoter region of multiple integrons. The LexA repressor protein is only released following extreme bacterial stress such as DNA damage caused by UV irradiation. The ssDNA formed following DNA damage and stress causes the activation of the SOS/RecA pathway which induces the autoproteolytic cleavage of LexA (Sundin and Weigand, 2007). Integron integrases become upregulated and promote genetic variation by capturing and excising gene cassettes in times of stress.

However no LexA binding motif appears to be present in the ILE *xerC* region. The *rulB* encoded polymerase V is also under the regulation of LexA as there is a binding domain present upstream of the *rulA* gene of the *rulAB* operon (Kim and Sundin, 2000). The SOS regulation of *rulB* may also regulate the ILE integrase genes. LexA repression of integron integrases also freezes the integrons' genetic content when the bacteria are in a steady environment (Cambray *et al.,* 2011).

Once the gene cassettes are incorporated into the integron their expression is under different regulation. Promoters within the integrase gene (Pc1) can drive the expression of gene cassettes, but promoter strength can be an issue. As the cassettes get further away from the Pc1/2 promoter their expression levels drop due to the increase in distance (Collis and Hall, 1995). This often leads to integrons containing no more than six separate gene cassettes if only Pc1/2 is present (Coleman and Holmes, 2005). Although promoters Pc1 and 2 function sufficiently for integrons containing less than six gene cassettes they do not facilitate gene expression for large chromosomal super-integrons that often contain over 100 gene cassettes. The super-integron in *Vibrio cholerae* contains 179 gene cassettes (Rowe-Magnus *et al.,* 1999). In larger integrons the gene cassettes further away from the integrase promoter are either transcriptionally silent and act as gene cassette reservoirs (Coleman and Holmes, 2005) or the gene cassettes contain their own internal promoters that can also regulate the expression of other gene cassettes downstream (Figure 5.1c) (Gillings, 2014). There have also been cassettes identified that have no ORFs and may be promoters in their own right (Holmes *et al.,* 2003). Some integrons may also overcome weak promoters by rearranging the gene cassettes so that when important genes are needed they are strongly expressed.

All of the expression studies in the literature have been done on 'true' class one, two and three integrons rather than integron-like elements (ILEs) found in *Pseudomonas syringae*. It is hypothesised that the *xerC* integrase gene present in all of the identified *Pseudomonas syringae* ILEs (Chapters 3 and 4) acts in a similar way to the integron integrase gene in terms of being regulated by the SOS response. This was further indicated as the *rulAB* promoter may control the expression of ILE integrase under the regulation of the LexA repressor protein, as a LexA binding site can be found upstream of *rulAB* (Jackson *et al.,* 2011). The *rulB* gene is disrupted by ILE insertion in *P. syringae* and other *Pseudomonas* species. This was why *rulB'* expression was also included in this set of experiments.

Another difference between ILEs and integrons is the expression of captured genes in the ILE variable end. All of the ILEs from *P. syringae* pathovars contain TTE gene(s) in the reverse orientation of 'true' integrons so that transcription begins at the 3′ end of the ILE and therefore cannot be under the regulation of the Pc1 or Pc2 promoters (Figure 5.1); (Rhodes *et al.,* 2014). In addition, there may be P_{int} promoters present in or close to the ILE *xerC* integrase gene as seen in the previous chapter by the identification of the -10 and -35 boxes. However this promoter would not express the TTE gene(s) as they are transcribed in the opposite direction. The TTE genes present on the ILEs may have their own promoters and regulation via a *hrp* box, although this is not known.

TTE gene expression was predominantly tested using plant apoplastic fluid and *in planta* growth. Three different cultivars of bean were used: Tendergreen (TG); Canadian Wonder (CW); Red Mexican (RM). These will all produce a HR reaction with *Ppi* 203, but disease will be seen with *Psy* 3023. *Pisum sativum* (Pea) cv. Kelvedon Wonder (KW) was also used which is

susceptible to *Ppi* 203 and a disease response is seen. This gene expression study was used to identify if there are any similarities in gene expression across the *xerC* integrases, other ILE integrases present and TTE genes from two characterised ILEs from *P. syringae* pv. *pisi* 203 and *P. syringae* pv. *syringae* 3023. The study would also allow comparison to previously studied integrons and if any indications can be deduced about how the genes in the ILE variable end are expressed when the bacterium are under stresses such as UV, sub-optimal temperatures, DNA damaging agents, bacterial conjugation, *in planta* stress and plant apoplastic fluid (fluid around the plant cells where *P. syringae* colonises).

The gene expression profiles were analysed using reverse-transcription quantitative PCR (RT-qPCR). RT-qPCR can be used to measure gene expression in relation to the base line gene expression of housekeeping genes, such as DNA gyrase subunit B (*gyrB*), citrate synthase (*cts*), glyceraldehyde-3-phosphate dehydrogenase A (*gapA*) and RNA polymerase sigma factor (*rpoD*) (Berge *et al.,* 2014). Although multiple housekeeping genes can be used when analysing bacterial gene expression, one housekeeping gene is sufficient. In many studies, including this one, *gyrB* was used as the housekeeping gene (Martens *et al.,* 2008; Teramoto *et al.,* 2007). DNA *gyrB* is an enzyme that promotes DNA supercoiling (Wigley *et al.,* 1991) and was constitutively expressed to the same Ct value for all tested conditions in this study. DNA *gyrB* has also been used in previous RT-qPCR studies on phytopathogenic Pseudomonads (Takle *et al.,* 2007). Using a housekeeping gene allows quantification and normalisation, to account for differences in bacterial strain gene expression.

The most commonly used fluorophore technologies are TaqMan, developed by Applied Biosystems, and SYBR Green. This study used TaqMan.

However, both technologies have their advantages and limitations. SYBR Green is produced by multiple manufacturers resulting in cheaper prices whereas TaqMan is more expensive. Although, TaqMan is up to 10,000 fold more sensitive than SYBR Green (Zhou *et al.,* 2017) and also works with predesigned assays whereas SYBR Green often requires experimental optimisation. SYBR Green also requires more downstream analysis than TaqMan. This is due to both technologies using different approaches to quantify gene expression and gene amplification. SYBR Green is a non-specific intercalating dye that accumulates within the DNA as it is amplified resulting in more dye for genes that have higher expression levels (Figure 5.2A). TaqMan is sequence-specific and uses primers and a probe designed from the gene of interest. The region amplified is usually short, ~30/40bp, with the probe situated between the two primers (Figure 5.2B). TaqMan probes are labelled with a fluorescent reporter and quencher, which are maintained in close proximity until hybridization to the target DNA occurs. The reporter fluorophore, fluoroscein (FAM) and fluorescent quencher, rhodamine (TAMRA) are bound to the 5′ and 3′ ends of the probe respectively (van Guilder *et al.,* 2008). With no amplification and low expression of the target gene the reporter and quencher remain in close proximity and no fluorescent signal is detected. Once amplification begins the reporter and quencher fluorophores are released from the probe and the reporter fluorescence is detected. Gene expression is measured by the quantity of RNA produced under defined conditions. The more RNA present the more the gene is being expressed. RNA is extracted and purified from the sample and reverse transcription forms copy DNA (cDNA) that can be quantified by RT-qPCR. This step is required as RNA is unstable and Taq polymerase is unable to recognise uracil in RNA (Alcamo, 1999).

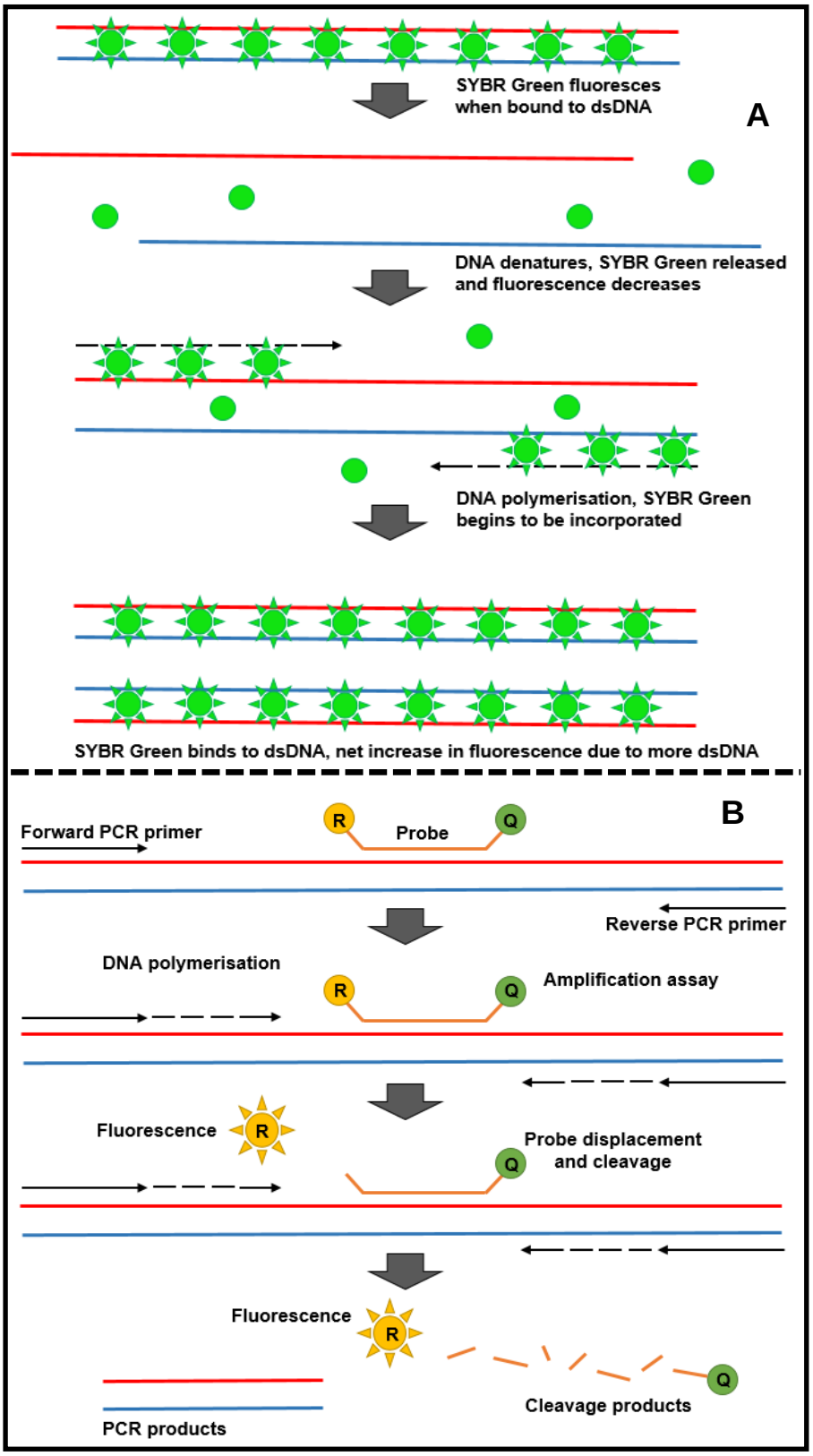


Figure 5.2: Comparison of RT-qPCR technologies; SYBR Green and TaqMan. Both technologies are based on the measure of fluorescence to quantify amplified DNA. SYBR Green (**A**) intercalates with dsDNA and as the target is amplified the SYBR Green accumulates, non-target dsDNA also accumulates SYBR Green resulting in further analysis needed. TaqMan (**B**) uses sequence specific primers and probes for detection. The probe contains a fluorophore reporter and a quencher resulting in a fluorescent signal only when the target is amplified; more copies of a gene the more fluorescence.

5.2: Results

5.2.1: Checking apoplastic fluid preparations.

5.2.1a: Confirmation that apoplastic fluid contains no cell lysate contamination via a malate dehydrogenase assay.

Plant apoplastic fluid was used as one of the stresses during the ILE gene expression research. It was important to check that the apoplastic fluid preparations were not contaminated by cell lysate as this would change the gene expression due to other enzymes present inside the cell. The check was performed via a malate dehydrogenase assay as malate dehydrogenase is present in the cell cytoplasm and its activity can be measured via the oxidisation of NADP to NAD^+ at 340nm. The results (Table 5.1) showed some enzyme activity for the apoplastic preparations (between 0.15-0.47 U/ mL), but this was far lower than the enzyme activity for the cell lysate preparations which ranged from 2.76-4.42 U/ mL. The positive malate dehydrogenase test which used standard malate dehydrogenase in place of the apoplastic fluid also showed higher levels of enzyme activity at 4.52 U/ mL, and the negative water only control was low at 0.02 U/ mL.

Table 5.1: Malate dehydrogenase assay on prepared plant apoplastic fluids and plant cell lysate preparations. Four different apoplastic fluids were checked for cell lysate contamination via a malate dehydrogenase assay. Cell lysate (CL) preparations were also included to ensure malate dehydrogenase was present inside the cells. Malate dehydrogenase (MDH) and water acted as the positive and negative controls respectively. Three replicates per preparation were measured. TG; Tendergreen bean, CW; Canadian Wonder bean, RM; Red Mexican bean.

Apo.	Av. △OD340/min	Activity U/mL ±S/D
TG	0.04	0.36 ±0.2
CW	0.02	0.15 ±0.1
RM	0.03	0.32 ±0.1
Pea	0.05	0.47 ±0.1
Water	0.00	0.02 ±0.0
TG CL	0.45	4.42 ±0.6
CW CL	0.40	3.88 ±0.3
RM CL	0.43	4.23 ±0.3
Pea CL	0.28	2.76 ±0.1
MDH	0.46	4.52 ±0.5

5.2.1b: Determining relative concentration of apoplastic fluids.

Before the extracted apoplastic fluids could be used for gene expression studies its' concentration had to be determined and normalised to the same starting concentration. This ensured any differences in gene expression when grown in apoplastic fluid was due to the stress and not due to some reactions containing a higher concentration of apoplastic fluid than others. To determine the concentration of the apoplastic fluid preparations indigo carmine was used to assay the relative concentration of protein in the samples (Table 5.2). The apoplastic fluid preparations were then standardised to the same starting concentration for all experiments.

Table 5.2: Relative concentrations of different apoplastic fluid preparations and normalisation data. The table shows the indigo carmine assay (IC) results from four different apoplastic fluid preparations (AWF) plus the infiltrate media (water) used as a control. Alongside the IC assay results are the results without any IC added. These values are then used in the following equation; Apoplast dilution factor = OD_{610}infiltrate / (OD_{610}infiltrate - OD_{610}AWF).

Apoplastic fluid concentration assay using Indigo Carmine (IC)					Norm. to a dil. factor of 1 in 5mL	
AWF Preps.	Av. OD610 without IC added	Av. OD610 with IC added	Difference between both OD610 measurements	Apo. fluid dil. factor	AWF vol. (uL)	H2O vol. (uL)
Infiltrate (H_2O)	0.000	0.106	0.106	N/A	N/A	N/A
Tendergreen	0.057	0.069	0.011	1.119	4467.08	532.92
Canadian Wonder	0.180	0.214	0.034	1.470	3401.25	1598.75
Red mexican	0.085	0.094	0.009	1.096	4561.13	438.87
Kelvedon Wonder	0.115	0.143	0.028	1.357	3683.39	1316.61

5.2.2: Expression of *P. syringae* pv. *pisi* 203 ILE genes.

Gene expression experiments were performed on the ILE in *P. syringae* pv. *pisi* 203. The ILE contains four genes, three integrases and a TTE gene, all of which were studied along with a recombinase gene, *rci,* which is outside the 3' end of the ILE and disrupted *rulB'* gene present at the 5' end of the ILE (Figure 5.3).

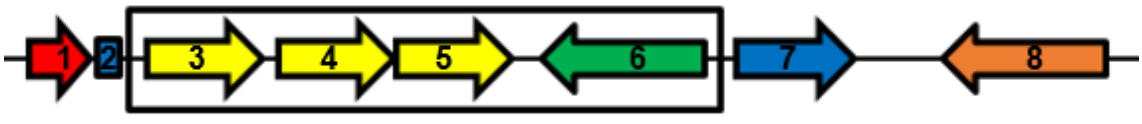


Figure 5.3: *P. syringae* pv. *pisi* 203 ILE genetic makeup. The ILE contains seven open reading frames along with the disrupted *rulB* gene. **1**; *rulA,* **2**; *rulB'*, **3**; *xerC*, **4**; *ORFE int*, **5**; *ORFD int*, **6**; *avrPpiA1*, **7**; *rulB''*, **8**; *rci.* Genes **2**, **3**, **4**, **5**, **6** and **8** were chosen for gene expression analysis.

The ILE gene expression profiles were recorded following various stresses and compared to no stress controls such as minimal media and no UV treatment.

Stresses included UV exposure, DNA damage, extreme temperatures and bacterial conjugation. All of the tests were replicated three times with independent biological replicates and standard errors of the mean were calculated based on these three measurements (Section 2.14). All significant differences were measured to $p<0.05$.

5.2.2a: Expression of *P. syringae* pv. *pisi* 203 ILE genes following inoculation into plant apoplastic fluid.

ILE gene expression was measured following six hours of bacterial inoculation in Tendergreen bean (TG), Canadian Wonder bean (CW), Red Mexican bean (RM) and Kelvedon Wonder (KW) pea apoplastic fluids and compared to a minimal media (MM) control. All of the genes tested showed variations in expression depending on plant apoplastic fluid used (Figure 5.4).

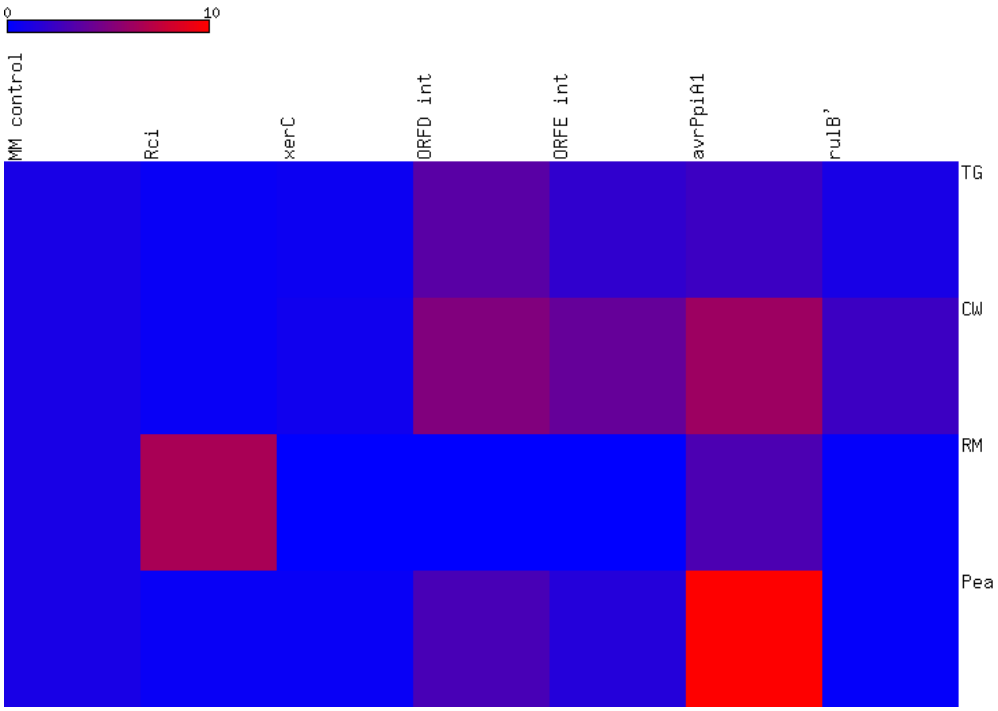


Figure 5.4: ILE gene expression heat map from *P. syringae* pv. *pisi* 203 when tested in plant apoplastic fluid. Six ILE genes from *Ppi* 203 tested for expression following bacterial growth in TG, CW, RM bean and KW pea apoplastic fluid for six hours. The blue/ red spectrum represents increasing gene expression from 0 to 10 times the expression of the MM control of one. (Full X-fold values available in Appendix IV).

The results show some of the conditions used caused a down regulation in gene expression compared to the MM control, such as the effect of KW pea apoplastic fluid on ILE *xerC* expression, which decreased to below 0.5 times the expression compared to MM and TG apoplastic fluid on *rci* expression which also decreased to below 0.5 times the expression compared to MM (Appendix IV). However some conditions did upregulate genes, such as *rci* when in RM apoplastic fluid led to an upregulation of between five and ten times compared to MM (Figure 5.4). The ILE integrase gene *xerC,* for example, was downregulated in all of the conditions when compared to MM. However, some genes were upregulated. This included both ORFD and ORFE integrases being upregulated on treatment with TG, CW and KW pea apoplastic fluid. The *rulB* gene was also upregulated between 1 and 5 times, but only in TG or CW apoplastic fluid.

The TTE, *avrPpiA1,* showed some interesting expression profiles (Figure 5.5). *AvrPpiA1* appears to be upregulated the most; over 10 times the expression compared to MM, when in KW pea apoplastic fluid which is the host plant of *Ppi* 203. The *avrPpiA1* gene was the only gene to upregulated in all of the apoplastic fluid. The TTE gene had an x-fold increase of 12.7 ± 2.3 in KW pea apoplast and 6.2 ± 2.1 fold increase in CW apoplast. TG and RM apoplast also showed an x-fold increase of 2.5 ± 0.7 and 3.00 ± 0.9 respectively (Figure 5.5).

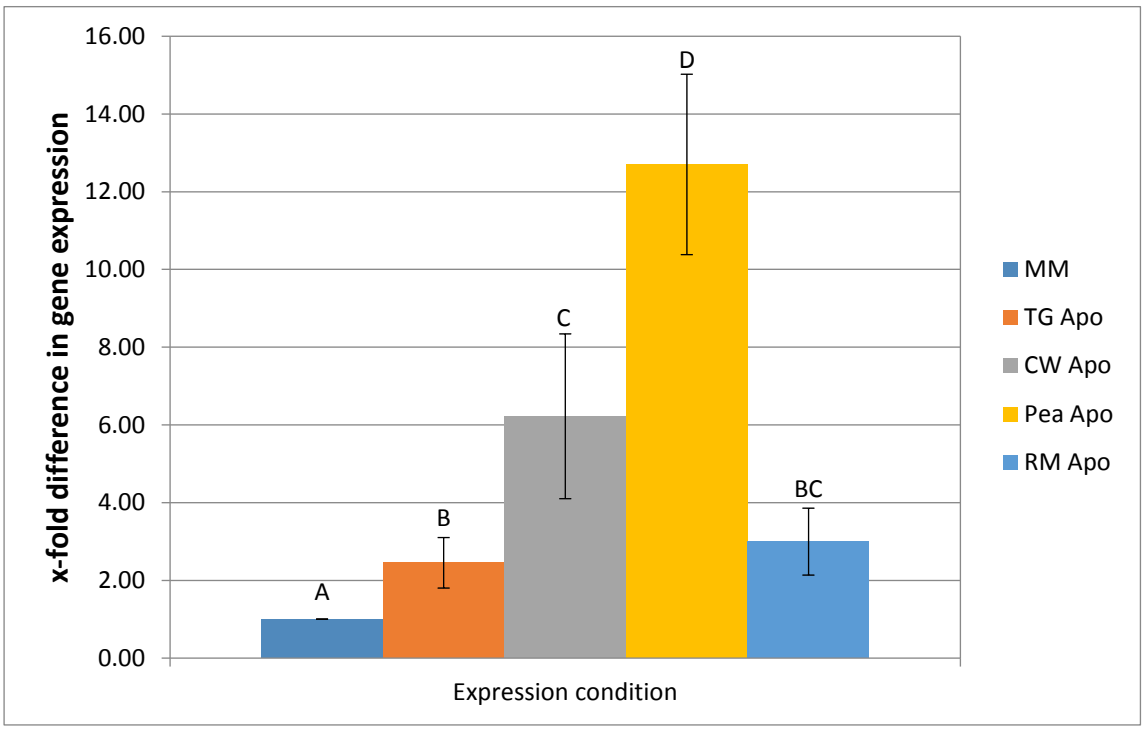


Figure 5.5: Gene expression of *avrPpiA1* from *P. syringae* pv. *pisi* 203 ILE in plant apoplast. Gene expression was measured via RT-qPCR following six hours of growth in plant apoplastic fluids. X-fold differences in gene expression were compared to the control MM condition and normalised against DNA *gyrB*. Three replicates per condition were tested and their mean values are displayed with standard error of the mean. (T-test and Tukey statistical analysis was used for significant differences [p<0.05] shown by letters above the bars).

5.2.2b: Expression of *P. syringae* pv. *pisi* 203 ILE genes following inoculation *in planta*.

The expression of ILE genes from *Ppi* 203 exhibited similar patterns *in planta* to the gene expression seen when the bacteria was incubated with apoplastic fluid. The TTE, *avrPpiA1* (Figure 5.6), showed the highest expression in KW pea *in planta* of 30.8 ± 5.9 times the expression of the control as *Ppi* 203 causes disease in KW pea. This was over twice the expression seen in the KW pea apoplastic fluid which was 12.7 ± 2.3 times. All three bean cultivars showed an increase in *avrPpiA1* expression with CW being the highest of the three at 6.0 ± 1.0 times the control expression. RM was slightly higher than the control at 1.3 ± 0.1 times, but this was not significant.TG showed 1.9 ± 0.3 times the expression

of the control. These values were lower than in KW pea as *Ppi* 203 causes HR on bean cultivars. This was expected as TTEs are crucial to plant colonisation.

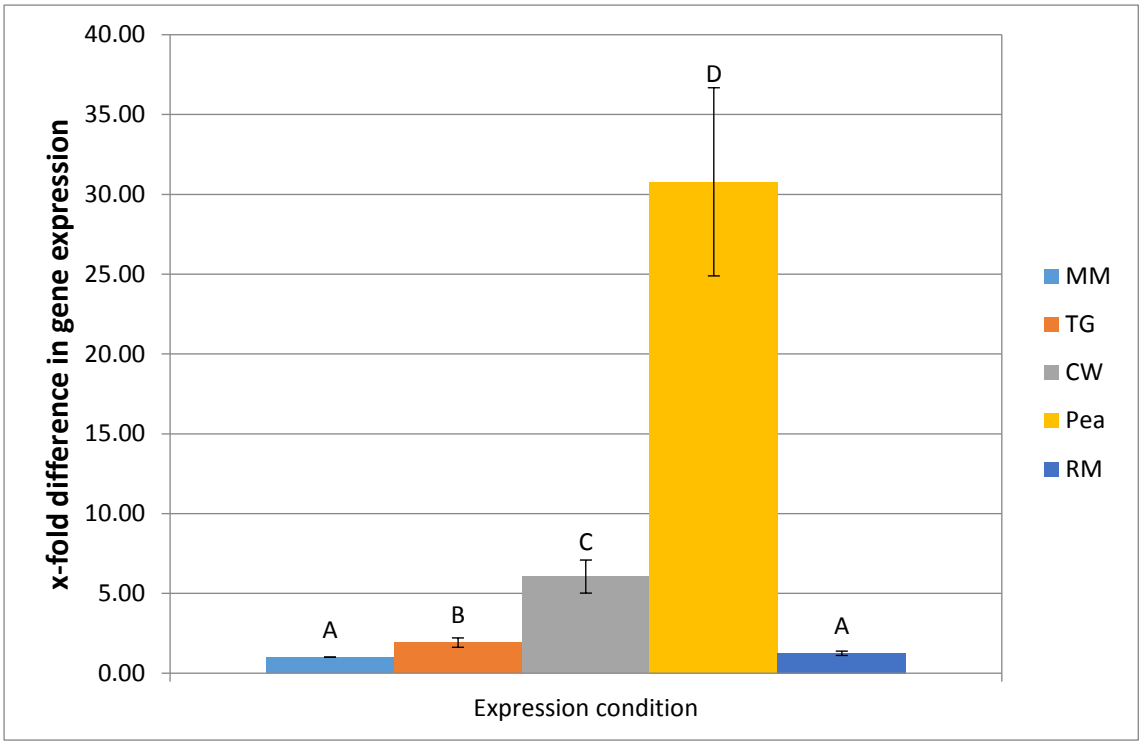


Figure 5.6: Gene expression of *avrPpiA1* from ILE *P. syringae* pv. *pisi* 203 *in planta*. Gene expression was measured via RT-qPCR following six hours of growth *in planta* of three bean cultivars, TG, CW and RM and in KW pea. X-fold differences in gene expression were compared to the control MM condition and normalised against DNA *gyrB*. Three replicates were tested and their mean values are displayed with standard error of the mean. (T-test and Tukey statistical analysis was used for significant differences [$p<0.05$] shown by letters above the bars).

The remaining ILE genes (Figure 5.7) were mostly downregulated with the *xerC* integrase gene being downregulated to below 0.5 times for all of the *in planta* conditions, this was also the case for the *rci* gene which was downregulated by all of the *in planta* conditions. The two further ILE integrase genes did show some upregulation in gene expression of between one and five times for TG, CW and KW pea, but was down regulated in RM. The *rulB'* region was only upregulated in RM, to 12.0 ± 2.6 times the control, but in all of the other

conditions it was downregulated to 0.1 ± 0.0 times for TG, 0.5 ± 0.2 times for CW and 0.5 ± 0.1 times for KW pea.

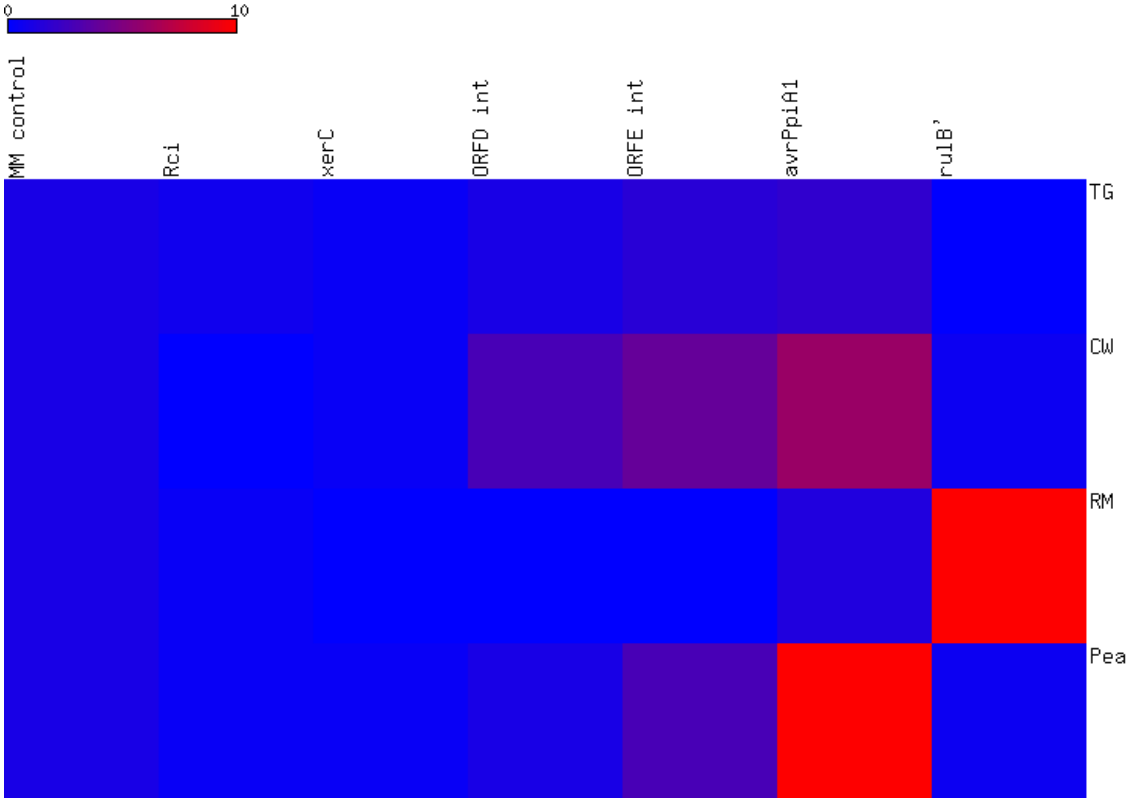


Figure 5.7: ILE gene expression from *P. syringae* pv. *pisi* 203 when tested *in planta*. Six ILE genes from *Ppi* 203 tested for expression following bacterial growth in TG, CW and RM beans and KW pea for six hours. The blue/ red spectrum represents increasing gene expression from 0 to 10 times the expression of the MM control (Full X-fold values available in Appendix IV).

5.2.2c: Expression of *P. syringae* pv. *pisi* 203 ILE genes following bacterial conjugation with *E. coli* DH5α.

Ppi 203 was subjected to conjugation between *Ppi* 203, *E. coli* DH5α and *E. coli* DH5α (pRK2013) (Figure 5.8). These results were interesting as five of the six ILE genes were downregulated compared to the MM control and *rci* was downregulated to 0.2 ± 0.0 times the expression. The three integrase genes at the 5' end of the ILE were all downregulated to similar values with both *xerC* and ORFD *int* being downregulated to 0.5 ± 0.1 and 0.40 ± 0.1 times expression respectively, which showed no significant difference between the two. The *rulB'*

region was also downregulated a similar amount as the integrase genes. The ORFE integrase was also downregulated, but not as much as the other integrases, to 0.6 ± 0.1 times expression. As showed with the apoplastic and *in planta* tests, *avrPpiA1* had the highest gene expression, although during conjugation it only reached 1.0 ± 0.1 times the expression of the MM control and was not significantly different from the MM control.

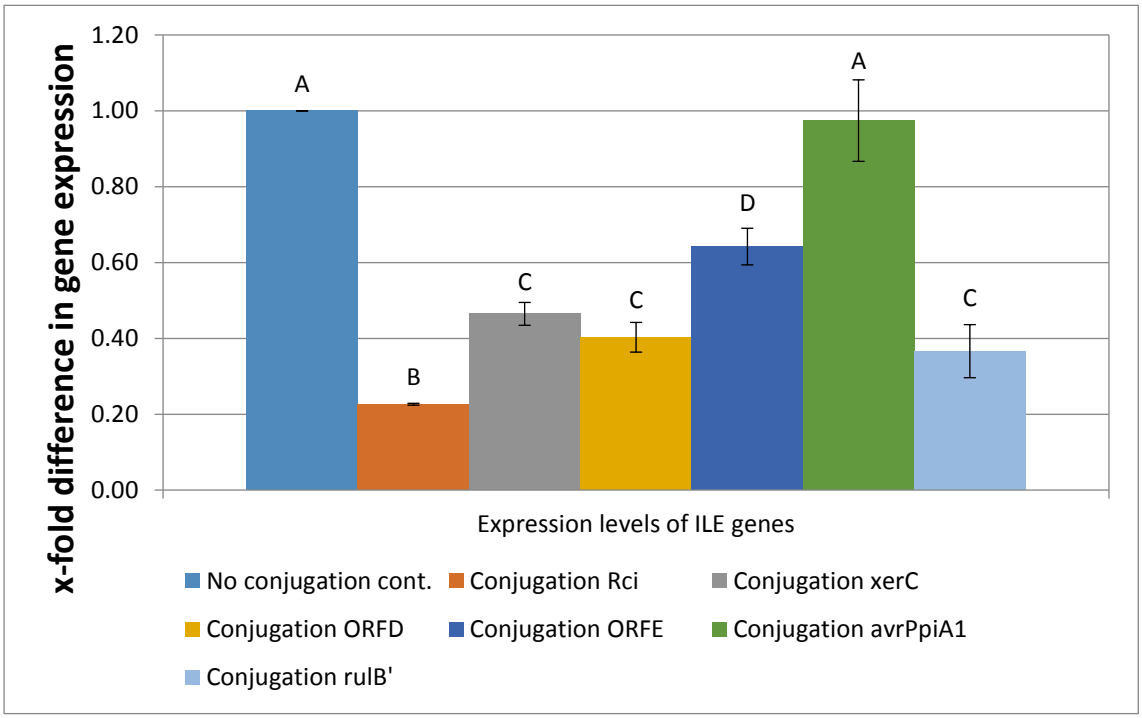


Figure 5.8: Gene expression of *P. syringae* pv. *pisi* ILE genes following 6 hours of conjugation with *E. coli* DH5α and *E. coli* DH5α (pRK2013). Gene expression was measured via RT-qPCR following six hours of bacterial conjugation with two *E. coli* strains. X-fold differences in gene expression were compared to the control MM condition and normalised against DNA *gyrB*. Three replicates per condition were tested and their mean values are displayed with standard error of the mean. (T-test and Tukey statistical analysis was used for significant differences [$p<0.05$] shown by letters above the bars).

5.2.2d: Effect of mitomycin C on bacterial growth.

The effect of mitomycin C (MMC) on bacterial growth was tested to observe any links between growth and gene expression and to also observe if complete cell death was caused by any of the concentrations used which would

have resulted in minimal or zero gene expression. Four different MMC concentrations were chosen, 0.05 µg/ mL, 0.1 µg/ mL, 0.5 µg/ mL and 1 µg/ mL. *Ppi* 203 and *Psy* 3023 (Figure 5.9) were both tested and viable cell counts were performed and colony forming units (CFU) / mL were calculated (see Section 2. 14.5).

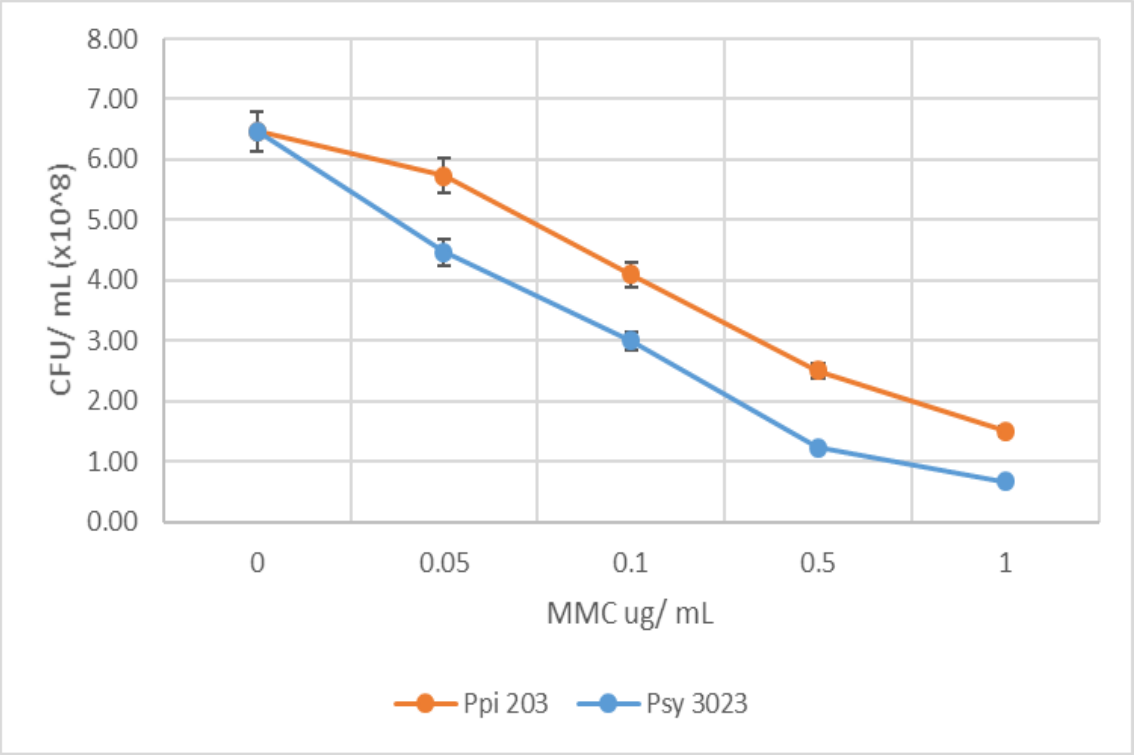


Figure 5.9: Viable counts of *Pseudomonas syringae* pv. *pisi* 203 and pv. *syringae* 3023 following growth in media containing mitomycin C for six hours. The cells were incubated in MM broth plus 0.05 µg/ mL, 0.1 µg/ mL, 0.5 µg/ mL and 1 µg/ mL MMC for six hours with subsequent growth on LB agar plates for 48 hours at 25ºC. A MM control was used as a control with no MMC present. All tests were performed in triplicate with standard deviation error bars shown.

As the *Ppi* 203 growth results show (Figure 5.9) the MMC did have an impact on cellular growth as with each increase in MMC concentration CFU/ mL decreased. However the highest concentration of MMC, 1 µg/ mL, did not cause complete cell death, but cellular growth was reduced. The growth of *Psy* 3023 in MMC shows a similar trend to *Ppi* 203 although the CFU/ mL values for *Psy* 3023

are lower for every MMC concentration. At the highest MMC concentration of 1 µg/ mL only one viable colony (data not shown) was present on two of the three plates and the third plate had no colonies present. This compared to *Ppi* 203 which had an average colony count of 15 at the highest MMC concentration. This difference may be due to more *Ppi* 203 cells being present at the start of the experiment.

These four concentrations of MMC were chosen as bacterial growth in MMC studies showed that these values caused a reduction in growth but do not kill all of the cells and therefore these concentrations caused a suitable bacterial stress (see Sections 5.2.2e and 5.2.3e for gene expression results).

5.2.2e: Expression of *P. syringae* pv. *pisi* 203 ILE genes following exposure to increasing concentrations of mitomycin C.

It was hypothesised that some of the ILE genes may be regulated by the SOS response following DNA damage. To test this *Ppi* 203 was incubated for six hours with MMC, a DNA damaging agent which creates DNA crosslinks. The gene expression profiles of MMC treated cells (Figure 5.10) showed that some ILE gene expression appears to increase as the dose of MMC increases

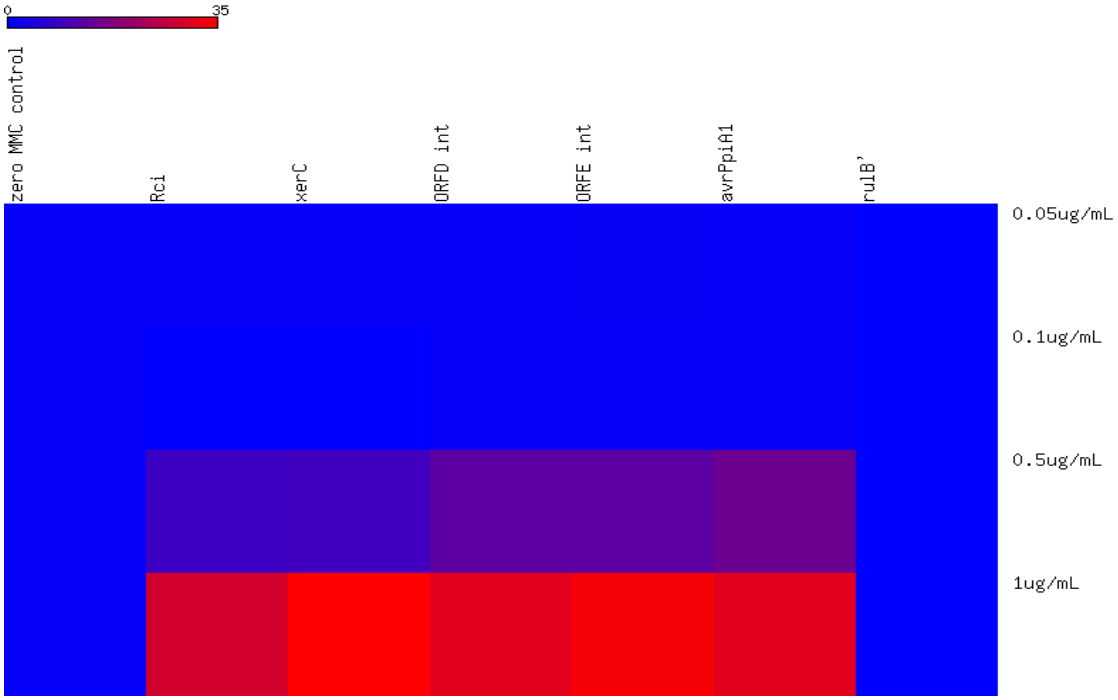


Figure 5.10: ILE gene expression from *P. syringae* pv. *pisi* 203 when incubated for six hours with the addition of mitomycin C. Six ILE genes from *Ppi* 203 tested for ILE gene expression following bacterial growth in media supplemented with either 0.05, 0.1, 0.5 or 1 µg/ mL MMC. Cells were incubated for six hours prior to expression analysis. The blue/ red spectrum represents increasing gene expression from 0 to 35 times the expression of the zero MMC control (Full X-fold values available in Appendix IV).

The expression of *rulB'* did not reach the same level as the MM control and stays at zero for all of the MMC doses except for 0.05 µg/ mL MMC where expression reaches 0.5 ± 0.1 times the control. However, for the other ILE genes, expression increases greatly with 0.5 µg/ mL MMC, with *xerC* at 9.2 ± 0.1 times the control expression, *avrPpiA1* at 14.9 ± 1.8 and *rci* at 8.4 ± 0.4 times the control expression. The expression increases again with 1 µg/ mL MMC with *xerC* at 39.3 ± 0.2 times the control expression, *avrPpiA1* at 30.7 ± 0.7 and *rci* at 28.7 ± 1.3 times the control expression.

5.2.2f: Expression of *P. syringae* pv. *pisi* 203 ILE genes following exposure to increasing doses of UVB irradiation.

The UV tests on ILE gene expression were performed following a similar hypothesis to that of the MMC experiments. UVB causes DNA damage and increased ssDNA which activates the SOS response. *Ppi* 203 was subjected to four different doses of UVB irradiation (302nm), 15, 30, 45 and 60 seconds along with a no UVB control test (Figure 5.11).

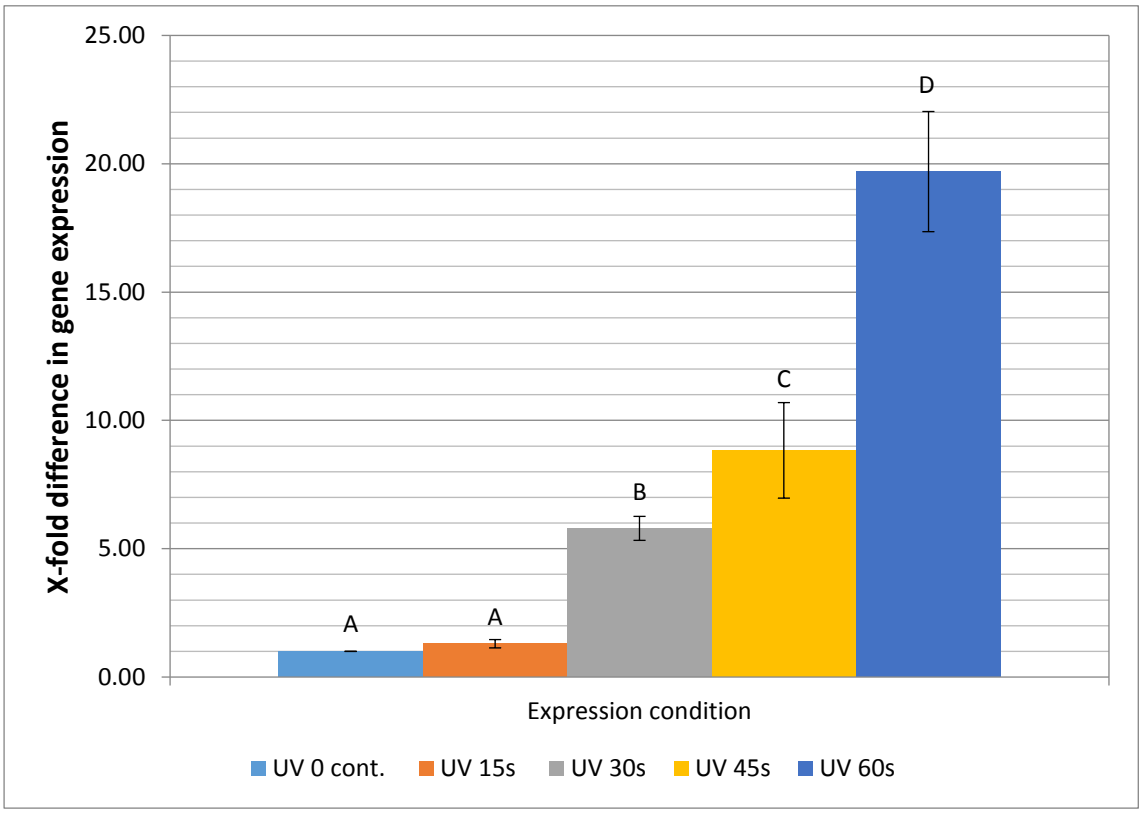


Figure 5.11: Gene expression of ILE *P. syringae* pv. *pisi* 203 *avrPpiA1* following UVB irradiation (302nm) for different exposure times. *Ppi* 203 cells were exposed to UVB irradiation for either 15, 30, 45 or 60 seconds and subsequently recovered in LB broth for six hours. The x-fold expression for *avrPpiA1* was normalised against the expression of *avrPpiA1* with no UVB irradiation. Three replicates per condition were tested and their mean values are displayed with standard error of the mean. (T-test and Tukey statistical analysis was used for significant differences [$p<0.05$] shown by letters above the bars).

Interestingly the gene expression of *avrPpiA1* increased in a dose dependent manner with the UVB exposure. The longer the bacterium were exposed to UVB irradiation the greater the expression of *avrPpiA1.* The expression of *avrPpiA1* started at 1.3 ± 0.2 times the control with only 15 seconds of UVB irradiation, but increased with 60 seconds of UVB irradiation to 19.7 ± 2.3 times the *avrPpiA1* gene expression of the no UVB control. This dose dependent response to UVB irradiation was also seen for some of the other ILE genes including the two integrase ORFD and E along with *xerC* and the *rci* gene (Figure 5.12). The *rulB'* expression levels were different to the other genes as *rulB'* had very low expression for the first three exposure levels with the values all equal to or below one. However, with the 60 seconds of UVB exposure the expression of *rulB'* increased to 5.4 ± 1.8 times the no UVB control expression level (Figure 5.12)

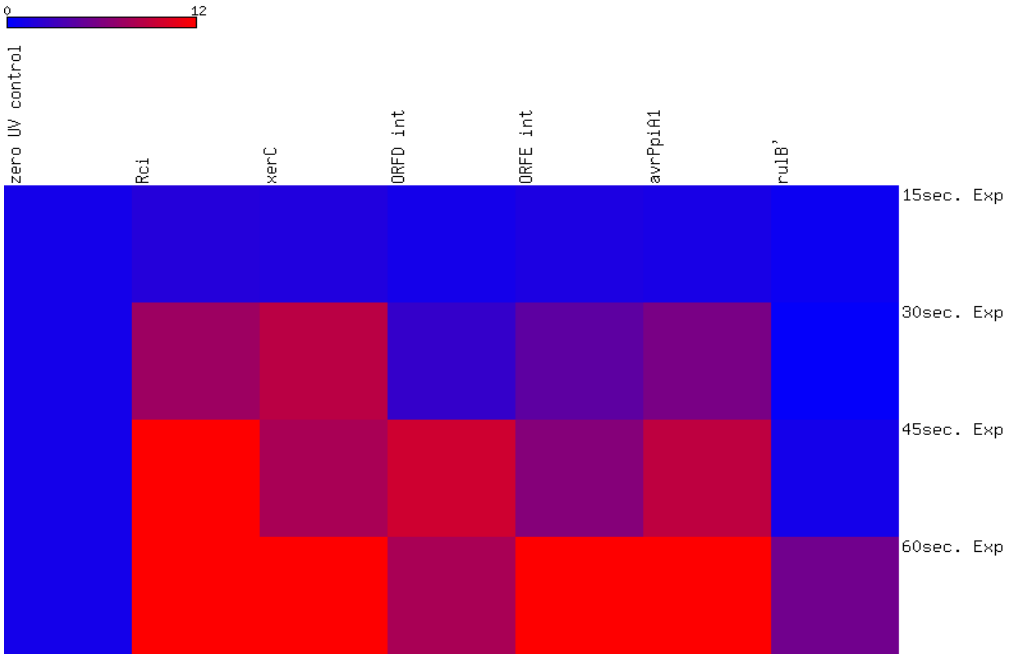


Figure 5.12: ILE gene expression from *P. syringae* pv. *pisi* 203 following exposure to UVB irradiation for 15, 30, 45 and 60 seconds. The heat map shows the relationship between increasing UVB exposure and gene expression of all four ILE genes plus *rci* and *rulB'*. The blue/ red spectrum represents increasing gene expression from 0 to 12 times the expression of the no UV control (Full X-fold values available in Appendix IV).

5.2.2g: Expression of *P. syringae* pv. *pisi* 203 ILE genes when stressed by sub-optimal growth temperatures.

ILE gene expression tests on *Ppi* 203 were performed using different sub-optimal temperatures during the growth of *Ppi* 203 over the course of six hours. The ILE gene expression results from this test did not show a lot of gene expression with many of the genes having lower expression levels than the bacterial cells grown at the optimal temperature of 25ºC (Figure 5.13). There were some slight exceptions to this rule which included *xerC* expression when grown at 4ºC. However, this was only an expression increase of 0.1 ± 0.1, from the 25ºC control of one. The ORFD integrase also showed some slight increases in gene expression when subjected to -80ºC and 4ºC as the expression levels were increased to 1.2 ± 0.0 and 1.3 ± 0.0 respectively.

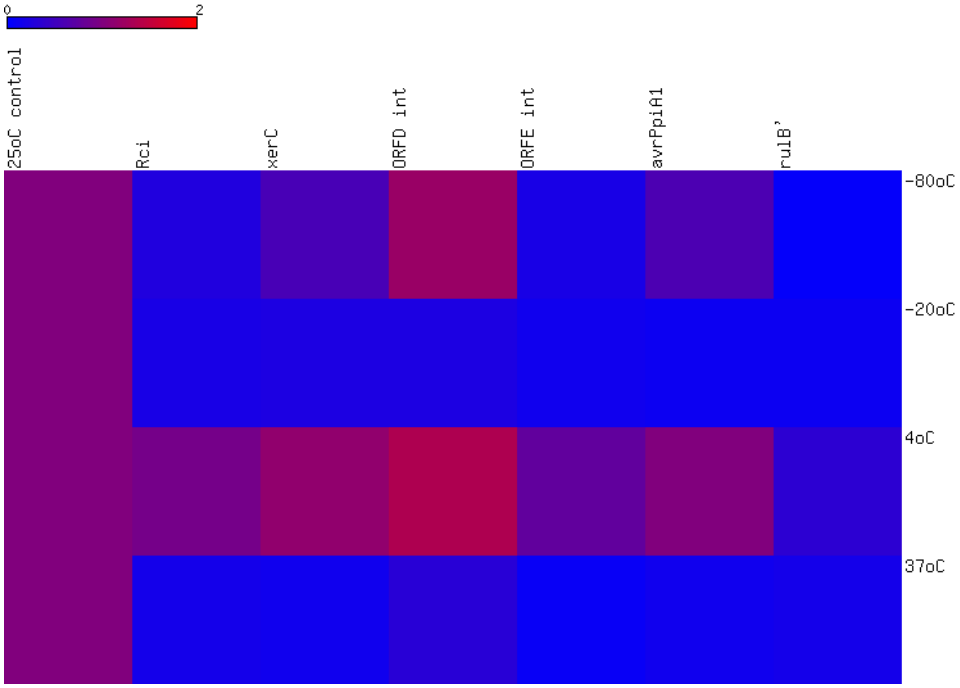


Figure 5.13: ILE gene expression from *P. syringae* pv. *pisi* 203 following growth at different sub-optimal temperatures. This shows that many of the genes were downregulated compared to the control and even the one that were upregulated were only upregulated to a maximum of 1.4 times the control expression level. If the threshold was higher than 2 like in previous heat maps then the grid would be entirely blue due to the low expression levels. The blue/ red spectrum represents increasing gene expression from 0 to 2 times the expression of the 25ºC control (Full X-fold values available in Appendix IV).

5.2.3g: Summary of ILE integrase and TTE gene expression from *P. syringae* pv. *pisi* 203.

The data for all of the gene expression levels of the ILE in *Ppi* 203 were pooled and used to create a gene expression heat map matrix to allow comparison of ILE gene expression under different conditions with the same scale (Figure 5.14). This is the same data as shown in the previous figures but all under the same scale bar.

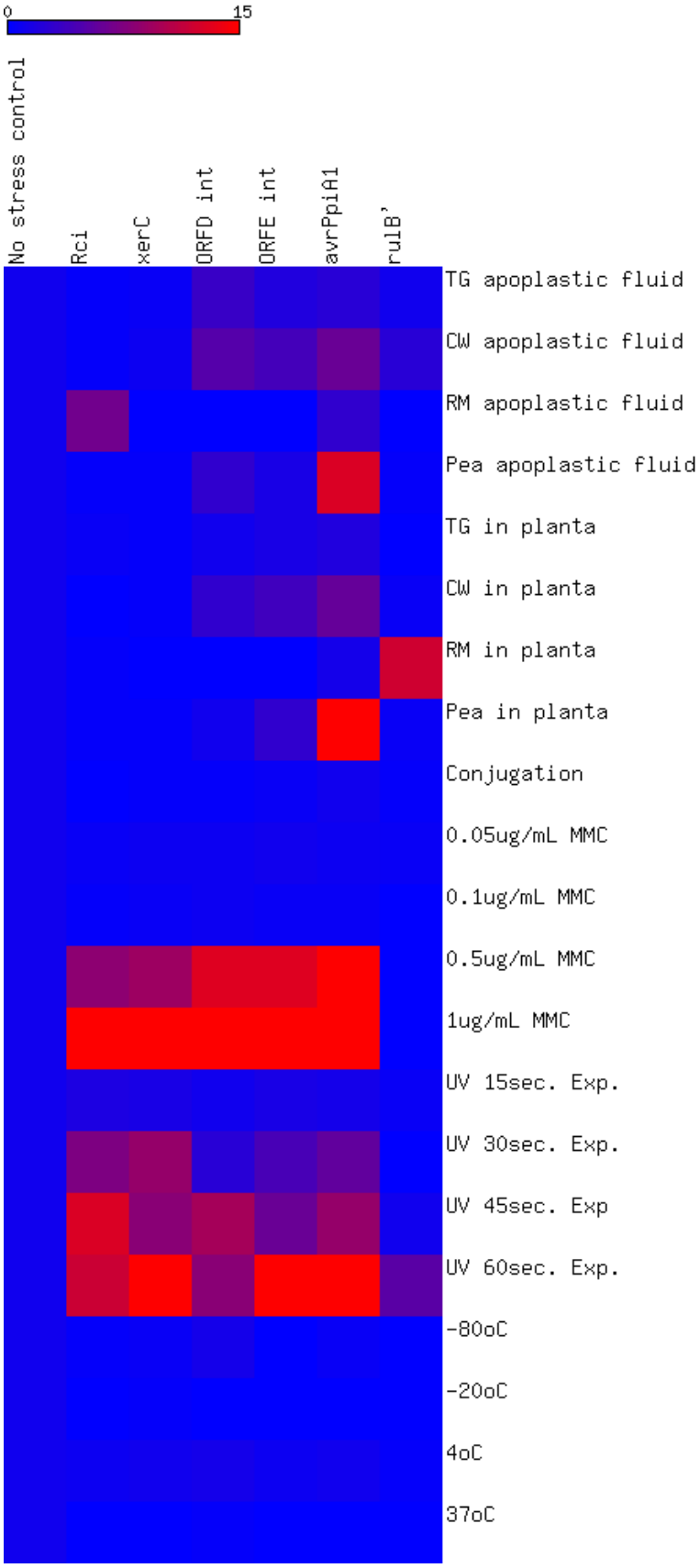


Figure 5.14: Heat map matrix of ILE *Ppi* 203 gene expression following various bacterial stresses. The matrix comprises the expression levels of six different ILE associated genes under six different types of bacterial stress, plant apoplastic fluid, *in planta,* bacterial conjugation with *E. coli,* mitomycin C, UVB irradiation and sub-optimal temperature. The blue/ red spectrum represents increasing gene expression from 0 to 15 times the expression of the control. (A full report of mean X-fold values is available in Appendix IV).

5.2.3: Expression of *P. syringae* pv. *syringae* 3023 ILE genes.

Following the gene expression studies on the ILE in *Ppi* 203 another ILE was selected for the same gene expression tests. The ILE identified in *P. syringae* pv. *syringae* (*Psy*) 3023 was chosen to compare ILE gene expression across two different systems. Similar genes are present in *Psy* 3023 and *Ppi* 203 as both contain three integrase genes at the 5' end and they both contain type three effector (TTE) genes. There are some differences between the ILEs as *Psy* 3023 contains two TTE compared to *Ppi* 203 only having one and the *Psy* 3023 ILE also contains a large transposon-like insertion making it 7.2kb longer. The genes studied from ILE *Psy* 3023 were the *xerC* integrase, the two TTE, HopH1 and HopAP1 and the smaller fragment of the disrupted *rulB* gene, *rulB'* (Figure 5.15). The expression of these genes was compared to *gyrB*. All significant differences were measured to $p<0.05$.

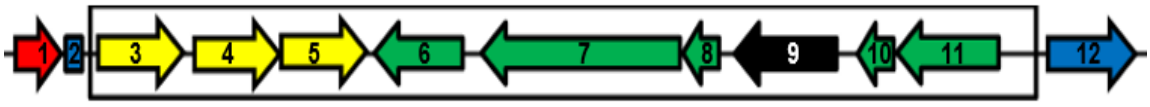


Figure 5.15: Genetic makeup of ILE from *P. syringae* pv. *syringae* 3023. The ILE in *Psy* 3023 contains 11 open reading frames plus the disrupted *rulB* gene. **1**; *rulA,* **2**; *rulB',* **3**; *xerC,* **4** and **5**; *int,* **6**; *hopH1,* **7**; transposase, **8**; recombinase, **9**; hypothetical, **10**; resolvase, **11**; *hopAP1,* **12**; *rulB"*. From these 12 regions **2**, **3**, **6** and **11** were chosen for gene expression studies.

5.2.3a: Expression of *P. syringae* pv. *syringae* 3023 ILE genes following inoculation into plant apoplastic fluid.

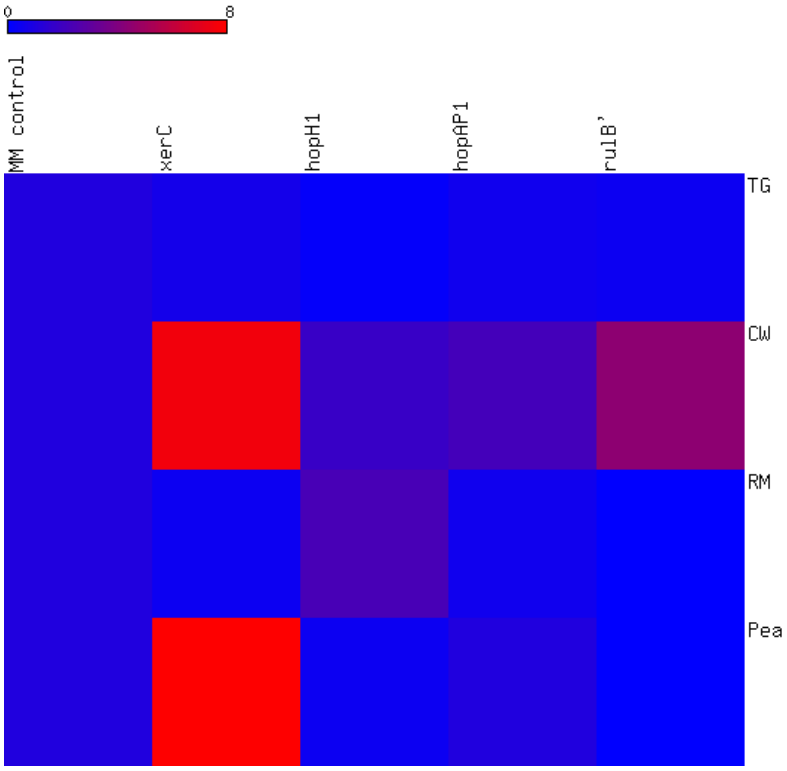


Figure 5.16: ILE gene expression heat map from *P. syringae* pv. *syringae* 3023 when grown in plant apoplastic fluid. Four ILE genes from *Psy* 3023 tested for expression following bacterial growth in bean and KW pea apoplastic fluid for six hours. The blue/ red spectrum represents increasing gene expression from 0 to 8 times the expression of the MM control (Full X-fold values available in Appendix IV).

ILE gene expression was measured following bacterial growth in plant apoplastic fluid for six hours with MM as the control. The expression of the first ILE gene, *xerC,* showed no increase in gene expression when TG and RM apoplastic fluid was tested. TG produced an x-fold difference of 0.7 ± 0.2 and RM was 0.5 ± 0.1. However, expression of *xerC* did increase when *Psy* 3023 was grown in both CW and KW pea apoplastic fluid. CW caused an increase in expression to 7.5 ± 3.7 times the MM control and KW pea increased expression by 11.6 ± 4.0 times. CW apoplastic fluid also increased the expression of the other genes tested. *HopH1* expression increased by 1.8 ± 0.4, *hopAP1* increased by 2.2 ± 0.8 and *rulB'* increased by 4.4 ± 1.8. The majority of the other genes

were either downregulated below the control expression level or were the same as the control. One exception of this was *hopH1* expression in RM apoplastic fluid where the expression of *hopH1* increased by 2.3 ± 1.0 times (Figure 5.16). This also showed uncoupled regulation between *rulB'* and *xerC* indicating they do not share a promoter.

5.2.3b: Expression of *P. syringae* pv. *syringae* 3023 ILE genes following inoculation *in planta*.

Following the apoplastic fluid expression tests it was important to assess if similar expression patterns are observed *in planta.* To do this *Psy* 3023 was inoculated into TG, CW, RM bean and KW pea plants and incubated for six hours. Some similarities were seen between the data for apoplastic fluid treatment and *in planta* studies (Figure 5.17). Changes in TTE expression may be due to *Psy* 3023 causing disease in bean cultivars but HR in the KW pea.

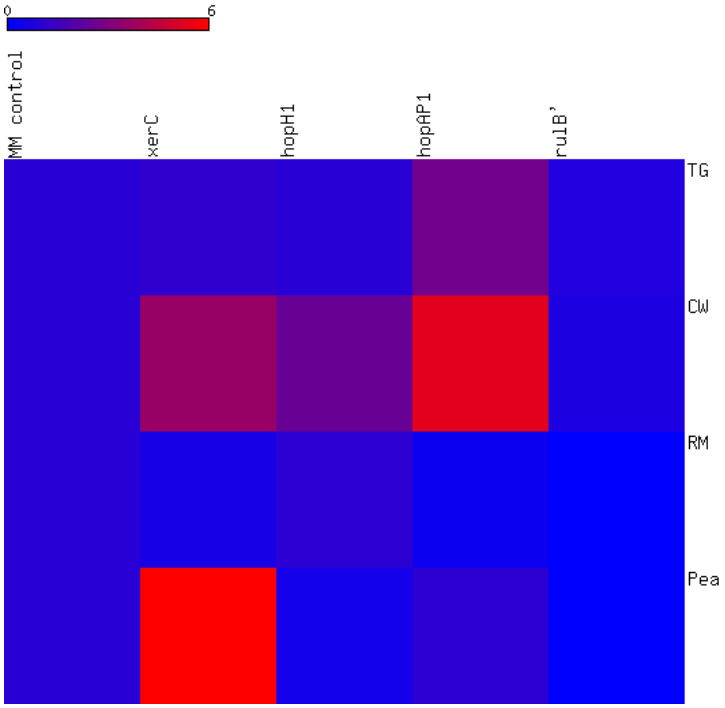


Figure 5.17: ILE gene expression from *P. syringae* pv. *syringae* 3023 when tested *in planta*. The four ILE genes from *Psy* 3023 had their gene expression profiles analysed bacterial growth in TG, CW and RM beans and KW pea for six hours. The blue/ red spectrum represents increasing gene expression from 0 to 6 times the expression of the MM control (Full X-fold values available in Appendix IV).

The *xerC* gene was also upregulated in both CW and KW pea *in planta* by 3.6 ± 0.7 and 14.5 ± 3.4 times the control expression. Similar to the apoplastic fluid expression studies, CW causes both TTE genes to become upregulated to 2.5 ± 0.3 for *hopH1* and 5.3 ± 1.0 for *hopAP1.* Some differences do occur with the *rulB'* expression, as *in planta* CW tests failed to produce the increase in *rulB'* expression seen in the CW apoplastic fluid. A final point about the *in planta* test was that TG shows slightly higher expression of all the genes compared to the TG apoplastic fluid tests.

5.2.3c: Expression of *P. syringae* pv. *syringae* 3023 ILE genes following bacterial conjugation with *E. coli* DH5α.

Bacterial conjugation between *Psy* 3023 and *E.coli* DH5α caused all of the genes on the ILE to be upregulated. The only gene to be downregulated by bacterial conjugation was the *rulB'* gene (Figure 5.18). This indicates that bacterial conjugation may have a role in ILE movement and dissemination in the case of *Psy* 3023 as this result was not seen with *Ppi* 203.

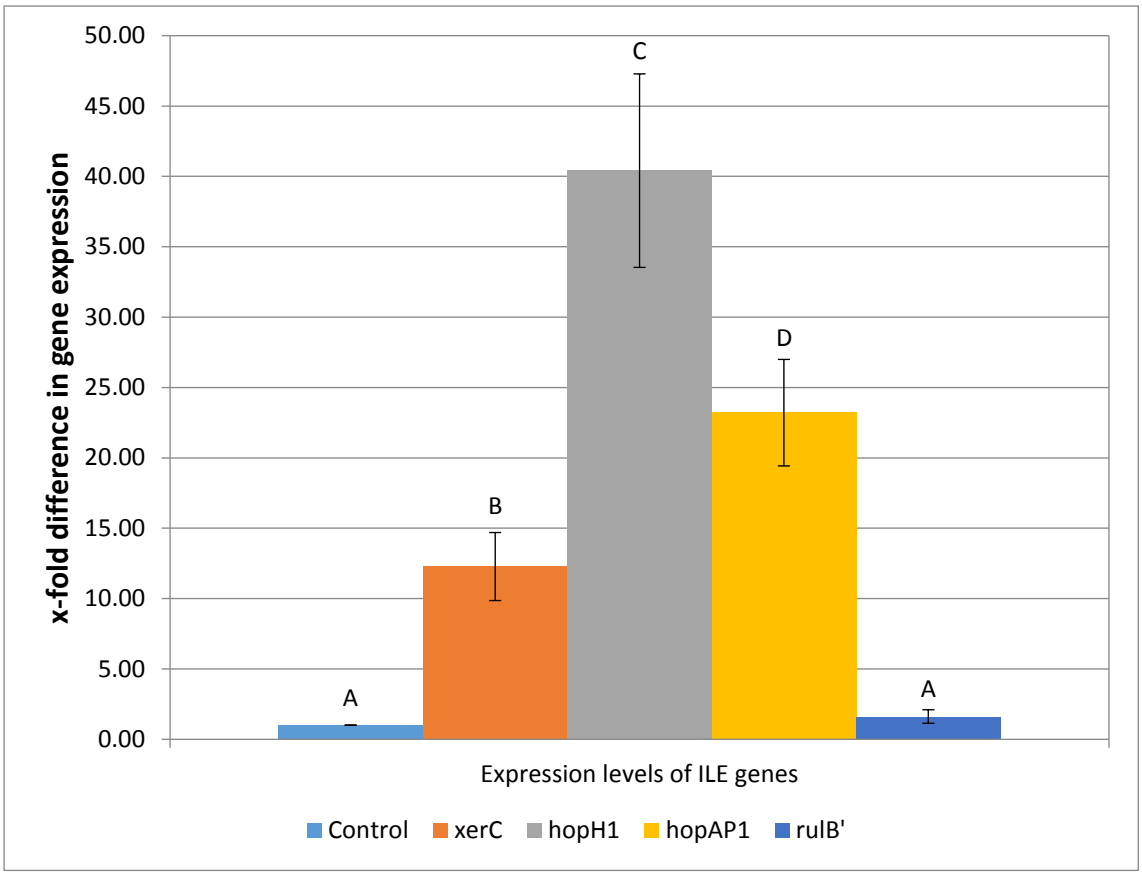


Figure 5.18: Gene expression of ILE *P. syringae* pv. *syringae* 3023 genes following 6 hours of conjugation with *E. coli* DH5α and *E. coli* DH5α (pRK2013) together. Gene expression was measured via RT-qPCR following six hours of bacterial conjugation with two *E. coli* strains. X-fold differences in gene expression were compared to the control MM condition without conjugation and normalised against DNA *gyrB*. Three replicates per condition were tested and their mean values are displayed with standard error of the mean. (T-test and Tukey statistical analysis was used for significant differences [$p<0.05$] shown by letters above the bars).

As figure 5.18 shows, the *xerC* gene present on the ILE was upregulated by conjugation to 12.3 ± 2.4 times the expression of the no conjugation control. *HopH1* and *hopAP1* were also upregulated by 40.4 ± 6.9 and 23.2 ± 3.8 times the control respectively. The *rulB'* gene was slightly upregulated to 1.6 ± 0.5, however this increase in gene expression was not significantly different from the no conjugation control.

5.2.3d: Expression of *P. syringae* pv. *syringae* 3023 ILE genes following exposure to increasing concentrations of mitomycin C.

The gene expression of the ILE in *Psy* 3023 was different from the gene expression seen in the *Ppi* 203 ILE. Whereas in *Ppi* 203 the ILE genes increased in expression as the MMC concentrations increased in a dose dependent manner, *Psy* 3023 did not respond in a dose dependent way (Figure 5.19). The overall expression values were also far lower than seen with *Ppi* 203.

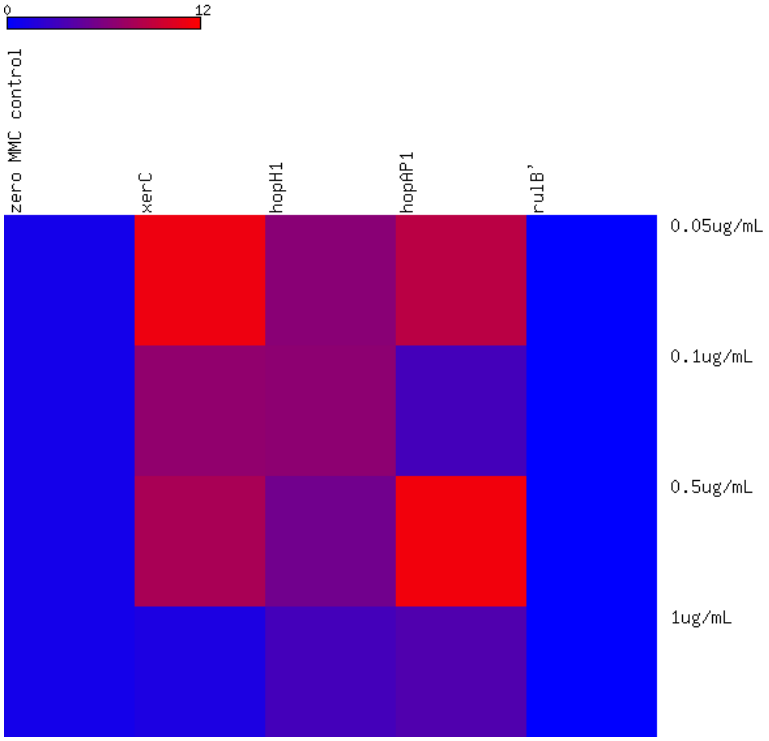


Figure 5.19: ILE gene expression from *Psy* 3023 when incubated for six hours with the addition of MMC. Six ILE genes from *Psy* 3023 tested for ILE gene expression following bacterial growth in media supplemented with either 0.05, 0.1, 0.5 or 1 µg/ mL MMC. Cells were incubated for six hours prior to expression analysis. The blue/ red spectrum represents increasing gene expression from 0 to 12 times the expression of the control containing no MMC (Full X-fold values available in Appendix IV).

Many of the *Psy* 3023 ILE genes were upregulated with the lowest concentration of MMC. *XerC* showed an increase in gene expression of 11.1 ± 2.4 with 0.05 µg/ mL, this did not significantly change when 0.1 µg/ mL and 0.5 µg/ mL MMC was used as *xerC* expression was 6.8 ± 1.4 and 8.0 ± 2.6

respectively, for the higher concentrations of MMC. However, a further increase in MMC concentration to

1 µg/ mL caused gene expression to decrease to 1.4 ± 0.3 which was not significantly different to the no MMC control. The pattern of increased gene expression with the lower MMC concentrations continued with both TTE genes, but their expression was also increased following treatment with 1 µg/ mL MMC. *HopH1* increased in expression to 3.3 ± 1.3 and *hopAP1* increased to 3.8 ± 1.0 (Figure 5.20). An interesting final point was that the expression of *rulB'* was effectively zero for all of the MMC concentrations tested, similar to the result seen from *Ppi* 203.

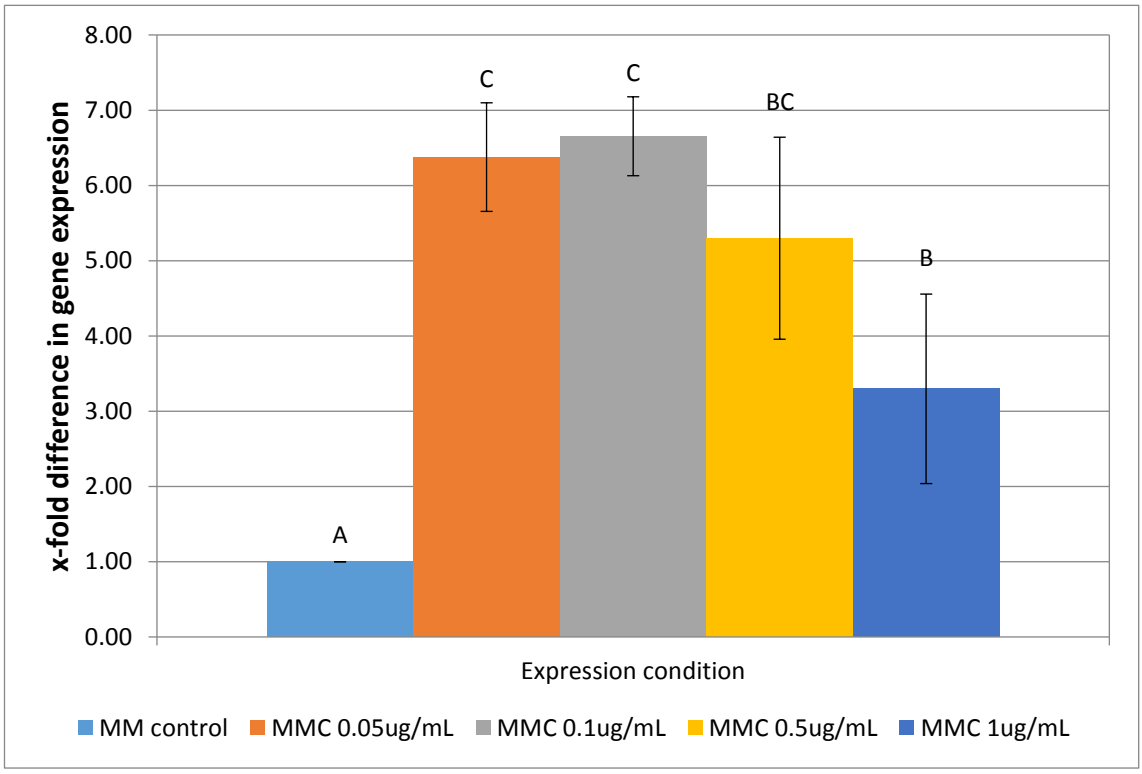


Figure 5.20: Type three effector, *hopH1,* expression from *Psy* 3023 when grown in media supplemented with increasing concentrations of mitomycin C. *Psy* 3023 was grown in media containing four different concentrations of MMC, 0.05, 0.1, 0.5 and 1 µg/ mL. Following this the expression of *hopH1* was analysed relative to a no MMC control. X-fold differences in gene expression were compared to the control MM condition and normalised against DNA *gyrB*. Three replicates per condition were tested and their mean values are displayed with standard error of the mean. (T-test and Tukey statistical analysis was used for significant differences [$p<0.05$] shown by letters above the bars).

5.2.3e: Expression of *P. syringae* pv. *syringae* 3023 ILE genes following exposure to increasing doses of UVB irradiation.

The expression of ILE *xerC* integrase in *Psy* 3023 appeared to follow the reverse pattern of what was seen in *Ppi* 203 when the cells were subjected to increasing UVB exposure (Figure 5.21). The overall ILE expression levels seen in *Psy* 3023 following UVB exposure were lower than those seen in *Ppi* 203. In *Ppi* 203 the *xerC* expression increased as the exposure time increased whereas in *Psy* 3023 the *xerC* expression peaked at 30 seconds of UVB exposure before falling. This is clearly shown in the x-fold differences in gene expression as with the lowest UVB exposure level of 15 seconds *xerC* expression increased 4.5 ± 2.0 fold. A similar value of 4.7 ± 0.7 was observed with 30 seconds of UVB exposure. However, with 45 seconds of exposure *xerC* expression decreased to 2.5 ± 0.8 with a further decrease to 1.2 ± 0.3 following 60 seconds of UVB exposure.

An inverse relationship between UVB exposure and gene expression was also present when analysing the *hopAP1* expression. *HopAP1* expression started at one with the no UVB control before peaking at 2.7 ± 1.1 with 15 seconds of UVB exposure and decreased in expression until it reached 0.5 ± 0.2 times the expression of the no UVB control with 60 seconds of UVB exposure. The two remaining genes, *hopH1* and *rulB'* both peaked at different UVB exposure levels. *HopH1* was upregulated with all of the UVB exposure levels with no significant differences between 15, 30 and 45 seconds. Although the peak in expression was seen with 45 seconds of UVB exposure as *hopH1* expression increased to 1.4 ± 0.3 times the no UVB control. The gene expression then decrease with 60 seconds UVB exposure to 0.2 ± 0.0. The *rulB'* gene also showed no dose dependent response to UVB, although the shortest exposure

did increase its expression 2.7 ± 0.5 fold and 30 seconds of UVB exposure decreased expression to 0.3 ± 0.0. Both 45 and 60 seconds resulted in no significant difference from the no UV control condition, this may be due to *Psy* 3023 being more susceptible to UVB irradiation at 30 seconds than *Ppi* 203 (Chapter 7). However, with 60 seconds of UVB exposure results suggest *Psy* 3023 was less susceptible than *Ppi* 203.

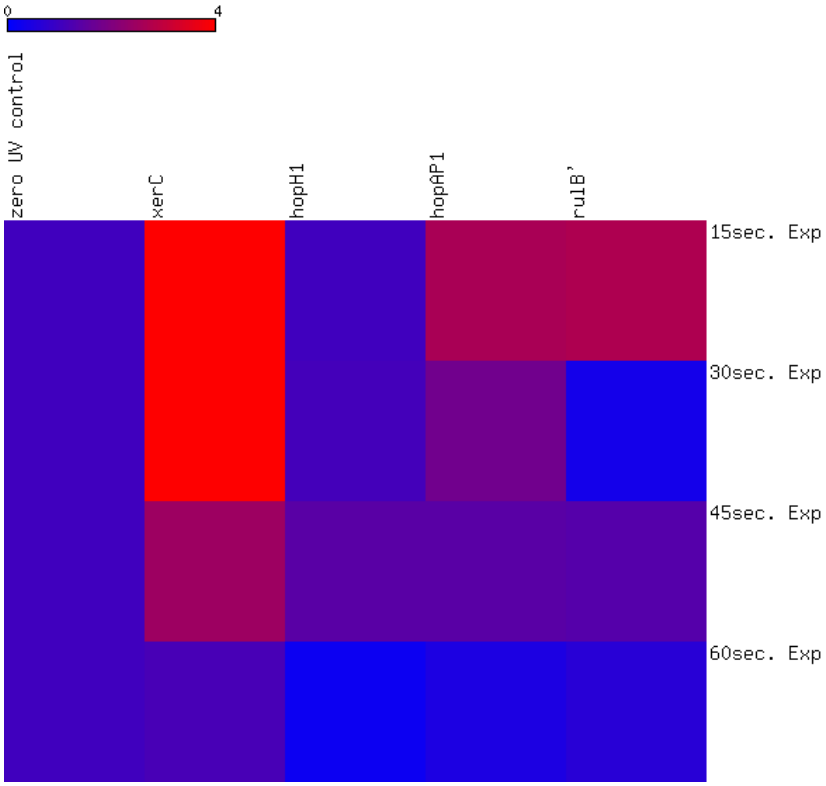


Figure 5.21: ILE gene expression from *P. syringae* pv. *syringae* 3023 following exposure to UVB irradiation for 15, 30, 45 and 60 seconds. The heat map shows the relationship between increasing UVB exposure and gene expression of all three ILE genes plus *rulB'*. The blue/ red spectrum represents increasing gene expression from 0 to 4 times the expression of the zero UVB control (Full X-fold values available in Appendix IV).

5.2.3f: Expression of *P. syringae* pv. *syringae* 3023 ILE genes when stressed by sub-optimal growth temperatures.

The effect of temperature on the gene expression of the *Psy* 3023 ILE was interesting as only one temperature produced an increase in gene expression across all four genes. The increase in gene expression was caused by incubating the cells at 4ºC for six hours. This temperature caused *xerC* expression to

increase by 4.8 ± 2.2 times the control, *hopH1* (Figure 5.22) and *hopAP1* to increase by 2.7 ± 0.4 and 2.4 ± 0.5 times the 25ºC control respectively, and *rulB'* to increase by 3.2 ± 0.7 fold. Some of the other temperatures did increase gene expression. For example, -20ºC increased the expression of *hopH1* by 1.8 ± 0.2 fold and -80ºC increased *rulB'* expression by 4.8 ± 0.9 fold. *HopH1* was downregulated 0.5 ± 0.1 fold when grown at 37ºC (Figure 5.23). These expression results also show how *rulB'* expression does not relate to *xerC* expression suggesting different regulation of the two genes.

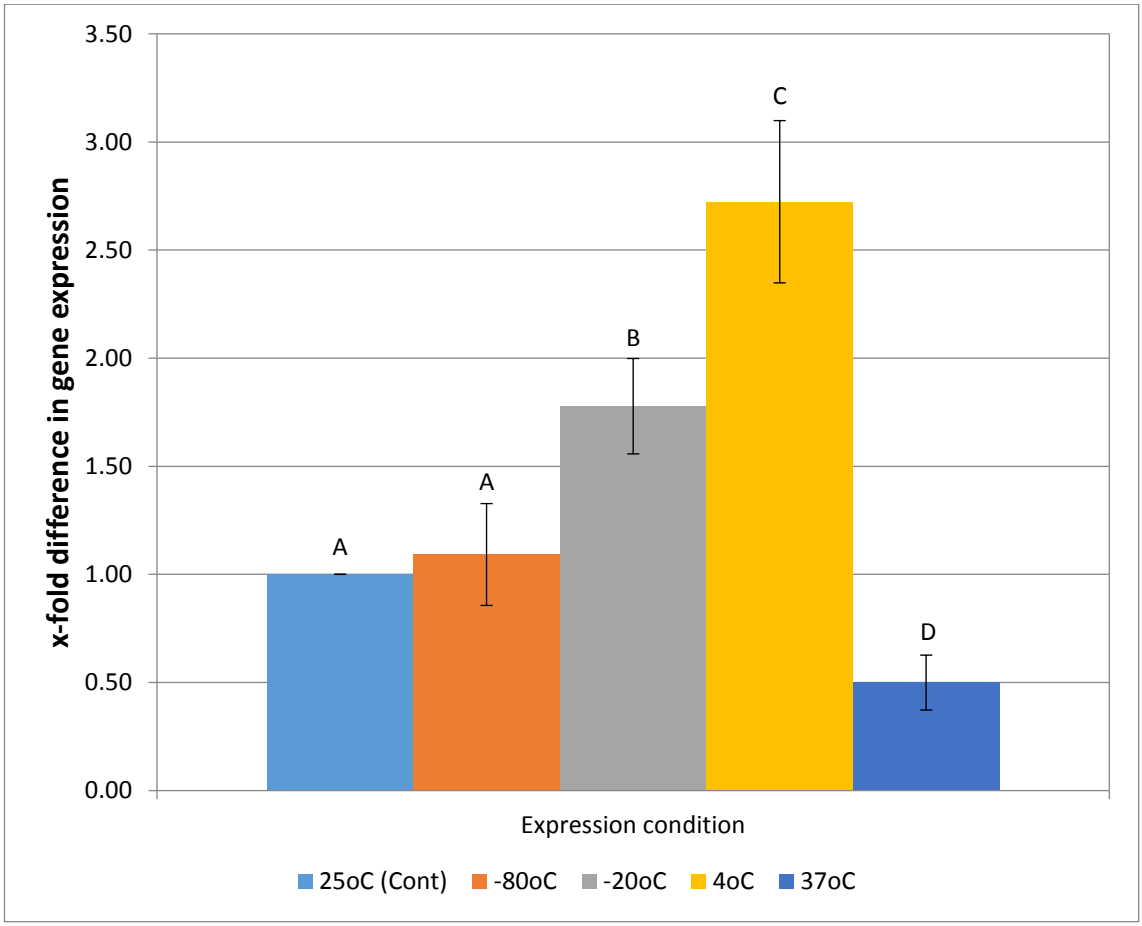


Figure 5.22: Gene expression of *P. syringae* pv. *syringae* 3023 ILE *hopH1* following growth at sub-optimal temperatures. *Psy* 3023 cells were grown for six hours at sub-optimal temperatures. The temperatures used were -80, -20, 4 and 37ºC. The x-fold expression of *hopH1* was normalised against the expression of *hopH1* when *Psy* 3023 was grown at its optimal temperature of 25ºC. Three replicates per condition were tested and their mean values are displayed with standard error of the mean. (T-test and Tukey statistical analysis was used for significant differences [$p<0.05$] shown by letters above the bars).

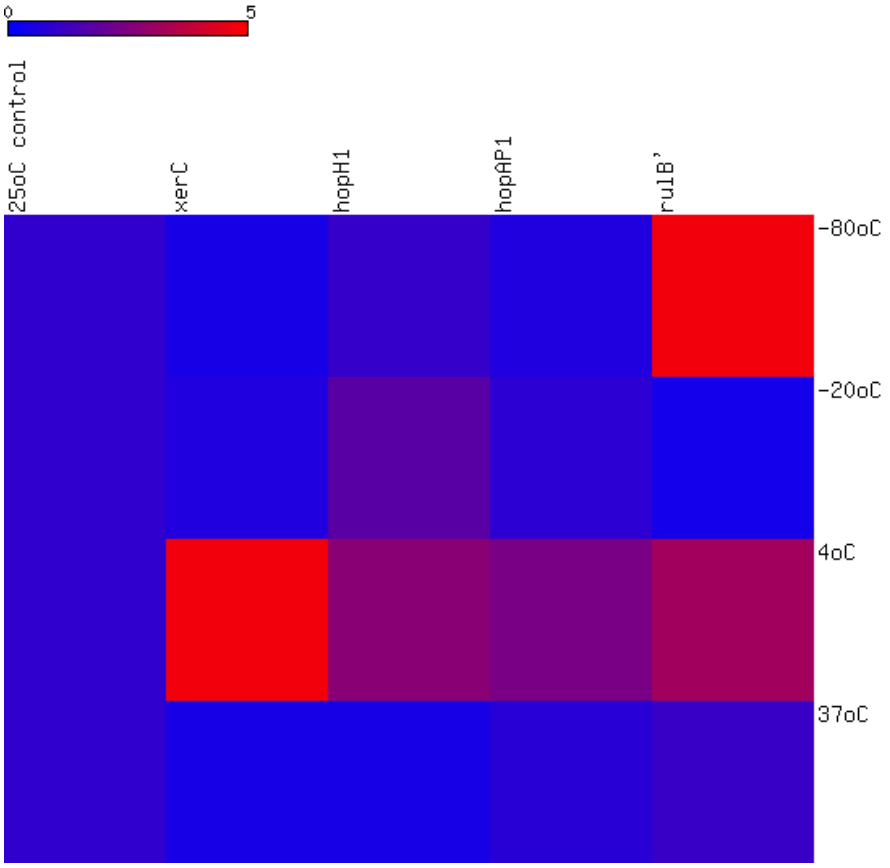


Figure 5.23: ILE gene expression from *P. syringae* pv. *syringae* 3023 following growth at different sub-optimal temperatures. The expression matrix shows how the genes were expressed in regard to the growth temperature. Many of the genes had low expression although 4ºC does show increases for all of the genes tested. The blue/ red spectrum represents increasing gene expression from 0 to 5 times the expression of the optimal 25ºC control (Full X-fold values available in Appendix IV).

5.2.3g: Summary of ILE integrase and TTE gene expression from *P. syringae* pv. *syringae* 3023.

The data from all of the conditions for all of the genes was compiled to make one overall figure for all of the gene expression in the ILE from *Psy* 3023 (Figure 5.24). This allows comparison between all of the different conditions but also between the two distinct ILEs on the same scale. This is the same data as shown in the previous figures but all under the same scale bar.

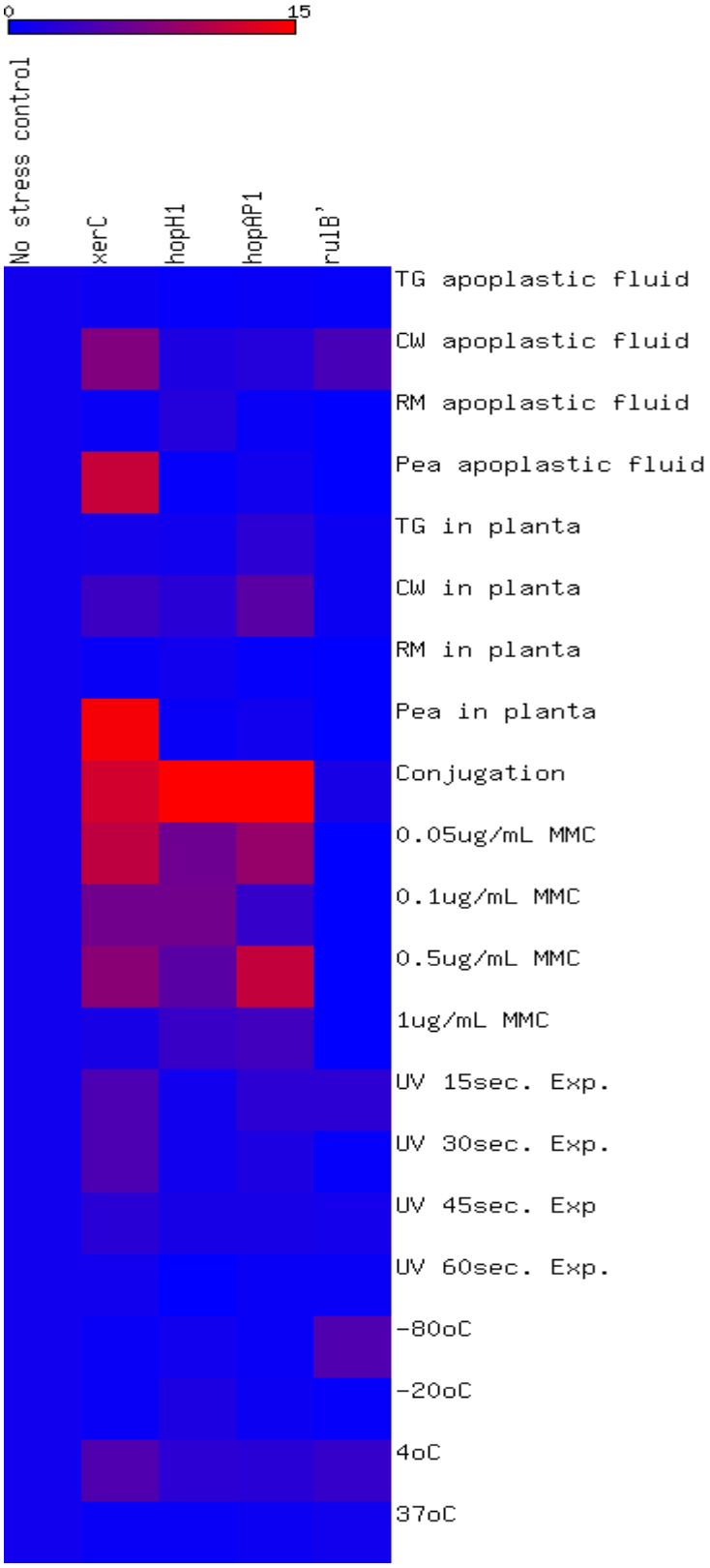


Figure 5.24: Heat map matrix of *P. syrinage* pv. *syringae* 3023 ILE gene expression following various bacterial stresses. The matrix comprises the expression levels of four different ILE associated genes under six different types of bacterial stress, plant apoplastic fluid, *in planta,* bacterial conjugation with *E. coli,* mitomycin C, UVB irradiation and sub-optimal temperature. The blue/ red spectrum represents increasing gene expression from 0 to 15 times the expression of the various no stress controls. (A full report of mean X-fold values is available in Appendix IV).

5.3: Discussion

Very little was known about the expression of genes in *P. syringae* ILEs when the bacteria are subjected to environmental stresses. It will be of interest to know if ILE genes are expressed and what triggers their expression. The expression of the integrase genes are important to study as they may indicate what causes the ILE to move from different *rulB* genes but also what stresses may increase the ILEs ability to capture exogenous genes. The expression of TTE genes is important as they play a role in disease progression.

To mimic environmental stresses that the bacteria may encounter they were subjected to six different stresses. These stresses included plant apoplastic fluid, *in planta* growth, bacterial conjugation, temperature changes, ultraviolet B irradiation (302nm) and DNA damaging agents such as mitomycin C. These different conditions may act upon ILE integrase and TTE gene expression differently. In order to deduce if the mechanisms of gene expression was consistent across different ILEs, two ILEs were studied. One ILE was present in *Ppi* 203 which is a pea pathogen and causes disease on the pea cultivar Kelvedon Wonder (KW) and the other was present in *Psy* 3023 which is a pathogen of bean. Both ILEs belong to strains of the *Pseudomonas syringae* species and both had similar genetic content with the presence of *xerC* integrase genes and TTE genes.

The first condition tested was bacterial gene expression in multiple plant apoplastic fluids. The apoplastic fluids included three *Phaseolus vulgaris* cultivars, Tendergreen (TG), Canadian Wonder (CW) and Red Mexican (RM) and apoplastic fluid from *Pisum sativum* cv. Kelvedon Wonder (KW). Before testing could begin using the prepared apoplastic fluid it had to be confirmed that no cell lysate was present in the preparations and that all of the preparations were

approximately the same concentration. If this was not the case then differences in gene expression could be due to higher concentrations of apoplastic fluid relative to each other, or that cell lysate was present contaminating the apoplastic fluid preparation.

To check for cell lysate contamination a malate dehydrogenase assay was performed as per O'Leary *et al.* (2016). This assay was performed as malate dehydrogenase is only present within the cytoplasm and mitochondria and therefore indicates cellular disruption (Musrati *et al.,* 1998). All of the apoplastic fluid preparations contained trace levels of malate dehydrogenase activity. However, these levels did not significantly differ between inoculation treatments. Apoplastic fluid concentration was also determined using an indigo carmine assay which readily binds strongly to –NH groups of proteins (Roberts *et al.,* 1998) to give a relative protein concentration that can be measured at 610nm (O'Leary *et al.,* 2014). The concentration assay was also performed with the apoplastic fluid preparations, but without any indigo carmine added to allow for the absorption of light by the apoplastic fluid, this is also the reason why the infiltrate media (sterile water) is tested. All of the dilutions were then standardised to the same starting concentration.

Interestingly an increase in TTE gene expression only occurred in *Ppi* 203 when grown in *Pisum sativum* KW apoplastic fluid which resulted in *avrPpiA1* being upregulated 12.7 ± 2.3 fold. However both TTE genes, *hopH1* and *hopAP1* showed some increase in expression when *Psy* 3023 was grown in CW and RM apoplastic fluid. This increase in TTE gene expression was expected with *Ppi* 203 growing in its host plants' apoplastic fluid as Boch *et al.* (2002) demonstrated that TTE genes are upregulated during plant colonisation. The *in vitro* tests using

apoplastic fluid mimics the plant colonisation process resulting in the upregulation of *avrPpiA1* (Rico *et al.,* 2009).

Following on from the apoplastic fluid tests both *Ppi* 203 and *Psy* 3023 were inoculated into TG, CW, RM bean and KW pea plants for *in planta* testing of ILE gene expression. This would allow any similarities or differences in ILE gene expression between *in vitro* apoplastic fluid and *in planta* tests to be observed. The *in planta* tests showed similar patterns in gene expression as the apoplastic fluid tests. The *avrPpiA1* gene in *Ppi* 203 was upregulated in KW pea along with both TTEs, *hopH1* and *hopAP1* being upregulated by CW as seen with the CW apoplastic fluid. These results were expected as *Ppi* 203 causes disease on KW pea and the HR on bean plants. The opposite is true for *Psy* 3023 as *Psy* 3023 causes' disease on bean with the HR being observed on KW pea plants. TTE genes would be upregulated during disease progression in response to host colonisation and suppression of host immune responses. Similarities were expected between *in planta* and apoplastic fluid preparations due to the bacteria residing within the extracellular space of the plant. Previous studies have also shown that both TTE genes (Rico and Preston, 2008) and integrase genes (Yu *et al.,* 2013) are upregulated in apoplastic fluid.

Although apoplastic fluid may mimic what happens to the bacterium inside the plant there are some processes that can only happen *in planta* that may affect gene expression including the HR which cannot occur in apoplastic fluid. One example of this is bacterium-plant cell contact that alters bacterial gene expression. It has been shown that *hrp* genes and subsequently the TTSS in *Ralstonia solanacearum* are specifically induced when the bacterium comes into contact with the surface of the plant cell (Aldon *et al.,* 2000). Although this work was not on *P. syringae* the same may be true and TTE genes may be induced by

plant cell contact as well as apoplastic fluid. The levels of gene expression *in planta* were lower than the apoplastic fluid tests. This could be because other factors inside the plant may reduce and alter gene expression such as plant cell contact (Aldon *et al,.* 2000). Another point to consider with further research would be having a longer incubation time up to 48 hours in apoplastic fluid rather than the six hours used. Studies have shown that within the first few hours of inoculation rapid changes occur to the bacterium including transcriptional, physiological and metabolic changes (O'Leary *et al.,* 2016). Allowing a longer incubation time would allow gene expression to settle and allow further expression studies into the complex relationship between bacterium and plant. However the gene expression 'window' may also be missed.

Bacterial conjugation was also tested as bacterial stress to observe any differences in ILE gene expression. Conjugation was chosen as previous ILEs in *P. fluorescens* have been shown to be mobile during conjugation events (Rhodes *et al.,* 2014) which may cause an increase in integrase gene expression. An increase in integrase gene expression was only observed in *Psy* 3023 and not *Ppi* 203. This may be due to the ILE in *Psy* 3023 being mobile whereas the *Ppi* 203 ILE may have become fixed in the chromosome (Chapter 6). There is also a link between ILEs and conjugative elements such as integrative conjugative elements (ICEs). ILEs share some similar features to ICEs. Both have been identified in disrupted *rulB* homologues (Maayer *et al.,* 2015). An interesting result from the conjugation tests was that both TTE genes, *hopH1* and *hopAP1* from *Psy* 3023, were upregulated beyond 20 times the control although *avrPpiA1* in *Ppi* 203 was not upregulated to the same scale as *hopH1* and *hopAP1.* Upregulation of virulence genes in bacteria following conjugation has been previously shown in *P. aeruginosa* (Grohmann, 2013).

UVB stress was used to study ILE gene expression as it is known to cause DNA breaks resulting in ssDNA molecules that activate the SOS response (Sundin and Weigand, 2007) which was hypothesised to regulate ILE integrase expression, in a similar manner to integrons, and the *rulAB* operon due to a LexA repressor binding site being present in the *rulA* promoter (Ramos, 2004). However the results showed that this was not true as *rulB'* and *xerC* have different expression patterns when subjected to UVB irradiation. This suggests that *xerC* is independently expressed from *rulB'* and may have its own regulatory domain. The *rulB* gene product, DNA polymerase V is under the regulation of the SOS response and is actively transcribed following chemical or radiation derived DNA damage (Krishna *et al.,* 2007; Bridges, 2005), (see Chapter 7 for more on the SOS response). Several genes that are regulated by the LexA repressor have been shown to be upregulated in response to UV irradiation in *E. coli* (Courcelle *et al.,* 2001). The *E. coli* study also reported that the *rulAB* homologue, *umuDC,* was upregulated following UV irradiation. However during the current research expression of *rulB'* was very low during UVB exposure. This could be due to methodical errors and needs to be repeated. There may have been some differences in UVB exposure due to the different glass petri dishes used. Glass absorbs UV irradiation and could therefore influence the results. This was controlled by using a new batch of petri dishes for this experiment but maintaining the same dishes throughout.

As the SOS response is activated by DNA damage the DNA damaging agent, MMC, was also tested along with UVB irradiation on both *Ppi* 203 and *Psy* 3023. MMC causes DNA damage via a different mechanism to UVB irradiation. Whereas UVB causes ssDNA breaks MMC causes DNA crosslinks between complementary strands and can be lethal to cells as only one crosslink per

genome can cause bacterial cell death (Tomasz, 1995; Szybalski and Lyer, 1964).

The results presented here showed that both DNA damaging events caused an increase in ILE *xerC* expression which increased as the dose of either UVB or MMC increased when *Ppi* 203 was tested. For example the highest concentration of MMC, 1 µg/ mL caused the expression of *xerC* to increase to 39.3 ± 0.1 fold, whereas the lower concentration of 0.5 µg/ mL caused a 9.2 ± 0.0 fold increase in expression. The same pattern was also observed when the cells were exposed to increasing UVB irradiation. As longer exposure times were used higher levels of *xerC* expression were reached. When the cells were exposed to 45 seconds of UVB irradiation the expression of *xerC* increased 8.0 ± 1.6 fold compared to an increase of 17.5 ± 2.1 fold with 60 seconds of UVB exposure. The TTE gene, *avrPpiA1,* also followed a similar pattern as its expression increased as the dosage of MMC and exposure time of UVB increased. This was hypothesised to be the reason why the ILE inserts into *rulB,* to share its promoter and regulation mechanisms, although there appears to be no relationship between *rulB'* expression and the expression of the other ILE genes so it is unlikely. The expression of the effector gene must be driven by another promoter. Although the SOS response may upregulate *rulAB* in its fully intact form the same appears not to be true for the truncated *rulB'* fragment as both UVB irradiation and MMC caused the *rulB'* fragment to be downregulated below both of the respective control tests. Again this needs to be repeated for certainty.

The ILE in *Psy* 3023 did not show the same dose-dependent response to either UVB irradiation or MMC as *Ppi* 203 did. The ILE *xerC* gene in *Psy* 3023 was upregulated to 4.5 ± 2.0 times the control with the lowest UVB exposure time of 15 seconds. The higher exposure times caused the expression of *xerC* to fall

until it reached 1.1 ± 0.3 times the control which was not significantly different to the control. A similar pattern was also seen with MMC on the expression of *xerC* as the lowest concentration provided the highest expression which decreased as the MMC concentration increased. These results may be due to *Psy* 3023 being more sensitive to UVB irradiation at lower dose levels than *Ppi* 203 (Chapter 7). The TTE gene *hopH1* on the ILE of *Psy* 3023 was also upregulated with the lower levels of UVB and MMC, but the highest levels of both UVB exposure and MMC caused *hopH1* expression to drop. *Psy* 3023 was more susceptible to MMC than *Ppi* 203 and may therefore be more susceptible to DNA damage overall. Bacterial growth studies of both *Ppi* 203 and *Psy* 3023 in MM containing MMC (see Section 5.2.2d) showed that *Psy* 3023 was more susceptible to MMC as at every MMC concentration; the colony counts and CFU/ mL for *Psy* 3023 were consistently lower than *Ppi* 203. The increased susceptibility of *Psy* 3023 could be due to a number of reasons. One reason may be that *Ppi* 203 has more DNA damage repair genes present and faster turnover of DNA repair compared to *Psy* 3023 (Zgur-Bertok, 2013). The susceptibility of *Psy* 3023 would cause the higher concentrations of MMC such as 1 µg/ mL to kill the majority of the cells, although not all of them. This may have caused the reduction in gene expression due to fewer cells being present to begin with compared to the *Ppi* 203 result which showed the dose dependent response to MMC.

The final bacterial stress tested to observe differences in ILE gene expression was suboptimal temperatures. The range of temperatures used included temperatures below the bacterium's normal laboratory growing temperature of 25^{o}C and also temperatures above the norm. When *Ppi* 203 was subjected to the suboptimal temperatures many of the genes were downregulated below the control for the very extreme temperatures of -80^{o}C,

-20ºC and 37ºC. However 4ºC provided an increase in gene expression for all the genes except *rulB'*. This was also the case with *Psy* 3023 with some exceptions such as *rulB'* was upregulated at 4ºC and also -80ºC while *hopH1* was upregulated at -20ºC. These differences in gene expression at different temperatures could be due to a number of molecular and physiological changes. On a molecular level temperature-mediated regulation occurs at transcription and translation events. This includes DNA supercoiling, changes in mRNA conformation and changes in protein conformation (Hurme and Rhen, 1998). As temperature increases, proteins begin to lose conformation and function resulting in denaturation, this includes many regulatory proteins including LexA which can no longer bind to its regulatory domain (Helene, 2012; Daniel *et al.,* 1996). DNA supercoiling is also affected by temperature changes and is thought to be a central regulator in many virulence genes (Dorman, 1991). This is due to temperature altering DNA topology and histone proteins causing the DNA to remain supercoiled resulting in additional DNA methylation and the repression of transcription and translation as the DNA is unavailable for the machinery to gain access (Hurme and Rhen, 1998). DNA supercoiling may also be influenced by the denaturation of topoisomerase and / or DNA gyrase. Topoisomerase is important for the unwinding of DNA from its supercoil form to allow transcription and translation (Champoux, 2001) and DNA gyrase can reverse supercoils from positive to negative supercoils (Wigley *et al.,* 1991). Temperature can also affect physiological parameters such as cell structure, membrane fluidity and membrane integrity. These often cause serious cell damage that are often lethal to the cells (Trevors *et al.,* 2012).

Overall it is clear that multiple different stresses have an impact on ILE gene expression from both *Ppi* 203 and *Psy* 3023. Apoplastic fluid and *in planta*

stress causes the TTE genes to be upregulated when in the bacterium's host plant and also the integrase genes are upregulated in some plants when it may be necessary for the ILE to either excise a TTE or gain a new TTE to avoid plant detection. There was also a clear link between DNA damage and upregulation in ILE gene expression in *Ppi* 203 when using UVB and MMC. The expression study also suggested that the ILE *xerC* gene does not share its regulatory domain with the *rulB* gene as the expression profiles between *rulB'* and *xerC* were substantially different. The *rulB'* fragment showed little differences in gene expression compared to the ILE genes. This may be due to it being disrupted by the ILE and further gene expression tests on an intact *rulB* and *rulA* would be useful to ascertain if there are any similarities in gene expression between *rulAB* and the ILE inserted into it and whether they share the same promoters and regulatory domains. An important point to also consider is that not all of the cells will contain an ILE insertion. This could influence the expression of ILE genes if an unknown factor was influencing the rate of ILE insertion and excision. A final point to mention is the possible unspecific amplification of other *xerC* genes within the genome that are similar to the ILE *xerC* gene and may cause differences in measured gene expression across the two different strains tested. However this was limited by choosing primers specific to the ILE *xerC* gene.

Chapter 6. Movement of integron-like elements between different *rulAB* systems.

6.1: Introduction

A key aspect of integrons is their ability to move between different genetic loci within a species and also between different bacterial species via horizontal gene transfer. This allows advantageous genes present on the integron to be utilised by other bacterial species promoting enhanced virulence and antibiotic resistance in some cases (Gillings, 2014). However integrons are not known to be independently mobile and require other genes to facilitate movement such as transposase and resolvase, encoding genes which facilitate integron movement by transposition (Partridge *et al.,* 2002).

As integron-like elements (ILEs) are similar to integrons it has been hypothesised that they may also be mobile within species and also between species of Pseudomonads. ILEs have previously been shown to be mobile in *P. fluorescens* by Rhodes *et al.* (2014). ILEs were identified in six environmental *P. fluorescens* strains and were denoted FH1 through FH6. All of the ILEs were chromosomally located into the *rulB* gene and harboured heavy metal resistance genes as they were isolated from a disused copper mine. Insertion into *rulB* was interesting as it encodes a type V DNA polymerase used for transcription following DNA damage which stalls the normal type III DNA polymerase replication fork (Hawver *et al*., 2015: Ohmori *et al*., 2001). ILEs may insert into *rulB* to disrupt DNA repair allowing self-preservation of the ILE. They may also insert into *rulB* due to its close proximity to the origin of replication on many plasmids.

Following isolation and characterisation the strains were transformed using the IncP-9 toluene-degrading plasmid pWW0 to attempt to cure native

plasmids from the six strains (Rhodes *et al.,* 2014; Pickup, 1989). Following the addition of pWW0 to the *P. fluorescens* strains FH1-FH6, an insertion event had occurred into the *rulB* gene on pWW0 which was later identified as the chromosomal ILE. This was the first occurrence of an ILE moving between two different genetic loci, although the mechanisms and genes required for ILE movement were not known. The same movement into *rulB* on pWW0 was observed for all six ILEs from *P. fluorescens.*

An ILE was identified by Arnold *et al.* (1999; 2000) in *P. syringae* pv. *pisi* 203. This ILE was also inserted into a *rulB* gene present in the chromosome as seen in *P. fluorescens.* The ILE in *Ppi* 203 was different to the *P. fluorescens* ILEs and contained TTE genes used for plant colonisation and disease progression. The ILE in *Ppi* 203 was never shown to be mobile.

The identification of new ILEs in different *P. syringae* pathovars (Chapters 3 and 4) raised the question of whether the *P. syringae* ILEs are mobile in a similar way to the ILEs in *P. fluorescens.* If this was true, virulence genes present on the ILEs may be disseminated across a range of *P. syringae* pathovars that infect many economically important crops and could be devastating to food and trade supplies. *P. syringae* pathovars have many host plants including; soya beans, peas, tobacco, tomatoes, kiwi fruit and *Prunus* sp. (Hirano and Upper. 2000). If the ILEs present in *P. syringae* were mobile it would be the first documented case of this occurring.

To test if the newly identified ILEs in *P. syringae* were mobile the first step was to transform the cells with pWW0 to see if any ILEs move into the *rulB* gene on pWW0, this would be the same mechanism as used in the *P. fluorescens* strains. *P. putida* PaW340 was used as the donor strain containing the native pWW0, *P. putida* PaW340 (pWW0::km^r) and the *P. syringae* strains were the

recipients following bacterial conjugation between the two strains. Electroporation of the pWW0 into the *P. syringae* strains was also tested, although it was hypothesised that bacterial conjugation was the trigger for ILE movement and this method would yield no ILE movement onto pWW0. It may also be bacterial stress that is responsible for ILE movement to promote bacterial survival in times of environmental stress (Cambray *et al.,* 2011). This method was tested on two ILE containing strains; *Ppi* 203 and *Psy* 3023.

The next stage of the experiment was to clone different *rulAB* operons into a broad host vector that would easily enter the recipient cells and facilitate further investigation into *rulAB* ILE integration. This was performed by PCR amplification of the desired *rulAB* operon which was then directly ligated into pCR2.1. The cloned gene was digested from pCR2.1 using restriction enzymes and re-ligated into the broad host range plasmid pBBR1MCS-2.

Having tested the pWW0 plasmid containing the *rulB* ILE insertion point on both the *P. fluorescens* and *P. syringae* borne ILEs different versions of *rulB* were tested. The *rulAB* operon from pWW0 was used along with the *rulAB* operon from the genomic island PPHGI-1, which is present in *Pph* 1302A. This would allow any differences between *rulB* genes from different species to be observed. The two *rulAB* operons were cloned into a broad host range vector, pBBR1MCS-2 (Kovach *et al.,* 1995) and these were both ligated and electroporated into *P. fluorescens* FH1, *Ppi* 203 and *Psy* 3023.

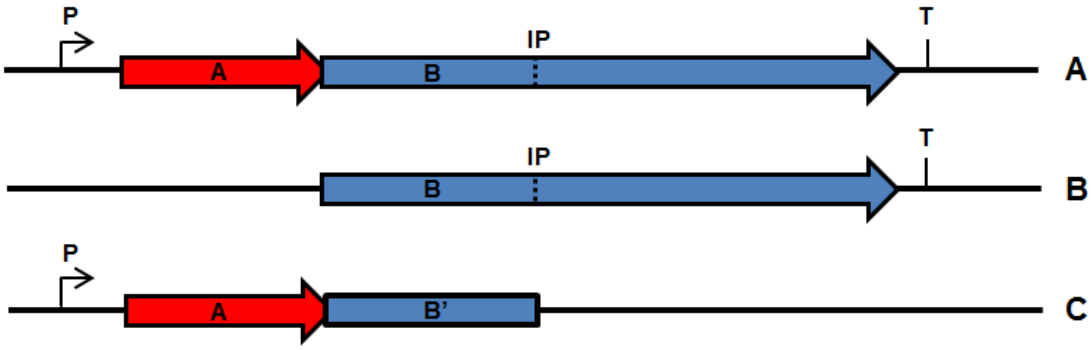


Figure 6.1: Versions of *rulAB* from pWW0 cloned into pBBR1MCS-2 to test ILE movement. Three versions of *rulAB* from pWW0 to assess the requirements of ILE movement from the chromosome of *P. fluorescens* FH1 to a plasmid borne *rulAB*. **A)** Intact *rulAB* operon with both promoter (P) and terminator (T) present. **B)** Just the *rulB* gene present without a promoter but terminator (T) is present. **C)** Full *rulA* gene with *rulB'* which is cut off at the point of ILE insertion, promoter (P) is present. IP refers to previously identified ILE insertion point.

Following on from the *P. syringae* ILE experiments the focus of ILE movement tests shifted towards the ILE present in *P. fluorescens* FH1 and what was required for the ILE to move. Three different versions of *rulAB* from pWW0 were created and cloned into pBBR1MCS-2 for testing in FH1 (Figure 6.1). The cloned constructs were also electroporated into FH1 to assess if conjugation was the trigger for ILE movement. The three *rulAB* versions consisted of the intact *rulAB* with promoter and terminator domains still present, to identify if any of the other genes present on pWW0 were required for ILE movement as they were no longer present in the construct. The next version was just *rulB*. If ILE insertion occurred then it would indicate that only a sequence target is needed as no promoter was present and the genes from pWW0 were still missing. The final *rulAB* version was *rulAB'-IP* which consisted of the entire *rulA* gene and part of *rulB* up to the previously identified ILE insertion point. Again the *rulA* promoter was present but not the *rulB* terminator. This construct was expected to show no ILE insertion as *rulB* is not active and the sequence site is missing. These *rulAB*

constructs derived from pWW0 would answer the question; is the intact functioning form of *rulAB* needed for ILE insertion in the case of *P. fluorescens*? A further question to consider was do ILEs circularise following excision from *rulB* on a chromosome or a plasmid? Mobile genetic elements such as genomic islands have been shown to circularise following excision. One example of this is PPHGI-1 (Godfrey *et al.,* 2011) which forms a circular intermediate following excision from the genome. This may also happen to mobile ILEs during ILE movement between *rulB* genes. To test this primers were designed (Section 2.17) that would only amplify a DNA product if the ILE had formed a circular intermediate.

6.2: Results

6.2.1: ILE movement using pWW0 as a basis for *rulB*.

The first step was to replicate the work of Rhodes *et al.* (2014) to observe the movement of the ILE from the chromosome of *P. fluorescens* into the *rulB* gene on pWW0 situated in *P. putida* PaW340. This was achieved by agar plate mating allowing conjugation to occur between the strains and the plasmid moving into FH1 whilst the ILE moved from the chromosome of FH1 onto the plasmid, pWW0. Cells without the plasmid were selected against using kanamycin as pWW0 confers kanamycin resistance and ILE movement was assessed using PCR with the GRrulAB-xerC primers (see Section 2.4.2). This showed any ILE insertion into the pWW0 *rulB* only as there was enough difference between the chromosomal *rulAB* and the pWW0 *rulAB* to allow specific primers to be developed (Figure 6.2).

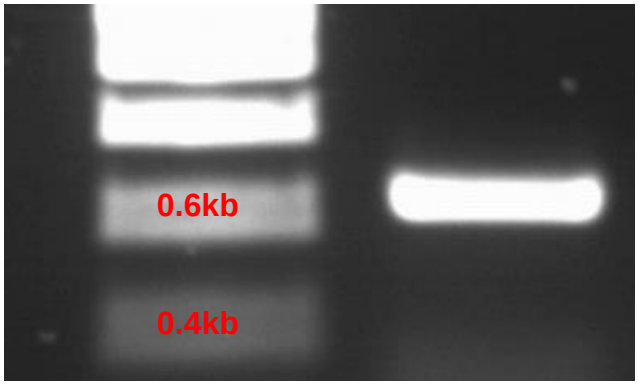


Figure 6.2: PCR of *rulAB-xerC* confirming ILE movement from FH1 chromosome into *rulB* on pWW0. The PCR confirms ILE movement onto pWW0 following conjugation between *P. fluorescens* FH1 and *P. putida* PaW340 (pWW0) using pWW0 *rulAB* derived primers. PCR product was expected to be 590bp which was shown. Hyperladder 1kb (Bioline, UK) was used.

The same experiment was tried with both *Ppi* 203 and *Psy 3023* using kanamycin as the selective agent to eliminate any cells without pWW0 and PCR was performed using the same primers. *XerC* is similar across all ILEs to allow primer binding. However, simple agar plate mating and conjugation failed to get

pWW0 inside both *P. syringae* strains. Electroporation of pWW0 into competent *Ppi* 203 and *Psy* 3023 cells was also tried using extracted and purified pWW0 (Figure 6.3). However conjugation and electroporation of pWW0 into *Ppi* 203 and *Psy* 3023 failed on multiple occasions.

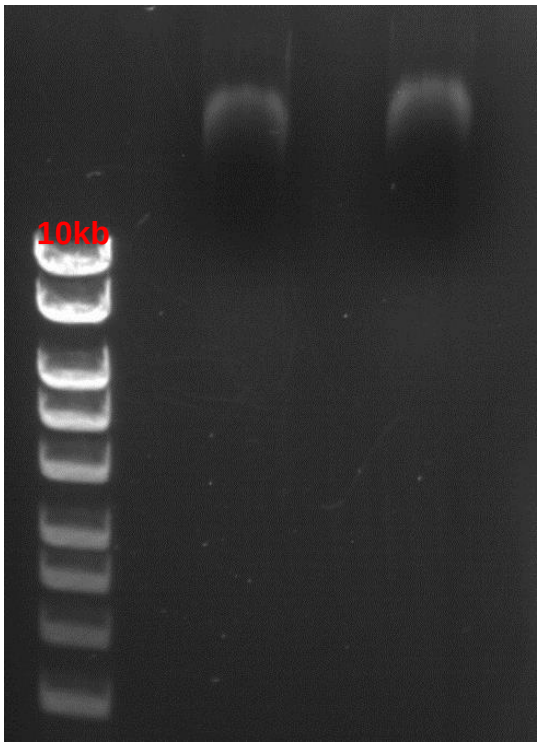


Figure 6.3: Midi plasmid preparation of pWW0 from *P. putida* PaW340. Following extraction and purification 10uL of the prepared pWW0 plasmid was run on an agarose gel along with Hyperladder 1kb. The smears at the top of gel show the plasmid with a high molecular weight, 117kb which is the prepared pWW0. Hyperladder 1kb (Bioline, UK) was used.

6.2.2: Cloning two different *rulAB* operons into pBBR1MCS-2.

The full version of pWW0 did not enter *Ppi* 203 or *Psy* 3023 to assess if the ILEs would move into the *rulB* gene as seen with the FH1 ILE. To overcome this the *rulAB* operon from pWW0 and also the *rulAB* operon from PPHGI-1 in *Pph* 1302A were cloned into the broad host range vector, pBBR1MCS-2.

6.2.2a: Empty cloning vectors electroporated into recipient strains.

Before the pBBR1MCS-2 plasmid containing the cloned *rulAB* operons were electroporated into the recipient strains empty vector controls were tested to ensure that *P. fluorescens* FH1, *Ppi* 203 and *Psy* 3023 could take up the plasmid and produce viable kanamycin resistant colonies (Figure 6.4).

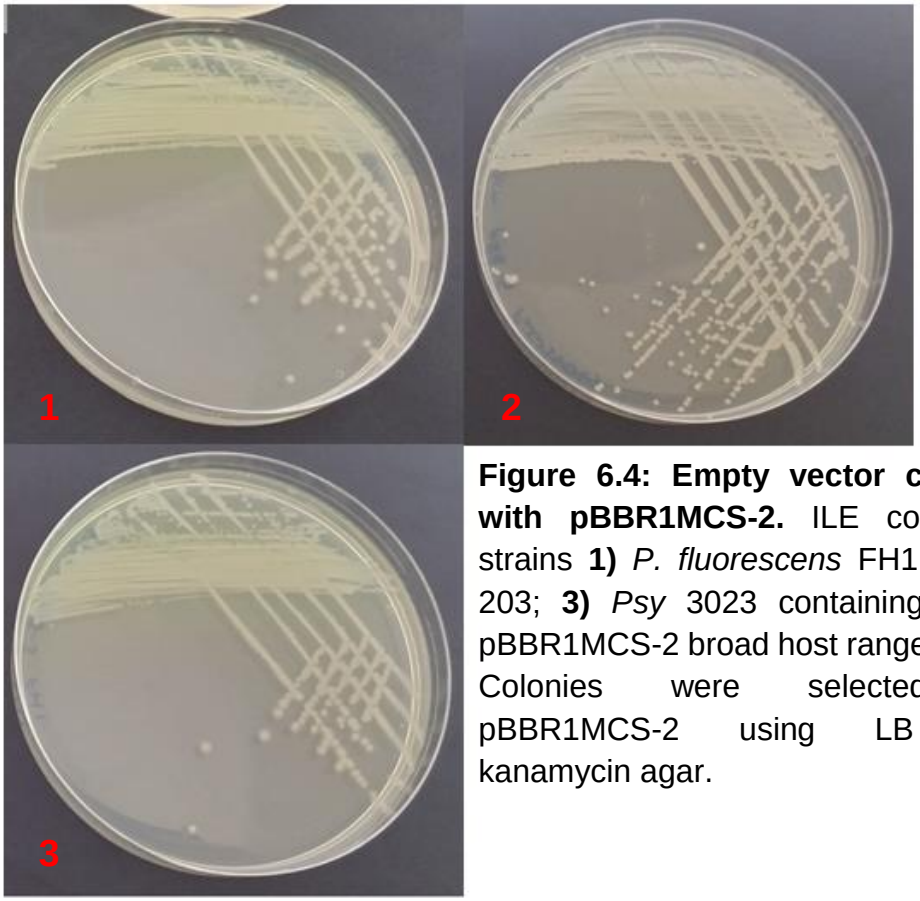


Figure 6.4: Empty vector controls with pBBR1MCS-2. ILE containing strains **1)** *P. fluorescens* FH1; **2)** *Ppi* 203; **3)** *Psy* 3023 containing empty pBBR1MCS-2 broad host range vector. Colonies were selected for pBBR1MCS-2 using LB plus kanamycin agar.

6.2.2b: Cloning of *rulAB* operons.

Following the successful electroporation of the empty vector into the three ILE containing strains the two *rulAB* operons from pWW0 and PPHGI-1 were first amplified via PCR before being cloned into pBBR1MCS-2 via pCR2.1 (Figure 6.5). The pBBR1MCS-2 plasmid was also restriction digested with the appropriate enzymes, *Spe1* and *Xba1* for pWW0 and *EcoR1* for PPHGI-1.

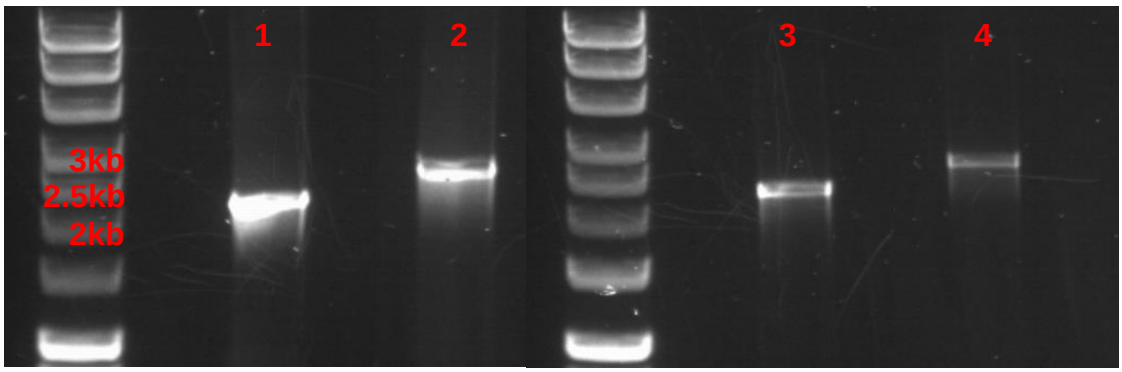


Figure 6.5: Cloning of pWW0 *rulAB* and PPHGI-1 *rulAB.* 1) pWW0 *rulAB* amplified (2200bp) along with **2)** PPHGI-1 *rulAB* (~2600bp). **(3-4)** shows the same fragments following ligation into pCR2.1 and subsequent restriction digestion to make the fragments ready for ligation into pBBR1MCS-2. Hyperladder 1kb (Bioline, UK) was used.

6.2.2c: Sub-cloning of *rulAB* operons into *P. fluorescens* FH1, *Ppi* 203 and *Psy* 3023.

The two *rulAB* operons were successfully ligated into pBBR1MCS-2 and then electroporated into *P. fluorescens* FH1, *Ppi* 203 and *Psy* 3023. Successful cloning and ligation into *P. fluorescens* FH1, *Ppi* 203 and *Psy* 3023 was determined by blue-white selection and resistance to kanamycin. Presence of both *rulAB* operons in all three strains was confirmed via PCR (Figure 6.6).

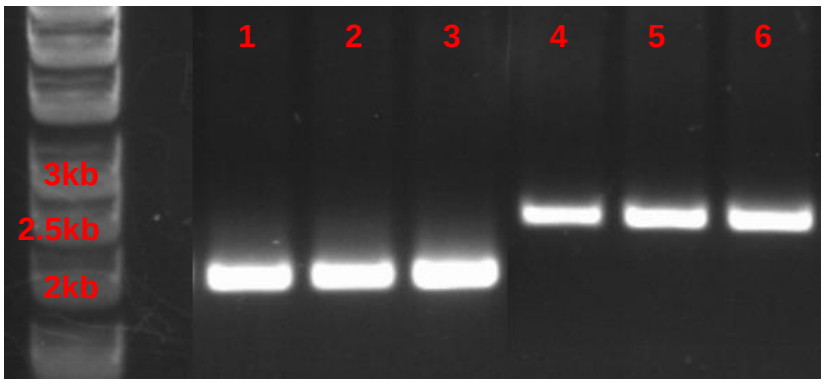


Figure 6.6: PCR amplification of extracted pWW0 and PPHGI-1 *rulAB* operons from *P. fluorescens* FH1, *Ppi* 203 and *Psy* 3023. Lanes **1-3** show that the pWW0 *rulAB* operon was present in *P. fluorescens* FH1, *Ppi* 203 and *Psy* 3023 in that order at a size of 2200bp and lanes **4-6** show that the PPHGI-1 *rulAB* operon was present in *P. fluorescens* FH1, *Ppi* 203 and *Psy* 3023 in that order at a size of ~2500bp. Hyperladder 1kb (Bioline, UK) was used.

6.2.2d: Identification of ILE movement into cloned constructs.

The next step was to assess if any of the ILEs from the three strains had moved into the cloned versions of *rulAB.* This was performed via PCR using the GRrulAB-xerC primers (Section 2.4.2) for the pWW0 *rulAB* and the 2015rulAB-xerC primers for the PPHGI-1 *rulAB* as it shares homology to the *Ppi* 203 *rulAB.* ILE movement was only observed with the FH1 ILE moving into the cloned pWW0 *rulAB* (Figure 6.7). Both *Ppi* 203 and *Psy* 3023 failed to move into the pWW0 *rulAB* and no ILEs were observed moving into the *rulAB* operon from PPHGI-1. This may be because it is a non-functional version as the fully function version of *rulAB* from PPHGI-1 proved lethal to the tested cells and could not be cloned. The fainter bands on the gel were later revealed by DNA sequencing to be non-specific binding of the primers.

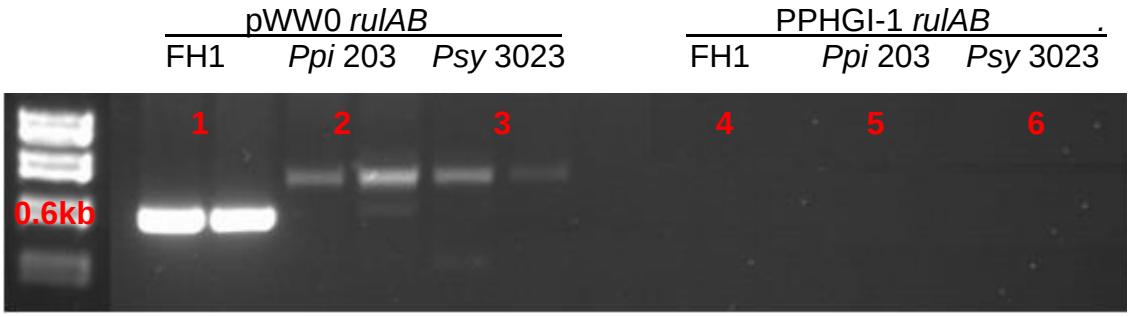


Figure 6.7: ILE movement from *P. fluorescens* FH1, *Ppi* 203 and *Psy* 3023 into cloned versions of *rulAB* from pWW0 and PPHGI-1. Lane **1** shows that the ILE from FH1 moved into the cloned version of pWW0 as the *rulB* – ILE *xerC* junction region amplified at 590bp. Lanes **2** and **3** were both negative with further analysis meaning the ILEs from *Ppi* 203 and *Psy* 3023 did not move into pWW0 *rulAB.* Lanes **4-6** were also negative showing that no ILEs moved into the *rulAB* operon derived from PPHGI-1. Hyperladder 1kb (Bioline, UK) was used.

6.2.3: Cloning different versions of the pWW0 *rulAB* operon looking for a specific ILE insertion target.

As the results from the previous experiments showed that only the ILE from *P. fluorescens* FH1 appeared to move into the cloned *rulAB* operon from

pWW0 it was decided to focus more on this system to identify what part of *rulAB* was needed for ILE movement and whether the ILE insertion targeted a specific sequence or if a functional *rulAB* operon was also a requirement. Three versions of pWW0 *rulAB* were constructed and ligated into pBBR1MCS-2 (Figure 6.1).

6.2.3a: Cloning of three different pWW0 *rulAB* versions.

The first step was to amplify the three *rulAB* versions from pWW0 via PCR (see Section 2.16.1) and confirm the sizes via gel electrophoresis (Figure 6.8). All of the amplified fragments were the correct expected size with the intact *rulAB* product at 2.2kb, just *rulB* at 1.5kb and *rulAB'-IP* at 0.8kb.

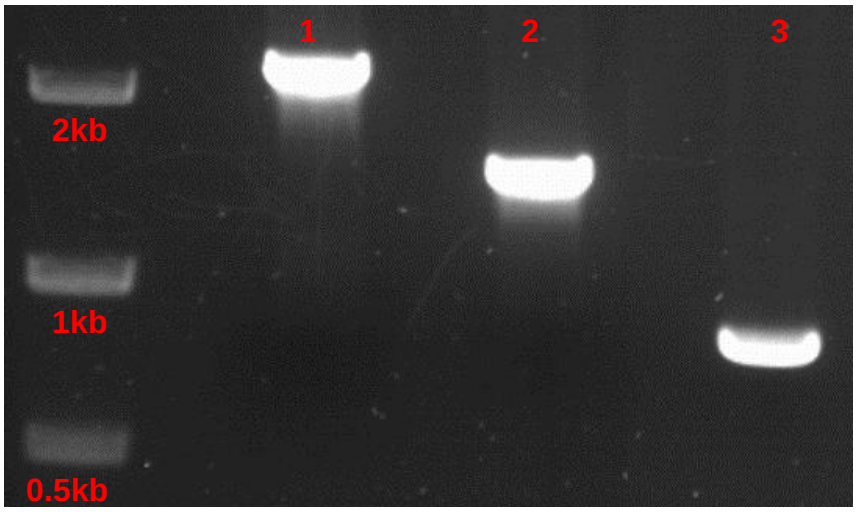


Figure 6.8: PCR amplification of three different pWW0 *rulAB* versions. The three versions of pWW0 *rulAB* were; **1)** fully intact *rulAB* with promoter and terminator, expected size was 2.2kb, **2)** *rulB* without *rulA*, only terminator present, expected size was 1.5kb, **3)** *rulAB'-IP,* intact *rulA* along with part of *rulB* up to the previously identified ILE insertion point, expected size was 0.8kb. Easyladder 1 (Bioline, UK) was used.

The amplified regions were then directly ligated into pCR2.1 following clean-up and chemically competent *E.coli* cells were transformed with the constructs. Following cellular growth and extraction of the pCR2.1 constructs the regions of interest were cut out via restriction enzyme digest (Figure 6.9) to facilitate ligation into pBBR1MCS-2.

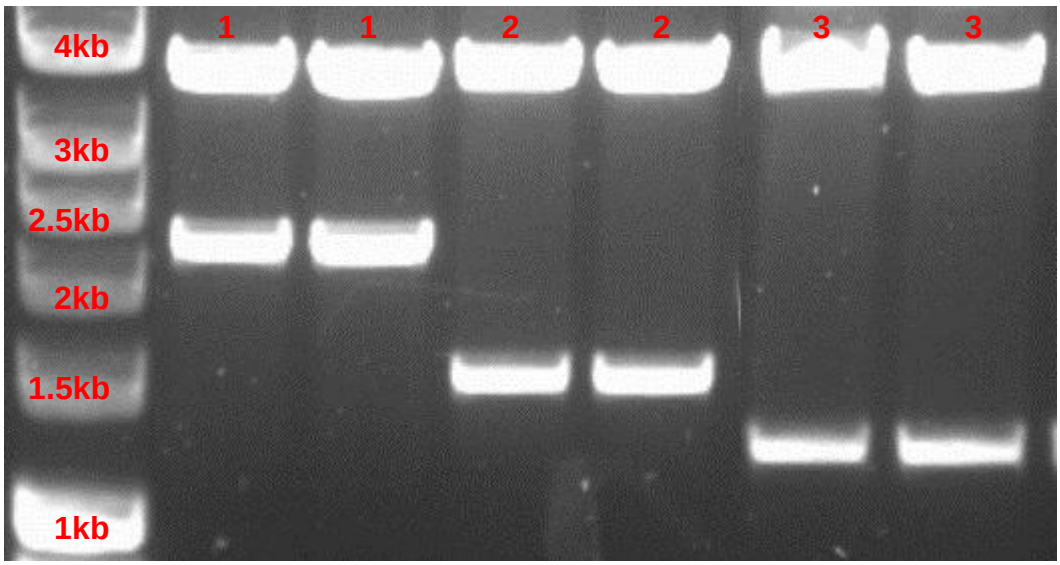


Figure 6.9: Restriction digest of pWW0 *rulAB, rulB* and *rulAB'-IP* regions from pCR2.1. The gel image shows the three *rulAB* regions in duplicate; **1)** fully intact *rulAB* with promoter and terminator, expected size was 2.2kb, **2)** *rulB* without *rulA*, only terminator present, expected size was 1.5kb, **3)** *rulAB'-IP,* intact *rulA* along with part of *rulB* up to the previously identified ILE insertion point, expected size was 0.8kb. The bands present at the top of the gel show the remaining pCR2.1 plasmid from the restriction digest, pCR2.1 without any insert is 3.9kb as shown on the gel. Hyperladder 1kb (Bioline, UK) was used.

6.2.3b Confirmation of successful ligation and transformation of pWW0 *rulAB* variants into pBBR1MCS-2 and subsequently into FH1.

The final part of the cloning procedure was to ligate the three *rulAB* versions into pBBR1MCS-2 and then electroporated the three constructs into *P. fluorescens* FH1 cells. Successfully ligated and transformed cells would be resistant to kanamycin and produce white colonies when grown in the presence of X-gal. These white resistant colonies were picked and the construct was extracted to allow PCR confirmation that the cloning had worked. The PCR used the original primers used to amplify the three regions (Section 2.16.1) and the three regions were successfully amplified (Figure 6.10) indicating that FH1 now contained the three different *rulAB* versions from pWW0. This allowed further testing on ILE movement into the constructs to begin.

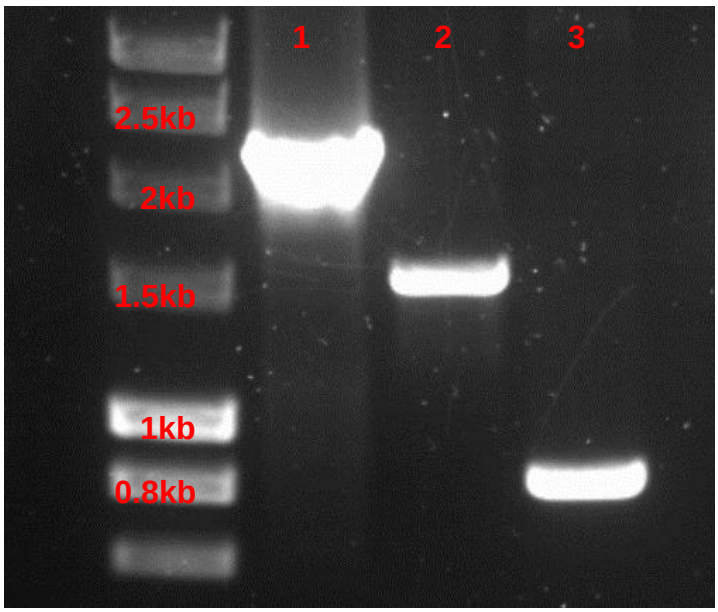


Figure 6.10: Amplified pWW0 *rulAB* regions from transformed *P. fluorescens* FH1 cells. The regions were successfully amplified following transformation and they were all the expected sizes. **1)** Intact *rulAB,* 2.2kb, **2)** just *rulB,* 1.5kb, **3)** *rulAB'-IP,* 0.8kb. Hyperladder 1kb (Bioline, UK) was used.

6.2.3c: Identification of any ILE movement from the chromosome of FH1 into the three constructs containing different pWW0 *rulAB* variants.

As previously shown in Section 6.2.2 the ILE on the chromosome of FH1 will move into an intact cloned version pWW0 *rulAB* on pBBR1MCS-2 following electroporation of pBBR1MCS-2 (pWW0*rulAB*::km^r). This was repeated along with the two new constructs; pBBR1MCS-2 (pWW0*rulB*::km^r) and pBBR1MCS-2 (pWW0*rulAB'-IP*::km^r) to assess what part of *rulAB* was required for ILE insertion into *rulB.* Following incubation of the FH1 cells containing the *rulAB* constructs, the constructs were extracted (see Section 2.3.2) and PCR (Figure 6.11) was performed using primers that would amplify a fragment from the start of *rulAB* into the start of the ILE, the 5' end, and primers that amplified a fragment from the 3' end of the ILE into the end of *rulB* (Section 2.16.6). This was repeated for all three variants of the pWW0 *rulAB.* The primers were still viable as the binding sites were still present.

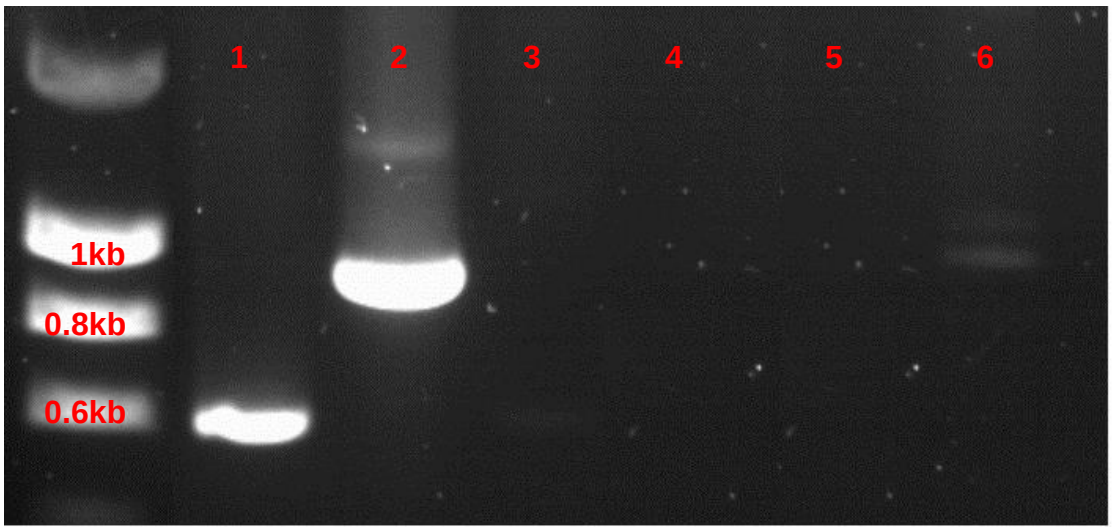


Figure 6.11: Identification of FH1 ILE movement into cloned versions of *rulAB, rulB* or *rulAB'-IP* present on pBBR1MCS-2. The amplified products were from the 5' end of the ILE out into *rulAB* and from the 3' end of the ILE out into *rulB.* This allowed confirmation that the entire ILE had moved into the construct. **1)** Shows that the 5' end of the ILE (*xerC*) is present in pBBR1MCS-2 (pWW0*rulAB*), expected size was 590bp. **2)** the 3' end of the ILE (*sulP*) present in pBBR1MCS-2 (pWW0*rulAB*). Bands **3-4** would show the same amplification but with a different 5' end primer if an ILE had inserted into pBBR1MCS-2 (pWW0*rulB*). However this was negative meaning no ILE movement. Bands **5-6** would have also show amplification if the ILE had moved into pBBR1MCS-2 (pWW0*rulAB'-IP*). However no ILE movement was seen. Hyperladder 1kb (Bioline, UK) was used.

As shown the ILE in FH1 moved into the intact form of *rulAB* from pWW0 as shown in previous tests. No ILE movement was seen into (pWW0*rulB*) or (pWW0*rulAB'-IP*) using the parameters of the experiment. Following the FH1 tests the constructs were also tested in *Ppi* 203 and *Psy* 3023 although no ILE movement was seen for any of the three constructs.

6.2.4 Comparison of the frequency of FH1 ILE movement into native *rulAB* on pWW0 and the cloned version on pBBR1MCS-2.

Once it was confirmed that the ILE present in *P. fluorescens* FH1 would move into a cloned version of *rulAB* from pWW0 it was interesting to identify if the ILE had a preference over the native pWW0 plasmid or the cloned version of *rulAB.* This may indicate other genes and factors on pWW0 that influence the

frequency of ILE movement. To test this multiple FH1 cells were transformed with either pWW0 or pBBR1MCS-2 (pWW0*rulAB*) and incubated for 24 hours. Following this the cells were lysed and PCR (Figure 6.12) was used to identify any ILE inserts across 282 biological replicates for FH1 (pBBR1MCS-2 (pWW0*rulAB*)) and the same number for FH1 (pWW0::km^r). From this the percentage of ILE movement was calculated. Each PCR contained a positive control with a known ILE insertion into *rulB* and a negative control which contained *rulAB* but no ILE.

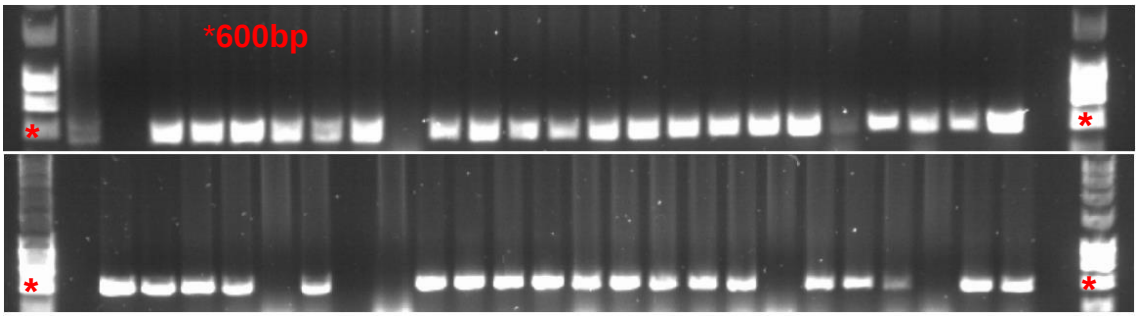


Figure 6.12: Examples of ILE movement frequency gels for both pBBR1MCS-2 (pWW0*rulAB*) and pWW0. The top gel image shows all of the positive amplifications for ILE movement into pBBR1MCS-2 (pWW0*rulAB*) for 24 out of the 282 replicates. The bottom gel image shows the same but for ILE insertion into the native pWW0 for 24 out of the 282 replicates. Expected size was 590bp. Hyperladder 1kb (Bioline, UK) was used.

As shown by the example gel images the frequency of ILE insertion is fairly high for both of the *rulAB* systems and following the analysis of all of the PCR amplifications there was only a small difference between the two systems as the pBBR1MCS-2 (pWW0*rulAB*) system showed a 59.2% ILE insertion frequency and the native pWW0 *rulAB* showed a 53.2% ILE insertion frequency. This was only a difference of 6% across the two systems and it appears the ILE has no preference over which *rulB* it inserts into. Both frequency values are not significantly different from each other at the 95% confidence interval. The complete analysis of ILE insertion frequency is in Table 6.1.

Table 6.1: Breakdown of ILE insertion frequency PCR test results into both pBBR1MCS-2 (pWW0*rulAB*) and pWW0. Each PCR consisted of 94 biological replicates plus two controls. Each PCR was also replicated three times resulting in 282 replicates. The frequency of ILE insertion is shown as a percentage at the bottom of the table.

PCR no.	Number of positive insertions in pBBR1MCS-2 (pWW0*rulAB)*	Number of positive insertion in pWW0	Total number of colonies tested
1	88	73	94
2	34	34	94
3	45	43	94
Total	167	150	282
Frequency of ILE movement	59.2% ± 30.3%	53.2% ± 21.7%	100%

6.2.5: Testing if ILEs circularise during movement to different *rulB* genes.

It was hypothesised that ILEs may form circular intermediates when moving from one *rulB* gene to another *rulB* gene. This was tested by PCR using primers (Table 2.6) that would only produce an amplification product if circularisation of the ILE had occurred. Six *P. syringae* strains containing an ILE were tested for the formation of ILE circular intermediates. These strains were *P. fluorescens* FH1 and FH4, *Ppi.* 203, *Psy.* 3023, *Psy.* B728a and *Pgy* 2411. The strains were subjected to stresses to increase the probability of ILE movement. These stresses were conjugation with *E. coli* DH5α and *E. coli* DH5α (pRK2013), cold stress, UV irradiation and growth in liquid minimal media. The first test was conjugation on three of the six strains, *P. fluorescens* FH1, *Ppi.* 203 and *Psy.* 3023 (Figure 6.13).

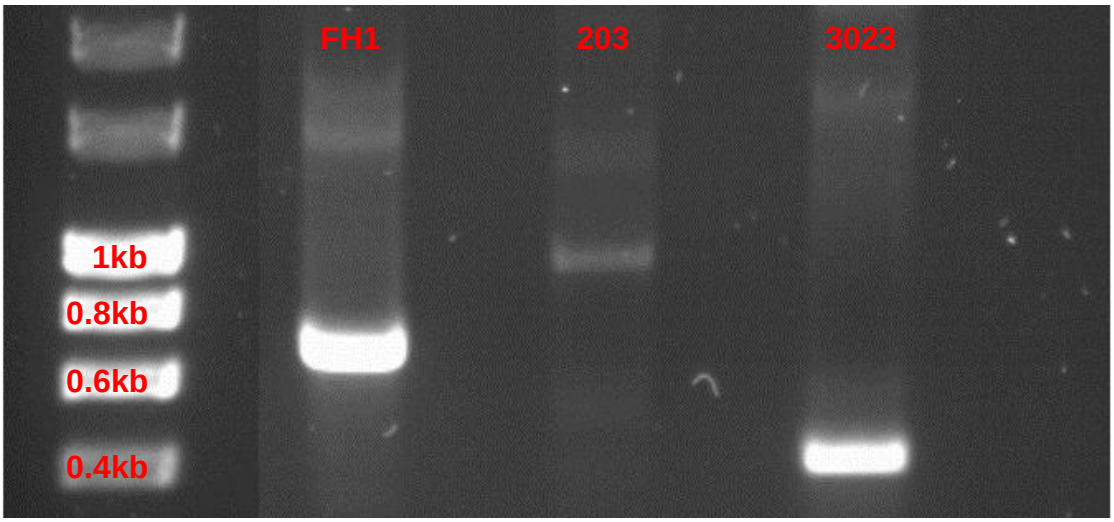


Figure 6.13: ILE circular intermediate PCR following conjugation of either *P. fluorescens* FH1, *Ppi.* 203 and *Psy.* 3023 with *E. coli* DH5α and *E. coli* DH5α (pRK2013). The results show a bright band for both FH1 and *Psy.* 3023, however *Ppi.* 203 had a much fainter band. The FH1 band is ~700bp, the faint *Ppi.* 203 band is ~1000bp and the *Psy.* 3023 band is smaller at ~400bp. Hyperladder 1kb (Bioline, UK) was used as a size marker.

The PCR products from the first circular intermediate tests (Figure 6.13) were sent for sequencing (Section 2.11). The sequencing returned with a gene not related to ILEs for all three PCR products. This indicated unspecific primer binding. The test was repeated with same outcome. The six ILE strains were then tested for circular intermediates following cold stress and UV irradiation stress (Section 2.17). This test provided faint PCR products (Figure 6.14) for *P. fluorescens* FH1, *Psy.* 3023 and *Psy.* B728a. These PCR products were also sequenced (Section 2.11) but also returned with a sequence not matching any of the ILEs.

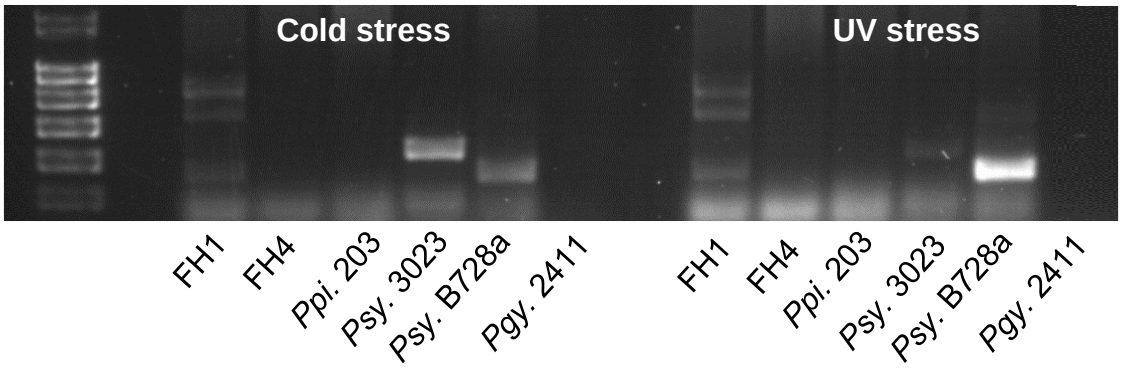


Figure 6.14: ILE circular intermediate PCR of *P. fluorescens* FH1, *P. fluorescens* FH4, *Ppi.* 203, *Psy.* 3023, *Psy.* B728a and *Pgy.* 2411 following cold stress and UV irradiation stress. The PCR products were faint for many of the strains that showed amplification. These strains were FH1, *Psy.* 3023 and *Psy.* B728a. There is a brighter band for *Psy.* B728a following UV stress. Multiple bands are present for some of the strains indicating unspecific primer binding. Hyperladder 1kb (Bioline, UK) was used as a size marker.

The results for the ILE circular intermediate tests require further work and hopefully future research will resolve some of these issues and potential identify ILE circular intermediates forming during ILE movement.

6.3 Discussion

ILE movement was only previously seen in the *P. fluorescens* FH1 strain when pWW0 was introduced for plasmid incompatibility testing (Rhodes *et al.,* 2014). During this experiment a piece of chromosomal DNA had moved into the *rulB* gene on pWW0. Following on from this ILEs were identified in *P. syringae* pathovars (Arnold *et al.,* 2000; Chapter 4) and it was hypothesised that these ILEs may also move into an exogenous *rulB* gene as the ILE also inserts into a chromosomal version of *rulB*. If the ILEs belonging to the *P. syringae* pathovars could move into plasmid borne *rulB* genes it may facilitate their dissemination to other bacteria and spread virulence genes present on the ILEs. This may lead to the development of new diseases on new host plants (Sarkar *et al.,* 2006). Further to identifying if the ILEs in *P. syringae* would move into *rulB,* cloning experiments were designed to identify what section of *rulAB* is required for successful ILE insertion. These experiments were carried out on FH1 but future work hopes to include the *P. syringae* ILEs with different *rulB* genes.

The first stage of the experiment was to replicate the ILE movement from the chromosome of FH1 into *rulB* on pWW0 and to also try the same method with the *P. syringae* strains containing an ILE. The FH1 ILE movement into pWW0 *rulB* was successfully carried out. However the two *P. syringae* ILE containing strains that were selected for testing, *Ppi* 203 and *Psy* 3023, could not be transformed with pWW0 using conjugation, heat shock or electroporation. This failure to transform *Ppi* 203 and *Psy 3023* with pWW0 may be due to three factors. The first may be that pWW0 is too large for the *P. syringae* strains. The pWW0 plasmid is 117kb without an ILE insert and many studies using a plasmid above 100kb often require specialised competent cells and transformation via electroporation for successful transformation (Addgene, 2017). Another point to

consider is the difference in transformation efficiency between different bacterial strains. A study on *E. coli* transformation showed that different strains can show up to a 30-fold difference in transformation efficiency (Sheng *et al.,* 1995). This may be why *P. fluorescens* readily takes up pWW0. The third possible reason the *P. syringae* strains cannot be transformed with pWW0 may be because a gene on pWW0 is lethal to the cells preventing growth post transformation.

To overcome these potential problems with the full form of pWW0 it was decided to clone the *rulAB* operon with promoter and terminator into the pBBR1MCS-2 broad host range vector. This would eliminate the two previous problems; the plasmid would be smaller, increasing transformation efficiency and any potentially lethal genes would no longer be present, providing pWW0 *rulAB* was not lethal to the *P. syringae* strains. Alongside cloning the pWW0 *rulAB* a *rulAB* operon was cloned from the PPHGI-1 genomic island in *P. syringae* pv. *phaseolicola* 1302A to assess any differences between *rulAB* operons. However the full version of *rulAB* from PPHGI-1 would not clone whereas a version of it missing the last 60bp of the operon would.

Cloning was attempted with the standard pCR2.1 cloning protocol and also direct PCR cloning into pBBR1MCS-2 using restriction site tagged primers (Brown, 2006; Dallas-Yang *et al.,* 1998).This led to the conclusion that the PPHGI-1 *rulAB* was lethal to *Ppi 203* and *Psy* 3023 which was an interesting and unexpected result. Once both the pWW0 and PPHGI-1 *rulAB* operons were cloned into pBBR1MCS-2 they were electroporated into FH1, *Ppi* 203 and *Psy* 3023. ILE movement into the cloned *rulB* genes was only observed for FH1 with the pWW0 *rulAB* present. This may be due to the attachment sites in the two different species being different and therefore not facilitating insertion. If the sequences are different then recombination at the attachment site will not occur.

This will require further work cloning a more similar *rulAB* operon. This was interesting as it meant that conjugation was not the only driving force behind ILE movement as previously thought. It had been previously hypothesised that the mechanisms of horizontal gene transfer such as conjugation may trigger the integration of ILEs into *rulB* through induction of the integron integrase (Rhodes *et al.,* 2014; Baharoglu *et al*., 2010; 2012; Cambray *et al*., 2011). It also meant that only *rulAB* was needed from pWW0 with all of the other plasmid genes apparently not required. Further work will be needed on the *P. syringae* ILEs using different *rulAB* operons from strains phylogenetically closer to the ILE containing strains.

Following on from the successful FH1 ILE mobility tests into the cloned *rulAB* from pWW0 the research continued to use FH1 as a model for ILE movement. Different versions of pWW0 *rulAB* were cloned to assess if the intact *rulAB* operon was required or if only a sequence target was sufficient for ILE insertion. To test this three versions of pWW0 *rulAB* were cloned: the intact version as used previously; just *rulB; rulAB'-IP* which consisted of the *rulA* gene but only up to the previously identified ILE insertion point. It was already known that the intact version would show ILE movement, the *rulB* version was unknown and the *rulAB'-IP* version was expected to show no ILE movement due to the whole insertion point being absent. The results showed that the ILE from FH1 only moved into the full intact version of *rulAB.* The lack of ILE movement into the other two versions of *rulAB* may be due to the lack of a functioning *rulAB* encoded DNA polymerase V or the lack of both a promoter and terminator to provide some form of transcribed protein, as promoters and terminator domains often interact with one another forming hairpins and loops during gene expression (Yang and Lewis, 2010; Meng *et al.,* 2001). The frequency of ILE insertion and selection into

either the cloned pWW0 *rulB* or the native *rulB* on pWW0 appears to remain the same across both systems (Table 6.1). However there were some similarities across all three replicate experiments in terms of the frequency of insertion and it was not as random as expected. This could be due to an artefact being present in the experiment that facilitates ILE insertion and selection in *rulB.* This artefact may have varied in the experimental preparations causing the pattern to be observed. Although this may alter the individual frequencies the overall conclusion that there is no bias between either system is still valid. Future work could include cloning the different versions of *rulAB* but including suitable promoters and terminator domains to see what effect this has on ILE insertion.

This still leaves the big question, why target *rulAB?* It has been hypothesised by Rhodes *et al.* (2014) that the ILEs may insert into *rulB* on pWW0 as it is near the origin of replication for pWW0 ensuring that it is one of the first regions transferred during conjugation and therefore increasing its chances of replication. However, another possibility is that ILEs insert into *rulB* to use its regulatory mechanisms to drive its own expression. The expression analysis results (Chapter 5) suggest that the ILE integrase gene, *xerC,* is not regulated by the LexA/SOS pathway as seen with integrons (Cambray *et al.,* 2011). However there does appear to be a promoter signal at the start of the ILEs studied (Figure 4.4). Another possibility is that the studied *rulB* gene from pWW0 whilst encoding a DNA polymerase V also exhibits some form of recombinase activity facilitating ILE integration and possibly ILE circularisation. There have been indications that the *rulB* encoded protein may function as a recombinase like protein as when *rulB* on PPHGI-1 is knocked out the genomic island loses the ability to form a circular intermediate (Lovell *et al.,* 2009). This was one of the reasons for carrying out circularisation tests on the ILEs.

The ILE circular intermediate tests did not provide any DNA sequence data at this time. However it was important to test if and when ILEs form circular intermediates during movement. Other MGEs including genomic islands such as PPHGI-1 (Godfrey *et al.,* 2011) form circular intermediates when excised from the genome. Forming a circular intermediate is beneficial as it more easily transferred between systems, may be self-replicative and a circular molecule is more stable and less likely to be degraded than a linear molecule.

There could be multiple reasons why the ILE circular intermediate tests failed including low specificity primers, the wrong conditions for ILE excision or simply that ILEs do not form circular intermediates. This area of research requires further investigation. The first step would be to design new primers to amplify across the circular intermediate if it forms (Figure 6.15).

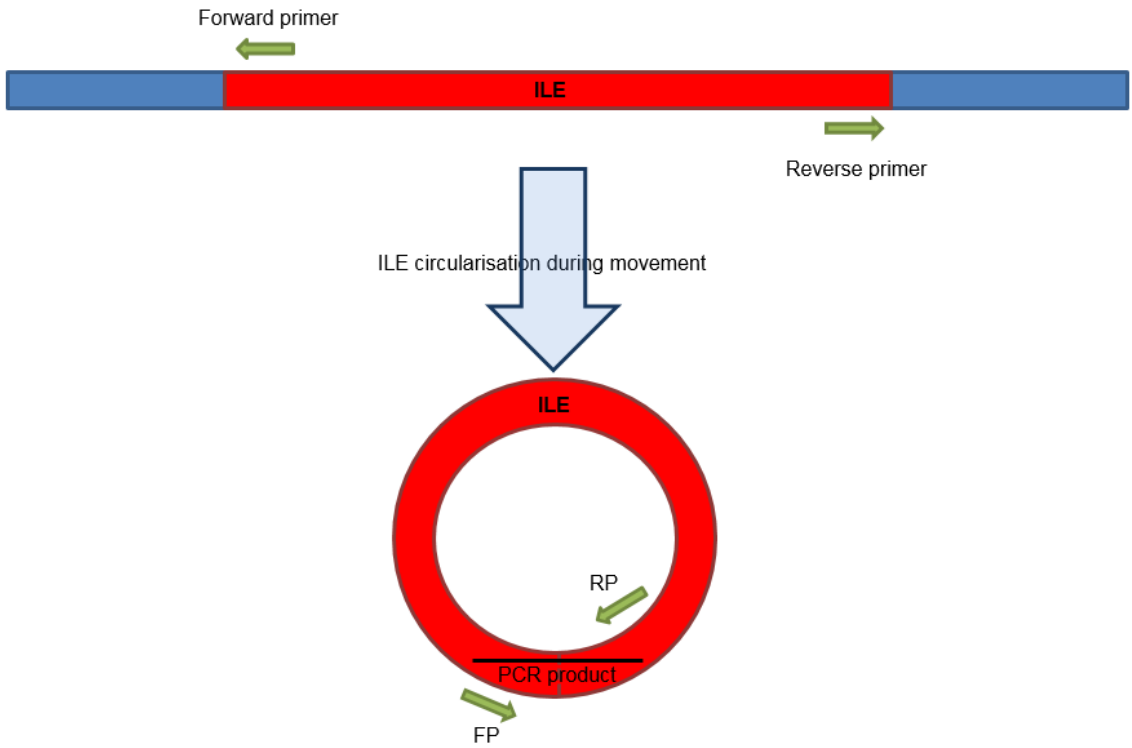


Figure 6.15: Generic outline of how ILE circularisation PCR tests work. Two primers are designed facing out from the ILE meaning no amplification will occur with the ILE in its linear form. If the ILE is excised from the genome and forms a circular intermediate molecule the primers will face each other and a PCR product will be produced.

It has been identified that in the case of the movement of *P. fluorescens* FH1 ILE it will only move into the *rulB* gene derived from pWW0, however only *rulAB* is required from pWW0 but it must be in its intact form. It also appears that the *P. syringae* ILEs are not mobile when using the same conditions. Future work should focus on investigating if the *P. syringae* ILEs are mobile by using different *rulAB* operons from different strains and species. As well as studying the relation between ILEs and *rulB* and if there is any regulatory benefit conferred by *rulAB* to the ILE.

Chapter 7. What effect does disruption of the *rulB* gene have on bacterial growth under ultra-violet radiation stress?

7.1: Introduction

One of the aims of this research was to ascertain the impact of the ILE disruption to the *rulB* gene on bacterial growth following ultra-violet irradiation (UVB). The *rulB* gene forms an operon with *rulA* which provides UVB tolerance to the bacterial cells by encoding an error-prone DNA polymerase V belonging to the gamma (ɣ) family of polymerases.

Bacteria have many mechanisms to deal with DNA damage caused by UVB radiation. These include the direct reversal of the damage by photoreactivation, removal of the damaged base by base excision repair and removal of a complete oligonucleotide by nucleotide excision repair (Goosen and Moolenaar, 2008). Other pathways are also used for DNA damage repair including homologous end-joining and nonhomologous end-joining (Shuman and Glickman, 2007). The primary pathway for the repair of DNA damaged by UVB irradiation is the direct reversal of pyrimidine dimers that form during UVB exposure. This is called photoreactivation and is catalysed by the enzyme, photolyase (Lledo and Lynch, 2009) (Figure 7.1). Photoreactivation uses photolyase to split the pyrimidine dimer that has formed on the DNA due to UVB irradiation. If the pyrimidine dimer was not rectified the replication of DNA would stop due to the polymerase III mechanism being unable to read the correct DNA base. Photolyase uses light at 300 nm to split the dimer apart using a chromophore and FADH to generate the energy required (Weber, 2005).

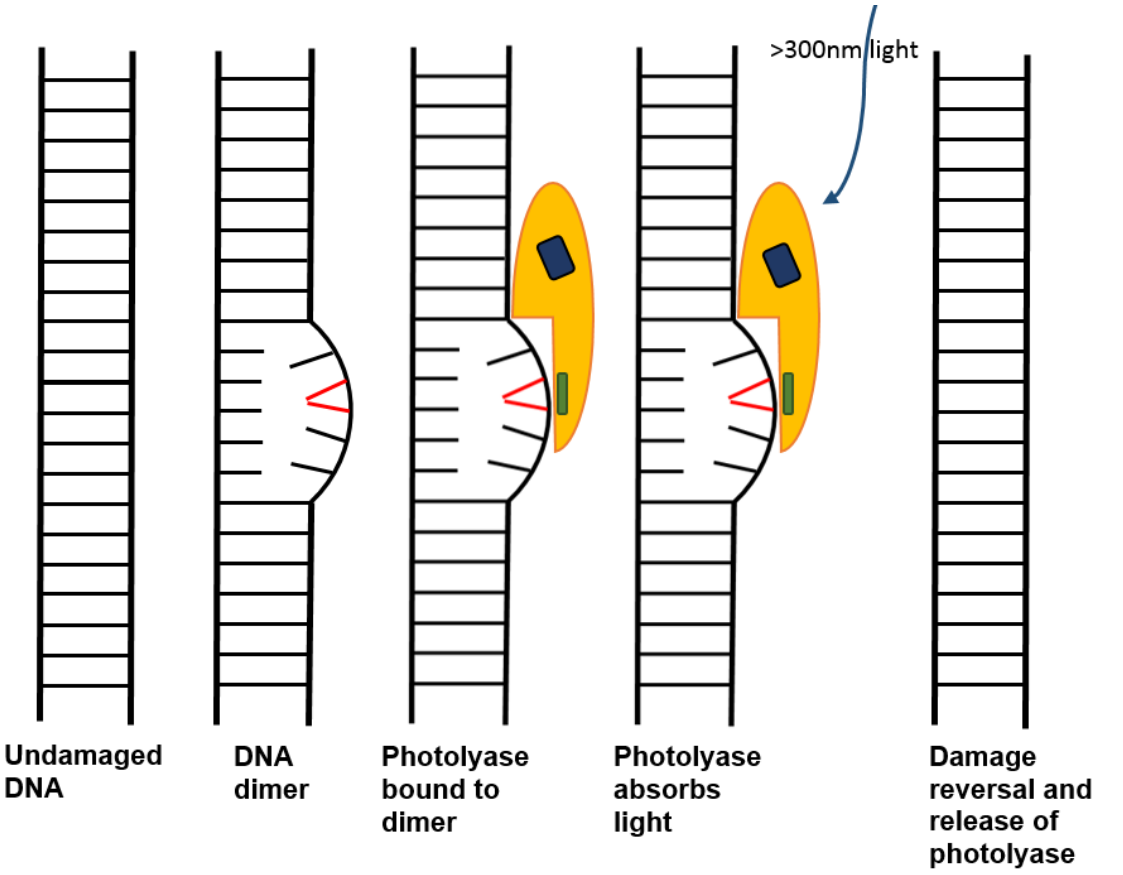


Figure 7.1: Direct reversal of pyrimidine dimers caused by UVB DNA damage. UVB irradiation causes the formation of pyrimidine dimers. Dimers are lesions caused when two consecutive bases on one strand bind together. Photolyase contains two chromophores responsible for absorbing light energy. One of the chromophores is FADH. The energy from FADH is then used to split the dimer apart, returning the DNA back to its undamaged state.

One of the other processes used to correct damaged DNA following either UVB irradiation or other DNA damaging events, such as cross-linking caused by Mitomycin C is a methyl-directed mismatch repair system. This system removes wrongly incorporated bases into the damaged DNA using the *mut* family of genes to correct gaps and mismatches in front of the stalled replication fork. The Mut proteins form a complex (MutS and MutL) at the hemi-methylated site, double stranded DNA is not entirely methylated due to the mismatch, Mut S and L cleave the strand with the incorrect base and the strand is resynthesized (Figure 7.2) (Fukui, 2010). One of the issues with this is that a defect in one of the Mut proteins

can cause mutation rates to be elevated up to 100-fold due to incorrect bases not being rectified (Sundin and Weigand, 2007).

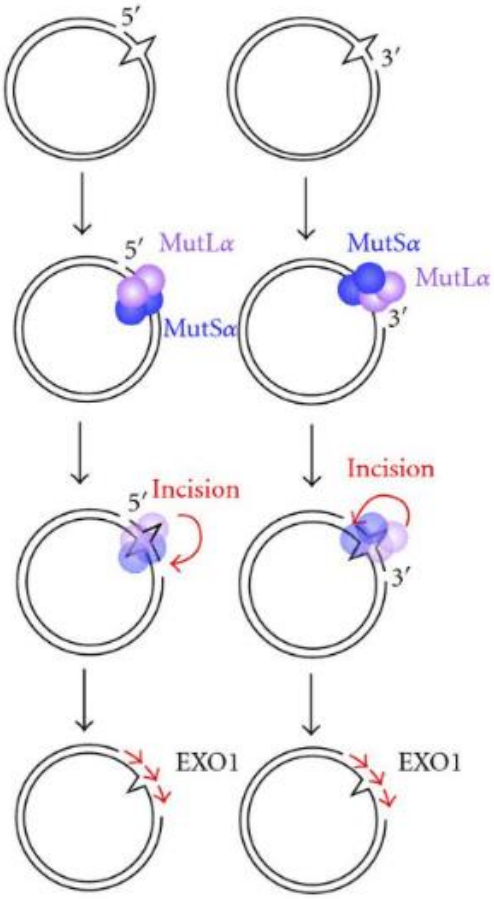


Figure 7.2: The Mut complex and DNA repair interaction. The MutS and L complex interacts with a DNA-looping mechanism. The Mut proteins cuts specifically on the non-methylated strand to remove mismatch base. Exonuclease regenerates the nucleotide gap. (Image from Fukui, 2010 used with the creative commons license 3.0; creativecommons.org/licenses/by/3.0/).

Another UVB damage repair mechanism is the *rulAB* system. The reason *rulB* is crucial to maintaining bacterial growth during high levels of UVB exposure is that *rulB* encodes a DNA polymerase V belonging to the error-prone Y family of polymerases which also includes DNA polymerase IV proteins (Hawver *et al*., 2015: Ohmori *et al*., 2001). Polymerase V can synthesise DNA across lesions caused by UVB irradiation and have the ability to maintain DNA replication following DNA damage. This results in enhanced bacterial survival (Stockwell *et*

al., 2013; Cazorla *et al*., 2008). DNA damaging agents, such as UVB irradiation causes the normal dsDNA to become fragmented into single-stranded DNA that becomes detached from the genome. These single-stranded breaks in the dsDNA cause the replication process to stall until the DNA break is fixed.

The DNA repair process involving *rulB* and its protein product DNA polymerase V is more error prone that the photoreactivation DNA damage repair system. Due to the ability of Y family polymerases introducing potential errors into the replication process their expression is highly regulated by the SOS pathway (Krishna *et al.,* 2007; Bridges, 2005). This pathway is activated when the DNA replication fork is stalled due to dsDNA breaks. The stall causes long stretches of ssDNA. The ssDNA activates the RecA pathway which induces the autoproteolytic cleavage of LexA allowing SOS genes to be expressed. LexA is a repressor of *rulB* and is activated by LexA cleavage. This results in the upregulation of the RulB DNA polymerase V which has the ability to bypass the DNA breaks and perform translesion DNA synthesis (Figure 7.3) (Bridges, 2005; Sundin and Weigand, 2007). However the nature of these error prone polymerases produces an increase in mutation rates whilst active. This could lead to a trade off in the form of bacterial survival versus mutations due to the relatively quick activation of DNA polymerase V, compared to the time for other repair mechanisms to resolve the DNA damage. The DNA lesions and replication fork stalls cause cell division to cease and if not rectified will lead to bacterial cell filamentation. Cells continue to elongate but do not divide and this leads to eventual cell death (Charpentier *et al.,* 2011).

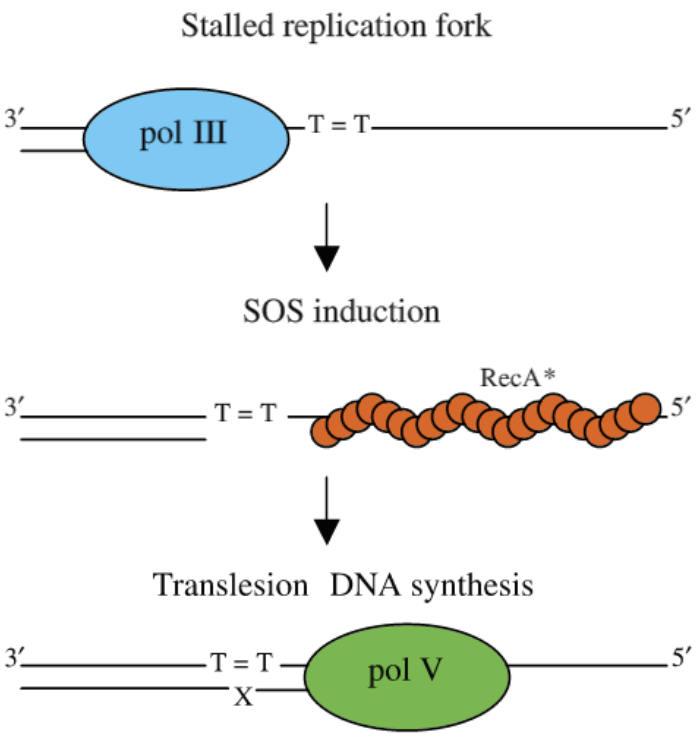


Figure 7.3: How DNA polymerase V works. Translesion synthesis occurring in response to DNA damage. Induction of the SOS regulon in response to DNA damage includes the expression of the error-prone DNA polymerases V (RulB). Pol. III is a DNA polymerase that is the primary enzyme used in DNA replication. (Image adapted from Sundin and Weigand, (2007). Used under agreement from Oxford University press, no. 4138330783037)

There are many homologues of the *rulAB* operon that encode similar DNA polymerase V enzymes. This includes *umuDC* first described in *E. coil* (Smith and Walker, 1998) and *rumAB* found in *Vibrio* species (Hochhut *et al.,* 2001). The operons encode resistance to UVB radiation across many different bacteria (Sundin *et al.,* 2000). Research has indicated that the *rulAB* homologue *umuDC* acts as a SOS response regulator when the *umuD* (*rulA*) protein is truncated to UmuDpR and has the same mode of action as LexA (Diaz-Magana *et al.,* 2015). Could the same also be true of *rulAB* and could *rulA* have some form of regulation function as well as *rulB* encoding a DNA polymerase V? If this is the case *rulAB* may regulate its own expression following the activation of the SOS response.

Members of the *Pseudomonas syringae* species need to be able to withstand high levels of UVB radiation as *P. syringae* populations are often located in the phyllosphere and more usually on the phylloplane (Lindow and Brandl, 2003). *P. syringae* colonises the leaf in high numbers (~$3x10^5$ cells) before entering the plant via the stomata (Melotto *et al.,* 2008: Hirano *et al.,* 1995). A functional *rulAB* operon is very important in maintaining high bacterial populations on the leaf surface (Sundin *et al.,* 2000). Although UVB levels experienced by *P. syringae* will be higher than other ecological and phytopathogenic bacteria, their exposure is lessened by *P. syringae* typically colonising the underside of the leaf prior to infection. However, the expression of *rulAB* in *P. syringae* is thought to contribute significantly to their virulence and ecological fitness due to the increased UVB tolerance (Tark *et al.,* 2005). Therefore any bacterial population colonising the plant surface and attempting to infect the host must be able to withstand UVB exposure even on the underside of the leaf. The *rulAB* determinant has been described in 14 pathovars of *P. syringae.* However the *rulAB* positive pathovars have a wide range of UVB radiation tolerance (Sundin *et al.,* 2000).

It had not been previously considered whether the disruption to *rulB* affects the primary role of *rulAB* in enhancing bacterial survival in mutagenic environments. However there have been studies into the presence and absence of *rulB* on bacterial fitness following UVB exposure (Zhang *et al.,* 2004). The results showed no significant difference in both bacterial growth rate and overall bacterial fitness between strains lacking a *rulAB* operon following UVB irradiation. Analysis of previously identified *rulB* ILE disruptions revealed a uniform distribution of insertions within the *rulB* coding sequence that result in the truncated *rulB* fragments, *rulB'* (~120bp) and *rulB"* (~1080bp), remaining in frame

(Jackson et al., 2011). This suggests that native DNA repair functions of *rulB* encoded DNA polymerase V may be maintained through one or both of the individual disrupted *rulB* gene fragments.

It was essential to ascertain the functionality of the *rulB* protein product in both forms, intact and disrupted by an ILE insertion. By assessing the functionality of *rulB* in terms of bacterial growth following UVB exposure the results would indicate three possibilities. Firstly that an intact *rulB* does not confer any benefit to the growth of the bacteria. Secondly that an intact *rulB* does confer a beneficial trait to *P. syringae* and finally the possibility that *rulB* functions in the same way regardless of its form and the fragmented *rulB* produces a functional protein similar to the full protein product. In order to come to a conclusion, all of the ILE positive *P. syringae* strains with a disrupted *rulB* gene were irradiated with UVB along with *P. syringae* strains containing an intact *rulB* gene. *Psy.* B86-17 was the control for intact *rulB* strains as its *rulAB* operon has been sequenced. *Ppi* 203 was the control for ILE disrupted *rulB* due to this disruption being sequenced (Arnold *et al.,*1999; 2000) (Table 7.1). There were also two other strains included as controls; *P. putida* PaW340 (pWW0::km^r) which had an intact, sequenced *rulB* gene and *P. fluorescens* FH1 (pWW0::km^r::ILE$_{FH1}$) which had a sequenced disrupted *rulB* gene. These were included to observe any differences in UVB tolerance between *Pseudomonas* species. The bacterial strains were subjected to 0, 30, 60 and 120 seconds of UVB exposure and their growth was recorded every hour for eight hours and again at 24 hours via optical density at 600 nm (OD_{600}).

Table 7.1: List of *Pseudomonas* strains used for the UVB stress growth assay along with their *rulB* determinant. ✓= presence of gene ✗ = absence of gene.

Strain	Intact *rulB*	Disrupted *rulB*
Pp. Paw340 (pWW0::km^r)	✓	✗
Psy. B86-17	✓	✗
Ppi. 1124B	✓	✗
Ppi. 1746A	✓	✗
Ppi. 1758B	✓	✗
Ppi. 1785A	✓	✗
Ppi. 1796A	✓	✗
Ppi. 1807A	✓	✗
Pph. 1302A	✓	✗
Pph. 1375A	✓	✗
Psy. 1150	✓	✗
Psy. 2242A	✓	✗
Pma. 1809	✓	✗
//	////////////////	//////////////////////
Pf. FH1 (pWW0::km$_r$::ILE$_{FH1}$)	✗	✓
Ppi. 203	✗	✓
Ppi. 202	✗	✓
Ppi. 223	✗	✓
Ppi. 283	✗	✓
Ppi. 288	✗	✓
Ppi. 390	✗	✓
Ppi. 1456A	✗	✓
Ppi. 1939	✗	✓
Ppi. 2889B	✗	✓
Psy. B728a	✗	✓
Psy. 3023	✗	✓
Pma. 1852A	✗	✓
Pma. 5422	✗	✓
Pma. 6201	✗	✓

7.2: Results

7.2.1: Assessing the impact of UVB irradiation on *Pseudomonas* strains that contain an intact *rulB* gene.

The growth of 13 strains containing intact versions of the *rulB* gene, selected from strains in Chapter 2, was recorded every hour for eight hours and then a final time point at 24 hours with two replicates per strain. The strains were subjected to UVB irradiation for zero, 30, 60 and 120 seconds with subsequent recovery in LB liquid media. A further exposure level was used at 300 seconds but this resulted in all of the strains failing to grow even after 48 hours so it was not included in future tests. The optical densities at a wavelength of 600nm (OD_{600}) were measured and used to produce a series of growth curves for the various UVB exposure times (Figures 7.4-7.7). The controls used were *P. putida* PaW340 (pWW0::km^r) and *P. syringae* pv. *syringae* B86-17, both have sequenced intact *rulB* genes and belong to the *Pseudomonas* genus.

7.2.1a: Growth of *Pseudomonas* strains with intact *rulB* following no UVB exposure.

Zero UVB exposure tests set the base line for growth of the strains containing an intact *rulB.* The data (Figure 7.4) showed that all of the strains increase in cell density over the time period, all reaching an OD_{600} >0.1 by eight hours and an OD_{600} range of 1.4 ± 0.0 to 2.4 ± 0.3 by 24 hours. These results showed that intact *rulB* strains will grow in the given conditions of the experiment and any differences in individual strain growth can be normalised using this data.

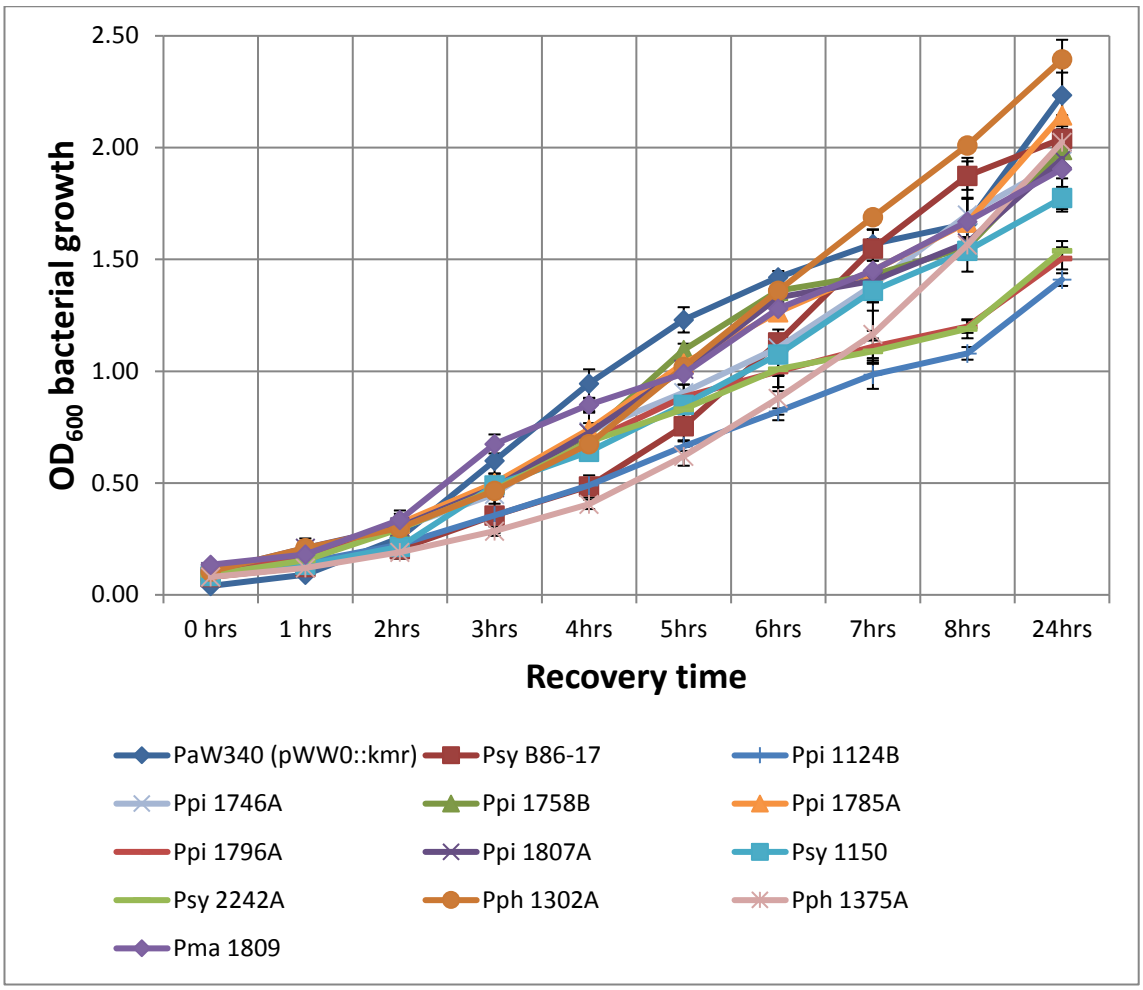


Figure 7.4: Growth of *Pseudomonas* strains containing an intact *rulB* gene without UVB (302nm) exposure. Cell density via optical density at 600nm was measured over the course of 24 hours. PaW340 (pWW0::km^r) and *Psy* B86-17 were used as control strains. Two replicates per strain were tested and their mean values are displayed with standard deviation error bars.

7.2.1b: Growth of *Pseudomonas* strains with intact *rulB* following 30 seconds of UVB exposure.

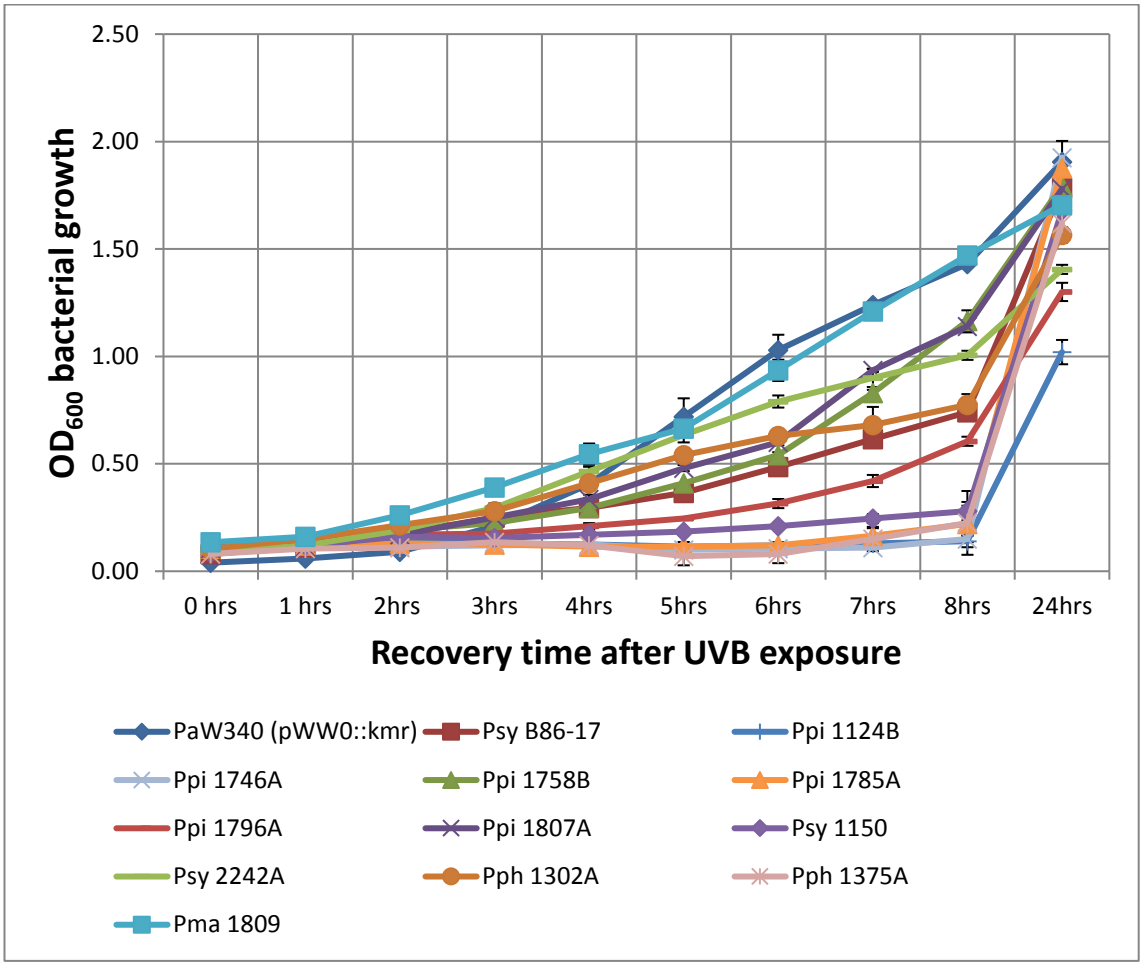


Figure 7.5: Growth of *Pseudomonas* strains containing an intact *rulB* gene following 30 seconds of UVB (302nm) exposure. Cell density via optical density at 600nm was measured over the course of 24 hours. PaW340 (pWW0::km^r) and *Psy* B86-17 were used as control strains. Two replicates per strain were tested and their mean values are displayed with standard deviation error bars.

Following 30 seconds of UVB exposure it was observed that some of the intact *rulB* strains had difficulty growing following the exposure for the first eight hours. *Ppi* 1124B only increased in OD_{600} from 0.1 ± 0.0 to 0.2 ± 0.1 over the eight hour period compared to other strains such as *Pma* 1809 which had a bigger growth increase from 0.1 ± 0.0 to 1.7 ± 0.00, indicating that some strains may be naturally more susceptible to UVB radiation that others. It was also clear that the growth of the strains was affected as the mean OD_{600} of all the strains with an

intact *rulB* at eight hours was down from 1.6 ± 0.1 with zero UVB exposure to 0.7 ± 0.1 with 30 seconds exposure and at 24 hours down from 1.9 ± 0.1 to 1.7 ± 0.1 respectively. However, all of the strains had grown to a density of >1.0 by the 24 hour time point.

7.2.1c: Growth of *Pseudomonas* strains with intact *rulB* following 60 seconds of UVB exposure.

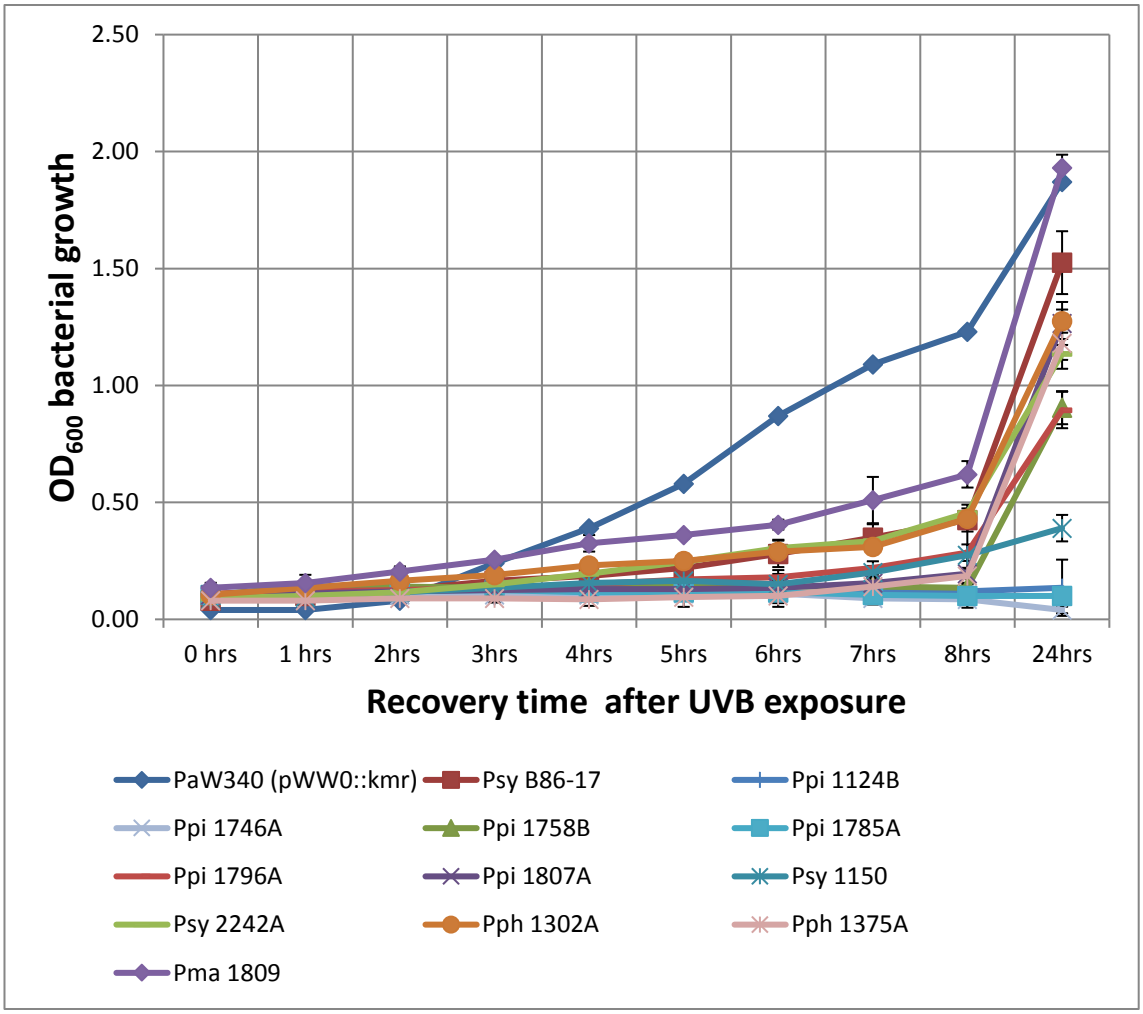


Figure 7.6: Growth of *Pseudomonas* strains containing an intact *rulB* gene following 60 seconds of UVB (302nm) exposure. Cell density via optical density at 600nm was measured over the course of 24 hours. PaW340 (pWW0::km^r) and *Psy* B86-17 were used as control strains. Two replicates per strain were tested and their mean values are displayed with standard deviation error bars.

With an increase in UVB exposure time to 60 seconds the cell density remained lower than the previous exposure levels of zero and 30 seconds with

strains only starting to recover by the eight hour time point. The control strain, PaW340 (pWW0::km^r), showed minor differences to the previous test, down from 1.4 ± 0.0 to 1.2 ± 0.0 after eight hours. However, the *P. syringae* control strain B86-17 was similar to the other *P. syringae* strains, meaning that UVB tolerance may be a species specific trait. When the mean growth density of all the intact *rulB* strains at eight hours was compared to the zero UVB exposure baseline the cell density had decreased by 1.2 from 1.6 ± 0.1 to 0.4 ± 0.1 and at 24 hours from 1.9 ± 0.1 to 1.0 ± 0.2, a decrease of 0.9. It was clear from Figure 7.6 that 60 seconds of UVB exposure had affected the cells much more than the 30 seconds exposure. For example three strains, *Ppi* 1746A, *Ppi* 1785A and *Ppi* 1124B failed to recover at all from 60 seconds of UVB exposure and the majority of the strains, 11 out of 13, were below an OD_{600} of 0.5 at eight hours.

7.2.1d: Growth of *Pseudomonas* strains with intact *rulB* following 120 seconds of UVB exposure.

Figure 7.7 highlights the impact of UVB radiation on bacterial cells. Nine out of the 13 strains with intact *rulB* failed to grow when subjected to 120 seconds of direct UVB exposure. None of the strains reached an OD_{600} of 0.3 at eight hours compared to all of the strains being above an OD_{600} of one with zero UVB exposure. This was also highlighted in the mean OD_{600} growth values both at eight hours and 24 hours. At eight hours the mean OD_{600} of all strains with an intact *rulB* with no UVB exposure was 1.6 ± 0.1 compared to 0.1 ± 0.0 with 120 seconds of UVB exposure. The same pattern was also true at 24 hours with the mean value decreasing from 1.9 ± 0.1 to 0.4 ± 0.1.

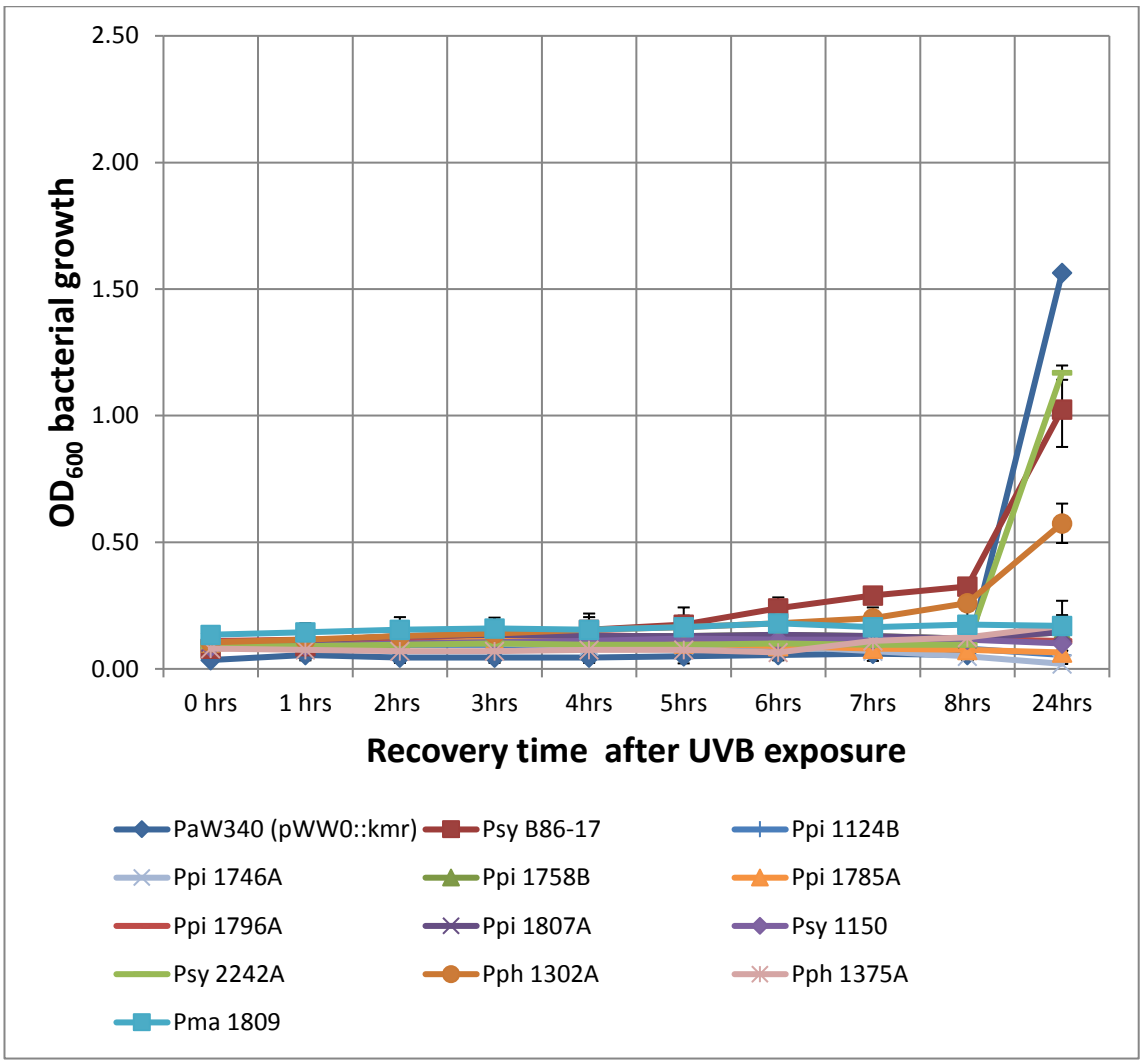


Figure 7.7 Growth of *Pseudomonas* strains containing an intact *rulB* gene following 120 seconds of UVB (302nm) exposure. Cell density via optical density at 600nm was measured over the course of 24 hours. PaW340 (pWW0::km^r) and *Psy* B86-17 were used as control strains. Two replicates per strain were tested and their mean values are displayed with standard deviation error bars.

7.2.2: Assessing the impact of UVB radiation on bacterial strains that contain an ILE disrupted *rulB* gene.

As *rulB* encodes a DNA polymerase V which aids in DNA replication following UVB DNA damage it was important to investigate what effect an ILE disruption to *rulB* had on bacterial growth. A selection of strains with ILE insertions into *rulB* were tested in the same way as the strains with an intact *rulB* gene (Section 7.2.1). Two controls were used for these experiments, *P.*

fluorescens FH1 (pWW0::km^r::ILE$_{FH1}$) and *P. syringae* pv. *pisi* 203 both have sequenced *rulB* disruptions caused by an ILE insertion. See Chapters 3 and 4 for further information on these control strains.

7.2.2a: Growth of *Pseudomonas* strains with a disrupted *rulB* gene following zero UVB exposure.

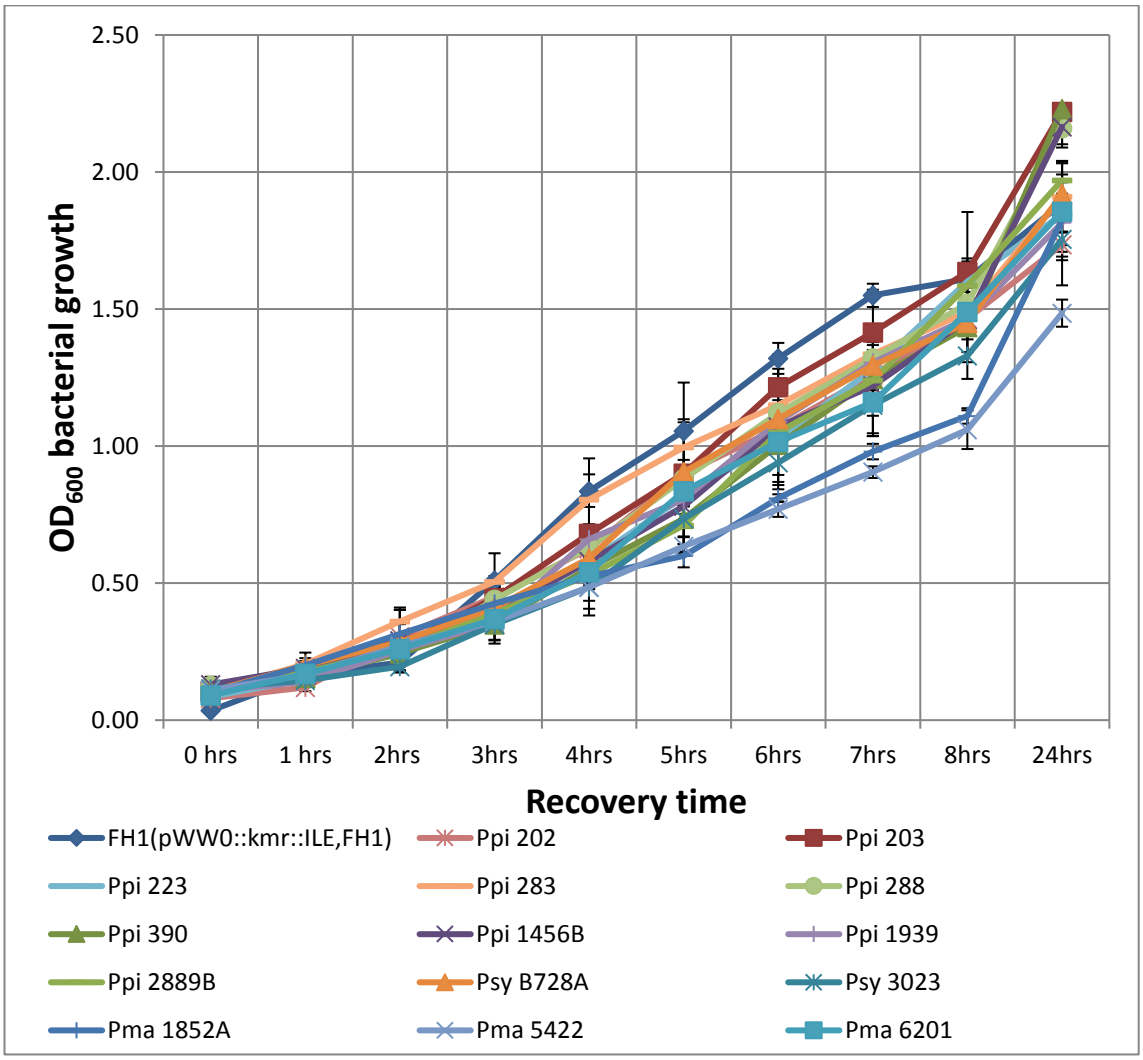


Figure 7.8: Growth of *Pseudomonas* strains containing a disrupted *rulB* gene without UVB (302nm) exposure. Cell density via optical density at 600nm was measured over the course of 24 hours. FH1 (pWW0::km^r::ILE$_{FH1}$) and *Ppi* 203 were used as control strains. Two replicates per strain were tested and their mean values are displayed with standard deviation error bars.

The baseline growth of the strains containing ILE insertions into *rulB* without UVB exposure was established to enable any differences in strain growth to be observed and to differentiate between any differences in individual strain

growth patterns. As Figure 7.8 shows, all of the strains followed the same growth pattern and also followed the same pattern as the intact *rulB* strains in Figure 7.4. There was individual growth variation present, for example at eight hours *Ppi* 203 was at an OD_{600} of 1.6 ± 0.0 compared to *Pma* 5422 which was at 1.0 ± 0.1. This was an important factor to consider when analysing any differences in growth due to the UVB exposure.

7.2.2b: Growth of *Pseudomonas* strains with a disrupted *rulB* gene following 30 seconds of UVB exposure.

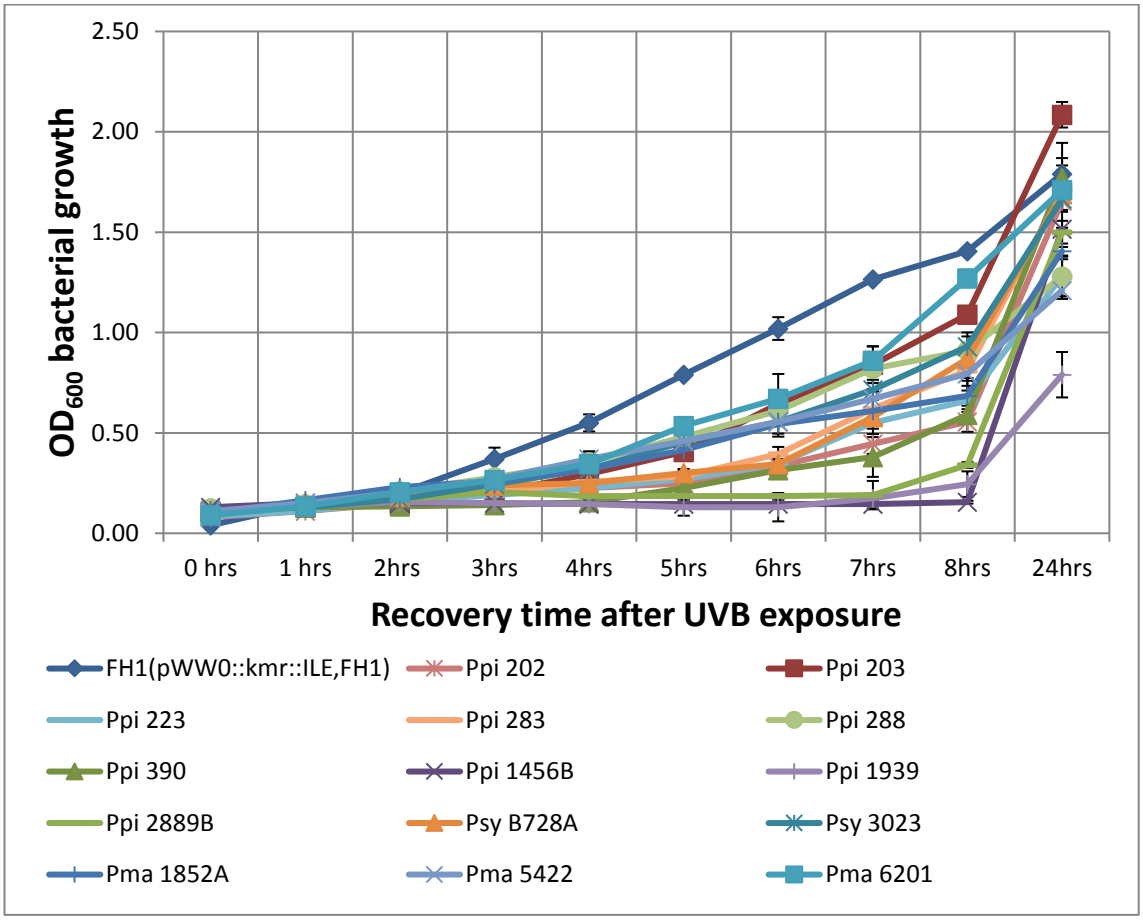


Figure 7.9: Growth of *Pseudomonas* strains containing a disrupted *rulB* gene following 30 seconds of UVB (302nm) exposure. Cell density via optical density at 600nm was measured over the course of 24 hours. FH1 (pWW0::km^r::ILE$_{FH1}$) and *Ppi* 203 were used as control strains. Two replicates per strain were tested and their mean values are displayed with standard deviation error bars.

From Figure 7.9 it is evident that 30 seconds of UVB exposure was enough to reduce the growth of the bacteria up to the eight hour time point. However as seen with the intact *rulB* graph (Figure 7.5) many of the strains regained higher levels of growth by the 24 hour point. Strain *Ppi* 1456B best showed this as its growth remained almost stationary between zero and eight hours increasing from 0.1 ± 0.0 to 0.1 ± 0.0. However, it reached an OD_{600} of 1.5 ± 0.0 after 24 hours, compared to the measurement with no UVB exposure of 2.2 ± 0.1. Another

interesting point was that the growth of the intact *rulB* strains at the same exposure level (Figure 7.5) looked very similar to the data in Figure 7.9. This was backed up by the mean OD_{600} values as the intact *rulB* average at eight hours with 30 seconds of UVB exposure was 0.7 ± 0.1 compared to the disrupted *rulB* value of 0.8 ± 0.1. The values were also similar at the 24 hour point with 1.6 ± 0.1 and 1.5 ± 0.1 respectively.

7.2.2c: Growth of *Pseudomonas* strains with a disrupted *rulB* gene following 60 seconds of UVB exposure.

The exposure level was increased to 60 seconds and this further reduced the bacterial growth (Figure 7.10). All of the strains were below an OD_{600} of 0.5 at the eight hour time point following 60 seconds of UVB exposure. The majority of the strains reached this OD_{600} or higher at eight hours following 30 seconds of UVB exposure. The strains were also below an OD_{600} of one at 24 hours following 60 seconds of UVB exposure except for the *P. fluorescens* control strain. However with 30 seconds of UVB exposure all strains except *Ppi* 1939 were above an OD_{600} of one.

These values are also significantly different to the no UVB baseline measurements where the average at eight hours was 1.4 ± 0.0 compared to the average after 60 seconds of UVB exposure at 0.2 ± 0.0. Interestingly this exposure level also showed some differences between the growth of bacteria containing the intact form of *rulB* and those with the disrupted form which was not observed at the 30 seconds exposure level. The average growth at 24 hours for bacteria with the intact *rulB* was 0.97 ± 0.18 compared to those with disrupted *rulB* value of 0.6 ± 0.1. These values were significantly different from each other (see details in Section 7.2.3).

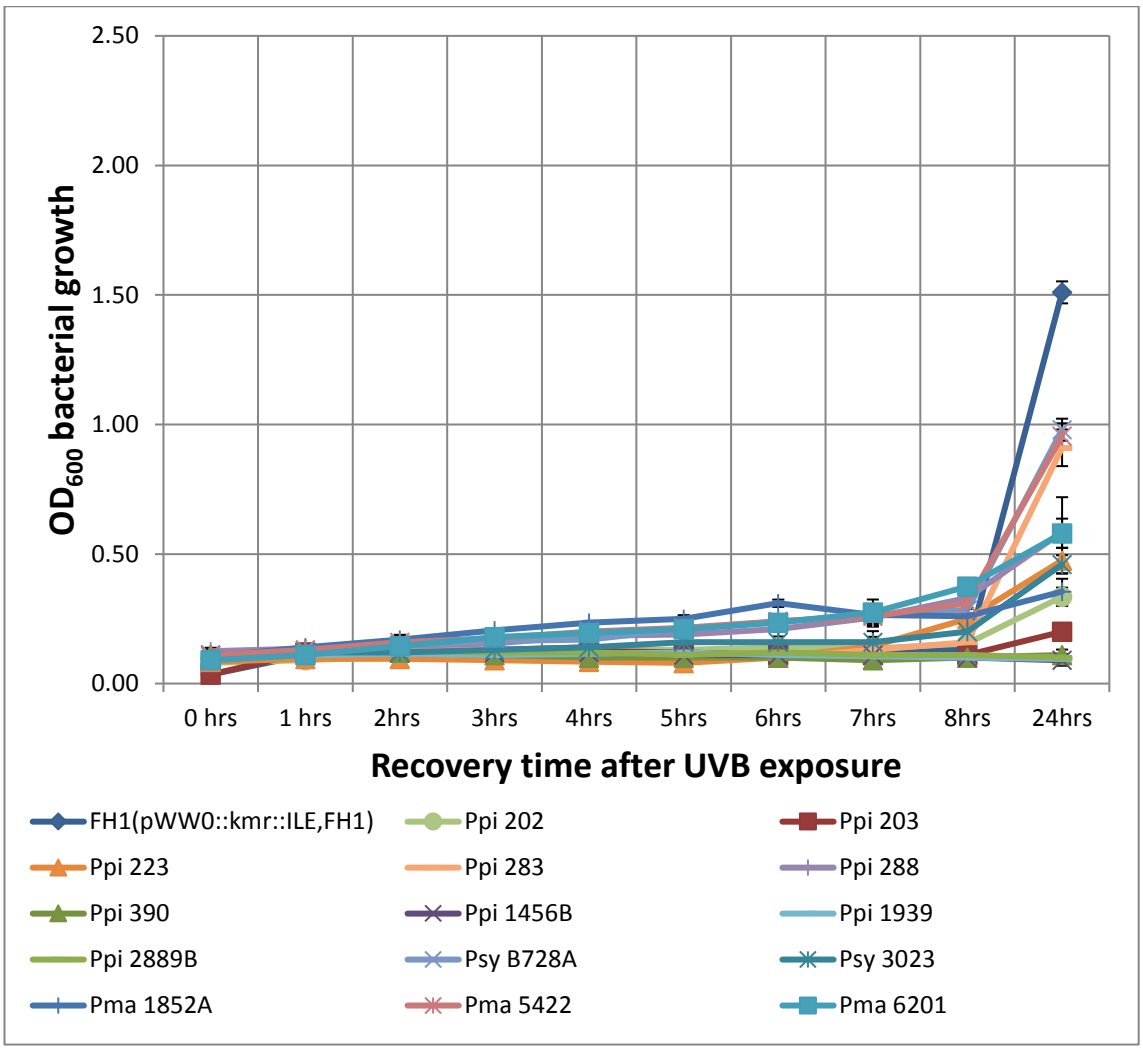


Figure 7.10: Growth of *Pseudomonas* strains containing a disrupted *rulB* gene following 60 seconds of UVB (302nm) exposure. Cell density via optical density at 600nm was measured over the course of 24 hours. FH1 (pWW0::km^r::ILE$_{FH1}$) and *Ppi* 203 were used as control strains. Two replicates per strain were tested and their mean values are displayed with standard deviation error bars.

7.2.2d: Growth of *Pseudomonas* strains with a disrupted *rulB* gene following 120 seconds of UVB exposure.

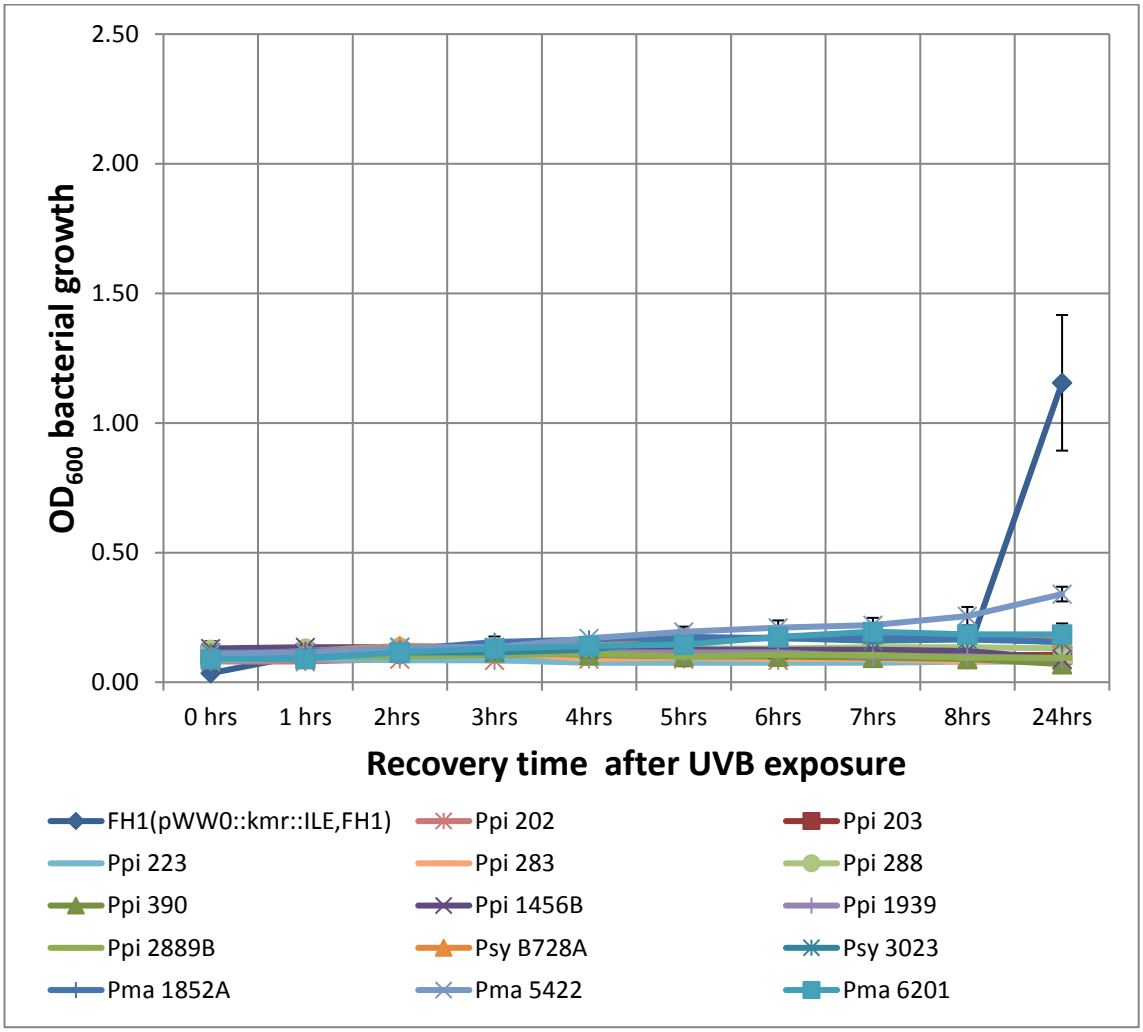


Figure 7.11: Growth of *Pseudomonas* strains containing a disrupted *rulB* gene following 120 seconds of UVB (302nm) exposure. Cell density via optical density at 600nm was measured over the course of 24 hours. FH1 (pWW0::km^r::ILE$_{FH1}$) and *Ppi* 203 were used as control strains. Two replicates per strain were tested and their mean values are displayed with standard deviation error bars.

The final exposure level of 120 seconds causes most of the strains' growth to remain stationary. This was also the same for the strains with a disrupted *rulB* gene with the exception of *Pma* 5422 which does have decreased growth compared to 60 seconds of UVB exposure. However the growth slightly increased by the 24 hour point to an OD_{600} of 0.34. The other exception was the

Pseudomonas fluorescens FH1 (pWW0::km^r::ILE$_{FH1}$) control. The OD_{600} average dropped from 0.2 ± 0.0 for 60 seconds of UVB exposure to 0.1 ± 0.0 for the 120 seconds exposure at the eight hour time point and from 0.6 ± 0.1 to 0.3 ± 0.1 respectively for the 24 hours point. Although there was a difference between the different exposure levels eight hours after exposure the mean growth values for intact and disrupted *rulB* were the same after eight hours with OD_{600} values of 0.13. However at the 24 hour time point there was a difference from 0.4 ± 0.1 for intact *rulB* to 0.3 ± 0.1 for the disrupted *rulB*.

7.2.3: Analysing the growth rates of *Pseudomonas* strains following UVB exposure.

Analysis of bacterial growth rates was performed. This would identify whether UVB irradiation affects growth rate and if the form of *rulB* confers any advantages or disadvantages following UVB exposure in terms of overall growth rate. The bacterial growth data from Sections 7.2.1 and 7.2.2 were normalised against the zero UVB exposure growth data for all of the strains for either intact *rulB* or disrupted *rulB.* This removed any variation in growth. These normalised OD_{600} growth values were transformed into natural log values (Ln) and linear regression plots were drawn using Graphpad Prism 7 software (available at: https://www.graphpad.com/scientific-software/prism/). The first data set was the linear regressions for the UVB exposure times with intact and disrupted *rulB* (Figure 7.12).

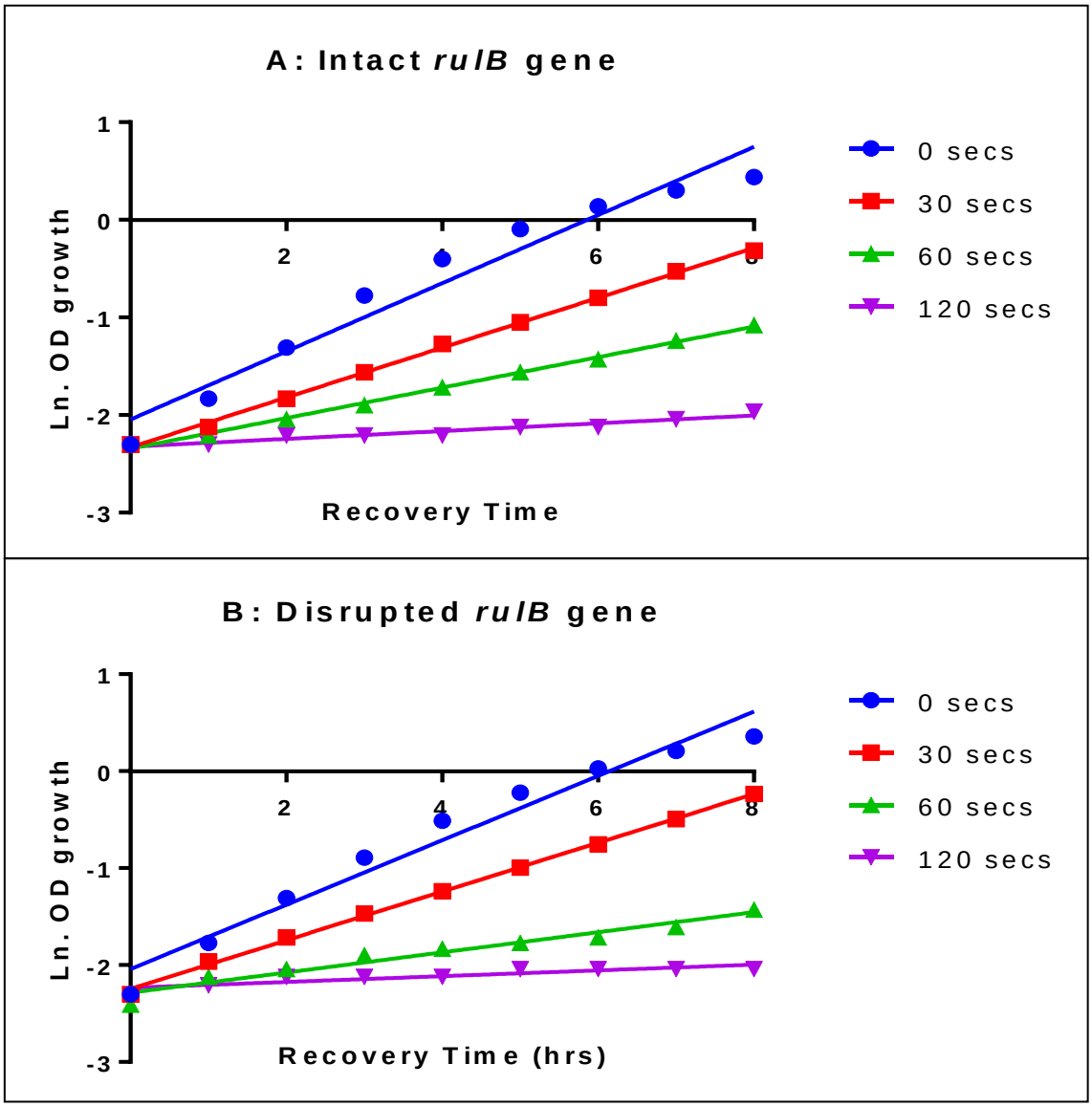


Figure 7.12: Linear regressions of bacterial growth over an eight hour time period following UVB irradiation. OD_{600} values transformed to Ln after UVB exposure and plotted against recovery time in LB broth. A) The linear regression of the normalised data obtained from strains containing an intact *rulB* gene. B) Linear regression plot of the normalised data obtained from strains containing a disrupted *rulB* gene.

The linear regression analysis (Figure 7.12) indicates the difference in growth rate following the four different UVB exposure times and indicates that with increasing UVB exposure growth rate diminishes over the course of eight hours. It is also apparent that regardless of the state of *rulB* the same pattern is observed, although this does not indicate that any significant differences have

been observed between intact and disrupted *rulB*. The analysis also gave the growth rate values for each slope (Table 7.2), allowing analysis of the rate. In the case of UVB exposure each exposure level had a significant effect on the growth rate of the bacteria. The values were significantly different within either the intact or disrupted *rulB* data set at the 95% CI. All of the slopes were significantly different from zero ($p<0.0001$).

Table 7.2: Growth rate values from linear regression analysis.

	Growth rate (95% CI)		*rulB* form cause
UV exposure time	**Intact *rulB***	**Disrupted *rulB***	**significant difference?**
0 seconds	0.28-0.42	0.28-0.39	✗
30 seconds	0.25-0.26	0.24-0.26	✗
60 seconds	0.15-0.16	0.08-0.13	✓
120 seconds	0.03-0.05	0.02-0.04	✗

The growth rates at the 95% CI overlap to some degree between intact and disrupted *rulB* strains. The exception being at the 60 second exposure time where there was a significant difference at the 95% CI meaning there was evidence to suggest that the disruption of *rulB* caused a slower growth rate after 60 seconds of UVB exposure. Disrupted *rulB* had a growth rate average of 0.1039 hr^{-1} compared to 0.1561 hr^{-1} for intact *rulB*. Therefore the disruption to *rulB* appeared to not have a significant impact on bacterial cell growth following zero, 30 and 120 seconds of UVB exposure. This was also observed when analysing the linear regression plots with both intact and disrupted *rulB* values overlaid (Figures 7.13).

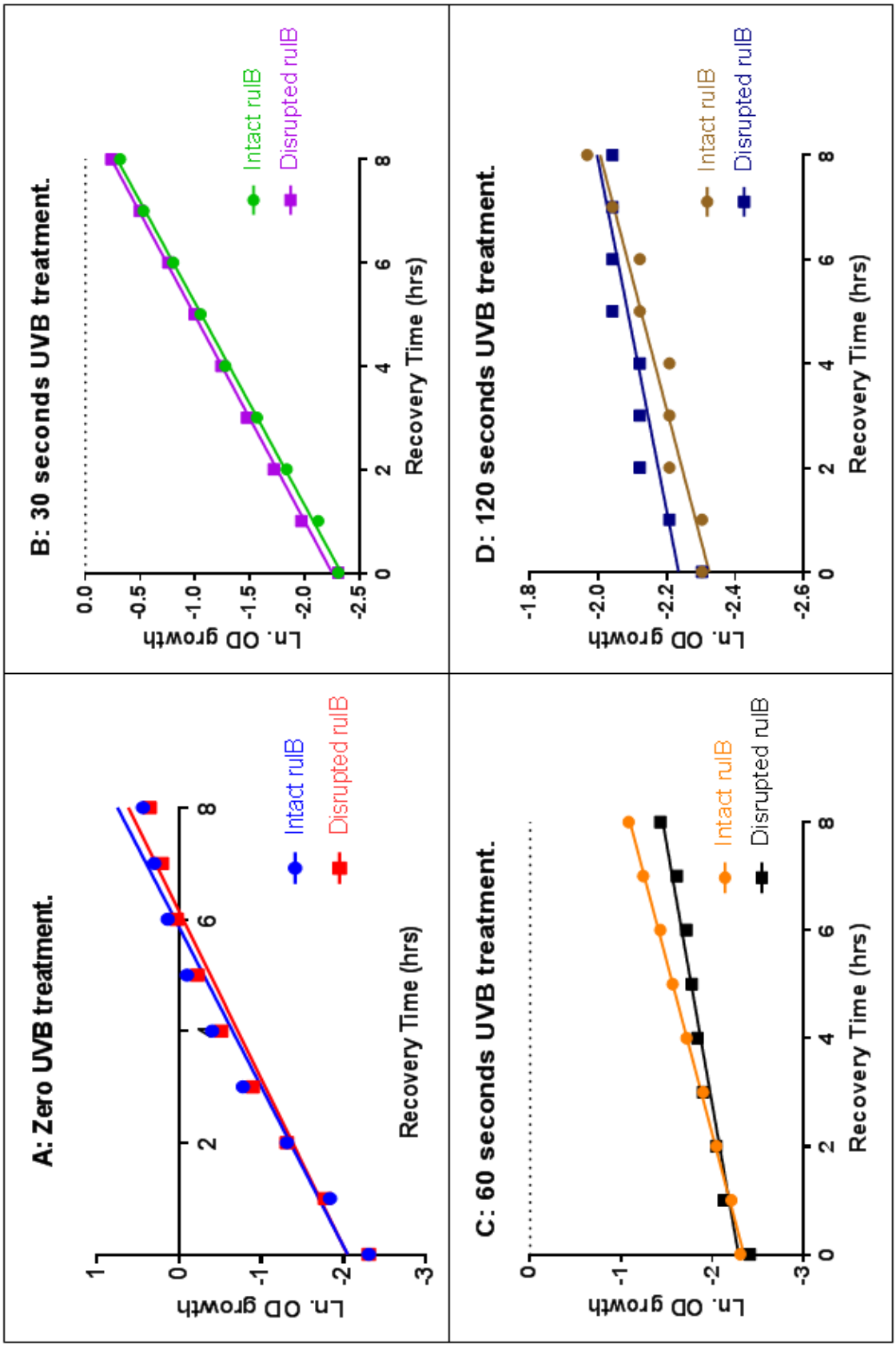


Figure 7.13: Linear regressions of normalised bacterial growth OD_{600} with intact and disrupted *rulB* plots overlaid. The plots have both *rulB* versions on to allow comparison of growth rates. A) No UVB exposure, normal cellular growth; B) Growth rate following 30 seconds of UVB exposure; C) 60 seconds of UVB exposure; D) 120 seconds of UVB exposure.

7.2.4: End point bacterial growth analysis following UVB exposure on strains containing either an intact or disrupted *rulB* gene.

The normalised data from the OD_{600} growth values used in Section 7.2.3 was used to analyse any statistical differences in end point growth eight hours after UVB exposure. Eight hours was chosen over 24 hours due to it being more representative of the relationship between UVB exposure and growth in the short time period where *rulB* may be more active. Table 7.3 shows the normalised growth values for the eight hour recovery time point for comparison between the two different *rulB* versions. The normalised data was also used to produce two growth curves to visualise the difference between exposure levels and also the differences between intact and disrupted *rulB* (Figures 7.14 and 7.15).

Table 7.3: Comparison of normalised OD_{600} data for intact and disrupted *rulB* genes at eight hours after UVB exposure. Data shown with ± SEM.

	Intact *rulB*	Disrupted *rulB*
Recovery time	**8 hours**	**8 hours**
Exposure level		
0s UVB	**1.56 ± 0.07**	**1.43 ± 0.04**
30s UVB	**0.72 ± 0.14**	**0.79 ± 0.08**
60s UVB	**0.35 ± 0.09**	**0.24 ± 0.03**
120s UVB	**0.13 ± 0.13**	**0.13 ± 0.01**

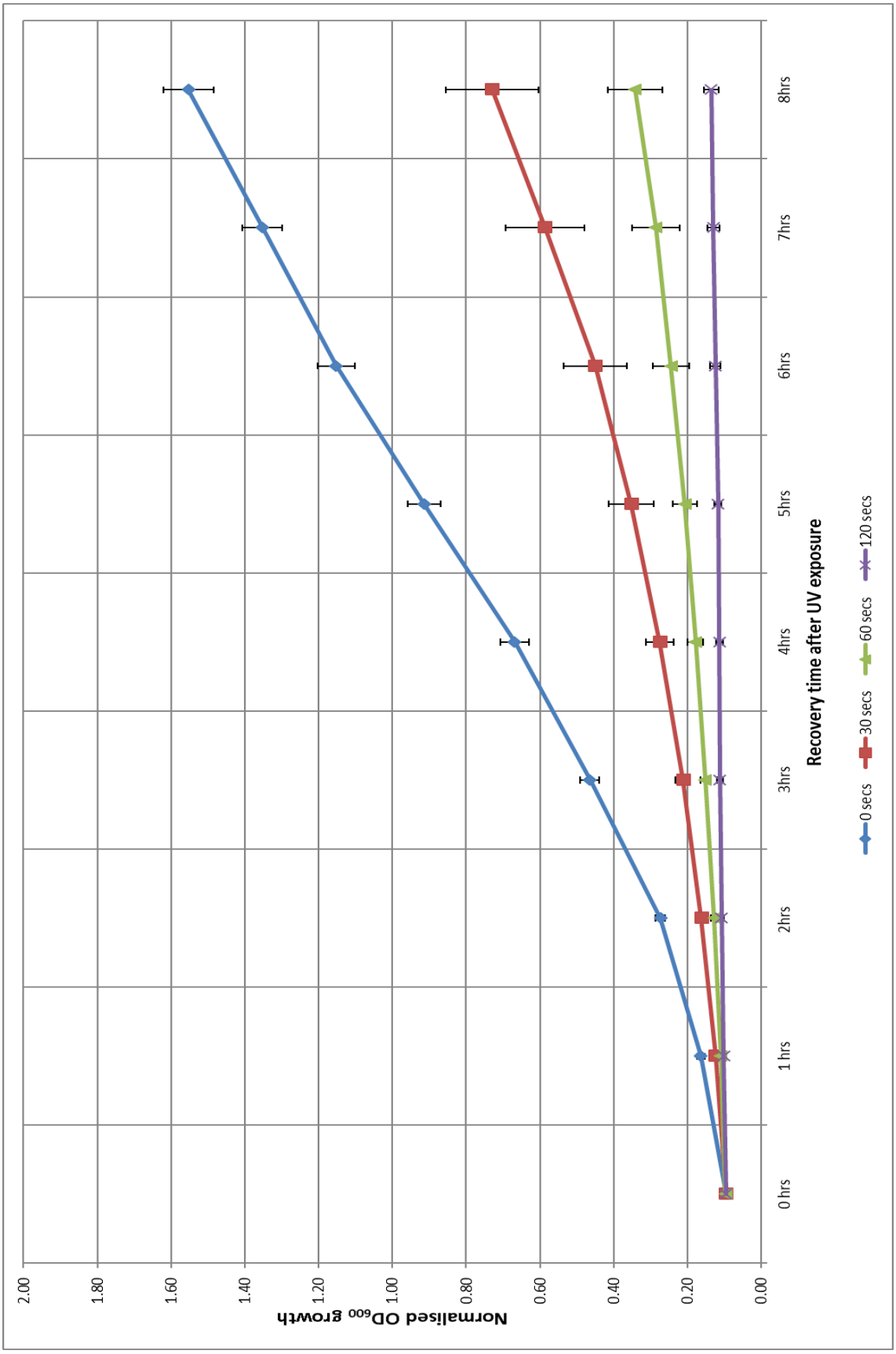


Figure 7.14: Normalised growth curve of combined *Pseudomonas* strains with an intact *rulB* gene following UVB exposure. Cell density via optical density at 600nm was measured over the course of eight hours. All of the OD_{600} growth values from 13 strains were normalised against zero UVB exposure with standard deviation error bars.

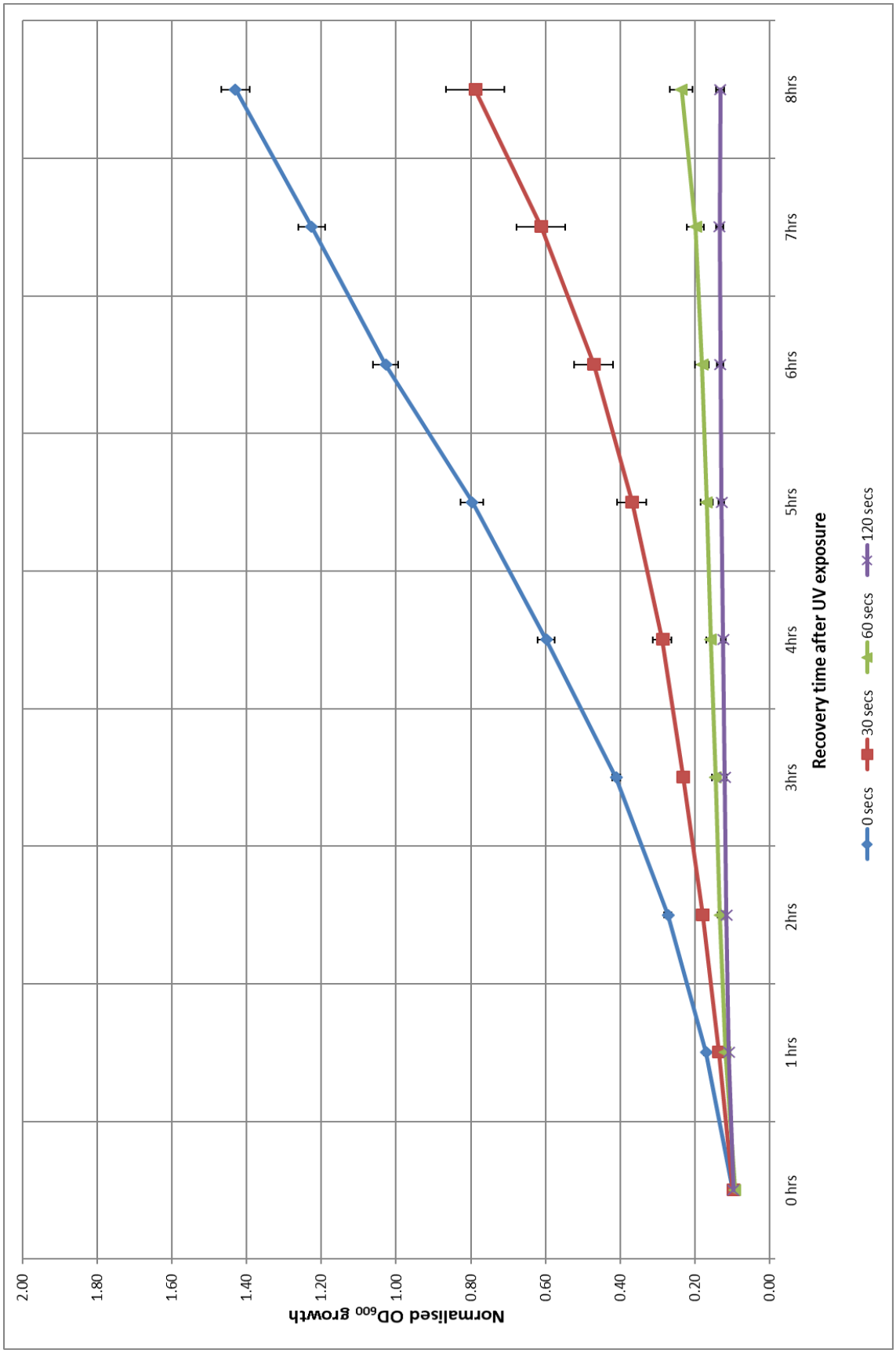


Figure 7.15: Normalised growth curve of combined *Pseudomonas* strains with a disrupted *rulB* gene following UVB exposure. Cell density via optical density at 600nm was measured over the course of eight hours. All of the OD_{600} growth values from 13 strains were normalised against zero UVB exposure with standard deviation error bars.

The normalisation of the data allowed a two-factor ANOVA with replication to be performed to analyse the results and observe if there were any significant differences between the amounts of UVB exposure on end point growth at eight hours and also if there was any difference in end point growth caused by the *rulB* gene being disrupted.

The ANOVA was carried out using the Minitab 17 statistical software (Table 7.4). The ANOVA took all of the growth values into consideration. The same number of strains for intact and disrupted *rulB* were used with a random number generator used to remove excess strains containing a disrupted *rulB*. The outcome was based on all of the values for exposure levels and *rulB* variation whilst comparing the two for any cross over interaction between them.

Table 7.4: Two-factor ANOVA of all strains tested for UVB tolerance with replication. This output shows the P values for any significant differences between different UVB exposure levels, differences between the intact and disrupted *rulB* genes and also the P value for the interaction between the two. Output from Minitab 17.

Anova: Two-Factor With Replication				
ANOVA				
Source of Variation	*SS*	*df*	*MS*	*P-value*
Diff. exposures	67.89258884	55	1.234410706	8.4675E-81
Intact vs. Disrupted *rulAB*	0.272304018	1	0.272304018	4.8465E-07
Interaction	5.668820982	55	0.103069472	5.31168E-26
Within	1.06805	112	0.009536161	
Total	74.90176384	223		

Using the ANOVA results the different exposure levels, 0, 30, 60 and 120 seconds were shown to produce OD_{600} growth values that were significantly different from each other. The P value for exposure level was $p=8.47 \times 10^{-81}$ which is less than 0.001, therefore it was concluded that at the 99.9% confidence

interval (CI) there was enough evidence to suggest that changing UVB exposure levels corresponds to a significant difference in end point growth at eight hours.

The ANOVA results also showed that the difference in *rulB,* whether it is intact or disrupted, does cause a significant difference in end point growth at eight hours. For intact versus disrupted *rulB* $p=4.85 \times 10^{-7}$ which was less than 0.001 meaning that it was concluded that at the 99.9% CI the form of *rulB* corresponds to a significant difference in end point growth at eight hours following UVB exposure.

The same result was obtained for the interaction between exposure level and *rulB* disruption with a P value of $p=5.31 \times 10^{-26}$ meaning that there was enough evidence to suggest that the interaction of the exposure level and the form of *rulB* corresponds to a significant difference in growth at the 99.9% CI. Strains with a disruption to *rulB* are at a significant disadvantage in terms of growth at eight hours after UVB irradiation. Strains with an intact *rulB* gene appear to grow better within the first eight hours of recovery following UVB irradiation.

7.3: Discussion

The *rulB* gene encodes a DNA polymerase belonging the gamma (ɣ) family polymerases (Hawver *et al.,* 2015) and is an error-prone DNA polymerase V. This polymerase V is a mutagenic DNA repair determinant (Zhang and Sundin, 2004) and has the ability to maintain gene expression following a stress event where the DNA is damaged and strand breaks occur such as those caused by UVB irradiation (Sundin and Weigand, 2007). UVB radiation is a major contributor to the formation of DNA replication stalls as direct UVB radiation causes the production of DNA photoproducts, such as pyrimidine dimers, that block DNA replication and RNA transcription which can be lethal to bacterial cells (Tark *et al.,* 2005). Previous studies have been conducted to assess the impact on cell growth and survival of *rulAB* or its' homologues in terms of the presence or absence of the gene (Cazorla *et al.,* 2008), but not whether the gene confers the same level of UVB tolerance if *rulB* is disrupted by an insert (Sundin and Murillo, 1999), in this case an ILE. The effect of the presence or absence of *rulAB* on UVB tolerance has been studied using *P. syringae.* The findings showed that *rulAB* positive strains had approximately a 10-fold increase in survival following UVB irradiation compared to strains lacking the *rulAB* operon. Research has also revealed that *rulAB* positive strains maintain a significantly larger epiphytic populations on the leaf surface (Sundin and Murillo, 1999). Research carried out with *P. fluorescens* also showed that when strains harbour the conjugative plasmid, pA506 they have a higher tolerance of UVB radiation. This was presumed to be due to the *rulAB* gene present on pA506 (Stockwell *et al.,* 2013).

To investigate whether the ILE disruption to *rulB* caused a significant difference to UVB tolerance and bacterial cell growth 13 strains with intact *rulB* genes and 15 strains with a disrupted *rulB* gene were subjected to UVB irradiation

for either 0, 30, 60 or 120 seconds with subsequent recovery time of 24 hours. To determine the effect of UVB irradiation the OD_{600} of each strain was measured at one hour intervals for the first eight hours and then again at 24 hours. Although *rulB* is disrupted it was not known if a functional protein was still produced from the disrupted *rulB* or if another error prone DNA polymerase is present in some of the strains.

The first set of results (Section 7.2.1) revealed that the strains with an intact *rulB* gene had their growth severely reduced during the first eight hours of recovery following UVB exposure presumably due to the DNA damaging nature of UVB radiation (Schuch *et al.,* 2017). However the effect of UVB damage was clearly linked to exposure time as the more UVB irradiation the greater the decline in growth compared to the no UVB irradiation control (Figure 7.4). This was likely due to the increased accumulation of DNA lesions and pyrimidine dimers that the bacteria was unable to repair leading to bacterial growth remaining stationary. This correlation between increasing UVB exposure time and decreasing bacterial cell growth has been shown before and is a dose dependant correlation (Charpentier *et al.,* 2011).

The effect of increasing UVB irradiation on bacterial growth was clearly shown. When subjected to 30 seconds of UVB exposure all of the strains regained a similar level of growth to the zero UVB controls after 24 hours of recovery. However the same was not true when the exposure was increased to 60 seconds (Figure 7.6) as four strains failed to regain the same level of growth and the growth of three strains remained stationary. The same pattern continued with 120 seconds of UVB exposure as only four strains increased in growth after 24 hours. A similar pattern was observed with the strains that contained a

disrupted *rulB* gene (Section 7.2.2). Increasing UVB exposure reduced the bacterial growth of all of the strains.

An interesting set of results was the normalised growth rates of the strains to observe any difference in overall growth rate between intact and disrupted *rulB.* Linear regression growth rates were constructed from the natural log of the OD_{600} values following normalisation. The growth rates in Figure 7.12 show that both *rulB* forms follow the same pattern; as UVB exposure increases, growth decreases and that as recovery time increases so does growth.

Statistical analysis on the growth rates for all of the exposure levels found that at the 95% CI, the form of *rulB* only has a significant impact on bacterial growth rate when the cells are exposed to UVB irradiation for 60 seconds. This may be due to the 30 second exposure level not being high enough to damage the DNA sufficiently to activate *rulB* expression. The 120 seconds exposure level appears to be the threshold for bacterial cell survival and may provide too much DNA damage for the cells to compensate. Future work could involve looking at exposure levels closer to 60 seconds to investigate if there is a cut-off point where disrupted *rulB* fails to cause a significant decrease in bacterial growth rate.

However the statistics on the normalised data (Figures 7.14 and 7.15) revealed that at the eight hour recovery time point there was enough evidence to suggest a significant difference in bacterial growth due to intact or disrupted *rulB*. This result was expected as over a relatively short period of time the strains with an intact *rulB* encoded DNA polymerase V may be able to bypass the UVB damaged DNA and continue DNA replication. The growth of the strains with a disrupted *rulB* stalls as the DNA is repaired using mechanisms such as photoreactivation, nonhomologous end-joining or homologous end-joining (Figure 7.16) (Shuman and Glickman, 2007; Alberts *et al.,* 2008), methyl-directed

mismatch repair (Fukui, 2010) or homologous recombination using the RecBCD pathway (Yu *et al.,* 1998), allowing DNA polymerase III to continue DNA replication. This also happens in the intact *rulB* strains but *rulB* provides another, more error prone solution to aid cellular growth following severe UVB induced DNA damage.

By 24 hours the majority of the DNA damage should have been resolved. DNA repair is estimated to take between 24 and 72 hours (Clancy, 2008) if the initial UVB irradiation did not kill the bacterial cells. This means that by 24 hours there may be no advantage to have a functioning *rulB* encoding the error prone DNA polymerase V.

Another point to consider was although other DNA repair mechanisms are present some strains may contain multiple copies of *rulB* or its homologues in a similar way to bacteria having multiple copies of housekeeping genes (Klappenbach *et al.,* 2000). This would give strains with multiple UVB tolerance genes or multiple DNA repair genes an advantage over other strains and would make the disputed *rulB* redundant. It may also be possible for a bacterium to utilise horizontal gene transfer as a mechanism for cells containing a disrupted *rulB* gene to acquire an intact version (Fall *et al.,* 2007).

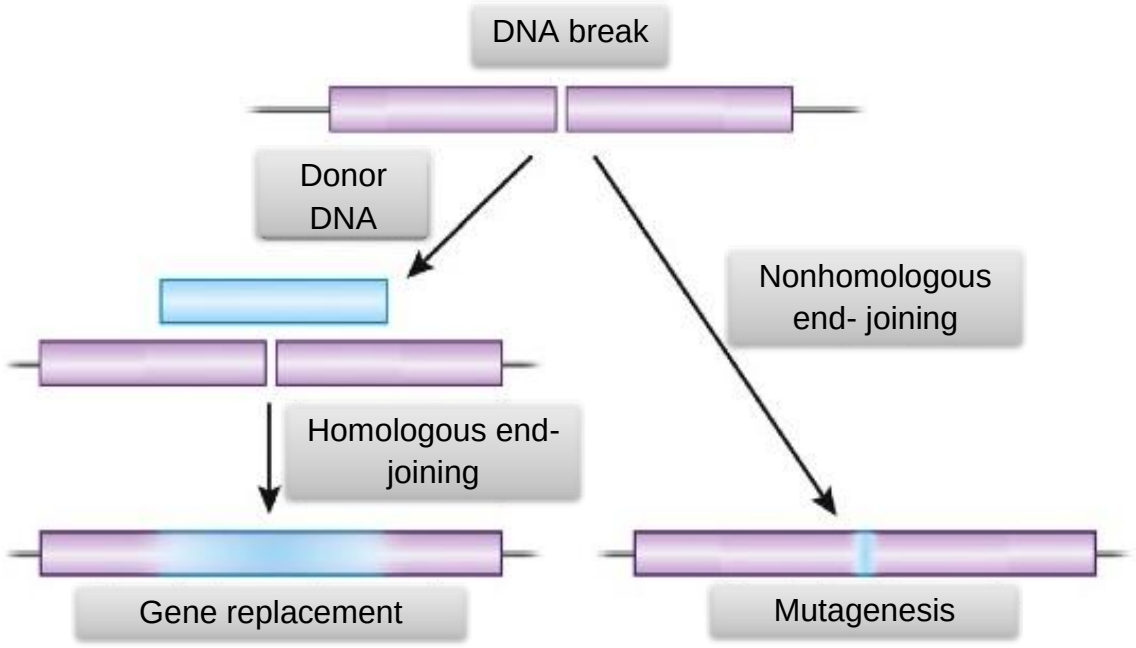


Figure 7.16: Nonhomologous and homologous end-joining to repair DNA lesions. The DNA break is caused by UVB irradiation and a section of DNA is lost. Two pathways are used to repair the break. The first is nonhomologous end-joining which joins the two ends back together causing alterations to the original DNA and may be deleterious to the bacterium. The second is homologous end-joining which uses a template DNA strand to copy. This method is far more specific but harder to accomplish. (Image obtained from: upload.wikimedia.org/wikipedia/commons/d/de/Repair_outcomes_of_a_genomi c _double-strand_break_for_ZFN_cleavage.jpg, used under the creative commons license 3.0; https://creativecommons.org/licenses/by/3.0/).

Although significant differences between intact *rulB* and disrupted *rulB* were observed at the end point growth analysis when looking at the bacterial growth rates the only significant difference between the two *rulB* versions was after 60 seconds of UVB exposure. All of the other exposure times indicated that there was no significant difference to bacterial growth rate influenced by a disrupted *rulB* gene caused by ILE insertion. This showed that ILE insertion into *rulB* may not be as detrimental to the bacteria as previously thought. However, bacterial growth rate is only one measure of overall bacterial fitness and not representative of the full impact a disrupted *rulB* gene may have on bacterial cells.

Chapter 8. General discussion and conclusions.

8.1: Discussion

The main aim of this research was to identify and characterise integron-like elements (ILEs) in a range of phytopathogenic *Pseudomonas syringae* strains and conduct further research into the ILEs' insertion into the DNA polymerase V encoding gene, *rulB.* Additional research aims included: characterising the genetic expression of the ILE integrase genes and the TTE genes present on the ILEs; investigating the relationship between ILE insertions into *rulB* and the ability of *rulB* to maintain its DNA damage repair capabilities and finally to assess what target the ILE requires for insertion into *rulB.* These aims, as set out in Section 1.11 were:

1- To determine the frequency of integron-like elements (ILEs) and *rulAB* genes in different plant pathogenic *Pseudomonas* bacteria.

2- To determine whether ILEs and ILE-captured genes are mobile.

3- To identify the conditions in which integron integrase and type three effectors (TTE) are expressed.

4- To characterise the ILE co-localisation with *rulB.*

5- To investigate if the disrupted RulB protein still conveys UV resistance.

The ILE research stemmed from previous work by Arnold *et al.* (2000) and Rhodes *et al.* (2014). Both identified sections of DNA that had inserted into and disrupted the DNA polymerase V encoding *rulB* gene crucial for DNA damage repair (Cazorla *et al.,* 2008). Rhodes *et al.* (2014) identified a chromosomal ILE from *P. fluorescens* FH1 that had inserted into a plasmid borne *rulB* gene on pWW0 whereas Arnold *et al.* (2000) identified an ILE in *P. syringae* pv. *pisi* 203, but no ILE movement was observed. There have been other instances of

genomic inserts into *rulB* homologues; for example integrative and conjugative elements (ICE) were identified in *Pantoea ananatis* and they also insert into the *umuC* gene of the *umuDC* operon which is a homologue of *rulAB* (Maayer *et al.,* 2015). It was hypothesised that ILEs and possible further genomic insertions into *rulB* homologues may be present in other pathovars and strains of *P. syringae* along with other bacterial species.

The first aim and the largest part of this research project was identifying and characterising potential ILEs from a range of *P. syringae* pathovars which included *pisi, syringae, tomato* and *glycinea.* This was primarily conducted via DNA hybridisation blots, but progressed onto PCR to identify ILEs and whole genome sequencing to characterise the content of the ILEs and the surrounding genome.

Using PCR to identify ILEs (Chapter 3) 164 *P. syringae* strains were screened for three distinct ILE associated regions. The first region was the *rulAB* gene as it appeared to be a hotspot for ILE insertion. The second was a *xerC* gene that appears to always be present at the 5' end of the ILE, although this position is not exclusive to ILEs, and the final region was the *rulB-xerC* junction region at the point of previously identified ILE insertion into *rulB.* The ILE screening began with DNA hybridisation using probes from *P. fluorescens.* This method worked well for the large sample size and was also able to detect the ILE regions even though many of the genes had slight but noticeable differences in their sequences to the probe. However PCR was used for a second screening as sequence analysis can be performed following PCR. The PCR primers could not use the same sequences as used for the hybridisation probes. This was due to the sequences not being homologous between the *Pseudomonas fluorescens* and *putida* strains used for the hybridisation probes and the *Pseudomonas*

syringae pathovars that were being screened. This was where the difference between DNA hybridisation and PCR was noticed. Hybridisation probes require >40%+ identity between them and the sample whereas PCR primers require >85% identity for amplification and the first 5 nucleotides need to be homologous to the target gene (Sommer and Tautz, 1989). Using PCR, 21 strains containing ILEs were identified from the 164 screened.

Following on from the identification of the ILEs they were characterised using a combination of PCR, using primers developed from *Ppi* 203 to identify ILEs that were the same as the *Ppi* 203 ILE, and whole genome sequencing to characterise ILEs that were different. The characterisation of the ILEs revealed that all of the 21 ILEs contained a 5' *xerC* integrase gene and TTE genes in the 3' variable end, although not all of the ILEs contained the same TTE genes (Chapter 4). For example *Ppi* 390 contained *avrPpiA1, Psy* 3023 contained *hopH1* and *hopAP1,* and *Pgy* 2411 contained *hopH1* and *hopC1.* TTEs play an important role in bacterial – plant interactions and therefore ILEs may also play a role in disseminating TTEs across multiple strains. ILEs may also have the ability to silence and excise certain TTEs if it is not beneficial to the bacterium or the plant may recognise the TTE and mount a HR. This could lead to ILEs containing TTEs contributing directly to the zig-zag theory of bacterial and plant co-evolution (Jones and Dangl, 2006) and act as a reservoir of TTEs for bacteria to pick and choose from. As a plant becomes resistant to a bacteria it can recognise a repertoire of TTEs. However bacteria have the ability to excise, silence or gain new TTEs that the plant host will no longer recognise. ILEs may be acting as a TTE reservoir facilitating disease progression.

Due to ILEs being potentially important to plant colonisation and plant disease symptoms, research was performed to assess how and when ILEs

become mobile and what their target site was within *rulB* and whether any other genes are required for ILE excision and insertion (Chapter 6). Previous studies by Rhodes *et al.* (2014) showed that the chromosomal ILEs present in *P. fluorescens* FH1-FH6 are mobile into the *rulB* gene present on the TOL plasmid pWW0, which is native to *P. putida.* The first stage to assess the mobility of ILEs was to repeat the experiment of Rhodes *et al.* (2014) and to also apply the same technique to the *P. syringae* strains containing ILEs. However, this did not work for the *P. syringae* ILEs due to pWW0 not inserting into the strains. The pWW0 plasmid is 117kb without an ILE insert and many studies using a plasmid above 100kb often require specialised competent cells and transformation via electroporation for successful transformation (Addgene, 2017). This may be the reason why the plasmid did not transfer into the *P. syringae* strains.

To overcome this *rulAB* from pWW0 and PPHGI-1, a genomic island from *P. syringae* pv. *phaseolicola* 1302A, were cloned into pBBR1MCS-2 to reduce the size of the plasmid and to also assess if any genes other than *rulAB* present on pWW0 are required for ILE insertion (see Section 6.2.5). The results were similar with ILE insertion into the cloned *rulB* only seen in *P. fluorescens* FH1. However, this did answer the question that only *rulAB* from pWW0 is required for ILE insertion in *P. fluorescens*. The other ILEs in *P. syringae* may require different conditions for movement or the ILEs may be fixed in the genome if they are beneficial. The ILE movement experiments progressed into looking at pWW0 and *P. fluorescens* FH1 in more detail to assess if the entire *rulAB* operon was required or if only *rulB* was required. The results indicated that the entire *rulAB* operon may be required for ILE movement in *P. fluorescens* FH1. The entirety of *rulAB* may be required as the ILE may use its regulatory mechanisms or its possible function as a recombinase gene. It has also been shown that *umuD*

(*rulA*) can regulate other SOS genes which may include self-regulation of the *umuDC* (*rulAB*) operon (Diaz-Magana *et al.,* 2015). The *rulB* gene from pWW0 whilst encoding a DNA polymerase V may also exhibit some form of recombinase activity facilitating ILE integration and possibly ILE circularisation. A similar interaction has been observed in the past as when *rulB* on PPHGI-1 is knocked out the genomic island loses the ability to form a circular intermediate (Lovell *et al.,* 2009). This was one of the reasons for carrying out circularisation tests on the ILEs. ILE circularisation tests were performed on *P. fluorescens* FH1, *Ppi* 203 and *Psy* 3023 with FH1 showing a positive result (Figure 6.13). However this area of research needs to be continued.

Analysing the expression of ILE genes may lead to a better understanding of when ILEs are active and also when they are more likely to be mobile. In order to observe ILE gene expression different conditions were chosen to stress the bacteria. The stresses included plant apoplastic fluid which mimics the bacteria's host environment, changes in temperature and also included two stresses, UVB irradiation and MMC, that would activate the SOS response as both treatments cause DNA damage (Krishna *et al.,* 2007; Charpentier *et al.,* 2011). Both UVB and MMC caused all of the ILE genes tested to be upregulated compared to their respective controls, except *rulB* which was only upregulated with the UVB treatment. This was an unexpected result as *rulB* would be expected to be upregulated during UVB exposure as *rulB* encodes a DNA polymerase V that has the ability to continue DNA synthesis following DNA damage and stalling of the replication fork (Tark *et al.,* 2005). The gene expression results also showed that unlike integrons the expression of the ILE integrase is uncoupled from the expression of *rulB* and *xerC* is unlikely to be under the regulation of the SOS

response. If *xerC* were under the same regulation as *rulB* it would be expected that a similar gene expression profile would be observed, this was not the case.

All of the identified ILEs are inserted into *rulB* and therefore disrupts the gene and causes a truncated protein product to be formed. Research was undertaken to assess if the disrupted *rulB* gene by an ILE was deleterious to bacterial survival following exposure to UVB irradiation. This research was important as *P. syringae* colonises the leaf surface which is exposed to high levels of UV radiation (Lindow and Brandl, 2003). Therefore it is important that the bacterial cells can withstand UV radiation and maintain a viable bacterial population. The disrupted *rulB* results were compared to strains with an intact *rulB* gene and the difference in bacterial survival between the two sets of strains was minimal with only one exposure level showing a slight but significant difference (Chapter 7). The lack of difference between the two groups tested could be due to the strains containing other genes that repair damaged DNA or even other copies of *rulB* elsewhere in the genome. Another possibility is that the truncated *rulB* gene is still active, but this requires further investigation.

8.2: Conclusions from research.

Conclusions can be drawn from this research on ILEs in *P. syringae* and their potential impact on bacterial evolution. The first conclusion to be drawn was that ILEs are abundant in *P. syringae* pathovars and that they all follow a similar genetic makeup. All of the ILEs identified contained a *xerC* integrase, at least one TTE gene and they were all inserted into *rulB*.

The second conclusion was that although all the ILEs were inserted into *rulB* the *P. syringae* ILEs appear to be not as mobile under the conditions tested as the *P. fluorescens* FH1 ILE. This was shown as the FH1 ILE moved into a

cloned version of *rulB* without any of the genes from pWW0 being present, however the entirety of *rulAB* may be required.

The expression analysis showed a variation in ILE integrase and TTE expression. One interesting result was that the ILE *xerC* integrase appears to not be under the regulation of the SOS response as differences in gene expression were observed between *xerC* and *rulB* (known to be under SOS regulation). The overall finding was that gene expression increases in times of stress. This is most likely because different beneficial genes are required to support the bacteria in stressful environments. These beneficial genes may be present in the environment and ILEs may be able to capture and express them. This also linked into the UVB irradiation work on the intact and disrupted *rulB* containing strains. This showed that although the ILE insertion disrupts the *rulB* gene and truncates the DNA polymerase V protein there was only a slight difference in growth between the two groups and this may indicate that other mechanisms of DNA damage repair are present to compensate for the loss of *rulB.*

8.3: Future research.

The current study left many questions unanswered. These questions include: do ILEs circularise during excision and integration; are the ILEs identified in *P. syringae* strains mobile as seen with the ILE in *P. fluorescens* FH1; will the ILEs act in a similar way to integrons and capture exogenous TTE genes to overcome plant resistance?

The ILE circularisation work has been started in this study, but requires further work with new primers and different stresses to observe any possible ILE circularisation. Future research on this could also include a time series

experiment to identify how long after excision an ILE circularises and whether or not the circularised form is always present in the cell.

The ILE mobility work has also been started but needs to be continued with the focus on the *P. syringae* ILEs and different versions of *rulAB*. This would also identify if these ILEs are true integrons. This question would be investigated by introducing various *rulB* genes from different bacterial strains and observing if ILE movement occurs during cell stress. The question would then be, is the fully functional *rulB* gene required or is it just a certain sequence site required for ILE insertion? This would be answered by cloning different versions of *rulB* to use as a capture target in the same way as used for the previous studies on *P. fluorescens* FH1 (Chapter 6).

The final question that is that will the ILEs behave the same as true integrons and capture exogenous genes from the environment to overcome plant resistance. It would be expected that a bacterial pathogen would undergo severe cell stress during plant interaction. This work could use the tomato pathogen, *Pto* DC3000, which encodes: (i) a potential ILE insertion site (*rulB* gene) and (ii) a suite of 28 key effectors for suppressing host resistance. *Pto* D28E lacks all 28 effectors compared to the wild-type and D28E is avirulent and unable to cause disease in tomato (Cunnac *et al.,* 2011). Two effectors *hopM1* and *avrPtoB* are solely able to convert this strain back to virulence. Therefore using D28E carrying *hopM1* in a series of experimental evolution experiments where sole *avrPtoB* and an ILE carrying *avrPtoB* are provided as genetic material for capture by the ILEs virulence on tomato will show that the ILE has captured the gene. This will encompass two mechanisms of ILEs, firstly will an ILE be incorporated to overcome host resistance and secondly will an existing ILE capture and express an effector?

References:

Addgene. (2017). *Plasmids 101: A desktop resource* [online]. 3rd ed. Cambridge, USA: Addgene. [Accessed 26/04/2017].

Alberts, B., Johnson, A., Lewis. J., Raff, M., Roberts, K. and Walter, P. (2002). DNA repair. In: *Molecular Biology of the Cell*. 4th edition. New York: Garland.

Alcamo, I.E. *DNA Technology: The Awesome Skill.* (1999). 2nd edition. New York: Harcourt Academic Press.

Aldon, D., Brito, B., Boucher, C. and Genin, S. (2000). A bacterial sensor of plant cell contact controls the transcriptional induction of *Ralstonia solanacearum* pathogenicity genes. *EMBO J.* 19(10): pp.2304-2314.

Alfano, J. R. and Collmer, A. (1997). The type III (Hrp) secretion pathway of plant pathogenic bacteria: trafficking harpins, Avr proteins, and death. *Journal of Bacteriology*. 179: pp.5655-5662.

Alfano, J. R., Charkowski, A.O., Deng, W.L., Petnicki, T., van Dijk, K. and Collmer, A. (2000). The *Pseudomonas syringae* Hrp pathogenicity island has a tripartite mosaic structure comprised of a cluster of Type III genes bounded by an exchangeable effector and conserved effector loci that contribute to parasitic fitness and pathogenicity in plants. *Proceedings of the National Academy of Sciences USA.* 97: pp.4856-4861.

Arnold, D.L., Brown, J., Jackson, R.W. and Vivian, A. (1999). A dispensable region of the chromosome which is associated with an avirulence gene in *Pseudomonas syringae* pv. *pisi*. *Microbiology*. 145: pp.135-141.

Arnold, D.L., Jackson, R.W. and Vivian, A. (2000). Evidence for the mobility of an avirulence gene, *avrPpiA1*, between the chromosome and plasmids of races of *Pseudomonas syringae* pv. *pisi. Molecular Plant Pathology.* 1(3): pp195-199.

Arnold, D.L., Jackson, R.W., Fillingham, A.J., Goss, S.C., Taylor, J.D., Mansfield, J.W. and Vivian, A. (2001). Highly conserved sequences flank avirulence genes: isolation of novel avirulence genes from *Pseudomonas syringae* pv*. pisi*. *Microbiology.* 147(5): pp. 1171-1182.

Arnold, D.L., Godfrey, S.A.C. and Jackson, R.W. (2009). *Pseudomonas syringae*: Genomics provides important insights to secretion systems, effector genes and the evolution of virulence. In: Jackson, R.W. (2009). *Plant Pathogenic Bacteria: Genomics and Molecular Biology.* Norfolk: Caister Academic Press. pp. 203-226.

Baharoglu, Z., Bikard, D. and Mazel, D. (2010). Conjugative DNA transfer induces the bacterial SOS response and promotes antibiotic resistance development through integron activation. *PLoS: Genetics.* 6(10).

Baharoglu, Z., Krin, E. and Mazel, D. (2012). Connecting environment and genome plasticity in the characterization of transformation-induced SOS regulation and carbon catabolite control of the *Vibrio cholerae* integron integrase. *Journal of Bacteriology* 194(7): pp.1659–1667.

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

Bardaji, L., Echeverria, M., Rodriguez-Palenzuela, P., Martinez-Garcia, P.M. and Murillo, J. (2017). Four gene essential for recombination define GInts, a new type of mobile genetic island widespread in bacteria. *Scientific Reports.* 10(7).

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

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

Bigeard, J., Colcombet, J. and Hirt, H. (2015). Signaling mechanisms in pattern-triggered immunity (PTI). *Molecular Plant.* 8: pp. 521-539.

Blakely, G., Colloms, S., May, G., Burke, M. and Sherratt, D. (1991). *Escherichia coli* XerC recombinase is required for chromosomal segregation at cell division. *The New Biologist*. 3(8): pp.789-798.

Boch, J., Joardar, V., Gao, L., Robertson, T.L., Lim, M. and Kunkel, B.N. (2002). Identification of *Pseudomonas syringae* pv. *tomato* genes induced during infection of *Arabidopsis thaliana. Molecular Microbiology*. 44: pp.73-88.

Bodey, G.P., Bolivar, R., Fainstein, V. and Jadeja, L. (1983). Infections caused by *Pseudomonas aeruginosa*. *Reviews of Infectious Diseases.* 5(2): pp.279-313.

Bradbury, J.F. (1986). *Guide to Plant Pathogenic Bacteria*. Kew, UK: International Mycological Institute.

Brent, R. and Ptashne, M. (1981). Mechanism of action of the *lexA* gene product. *Proceedings of the National Academy of Sciences USA.* 78(7): pp.4204-4208.

Bridges, B.A. (2005). Error-prone DNA repair and translesion synthesis: focus on the replication fork. *DNA Repair*. 4(5): pp.618-9.

Brown, T.A. (2006). *Gene cloning and DNA analysis: An introduction.* 5th ed. Oxford: Blackwell publishing.

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

Cambray, G., Sanchez-Alberola, N., Campoy, S., Guerin, E., Da Re, S., Gonzalez-Zorn, B., Ploy, M.C., Barbe, J., Mazel, D. and Erill, I. (2011) Prevalence of SOS-mediated control of integron integrase expression as an adaptive trait of chromosomal and mobile integrons. *Mobile DNA*. 2(6).

Cazorla, F.M., Codina, J.C., Abad, C., Arrebola, E., Tores, J.A., Murillo, J., Perez-Garcia, A. and de Vicente, A. (2008). 62-kb plasmids harboring *rulAB* homologues confer UVB-tolerance and epiphytic fitness to *Pseudomonas syringae* pv. *syringae* mango isolates. *Microbial Ecology.* 56(2): pp.283-291.

Champoux, J.J. (2001). DNA topoisomerases: structure, function, and mechanism*. Annual Review of Biochemistry*. 70: pp.369–413.

Charpentier, X., Kay, E., Schneider, D. and Shuman, H.A. (2011). Antibiotics and UVB radiation induce competence for natural transformation in *Legionella pneumophila. Journal of Bacteriology.* 193(5): pp.1114-1121.

Chen, Z., Agnew, J.L., Cohen, J.D., He, P., Shan, L., Sheen, J. and Kunkel, B.N. (2007). *Pseudomonas syringae* type III effector AvrRpt2 alters *Arabidopsis thaliana* auxin physiology. *Proceedings of the National Academy of Sciences USA.* 104(50): pp.20131-20136.

Clancy, S. (2008). DNA Damage & Repair: Mechanisms for Maintaining DNA Integrity. *Nucleic Acid Structure and Function: Nature Education.* 1(1): pp.103.

Clark, C.A., Purins, L., Kaewrakon, P., Focareta, T. and Manning, P,A. (2000). The *Vibrio cholerae* O1 chromosomal integron. *Microbiology.* 146(10): pp.2605-2612.

Coenye, T. and Vandamme, P. (2003). Diversity and significance of *Burkholderia* species occupying diverse ecological niches. *Environmental Microbiology.* 5: pp.719-729.

Coleman, N.V. and Holmes, A.J. (2005). The native *Pseudomonas stutzeri* strain Q chromosomal integron can capture and express cassette-associated genes. *Microbiology.* 151(6): pp.1853-1864.

Courcelle, J., Khodursky, A., Peter, B., Brown, P.O. and Hanawalt, P.C. (2001). Comparative gene expression profiles following UV exposure in wild type and SOS-deficient *Escherichia coli. Genetics.* 158: pp.41-64.

Cunnac, S., Lindeberg, M. and Collmer, A. (2009). *Pseudomonas syringae* type III secretion system effectors: repertoires in search of functions. *Current Opinions in Microbiology.* 12: pp.53-60.

Cunnac, S., Chakravarthy, S., Kvitko, B.H., Russell, A.B., Martin, G.B. and Collmer, A. (2011). Genetic disassembly and combinatorial reassembly identify a minimal functional repertoire of type III effectors in *Pseudomonas syringae*. *Proceedings of the National Academy of Science USA.* 108(7): [E-publication]

Dallas-Yang, Q., Jiang, G. and Sladek, F.M. (1998). Digestion of terminal restriction endonuclease recognition sites on PCR products. *Benchmarks; Biotechniques.* 24(4): pp.582-584.

Daniel, R.M., Dines, M. and Petach, H.H. (1996). The denaturation and degradation of stable enzymes at high temperatures. *Biochemical Journal.* 317: pp.1-11.

De Torres-Zabala, M., Truman, W., Bennett, M.H., Lafforgue, G., Mansfield, J.W., Rodrigues Egea, P., Bogre, L. and Grant, M. (2007). *Pseudomonas syringae* pv. *tomato* hijacks the *Arabidopsis* abscisic acid signalling pathway to cause disease. *EMBO J.* 7: pp.1434-1443.

Denny, T.P., Gilmour, M.N. and Selander, R.K. (1988). Genetic Diversity and Relationships of two pathovars of *Pseudomonas syringae. Microbiology.* 134(7): pp.1949-1960.

Diaz-Magana, A., Alva-Murillo, N., Chavez-Moctezuma, M.P., Lopez-Meza, J.E., Ramirez-Diaz, M.I. and Cervantes, C. (2015). A plasmid-encoded UmuD homologue regulates expression of *Pseudomonas aeruginosa* SOS genes. *Microbiology.* 161(7): pp.1516-1523.

Ditta, G., Stanfied, S., Corbin, D. and Helinski, D.R. (1980). Broad host range DNA cloning system for Gram-negative bacteria: Construction of a gene bank of *Rhizobium meliloti. Proceedings of the National Academy of Science USA.* 77(12): pp. 7347-7351.

Dodds, P.N. and Rathjen, J.P. (2010). Plant immunity: towards an integrated view of plant–pathogen interactions. *Nature.* 11: pp. 539-548.

Domingues, S., da Silva, G.J. and Nielsen, K.M. (2012). Integrons: Vehicles and pathways for horizontal dissemination in bacteria. *Mobile Genetic Elements.* 2(5): pp.211-223.

Dorman, C.J. (1991). DNA supercoiling and environmental regulation of gene expression in pathogenic bacteria. *Infection and Immunity.* 59(3): pp.745-749.

Ebrahim, S., Usha, K. and Singh, B. (2011). Pathogenesis Related (PR) proteins in plant defense mechanism. *Science against microbial pathogens.* pp.1043-1054.

Erbs, G. and Newman, M.A. (2009). MAMPs/PAMPs-Elicitors of Innate Immunity in Plants. In: Jackson R.W. (2009). *Plant Pathogenic Bacteria: Genomics and Molecular Biology.* Norfolk: Caister Academic Press, pp.227-240.

Escobar, M.A. and Dandekar. A.M. (2003). *Agrobacterium tumefaciens* as an agent of disease. *Trends in Plant Science.* 8(8): pp.380-86.

Fall, S., Mercier, A., Bertolla, F., Calteau, A., Gueguen, L., Perriere, G., Vogel, T.M. and Simonet, P. (2007). Horizontal gene transfer regulation in bacteria as a "spandrel" of DNA repair mechanisms. *PloS One*. 2(10): pp. 1055.

Farrugia, D.N., Elbourne, L.D., Mabbutt, B.C. and Paulsen, I.T. (2015). A novel family of integrases associated with prophages and genomic islands integrated within the tRNA-dihydrouridine synthase A (dusA) gene. *Nucleic Acids Research*. 43(9): pp.4547-4557.

Felix, G., Duran, J.D., Volko, S. and Boller, T. (1999). Plants have a sensitive perception system for the most conserved domain of bacterial flagellin. *The Plant Journal.* 18(3): pp. 265-276.

Fritig, B. and Legrand, M. (1993). *Mechanisms of Plant Defence Responses.* Holland: Kluwer Academic Publishers.

Fry, W.E. (1982). *Principles of Plant Disease Management*. New York: Academic Press.

Fu, Z.Q., Guo, M., Jeong, B., Tin, F., Elthon, T.E., Cerny, R.L., Staiger, D. and Alfano, J.R. (2007). A type III effector ADP-ribosylates RNA-binding proteins and quells plant immunity. *Nature.* 447: pp.284-289.

Fukui, K. (2010). DNA mismatch repair in eukaryotes and bacteria. *Journal of Nucleic Acids.* 2010.

Garnier, M., Jagoueix-Eveillard, S., Cronje, P.R., Le Roux, H.F. and Bové, J.M. (2000). Genomic characterization of a liberibacter present in an ornamental rutaceous tree, *Calodendrum capense*, in the Western Cape Province of South Africa. Proposal of *'Candidatus Liberibacter africanus* subsp. *capensis'*. *International Journal of Systematic and Evolutionary Microbiology.* 50(6): pp.2119-2125.

Gerlach, R.G. and Hensel, M. (2007). Protein secretion systems and adhesins: The molecular armory of Gram-negative pathogens. *International Journal of Medical Microbiology.* 297(6): pp.401-415.

Gillings, M.R. (2014). Integrons: past, present, and future. *Microbiology and Molecular Biology Reviews.* 78(2): pp.257-277.

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

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

Goosen, N. and Moolenaar, G.F. (2008). Repair of UVB damage in bacteria. *DNA Repair.* 7(3): pp.353-379.

Greated, A., Lambertsen, L., Williams, P.A. and Thomas, C.M. (2002). Complete sequence of the IncP-9 TOL plasmid pWW0 from *Pseudomonas putida*. *Environmental Microbiology*. 4 (12): pp.856-871.

Green, S., Studholme, D.J., Laue, B.E., Dorati, F., Lovell, H., Arnold, D., Cottrell, J.E., Bridgett, S., Blaxter, M., Huitema, E., Thwaites, R., Sharp, P.M., Jackson, R.W. and Kamoun, S. (2010). Comparative Genome Analysis Provides Insights into the Evolution and Adaptation of *Pseudomonas syringae pv. aesculi* on *Aesculus hippocastanum. PLoS ONE.* 5(4).

Greenburg, J.T. and Vinatzer, B.A. (2003). Identifying type 3 effectors of plant pathogens and analysing their interaction with plant cells. *Current Opinion in Microbiology.* 6: pp.20-28.

Grohmann, E. (2013). Mating cell-cell channels in conjugating bacteria. In: Grohmann, E. (2013). *Madame Curie Bioscience Database.* Texas: Landes Bioscience.

Guerin, E., Cambray, G., Sanchez-Alberola, N., Campoy, S., Erill, I., Da Re, S., Gonzalez-Zorn, B., Barbé, J., Ploy, M.C. and Mazel, D. (2009). The SOS response controls integron recombination. *Science.* 324: pp.1034.

Gurevich, A., Saveliev, V., Vyahhi, N. and Tesler, G. (2013). QUAST: quality assessment tool for genome assemblies. *Bioinformatics* 29(8): pp. 1072-1075.

Gyohda, A., Funayama, N. and Komano, T. (1997). Analysis of DNA inversions in the shufflon of plasmid R64. *Journal of Bacteriology*. 179(6): pp.1867-1871.

Haas, D. and Keel, C. (2003). Regulation of antibiotic production in root-colonizing *Pseudomonas* spp. and relevance for biological control of plant disease. *Annual Review of Phytopathology.* 41: pp.117-153.

Hall, R.M. and Collis, C.M. (1995). Mobile gene cassettes and integrons: capture and spread of genes by site-specific recombination. *Molecular Microbiology.* 15(4): pp.593-600.

Hausladen, A. and Stamler, J.S. (1998). Nitric oxide in plant immunity. *Proceedings of the National Academy of Science USA.* 95(18): pp. 10345-10347.

Hawver, L.A., Tehrani, M., Antczak, N.M., Kania, D., Muser, S., Sefcikova, J. and Beuning, P.J. (2015). Point mutations in *Escherichia coli* DNA *pol V* that confer resistance to non-cognate DNA damage also alter protein-protein interactions. *Mutation Research*. 780: pp.1-14.

Helene, C., Charlier, M., Montenay-Garestier, T. and Laustriat, G. (2012). *Trends in Photobiology.* London: Plenum Publishing.

Hirano S.S., Rouse D.I., Clayton M.K. and Upper C.D. (1995). *Pseudomonas syringae* pv. *syringae* and bacterial brown spot of bean: a study of epiphytic phytopathogenic bacteria and associated disease. *Plant Disease* 79: pp.1085-1093.

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

Hochhut, B., Lotfi, Y., Mazel, D., Faruque, S.M., Woodgate, R. and Waldor, M.K. (2001). Molecular analysis of antibiotic resistance gene clusters in *Vibrio cholerae* O139 and O1 SXT constins. *Antimicrobial Agents and Chemotherapy*. 45(11): pp.2991-3000.

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

Hurme, R. and Rhen, M. (1998). Temperature sensing in bacterial gene regulation – what it all boils down to. *Molecular Microbiology.* 30(1): pp.1-6.

Hutchison, M.L., Tester, M.A. and Gross, D.C. (1995). Role of biosurfactant and ion channel-forming activities of syringomycin in transmembrane ion flux: a model for the mechanism of action in the plant-pathogen interaction. *Molecular Plant-microbe Interactions.* 8(4): pp.610-620.

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

Jackson, R.W. (2009). *Plant Pathogenic Bacteria*. Norfolk: Caister Academic Press.

Jackson, R.W., Vinatzer, B., Arnold, D.L., Dorus, S. and Murillo, J. (2011). The influence of the accessory genome on bacterial pathogen evolution. *Mobile Genetic Elements*. 1(1): pp.55-65.

Jackson, R.W., Vinatzer, B., Arnold,D.L., Dorus, S. and Murillo, J. (2011). The influence of the accessory genome on bacterial pathogen evolution. *Mobile Genetic Elements.* 1(1): pp.55-65.

Jin, Q., Thilmony, R., Zwiesler-Vollick, J. and He, S.Y. (2003). Type III protein secretion in *Pseudomonas syringae*. *Microbes and Infection.* 5(4): pp.301-310.

Jones, J.D.G. and Dangl, J. L. (2006). The plant immune system. *Nature*. 444: pp. 323-329.

Jove, T., Da Re, S., Denis, F., Mazel, D. and Ploy, M.C. (2010). Inverse correlation between promoter strength and excision activity in class 1 integrons. *PLoS Genetics*. 6(1).

Kim, J.J. and Sundin, G.W. (2000). Regulation of the rulAB mutagenic DNA repair operon of *Pseudomonas syringae* by UV-B (290 to 320 nanometers) radiation and analysis of rulAB-mediated mutability *in vitro* and *in planta*. *Journal of Bacteriology*. 182(21): pp.6137-6144.

Klappenbach, J.A., Dunbar, J.M. and Schmidt, T.M. (2000). rRNA operon copy number reflects ecological strategies of bacteria. *Applied and Environmental Microbiology*. 66(4): pp.1328-1333.

Kovach, M.E., Elzer, P.H., Hill, D.S., Robertson, G.T., Farris, M.A., Roop, R.M. and Peterson, K.M. (1995). Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. *Gene.* 166(1): pp.175-176.

Krishna, S., Maslov, S. and Sneppen, K. (2007). UVB-induced mutagenesis in *Escherichia coli* SOS response: a quantitative model. *PloS: Computational Biology.* 3(3): pp. 41.

Kvitko, B.H., Ramos, A.R., Morello, J.E., Oh, H.S. and Collmer, A.J. (2007). Identification of harpins in *Pseudomonas syringae* pv. *tomato* DC3000, which are functionally similar to HrpK1 in promoting translocation of typre 3 secretion system effectors. *Journal of Bacteriology.* 189: pp.8059-8072.

Lévesque, C., Brassard, S., Lapointe, J. and Roy. P.H. (1994). Diversity and relative strength of tandem promoters for the antibiotic-resistance genes of several integrons. *Gene.* 142(1): pp.49-54.

Li, Q.Q., Feng, J.X., Tang, J.L., Lin, W., Duan, C.J., Ye, Y.F. and Luo, K. (2005). *Siraita grosvenorii* is a new host of *Ralstonia solanacearum* in China. *Plant Pathology.* 54(6): pp. 811.

Lindow, S.E., Arny, D.C. and Upper, C.D. (1978). Distribution of ice nucleation-active bacteria on plants in nature. *Applied and Environmental Microbiology.* 36(6): pp. 831-838.

Lindow, S.E. and Brandl, M.T. (2003). Microbiology of the Phyllosphere. *Applied and Environmental Microbiology.* 69(4): pp. 1875-1883.

Lledo, J.I.L. and Lynch, M. (2009). Evolution of mutation rates: Phylogenomic analysis of the photolyase/ cryptochrome family. *Molecular Biology and Evolution.* 26(5): pp.1143-1153.

López-Millán, A.F., Morales, F., Abadía, A. and Abadía, J. (2001). Iron deficiency-associated changes in the composition of the leaf apoplastic fluid from field-grown pear (*Pyrus communis*) trees. *Journal of experimental botany.* 52(360); pp.1489-1498.

Lovell, H.C., Mansfield, J.W., Godfrey, S.A.C., Jackson, J.W., Hancock, J.T. and Arnold, D.A. (2009). Bacterial evolution by genomic island transfer occurs via DNA transformation *in planta. Current biology.* 19(18): pp.1586-1590.

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

Luna, E., Pastor, V., Robert, J., Flors, V., Mauch-Mani, B. and Ton, J. (2010). Callose Deposition: A Multifaceted Plant Defense Response. *Molecular Plant-microbe Interactions.* 24(2): pp.183-193.

Maayer, P., Chan, W.Y., Martin, D.A., Blom, J., Venter, S.N., Duffy, B., Cowan, D.A., Smits, T.H. and Coutinho, T.A. (2015) Integrative conjugative elements of the ICEPan family play a potential role in *Pantoea ananatis* ecological diversification and antibiosis. *Frontiers in Microbiology*. 6: pp.576.

Madigan, M., Martinko, J., Stahl, D. and Clark, D. (2012). *Biology of Microorganisms.* 13th ed. San Francisco: Pearson.

Maloy, O.C. (2005). Plant Disease Management. In: The *Plant Health Instructor.* Washington: Online Article.

Mansfield, J., Genin, S., Magori, S., Citovsky, V., Sriariyanum, M., Ronald, P., Dow, M., Verdier, V., Beer, S.V., Machado, M.A., Toth, I., Salmond, G. and Foster, G.D. (2012). Top 10 plant pathogenic bacteria in molecular plant pathology. *Molecular Plant Pathology.* 13(6): pp. 614-629.

Manoil, C. (2000). Tagging exported proteins using *Escherichia coli* alkaline phosphatase gene fusions. *Methods in Enzymology*. 326: pp.35-47.

Martens, M., Dawyndt, P., Coopman, R., Gillis, M., De Vos, P. and Willems, A. (2008). Advantages of multilocus sequence analysis for taxonomic studies: a case study using 10 housekeeping genes in the genus *Ensifer* (including former *Sinorhizobium*). *International Journal of Systematic and Evolutionary Microbiology*. 58(Pt 1): pp.200-214.

Melotto, M., Underwood, W. and He, S.Y. (2008). Role of stomata in plant innate immunity and foliar bacterial diseases. *Annual Reviews in Phytopathology.* 46: pp.101-122.

Meng, W., Belyaeva, T., Savery, N.J., Busby, S.J., Ross, W.E., Gaal, T., Gourse, R.L. and Thomas, M.S. (2001). UP element-dependent transcription at the *Escherichia coli* rrnB P1 promoter: positional requirements and role of the RNA polymerase alpha subunit linker. *Nucleic Acids Research.* 29(20): pp.4166-4178.

Mohr, T.J., Lui, H., Yan, S., Morris, C.E., Castillo, J.A., Jelenska, J. and Vinatzer, B.A. (2008).Naturally occurring non-pathogenic isolates of the plant pathogen *Pseudomonas syringae* lack a Type III secretion system and effector gene orthologues. *Journal of Bacteriology.* 190(8): pp.2858-2870.

Musrati, R.A., Kollárová, M., Mernik, N. and Mikulásová, D. (1998). Malate dehydrogenase: distribution, function and properties. *General Physiology and Biophysics.* 17(3): pp.193-210.

Neale, H.C., Laister, R., Payne, J., Preston, G., Jackson, R.W. and 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. *Environmental Microbiology*. 18(11): pp.4144-4152.

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

Nimchuk, Z.L., Fisher, E.J., Desveaux, D., Chang, J.H. and Dangl, J.L. (2007). The HopX (AvrPphE) family of *Pseudomonas syringae* type III effectors require a catalytic triad and a novel N-terminal domain for function. *Molecular Plant-Microbe Interactions.* 20(4): pp.346-357.

Nomura, K., Melotto, M. and He, S.Y. (2005). Supression of host defense in compatible plant-*Pseudomonas syringae* interactions. *Current Opinion in Plant Biology.* 8: pp.361-638.

Nürnberger, T., Brunner, F., Kemmerling, B. and Piater, L. (2004). Innate immunity in plants and animals: striking similarities and obvious differences. *Immunology Reveiws.* 198: pp. 249-266.

Ohmori, H., Friedberg, E.C., Fuchs, R.P., Goodman, M.F., Hanaoka, F., Hinkle, D., Kunkel, T.A., Lawrence, C.W., Livneh, Z., Nohmi, T., Prakash, L., Prakash, S., Todo, T., Walker, G.C., Wang, Z. and Woodgate, R. (2001). The Y-family of DNA polymerases. *Molecular Cell*. 8(1): pp.7-8.

O'Leary, B.M., Rico, A., McCraw, S., Fones, H.N. and Preston, G.M. (2014). The infiltration-centrifugation technique for extraction of apoplastic fluid from plant leaves using *Phaseolus vulgaris* as an example. *Journal of Visualized Experiments: JoVE*. (94).

O'Leary, B.M., Neale, H.C., Geilfus, C.M., Jackson, R.W., Arnold, D.L. and Preston, G.M. (2016). Early changes in apoplast composition associated with defence and disease in interactions between *Phaseolus vulgaris* and the halo blight pathogen *Pseudomonas syringae* pv. *phaseolicola*. *Plant, cell and environment.* 39(10): pp. 2172-2184.

Özen, A.I. and Ussery, D.W. (2012). Defining the *Pseudomonas* Genus: Where do we draw the line with Azotobacter? *Microbial Ecology.* 63(2): pp.239-248.

Pacific Bioscienes, (2017). *PacBio.* Available from: http://www.pacb.com/ [Accessed 07/02/2017].

Pai, S., Bedford, L., Ruramayi, R., Aliyu, S.H., Sule, J., Maslin, D. and Enoch D.A. (2016). *Pseudomonas aeruginosa* meningitis/ventriculitis in a UK tertiary referral hospital. *Quarterly Journal of Medicine.* 109(2). [E-publication].

Palleroni, N.J. (1984). Genus I. *Pseudomonas migula*. *Bergey's Manual of Systematic Bacteriology*, 1: pp. 141–199.

Palleroni, N.J. (1992). Introduction to the Pseudomonadaceae. In: *The Prokaryotes, A Handbook on the Biology of Bacteria,Ecophysiology, Isolation, Identification and Applications, Vol. III*. 2nd ed. New York: Springer, pp. 3071–3085.

Partridge, S.R., Brown, H.J. and Hall, R.M. (2002). Characterization and movement of the class 1 integron known as Tn*2521* and Tn*1405. Antimicrobial Agents and Chemotherapy.* 46(5): pp.1288-1294.

Pavlidis, P. and Noble W.S. (2003) Matrix2png: A Utility for Visualizing Matrix Data. *Bioinformatics.* 19: pp.295-296.

Pickup, R. (1989) Related plasmids found in an English Lake District stream. *Microbial Ecology*. 18: pp.211–220.

Pitman, A.R., Jackson, R.W., Mansfield, J.W., Kaitell, V., Thwaites, R. and Arnold, D.L. (2005). Exposure to host resistance mechanisms drives evolution of bacterial virulence in plants. *Current Biology.* 15: pp.2230-2235.

Ploy, M.C., Lambert, T., Couty, J.P. and Denis, F. (2000). Integrons: an antibiotic resistance gene capture and expression system. *Clinical Chemistry and Laboratory Medicine.* 38(6): pp.483-487.

Preston, G.M., Studholme, D.J. and Caldelari, I. (2005). Profiling the secretomes of plant pathogenic Proteobacteria. *FEMS Microbiology Review.* 29(2): pp.331-360.

Ramos, J.L. (2004). *Virulence and Gene Regulation.* UK: Springer Publishing.

Rhodes, G., Bosma, H., Studholme, D., Arnold, D.,L., Jackson, R.,W. and Pickup, R.,W. (2014). The rulB gene of plasmid pWW0 is a hotspot for the site-specific insertion of integron-like elements found in the chromosomes of environmental *Pseudomonas fluorescens* group bacteria. *Environmental Microbiology*. 16(8): pp.2374.

Rico, A. and Preston, G.M. (2008). *Pseudomonas syringae* pv. *tomato* DC3000 uses constitutive and apoplast-induced nutrient assimilation pathways to catabolize nutrients that are abundant in the tomato apoplast. *Molecular Plant Microbe Interactions*. 21(2): pp.269-282.

Rico, A., Jones, R. and Preston, G.M. (2009). Adaption to the plant apoplast by plant pathogenic bacteria. In: Jackson, R.W. (2009). *Plant Pathogenic Bacteria*. UK: Caister Academic Press. pp.63-89.

Roberts, E.L., Burguieres, S. and Warner, I. (1998). Spectroscopic studies of indigo carmine dye in organized media. *Applied Spectroscopy.* 52(10): pp.1305-1313.

Roche Applied Sciences (2014) *Introduction to the DIG system and the DIG application manual.* Penzberg, Germany: Roche

Rouxel, T. and Balesdent, M.H. (2010). Avirulence Genes. *Encyclopedia of life sciences.* [E-publication].

Rowe-Magnus, D.A., Guerout, A.M. and Mazel, D. (1999). Super-integrons. *Research in Microbiology*. 150(9-10): pp.641-651.

Rutherford, K., Parkhill, J., Crook, J., Horsnell, T., Rice, P., Rajandream, M.A. and Barrell, B. (2000). *Bioinformatics* 16(10): pp.944-5.

Sarkar, S.F., Gordon, J.S., Martin, G.B. and Guttman, D.S. (2006). Comparative genomics of host-specific virulence in *Pseudomonas syringae. Genetics.* 174(2): pp.1041-1056.

Schaad, N.W., Jones, J.B. and Chun, W. (2001). *Laboratory guide for the identification of plant pathogenic bacteria (3th ed.)*. Paul, Minnesota, USA: Phytopathological Society St.

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

Schuch, A.P., Moreno, N.C., Schuch, N.J., Menck, C.F. and Garcia, C.C. (2017). Sunlight damage to cellular DNA: Focus on oxidatively generated lesions. *Free Radical Biology and Medicine*. [E-publication ahead of print].

Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation*. Bioinformatics* 30(14): pp. 2068-9.

Sheng, Y., Mancino, V. and Birren, B. (1995). Transformation of *Escherichia coli* with large DNA molecules by electroporation. *Nucleic Acids Research.* 23(11): pp. 1990-1996.

Shenge, K.C., Mabagala, R.B., Mortensen, C.N., Stephan, D. and Wydra, K. (2007). First Report of Bacterial Speck of Tomato Caused by *Pseudomonas syringae* pv. *tomato* in Tanzania. *American Phytopathological Society.* 91(4): pp. 462.

Shuman, S. and Glickman, M.S. (2007). Bacterial DNA repair by non-homologous end joining. *Nature Reviews: Microbiology.* 5(11): pp.852-861.

Silby, M.W., Winstanley, C., Godfrey, S.A.C., Levy, S.B. and Jackson, R.W. (2011). *Pseudomonas* genomes: diverse and adaptable. *Federation of European Microbiological Societies.* 35: pp. 652-680.

Smith, E.F. (1905). *Bacteria in relation to plant diseases.* Washington D.C.: Carnegie institution of Washington.

Smith, B.T. and Walker, G.C. (1998). Mutagenesis and more: *umuDC* and the *Escherichia coli* SOS response. *Genetics*. 148(4): pp.1599-1610.

Sommer, R. and Tautz, D. (1989). Minimal homology requirements for PCR primers. *Nucleic Acids Research*. 17(16): pp.6749.

Staskawicz, B.J. and Panopoulos, N.J. (1980). Phaseolotoxin transport in *Escherichia coli* and *Salmonella typhimurium* via the oligopeptide permease. *Journal of Bacteriology*. 142: pp. 474-479.

Staskawicz, B.J., Dahlbeck, D. and Keen, N.T. (1984). Cloned avirulence gene of *Pseudomonas syringae* pv. *glycinea* determines race-specific incompatibility on *Glycine max* (L.) Merr. *Proceedings of the National Academy of Science USA.* 81: pp. 6024-6028.

Stavrinides, J. (2009). Origin and Evolution of Phytopathogenic Bacteria. In: Jackson, R.W. (2009). *Plant Pathogenic Bacteria: Genomics and Molecular Biology.* Norfolk: Caister Academic Press, pp. 1-36.

Stevens, C., Bennett, M.A., Athanassopoulos, E., Tsiamis, G., Taylor, J.D. and Mansfield, J.W. (1998). Sequence variation in alleles of the virulence gene *avrPphE.R2* from *Pseudomonas syringae* pv. *phaseolicola* lead to loss of recognition of the AvrPphE protein within bean cells and a gain in cultivar-specific virulence. *Molecular Microbiology.* 29: pp. 165-177.

Stockwell, V.O., Davis, E.W., Carey, A., Shaffer, B.T., Mavrodi, D.V., Hassan, K.A., Hockett, K., Thomashow, L.S., Paulsen, I.T. and Loper, J.E. (2013). pA506, a conjugative plasmid of the plant epiphyte *Pseudomonas fluorescens* A506. *Applied and Environmental Microbiology.* 79(17): pp. 5272-5282.

Stokes, H.W. and Hall, R.M. (1989). A novel family of potentially mobile DNA elements encoding site-specific gene-integration functions: integrons. *Molecular Microbiology.* 3(12): pp.1669-1683.

Sundin, G.W. and Murillo, J. (1999). Functional analysis of the *Pseudomonas syringae rulAB* determinant in tolerance to ultraviolet B (290-320 nm) radiation and distribution of *rulAB* among *P*. *syringae* pathovars. *Environmental Microbiology.* 1(1): pp.75-87.

Sundin, G.W., Jacobs, J.L. and Murillo, J. (2000). Sequence diversity of *rulA* among natural isolates of *Pseudomonas syringae* and effect on function of *rulAB*-mediated UVB radiation tolerance. *Applied and Environmental Microbiology.* 66(12): pp.5167-5173.

Sundin, G.W. and Weigand, M.R. (2007). The microbiology of mutability. *FEMS Microbiology Letters*. 277(1), pp.11-20.

Suzuki, A. and Takikawa, Y. (2004).Cloning of a specific DNA region of white top pathogen of pea and its detection by polymerase chain reaction. *Journal of General Plant Pathology.* 70: pp.174-180.

Szybalski, W. and Lyer, V.N. (1964). Crosslinking of DNA by enzymatic or activated mitomycins and porfiromycins, bifunctionally ‘alkylating’ antibiotics. *Federation Proceedings.* 23: pp.946-957.

Takle, G.W., Toth, I.K. and Brurberg, M.B. (2007). Evaluation of reference genes for real-time RT-PCR expression studies in the plant pathogen *Pectobacterium atrosepticum. BMC Plant Biology.* 7(50).

Tark, M., Tover, A., Tarassova, K., Tegova, R., Kivi, G., Horak, R. and Kivisaar, M. (2005). A DNA polymerase V homologue encoded by TOL plasmid pWW0 confers evolutionary fitness on *Pseudomonas putida* under conditions of environmental stress. *Journal of Bacteriology*. 187(15): pp.5203-5213.

Taylor, R.G., Walker, D.C. and McInnes, R.R. (1993). *E. coli* host strains significantly affect the quality of small scale plasmid DNA preparations used for sequencing. *Nucleic Acids Research.* 21(7): pp.1677-1678.

Teramoto, K., Sato, H., Sun, L., Torimura, M., Tao, H., Yoshikawa, H., Hotta, Y., Hosoda, A. and Tamura, H. (2007). Phylogenetic classification of *Pseudomonas putida* strains by MALDI-MS using ribosomal subunit proteins as biomarkers. *Analytical Chemistry*. 79(22): pp.8712-8719.

Thakur, S., Weir, B.S. and Guttman, D.S. (2016). Phytopathogen genome announcement: Draft genome sequences of 62 *Pseudomonas syringae* type and pathotype strains. *Molecular Plant-Microbe Interactions.* 29(4): pp.243-246.

Thermo Fisher Scientific (2008). *T009 Technical Bulletin: NanoDrop 1000 & 8000.* Paisley, UK: Thermo Fisher Scientific.

Tomasz, M. (1995). Mitomycin C: small, fast and deadly (but very selective). *Chemistry and Biology.* 2: pp.575-579.

Toyobo Co., Ltd. (2010). *Toyobo enzymes: malate dehydrogenase.* Available from: toyobo-global.com/seihin/xr/enzyme/pdf_files/209_212MAD_211.pdf. [Accessed 24/05/2017].

Trevors, J.T., Bej, A.K., Mojib, N., van Elsas, J.D. and van Overbeek, L. (2012). *Extremophiles*. 16(2): pp.167-176.

Tsiamis, G., Mansfield, J.W., Hockenhull, R., Jackson, R.W., Sesma, A., Athanassopoulos, E., Bennett, M.A., Stevens, C., Vivian, A., Taylor, J.D. and Murillo J. (2000). Cultivar-specific avirulence and virulence functions assigned to *avrPphF* in *Pseudomonas syringae* pv. *phaseolicola*, the cause of bean halo-blight disease. *EMBO J.* 19(13): pp.3204-3214.

Turner, J.G. (1989). Development of the chlorotic symptom caused by Tabtoxin. *Phytotoxins and plant pathogenesis.* 27: pp.219-238.

van Baarlen, P., van Belkum, A., Summerbell, R.C., Crous, P.W. and Thomma, B.P.H.J. (2007). Molecular mechanisms of pathogenicity: how do pathogenic microorganisms develop cross-kingdom host jumps? *Federation of European Microbiological Societies.* 31: pp.239-277.

van Guilder, H, D., Vrana, K.E. and Freeman, W.M. (2008). Twenty-five years of quantitative PCR for gene expression analysis. *Biotechniques.* 44(5): pp.619-626.

Vivian, A. and Arnold, D.L. (2000). Bacterial effector genes and their role in host-pathogen interactions. *Journal of Plant Pathology.* 82(3): pp.163-178.

Weber, S. (2005). Light-driven enzymatic catalysis of DNA repair: a review of recent biophysical studies on photolyase. *Biochimica et Biophysica Acta (BBA) – Bioenergetics.* 1707(1): pp.1-23

Wigley, D.B., Davies, G.J., Dodson, E.J., Maxwell, A. and Dodson, G. (1991). Crystal structure of an N-terminal fragment of the DNA gyrase B protein. *Nature*. 351(6328): pp.624–629.

Williams, K.P., Gillespie, J.J., Sobral, B.W.S., Nordberg, E.K., Snyder, E.E., Shallom, J.M. and Dickerman, A.W. (2010). Phylogeny of Gammaproteobacteria. *Journal of Bacteriology*. 192 (9): pp. 2305–2314

Yang, X. and Lewis, P.J. (2010). The interaction between bacterial transcription factors and RNA polymerase during the transition from initiation to elongation. *Transcription.* 1(2): pp.66-69.

Young, J.M. (2008). An overview of bacterial nomenclature with special reference to plant pathogens. *Systematic and Applied Microbiology.* 31(6-8): pp. 405-424.

Yu, M., Souaya, J. and Julin, D.A. (1998). The 30-kDa C-terminal domain of the RecB protein is critical for the nuclease activity, but not the helicase activity, of the RecBCD enzyme from *Escherichia coli*. *Proceedings of the National Academy of Sciences USA*. 95(3): pp.981-986.

Yu, X., Lund, S.P., Scott, R.A., Greenwald, J.W., Records, A.H., Nettleton, D., Lindow. S.E., Gross, D.C. and Beattie, G.A. (2013). Transcriptional responses of *Pseudomonas syringae* to growth in epiphytic versus apoplastic leaf sites. *Proceedings of the National Academy of Sciences: USA.* 110(5): pp.425-434.

Zgur-Bertok, D. (2013). DNA damage repair and bacterial pathogens. *PloS Pathogens.* 9(11).

Zhang, S. and Sundin, G.W. (2004). Long-term effect of mutagenic DNA repair on accumulation of mutations in *Pseudomonas syringae* B86-17. *Journal of Bacteriology*. 186(22): pp.7807-7810.

Zhou, J.M. and Chai, J. (2008). Plant Pathogenic bacterial type 3 effectors subdue host responses. *Current Opinion in Microbiology.* 11: pp.179-185.

Zhou, X., Zhang, T., Song, D., Huang, T., Peng, Q., Chen, Y., Li, A., Zhang, F., Wu, Q., Ye, Y. and Tang, Y. (2017). Comparison and evaluation of conventional RT-PCR, SYBR green I and TaqMan real-time RT-PCR assays for the detection of porcine epidemic diarrhoea virus. *Molecular and Cellular Probes* [E-publication].

Appendix I: Bacterial strains & plasmids used during research.

Bacterial strain	Resistance	Reference
Pseudomonas fluorescens		
Pseudomonas fluorescens strain FH1	N/A	Rhodes *et al.*, 2014
Pseudomonas fluorescens strain FH4	N/A	Rhodes *et al.*, 2014
Pseudomonas fluorescens strain FH1 (pWW0::km^r::lLE$_{FH1}$)	Km	Rhodes *et al.*, 2014
Pseudomonas fluorescens strain FH4 (pWW0::km^r::lLE$_{FH4}$)	Km	Rhodes *et al.*, 2014
Pseudomonas putida		
Pseudomonas putida strain PaW340	Stm	Rhodes *et al.*, 2014
Pseudomonas putida strain PaW340 (pWW0::km^r)	Stm, Km	Rhodes *et al.*, 2014
Pseudomonas putida strain PaW340 (pWW0::km^{r}Δ*rulAB*)	Stm, Km	Rhodes *et al.*, 2014
Pseudomonas syringae* pv. *pisi		
Pseudomonas syringae pv. *pisi* race 1 strain 277	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 1 strain 299A	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 1 strain 379	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 1 strain 456A	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 1 strain 461	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 1 strain 862A	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 1 strain 2491B	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 1 strain 4461	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 2 strain 202	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 2 strain 203	N/A	Arnold *et al.*, 2000
Pseudomonas syringae pv. *pisi* race 2 strain 223	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 2 strain 278	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 2 strain 285	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 2 strain 288	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 2 strain 374A	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 2 strain 390	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 2 strain 1124B	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 2 strain 1517C	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 2 strain 1576A	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 2 strain 1577	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 2 strain 1759	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 2 strain 1842A	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 2 strain 1924	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 2 strain 1939	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 2 strain 2889B	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 3 strain 222	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 3 strain 283	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 3 strain 870A	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 3 strain 895A	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 3 strain 1125	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 3 strain 1214	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 3 strain 1216	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 3 strain 1380A	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 3 strain 1441	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 3 strain 1554A	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 3 strain 1892	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 3 strain 2183A	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 3 strain 2186A	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 3 strain 2191A	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 3 strain 2817A	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 3 strain 4411	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 3 strain 4574	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 4 strain 1452	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 4 strain 1456A	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 4 strain 1456B	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 4 strain 1456C	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 4 strain 1456D	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 4 strain 1456E	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 4 strain 1456F	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 4 strain 1525	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 4 strain 1528	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 4 strain 1758B	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 4 strain 1811A	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 4 strain 1812A	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 4 strain 2171A	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 4 strain 5143	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 5 strain 974B	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 5 strain 2301C	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 5 strain 2532A	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 6 strain 1683	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 6 strain 1688	N/A	HRI cuture collection

Pseudomonas syringae pv. *pisi* race 6 strain 1704B	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 6 strain 1745A	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 6 strain 1746A	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 6 strain 1755A	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 6 strain 1759	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 6 strain 1785A	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 6 strain 1796A	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 6 strain 1797A	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 6 strain 1804A	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 6 strain 1807A	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 6 strain 1842B	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 6 strain 1842C	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 6 strain 1842D	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 6 strain 1942	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 7 strain 1691	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 7 strain 2491A	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 7 strain 4298	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 7 strain 4300	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 7 strain 4409	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 7 strain 4426	N/A	HRI cuture collection
Pseudomonas syringae pv. *pisi* race 7 strain 4466	N/A	HRI cuture collection
Pseudomonas syringae* pv *phaseolicola		
Pseudomonas syringae pv. *phaseolicola* strain 103	N/A	Tsiamis *et al.,* 2000
Pseudomonas syringae pv. *phaseolicola* race 1 strain 1281A	N/A	Tsiamis *et al.,* 2000
Pseudomonas syringae pv. *phaseolicola* race 6 strain 1299A	N/A	Tsiamis *et al.,* 2000
Pseudomonas syringae pv. *phaseolicola* race 3 strain 1301A	N/A	Tsiamis *et al.,* 2000
Pseudomonas syringae pv. *phaseolicola* race 4 strain 1302A	N/A	HRI cuture collection
Pseudomonas syringae pv. *phaseolicola* race 5 strain 1375A	N/A	Tsiamis *et al.,* 2000
Pseudomonas syringae pv. *phaseolicola* race 6 strain 1448A	Rif	Tsiamis *et al.,* 2000
Pseudomonas syringae pv. *phaseolicola* race 7 strain 1449A	N/A	Tsiamis *et al.,* 2000
Pseudomonas syringae pv. *phaseolicola* race 7 strain 1449B	N/A	Tsiamis *et al.,* 2000
Pseudomonas syringae pv. *phaseolicola* race 8 strain 2656A	N/A	Tsiamis *et al.,* 2000
Pseudomonas syringae pv. *phaseolicola* race 9 strain 2709A	N/A	Tsiamis *et al.,* 2000
Pseudomonas syringae* pv. *syringae		
Pseudomonas syringae pv. *syringae* strain B728A	N/A	HRI cuture collection
Pseudomonas syringae pv. *syringae* strain 100	N/A	HRI cuture collection
Pseudomonas syringae pv. *syringae* strain 1142	N/A	HRI cuture collection
Pseudomonas syringae pv. *syringae* strain 1150	N/A	HRI cuture collection
Pseudomonas syringae pv. *syringae* strain 1212	N/A	HRI cuture collection
Pseudomonas syringae pv. *syringae* strain 1282-8	N/A	HRI cuture collection
Pseudomonas syringae pv. *syringae* strain 1338A	N/A	HRI cuture collection
Pseudomonas syringae pv. *syringae* strain 2242A	N/A	HRI cuture collection
Pseudomonas syringae pv. *syringae* strain 2673C	N/A	HRI cuture collection
Pseudomonas syringae pv. *syringae* strain 2675C	N/A	HRI cuture collection
Pseudomonas syringae pv. *syringae* strain 2676C	N/A	HRI cuture collection
Pseudomonas syringae pv. *syringae* strain 2677C	N/A	HRI cuture collection
Pseudomonas syringae pv. *syringae* strain 2682C	N/A	HRI cuture collection
Pseudomonas syringae pv. *syringae* strain 2692C	N/A	HRI cuture collection
Pseudomonas syringae pv. *syringae* strain 2703C	N/A	HRI cuture collection
Pseudomonas syringae pv. *syringae* strain 2732A	N/A	HRI cuture collection
Pseudomonas syringae pv. *syringae* strain 3023	N/A	HRI cuture collection
Pseudomonas syringae* pv. *maculicola		
Pseudomonas syringae pv. *maculicola* strain M4	N/A	HRI cuture collection
Pseudomonas syringae pv. *maculicola* strain 65	N/A	HRI cuture collection
Pseudomonas syringae pv. *maculicola* strain 1809A	N/A	HRI cuture collection
Pseudomonas syringae pv. *maculicola* strain 1813	N/A	HRI cuture collection
Pseudomonas syringae pv. *maculicola* strain 1820	N/A	HRI cuture collection
Pseudomonas syringae pv. *maculicola* strain 1821A	N/A	HRI cuture collection
Pseudomonas syringae pv. *maculicola* strain 1838A	N/A	HRI cuture collection
Pseudomonas syringae pv. *maculicola* strain 1846A	N/A	HRI cuture collection
Pseudomonas syringae pv. *maculicola* strain 1848B	N/A	HRI cuture collection
Pseudomonas syringae pv. *maculicola* strain 1852A	N/A	HRI cuture collection
Pseudomonas syringae pv. *maculicola* strain 1855A	N/A	HRI cuture collection
Pseudomonas syringae pv. *maculicola* strain 5422	N/A	HRI cuture collection
Pseudomonas syringae pv. *maculicola* strain 5429	N/A	HRI cuture collection
Pseudomonas syringae pv. *maculicola* strain 6201A	N/A	HRI cuture collection
Pseudomonas syringae pv. *maculicola* strain 6319A/1	N/A	HRI cuture collection
Pseudomonas syringae pv. *maculicola* strain 6328A/1	N/A	HRI cuture collection

Pseudomonas syringae* pv. *tomato		
Pseudomonas syringae pv. *tomato* strain DC3000	N/A	HRI cuture collection
Pseudomonas syringae pv. *tomato* strain 19	N/A	HRI cuture collection
Pseudomonas syringae pv. *tomato* strain 119	N/A	HRI cuture collection
Pseudomonas syringae pv. *tomato* strain 138	N/A	HRI cuture collection
Pseudomonas syringae pv. *tomato* strain 1108	N/A	HRI cuture collection
Pseudomonas syringae pv. *tomato* strain 2944	N/A	HRI cuture collection
Pseudomonas syringae pv. *tomato* strain 2945	N/A	HRI cuture collection
Pseudomonas syringae pv. *tomato* strain 6034	N/A	HRI cuture collection
***Pseudomonas syringae* pv. antirrhini**		
Pseudomonas syringae pv. *antirrhini* strain 152E	N/A	HRI cuture collection
Pseudomonas syringae pv. *antirrhini* strain 4303	N/A	HRI cuture collection
Pseudomonas syringae* pv. *lachrymans		
Pseudomonas syringae pv. *lachrymans* strain 789	N/A	HRI cuture collection
Pseudomonas syringae pv. *lachrymans* strain 3988	N/A	HRI cuture collection
Pseudomonas syringae* pv. *glycinea		
Pseudomonas syringae pv. *glycinea* strain 1139	N/A	HRI cuture collection
Pseudomonas syringae pv. *glycinea* strain 2411	N/A	HRI cuture collection
Pseudomonas syringae pv. *glycinea* strain 3318	N/A	HRI cuture collection
Pseudomonas asperii		
Pseudomonas asperii strain 959	N/A	HRI cuture collection
Pseudomonas asperii strain 1947	N/A	HRI cuture collection
Pseudomonas caricapapayae		
Pseudomonas caricapapayae strain 1873	N/A	HRI cuture collection
Pseudomonas caricapapayae strain 3080	N/A	HRI cuture collection
Pseudomonas caricapapayae strain 3439	N/A	HRI cuture collection
Pseudomonas savastanoi		
Pseudomonas savastanoi strain 639	N/A	HRI cuture collection
Pseudomonas savastanoi strain 2716	N/A	HRI cuture collection
Pseudomonas savastanoi strain 3334	N/A	HRI cuture collection
Pseudomonas corrugata		
Pseudomonas corrugata strain 2445	N/A	HRI cuture collection
Pseudomonas corrugata strain 3056	N/A	HRI cuture collection
Pseudomonas corrugata strain 3316	N/A	HRI cuture collection
Pseudomonas cichorii		
Pseudomonas cichorii strain 907	N/A	HRI cuture collection
Pseudomonas cichorii strain 943	N/A	HRI cuture collection
Pseudomonas cichorii strain 3109A	N/A	HRI cuture collection
Pseudomonas cichorii strain 3109B	N/A	HRI cuture collection
Pseudomonas cichorii strain 3283	N/A	HRI cuture collection
Pseudomonas marginalis		
Pseudomonas marginalis strain 247	N/A	HRI cuture collection
Pseudomonas marginalis strain 949	N/A	HRI cuture collection
Pseudomonas marginalis strain 2380	N/A	HRI cuture collection
Pseudomonas marginalis strain 2644	N/A	HRI cuture collection
Pseudomonas marginalis strain 2645	N/A	HRI cuture collection
Pseudomonas marginalis strain 2646	N/A	HRI cuture collection
Pseudomonas marginalis strain 3210	N/A	HRI cuture collection
***Pseudomonas* used during cloning**		
P. putida PaW340 (pBBR1MCS-2::pWW0 rulAB)	Km	This study
P. putida PaW340 (pBBR1MCS-2::pWW0 rulB)	Km	This study
P. putida PaW340 (pBBR1MCS-2::pWW0 rulAB'-lP) pWW0 rulAB upto lLE insertion point	Km	This study
P. putida PaW340 (pBBR1MCS-2::PphGI rulAB)	Km	This study
P. flourescens FH1 (pBBR1MCS-2::pWW0 rulAB)	Km	This study
P. flourescens FH1 (pBBR1MCS-2::pWW0 rulB)	Km	This study
P. flourescens FH1 (pBBR1MCS-2::pWW0 rulAB'-lP) pWW0 rulAB upto lLE insertion point	Km	This study
P. flourescens FH1 (pBBR1MCS-2::PphGI rulAB)	Km	This study
Ppi 203 (pBBR1MCS-2::pWW0 rulAB)	Km	This study
Ppi 203 (pBBR1MCS-2::PphGI rulAB)	Km	This study
Psy 3023 (pBBR1MCS-2::pWW0 rulAB)	Km	This study
Psy 3023 (pBBR1MCS-2::PphGI rulAB)	Km	This study
E.coli		
DH5α, non pathogenic *E.coli* used during cloning	N/A	Taylor et al., 1993
DH5α (pBBR1MCS-2::pWW0 rulAB)	Km	This study
DH5α (pBBR1MCS-2::pWW0 rulB)	Km	This study
DH5α (pBBR1MCS-2::pWW0 rulAB'-lP) pWW0 rulAB upto lLE insertion point	Km	This study
DH5α (pBBR1MCS-2::PPHGI rulAB)	Km	This study

Plasmids	Resistance	Reference
pWW0::km^r (large self-transmissible plasmid with heavy	Km	Rhodes *et al.*, 2014
metal resistance)		
pWW0::km^r (large self-transmissible plasmid with heavy metal	Km	Rhodes *et al.*, 2014
resistance, also has the integron inserted into rulAB from Pf. FH1)		
pRK2013, kanamycin resistance helper vector from E. coli DH5α	Km	Ditta, 1980
pBBR1MCS-2, kanamycin resistance broad host range vector	Km	Kovach, 1995
pBBR1MCS-5, gentamicin resistance broad host range vector	Gm	Kovach, 1995
pCR2.1, cloning vector for PCR products	Km	Invitrogen circa 1999

Appendix II: Chemical composition of media, buffers and solutions.

KB Media

15.2 g *Pseudomonas* powder, 400 mL distilled water, 4 g glycerol. Autoclaved.

LB agar

10 g LB powder per 400 mL distilled water, 5.5 g bacteriological agar per 400 mL. Autoclaved.

LB Broth

50 g LB powder, 2 L distilled water, dispensed into 10 mL aliquots. Autoclaved.

NA Media

11.2 g nutrient agar powder, 400 mL distilled water, bring to the boil. Autoclaved.

M9 Minimal Media agar

4.2 g M9 powder, 400 mL distilled water, 5.5 g bacteriological agar. Autoclaved. Then add 1 mL 1 M magnesium sulphate ($MgSO_4$), 10 mL 20% glucose ($C_6H_{12}O_6$).

M9 Minimal Media broth

4.2 g M9 powder, 400 mL distilled water. Autoclaved.
Then add 1 mL 1 M magnesium sulphate ($MgSO_4$), 10 mL 20% glucose ($C_6H_{12}O_6$).

¼ Ringers solution

1 tablet per 500 mL of distilled water. Autoclaved.

Loading Dye for PCR samples

25 mg bromophenol blue, 25 mg xylene cyanol, 3.3 mL glycerol, 6.7 mL distilled water.

50x TAE

242 g Tris base dissolved in 750 mL deionized water. Add 57.1 mL acetic acid and 100 mL of 0.5 M EDTA (pH 8.0) and adjust the solution to a final volume of 1 L with distilled water.

1xTAE

Dilute 20 mL 50x TAE with 1980 mL distilled water to make a 2 L stock.

2x Saline-Sodium Citrate (SSC) buffer

Dilute 20x SSC (Roche, UK) with distilled water with a ration of 1:10. Eg. 50 mL 20x SSC in 450 mL distilled water.

Low Stringency Buffer (L.S.B.)

2x SSC containing 0.1% w/v Sodium Dodecyl Sulfate (SDS).

High Stringency Buffer (H.S.B.)

0.5x SSC containing 0.1% w/v SDS.

Washing Buffer

0.1 M maleic acid, 0.15M sodium chloride (NaCl), pH 7.5 with sodium hydroxide (NaOH). Autoclaved. Add 0.3% v/v Tween 20.

Maleic Acid Buffer

0.1 M maleic acid, 0.15 M NaCl, pH 7.5 with NaOH. Autoclaved.

Blocking Solution

Dilute 10x blocking solution (Roche, UK) 1:10 with maleic acid buffer. Prepare fresh.

Antibody Solution

Centrifuge Anti-Digoxigenin-AP (Roche, UK) for 5 mins at 10,000 rpm before each use. Pipette the necessary volume carefully from the surface. Dilute the Anti-Digoxigenin-AP 1:10,000 in blocking solution. Prepare fresh (2 hours, 4^{o}C).

Detection Buffer

0.1 M Tris-HCl, 0.1 M NaCl, pH 9.5. Autoclaved.

Isopropyl β-D-1-thiogalactopyranoside (IPTG)

25 mg of IPTG powder (Sigma) to 1 mL of sterile water. 800 µL per 400 mL agar.

5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal)

20 mg of x-Gal (Sigma) to 1 mL of Dimethylformamide (DMF). 800 µL per 400 mL agar.

Rifampicin

10 mg/ mL of methanol.

Kanamycin

10 mg/ mL of water, filter sterilise.

Streptomycin

200 mg/ mL of water, filter sterilise.

1x TE buffer pH8

Stock solution of 1 M Tris-HCl: 12.1 g in 100 mL SDW and pH 8.0 with HCl
Stock solution of 0.5 M EDTA: 18.6 g in 100 mL SDW and pH 8.0 with NaOH
1 x TE; Mix: 100 mL 1 M Tris-HCl and 20 mL 0.5 M EDTA

Appendix III: Sequence alignment of conserved ILE junction.

```
CLUSTAL O(1.2.2) multiple sequence alignment


Ppi203          ---------------TCCGCAGCT-ATGACGCCAGCCATTTGTGAAAATGGGGGCTCCGT
Ppi202          ----------------CGCAGCT-ATGACGCCAGCCATTTGTGAAAATGGGGGCTCCGT
Ppi223          ----------------GCCCAGCT-ATGACGCCAGCCATTTGTGAAAATGGGGGCTCCGT
Ppi283          ----------------CCGCAGCT-ATGACGCCAGCCATTTGTGAAAATGGGGGCTCCGT
Ppi288          -----------------CCCAGCT-ATGACGCCAGCCATTTGTGAAAATGGGGGCTCCGT
Ppi374A         -----------------CGCAGCT-ATGACGCCAGCCATTTGTGAAAATGGGGGCTCCGT
Ppi390          -----------------CGCAGCT-ATGACGCCAGCCATTTGTGAAAATGGGGGCTCCGT
Ppi1452         ------------------GTAGCT-ACGACGCCAGCCATTTGTGAAAATGGGCGCTCCCT
Ppi1456A        ------------------TAGCT-ACGACGCCAGCCATTTGTGAAAATGGGCGCTCCCT
Ppi1456B        --------------AAAGTAGCT-ACGACGCCAGCCATTTGTGAAAATGGGCGCTCCCT
Ppi1456C        ------------------TAGCT-ACGACGCCAGCCATTTGTGAAAATGGGCGCTCCCT
Ppi1456D        ----------------CCGTAGCT-ACGACGCCAGCCATTTGTGAAAATGGGCGCTCCCT
Ppi1456E        ---------------GACGTAGCT-ACGACGCCAGCCATTTGTGAAAATGGGCGCTCCCT
Ppi1456F        -----------------GTAGCT-ACGACGCCAGCCATTTGTGAAAATGGGCGCTCCCT
Ppi1939         ----------------------T-CTGACGCCAGCCATTTGTGAAAATGGGGGCTCCGT
Ppi2889B        --------------AACGTAGCT-ACGACGCCAGCCATTTGTGAAAATGGGCGCTCCCT
PsyB728a        ----------------------T-CTGACGCCAGCCATTTGTGAAAATGGGGGCTCCGT
Psy3023         GACGGTGCGTCATCGCCCGTAGCTACGACGCCAAGCCATTTGTGAAGATGGGGGCTCCCT
Pma1852         ---------------GACGGAGCT-ACGACGCCAACCGTTTGTGAAAATGGGCGCGCCCT
Pma5422         ----------------CCGGAGCT-ACGACGCCAACCGTTTGTGAAAATGGGCGCGCCCT
Pma6328A        --------------------------CGACGCCAACCGTTTGTGAAAATGGGCGCGCCCT
Pgy2411         --------------------------CGACGCCAACCGTTTGTGAAAATGGGCGCGCCCT
                                          * * ** ******** ***** ** ** *

Ppi203          ATTTCCAGATCAAAGAGGTGCTGCGCCGCAATGGCATCAAGGTGTTCAGCAGCAACTACG
Ppi202          ATTTCCAGATCAAAGAGGTGCTGCGCCGCAATGGCATCAAGGTGTTCAGCAGCAACTACG
Ppi223          ATTTCCAGATCAAAGAGGTGCTGCGCCGCAATGGCATCAAGGTGTTCAGCAGCAACTACG
Ppi283          ATTTCCAGATCAAAGAGGTGCTGCGCCGCAATGGCATCAAGGTGTTCAGCAGCAACTACG
Ppi288          ATTTCCAGATCAAAGAGATGCTGCGCCGCAATGGGATCAAGGTGTTCAGCAGCAACTACG
Ppi374A         ATTTCCAGATCAAAGAGGTGCTGCGCCGCAATGGCATCAAGGTGTTCAGCAGCAACTACG
Ppi390          ATTTCCAGATCAAAGAGGTGCTGCGCCGCAATGGCATCAAGGTGTTCAGCAGCAACTACG
Ppi1452         ATTTTCAGATCAGAGAGGTGTTGCGCCGCAACGGCATCAAGGTGTTCAGCAGCAACTATG
Ppi1456A        ATTTTCAGATCAGAGAGGTGTTGCGCCGCAACGGCATCAAGGTGTTCAGCAGCAACTATG
Ppi1456B        ATTTTCAGATCAGAGAGGTGTTGCGCCGCAACGGCATCAAGGTGTTCAGCAGCAACTATG
Ppi1456C        ATTTTCAGATCAGAGAGGTGTTGCGCCGCAACGGCATCAAGGTGTTCAGCAGCAACTATG
Ppi1456D        ATTTTCAGATCAGAGAGGTGTTGCGCCGCAACGGCATCAAGGTGTTCAGCAGCAACTATG
Ppi1456E        ATTTTCAGATCAGAGAGGTGTTGCGCCGCAACGGCATCAAGGTGTTCAGCAGCAACTATG
Ppi1456F        ATTTTCAGATCAGAGAGGTGTTGCGCCGCAACGGCATCAAGGTGTTCAGCAGCAACTATG
Ppi1939         ATTTCCAGATCAAAGAGGTGCTGCGCCGCAATGGGATCAAGGTGTTCAGCAGCAACTACG
Ppi2889B        ATTTTCAGATCAGAGAGGTGTTGCGCCGCAACGGCATCAAGGTGTTCAGCAGCAACTATG
PsyB728a        ATTTCCAGATCAAAGAGGTGCTGCGCCGCAATGGCATCAAGGTGTTCAGCAGCAACTACG
Psy3023         ATTTTCAGATCAAAGAGGTGTTGCGCCGCCACGGCATAAAGGTGTTCAGCAGCAACTACG
Pma1852         ATTTTCAGATTAAGGATGTACTGAAGCGAAACGGTATCAAGGTTTTCAGCAGCAACTATG
Pma5422         ATTTTCAGATTAAGGATGTACTGAAGCGAAACGGTATCAAGGTTTTCAGCAGCAACTATG
Pma6328A        ATTTTCAGATTAAGGATGTACTGAAGCGAAACGGTATCAAGGTTTTCAGCAGCAACTATG
Pgy2411         ATTTTCAGATTAAGGATGTACTGAAGCGAAACGGTATCAAGGTTTTCAGCAGCAACTATG
                **** ***** *  **  *  **   **  * ** ** ***** ************** *

Ppi203          CGCTTTAGTAAGTTAGGTGGAATGCTTCTGGGCTACGCTGATTCTGTCGACGTTCTGGGG
Ppi202          CGCTTTAGTAAGTTAGGTGGAATGCTTCTGGGCTACGCTGATTCTGTCGACGTTCTGGGG
Ppi223          CGCTTTAGTAAGTTAGGTGGAATGCTTCTGGGCTACGCTGATTCTGTCGACGTTCTGGGG
Ppi283          CGCTTTAGTAAGTTAGGTGGAATGCTTCTGGGCTACGCTGATTCTGTCGACGTTCTGGGG
Ppi288          CGCTTTAGTAAGTTAGGTGGAATGCTTCTGGGCTACGCTGATTCTGTCGACGTTCTGGGG
Ppi374A         CGCTTTAGTAAGTTAGGTGGAATGCTTCTGGGCTACGCTGATTCTGTCGACGTTCTGGGG
Ppi390          CGCTTTAGTAAGTTAGGTGGAATGCTTCTGGGCTACGCTGATTCTGTCGACGTTCTGGGG
Ppi1452         CACTCTAGTAAGTTAGGTGGAATGCTTCTGGGCTACGCTGATTCTGTCGACGTTCTGGGG
Ppi1456A        CACTCTAGTAAGTTAGGTGGAATGCTTCTGGGCTACGCTGATTCTGTCGACGTTCTGGGG
Ppi1456B        CACTCTAGTAAGTTAGGTGGAATGCTTCTGGGCTACGCTGATTCTGTCGACGTTCTGGGG
Ppi1456C        CACTCTAGTAAGTTAGGTGGAATGCTTCTGGGCTACGCTGATTCTGTCGACGTTCTGGGG
Ppi1456D        CACTCTAGTAAGTTAGGTGGAATGCTTCTGGGCTACGCTGATTCTGTCGACGTTCTGGGG
Ppi1456E        CACTCTAGTAAGTTAGGTGGAATGCTTCTGGGCTACGCTGATTCTGTCGACGTTCTGGGG
Ppi1456F        CACTCTAGTAAGTTAGGTGGAATGCTTCTGGGCTACGCTGATTCTGTCGACGTTCTGGGG
Ppi1939         CGCTTTAGTAAGTTAGGTGGAATGCTTCTGGGCTACGCTGATTCTGTCGACGTTCTGGGG
Ppi2889B        CACTCTAGTAAGTTAGGTGGAATGCTTCTGGGCTACGCTGATTCTGTCGACGTTCTGGGG
PsyB728a        CGCTTTAGTAAGTTAGGTGGAATGCTTCTGGGCTACGCTGATTCTGTCGACGTTCTGGGG
Psy3023         CGCTCTAGTAAGTTAGGTGGAATGCTTTTGGACTACGCTGATTCTGTCGACGTTCTGGGG
Pma1852         CGCTTTAGTAAGTTAGGTGGAATGCTTCTGGGCTACGCTGATTCTGTCGACGTTCTGGGG
Pma5422         CGCTTTAGTAAGTTAGGTGGAATGCTTCTGGGCTACGCTGATTCTGTCGACGTTCTGGGG
Pma6328A        CGCTTTAGTAAGTTAGGTGGAATGCTTCTGGGCTACGCTGATTCTGTCGACGTTCTGGGG
```

```
Pgy2411         CGCTTTAGTAAGTTAGGTGGAATGCTTCTGGGCTACGCTGATTCTGTCGACGTTCTGGGG
                * ** ********************** *** ***************************

Ppi203          GATCGCCCATGGCCTCTACTGCCTCTTATGAGCACGTGCTTAACCGCTATCCTGACGTTC
Ppi202          GATCGCCCATGGCCTCTACTGCCTCTTATGAGCACGTGCTTAACCGCTATCCTGACGTTC
Ppi223          GATCGCCCATGGCCTCTACTGCCTCTTATGAGCACGTGCTTAACCGCTATCCTGACGTTC
Ppi283          GATCGCCCATGGCCTCTACTGCCTCTTATGAGCACGTGCTTAACCGCTATCCTGACGTTC
Ppi288          GATCGCCCATGGCCTCTACTGCCTCTTATGAGCACGTGCTTAACCGCTATCCTGACGTTC
Ppi374A         GATCGCCCATGGCCTCTACTGCCTCTTATGAGCACGTGCTTAACCGCTATCCTGACGTTC
Ppi390          GATCGCCCATGGCCTCTACTGCCTCTTATGAGCACGTGCTTAACCGCTATCCTGACGTTC
Ppi1452         GATCGCCCATGGCCTCTACTGCCTCTTATGAGCACGTGCTTAACCGCTATCCTGACGTTC
Ppi1456A        GATCGCCCATGGCCTCTACTGCCTCTTATGAGCACGTGCTTAACCGCTATCCTGACGTTC
Ppi1456B        GATCGCCCATGGCCTCTACTGCCTCTTATGAGCACGTGCTTAACCGCTATCCTGACGTTC
Ppi1456C        GATCGCCCATGGCCTCTACTGCCTCTTATGAGCACGTGCTTAACCGCTATCCTGACGTTC
Ppi1456D        GATCGCCCATGGCCTCTACTGCCTCTTATGAGCACGTGCTTAACCGCTATCCTGACGTTC
Ppi1456E        GATCGCCCATGGCCTCTACTGCCTCTTATGAGCACGTGCTTAACCGCTATCCTGACGTTC
Ppi1456F        GATCGCCCATGGCCTCTACTGCCTCTTATGAGCACGTGCTTAACCGCTATCCTGACGTTC
Ppi1939         GATCGCCCATGGCCTCTACTGCCTCTTATGAGCACGTGCTTAACCGCTATCCTGACGTTC
Ppi2889B        GATCGCCCATGGCCTCTACTGCCTCTTATGAGCACGTGCTTAACCGCTATCCTGACGTTC
PsyB728a        GATCGCCCATGGCCTCTACTGCCTCTTATGAGCACGTGCTTAACCGCTATCCTGACGTTC
Psy3023         GATTGCCCATGGCCTCTACTGCCTCTTATGAGCACGTGCTTCACCGCTACCCCAGCGTTC
Pma1852         GATCGCCCATGGCCTCTACTGCCTCTTATGAGCACGTGCTTAACCGCTATCCTGACGTTC
Pma5422         GATCGCCCATGGCCTCTACTGCCTCTTATGAGCACGTGCTTAACCGCTATCCTGACGTTC
Pma6328A        GATCGCCCATGGCCTCTACTGCCTCTTATGAGCACGTGCTTAACCGCTATCCTGACGTTC
Pgy2411         GATCGCCCATGGCCTCTACTGCCTCTTATGAGGACGTGCTTCACCGCTACCCCAGCGTTC
                *** ************************** ******** ******* **   *****

Ppi203          AGGAATGGCTGGCGCTGCTCGGTAACCTGGGAAGAGCGTCTGCTACCTTGGAGGCGTATG
Ppi202          AGGAATGGCTGGCGCTGCTCGGTAACCTGGGAAGAGCGTCTGCTACCTTGGAGGCGTATG
Ppi223          AGGAATGGCTGGCGCTGCTCGGTAACCTGGGAAGAGCGTCTGCTACCTTGGAGGCGTATG
Ppi283          AGGAATGGCTGGCGCTGCTCGGTAACCTGGGAAGAGCGTCTGCTACCTTGGAGGCGTATG
Ppi288          AGGAATGGCTGGCGCTGCTCGGTAACCTGGGAAGAGCGTCTGCTACCTTGGAGGCGTATG
Ppi374A         AGGAATGGCTGGCGCTGCTCGGTAACCTGGGAAGAGCGTCTGCTACCTTGGAGGCGTATG
Ppi390          AGGAATGGCTGGCGCTGCTCGGTAACCTGGGAAGAGCGTCTGCTACCTTGGAGGCGTATG
Ppi1452         AGGAATGGCTGGCGCTGCTCGGTAACCTGGGAAGAGCGTCTGCTACCTTGGAGGCGTATG
Ppi1456A        AGGAATGGCTGGCGCTGCTCGGTAACCTGGGAAGAGCGTCTGCTACCTTGGAGGCGTATG
Ppi1456B        AGGAATGGCTGGCGCTGCTCGGTAACCTGGGAAGAGCGTCTGCTACCTTGGAGGCGTATG
Ppi1456C        AGGAATGGCTGGCGCTGCTCGGTAACCTGGGAAGAGCGTCTGCTACCTTGGAGGCGTATG
Ppi1456D        AGGAATGGCTGGCGCTGCTCGGTAACCTGGGAAGAGCGTCTGCTACCTTGGAGGCGTATG
Ppi1456E        AGGAATGGCTGGCGCTGCTCGGTAACCTGGGAAGAGCGTCTGCTACCTTGGAGGCGTATG
Ppi1456F        AGGAATGGCTGGCGCTGCTCGGTAACCTGGGAAGAGCGTCTGCTACCTTGGAGGCGTATG
Ppi1939         AGGAATGGCTGGCGCTGCTCGGTAACCTGGGAAGAGCGTCTGCTACCTTGGAGGCGTATG
Ppi2889B        AGGAATGGCTGGCGCTGCTCGGTAACCTGGGAAGAGCGTCTGCTACCTTGGAGGCGTATG
PsyB728a        AGGAATGGCTGGCGCTGCTCGGTAACCTGGGAAGAGCGTCTGCTACCTTGGAGGCGTATG
Psy3023         AGGAATGGCTGACACTGCTCGATAACCTGGGAAGAGCGCCGGCTACTTTGGATGCGTACG
Pma1852         AGGAATGGCTGGCGCTGCTCGGTAACCTGGGAAGAGCGTCTGCTACCTTGGAGGCGTATG
Pma5422         AGGAATGGCTGGCGCTGCTCGGTAACCTGGGAAGAGCGTCTGCTACCTTGGAGGCGTATG
Pma6328A        AGGAATGGCTGGCGCTGCTCGGTAACCTGGGAAGAGCGTCTGCTACCTTGGAGGCGTATG
Pgy2411         AGGAATGGCTGGCGCTGCTCGGTAACCTGGGAAGAGCACCGGCTACCTTGGATGCCTACG
                *********** * ******* ***************  * ***** ***** ** ** *

Ppi203          GCCGGGGGTTGGCGCACTACCTGCTCCACTGCGAGGCCTCCGGTCTGGAGGCTGAATCCA
Ppi202          GCCGGGGGTTGGCGCACTACCTGCTCCACTGCGAGGCCTCCGGTCTGGAGGCTGAATCCA
Ppi223          GCCGGGGGTTGGCGCACTACCTGCTCCACTGCGAGGCCTCCGGTCTGGAGGCTGAATCCA
Ppi283          GCCGGGGGTTGGCGCACTACCTGCTCCACTGCGAGGCCTCCGGTCTGGAGGCTGAATCCA
Ppi288          GCCGGGGGTTGGCGCACTACCTGCTCCACTGCGAGGCCTCCGGTCTGGAGGCTGAATCCA
Ppi374A         GCCGGGGGTTGGCGCACTACCTGCTCCACTGCGAGGCCTCCGGTCTGGAGGCTGAATCCA
Ppi390          GCCGGGGGTTGGCGCACTACCTGCTCCACTGCGAGGCCTCCGGTCTGGAGGCTGAATCCA
Ppi1452         GCCGGGGGTTGGCGCACTACCTGCTCCACTGCGAGGCCTCCGGTCTGGAGGCTGAATCCA
Ppi1456A        GCCGGGGGTTGGCGCACTACCTGCTCCACTGCGAGGCCTCCGGTCTGGAGGCTGAATCCA
Ppi1456B        GCCGGGGGTTGGCGCACTACCTGCTCCACTGCGAGGCCTCCGGTCTGGAGGCTGAATCCA
Ppi1456C        GCCGGGGGTTGGCGCACTACCTGCTCCACTGCGAGGCCTCCGGTCTGGAGGCTGAATCCA
Ppi1456D        GCCGGGGGTTGGCGCACTACCTGCTCCACTGCGAGGCCTCCGGTCTGGAGGCTGAATCCA
Ppi1456E        GCCGGGGGTTGGCGCACTACCTGCTCCACTGCGAGGCCTCCGGTCTGGAGGCTGAATCCA
Ppi1456F        GCCGGGGGTTGGCGCACTACCTGCTCCACTGCGAGGCCTCCGGTCTGGAGGCTGAATCCA
Ppi1939         GCCGGGGGTTGGCGCACTACCTGCTCCACTGCGAGGCCTCCGGTCTGGAGGCTGAATCCA
Ppi2889B        GCCGGGGGTTGGCGCACTACCTGCTCCACTGCGAGGCCTCCGGTCTGGAGGCTGAATCCA
PsyB728a        GCCGGGGGTTGGCGCACTACCTGCTCCACTGCGAGGCCTCCGGTCTGGAGGCTGAATCCA
Psy3023         GCAGGGGATTGGCGCATTACTTGCTCCACTGCGAAGCCTCCGGTCTGGAGGCTGAATCCA
Pma1852         GCCGGGGGTTGGCGCACTACCTGCTCCACTGCGAGGCCTCCGGTCTGGAGGCTGAATCCA
Pma5422         GCCGGGGGTTGGCGCACTACCTGCTCCACTGCGAGGCCTCCGGTCTGGAGGCTGAATCCA
Pma6328A        GCCGGGGGTTGGCGCACTACCTGCTCCACTGCGAGGCCTCCGGTCTGGAGGCTGAATCCA
Pgy2411         GCAGGGGATTGGCGCATTACTTGCTCCACTGCGAAGCCTCCGGTCTGGAGGCTGAATCCA
                ** **** ******** *** ************* ************************
```

```
Ppi203          TCACATTTGAGCAAGTCACGCTCTACATCCGTCGGCTACAGCCCGGGCAAGAAAACGCGG
Ppi202          TCACATTTGAGCAAGTCACGCTCTACATCCGTCGGCTACAGCCCGGGCAAGAAAACGCGG
Ppi223          TCACATTTGAGCAAGTCACGCTCTACATCCGTCGGCTACAGCCCGGGCAAGAAAACGCGG
Ppi283          TCACATTTGAGCAAGTCACGCTCTACATCCGTCGGCTACAGCCCGGGCAAGAAAACGCGG
Ppi288          TCACATTTGAGCAAGTCACGCTCTACATCCGTCGGCTACAGCCCGGGCAAGAAAACGCGG
Ppi374A         TCACATTTGAGCAAGTCACGCTCTACATCCGTCGGCTACAGCCCGGGCAAGAAAACGCGG
Ppi390          TCACATTTGAGCAAGTCACGCTCTACATCCGTCGGCTACAGCCCGGGCAAGAAAACGCGG
Ppi1452         TCACATTTGAGCAAGTCACGCTCTACATCCGTCGGCTACAGCCCGGGCAAGAAAACGCGG
Ppi1456A        TCACATTTGAGCAAGTCACGCTCTACATCCGTCGGCTACAGCCCGGGCAAGAAAACGCGG
Ppi1456B        TCACATTTGAGCAAGTCACGCTCTACATCCGTCGGCTACAGCCCGGGCAAGAAAACGCGG
Ppi1456C        TCACATTTGAGCAAGTCACGCTCTACATCCGTCGGCTACAGCCCGGGCAAGAAAACGCGG
Ppi1456D        TCACATTTGAGCAAGTCACGCTCTACATCCGTCGGCTACAGCCCGGGCAAGAAAACGCGG
Ppi1456E        TCACATTTGAGCAAGTCACGCTCTACATCCGTCGGCTACAGCCCGGGCAAGAAAACGCGG
Ppi1456F        TCACATTTGAGCAAGTCACGCTCTACATCCGTCGGCTACAGCCCGGGCAAGAAAACGCGG
Ppi1939         TCACATTTGAGCAAGTCACGCTCTACATCCGTCGGCTACAGCCCGGGCAAGAAAACGCGG
Ppi2889B        TCACATTTGAGCAAGTCACGCTCTACATCCGTCGGCTACAGCCCGGGCAAGAAAACGCGG
PsyB728a        TCACATTTGAGCAAGTCACGCTCTACATCCGTCGGCTACAGCCCGGGCAAGAAAACGCGG
Psy3023         TCACATTTGAGCAAGTCACGCTCTACATCCGTCGGCTACTGCCCGGGCAAGAAAACGCGG
Pma1852         TCACATTTGAGCAAGTCACGCTCTACATCCGTCGGCTACAGCCCGGGCAAGAAAACGCGG
Pma5422         TCACATTTGAGCAAGTCACGCTCTACATCCGTCGGCTACAGCCCGGGCAAGAAAACGCGG
Pma6328A        TCACATTTGAGCAAGTCACGCTCTACATCCGTCGGCTACAGCCCGGGCAAGAAAACGCGG
Pgy2411         TCACATTTGAGCAAGTCACGCTCTACATCCGTCGGCTACAGCCCGGGCAAGAAAACGCGG
                ************************************ ********************

Ppi203          TGGCGAATTCGACCTTGCACCAGCGCCTCACCGCGATCCGCCTGTGGTACGACCACCTGG
Ppi202          TGGCGAATTCGACCTTGCACCAGCGCCTCACCGCGATCCGCCTGTGGTACGACCACCTGG
Ppi223          TGGCGAATTCGACCTTGCACCAGCGCCTCACCGCGATCCGCCTGTGGTACGACCACCTGG
Ppi283          TGGCGAATTCGACCTTGCACCAGCGCCTCACCGCGATCCGCCTGTGGTACGACCACCTGG
Ppi288          TGGCGAATTCGACCTTGCACCAGCGCCTCACCGCGATCCGCCTGTGGTACGACCACCTGG
Ppi374A         TGGCGAATTCGACCTTGCACCAGCGCCTCACCGCGATCCGCCTGTGGTACGACCACCTGG
Ppi390          TGGCGAATTCGACCTTGCACCAGCGCCTCACCGCGATCCGCCTGTGGTACGACCACCTGG
Ppi1452         TGGCGAATTCGACCTTGCACCAGCGCCTCACCGCGATCCGCCTGTGGTACGACCACCTGG
Ppi1456A        TGGCGAATTCGACCTTGCACCAGCGCCTCACCGCGATCCGCCTGTGGTACGACCACCTGG
Ppi1456B        TGGCGAATTCGACCTTGCACCAGCGCCTCACCGCGATCCGCCTGTGGTACGACCACCTGG
Ppi1456C        TGGCGAATTCGACCTTGCACCAGCGCCTCACCGCGATCCGCCTGTGGTACGACCACCTGG
Ppi1456D        TGGCGAATTCGACCTTGCACCAGCGCCTCACCGCGATCCGCCTGTGGTACGACCACCTGG
Ppi1456E        TGGCGAATTCGACCTTGCACCAGCGCCTCACCGCGATCCGCCTGTGGTACGACCACCTGG
Ppi1456F        TGGCGAATTCGACCTTGCACCAGCGCCTCACCGCGATCCGCCTGTGGTACGACCACCTGG
Ppi1939         TGGCGAATTCGACCTTGCACCAGCGCCTCACCGCGATCCGCCTGTGGTACGACCACCTGG
Ppi2889B        TGGCGAATTCGACCTTGCACCAGCGCCTCACCGCGATCCGCCTGTGGTACGACCACCTGG
PsyB728a        TGGCGAATTCGACCTTGCACCAGCGCCTCACCGCGATCCGCCTGTGGTACGACCACCTGG
Psy3023         TGGCGAGTTCGACCTTGCACCAGCGCCTCACCGCGATCCGCCTGTGGTACGACCACCTGG
Pma1852         TGGCGAATTCGACCTTGCACCAGCGCCTCACCGCGATCCGCCTGTGGTACGACCACCTGG
Pma5422         TGGCGAATTCGACCTTGCACCAGCGCCTCACCGCGATCCGCCTGTGGTACGACCACCTGG
Pma6328A        TGGCGAATTCGACCTTGCACCAGCGCCTCACCGCGATCCGCCTGTGGTACGACCACCTGG
Pgy2411         TGGCGAATTCGACCTTGCACCAGCGCCTCACCGCGATCCGCCTGTGGTACGACCACCTGG
                ****** **************************************************

Ppi203          TGTTTCAGGGGCGTTGCGCACAGAATCCGGTACCTCGCGGCCAGCACGGCCGCTTATGTC
Ppi202          TGTTTCAGGGGCGTTGCGCACAGAATCCGGTACCTCGCGGCCAGCACGGCCGCTTATGTC
Ppi223          TGTTTCAGGGGCGTTGCGCACAGAATCCGGTACCTCGCGGCCAGCACGGCCGCTTATGTC
Ppi283          TGTTTCAGGGGCGTTGCGCACAGAATCCGGTACCTCGCGGCCAGCACGGCCGCTTATGTC
Ppi288          TGTTTCAGGGGCGTTGCGCACAGAATCCGGTACCTCGCGGCCAGCACGGCCGCTTATGTC
Ppi374A         TGTTTCAGGGGCGTTGCGCACAGAATCCGGTACCTCGCGGCCAGCACGGCCGCTTATGTC
Ppi390          TGTTTCAGGGGCGTTGCGCACAGAATCCGGTACCTCGCGGCCAGCACGGCCGCTTATGTC
Ppi1452         TGTTTCAGGGGCGTTGCGCACAGAATCCGGTACCTCGCGGCCAGCACGGCCGCTTATGTC
Ppi1456A        TGTTTCAGGGGCGTTGCGCACAGAATCCGGTACCTCGCGGCCAGCACGGCCGCTTATGTC
Ppi1456B        TGTTTCAGGGGCGTTGCGCACAGAATCCGGTACCTCGCGGCCAGCACGGCCGCTTATGTC
Ppi1456C        TGTTTCAGGGGCGTTGCGCACAGAATCCGGTACCTCGCGGCCAGCACGGCCGCTTATGTC
Ppi1456D        TGTTTCAGGGGCGTTGCGCACAGAATCCGGTACCTCGCGGCCAGCACGGCCGCTTATGTC
Ppi1456E        TGTTTCAGGGGCGTTGCGCACAGAATCCGGTACCTCGCGGCCAGCACGGCCGCTTATGTC
Ppi1456F        TGTTTCAGGGGCGTTGCGCACAGAATCCGGTACCTCGCGGCCAGCACGGCCGCTTATGTC
Ppi1939         TGTTTCAGGGGCGTTGCGCACAGAATCCGGTACCTCGCGGCCAGCACGGCCGCTTATGTC
Ppi2889B        TGTTTCAGGGGCGTTGCGCACAGAATCCGGTACCTCGCGGCCAGCACGGCCGCTTATGTC
PsyB728a        TGTTTCAGGGGCGTTGCGCACAGAATCCGGTACCTCGCGGCCAGCACGGCCGCTTATGTC
Psy3023         TTTTTCAGGGGCGTTGCGCACGGAATCCGGTACCTCGCGGTCAGCACGGCCGCTTGTGTC
Pma1852         TGTTTCAGGGGCGTTGCGCACAGAATCCGGTACCTCGCGGCCAGCACGGCCGCTTATGTC
Pma5422         TGTTTCAGGGGCGTTGCGCACAGAATCCGGTACCTCGCGGCCAGCACGGCCGCTTATGTC
Pma6328A        TGTTTCAGGGGCGTTGCGCACAGAATCCGGTACCTCGCGGCCAGCACGGCCGCTTATGTC
Pgy2411         TGTTTCAGGGGCGTTGCGCACAGAATCCGGTACCTCGCGGCCAGCACGGCCGCTTATGTC
                * ****************** ****************** ************** ****

Ppi203          AGGTCCCTGGACATTCAGGCTTCGTAAGAGGGCTGGTACCTCGCTTGATAAAGCTGCCCG
Ppi202          AGGTCCCTGGACATTCAGGCTTCGTAAGAGGGCTGGTACCTCGCTTGATAAAGCTGCCCG
Ppi223          AGGTCCCTGGACATTCAGGCTTCGTAAGAGGGCTGGTACCTCGCTTGATAAAGCTGCCCG
```

```
Ppi283          AGGTCCCTGGACATTCAGGCTTCGTAAGAGGGCTGGTACCTCGCTTGATAAAGCTGCCCG
Ppi288          AGGTCCCTGGACATTCAGGCTTCGTAAGAGGGCTGGTACCTCGCTTGATAAAGCTGCCCG
Ppi374A         AGGTCCCTGGACATTCAGGCTTCGTAAGAGGGCTGGTACCTCGCTTGATAAAGCTGCCCG
Ppi390          AGGTCCCTGGACATTCAGGCTTCGTAAGAGGGCTGGTACCTCGCTTGATAAAGCTGCCCG
Ppi1452         AGGTCCCTGGACATTCAGGCTTCGTAAGAGGGCTGGTACCTCGCTTGATAAAGCTGCCCG
Ppi1456A        AGGTCCCTGGACATTCAGGCTTCGTAAGAGGGCTGGTACCTCGCTTGATAAAGCTGCCCG
Ppi1456B        AGGTCCCTGGACATTCAGGCTTCGTAAGAGGGCTGGTACCTCGCTTGATAAAGCTGCCCG
Ppi1456C        AGGTCCCTGGACATTCAGGCTTCGTAAGAGGGCTGGTACCTCGCTTGATAAAGCTGCCCG
Ppi1456D        AGGTCCCTGGACATTCAGGCTTCGTAAGAGGGCTGGTACCTCGCTTGATAAAGCTGCCCG
Ppi1456E        AGGTCCCTGGACATTCAGGCTTCGTAAGAGGGCTGGTACCTCGCTTGATAAAGCTGCCCG
Ppi1456F        AGGTCCCTGGACATTCAGGCTTCGTAAGAGGGCTGGTACCTCGCTTGATAAAGCTGCCCG
Ppi1939         AGGTCCCTGGACATTCAGGCTTCGTAAGAGGGCTGGTACCTCGCTTGATAAAGCTGCCCG
Ppi2889B        AGGTCCCTGGACATTCAGGCTTCGTAAGAGGGCTGGTACCTCGCTTGATAAAGCTGCCCG
PsyB728a        AGGTCCCTGGACATTCAGGCTTCGTAAGAGGGCTGGTACCTCGCTTGATAAAGCTGCCCG
Psy3023         AGGTCCCTGGACATTCAGGCTTTGTGAGAGGGCTGGTACCTCGCTTGATAAAGCTGCCCG
Pma1852         AGGTCCCTGGACATTCAGGCTTCGTAAGAGGGCTGGTACCTCGCTTGATAAAGCTGCCCG
Pma5422         AGGTCCCTGGACATTCAGGCTTCGTAAGAGGGCTGGTACCTCGCTTGATAAAGCTGCCCG
Pma6328A        AGGTCCCTGGACATTCAGGCTTCGTAAGAGGGCTGGTACCTCGCTTGATAAAGCTGCCCG
Pgy2411         AGGTCCCTGGACATTCAGGCTTCGTAAGAGGGCTGATACCTCGCTTGATAAAGCTGCCCG
                ********************** ** ********* ************************

Ppi203          ACATCCCCACGGATGAGCAGTGGCGTTACTTTCTGAGTATCGCGGCCAGGTCGTCTATTC
Ppi202          ACATCCCCACGGATGAGCAGTGGCGTTACTTTCTGAGTATCGCGGCCAGGTCGTCTATTC
Ppi223          ACATCCCCACGGATGAGCAGTGGCGTTACTTTCTGAGTATCGCGGCCAGGTCGTCTATTC
Ppi283          ACATCCCCACGGATGAGCAGTGGCGTTACTTTCTGAGTATCGCGGCCAGGTCGTCTATTC
Ppi288          ACATCCCCACGGATGAGCAGTGGCGTTACTTTCTGAGTATCGCGGCCAGGTCGTCTATTC
Ppi374A         ACATCCCCACGGATGAGCAGTGGCGTTACTTTCTGAGTATCGCGGCCAGGTCGTCTATTC
Ppi390          ACATCCCCACGGATGAGCAGTGGCGTTACTTTCTGAGTATCGCGGCCAGGTCGTCTATTC
Ppi1452         ACATCCCCACGGATGAGCAGTGGCGTTACTTTCTGAGTATCGCGGCCAGGTCGTCTATTC
Ppi1456A        ACATCCCCACGGATGAGCAGTGGCGTTACTTTCTGAGTATCGCGGCCAGGTCGTCTATTC
Ppi1456B        ACATCCCCACGGATGAGCAGTGGCGTTACTTTCTGAGTATCGCGGCCAGGTCGTCTATTC
Ppi1456C        ACATCCCCACGGATGAGCAGTGGCGTTACTTTCTGAGTATCGCGGCCAGGTCGTCTATTC
Ppi1456D        ACATCCCCACGGATGAGCAGTGGCGTTACTTTCTGAGTATCGCGGCCAGGTCGTCTATTC
Ppi1456E        ACATCCCCACGGATGAGCAGTGGCGTTACTTTCTGAGTATCGCGGCCAGGTCGTCTATTC
Ppi1456F        ACATCCCCACGGATGAGCAGTGGCGTTACTTTCTGAGTATCGCGGCCAGGTCGTCTATTC
Ppi1939         ACATCCCCACGGATGAGCAGTGGCGTTACTTTCTGAGTATCGCGGCCAGGTCGTCTATTC
Ppi2889B        ACATCCCCACGGATGAGCAGTGGCGTTACTTTCTGAGTATCGCGGCCAGGTCGTCTATTC
PsyB728a        ACATCCCCACGGATGAGCAGTGGCGTTACTTTCTGAGTATCGCGGCCAGGTCGTCTATTC
Psy3023         ACATCCCCACGGATGAGCAGTGGCGTTACTTTCTCAGCATCGCGGCCAGGTCGTCTATTC
Pma1852         ACATCCCCACGGATGAGCAGTGGCGTTACTTTCTGAGTATCGCGGCCAGGTCGTCTATTC
Pma5422         ACATCCCCACGGATGAGCAGTGGCGTTACTTTCTGAGTATCGCGGCCAGGTCGTCTATTC
Pma6328A        ACATCCCCACGGATGAGCAGTGGCGTTACTTTCTGAGTATCGCGGCCAGGTCGTCTATTC
Pgy2411         ACATCCCCACGGATGAGCAGTGGCGATACTTTCTGAGTATCGCGGCCAGGTCGTCTATTC
                ************************ ******** ** **********************

Ppi203          GTGATCGGCTTATGCTCTCGTTGGCGTACTGCGGTGCGCTCCGTCGTGCCGAACTGGTGG
Ppi202          GTGATCGGCTTATGCTCTCGTTGGCGTACTGCGGTGCGCTCCGTCGTGCCGAACTGGTGG
Ppi223          GTGATCGGCTTATGCTCTCGTTGGCGTACTGCGGTGCGCTCCGTCGTGCCGAACTGGTGG
Ppi283          GTGATCGGCTTATGCTCTCGTTGGCGTACTGCGGTGCGCTCCGTCGTGCCGAACTGGTGG
Ppi288          GTGATCGGCTTATGCTCTCGTTGGCGTACTGCGGTGCGCTCCGTCGTGCCGAACTGGTGG
Ppi374A         GTGATCGGCTTATGCTCTCGTTGGCGTACTGCGGTGCGCTCCGTCGTGCCGAACTGGTGG
Ppi390          GTGATCGGCTTATGCTCTCGTTGGCGTACTGCGGTGCGCTCCGTCGTGCCGAACTGGTGG
Ppi1452         GTGATCGGCTTATGCTCTCGTTGGCGTACTGCGGTGCGCTCCGTCGTGCCGAACTGGTGG
Ppi1456A        GTGATCGGCTTATGCTCTCGTTGGCGTACTGCGGTGCGCTCCGTCGTGCCGAACTGGTGG
Ppi1456B        GTGATCGGCTTATGCTCTCGTTGGCGTACTGCGGTGCGCTCCGTCGTGCCGAACTGGTGG
Ppi1456C        GTGATCGGCTTATGCTCTCGTTGGCGTACTGCGGTGCGCTCCGTCGTGCCGAACTGGTGG
Ppi1456D        GTGATCGGCTTATGCTCTCGTTGGCGTACAGCGGTGCGCTCCGTCGTGCCGAACTGGTGG
Ppi1456E        GTGATCGGCTTATGCTCTCGTTGGCGTACTGCGGTGCGCTCCGTCGTGCCGAACTGGTGG
Ppi1456F        GTGATCGGCTTATGCTCTCGTTGGCGTACTGCGGTGCGCTCCGTCGTGCCGAACTGGTGG
Ppi1939         GTGATCGGCTTATGCTCTCGTTGGCGTACTGCGGTGCGCTCCGTCGTGCCGAACTGGTGG
Ppi2889B        GTGATCGGCTTATGCTCTCGTTGGCGTACTGCGGTGCGCTCCGTCGTGCCGAACTGGTGG
PsyB728a        GTGATCGGCTTATGCTCTCGTTGGCGTACTGCGGTGCGCTCCGTCGTGCCGAACTGGTGG
Psy3023         GTGATCGGCTTATGCTCTCGTTGGCGTACTGCGGTGCGCTCCGTCGTGCCGAACTGGTGG
Pma1852         GTGATCGGCTTATGCTCTCGTTGGCGTACTGCGGTGCGCTCCGTCGTGCCGAACTGGTGG
Pma5422         GTGATCGGCTTATGCTCTCGTTGGCGTACTGCGGTGCGCTCCGTCGTGCCGAACTGGTGG
Pma6328A        GTGATCGGCTTATGCTCTCGTTGGCGTACTGCGGTGCGCTCCGTCGTGCCGAACTGGTGG
Pgy2411         GTGATCGGCTTATGCTCTCGTTGGCGTACTGCGGTGCGCTCCGTCGTGCCGAACTGGTGG
                **************************** ****************************

Ppi203          CCCTTAGAATCGAAGACCTGGATCTCGCCCATCGACTCATTTCAGTGCGCGCAGAAACGA
Ppi202          CCCTTAGAATCGAAGACCTGGATCTCGCCCATCGACTCATTTCAGTGCGCGCAGAAACGA
Ppi223          CCCTTAGAATCGAAGACCTGGATCTCGCCCATCGACTCATTTCAGTGCGCGCAGAAACGA
Ppi283          CCCTTAGAATCGAAGACCTGGATCTCGCCCATCGACTCATTTCAGTGCGCGCAGAAACGA
Ppi288          CCCTTAGAATCGAAGACCTGGATCTCGCCCATCGACTCATTTCAGTGCGCGCAGAAACGA
Ppi374A         CCCTTAGAATCGAAGACCTGGATCTCGCCCATCGACTCATTTCAGTGCGCGCAGAAACGA
```

```
Ppi390          CCCTTAGAATCGAAGACCTGGATCTCGCCCATCGACTCATTTCAGTGCGCGCAGAAACGA
Ppi1452         CCCTTAGAATCGAAGACCTGGATCTCGCCCATCGACTCATTTCAGTGCGCGCAGAAACGA
Ppi1456A        CCCTTAGAATCGAAGACCTGGATCTCGCCCATCGACTCATTTCAGTGCGCGCAGAAACGA
Ppi1456B        CCCTTAGAATCGAAGACCTGGATCTCGCCCATCGACTCATTTCAGTGCGCGCAGAAACGA
Ppi1456C        CCCTTAGAATCGAAGACCTGGATCTCGCCCATCGACTCATTTCAGTGCGCGCAGAAACGA
Ppi1456D        CCCTTAGAATCGAAGACCTGGATCTCGCCCATCGACTCATTTCAGTGCGCGCAGAAACGA
Ppi1456E        CCCTTAGAATCGAAGACCTGGATCTCGCCCATCGACTCATTTCAGTGCGCGCAGAAACGA
Ppi1456F        CCCTTAGAATCGAAGACCTGGATCTCGCCCATCGACTCATTTCAGTGCGCGCAGAAACGA
Ppi1939         CCCTTAGAATCGAAGACCTGGATCTCGCCCATCGACTCATTTCAGTGCGCGCAGAAACGA
Ppi2889B        CCCTTAGAATCGAAGACCTGGATCTCGCCCATCGACTCATTTCAGTGCGCGCAGAAACGA
PsyB728a        CCCTTAGAATCGAAGACCTGGATCTCGCCCATCGACTCATTTCAGTGCGCGCAGAAACGA
Psy3023         CCCTTAGAATCGAAGACCTGGATCTCGCCCATCGACTCATTTCAGTGCGCGCAGAAACGA
Pma1852         CCCTTAGAATCGAAGACCTGGATCTCGCCCATCGACTCATTTCAGTGCGCGCAGAAACGA
Pma5422         CCCTTAGAATCGAAGACCTGGATCTCGCCCATCGACTCATTTCAGTGCGCGCAGAAACGA
Pma6328A        CCCTTAGAATCGAAGACCTGGATCTCGCCCATCGACTCATTTCAGTGCGCGCAGAAACGA
Pgy2411         CCCTTAGAATCGAAGACCTGGATCTCGCCCATCGACTCATTTCAGTGCGCGCAGAAACGA
                ************************************************************

Ppi203          CCAAAGGCCGACGCAGCCGTGTCGTGTGCTACAGCCCTGACATTGCGCCGATACTTGGAA
Ppi202          CCAAAGGCCGACGCAGCCGTGTCGTGTGCTACAGCCCTGACATTGCGCCGATACTTGGAA
Ppi223          CCAAAGGCCGACGCAGCCGTGTCGTGTGCTACAGCCCTGACATTGCGCCGATACTTGGAA
Ppi283          CCAAAGGCCGACGCAGCCGTGTCGTGTGCTACAGCCCTGACATTGCGCCGATACTTGGAA
Ppi288          CCAAAGGCCGACGCAGCCGTGTCGTGTGCTACAGCCCTGACATTGCGCCGATACTTGGAA
Ppi374A         CCAAAGGCCGACGCAGCCGTGTCGTGTGCTACAGCCCTGACATTGCGCCGATACTTGGAA
Ppi390          CCAAAGGCCGACGCAGCCGTGTCGTGTGCTACAGCCCTGACATTGCGCCGATACTTGGAA
Ppi1452         CCAAAGGCCGACGCAGCCGTGTCGTGTGCTACAGCCCTGACATTGCGCCGATACTTGGAA
Ppi1456A        CCAAAGGCCGACGCAGCCGTGTCGTGTGCTACAGCCCTGACATTGCGCCGATACTTGGAA
Ppi1456B        CCAAAGGCCGACGCAGCCGTGTCGTGTGCTACAGCCCTGACATTGCGCCGATACTTGGAA
Ppi1456C        CCAAAGGCCGACGCAGCCGTGTCGTGTGCTACAGCCCTGACATTGCGCCGATACTTGGAA
Ppi1456D        CCAAAGGCCGACGCAGCCGTGTCGTGTGCTACAGCCCTGACATTGCGCCGATACTTGGAA
Ppi1456E        CCAAAGGCCGACGCAGCCGTGTCGTGTGCTACAGCCCTGACATTGCGCCGATACTTGGAA
Ppi1456F        CCAAAGGCCGACGCAGCCGTGTCGTGTGCTACAGCCCTGACATTGCGCCGATACTTGGAA
Ppi1939         CCAAAGGCCGACGCAGCCGTGTCGTGTGCTACAGCCCTGACATTGCGCCGATACTTGGAA
Ppi2889B        CCAAAGGCCGACGCAGCCGTGTCGTGTGCTACAGCCCTGACATTGCGCCGATACTTGGAA
PsyB728a        CCAAAGGCCGACGCAGCCGTGTCGTGTGCTACAGCCCTGACATTGCGCCGATACTTGGAA
Psy3023         CCAAAGGCCGACGCAGCCGTGTCGTGTGCTACAGCCCTGACATTGCGCCGATACTTGGAA
Pma1852         CCAAAGGCCGACGCAGCCGTGTCGTGTGCTACAGCCCTGACATTGCGCCGATACTTGGAA
Pma5422         CCAAAGGCCGACGCAGCCGTGTCGTGTGCTACAGCCCTGACATTGCGCCGATACTTGGAA
Pma6328A        CCAAAGGCCGACGCAGCCGTGTCGTGTGCTACAGCCCTGACATTGCGCCGATACTTGGAA
Pgy2411         CCAAAGGCCGACGCAGCCGTGTCGTGTGCTACAGCCCTGACATTGCGCCGATACTTGGAA
                ************************************************************

Ppi203          CGCATCTTCATGCCCTTCGTTTGGCCGGTTGGTCGAAAGGAGCCCTGTTTCGATCCGAGT
Ppi202          CGCATCTTCATGCCCTTCGTTTGGCCGGTTGGTCGAAAGGAGCCCTGTTTCGATCCGAGT
Ppi223          CGCATCTTCATGCCCTTCGTTTGGCCGGTTGGTCGAAAGGAGCCCTGTTTCGATCCGAGT
Ppi283          CGCATCTTCATGCCCTTCGTTTGGCCGGTTGGTCGAAAGGAGCCCTGTTTCGATCCGAGT
Ppi288          CGCATCTTCATGCCCTTCGTTTGGCCGGTTGGTCGAAAGGAGCCCTGTTTCGATCCGAGT
Ppi374A         CGCATCTTCATGCCCTTCGTTTGGCCGGTTGGTCGAAAGGAGCCCTGTTTCGATCCGAGT
Ppi390          CGCATCTTCATGCCCTTCGTTTGGCCGGTTGGTCGAAAGGAGCCCTGTTTCGATCCGAGT
Ppi1452         CGCATCTTCATGCCCTTCGTTTGGCCGGTTGGTCGAAAGGAGCCCTGTTTCGATCCGAGT
Ppi1456A        CGCATCTTCATGCCCTTCGTTTGGCCGGTTGGTCGAAAGGAGCCCTGTTTCGATCCGAGT
Ppi1456B        CGCATCTTCATGCCCTTCGTTTGGCCGGTTGGTCGAAAGGAGCCCTGTTTCGATCCGAGT
Ppi1456C        CGCATCTTCATGCCCTTCGTTTGGCCGGTTGGTCGAAAGGAGCCCTGTTTCGATCCGAGT
Ppi1456D        CGCATCTTCATGCCCTTCGTTTGGCCGGTTGGTCGAAAGGAGCCCTGTTTCGATCCGAGT
Ppi1456E        CGCATCTTCATGCCCTTCGTTTGGCCGGTTGGTCGAAAGGAGCCCTGTTTCGATCCGAGT
Ppi1456F        CGCATCTTCATGCCCTTCGTTTGGCCGGTTGGTCGAAAGGAGCCCTGTTTCGATCCGAGT
Ppi1939         CGCATCTTCATGCCCTTCGTTTGGCCGGTTGGTCGAAAGGAGCCCTGTTTCGATCCGAGT
Ppi2889B        CGCATCTTCATGCCCTTCGTTTGGCCGGTTGGTCGAAAGGAGCCCTGTTTCGATCCGAGT
PsyB728a        CGCATCTTCATGCCCTTCGTTTGGCCGGTTGGTCGAAAGGAGCCCTGTTTCGATCCGAGT
Psy3023         CGCATCTTCATGCCCTTCGTCTGGCCGGTTGGTCGAAAGGAGCCCTGTTTCGATCCGAGT
Pma1852         CGCATCTTCATGCCCTTCGTTTGGCCGGTTGGTCGAAAGGAGCCCTGTTTCGATCCGAGT
Pma5422         CGCATCTTCATGCCCTTCGTTTGGCCGGTTGGTCGAAAGGAGCCCTGTTTCGATCCGAGT
Pma6328A        CGCATCTTCATGCCCTTCGTTTGGCCGGTTGGTCGAAAGGAGCCCTGTTTCGATCCGAGT
Pgy2411         CGCATCTTCATGCCCTTCGTCTGGCCGGTTGTTCGAAAGGAGCCCTGTTTCGATCCGAGT
                ******************* ********** ****************************

Ppi203          CTGATCGCAATCGAGGTTCGGCACTCACGCGGTGGACGTGGAGTAAAACGGTAGAAAGAT
Ppi202          CTGATCGCAATCGAGGTTCGGCACTCAC--------------------------------
Ppi223          CTGATCGCAATCGAGGTTCGGCACTCACGCGGTGGACGTGGAGTAAAACGGTAGAAAGAT
Ppi283          CTGATCGCAATCGAGGTTCGGCACTCACGC------------------------------
Ppi288          CTGATCGCAATCGAGGTTCGGCACTCACGCGGTGGACGTGGAGTAAAACGGTAGAAAGAT
Ppi374A         CTGATCGCAATCGAGGTTCGGCACTCACGCGGTGGACGTGGAGTAAAACGG---------
Ppi390          CTGATCGCAATCGAGGTTCGGCACTCACGCGGTGGACGTGGAGTAAAACGGTAGAAAGAT
Ppi1452         CTGATCGCAATCGAGGTTCGGCACTCACGCGGTGGACGTGGAGTAAAAC-----------
Ppi1456A        CTGATCGCAATCGAGGTTCGGCACTCACGCGGTGGACGTGGAGTAAAACG----------
```

```
Ppi1456B        CTGATCGCAATCGAGGTTCGGCACTCACGCGGTGGACGTGGAGTAAAACGGTAGAAAGAT
Ppi1456C        CTGATCGCAATCGAGGTTCGGCACTCACGCGGTGGACGT--------------------
Ppi1456D        CTGATCGCAATCGAGGTTCGGCACTCACGCGGTGGACGTGG------------------
Ppi1456E        CTGATCGCAATCGAGGTTCGGCACTCACGCGGTGGACGTGGAGTAAAACGGTAGAAAGAT
Ppi1456F        CTGATCGCAATCGAGGTTCGGCACTCACGCGGTGGACGGGGAG----------------
Ppi1939         CTGATCGCAATCGAGGTTCGGCACTCACGCGGTGGACG---------------------
Ppi2889B        CTGATCGCAATCGAGGTTCGGCACTCACGCGGTGGACGTGGAGTAAAACGGTAGAAAGAT
PsyB728a        CTGATCGCAATCGAGGTTCGGCACTCACGCGGTGGACGTGGAGTAAAAC-----------
Psy3023         CTGATCGCAATCGAGGTTCGGCACTCACGCGGTGGACGTGGAGTAAAACGGTAGAAAGAT
Pma1852         CTGATCGCAATCGAGGTTCGGCACTCACGCGGTGGACGTGGAGTAAAACGGTAGAAAGAT
Pma5422         CTGATCGCAATCGAGGTTCGGCACTCACGCGGTGGACGTGGAGTAAAACGGTAGAAAGAT
Pma6328A        CTGATCGCAATCGAGGTTCGGCACTCACGCGGTGGACGTGGAGTAAAACGGTAGAAAGAT
Pgy2411         CTGATCGCAATCGAGGTTCGGCACTCACGCGGTGGACGTGGAGTAAAACGGTAGAAAGAT
                ***************************
```

Ap.3.1 Sequence alignment of conserved ILE end from *rulB'* into ILE *xerC/D.*

Multiple alignment of DNA spanning from *rulB'* into the ILE xerC/D gene (conserved end of the ILE.) The highlighted sequence shows the region between the end of *rulB'* and the start of *xerC/D* on the ILE. Stars represent complete identity between all of the sequenced strains at that loci. Analysis was carried using ClustalOmega software available at http://www.ebi.ac.uk/Tools/msa/clustalo/.

Appendix IV: Mean X-fold values of gene expression from *Ppi* 203 and *Psy* 3023.

The values below are the X-fold increases/ decreases in gene expression for ILE and ILE associated genes from *Ppi* 203 and *Psy* 3023. These values were used to construct the heat map figures in Chapter 5.

P. syringae pv. *pisi* 203:

Gene	Control	*Rci*	*xerC*	*ORFD int*	*ORFE int*	*avrPpiA1*	*rulB'*
Condition							
TG apoplastic fluid	1	0.38	0.50	3.47	1.90	2.45	1.04
CW apoplastic fluid	1	0.35	0.72	5.06	4.02	6.22	2.37
RM apoplastic fluid	1	6.70	0.00	0.01	0.00	3.00	0.18
Pea apoplastic fluid	1	0.38	0.46	2.92	1.43	12.70	0.30
TG in planta	1	0.64	0.38	1.01	1.60	1.92	0.13
CW in planta	1	0.11	0.41	2.94	3.97	6.06	0.51
RM in planta	1	0.41	0.00	0.00	0.00	1.25	11.99
Pea in planta	1	0.42	0.43	1.04	2.83	30.78	0.49
Conjugation	1	0.23	0.46	0.40	0.64	0.97	0.37
0.05ug/mL MMC	1	0.70	0.76	0.91	1.10	0.86	0.48
0.1ug/mL MMC	1	0.45	0.52	0.72	0.59	0.55	0.00
0.5ug/mL MMC	1	8.41	9.18	12.92	13.07	14.96	0.00
1ug/mL MMC	1	28.67	39.26	31.17	33.48	30.65	0.01
UV 15sec. Exp.	1	1.86	1.58	0.97	1.44	1.30	0.70
UV 30sec. Exp.	1	7.33	8.69	2.57	4.43	5.79	0.21
UV 45sec. Exp	1	12.68	7.97	9.72	6.24	8.83	1.01
UV 60sec. Exp.	1	11.87	17.45	7.99	15.71	19.69	5.36
-80oC	1	0.28	0.57	1.20	0.19	0.61	0.04
-20oC	1	0.20	0.24	0.22	0.14	0.11	0.11
4oC	1	0.92	1.14	1.35	0.76	1.01	0.36
37oC	1	0.18	0.15	0.32	0.07	0.14	0.18

***P. syringae* pv. *syringae* 3023:**

Gene	Control	*xerC*	*hopH1*	*hopAP1*	*rulB'*
Condition					
TG apoplastic fluid	1	0.72	0.24	0.52	0.45
CW apoplastic fluid	1	7.51	1.84	2.17	4.43
RM apoplastic fluid	1	0.49	2.25	0.56	0.01
Pea apoplastic fluid	1	11.56	0.45	1.00	0.02
TG in planta	1	1.20	1.03	2.68	0.79
CW in planta	1	3.59	2.50	5.25	0.74
RM in planta	1	0.63	1.07	0.34	0.01
Pea in planta	1	14.51	0.55	1.07	0.09
Conjugation	1	12.28	40.41	23.21	1.62
0.05ug/mL MMC	1	11.08	6.38	8.79	0.01
0.1ug/mL MMC	1	6.79	6.65	3.19	0.01
0.5ug/mL MMC	1	8.04	5.30	11.40	0.04
1ug/mL MMC	1	1.42	3.30	3.76	0.03
UV 15sec. Exp.	1	4.51	1.05	2.66	2.73
UV 30sec. Exp.	1	4.65	1.07	1.77	0.35
UV 45sec. Exp	1	2.45	1.43	1.41	1.35
UV 60sec. Exp.	1	1.17	0.23	0.49	0.64
-80oC	1	0.48	1.09	0.68	4.76
-20oC	1	0.66	1.78	0.87	0.40
4oC	1	4.76	2.72	2.39	3.15
37oC	1	0.53	0.50	0.79	1.16